Master Thesis
Software Engineering
Thesis no: MSE-2012-115
March 2013

Analytics-based Software Product Planning

Farnaz Fotrousi
Katayoun Izadyan

School of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona
Sweden

This thesis is submitted to the School of Engineering at Blekinge Institute of Technology in partial fulfillment of the requirements for the degree of Master of Science in Software Engineering. The thesis is equivalent to 30 weeks of full time studies.

Contact Information:
Author(s):

Farnaz Fotrousi
E-mail: farnaz@farnazfotrousi.com

Katayoun Izadyan
E-mail: katti_izadian@yahoo.com

University advisor(s):
Dr. Samuel Fricker
School of Engineering, BTH

School of Computing
Blekinge Institute of Technology
SE-371 79 Karlskrona
Sweden

Internet : www.bth.se/com
Phone : +46 455 38 50 00
Fax : +46 455 38 50 57

ABSTRACT

Context. Successful software product management concerns about developing right software products for right markets at the right time. The product manager, who carries responsibilities of planning, requires but does not always have access to high-quality information for making the best possible planning decisions. The following master thesis concentrates on proposing a solution that supports planning of a software product by means of analytics.

Objectives. The aim of the master thesis is to understand potentials of analytics in product planning decisions in a SaaS context. This thesis focuses on SaaS based analytics used for portfolio management, product roadmapping, and release planning and specify how the analytics can be utilized for planning of a software product. Then the study devises an analytics-based method to enable software product planning.

Methods. The current study was designed with a mixed methodology approach, which includes the literature review and survey researches as well as case study under the framework of the design science. Literature review was conducted to identify product planning decisions and the measurements that support them. A total of 17 interview based surveys were conducted to investigate the impact of analytics on product planning decisions in product roadmapping context. The result of the interviews ended in an analytics-based planning method provided under the framework of design science. The designed analytics-based method was validated by a case study in order to measure the effectiveness of the solution.

Results. The identified product planning decisions were summarized and categorized into a taxonomy of decisions divided by portfolio management, roadmapping, and release planning. The identified SaaS-based measurements were categorized into six categories and made a taxonomy of measurements. The result of the survey illustrated that importance functions of the measurement-categories are not much different for planning-decisions. In the interviews 61.8% of interviewees selected "very important" for "Product", 58.8% for "Feature", and 64.7% for "Product healthiness" categories. For "Referral sources" category, 61.8% of responses have valuated as "not important". Categories of "Technologies and Channels" and "Usage Pattern" have been rated majorly "important" by 47.1% and 32.4% of the corresponding responses. Also the results showed that product use, feature use, users of feature use, response time, product errors, and downtime are the first top measurement-attributes that a product manager prefers to use for product planning. Qualitative results identified "product specification, product maturity and goal" as the effected factors on analytics importance for product planning and in parallel specified strengths and weaknesses of analytical planning from product managers' perspectives. Analytics-based product planning method was developed with eleven main process steps, using the measurements and measurement scores resulted from the interviews, and finally got validated in a case. The method can support all three assets of product planning (portfolio management, roadmapping, and release planning), however it was validated only for roadmapping decisions in the current study. SaaS-based analytics are enablers for the method, but there might be some other analytics that can assist to take planning decisions as well.

Conclusion. The results of the interviews on the roadmapping decisions indicated that different planning decisions consider similar importance for measurement-categories to plan a software product. Statistics about feature use, product use, response time, users, error and downtime have been recognized as the most important measurements for planning. Analytics increase knowledge about product usability and functionality, and also can assist to improve problem handling and client-side technologies. But it has limitations regarding to receiving formed-based customer feedback, handling development technologies and also interpreting some measurements in practice. Immature products are not able to use analytics. To create, remove, or enhance a feature, the data trend provides a wide view of feature desirability in the current or even future time and clarifies how these changes can impact decision making. Prioritizing features can be performed for the features in the same context by comparing their measurement impacts. The analytics-based method covers both reactive and proactive planning.

Keywords: Software product management, product planning, analytics, decision-making, SaaS

Table of Contents

Table of Figures

Table of tables

Table of Equations

1. INTRODUCTION

Software market has evolved from primary developing software as a customized product to developing software as a standard product. During this evolution, the role of product manager and product management function have been emerged within product software companies [1]. Product management is the discipline and business process of managing a product from its inception to the market or customer delivery to produce values for a business [2]. Successful product management aims at developing right products for the right markets at the right time [2]. The success of the developed product relies on different factors, internal, and external stakeholders who are involved in providing a software product plan [1][2][3], and also by the product manager who carries the responsibilities of planning [2].

The trend of changing the software delivery from packaged product to SaaS (Software as a Service) delivery model [4], implies changing in product planning.

Faster release of new features [5], ease of developing more features upon request [5], saving costs [6] in addition to facilitating data collection, provide more rationale for importance of product planning for a SaaS-based application. Small releases in short periods are followed in planning a SaaS product [5].

On the other hand, while planning the development of a software product, a product manger wishes to have high-quality information, but doesn't always have access. That information is used for evaluating and prioritizing the requirements [7] and specifying the scope and timing of releases [8][9]. Today, company-internal stakeholders, focus groups with customer, and reports about user complaints provide inputs for such decision-makings [1]. Information from company-internal stakeholders suffers from accuracy problems because each such stakeholder only represents an intermediary to the market. Also, that person biases the information with his own interests, which may deviate from real market needs. The few customers that participate in focus groups are hardly representative of a whole market. From complaints, problems can be derived in product use, but not whether features are being used or attractive.

SaaS delivery model enables large-scale monitoring of software use [5], and hence can support product planning with first hand information about market needs and attractiveness of software and its features. SaaS providers have access to large amount of useful user data, which supports gaining automatic feedback for analytics.

Analytics uses data to enhance the quality of information by reflecting real market and customer needs, both quantitatively and qualitatively. This information makes moving from an ad hoc intuitive planning toward a more logical planning. By using analytics, potential features for enhancing a product can be identified. Analysis of the past, current, or future state of related data and corresponding trends provide the information that leads taking more suitable and logical planning decisions. The information can support both proactive and reactive planning. From a product planning perspective, finding right data to be used for right decisions, selecting right piece of data from huge amount of collected data and the interpretation of their values are challenges of using analytics. Another limitation of using analytics returns to collecting them. The product that has not implemented yet cannot provide the opportunity to be monitored for analytics, which may lead to devise another strategy such as prototyping for collecting the data.

Ignoring the current accurate and representative data about product or feature attractiveness can lead to wrong product planning decisions. By ignoring analytics, decisions will be made based on opinions and will not reflect the real customer's requirements, specially in the bespoke products which customer are known and can be involved [8]. Also, some planning decisions will rely on intermediaries' interests and blur real market demands. Using a limited number of stakeholders decreases the sampling population and hence, reduces quality and quantity of market indicators. Lack of continuous monitoring of customers, prevents product managers to perceive trends of feature's attractiveness and changes in market conditions. Thus, the company will not be able to adjust its offerings in a timely manner.

Understanding the effect of analytics on decisions of product planning and proposing a solution for analytics-based planning are the aims of the current study. For this purpose, different decisions related to core assets of product planning such as portfolio management, roadmapping, and release planning were identified by a literature review. SaaS-based measurements were specified by the literature review as well. Then, an interview-based survey was conducted to analyze the impact of analytics on planning. During the interviews, all the interviewees selected roadmapping as the type of decision-making that they were involved most. So the study results were concluded for the roadmapping decisions. Finally, a solution was proposed to apply analytics for product planning, which was validated within a case.

So, the contributions of the study were as follows:

- Identify and classify product planning decisions.
- Identify and classify SaaS-based measurements.
- Identify the importance of analytics in making product planning decisions.
- Proposed an analytics-based method, which will be validated within a case.

The current master thesis is divided into eight chapters. The thesis starts with "Introduction" chapter and then "Background, and Related Work" chapter presents the research background in addition to the taxonomy of decisions and taxonomy of measurements as the related work. Chapter 3 concentrates on "Research Methodology". Interview-based survey is considered in chapter 4, which focuses on impacts of analytics on product planning decisions through an empirical study. Chapter 5 provides an analytics-based product planning method and its validated results within a case. It follows the design-science research guidelines [10] and provides a variable artifact in a form of a method that is demonstrated by well-executed evaluations. This design-science research has technology-oriented as well as management-oriented audiences. Discussions, Lesson learned, Conclusion and References are presented in chapters 6 to 9 respectively.

2. BACKGROUND AND RELATED WORK

Product planning involves processes through which a product is conceived, brought to market and managed across its life cycles [11]. Well development of planning in an organization can reduce resources' costs and increase revenues and profits [11]. Planning of software product is considered as a process area of software product management [1][2][3][6][12], that serves the required information for product planning. A product manager who is chiefly responsible for product planning task [6], wishes for high quality information that are not biased by personal interest and opinions of customers, partners and company internal stakeholders. The information assists him to understand real customer and market needs with large sample population, find out attractiveness and drawbacks of current product, and guide sales to achieve a well-designed product plan.

Analytics has the potential to address the product manager's wishes. It provides measurements that are not affected by power and politics for guiding sales and marketing [13], for informing usability, reliability, and quality of service engineering [13, 14], and to support quality assurance. Despite their importance, analytics have not been used yet for guiding product planning. It is unclear whether and how analytics can be used to evaluate and prioritize the development and evolution of product requirements/features and which SaaS analytics should be used for that purpose.

Specifying the effect of analytics on decisions of product planning in a SaaS context product and devising a method to enable analytics-based product planning are the aims of the current study. For this purpose, product planning decisions and SaaS based analytics should be extracted from literature and categorized. This section will present the literature behind the study: Section 2.1 focuses on core assets of product planning and introduces the decisions, which product managers made in those areas. Then other approaches of product planning are identified in order to compare them later with our proposed analytics based method. Section 2.2 expands on definitions for analytics and SaaS-based measurement attributes suit for product planning. Then it introduces other measurement solution presented in literature and the differences with the selected attributes. Section 2.3 provides a summary of the SaaS opportunities for product planning.

2.1 Software Product Planning

When a product is planned, decisions are taken at four levels in a company: portfolios, roadmaps, release plans and requirements, which are identified through the process of portfolio management, product roadmapping and release planning [1][3]. Portfolio management deals with decision-making for existence of product(s) by considering the market trends and development strategies [8] for the levels of the company or business unit [14]. Product roadmapping addresses features in different releases of the product, specifies major technology areas [3][15] and simplifies release planning [3]. It captures long-term plans for product evolution and provides a bridge between management, market and product development, which specifies product positioning and development aspects [16] in order to link business view to requirements through high level definitions of the future features [17]. Release planning deals with requirements of each release [1], which involves requirements elicitation and allocation of the prioritized requirements to development projects [3]. Release planning scopes development projects [8] and addresses the process of deciding which requirement of an evolving software system should be assigned to which release [18][19]. Different factors are known as criteria for deciding whether a requirement is included in a specific release [7] or not.

Product planning contains strategic, tactical and operational activities [20]. A product manager, who is chiefly responsible for product planning tasks [6], is mostly business oriented and involves in strategic and tactical aspects in comparison with a project manager [20]. Activities in portfolio management and roadmapping are in the strategic and tactical levels, while release planning is mainly in the operational level [21].

Decisions that product managers make during the product planning processes are about creation, change, deletion, and allocation. At the portfolio level, these decisions concern products and at the roadmap level they concern features that these products consist of. The requirements that the features contain are concerned in the release level. Confirm a technology for such product, features, and requirements and also prioritizing features and requirements, support such decision-making. Table 1 gives an overview of decisions that are reported to be made in planning of a software product through the taxonomy of product planning decisions. The taxonomy that was extracted from literature, through the current study, validated with several product managers via interviews to contain academic credibility to be relied on. The literature review execution process have been presented in [Appendix A.2]

In literatures, product manager takes decisions about planning processes from different prospective such business-driven perspective (i.e. ”assess feature business value”, “determine cost”) [43], capability perspective (i.e. “requirements gathering”, “requirements prioritization”, “release definition”) [36] or release planning perspective (i.e. “elicit requirement”, “specify problem”, ”estimate resource”, ”estimate alternatives plans”, and “implementation ”) [35]. The decision taxonomy includes all these decision components but in a higher level. For a single decision in the taxonomy, several minor decisions are discussed and confirmed. As an instance, decisions about initial time estimation in roadmapping, estimated efforts for requirements, available alternatives of the plan [35] are placed under “allocate feature to releases” decision in the taxonomy.

Table 1: Taxonomy of product planning decisions

Category	Decision	References
Portfolio Management	Create a new product	[2][22][23]
	Remove an old product	[2]
	Confirm a new technology for developing a product	[22]
Roadmapping	Create a new feature for the current product	[24]
	Remove a feature from the current product	[25]
	Enhance a feature(s) in the current product	[2]
	Prioritize features in the current product	[24][26][27] [28]
	Allocate resources	[15] [16] [24] [27][28]
	Allocate features to releases	[15][24][26][27][28][29]
	Confirm a new technology for developing a feature(s)	[15] [16] [24][27] [30] [31]
Release Planning	Create new requirements for a feature(s)	[1] [32][33] [34] [35] [36] [37]
	Change a requirement(s) in a feature	[32][31][37]
	Remove a requirement from a feature(s)	[38]
	Prioritize requirements	[1][31] [36] [39] [40]
	Allocate resources	[32] [33] [41][34] [38] [42]
	Allocate requirements to releases	[1] [15] [18] [33][31] [34] [35] [36][42] [40]
	Confirm a new technology for set of requirements	[34]

Product plan models [15][16][44], adapted in several details based on product and reason of utilizing the plan. The comprehensive models integrate market, products, technology, people, and processes [22]. The standard T-plan [45] framework, is one of these models is used for more than one decade and supports product planning. This framework consists of four workshops that involve activities of identifying market and business drivers, conceptualizing the product(s), and identifying current and future technologies which all lead to take the activity of constructing the roadmap in a time-based manner [45][46]. Later on, this model was adapted by R. Phaal and G. Muller [44] with a schematic multi-layered roadmap in 3 main perspectives: “commercial and strategic” (i.e. market, business), “design, development, and production” (i.e. product, service, system), and “technology and research“ (i.e. technology, science, and resources perspectives) as it have presented in Figure 1.

The perspective can be also supported with the high level planning decisions of those shown in Table 1.

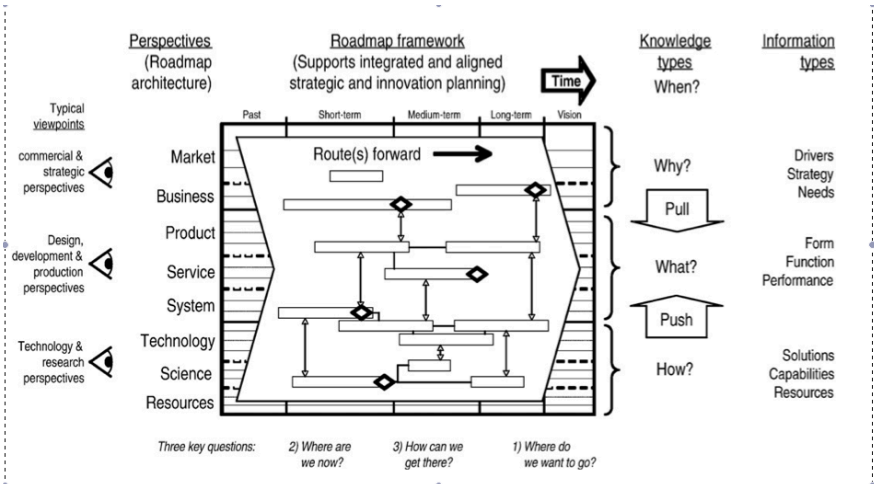


Figure 1: Schematic multi layer roadmap, aligning strategy (R. Phaas and G. Muller, 2009 [44])

Product manager involves two types of planning: reactive and proactive. In reactive product planning, the main focus of product manager is to keep the business going. Therefore any action or change should be taken to respond to opportunities and threats related to a product. Within the reactive planning, software product manager regularly analyzes the soft measures in comparison with the product plan and performs an action for significant deviation from planned measures [6]. In a proactive planning, a product manager upgrades the product with new features or proposes the release of new products. These kinds of decisions are based on predictions of the product's future state with the aim of solving problems and satisfying customers but also can be result of technology-push [47]. In both approaches stakeholder's feedback provide broader perspective on current or future requirements of stakeholders and market requirements.

Different approaches have been studied for product planning but none of them refers to analytics as a potential input for product planning. The approaches of Ad-hoc planning, systematic planning (e.g. Cost-Value approach) or hybrid approaches (e.g. Evolve*) [8] concentrate on techniques to be used for requirements or features selection and prioritization [48][49][32], in addition to allocate them to releases [24][33][26]. In all these techniques, stockholders' preferences specify the value of features, which can be considered as a notable challenge.

Another approach for planning addresses understanding the value of requirements or features for planning. Feature tree model eases the planning of an evolving software release to reduce the complexity of planning, increase the trust, and help decision-makings to what and when to implement [15]. A decision framework based on combination of feature, times, and cost is provided to fulfill the main software product planning goal which is maximizing the product's value within available resources [43]. With similar goal another model [41] integrates the degree of stakeholders' satisfaction with the value of features and their urgency levels. However identifying feature's value and urgency levels could be facilitated with more accuracy by monitoring of the product use.

Addressing the product planning challenges can be helped by data about product use. Twelve challenges have been recognized during planning which are categorized as human-oriented and system-oriented categories [9]. The challenges are mainly studied on release planning but most of them are also shown up in portfolio management and roadmapping. The following challenges have been seen as the samples: Foreseen feature release, prioritization of requirements and features, supporting old release, project monitoring, stakeholders' involvement, and interdependency among systems. Data about how product is used, which features is more attractive and what are the user's interests can support the challenges of product planning in the form of analytics.

Monitoring data from an online product enables the assessment of products, features, and requirements to understanding stockholders' preferences and provide essential data for guiding the planning in both perspectives.

2.2 Analytics

Analytics is an extensive use of data, statistics and quantitative analysis [50] to obtain insight and actionable information [51] for a data-driven decision making [52][53]. Analytics evaluates the monitored data in past, present and future to extract important information in order to assist decision makers [52] to take a better decision as the main goal [54].

Analytics is important for product planning because they create insights about customers' preferences and show what attracts them to use the product, what encourages them to do activities for more value for business, and what keeps them as royal customers [55]. If software firms can see precisely how customers are satisfied with the product when they're using the product, where they're running into difficulty, and how to engage and retain them, the software product planning can be scheduled well. Analytics enhances the quality of information by reflecting real market needs [56] and can be a facilitator for product managers in their decision makings.

Analytics can be performed from multiple sources such as web [57]. Web analytics is the variety assessments of data for a general understanding of the visitors' experiences online for the product use [55]. Detailed statistics can be collected from different sources such as web traffic, web-based transactions, web server performance, usability studies and user submitted information. Well-established tools have to be provided to measure them through server requests, JavaScript tags and client's cookies [58] (while the permission from clients is being received according to cookies law).

Analytics provides essential data as a source of measurement for understanding customer, user, and product use. There are variety of proposed measurements for analytics [55][58][13][59][60][61][62]. Analytical tools such Google analytics, Piwik, Yahoo web analytics, Stat Counter, New relics and Woopra also support the measurements.

Analytics literature conceptualized a SaaS application as a product that consists of different features rather than pages. These features are accessed with different usage patterns from several sources while specific technologies or channels are used for the access and the product healthiness is being achieved. Analytics measures attributes of theses conceptual elements. These attributes are categorized and organized as a taxonomy of measurement-attributes presented in Table 2. The entities that are measured through the study will be referred as the measurement-categories that each includes corresponding measurement attributes. Interpretations of the measurements have been presented in Appendix A.1. This taxonomy has been developed through a literature review execution, which can be found in details in Appendix A.2.

Table 2: Taxonomy of measurements for SaaS-based applications

Measurement-categories (Measurement Entity)	Measurement-attributes
Product (i.e. Value of the product from user's perspective)	Product use Overall amount of users Time between visits Duration of using the product New users Returning users
Feature (i.e. Value of the features from user's perspective)	Users that use a feature Feature use Duration of using a feature Entrance feature Exit feature Bounce
Usage Pattern (i.e. Usage pattern of the product)	Click activity Depth of use Click stream/path
Referral Sources (i.e. Referral sources for product use)	Referrers Location/ISP per use Search engines and keywords Campaigns
Technologies and Channels (i.e. Technology and channels used to access the product)	Languages Browsers Operating Systems Plugins Screen resolution
Product Healthiness	Errors Downtime Response time Throughput DOS attack Worm attacks

Being simple SaaS-based measurement-attributes, none-confidential data of organizations and supported by general analytical tools were the criteria for filtering the attributes. So other similar categorizations for analytics [13][55] needed customization to support the exclusion criteria as well as bringing the product planning concept to the center of attentions. Although the taxonomy is in different perspective but it mainly support other categorization as well.

The nine categories introduced at [13], are supported by the taxonomy. Some categories such as "how healthy is my infrastructure?" can directly be mapped to "Product healthiness" category of the taxonomy. Data about "where is my traffic coming" category is presented in "referral sources" category and data about "how well did the visitors benefit from my business" is covered in the "usage pattern" category of the taxonomy. The category of "what is working best (worst)?" includes the measurements of both "product" and "feature" categories.

The analytics categorization from perspectives of customer "reach", "acquisition", "conversion" and "retention" [55] are covered by the analytics taxonomy in this study. The measurements under the categories of "reach" and "acquisition" are supported by the taxonomy through the categories of "product", "feature" and "referral sources". The "retention" category defines complex measurements, which are a mixture of "product", "referral source", and "usage pattern" measurements. The measurements related to "returning visitors' activities" which is defined under "retention" category can be calculated by considering measurement-attributes of "returning users" from "product" and "click activity" from "usage pattern" categories in the taxonomy. The "usage pattern" category also supports the measurements corresponding to "conversion" category from [55] study.

Companies can use analytics in several areas further than web analytics. Business analytics provides better insights particularly from operational data stored in transactional systems to achieve business effectiveness [57]. For planning, organizations use analytics in external areas such as potential and existing competitors, suppliers, customers and substitutes (Porter's five force), sales, market, price optimization and internal areas of processes, operations, and resource management [63][57].

Analytics is also used to inform usability, reliability, and quality of service engineering decisions. In usability engineering data about page use frequency, users, click paths, and events are stored in web-log files and combined with user feedback [64][65]. These measurements are usually collected during user testing or after a release. The data is used to evaluate the acceptance of a solution for given classes of users. One of the established methods used for analytics-based usability engineering is A/B testing, which allows comparing the attractiveness of two alternative designs [66].

In reliability engineering data about product availability, probability of fail and fail safe are used for both software and hardware [67][68]. These measurements are usually collected during entire product life cycle, including planning, development, testing, manufacturing, operation and maintenance. The data is used to cope with the probability of failure for features, components and products and is known as the heart of risk analysis and quality assurance. Reliability engineering depends on probabilistic methods such fault tree to predict whether the reliability is fulfilled [67].

In Quality of Service (QoS) analytics measures availability, duration, performance, security, response time and throughput [69][70]. These measurements are usually collected during and after release and widely drawn attention in network, multimedia, distributed and real time products [71][72]. The data is used to ensuring a high-quality combination of multiple quality attributes. The QoS approaches such as QoS computation model address different perspectives of quality attributes and domain specific criteria [70]. Also, QoS supports performance engineering by analytics of workload intensities, delay, loss ratio and throughput to assure meeting performance objectives in all software engineering activities and analyses [72][73].

Even though needs of product planning for analytics overlap, but differ in some important aspects from other web-engineering domain. Conceptualizing a product consist of features instead of web application that has pages[13] and looking into use cases and quality attributes inside a page as a feature requires a new perspective. Product managers are interested to look insight about customer preferences in order to create business value [6] into the plan is another motivation for the study. How product managers would use analytics, has not been established yet and is the subject of the current research.

Although relying on just web analytics may not provide data for all aspects in product planning (e.g. competitors, suppliers), the amounts of data that are exploding in a SaaS-based product provide the good opportunity to gain insight about customer preferences and products.

2.3 SaaS opportunities

Online businesses delivered through the SaaS model are gradually gaining attentions [74]. SaaS model removes the necessity of installing a software application on user's local devices and delivers it through a thin client interface such as a web browser by hosting it via the Internet [5][56][75] that provides the opportunity of having thousands of users. Therefore this service reduces the difficulty of product maintenance and the purchase expenses by on-demand pricing [76]. A SaaS application, which is also known as a type of hosted application, have improved some weak points of on-premise software such as high operational costs, high subscriptions costs [77] and long releases [5] for the product. It is one delivery mode of cloud computing with the abilities of providing dynamic scale for applications, resources and resource utilization monitoring [78].

The SaaS model is particularly attractive for supporting the software product planning. It enables large scale monitoring of software use [5] and provides information about market needs and attractiveness of product and its features. Characteristics such as centralizing feature updating[79], faster release of new features [5], ease of developing more features upon request [79] , cost savings [6] in addition to facilitating analytics collection provide more rationale to concentrate on utilizing SaaS-based analytics for planning of a software product in the current study.

For a SaaS-based product, delivering high quality product, which convinces subscriber to renew their subscription is the major goal. So the measurements related to productivity, performance and usability are the most important metrics in this application type [13]. But planning a SaaS-based product requires more study to answer questions about: what can analytics do for planning decisions? How can they support the product planning decisions? Can effective measurements be collected for product planning? If the measurements can be collected another question is arisen: What is an effective way? These questions will clarify the problem statement of the current research.

3. RESEARCH METHODOLOGY

3.1 Aims and Objectives

Specifying the effect of analytics on decisions of product planning in a SaaS context product and devising a method to enable analytics based product planning are aims of the current study. This thesis will focus on measurements used for portfolio management, product roadmapping and release planning, and evaluates how the measurements can be utilized for planning decisions. Aims and objectives are summarized as follow:

- Describe how the product planning decisions can be informed by analytics.
- Propose and evaluate an effective analytics-based product planning solution.

3.2 Expected Outcomes

The outcomes of the project will cover the following points:

- A list of weighted measurements for planning decisions. This will lead product manager to take the planning decisions based on the importance levels of information extracted from measurements.
- An analytics-based method for software product planning, which will be developed by considering the values of SaaS-based measurement-attributes on planning decisions. The method will allow a product manager to transform the information of measurements to a recommendation for a planning decision.

3.3 Research Questions

The Table 3 presents research questions and their aims for the current study.

Table 3: Research questions and their aims

Research Question	Aim
RQ1: How does analytics support planning of a SaaS product? **RQ1.1: Would product managers use different analytics for different type of product planning decisions?** **RQ1.2: What analytics do product managers prefer for a given product planning decision?** **RQ1.3: Why does a product manager consider analytics and why not?**	To provide the association between analytics and decision-making of product planning
RQ2: What is an effective method to transform measurements of SaaS use into recommendation for product planning?	To provide analytics-based method for taking product planning decision.

3.4 Research Processes

The study was conducted through an interview-based survey [80] to answer RQ1, and then a method was constructed in an overall framework of a design science research [85], which was evaluated through a case study research [87] (to answer RQ2). Figure 2 presents an overview of performing the study. In the starting point of the study, literature review was conducted to find a taxonomy of product planning decisions [Table 1] as well as taxonomy of measured attributes [Table 2], applicable in a SaaS-based application They are considered as the inputs for the rest of the study. By conducting an interview-based survey, the associations of measurement-categories and measurement-attributes with planning decisions were identified (RQ1.1 was answered) and then the overall value of each measurement for product planning was evaluated (answer RQ1.2). After that, benefits behind using analytics by a product manager were also clarified (answering RQ1.3). All answers for RQ1.1, RQ1.2 and RQ1.3, answered the RQ1 as well. Experiment study was another alternative, however it was dismissed, as answering the research question needs factual information from groups of practitioners (the more the better) with product management knowledge and generalized it, while experiment has the limitation respectively.

At the next step, an analytics-based planning method was developed by design science research method, which applies measurements weights as inputs and generates outputs as high quality information for product manager to take decisions (RQ2 was answered). There were other alternatives for design science research such as action research, however it was ignored as it focuses on changes while this study needed concentrating on design artifacts.

The proposed method was validated by conducting a case study in a software company. Experimentation could be an alternative for that evaluation, however that approach was dismissed because of several reasons. First of all, if it was supposed to be performed by real product managers, there was a limitation of finding experiment's subjects, since the product managers are usually too busy to spend several hours for conducting an experiment (i.e. monitoring the product for months, finding related measurements one by one and interpreting each change). If the experiment supposed to be performed by students who had product planning knowledge, the results had shortage of industrial experience for a real product. So we might not be able to generalize it to the real situation. Also the experiment would look at the artificial environment while a case study focuses on available phenomenon within real-life context and is the preferred strategy to answer "how" question.

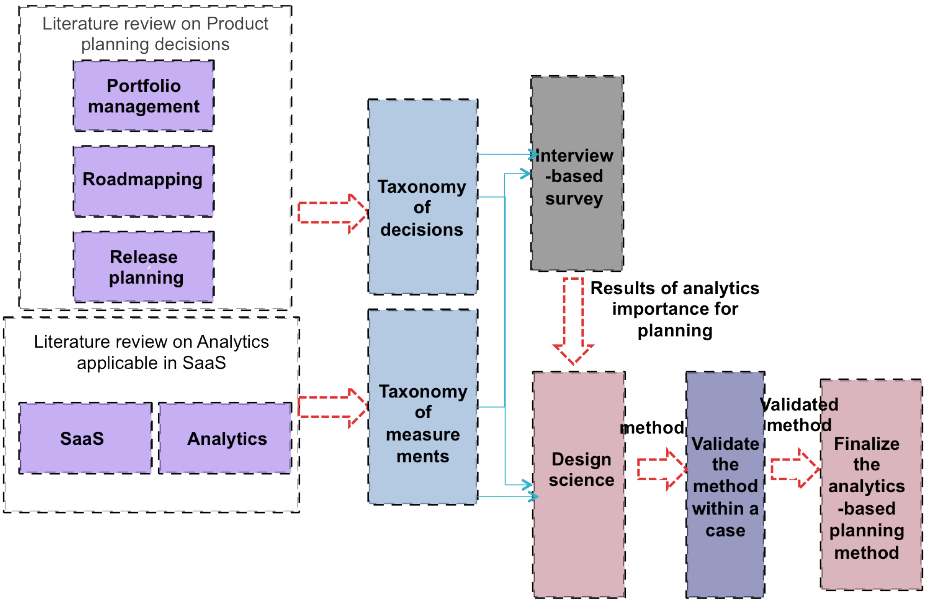


Figure 2: The overall processes of the current research

3.4.1 Interview-based Survey

The survey was conducted with the purpose of identifying the impact of analytics on product planning decisions, which is done by analyzing the effects of measurement-categories and measurement-attributes on the decisions. Extracted taxonomy of decisions and taxonomy of measurement-attributes were the inputs of the interview-based survey.

The survey was conducted based on the guideline for eleven stages of survey research process [81]:

Stage 1: Identify factors of the study and the method of the research

The goal of this study was to understand the effect of SaaS-based analytics on decisions related to product planning.

Data collection was performed by means of phone interview-based survey to prevent misinterpretation of the questions. It was a semi-structured interview-based on structured questionnaire to find out unspecified helpful information [Appendix B.1].

To avoid disadvantages of telephone survey related to lack of visual material and avoid complexity, the screen of the interviewer's computer that presents the questionnaire was shared with interviewees through web-based screen sharing applications.

Stage 2: Determining the research schedule and budget

For conducting interviews, a timetable of 45 days was established which also considered unpredictable delays. As it was a thesis, no budget has been considered for the research.

Stage 3: Establishing an information base

Before designing the survey instrument, a study was conducted to cover the decisions of product planning and SaaS-based measurements. The information in the form of taxonomy of planning decisions and taxonomy of SaaS-based measurements were the inputs for constructing the survey instrument. The taxonomy of decision provides three main decision categories named Portfolio management, roadmapping and release planning with their corresponding decisions [Table 1]. Also the taxonomy of measurement presents six main measurement categories with their related measurement attributes [Table 2]. The information was considered as inputs for the survey. All interviewees were asked about the impact of measurement categories and attributes on product planning decisions.

Stage 4: Sampling

A list of participants as the working population (sampling frame) for the interview was identified including product managers, other professionals or managers who were involved in product planning inside Sweden. For this purpose, interview invitations were sent to product managers and professionals, who were introduced by "Tolpagorni Product Management AB" and were active at the "International Software Product Management Association (ISPMA)". "Tolpagorni Product Management AB" is a Swedish specialist company in Product management that offers services include consulting, coaching, workshops and conferences, both public and in-house trainings. Also after each interview, interviewees were asked to recommend another professional product manager who might be eager to take part in the interview.

Stage 5: Determining the sample size

The interviews were conducted with the population of 17 interviewees who were proficient in product management. In the first phase, the whole interview plan was started by 12 pre-scheduled interview sessions, because the result of previous studies show that the data saturation usually occurs within the first twelve interviews [82]. However after first data analysis, since the data were not sufficient for categorical analysis of decision types, the second phase of interviews was scheduled with five more interviewees. The total number of 17 interviews fulfilled the initial requirements for data analysis, and due to time limitation conducting more interviews for improving the results was considered as future researches.

Stage 6: Designing the questions

The questionnaire was started with questions about context facets of the product, organization (company size and development team size) and people (role and experience). Context of the study has a large effect on drawing a conclusion when study evidences are integrating [83], since different researches might concentrate on the same study while ending in different conclusions due to domain differences of people skill, organization size, cultures, and people roles [83].

Questions about product planning formed the core of the interview, which was followed up within two parts: "Planning Decisions" and "Analytics" [Appendix B.1]. First interviewees were questioned about a product that they have planned and are most satisfied with [Appendix B.1, "Demographic and Decisions" section]. Then questions were asked about their taken planning decisions in that particular product [Appendix B.1, "Demographic and Decisions" section]. Later on, the impacts of measurements on product planning decisions were considered [Appendix B.1, "Categories of analytics" section]. The third part of the planning questions concentrated on the importance level of measurement-attributes and measurement-categories [Appendix B.1, "Analytics" section].

As it can be seen in [Appendix 3.1], most questions (except question number 1,3,4,5 and 6) had multiple-choice answers or drag-and-drop answers. The interviewees could select the answers among those pre-defined answers, which made our quantitative results.

However to avoid incompleteness of results, each question was accompanied with one open-text box answer. This allowed the interviewees to fill their desired response, which might not be included in pre-defined answers.

Stage 7: Pretesting the survey instrument

Interviews were piloted by population of two product managers and two students who had product planning knowledge. The product managers helped us in testing and improving the content of survey. They also took part in main interviews, as their experiences were valuable. The survey was piloted for 20 days. After initial testing and several refinements, the second phase was started. The main goal of piloting the interview was to be sure of selecting appropriate questions and multiple-choice answers that could fulfill the goal of RQ1 research question.

Stage 8: Selecting and training interviewees

General training about definitions of portfolio management, roadmapping, and release planning were provided for interviewees and then they were instructed with an overview about various types of questions to be asked.

Stage 9: Implementing the survey

The interviews were fulfilled by calling through the mobile phone or the computer telecommunication programs and the timetable was maintained strictly. Before each interview an access to online survey questions were provided for each interviewee with the help of Survey Gizmo tool. Therefore the interviewees could easily follow up questions that were asked by the interviewers.

Stage 10: Coding the completed questionnaires and computerizing the data

Survey Gizmo tool was considered for developing interviews questions, and the answers were entered to the computer for data processing using the tool and in parallel, the interviews were recorded after getting permissions from interviews in the sake of future reference.

Stage 11: Analyzing the data and preparing the final report

The recorded answers by the Gizmo tool were exported to a tabular format to prepare it for quantitative analysis and the recorded voices were transcribed for qualitative analysis.

Understanding any existing relation between product planning decisions and the category of measurements was interesting for answering RQ1.1. This required quantitative data collected from "Categories of analytics" section of the survey instrument and was fulfilled by categorical analysis. The analysis was conducted using Kruskal-Wallis tests to accept or reject the hypothesis corresponding the objective of the RQ1.1. Since non-normal data was the pre-requisite for using the Kruskal-Wallis tests, normality was checked by Kolmogorov-Smirnov test.

Descriptive analysis was conducted to show which measurement attributes had more value for a given product planning decision (to answer RQ1.2). Quantitative data collected from the "Categories of analytics" and "Analytics" sections of the instrument were able to fulfill the objective. It was performed by calculating scores of measurement-categories and measurement-attributes inside each category, and then specifying the value of measurement-attributes amongst the all. The measurement-attributes were categorized using independent T-test to specify the group of measurement-attributes that are more preferable for product planning. Confidence intervals of the values were calculated using T-test.

To analyze qualitative data resulted from the interviews and specify the advantages of using analytics for product mangers and reasons behind their selections for analytics importance, content analysis was selected [84]. Through the interviews, interviewees were questioned about their arguments for their particular answers. For instance for "categories of analytics" section, if they believed that one category had "important" value on one decision, they were asked about the reason behind that. These arguments were analyzed by means of the content analysis method.

3.4.2 Method Construction and Case Study based Evaluation

One important contribution of this thesis is to design an analytical-supported method for planning a software product. The method was formulated using the values of measurement attributes for product planning, which is the answer of RQ1.2. This study constructs a method in the overall framework of the design science and then uses the case study research method to demonstrate and evaluate the method's usefulness. The design science approach introduces a set of activities for designing an innovative artifact in order to enable researchers for better understanding of the problems and use improvement feedback after evaluations [10]. In the following, seven guidelines of design science in information systems [85] will be applied to construct the analytics-based method and then method's demonstration and validation are performed by a case study research.

Objective of the method

The main objective is to construct a method that describes how product-planning decisions can be made by the help of analytics.

Design and development

This activity involves defining functionality and architecture of a new artifact, which is defined as an analytics-based method in this study. As a software application solution, the analytics-based method [Figure 3] is proposed for planning of a software product and supported by pre-rated measurements [section 4.1.2.2.2] and a set of instance decisions (from taxonomy of decisions in section 2.1). In this method, related measurements to the features are monitored, changes in the measurement values are evaluated, and then positive or negative impacts on the related decision are traded off. The output provides supportive information for the product manager to evaluate the decision from product perspective. To identify the final decision and its alternatives, this output usually does not suffice and it might be needed considering other criteria such as availability of resources, market goals, and competitors [7], which does not recite in the scope of the method. The proposed method also collects feedback from the previous experiences of product managers.

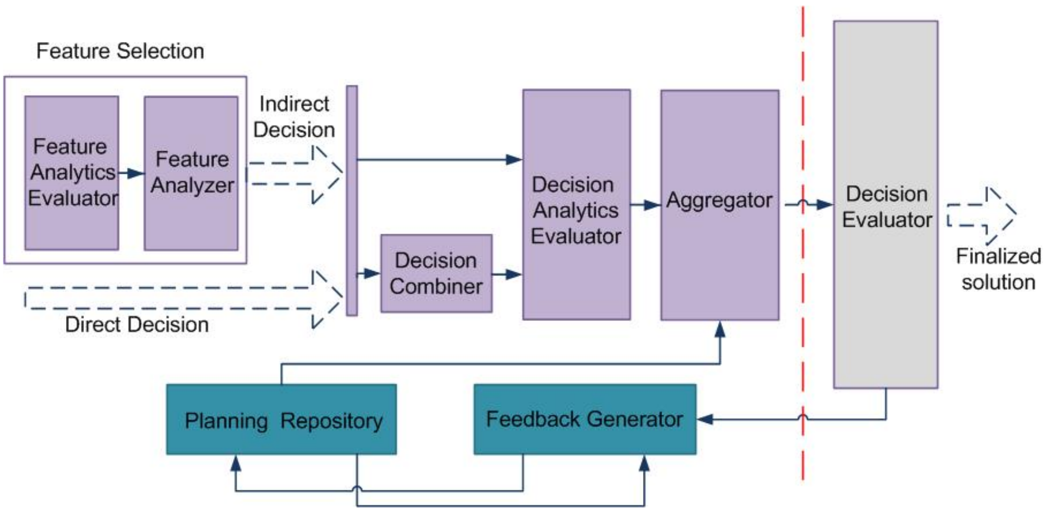


Figure 3: Overall view of analytics-based method to support product planning decisions

The method provides two types of analytics support: feature selection and planning decision supports. The first type considers monitoring feature-related analytics to determine higher priority features and the second type uses analytics deeply to study a planning decision in detail and provide decisions support information in the form of a recommendation indicating whether the decision should be taken or not. A product manager or a professional involves in product planning are actors of the method. The method is designed based on the following main components illustrated in Figure 3:

Feature Analytics Evaluator
Input: No input.
Objective: This component concerns the evaluation process of analytics for features. Studying different sources of information related to a SaaS-based product, determines a short list of features to be analyzed. The main goal of this component is to identify measurable attributes for a feature and then monitor and observe changes of the measurements.
Output: A list of observed data for each feature is the output.

Feature Analyzer
Input: Data about observed measurements related to a feature(s)
Objective: The objective is to analyze the information extracted from measurements and select higher priority features to be decided about. By studying changes of measurement values and considering external factors such as strategies defined for the product, a feature priority is analyzed by the product manager in order to select the feature to be decided.
Output: A product planning decision(s) for a higher priority feature(s) is the output, which is selected by the product manager. As these decisions are provided after a process, they called indirect decisions.

Decision Combiner
Input: two/more direct/indirect product planning decisions.
Objective: This component initially takes a set of instance decisions as the input and identifies them as a combined decision. For examples, "Should the upload size be increased to 10MB?" is a simple instance decision and "Which action amongst, Create English version for UI and Create Chrome support for UI, has more priority? " is a compound decision which includes two simple instance decisions. Another example that addresses a compound decision is "Should the Wiki feature be enhanced or removed or kept it unchanged?" that includes two simple decisions of "Should the Wiki feature be enhanced" and "Should the Wiki feature be removed?".
Output: A combined decision is the output.

Decision Analytics Evaluator
Input: A product planning decision is the input of the component.
Objective: This component concerns the evaluation process of analytics for decisions with the goal of observing changes of other measurements (more than those have been observed in the "Feature-Analytics Evaluator" component) and measure their positive or negative impact on a decision.

This impact is discussed based on the current data or predicted data. For each instance decision, the positive value of a measurement implies that it satisfies the decision. A compound decision is broken into simple instance decisions and for each, the impacts of the measurements are being analyzed.
Output: a list of measurements and their impacts on the decision would be the output.

Aggregator
Input: The measurements evaluation is the input for this component from "Decision Analytics Evaluator" component. A predefined ranked list of measurements is another input from "Planning Repository" component to weight the measurements for the planning decision. This weighted list is learnt from a crowd of product managers (through the study discussed in section 4.1.2.2.2) or stored by the responsible product manager in

his previous experiences. It is also possible for a product manager to generate the ranked list when the proposed ranks are not satisfactory enough.
Objective: The goal of this component is to aggregate the impacts of measurements for each instance decision, based on the overall measurement weights to propose a recommendation to take the decision or not.
Output: This component, which is also the method's output, generates a recommendation for the product manager about the product planning decision.

Decision Evaluator
Input: The recommendation about taking the product planning decision is the input for the component.
Objective: This component examines the recommendation in compare with other criteria such as resources, competitors, and market to evaluate and identify the final answer. Also the product manager can define alternatives to the solution. As there are many different theories and techniques for decision-making, this component has been excluded from the current study.
Output: The final decision is the output, which is the result of decision-making process.

Feedback generator
Input: The decision and the list of observed measurements are the inputs.
Objective: The aim of the component is to receive product manager feedback about the measurement weights. After passing an enough period of time since decision-making (decided by the product manager), the product manager will compare the current states of measurements with their states in which the decision was taken, then will be able to rank the measurements based on the feedback from previous decision-making process.
Output: An update version of measurement weights would be the output. The result will be stored in "Planning Repository" component for future references.

Planning Repository
Input: The new list of weighted measurements is the input.
Objective: This repository save and restore weighted measurements which initially has been calculated from the previous study (section 4.1.2.2.2), but can be updated using feedback from ongoing such decision-making experiences.
Output: The retrieved list of weighted measurements from the repository and send it as an input to the "Analytic Aggregator" and "Feedback generator" components..

Demonstration and evaluation

The application of the analytics-based method was presented and then evaluated using instances of problems defined during a case study research. To achieve a flexible design, the demonstration and evaluation followed generate/test cycles [85], where it defined alternatives for generating the method and testing it against requirements and/or constraints.

Communication and Contribution

The current design science research provided a rigorous method in both construction and evaluation of the design artifact. Iteration was central for the design. The result of the conducted case study will show the effectiveness of the designed artifact and its contribution in literature.

3.4.2.1 Case Study

Designing the case study:

The objective of conducting the case study is to demonstrate and evaluate the proposed method while using the design science approach as the framework. The study was an exploratory evaluation study carried out for a product that has been developed at the Zurich University of the Arts available for its students, faculty members and staffs. The product is a SaaS-based application which provides a collaborative platform for media archives used for content sharing and managing media such as text, sound, pictures and movies to be uploaded and archived [86].

The design of the case study is defined by components that indicate what data are collected and what is to be done after data collection [87]. The study **proposition** indicates that the analytics-based method is effective for making different decisions in product planning. It helps to answer the "How" research question related to evaluation of the proposed analytics-based method. The **unit of analysis** in this study is product-planning decisions that are made by a product manager or professional involves in product planning of a SaaS-based software. The **theory** that is developed in this study says:

"SaaS-based measurements can be effectively used for making different decisions in product planning by transforming them to applicable recommendations in the software plan."

This theory will be tested through the designed analytics-based method [Figure 3: Overall view of analytics-based method to support product planning decisionsFigure 3] using the case study research. Holistic single-case is considered as the design type of the case study.

There is a logic that **links measurements to the proposition**. Related measurement data that are collected during the case study provide a positive or negative recommendation for taking the right decision. The product manager of the case and another professional product manager who was not familiar with the case, evaluates the effectiveness of this link through a stream of questions encompassing the following criteria:

- Measurement selection
- Measurement interpretation
- Analysis of the feature selection
- Comparison process for measurements
- Method's output
- Prioritization of alternative decisions
- Trade-off between decisions
- Feedback from implemented decisions
- Uncertainties handling
- Effectiveness of the method.
- Applicability of the method in organizations

Evaluating these criteria provides interpretation for the case study regarding effectiveness of the analytics-based method.

Conducting the case study- Preparing for data collection:

The first preparation step for the case study was collecting all SaaS-based measurements from product use for particular time period. While the product of the case was running, all the measurements were gathered by two web-based analytical tools: "Piwik" and "New Relic". The tools were installed and configured within the organization to gather measurement data automatically by monitoring the product.

Piwik is a GPL licensed web analytics software package that provides detailed reports on an online product [88] about visitors, pages visits, the way of accessing product and so on. It is a PHP MySQL software that stores data in a rational database.

New Relic tool provides performance analytics for SaaS-based applications in Ruby, Java, PHP, .Net and Python [89]. It tracks customer experiences from a click until a page is loaded and measures data with performance management's perspective. Data are stored in a log file and data can be presented in an interactive user interface.

To increase flexibility of extracting several various data, the collected data were converted to .CSV (comma separated version) format files that store tabular data. A parser was implemented to generate CSV file from the New Relic log file [Appendix C.4] and Piwik data were converted directly by MySQL Workbench tool.

Another preparation was related to study the available features of the product, which could be extracted from the provided feature tree [15] . Also other possible features to be created in the product were listed by comparing with common features of similar media archive software applications.

Before conducting the case study, previous findings of related study area were presented to the product manager of the case who finally evaluated the results of the case study. The presentation helps to increase the specific domain knowledge of the subject about terminologies, basic concepts and issues. Final preparation provided interview questions to be asked from product managers, through semi-structured interviews about the effectiveness of the proposed method. The interview questions are presented in Appendix C.2.

Conducting the case study-Collecting the data:

The evidences needed for the case study were gathered from source of archival records, direct observation, documentation, and interview. Archival records of the product use were served in terms of related measurements for a specific period of time. Direct observation means monitoring the use of product to find out trends or deviations. It provided additional information about the product use that was not available in archival records with highly quantitative data. Feature tree document was the other source being used during the case study that showed the features of the product and provided guidance about when to implement unavailable feature [15]. Final conclusion about the effectiveness of the method was identified by the aid of interviews with two product managers. One is the product manager of the case and the other product manager is not familiar with the case.

In order to organize the evidences, collections of data were maintained in a tabular format to facilitate tracking of the case. The procedure of data collection involved the following steps:

1. Form the link between planning decisions and features by studying the list of available features in the product (documentation source). The main strategy was to study and cover all decisions specified for roadmapping (section 2.1).
2. For a feature from the above list, interpret the link between the selected feature and related measurements. The criteria that are presented as the qualitative results of interview-based survey in section 0, might be considered to identify relevancy of the measurements.
3. Observe the measurement-attributes for specific time point(s) or time duration (archival records and direct observation sources). The time points may be considered regarding specific events. Then the measurements are transformed to recommendations for the decision.
4. Collect product managers points of view about the effectiveness of the procedure for making the decision through two semi-structured interviews (interview source) in order to provide final answer for the research question. While all recommendations for the product manager of the case were specified, two interviews with the product

of the case and another professional product manager who was not familiar with the case were conducted.

Conducting the case study-Analyzing the data:

Analysis includes examining, categorizing, tabulating and testing the data [87]. This study follows the strategy of theoretical proposition that forms the data collection and guides the case study analysis.

Analysis of the case study was conducted in two phases, which applied pattern matching as the analysis approach. In the first phase, a product planning decision was analyzed to investigate whether related measurements to a feature can support decision making in product planning. In the second phase, the theory of the case study was analyzed to answer the research question and link the data to the theoretical proposition of the study. It concerns the effectiveness of the analytics-based method using criteria discussed in the section related to designing the case study [section 3.4.2.1].

In the first analysis, measurements were variables, which produced a pattern to support a feature for making decisions. If related measurement-attributes were not accessible or collectable, or a decision was not supported by measurements the theory should discussed the circumstances. Those patterns confirmed that the related theory provided recommendation for decision-making and then permitted the related decision to be analyzed in the second phase. In the process of transforming the measurements to recommendation, other variables such as planning decisions, measuring function, impact function, and measurement weights were involved. The analysis also studied possible alternatives for these variables.

The decisions discussed in the first analysis, made the proposition for the second phase of analysis. The proposition was defined as the effectiveness of making the decision through the proposed method. A pattern of variables that specifies the criteria discussed in section 3.4.2.1, was formed to investigate the proposition. These criteria were set during the interview with the product manager of the case. For each proposition, if the interview's outcome didn't show the patterns as predicted, then the proposition had to be challenged.

4. INTERVIEW-BASED SURVEY

4.1 Analysis and Results

In this section, RQ1 research question will be answered. Demographic results are presented in section 5.1.1 and then quantitative and qualitative results are outlined in 5.1.2 and 5.1.3 sections. Section 5.1.2.1 will address answering RQ1.1 for existence of any relation between analytics and different type of product planning decisions. Then RQ1.2 will be answered in section 5.1.2.2 to show which measurement attributes has more value for a given product planning decision. The strength and weakness of analytics (the answer for RQ1.3) will be presented in section 5.1.3.1.

4.1.1 Demographic Results

Demographic results illustrates the distribution of answers among interviewees from different perspectives of product, people, and organization [83]. The interviews' results achieved 82.4 % from product managers, 11.8% from chief technology officers (CTO), and 5.9 % from chief executive officers (CEO) with average of 7.5, 9 and 12 years of experience respectively. These results reflect the professional's ideas in product planning with different experience. Knowledge of product managing was an important requirement for the interview. So if the interviewees had product managing experience, product manager role was assigned to them, otherwise other higher roles were considered. Table 4 presents the distribution of interviewees' roles and experiences.

Table 4: Interviewees roles and experiences

Role	Experience range (years)	Experience average	Population
Product manager	2<<15	7.57	82.4
CEO	11<<12	12	5.9
CTO	8<<10	9	11.8

Through the interview, it was asked to consider a product, which they have planned and are most satisfied with and then answer the questions. The responses were categorized by product type based on the taxonomy presented by Forward and Lethbridge [90]. Figure 4 illustrates the distribution of product types. Most of the products were assigned to Consumer-Oriented-Software (41.2%) and fewer products fell into Design-and-Engineering category (5.9%). This categorization was considered for further analysis of product planning responses. Figure 5 shows that 41.2% of the products were new products and 58.8% of the responses were evolutionary products. The distribution of interviews among different product type magnifies the difference of product characteristics on interview results.

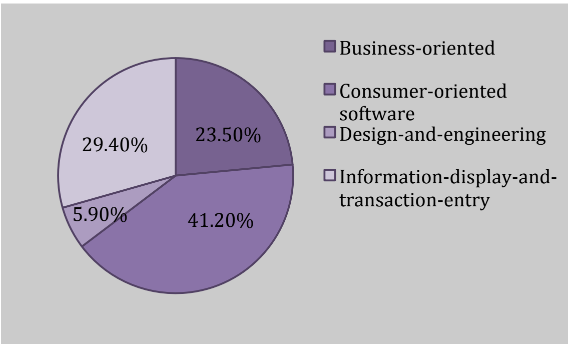


Figure 4: Product type distributions

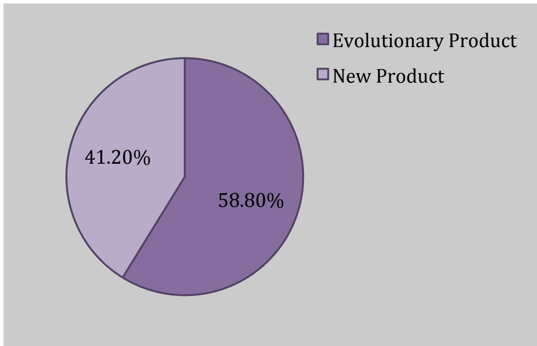


Figure 5: New-evolutionary distributions

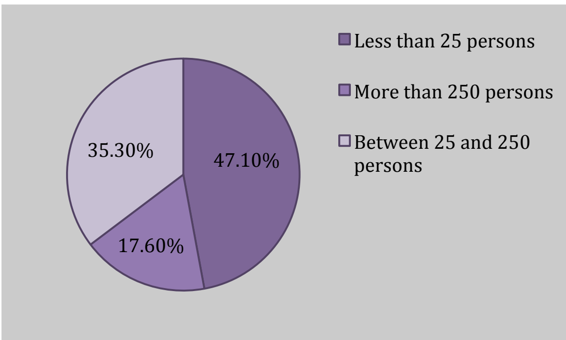


Figure 6: Organization size distributions

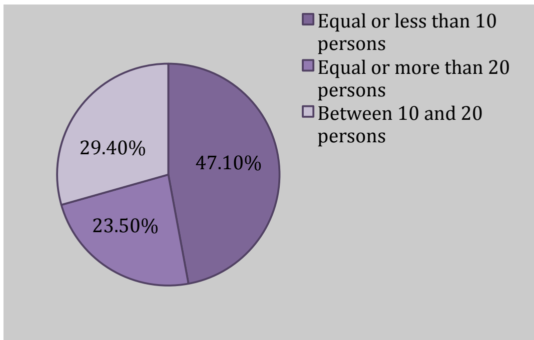


Figure 7: Development team size distributions

The interview responses were also analyzed from organizational perspective. Organizational culture differed between small, medium, and large companies, for example in flexibility, competitiveness, choices, and economy [91]. These differences affected empirical results. Figure 6 illustrates the organization size of interviewees. Three categories were considered for organization size: equal or less than 25, between 25 and 250, and equal and more than 250 persons. The size of the development team was also interesting to understand the scope of planning decisions, which are presented in Figure 7. Most of the interviewees belonged to small companies (47.10%) with the development team of less than 10 persons (47.10%). Conducting the interviews among different size of organization reflected the professional's perspectives in different organization scale and increased the reliability and generalization of results.

4.1.2 Quantitative Analysis and Results

During the interviews, all the interviewees selected roadmapping with features as the type of decision making they were involved in most. So this section will study the effect of analytics on roadmapping decisions. Each interviewee selected importance of measurement categories for two product planning decisions through the 4-step Likert scale. Then for the first decision, the importance of measurement-attributes in each measurement-category was sorted. The collected data contribute to answer whether the interviewees selected different measurements for different product planning decisions and which measurements are more useful. They will be identified in the following sub sections.

4.1.2.1 Relation between Measurement-category and Decision type

This section will study whether a product manager uses different analytics for different type of product planning decisions. Appendix B.2.3 shows the contingency tables including the frequency distribution of the interviewees' answers for importance of measurement categories

Understanding the relation between product planning decisions and measurement-categories was fulfilled by categorical analysis. For analyzing categorical, methods such as chi-squared, fisher's exact and Anova tests are more popular [92]. However, involving more than 2 groups of samples directed to concentrate on Anova and Kruskal-Wallis tests. For more than two groups of samples, they are commonly used techniques to compare groups of measurement data to study if they are from the same distribution in each group. These tests are used when there is one nominal variable and one measurement variable. Anova assumes a normal distribution, while Kruskal-Wallis is a non-parametric method that doesn't have normality assumption for data distribution.

In the first step, data for measurement-category were grouped based on decision types and then in another test, grouped by product types. Normality was checked inside each group using Kolmogorov-Smirnov test. For all groups in the test, results were significant (Lower than or equal to 0.05) which means the data was not normally distributed [Appendix B.2.2].

Since the distributions were not shaped normally for all the data of groups, Kruskal-Wallis test [93] was selected to detect the difference between three or more independent groups of samples [94]. Kruskal-Wallis test is involved substituting ranks for measurement and then calculating H statistic test. Sample size of less than 5 in each group, may be directed to unreliable results. For each Kruskal-Wallis test, the null hypothesis indicates that all distribution functions in groups are equal, while the alternate hypothesis defines that at least one of the populations tends to have larger values than at least one of the other populations. The Kruskal-Wallis result of less than 0.05 indicates the rejection of the null hypothesis.

The following hypothesis was proposed and tested quantitatively:

H0: There is no difference between each measurement-category value for different product planning decisions.

H0 will be accepted when all sub-hypothesis for each measurement-category are accepted. All sub-hypotheses are presented in Appendix B.2.1.

The H0 hypothesis was tested using Kruskal-Wallis test. The results are presented in Table 5. This table illustrates the Chi-square value (Kruskal-Wallis H), the degrees of freedom, and the significance level. “Asymp. Sig.” which is the abbreviation for asymptotic significance, is equivalent to P value.

Table 5: Krushkal-Wallis test for measurement-categories grouped by product planning decisions

	Product	Feature	Usage pattern	Referral sources	Technologies and channels	Product healthiness
Chi-Square	1.047	2.487	3.355	7.347	4.871	3.109
Df	5	5	5	5	5	5
Asymp. Sig.	.959	.778	.645	.196	.432	.683

Since “Asymp. Sig.” for all hypotheses was more than 0.05, there was no significance to reject all sub-hypotheses and thus the $H0$ was not rejected as well. It means that distributions functions of measurement-categories are not different for planning-decisions. In another meaning, the product managers didn’t use different analytics for different type of product planning decisions that is the answer of RQ1.1.

4.1.2.2 Measurement Support for Product Planning

This section is aimed to specify measurements that a product manager would prefer for product planning. To achieve that, firstly values of measurement categories for product planning decisions were specified and then each measurement attribute was scored inside each measurement-category and after that scored amongst all attributes by considering the value of the corresponding category.

4.1.2.2.1 Scores of Measurement-Categories

As there were not significant differences between categories of measurements for product planning decisions, therefore a descriptive analysis conducted for measurement-categories regardless of the decision type. The Figure 8 to Figure 13 present the distribution of answers for the importance level of measurement-categories. The x-axis has the values of 0 to 4, which shows the importance level (The values of 0 to 4 implicate “no idea”, “no important”, “less important”, “important”, and “very important” respectively which are points allocated by respondents) and y-axis specifies the percentage of the distribution.

According to descriptive analysis of measurement-categories, the “Product”, “Feature”, and “Product healthiness” categories are “very important” in the product planning decisions while “Referral sources” has “no importance” value.

61.8% of the responses have been selected as “very important” for “Product” [Figure 8], 58.8% for “Feature” [Figure 9] and 64.7% for “Product healthiness”[Figure 13] categories. For “Referral sources” category [Figure 11], 61.8% of responses have been chosen as “not important”. Categories of “Technologies and channels” [Figure 12] and “Usage pattern” [Figure 10] have been rated majorly “important” regarding 47.1% and 32.4% of the corresponding responses.

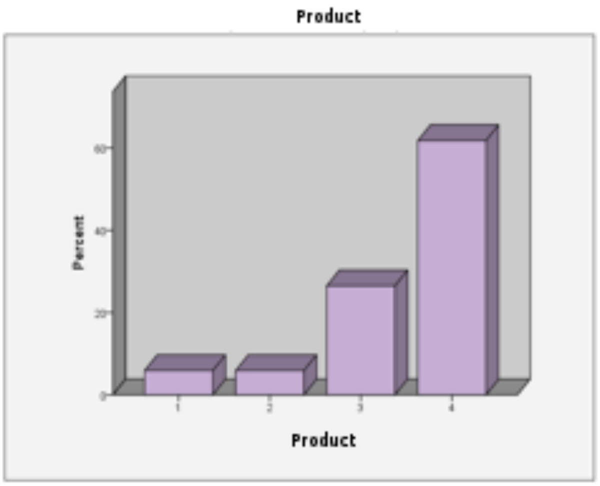


Figure 8: Distribution of "Product" category rates

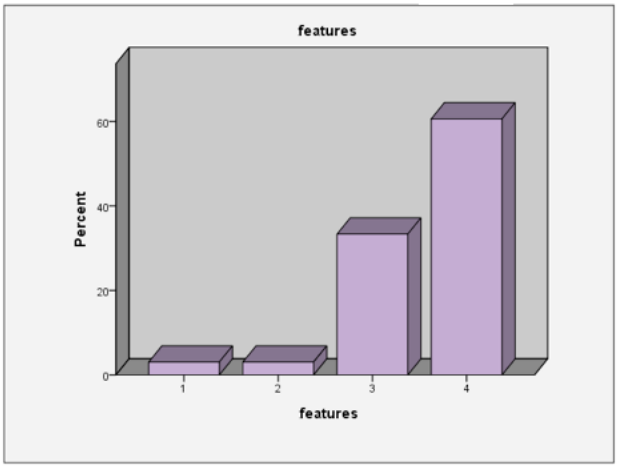


Figure 9: Distribution of "Feature" category rates

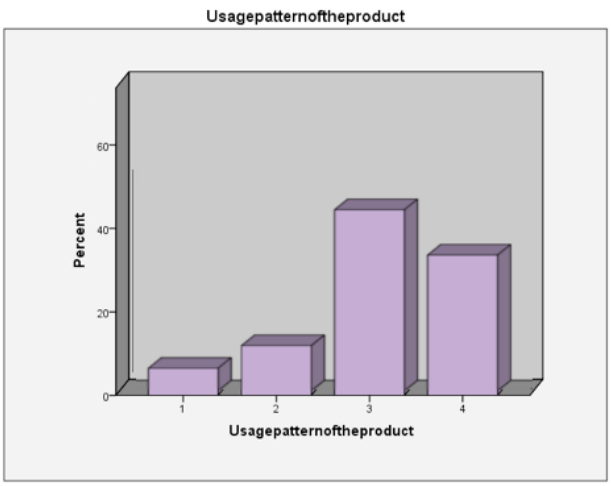


Figure 10: Distribution of "Usage pattern" category rates

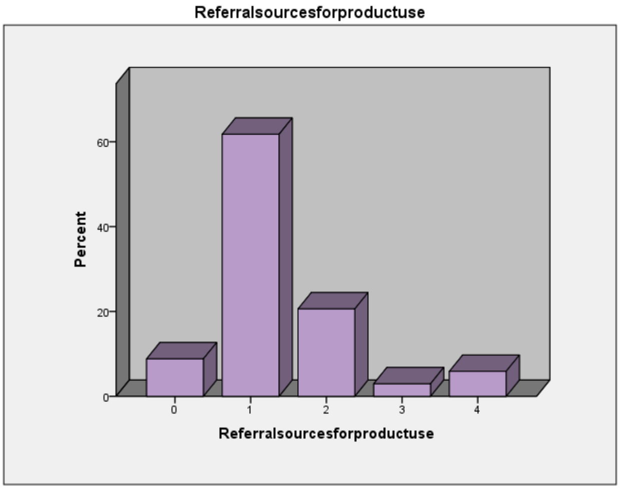


Figure 11: Distribution of "Referral source" category rates

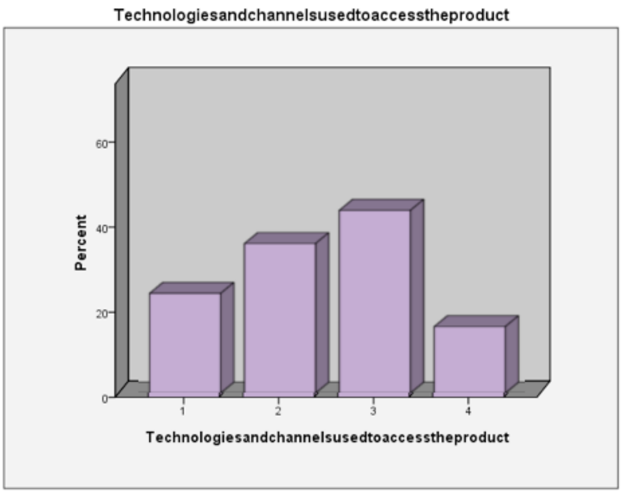


Figure 12: Distribution of "Technologies and channels" category rates

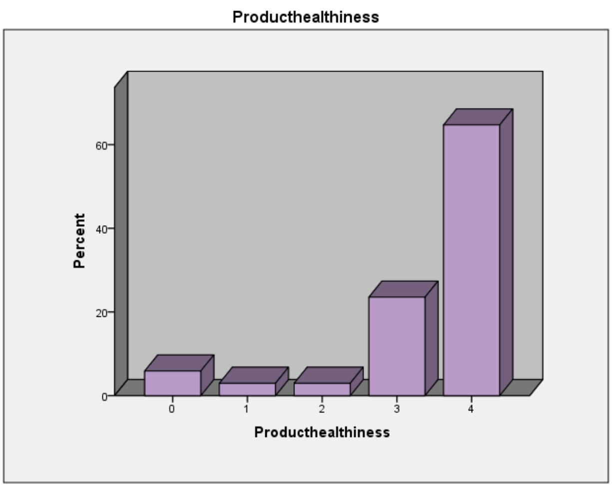


Figure 13: Distribution of "Product healthiness" category rates

Statistics in Table 6 present the numbers of valid and missing samples, mean, variance, confidence interval of mean, maximum and minimum values for each measurement-category. The mean statistics is calculated by the sum of multiplying the importance level (0 to 4) and frequencies of instances in that level divided by number of valid instances. The mean that is also called score of measurement-category is interpreted as the importance level of the category for product planning. The missing samples refer to those samples that product managers were not able to judge about the importance level of the category. Estimated confidence interval for the mean has been calculated by T-test with the state of 95% confidence level where the data was normally distributed.

Table 6: Statistics for measurement-categories

		Product	Feature	Usage pattern	Referral sources	Technologies and channels	Product healthiness
N	Valid	34	33	34	31	30	32
	Missing	0	1	0	3	4	2
Mean		3.44	3.52	3.12	1.48	2.43	3.59
Variance		.739	.508	.713	.725	.944	.507
Confidence interval		±0.30	±0.25	±0.29	±0.31	±0.36	±0.26
Minimum		1	1	1	1	1	1
Maximum		4	4	4	4	4	4

The results show that although the mean of the answers for categories has different range, all categories have been ranked "very important" and "not important" from at least one respondent's perspective.

4.1.2.2.2 Scores of Measurement Attributes

In the interview, measurement-attributes were ranked inside the related categories. Figure 14 is a sample, which depicts the percentage frequency of the rates for "product use" measurement-attribute that indicates, 66.7% of respondents have answered it as the highest rank. The figures in the Appendix B.2.3 presents the bar charts related to different measurement-categories.

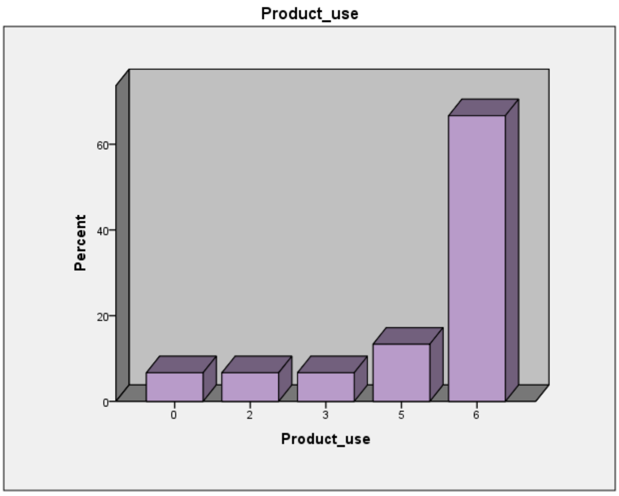


Figure 14: Percentage frequency of rates for "product use" measurement-attribute

To compare the importance level of measurement-attributes inside the corresponding category, the importance means of measurement-attributes for each category were calculated separately. Table 7 presents the scores of measurement-attributes in the "Product" category. The "Mean score" column calculates the mean value of a measurement-attribute inside category. The mean is calculated by the sum of multiplying the ranked level (i.e. It is between 1 and 6 as 6 items were available in the product category) and frequencies of instances in that level divided by number of valid instances. In the table, The "Score of the related category" column indicates the score of corresponding category among other measurement-category. Scaling the values of "Mean score" and "Score of the related category" down to fit the range between 0 and 1 and then multiply them, resulted the values in "Score of the attribute" column. These values show the importance level of measurement attributes amongst all. For all the scores, confidence intervals (with confidence level of 95%) have been also calculated using T-test with the normally distributed data (e.g. the mean value of 5.00±0.65 shows the confidence interval of 0.65 for the mean value of 5.00).

Table 7: Scores of measurement-attributes for "Product" category

Measurement-Category	Measurement-attributes	Mean score	Score of the related Category	Score of the attribute
Product	Product use	5.00±0.68	3.44±0.30	0.72±0.16
	Overall amount of users	3.33±0.82	3.44±0.30	0.48±0.16
	Time between visits	1.60±0.65	3.44±0.30	0.23±0.11
	Duration of using the product	3.13±0.73	3.44±0.30	0.45±0.14
	New users	1.60±0.57	3.44±0.30	0.23±0.10
	Returning users	2.80±0.81	3.44±0.30	0.40±0.15

To compare measurements across the categories, "Score of the attribute" column is considered as a weighting factor. Appendix B.2.5 illustrates more details for other measurements categorized by measurement-category. The score for each measurement-attribute is presented in Figure 15 by a descending order, which means that the measurement-attributes at the top levels of the figure are weighted as the important ones. By applying the confidence interval a range of possible measurement scores were also recognizable for each attribute in Figure 15, in which minimum and maximum scores are presented as the top and bottom bars of the main measurement.

Due to limitation of the interviews' session times, the study had to be abstained from collecting measurements, which their corresponding categories were answered with "no-important" or "less-important". The analysis of the missed values shows that 88.2% of answers for the measurement-attributes in "Referral sources" category are missed which means the corresponding categories were not labeled with "important" or "very important" [Appendix B.2.5].

The measurement attributes were categorized into 4 groups in different importance levels for product planning. The categorization was performed using independent samples T-test. The categories that are labeled from group1 to group4 were statistically different from each other. The groups can be seen in Figure 15. From the organized groups, measurement-attributes that belong to the first group are more preferable for product planning than attributes of the other groups. Measurement attributes in the second and third groups are less preferable respectively. For the attributes of the group 4, product managers would prefer to use them with low priority or at least would not able to judge about their importance.

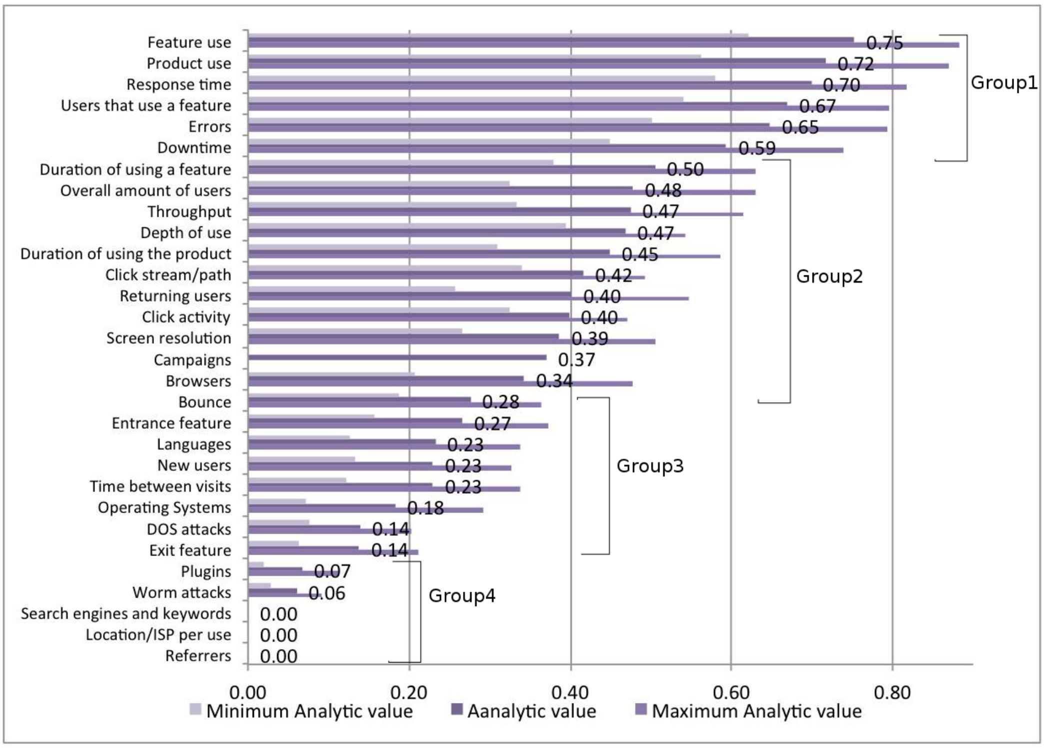


Figure 15: Measurement Scores

The groups were specified using a couple of independent sample T-tests defined through the following procedure:

For conducting the tests, dependent variables are the scores of attributes sorted in a descending order and independent variable is labels of two groups (e.g. group1 and group2) that will be formed during each test. From the top of the measurement list, samples (initially two samples) were selected in order and organized in the one group and next sample of the list is assigned to the second group. After checking normality of each group, T-test is executed by the hypothesis of "There is no difference between the two population". P>0.05 confirmed the sample of the second group belongs to the first group and then for the next test it will be assigned to the first group. P<0.05 indicates the second group (with only one member) significantly differs from the first group which means a new group with only one member should be formed. Then the next test is similarly performed with the new group. The procedure is continued until all measurement attributes are assigned to one group. The Appendix B.2.4 in Table 36 illustrates the p-value of each test and specifies when a new group is formed.

The study fulfilled the required assumptions for conducting T-test: For each test, normality of dependent variable (i.e. score of attribute) was confirmed and no significant outlier was found among them. A categorical variable was independent because a measurement-attribute in one group was not member of the other group. Homogeneity of variances was another requirement that was checked by Levene's test.

The section concludes that product use, feature use, users of feature use, response time, product errors, and downtime are the first top measurement-attributes that a product manager prefer to use for product planning RQ1.2.

4.1.3 Qualitative Analysis and Results

The goal of this section is to specify the advantages/disadvantages of using analytics for product mangers and clarify reasons behind their selections for analytics importance (address RQ1.3). So during the analysis strengths and weaknesses of using analytics for product planning as well as influential factors on analytics importance will be specified during a qualitative analysis. For analyzing qualitative data resulted from interviews, content analysis has been selected. Through the interviews, interviewees were questioned about their arguments for particular answers. These arguments were analyzed by the means of the content analysis method. Another alternative such as grounded theory is also widely used for analyzing qualitative data but had some limitations for the current interviews. Grounded theory requires simultaneous collection and analysis of data resulted from the first interview and then using them as an input for further interviews [95]. Within the grounded theory, one must start the data analysis with hypothetical theory and then try to prove it [84][95].

Narrative analysis was another alternative for analyzing the interviews results [96]. Narrative analysis obtains information, which is not usually applicable by other methods and was not important for the current study. As the examples, the following information can be mentioned: in-depth understanding of interviewees' subjective experiences, modes of thought, emotional characteristic, and cultural characteristic [97].

Hsieh and Shannon[84] presented three content analysis approaches based on coding differences: conventional, direct, and summative content analysis. According to their results, conventional content analysis was well suited for analyzing the interviews results, because there were few theories or literature about the interview's phenomenon, which could not rely on. Therefore, using perceived categories were avoided; instead the categories were formed from the collected data.

Satu Elo and Helvi Kyngas presented content analysis in inductive and deductive ways [98]. Again, because there was limit knowledge on the phenomenon, the inductive type was considered for the current interviews. Inductive content analysis was applied for the interviews in the three phases of preparation, organizing, and reporting, which are illustrated in Figure 16.

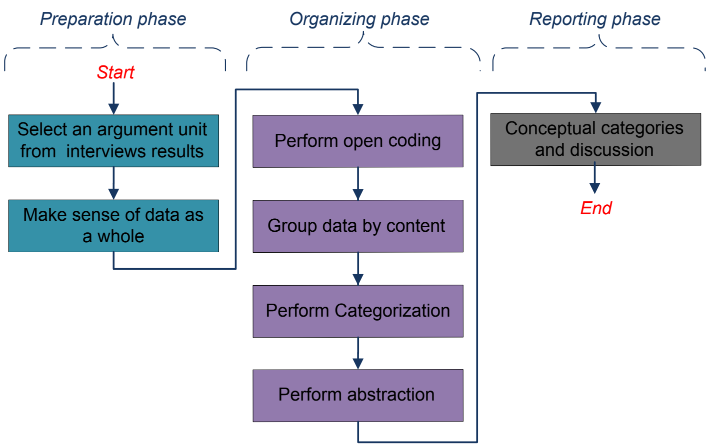


Figure 16: Content analysis phases

First of all, one unit of argument was selected, then tried to read through the content and write down headings that describe all aspects of the arguments. This processed were repeated for all arguments one by one. Later on, all headings were gathered and categories were generated. Then categories were grouped under higher headings to reduce the number of collapsing and similar categories. Categorization provided a mean of interpreting the phenomenon, increasing understandability, and facilitating decision making ability [98]. At the end of the content analysis, abstraction was performed which led to general descriptions and further discussions based on the categories. Figure 17 presents parts of content analysis for the interviews.

Through the interview, interviewees were questioned about their argumentation of particular answer. The questions of "How can the 'X' measurements affect the 'Y' decision?" and "What are the basis for your selection?" were orally asked from the interviewees. By repeating the content analysis process for all the interviewees' arguments, three influence factors were recognized that affect the importance of measurements: The "product specification" (e.g. customer type, access type, network type, users' numbers), "maturity" and "Product goal" factors. These factors, related discussions in addition to some related quotes from interviewees are shown in Table 8. Also the strengths and weaknesses of using analytics were another results extracted from the qualitative analysis that will be discussed in section 0.

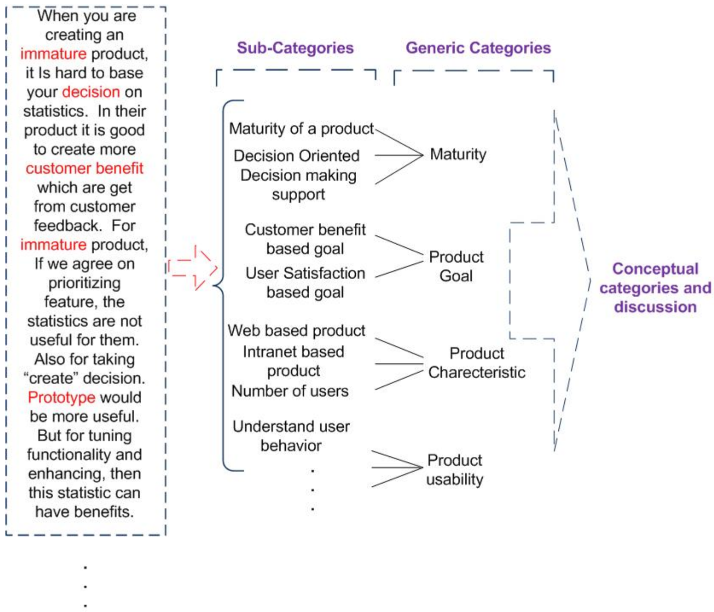


Figure 17: An example of content analysis process

4.1.3.1 Strengths and Limitations of using analytics for product planning decisions

According to the interviewees, analytics increases a product manager's knowledge about Usability. It is important to understand the product usability, especially in understanding the user behavior, ease of use, and identifying popular features. Also analytics improve learnability of different usage patterns, which can improve decision-making quality in general. Analytics improves problem handling as well. Users won't accept faulty product and analytics helps to identify problems as soon as possible and devise a replacement in the time of failure. Improve handling the client-side technologies are done by the analysis as well.

The limitations of using analytics address the limitation in receiving customer feedback and end-user feedback. The feedback can be received directly or through survey forms, help to take better decisions, which cannot be replaced by analytics. Another limitation refers to remaining some measurements in theory, which is difficult to be applied in a real product. Measurements such as "Depth of use" and "Click path" are interesting to know, but it might be difficult for some companies to collect and interpret them in a real product. The next limitation indicates that for an immature product, analytics is not so much helpful. Within the first release of a product, when a product is hardly mature, data collected through measurements cannot help decision making so much. But after a release or providing a prototype those analytics can be important. Analytics is not important for the development technology aspect. Technology has two aspects in planning a SaaS-based product. One aspect is related to the technology that should be considered for product development, and the other one is the technology that is required in client side for running a product. For finding development technologies analytics is not important but after the release of the technology, product popularity is monitored to find the effect of implementing the technology on planning.

Table 9 gives an overview of the product managers' detailed reflections about strengths and limitations of analytics for product planning. For each, the related cause-effect from some important interviewees' quotes and the corresponding analysis codes are presented.

Table 8: Influence factors on importance of measurements for planning

Influence factors	Discussion	Quotes from interviewees
Product specification	By considering specification of product, impact of analytics on decisions and also importance levels of analytics are different. Product specification differs regarding the number of users, customer type, access type (web based, client side installation), network type (Internet, intranet), level of quality, and delivered services.	“Referral source is less important, as we knew all the traffic; there is no difference for us.” “Technology and channel data is less important, we have to support all browsers and cover related technology as it is a web based product.” “New user is not important because we are dealing with available users not new users.” Referral source is not importance since we sell product to organization not end users so they do not care where the customers are coming from.” “Referral source is not important, because users are all over the world as they use their mobile phone.” “Product healthiness is very important. If we can not achieve desire reliability and performance we can go home.” Technology and channel is very important because the product is a web-based tool.” “Technology and channel is important as we try to support more browsers in their product.”
Maturity	Maturity of a product has considerable dependency on the importance of analytics. The impacts of analytics on decisions are different based on the maturity level of product. Analytics can impact all decisions in a mature product, however in an immature product, for some decisions such as "Create a new feature" analytics has no use. For decisions such as prioritization, first the maturity level of product should be considered. Within the first release of a product, when a product is hardly mature, analytics cannot help decision making so much. Afterward the related measurements can be important which fulfills a goal of particular release.	“When you are creating an immature product, it is hard to base your decision based on these kinds of statistics. Instead of analytics for creating decision and for immature product, we create prototype, test prototype. But for tuning functionality and enhancing, then this statistic can have benefits.”
Product goal	Analytics importance is different based on a product goal (i.e. customer-centric, user satisfaction-centric, focus group-centric, or market-centric)	“Product value is very important. Having value for users is the reason of creating the product. We are building the product to give value to the users.” “The goal is to increase web users, if product use is not too many then action should be taken to find the reason or update features.” “Product value is very important. Modules that people don’t use, there is no use to hear that. Good to know what they use.” “Product value is very important, all that reflect the users idea is valuable.” “In our product it is good to create more customer benefits “

Table 9: Strengths and Limitations of using analytics

Strength					
Strength	**Cause**	**Effect**	**Quotes From interviewees**	**Related Codes**	
Increase knowledge about product usability and functionality	Applying measurements in categories of "product", "feature", "usage pattern"	Improve understanding the user behavior, ease of use, and identifying popular features and tuning functionalities. Increase learnability of different usage patterns, which can improve decision-making quality in general.	Usage pattern is very important because "we want to know how product is being used. We don't have this information now but we need to learn it to have any evolution of the product." "Product use is very important to monitor the popularity level of product during time period." Product measurements are very important because "if you don't know how much it use you don't know it has worth or not"	Understanding product use	Product usability
			By Usage pattern measurements "you can measure to understand how user interact with the service." Usage pattern is very important as "it is good to have history of users' activities"	Understanding user's behavior	
			Feature value is very important because "I wants to know what are important features" Value of features is very important because "for example statistics about feature use gives the overall usage about features. Bounce rate is important because shows how many users of community cannot find the features. Bounce (frequency) or time duration both can provide sufficient picture of finding features." Feature value " is important to know which feature is used. If no one use a specific feature, then you can remove it" Usage pattern is important as "This provides further and assisting information for value of feature from user perspective."	Feature popularity	
			Usage patterns It could be nice to see different usage pattern. The product many be used 200 times, it would be interesting to see what could be learn from pattern."	Learnability	
			Usage pattern is "Important as they are trying a lot of efforts to make the product as easy use as possible. And they can understand the problem that make the product hard to use"	Understanding ease of use	
			Product value is important "for tuning functionality and enhancing"	Understanding functionality	
Improve problem handling	Applying "Product healthiness" measurements	Identifying problem as soon as possible facilitate devising a replacement in the time of failure.	Product healthiness is important as "These statistics have lot of benefit for decisions. The errors can be seen very quickly and repaired in each month release." Product healthiness is important "since we see the product has problems, we need to work with"	Identifying problems	Problem handling
			Product healthiness is very important as "If the product has failure we will understand quickly and will replace with new product"	Product replacement	
Improve client- side technologies	Using measurements of Product, Feature, Product	Improve client-side technology, facilitate technology adaptability and identify popular technologies	Product healthiness is important" as comparing performance and reliability of the used technology with previous one is interesting for decision. Also it has a good relation with new technology" Product and Feature measurements are important as "There were always balancing between technology risks and the value of perspective feature (or product). So if a technology had a big risk and the feature was not very valuable then they would avoid the technology"	Supporting Technology	Technology handling – client side

	healthiness		“Technologies may have more impact on product healthiness. So if the product healthiness doesn’t have good situation the they ignore the technology”		
			Technology and channels measurements are “ important because we can adapt user interfaces”	Technology adaptability	
			“Technology and channels is a tricky category, what do you mean by technology? Technology that used for development or technology that is related to users. They are different with each other. For development part analytics is not important, although for user side that plays important role” “Confirming technology has 2 sides, one is back side which is not related to users, and analytic do not affect that, actually it is internal technology. Analytics is important depend on feature, whether is used before are not.”	Technology perspective	

Limitations

Limitations	Cause	Effect	Quotes From interviewees	Related Codes
Analytics might remain in Theory.	Applying measurements of usage pattern such as "Depth of use" and "Click path"	Might not be easy to collect them or be difficult to interpret them. So they remain in theory.	Usage pattern “is important in theory but not for our product, because in reality we do not have access to such data” Usage pattern measurements such as “Click stream is good but hard to extract meaningful data from them. They are too much data and difficult to be interpreted.” “Users use the product differently, sometimes having more depth of use doesn’t mean it is better. Some may reach his goal in 3 depths, some may need more! So interpreting the depth can be difficult.”	Theory based
Limited in Receiving customer feedback	Applying just analytics	Cannot replace customers and end-users feedback. There is a need to have direct or survey-based interaction with customer or end-user	“The number of users and duration of using are not good measures to say value of product but considering other measures (such as having interviews with them to ask them considering qualitative investigations beside quantitative results) will make it important “ “Measuring fun of the game is important which is difficult to be measured with these measurements. Having feedback from customers and also using forums that show their requirement helps product managers in their decisions.” “For a product it is good to create more customer benefit which is get from interview with customers and customer feedback by their service organizations.”	Feedback
Limited for immature product	Being an immature product	Measurements cannot be helpful unless a prototype is created.	“When you are creating a product and immature product, it is hard to base your decision based on these kinds of statistics. For an immature product, they create prototype, test prototype”	Maturity
Limited in technology handling (Development side)	Collecting the measurements from a web application	Cannot be important for development technology aspect.	“Technology is a tricky category, what do you mean by technology? Technology that used for development or technology that is related to users. They are different with each other. For development part analytics is not important, although for user side that plays important role” “Confirming technology has 2 sides, one is back side which is not related to users, and analytic do not affect that, actually it is internal technology. Analytics is important depend on feature, whether is used before are not.” Technology measurements“ is not important for enhancement although supported technology can show the possibility of enhancement” “Less important, we have to support all browsers and cover related technology as it is a web based product.”	Development technology

4.1.4 Synthesis

The purpose of this part is to summarize the major findings and implications of the quantitative and qualitative results.

The “Product healthiness” was the most important measurement-category with the 64.7% of the selections. The interviewees believed that users do not accept faulty product. They mentioned that it is not important how good the product is when the users face with several running problems, errors, and long response time. The product healthiness is a minimum level of user expectation that plays a great role in the success of a product. However two measurement-attribute of “worm attack” and “Dos attack” had considerably less importance in comparison with other measurement in this category (6% and 14% respectively), as the interviewees believed these problems have been rarely occurred.

The “Product” was the next “very important” category with 61.8% of the responses. Three factors were mentioned by interviews that can affect the importance of this category for product planning: Product goal, product usability, and product maturity. The value of “Product” measurements for decisions differed from the goals of products i.e. customer based goal, user-based goal, or market-based goal. “Product” category seemed interesting to the interviewees as they believed that its measurements could help a product manager in understanding to what extent a product is user-friendly. Maturity of a product was mentioned by several interviewees: “Product” measurements can impact all decisions in a mature product, however in an immature product, for some decisions, such as "Create a new feature", analytics has no use. In this category, “New users” had less value among other measurements (23 %), because most of the interviewees believed that new users could be treated similarly as current users and returning users.

“Feature” category was labeled as an “important” category by 58.8% of responses as it illustrates how successful the product features are. 75% of interviewees believed that “feature use” has an important value in product planning and decision-making as it illustrates the feature popularity among users and can be considered in planning new features that might have appeal for the users. “Entrance feature” and “Exit feature” had less value in this category, i.e. 27% and 14% respectively, because of the product type. Some products had one entrance and one exits, some had different exits, which were not important, from which the users exit, because it varies based on the users required service.

“Usage pattern” category was rated majorly “important” (47.1%). The reason behind this selection was mostly the access to usability information of a product. The interviewees believed that measurements in this category could help a product manager to understand user behaviors. Also these measurements increase learnability of different usage patterns, which can improve decision-making quality in general. The value distributions among measurements in this category were almost similar (42%, 47%, and 40%). Some, who did not rank these measurements as important, mentioned that these are theory-based measurements, they can be interesting but in practice they are hardly accessible.

“Technology and channel” was rated as “important” by 32.4% of responses. The interview results illustrated interviewees’ uncertainty in “technology” terminology. Two different perspectives of technology were mentioned by responses. One aspect was related to the technology that should be considered for product development, other was technology that is required in client side for running a product. The importance of this category was different based on the interviewees’ perspectives. The importance levels of measurements in this category were directly depending on the type of product. For instance, 34% of responses, which were about web-based products, believed that “Browser” measurement-attribute had higher value than other measurements in this category. Product mangers of those 18% of products, which were depended on a particular operating system, mentioned “Operating systems” as an important one. “Language” was also depended on product type and 23% of the interviewees found it interesting, but some said that it is no matter because all languages should be supported or it is an English-based product.

"Referral sources" category had less importance among other categories. 61.8% of the responses mentioned it as a "not important" category because of their product types. Most of the interviewees' product had specific customers or particular users, so they already knew from where their traffic comes. Also some products were intranet based so the referral sources of products were obviously clear.

As it was mentioned above, several factors impacted the answers of the interviews. The responses were different based on the product specification, maturity level and product goal.

4.2 Interpretation

The interviews results aim at answering RQ1, i.e. "How do the SaaS-based measurements affect the product planning decisions?". The answer provides the association between the measurements and decision-making of product planning. The results indicate that different planning decisions consider the importance of measurement-categories in a similar way. The interpretation indicates that if one set of measurements has been recognized important for planning, it can be used importantly for all decisions regardless of their types (answer to the RQ1.1). Various values are defined for different measurement-categories to take the decision. Among all categories, "Product", "Feature", and "Product healthiness" were mostly considered as the "very important" categories of measurements, while "Referral sources" was mostly labeled as a "not important" category. The measurements were scored among all, based on the importance inside the category and between the categories. The scores can be applied in order to attach value to measurements attributes when decision-making involves trade-off between the measurement values. These scores have been presented in Figure 15. Feature use, product use, users that use a feature, response time, error and downtime are the measurement attributes that product managers prefer the most to use for product planning (answer to the RQ1.2).

During the interview we found some clue that we need more consideration when using the overall scores of measurements. The justification of the interviewees for assigning a value to a measurement showed that different factors such as product characteristics, product maturity and product goal have affected on their selection. Investigating about these factors needs a further study in the future. Also the study concluded that analytics improves knowledge about product usability, functionality, problem handling and client-side technologies and has limitations regarding to receiving formed-based customer feedback, handling development technologies and also interpreting some measurements in practice (answer to the RQ1.3).

4.3 Summary and Discussion

4.3.1 Quantitative Study – Summary and Discussion

The study of the relation between planning decisions and measurement-categories (Product, Feature, Pattern Usage, Referral sources, Technology and channels, and product healthiness) showed that distribution functions of measurement-categories are not different for planning-decisions.

The results showed that measurement-categories of "Product", "Feature", and "Product healthiness" are the most effective categories for taking product planning decisions while "Referral sources" doesn't have too much effect. The study of the importance level of measurements for product planning decision has been conducted as well. The results showed that the measurements of "product use, feature use, users of feature use, response time, and product errors" are the first top measurement-attributes for product planning. The result correlates with the study that shows in a SaaS-based product, measurements related to usability, performance and productivity have been recognized as the most important metrics [13].

In traditional planning, values of features have been always important criteria. Stakeholders might specify a feature value by simply assigning a number, based on the assumed impact of the feature, on the overall product [43][41]. The values can suffice to prioritizing the features over each other [26], which means there would be no absolute valuation, but each feature is relatively situated amongst two other features. This correlates with the study results that "product use" and "feature use" measurements are the two most important measurements for product planning that might assist product managers to assign feature values by monitoring the measurements.

4.3.2 Qualitative Study – Summary and Discussion

The qualitative study clarified the reasons for selecting the importance level of measurements by the interviewees. Understanding the usability and errors of products were two important reasons for considering "Product", "Feature", and "Product healthiness" as the important measurement-categories. Measurements related to "Technologies and channels" category were considered none important when development technology aspects were being discussed, however they could facilitate client-side technologies.

Qualitative analysis of the study presented some factors that were mentioned as the affected parameters on the selection of the measurements importance. These factors could be related to the factors of product specifications (e.g. customer type, access type, network type, users' numbers), product maturity and product goals. The factors correlates with situational factors [99] discussed for the SPM area. Details of the affected factors and the way of affection should be studied as a future work. It is interpreted that the measurement-categories are mostly affected by other factors rather than decisions.

The interviewees mentioned customer feedback as an important data for their decisions. However the corresponding measurement attributes were excluded in the study, as they couldn't be supported in the analytical tools.

4.4 Validity Threats

The interview-based survey had both qualitative and quantitative perspectives. So the threats were considered from the both perspectives. The validity issues were considered in 6 categories of description, interpretation, reliability, construct validity, external validity, and conclusion validity [19][100][101].

Description: For providing valid description of what has been heard through interviews, there was an initial threat of inaccuracy or incompleteness of data. Audio recording of the interviews mitigated this threat. Also the investigator triangulation was used, in which the interviewers peer checked and reviewed the responses and took extensive notes independently. Also the online survey Gizmo tool provided the capability of inserting the interviewees' responses to the database in real time.

Interpretation: There was a threat of not having a valid interpretation from interviewees' arguments. The threat was about extracting the categories and meaning from what happened and learnt in the interviews. The categories that were achieved from qualitative results were quite subjective that means each interviewer should present how the interpretations were formed. This threat was mitigated by conducting multiple meetings among the interviewers, in which all the transcripts from the interviews were peer reviewed and commented. Furthermore, qualitative analysis method was used that provided the capability of tracing the route by which the interviewers came to some certain interpretations.

Reliability: Standard research instruments can increase the reliability of formal design research. However some parts of the interview-based survey had quite flexible design, which limited the formal reliability testing. The main reliability threat was about transcriptions errors [102]. Inaccurate punctuation and mistyped words were threats for changing the entire meaning and interpretation. Audio reviewing decreased this threat to a considerable extend.

Construct validity: Construct validity considers whether the method measured what was expected to measure or not. The quantitative data that provided by the product manager's answers were subjective, because they might interpret the questions differently. For increasing understandability, all questions were documented with “Survey gizmo” tool. Required access to online questionnaire sent to interviewees before each interview to be able to follow questions online. Also for decreasing the impact of misinterpretation, definitions of important concepts were provided in an introduction page and corresponding survey questions. An open discussion for each question allowed interviewees to ask about their uncertainties.

Conclusion validity: In order to achieve more reliable result and ensure that the online interview tool and the posed questions have high quality, several pilot interviews had been organized to avoid having low quality questions and layout. Also there was a threat that the quantity of interviewees might affect the results. Therefore the mix interview type was considered, hence all the questions had both qualitative and quantitative results and all the interviewees were asked about their motivations for answers. Furthermore all interviewees had product planning experiences and were well familiar with product planning decisions. Therefore we believed that the interviewers were experienced enough to answer the questions. The interview itself was designed to select 2 random answers by interviewees for selecting the planning decision at the beginning, and then follow the rest based on them. However the risk of missed conclusion, due to lack of enough response for each decision, was mitigated by conducting multiple extra interviews after achieving the result of the initial analysis.

External validity: The main threats with the interview were reliability and generalization. The interview was conducted amongst different size of organizations: small, medium, and large. Also different types of products were questioned about to avoid the similarity due to product type. The results of the interviews have the capacity of being generalized for different types of products and organizations. As 47.10% of the interviewees were product managers in small size organizations, it might be required to conduct more interviews in medium and large organizations to improve the quality of the results in future research.

5. METHOD CONSTRUCTION - CASE STUDY BASED EVALUATION

The analytics-based method as the artifact of the design-science study is designed and presented in the section 3.4.2. This method is demonstrated and evaluated using a case study. Demonstration is the pre-requisite of the evaluation which is presented in section 5.1. The context of the case study will be discussed in section 5.2 and the evaluation in section 5.3. The method evaluation is performed in two phases. The first phase of the study demonstrates the utility of the approach through some examples to show if measurements related to a feature can support decision making in product planning and the second phase investigates the effectiveness of the analytics-based method which supports planning decisions of a software product.

5.1 Decision Making in a Case

This section demonstrates using the analytics-based method (section 3.4.2) through the case study. The configuration of right analytical tools is the pre-requisite of applying the method in order to enable monitoring of the product and collecting the measurements of the product use. Also a resource with product management knowledge is required to perform the method.

Overall product monitoring usually works as a trigger to utilize the analytics-based method. Dissatisfaction of product success or decrease in product performance is mostly clarified by overall product monitoring. So new decisions are proceed to overcome the problem. In the current method, monitored data facilitate taking the decisions.

The following processes elaborate on the defined components of the method [Figure 3] that have been taken from identifying a problem and mapping it to a decision, to propose a recommendation for making the decision:

1. **Define an instance decision** is the first step. The effects of analytics on six roadmapping decisions (in section 2.1) will be studied in this process. These decisions are generally talk about features, but to link them to specific features, instance decisions are defined (e.g. Should Wiki feature be enhanced?). Selecting a right feature to be decided requires studying source of information about product's features (documentation source). There are also situations that consider taking a pre-defined decision directly without studying about the features. For instance the decision can be concluded earlier from the customer's feedback.

2. **Make connection between the feature and the measurements (Specify measurements),** which are informative about the feature or even the whole product (i.e. "the number of visits per month" is a sample statistic for "Product use" measurement-attribute). This is a simple sample statistic. To facilitate the process of selecting sample statistics, some patterns are recommended that will be explained in section 5.1.1.3.
   While it will be useful for a decision to find out a wider range of related measurement-attributes and observe their changes, for initial determination of features that has potential for decision-making, monitoring the most related measurements might be enough to select features to be decided. If a feature passes through this step, then there is an opportunity to look closer at its other related measurements in later steps.

3. **Observe** measurement-attributes for the time periods(s) within which, the measurement values will be compared. (For example, the last June and the current month are samples of two time periods). This time period can also be referred to future in case of forecasting data. Based on the nature of sample statistics, there can be one, two or even more comparison time period(s). The measurement values are collected using web analytical tools that run for a certain period of time. Sometimes, the process of calculating values needs interpretation and extra calculations of the available data, as the tool does not provide values directly.

4. **Analyze the importance of the feature (Feature analysis)** by studying values and changes of measurements and considering external factors (such as defined strategies of the product). Monitoring the changes of measurements for a feature, might notify the decision-maker about outliers or deviations from the plan, which suggests taking a related decision. This analysis is based on intuitive interpretation. If the changes of the measurements do not confirm taking a planning decision about the feature, then the potentiality of another feature will be studied. In this step the non-formal interpretation of the product manager is enough. If the feature is not recognized as an important one the step 1 to 4 can be repeated.

5. **Specify extended measurement-attributes** more than those have been studied in step 2. For a decision all related sample statistics of corresponding measurement-attributes are identified.

6. **Observe extended measurement-attributes** of steps 5. The sample statistics of measurement-attributes are observed for the time periods(s) within which, the measurement values will be compared.

7. **Define a comparison function (Decision analysis).** Comparing value of measurements requires a comparison function, which is defined based on a product 's goal. This function is defined by propounding a question goal based on the sample statistic. As an example, for "index-page bounce" this question can be raised as: "Is the value of index-page bounce rate for the recent 6 months more than 20%? ". Positive answer indicates that the measurements support the decision and negative one shows its negative affirmation on the decision. The output presents the supportive level of the measurement for the decision. The comparison functions used in the documents have been presented in section 5.1.1.4.

8. **Confirm measurement weights,** which were initially assigned from the previous feedback of the current product or are studied in the previous research (from the interview study explained in section 4.1.2.2.2). It is used to show the importance level of one measurement-attribute amongst the whole. If the proposed weights do not satisfy the importance for the feature in the specific product, the weight (as the level of importance) can simply be updated by scaling them with the values of 1 to 4 (corresponding values for not-important, less-important, important, and very-important similar to the previous interview study) with the same distance and resolution of 0.5. Different factors can affect on importance level of one measurement-attribute for a feature as studied in section 0 but measuring the importance level based on these factors needs to be studied as a future research. Until then, direct contribution of product managers in confirming or defining the level is required to achieve more effective results.

9. **Calculate comparison values of the measurements,** to make them comparable. It aggregates the measurement values with their weights for each instance decision (usually by multiplying the weight of the measurement by its impact value). Then this step visualizes the measurements through a bar chart with positive and negative elements that helps the trade-off between the results. The visualization provides general overview about data to conclude easily by eye tracking instead of working with numbers, which might be preferable for the method's user. In some situations of minor difference between data, comparing the numbers by a function is inevitable.

10. **Pre-evaluation of the decisions,** by using data provided from the previous step. Here the initial trade-off might be applied for some decisions but it is mainly done in "Decision Evaluator" to find a final solution by comparing the product criterion (the

output of the analytics-based method presented in step 9) with other criteria [7] (i.e. resources, market, competitors). Positiveness of data in the bar chart implies recommendation for taking the decision and negativeness disapproves the decision. If the visualization was not enough to conclude from the data and distribution of them in positive and negative side of the chart was similar which could not be caught by eye tracking, data are examined using a mean function to show the level of positiveness and negativeness. Using an impact function presented in section 5.1.1.5, such as arithmetic mean or harmonic mean assists when there are not strong justifications for the choices.

11. **Collecting feedback** and recording new ranks for measurements based on experiences of the previous decisions. After implementing the feature that had been decided, there is a possibility that the monitored data have not been altered as expected or the decision has not satisfied the problems. Alternatively data might have improved more than expectation. This may result in revising the importance of the measurements. The process of updating the ranks is done by filling a questionnaire form, which assigns a Likert scale of 1 to 4 (corresponding values for not-important, less-important, important, and very-important) by considering the distance of 0.5 between scores to increase resolution of data. The new ranks list will be stored in the repository for future references.

According to the analytics-based method in Figure 3, steps 2 and 3 are mapped to "Feature-Analytics Evaluator" and step 4 to "Feature Analyzer" components. "Decision Combiner" is supported by step 1, while a combined decision is defined. Steps 5 to 7 are part of "Decision-Analytics Evaluator" component and steps 8 to 10 are defined inside "Aggregator" component. "Feedback generator" component is also supported by step 11.

To make the process clearer, Figure 18 presents a general overview about states extracted from the above processes. The states involve in feature selection, decision analysis, recommendation and feedback of the analytics-based method, which are the main activity sets in Figure 18.

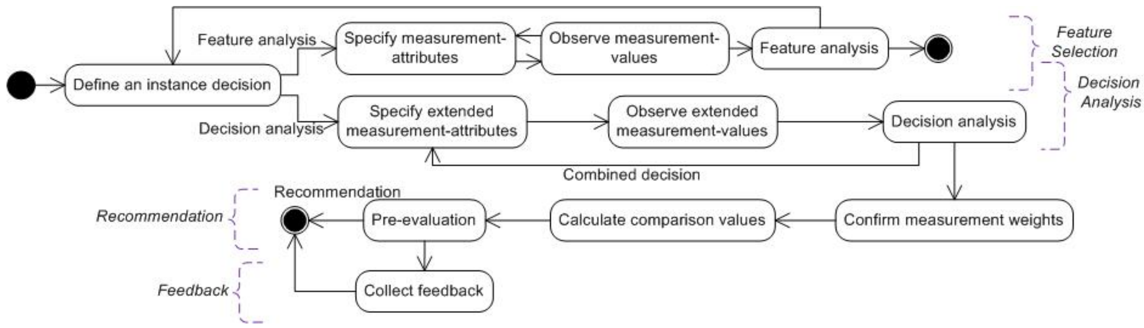


Figure 18: A state chart of analytics-based method

5.1.1 Variability of the method

During the case demonstration, decisions type, measurement type, patterns for sample statistics, comparison functions and impact functions were discussed as different variables. The variability of the method can be caused because of:

5.1.1.1 Decision Type

Three types of product planning decisions are supported by the method: simple decision, multiple decision and compound decision. A product manager can think about a single decision when she/he wants to add, remove or delete a feature. “Should the English version of UI (User interface) be created?” is an example of a simple decision. Sometimes a product manager needs to prioritize between two or more decisions for an instance decision about selecting between “Should Chrome support be created? ” and “Should English version of UI be created?” is named as a multiple decision. Compound decisions include two or more closely related simple decisions that just one decision should be suggested. For instance decision about selecting between ”Should the Wiki feature be removed?” or “Should the Wiki feature be enhanced?” is a compound decision. Decision types are considered in step 1, when an instance decision(s) should be defined. If the decision is related to a combined decision (multi or compound decision), step 5 to 7 are followed separately for each sub decision.

5.1.1.2 Measurement Type

Measurements can be considered in two forms of simple or combined depends on the selected decision. Simple measurements are those, which can be used independently such as “number of visitors”. Sometimes the sample statistic related to a feature is composed of two or even more statistics as a combined sample statistic (i.e. the bounced rate of Google Chrome users). The measurement types are considered in steps 2, 3, 5, and 6 of the method.

5.1.1.3 Pattern for Sample Statistics

To facilitate the process of selecting sample statistics, statistics patterns are used for each measurement. These patterns work as a guideline to provide a view about how the sample statistics can be looked like. Table 10 presents the patterns for only “Overall amount of users” measurement-attribute. The entire pattern is provided in Appendix C.1. The sample statistics are considered in steps 2 and 5 of the method.

Table 10: Sample of simple statistics pattern

Measurement-category	Measurement-attribute	Patterns for sample statistics
Product	Overall amount of users	Number of unique visitors per a year/ month/day
		Minimum/ maximum/ median number of unique visitors per a year/month/ day
		Average of unique visitors per a year/ month/ day
		Maximum/minimum/ median number of unique visitors per a week in one calendar month or year
		Variance of unique visitors per a day/week in one calendar month

The above sample statistics can be differently selected base on product type, product manager perspectives (optimistic, pessimistic or moderate), tool accessibility, or decision type.

5.1.1.4 Comparison Function

To compare previous value of a measurement with a new value, we used a step function to measure if the changes satisfy the decision. The conditions of this function can be changed based on product manager's consideration. The following pattern is recommended:

- Increase/ Decrease
- X percentage increase/ decrease
- X time increase/ decrease
- Drastic change in value
- Reaching specific value

In this case we evaluated two types of comparison functions. In one function, the output had a discrete value of -1 or 1 [Figure 19]. +1 implies decision satisfaction by the measurements and -1 indicates vice versa.

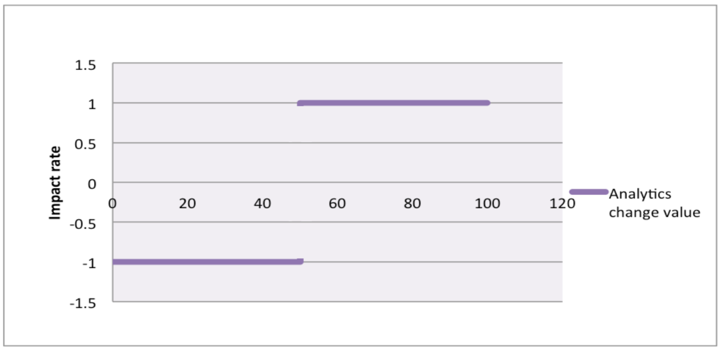


Figure 19: The first applied comparison function

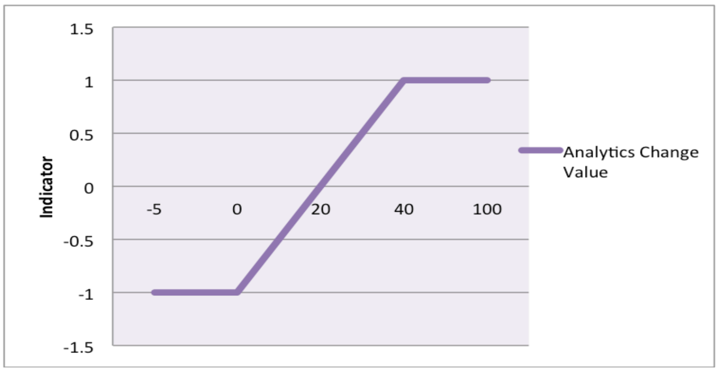


Figure 20: The second applied comparison function

The desired level of measurement values to confirm a decision might not be a discrete variable but a continuous one. Therefore, we also used a continuous function [Figure 20] that presents high-resolution data and ranged any float value between -1 and 1, which is defined based on the product manager's considerations. This function can be presented by another alternative functions as well. The slope of lines for values ranging from -1 to 0 is different with 0 to +1 range [Figure 21]. Based on the nature of measurement values, other curves such as a sigmoid graph can be also used [Figure 22]. More advanced research will provide more insight to select the more suitable functions.

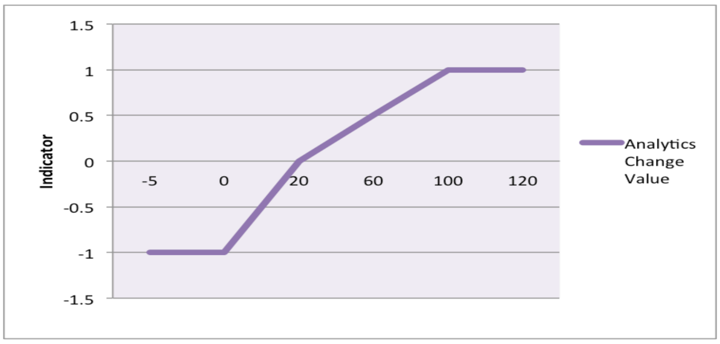


Figure 21: A comparison function (The first alternative)

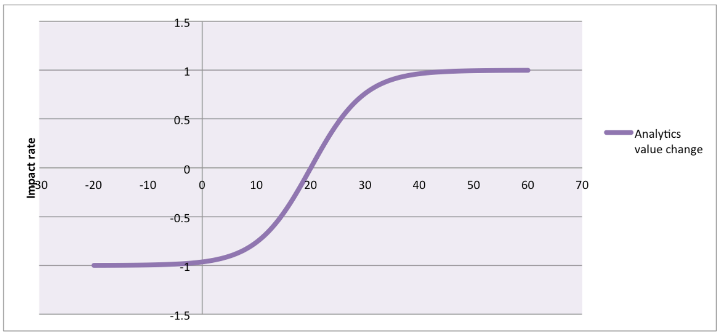


Figure 22: A comparison function (The second alternative)

5.1.1.5 Impact Function

A list of positive and negative confirmation of measurements' changes is essential for our conclusion. An impact function can support decision-making process, although it can be accomplish by considering other factors. Arithmetic mean function was used to calculate an average of impacts by Equation 1, when we couldn't decide from the bar chart alone using our naked eye. There are also other alternatives for the mean formula.

$$A = \frac{\sum_{i=1}^{n} x_i}{n}$$

Equation 1: Arithmetic mean

$$H = \frac{n}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

Equation 2: Harmonic mean

$$G = \sqrt[n]{\prod_{i=1}^{n} x_i}$$

Equation 3: Geometric mean

Arithmetic mean is the most commonly used type of average functions. Sometimes the results can be biased due to big outliers in which case harmonic mean works better instead. Harmonic mean [Equation 2] tends strongly toward the least value of a list to be pessimistic over the data. . Geometric mean [Equation 3] is another type of mean function that shows the tendency to the center. Whenever the data has interrelated, the geometric mean is suggested, but it also presents some limitations. It cannot be computed when both negative and positive values are available in the list or most of the values are zero, while in the analytics-based algorithm, the impact indicator variable, which is a parameter in the formula, can be both positive and negative. This leads to rejection of the geometric mean's suitability for the algorithm.

Selecting the right average formula to use weighted arithmetic mean or harmonic mean can be depended on how impact indicators are distributed and how product manger is pessimistic. If there are outliers in the value list or product manager prefers to decide pessimistically, using harmonic mean function is recommended.

5.2 Case Study Context

5.2.1 Organization

The study was carried out in an organization that developed innovative software as a service for managing different media such as movies, sound, pictures, and text [46]. A product manager and a project manager were responsible for a team of up to five developers. They reported to an internal steering committee that was managed by the development organization, of both the product-owning organization, and departments that used the solution. The organization was small with many responsibilities, which were distributed among a few professionals.

5.2.2 Product

The organization was characterized with a SaaS-based product that was developed at the Zurich University of the Arts available for its students, faculty members and staffs. The product is a collaborative platform for media archives used for content sharing and managing different media to be uploaded and archived [86]. The product provides different features and is rapidly growing.

5.2.3 Information Model

Information about features was organized in a feature tree model [15], which was based on requirements' dependencies. In this solution, a feature was a name for group of requirements that were implemented in the same development. Feature tree included several features and sub-features with AND, OR, and REQUIRE dependencies.

5.2.4 Product Planning

The organization desired to improve its development approach by enhancing the product perspective. The short and long term planning was considered to increase productivity of the limited resources and improve the product quality. Recently they changed their traditional product planning approach to a feature-driven release planning and applied it together with roadmapping to cover timing and resource limitations. Samuel Fricker [15] recommended the feature tree as a solution for release planning. Following planning process was repeated for features within the tree:

Evaluation: The product manager evaluated features with respect to stakeholders.
Identification: The product manager identified all features that need to be implemented and defined critical features for the next release.
Prioritization: The product manager prioritized the severity of stakeholders.
Preparation: As a preparation, the product manager built a pilot project with the prioritized stakeholder. A roadmap was defined to provide the context for release planning. Then all potential features for the roadmap were confirmed by the product

manager together with steering committee. Later on specification for immediate features was refined by the help of the product manager, user experience designer and development team.

Implementation: The feature specification, which included details of requirements, was handed over to development team to be implemented. The product manager supported the development team with internal clarifications.

Feedback: After implementation, the pilot project became available to the existing users in order to collect feedback. Feedback was collected in the pilot project continuously. The feedback fed improvement that was needed to be done for implemented features or specified features.

The organization used analytical tools (Piwik and New Relic) for monitoring the product and collecting data that describe the current situation of the product or features. The data that they collected was relevant but was not enough for increasing the quality of product planning decisions.

5.3 Method Evaluation and Results

The analytics-based method was evaluated using the cases study. To show utility and effectiveness of the artifact's evaluation [85], the case study analysis was conducted in two phases: several examples in different conditions were examined to fulfill utilization goal in the first phase of the case study analysis [section 5.3.1] that was about analytics-support for decision-making. The second phase [section 5.3.2] investigated the effectiveness of the proposed method that also included interviews with product managers. Based on these evaluations, section 5.3.3 will discuss limitations of the analytics-based method.

5.3.1 The Utilization of the Analytics-based Method

The strategy followed for the evaluation of the method covered all types of decisions (discussed in section 2.1), which could be related to several features of the product. Alternatives of variables (time durations, comparison functions, and impact functions) have been studied to evaluate analytics-support for decisions.

Naturally planning decisions incorporate with features. So, a feature should be identified that was extracted from available feature list of the product, or from common feature list of similar product in the media archive context, or the feature was selected as a request from the product manager of the case.

Table 11 presents different practices in the case study to evaluate the proposed method. The "evaluation" column shows if running the method could support the practice.

Table 11: Practices of using the method in the case study

No.	Practices	Feature	Instance decisions	Evaluation	Details of practice
1	Simple Decision	Internationalization	- Should English version for UI be created?	Supported	Table 38
2	Simple Decision	Chrome Support	- Should Chrome support be created?	Supported	Table 13
3	Simple Decision	Automatic Delete	- Should Automatic delete be created?	Not supported (Limitation of tool support for action inside pages)	

No.	Practices	Feature	Instance decisions	Evaluation	Details of practice
4	Simple Decision	Upload PDF larger than 20MB	- Should the upload size for PDF file be enhanced to 20MB?	Not supported (Limitation of tool support for some measurements: Http error particular for uploading PDF, uploaded file size)	
5	Simple Decision	Wiki	- Should Wiki feature be removed?	Supported	Table 39
6	Simple Decision	Wiki	- Should Wiki feature be enhanced?	Supported	Table 40
	Feature Selection	Top level features of feature tree		Supported	Table 14
7	Simple Decision	Facebook Share	- Should the feature of sharing medias in Facebook be created?	Not supported (Limitation of tool support for action inside pages)	
8	Simple Decision	Individual Permission of the group	- Should eliminating permission of an individual from a group, be removed?	Not supported (Limit access to application-measurements: Permissions assigned to user and resources, number of visits with particular permission)	
9	Multi Decision	Internationalization and Chrome support features	- Should English version for UI be created? - Should English version for UI be created?	Supported	Compare Table 13 and Table 38
10	Compound Decision	Wiki feature	- Should Wiki feature be removed? Or - Should Wiki feature be enhanced?	Supported	Compare Table 39 and Table 40

5.3.1.1 Examples of Method Execution

Example 1

This example demonstrates how the method works for a sample decision of ”Should Chrome support be created?”.

Method trigger: Product analysis has already identified a trend of decrease in the number of visits for the last five months, which is presented in Figure 23 (i.e. dashed line presents the trend). Different features can be studied by the goal of finding a solution and take effective planning decisions to reverse the decreasing trend.

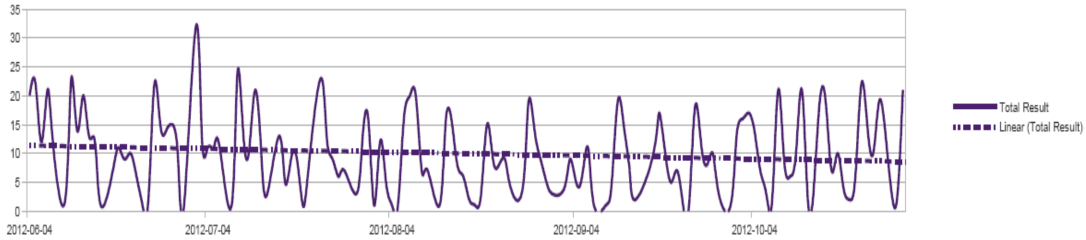


Figure 23: Visits Trend

It was performed as a part of the case study based on the processes explained in section 5.1. These processes are outlined in Figure 24.

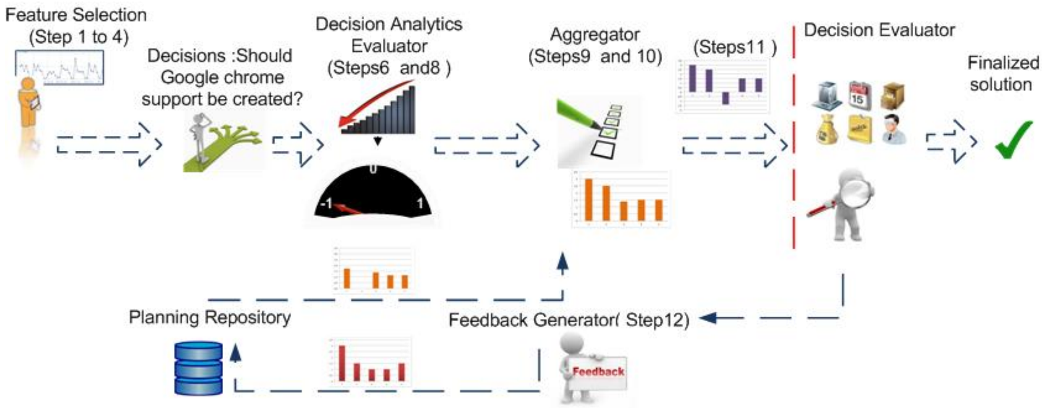


Figure 24: How the analytics-based method works

Step1: Defining an instance decision

Firefox browser is the only browser that product of the case supports it formally. In order to improve the product, the decision of "Should the browser support be enhanced?" is initially defined to be evaluated.

Step 2: Making connection between the feature and the measurements (Specify measurements)

For the "browsers support" feature, data about visits with different browsers such Chrome and Safari can be useful. So "Percentage of visits with Chrome browser" and "Trend of using Chrome browser" were selected from the list of sample statistics to be observed. All involved sample statistics are presented in Table 12.

Table 12: Observed measurements for "Browsers support" feature

Sample statistics	X1	Duration 1
Percentage of visits with Chrome browser	19.16	From June to Oct 2012
Percentage of visits with Safari browser	6.81	From June to Oct 2012
Percentage of visits with Opera browser	0.52	From June to Oct 2012
Percentage of visits with Internet Explorer browser	0.45	From June to Oct 2012
Trend of Firefox use	236 236 159 120 222	From June to Oct 2012
Trend of Chrome use	35 47 64 48 62	From June to Oct 2012

Definitions of columns and their mapping to the process steps of the method [section 5.1] have been shown as follows:

Sample statistics: An instance for a measurement-attribute (related to Step 2).
X1: The observed value of the sample statistic for duration of "Duration1" column (related to Step 3).
Duration1: the time periods(s) within which, the measurement values will be compared (related to Step 3).

Step 3: Observe measurements

Monitoring of browser-use reveals that 19.16% have been visited using Chrome browser, which was the most visited browser after Firefox [Figure 25]. So feature selection was looped with "Chrome support" feature.

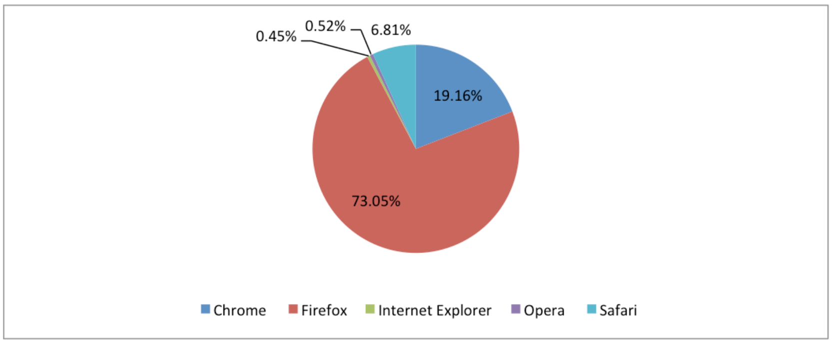


Figure 25: Distribution of browsers used in the case

Trend of the measurement shows Chrome-browser-use has an increasing trend in the case, while Mozilla-Firefox-use shows a decreasing one. The red line in Figure 26 depicts the trend (Mozilla-Firefox-use is the upper diagram) in the graph, which also includes forecasting of November and December 2012.

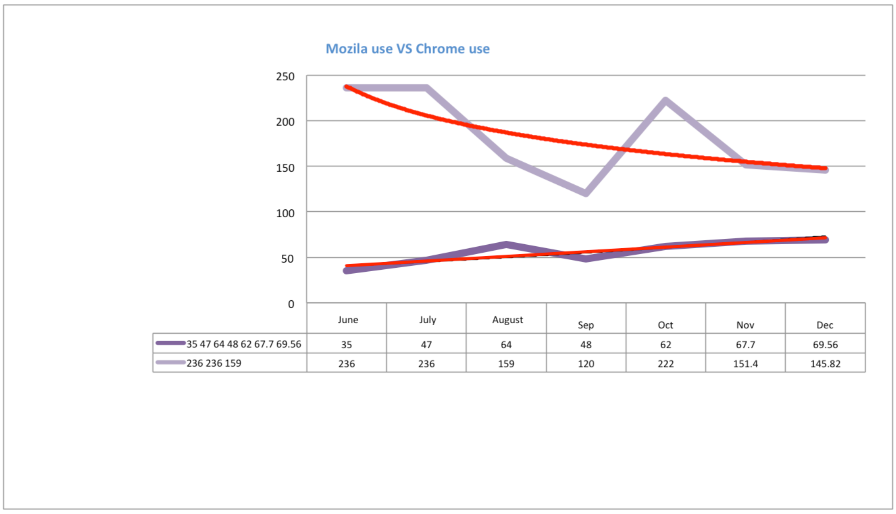


Figure 26: Chrome use trend vs. Firefox Mozilla use in the case

Step 4: Analyze the importance of the feature (Feature analysis)

The world trends of browsers have also considered an increasing trend for Chrome-support [Figure 27], which shows the correlation with the results from previous step. As a result of feature analysis, the decision of “Should Chrome-support feature be created?” was selected to be investigated through the proposed method (Step4).

Step 5: Specified extended measurement

To make decision about creating Chrome-support feature, more complete measurements of “product-use”, ”bounce rate”, and ”click activity” were candidate to be integrated with the “browser” measurement. The data provide insight about the requested feature and thus form a pattern of measurements for the specific feature. In order to collect data, instance statistics should be defined to make them comparable. Table 13 presents the evaluation of data for taking the decision.

The column “sample statistics” corresponds to the instance statistics variables (i.e. “Percentage product use with Chrome browsers” is a sample statistics which integrate statistics about “product use” and “browsers” measurements).

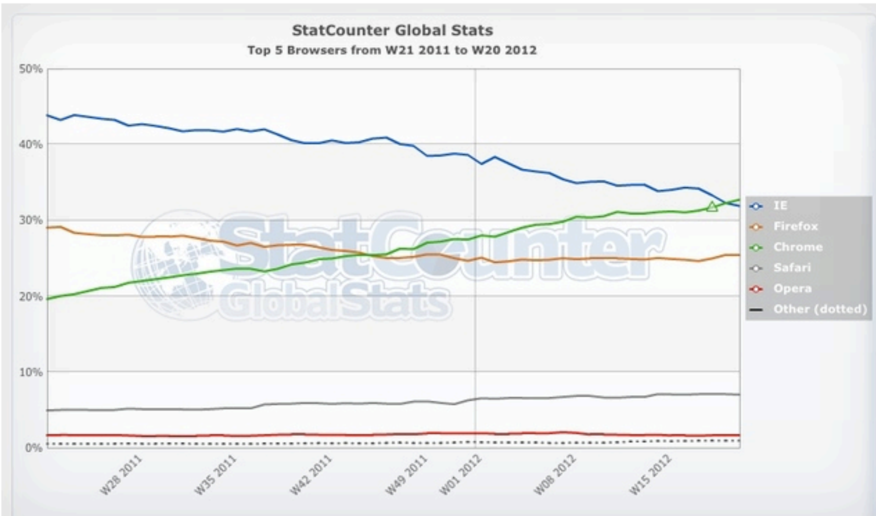


Figure 27: Trends of browsers according to the StatCounter's report.

Step 6: Observe extended measurements

This step is related to observing sample statistics for the defined time periods (“Duration1” and “Duration2” column), which are presented in column X1, and X2 of Table 13.

The time points might be defined by considering/eliminating specific event or exogenous factors such as a network failure (as an exogenous factor) in 2 months.

Step 7: Define a comparison function (Decision analysis)

Answering to the question of “How do the measurements confirm the decision?” directs to specifying if data-changes during a time period has positive or negative confirmation on taking the decision. It was defined by a comparison function. In Table 13, column “Comparison Function” corresponds to the function. The question of “Are the percentage of Chrome users more than 20% of total users from June to October?” defines “Function” variable for “Percentage product use with Chrome browsers” sample statistic.

The function can be simple function, which generates value of +1 and -1 as the output [Figure 19] or be a proportional function to calculate the degree of negativeness and positiveness [Figure 20]. This output is presented in “Indicator” column. The slope of the line in the chart specifies the indicator.

Table 13: Analytics for “Should Chrome support be created?" decision

Measure ment-attributes	Sample statistics	X1	Duration 1	X2	Duration 2	Function’s boundary	Indicator	Recorded rate	Decided rate	Total value
Product use	Percentage amount of use with Chrome browser	19.16%	From June to October			>20%	-0.04	3.58	3.58	-0.15
Browser								1.71		
Overall amount of users	Percentage of visitors using Chrome browser	20.68%	From June to October			>20%	0.03	2.39	3.00	0.10
Browser								1.71		
Returning users	Percentage of return users with Chrome support	21 %	From June to October			>10%	1.00	2.39	3.50	3.50
Browser								1.15		
Duration of using the product	Duration of using product with Chrome browser over all browser	9.51 %	From June to October			>10%	-0.05	1.15	2.00	-0.10
Browser										
Bounce	Percentage of bounced visit with Chrome browsers over total bounces	29 %	For July and August	32 %	For September and October	>	-0.40	1.38	1.71	-0.68
Browser								1.71		
Click activity	Number of total action with Chrome browser per month	271 340 463 262 507				Ascending trend	1.00	1.99	3.00	3.00
Browser								1.71		

The value of 19.16% for the sample statistic (Percentage product use with Chrome browsers) generates -1 as the output (as it is less than 20%), which interprets negative confirmation of the measurement on the decision. To make this level more precise, a proportional function is applied [Equation 4]. The function usually has more than one part, which are differentiated by conditions. (i.e. if x>20% then the decision is confirmed, otherwise it is not, where x is a variable for the measurement value related to the time duration). This condition is defined based on measurement values rated to one time duration or the proportion of their changes in two time durations. Different patterns for defining the functions are presented in section 5.1.1.4. So based on the defined function, "indicator" variable is assigned by the value of -0.04 as the output of the function. The motivation for using the proportion function in the example indicates the difference between 19.16% and 20% (function's boundary) is not too much, and it is not fair to use simple function (to give -1) for both 19.16% and 2%.

$$f(x) = \begin{cases} -1 & x < 0 \\ \dfrac{x-20}{20} & 0 < x < 40 \\ 1 & x \geq 40 \end{cases}$$

Equation 4: An example for comparison function

Sometimes a sample measurement variable uses a trend function. Table 13 shows an increase trend function for "Number of total action with Chrome browser". So, the indicator variable will receive value of +1.

Figure 28: The trend of total actions with Chrome browser in 5 months

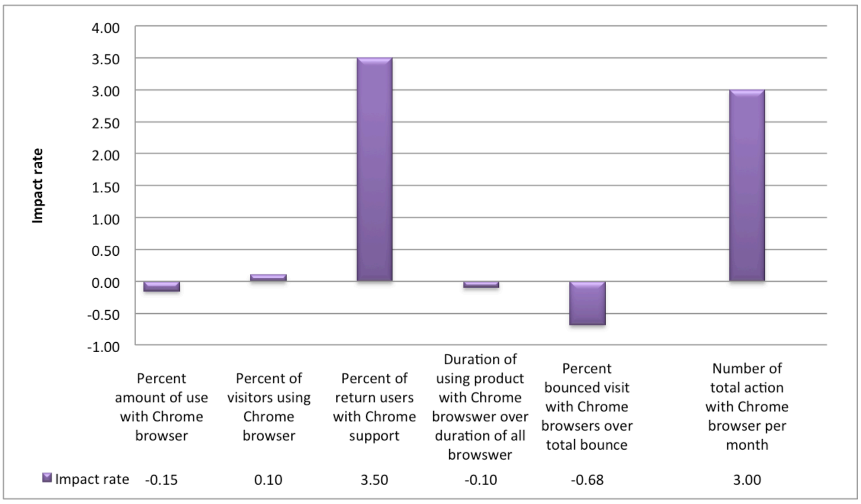
Step 8: Confirm measurements' weights

By specifying the indicator related to a decision, they are aggregated based on the measurements' weights. In Table 13 "Recorded Rate" and "Decided Rate" columns define the weights. Recorded rates are the measurements' weights extracted from the interview-based survey that show importance levels of measurements [section 4.1.2.2.2]. Inappropriateness of the weights by considering criteria related to product/feature, made the user of the method to suggest an updated weight. The criteria were concluded from the qualitative analysis of the interview-based survey in section 5.1.3. If sample statistic relates to more than one measurement-attribute, then one weight is chosen from the recommended weights or an updated weight is assigned.

New weights are in the range of 1 to 4 (not-important, less-important, important, and very-important) by resolution of 0.5 (i.e. it includes 1.5, 2.5, and 3.5). "Decided Rate" column shows the confirmed weights.

Step 9: Calculate comparison values of the measurements

"Total Value" column presents the weighted indicators that multiplies the "indicator" column and decided rate columns. The total values were visualized in Figure 29 for evaluation of the method's user. Alternatively arithmetic mean function used to compare numerically the differences, which confirms the same recommendation. It is more useful when it couldn't be distinguished by naked eye (Step 8).



Decision: Should the Google Chrome support be created?
Recommendation: Decision is recommended to be taken.

Figure 29: The measurements' impact for "Should the Google Chrome support be created?" decision

Step 10: Pre-evaluation of the decisions

Visualizing the "Total Value" through the measurements' impact chart Figure 29 depicts the positive impact of the measurements, which recommends taking the decision from the product manager's perspective.

Step 11: Collect Feedback

This step receives feedback after implementation of the decision. The interview with a product manager revealed that some measurements needed to be updated. The weights of measurements (i.e. product use, overall amount of users, returning users, duration of using the product, bounce, click activity, browsers) were updated with new weights (2, 4, 3, 2, 4, 2, 3). In comparing with previous weights, "product use" has been recognized less important while "overall amount of users" was assigned with more weight, but the weight of "duration of using the product" didn't have any changes. These weights are maintained in storage for future usage.

Applying the method has been found useful to facilitate making the defined decision. The overall monitoring of the product illustrated unsatisfactory progress within the last 6 months. The result of monitoring was a trigger for applying the method. Within the case, the method brought detailed information about using of particular browser such as visit trend, product use, and click activities. This information together with comparison function brought a wide perspective to identify benefit of adding Google chrome feature. The statistics about different browsers were easily accessible by analytical tools, which increased the adaptability of the method. The measurement values provided by survey results assist deciding about measurements' impact. However, the product manager was also able to change the measurement values based on the product and decision types.

Example 2:

Step1: Defining (an) instance decision(s)

Feature tree used as one source of information for data collection. Enhancing the product was the main goal to extend more desirable features. So enhancing "Media Entry", "Media Resource", "Upload" and "Media set" feature sets in top level of the tree was studied separately.

Step 2: Making connection between the feature and the measurements (Specify Measurements)

For all these feature sets, "feature use" was recognized as an effective measurement attribute (Step2). Selecting just one measurement doesn't provide a direct indicator to compare features in the right way unless the product manager can be sure to be able to conclude from the results. The product manager's interpretation in selecting more effective measurements and then concluding from the results has important role in making the connection.

Step 3: Observe measurements

Observing "Number of page view" for each feature (Step3) can interpret that "Media Entry" feature might be an important feature to concentrate on (with 7369 page views) and "Media-Resource" feature was less appealing among users (1099 page views) in comparison with "Media Entry", "Media Set" and Upload features [Table 14].

Table 14: Observed measurements for top level features of the feature-tree

Sample statistics	X1	Duration 1
Number of "Media Entry" feature use	7369	From June to Oct 2012
Number of "Media sets" feature use	6265	From June to Oct 2012
Number of "upload" feature use	1730	From June to Oct 2012
Number of "Media resources" feature use	1099	From June to Oct 2012

Definitions of columns and their mapping to the process steps of the method [section 5.1] have been shown as follows:

Sample statistics: An instance for a measurement (related to Step 2).

X1: The observed value of the sample statistic for duration of "Duration1" column (related to Step 3).

Duration1: the time periods(s) within which, the measurement values will be compared (related to Step 3).

Step 4: Analyze the importance of the feature (Feature analysis)

So a sub-feature of "Media-Resource" was selected to be enhanced. Creating the Facebook-share feature had chosen to be studied. As the feature was not available in the product, related measurements could not be collected. So the method was stopped in this step and implementing a prototype was recommended to monitor it for a structured decision.

In this example, applying the method has been found useful to facilitate selection of high popularity features. For increasing the popularity and quality of the product from users' perspectives, a sub-feature of "Media-Resource" was selected to be studied. As the recommended decision was dealt with a new feature in the product, a pilot project was suggested in order to study the decision impact in time period. By piloting the decision, impacts of adding new feature of "Facebook share" can be deeply analyzed based on users' behavior and the final decision can be clarified.

5.3.1.2 Analytics-Support for Decision-making

The proposed method is evaluated with different conditions and initial data. Based on the case study's design, different patterns of variables specify this supportiveness and those, which cannot support, will explain the circumstances. The results of the analysis have been discussed in the following sub sections.

Decision Coverage

Different types of planning decisions such as create, remove, enhance a feature, prioritize and allocate features to releases, and confirm a new technology were tested in the case. A planning decision could be simple or multi decision. Prioritizing between decisions and allocating to a release are multi-decisions including at least two simple decisions. During the evaluation, a compound decision was tested too. A compound decision is made of two or more simple decisions, which the product manager decides among them. As an example, deciding between "Should Wiki feature be removed?" or "Should Wiki feature be enhanced?" is a compound decision that contains two simple instance decisions.

Table 11 presents the details of decisions that were evaluated in the case, and Appendix C.2 shows how the method was applied for the evaluations.

Studying of the case showed that in the method use, "confirm a new technology" decision plays the same role of "create a new feature" decision, when the feature is equivalent with the corresponding technology. So "confirm a technology" decision was renamed to "create a technology/technology support" decision.

The evaluation of the method confirms that among the road-mapping decisions discussed in section 2.1, "prioritize features in the current product" and "allocate features to releases" are taken when the other instance decisions have been made. In the other word prioritizing decisions doesn't use analytics directly. Other decisions (create, remove, and enhance) are prerequisite for the "prioritize" decision and it is the prerequisite of "allocate" decision.

Analytics Supportiveness

The features for the study have been specified through applying four ways: studying the Madek's features from the application directly, interpreting the feature trees [15], exploring similar media archive applications in the market, receiving recommendation for a feature from the product manager of the case, and monitoring the whole product. By selecting a feature, an interpretation was also performed to determine which sample statistics were related to the feature and how the data could be collected.

One challenging part of the case study was related to finding appropriate sample statistics and collecting them. During the case study, for some features the studied SaaS-based measurements were not supportive [presented in Table 11], while more operational data stored in database could be helpful. (e.g. To remove particular permission of an individual user from a group, effective measurements could be about the permissions assigned to users and resources in addition to number of visits with that specific permission). Also some data couldn't be collected through the analytical tool (e.g. feature related to deleting a resource has implemented as an action inside a page and no URL captures it which is not supported by the Piwik tool). This made to consider circumstances for analytics' supportiveness in the study. These problems refer to limitations of data collection while it is important phase.

There was a case that measurements were collected with some difficulties, and it requires extra interpretation to be collected. For example, "Search" parameter at the end of visited URLs indicates that Search feature was used. So this interpretation implies counting the URLs containing the “search” parameter shows the occurrences of the search features.

The above examples show that there is necessity of setting up an environment to collect right measurements related to features of a product before start of using the analytics-based method.

Alternatives and Boundaries

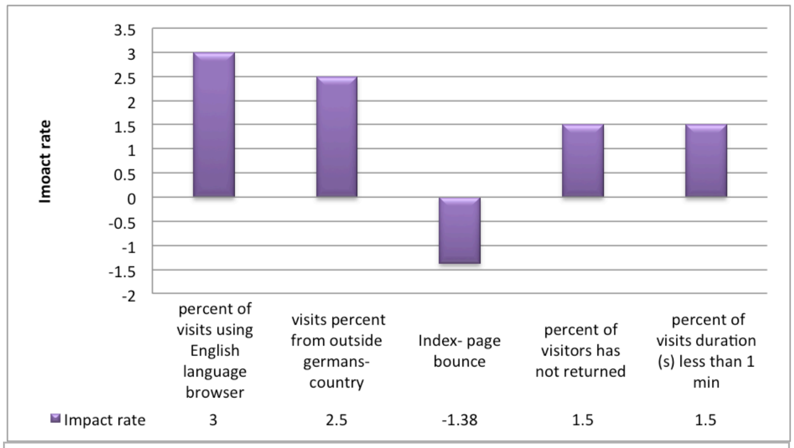
During the case study different defined variables for sample statistics, time durations, comparison functions, and impact functions have been evaluated:

- Sample statistics: simple/ compound
- Time points: past/ present/ future
- Comparison functions: trend function, function’s boundary
- Impact functions: visualizing / arithmetic mean/ harmonic mean

Boundaries have been checked as a part of the evaluation. Assigning +1 or -1 to all indicators, and +4 or -4 to all weights still the method worked properly. Zero value inside the “impact bar chart” implies the existence of zero value for the measurement indicator that interprets no considerable change of the measurement value for the specified duration.

Outputs of the method

The method provides information about measurements’ impact on a decision from product’s perspective. The output implies a recommendation as a guideline to make single, multi-decisions (for prioritization) or compound decisions from product’s perspective but final decision is taken by comparing the product criterion with other criteria. Figure 29 to Figure 32 show the outputs of the method for single decisions but for prioritization and deciding a compound decision, more interpretation was required.



Decision: Should an English version of UI be created?
Recommendation: Decision is recommended to be taken.

Figure 30: The measurements’ impact for “Should an English version of UI be created?” decision

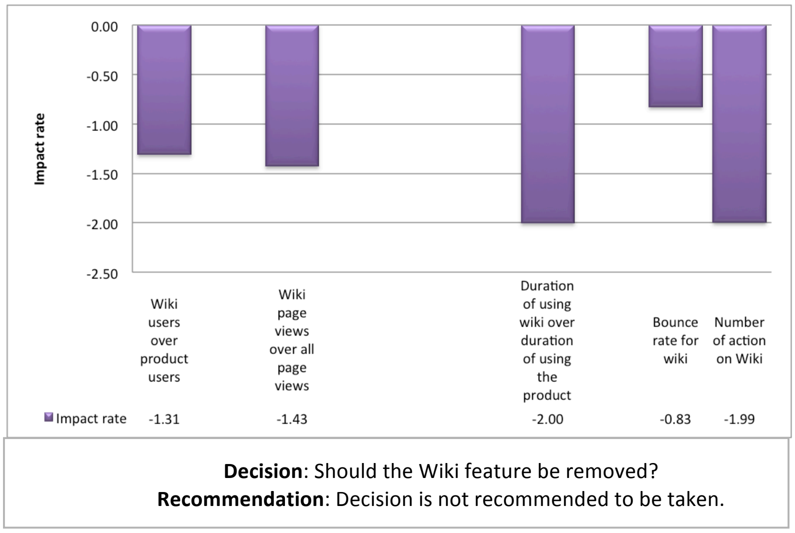


Figure 31: The measurements' impact for "Should the Wiki feature be removed?" decision

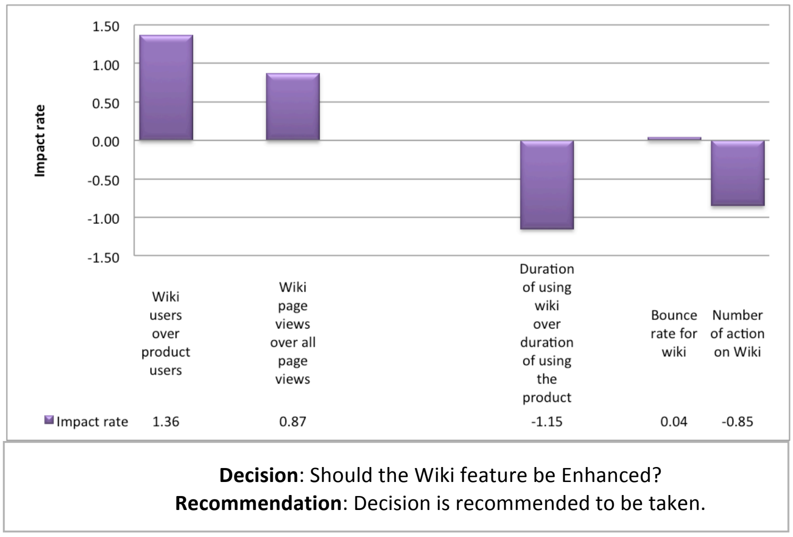


Figure 32: The measurements' impact for "Should the Wiki feature be enhanced?" decision

Prioritization between decisions [between Figure 29 and Figure 30] was conducted by comparing their impact diagrams. Arithmetic mean used to compare the differences, as it couldn't be distinguished by naked eye. The numbers 0.94 and 0.84 for the arithmetic means interpreted more priority for Chrome support in comparison with internationalization.

Impact diagrams in Figure 31 and Figure 32, demonstrates the recommendation of enhancing wiki instead of removing it, as a compound decision. Comparing arithmetic mean of -1.5 and 0.05 for those figures shows the importance of taking the enhancing decision instead of removing decision. Negative arithmetic mean for remove-wiki decision doesn't recommend it to be taken.

5.3.2 The Effectiveness of Analytics-based Method

Effectiveness is "the degree to which something is successful in producing a desired result" according to the Oxford dictionary definition. Through the case study, the method was evaluated by means of a real case in terms of applicability. Data collection and analysis steps were presented in a tabular format and process steps were linked to facilitate traceability. Then the analytical outputs from the real application were evaluated by the product manager of the case in order to specify the degree in which these decisions can be applicable for the planning of a new release.

After that the evaluation was continued through two interviews with product managers: the product manager of the case and another professional product manager who was unfamiliar with the case. The main goal of the interview was to specify how well different processes in the method contribute in providing the desired results.

These interviewees started by a presentation about the method's processes and outputs, followed with questions about selecting measurements, their interpretations, analysis in addition to final solutions. Related questions have been presented in Appendix C.2.

The product manager of the case investigated the outcomes of the method and confirmed that the outcomes were considerable: The decision related to "supporting the chrome browser" had been recently applied in the new version of the product and the decision related to "internationalization of UI" might be considered in future, when their strategy are changed to support internationalization environment. The outcomes for enhancing and removing "wiki" features were acceptable. Also un-popularity of "meta-context" features was found interesting for the product manager to be considered it in the next releases.

The results of evaluating the method's effectiveness through the interviews and the interviewees' quotes are presented in Table 15.

Table 15: Effectiveness evaluation of the method

Evaluation criteria	Evaluation of the Product Manager 1 (the case)	Quotes of the Product Manager1	The Evaluation of product manager 2	Quotes of the Product Manager2	Interpretation
Measurement selection	Easy	"Through the taxonomy I could easily select the related measurements for the decision but without the taxonomy it has challenges"	Can be both easy or difficult	"For some features the measurement are clear and simple; however there might be features that require contemplation and extra time. Sometimes, it might be required to mix measurements and consider them in more details."	Between easy to Difficult The list of measurements can be helpful but how to combine the measurements has some difficulties
Measurement interpretation	Easy to difficult	"It depends on the type of the measurement."	Difficult	"Measurements interpretations are always difficult. Because there might be other external factors that make the interpretation complicated: environmental, seasonal, economical, political factors or etc."	Between easy to difficult It depends to how the measurement is simple or compound and which external factor should be considered and how.
Analysis of feature selection by measurements	Strong	"If the related measurements can be collected by automatic tool then it is strong. There was a problem that in the previous system observing most of measurements depended on URLs and different features could be related to that URLS, but for the new release we have changed it completely specially in UI. Now it is much easier to relate features to URLs and observe feature measurements"	Range between very strong and strong	"Measurements are useful for decision-making and feature analysis. They can strongly facilitate analysis of an importance feature."	Range between strong and very strong. But the tool should be able to collect measurements related to features but now the implementation style may prevent it.
Comparison process for measurements	Medium	"Defining goal for comparison need some challenges "	In the range of Medium till Easy	"It is very easy because giving weight for comparing the measurements is initiative, so this reduces the complexity."	Range between medium and easy Comparison provide result logically although requires learning

Evaluation criteria	Evaluation of the Product Manager 1 (the case)	Quotes of the Product Manager1	The Evaluation of product manager 2	Quotes of the Product Manager2	Interpretation
Method's output	Strong	"The output presentation should be improved. The table in the graph is not so much understandable, so it is better to combine it with the decision and recommendation."	Strong	"The method provides the strong contribution for product planning decisions."	Strong Needs improve presentation which was applied
Prioritizing alternative decisions	Medium	"Different factors are used to support prioritization. For example stakeholder group may need a features, so it has priority. When a stakeholder group needs two features in the same context, measurements help to prioritize them."	Strong	"The method is strongly support prioritization between alternative decisions."	Range between Medium to Strong When feature are in the same context, measurement is helpful.
Trade-off between decisions	Very strong		Very strong	"The method is very strongly support trade-off between alternative decisions."	Very Strong
Feedback from implemented decision	Between useful and very useful		Useful	"The feedback from previous monitored decisions is obviously useful, because it expand product planning vision."	Range between useful and very useful It provides expandable vision in planning.
Uncertainties handling	Medium	"There are lots of uncertainties"	Strong	"The method strongly helps uncertainties- handling for internal factors. However it might have less effect in handling close-level of uncertainties."	Range between medium to strong There are lots of uncertainties and the close-level uncertainties receive less impact by the method.
Method effectiveness	Strong	"There is a limitation of tools to be closer to features. If features are not recognized by URL, a customized tool is required"	Strong	"The method strongly facilitates decision-making by considering the internal factors. However it might not be helpful in every situation, specially where the external factors have stronger impact on planning decisions than internal ones."	Strong The method might not be helpful in the situation, where the external factors have stronger impact on planning decisions than internal ones. Tools also should be close to features.
Applicability of the method in organizations	Easy to Difficult	"For the small online project is easy but for whole organization has different measurements that need more efforts and requires a specialist to interpret the data. "	Easy	"It can be easily implemented if an analytical tool is available."	Range between Easy to difficult For an online product is easy while an analytical tool has integrated.

5.3.3 Limitations of the Analytics-based Method

The study presented that decisions about "prioritize features in the current product" and "allocate features to releases" are not supported directly by analytics. Analytics is helpful to make creating, removing, or enhancing decisions individually, and then to prioritize features, utilize the impact chart (or data). Prioritization features and comparing the impact bar chart are meaningful when the features are in the same context. For the features with different corresponding measurements, the method cannot be helpful for prioritization of those features. Supporting analytics for "Allocate Features to Releases" decision is limited to the support of "Prioritize Feature" decision, as a pre-requisite decision. However other external factors (i.e. release time, resources) have the main impact on allocating features to release.

For using the analytics-based method, there is necessity of having right measurements that are collectable. The measurements are recommended not be limited to web level, while application level measurements available in the organization's database can be helpful for some features. Setting up the right environment initially is required to collect appropriate measurements related to a feature/decision. A customized analytical tool that is implemented for a business may be able to collect web analytics and business analytics together to support those features with the operational data stored in business databases. On the other hand, the general available web analytical tools support collecting data about web pages while features are not always equivalent with a webpage and it can be presented as a part of a web page or even in several webpages. So available web analytical tools have limitations of collecting data about all types of features, which again confirm the importance of instruments for using the method.

Another limitation returns to maturity of product. When a product is newly created, there is no user experience for the product use, unless measurements are collected from similar product. Developing a prototype before finalizing the plan can provide the opportunity of using the valuable information for planning.

5.4 Interpretation

The proposed analytic-based method aims at answering RQ2, which propose an effective way for analytics-based product planning. The method works as a decision helper for a product manager. The collected data provides product managers a wide view about changes of measurement values for a feature and their impacts on his decision. In other words, it gives a way of thinking and analyzing all criteria together, but doesn't provide him a determined decision answer. In this method, analytics guides product managers, but not as the only decision-making tool. The method was validated by a case study.

The utility of the method is for planning any software product that is SaaS-based and a set of measurements has been already collected from. Although it can be generalized for various types of web applications, but the proposed measurements in current study are most meaningful in the SaaS context.

Decisions about creating, removing, and enhancing a feature is supported by analytics while prioritizing decision utilizes the impact information for its simple decisions when they are in the same context. Allocation decision utilizes the prioritization output in addition to other external factors to be determined.

Setting up an environment for collecting right measurements with an appropriate data gathering mechanism is the pre-requisite of using the analytics–based method. An instrument might require to be implemented, because the general available analytical tools support just those features equivalent with a webpage not those that is presented as a part of a webpage or multi pages.

The results from evaluating the method effectiveness showed positive feedback from the product managers. The product managers believed that the method can be effectively considered for prioritizing planning decisions, and selecting between alternative decisions based on corresponding measurements when features are in the same context. They believed the method could be easily performed if an organization has access to an analytical tool specially for an online product. All the steps of the method were questioned in the interviews and the results showed that interpreting the measurements was the most challenging step. The list of measurements can be helpful but the difficulties referred to the way of combining the measurements. In fact the measurement interpretation depends on how much the measurement is simple or compound and which external factor should be considered and how. Regarding comparison between results, they mentioned the comparison provides results logically although it sometimes requires learning. The feedback part also seemed so interesting for them because they believed it provides expandable vision in planning.

The method had some limitations in their point of views, where one was related to situations that the a decision needs to be taken by considering external factors, and in that case the output of the method will not have enough credibility. They believed external influential factors in organizations make interpretation complex. The method might not be helpful in every situation, specially where the external factors have stronger impact on planning decisions than internal ones. There are lots of uncertainties and the close-level uncertainties receive less impact by the method. It is also a limitation for selecting a tool that is close to features of a product and be able to collect measurements related to features.

5.5 Summary and Discussion

The proposed method, tries to shift the mode of product planning decisions from the intuition-based mode to a more data-driven mode, where inevitably the intuition is also involved but it is applied in smaller granules of decision. So it is clearer and less random. In this proposed method, the value of the features and their priority is evaluated in a process, which is rationalized by analytics. This process tries to segregate product-planning steps and then aggregate them into the final decision.

The cost of using the method is calculated by the cost of providing analytical tools used to collect data automatically, and the cost of connecting the tools to the product. The tools can be bought, developed, or even downloaded freely. Also the costs of human involvement in manual tasks should be considered.

In the analytics-based method, analytics supports both feature selection and decision analysis. For decision analysis, all simple measurements are involved to make feature-related measurements that are evaluated in terms of positive and negative effects on the decision. But for feature analysis, most effective measurements are involved in the process, which are specified based on product manager's interpretations. Although the product manager can interpret and conclude from one measurement in feature selection, more measurements involvements in the process conclude more decrease in misinterpretation. For instance, increasing the feature use can have more reasons than increasing feature popularity. Badly implemented feature might have high use rates at first but decrease in time duration. Increasing of the bounced rate and exit rate besides increasing the feature use might indicate a problem such as a badly implemented feature.

A single measurement might be insufficient reaching the right conclusion. A rarely used feature doesn't necessarily mean that it is less important than the other features. But considering it combined with other measurements can provide better insight, although it finally requires the product manager's interpretation. Therefore using a list of measurements is recommended to decrease misinterpretations when the product manager might also be confused from the results.

Although the analytical methods such as A/B testing (for usability engineering) [66], fault tree (for reliability engineering)[67] or QoS computation model (for QoS analysis) [70] can provide data to support product planning partially but using the analytics-based method put all measurements together and provide an overview to generate a recommendation for planning the product.

The analytics-based method can also be supportive for some of available planning approaches. Feature tree is an approach for product planning which reduces the complexity of planning by providing a general overview about what to implement and when to implement in a form of a tree [15]. Feature hierarchy shows a sub-feature as the child of the parent feature, where it should be developed after the parent. By this approach, the evolution is planned and the development progress is reviewed by the aid of color codes allocated to features boxes of the tree (i.e. “Next major release”, “After next major release” and “Not yet implemented”). The proposed analytics-based method has a strong correlation with the approach that their integration facilitates the product planning and improves its accuracy.

The analytics-based method is useful for feature tree approach to identify and prioritize feature developments (which are product planning steps), allocate the color codes to them and provide the feature tree. Monitoring measurements related to features of the same level and branch in the tree provides more comparison data to define which branch of tree should be developed with more priority. As an example, comparing the number of users for those features, feature use, the bounced rate and etc. can discuss the priority to concentrate on the feature with less desirability and improve it, or develop the sub-features related to the more desirable feature. The prioritization strategy depends on the nature of the feature and product manager’s interpretation, which utilizes the method as a feature selection procedure. Also it is possible to consider it as a prioritization decision and use the whole procedure of the analytics-based method to apply method’s recommendation and select between branches of the feature tree. When strategy is defined, for the selected features, creating, enhancing or removing decisions are defined and evaluated through the method with a prioritization decision after all. This will facilitate providing data-driven feature tree.

Another approach to product roadmapping [16] for small organization identified four-steps process:

1. Define strategic mission/vision and outline product vision
2. Scan the environment
3. Revise and distil the product vision as product roadmaps.
4. Estimate product life cycle and evaluate the mix of development efforts planned

The analytics-based approach is a complementary approach to provide more information in these processes. Monitoring of measurements can be used as a guideline to shape product vision (process1) and identify trends in the product environment and potential customers (process 2). Although concentration of the study was for product’s perspective, generalization will encompass other factors such as competitors and markets. Product’s requirements are defined through the related decision evaluation in the method to provide the roadmap (process 3) and track the analytical history of the product use to specify the rational behind roadmap evolution. Although recommendation from prioritization decision in the method can be used to specify the releases, but the activities of the process4 are mostly not supported by the proposed method.

Standard T-plan [45] as an ad-hoc plan, discusses road-mapping activities in three layers of market/business, product/service, and technology, in addition to time-basis layer. By integrating it with R. Phaal and G. Muller’s model [44], another layer for constructing the “roadmap” is formed that handles the aspects of resources and time frames. This model will clarify the position of the analytics-based method in a traditional product planning. Figure 33 presents the four layers of traditional road-mapping activities and relates it to the analytics-based method.

The analytics-based method supports "Product/Service" activities for the decisions about features in an affirmative or non-affirmative way while its support for "Technology" can be evaluated partially. The two perspectives of technology that were achieved in the interview and concluded various valuations for the corresponding measurement-category, correlate user-centric and technology-centric approaches [103]. User-centric approach is being more supported in the method as the more weighted measurements are mostly related to end-users' behaviors.

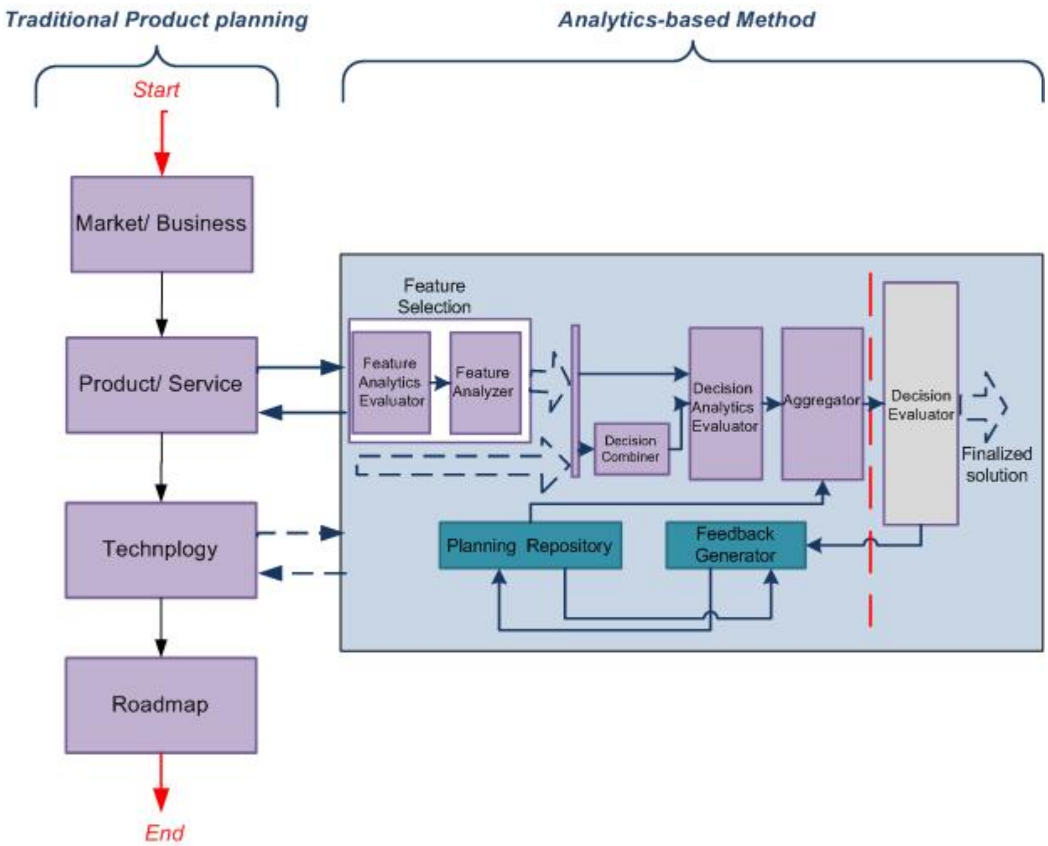


Figure 33: The position of Analytics-based method in traditional product planning

Most of other different planning approaches concentrate on techniques of requirement engineering and prioritization in a product plan [6][24][33][26][38]. However none of the product planning approaches has considered the analytics-based feedback. The proposed method in this study doesn't replace these techniques but instead works aligned with them by generating more supportive data. The main challenge of these techniques is about finding the right value of features or requirements. Although the proposed solution does not identify the real value of the features, but the comparative data and small-granule reasons are good supportive data to feed these techniques.

Generalized characteristic of the method shows its potential for the possibility of supporting internal and external strategic planning [63], although it was excluded from the current study.

Amongst road-mapping decisions discussed in section 2.1 "create a new feature", "enhance a feature", and "remove features" are supported directly by analytics in the proposed solution. "Prioritize features in the current product" is supported conditionally for the features in the same context while "allocate features to the releases" utilizes the prioritization output and "allocate resources" decisions is not supported. Selection or prioritization of decisions is based on the degree of affordances of the corresponding measurements. "Confirm technology" decision could be addressed by other decision about creating or enhancing a feature, which means a technology-support could be interpreted as a typical product feature. For instance, "Creating chrome support" is a "create a new feature" decision with the technology support as a feature. So it can be dismissed from decisions discussed in section 2.1.

There of course exist limitations to the effectiveness of this method. The proposed method is subject to different degrees of effectiveness regarding the degrees of similarities in the measurements related to two different decisions. In this sense, the more similar measurements in two different decisions mean the more degree of comparability and hence more effective method. Also the proposed method does not directly and specifically suggest feature/resource allocation for each release as the decisions involve other important considerations such as the available resources and proportion of the timeframe.

5.6 Validity Threats

Like any other research, current study has been challenged by some threats to validity. The main threats were trustworthiness and completeness of the proposed method. The method was tested with boundary conditions for different weights of measurement-categories and measurement-attributes within the case. It was also tested with all decision types in roadmapping. Different product manager perspectives (optimistic, pessimistic and moderate) were taken into account in reliability analysis. The method was deeply studied within the case study. Yik [87] recommends four tests for establishing the quality of empirical research specially in case studies: construct validity, internal validity, external validity and reliability. The testes were considered for mitigating threats of the case study as follows:

5.6.1 Construct Validity

For case studies, there is always a risk that investigator does not develop operational measures and uses subjective judgment to collect data [87] . The construct validity of the case was increased by the following three tactics:

Use multiple sources of evidence: Precious advantage in case study is gained when more than one source is considered for collecting data [87]. As mentioned in section 3.4.2.1, evidences for the case study came from different sources of archival records and direct observations such as Piwik tool (as an opensource web analytics software), Piwik databases (were accessible through SCP(Secure Contain and Protect server) server) , and NewRelic tool (web application performance management tool). The Feature tree document was another source being used during the case study, showed the available features of the product and provided guidance about when to implement the unavailable features.

Establish chain of evidence: Data traceability increases the reliability of information in a case study [87]. In order to organize the evidences and facilitate traceability, data collection and analysis steps were maintained by Microsoft Excel tool in a tabular format. Also version management was considered to present the evidence order.

Validate the collected data through the interviews: The evidence from the case was supposed to be validated by the product manager of the company through the interviews. However, two professional product managers conducted the validation: the product manager of the case and the other product manager who was unfamiliar with the case. The validation improved the accuracy of the collected data and analysis, hence increased construct validity of the case. To improve the accuracy in validation results, training sessions were held for presenting what was done during the method with concrete examples and what are the case's specifications (for unfamiliar product manager with the case).

5.6.2 External Validity

External validity identifies domains within which a case study's results can be generalized [87]. As it was a single case study, there is a threat that the obtained results can only be applicable for a particular company. There was no possibility of applying the case in other companies and replicating the findings because of the vast analysis boarder and time pressure. For mitigating the external validity threat, a theory was considered which is explained in section 3.4.2. Using a theory increases the probability of achieving same results in different domains which increases the external validity [87]. Furthermore the context of the company is explained in detail, which strengthens the generalization by facilitating the ability of mapping the company to other organizations with similar context. The analytics-based method is well suited for other software contexts as well, when the product manager selects sample measurements from the proposed patterns of measurements, functions based on decisions, and reconsiders the weights based on particular product type and situational factors.

5.6.3 Internal Validity

Internal validity makes a causal effect relationship and distinguish spurious relations [104], The study interfered with some confounding factors. One confounding factor involves problem on learning the method and applying it by the product manager. To avoid interfering the effect of learning for using the method with benefits from evaluating, we separated the activities aimed at learning from activities for evaluating. In our case, learning phase took four months for us to show how the method is applicable in the case. Then the evaluation period was started by product managers. Also, for this step training activities for the method and evaluation activities were separated to mitigate this validity threat.

Another confounding factor was related to rarely used features. When a feature is rarely used, it doesn't mean that it has less important among others. The effect was mitigated by applying product managers' perspective about feature values and using the product specifications. While using the analytics-based method, this perspective can be applied on the defined goal for each feature's measurement.

Badly implemented feature is another confounding factor, which is mitigated by using other feature-related measurements. It will facilitate finding problems while the collected measurement values do not show the same direction.

5.6.4 Reliability

The reliability illustrates that if another investigator redoes the case, same results will be obtained. Since a main challenge with the case study is reliability [105], the results were examined based on repeatability and stability in three steps: first, the degree to which a measurement repeatedly performs the same, second, the measurement of stability over time, and third, the measurement similarity within the particular time periods. To do so, the case was designed in detail, documented in every step including data collection, theory, and analysis, which are mentioned in section 5.1. Then for some random decisions, follow the documented process again to check the repeatability and stability for the case.

Another issue indicates that for some measurement-attributes there was a risk of unreliable weighting subjected to interview limitations. This threat was mitigated by considering alternative arguments in the analytics-based method.

6. DISCUSSION

6.1 Strengths and weakness of an analytical product Planning approach

According to the study, analytics-based product planning has some strengths: It shifts the mode of product planning decisions from the intuition-based mode to a more data-driven mode which increase usability and reliability of the product.

The limitations of the approach address difficulties in finding/implementing right tools, and interpreting the observed data. It is important to collect right measurements related to a feature through a suitable tool. Collecting the measurements is possible just for a mature product. To interpret analytics, the measurements need to be combined with observations about external factors such as environmental, seasonal, economical, political factors, which all can have some difficulties.

For some product features the related measurements are clear and simple, however there might be features that require contemplation and extra time. Sometimes, it might be required to mix measurements and consider them in more details. Some planning decisions such as "allocating resources" cannot be supported by measurements, as related measurements were not included in the taxonomy.

6.2 Potential Opportunities of Analytics-based Method for Planning

This analytics-based method supports product management decision making in terms of providing a way of taking into consideration the trends of product usage and reacting to irregularities, outliers or deviations in visitors' behaviors and considering their prediction to turn operational data into strategic information [59]. Upstream scoping and downstream change decisions [37] are supported in the method. This provides supportive solutions for some challenges of product planning. Foreseen feature release, prioritization of requirements and features, and project monitoring are in the list of 12 challenges of product planning mentioned by A. S. Danesh and R. Ahmad [9], which can be overcame through the analytical solution by analysis of data in order to discover new and meaningful data patterns.

Analytics-based method supports both reactive and proactive planning. In reactive product planning, a software product manager regularly analyzes the soft measures in comparison with the product plan and performs an action for significant deviation from planned measures [6]. In a proactive planning, a product manager upgrades the product with new features or proposes the release of new products, which are based on predictions of the product's future state with the aim of solving problems and satisfying customers.

For strategic planning of a software product, initially strategic mission and vision are defined to reflect organizational goal in the plan. Market opportunities, business rules, and their drivers in addition to financial goals and targeted customers are identified. Product vision is outlined and product features accompanied with the key prioritizations of their development are identified. Technology solutions are defined, and the releases and timeframes are decided based on available resources to form the development efforts and finally provide the roadmap. Specifying product outline, product features and their prioritization are supported by SaaS-based supportive measurements through the proposed method. Technology solutions for client side can also be supported but analytics cannot support technologies related to development such as programming languages.

Although the SaaS-based analytics provides information for some activities such as defining organizational goal and business rules affected by external or internal factors (e.g. competitors, market, resource), but the provided information is not sufficient. The activities cannot be mainly supported by web analytics, while business analytics can be supportive for them [63]. SaaS based analytics helps to take decision from product's perspective but external factors will effect on making the final decision.

6.3 Analytics-based Approach versus Other Approaches

The study proposed an analytics-based method that applied different types of measurements to assist a product manager to find vital features and make a planning decision from product's perspective. The method was presented in eleven processes that involve selection of measurement-attributes (related to a feature), observing measurement values, evaluating all measurement together, and providing one recommendation for the decision from the product's perspective, using the values of measurements aggregately. There is also the possibility of receiving a feedback about the measurements' weights after implementing the decision.

The SaaS-based delivery model provides facilities to gather a new range of detailed, usable and real-time measurements which are more reliable information for the taking the planning decision that leading to improved quality and efficiency of the product [106]. Also for taking planning decisions, one might rely on trends of data changes and evaluate which and how the new product planning could seize the opportunities reflected in the data and their trend lines.

Feature tree is an approach for product planning which reduces the complexity of planning by providing a general overview about what to implement and when to implement [15]. The proposed method has a strong correlation with the approach and their integration facilitates the product planning and improves its accuracy. The analytics-based method is useful for feature tree approach to identify and prioritize feature developments (which are product planning steps), allocate the color codes to them and provide the feature tree.

Another approach to product roadmapping is presented by [16] for small organization in four-steps process. The analytics-based approach is a complementary approach to provide more information in these processes. Monitoring of measurements can be used as a guideline to shape product vision (process 1) and identify trends in the product environment and potential customers (process 2). Although recommendation from prioritization decision in the method can be used to specify the releases, but the activities of the process 4 are mostly not supported by the proposed method.

Product/service, and technology layers in Standard T-plan [45] (as an ad-hoc plan) are supported in the analytics-based approach by providing high quality information to perform corresponding activities. The relation of the proposed method and traditional road-mapping activities can be seen in Figure 33.

Most of other different planning approaches concentrate on techniques of requirement engineering and prioritization in a product plan [6][24][33][26][38]. However none of these product planning approaches has considered the analytics-based feedback while considering might improve the quality which can be considered as a future work.

6.4 Planning Decision Support in the Analytics-based Approach

Product planning decisions have been considered in different studies from different prospective. J. Momoh and G.Ruhe mentioned five-steps release planning decision-process done by a product manager: “elicit requirement”, “specify problem”, ”estimate resource”, ”estimate alternatives plans”, and “implementation ” [35]. A. Nejmeh and I. Thomas looked at product planning decisions from business-driven perspective: “assess feature business value”, “determine cost” and etc. [43]. I van de Weerd, W.Bekkers and S.Brinkkemper categorized software product planning process with capability perspective: “requirements gathering”, “requirements prioritization”, “release definition” and etc. [36]. The decision taxonomy includes all these decision components but in a different level. The taxonomy concentrates on higher level of decisions while the decisions in the previous study were mostly in a lower level, which two or more can make a higher one together.

Amongst road-mapping decisions discussed in the taxonomy, “create a new feature”, “enhance a feature”, and “remove features” are supported directly by analytics in the proposed solution. “Prioritize features in the current product” is supported conditionally and “allocate features to the releases” decision utilizes the prioritization results, while “allocate resources” decision is not supported. Selection or prioritization of decisions is based on the degree of affordances of the corresponding measurements. “Confirm a new technology” decision can be addressed by other decisions about creating or enhancing a feature, which means a technology-support can be interpreted as a typical product feature.

6.5 Measurements Support in the Analytics-based Approach

The taxonomy of measurements consist of 6 categories: “Product”, “Feature”, “Usage pattern”, “Referral sources”, “Technologies and channels”, and “Product healthiness”. The categorization was defined based on the definitions of measurements, product planning definitions by inspiration from references of web analytics. The taxonomy conceptualizes a web application as a product, which consists of features. General web analytical tools collect data about page use, which might not be mapped to the feature use, as a feature might be a part of a page or even multi pages or might not be collectable through a server request.

So, setting up the right environment for collecting right measurements is a pre-requisite of using the analytics-based method. There is necessity of selecting right measurements for a decision and make it available to be collected. So attempting to adapt analytics strategy inside the organization is a challenge part of using web analytics [107]. Providing a right instrument by configuring or customizing available tools or implementing a new tool should be concerned before applying the analytical-based approach.

Organization can use analytics in several areas of their strategic planning (e.g. Porter’s five force factors as the external factors and process, operations, and resources as the internal factors) [63]. The current study has investigated practical usage of data in the operational and customer’s perspective, but the proposed platform is flexible enough to support different aspects of planning. It only needs to study the measurements in other analytics more than web analytics and in parallel, benefit from the characteristic of SaaS based product delivery model for collecting the corresponding web analytics. Evaluating all type of analytics together will hopefully provide more supports for the areas of strategic planning which can be studied as a future work.

6.6 The Importance of Analytics for Software Product Planning

The results of interview-based survey showed measurement-categories of "Product", "Feature", and "Product healthiness" are the most important categories for taking product planning decisions while "Referral sources" doesn't have too much effect. The study of the importance level of measurements for product planning decision has been conducted as well. The results show that "product use", "feature use", "users of feature use", "response time", "product error", and "down time" are the first top measurements for product planning. The result correlates with the study that shows in a SaaS-based software, measurements about usability, performance and productivity have been recognized as the most important data [13].

In traditional planning, values of features have been always important criteria. Stakeholders might specify a feature value by simply assigning a number, based on the assumed impact of the feature, on the overall product [43][41]. The values can suffice to prioritizing the features over each other [26], which means there would be no absolute valuation, but each feature is relatively situated amongst two other features. This correlates with the study results that "product use" and "feature use" measurements are the two most important measurements for product planning that might assist product managers to compare the priority of decisions when features are in the same context. Qualitative analysis of the survey results presented some factors that were mentioned as the affected parameters on the selection of the measurements' importance. These factors could be related to the factors of product specifications (e.g. customer type, access type, network type, users' numbers), product maturity and product goals. The factors correlates with situational factors [99] discussed for the SPM area.

7. TEN LESSONS LEARNED

1. Different planning decisions consider the importance of measurement-categories in a similar way. The interpretation indicates that if one set of measurements has been recognized important for planning, it can be used importantly for all decisions regardless of their types.
2. Some measurements have more value for product planning. "Product", "Feature" and "Product healthiness" are important measurement-categories for planning. Measurement-attributes of "product use", "feature use", "response time", "users of feature use", "product errors", and "down time" have been recognized as the first top measurements for product planning.
3. Amongst roadmapping decisions, decisions about "creating a new feature", "enhancing a feature", "removing features" and "prioritizing features in current product" are supported by the SaaS-based analytics method.
4. Setting up the right environment for collecting right measurements is a pre-requisite of using the analytics-based method.
5. For newly created product, the analytics-based method cannot be used effectively. As there is not enough user experience for product use, the suitable data are not available to be used for planning. Developing a prototype before finalizing the plan can provide the opportunity of using the valuable information for planning.
6. To support all aspects of product planning by analytics, SaaS-based analytics might not be sufficient. There are activities of product planning that SaaS-based measurements cannot provide sufficient information for them while business analytics might be supportive.

7. In the analytics-based method, product criterion is just one factor to accept the analytics-based recommendation for decision making. Final solution is achieved after trade off between product criterion (achieved in using the analytics-based method) and other criteria (e.g. resources, market, competitors).
8. The more measurements are applied for analyzing a feature/decision, the more decrease in misinterpretation would be gained.
9. The analytics-based method might not be so much helpful in the situation, where the external factors (e.g. customer request, market needs) have stronger impact on planning decisions than internal ones (i.e. product criterion is specified through the method's recommendation).
10. The analytics-based method in this study doesn't replace most of available product planning techniques but instead works aligned with them by generating more supportive information.

8. CONCLUSION

Products are the artifacts to satisfy the customers' needs, and hence product managers require bringing the voice of market and customer to the product planning processes, where this happens effectively through a data-driven endeavor of sensing and understanding the requirements. Different types of measurements assist a product manager in product planning, where each might be gathered through a different channel and process. In software product management, cloud computing and SaaS-based product delivery have provided opportunity of improving quality and efficiency of product by gathering a new range of detailed, usable and real-time user-related data, which it was not possible to get before.

The research was conducted to show how analytics assists product managers and provide them the right information for software proactive and reactive planning. The literature review of the study specified a taxonomy of SaaS-based measurements in six categories: "Product", "Feature", "Usage pattern", "Referral sources", "Technologies and channels", and "Product healthiness", in addition to a taxonomy of planning decisions taken in portfolio management, road-mapping and release planning. Later on, the research got focused on the roadmapping as the interviewees were experienced more. An interview-based survey research was conducted with the professionals in the product management area to understand the effect of analytics on planning decisions in a software product.

The results of an interview-based survey (address RQ1) illustrated that distribution function of measurement-categories are not different for planning decisions of a software product. "Product", "Feature" and "Product healthiness" were recognized as "very important" categories while "Referral sources" category was chosen as "not important" (address RQ1.1). Then a list of overall scored measurements was calculated which presents that the measurements of "product use, feature use, users of feature use, response time, product errors, and downtime" are the first top measurement-attributes that a product manager prefers to use for product planning (address RQ1.2). During the qualitative analysis of the interview-based survey, we found some clue that we need more consideration when using the overall scores of measurements. The justification of interviewees for assigning a value to a measurement shows that different factors such as product characteristic, product maturity and product goal have affected on their selection. Investigating about these factors needs a further study in a future work. The qualitative analysis also identified the strengths and weakness of using analytics for planning. Analytics increases knowledge about product usability and functionality, and also can assist to improve problem handling and client-side technologies. But it has limitations regarding to receiving formed-based customer feedback, handling development technologies and also interpreting some measurements in practice. Immature products are not able to use analytics too (address RQ1.3).

An analytics-based method was proposed to support planning of a software product in a design science research (address RQ2). It designed with eleven main processes to transform measurements to a recommendation for product planning. It was evaluated with different conditions and initial data: different types of decisions (create, remove, enhance a feature, prioritize, and allocate features to releases), comparison functions, current and predicted data, simple and compound measurements, simple and compound decisions.

Necessity of having the right and available measurements is an important finding. Setting up the right environment for collecting right measurements before running the method is a need. Also the proposed method is appropriate when product is mature enough to collect measurements from the user's behavior. The analytics-based method covers both reactive and proactive planning.

By dissecting the processes in the bespoke product planning and finding out the decision structures, close attentions were paid to involve analytics in the analytics-based planning method. It enhances a product manager's intuitions and helps him to find out the rationales in his decisions and communicate them better.

In the bespoke product roadmapping, market and business drivers are specified. Product or service vision is outlined and their features are identified and prioritized. Technology solutions are defined and releases are decided based on available resources and timeframes. Finally the roadmap is created. Analytics-based method can be integrated in the steps to provide recommendation for the involved decisions by collecting the effective corresponding measurements. The more measurements are applied for analyzing a feature/decision, the more decrease in misinterpretation would be gained. However, The SaaS-based analytics might not be sufficient. There are activities of product planning that SaaS-based measurements cannot provide sufficient information for them while business analytics might provide more supportive data.

To create, remove or enhance a feature, analytics provides a wide view of data changes to transform them to a recommendation for planning. Prioritizing features is occurred by comparing measurements' impacts for the corresponding decisions, when the features are in the same context. product criterion is just one factor to accept the analytics-based recommendation for decision making. Final solution is achieved after trade off between product criterion (achieved in using the analytics-based method) and other criteria (e.g. resources, market, competitors).

8.1 Limitations of the Study

Through the study, some limitations have been encountered that may influence the study results. One limitation is related to scope of the thesis that made us to narrow it down:

Interviewees of the study had chosen roadmapping as the most involved planning asset, which made us to concentrate on the effect of analytics on roadmapping and prevented us from discussing the results in the portfolio management and release planning. Interviewing with product managers who were usually busy enough, and the limitation of having the interview duration for less than one hour, obliged us to focus on just road-mapping as one of core assets in the product planning.

This study concentrates on web analytics that related measurements were collected with the purpose of understanding and optimizing web usage in a SaaS-based application. Web analytics lack business data and we excluded the study of feature decisions with the confidential data of the organization, as we had no access to them in the case study. So this study was not able to support using analytics for market trends and financial benefits.

The responses from 17 interviews couldn't be considered a large amount of data to be classified and mined. More responses could provide more accurate scores of measurements for planning and results of the dependency relationship between product types and analytics would be more reliable.

There are some limitations subject to the case study. Limitations of the analytical tool and collecting suitable measurements made us ignore addressing some decisions in the proposed method, as we couldn't access the required data. The Piwik tool couldn't collect data about all features (e.g. it doesn't support client-side actions such as tabs), and some features could be supported by application measurements inside the Madek's database. There were some situations that the related measurements were not defined for the product. For instance no campaign were defined for the product, so no data were collected for the related measurement. Another limitation in our study comes from the limited duration of collecting data (since June 2012). It prevented us to access large amount of data to study data trends for the organization or compare current data with similar time points in the previous year. Also in this duration, we didn't recognized major updates to map related decision-making process to changes of measurement values in order to study the changes of data after applying the decision.

The last limitation is about real evaluation of the feedback component in the proposed method. The pre-requisite for this component is implementing the features when deciding about. Limited time for the thesis prevented us to test it practically although it was tested by unreal feedback from the product manager of the case.

8.2 Future Work

In this section the following recommendations for future work are made:

- One recommendation for future is to extend this research and support the effect of analytics on portfolio management and release planning as well. The scope of this study was limited to investigating the effect of analytics on roadmap planning. Although the proposed method has been generalized, more detailed study about portfolio management and release planning may achieve different results specifically. Conducting an experiment will provide insight into causes and effects, also it presents how decision results are changed when measurements are manipulated. Also this study is confined under limitations to access some measurements, which were unavailable or could be related to confidential business data. The data related to financial aspects, campaigns, subscriptions and user feedback is recommended to be included in a future study.
- Another suggestion for future study is to understand situational factors that affect importance level of measurements for planning. During the interview in our qualitative study, we noticed some interviewees had chosen the importance of measurements based on some factors. Product goal (i.e. customer-centric, user-centric market-centric), and product characteristic (i.e. different product with different customer type, access type, network type) and product maturity can be mentioned as their selection justifications. Although the general categorization of product type did not indicate any difference in rating the measurements but interviewees' justifications may contradict this statement. The inconsistency between results recommends including this study in a further research.
- Furthermore, developing a bundle measurements library is recommended. During the case study we understood that for different types of features, different measurements were involved. Providing a bundle measurements library that specifies a list of important measurements for specific instance features will create added value for a future study.
- Finally, more advanced research will provide more knowledge to select the most suitable comparison function for the proposed method based on the product type and effective measurements. Studying on the suitable functions will increase the accuracy of indicators for making a decision.

9. References

[1] I. van de Weerd, S. Brinkkemper, R. Nieuwenhuis, J. Versendaal, and L. Bijlsma, “Towards a reference framework for software product management,” in *14th IEEE International Conference on Requirements Engineering*, 2006, pp. 319–322.

[2] C. Ebert, “The impacts of software product management,” *Journal of Systems and Software*, vol. 80, no. 6, pp. 850–861, 2007.

[3] S. A. Fricker, “Software Product Management,” in *Software for People*, A. Maedche, A. Botzenhardt, and L. Neer, Eds. Springer Berlin Heidelberg, 2012, pp. 53–81.

[4] M. A. Cusumano, “The changing software business: Moving from products to services,” *Computer*, vol. 41, no. 1, pp. 20–27, 2008.

[5] V. Choudhary, “Software as a service: Implications for investment in software development,” in *40th Annual International Conference on System Sciences (HICSS 2007)*, 2007, p. 209a–209a.

[6] H. B. Kittlaus and P. N. Clough, *Software Product Management and Pricing: Key Success Factors for Software Organizations*. Springer-Verlag New York Inc, 2009.

[7] S. Barney, A. Aurum, and C. Wohlin, “A product management challenge: Creating software product value through requirements selection,” *Journal of Systems Architecture*, vol. 54, no. 6, pp. 576–593, 2008.

[8] M. Svahnberg, T. Gorschek, R. Feldt, R. Torkar, S. B. Saleem, and M. U. Shafique, “A systematic review on strategic release planning models,” *Information and Software Technology*, vol. 52, no. 3, pp. 237–248, 2010.

[9] A. S. Danesh and R. Ahmad, “Software release planning challenges in software development: An empirical study,” *African Journal of Business Management*, vol. 6, no. 3, pp. 956–970, 2012.

[10] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, “A design science research methodology for information systems research,” *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, 2007.

[11] K. B. Kahn, *Product Planning Essentials*. M.E. Sharpe, 2011.

[12] “SPMBOK « International Software Product Management Association.” [Online]. Available: http://ispma.org/spmbok/. [Accessed: 12-Mar-2012].

[13] A. Croll and S. Power, *Complete web monitoring*. O’Reilly Media, 2009.

[14] R. G. Cooper, S. J. Edgett, and E. J. Kleinschmidt, “New product portfolio management: practices and performance,” *Journal of product innovation management*, vol. 16, no. 4, pp. 333–351, 1999.

[15] S. Fricker and S. Schumacher, “Release Planning with Feature Trees: Industrial Case,” in *Requirements Engineering: Foundation for Software Quality*, vol. 7195, B. Regnell and D. Damian, Eds. Springer Berlin / Heidelberg, 2012, pp. 288–305.

[16] J. Vähäniitty, C. Lassenius, and K. Rautiainen, “An approach to product roadmapping in small software product businesses,” presented at the Quality Connection - 7th European Conference on Software Quality (ECSQ2), 2002, pp. 12–13.

[17] L. Lehtola, M. Kauppinen, and S. Kujala, “Linking the business view to requirements engineering: long-term product planning by roadmapping,” in *13th*

IEEE International Conference on Requirements Engineering, 2005, pp. 439–443.
[18] G. Du, J. McElroy, and G. Ruhe, “Ad hoc versus systematic planning of software releases–a three-staged experiment,” *Product-Focused Software Process Improvement*, pp. 435–440, 2006.
[19] G. C. van de Weerd, “Advancing in software product management: An incremental method engineering approach,” *SIKS Disseration Series*, no. 2009–34, 2009.
[20] C. Manteli, I. van de Weerd, and S. Brinkkemper, “An empirical research on the relationships between software product management and software project management,” *SciencesNew York*, vol. January, p. 81, 2010.
[21] I. van de Weerd, S. Brinkkemper, and J. Versendaal, “Incremental method evolution in global software product management: A retrospective case study,” *Information and Software Technology*, vol. 52, no. 7, pp. 720–732, 2010.
[22] B. A. Nejmeh and I. Thomas, “Business-driven product planning using feature vectors and increments,” *Software, IEEE*, vol. 19, no. 6, pp. 34–42, 2002.
[23] I. Weerd, W. Bekkers, and S. Brinkkemper, “Developing a maturity matrix for software product management,” *Software Business*, pp. 76–89, 2010.
[24] J. Momoh and G. Ruhe, “Release planning process improvement—an industrial case study,” *Software Process: Improvement and Practice*, vol. 11, no. 3, pp. 295–307, 2006.
[25] G. Muller, “Roadmapping,” Phillips research, Eindhoven, The Netherlands, 1999.
[26] J. Jiao and Y. Zhang, “Product portfolio planning with customer-engineering interaction,” *Iie Transactions*, vol. 37, no. 9, pp. 801–814, 2005.
[27] O. Saliu and G. Ruhe, “Supporting software release planning decisions for evolving systems,” in *29th Annual IEEE/NASA Workshop on Software Engineering*, 2005, pp. 14–26.
[28] E. Bjarnason, K. Wnuk, and B. Regnell, “Overscoping: Reasons and consequences—A case study on decision making in software product management,” in *Fourth International Workshop on Software Product Management (IWSPM).*, 2010, pp. 30–39.
[29] A. Ngo-The and G. Ruhe, “A systematic approach for solving the wicked problem of software release planning,” *Soft Computing-A Fusion of Foundations, Methodologies and Applications*, vol. 12, no. 1, pp. 95–108, 2008.
[30] A. Tang, T. Boer, and H. Vlient, “Building roadmaps: a knowledge sharing perspective,” in *Sixth Workshop SHAring and Reusing architectural Knowledge*, New York, NY, USA, 2011, pp. 13–20.
[31] V. Heikkila, A. Jadallah, K. Rautiainen, and G. Ruhe, “Rigorous support for flexible planning of product releases-a stakeholder-centric approach and its initial evaluation,” in *43rd Hawaii International Conference onSystem Sciences (HICSS)*, University fo Hawai, Manoa, Honolulu, Hawai, 2010, pp. 1–10.
[32] M. L. Garcia and O. H. Bray, “Fundamentals of technology roadmapping,” *Sandia Nat. Laboratories, Albuquerque, New Mexico*, 1997.
[33] S. Lee, S. Kang, Y. S. Park, and Y. Park, “Technology Roadmapping for R&D planning: The case of the Korean parts and materials industry,” *Technovation*, vol. 27, no. 8, pp. 433–445, 2007.
[34] S. Fricker and S. Schumacher, “Variability-Based Release Planning,” *Software Business*, vol. 80, pp. 181–186, 2011.

[35] B. Regnell, L. Karlsson, and M. Host, "An analytical model for requirements selection quality evaluation in product software development," in *11th IEEE International Conference on Requirements Engineering*, 2003, pp. 254–263.
[36] A. Al-Emran and D. Pfahl, "Operational planning, re-planning and risk analysis for software releases," *Product-Focused Software Process Improvement*, pp. 315–329, 2007.
[37] L. Karlsson and B. Regnell, "Introducing tool support for retrospective analysis of release planning decisions," *Product-Focused Software Process Improvement*, vol. 4034, pp. 19–33, 2006.
[38] K. Wnuk, R. B. Svensson, and B. Regnell, "Investigating Upstream versus Downstream Decision-Making in Software Product Management," in *Third International Workshop on Software Product Management (IWSPM).*, 2009, pp. 23–26.
[39] B. Regnell and K. Kuchcinski, "Exploring Software Product Management decision problems with constraint solving-opportunities for prioritization and release planning," in *Fifth International Workshop on Software Product Management (IWSPM).*, 2011, pp. 47–56.
[40] T. AlBourae, G. Ruhe, and M. Moussavi, "Lightweight Replanning of Software Product Releases," in *International Workshop on Software Product Management*, 2006.
[41] H. C. Benestad and J. E. Hannay, "A comparison of model-based and judgment-based release planning in incremental software projects," in *33rd International Conference on Software Engineering*, 2011, pp. 766–775.
[42] G. Ruhe and M. O. Saliu, "The art and science of software release planning," *Software, IEEE*, vol. 22, no. 6, pp. 47–53, 2005.
[43] T. AlBourae, G. Ruhe, and M. Moussavi, "Lightweight replanning of software product releases," in *International Workshop on Software Product Management (IWSPM'06).*, 2006, pp. 27–34.
[44] R. Phaal and G. Muller, "An architectural framework for roadmapping: Towards visual strategy," *Technological Forecasting and Social Change*, vol. 76, no. 1, pp. 39–49, 2009.
[45] R. Phaal, C. J. P. Farrukh, and D. R. Probert, "Fast-start technology roadmapping," in *9th International Conference on Management of Technology (IAMOT 2000)*, 2000, pp. 21–25.
[46] R. Phaal, C. J. P. Farrukh, and D. R. Probert, "Technology roadmapping–A planning framework for evolution and revolution," *Technological forecasting and social change*, vol. 71, no. 1–2, pp. 5–26, 2004.
[47] T. Gorschek, S. Fricker, K. Palm, and S. Kunsman, "A lightweight innovation process for software-intensive product development," *IEEE software*, vol. 27, no. 1, pp. 37–45, 2010.
[48] J. Karlsson and K. Ryan, "A cost-value approach for prioritizing requirements," *Software, IEEE*, vol. 14, no. 5, pp. 67–74, 1997.
[49] G. Ruhe, A. Eberlein, and D. Pfahl, "Trade-off analysis for requirements selection," *International Journal of Software Engineering and Knowledge Engineering*, vol. 13, no. 04, pp. 345–366, 2003.
[50] T. H. Davenport and J. G. Harris, *Competing on Analytics: The New Science of Winning*. Harvard Business Press, 2007.
[51] D. Zhang, Y. Dang, J.-G. Lou, S. Han, H. Zhang, and T. Xie, "Software analytics as a learning case in practice: Approaches and experiences," in

International Workshop on Machine Learning Technologies in Software Engineering, 2011, pp. 55–58.
[52] R. P. Buse and T. Zimmermann, "Analytics for software development," in *FSE/SDP workshop on Future of software engineering research*, 2010, pp. 77–80.
[53] R. P. L. Buse and T. Zimmermann, "Information needs for software development analytics," in *34th International Conference on Software Engineering (ICSE)*, 2012, pp. 987–996.
[54] T. H. Davenport, J. G. Harris, and R. Morison, *Analytics at work: smarter decisions, better results*. Harvard Business Press, 2010.
[55] E. T. Peterson, *Web analytics demystified: a marketer's guide to understanding how your web site affects your business*. Celilo Group Media, 2004.
[56] M. Godse and S. Mulik, "An approach for selecting software-as-a-service (SaaS) product," in *IEEE International Conference on Cloud Computing (CLOUD'09)*, 2009, pp. 155–158.
[57] R. Kohavi, N. J. Rothleder, and E. Simoudis, "Emerging trends in business analytics," *Communications of the ACM*, vol. 45, no. 8, pp. 45–48, 2002.
[58] A. Kaushik, *Web Analytics: An Hour A Day (W/Cd)*. Wiley-India, 2007.
[59] R. Bose, "Advanced analytics: opportunities and challenges," *Industrial Management & Data Systems*, vol. 109, no. 2, pp. 155–172, 2009.
[60] A. Kaushik, *Web analytics 2.0: the art of online accountability & science of customer centricity*. Sybex, 2009.
[61] B. Clifton, *Advanced web metrics with Google Analytics*. Sybex, 2012.
[62] J. Burby and S. Atchison, "Actionable Web Analytics," *Sybex, Indianapolis*, 2007.
[63] K. P. Shung and M. C. Junyu, "Application of Analytics in Business Strategy," *Business Intelligence Journal*, vol. 5, no. 1, pp. 190–193, 2012.
[64] M. Kuniavsky, *Observing the user experience: a practitioner's guide to user research*. Morgan kaufmann, 2003.
[65] A. Holzinger, "Usability engineering methods for software developers," *Communications of the ACM*, vol. 48, no. 1, pp. 71–74, 2005.
[66] E. Dixon, E. Enos, and S. Brodmerkle, "A/B Testing," U.S. Patent App. 13/111,505May-2011.
[67] M. Houtermans, T. V. Capelle, and M. Al-Ghumgham, "Reliability Engineering & Data Collection," in *Second International Conference on Systems (ICONS'07)*, 2007, pp. 42–42.
[68] S. Banerjee, H. Srikanth, and B. Cukic, "Log-based reliability analysis of software as a service (saas)," in *Software Reliability Engineering (ISSRE), 2010 IEEE 21st International Symposium on*, 2010, pp. 239–248.
[69] D. A. Menasce, "QoS issues in Web services," *Internet Computing, IEEE*, vol. 6, no. 6, pp. 72–75, 2002.
[70] Y. Liu, A. H. Ngu, and L. Z. Zeng, "QoS computation and policing in dynamic web service selection," in *13th international World Wide Web conference on Alternate track papers & posters*, 2004, pp. 66–73.
[71] R. Rajkumar, C. Lee, J. Lehoczky, and D. Siewiorek, "A resource allocation model for QoS management," in *The 18th IEEE Real-Time Systems Symposium.*, 1997, pp. 298 –307.

[72] J. Shaikh, M. Fiedler, and D. Collange, “Quality of Experience from user and network perspectives,” *Annals of Telecommunications*, vol. 65, no. 1, pp. 47–57, 2010.
[73] M. Woodside, G. Franks, and D. C. Petriu, “The Future of Software Performance Engineering,” in *Future of Software Engineering (FOSE ’07).*, 2007, pp. 171 –187.
[74] R. Buyya, C. S. Yeo, and S. Venugopal, “Market-oriented cloud computing: Vision, hype, and reality for delivering it services as computing utilities,” in *10th IEEE International Conference on High Performance Computing and Communications (HPCC’08)*, 2008, pp. 5–13.
[75] P. Mell and T. Grance, “The NIST definition of cloud computing (draft),” *NIST special publication*, vol. 800, p. 145, 2011.
[76] L. Wang, G. Von Laszewski, A. Younge, X. He, M. Kunze, J. Tao, and C. Fu, “Cloud computing: a perspective study,” *New Generation Computing*, vol. 28, no. 2, pp. 137–146, 2010.
[77] S. L. Saini, D. K. Saini, J. H. Yousif, and S. V. Khandage, “Cloud Computing and Enterprise Resource Planning Systems,” in *The World Congress on Engineering.*, 2011, pp. 681–684.
[78] C. Weinhardt, A. Anandasivam, B. Blau, N. Borissov, T. Meinl, W. Michalk, and J. Stö\sser, “Cloud computing–a classification, business models, and research directions,” *Business & Information Systems Engineering*, vol. 1, no. 5, pp. 391–399, 2009.
[79] G. Verma, “Software as a Service,” *Journal of Information Technology.*, vol. 9, no. 1, p. 32, 2011.
[80] L. Gideon, *Handbook of Survey Methodology for the Social Sciences*. Springer, 2012.
[81] A. R. Hevner, S. T. March, J. Park, and S. Ram, “Design science in information systems research,” *Mis Quarterly*, vol. 28, no. 1, pp. 75–105, 2004.
[82] R. K. Yin, *Case study research: Design and methods*, vol. 5. Sage Publications, Incorporated, 2008.
[83] L. M. Rea and R. A. Parker, *Designing and conducting survey research: A comprehensive guide*, 3rd ed. Jossey-Bass, 2005.
[84] G. Guest, A. Bunce, and L. Johnson, “How Many Interviews Are Enough? An Experiment with Data Saturation and Variability,” *Field Methods*, vol. 18, no. 1, pp. 59–82, Feb. 2006.
[85] K. Petersen and C. Wohlin, “Context in industrial software engineering research,” in *3rd International Symposium on Empirical Software Engineering and Measurement*, 2009, pp. 401–404.
[86] H. F. Hsieh and S. E. Shannon, “Three approaches to qualitative content analysis,” *Qualitative health research*, vol. 15, no. 9, pp. 1277–1288, 2005.
[87] “Medienarchiv der Künste.” [Online]. Available: http://medienarchiv.zhdk.ch/. [Accessed: 09-Sep-2012].
[88] “Piwik - Web analytics - Open source.” [Online]. Available: http://piwik.org/. [Accessed: 09-Sep-2012].
[89] “New Relic : Web Application Performance Management (APM).” [Online]. Available: https://newrelic.com/. [Accessed: 09-Sep-2012].
[90] A. Forward and T. C. Lethbridge, “A taxonomy of software types to facilitate search and evidence-based software engineering,” in *Conference of the center*

for advanced studies on collaborative research: meeting of minds, New York, NY, USA, 2008, pp. 14:179–14:191.

[91] J. H. Gray, I. L. Densten, and J. C. Sarros, “Size matters: organisational culture in small, medium and large Australian organisations,” *Journal of Small Business and Entrepreneurship*, vol. 17, no. 1, pp. 31–46, 2003.

[92] A. Agresti, *An introduction to categorical data analysis*, Second Edition. Wiley-Interscience, 2007.

[93] W. H. Kruskal and W. A. Wallis, “Use of ranks in one-criterion variance analysis,” *Journal of the American statistical Association*, vol. 47, no. 260, pp. 583–621, 1952.

[94] G. W. Corder and D. I. Foreman, *Nonparametric statistics for non-statisticians: a step-by-step approach*. Wiley, 2009.

[95] K. Backman and H. A. Kyngäs, “Challenges of the grounded theory approach to a novice researcher,” *Nursing & Health Sciences*, vol. 1, no. 3, pp. 147–153, 1999.

[96] M. Sandelowski, “Combining Qualitative and Quantitative Sampling, Data Collection, and Analysis Techniques in Mixed-Method Studies,” *Research in nursing & health*, vol. 23, no. 3, pp. 246–255, 2000.

[97] C. P. Smith, “Content analysis and narrative analysis,” *Handbook of research methods in social and personality psychology*, pp. 313–335, 2000.

[98] S. Elo and H. Kyngäs, “The qualitative content analysis process,” *Journal of advanced nursing*, vol. 62, no. 1, pp. 107–115, 2008.

[99] W. Bekkers, I. van de Weerd, S. Brinkkemper, and A. Mahieu, “The influence of situational factors in software product management: an empirical study,” in *Second International Workshop on Software Product Management (IWSPM’08).*, 2008, pp. 41–48.

[100] C. Robson, *Real world research: a resource for social scientists and practitioner-researchers*, vol. 2. Blackwell Oxford, 2002.

[101] D. S. Cruzes and T. Dyb\aa, “Research synthesis in software engineering: A tertiary study,” *Information and Software Technology*, vol. 53, no. 5, pp. 440–455, 2011.

[102] K. L. Easton, J. F. McComish, and R. Greenberg, “Avoiding common pitfalls in qualitative data collection and transcription,” *Qualitative Health Research*, vol. 10, no. 5, pp. 703–707, 2000.

[103] J. Kim and Y. Park, “Identifying a new service opportunity from potential needs: User-centric service map,” in *2010 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)*, 2010, pp. 357–361.

[104] B. Kitchenham, S. Linkman, and D. Law, “DESMET: a methodology for evaluating software engineering methods and tools,” *Computing & Control Engineering Journal*, vol. 8, no. 3, pp. 120–126, 1997.

[105] N. Golafshani, “Understanding reliability and validity in qualitative research,” *The qualitative report*, vol. 8, no. 4, pp. 597–607, 2003.

[106] A. Maglyas, U. Nikula, and K. Smolander, “What do we know about software product management?-a systematic mapping study,” in *Software Product Management (IWSPM), 2011 Fifth International Workshop on*, 2011, pp. 26–35.

[107] A. Phippen, L. Sheppard, and S. Furnell, “A practical evaluation of Web analytics,” *Internet Research*, vol. 14, no. 4, pp. 284–293, 2004.

[108] B. A. Kitchenham and S. Charters, "Guidelines for performing systematic literature reviews in software engineering," 2007.
[109] J. W. Creswell, *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications, 2009.
[110] L. A. Goodman, "Snowball Sampling," *The Annals of Mathematical Statistics*, vol. 32, no. 1, pp. 148–170, Mar. 1961.

APPENDIXES

Appendix A

Appendix A.1

Measurement Interpretations

The following table provides one sample of measurements interpretation:

Table 16: Measurements Interpretations

Measurement-attributes	**Interpretation**
Product use	Statistics about times that a product has been used in a time period
Overall amount of users	Statistics about users that use a product in a time period
Time between visits	Statistics about time between different visits done by product users in a time period
Duration of using the product	Statistics about duration that users used a product in a time period
New users	Statistics about users with first visit from a product in a time period
Returning users	Statistics about users that return to a product and are not new users in a time period
Users that use a feature	Statistics about users that use particular feature in a time period
Feature use	Statistics about times that features are used in the product in a time period
Duration of using a feature	Statistics about duration that a feature has been used in a time period
Entrance feature	Statistics about times that a particular feature served as an entrance to the online product in a time period
Exit feature	Statistics about times that particular feature was the last one viewed by users in a time period
Bounce	Statistics about single feature that has entrance and exit by a particular user and afterward the user won't return to the product
Click activity	Statistics about click on a particular feature done by users in a time period
Depth of use	Statistics about particular number of features that are used by users in a time period
Click stream/path	Statistics about particular sequence of features that are used by users in a time period
Referrers	Statistics about sources that users are directed to a product in a time period
Location/ISP per use	Statistics about geographical locations of each product use in a time period
Search engines and keywords	Statistics about search engine that users use and the keywords that they enter in a time period
Campaigns	Statistics about how efficient various marketing campaigns are in bringing visitors to a product in a time period
Languages	Statistics about particular language that users use in a time period
Browsers	Statistics about particular browser that users use in a time period
Operating Systems	Statistics about particular operating system that users use in a time period
Plugins	Statistics about particular plugins that users use in a time period
Screen resolution	Statistics about particular screen resolution that users use in a time period
Errors	Statistics about particular errors that users face with in a time period
Downtime	Statistics about period of time that the product is not in operation
Response time	Statistics about length of time taken by system to respond to a request
Throughput	Statistics about the amount of work that can be performed in a given time period
DOS attack	Statistics about the number of attempts that make a product unavailable for users
Worm attacks	Statistics about the number of worm attempts that make a product unavailable for users

Appendix A.2

Literature review Execution

Literature review was conducted to identify taxonomies of software product planning decisions and measurements in a SaaS application. This appendix shows how the literature review for the above results was executed.

Product Planning Decisions - Literature Review Execution

The literature review for the study could be selected among systematic literature review, systematic mapping and tertiary review. This study didn't have strong needs for in-depth studies revision, synthesis of data and aggregating evidence. Also it had a limitation about time allocation for the study, as the literature review was not the main research study. Therefore, tertiary review was considered as the most suitable technique. It requires fewer resources and is appropriate for providing an overview, background and related work which were the inputs for our research questions [108]. Tertiary review is suitable if existing evidence derives from high quality literature and extracting data is straightforward [108]. Therefore the literature review was fulfilled in 7steps as follows [109] and one additional step for snowball sampling [110] **:**

Step1- As the first step, keywords were identified based on the goal of the research question RQ1. More keywords were identified after preliminary study of research papers. Search strings were constructed by important keywords which included "product planning", "software" and "decision making", and core assets of product planning (portfolio management, roadmapping, and release planning) as essential parts.

Step2- IEEE Xplore, Engineering Village and Google Scholar were selected as search databases. The search strings were searched through the abstracts and also titles, while meta data was avoided because the keywords sometimes matched with the name of conferences, completely irrelevant with the question.

Step3- Eliminating the irrelevant search results based on exclusive criteria, which was done automatically by the search engine. Search results were limited to journal articles, conference papers, and workshop papers. The papers that were selected were those that peer reviewed and were only in English. Availability of full texts was a primary criterion for selecting documents. Publication was filtered based on software product planning, product management, and requirement engineering to be close to the research area.

Step4- The search results were examined with respect to the title relevancy.

Step5- Zotero was used as a reference management for categorizing relevant search results. Search results were examined through abstracts and conclusions to identify their content relevancy.

Step6- Selected papers were browsed through the content, discussion and conclusion. Final decision was made based on their content relevancy and mentioning at least one planning decision. Table 17 presents search strings and number of results in each database with total unique results of 29. As it can be seen in Table 17, through the databases IEEE Xplore, Engineering village and Google scholar 8, 5 and 26 related documents were selected respectively. As the total number of 39 documents (8+5+26) contained some duplicated documents, the final results decreased to total number of 29 unique documents after removing the duplications.

Step7- The decisions founded in the selected papers were summarized in an Excel file. Then categorization was applied to avoid similarity between results. All the results were organized and incorporated in the decision taxonomy of product planning.

Step8- Snowball sampling was performed like chain-referral for documents that extracted from Step6. All 29 documents, which were resulted from step 7 were deeply studied and analyzed. From them, 14 documents were selected as the seeds for snowball, which were generally fundamental in SPM and widely cited or the authors were active in SPM and

ISPMA (International Software Product Management Association). The snowball sampling was performed in one wave and for each reference list of results (seeds and first-wave results) step 3 to step 7 were repeated to find out most relevant articles. Due to large number of references for the seeds, one wave was recognized sufficient for the study. From 14 articles as the seeds, 215 articles were resulted. By performing step 3 to step 7, 4 more documents were added to the results of keyword search (29 documents) and then the final result reached total unique documents of 33.

Table 17: Literature review results for planning decisions

Key words	Database	Step 3	Step 4	Step 5	Steps 6 and 7
(("Document Title":((software) AND (("product management")OR("product planning")OR ("product plan")OR("product manager")OR("roadmapping")OR("roadmap")OR ("release planning")OR("portfolio management").RB&LB.("decision") OR(decision-making)OR("decision making")))) OR "Abstract":((software) AND (("product management")OR("product planning")OR ("product plan")OR("product manager")OR("roadmapping")OR("roadmap")OR ("release planning")OR("portfolio management").RB&LB.("decision") OR(decision-making)OR("decision making"))))	IEEEXplore	302	27	12	8
((software) AND (("product management") OR ("product planning") OR ("product plan") OR ("product manager") OR ("roadmapping") OR ("roadmap") OR ("release planning") OR ("portfolio management")) AND (("decision") OR (decision-making) OR ("decision making")))	Engineering Village	289	40	15	5
(("software product management") OR ("software product planning") OR("software product plan")) AND ("decisions")	Google Scholar	532	81	39	26
Total number of documents:	33				

The above steps extracted a list of product planning decisions, which were categorized through a development process based on a "loop" method. The development process [Figure 34] for finding the taxonomy started with the input of product planning decisions extracted from literature. First, each decision was allocated to its related category (Portfolio management, roadmapping or release planning). If it was a new decision, it was added to related category of the taxonomy if not, it was neglected. This process was repeated until the saturation was occurred for extracted decisions. At the end, the taxonomy was organized and finalized. All product planning decisions were extracted, organized and finalized in the taxonomy based on decisions for portfolio management, product roadmapping and release planning [Table 1].

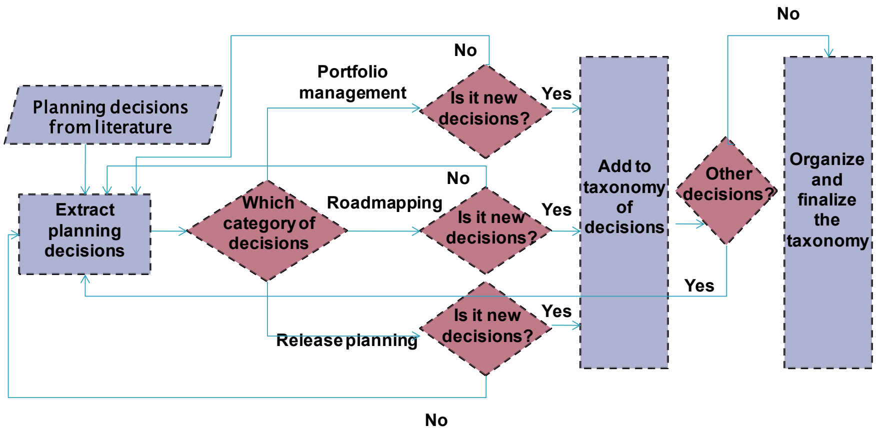


Figure 34: Development process for taxonomy of decisions

Analytics in Product Planning- Literature Review Execution

The similar literature review process was performed for identifying measured attributes that support product-planning decision with minor changes. Keywords were identified and different search strings were constructed by “SaaS”, “Software as a Service”, “analytics”, and “Product planning” keywords. Then a search strategy that is mentioned in previous section was performed. As it can be seen in Table 18, through the databases of IEEE Xplore, Engineering village and Google scholar, 4, 3 and 5 documents were selected respectively. As the total number of 12 documents (4+3+5) contained some duplicated documents, the final results decreased to total number of 7 unique documents after removing the duplications. Snowball sampling (Step 8) was not considered for analytics review, because extracted data that were achieved by the keyword search redirected us from identifying actual measurements and instead studying analytical tools were considered as an alternative research strategy. So Step8 from the search strategy was omitted. To prevent missing any measurement, another extra search strategy was considered. Analytical tools such as Google analytics, Piwik, Yahoo web analytics, StatCounter, New relics, and Woopra were installed and studied closely to extract measurements that can support decision-making.

Table 18: Literature review results for SaaS-based measurements

Key words	Database	Step 3	Step 4	Step 5	Step 6
("web analytics" AND ("SaaS" OR " Software as a Service" OR "saas" OR "software as a service")))	IEEEXplore	382	51	12	4
(((((((SaaS) WN KY) OR ((Software as a Service) WN KY)) OR ((saas) WN KY)) OR ((software as a service) WN KY)) OR ((software product) WN KY)) AND ((analytics) WN KY))	Engineering Village	134	40	9	3
("web analytics" AND ("SaaS" OR " Software as a Service" OR "saas" OR "software as a service")))	Google Scholar	240	91	33	5

For developing the taxonomy of the measurement-attributes a development process based on a “loop” method, was designed [Figure 35]. First, on the extracted measurement-attributes from literature and analytics-based applications, exclusion criteria were applied to filter. Then for each measurement-attribute, decision was made that if it was a new measurement-attribute, it was added to a new or previous category, if not it was neglected. Basically each category was formed when one or more measurement-attributes couldn't be assigned to available categories. This process was repeated for all extracted measurement-attributes. Then the taxonomy was finalized.

The collected attributes were filter according to the following criteria:

- SaaS-based simple measurement-attributes (i.e. For a combined attribute, their simple attributes have been selected)
- Is not confidential data of the organization (e.g. financial data),
- Quantitative data that are supported by the available analytical tools (e.g. exclude form-based measurements such as end-users feedback)
- No extra implementation is required by organizations to collect it (e.g. Customer subscriptions).

The performed development process is illustrated in Figure 35.

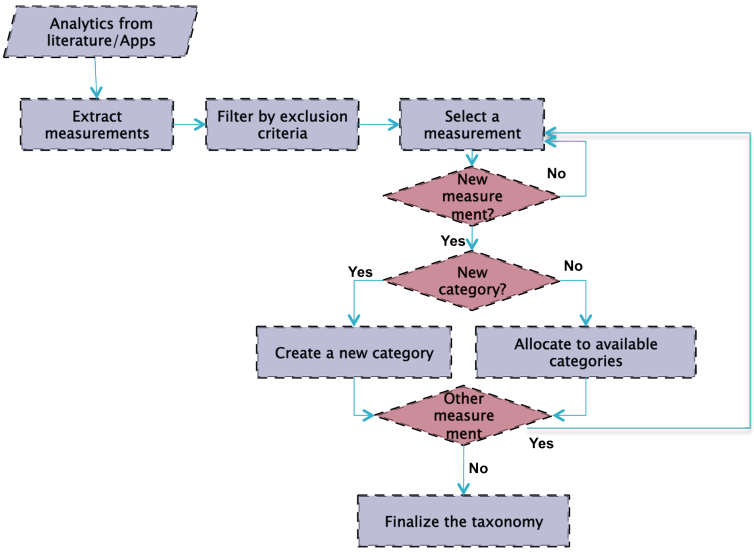


Figure 35: Development process for taxonomy of measurements

Appendix B

Appendix B.1

Interview-based Survey Questionnaire

Analytics-based Software Product Planning

Demographics-Decisions

Dear Practitioner,

While planning development of software products, product managers can support decisions about products in portfolio management, features in roadmapping and requirements in release planning. Online analytics are collected by monitoring of software use and can support the decision-making. Together with you, we would like to evaluate how analytics can affect planning decisions.

Who should complete the survey? Product managers and other professionals who are involved in product planning.

What questions are asked? Questions about one of your product planning experiences and your opinions about possible use of analytics. It will only take you approximately 10 minutes to complete the survey. You can include your contact information at the end of the survey, in the case of your interest in receiving survey results.

What should you do if you have any question related the survey? In the case that you have any question, please do not hesitate to contact us:

Persons in charge:

Farnaz Fotrousi (Fafo10@student.bth.se)

Katayoun Izadyan (Kaia10@student.bth.se)

Supervisor:

Dr. Samuel A. Fricker (samuel.fricker@bth.se)

Thanks in advance

Demographics Questions:

In the following question we would like to be familiar with your previous experiences in software product planning.

1. For how many years have you been involved in such decision making?

2. What is the size of your current organization?

- Under 25 employees
- 25 to 250 employees
- More than 250 employees

Software Product Planning questions:

In the following questions, think about a product that you were planning during last 2 years and you are most satisfied with.

3. What was your role(s)?

4. What is the type of software product that you have selected?

5. Is it a new product or an evolutionary one?

6. How many persons were involved in development of the product?

7. Which kind of decision making were you mostly involved? *

- Portfolio management
- Product roadmapping
- Release planning

Analytics-based Software Product Planning

Online Feedback

Categories of analytics

In the following questions, we expect to find out the category of analytics that are helpful for decisions such as those you have indicated before. With analytics, you can monitor product use and collect information from an online prototype. Please rank the importance of information for the following decisions:

How important is the following information to decide about <u>creating a new feature for the current product</u>?

	No Idea	Not important	Less important	Important	Very Important
Value of the product from user's perspective (i.e. statistics about product use, overall number of users, duration of using the product) *	○	○	○	○	○
Value of a feature from user's perspective (i.e. statistics about features use, duration of using a feature) *	○	○	○	○	○
Usage pattern of the product (i.e. statistics about depth of use, click activities) *	○	○	○	○	○
Referral sources for product use (i.e. statistics about traffic source location/ISP) *	○	○	○	○	○
Technologies and channels used to access the product (i.e. operating system, browser) *	○	○	○	○	○
Product healthiness (i.e. reliability, performance and security) *	○	○	○	○	○

How important is the following information to decide about <u>enhancing a feature(s) in the current product</u>?

	No Idea	Not important	Less important	Important	Very Important
Value of the product from user's perspective (i.e. statistics about product use, overall number of users, duration of using the product) *	○	○	○	○	○
Value of a feature from user's perspective (i.e. statistics about features use, duration of using a feature) *	○	○	○	○	○
Usage pattern of the product (i.e. statistics about depth of use, click activities) *	○	○	○	○	○
Referral sources for product use (i.e. statistics about traffic source location/ISP) *	○	○	○	○	○
Technologies and channels used to access the product (i.e. operating system, browser) *	○	○	○	○	○
Product healthiness (i.e. reliability, performance and security) *	○	○	○	○	○

Analytics-based Software Product Planning

Analytic

Analytics

In this section, we expect to understand the importance of analytics for the product. Please rank the analytics based on their importance and move them to right-hand list. You can leave unimportant ones on the left-hand list.

Please rank the analytics that present the value of product from user's perspective.

Drag items from the left-hand list into the right-hand list to order them.

- Statistics about product use
- Statistics about overall amount of users
- Statistics about time between visits
- Statistics about duration of using the product
- Statistics about new users
- Statistics about returning users

Please rank the analytics that present the value of the feature from user's perspective.

Drag items from the left-hand list into the right-hand list to order them.

- Statistics about users that use a feature
- Statistics about feature use (Total number of times that features are used in the product)
- Statistics about duration of using a feature
- Statistics about entrance feature (The number of times a particular feature served as an entrance to the online product)
- Statistics about exit feature (The number of times that particular feature was the last one viewed by users)
- Statistics about bounce (Number of single feature visits of the product)

Please rank the analytics that present usage pattern of the product.

Drag items from the left-hand list into the right-hand list to order them.

- Statistics about click activity (i.e. hyperlink, share, like)
- Depth of use (Number of features, users viewed)
- Click stream/path (The sequence of clicking features in the product)

Please rank the analytic that present referral sources of the product.

Drag items from the left-hand list into the right-hand list to order them.

- Statistics about referrers (Sources that users are directed to the product)
- Statistics about location/ISP per use (Geographical locations of each product use)
- Statistics about campaigns (how efficient various marketing campaigns are in bringing visitors to your website)
- Statistics about search engines and keywords

Please rank the analytic that present <u>technologies and channels used to access the product.</u>.

Drag items from the left-hand list into the right-hand list to order them.

- Languages
- Browsers
- Operating Systems
- Plugins
- Screen resolution

Please rank the analytic that present <u>product healthiness.</u>

Drag items from the left-hand list into the right-hand list to order them.

- Statistics about errors (i.e. http error)
- Downtime (The period of time that the product is not in operation)
- Response time (The length of time taken by system to respond to a request)
- Throughput (The amount of work can be performed in a given time period)
- Statistics about DOS attacks (Number of attempts that make a product unavailable for users)
- Statistics about worm attacks

Analytics-based Software Product Planning

Comment - Contact

General Feedback

Do you think of any other analytics that can help a product manager in decision-making or any comments that can help us in improving our survey? please let us know.

9. If you would like to receive the survey results, please provide your contact information.

Name

Email

Analytics-based Software Product Planning

Thank You!

Thank you for taking our survey. Your response is very important to us.

Appendix B.2

Appendix B.2.1

Quantitative Hypotheses

Table 19: Quantitative test hypotheses

Hypotheses
H0: There is no difference between each measurement-category value for different product planning decisions. $H_1 0$: There is no difference between the values of “Product value” measurement-category value for different product planning decisions. $H_2 0$: There is no difference between values of “Value of a feature from users’ perspective” measurement-category value for different product planning decisions. $H_3 0$: There is no difference between values of “Usage pattern” measurement-category value for different product planning decisions. $H_4 0$: There is no difference between values of “Referral sources” measurement-category value for different product planning decisions. $H_5 0$: There is no difference between importance of “Technologies and channels” measurement-category value for different product planning decisions. $H_6 0$: There is no difference between values of “Product healthiness” measurement-category value for different product planning decisions.

Appendix B.2.2

Normality Test Results

Normality test of measurement-category for selected decisions

Table 20: Normality test for "Create a new feature for the current product" decision

Selected_Decision = Create a new feature for the current product [b]

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.330	5	.079	.735	5	.021
Feature	.330	5	.079	.735	5	.021
Usage pattern	.221	5	.200*	.902	5	.421
Technologies and channels	.287	5	.200*	.914	5	.490
Product healthiness	.367	5	.026	.684	5	.006

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

b. Referralsourcesforproductuse is constant. It has been omitted.

Table 21: Normality test for "Enhance a feature in the current product" decision

Selected_Decision = Enhance a feature(s) in the current product

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.376	6	.008	.666	6	.003
Feature	.319	6	.056	.683	6	.004
Usage pattern	.319	6	.056	.683	6	.004
Referral sources	.202	6	.200*	.853	6	.167
Technologies and channels	.254	6	.200*	.866	6	.212
Product healthiness	.277	6	.168	.800	6	.059

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 22: Normality test for "Remove a feature from the current product" decision

Selected_Decision = Remove a feature from the current product

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.367	5	.026	.684	5	.006
Feature	.473	5	.001	.552	5	.000
Usage pattern	.300	5	.161	.833	5	.146
Referral sources	.300	5	.161	.883	5	.325
Technologies and channels	.221	5	.200*	.902	5	.421
Product healthiness	.473	5	.001	.552	5	.000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 23: Normality test for "Allocate features to releases" decisions

Selected_Decision = Allocate features to releases

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.371	8	.002	.724	8	.004
Feature	.391	8	.001	.641	8	.000
Usage pattern	.263	8	.109	.827	8	.056
Referral sources	.250	8	.150	.860	8	.120
Technologies and channels	.222	8	.200*	.912	8	.366
Product healthiness	.330	8	.010	.628	8	.000

*. This is a lower bound of the true significance.

Lilliefors Significance Correction

Table 24: Normality test for "Prioritize features in the current product" decision

Selected_Decision = Prioritize features in the current product

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.349	5	.046	.771	5	.046
Feature	.231	5	.200*	.881	5	.314
Usage pattern	.473	5	.001	.552	5	.000
Referral sources	.330	5	.079	.735	5	.021
Technologies and channels	.237	5	.200*	.961	5	.814
Product healthiness	.367	5	.026	.684	5	.006

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Table 25: Normality test for "Confirm a new technology for developing a feature" decision

Selected_Decision = Confirm a new technology for developing a feature(s)[b]

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.367	5	.026	.684	5	.006
Feature	.318	5	.109	.701	5	.010
Usage pattern	.231	5	.200*	.881	5	.314
Technologies and channels	.330	5	.079	.735	5	.021
Product healthiness	.473	5	.001	.552	5	.000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

b. Referralsourcesforproductuse is constant. It has been omitted.

Normality test of measurement-category for Product type

Table 26: Normality test for "Information display and transaction entry" application type

Application_Type = Information_display_and_transaction_entry

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.433	10	.000	.594	10	.000
Feature	.524	10	.000	.366	10	.000
Usage pattern	.336	10	.002	.784	10	.009
Referral sources	.360	10	.001	.731	10	.002
Technologies and channels	.370	10	.000	.686	10	.001
Product healthiness	.333	10	.002	.678	10	.000

a. Lilliefors Significance Correction

Table 27: Normality test for "Consumer oriented software" application type

Application_Type = Consumer-oriented software

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.470	14	.000	.532	14	.000
Feature	.350	14	.000	.731	14	.001
Usage pattern	.306	14	.001	.773	14	.002
Referral sources	.430	14	.000	.641	14	.000
Technologies and channels	.199	14	.138	.834	14	.014
Product healthiness	.352	14	.000	.600	14	.000

a. Lilliefors Significance Correction

Table 28: Normality test for "Business oriented" application type

Application_Type = Business-oriented

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Product	.391	8	.001	.641	8	.000
Feature	.375	8	.001	.706	8	.003
Usage pattern	.327	8	.012	.810	8	.037
Referral sources	.228	8	.200*	.835	8	.067
Technologies and channels	.325	8	.013	.665	8	.001
Product healthiness	.371	8	.002	.724	8	.004

Lilliefors Significance Correction

Appendix B.2.3

Contingency analysis

Contingency analysis for planning decisions and measurement-categories

This section presents contingency tables for different decisions and measurement-categories. They present the distribution of answers for different importance level of measurement-categories (from 0 to 4 indicates "no idea", "not important", "less important", "important" and "very important") for different decision types.

Table 29: Contingency table for selected decisions and "Product" measurement-category

Decisions		Product 1	2	3	4	Total
Create	*f*	0	1	0	1	3
	%	20	0	20	60	100
Enhance	*f*	0	0	0	3	3
	%	16.66	0	16.66	66.66	100
Remove	*f*	0	0	0	1	4
	%	0	0	40	60	100
Allocate	*f*	0	0	0	3	5
	%	0	12.5	25	62.5	100
Prioritize	*f*	0	1	1	3	5
	%	0	20	20	60	100
Confirm technology	*f*	0	0	2	3	5
	%	0	0	40	60	100
Total	*f*	2	2	9	21	34
	%	5.9	5.9	26.5	61.7	100

Table 30: Contingency table for selected decisions and "Feature" measurement-category

		Feature 0	1	2	3	4	Total
Create	*f*	0	1	0	1	3	5
	%	0	20	0	20	60	100
Enhance	*f*	0	0	0	3	3	6
	%	0	0	0	50	50	100
Remove	*f*	0	0	0	1	4	5
	%	0	0	0	20	80	100
Allocate	*f*	0	0	0	3	5	8
	%	0	0	0	37.5	62.5	100
Prioritize	*f*	0	0	1	2	2	5
	%	0	0	20	40	40	100
Confirm technology	*f*	1	0	0	1	3	5
	%	20	0	0	20	60	100
Total	*f*	1	1	1	11	20	34
	%	2.9	2.9	2.9	32.4	58.9	100

Table 31: Contingency table for planning decisions and "Usage Pattern" measurement-category

		Usage pattern 1	2	3	4	Total
Create	*f*	1	1	1	2	5
	%	20	20	20	40	100
Enhance	*f*	0	0	3	3	6
	%	0	0	50	50	100
Remove	*f*	1	0	2	2	5
	%	20	0	40	40	100
Allocate	*f*	0	1	4	3	8
	%	0	12.5	50	37.5	100
Prioritize	*f*	0	1	4	0	5
	%	0	20	80	0	100
Confirm technology	*f*	0	1	2	2	5
	%	0	20	40	40	100
Total	*f*	2	4	16	12	34
	%	5.9	11.8	47.1	35.2	100

Table 32: Contingency table for planning decisions and "Referral sources" measurement-category

		Referral sources 0	1	2	3	4	Total
Create	*f*	0	5	0	0	0	5
	%	0	100	0	0	0	100
Enhance	*f*	2	2	2	0	0	6
	%	33.3	33.3	33.3	0	0	100
Remove	*f*	1	3	1	0	0	5
	%	20	60	20	0	0	100
Allocate	*f*	0	3	3	1	1	8
	%	0	37.5	37.5	12.5	12.5	100
Prioritize	*f*	0	3	1	0	1	5
	%	0	60	20	0	20	100
Confirm technology	*f*	0	5	0	0	0	5
	%	0	100	0	0	0	100
Total	*f*	3	21	7	1	2	34
	%	8.9	61.8	20.5	2.9	5.9	100

Table 33: Contingency table for planning decisions and "Technologies and channels" measurement-category

		Technologies and channels					
		0	**1**	**2**	**3**	**4**	**Total**
Create	*f*	1	1	0	2	1	5
	%	20	20	0	40	20	100
Enhance	*f*	0	2	3	1	0	6
	%	0	33.3	50	16.7	0	100
Remove	*f*	2	1	1	1	0	5
	%	40	20	20	20	0	100
Allocate	*f*	0	1	3	2	2	8
	%	0	12.5	37.5	25	25	100
Prioritize	*f*	0	1	1	2	1	5
	%	0	20	20	40	20	100
Confirm technology	*f*	1	0	1	3	0	5
	%	20	0	20	60	0	100
Total	*f*	4	6	9	11	4	34
	%	11.8	17.6	26.4	32.2	11.8	100

Table 34: Contingency table for planning decisions and "Product healthiness" measurement-category

	Product healthiness					
	0	**1**	**2**	**3**	**4**	**Total**
Create	0	0	0	2	3	5
	0	0	0	40	60	100
Enhance	1	1	0	1	3	6
	16.7	16.7	0	16.7	50	100
Remove	0	0	1	0	4	5
	0	0	20	0	80	100
Allocate	1	0	0	2	5	8
	12.5	0	25	0	62.5	100
Prioritize	0	0	0	2	3	5
	0	0	0	40	60	100
Confirm technology	0	0	0	1	4	5
	0	0	0	20	80	100
Total	2	1	1	8	22	34
	5.9	2.9	2.9	23.5	64.7	100

Appendix B.2.4

Scores of measurement attributes

The following table shows the overall scores of measurement attributes. Each measurement attribute used to be scored inside a measurement-category and here this table shows the scores of measurement attributes amongst all by considering the value of the corresponding category. The "N" column indicates number of available answer for the attributes, "Mean" and "Variance" columns show the mean and variance values of a measurement-attribute, and "Missed" column presents the percent of missed answers for the record. "Score of the related category" column indicates the score of corresponding category among other measurement-category. "Score of the attribute" column, which is calculated by multiplying the values of "Mean" and "Score of the related category" column, shows the importance amongst all measurements. For all these scores, confidence intervals (with confidence level of 95%) have been also calculated using T-test distribution (e.g. the mean value of 5.00±0.65 shows the confidence interval of 0.65 for the mean value of 5.00).

Table 35: weights for measurements

Measurement-Categories	Measurement-attributes	N	Mean	Variance	Missed	Confidence Interval	Score of the category	Score of the attribute
Product value	Product use	30	5.00	3.310	11.8	0.68	3.44	0.72±0.16
	Overall amount of users	30	3.33	4.782	11.8	0.82	3.44	0.48±0.16
	Time between visits	30	1.60	3.007	11.8	0.65	3.44	0.23±0.11
	Duration of using the product	30	3.13	3.844	11.8	0.73	3.44	0.45±0.14
	New users	30	1.60	2.317	11.8	0.57	3.44	0.23±0.10
	Returning users	30	2.80	4.717	11.8	0.81	3.44	0.40±0.15
Value of a feature from user's perspective	Users that use a feature	32	4.56	2.577	5.9	0.58	3.52	0.67±0.13
	Feature use	32	5.13	2.435	5.9	0.56	3.52	0.75±0.14
	Duration of using a feature	32	3.44	3.222	5.9	0.65	3.52	0.50±0.13
	Entrance feature	32	1.81	3.125	5.9	0.64	3.52	0.27±0.11
	Exit feature	32	0.94	1.609	5.9	0.46	3.52	0.14±0.08
	Bounce	32	1.88	1.88	5.9	0.49	3.52	0.28±0.09
Usage pattern	Click activity	30	1.53	1.361	11.8	0.44	3.12	0.40±0.08
	Depth of use	30	1.80	1.269	11.8	0.42	3.12	0.47±0.08
	Click stream/path	30	1.60	1.490	11.8	0.46	3.12	0.42±0.08
Referral sources	Referrers	4	0	0	88.2	Not applicable	1.48	0.00
	Location/ISP per use	4	0	0	88.2		1.48	0.00
	campaigns	4	4	0	88.2		1.48	0.37
	search engines and keywords	4	0	0	88.2		1.48	0.00
Technologies and channels	Languages	22	1.91	3.325	35.3	0.81	2.43	0.23±0.11
	Browsers	22	2.82	4.918	35.3	0.98	2.43	0.34±0.14
	Operating Systems	22	1.50	4.357	35.3	0.93	2.43	0.18±0.12
	Plugins	22	.55	.831	35.3	0.40	2.43	0.07±0.05
	Screen	22	3.18	3.013	35.3	0.77	2.43	0.39±0.13

Measurement-Categories	**Measurement-attributes**	**N**	**Mean**	**Variance**	**Missed**	**Confidence Interval**	**Score of the category**	**Score of the attribute**
	resolution							
Product healthiness	Errors	30	4.33	3.609	11.8	0.71	3.59	0.65±0.15
	Downtime	30	3.97	3.757	11.8	0.72	3.59	0.59±0.15
	Response time	30	4.67	1.747	11.8	0.49	3.59	0.70±0.12
	Throughput	30	3.17	4.075	11.8	0.75	3.59	0.47±0.15
	DOS attacks	30	.93	1.030	11.8	0.38	3.59	0.14±0.07
	Worm attacks	30	.40	.248	11.8	0.19	3.59	0.06±0.03

T-test for grouping the measurement-attributes

In this section two top tables present 28 independent T-tests and two groups were formed for the independent variable of each test. “Group1” and “Group2” columns specify the “attributes no.” available in table at the bottom of the page. Looking at “Group1” column, for example “1-6” means the attributes no. 1 till 6 are included in the first group. “P-value” column shows the result of each T-test. As it has highlighted in Table 36, p<0.05 indicates the groups have significant difference and their members belong to different groups. As an instance, the attribute no. 6 and 7 belong to a new group, because their p-values are less than 0.05.

Table 36: T-tests for grouping measurement-attributes

Group1	Group2	P-value
1-2	3	0.407
1-3	4	0.208
1-4	5	0.209
1-5	6	0.68
1-6	7	0.032
1-6	8	0.022
7-8	9	0.454
7-9	10	0.529
7-10	11	0.154
7-11	12	0.053
7-12	13	0.079
7-13	14	0.188
7-14	15	0.188
7-15	16	0.13
7-16	17	0.072

Group1	Group2	P-value
7-17	18	0.021
7-18	19	0.015
18-19	20	0.121
18-20	21	0.43
18-21	22	0.5
18-22	23	0.067
18-23	24	0.054
18-24	25	0.164
18-25	26	0.041
18-25	27	0.032`
26-27	28	0.084
26-28	29	0.426
26-29	30	0.497

Attribute no.	Measurement-attribute
1	Feature use
2	Product use
3	Response time
4	Users that use a feature
5	Errors
6	Downtime
7	Duration of using a feature
8	Overall amount of users
9	Throughput
10	Depth of use
11	Duration of using the product
12	Click stream/path
13	Returning users
14	Click activity
15	Screen resolution

Attribute no.	Measurement-attribute
16	Campaigns
17	Browsers
18	Bounce
19	Entrance feature
20	Languages
21	Time between visits
22	New users
23	Operating Systems
24	DOS attacks
25	Exit feature
26	Plugins
27	Worm attacks
28	Referrers
29	Location/ISP per use
30	Search engines and keywords

Appendix B.2.5

Frequencies of measurements

The following tables (Figure 36 to Figure 65) show the frequencies of measurements rates that are assigned during the interview-based survey by interviewees. The measurement belongs to a measurement-category that are specified by “measurement-category X: measurement-attribute Y” in the captions of the tables. X refers to the measurement-category and Y refers to the measurement-attribute. In the tables x-axis shows the distribution of measurements rates: higher number shows the higher rate.

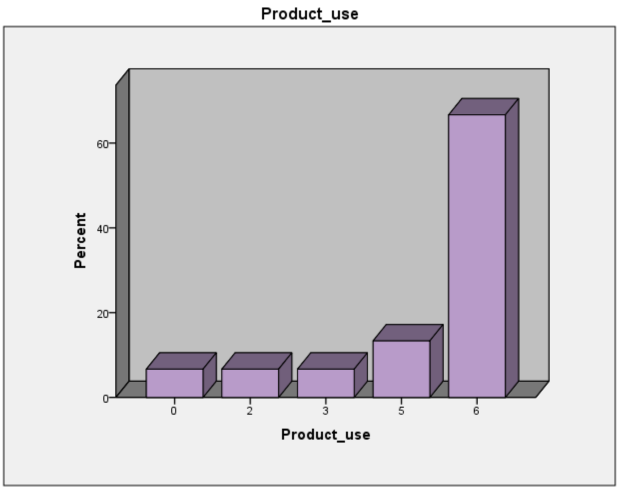


Figure 36: Percentage frequencies of rates for "Product: Product use”.

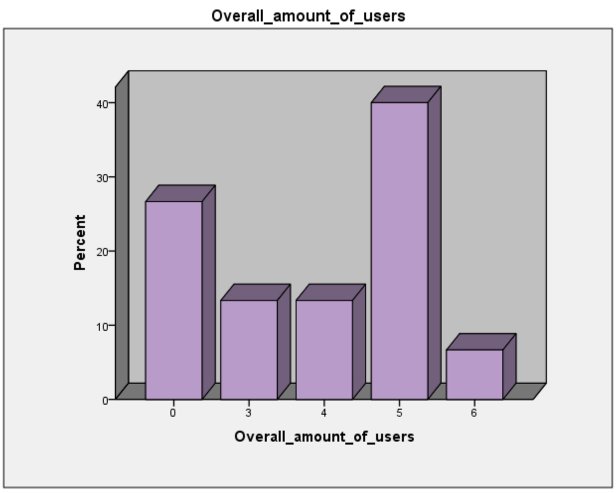


Figure 37: Percentage frequencies of rates for "Product: Overall amount of users".

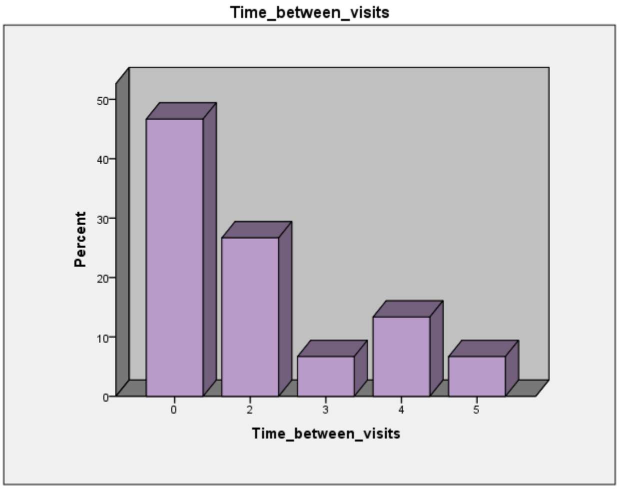


Figure 38: Percentage frequencies of rates for "Product: Time between visits".

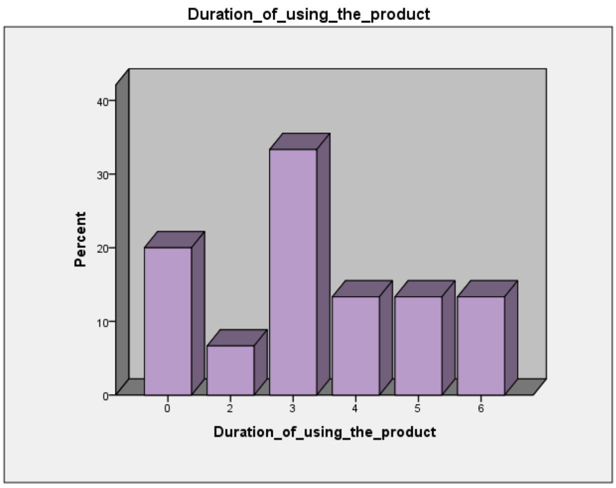


Figure 39: Percentage frequencies of rates "Product: Duration of using the product".

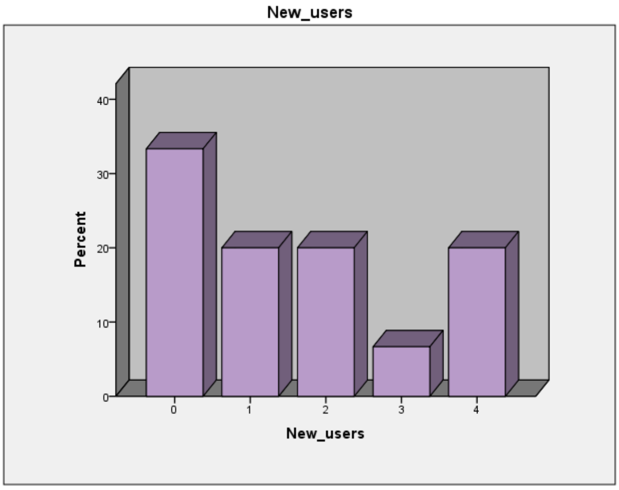


Figure 40: Percentage frequencies of rates for "Product: New users".

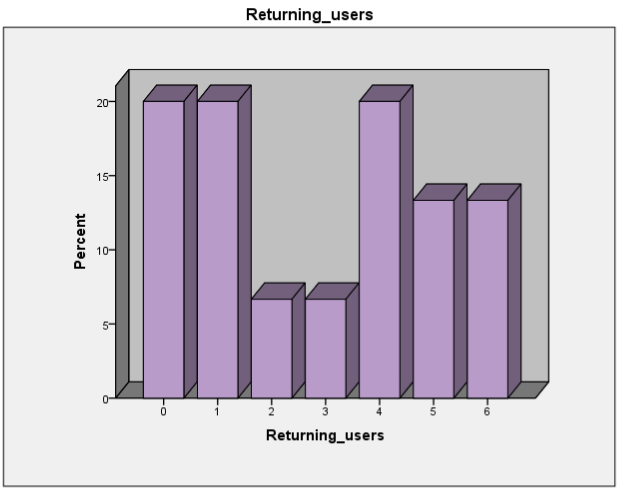


Figure 41: Percentage frequencies of rates for "Product: Returning users".

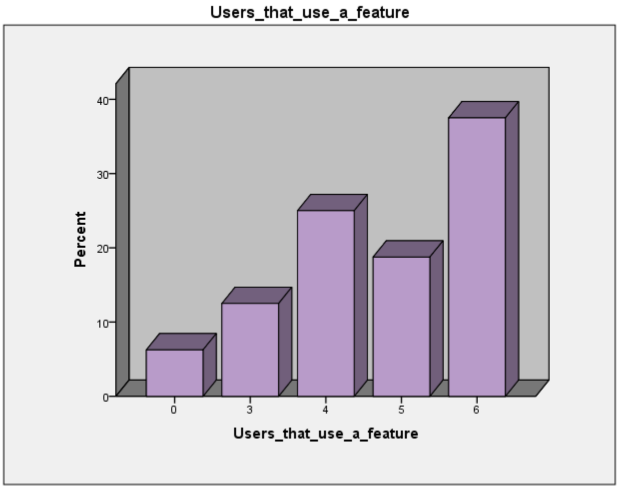


Figure 42: Percentage frequencies of rates for "Feature: Users tat use a feature"

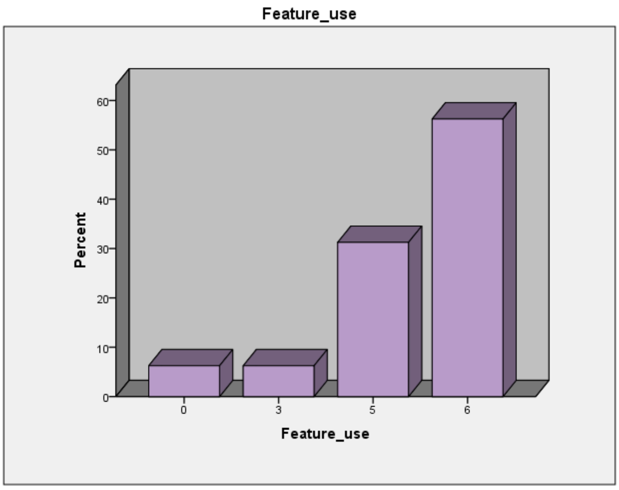


Figure 43: Percentage frequencies of rates for "Feature: Feature use"

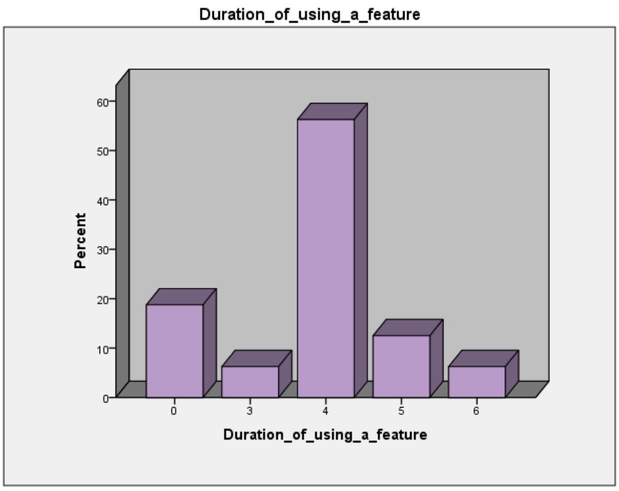


Figure 44: Percentage frequencies of rates for "Feature: Duration of using a feature"

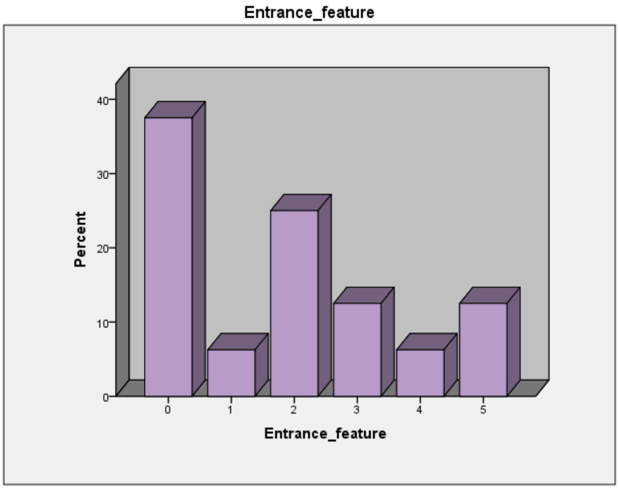


Figure 45: Percentage frequencies of rates for "Feature: Entrance feature"

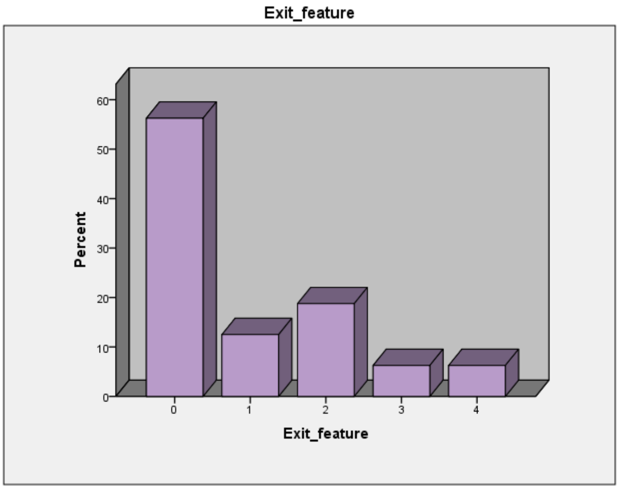


Figure 46: Percentage frequencies of rates for "Feature: Exit feature"

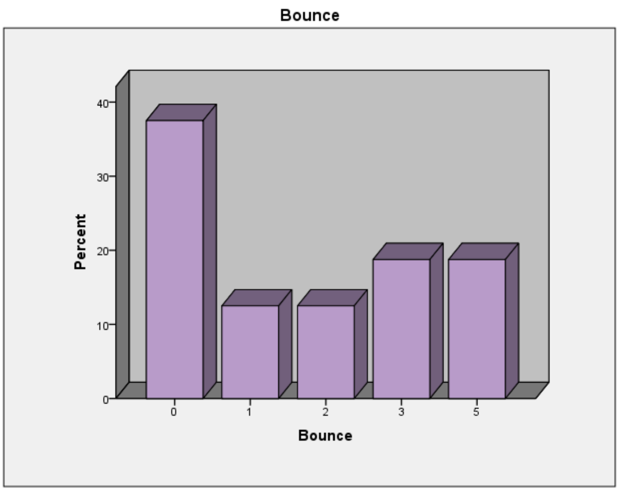


Figure 47: Percentage frequencies of rates for "Feature: Bounce"

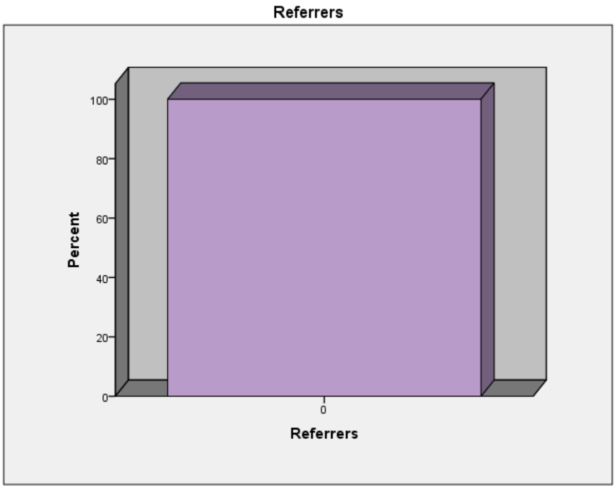


Figure 48: Percentage frequencies of rates for "Referral sources: Referrers"

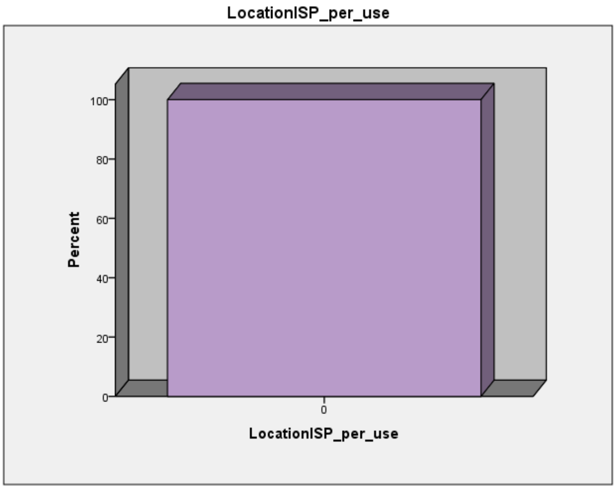


Figure 49: Percentage frequencies of rates for "Referral sources: Location/ISP per use"

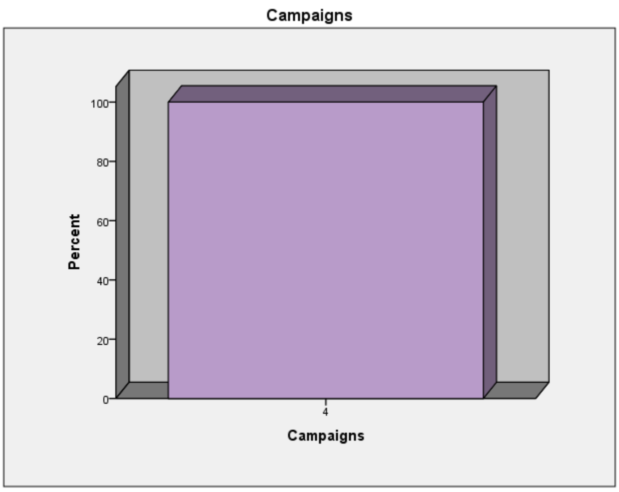


Figure 50: Percentage frequencies of rates for "Referral sources: Campaigns"

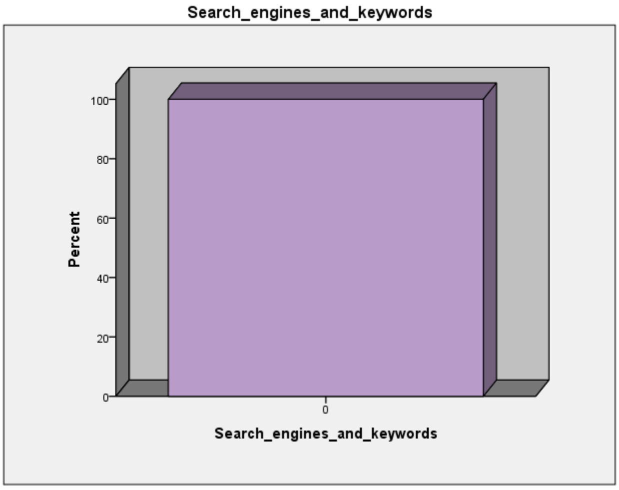


Figure 51: Percentage frequencies of rates for "Referral sources: Search engines and keywords"

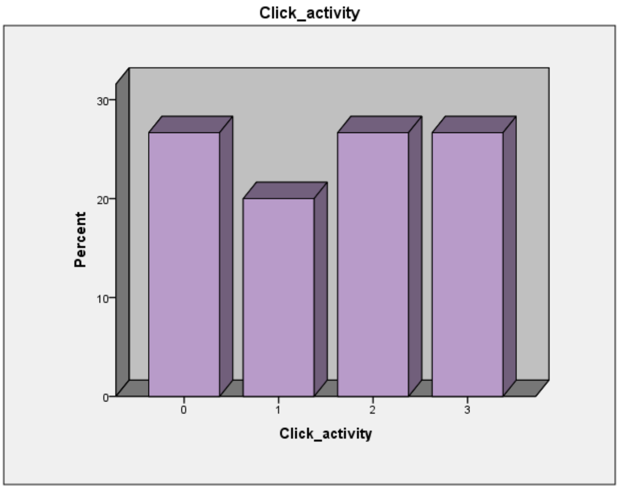


Figure 52: Percentage frequencies of rates for "Usage pattern use: Click activity"

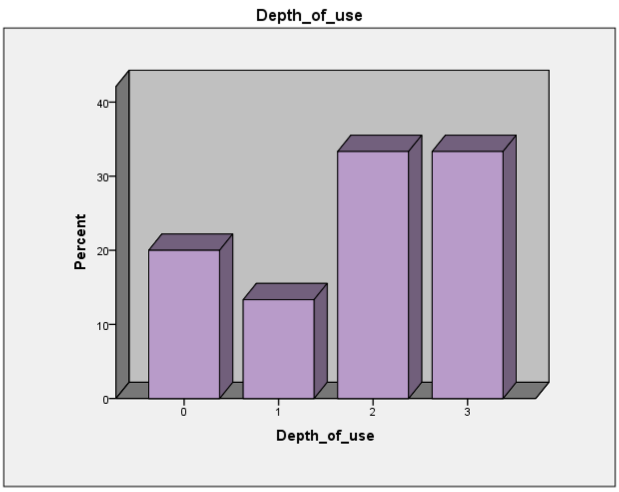


Figure 53: Percentage frequencies of rates for "Usage pattern use: Depth of use"

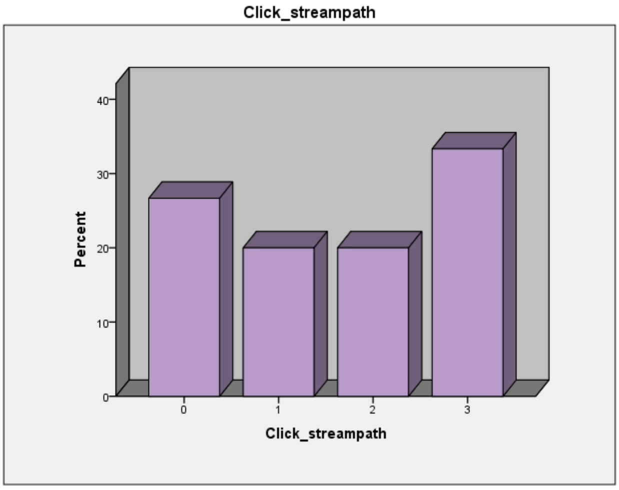


Figure 54: Percentage frequencies of rates for "Usage pattern use: Click stream path"

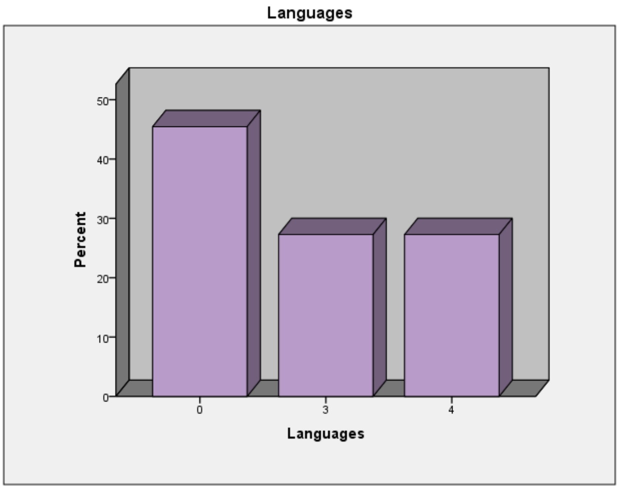


Figure 55: Percentage frequencies of rates for "Technologies and channels: Language"

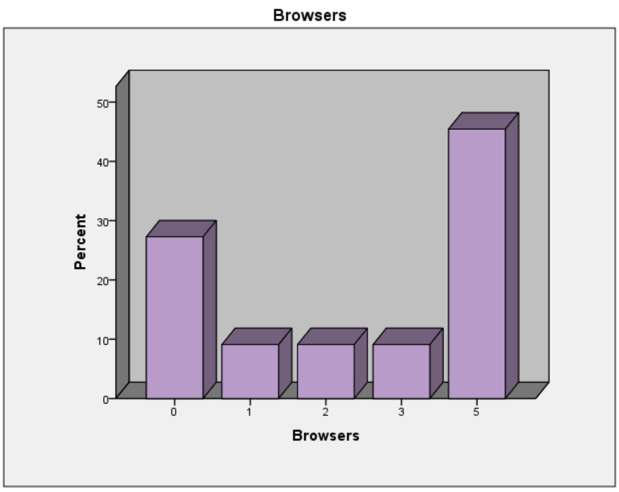


Figure 56: Percentage frequencies of rates for "Technologies and channels: Browsers"

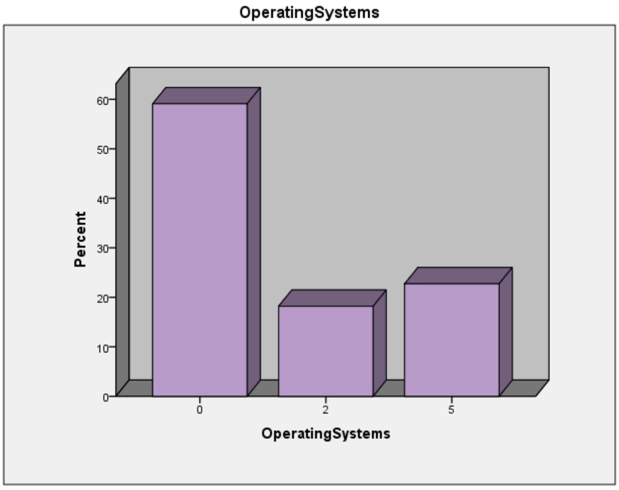


Figure 57: Percentage frequencies of rates for "Technologies and channels: Operating system"

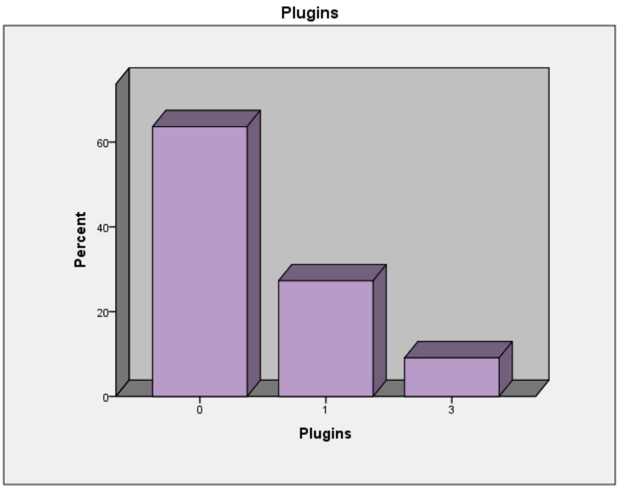


Figure 58: Percentage frequencies of rates for "Technologies and channels: Plugins"

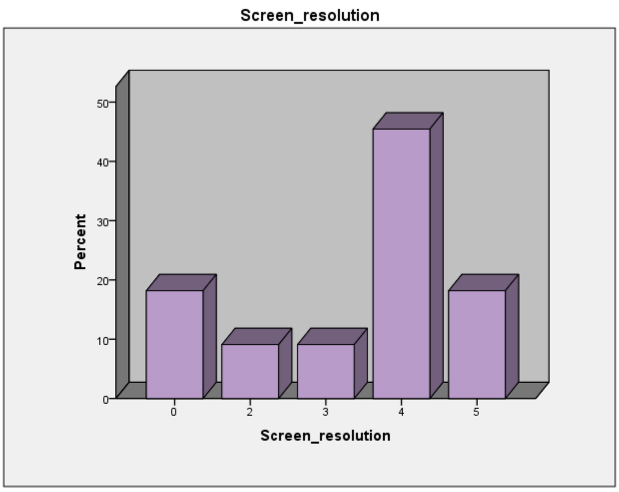


Figure 59: Percentage frequencies of rates for "Technologies and channels: Screen resolution"

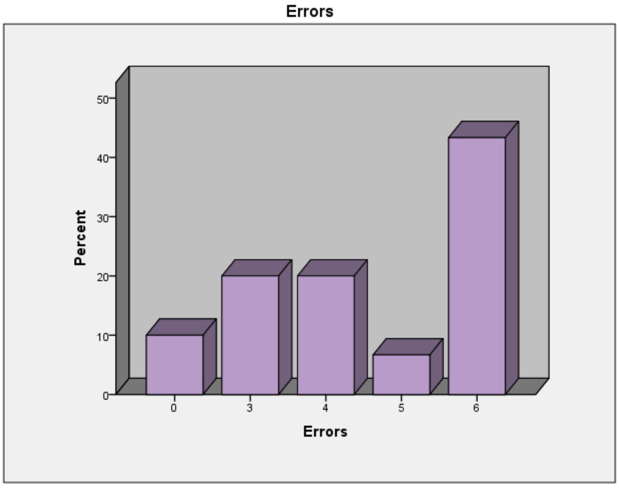


Figure 60: Percentage frequencies of rates for "Product healthiness: Errors"

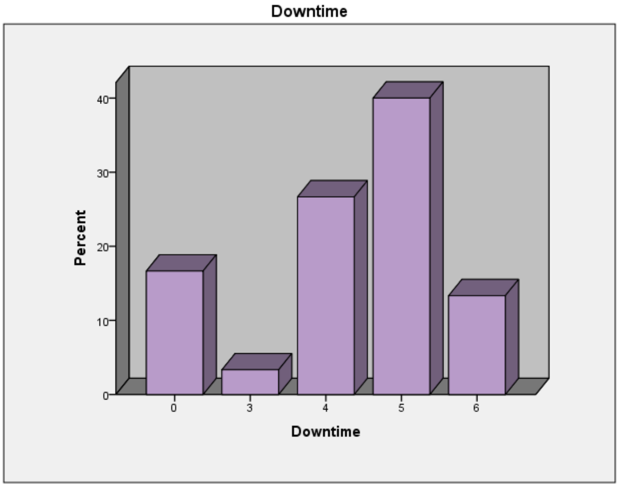


Figure 61: Percentage frequencies of rates for "Product healthiness: Downtime"

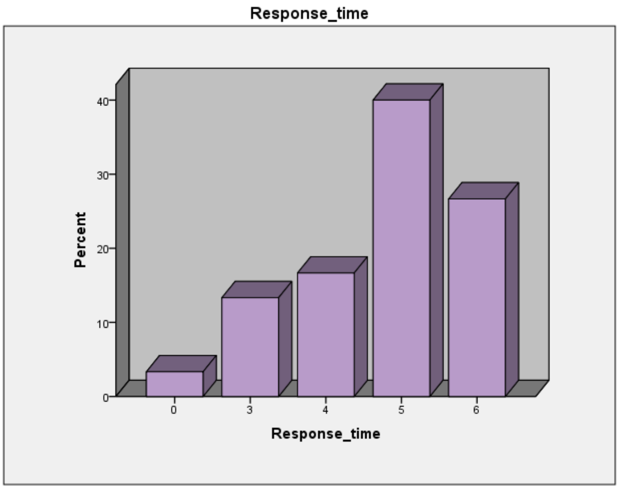


Figure 62: Percentage frequencies of rates for "Product healthiness: Response time"

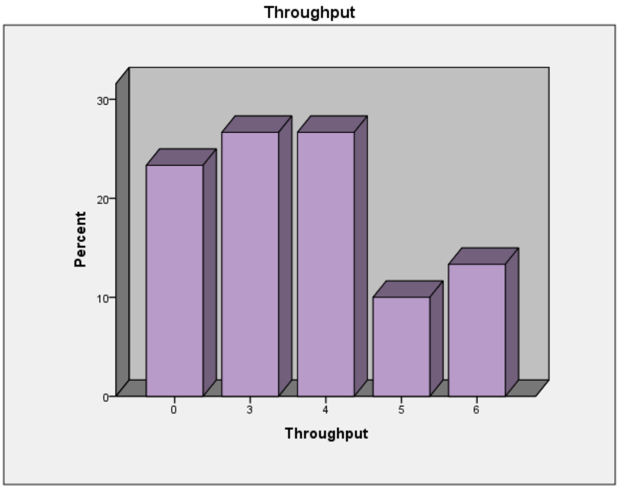


Figure 63: Percentage frequencies of rates for "Product healthiness: Throughput"

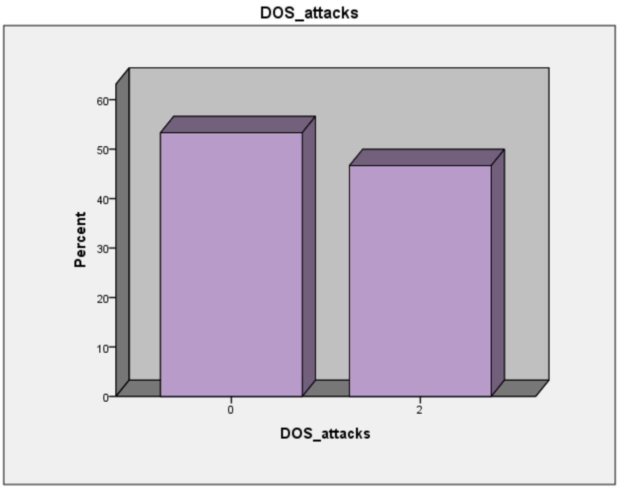


Figure 64: Percentage frequencies of rates for "Product healthiness: DOS attacks"

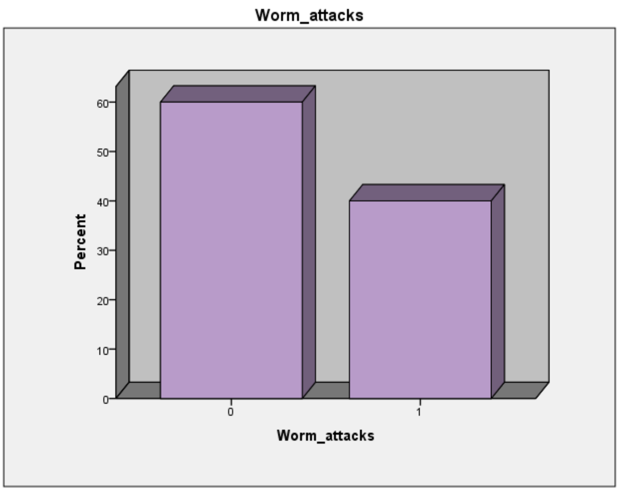


Figure 65: Percentage frequencies of rates for "Product healthiness: Worm attacks"

Appendix C

Case study

Appendix C.1

Pattern for sample statistics

Table 37: Patterns for sample statistics

Measurement - categories	Measurement -attributes	Sample
Product	Product use	Number/ percentage of visits per a time period
		Minimum/ maximum/ median number of visits per a time period
		Average of visits per a time period
		Variance of visits per a time period
		Product use trend within a time period
	Overall amount of users	Number of unique visitors per a time period
		Minimum/ maximum/ median number of unique visitors per a time period
		Average of unique visitors per a time period
		Variance of unique visitors per a time period
		Unique visitors trend within a time period
	Returning user	Number/ percentage of returning visitors per a time period
		Minimum/ maximum/ median number of returning visitors per a time period
		Average of returning visitors per a time period
		Variance of returning visitors per a time period
		Returning visitors trend within a time period
	Time between visits	Number/ percentage of visits with more than x days/weeks/ months since last visit
		Average of visits with more than x days/weeks/ months since last visit
		Maximum/ Minimum/ median of visits with more than x days/weeks/ months since last visit
	Duration of using the product	Average visit duration per a time period
		Minimum/ maximum/ median visit duration per a time period
		Using duration trend within a time period
	New users	Number/ percentage of new users per a time period
		Minimum/ maximum/ median number of new users per a time period
		Average of new users per a time period
		Variance of new users per a time period
		New users trend within a time period
Feature	Users for a feature	Number percentage of unique pageviews per a time period
		Minimum/ maximum/ median number of unique pageviews per a time period for "X" page
		Average of unique pageviews per a time period for "X" page
		Variance of unique pageviews per a time period for "X" page
		Unique pageviews trend within a time period for "X" page
	Feature use	Number/ percentage of pageviews per a time period for "X" page
		Minimum/ maximum/ median number of pageviews per a time period for "X" page
		Average of pageviews per a time period for "X" page
		Variance of pageviews per a time period for "X" page
		Pageviews trend within a time period for "X" page
	Duration of using a feature	Average/ percentage duration of using "X" feature per a time period
		Minimum/ maximum/ median duration of using "X" feature per a time period
		Using duration trend within a time period for "X" page
	Entrance feature	Number/ percentage of entrances per a time period for "X" page
		Minimum/ maximum/ median number of entrances within a time period for

Measurement - categories	Measurement -attributes	Sample
		"X" page
		Average of entrances per a time period for "X" page
		Variance of entrances per a time period for "X" page
		Entrance trend within a time period for "X" page
	Exit feature	Number/ percentage of exits per a time period for "X" page
		Minimum/ maximum/ median number of exits within a time period for "X" page
		Average of exits per a time period for "X" page
		Variance of exits per a time period for "X" page
		Exit trend within a time period for "X" page
	Bounce	Number/ percentage of bounce visit per a time period for "X" page
		Minimum/ maximum/ median number of bounce visit within a time period for "X" page
		Average of bounce visit per a time period for "X" page
		Variance of bounce visit per a time period for "X" page
		Bounce visit trend within a time period for "X" page
	Click activity	Number/ percentage of visits/ users with number/ minimum/ maximum of "X" clicks on "Y" feature per a time period
		Minimum/ maximum/ median number of visits/ users with number/ minimum/ maximum of "X" clicks on "Y" feature per a time period
		Average of visits/ users with number/ minimum/ maximum of "X" clicks on "Y" feature per a time period
		Variance of visits/ users with number/ minimum/ maximum of "X" clicks on "Y" feature per a time period
		Visits/ users trend with number/ minimum/ maximum of "X" clicks on "Y" feature per a time period
	Depth of use	Number/ percentage of visits/ users with number/ minimum/ maximum of ("X" depth of use, depth of use > "X" or Depth of use < "X") per a time period
		Minimum/ maximum/ median number of visits/ users with number/ minimum/ maximum of ("X" depth of use, depth of use > "X" or Depth of use < "X") per a time period
		Average of visits/ users with number/ minimum/ maximum of ("X" depth of use, depth of use > "X" or Depth of use < "X") per a time period
		Variance of visits/ users with number/ minimum/ maximum of ("X" depth of use, depth of use > "X" or Depth of use < "X") per a time period
		Visits/ users trend with number/ minimum/ maximum of ("X" depth of use, depth of use > "X" or Depth of use < "X") per a time period
	Click stream/path	Number/ percentage of visits/ users with exact/ minimum/ maximum sequential clicks of "X, Y, Z, W, ..." per a time period
		Minimum/ maximum/ median number of visits/ users with exact/ minimum/ maximum sequential clicks of "X, Y, Z, W, ..." per a time period
		Average of visits/ users with exact/ minimum/ maximum sequential clicks of "X, Y, Z, W, ..." per a time period
		Variance of visits/ users with exact/ minimum/ maximum sequential clicks of "X, Y, Z, W, ..." per a time period
		Visits/ users trend with exact/ minimum/ maximum sequential clicks of "X, Y, Z, W, ..." per a time period
Referral sources for product use	**Referrers**	Number/ percentage of direct/website/search engine entries per a time period
		Minimum/ maximum/ median number of direct/website/search engine entries per a time period
		Average of direct/website/search engine entries per a time period
		Variance of direct/website/search engine entries per a time period
		Visits/ users trend with direct/website/search engine entries per a time period
	Location/ISP per use	Number/ percentage of visits/ users per a time period from "X" location
		Minimum/ maximum/ median number of visits/ users per a time period from "X" location
		Average of visits/ users per a time period from "X" location
		Variance of visits/ users per a time period from "X" location
		Visits/ users trend per a time period from "X" location

Measurement - categories	Measurement -attributes	Sample
	Campaigns	Number/ percentage of visits/ users per a time period redirected from "X" campaign
		Minimum/ maximum/ median number of visits/ users per a time period redirected from "X" campaign
		Average of visits/ users per a time period redirected from "X" campaign
		Variance of visits/ users per a time period redirected from "X" campaign
		Visits/ users trend per a time period redirected from "X" campaign
	Search engines and keywords	Number/ percentage of visits/ users per a time period redirected from "X" search engine or used "Y" keyword
		Minimum/ maximum/ median number of visits/ users per a time period redirected from "X" search engine or used "Y" keyword
		Average of visits/ users per a time period redirected from "X" search engine or used "Y" keyword
		Variance of visits/ users per a time period redirected from "X" search engine or used "Y" keyword
		Visits/ users trend per a time period redirected from "X" search engine or used "Y" keyword
Technologies and channels used to access the product	Languages	Number/ percentage of visits/ visitors per a time period with "X" language
		Minimum/ maximum/ median number of visits/ visitors per a time period with "X" language
		Average of visits/ visitors per a time period with "X" language
		Variance of visits/ visitors per a time period with "X" language
		Visits/ visitors trend within a time period with "X" language
	Browsers	Number/ percentage of visits/ visitors per a time period with "X" browser
		Minimum/ maximum/ median number of visits/ visitors per a time period with "X" browser
		Average of visits/ visitors per a time period with "X" browser
		Variance of visits/ visitors per a time period with "X" browser
		Visits/ visitors trend within a time period with "X" browser
	Operating system	Number/ percentage of visits/ visitors per a time period with "X" operating system
		Minimum/ maximum/ median number of visits/ visitors per a time period with "X" operating system
		Average of visits/ visitors per a time period with "X" operating system
		Variance of visits/ visitors per a time period with "X" operating system
		Visits/ visitors trend within a time period with "X" operating system
	Plugins	Number/ percentage of visits/ visitors per a time period with "X" plugins
		Minimum/ maximum/ median number of visits/ visitors per a time period with "X" plugins
		Average of visits/ visitors per a time period with "X" plugins
		Variance of visits/ visitors per a time period with "X" plugins
		Visits/ visitors trend within a time period with "X" plugins
	Screen resolution	Number/ percentage of visits/ visitors per a time period with "X" screen resolution
		Minimum/ maximum/ median number of visits/ visitors per a time period with "X" screen resolution
		Average of visits/ visitors per a time period with "X" screen resolution
		Variance of visits/ visitors per a time period with "X" screen resolution
		Visits/ visitors trend within a time period with "X" screen resolution
Product healthiness	Errors	Number/ percentage of product logs/ visits/ visitors per a time period with "X" type of error
		Minimum/ maximum/ median number of product logs/ visits/ visitors per a time period with "X" type of error
		Average of product logs/ visits/ visitors per a time period with "X" type of error
		Variance of product logs/ visits/ visitors per a time period with "X" type of error
		Product logs/ visits/ visitors trend within a time period with "X" type of error
	Downtime	Number/ percentage of downtime equals to "X", downtime > "X" or downtime < "X" within a time period

Measurement - categories	Measurement -attributes	Sample
		Minimum/ maximum/ median number of downtime equals to "X", downtime > "X" or downtime < "X" within a time period
		Average of downtime equals to "X", downtime > "X" or downtime < "X" within a time period
		Variance of downtime equals to "X", downtime > "X" or downtime < "X" within a time period
		Downtime trend within a time period when (downtime equals to "X", downtime > "X" or downtime < "X")
	Response time	Number/ percentage of visits/ visitors per a time period with (response time equals to "X", response time > "X" or response time < "X") for "X" feature or product
		Minimum/ maximum/ median number of visits/ visitors per a time period with (response time equals to "X", response time > "X" or response time < "X") for "X" feature or product
		Average of visits/ visitors per a time period with (response time equals to "X", response time > "X" or response time < "X") for "X" feature or product
		Variance of visits/ visitors per a time period with (response time equals to "X", response time > "X" or response time < "X") for "X" feature or product
		Visits/ visitors trend within a time period with (response time equals to "X", response time > "X" or response time < "X") for "X" feature or product
	Throughput	Number/ percentage of times that (throughput equals to "X", throughput > "X" or throughput < "X") for "X" feature or product within a time period.
		Minimum/ maximum/ median number of times that(throughput equals to "X", throughput > "X" or throughput < "X") for "X" feature or product within a time period.
		Average of times that (throughput equals to "X", throughput > "X" or throughput < "X") for "X" feature or product within a time period.
		Variance of times that (throughput equals to "X", throughput > "X" or throughput < "X") for "X" feature or product within a time period.
		Throughput trend within a time period that (throughput equals to "X", throughput > "X" or throughput < "X") for "X" feature or product within a time period.
	DOS attacks	Number/ percentage of DOS attacks per a time period
		Minimum/ maximum/ median number of DOS attacks per a time period
		Average of DOS attacks per a time period
		Variance of DOS attacks per a time period
		DOS attacks trend within a time period
	Worm attacks	Number/ percentage of worm attacks per a time period
		Minimum/ maximum/ median number of worm attacks per a time period
		Average of worm attacks per a time period
		Variance of worm attacks per a time period
		Worm attacks trend within a time period

Appendix C.2

Method effectiveness Questionnaire

1) How did you select the measurements for the features? (Very difficult, difficult, easy, very easy, no idea)
2) How were the selected measurements interpreted? (Very difficult, difficult, easy, very easy, no idea)
3) How much did the measurements assist analyzing the importance of the features? (Very strong , strong, medium, weak, very weak, no idea)
4) How did you compare the measurements for particular decision? (Very difficult, difficult, easy, very easy, no idea)
5) How much do you evaluate the method output? (Very strong, strong, medium, weak, very weak, no idea)
6) How much did the method facilitate prioritizing alternative decisions? (Very strong, strong, medium, weak, very weak, no idea)
7) How much did the method facilitate trade-off between decisions? (Very strong, strong, medium, weak, very weak, no idea)
8) Do you think the feedback from previous monitored decisions is useful? (Very useful, useful, medium, not useful, no idea)
9) How much did the method facilitate the uncertainties about the decisions? (Very strong, strong, medium, weak, very weak, no idea)
10) How much do you evaluate the effectiveness of the method? (Very strong, strong, medium, weak, very weak, no idea)
11) How much do you think the method can be applicable in your organization? (Very difficult, difficult, easy, very easy, no idea)

Appendix C.3

Results for planning-decisions analysis

The following tables present the extracted data during evaluation of the analytics-based method for a specific planning decision. Definitions of columns and their mapping to the process steps of the method [section 5.1] have been shown as follows:

Sample statistic: An instance for a measurement (related to Step 5).

X1: The observed value of the sample statistic for duration of “Duration1” column (related to Step 6).

X2: The second observed value of the sample statistic for duration of “Duration2” column (Step 6).

Duration1: The first time periods(s) within which, the measurement values will be compared. (related to Step 6).

Duration2: The second time periods(s) within which, the measurement values will be compared. (related to Step 6).

Function Goal: A goal question based on the measurement, which defines comparison function (related to Step 7).

Function’s Boundary: It specifies conditions of a comparison function. A comparison function receives an measurement value (s) as an input(s) and compares them based on the defined goal (Conditions in the function). The output is the positive or negative effect of measurements on the decision (related to Step 7).

Indicator: It is the output of comparison function and shows if measurements support the decision positively or has negative affirmation on the decision (related to Step 7).

Recorded Rate: The weight for a measurement extracted from the interview-based survey that shows importance levels of the measurement among others (related to Step 8)

Decided Rate: Decided weights for measurements. It can be new one by considering criteria related to product/feature or recorded rate that previously have been specified (related to Step 8).

Total value: It is the weighted indicator that calculates by multiplying the “indicator” column and “Decided rate” column (related to Step 8).

At the bottom of the tables three variables are defined:

Impact Function: This function is Arithmetic mean or Harmonic functions to calculate the mean of Total value for comparison (related to Step 9)

Impact Function Output: This is the output of impact function (related to Step 9)

Pre-evaluation: This shows initial solution for the decision (related to Step 10)

Decision: Should English version of the product be created?

Table 38: Analytics for "Should English version of the product be created?" decision.

Measurement-attributes	Sample statistic	Function Goal	X1	Function's boundary	Indicator	Recorded rate	Decided rate	Total value
Product use Language	Percentage of visits using English language browser	Is the percentage of visits using English language browser more than 10%?	13.50%	>10%	1	1.71	3	3
Product use Location	Percentage of visits from outside Germans-country	Is the percentage of visits from outside Germans-country more than 10% ?	12.50%	>10%	1	0.00	2.5	2.5
Bounce	Index- page bounce	Is the Index-page bounce more than 20% of all Index-page visits?	13%	>20%	-1	1.38	1.38	-1.38
Returning user	Percentage of visitors has not returned	Is the percentage of visitors has not returned more than 20%?	21%	>10%	1	1.15	1.5	1.5
Duration of using the product	Percentage of visits duration (s) less than 1 min	Is the percentage of visits duration (s) less than 1 min more than 10%?	26%	>10%	1	1.15	1.5	1.5

Comparison functions: Use one variable function with the output of 1 and -1.
Impact Function: Arithmetic Mean
Impact Function Output: 0.83
Pre-evaluation: Taking the decision is recommended

Decision: Should Wiki feature be removed?

Table 39: Analytics for "Should Wiki feature be removed?" decision

Measurement-attributes	Sample statistics	Function Goal	X1	Duration 1	X2	Duration2	Indicator	Recorded rate	Decided rate	Total value
Users for a feature	Wiki users over product users	Has the metric decreased 80% in two recent months?	13%	For July and August	6%	For September and October	-0.33	3.34	4	-1.31
Overall amount of users								2.39		
Feature use	Wiki page views over all page views	Has the metric decreased 80% in two recent months?	19%	For July and August	10%	For September and October	-0.41	3.76	3.5	-1.43
Product use								3.58		
Duration of using a feature	Duration of using wiki over duration of using the product	Has the metric decreased 50% in two recent months?	2 min 47 sec	For July and August	3 min 39 sec	For September and October	-1	2.52	2	-2.00
Duration of using the product								1.15		
Bounce	Bounce rate for wiki	Has the metric decrease 80% in two recent months?	28%	July and August	19%		-0.6	1.38	1.38	-0.83
Click activity	Number of action on Wiki	Has the metric decreased 50% in two recent months?	12	July and August	10	For September and October	-0.996	1.99	2	-1.99

Impact Function: Arithmetic Mean
Impact Function Output: -1.51
Pre-evaluation: Taking the decision is not recommended

Decision: Should the Wiki feature be enhanced?

Table 40: Analytics for "Should the Wiki feature be enhanced?" decision

Measure ment-attributes	Sample statistics	Function Goal	X1	Duration 1	X2	Duration 2	Indicator	Recorded rate	Decided rate	Total value
Users for a feature	Wiki users over product users	Has the metric decreased 30% in two recent months?	13 %	July and August	6%	September and October	0.34	3.34	4	1.36
Overall amount of users								2.39		
Feature use	Wiki page views over all page views	Has the metric decreased 30% in two recent months?	19 %	July and August	10 %	For September and October	0.25	3.76	3.5	0.87
Product use								3.58		
Duration of using a feature	Duration of using wiki over duration of using the product	Has the metric decreased 30% in two recent months?	2 min 47 sec	July and August	3 min 39 sec	September and October	-1	2.52	1.15	-1.15
Duration of using the product								1.15		
Bounce	Bounce rate for Wiki	Has the metric decreased 30% in two recent months?	28 %	July and August	19 %	September and October	0.03	1.38	1.38	0.04
Click activity	Number of action on Wiki	Has the metric decreased 30% in two recent months?	12	July and August	10	September and October	-0.43	1.99	2	-0.85

Impact Function: Arithmetic Mean
Impact Function Output: 0.05
Pre-evaluation: Although the output is positive, it is close to Zero. So there is no recommendation for the decision

Appendix C.4

Parser for the New Relic log

```
import java.io.*;
importjava.util.*;
importjava.util.regex.Matcher;
importjava.util.regex.Pattern;

public class parser-newrelic{
public static void main(String[] args) throws Exception
  {
    Scanner s = new Scanner(new FileReader(new File("/Users/production.log")));
FileWriterfw = new FileWriter("/Users/output-production.csv");
PrintWriter pw = new PrintWriter(fw);
    Matcher m;
    Pattern p;
    String last="";
intarraylong=18;
String[] field= new String[arraylong];
while (s.hasNextLine()) {
      String line = s.nextLine();
if ("--- START ---".equals(line)) {
      } else if ("--- END ---".equals(line)) {
      } else if (line.startsWith("Started")) {
                        for (inti=0;i<arraylong;i++)
                        { field[i]=""; }
p = Pattern.compile(
        "Started ([^ ]+) \"([^ ]+)\" for ([^ ]+) at ([^ ]+) ([^ ]+) ([^ ]+)");
m = p.matcher(line);
if (m.find()) {
field[0]=m.group(1);
field[1]=m.group(2);
field[2]=m.group(3);
field[3]=m.group(4);
field[4]=m.group(5);
field[5]=m.group(6);
       }
last="Started";
      }  else if (line.startsWith("  Processing")) {
p = Pattern.compile(
        "  Processing by ([^ ]+) as ([^ ]+)");
m = p.matcher(line);
if (m.find()) {
                        field[6]=m.group(1);
                        field[7]=m.group(2);
        }
last="Processing";
      } else if (line.startsWith("Sent")) {
p = Pattern.compile(
        "Sent data  \\(([^ ]+)\\)");
m = p.matcher(line);
if (m.find()) {
                        field[8]=m.group(1);
        }
last="Sent";
      }  else if (line.startsWith("Redirected")) {
p = Pattern.compile(
        "Redirected to ([^ ]+)");
m = p.matcher(line);
if (m.find()) {
                        field[9]=m.group(1);
        }
last="Redirected";

      }  else if (line.startsWith("Completed")) {
p = Pattern.compile(
```

```
        "Completed ([^ ]+) ([^ ]+) in ([^ ]+) \\(Views: ([^ ]+) \\| ActiveRecord: ([^ ]+)\\)");
m = p.matcher(line);
if (m.find()) {
                field[10]=m.group(1);
                field[11]=m.group(2);
                field[12]=m.group(3);
                field[13]=m.group(4);
                field[14]=m.group(5);
        }
last="Completed";

    }  else if (line.startsWith("Served")) {
p = Pattern.compile(
        "Served asset ([^ ]+) - (.*) \\(([^ ]+)\\)");
m = p.matcher(line);
if (m.find()) {
                field[15]=m.group(1);
                field[16]=m.group(2);
                field[17]=m.group(3);

        }
last="Served";

    } else if (line.isEmpty() && (last.equals("Started")||last.equals("Completed")||last.equals("Served"))){
                for (inti=0;i<arraylong;i++){
                                pw.print(field[i]);
                                if (i==arraylong-1) pw.print("\n"); else pw.print(",");
                                last="";
                 }
    }

   }
pw.flush();
pw.close();
fw.close();
System.out.println("CSV file is generated");

}

 }
```