Aalto University
School of Business

Evaluating the Value of Web Metrics

Information Systems Science

Master's thesis

Tommi Riihimäki

2014

Department of Information and Service Economy
Aalto University
School of Business

Aalto University
School of Business

EVALUATING THE VALUE OF WEB METRICS

Master's Thesis
Tommi Riihimäki
13.3.2014
Information and Service Management

Approved in the Department of Information and Service Economy ___________
and awarded the grade _______________

AALTO UNIVERSITY SCHOOL OF BUSINESS ABSTRACT

Department of Information and Service Economy 13.03.2014

Master's Thesis

Tommi Riihimäki

ABSTRACT

Objectives of the Study

The unique measurability of websites allows the collection of detailed information about the behavior and characteristics of website visitors. This thesis examines the value of different web metrics based on the behavior of website visitors. The objective is to develop and test a method for identifying key metrics that are the most valuable for site developers to follow. The key web metrics are expected to contain the most useful and relevant information about the website visitors. The value of the web metrics is evaluated by examining the relationships of the metrics towards website conversions. This thesis also suggests how different web metrics can be analyzed to reveal important characteristics of site visitors and how web metrics can be used to evaluate the effectiveness of the different aspects of a website.

Academic background and methodology

This thesis draws from the field of web analytics. Based on the previous work on the field, a new framework for the role of web metrics analysis in website development is presented. The framework forms the basis for the case study conducted in this thesis. The case study examines one corporate website by following fourteen different web metrics during a period of six months. The relationship analysis between the web metrics and conversions is conducted using correlation and regression analyses. The objective of the case study is to test if meaningful key metrics can be identified using the method proposed in this thesis.

Findings and conclusions

The case study conducted in this thesis identifies two key web metrics for the site under examination: search engine traffic and the rate of return visits. Search engine traffic was selected as a key metric based on the relationship analysis and the rate of return visits was chosen based on the examination of the visitor characteristics. The successful identification of relevant key metrics suggests that the framework proposed in this study can be used as the basis for web metrics analysis. Based on the results of the case study, this thesis proposes that in addition to aggregated metrics, segmented web metrics are needed in order to achieve a more diverse view of the web users.

Keywords

Web analytics, web metrics, clickstream data

AALTO-YLIOPISTON KAUPPAKORKEAKOULU TIIVISTELMÄ
Tieto- ja palvelutalouden laitos 13.03.2014
Pro Gradu -tutkielma
Tommi Riihimäki

ABSTRAKTI

Tutkimuksen tavoitteet

Verkkosivujen ainutlaatuinen mitattavuus mahdollistaa yksityiskohtaisen tiedon keräämisen verkkosivujen vierailijoiden käyttäytymisestä ja ominaisuuksista. Tämä tutkielma tarkastelee eri vierailijoiden käyttäytymiseen perustuvien web-mittareiden arvoa. Tarkoituksena on kehittää ja testata mallia verkkosivujen kehittäjien kannalta arvokkaimpien avainmittarien löytämiseen. Avainmittarien odotetaan sisältävän kaikista hyödyllisintä ja merkityksellisintä tietoa sivujen käyttäjistä. Web-mittarien arvoa arvioidaan tarkastelemalla mittarien yhteyttä verkkosivujen konversioihin. Tämä tutkielma ehdottaa myös keinoja, miten web-mittareita voidaan analysoida tärkeiden ominaisuuksien paljastamiseksi vierailijoista ja miten web-mittareita voidaan hyödyntää verkkosivujen eri osa-alueiden tehokkuuden arvioimisessa.

Kirjallisuuskatsaus ja metodologia

Tämä tutkielma pohjautuu web-analytiikkaan. Tutkielma esittää uuden aikaisempaan alan tutkimukseen perustuvan mallin web-mittarien analyysin roolista verkkosivujen kehityksessä. Malli muodostaa pohjan tutkielmassa toteutetulle tapaustutkimukselle. Tapaustutkimus tarkastelee yhden yrityksen verkkosivuja seuraamalla neljäätoista eri web-mittaria kuuden kuukauden ajan. Web-mittareiden ja konversioiden yhteyttä tarkastellaan korrelaatio- ja regressioanalyysillä. Tutkielman tarkoituksena on testata, onko tutkielmassa esitettyä mallia käyttämällä mahdollista löytää merkitseviä avainmittareita.

Tulokset ja päätelmät

Tutkielmassa toteutettu tapaustutkimus identifioi kaksi avainmittaria tutkitulle verkkosivulle: hakukoneiden tuottama liikenteen määrä ja palaavien vierailujen suhde uusiin vierailuihin. Hakukoneiden tuottaman liikenteen määrän valinta avain-mittariksi perustuu suhdeanalyysin ja palaavien vierailujen suhteen valinta perustuu vierailijoiden ominaisuuksien tarkasteluun. Olennaisten avainmittarien löytäminen viittaa siihen, että tutkielmassa esitettyä mallia voidaan käyttää web-mittarien analyysin pohjana. Tutkielma esittää tapaustutkimuksen tulosten perusteella, että laajemman käyttäjäkuvan saamiseksi kokonaismittarien lisäksi tulee seurata myös segmentoituja mittareita.

Avainsanat

Web-analytiikka, web-mittarit, käyttäjäseuranta

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

The Internet presence of companies has become a crucial instrument of corporate communication and electronic business (Fasel and Zumstein, 2009). Web has become the leading influence in consumer purchasing choices and the most used source of information (Fleishman-Hillard, 2012). With the growing importance of the Internet, the monitoring and optimization of websites and online marketing have become vital tasks for corporations (Fasel and Zumstein, 2009). A company website is an important branding tool and provides direct benefits in terms of e-commerce sales and indirect benefits in terms of offering information and services to various stakeholders (Welling and White, 2006). The potential benefits of e-commerce websites are widely accepted but companies lack systematic and comprehensive methods for measuring and quantifying these benefits (Merwe and Bekker, 2003).

From a supplement media for other channels, corporate websites have risen as an important business channel in their own term and for many companies websites have become the most important channel to reach out for their customers (Phippen et al., 2004). In many ways, corporate websites are unique in comparison to other channels. Unlike other media, like television or newspapers, web is highly measurable (Tonkien et al., 2010). Every click, every visit, and every page view can be counted and companies have access to massive amounts of visitor data. Multiple tools, both free and commercial, have made the recording of website users' behavior fast and easy. Web analytics, the monitoring and evaluation of website usage, has emerged as an active field of research within business intelligence (Chen et al., 2012).

The large amounts of raw user data have introduced another problem: it has become difficult to identify and separate important information from less meaningful data (Phippen et al., 2004). The diverse nature of different websites has further complicated the problem. A website of a public library may have very different objectives compared with an online retailer. It is challenging to say what kind of visitor behavior is the most beneficial: should a visitor view as many different pages as possible and see a lot of site content or would it be better that a site visit is fast and efficient? Despise all the available tools and data, the evaluation of website success is still often based on subjective views and opinions (Kaushik, 2010). Many organizations lack the tools to objective evaluate the performance of their websites. Especially with business-to-business websites that are not generating any direct sales, demonstrating success and return-on-investment has been cited as the top challenge in website management (Weitz and Rosenthal, 2010).

This thesis is proposing a new framework for the role of web metrics analysis in website development. The thesis is conducting a case study to test if the framework can be used to discover meaningful information about website visitors that can be utilized in website development and evaluation. The case study evaluates the value of different web metrics by examining their relationships towards website conversions. The web metrics are also analyzed to see what kind of visitor characteristics the metrics can reveal. The objective of the web metrics analysis is to identify a set of key web metrics. The key metrics are expected to contain the most useful information about the site visitors and about the website itself.

1.1 Web analytics

Web Analytics Association (2008) defines web analytics as the “measurement, collection, analysis and reporting of Internet data for the purposes of understanding and optimizing Web usage”. Web analytics takes advantage of the unique measurability of web sites. Web analytics attempts to gain understanding about website visitors and use this knowledge to improve the effectiveness of websites.

The field of web analytics got started in the early 1990s when public web began to expand (Cooper, 2012). During the past twenty years, web analytics has become a vital part of website development. From the top 500 retail websites, 91% are implementing web analytics to gather information about site users (Hamel, 2012). In spite of being used by practitioners already for two decades, Buckin and Sismeiro (2009) are stating that web analytics is still in the early growth phase of its life cycle. The field keeps evolving and more advanced methods are continuously introduced.

When an Internet user enters a certain website, her actions and behavior on the site can be monitored and recorded by the site owners. Web browsing is based on the exchange of information between a visitor’s web browser and a host web server. Different tools and techniques exist to gather and compile this information exchange. Web analytics tools like Google Analytics have offered ways to all kinds of organization to implement web analytics, no matter their size or resources.

According to Nakatani and Chuang (2011), web analytics is used to “understand online customers and their behaviors, design actions influential to them, and ultimately foster behaviors beneficial to the business and achieve the organization’s goal.” Nakatani and Chuang emphasize that the goal of web analytics is not just to optimize web pages; the ultimate goal is to support the achievement of overall objectives of an organization. As Internet presence of corporations is automatically expected, many companies have set up their websites without clear goals or strategy, just to fulfill the need to establish online presence (Welling and White, 2006). Web analytics can help companies to fully define the purpose and objectives of their websites.

1.2 Benefits of web analytics

The use of web analytics has many potential benefits. Having access to information about the site users helps site developers to make appropriate decisions about the website. The more knowledge available about the visitors, the easier it is to satisfy their needs. Phippen et al. (2004) are stating that web analytics is crucial to the success of websites. Based on their study, companies that are adopting web analytics are reaping multiple benefits and finding invaluable information about their customers.

Web analytics has a powerful capability of providing extensive and comprehensive statistics about users’ behavior on websites (Wang et al., 2011). Web analytics can offer increased website visibility and greater user satisfaction (Plaza, 2010). Measurement of website usage provides a valuable source of customer-centric information about the popularity of a site (Budd, 2012). Web analytics is important for e-commerce websites aimed for consumers but also important for business-to-business websites. Wilson (2010) is stating that a visit to a website is an important component of the buyer-seller relationship. The behavior of visitors needs attention so that a website can become a more productive source of customer leads, customer satisfaction, purchase-related activities, and brand equity. Web analytics can be used to understand visitors’ website usage and navigation patterns and to provide B2B marketing managers with an insightful mechanism for improving website performance (Wilson, 2010). Web analytics plays a vital part of website performance evaluation by offering objective information about the visitor behavior. (Kaushik, 2010)

Web analytics tools provide site owners a number of web metrics that depict users' activities on websites. The tool can tell how many users visited a site during a certain time period, what pages did they view, what kind of media they consumed, what kind of material they downloaded, et cetera (Wang et al., 2011). Ghandour et al. (2010) examined the relationship between web metrics and financial performance. The study found many significant positive correlations between website usage and financial performance. Companies with perceived success in their website usage were also successful when evaluated with financial metrics.

1.3 Challenges of web analytics

Web analytics tools enable the collection of very detailed information about the visitors, including the exact amount of time spend on a certain page, the physical location of a visitor, and the full navigation path of a user. However, is has proven to be difficult to make broadly accepted conclusions about the raw visitor data offered by the tools. Phippen et al. (2004) are saying that "Over time, businesses have begun to find the use of basic metrics such as hits and pages views to be woefully inadequate for assessing the success of Websites, due to the fact that their simplistic and ambiguous nature can induce misleading conclusions."

Kaushik (2010) gives the following example: any web analytics tool will show you what pages on your site are most frequently visited. But based on pure clickstream data, it is difficult to say if these pages really have the content that the visitors are the most interested in. The reason for frequent views might just be a misleading navigation system or unbalanced search engine optimization for different pages. What is further complicating the problem is the diversity of website visitors. Websites are serving multiple stakeholder groups at the same time, different groups having different reasons for visiting a site. Wang et al. (2011) are stating that the average behavior of all site visitors can sometimes be misleading.

Welling and White (2006) conducted 25 interviews with companies of different sizes. Only one of the companies present in the study was attempting to connect website usage with sales or other business goals. The authors state that most companies are doing a poor job of measuring the performance of their websites. Also Cooper (2012) states that there is significant under-exploitation with the possibilities that web analytics can provide.

Plaza (2010) is saying that web analytics efforts often offer too broad and non-strategic information. Colleting and identifying relevant information can be very cumbersome and time consuming. But Plaza is stating that “However, the adoption of key metrics can contribute to reducing time and costs of finding relevant information about a website’s performance.” Wang et al. (2011) are saying that fewer studies have been conducted to examine the relationship among various metrics. The authors state that the interpretation of web metrics may be oversimplified when the effect of intervening variables is not considered. Web metrics should not be examined separately but together with other metrics.

1.4 Definitions

This chapter defines the most important terms used in this thesis. At first, the most basic components and aspects of websites are defined. The definitions in this chapter are taken from Web Analytics Association (2008). After the basic terms, web metrics are shortly discussed. At the end of this chapter, the concept website conversion is defined.

1.4.1 Basic terms

A website consists of multiple *pages*. A page is a definable unit of content in web that can be separated from other pages. Based on the definition, content like flash animations and media files may also be defined as pages even though they differ from traditional pages. *A visit* is an interaction, by an individual, with a website consisting of viewing one or more pages. A single *visitor* can open multiple visits that occur during different times. If a visitor has not visited the site before, the visit he will conduct is called a new visit. If a visitor has been on the site before, his visit is called a return visit. In this thesis, the word *user* is used as a synonym for a visitor. (Web Analytics Association, 2008)

Clickstream data is the primary data source for web analytics. Clickstream data refers to the data collected from users’ activities and behavior on the site (Web Analytics Association, 2008). Every click and interaction on the site become clickstream data that can be recorded with different tools. Clickstream data includes for example the total number of the visitors, the number of page views of all separate pages, the popularity of different traffic sources, and all the product purchasing information from online shops. It is important to remember that clickstream data does not have to be the only data source for web analytics. Other data sources can include for example visitor surveys and questionnaires (Kaushik, 2007).

A traffic source is the origin or the source for a visit (Web Analytics Association, 2008). Every visit has an origin. Traffic sources are usually divided into three groups: search engines, referrals or links from another sites, and direct visits of typing the site URL directly to a browser or opening a bookmark. Following traffic sources is important because it tells about a site’s visibility with search engines and can reveal important partner organizations or affiliate websites (Wang et al., 2011).

Search engine optimization (SEO) means influencing a website’s ranking with search engines. The objective is to make a site’s ranking as high as possible when a search engine user is making a search with certain key words related to the site. SEO is crucial for any website that wishes to attract large numbers of visitors. Web metrics can be used the measure the efficiency of SEO efforts. (Evans, 2007)

1.4.2 Web metrics

The metrics that are generated from clickstream data are called *web metrics*. Some typical web metrics include the amount of new visitors during a certain period and the average number of page views by visitors. Different web metrics reveal different aspects of visitors’ behavior in a website.

Web Analytics Association (2008) defines three different types of web metrics: counts, ratios, and KPIs. A count is the basic unit of measure, a single number. For example, the total number of visits is a count. A ratio is typically a count divided by another count. For example, the average number of page views per visit is a ratio. KPI (key performance indicator) is a count or a ratio that is infused with business strategy and provides meaningful information.

A web metric can be aggregated, segmented, or individual. An aggregated metric relates to the total site traffic for a certain period of time. A segmented metric is a subset of the site traffic for a defined period of time that is filtered in some way, for example based on the visitor type like new visits or return visits. An individual metric relates to the activity of a single visitor for a defined period of time. (Web Analytics Association, 2008)

1.4.3 Conversions

Welling and White (2006) defines website performance as the extent to which a website supports a company to achieve its business objectives, whether they be financial, behavioral, or strategic. Potts (2007) is saying that websites have been traditionally built to facilitate sales but their role has evolved to include functions like customer support, value-adding services, and information sharing to stakeholders such as government agencies and investors. In order to objectively evaluate the achieving of the overall objectives, the objectives should be broken down into more specific goals a company wants its visitors to complete while visiting a website.

A conversion happens when a visitor is completing a specific target action of a website (Web Analytics Association, 2008). A conversion can be for example the ordering a product from an online shop, viewing a specific page, or downloading promotional material. The target actions are called conversion goals. A website can have many different conversion goals. When a visitor completes at least one conversion goal, the visitor is said to be converted.

Tonkin et al. (2010) divides conversion goals into two categories: transaction goals and engagement goals. Transactional goals have direct monetary value, like purchasing a product or becoming a lead by registering an account. Engagement goals relate to a threshold or interaction without direct monetary value, like viewing certain content or spending a desired amount of time on a site. The conversion goals must fully align with business objectives. Tonkin et al. (2010) present three properties for a good conversion goal: the goals need to be measurable, they need to correspond to bottom-line business objectives, and they can be connected to marketing efforts.

In addition to the total number of conversions, *a conversion rate* can be calculated for a website. A conversion rate is calculated by dividing the number of converted visitors by the total number of site visitors. The rate of 100% means that every visitor is completing one or more conversion goals and 0% means that none of the visitors is converting. The total number of conversion can be used as an indicator for the overall performance of a site while the conversion rate tells about the quality of a single visit.

1.5 Research question

A major challenge with web analytics has been the discovering of meaningful information among large amounts of data. This thesis proposes a framework for identifying relevant information by analyzing web metrics together with website conversions. The main research question for this thesis draws closely from the proposed importance of adopting key metrics by Plaza (2010) and the suggested need to analyze the connections between different metrics by Wang et al. (2011). The main research question for the thesis is:

Which web metrics are most closely connected with website conversions?

To test the proposed model, a case study that examines multiple web metrics collected from a corporate website is conducted. The case study evaluates which web metrics have the closest relationship with website conversions. By assessing the relationships with conversions, the objective is to identify a smaller set of key web metrics that are most valuable for the website developers to follow. The relationships are examined using correlation analysis and regression analysis.

The supporting research question for the thesis is:

What information different web metrics reveal about the characteristics of website visitors?

In addition to the relationship analysis, this thesis proposes how web metrics can be analyzed to gain insights about the characteristics and behavior of website visitors. The case study conducted in this thesis will observe what kind of information different web metrics reveal about the visitors and about the website itself. The main and supporting research questions are closely related with each other. The web metrics with the closest connections to website conversions are expected to contain the most important and relevant information about the website visitors.

1.6 Structure of the thesis

This thesis is structured in the following way: after an introduction for the thesis, the field of web analytics is introduced. The potential benefits of web analytics and the current challenges of the field are briefly discussed. After this, the most important terms used in the thesis are defined. Based on the current challenges of web analytics, the research question for this thesis is presented at the end of the first chapter.

The second chapter is a literature review of the field of web analytics. At first, the most common general approaches towards website development and evaluation are discussed. After this, a set of previous studies dealing with web analytics and clickstream data are summarized. Previous frameworks for web analytics are discussed at the end of the second chapter.

The third chapter introduces the new framework for the role of web metric analysis in website development and evaluation. Correlation and regression analyses, the methods used to evaluate the relationships between the web metrics and conversion, are shortly discussed. The chapter also introduces Google Analytics, the web analytics tool used to calculate the web metrics.

The fourth chapter presents the case study conducted in this thesis. At first, the website examined in the study is introduced. After this, the web metrics examined in the study and the conversion goals of the website are discussed. At the end of the chapter, hypotheses for the relationships between different web metrics and conversions are presented.

The fifth chapter presents the results of the case study. After a brief overview of the site usage, the web metrics are analyzed by groups. Based on the web metrics analysis, the key metrics for the website are identified. The fifth and last chapter presents the conclusion for the thesis.

2. LITERATURE REVIEW

This chapter gives a brief literature review on the field of web analytics. At the beginning of the chapter, the most common general approaches towards website development and evaluation are summarized. After this, a set of previous studies dealing with web analytics and clickstream data are discussed. The last part of the chapter introduces earlier frameworks for web analytics and some connections between the different models are identified.

2.1 Website development and evaluation

The development and evaluation of websites has been a popular subject of academic studies. Tan et al. (2009) are stating that websites can be examined from many different viewpoints and site effectiveness assessment always depends on the perspective of the evaluator. The authors propose that website effectiveness can be evaluated from user-related, function-related, or investor-related view. User-related models concentrates on user-focused factors like website usability and customer satisfaction. Function-related models examine the architectural design and technical quality of websites. Investor–related models concentrate on the operational performance of websites and evaluate how well a website is supporting the overall business objectives of a company.

Based on the examination of previous research, Chiou et al., (2010) proposes that there are three common approaches to website evaluation and development: IS-approach, marketing-approach, and combined-approach. Studies using IS-approach concentrates on technical factors, such as ease of use, visual design, information quality, and site navigation structure. Many studies with IS-approach present heuristics for good site design and user-friendly interface. IS-approach is closely linked with user-related and function-related models proposed by Tan et al. (2009). Marketing-approach focuses on the commercial aspect of websites and examines web users as potential customers. Marketing-approach is connected with investor-related view presented by Tan et al. (2009). Combined-approach studies combine both IS and marketing elements on their models.

User-related view and IS-approach towards websites have been used for example in the study of website design elements by Palmer (2002). The study proposes five website design elements that are related to website usability. The elements include navigation structure, site content, interactivity, responsiveness, and loading time of a site. Website developers need to pay attention to all these design elements and find appropriate metrics to measure the performance of the elements. Investor-related view and marketing-approach are applied for example in the study of the decision-making process of online customers by Soonsawad (2013). The study examines the relationship between website components and the customer decision processes. The objective is to gain understanding how to turn site visitors into purchasers.

Tan et al. (2009) states that the IS success model of DeLone and McLean (1992) has been often used as a theoretical background for website evaluation studies. For example, Schaupp et al. (2009) presents a model of website success that is an extension of the original IS success model. The model presents three elements (subjective norms, information quality, and system quality) that are related to two different website success measures (individual impacts and website satisfaction). Other common theoretical frameworks include the Technology Acceptance Model (TAM) by Davis (1989) and the field of human-computer interaction (HCI). These theoretical backgrounds are connected more closely with the IS-approach of website development. The studies with marketing-approach draws more from the field of marketing than information system science. For example, the 4S web-marketing mix model by Constantinides (2002) is based on the 4Ps (product, price, place, and promotion) marketing mix by Borden (1964). The four dimensions of the web-marketing mix by Constantinides are scope (includes strategy and objectives), site (contains website browsing experience), synergy (includes integration with other marketing channels), and system (contains the technological aspects of a website).

No matter which approach or viewpoint is applied towards website development and evaluation, site developers need to objectively measure the effectiveness of chosen site elements. Web analytics can be used to measure both user-related, function-related, and investor-related factors. Web analytics can offer metrics that have a clear usability focus, such as the average time on page and the number of page views. It is also possible to produce metrics with technical focus like the average page loading time. Conversions play an essential part of web analytics. Conversions have a strong commercial focus and can be used to measure the success of marketing efforts and operational performance.

2.2 Previous studies of web analytics

Web analytics and clickstream data have been the focus of multiple studies. Bucklin and Sismeiro (2009) divides the studies of website usage behavior examined with clickstream data into three different research themes. The first research theme includes studies of how users browse and navigate within websites, how their behavior change when they visit a site multiple times, and how they respond to site design and structure. The second theme is about how clickstream data is used to evaluate online advertising methods, including banner advertising and paid search results. The third research theme includes studies that examine how clickstream data can be used to analyze the behavior of online shoppers and how online purchases can be predicted. The literature review in this chapter will concentrate on the first theme, browsing behavior.

Wang et al. (2011) used Google Analytics to examine an educational website. The objective of the study was to identify important behavior characteristics of different visitors. Based on the traffic sources, three visitor segments were created: search engine traffic, direct traffic, and referral traffic. The study also examined whether the users of the website behaved differently during weekdays and weekends. The study followed three web metrics: the number of visits, pages viewed per visit, and the average time on site. The study used the multivariate analysis of variance to examine the relationship between the web metrics, different traffic sources, and the day of the week. Significant differences with page views and time on site were discovered between visitors from different traffic sources. Visitors from direct traffic stayed a significantly longer time on the site and viewed more pages than visitors from other sources, especially during weekdays. Visitors acquired through search engines viewed the least amount of pages and spent the shortest amount of time on the site. In general, visitors viewed more pages and spent more time on the site during weekdays than during weekends. From the total site traffic, 60 percent was direct traffic. Based on the high portion of direct traffic and results from the relationship analysis, the study suggests that the site has a largely purposeful and loyal user-base. (Wang et al., 2011)

Plaza (2009) used Google Analytics to examine different traffic sources for an academic website. When Wang et al. (2011) were using traffic sources only as a segmentation method, Plaza made use of web metrics to assess the effectiveness of the traffic sources. Plaza evaluated the effectiveness based on return visit behavior and the length of a visit. Based on the time series analysis of the study, return visitors navigate deeper into the website than new visitors. The study is stating that the more pages viewed and the more time spent on a site, the more valuable a visit is. Return visitors are hereby more desired than new visitors. The traffic source that nurtures the most return visits is the most effective source. For this particular website, most return visitors entered the site by typing the site’s URL directly to their browser or opening a bookmark, making direct visits the most effective traffic source. The study recommends promoting direct site visits over the other traffic sources for the site. The study also noticed a relationship between bounce rate and return visits rate: the lower the bounce rate, the higher the rate of return visits. (Plaza, 2009)

Pakkala et al. (2012) used Google Analytics to examine three food composition websites based on different European countries. The study period was five months. The study followed multiple web metrics, including, for example, the bounce rate, the depth of visit, the rate of return visits, and visitor loyalty. By analyzing the search engine keywords used to enter the sites and user amounts, the study discovered that the audience of the sites evolved from food composition oriented professionals towards more general audience. Based on the content, the study divided the pages of the sites into different groups and examined the number of page views and time spend on pages by groups. Even though the visitors viewed many navigation and search pages, the study proposes that navigation was not a problem on the sites because the time spent on these pages was low. Dividing pages into groups differs from the approach of Plaza (2009) who proposed that page stickiness is always a desired quality, no matter the page. The study states that access through search engines can be considered a good success indicator for the websites. Search engine traffic is clearly linked with site popularity. The study talks about the snowball effect with websites: the more traffic you have, the more traffic you will gain. (Pakkala et al., 2012)

Singh et al. (2011) used a set of fifteen web metrics to analyze a university website over a period of five months. The study focused on how to utilize web metrics to acquire as many visitors to the site as possible. Metrics like the number of page views, the number of files accessed, and number of entry pages were examined. The study identified the periods with the least amount of traffic and pointed out the need to find ways to acquire more visitors during the slow periods. The study also examined what are the most common exit pages (the page last visited by a visitor before leaving the site). It was proposed that these pages are driving away the visitors and efforts to improve the pages are needed. (Singh et al., 2011)

Arendt and Wagner (2010) used Google Analytics to examine the usage of a university library website. The purpose was to study if and how clickstream data can be used in site redesign.

Based on the analysis of the most popular pages, the visibility of these pages within the site and especially on the front page was increased to make sure it is easy for the visitors to navigate to these popular pages. By analyzing user navigation paths, some of the links on the front page were discovered to be confusing. During the redesign, these links were made more distinct from each other. The study states that clickstream data "provided facts to assist in the decision-making process rather than relying on staff members' opinions and conjecture alone". The study also states that the collected data sometimes showed conflicting patterns between different metrics and utilizing these metrics with site redesign was challenging. Not all of the metrics used in the study were meaningful. The use of the site decreased sharply during academic break times but this information did not help to improve the site. (Arendt and Wagner, 2010)

Chiang et al. (2010) examined how different web metrics are related to each other and to the amount of Web 2.0 features used on the sites under evaluation. The results of the study indicate that the increasing number of visitors will result in even more new and return visits. When a website becomes more popular, a single visitor is more likely to visit the site more frequently. This result is similar to the snowball effect proposed by Pakkala et al. (2012). The total number of visits was positively associated with the average number of page views: the more visits, the more page views by a single user. The average number of page views was in turn negatively related to the average time on page: the more page views, the shorter the time spent on pages. The study states that shorter time on pages implicates more efficient navigation structure and clear site content. Websites that provide an easy and fast browsing experience attract site visitors to navigate deeper. This is a contradiction towards the approach of Plaza (2009) who proposed that longer time on site is always more desired than short visits. The more Web 2.0 features, the higher the average page views and lower the average time spent on pages. Web 2.0 includes features like user reviews, blogs, video sharing, and social networking. The study underlines that it is crucial for especially smaller companies with less knows brands to continuously increase the visibility and popularity of their websites. (Chiang et al., 2010)

Budd (2012) used Google Analytics to evaluate the effectiveness of a website by investigating different traffic sources. Unlike Plaza (2009) who evaluated the value of traffic sources based on the amount of return visits they generate, Budd linked traffic sources to website conversions and investigated what is the most effective traffic source to achieve conversions. For the site examined in the study, Google search engine was the traffic source that produced the biggest number of total conversions but did not have the highest comparative conversion rate. Another search engine had much lower total amount of conversions but had significantly higher conversion rate. The author recommends concentrating on promoting the traffic sources with highest conversion rates. (Budd, 2012)

Weitz and Rosenthal (2010) examined the value of a business-to-business website by analyzing the relationship of web metrics, like the number of visits and time per page, and financial measures, like sales, profits, and amount of goods sold. Data was examined over a period of 15-months that included an email promotion campaign. The campaign drastically increased the number of visits but the study did not find a relationship between increasing visits and sales revenues. As the number of visits increased, the average page views and average time on site decreased. The average number of inquiries via website stayed the same. Based on the web metrics, the study states that the new visitors acquired through the campaign only took a quick look of the site, viewed a small number of pages, and did not complete any conversion goals of the site. (Weitz and Rosenthal, 2010)

Wilson (2010) analyzed the user amounts and navigation paths of a B2B website by examining clickstream data. The purpose of the study was to examine if conversion rates can be improved by making certain changes to the site. The study conducted three field experiments by changing some aspects of the site and keeping other aspects the same. The changes included adding additional information, announcing free shipping in different parts of the site, improving in-site navigation, and reducing the steps in the checkout process. The study then examined how the changes affect the navigation paths towards conversions. By making these changes, the site owners were able to increase the conversion rate and to see which changes were the most effective. The study proposes that this demonstrates that clickstream data and a web analytics software provide a useful combination of tools that can be used to improve a site's conversion rate and enhance the performance of B2B websites. (Wilson, 2010)

2.3 Frameworks for web analytics

Figure 1 introduces the general framework of *the major components of an online marketing system* by Tonkin et al. (2012). The model emphasizes that a website does not exist in isolation but is part of an interconnecting system that is used to advance a company's business goals. The purpose of web analytics is to measure and analyze all the key components of e-commerce in order to facilitate the achieving of the overall business objectives.

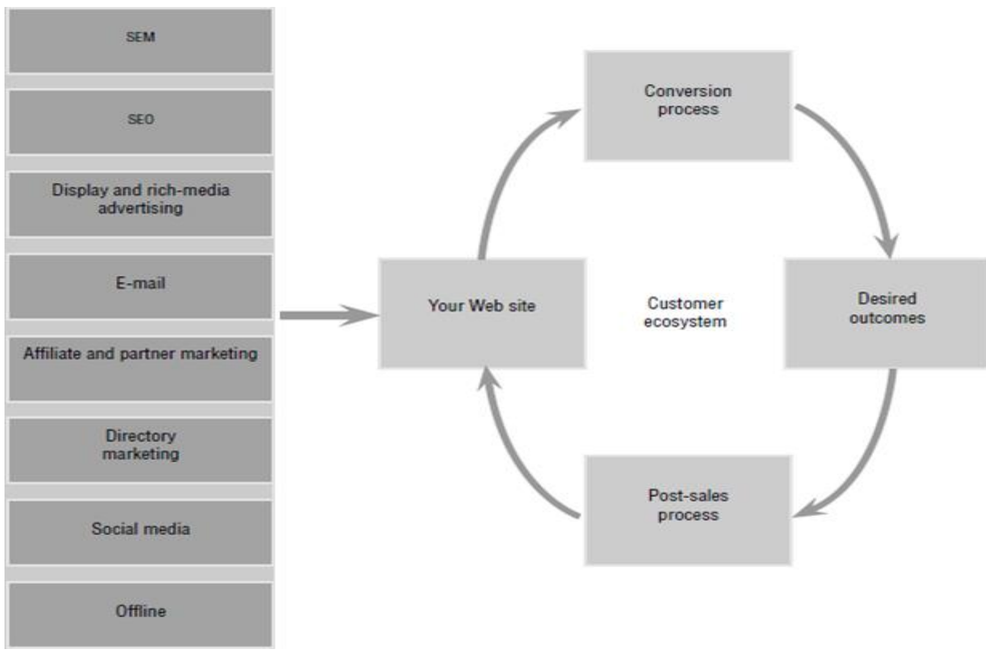

Figure 1. The major components of an online marketing system (Tonkin et al., 2012)

On the left side of the framework, there are eight *inbound marketing channels*. These channels represent ways for engaging with potential customers and attracting new visitors to the site. Websites are directly linked to these acquisition channels. Web analytics can be used to analyze the relationship between the channels and a website. With web analytics, it is possible to quantify the value of different channels and calculate measurements like return on investment. On the right side of the framework, there is the *customer ecosystem,* the virtuous cycle of acquiring new customers and converting them to repeat buyers. This ecosystem includes the two most important processes for analytical marketing: the *conversion process* and the *post-sales process.* The conversion process is about achieving the conversion goals of the site that are directly linked with business objectives. The post-sales process includes the activities for maximizing repeat business. Strengthening the customer ecosystem by increasing the amount of repeat buyers and increasing the conversion rate of the visitors is a key value proposition of web analytics. (Tonkin et al., 2012)

Figure 2 presents *the Trinity approach* by Kaushik (2007). The objective of the approach is to use web analytics to achieve actionable insights and metrics. Instead of mere reporting, the goal is to achieve genuine understanding about visitors that can drive strategic differentiation and sustainable competitive advantage.

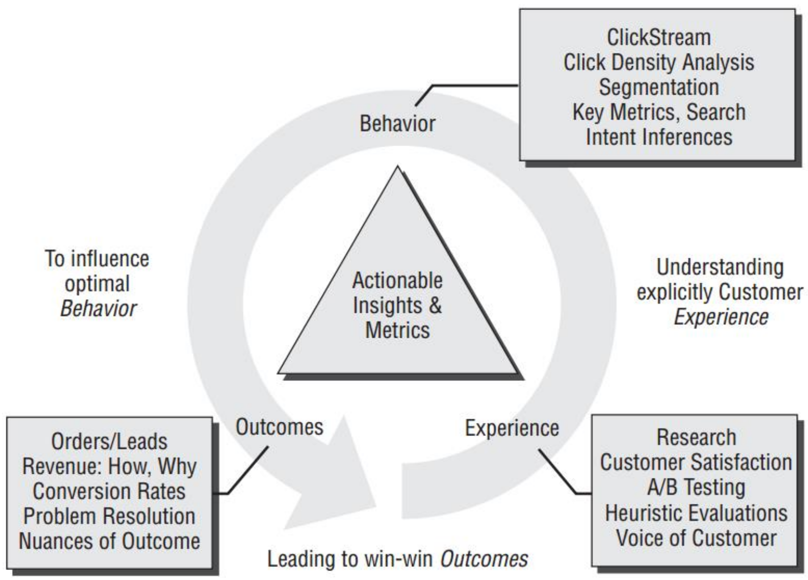

Figure 2. Trinity approach (Kaushik, 2007)

The Trinity framework has three components, first component being *behavior analysis*. The main data source for behavior analysis is clickstream data but raw clickstream data of visitor behavior does not offer much information. However, when clickstream data is used for click density analysis, visitor segmentation, and identifying key metrics, it is possible to gain valuable understanding about visitor behavior. Second component, *outcome analysis*, is about how well the website is achieving its goals. Every website needs to have clearly defined objectives. For pure online retailers, outcomes can be measured with sales revenues or purchases per visit. For websites that are not generating any direct sales, outcomes can be the number of lead generated, customer problems solved, or the amount of product information shared. The achieving of these objectives can be measured by conversions. Third component of the Trinity framework is *experience analysis*. The objective of experience analysis is to get into the heads of visitors and gain insight about the reasons for their actions and behavior. Behavior analysis is about *what* the visitors do while experience analysis is about *why* the visitors do the things they do. Experience analysis tools include for example visitor surveys, A/B testing methodology, and lab usability testing. (Kaushik, 2007)

Both Tonkin et al. (2012) and Kaushik (2007) are connecting website usage with website conversions. In order to increase the number of conversions, understanding about the visitor behavior is needed. Both models are stating that conversions do not need to be transactional goals that are measured with financial measures. The conversion goals can also include engagement goals that are related to driving interest and awareness, for example the sharing of promotional material through a site.

Zheng et al. (2012) approach web analytics and website performance using the concept of web attention. During the time of information overflow, a successful website needs to efficiently catch the attention of the site visitors. The authors present a framework in which website performance can be directly evaluated by examining different web metrics.

.

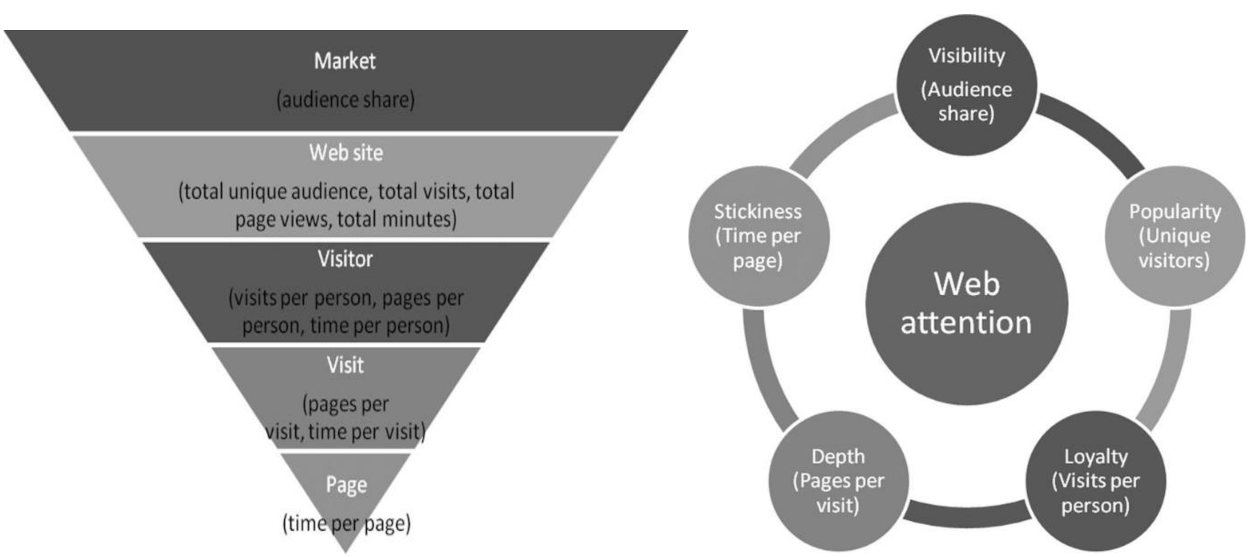

Figure 3. Levels of analysis & Five-dimensional model of web attention (Zheng et al., 2012)

The framework is based on five different levels of analysis: *per market, per website, per visitor, per visit*, and *per page* (Figure 3). Per visit is on the higher level than the per page because one visit can consists of multiple page views, per visitor is on the higher level on the hierarchy than per visit because one visitor can pay multiple visits to the site, et cetera. Based on the different levels of analysis, a framework of *five-dimensional model of web attention* is presented. The purpose of the framework is to present a comprehensive model that integrates the multiple dimensions of web usage. The model aims to serve as an effective tool for capturing the multiple dimensions of web attention and to distinguish and benchmark different websites by their performance across these dimensions. The model proposes five distinct dimensions of web attention: *visibility, popularity, loyalty, depth*, and *stickiness.* Each dimension tries to capture one level of analysis and each dimension can be measured with a certain web metric. Because of the complexity of online attention, none of the measures, when used independently, can provide a comprehensive view of a website usage website but a multidimensional model is needed. (Zheng et al., 2012)

To evaluate the effectiveness of a website, Peacock (2002) presents a *hierarchy of web user needs* (Figure 4). Just like with the framework of Zheng et al. (2012), the hierarchy of web user needs proposes that website effectiveness can be evaluated by examining clickstream data. Many connections between these two models can be seen. The hierarchy of web user needs has four tiers that map a visitor’s navigation through a website. At each level, there are a set of log diagnostics (web metrics) which can be used to measure visitor satisfaction.

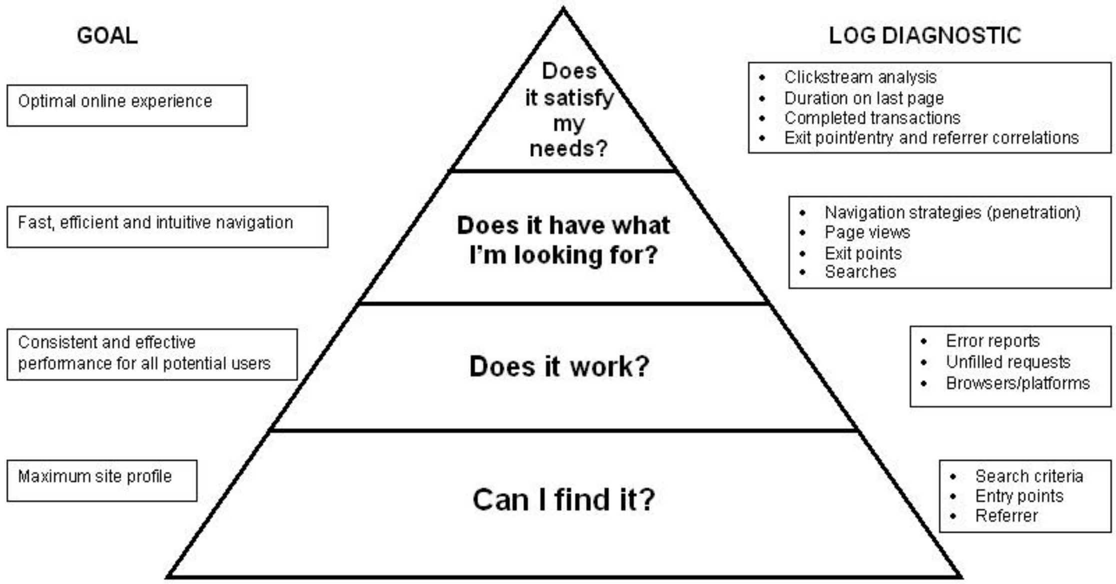

Figure 4. Hierarchy of web user needs (Peacock, 2002)

Level 1 of the hierarchy is concerned with the traffic sources of a website. For this level, direct traffic, links to the site, and search engine traffic are analyzed together with the popularity of entry pages. The goal is to maximize the number of desired visitors. This level is closely connected with the dimensions of popularity and visibility by Zheng et al. (2012). Level 2 deals with the technical aspects of a site, including site loading speed and reliability when visited using different browsers. The objective is to make sure a site is working properly during every visit. Level 3 examines the efficiency of the internal navigation of a site. The goal is to make sure that the visitors will find the information they need fast and easy. The web metrics used for this level match with the dimensions of stickiness and depth by Zheng et al. (2012). Level 4 analyses the overall effectiveness of a website. When a visitor enters a site, he has some kind of need. The objective of a website is to satisfy that need. The model suggests that the correlation of the exit pages with the original referrer or search terms can be used to assess if the needs of a visitor have been achieved. For example, if a visitor entered the page using a search term "jobs" or through a link from a job listing website, it can be examined how fast the visitor found the careers section of the site and if the visitor was satisfied with the section. The visitor satisfaction and the dimension of loyalty by Zheng et al. (2012) share many similarities. (Peacock, 2002)

Figure 5 presents the framework of *the relations between web and e-business metrics* by Fasel and Zumstein (2009). The model is targeted especially for online retailers but can also be applied to other types of websites. This framework has many similarities with the previous models introduced in this chapter.

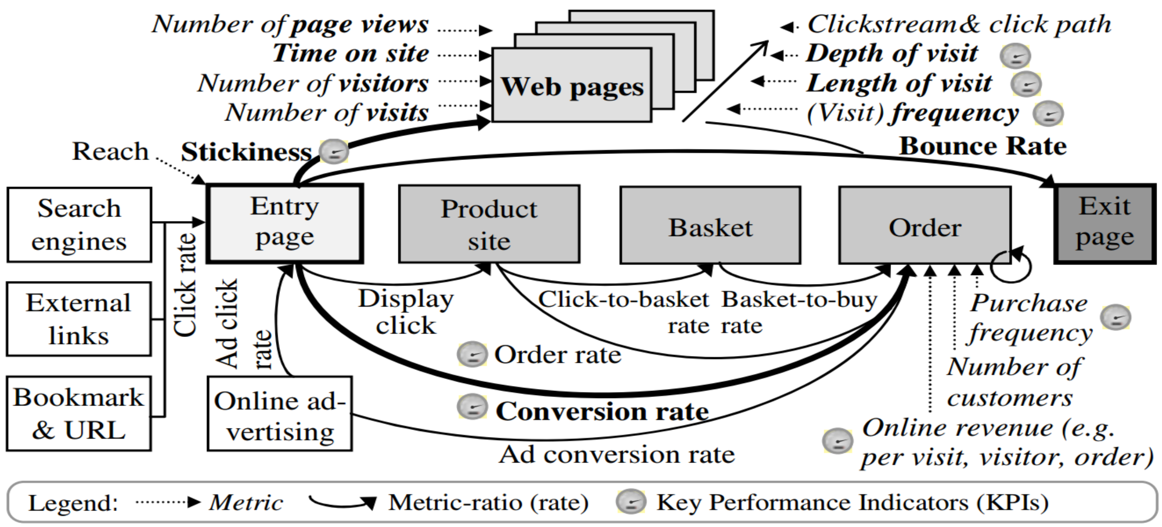

Figure 5. The relations between web and e-business metrics (Fasel and Zumstein, 2009)

The framework begins from the different traffic sources that lead to an entry page of a site. The traffic sources are associated with the inbound marketing channels of Tonkin et al. (2012) and with the first level of the hierarchy of web user needs by Peacock (2002). The traffic sources can be analyzed to see which of the channels are working the most efficiently. After an entry page, a visitor then navigates through a site and leaves from an exit page. If the entry page and the exit page are the same, this is measured by the bounce rate. Based on visitor behavior, different web metrics can be generated. Metrics include for example the number of visitors, the depth of a visit, the length of a visit, and visit frequency. The framework is utilizing the same web metrics that Zheng et al. (2012) is using to measure stickiness, depth, loyalty, and popularity. Based on the metrics, different ratios or rates can be calculated, the most important being the conversion rate. The framework proposes different KPIs that are more important than standard metrics, some being metrics and some being ratios. Web metrics and e-business metrics can be combined into KPIs. The mixing of these metrics can help to connect website objectives into overall business objectives. This focus on objectives is closely associated with the outcome analysis by Kaushik

(2007). These combined KPIs include, for example, online revenue per visit and online revenue per unique visitor. The framework also presents the navigation funnel for a product purchase, starting from a product page, then advancing to a shopping basket, and finally arriving at an order confirmation page. The steps of this navigation funnel can be analyzed with different ratios to see in which phase the visitors are dropping out of the funnel. (Fasel and Zumstein, 2009)

Singh et al. (2011) presents a practical model for web analytics that concentrates on how organizations can implement web analytics in practice. Figure 6 presents the framework for *web analytics process.* The authors say that "web analytics is not a technology to produce reports; it is a process of virtuous cycle for website optimization."

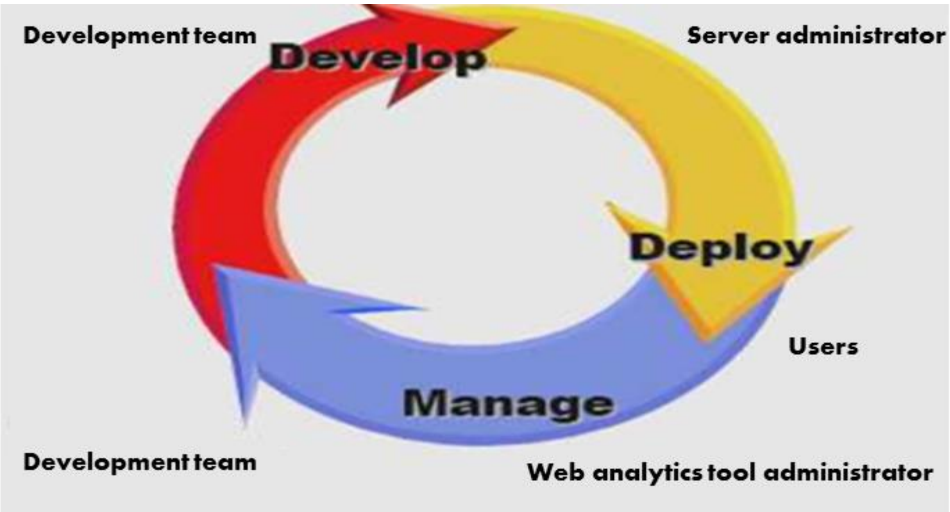

Figure 6. Web analytics process (Singh et al., 2011)

This framework proposes that three different teams are needed for efficient web analytics. The first team, *web analytics tool administrator team*, controls and uses the selected web analytics tools to collect and report the data collected of the website users' behavior. *The development team* use the data provided by administrator team to analyze and understand website visitors. This analysis is the basis for the continuous development and improvement of a website. *The server administrator team* deploys new web pages and applications produced by the development team. The team makes sure that a website works fast and without any technical problems. (Singh et al., 2011)

3. RESEARCH FRAMEWORK AND METHODS

This chapter presents the research framework, quantitative research methods, and research tools used in the thesis. At first, the new framework based on previous research is introduced. After this, correlation analysis and regressions analyses are briefly discussed. At the end of the chapter, web analytics tools are discussed in general and the web analytic tool used in this thesis, Google Analytics, is introduced.

3.1 Framework for the study

Figure 7 presents the research framework for *the role of web metrics analysis in website development*. The framework is used in the case study conducted in this thesis. The objective of the case study is to examine if this framework can be utilized to discover useful information about the website visitors and about the website itself.

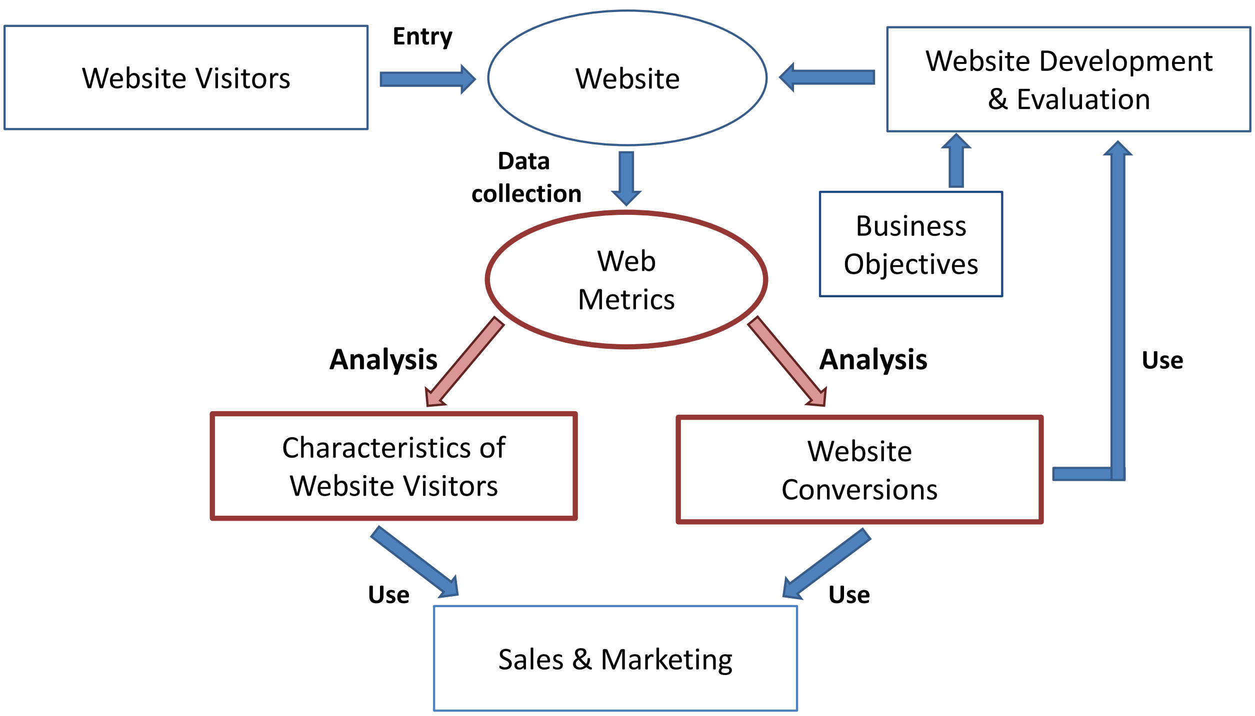

Figure 7. Research framework: The role of web metrics analysis in website development

Websites are subject to continuous development and evaluation. The objectives of a company's website should be aligned with the overall business objectives of the company. When visitors enter a website, their actions on the site can be closely monitored. Based on the browsing behavior of visitors, different web metrics can be calculated. A common approach to analyze web metrics is to examine their connection with website conversions. For example, the frameworks by Tonkin et al. (2012), Fasel and Zumstein (2009), and Kaushik (2007) are linking visitor behavior to website conversions. Another approach to analyze web metrics is to examine what kind of characteristics the web metrics reveal about the visitors. For example, Plaza (2009), Pakkala et al. (2012), and Wang et al. (2011) used web metrics to examine various visitor types, to analyze their behavior, and to make conclusions what different visitor characteristics tell about the effectiveness of websites. The objective of both these approaches is to achieve relevant and useful information about the visitors that can be used for site development and site evaluation. The objective is also to gain insights about the visitors that can be used for sales and marketing purposes. The web metrics that reveal the most important information about site visitors and about a website itself can be considered the most valuable metrics. This thesis is concentrating on the central part of the framework, the web metrics analysis.

3.2 Research methods

The framework is evaluating the relationship between web metrics and website conversions. In statistics, two common methods of examining the relationship between two variables are correlation analysis and regression analysis. Correlation analysis can be mainly used to assess the strength of the relationship between different variables. Regression analysis can be used to assess if the values of some variable or variables can be used to predict the values of some other variable.

3.2.1 Correlation analysis

Correlation analysis can be used to assess the strength of the association between different variables. Correlation analysis also indicates the direction of the relationship, which can be positive or negative. Correlation analysis does not offer any indication if the values of one variable can be used to estimate the values of another variable. With correlation, there is no clear independent or dependent variable. Being the most widely used index of correlation, *Pearson product-moment correlation coefficient* is chosen to be used in this thesis. The coefficient r can have a value between -1 and 1. 1 indicates perfect positive correlation and -1 indicates perfect negative correlation. A coefficient of zero indicates no linear relationship between the variables. (Kirk, 2008)

3.2.2 Regression analysis

The regression analysis examines if one or more independent variables (also called predicators) can be used to estimate the values of a dependent variable. If the values of only one predicator are used, the analysis is called simple regression, and if two or more predicators are used, the analysis is called multiple regression. Linear regression models assume that the relationship between each independent variable and dependent variable is linear. The objective of linear regression is to find the regression line of best fit. The line of best fit means the regression line that has the lowest discrepancy between the values given by the line and the actual values of the dependent variable. (Field, 2009)

In this thesis, multiple linear regression is used. SPSS Statistics is used to calculate the regressions. In the study, the following information gained with regression analyses is presented: regression coefficient B-values, the standard error for coefficient B-values (*SE B*), the standardized beta (β) values, p-values of predicators, and the square of the multiple correlation coefficient (R^2). The B-value tells how much the dependent variable changes if the independent variable increases by one unit. The standard error for the B-value (*SE B*) indicates to what extent the B-values will vary across different samples. The standardized β-value tells the number of standard deviations that the outcome will change as a result of one standard deviation change in the predictor. The values can be used to evaluate the importance of a predicator. Another variable to evaluate the importance of a variable is the p-value of a predicator. If the p-value is less than 0,05, the predicator has a genuine effect on the outcome. The value of R^2 indicates how much of the variance in the outcome can be predicted from the independent variables. The value is between zero and one, one indication that 100% of the variance can be explained from the variables. (Field, 2009)

Multicollinearity means high intercorrelation between the independent variables themselves. High intercorrelation means that the variables contain the same information and using these variables together as predictors with multiple regression analysis is problematic. Strong correlation between the variables can have a large effect on the results of regression analysis so the possible multicollinearity must be taken account when using regression analysis. (Morgan et al., 2004)

3.3 Research tools

3.3.1 Web analytics tools

Web analytics tools collect, process, and store clickstream data and present the data as meaningful information. Many different web analytics tools are available in the market, covering different price ranges and levels of sophistication. The market includes multiple free software which makes the threshold to start web analytics efforts low. Web analytics has become a standard practice for the retail industry and the popularity of web analytics tools keeps increasing (Hamel, 2012b). The four most popular solutions are Google Analytics, Adobe Analytics, IBM Digital Analytics, and WebTrends (Farina, 2013). Tonkin et al. (2010) are stating that web analytics tools can form the backbone of a company's online measurement strategy and serve as the most important tool for understanding the performance of a website.

There are multiple ways to collect clickstream data of visitor behavior. One way to classify web analytics tools is on the basis of data collection methods. Most common data collection methods are web server transaction log analysis and web page tagging. Web log analysis was one of the first ways to collect clickstream data. When a visitor's browser requests a web page from a web server, the request-related data (requestor's IP-address, time of visits, etc.) is recorded in a transaction log file. Web log analysis software can then be used to analyze the log file. Web page tagging collects data with invisible JavaScript code inserted in web pages. The code collects information about the visit and submits the data to a web data collection center or in-house database. The data is then integrated and reported through a web analytics tool. Page tagging is currently the most popular method for data collection. (Nakatani and Chuang, 2011)

3.3.2 Google Analytics

Google Analytics (GA) is Google's web analytics tool that became free for everyone to use in 2006. Since its launch, GA has become very popular with both companies and non-profit organizations. From Fortune 500 companies, 63% are using Google Analytics (Farina, 2013). Google also offers a tool with additional features called Google Analytics Premium, which is available for an annual fee. GA is a web-based system that does not require any software installation. All the data is stored on Google's own servers.

Google Analytics can be used to calculate tens of different web metrics of site visitors. The tool can also be used to further analyze web metrics and to discover site visitor characteristics. GA can also record the website conversions. The tool does not record individual information about the visitors. The visitor data is aggregated and it is not possible to single out the behavior and characteristics of one user. The IP addresses of the visitors are not available, mainly due to potential issues with visitor privacy.

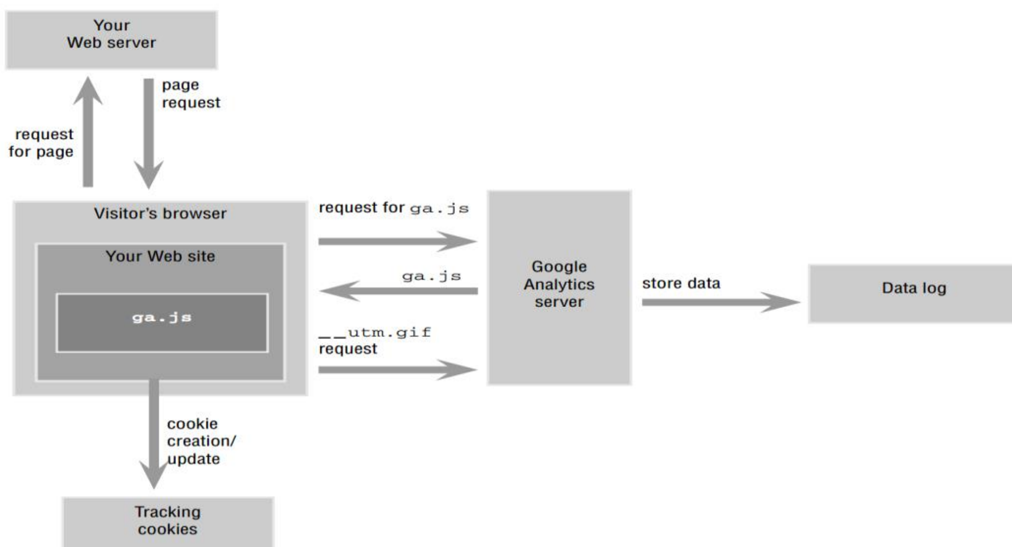

Figure 8. Data collection method of Google Analytics (Tonkin et al., 2010)

Figure 8 summarizes the data collection method of GA that is based on web tagging with Google Analytics Tracking Code (GATC). GATC consists of a string of JavaScript code that is placed on each page of the site. As a visitor's browser loads a page, GATC is automatically executed and the tracking code request a file named ga.js from the nearest server of Google. This ga.js file is a set of instructions that GA uses to track the visitors. After the ga.js file has been delivered to the browser, the second section of the tracking code begins to collect data about the visitor behavior and characteristics. Once all this information has been collected, the tracking code sets (or updates if the user has visited the site before) a number of cookies on the visitor's browser. After the cookies have been set, GATC sends all the collected data back to the Google Analytics servers by requesting an invisible image file named __utm.gif. After receiving this request, GA servers stores the visitor data into data logs. (Tonkin et al., 2010)

The results given by Google Analytics are not fully accurate. GA does not track visitors who have set their web browsers to block first-party cookies or disabled JavaScript. If a visitor removes the cookies from her web browser, she will appear as a new visitor next time when she visits the site. If a visitor uses a different browser or a different computer, he will appear as a different visitor. Not all the visitors of a website are human. Non-human visitors include different robots and crawlers that scan websites for various purposes. GA has ways to separate robots from human visitors but some visitors that GA identifies as human are actually robots. These challenges are not affecting only Google Analytics but other web analytics tools are facing the same difficulties. This thesis does not assess the potential inaccuracies with the collected clickstream data. The collected clickstream data is treated as it would be fully accurate and all the visitors are human visitors. (Tonkin et al., 2010)

4. THE CASE STUDY

This chapter introduces the case study conducted for this thesis. The objective of the case study is to test if the proposed framework for the role of web metrics analysis in website development can be used to identify relevant key web metrics. At the beginning of this chapter, the website examined in the study is introduced. After this, the methods for web metrics analysis used in the study are discussed. After the research methods, the web metrics and conversions goals chosen for the study are defined. Hypotheses for the connections between web metrics and conversions are presented at the end of this chapter.

4.1 Website examined in the study

This study examines a website that offers second-hand telecommunications equipment used in wireless mobile phone networks. The equipment offered through the site can be used as spare parts or as expansion components for GSM and 3G networks. The target audience for the website is mobile phone operators of all sizes around the world. The main focus group is operators who are looking for an inexpensive way to expand their old networks while making large investments to new LTE networks. The site is also serving various resellers, brokers, and other companies who are offering used equipment to mobile operators. The site was launched during the latter half of 2012.

The objective of the website is to promote the products offered and generate leads of potential customers. The site does not have an online shop and the objective of the website is not to make direct sales. The objective is to raise awareness of the company as a new supplier of used telecom equipment. The owner of the website is a mid-sized telecom asset management and consultancy company that is operating mainly in Southeast Asia but equipment sales are targeted to worldwide audience.

The study examines the clickstream data about the visitors’ behavior during a period of six months, from the 1st of June 2013 until the end of November 2013. When the data collection period is only six months, possible seasonal and cyclical factors that can influence the behavior cannot be evaluated and thus not discussed in this thesis. During the study period, the website was under continuous development and new content was published rapidly especially during June, July, and August.

4.2 Methods for web metrics analysis

The case study follows the framework of the role of web metrics analysis in website development. The study concentrates on the central part of the framework, the web metrics analysis. The web metrics are analyzed from two different viewpoints. The first approach is to analyze the relationships between the chosen web metrics and website conversions. The closer the relationship, the more valuable the web metric is proposed to be. The web metric values and the amount of conversions are calculated on daily basis. It is possible to pair the daily values of web metrics and the daily amount of conversions. This enables the use of correlation and regression analyses to evaluate the levels of relationships. The relationships of web metrics are examined towards the total number of conversions and towards the conversion rate. SPSS Statistics is used to conduct the analyses. The second approach is to use the web metrics to analyze the characteristics of visitors based on their browsing behavior. The metrics that contain the most meaningful characteristics are considered the most valuable. The analysis features of Google Analytics are used to examine the web metrics in greater detail.

Based on the web metrics analysis, the goal is to identify a smaller set of key web metrics that are the most important for the site owners to follow. The key metrics are expected to contain the most valuable information about the site visitors and about the website itself. If relevant key web metrics can be identified for the site under examination in the case study, the proposed framework can be considered useful for organization trying to improve their websites.

4.3 Web metrics and conversion goals

A set of fourteen simple web metrics are chosen to be examined in the study. The metrics are calculated with Google Analytics. When selecting the web metrics, the criterion for the selection was to choose the most common metrics that are frequently used in web analytics study books and in academic research. The metrics include both counts and ratios. Table 1 lists and defines the web metrics examined in the study. The selected web metrics are compared towards the total number of conversions and towards the conversion rate. The website has three different conversion goals that are calculated together as the total number of conversions. Table 2 presents the conversion goals used for the site.

Table 1. The web metrics used in the study

#	Web Metric	Definition	Type
1	New Visits	The number of visits conducted by visitors who visited the site the first time.	Count
2	Return Visits	The number of visits conducted by visitors who have visited the site before. New Visits + Return Visits = Total Number of Visits	Count
3	Rate of Return Visits	The rate of how many of the visits are return visits. Rate of Return Visits = Return Visits / Total Visits	Ratio
4	Average Page Views	The average number of pages viewed during a visit. The repeated views of a page by the same user are counted.	Ratio
5	Total Page Views	The total number of pages viewed. The repeated views of a page by the same user are counted.	Count
6	Unique Page Views	The number of visits during which the specified page was viewed at least once.	Count
7	Average Time on Page	The average amount of time visitors spent viewing a single page.	Ratio
8	Bounce Rate	The percentage of single-page visits.	Ratio
9	Search Engine Traffic	The number of visit generated by search engines.	Count
10	Referral Traffic	The number of visit generated by referrals or links from other web sites.	Count
11	Direct Traffic	The number of visit occured when the site's URL is directly typed into a browser or a bookmark of the site is clicked.	Count
12	Impressions on Google	The number of times the site appeared on Google search results pages viewed by a user of Google search engine.	Count
13	Clicks on Google	The number of clicks to the site on Google search results pages.	Count
14	Average Position on Google	The average ranking of the site on Google search results pages.	Ratio

Table 2. The conversion goals of the website

#	Conversion Goal	Definition	Type
1	Catalogues Downloaded	The number of visitors who downloaded at least one product catalogue during one day.	Count
2	RFQs Sent	The number of Request For Quotations sent through the website.	Count
3	Contact Forms Sent	The number of contact forms sent through the website.	Count

Based on their resemblance, the web metrics are divided into five different groups (Figure 9). The groups are new and return visits, page views, time and bounce rate, traffic sources, and search engine optimization (SEO) metrics. All the groups contain three web metrics except the third group, time and bounce rate, which contains only two metrics. The case study will analyze the web metrics by groups.

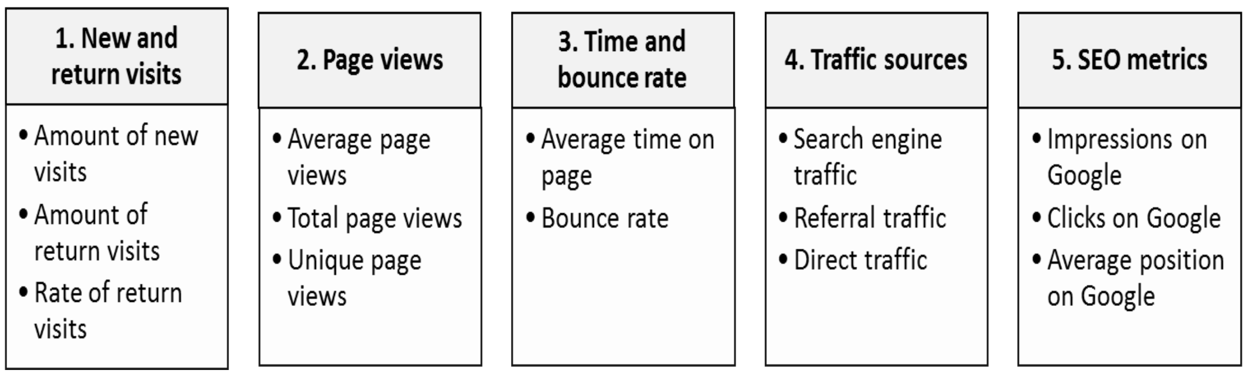

Figure 9. Web metrics by groups

4.4 Hypothesis

The relationships of the web metrics are examined towards the number of total conversions and the conversion rate. Some web metrics are expected to be correlated more closely with the total number of conversions and some more closely with the conversion rate. Clear differences with the correlations are expected between different web metrics. Only some of the metrics are expected to have a significant relationship with both the total number of conversions and the conversion rate. *Multicollinearity* is expected between the different web metrics used in this study so the potential intercorrelations between metrics need to be considered when conducting multiple regressions analyses.

It is expected that the amounts of new visits and return visits are closely correlated with the total number of conversions: the more traffic on the site, the more conversions. It is not expected that there will be a strong relationship between traffic levels and the conversion rate. As the results by Budd (2012) are indicating, increasing amount of visits does not automatically mean that an average visitor will also complete conversion goals more often. However, a different scenario is also possible. Returning visitors have been said to be more valuable than new visitors (Plaza, 2009). It is possible that if the rate of return visits is increasing, the conversion rate will also increase. Based on the results of Chiang et al. (2010), if the number of total visits will increases, the rate of return visits is expected to increase. This would mean a positive relationship between total traffic levels and the conversion rate.

The total number of page views and the unique page views are expected to be correlated with the total number of conversions. With the unique page views, the relationship is expected to be slightly stronger because unique content could be more valuable than content that is already seen by a visitor before. The average number of page views is expected to be correlated with the conversion rate. The framework of Zheng et al. (2012) proposes that average page views can be used to measure the depth of visits. The deeper an average visitor navigates to the site, the more likely he is expected to be converted. The average number of page views has also been criticized for having high potential of being misleading because of the different nature of websites (Kaushik, 2010). For some websites, relevant content is distributed on many different pages and many page views are preferred. But for some sites, an efficient visit that consists only of a few pages is desired. The study by Park and Chung (2009) proposed that the lower the number of page views, the more likely a user of a travel service website is to make a purchase. A visitor's interest towards specific content means fewer page views and increased change for a visitor to be converted.

The average time on page is expected to have a minor relationship with the total number of conversion and the conversion rate. Zheng et al. (2012) proposes that average time per page can be used to measure the stickiness of a site. The longer a visitor is viewing site content, the more likely he is expected to be converted. It is also possible that the opposite turns out to be true. Chiang et al. (2010) connected a shorter time on site with efficient site content and structure. The faster a visitor can find the information she needs, the more likely she is to be converted. This would mean that a small negative correlation would exist between the conversion rate and the average time on page. The bounce rate is expected to be strongly correlated with the conversion rate. According to Kaushik (2010), the bounce rate is a simple and instant measure of success. A high bounce rate is expected to be connected with a low conversion rate.

It is expected that search engine traffic is the most common traffic source and thus most closely correlated with the total number of conversions. It is expected that the number of visitors who are coming to the site directly is closely correlated with the conversion rate. Plaza (2009) and Wang et al. (2011) have stated that direct visits are more desired than visits through links or search engines. The visitors who come to the site directly have some previous knowledge of the site and they have a specific reason to visit the site.

Impressions and clicks on Google search engine are expected to have strong positive correlation with the total number of conversion. These metrics are closely associated with search engine optimization. The framework of Tonkin et al. (2012) is stating SEO as one of the inbound marketing channels of driving site awareness. Better site visibility with search engines is expected to result in increased traffic. Jansen and Spink (2006) say that 73 percent of search engine users never look beyond the first page of search engine results so the average position on Google is expected to have a strong negative correlation with the total number of conversions.

5. RESULTS OF THE CASE STUDY

This chapter presents the results of the case study. At the beginning of the chapter, an overview of the site usage is given. After this, the fourteen web metrics are examined and their relationships towards the total number of conversions and towards the conversion rate are evaluated. The five groups of web metrics are examined one group at a time. Based on the web metrics analysis, the study identifies the key metrics that are the most valuable to follow.

5.1 Overview of the site usage

The data collection period was six months, from June 1st to November 30th, 2013. Figure 10 presents an overview of the site visitors. The graph in the figure shows the daily number of total visits during the study period. The figure is taken from Google Analytics. The average number of monthly visits was increasing until October, which was the busiest month during the study period. The busiest day was November 14th with 120 visits. Compared with the potential world-wide customer base, the amount of visits has remained modest.

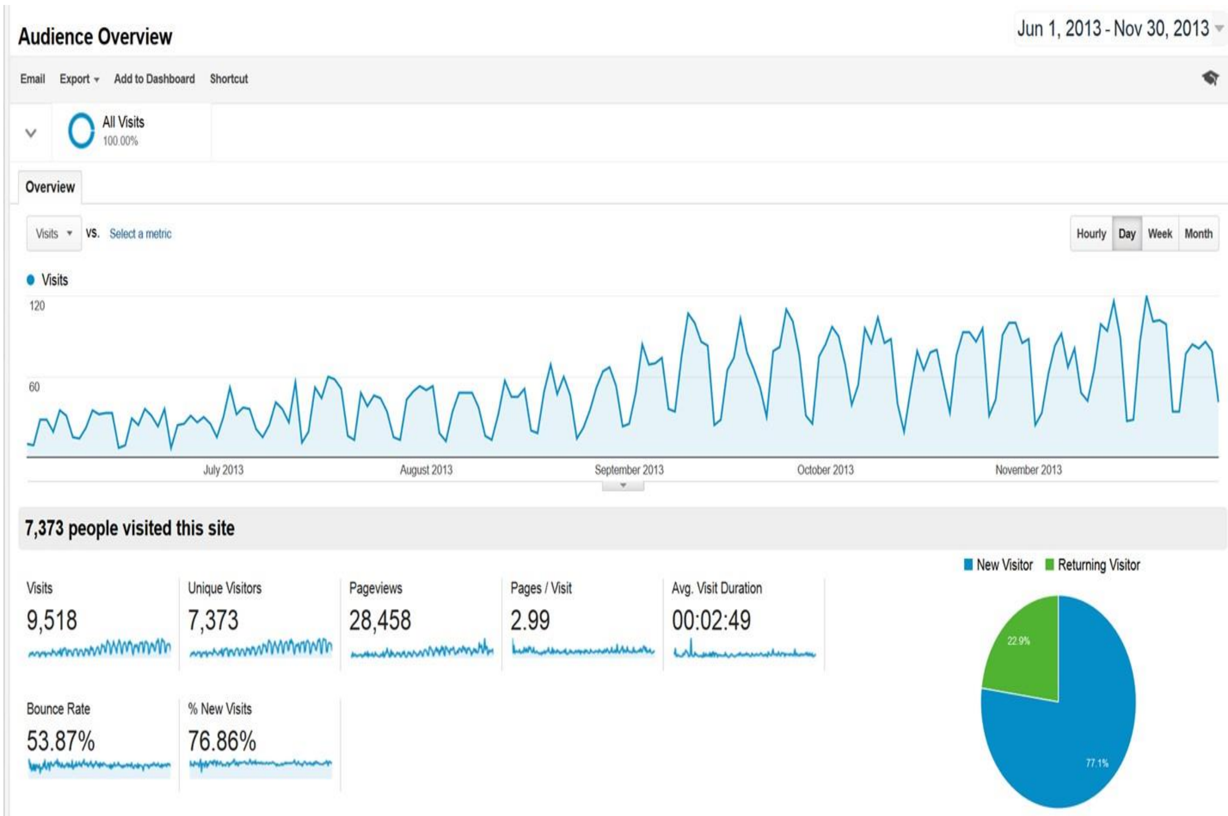

Figure 10. Audience overview, June 1st to November 30th (Google Analytics)

Figure 11 presents the daily amount of conversions during the study period. Just like with the daily visits, the average number of monthly conversions was growing until October. The day with the most conversions was November 12th with 80 conversions. Figure 12 presents the daily conversion rate during the study period. There was an increasing trend with the conversion rate until October, the average conversion rate of October being 67%.

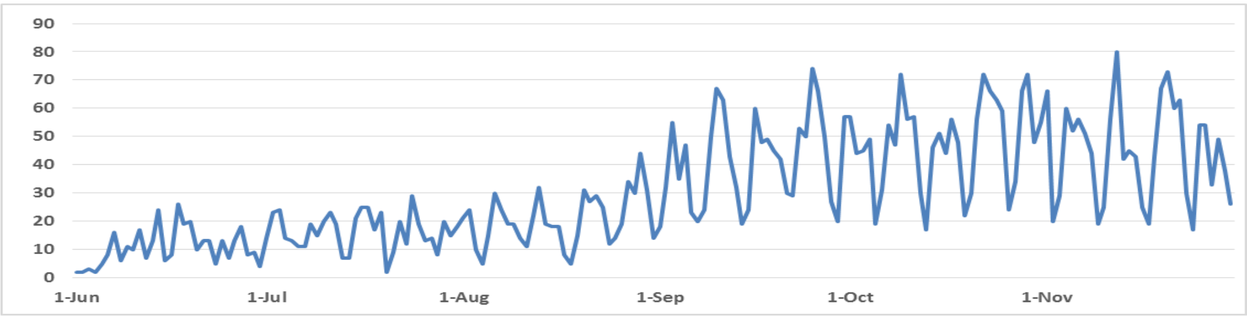

Figure 11. Daily amount of conversions, June 1st to November 30th

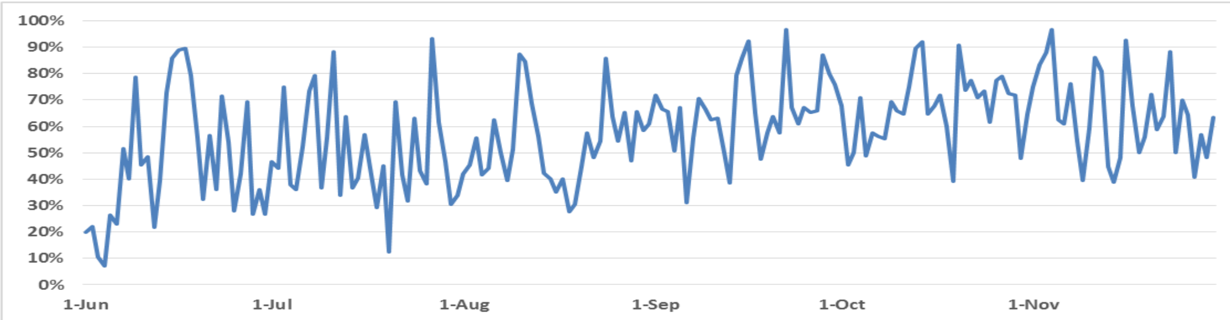

Figure 12. Daily conversion rate, June 1st to November 30th

Table 3. Monthly visits and conversions

MONTH	# OF TOTAL VISITS	# OF CONVERSIONS	CONVERSION RATE
JUNE	722	318	46 %
JULY	1113	510	50 %
AUGUST	1246	639	53 %
SEPTEMBER	1990	1252	66 %
OCTOBER	2285	1490	67 %
NOVEMBER	2162	1339	65 %

The rising amount of conversions means that the site has been fulfilling its objectives more effectively, i.e., the performance of the site has been increasing. The site has successfully increased both the total number of conversions and the conversion rate. Based on the conversions, it can be said that the site owners have been developing the site in a right direction. However, only following the amount of conversions does not offer insights why has the amount decreased or increased. It is necessary to examine how the visitor behavior has changed and what kind of behavior is most closely associated with conversions.

It is worrying to see that during the last month of the study period, both the total number of conversions and the conversion rate were lower than during the previous month. This breaks the promising trend of increasing conversions. The lower numbers might be just normal monthly fluctuation and the increasing trend will continue during the next month. However, if the conversions keep decreasing also during the following months, it is a cause for serious concern.

There is a large difference with the daily number of visits depending on the weekday of the visit (Table 4). The number of visits during Saturdays and Sundays is clearly lower than during other days. During the busiest day (Wednesday) the average number of visits is three times higher as during the slowest day (Sunday). Also the average number of conversions is lower during weekends. The fluctuations with the conversion rate are much more random than with the total amount of conversions. Sundays have the largest average conversion rate and Wednesdays have the lowest. The fluctuation remains high during the whole study period. The differences with the user amounts raise the question if different kinds of users visit the site between weekends and weekdays. The site is targeted to corporate buyers who might be visiting the page mainly on workdays. The study will also examine the data with Saturdays and Sundays excluded to see whether any differences with the full data could be found.

Table 4. Site traffic during different weekdays

WEEKDAY	AVERAGE # OF DAILY VISITS	AVERAGE # OF DAILY CONVERSIONS	AVERAGE DAILY CONVERSION RATE
MONDAY	54	35	62 %
TUESDAY	67	40	56 %
WEDNESDAY	68	37	52 %
THURSDAY	65	36	54 %
FRIDAY	62	33	52 %
SATURDAY	27	16	61 %
SUNDAY	23	16	67 %

5.2 The results for web metrics analysis

Table 5 presents the Pearson correlations between the web metrics and conversions. The significant correlations between the web metrics and conversions are highlighted. Most of the web metrics are significantly correlated with the total number of conversions. But only one of the metrics, impressions on Google, is significantly correlated with the conversion rate. Because of the large differences with the amounts of visits during weekdays and weekends, correlations are also calculated with data that excludes Saturdays and Sundays. Correlations for working days only are presented in Table 6. There are clear differences with the correlations whether weekends are included or excluded, especially with the conversion rate. When weekends are excluded, seven of the web metrics are significantly correlated with the conversion rate. Differences with correlations mean that the behavior of visitors varies between weekdays and weekends.

Table 5. Pearson correlations: web metrics and conversions (incl. weekends)

	NEW _VISITS	RETURN _VISITS	RATE_OF _RETURN	AVERAGE _PAGES	TOTAL _PAGES	UNIQUE _PAGES	AVERAGE _TIME	BOUNCE _RATE	TRAFFIC _SENGINE	TRAFFIC _REFERRAL	TRAFFIC _DIRECT	IMPRES SIONS	CLICKS	AVERAGE_ POSITION	TOTAL _CONVERSIONS	CONVERSION _RATE
NEW_VISITS	1	,824**	-,010	-,056	,883**	,917**	-,047	,040	,971**	,664**	,639**	,826**	,860**	-,309**	,896**	,068
RETURN_VISITS		1	,435**	,002	,830**	,860**	,035	-,008	,878**	,625**	,610**	,783**	,784**	-,202**	,807**	,017
RATE_OF_RETURN			1	,187*	,136	,142	,299**	-,203**	,101	,089	,098	,252**	,241**	,149	,068	-,100
AVERAGE_PAGES				1	,308**	,228**	,533**	-,423**	-,060	-,010	,047	,113	,030	-,151*	-,021	,010
TOTAL_PAGES					1	,988**	,170*	-,131	,869**	,612**	,655**	,760**	,755**	-,317**	,824**	,066
UNIQUE_PAGES						1	,122	-,108	,907**	,637**	,655**	,789**	,798**	-,310**	,849**	,058
AVERAGE_TIME							1	-,407**	-,040	,049	-,003	,143	,102	-,073	,012	,105
BOUNCE_RATE								1	,015	,010	,081	-,120	-,096	,009	-,008	-,097
TRAFFIC_SENGINE									1	,586**	,527**	,844**	,891**	-,314**	,912**	,098
TRAFFIC_REFERRAL										1	,384**	,558**	,549**	-,125	,613**	,045
TRAFFIC_DIRECT											1	,462**	,399**	-,108	,472**	-,166*
IMPRESSIONS												1	,900**	-,250**	,879**	,212**
CLICKS													1	-,278**	,879**	,146
AVERAGE_POSITION														1	-,292**	-,069
TOTAL _CONVERSIONS															1	,415**
CONVERSION _RATE																1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

Table 6. Pearson correlations: web metrics and conversions (excl. weekends)

	NEW _VISITS	RETURN _VISITS	RATE_OF _RETURN	AVERAGE _PAGES	TOTAL _PAGES	UNIQUE _PAGES	AVERAGE _TIME	BOUNCE _RATE	TRAFFIC _SENGINE	TRAFFIC _REFERRAL	TRAFFIC _DIRECT	IMPRES SIONS	CLICKS	AVERAGE_ POSITION	TOTAL _CONVERSIONS	CONVERSION _RATE
NEW_VISITS	1	,746**	-,140	,016	,851**	,892**	,077	,114	,962**	,650**	,535**	,778**	,780**	-,507**	,878**	,335**
RETURN_VISITS		1	,496**	,084	,781**	,815**	,139	,076	,822**	,603**	,521**	,706**	,650**	-,360**	,750**	,209*
RATE_OF_RETURN			1	,114	,076	,079	,125	-,062	,001	,071	,114	,104	,046	,112	-,026	-,226**
AVERAGE_PAGES				1	,459**	,360**	,680**	-,554**	,009	,005	,147	,123	,089	-,124	,032	-,056
TOTAL_PAGES					1	,986**	,360**	-,116	,840**	,554**	,578**	,673**	,642**	-,473**	,789**	,274**
UNIQUE_PAGES						1	,294**	-,074	,883**	,595**	,578**	,704**	,694**	-,478**	,815**	,272**
AVERAGE_TIME							1	-,500**	,096	,028	,086	,141	,107	-,099	,124	,083
BOUNCE_RATE								1	,109	,044	,073	-,089	-,114	-,053	,088	,081
TRAFFIC_SENGINE									1	,571**	,395**	,808**	,834**	-,509**	,902**	,368**
TRAFFIC_REFERRAL										1	,309**	,523**	,499**	-,339**	,590**	,194*
TRAFFIC_DIRECT											1	,290**	,146	-,137	,353**	-,054
IMPRESSIONS												1	,861**	-,420**	,839**	,474**
CLICKS													1	-,510**	,822**	,463**
AVERAGE_POSITION														1	-,503**	-,285**
TOTAL _CONVERSIONS															1	,675**
CONVERSION _RATE																1

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

5.2.1 New and return visits

A large majority, 77% of the visits to the site are new visits (Figure 13). Figure 14 presents the daily number of new and return visits. Traffic levels for both visit types have been increasing during the data collection period. The amount of new visits and return visits have a strong positive correlation with each other. The percentage of return visits (the rate of return visits) has stayed low. The rate has fluctuated between 15-35%, without any clear increasing or decreasing trend.

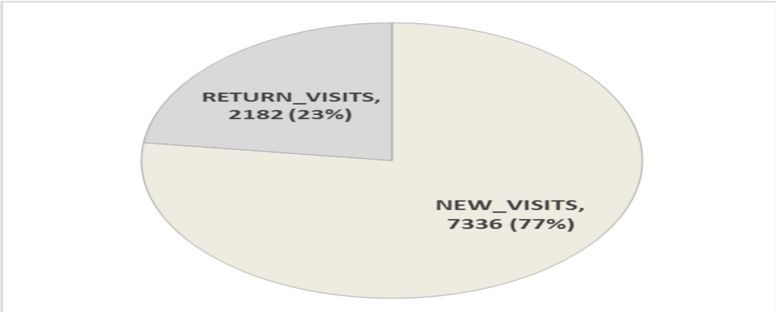

Figure 13. Total amount of new and return visits

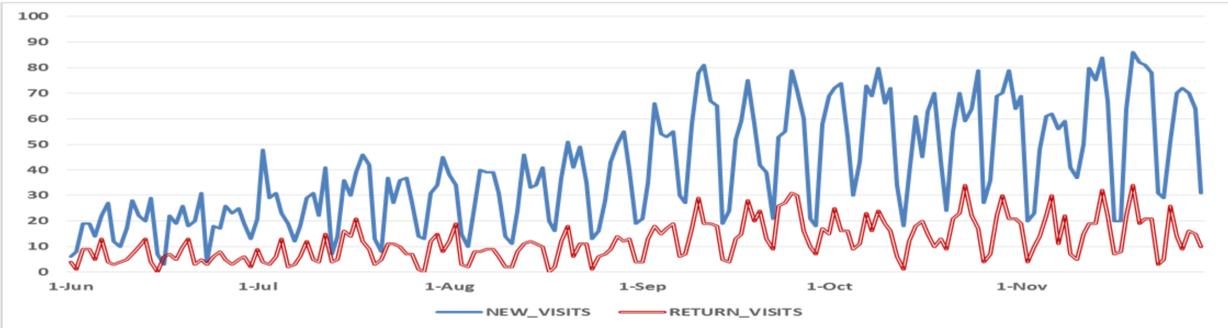

Figure 14. Daily number of new and return visits

As expected, the amounts of both new visits and return visits are strongly positively correlated with the total number of conversions (Table 7). The more traffic there has been, the higher the total number of conversions. However, the relationship with the conversion rate is unclear. When Saturdays and Sundays are included, there is no correlation between new and return visits and the conversion rate. When weekends are excluded, there exists small positive correlations between traffic levels and the conversion rate, the correlation being stronger with new visits. The rate of return visits is not correlated with the total number of conversions. When weekends are excluded, a small but significant negative correlation between the rate of return visits and the conversion rate is visible.

Table 7. Pearson correlations: visits and conversions

PEARSON CORRELATIONS	SATURDAYS AND SUNDAYS INCLUDED		SATURDAYS AND SUNDAYS EXCLUDED	
	TOTAL CONVERSIONS	CONVERSION RATE	TOTAL CONVERSIONS	CONVERSION RATE
NEW VISITS	,896**	,068	,878**	,335**
RETURN VISITS	,807**	,017	,750**	,209*
RATE OF RETURN VISITS	,068	-,100	-,026	-,226**

***Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).**

The relationship between new and return visits and the total number of conversions seems clear. The rate of return visits is not connected with the total amount of conversions. The relationship of these metrics with the conversion rate is much more uncertain. Regression analysis is used to see if any additional information can be achieved. Table 8 shows the results of multiple regression analysis when new visits and the rate of return visits are analyzed towards the conversion rate. The amount of return visits is excluded from the analysis because the metric is strongly correlated with both new visits and the rate of return visits.

Table 8. Regression analysis: new visits, rate of return visits, & conversion rate (excl. weekends)

DEPENDENT VARIABLE: CONVERSION RATE			
INDEPEND. VARIABLE	B	SE B	β
CONSTANT	,536	,066	
NEW VISITS	,003	,001	,310**
RATE OF RETURN VISITS	-,476	,216	-,183*

Note: R^2 = ,145 *p < ,05 **p < ,001

The results suggest that it is important to follow both new visits and return visits because of their strong correlation with the total number of conversions. Based on the regression analysis, the amount of new visits can also be used as a predicator towards the conversion rate. Both correlation and regression analyses suggest that the rate of return visits has a negative relationship with the conversion rate. This contradicts with the hypothesis that expected a positive relationship with the conversion rate. New visitors seem to be more likely to convert than returning visitors, making new visitors a more valuable visitor type. Table 9 shows the bounce rate, the average page views, and the average time on page for new and return visits. Plaza (2009) stated that returning visitors are more valuable because they spend a longer time on the site and view more pages. Just like Plaza proposed, the return visitors did spend a longer time on the site and viewed more pages but this has not facilitated return visitors to convert more often than new visitors.

Table 9. Behavior characteristics of new and returning users (Google Analytics)

Visitor Type	Behavior		
	Bounce Rate	Pages / Visit	Avg. Visit Duration
New Users	54.18% Site Avg: 53.87% (0.59%)	2.94 Site Avg: 2.99 (-1.79%)	00:02:36 Site Avg: 00:02:49 (-7.51%)
Returning Users	52.80% Site Avg: 53.87% (-1.99%)	3.17 Site Avg: 2.99 (6.02%)	00:03:31 Site Avg: 00:02:49 (25.25%)

The reason for contrary results between this study and previous research might be that the site examined in this study has not been able to attract a large enough group of loyal visitors who frequently visit the site. Figure 15 shows the count of visits based on how many times a visitor has visited the site before. The majority of the users visited the site only once and most of the return visits are from users who have visited the site just one time before. Only 14% of the visitors have visited the site three times or more. It seems that a larger group of loyal visitors is needed in order to see the differences between new and returning users more clearly.

Count of Visits	Visits	Pageviews
1	7,336	21,541
2	1,022	3,538
3	356	1,147
4	174	538
5	98	300
6	64	175
7	45	166
8	37	65
9-14	112	284
15-25	81	181
26-50	89	211
51-100	60	205
101-200	44	107

Figure 15. Count of visits by all visitors (Google Analytics)

5.2.2 Page views

The number of page views has increased closely together with the increased site traffic. The total number of page views and the unique number of page views are closely correlated with new the amount of new visits and return visits. The total page views and the unique page views are also almost perfectly positively correlated with each other. Unlike the total amounts of page views, the average number of page views has not increased during the study period. The daily average page views per visit has been around three pages and no trend in any direction is visible (Figure 16).

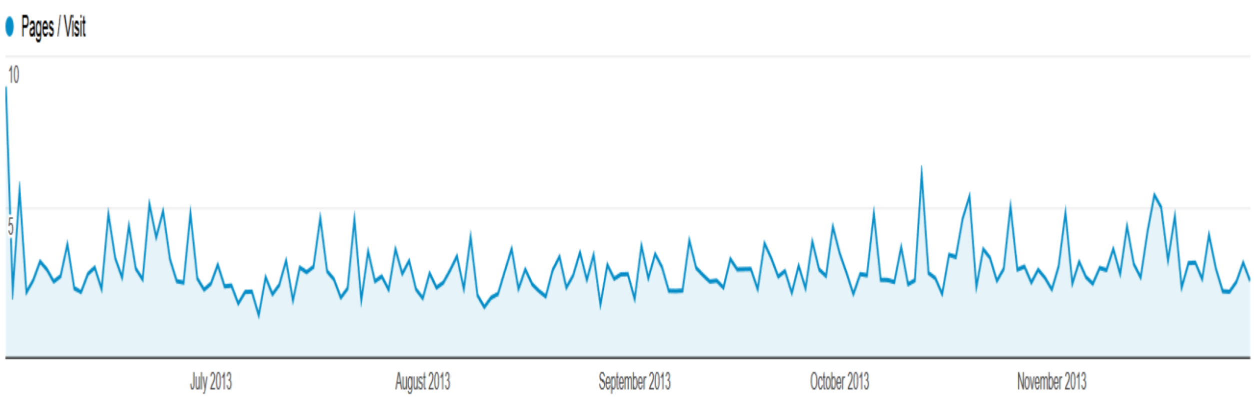

Figure 16. Daily average number of page views during a visit (Google Analytics)

The total and unique page views are both closely positive correlated with the amount of total conversions (Table 10). When weekends are excluded, there exists small but still significant positive correlations also with the conversion rate. The average number of page views does not have any correlation with conversions, no matter if weekends are included or excluded.

Table 10. Pearson correlations: page views and conversions

PEARSON CORRELATIONS	SATURDAYS AND SUNDAYS INCLUDED		SATURDAYS AND SUNDAYS EXCLUDED	
	TOTAL CONVERSIONS	CONVERSION RATE	TOTAL CONVERSIONS	CONVERSION RATE
TOTAL PAGE VIEWS	,824**	,066	,789**	,274**
UNIQUE PAGE VIEWS	,849**	,058	,815**	,272**
AVERAGE PAGE VIEWS / VISIT	-,021	,010	,032	-,056

***Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).**

Even though total page views and unique page views are closely correlated with the total number of conversions, it seems that following these metrics might not be very useful. Following the page view metrics in addition to new and return visits does not offer much additional information. Based on the hypothesis, unique page views would be more closely correlated with conversions but the results suggest that there are no differences between the total number of page views and the unique page views. The views of unique pages do not seem to be more valuable than the views of content that has already been seen by a user.

The average number of page views per visit does not have any correlation with conversions. It seems that the daily average value does not offer much information but it cannot be said that this metric would be fully meaningless. Using the average number of page views as a method for visitor segmentation could reveal relevant information about the site. Figure 17 shows the number of page views for all the visits. Most of the visits consists of only one or two page views. Following the behavior and navigation paths of visitors who navigate more deeply into the site might offer information about why most of the visitors leave the site just after one or two pages but some visitors keep browsing on the site.

Page Depth	Visits	Pageviews
1	5,127	5,127
2	1,540	3,080
3	867	2,601
4	518	2,072
5	317	1,585
6	212	1,272
7	179	1,253
8	125	1,000
9	106	954
10	94	940
11	59	649
12	55	660
13	42	546
14	38	532
15	34	510
16	24	384
17	21	357
18	15	270
19	12	228
20+	133	4,438

Figure 17. Depth of visits (Google Analytics)

Table 11 shows the ten most popular landing pages of the site. There are differences with the average pages per visit based on what was the first page that a visitor viewed. The average number of page views per visit can be used to compare the different pages of the site and examine their differences. The most popular landing page is the main page for the site (number 1 on the table). Despite being the most popular landing page, it is interesting to see that only 10% of the visitors enter the site through the main page. It is vital to make sure that visitors can easily begin to navigate through the site also from other pages than from the main page.

Table 11. Ten most popular landing pages of the site (Google Analytics)

Landing Page	Acquisition			Behavior
	Visits	% New Visits	New Visits	Pages / Visit
	9,518 % of Total: 100.00% (9,518)	77.08% Site Avg: 76.86% (0.27%)	7,336 % of Total: 100.27% (7,316)	2.99 Site Avg: 2.99 (0.00%)
1.	941	55.26%	520	3.00
2.	394	88.32%	348	2.85
3.	382	80.89%	309	1.85
4.	309	75.40%	233	2.88
5.	267	84.27%	225	2.73
6.	259	85.71%	222	3.20
7.	245	69.39%	170	2.61
8.	227	78.41%	178	2.96
9.	224	89.29%	200	1.93
10.	223	81.61%	182	2.89

It could be useful to examine the page views of single pages to see if any large fluctuations can be found. Figure 18 shows the daily number of page views for one sample page. The daily page views for this page have been constantly changing and a few spikes of daily views are clearly visible. Examining if large increases or decreases with the page views of a single page will have a connection with conversions might reveal information about the value of a page. If a connection towards conversions can be found, this indicates that a page has valuable content.

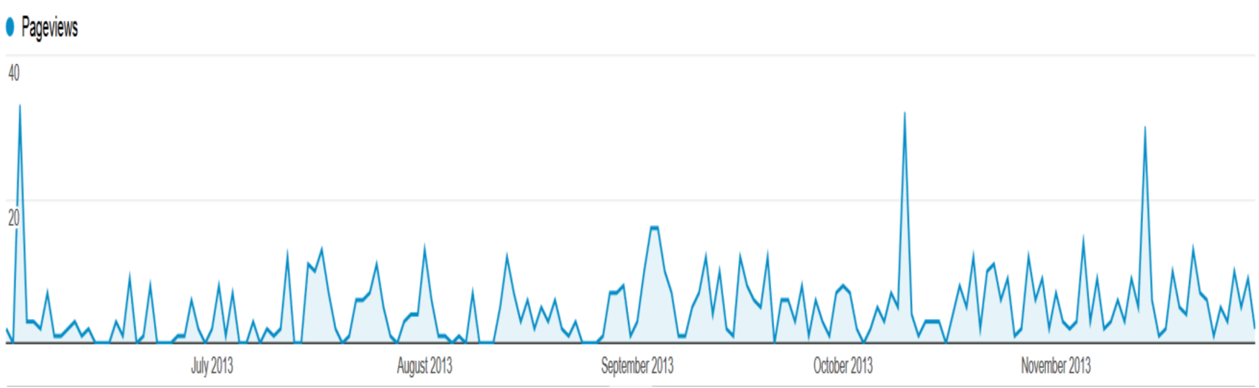

Figure 18. Daily page views for a sample page (Google Analytics)

5.2.3 Time and bounce rate

Figure 19 shows the daily average time on page and the daily average bounce rate. No clear trends with these metrics are visible. During the study period, the average bounce rate was 54%. The average time on page fluctuated notably during the first months but got more stable during the latter half of the study period, daily average for the last three months being three minutes on page.

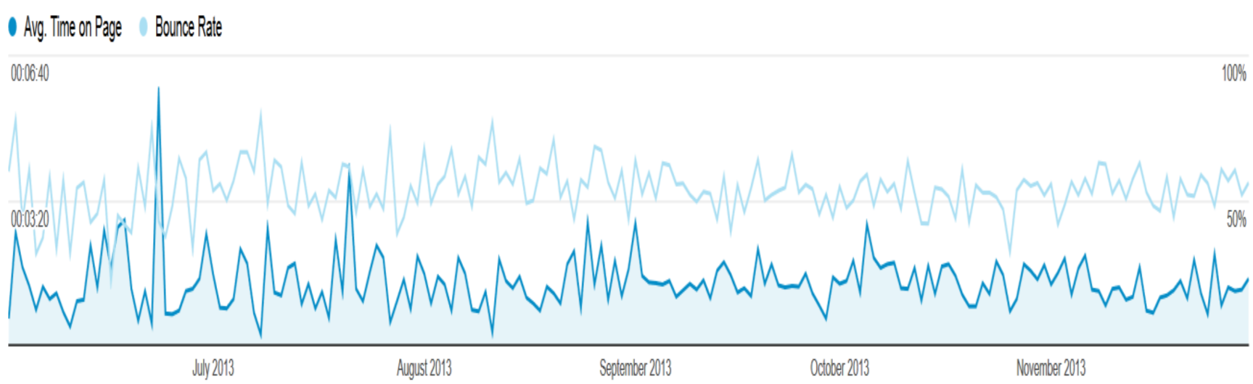

Figure 19. Daily average time on page and bounce rate (Google Analytics)

The average time on page and the bounce rate are negatively correlated with each other. Both metrics are correlated with the average number of page views, the average time on page positively and the bounce rate negatively. The average time on page and the bounce rate are not correlated with the total number of conversions nor the conversion rate (Table 12). As expected, regression analysis proposes that neither of these metrics can be used as predictors for the total number of conversions (Table 13). The *p*-values are under 0,05 but the value of R^2 is only 0,045 which means the amount of variance that these metrics can explain is minimal.

Table 12. Pearson correlations: average time on page, bounce rate, & conversions

PEARSON CORRELATIONS	SATURDAYS AND SUNDAYS INCLUDED		SATURDAYS AND SUNDAYS EXCLUDED	
	TOTAL CONVERSIONS	CONVERSION RATE	TOTAL CONVERSIONS	CONVERSION RATE
AVERAGE TIME ON PAGE	,012	,105	,124	,083
BOUNCE RATE	-,008	-,097	,088	,081

***Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).**

Table 13. Regression analysis: average time on page, bounce rate, & conversions (excl. weekends)

DEPENDENT VARIABLE: TOTAL AMOUNT OF CONVERSIONS			
INDEPEND. VARIABLE	**B**	**SE B**	**β**
CONSTANT	-,491	16,142	
AVERAGE TIME ON PAGE	,061	,028	,224*
BOUNCE RATE	49,762	24,963	,200*

Note: R^2 = ,045 *p < ,05

It seems that following the average values of these web metrics does not offer valuable insights about the visitor behavior. The results contradict with the hypothesis that expected both the average time on page and the bounce rate to be correlated with the conversion rate. Even though no relationships with conversions are visible with the average values, one cannot make a conclusion that these metrics would be meaningless. Table 14 shows the ten most visited pages on the site. Examining the average time on page and the bounce rate at page level reveals differences with these metrics between different pages. Comparing pages with low bounce rates or low average time on page to pages with high bounce rates or high average time on page might reveal insights how to improve the pages.

Table 14. Ten most visited pages on site (Google Analytics)

Page	Pageviews	Unique Pageviews	Avg. Time on Page	Entrances	Bounce Rate	% Exit
	28,458 % of Total: 100.00% (28,458)	20,233 % of Total: 100.00% (20,233)	00:01:25 Site Avg: 00:01:25 (0.00%)	9,518 % of Total: 100.00% (9,518)	53.87% Site Avg: 53.87% (0.00%)	33.45% Site Avg: 33.45% (0.00%)
1.	**1,399**	1,116	00:01:58	941	49.52%	45.03%
2.	**1,195**	764	00:01:12	132	40.91%	19.50%
3.	**1,123**	676	00:01:43	309	54.69%	29.83%
4.	**992**	690	00:00:57	207	43.96%	25.81%
5.	**887**	531	00:01:29	195	43.59%	25.48%
6.	**828**	544	00:01:39	394	51.02%	43.36%
7.	**808**	544	00:01:26	223	57.40%	33.17%
8.	**779**	472	00:01:18	210	44.76%	27.21%
9.	**712**	480	00:01:23	209	48.33%	30.76%
10.	**704**	553	00:02:21	382	68.06%	54.83%

When comparing the pages, it is important to remember that not all the pages are similar. For the main page of a site, the average time on page might not be so relevant but following the bounce rate could be crucial. It is important that the visitors will not see only the front page but also navigate deeper to the site to see the actual site content. However, the visitors do not need to spend a long time on the front page. For pages with detailed product information, the bounce rate can be high if the average time on page is longer. If a visitor arrives at the exact page he was looking for and spends a long time examining the product details, a page is efficiently designed even if the visitor will only see this one page. Pages can be grouped to different segments based on their content and different key metrics can be defined for different segments.

Figure 20 shows the number of visits grouped by the visit duration. The majority of the visits lasted only ten seconds or less. Examining monthly data reveals a positive correlation with the visits that lasted 60 seconds or more and the total number of monthly conversion. Segmenting the visits by duration and examining the browsing patterns of different segments might reveal usable information about the site visitors.

Visit Duration	Visits	Pageviews
0-10 seconds	5,347	5,653
11-30 seconds	510	1,178
31-60 seconds	606	1,678
61-180 seconds	1,162	4,564
181-600 seconds	1,105	6,938
601-1800 seconds	654	5,737
1801+ seconds	134	2,710

Figure 20. Visits grouped by visit duration (Google Analytics)

5.2.4 Traffic sources

Figure 21 shows the number of visits from different traffic sources during the data collection period. Search engines were the most common source for visits, generating 7 406 visits or 78% of the visits. Referral traffic generated 9% of the visits and 13% of the visitors arrived to the site directly. All the traffic sources are positively correlated with each other. As traffic has increased from one source, so has the number of visits from other sources.

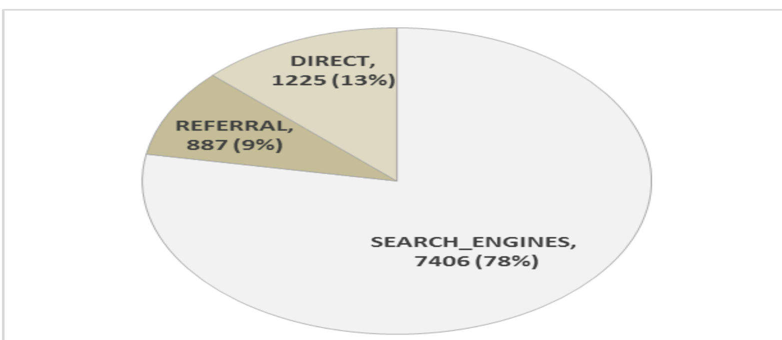

Figure 21. Number of visits by traffic sources

Table 15 shows the acquisition channels that have generated the most visits. Google search engine was the dominant source for visits, generating 73% of the visits to the site. The amount of visits from other search engines is marginal, only five percent of the visits. Google.com is generating the most referral traffic. Referrals from Google.com mean traffic generated by the services of Google other than the search engine, for example Google Groups, Google Base, or Google Maps.

Table 15. Most common acquisition channels (Google Analytics)

Source / Medium	Acquisition			Behavior		
	Visits	% New Visits	New Visits	Bounce Rate	Pages / Visit	Avg. Visit Duration
	9,518 % of Total: 100.00% (9,518)	77.08% Site Avg: 76.86% (0.27%)	7,336 % of Total: 100.27% (7,316)	53.87% Site Avg: 53.87% (0.00%)	2.99 Site Avg: 2.99 (0.00%)	00:02:49 Site Avg: 00:02:49 (0.00%)
1. google / organic	6,934	78.17%	5,420	53.85%	3.02	00:02:52
2. (direct) / (none)	1,225	76.65%	939	56.00%	2.90	00:02:40
3. yandex / organic	234	65.38%	153	56.41%	3.12	00:02:03
4. google.com / referral	105	86.67%	91	45.71%	3.28	00:03:21
5. yahoo / organic	96	85.42%	82	47.92%	3.11	00:02:50
6. google.fr / referral	86	81.40%	70	23.26%	4.35	00:03:45
7. bing / organic	57	87.72%	50	66.67%	2.23	00:02:32
8. vk.com / referral	46	2.17%	1	82.61%	1.17	00:02:25
9. google.co.in / referral	43	79.07%	34	34.88%	3.35	00:02:48
10. forum-volgograd.ru / referral	41	92.68%	38	92.68%	1.15	00:00:04

As expected, the rising number of visits from all the traffic sources is closely positively correlated with the total number of conversions (Table 16). The correlation between search engine traffic and total conversions is especially strong. When Saturdays and Sundays are included, there is no positive correlation between any of the traffic sources and the conversion rate. It is surprising to see that there exists a small negative correlation between direct traffic and the conversion rate. This means that the more direct visits to the site, the less likely an average visitor has been to convert. When weekends are excluded, search engine traffic and referral traffic are positively correlated also with the conversion rate. Direct traffic does not have a significant positive nor negative correlation towards the conversion rate.

Table 16. Pearson correlations: traffic sources and conversions

PEARSON CORRELATIONS	SATURDAYS AND SUNDAYS INCLUDED		SATURDAYS AND SUNDAYS EXCLUDED	
	TOTAL CONVERSIONS	CONVERSION RATE	TOTAL CONVERSIONS	CONVERSION RATE
SEARCH ENGINE TRAFFIC	,912**	,098	,902**	,368**
REFERRAL TRAFFIC	,613**	,045	,590**	,194*
DIRECT TRAFFIC	,472**	-,166*	,353**	-,054

***Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).**

Table 17 shows the regression analysis results when the traffic sources are analyzed at the same time towards the conversion rate. The results suggest that search engine is a good predicator, direct traffic is a minor predicator, and referral traffic cannot be used as a predicator for the conversion rate. The β-value of direct traffic is negative, which suggest a negative relationship between direct traffic and the conversion rate. Because of the intercorrelation between the traffic sources, these results should be observed with caution.

Table 17. Regression analysis: traffic sources and conversion rate (excl. weekends)

DEPENDENT VARIABLE: CONVERSION RATE			
INDEPEND. VARIABLE	B	SE B	β
CONSTANT	,433	,036	
SEARCH ENGINE TRAFFIC	,004	,001	,459***
REFERRAL TRAFFIC	,0002	,004	,005*
DIRECT TRAFFIC	-,008	,003	-,237**

Note: R^2 =0,183 *p > ,1 **p < ,01 ***p < ,001

Regression analysis for the traffic sources is also done towards the total number of conversions (Table 18). The results suggest that search engine traffic is the dominant predictor. Referral traffic is a minor predictor and direct traffic cannot be used to predict the total amount of conversions. Because of the intercorrelation, too wide-ranging conclusions can’t be made from this analysis.

Table 18. Regression analysis: traffic sources and total amount of conversions (incl. weekends)

DEPENDENT VARIABLE: TOTAL AMOUNT OF CONVERSIONS			
INDEPEND. VARIABLE	**B**	**SE B**	**β**
CONSTANT	-1,255	1,208	
SEARCH ENGINE TRAFFIC	,723	,034	,854***
REFERRAL TRAFFIC	,612	,184	,123**
DIRECT TRAFFIC	-,098	,137	-,025*

Note: R^2 = ,842 *p > ,1 **p < ,01 ***p < ,001

The hypothesis predicted that direct visits would be the most valuable visit type. The results gained from the analysis contradict with the hypothesis. The results suggest that search engine traffic is the most valuable source for visits, direct traffic being the least valuable traffic source. Table 19 presents visitor behavior characteristics grouped by traffic sources. Small differences exist between the groups. Referral traffic has the lowest bounce rate and the longest average visit duration.

Table 19. Visitor behavior grouped by traffic sources (Google Analytics)

Default Channel Grouping	Behavior		
	Bounce Rate	Pages / Visit	Avg. Visit Duration
	53.87% Site Avg: 53.87% (0.00%)	2.99 Site Avg: 2.99 (0.00%)	00:02:49 Site Avg: 00:02:49 (0.00%)
1. Organic Search	54.33%	2.98	00:02:51
2. Direct	54.99%	2.98	00:02:32
3. Referral	49.33%	3.03	00:02:55

The amount of referral traffic to the site is very low. Referral traffic has potential to be an important source of visits but currently the website owners have not been able to establish effective links to the site. Examining the referral sources does not reveal any trading websites or partner sites that would be generating steady referral traffic to the site. Realizing the unused potential of referral traffic could have a major impact on the total number of conversions. Establishing links on the correct partner sites could also increase the conversion rate considerably.

5.2.5 Search engine optimization metrics

The search engine optimization (SEO) web metrics include impressions, clicks, and the average position on Google search engine. Impressions and clicks are closely correlated with the amounts of new and return visits. Figure 22 shows the amount of daily clicks from Google search engine starting from the 5th of July. As expected, all the SEO metrics are strongly correlated with search engine traffic. The SEO metrics are also closely correlated with each other.

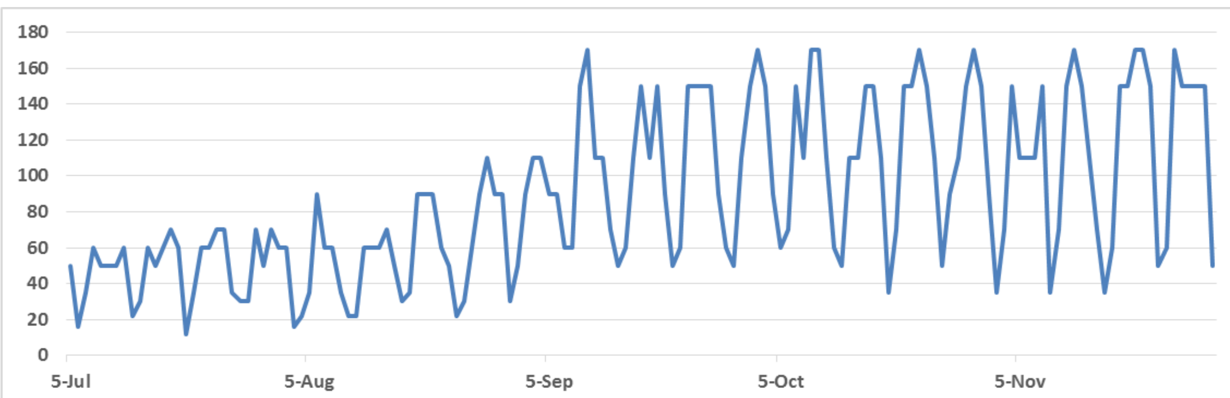

Figure 22. Daily clicks from Google search engine, July 5th to November 30th

Impressions, clicks, and the average position are strongly correlated with the total number of conversions (Table 20). When weekends are included, only impressions are significantly correlated with the conversion rate. When weekends are excluded, all the SEO metrics are significantly correlated with the conversion rate. As expected, the correlation with average position is negative.

Table 20. Pearson correlations: SEO metrics and conversions

PEARSON CORRELATIONS	SATURDAYS AND SUNDAYS INCLUDED		SATURDAYS AND SUNDAYS EXCLUDED	
	TOTAL CONVERSIONS	CONVERSION RATE	TOTAL CONVERSIONS	CONVERSION RATE
IMPRESSIONS	,879**	,212**	,839**	,474**
CLICKS	,879**	,146	,822**	,463**
AVERAGE POSITION	-,292**	-,069	-,503**	-,285**

***Correlation is significant at the 0.05 level (2-tailed). **Correlation is significant at the 0.01 level (2-tailed).**

Table 21 shows the regression analysis results when impressions and the average position are analyzed together towards the conversion rate. The results suggest that the amount of impressions can be used to predict the conversion rate. The average position does not seem to be a significant predictor. Table 22 shows the results when clicks and the average position are analyzed together towards the conversion rate. Based on the results, clicks can be used as a predicator. The results for the average position are similar with the previous analysis: the average position does not predict the conversion rate. Impressions and clicks are not analyzed at the same time because of the high intercorrelation between the metrics.

Table 21. Regression analysis: impressions, average position, & conversion rate (excl. weekends)

DEPENDENT VARIABLE: CONVERSION RATE			
INDEPEND. VARIABLE	**B**	**SE B**	**β**
CONSTANT	,553	,109	
IMPRESSIONS	,00005	,00001	,414**
AVERAGE POSITION	-,002	,001	-,142*

Note: R^2 = ,241 *p > ,1 **p < ,001

Table 22. Regression analysis: clicks, average position, & conversion rate (excl. weekends)

DEPENDENT VARIABLE: CONVERSION RATE			
INDEPEND. VARIABLE	**B**	**SE B**	**β**
CONSTANT	,521	,120	
CLICKS	,001	,0004	,407**
AVERAGE POSITION	-,001	,001	-,108*

Note: R^2 = ,223 *p > ,1 **p < ,001

The results suggest that following the SEO metrics is valuable. As discussed in the previous chapter, search engines (consisting almost solely of Google search engine) were the most common traffic source for the site. It is safe to say that the improved search engine visibility has been one of the main reasons for the increased number of total conversions. The SEO efforts of the website owners have been effective during the first half of the data collection period. But since August, the amount of impressions or clicks have not been increasing any longer. It seems that the previous SEO efforts do not work any longer and there is a need for new kinds of methods.

Search engine optimization can be said to contain two elements: optimizing the site for search engines' algorithms and optimizing the site for human visitors. The increasing amount of impressions and better site ranking means that the site is better optimized for search engines. The increasing amount of clicks indicates that the page is better optimized for human visitors. Figure 23 shows the daily average click through rate (CTR) for the site. CTR is calculated by dividing clicks with impressions. The ratio indicates how effectively the page previews on search engines are attracting visitors. CTR can also be used to compare how well different pages are optimized. Table 23 presents the impressions, clicks, and CTR for the ten most popular landing pages during November. The previews of pages with low CTR needs to be improved. It can also be useful to examine pages with the highest click through rates and check the amount of impressions for these pages. If a page has high CTR and low number of impressions, it means that the page has attractive content but the page is poorly optimized for the algorithms of search engines.

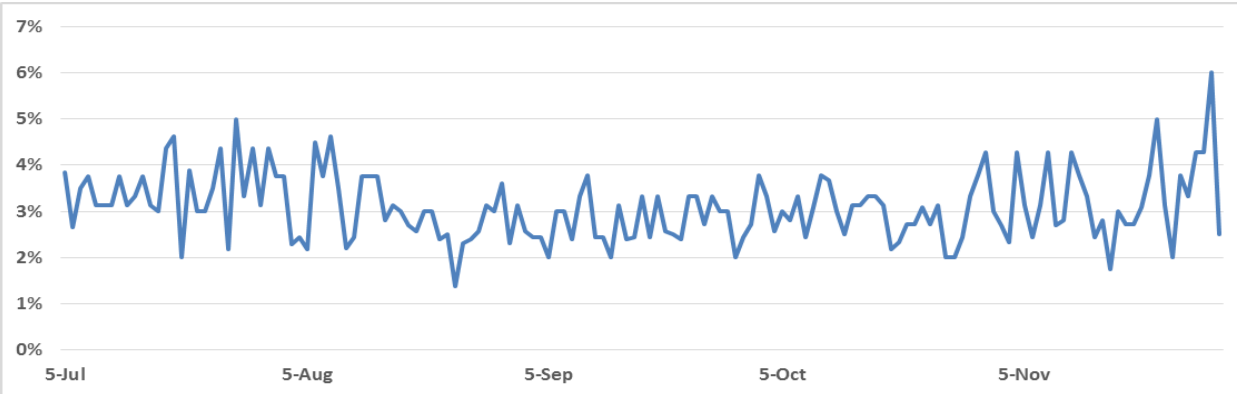

Figure 23. Daily click through rate, July 5th to November 30th

Table 23. Ten most popular landing pages during November (Google Analytics)

Landing Page	Impressions	Clicks	Average Position	CTR
1.	3,000	170	25	5.67%
2.	3,000	150	41	5.00%
3.	2,500	70	42	2.80%
4.	2,500	110	24	4.40%
5.	2,000	110	19	5.50%
6.	2,000	35	61	1.75%
7.	2,000	30	48	1.50%
8.	1,600	70	42	4.38%
9.	1,600	16	67	1.00%
10.	1,600	70	27	4.38%

Increased search engine visibility does not automatically lead to a higher conversion rate. The relationship between impressions and the conversion rate also indicates if the site is optimized for the correct key words. A website should not just try to maximize the amount of visitors but maximize the amount of *desired* visitors who are interested in the content of the site and more likely to convert. A close relationship between the SEO metrics and the conversion rate can indicate that SEO efforts for a website are implemented for the correct key words.

5.3 The key web metrics

The case study aims to identify a set of key web metrics that are the most valuable for the site developers to follow. Table 24 divides the fourteen web metrics followed in the study into four different groups based on their relationship with the total number of conversion and the conversion rate. The relationship with the total number of conversion indicates how closely a metric is connected with the overall performance of the site. The relationship with the conversion rate indicates if a metric is associated with the quality of a single visit.

Table 24. The relationships between web metrics and conversions

Group	Web Metric	Total Number of Conversions	Conversion Rate
1	Search Engine Traffic	Strong connection	Strong connection
1	Impressions on Google	Strong connection	Strong connection
1	Clicks on Google	Strong connection	Strong connection
2	New Visits	Strong connection	Minor connection
2	Total Page Views	Strong connection	Minor connection
2	Unique Page Views	Strong connection	Minor connection
3	Return Visits	Strong connection	No connection
3	Referral Traffic	Strong connection	No connection
3	Direct Traffic	Minor connection	Minor connection
3	Average Position on Google	Minor connection	Minor connection
3	Rate of Return Visits	No connection	Minor connection
4	Average Page Views	No connection	No connection
4	Average Time on Page	No connection	No connection
4	Bounce Rate	No connection	No connection

The web metrics in Group 1 (search engine traffic, impressions on Google, and clicks on Google) have a strong connection with both the total number of conversions and the conversion rate. Group 2 includes web metrics (new visits, total page views, and unique page views) that have a strong connection with the total number of conversions and a minor connection with the conversion rate. Group 3 includes metrics (return visits, referral traffic, direct traffic, the average position on Google, and the rate of return visits) that have either a strong connection with the total number of conversions and no connection with the conversion rate or a minor connection with both the total amount of conversions and the conversion rate. The web metrics in Group 4 (average page views, average time on page, and bounce rate), do not have any noticeable connection with the conversions.

Based on the level of the relationship towards the conversions, the web metrics in Group 1 and Group 2 seem to be the most valuable to follow. However, all of the web metrics in these groups are strongly correlated with each other. There is a risk that the metrics do not contain enough unique information from each other for a detailed daily monitoring and analysis of the values of all these metrics to be constructive. Intercorrelated metrics are likely to reveal information only about some characteristics of the visitors and do not give a comprehensive representation of visitor behavior. It is not desired to identify key metrics that have a close connection with each other.

Based on the relationship analysis and characteristics of the website visitors, two key web metrics for the site are proposed: *search engine traffic* and *the rate of return visits.*

Based on the results of the relationship analysis, search engine traffic is selected as the first key metric. Search engines have been the dominant traffic source for the site. Any large changes with this metric will surely have a vast effect on the total number of conversions and smaller but still significant effect on the conversion rate. Just by looking at this metric, the site developers can check if everything is as usual with the site. Following the trends of this metric tells in what direction the usage of the site is going. Search engine traffic is also closely connected with search engine optimization and a good indicator how well the SEO efforts are working, even though the actual SEO metrics reveal more details about the different aspects of the optimization. Because of the intercorrelation between the metrics with the closest connections to conversions, only one key metric is selected based on the relationship analysis.

The second key metric, the rate of return visits, is not strongly correlated with conversions. The selection of this key metric is not based on the relationship analysis but the metric is chosen based on the examination of visitor characteristics. The rate of return visits is representing a major weakness with the website. The site needs to start attracting loyal visitors who frequently visit the site. The objective of the site is to become a valuable marketing channel that shares information about the new product offerings with both new and old customers. This objective is not achieved if only a few visitors will visit the site multiple times. The site owners need to pay close attention to this metric. Changes to the site are necessary and the site developers need to follow how the changes will affect the rate of return visits.

6. CONCLUSION

6.1 Summary of the thesis

The main research question for the thesis was to examine which web metrics are most closely connected with website conversions. The supporting research question was to analyze what information different web metrics reveal about the characteristics of website visitors. The research questions formed the basis for the new research framework presented in this thesis. The research framework for the role of web metrics in website development is used in the case study conducted in this thesis. The objective of the case study was to test if the new model can be utilized to identify meaningful information about website visitors that can be used for website development and marketing purposes. The case study aims to identify a set of key web metrics that contain the most valuable information about the site visitors and about the website itself.

The case study of this thesis examined fourteen different web metrics collected from one corporate website during a period of six months. The study evaluated the relationships of the web metrics towards website conversions using correlation and regression analyses. The relationships were evaluated towards both the total number of conversions and the conversion rate. The study also examined what kind of characteristics the web metrics reveal about the site visitors and used the web metrics to evaluate the effectiveness of different aspects of the website.

The case study proposed two key web metrics for the site under examination, search engine traffic and the rate of return visits. Search engine traffic was selected based on the relationship analysis of the study. The study identified many strongly intercorrelated web metrics with strong relationships to conversions. Search engine traffic was selected among these metrics for its capability to cover many aspects about the site, like the amount of site traffic and search engine optimization. The rate of return visits was selected based on the examination of user characteristics. This metric represent one of the major weaknesses of the site, the low amount of return visits. The site developers need to start following this metric closely.

In addition to poor visitor loyalty, the web metric analysis revealed many other important visitor characteristics. The behavior of the users who visited the site during weekdays was different from the users who visited the site during weekends (Saturdays and Sundays). It is likely that different kind of visitors visit the site between weekdays and weekends. The site stickiness and the depth of the visits were low. Most of the visitors viewed only one or two pages and spent less than 60 seconds on the site. Google search engine was the dominant source of visits. The fact that one channel is generating almost all of the visits can be risky. Any changes with Google's search algorithm may have a large effect towards the site's traffic levels and the number of conversions. The use of referrals or links to attract visitors is severely underutilized. It is important for the site developers to diversify the site's inbound channels.

6.2 Practical implications

The successful identification of relevant key web metrics suggests that the framework presented in this thesis can be used by organizations to reveal useful information about website visitors. The web metrics analysis that includes evaluating the relationships between web metrics and conversions and discovering important characteristics from detailed metric examination seems to be a useful method for companies trying to improve their websites. It is not necessary to select the same web metrics and conversion goals that are examined in this thesis to utilize the proposed framework. Organizations can choose any metrics and conversions and still use the same methods for web metrics analysis that are used in this thesis.

Google Analytics proved to be an efficient tool for calculating and following web metrics. The time it took to configure Google Analytics to run with the website was minimal and exporting the data from the tool into SPSS Statistics went without problems. The features of Google Analytics made it possible to examine the web metrics in more detail and reveal valuable visitor characteristics.

The web metrics used in the case study are all aggregated metrics that are calculated for the whole visitor base. It seems that the aggregated values of many metrics, like the average time on page, offer too general information about the visitors. Aggregated web metrics also seem to be often correlated with each other which indicates that the metrics are containing the same kind of information. In order to find useful insights with these metrics, the metrics need to be calculated for visitor segments. Segmented metrics have potential to offer a more diverse view of the web users. Visitor segmentation can be done by combining different web metrics. For example, site developers can calculate the amount of return visitors acquired through referrals who spend more than five minutes on the site and then calculate the conversion rate for this segment. The objective is to identify segments with meaningful differences and examine if the conversion rates differ between the segments. It can also be helpful to try to identify segments with the highest and lowest conversion rates. By examining the connections towards conversions, it is possible to make statements about the value of different customers. When segments are identified, the next step is to try to transform the behavior characteristics revealed by web metrics into real-life customer features.

6.3 Limitations and further research

A major limitation to this thesis is that the clickstream data for the case study is collected only from one website. For other sites, the most valuable web metrics are likely to be different. Every website needs to define their own key metrics. The method used in this thesis to find the key web metrics can be applied to any website.

Another limitation is that this thesis does not examine the relationship between the selected conversion goals and the overall business objectives of the company that runs the website. To confirm if relevant conversion goals are selected, the relationship between the amount of conversions could be compared with financial measures, like sales volumes or profits. But for websites that are not generating any direct sales, this comparison can be problematic. It can be difficult to isolate the effect of a website from other marketing channels and from other factors on the overall profitability of a company. Despite the challenges with choosing the correct conversions, it is still necessary to have conversion goals for a website. Conversions are the only objective way to measure website success and performance.

In this thesis, all the data of website usage was collected from clickstream data. Also the previous research in web analytics is concentrated on clickstream data and web metrics. It is important to remember that there are also other ways to collect data about websites and their visitors. It is possible to combine web metrics with customer surveys and experiments. One interesting research subject might to examine the clickstream data collected from a group of volunteer visitors and then ask their reasons and motives for their browsing behavior. Connecting web metrics with other data sources can give more information about the value of the web metrics.

REFERENCES

Arendt, J. & Wagner, C. (2010) "Beyond Description: Converting Web Site Usage Statistics into Concrete Site Improvement Ideas", *Journal of Web Librarianship,* Vol. 4, No. 1, pp. 37-54.

Bucklin, R. & Sismeiro, C. (2009) "Click Here for Internet Insight: Advances in Clickstream Data Analysis in Marketing", *Journal of Interactive Marketing*, No. 23, pp. 35-48.

Chen, H., Chiang, R. & Storey, V. (2012) "Business Intelligence and Analytics: From Big Data to Big Impact", *MIS Quarterly*, Vol. 36, No. 4, pp. 1165-1188.

Chiang, I., Huang, C. Y. & Huang, C. W. (2010) "Traffic metrics and Web 2.0-ness", *Online Information Review*, Vol. 43, No. 1, pp. 115-126.

Chiou, W., Lin, C. & Perng, C. (2010) "A strategic framework for website evaluation based on a review of the literature from 1995–2006", *Information & Management*, Vol. 47, No. 5-6, pp. 282-290.

Constantinides, E. (2002) "The 4S Web-Marketing Mix model", *Electronic Commerce Research and Applications*, Vol. 1, No. 1, pp. 57-76.

Cooper, A. (2012) "A Brief History of Analytics", *JISC CETIS Analytics Series*, Vol. 1, No. 9.

Davis, F. (1989) "Perceived Usefulness, Perceived Ease of Use, and User Acceptance of Information Technology," *MIS Quarterly*, Vol. 13, No. 3, pp. 319-339.

DeLone, W. & McLean E. (1992), "Information systems success: The quest for the dependent variable", *Information Systems Research*, Vol. 3, No. 1, pp. 60-95.

Evans, M. (2007) "Analyzing Google rankings through search engine optimization data", *Internet Research*, Vol. 17, No. 1, pp. 21-37.

Farina, F. (2013) Online. Available at: http://www.e-nor.com/blog/google-analytics/google-analytics-solidifies-lead-in-fortune-500-adoption-in-2013. [21.10.2013].

Fasel, D. & Zumstein D. (2009) "A Fuzzy Data Warehouse Approach for Web Analytics", *Visioning and Engineering the Knowledge Society. A Web Science Perspective*, Vol. 5736, pp. 276-285.

Field, A. (2009) Discovering Statistics using SPSS, SAGE Publications Ltd., Thousand Oaks, California.

Fleishman-Hillard (2012), Online. Available at: http://fleishmanhillard.com/2012/01/news-and-opinions/2012-digital-influence-index-shows-internet-as-leading-influence-in-consumer-purchasing-choices/, [31.1.2012].

Ghandour, A., Benwell G. & Deans, K. (2010) "The Relationship Between Website Metrics and the Financial Performance of Online Businesses", *ICIS 2010 Proceedings*, Paper 27. Available at: http://aisel.aisnet.org/icis2010_submissions/27

Hamel, S. (2012), Online. Available at: www.online-behaviour.com/googleanalytics/myths.

Hamel, S. (2012b), Online. Available at: http://blog.webanalyticssolutionprofiler.com/blog/who-runs-web-analytics-in-the-top-500-retail-web-site-december-2011/. [15.2.2012].

Jansen, B. & Spink, A. (2006) "How are we searching the World Wide Web? A comparison of nine search engine transaction logs", *Information Processing & Management,* Vol. 42, No. 1, pp. 248-263.

Kaushik, A. (2007) Web Analytics – An Hour a Day, Wiley Publishing, Inc., Indiapolis, Indiana.

Kaushik, A. (2010) Web Analytics 2.0 – The Art of Online Accountability & Science of Customer Centricity, Wiley Publishing, Inc., Indiapolis, Indiana.

Kirk, R. (2008) Statistics – An Introduction, Thomson Wadsworth, Belmont, California.

Merwe, R. & Bekker, J. (2003) "A framework and methodology for evaluating e-commerce Web sites", *Internet Research: Electronic Networking Applications and Policy*, Vol. 13, No. 5, pp. 330-341.

Morgan, A., Leech, N., Gloeckner, G. & Barrett, K. (2004) "SPSS for Introductory Statistics – Use and Interpretation", Lawrence Erlbaum Associates, Inc., Mahwah, New Jersey.

Nakatani, K. & Chuang, T. (2011) "A web analytics tool selection method: an analytical hierarchy process approach", *Internet Research,* Vol. 21, No. 2, pp. 171-186.

Pakkala, H., Presser, K. & Christensen, T. (2012) “Using Google Analytics to measure visitor statistics: The case of food composition websites”, *International Journal of Information Management,* Vol. 32, No. 6, pp. 504–512.

Palmer, J., (2002) “Web Site Usability, Design, and Performance Metrics”, *Information Systems Research,* Vol. 13, No. 2, pp. 151-167.

Park, J. & Chung, H. (2009) “Consumers’ travel website transferring behaviour: analysis usingclickstream data-time, frequency, and spending”, *The Service Industries Journal,* Vol. 29, No. 10, pp. 1451-1463.

Peacock, D. (2002) “Statistics, Structures & Satisfied Customers: Using web log data to improve site Performance”, *Museums and the Web,* Available at: http://www.museumsandtheweb.com/mw2002/papers/peacock/peacock.html.

Phippen, A., Sheppard, L. & Furnell S. (2004) “A Practical Evaluation of Web Analytics”, *Internet Research,* Vol. 14, No. 4, pp. 284-293.

Plaza, B. (2009), “Monitoring web traffic source effectiveness with Google Analytics” *Aslib Proceedings: New Information Perspectives,* Vol. 61, No. 5, pp. 474-482.

Potts, K. (2007), Web Design and Marketing Solutions for Business Websites, Springer-Verlag Inc., New York.

Schaupp, L., Belanger, F. & Fan, W. (2009) “Examining the Success of Websites Beyond E-Commerce: An Extension to the IS Success Model” *Journal of Computer Information Systems,* Vol. 49, No. 4, pp. 42-52.

Singh, S., Makkar, A. & Singh, N. (2011) “A metrics based approach to analyze web usage pattern”, *International Journal of Computers & Technology*, Vol. 1, No. 1, pp. 1-6.

Soonsawad, P. (2013) “Developing a New Model for Conversion Rate Optimization: A Case Study”, *International Journal of Business and Management,* Vol. 8, No. 10, pp. 41-51.

Tan, F., Tung, L. & Xu, Y. (2009) ”A Study of Web-Designers’ Criteria for Effective Business-to-Consumer (B2C) Websites Using the Repertory Grid Technique”, *Journal of Electronic Commerce Research,* Vol. 10, No. 3, pp. 155-177.

Tonkin, S., Whitmore, C. & Cutroni J. (2010) Performance Marketing with Google Analytics, Wiley Publishing, Inc., New Jersey.

Wang, X., Shen, D., Chen, H. & Wedman, L. (2011) “Applying web analytics in a K-12 resource inventory”, *The Electronic Library,* Vol. 29, No. 1, pp. 20-35.

Web Analytics Association (2008), Online. Available at: http://www.digitalanalyticsassociation.org/Files/PDF_standards/WebAnalyticsDefinitions.pdf.

Weitz, R. & Rosenthal, D. (2010) “Valuing a B2B Website: A Case Study of an Industrial Products Supplier” *Journal of Business Case Studies,* Vol. 6, No. 5, pp. 59-64.

Welling R. & White L. (2006) “Web site performance measurement: promise and reality”, *Managing Service Quality,* Vol. 16, No. 6, pp. 654-670.

Wilson R. D. (2010) “Using clickstream data to enhance business-to-business website performance”, *Journal of Business & Industrial Marketing*, Vol. 25, No. 3, pp. 177-187.

Zheng, N., Chyi, H. & Kaufhold, K. (2012) “Capturing “Human Bandwidth”: A Multidimensional Model for Measuring Attention on Web Sites”, *The International Journal on Media Management*, Vol. 14, No. 2, pp. 157-179.

APPENDICES

Appendix 1. Sample of data (From SPSS Statistics)

	NEW_VISITS	RET_VISITS	RETURN_RATE	AVE_PAGES	TOTAL_PAGES	UNI_PAGES	AVE_TIME	BOUNCE	TRA_SENGINE	TRA_REFER	TRA_DIRECT	IMPRESSIONS	CLICKS	AVE_POSIT	CONVERSIONS	CONVERS_RATE
120	21	10	,323	2,94	91	66	106,97	,4516	27	2	2	2000	60	99,12	27	,8710
121	18	7	,280	2,72	68	47	63,92	,5200	24	1	0	2500	50	70,79	20	,8000
122	58	17	,227	4,36	327	191	316,75	,4400	65	5	5	4500	110	76,93	57	,7600
123	69	15	,179	3,48	292	198	209,50	,5476	69	6	9	5500	150	81,05	57	,6786
124	72	25	,258	2,82	274	196	162,34	,4742	75	15	7	4500	170	73,30	44	,4536
125	74	16	,178	2,12	191	142	131,60	,5000	74	6	10	4500	150	78,17	45	,5000
126	53	16	,232	2,78	192	133	130,09	,5652	51	9	9	3500	90	84,43	49	,7101
127	30	9	,231	2,74	107	82	291,18	,5897	29	8	2	2000	60	97,04	19	,4872
128	43	11	,204	4,80	259	162	460,81	,4815	25	21	8	2500	70	98,28	31	,5741
129	73	23	,240	2,59	249	194	170,80	,5729	62	18	16	4500	150	96,79	54	,5625
130	69	16	,188	2,59	220	169	178,65	,5294	70	5	10	4500	110	93,03	47	,5529
131	80	24	,231	2,52	262	193	173,71	,5577	84	8	12	5500	170	99,75	72	,6923
132	66	19	,224	3,68	313	201	212,40	,4706	67	12	6	4500	170	70,92	56	,6588
133	72	16	,182	2,44	215	170	113,25	,6364	71	9	8	3000	110	81,90	57	,6477
134	34	6	,150	2,58	103	76	169,30	,5250	30	3	7	2000	60	78,69	30	,7500
135	18	1	,053	6,11	116	62	319,16	,4211	15	0	4	2000	50	85,53	17	,8947
136	38	12	,240	2,82	141	100	200,02	,4200	41	6	3	3500	110	99,03	46	,9200
137	61	18	,228	2,65	209	154	116,57	,5443	63	7	9	3500	110	76,17	51	,6456
138	45	20	,308	2,12	138	111	122,05	,5385	51	9	5	4500	150	82,37	44	,6769
139	63	15	,192	3,42	267	195	270,60	,5128	70	4	4	4500	150	75,54	56	,7179
140	70	10	,125	3,35	268	200	225,75	,4375	66	7	7	3500	110	87,18	48	,6000
141	43	13	,232	4,63	259	146	253,91	,6071	40	8	8	1600	35	60,60	22	,3929
142	24	9	,273	5,36	177	131	235,79	,4242	28	3	2	3000	70	89,01	30	,9091
143	55	21	,276	2,36	179	138	72,75	,5526	66	5	5	5500	150	90,47	56	,7368
144	70	23	,247	3,61	336	215	223,67	,5269	75	7	11	5500	150	76,06	72	,7742
145	59	34	,366	3,32	309	218	165,88	,5269	82	7	4	5500	170	76,83	66	,7097
146	64	22	,256	2,56	220	156	181,02	,5116	68	7	11	5500	150	92,84	63	,7326
147	79	17	,177	2,96	284	194	190,88	,4688	83	6	7	3500	110	70,86	59	,6146
148	27	4	,129	5,03	156	113	189,94	,3226	26	2	3	2500	50	75,15	24	,7742
149	36	7	,163	2,93	126	92	123,84	,5349	32	5	6	4500	90	77,42	34	,7907
150	69	22	,242	3,03	276	201	227,45	,5714	68	9	14	4500	110	62,64	66	,7253

Data View Variable View

Appendix 2. Descriptive statistics of the web metrics and conversions

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean		Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic
NEW_VISITS	183	83	3	86	40,09	1,606	21,732
RET_VISITS	183	34	0	34	11,92	,578	7,815
RETURN_RATE	183	,667	,000	,667	,22412	,006254	,084608
AVE_PAGES	183	7,58	1,42	9,00	3,0117	,07020	,94967
TOTAL_PAGES	183	549	16	565	155,51	6,918	93,585
UNI_PAGES	183	355	13	368	110,56	4,732	64,015
AVE_TIME	183	1027,06	6,08	1033,14	169,9861	7,59910	102,79874
BOUNCE	183	,5694	,2222	,7917	,537302	,0066006	,0892918
TRA_SENGINE	183	89	4	93	40,47	1,703	23,036
TRA_REFER	183	21	0	21	4,85	,289	3,905
TRA_DIRECT	183	32	0	32	6,69	,371	5,015
IMPRESSIONS	149	4900	600	5500	2880,54	116,172	1418,063
CLICKS	149	158	12	170	88,35	3,791	46,273
AVE_POSIT	170	193,30	55,59	248,89	86,3566	1,50340	19,60190
CONVERSIONS	183	78	2	80	30,32	1,442	19,505
CONVERS_RATE	183	,8963	,0714	,9677	,577076	,0140428	,1899680
Valid N (listwise)	149						