ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ
Διοίκηση Επιχειρήσεων με κατεύθυνση Ηλεκτρονικό Επιχειρείν
MBA in e-business

Search Ranking Factors and Rank Correlations for Greek Ferry Routes Websites

Google | Greece | 2016

Άγγελος Μορφοβασίλης | Άννα Ψύλλα
Επιβλέπων καθηγητής:**Ψαρομήλιγκος Ιωάννης**

Contents

Abstract

The new technologies and the internet has been part of our everyday life and the number of internet users has been increased rapidly in the last decade. It is widely recognized that tourism industry has been extremely influenced by technological developments. Greece is a well-known tourist destination for several decades and the Greek economy heavily relies on the tourism industry. Especially in the summer period it is estimated that more than 23 million of people every year visit Greece and Greek islands. The combination of these two facts has bring many evolutionary changes in the way tourist offices operate. The electronic ticket has become an integral part of the Greek ferry industry and the majority of the travelers choose to book their ferry tickets online. As a result, the Google search results are of vital importance for tourist websites and the final search ranking can lead a website to success or failure. The aim of this thesis is to examine all the different factors that have a small or a large impact on the final Google search rankings. The survey which was made was focusing to Greek ferry route websites. From the survey's results, some useful conclusions were drawn about the factors, in which the websites need to pay more attention, if they want to improve their rankings in the Google search machine.

List of Tables

List of Figures

List of Abbreviations

WWW	World Wide Web
SEM	Search Engine Marketing
SERPs	Search Engine Results Pages
SEO	Search Engine Optimization
EB	Exabyte
ZB	Zettabyte
FTP	File Transfer Protocol
URL	Uniform Resource Locator
HTTPS	Hypertext Transport Protocol Security
SSL	Secure Sockets Layer
IP	Internet Protocol
CDN	Content Delivery Network
EMD	Exact Match Domain
CTR	Click Through Rate
CF	Citation Flow
TF	Trust Flow
UX	User Experience

Chapter 1: Introduction and outline

1.1 Introduction

The role of the Internet and the World Wide Web in our everyday life has revolutionized the way we seek and gain access to information. One way to find the desired information is through the search engines. Each website, depending on the content and the way it is structured, holds a specific position in the engine results list. Search Engine Marketing (SEM) is a form of web marketing and was created by the ever-increasing competition and the need of the websites to hold a higher position on that list. SEM aims to promote websites and increase their visibility on search engine results pages (SERPs) through Search Engine Optimization (SEO).

Search Engine Optimization is a structured method that improves the ranking of a web page in search engines' SERPs for specific keywords or key phrases that interferes with the structure and content of a web page so as to be as friendly as possible to search engines. Websites also are using tools for optimizing search engine results, as the correct and regular use of these tools advise about the errors that need to be corrected on a website.

The use of the Internet for tourist purposes by consumers is now taken for granted, whether it is information search or purchase of services and products. As a result, the corresponding pages offering tourist products could not deviate from this trend. In the need of survival in difficult economic conditions and aiming to increase their profits, tourist businesses are trying to find solutions to harmonize with this new trend. Businesses that will be successful in the internet area and attract more customers, will be able to achieve their goals and that is why the corresponding websites are looking for ways to emerge from their competitors.

1.2 Research Objectives

The aim of this thesis is to highlight the importance of internet and Google search rankings in the tourist industry. The research topic is based on the fact that successful shipping companies of the future will be those, that are able to meet the needs of customers and satisfy their desires in the best possible way. In order to

achieve that, customers first have to find and visit the corresponding website that will provide them with these services. Greek ferry root websites from their side, should find the appropriate way to improve their position on google search results in order to increase their traffic and attract more customers. Based on the «Search Ranking Factors and Rank Correlations» by Google U.S. 2015, the research which was made has thoroughly analyzed all the 43 factors that it is supposed to affect the search results of Google's search engine. Finally, the research presents the factors that appear to have a real impact on the Google search results and to which factors the websites should emphasize in order to improve their rankings.

1.3 Outline of the thesis

The thesis is structured and presented as follows:

Chapter 2: Internet and the evolution of search engines

This chapter focuses on the evolution of the internet and the search engines. It presents some interesting statistics about the internet users worldwide and shows the recorded increase of them in Greece the last decade. This chapter also presents the importance of search engines and shows the history of search engines from 1990 until today.

Chapter 3: Greek Tourism Industry

In the beginning of this chapter there are presented some historical information about Greek tourism. In the following sub chapters, it is explained how internet affects tourism industry and how Greek maritime sector applies the electronic ticket. Special reference is made to the company that first developed the integrated distribution system for electronic reservations on the coastal shipping and to one of the most successful internet tourist offices in Greece. Lastly, it presents the benefits of a good web site and the advantages and disadvantages of the electronic ticket.

Chapter 4: Search Ranking Factors

This chapter presents and analyze each one of the 43 factors that are going to be tested in the 5th chapter. These factors based on their content are divided into the following categories: Technical, user experience, content, social networks, backlinks, mobile and alternative sources.

Chapter 5: Greek ferry routes websites case study

This chapter, presents the results of the survey made for the Greek ferry routes websites. In the beginning, it describes the methodology and the specific parameters followed in order the survey results to be as reliable as possible. In the rest of the chapter, it is presented the results of the survey with individual graph for each factor and description about the correlation value. Lastly, the chapter closes with the conclusions made regarding the survey's results.

Chapter 2: Internet and the evolution of search engines

2.1 Introduction

Every day millions of people around the world use the Internet to search and retrieve information on a wide range of topics. This information is displayed in various formats, such as text, picture, sound or video. All this activity is possible, because thousands of individual networks are connected to the internet and share information. Never before, so much information from a wide variety of sources and in so many forms were available to the public. At present, there are 3.731.973.423 internet users worldwide. Asia is in the first place with the most internet users worldwide, holding a percentage of 50,2% and is followed by Europe with 17,1%. The following tables summarize the number of internet users worldwide. Also, internet growth statistics are represented for the years 2000 to 2017.

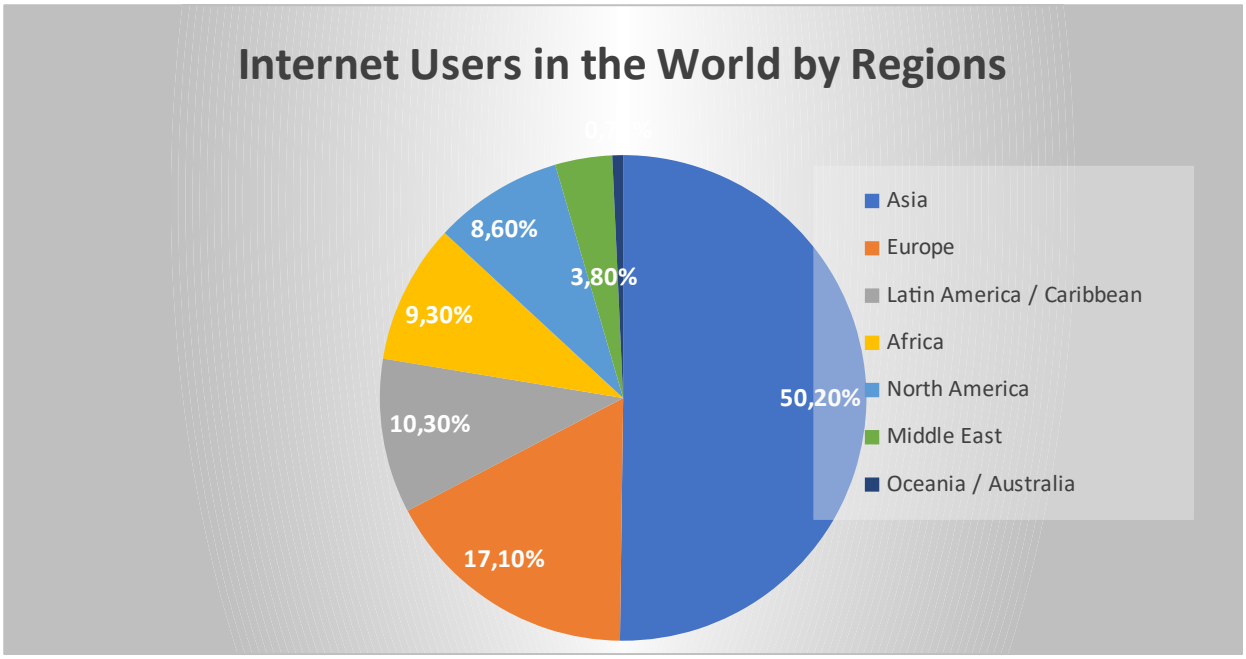


Table 2.1.1: Internet users in the world by region (source: internet world stats)

World internet usage and population statistics

World Regions	Population (2017 Est.)	Population % of World	Internet Users	Penetration Rate (% Pop.)	Growth 2000 - 2017	Users % Table
Africa	1,246,504,865	16.6 %	345,676,501	27.7 %	7,557.2%	9.3 %
Asia	4,148,177,672	55.2 %	1,873,856,654	45.2 %	1,539.4%	50.2 %
Europe	822,710,362	10.9 %	636,971,824	77.4 %	506.1%	17.1 %
Latin America / Caribbean	647,604,645	8.6 %	385,919,382	59.6 %	2,035.8%	10.3 %
Middle East	250,327,574	3.3 %	141,931,765	56.7 %	4,220.9%	3.8 %
North America	363,224,006	4.8 %	320,068,243	88.1 %	196.1%	8.6 %
Oceania / Australia	40,479,846	0.5 %	27,549,054	68.1 %	261.5%	0.7 %
WORLD TOTAL	7,519,028,970	100.0 %	3,731,973,423	49.6 %	933.8%	100.0 %

Table 2.1.2: World internet usage and population statistics (source: internet world stats)

In Greece, according to Hellenic Statistical Authority, for the year 2016 was recorder an increase of 1.5% in the households having internet access compared with 2015. More specifically, 7 out of 10 households have internet access at home (69.1%) and

during the last 5 years (2011 – 2016) the percentage of the increase reached 37.6% respectively [CITATION Hel16 \p 1 \l 1032]. Longitudinally, the evolution of households having internet access are depicted in the figure 2.1.1. Using Internet services at least once a week is considered as "Regular use" and it is recorded for 95.6% of the persons having used the internet in the first quarter of 2016, presenting an increase of 1.4% compared with 2015 (94.3%)[CITATION Hel16 \p 3 \l 1033].

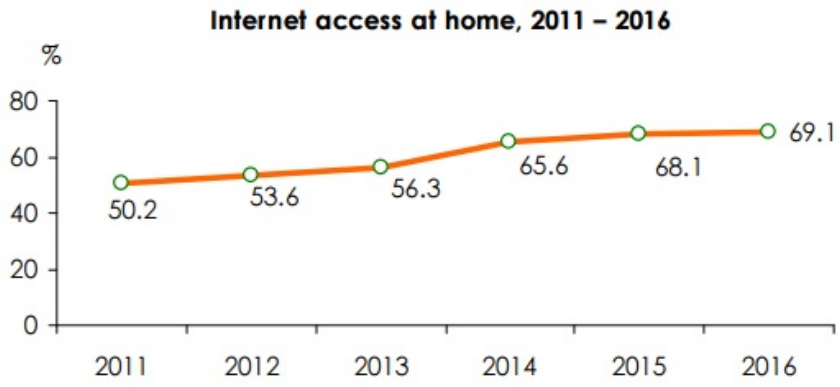


Figure 2.1.1: Internet access at home, 2011 – 2016 (source: Hellenic Statistical Authority press release 2016)

Correspondingly increased rates are also recorded for users with mobile internet access and ubiquitous connectivity. For the first quarter of 2016, 68.1% of the persons having used the internet, were connected using a smart phone, a portable PC (laptop, notebook, netbook or tablet) or other mobile device (PDA, e-book reader, MP3 player, etc.), thus recording an increase of 2.6% compared with the same period of 2015. The share of population accessing the internet on the go, as a percentage of population having accessed the internet, since 2010, is depicted in the figure 2.1.2.

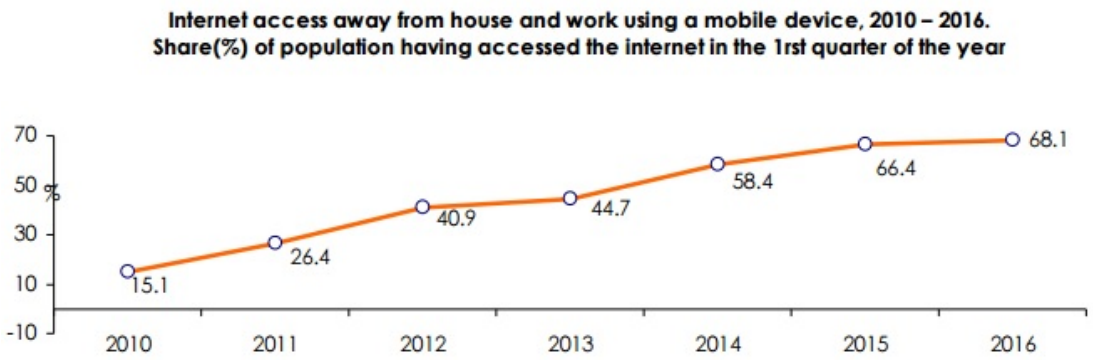


Figure 2.1.2: Internet access on the go using a mobile device 2010 - 2016(source: Hellenic Statistical Authority press release 2016)

2.2. Internet Vs World Wide Web

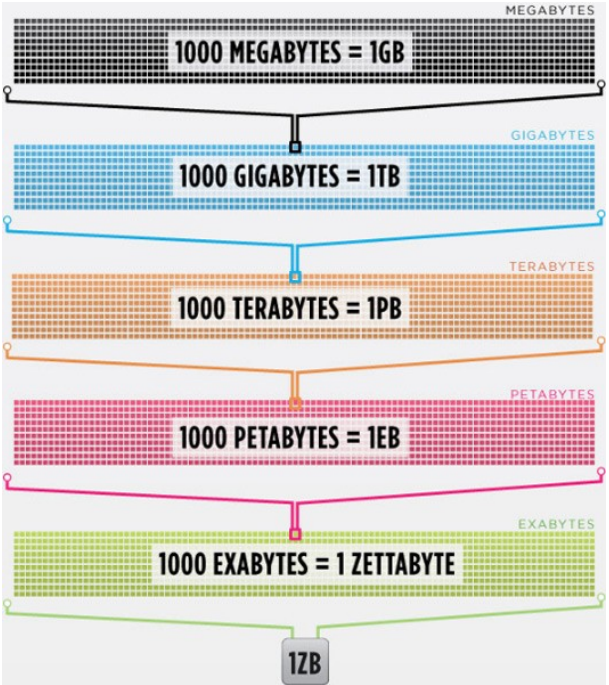
Many people use the terms Internet and World wide web (WWW) interchangeably, but in fact these two terms are not synonymous and this usage is technically incorrect.

The internet is a massive network of networks, a networking infrastructure. It connects millions of computers together globally, forming a network in which any computer can communicate with any other computer as long as they are both connected to the internet [CITATION Bea16 \l 1033]. The WWW is a way of accessing information over the medium of the internet via HTTP. It is just one of the connection protocols that are available in the internet. The access to the WWW is made through browsers and everyone can visit various web sites and view its online content[CITATION Too \l 1033].

2.3 The amount of information

From all the thing mentioned above, Internet and World Wide Web (www) have become an integral part of our everyday life and have brought us revolution in the way we seek and gain access to information. So how much information does the internet hold?

According to a survey published on July 2015, there were at least 4.66 billion web pages online. However, this calculation covers only the searchable web and not the Deep web [CITATION van16 \l 1033]. In 2014, researchers published a study in the journal Supercomputing Frontiers and Innovations estimating the storage capacity of the internet at 10^{24} bytes, or one million exabytes (EB).



One way to estimate the communication capacity of the internet is to measure the traffic

moving through it. According to Cisco's Visual Networking Index initiative, the internet is now in the zettabyte (ZB) era. A zettabyte equals one sextillion bytes, or 1.000 exabytes. By the end of 2016, global internet traffic has reached 1,1 ZB per year and by 2021, global traffic is expected to hit 3,3 ZB per year. One zettabyte is the equivalent of 36.000 years of high definition video [CITATION Cis17 \l 1033].

Figure 2.3: Multiple byte counting units

While the amount of information on the internet continuously growing, it becomes even harder for the user to find a suitable website, which will contain all the useful information who is looking for. As a result, the most effective way to search for any kind of information on the internet, is by using search engines. With this method, retrieving information about persons, businesses or services is achieved within a few seconds and as a result the person who performs the search, is easily led to the appropriate and suitable website.

2.4 Search Engines

The vast amount of information available on the Internet can be dizzying. Fortunately, there are tools available that will sort through the mass of information. Search engines collect information from Web sites and then, more or less, just dump that information into a database. In fact, search engines are just massive databases in which information from Internet documents are stored. The information in these databases is collected using a computer program (called a "spider" or a "robot") that scans the internet and gathers information about individual documents. These special programs work automatically to find documents or they are asked by a creator of a Web site to visit the site to be included in a database.

When a search engine is used, the order in which the results are listed also varies between search engines. Many search engines list the results using relevance ranking. Factors such as:

- how often your search terms are on the Web page
- where they are located on the page
- how many other Web pages link to the page

influence how high on the list of hits a page is listed. Many search engines allow Web sites to pay to have their pages listed higher in the results.

There are hundreds of these search engines available on the Web, but they all work in unique ways to collect and organize the information found. The information from Web sites might be gathered from all the words in a site, just the first few sentences in the body of a site, or only from the title or metatags (hidden descriptors of a site's content). Different search engines collect different information and this is the main reason why different results occur for the same search, from different search engines [CITATION Uni \l 1033].

2.5 The importance of Search Engines

Internet users consider search engines to be the central reference for exploring online content. Search engines transfer users to websites and information that they could not access otherwise. Search engines are so useful among the internet users because by typing key words or phrases, they appear a list of web pages that are as close as possible to the user searchable words.

The last few years, search engines have been used widely by the consumers worldwide, making them an integral part of consumption and advertising. A typical example is the case of the United States and the United Kingdom in 2006, where the total internet advertising cost exceeded the total radio advertising cost. Internet has continued to grow in share and significance when compared to other U.S. ad-supported media. In 2011, internet advertising surpassed cable television*, while in 2013, internet advertising exceeded broadcast television**[CITATION PwC16 \l 1033].

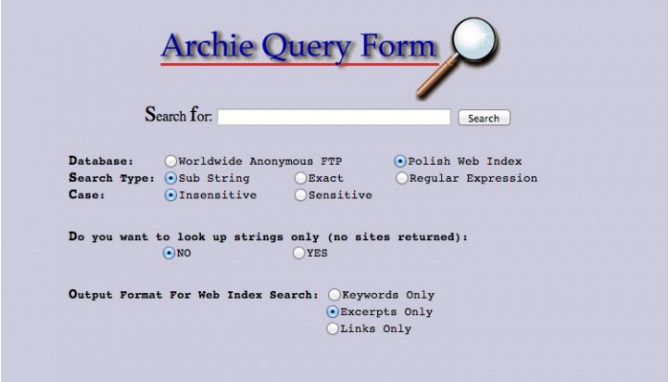
Internet advertising revenue growth outpaced other media outlets over the past five years. In every year since 2010, the annual growth rates of Internet advertising have exceeded those of other advertising media. Internet advertising has experienced double-digit annual growth in every year except 2009; no other media has experienced double-digit growth in any year [CITATION PwC16 \l 1033].

**Cable Television includes National Cable Networks and Local Cable television advertising revenue.*

***Broadcast Television includes Network and Syndicated and Spot television advertising revenue.*

2.6 History of Search Engines

In our days, the most well-known search engines hold the highest positions in the list of the most visible and valuable technology companies worldwide. However, it wasn't always this way. Today's search engines come from humble beginning and search has come a long way in the last few decades. But how search engines acquired their current form? What are the origins of this technology, which is not even 20 years old?



Archie (1990). The first tool for searching on the internet was Archie, the name stands for "archive". The application created by Alan Emtage at the early date of 1990, while he was a computer science student at McGill University in Montreal.

The program downloaded the directory listings of all the files located on public anonymous FTP (File Transfer Protocol) sites, creating a searchable database of file names; however, Archie did not index the contents of these sites since the amount of data was so limited it could be readily searched manually [CITATION ill02 \l 1033].

Gopher, Veronica & Jughead (1991). In 1991 Mark McCahill created Gopher at the University of Minnesota. Gopher was a menu system that simplified locating and using Internet resources. The user simply selected the desired item he wanted, from the menu. Gopher was a protocol system, which allowed server based text files to be hierarchically organized and easily viewed by end users who accessed the server using Gopher applications on remote computers. Based on Gopher's philosophy two new search programs Gopher came into the scene, Veronica and Jughead. Like Archie, they searched the file names and titles stored in Gopher index systems.

Veronica (Very Easy Rodent-Oriented Net-wide Index to Computerized Archives) provided a keyword search of most Gopher menu titles in the entire Gopher listings. On the other hand, Jughead (Jonzy's Universal Gopher Hierarchy Excavation and Display) was obtaining menu information from specific Gopher servers [CITATION Sey11 \l 1033].

World Wide Web Wanderer, Aliweb, Jump Station (1993). In June 1993, Matthew Gray (who now works for Google) produced what was probably the first web robot, the Perl-based World Wide Web Wanderer, and used it to generate an index called 'Wandex'. The purpose of the Wanderer was to measure the size of the World Wide Web. The Wanderer with its particular specifications, had the potential to become a general-purpose WWW search engine.

At the end of the same year (November 1993) another search engine called Aliweb is appearing. Aliweb provided a tool allowing users to just keep track of the services they provide, in such a way that automatic programs could simply pick up their descriptions, and combine them into a searchable database.

Jump Station search engine by using a web robot was able to find web pages and to build its index. As interface, used a web form for its query program. As a result, it was the first WWW resource-discovery tool, which was combining the three essential features of a web search engine (crawling, indexing, and searching). However, the meager resources available on the running platform, made its indexing and hence searching, limited to the titles and headings found in the web pages the crawler encountered [CITATION Sey11 \l 1033].

WebCrawler, Lycos (1994). It was the first Web search engine to provide full text search. Unlike with its predecessors, users were able to search for any word in any webpage, which has become the standard for the majority of search engines since. At first, WebCrawler was a desktop application and not a Web service as it is today. WebCrawler appeared with a database containing

pages from just over 4000 different Web sites. It was also the first one to be widely known by the public. Also in 1994, Lycos (which started at Carnegie Mellon University) was launched and became a major commercial endeavor [CITATION Sey11 \l 1033].

MetaCrawler, Alta Vista (1995). MetaCrawler was the first Meta - Search engine that provided search results by multiple search engines, rather than a single search engine algorithm. Daniel Dreilinger at Colorado State University developed Search Savvy, which let users searched up to 20 different search engines at once and a number of directories.

Alta Vista was one of the most popular to the public, but its popularity waned with the rise of Google. The two key participants who created the engine were Louis Monier, who wrote the crawler, and Michael Burrows, who wrote the indexer. AltaVista was backed by the most powerful computing server available. Became the fastest search engine and could handle millions of hits a day without any degradation. Introduced one key change with the inclusion of a natural language search. Users could type in a phrase or a question and get an intelligent response. For instance, “Where is London?” without getting a million-plus pages referring to “where” and “is” [CITATION Sey11 \l 1033].

Excite, Dogpile, HotBot, Ask Jeeves (1996). Netscape was looking to give a single exclusive search engine. Resultantly five major Search Engines were Yahoo!, Magellan, Lycos, Infoseek, and Excite joined the deal. Excite was one of the major Internet portals of the 1990s (along with Yahoo!, Lycos and Netscape) and in general one of the most recognized brands on the Internet.

Dogpile was developed by Aaron Flin it began operation in November 1996 and was a metasearch site. It searched multiple engines, filtered for duplicates and then

presented the results to the user. Inktomi software was incorporated in the widely-used HotBot search engine, which displaced AltaVista, as the leading web-crawler-based search engine, and which was in turn displaced by Google [CITATION Sey11 \l 1033].

HotBot is one of the early Internet search engines and was launched in May 1996. In contrast with its competitors, it was updating its search database in a higher frequency. It managed to be one of the most powerful search engines of its day, with a spider capable of indexing 10 million pages per day. This meant HotBot not only had the most up to date list of available new sites and pages, but was also capable of re-indexing all previously indexed pages to ensure they were all up to date as well.

At the same year, Garrett Gruener and David Warthen in Berkeley, California founded Ask Jeeves (Ask). The original idea behind Ask Jeeves was to allow users to get answers to questions posed in everyday, natural language, as well as traditional keyword searching. The current Ask.com still supports this, with added support for math, dictionary, and conversion questions.

Google (1998). Google largely owes its success to a patented algorithm called PageRank, that helps rank web pages that match a given search string. This innovative algorithm, analyses human generated links, assuming that web pages linked from many important pages are themselves likely to be important. Google algorithm computes a recursive score for pages, based on the weighted Page Rank sum of the pages linking to them. In addition to PageRank, Google over the years has added many other secret criteria for determining the ranking of pages on result lists. The exact

percentage of web pages that Google indexes are not known, as it is very hard to actually calculate. What is more, Google not only indexes web pages but also takes snapshots of other file types such as PDF, Word and Excel documents, spreadsheets, plain text files, etc.[CITATION Sey11 \l 1033].

AlltheWeb (1999). AlltheWeb is an Internet search engine that owes its existence to the Tor Egge's doctorate thesis, at the Norwegian University of Science and Technology. AlltheWeb was used primarily as a show piece site for FAST's enterprise search engine. A fresher database, more advanced search features, search clustering and a completely customizable look are some of the advantages AlltheWeb had over Google.

Teoma (2000). Professor Apostolos Gerasoulis and his colleagues at Rutgers University in New Jersey founded Teoma in 2000.Teoma was completely different from the others search engines, because of its link popularity algorithm. Unlike PageRank Algorithm by Google, Technology of Teoma analyzed links in context, to rank a web page's importance within its specific subject. For instance, a web page about "music" would rank higher if other web pages about "music" link to it.

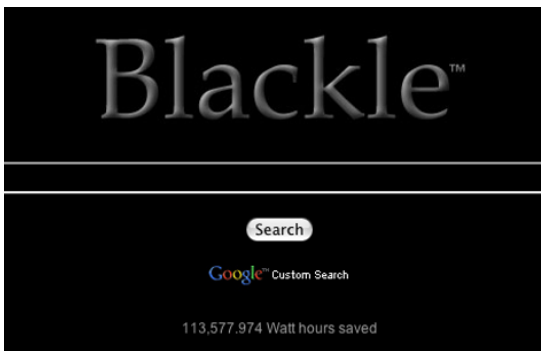
Yahoo! Search (2004). In the late 1990s, Yahoo! had been evolved into a full-fledged portal with a search interface. In 2003, the company decided to purchase Overture Services, Inc., which owned the AlltheWeb and AltaVista search engines. Despite the fact that Yahoo! owned multiple search engines, they preferred to keep using Google's search engine for its results. In 2004, it launched its own search engine based on the combined technologies of its acquisitions and Yahoo! Search became its own web crawler-based search engine, with a reinvented crawler called Yahoo! Slurp [CITATION Sey11 \l 1033].

MSN Search (2005). Initially, Microsoft comprised a search engine, index, and

web crawler. MSN Search first launched in 1998 and used search results from Inktomi. In 1999, launched a version which displayed listings from Looksmart blended with results from Inktomi. Only for a short time in 1999 were used results from AltaVista. In 2004, Microsoft upgraded MSN Search to provide its own self-built search engine results, the index of which was updated weekly and sometimes daily. The upgrade took place in February of 2005 and in addition to this, the company started providing its search results to other search engine portals, in an effort to better compete in the market. In June 1, 2009 Microsoft's rebranded search engine and Bing was launched [CITATION Sey11 \l 1033].

Wikiseek, Sproose & Blackle (2007). Wikiseek was a search engine that indexed Wikipedia pages and pages that were linked to from Wikipedia articles. The search engine was founded by Palo Alto and was officially launched on January 17, 2007 [CITATION Sey11 \l 1033].

Bob Pack in 2007 founded the first consumer search engine called Sproose. That particular search engine provides web search results from partners including MSN, Yahoo! and Ask. Sproose intends to have better-quality results than algorithmic search engines because its users have the ability to influence the ranking order. This is achieved by voting positively for websites (which moves them up in the order of search results) and deleting bad or spam results.



Blackle is a website powered by Google Custom Search, which aims to save energy by displaying a black background and using grayish-white font color for search results. The concept behind Blackle is that computer monitors consume less energy by displaying much darker colors[CITATION Sey11 \l 1033].

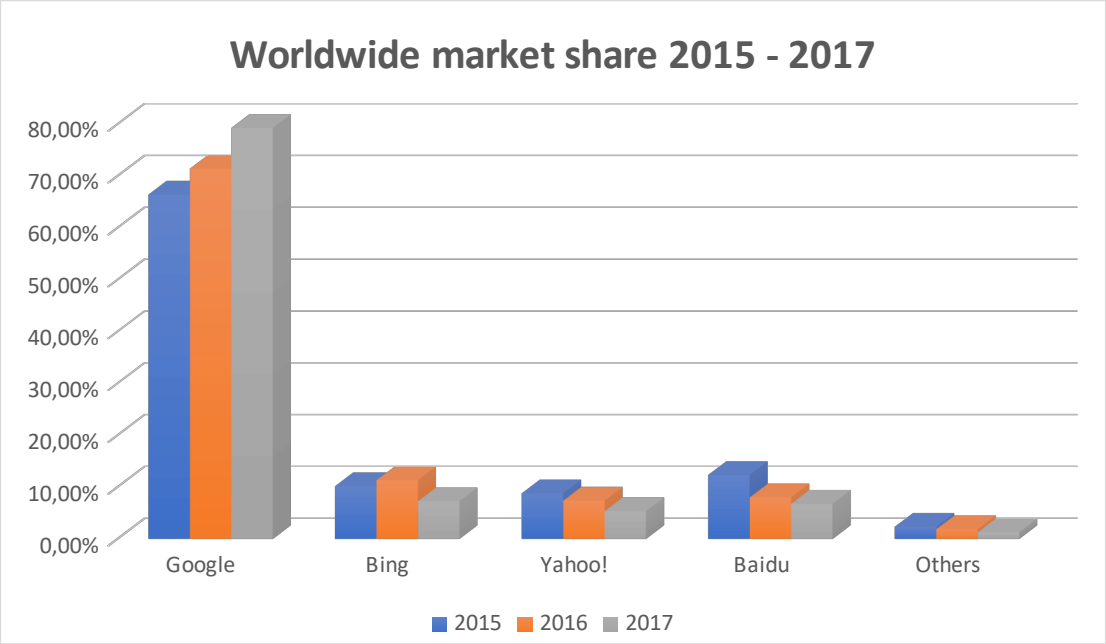
2.7 Most popular search engines in the world and market share for 2017

So, which are the top 10 most popular search engines in the world? Besides Google and Bing there are many other search engines which may be unknown to the majority of people, but still serve millions of search queries per day. It is estimated that 6.586.013.574 searches are made on a daily basis [CITATION All17 \l 1033].

The table 2.7.1 represents the 10 most popular search engines as derived from eBizMBA Rank which is a continually updated average of each website's U.S. Traffic Rank, from Quantcast and Global Traffic Rank from both Alexa and SimilarWeb. The list is ordered by most to least popular worldwide and except from the ranking, it is also represented the estimated unique monthly visitors [CITATION eBi17 \l 1033].

Top 10 Most Popular Search Engines 2017					
Search Engine	Estimated Unique Monthly Visitors	EBizMBA Rank	Quantcast Rank	Alexa Rank	SimilarWeb Rank
Google	1.800.000.000	1	1	1	1
Bing	500.000.000	33	8	40	43
Yahoo!	490.000.000	43	8	56	67
Baidu	480.000.000	54	150	4	9
Ask	300.000.000	205	329	110	177
Aol Search	200.000.000	273	350	276	194
DuckDuckGo	150.000.000	392	421	505	251
WolframAlpha	35.000.000	1878	1773	1817	2044
Yandex	30.000.000	2190	3228	2120	1221
WebCrawler	25.000.000	2955	1137	2289	5438

Table 2.7.1: Top 10 most popular search engines for 2017



According to www.netmarketshare.com, which is one of the most reliable websites on statistics for internet technologies, the market share for search engines from 2015 to 2017 is divided as follows [CITATION net17 \l 1033]:

Figure 2.7.1: Search engines worldwide market share 2015 – 2017, (source: netmarketshare.com)

	2015	2016	2017
Google	66,41%	71,41%	79,26%
Bing	10,16%	11,32%	7,27%
Yahoo!	8,76%	7,39%	5,35%
Baidu	12,33%	8%	6,80%
Others	2,34%	1,88%	1,32%

Table 2.7.2: Search engines worldwide market share 2015 – 2017, (source: netmarketshare.com)

- Google's search engine is at the top of the market share ranking with a significant difference from the others search engines and its percentage is slightly lower than 80%. In the second place comes Bing with 7,27%, while in the third place is Baidu search engine, that is basically used by the Chinese, with 6.8%, Yahoo! collects a rate of 5,35% and all the others share the amount of 1,32%.

- It is easily observed that over the last three years, google facing a significant improvement and is growing the gap with its competitors.
- The search engine with the biggest losses over the last three years is Baidu, which has a drop of 5,53%

Internet users in Greece do not differ in anything with the rest of the world and according to http://gs.statcounter.com, the 97,32% trust and use google for their searches, while Bing and Yahoo! come to second and third place with 1,22% and 1,09% respectively [CITATION sta17 \l 1033].

In conclusion, internet users consider search engines as the central reference point for exploring online content. The results of their queries, transfer them to websites and information that they could not access otherwise. From the analysis made, Google, Bing, Baidu and Yahoo are dominating the search engine market and have been used by most people in the world, but the largest market share for 2017 is belonging to Google with 79,26%.

Chapter 3: Greek Tourism Industry

As presented in chapter 2, the spread of the Internet is an indisputable fact and makes it one of the main ways of information and communication. It is therefore necessary for the tourism industry to adapt it to its own needs and to use it for its advance. With a strategic use of internet, countries based on tourism have an amazing opportunity to promote it worldwide and to boost their tourism market.

Greece is a well-known tourist destination for several decades and the Greek economy heavily relies on the tourism industry. Over the years, Greece was developed as a major tourist destination and became able to successfully fulfil all the high demands, of tourist requirements. In our days, Greece is established as one of the basic choices of tourists traveling either individually or organized. Specifically, in 2015 more than 23 million of people visited Greece [CITATION Hel161 \l 1033]. If it

is considered that the impact of the economic crisis is evident in society, the rise of tourism is a very encouraging element and a way of halting the economic crisis.

3.1 History of tourism in Greece

The start of tourism development in Greece began in the early 1950s. The country was at the initial stage of development as a tourist destination, without developed tourist infrastructure and without the required promotion and advertising. However, the number of tourists visiting the country was continuous rising and Greece gained reputation as a tourist destination in the rest of Europe. The increase in the number of visitors, has greatly prompted by the devaluation of the drachma (national currency) in 1953, which turned Greece into a very cheap tourist destination.

At the same time, the establishment of the Hellenic Tourism Organization (EOT) takes place and the first large investments in tourism infrastructure are being launched. The government in an attempt to contribute to the growth of tourism, financial helps the developing tourism businesses. All these have resulted in the constant increase of tourists visiting the country at least until the mid-1960s. In addition to this, by the middle of the same decade, growth has also been observed in domestic tourism. Until then the concept of tourism was mainly connected with residents of other countries (if not entirely to a large extent), tourism is now starting from the permanent residents of Greece. Investments in tourism development and tourism infrastructure were appearing, resulting in better exploitation of the country's tourism benefits. Thus, tourism became very important, as it started to affect factors such as employment and in general the economic growth of the country.

In the following decades, the 1970s and 1980s, Greece has consolidated as an international tourist destination. The available tourist infrastructures, offering a complete tourist product, that fulfills the needs and requirements of tourists. For these decades, tourist development was about certain and limited areas. The decade

of 90s was the beginning of the expansion and other sub-regions developed and became holiday resorts across Greece.

Today in Greece there are more than 9,500 hotel units and tourist accommodation while the geographical coverage is expanded in the whole country [CITATION Hel162 \l 1032]. This is also due to the fact that Greece is made up of too many islands and have been officially recorder more than 6,000 islands [CITATION vis17 \l 1033].

3.2 The first tourist offices and their evolution

The first major tourist agent mentioned in the history of traveling is Thomas Cook in England. In 1841, he organized a day trip from Leicester to Loughborough for 540 people. In 1845, he became the first official tourist agent in Europe [CITATION tho17 \l 1033]. In USA, the first official tourist office was founded in 1915 by American Express. In Greece, the first tourist offices operated in the early 1950s. The main initial activities had to do with migratory tourism and they also operated as representatives of foreign airlines companies.

3.3 Tourist agent's fee

The main characteristic of the tourist agent's payment, for the provided services, is that he is not paid exclusively by the customer. In his payment, it is also calculated the commissions paid to him by the airline and shipping companies, hotels, tour operators, car rental companies etc. As it is clearly understood, the way in which the tourist agent is paid, limits his ability to provide discounts or reduced prices to his customers, as the products and services he offers are on behalf of other companies. Especially in our days, the commissions provided by the airline and shipping companies are exceptionally low, about an average of 1% and 5% respectively

[CITATION Ama07 \l 1033]. As a result, the tour operators often are forced to put an extra commission on air tickets about 5 to 15 euros for the internal flights and 10 to 20 euros for abroad flights according to the destination and the office.

3.4 How internet affect tourism

In order to understand how important is internet in the tourism industry it is worth mentioning the survey carried out in November 2016 by the Hellenic Statistical Authority. According to this survey, data were published on the reasons why Greeks are using the internet on a daily basis. In the list with the most common internet activities the answer "Using services related to travel or travel related accommodation" was given in a percentage of 39,9% [CITATION Hel16 \p 5 \l 1033].

Today, the use of internet from the entire spectrum of the tourism industry is considered necessary as it provides the following possibilities:

- Direct access to multiple sources of information on a global scale.
- Exchange of messages between interested parties (businesses and Consumers) quickly and economically.
- Direct and effective customer support.
- Worldwide promotion of products, services and information
- Remote Selling of products and services over the web
- Improvement of business strategy procedures
- Significant reduction in the communication cost with customers, suppliers or intermediaries.
- Implementation and promotion of alternative advertising scenarios, depending on the customers.

All of the above activities ultimately work for the benefit of the final consumer, who has all the information available, in order to find the best solution. Through internet, the consumer will have access to whatever business he wants and will be able to close remotely all the pending issues, from a single ticket to the design of a complete holiday package.

3.5 E-tickets in the Greek maritime sector

In sea transport and especially in coastal shipping the electronic ticket is not used in the same way as in air carriers. The passenger is not able to buy a ticket, from the shipping company's site or from a tourist office, and go straight to the port in order to board and travel with the ferry to his destination. In contrast, it is essential for each passenger to have the ticket with him, so that the check could be carried out during boarding. So, when referring to the e-ticket for ferry boats, we refer to the ability of a passenger to buy only the reservation code and the ticket number. In other words, to book his ticket with an electronic way from a remote place and then he is obliged to pick up his ticket at some other time.

The electronic ticket to the maritime transport is a type of prepaid ticket which is called PTA (prepaid ticket in advance). The PTA method works as follows: passenger visits the ferry company's website and after choosing the desired route, purchases the ticket by paying with a credit card. Then, he has to go to the company's headquarters or to the central agency office of the company, in order to issue them and after this to be able to travel. By this way, the PTA reservation method can be considered as electronic ticket.

As it is easy to understand the e-ticket of the shipping companies does not have the advantages of the airline e-ticket, which means that the passenger does not benefit to a large extent from its use. Do not forget that one of the benefits of an electronic ticket is to save time and in the case of shipping e-ticket the passenger is benefited only by avoiding the transaction and in some cases saves money as some companies offer additional services with the purchase of e-ticket.

However, the development of e-tickets to maritime transportation is still at the beginning and so there is a belief that in the coming years, there will be some steps forward so that the user will benefit from its use as it is happening with the air tickets respectively.

3.6 The company FORTHcrs

Is a Greek company for the research, development and marketing of software and e-

administration products. Was established in December 1999 and is a member of the FORTHnet group. FORTHcrs is a Greek ticketing company with major projects and successful integrated interventions in the fields of tourism, transport and entertainment [CITATION for17 \l 1033]. In the summer of 2001, designed, created and put into operation the integrated distribution system for electronic reservations on the coastal shipping, which was called SeaConnect. At the same time, developed a system for the management of reservations for coastal shipping companies, paying attention to cover the special needs of the Greek islands. This system is currently used from the majority of Greek shipping companies and it is also adopted by some foreign companies.

The following projects in the transport area have been implemented by FORTHcrs:

- Reservations and ticketing systems for buses, trains and airplanes
- Entry control systems (check-in)
- Voice information portals
- On-line booking - internet applications
- Information kiosk
- Automated ticketing machines
- Destination management systems

SeaConnect is an electronic reservation distribution system linking the reservation systems of 45 Greek shipping companies to a single platform through the application OpenSeas™ for reservation and ticketing. This system is one of the top examples of electronic B2B intervention in Greek tourism.

With OpenSeas™ software more than 7.000 tourist offices in Greece and abroad, have access to all companies with which they cooperate, with the same user-friendly platform. Through the FORTHcrs platform more than 24.000.000 tickets are reserved annually and it is estimated that in the peak demand period the reservations are more than 160.000 tickets daily for all Greek shipping companies. Except from the reservations mentioned above, at the same time the application service multiple queries from users about availability and shipping routes [CITATION for171 \l 1033].

The operation of SeaConnect system has brought many benefits to the shipping industry. A new tourist office can easily extend its co-operation with other shipping companies, as well as for a any new shipping company to acquire immediately a sales network, with reservation and ticketing system, without having experience and the need to invest in equipment. With the OpenSeas™ application, all the tourist offices have uninterrupted operation, quick response and customer service (direct printing of a ticket to thermal printer). In addition, many tour operators use the software OpenSeas on site, that gives them the opportunity to sell ferry tickets from their websites.

What is more, all the travel and information content that passengers and tourists need, in order to plan their holidays and travel to the big and small islands of Greece, is available on the website www.openseas.gr. This content was missing from the Greek Internet and has significant number of visits both from Greece and other foreign countries.

SeaOnLine is a management reservation system for shipping companies. It is a software platform consisting of a set of specialized applications for maritime sector. These applications provide a fully customized operating environment for the handling of reservations and ticketing. The main advantages of this platform are the high configuration level (depending on the needs of each customer), its credibility and flexibility in terms of defining different trade policies, the ability to cooperate with external systems such as distribution systems and ERP systems. Every shipping company using it, acquires a complete set of operations such as:

- Route management (one-way, return, cyclic, random etc.)
- Management of ship structures of all types (conventional, high speed, with or without vehicle space etc.)
- Information and statistics (data for the Minister for Mercantile Marine and Island Policy)
- Group management for passengers and vehicles

- Usage of fare quotation. With this application, the calculation of the discount when ticketing is made automatically. Companies offers are automatically detected, based on characteristics of the entire booking.
- Managing pricelists and pricing policies
- Handling of trucks and garages
- Member card management and historical data
- Check in and check out system

The software platform is constantly evolving and currently allows 48 shipping companies to manage their commercial policy, reservations and tickets. The SeaOnLine platform is operating in FORTHcrs facilities and also provide services for foreign shipping companies such as Maritime Way in Italy, Montenegro Lines in Serbia/ Montenegro, and Bodrum Express in Turkey. All these companies do not need to provide hardware and software to operate electronic reservation systems.

3.7 The travel agency Airtickets.gr

The Airtickets company is a pioneer in the field of travel agencies since it is one of the first companies in the field in our country to operate on the Internet. Began its operation in 2000 by providing traditional travel agency services such as booking / selling air tickets, renting rooms, cars, organizing excursions and holiday packages. The company recognized the strength of the internet as a promising distribution channel for the travel agency sector. Its goal was to provide services to its customers 24 hours a day and services similar to those provided by a travel agency. This includes the ability to handle an order/ booking of a ticket, room or vacation package reliably and securely. Also, its website should be comprehensive, giving plenty of information to the user and the ability to compare and choose between different prices, companies, destinations, airports, dates, hotels and holiday packages. One of the main reasons for success was the useful design of the webpage, that made it easy to navigate even for a novice visitor without special Internet knowledge. Finally, due to the customer-centric nature of the branch, it was

very important to collect and organize customers data in a database that offered them a tool for marketing and future sales promotion of the company. Under this philosophy, the page was created and took its final form which is very user-friendly, comprehensive and functional website that provides online a wide range of tourist information. More specifically, the user has the ability to:

- See all the cities that can travel, as well as the corresponding airport codes
- To be informed about 20,000 flights to more than 150 airlines companies around the world.
- Has an extensive presentation of flights based on their prices starting from the most economical
- To cancel a reservation which has already be made
- To choose between a variety of organized guided tours
- To choose the form of payment, either through a credit card, deposit into a bank account or paypal
- To be informed about the transaction security provided by the website
- To be informed about the weather forecast for each destination

The benefits of airtickets.gr activity on the Internet have become visible in a very short period of time and today the company is one of the most recognizable online travel agents with more than 70,000 visits per day. Except for our country, the company is already active in 15 other countries and Chinese market has been the last addition [CITATION air17 \l 1033].

3.8 The benefits of a good web site

Every ferry company in Greece, irrespective of its size, has its own web site. In order customers to search for information and book online, companies' websites should be user friendly and meet the consumers' needs. Until the process of buying a ticket is completed, a customer makes several decisions: to enter a website, to navigate, to purchase, and to pay. This environment must be designed so that it can support customer decision making. To retain customers, the web site must be pleasing and enjoyable, and create a task with natural flow. Customers have different needs and

motivation, which affect decision making. It is therefore important to customize the design of the website environment.

Throughout the transaction, the user is only one click away from leaving. Unlike physical stores, there is no sales person and no social pressure to stay. It is easy to find substitute stores and compare prices. As figure 3.8.1 illustrates, there are many dependent variables that affect the decision to purchase. A successful purchase comes about as a consequence of a series of five decisions: the decision to visit, the decision to navigate, the decision to buy, the decision to pay, and the decision to keep. Each of these should be supported by information that can facilitate efficient decision-making. The most resolute decision is the decision to pay [CITATION Hel00 \l 1033].

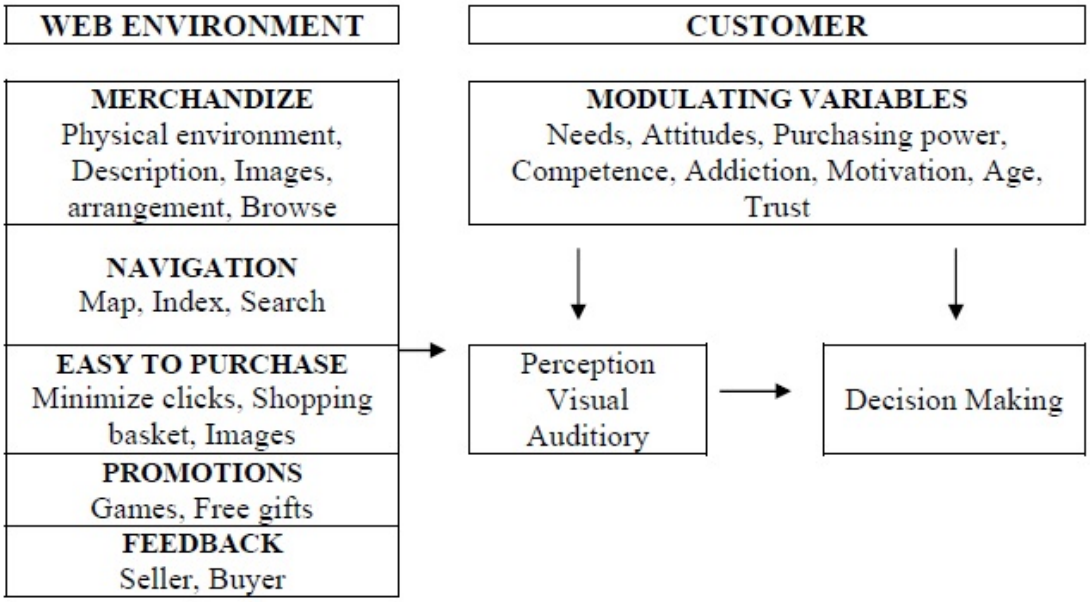


Figure 3.8.1: A system model for human factors research in e-commerce

The success of eCommerce is highly related to ease of use. According to the study "The benefits of good web design" it was shown that 39% of shoppers in a user test failed in their buying attempts because sites were too difficult to navigate [CITATION usa17 \l 1033]. Some general principles for design of eCommerce environments include:

- Simplicity: Do not compromise usability to gain function.
- Support: User should be assisted with proactive assistance.
- Visibility: Make objects visible and intuitive to control.
- Reversible action: Make actions reversible.
- Feedback: Provide visible feedback on all user actions.
- Accessibility: Make all objects accessible at all times.
- Personalization: Allow the user to customize the interface.

A well-designed web site combines many design elements: Good layout, elective graphics, exciting content and easy navigation. Many developers of web pages have been mistakenly guided by the challenges of technology rather than by user needs. A common mistake, however, is the use of dynamic graphics. They may look spectacular, but they are actually distracting, and the download time is often prohibitive [CITATION Del14 \l 1033]. Another mistake made by sellers on the Internet was their misjudgment regarding the necessity to continually update their site. Post-purchase activity involves consumers returning to the sellers' site with queries, for new information, and to repurchase. Such buyers demand to be treated to new information at every visit [CITATION Pep98 \l 1033].

A main objective of the companies is to achieve repeat visits to their web sites. The most important factors which drive repeated visits to websites are the high-quality content, the ease of use, the speed of the download and the provision of up-to date information.

3.9 Advantages and disadvantages of the electronic ticket

The e-ticket offers many benefits to travelers, travel agents and companies. It greatly reduces the use of paper, reduces the problem when some passengers lose their ticket, permits changes at the last minute, gives the opportunity to check in 30 minutes before the flight, simplifies the ticket payment process and at the Airports it does not create passenger congestion in the check in desks. All the advantages of electronic tickets are listed below in detail:

- Direct ticket purchase without visiting a travel agency (for airport tickets only as for the ferry tickets the passengers have to receive it after the reservation)
- Route check and availability of seats at any time.
- Money and cost saving. An e-ticket in the case of early booking is more economical than the regular ticket.
- Additional offers (bonus & miles) in case of an electronic ticket purchase (This applies to certain companies and for specific periods of time).
- Time saving. The online ticket can be booked from any computer with an internet connection, at any time of the day.
- The passenger can travel even in case of theft or loss of his ticket, since the reservation code is stored in the electronic data bases of each company.
- In case of theft or ticket loss, the passenger is not charged with ticket reissue
- In the case of an electronic ticket, changes can be made to the route, up to 30 minutes prior to the start of the flight check
- The e-ticket ensures a quicker check-in at the airport and therefore less trouble and delay

Despite the many advantages, there are also some problems related to the implementation of this new service. Thus, we can mention the following disadvantages:

- Security issues during payment (fraud using credit cards)
- E-commerce problems (using personal information for promoting other products)
- Difficulty in changing or canceling the ticket (procedures are complicated and time-consuming in case of emergency changes)
- Difficulty in purchasing in case of personal computer failure or network connection
- Difficulty in purchasing in case of company's system failure.

To conclude, the elements that characterize electronic tickets are the comfort and flexibility that they provide to passengers, less anxiety, shorter waiting times, usage of self-service machines. Companies, also benefit from e-tickets by reducing

operating costs from issuing, distributing, storing and booking tickets. In any case, the advantages over a regular ticket are many and surely overcome its disadvantages.

Chapter 4: Search Ranking Factors

Search engines work by using algorithms to evaluate websites by topic and relevance. This evaluation is used to structure pages in the search engine index, which ultimately results in user queries displaying the best possible ranking of the results display. The criteria for the evaluation of web pages and to produce this ranking are generally referred to as ranking factors.

The term “Ranking Factors” describes the criteria applied by search engines when evaluating web pages in order to compile the rankings of their search results. Ranking factors can relate to a website’s content, technical implementation, user signals, backlink profile or any other features the search engine considers relevant.

Search Engine Marketing (SEM) is a form of web marketing and was created by the ever-increasing competition and the need of the websites to hold a higher position on that list. SEM aims to promote websites and increase their visibility on search engine results pages (SERPs) through Search Engine Optimization (SEO). Understanding ranking factors is a prerequisite for effective search engine optimization.

4.1 Description of the factors

In this part of the study it is represented and analyzed 43 different factors that can affect the results of a search, in a search engine. These factors based on their content are divided into the following categories: Technical, user experience, content, social networks, backlinks, mobile and alternative sources.

4.1.1 Technical

This chapter is concerned with on-page factors that are primarily technical and not directly linked with a page's content.

#1. Title tag

A title tag is an HTML element that specifies the title of a web page. Title tags are displayed on search engine results pages (SERPs) as the clickable headline for a given result, and are important for usability, SEO, and social sharing. The title tag of a web page is meant to be an accurate and concise description of a page's content. Google typically displays the first 50 - 60 characters of a title tag. Title tags are a major factor

in helping search engines understand what your page is about, and they are the first impression many people have of your page. Title tags are used in three key places: (1) search engine results pages (SERPs), (2) web browsers, and (3) social networks. Your title tag determines (with a few exceptions) your display title in SERPs, and is a search visitor's first experience of your site. Even if your site ranks well, a good title can be the make-or-break factor in determining whether or not someone clicks on your link.

#2. Meta Description

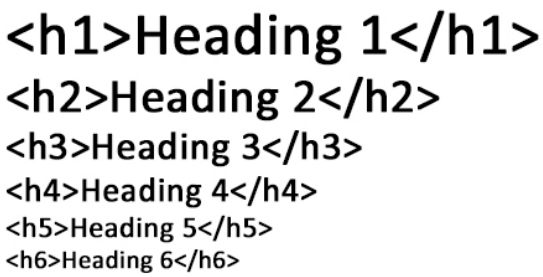
Meta descriptions are HTML attributes that provide concise summaries of webpages. They commonly appear underneath the blue clickable links in a search engine results page (SERP). Meta descriptions can be any length, but search engines generally

truncate snippets longer than 160 characters. It is best to keep meta descriptions long enough that they're sufficiently descriptive, but shorter than that 160-character limit. Meta description tags, while not tied to search engine rankings, are extremely important in gaining user click-through from SERPs. These short paragraphs are a

webmaster's opportunity to "advertise" content to searchers, and searchers' chance to decide whether the content is relevant and contains the information they're seeking from their search query. A page's meta description should intelligently (read: in a natural, non-spam way) employ the keywords that page is targeting, but also create a compelling description that a searcher will want to click. It should be directly relevant to the page it describes, and unique from the descriptions for other pages.

#3. H1

The h1 is an HTML tag that indicates a heading on a website. It is typically the largest and most important tag on an individual page of a website. It may be used to introduce the content of that page, to name a topic or to simply reflect the name of the page itself.



Websites choose to use their H1 tags differently. Some make them very long, writing them out like full sentences. Others may limit them to one or two very simple words that match the actual title of the page. By using the H1 tag websites have the opportunity to include keywords most relevant to the content of the page. While H1 tags may not directly impact SEO, it's still valuable to spend time optimizing your H1 tags because of the indirect benefits they provide. For example, including H1 tags on your website can improve user experience, which can help to improve your rankings in search results.

The combination of relevant keywords at the top of the page and the presence of desired information in this location makes the H1 tag so important. It may not be the strongest thing you can do for search engine optimization, but H1 tags are still very important for SEO because they tell both search engines and website visitors what the content of pages will be about.

#4. H2

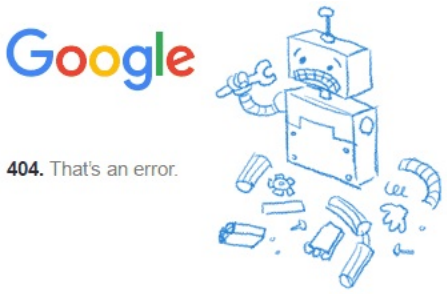
If Header 1 text is your book title, then Header 2 is your chapter heading, which give readers a clearer picture of what your content covers, and allow them to quickly find the specific information they want. A site that lacks H2 becomes *significantly* more difficult for the readers to quickly judge what the overall subject is, and to find the specific piece of information they want. If there is no H1 and H2 text to guide them, it is very likely to lose patience and leave.

#5. Alt Attribute

An alt tag, also known as "alt attribute" and "alt description," is an HTML attribute applied to image tags to describe the appearance and function of an image on a

```
img src="pancakes.png" alt="Stack of blueberry pancakes with powdered sugar">
```

page. Alt attributes is very useful for an image as it will be displayed in place of an image if an image cannot be loaded. What is more, alt attributes provide better image descriptions to search engine crawlers, helping them to index an image properly. Search engines and other robots cannot interpret images, but images can play a crucial part in how people interpret a particular web page. Alt tags solve for this by providing text which is read by the search engines. When Googlebot or other search engine crawlers inspect a page, images with properly formatted alt text contribute to how the page is indexed and where it ranks.



#6. Broken Links

When web developers delete a linked page, place an incorrect URL in a text link, or change a URL address without correctly updating the URL, a broken link is created. In other words, when searchers (and/or web crawlers) browse through a site and click on a dead link, they will be directed to a page with a 404 error. One of the ways search engines rank websites is by crawling their links. If there are dead links on a website, it can stop search engine spiders from crawling this site and, thus, indexing it. To sum up, the more broken links a site has the worse the site will be ranked.

#7. WWW Resolve

Duplicate content is content that appears on the Internet in more than one place. That “one place” is defined as a location with a unique website address (URL). If the same content appears at more than one web address, we have duplicate content. Sometimes duplicate content impact search engine rankings. When there are multiple pieces of, as Google calls it, "appreciably similar" content in more than one location on the Internet, it can be difficult for search engines to decide which version is more relevant to a given search query.

If a site has separate versions at "www.site.com" and "site.com" (with and without the "www" prefix), and the same content lives at both versions, it has effectively created duplicates of each of those pages. The same applies to sites that maintain versions at both http:// and https://. If both versions of a page are live and visible to search engines, we may run into a duplicate content issue. Fixing duplicate content issues all comes down to the same central idea: specifying which of the duplicates is the "correct" one.

Whenever content on a site can be found at multiple URLs, it should be canonicalized for search engines. The three main ways to do this is by using a 301 redirect to the correct URL, the rel=canonical attribute, or using the parameter handling tool in Google Search Console.

#8. IP Canonicalization

Canonicalization refers to how a website can use different URLs for the same piece of content (usually the entire web page). For example, every website resolves to a particular IP address of the hosting machine. It means, the website can be accessed via its domain name and at the same time it can be accessed using the *IP address* as well. The canonicalization problem occurs, when a website's IP address and domain name display the same page, but do not resolve to the same URL. This behavior is not search engine friendly, because the spiders are unsure as which URL is the correct one. The search engines do not like duplicate contents and so the IP canonicalization is very important in SEO point of view.

#9. Robots.txt

Robots.txt is a text file, webmasters create to instruct web robots (typically search engine robots) how to crawl pages on their website. The robots.txt file is part of the robot's exclusion protocol (REP), a group of web standards that regulate how robots crawl the web, access and index content, and serve that content up to users. In practice, robots.txt files indicate whether certain user agents (web-crawling software) can or cannot crawl parts of a website. These crawl instructions are specified by "disallowing" or "allowing" the behavior of certain (or all) user agents. After arriving at a website but before spidering it, the search crawler will look for a robots.txt file. If it finds one, the crawler will read that file first before continuing through the page. This file contains information about *how* the search engine should crawl. If the robots.txt file does *not* contain any directives that disallow a user-agent's activity (or if the site doesn't have a robots.txt file), it will proceed to crawl other information on the site.

#10. XML Sitemap

This protocol helps Google and other main search engines to easily understand your website structure while crawling it. It was first introduced by Google in 2005, with MSN and Yahoo offering their support to the protocol a year later. Sitemaps are

known as URL inclusion protocols**,** as they advise search engines on what to crawl. It comes in opposition to robots.txt files that are an exclusion protocol as it tells search engines what not to crawl. A benefit of having a XML sitemap is that gives to search engines page priority and thus crawl priority. By adding a tag on a XML sitemap saying which pages are the most important, bots will first focus on this priority pages.

#11. URL Rewrite

A website "www.mysite.com" hosted by a Web server and has a root folder. This root folde**r** is a file like all the ones anyone could see on his computer. By default, when a browser requests from a Web server the content of the URL "http://www.mysite.com/folder1/file1.html", the server looks if there is a file called "file1.html" inside the sub-folder "folder1". If this file is found, the content of the file is returned to the browser. This basic explanation therefore establishes a direct link between a URL and the path of a traditional computer file. For practical reasons, the Web servers pass through a supplementary step that enables them to separate the URL seen by the visitor and the physical path of the correspondent resource. This additional step is called URL Rewriting. Therefore, the URL Rewriting is a set of rules created by the webmaster in order to transform input URLs into different ones.

There are 2 main reasons to rewrite a URL: the first has to do with SEO and the second is because search engines like URLs without prolonged query strings. For instance, the URL abcdefg.com/4/basic.html is simpler to index than http://www.abcdefg.com/cfi-bn/geon.ploi?id=4&view=basic. Many search engines can identify the first URL whereas the 2nd one could “confuse” the search spiders. This could even result in missing out on important information that exists in the URL. This can be a significant barrier for enhancing the ranking of your internet site. With clean URLs, the folder names are more easily distinguishable and can develop the links to the actual keywords.

#12. Underscores in the URLs

The recurring question in SEO URLs is whether to use a hyphen (these-are-hyphens), or dash, or an underscore (these_are_underscores) to separate your words. Does Google recognize both of them as word separators? The short answer is that we should always use a hyphen for our SEO URLs. Google treats a hyphen as a word separator, but does not treat an underscore that way. Google treats and underscore as a word joiner — so red_sneakers is the same as redsneakers to Google. This has been confirmed directly by Google themselves, including the fact that using dashes over underscores will have a (minor) ranking benefit.

#13. Blocking Factors

Flash movies, which are a common way to incorporate information and professional visual design into a Web page, can cause problems for SEO as they do not conform to code conventions used by search engines and are not easily read or understood by Web-searching algorithm. This has particular ramifications for programs such as Google's Web crawling algorithms, which read and archive Web page data by "crawling" through hyperlinks in web pages and ranking data based on site inter-linking. Since Flash does not adhere to certain HTML standards, Google crawlers cannot digest the information included in Flash objects the same way.

In October 2014, Google announced that it would expand its warning system to help mobile users navigate away from Flash-heavy websites. The warning informs users on mobile devices that Flash sites “may not work” on their devices. If your website attracts a large mobile audience and uses Flash, Google’s warning is likely to have a serious negative impact on the amount of people that click through to your website, reducing your traffic from search.

Google, Apple and numerous other technology companies have made the message clear: Flash, while useful technology early in the web, is now more of a hindrance than a help from a usability and search perspective.

#14. Custom 404 Page

The 404 page is one chance to turn a negative situation into a positive one and every website should take advantage of it. Some people are under the misconception that they should overlook 404 pages because they are the last place you want your users to end up on your website. However, not having a custom 404 page can be detrimental to the user experience of a website and reflect poorly on the search rankings.

The 404-error message is appearing when the server has not found anything matching the request URL. No indication is given of whether the condition is temporary or permanent. Often webmasters will display a text 404 error but the response code is a 200. This tells search engine crawlers that the page has rendered correctly and many times the webpage will get erroneously indexed.

For SEO, a custom 404 page gives you the opportunity to put valuable links in front of your users that would have otherwise been lost with a default HTTP error page. These links could be to your homepage, your FAQ or any other relevant links that will keep them on your website when they could have otherwise left.

Web optimized 404 errors pages should contain:

- notification that the user has reached a page that does not exist
- a search box
- an easy-to-understand navigation system so the user can potentially find what they were originally looking to access
- a link to the home page

#15. Language

The hreflang tag (also referred to as rel=”alternate” hreflang=”x”) tells search engines which country a specific page is intended for, which language it is written in and provides alternatives for the same page intended for other countries and languages. It can also provide information on the default country and language if the

search engine is unable to decide which version to present (or if the page is not targeting a specific country or language). This signal will help Google increase the relevance of content it serves to users in various regions, helping reduce bounce rates, increase dwell time and, ultimately, boost conversion rates.

For example, if you create a Spanish-language version of your English-language homepage, you would tag it as "Español" by using hreflang=”es” so that searchers with an IP address that a search engine has reason to believe is in a Spanish-speaking country are served that page in Spanish instead of the English version.

Moreover, hreflang tags are important as they are giving solution to the problem with the publication of duplicate content. For instance, if a page on a website exists as a duplicate (or near duplicate) in English for the UK, US and Canada then providing hreflang annotations to inform search engines that these pages are not duplicates but intended for different audiences.

Hreflang attributes may not help you increase traffic; instead, the goal of using them is to serve the right content to the right users. They help search engines swap the correct version of the page into the SERP based on a user's location and language preferences.

#16. Structured Data Markup

Structured data is code. It’s a piece of code that can be put on a website. It is code in a specific format, written in such a way that search engines understand it. Search engines read the code and use it to display search results in a specific way.

Imagine a website with a lot of recipes. If structured data are added to a page with a recipe, the result in the search engines will change. It will be much “richer” regarding content that is shown. For this reason, we call these results rich snippets. This is what a rich snippet looks like:

Besides the title, the URL and the description of the search result, you can see how long it will take to make the absolute best ever lasagna. And, you'll see how many calories the lasagna contains.

There are all kinds of structured data. Structured data is always a code format. There is structured data for books, for reviews, for movies, and for products in your online store, for instance. In all cases, structured data add more details to a snippet in the search results.

Unfortunately, Google does not always create a rich snippet of a page, even if it has added the structured data. There are no guarantees. With structured data, websites "talk" to the search engines informing them which ingredients there are in a recipe, how long the preparation time is, and how many calories the dish will contain. Google will be able to grasp all that information instantly and can decide to show it in the search results.

So structured data is a tool to give Google detailed information about a page on a website. Google then will be able to use this information to create informative (rich) search results.

The big search engines have developed a project called Schema.org. On Schema.org can be found all the structured data markup supported by the search engines. This makes Schema.org a large collection of pieces of code. For instance, a website selling t-shirts, could show what color t-shirts it sells and what sizes it offers in its snippet.

#17. SSL Secure

The "s" at the end of the "http" part of a URL means the website is secure. HTTPS (Hypertext Transport Protocol Security) sites include the SSL 2048-bit key and can protect a site connection through authentication and encryption. When installed on a web server, an

SSL certificate activates the padlock and the https protocol and allows secure connections from a web server to a browser.

Secure websites can protect a user's connection by securing information in three layers:

- **Encryption** ensures that a user's activity cannot be tracked or their information stolen
- **Data integrity** prevents files from being corrupted as they're transferred
- And **authentication** protects against attacks and builds user trust

In 2014, Google rolled out updated algorithms across the board in favor of HTTPS websites. Then, it was a lightweight component within the overall ranking algorithm and HTTPS sites experienced only minor ranking increases. But Google indicated that they may strengthen the signal in the future. In 2015, Google stated that their HTTPS ranking boost may serve as a tie breaker if the quality signals for two different search results are equal in everything else [CITATION McG15 \l 1033]. Meaning, if your website is equal to your competitor's website in terms of speed, title tags, content freshness, etc. but your competitor's website is HTTPS and yours is not, Google will most likely rank theirs ahead of yours. Google in general, has encouraged webmasters to make the migration to a secure site for a while now and has been giving an increasing amount of weight in ranking boosts to websites that are HTTPS.

#18. Page Size (Mb)

Page size, also called page weight, refers to the overall size of a particular web page. A page size includes all the files that are used to create the web page: the HTML document, any included images, style sheets, scripts, and other media. Considering a recent study from Soasta reported that the average weight of a page is now above 2 MB. It also stated that average page size increases by approximately 8% every 6 months [CITATION Eve15 \l 1033]. The smaller the file

size of a page, the faster it will load for anyone who requests it. A delay beyond two seconds can cause viewers to abandon a page and for every additional second after that, the abandonment rate increases by 5.8% [CITATION SSh12 \l 1033]. On the other hand, the weight of a page will vary depending on the company or industry. An ecommerce website with a wide variety of photos, it's likely to have a larger page weight. In general, everyone should aim to be at or below the average. In the following graph, it is represented the average page size from 2011 until 2015 with a breakdown of how all those kilobytes are separated. The most notable ones are images, scripts and video content.

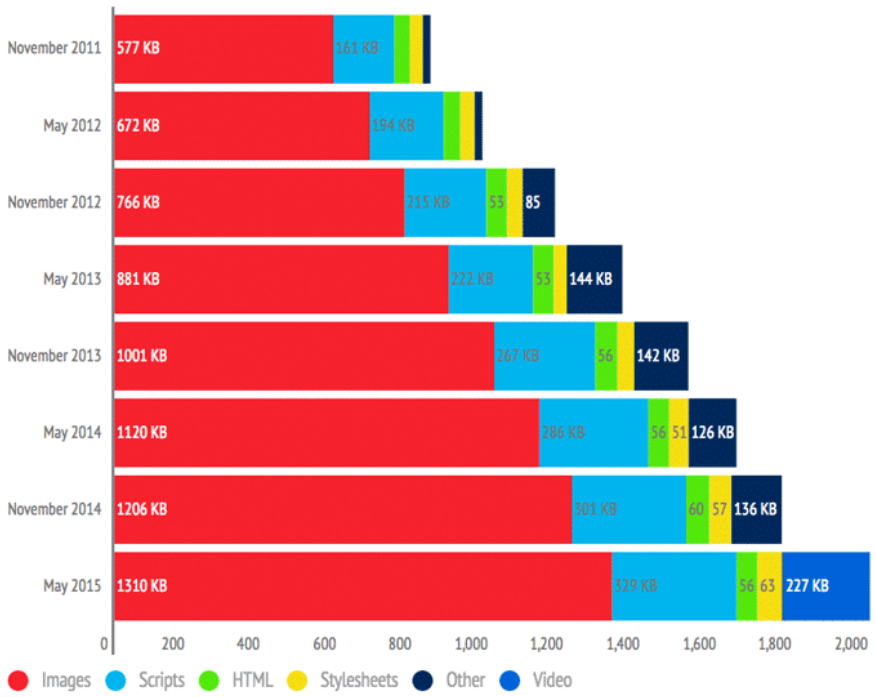


Figure 4.1.1: Average page size from 2011 to 2015, how kilobytes are separated

When it comes to page weight, images are one of the largest contributors. In order to downsize a webpage, it is good to start by resizing any unnecessarily large images. When uploading photos, keep in mind that the image dimensions should never be greater than the size of the container. If the container has a maximum width of 500 pixels, there's usually no reason to upload a 1,000 pixel-width image. In case a website wants to offer a high resolution image people could share and download it is recommended uploading a picture that is exactly the width of

the container and hyperlinking it to the high-resolution version to avoid weighing down your page.

4.1.2 User experience

The user experience is something Google is always looking to improve and for this reason the company is dedicated to understand how and why a positive experience is achieved so they can better analyze and reward that outcome in the search rankings. Below are represented some factors that are able to improve user experience and additionally search engine rankings. These factors are primarily aspects of design and usability. User experience is related to on-page optimization and fits somewhere between technical and content.

#19. Google Analytics

Google Analytics is a free Web analytics service that provides statistics and basic analytical tools for SEO and marketing purposes. The service is available to anyone with a Google account. Google bought Urchin Software Corporation in April 2005 and used that company's Urchin on Demand product as the basis for its current service.

Here are some benefits of using Google Analytics during an optimization campaign:

- Able to find out how your visitors locate your website. Other than the core set of keywords that you are optimizing, you will also be able to find out what other keywords your visitors type in to find your website. During the initial optimization campaign, the number of keywords may be little. But as time goes by, you will start getting more keywords being listed on organic listings, due to your overall optimization effort.
- Able to identify which pages and links your visitors click the most. You will be able to know which are the popular pages and links, and measure whether your optimization campaign is directing the traffic to the correct pages.

- Visitor segmentation. With this, you will be able to know how many new visitors that your search engine optimization campaign brings to you. You can segment your analytic result by new/returning visitors, geography and referral sources.

#20. Directory Browsing

A directory is simply a web site that contains a categorized listing of links from around the web. They aid surfers to locate the 'best' and most informative links for a particular category. For example, a category may be called 'Home and Garden' and in this category, there is list of links about home improvement and gardening. Directories consist of a collection of categories into which links are separated. Categories can have sub-categories to make the division of links more specific.

Directories are important tools in building link popularity and as a result help improve search engine ranking. They are an excellent source of inbound, one-way links, which are the most powerful types of links to help build link popularity. Getting listed in quality directories such as Yahoo can be more beneficial for your link popularity than a lot of links from smaller, newer directories.

#21. Desktop Pagespeed

PageSpeed Insights measures the performance of a page for mobile and desktop devices. It fetches the URL twice, once with a mobile user-agent, and once with a desktop user-agent. PageSpeed Insights checks to see if a page has applied common performance best practices and provides a score, which ranges from 0 to 100 points, and falls into one of the following three categories:

- *Good*: The page applies most performance best practices and should deliver a good user experience.
- *Needs work*: The page is missing some common performance optimizations that may result in a slow user experience. Please investigate the recommendations below.

- *Poor*: The page is not optimized and is likely to deliver a slow user experience. Please prioritize and apply the recommendations below.

PageSpeed Insights is being continually improved, to account for new performance best practices, and the provided score will change over time. A high score is correlated with a fast user experience but does not guarantee it.

PageSpeed Insights measures how the page can improve its performance on:

- *Time to above-the-fold load*: Elapsed time from the moment a user requests a new page and to the moment the above-the-fold content is rendered by the browser.
- *Time to full page load*: Elapsed time from the moment a user requests a new page to the moment the page is fully rendered by the browser.

However, since the performance of a network connection varies considerably, PageSpeed Insights only considers the network-independent aspects of page performance: the server configuration, the HTML structure of a page, and its use of external resources such as images, JavaScript, and CSS. Implementing the suggestions should improve the relative performance of the page. However, the absolute performance of the page will still be dependent upon a user's network connection and for this reason Google does not use the PageSpeed Insights tool to determine search engine ranking.

#22. Load Time (sec)

Page speed can be described in either "page load time" (the time it takes to fully display the content on a specific page) or "time to first byte" (how long it takes for your browser to receive the first byte of information

from the web server). Google has indicated site speed (and as a result, page speed) is one of the signals used by its algorithm to rank pages. And research has shown that Google might be specifically measuring time to first byte as when it considers page speed. In addition, a slow page speed means that search engines can crawl fewer pages using their allocated crawl budget, and this could negatively affect your indexation.

Page speed is also important to user experience. Pages with a longer load time tend to have higher bounce rates and lower average time on page. Longer load times have also been shown to negatively affect conversions.

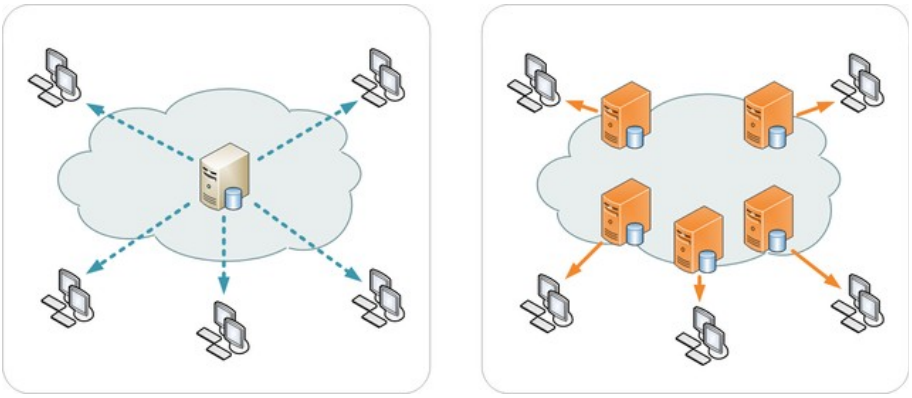
in 2010, Google announced website speed would begin having an impact on search ranking [CITATION Goo10 \l 1033]. Unfortunately, the exact definition of "site speed" remained open to speculation. The mystery widened further when Google's Matt Cutts announced that slow-performing mobile sites would soon be penalized in search rankings as well [CITATION Sch13 \l 1033]. Clearly Google is increasingly acting upon what is intuitively obvious: A poor performing website results in a poor user experience, and sites with poor user experiences deserve less promotion in search results.

#23. Ping

Ping is a computer network administration software utility used to test the reachability of a host on an Internet Protocol (IP) network. It measures the round-trip time for messages sent from the originating host to a destination computer that are echoed back to the source. A roundtrip network latency of no more than 100ms is recommended. If a large number of users coming from another continent, network latency may be as high as 200ms.

So, if ping numbers are higher than this, the solution to improving latency is pretty simple: Reduce the "distance" between your content and your visitors. If your servers are in Atlanta, but your users are in Sydney, you don't want your users to request content half way around the world. Instead, you want to move that content as close to your users as possible. Fortunately, there is an easy way to do this: move your static content into a Content Delivery Network (CDN). CDNs automatically

replicate your content to multiple locations around the world, geographically closer to your users. So now if you publish content in Atlanta, it will automatically copy to a server in Sydney from which your Australian users will download it. As you can see in the diagram below, CDNs make a considerable difference in reducing the distance of your user requests, and hence reduce the latency.



4.1.3 Content

When it comes to search rankings, the importance of good quality, relevant content cannot be understated. The aim is to give a clearer insight into which aspects of content in particular can improve the overall ranking of your site. As the trend, away from keywords and towards relevant content continues, high-ranking sites are shifting their focus from using keywords based on search queries to trying to understand the user's intention as a whole

#24. Keyword in domain

An exact match domain (EMD) is a domain name that includes the keyword phrase itself. Some people call them Keyword Domains. For example, a business which wanted to rank for the keyword "awesome cheap widgets", if chose an EMD that would look like this: www.awesomecheapwidgets.com

For years SEOs and search marketers have often purchased domains that contained their targeted keywords in order to increase CTRs (Click-through rate) and to help

gain higher rankings on Google, Yahoo and other search engines. However, nowadays and especially after some Google updates, EMDs are no longer as helpful as they once were. In 2012 Google noticed that a lot of websites were trying to improve their SEO by choosing exact-match domains and with this way they managed to win first-page ranking. In response, Google was forced to take action and in September 2012 launched the Exact Match Domain Update. This update is a filter that prevents poor quality sites from ranking well simply because they had words that match search terms in their domain names.

While there are lots of benefits having targeted words in your domain, they are not as important as they once were for SEO. In some cases, they can even hurt your search rankings if implemented incorrectly. The reason why domain names exist is to make web addresses easier for humans. Instead of having a URL like 134.27.123.12, a domain name lets you register something that people will actually be able to remember. This is why websites should avoid getting trapped into focusing 100% on search engines and take into account how humans will read their domain name when seeing it in search results or as a link on a website. As long as a website has great content visitors will want to come back again and again and the search engines will eventually follow.

#25. Keyword in Sub – Directory

A URL is human-readable text that was designed to replace IP addresses, that computers use to communicate with servers. They also identify the file structure on the given website. A URL consists of a protocol, domain name, and path (which includes the specific subfolder structure where a page is located). There are 2 main benefits of putting keywords in URL:

- Improved user experience. A well-crafted URL provides both humans *and* search engines an easy-to-understand indication of what the destination page will be about.
- Improve rankings. URLs are a minor ranking factor search engine use when determining a particular page or resource's relevance to a search query.

However, they do give weight to the authority of the overall domain itself and so, keywords used in a URL can also act as a ranking factor.

#26. Domain Age

Domain age is a minor SEO factor for your Google ranking. Among the hundreds of weighting factors that Google considers in determining how to rank search engine results is the age of your domain or in other words, how long that domain has been around.

It is worth noting that domain age is not referring just to how long a given domain name exists, but instead how long it has been since Google first indexed that domain, or saw a link to the domain. Just because a domain has been registered for 10 years does not mean that Google considers it 10 years old. In other words, a domain that has been registered for 10 years but has no actual site up, or nothing that Google has ever found, is the same as a domain that you bought yesterday.

For many years there was a belief that that older domains do better in search results and are given extra favor due to their age. The idea is that an older domain that has been established has earned a greater level of trust with Google than a younger domain that is not as established. According to Matt Cutts, Google Engineer, this is true to a point, but domain age is an insignificant factor that really carries very little weight in the Google algorithm [CITATION rap17 \l 1033].

4.1.4 Social networks

The correlations of social signals with rankings have remained practically unchanged at a high level. The following still applies: Top ranking URLs have more social signals – this factor increases exponentially in the top places.

#27. Facebook activity

Google Visibility and Facebook Likes

When someone click the Like button on a web page it does not show up on his profile. Google has no access to see what a particular person has Liked. This means

that Google cannot tell when a respected authority has endorsed something via a Like or not. So, while Google can load the total number of Likes a page has, it cannot evaluate the quality of those Likes, making the information useless. In fact, every single one of those Likes could have been purchased on Fiverr. Stone Temple (noted SEO social media source) ran two Facebook tests regarding likes and visibility. The test involves driving a large amount of likes to two different pages situated in three different web domains. They ran six tests. At least 50 of the added likes came from people they knew, while the rest were bought from Fiverr. The result of the test revealed something interesting: none of the pages were indexed or crawled by Google [CITATION Eng13 \l 1033].

Google and Shared Facebook Content

Social media sharing is considered relatively important to social media SEO. Sharing exposes content to other users who would have not seen that content, hereby bolstering the visibility of the original source. That is the reason why shares are also referred to, as social media signals. It does not, however, mean Facebook shares count toward bolstering Google search visibility. Stone Temple has also tested if shared content can affect in some way the Google search ranking. They asked about 50 of their followers to share different test pages. As with the Like test, there was no sign of the pages getting crawled or indexed. However, the number of signals (shares) was small, and the profiles, while real, and public, were not highly authoritative. This may have led to the pages not getting crawled or indexed.

Indexing Facebook Posts

Google have already index as much as 1.87 billion posts however this doesn't seem to cover the entire number of Facebook posts! What is more, not all profiles are set to public. Stone Temple decided to check 85 popular Facebook profiles to learn just

how many Facebook posts are indexed. All these profiles have large numbers of likes and strong PageRanks, making them pretty authoritative to the search engine. They looked at the following aspects: their last 10 posts and 10 posts that were 3 months or older. They also observed 10 posts that were 6 months or older and 10 posts that were 12 months or older. They also tracked the actual content of posts to see whether it impacted how Google indexed them. The results ultimately covered 3246 posts. According to their results, Google doesn't index all posts and shares. Less than 60 percent of posts are actually indexed. The results also shown that at least 85 percent of posts with links are indexed and posts are not more likely to be indexed if they are new.

All in all, Google does not appear to consider Facebook as an important indexing, ranking or discovery factor. For sure, Google does not use the Like data or the Facebook shared links. On the other hand, Google index 59% of shares on prominent profiles and interestingly enough, they do index 85% of the posts that contain links.

#28. Google+ activity

Google plus tries to integrate the social signals they can extract from Google plus into their search engine rankings. Intentional or not, the engineers who made Google+ built it for SEO. The impact of Google+ on the search engines is clear to be seen, and still evolving, which is characteristic of the SEO industry as a whole. Consider the factors that make sharing content on Google+ far different than sharing on other social networks:

Posts are crawled and indexed almost immediately. One of the original goals for Google+ was using it to power real-time search after Twitter cut off Google's access to its data in 2011. Since then, Google has been using Google+ to discover new content, and many web professionals have discovered that URLs shared on Google+

are crawled and indexed very quickly. Comparing this to Facebook, in which because of privacy settings and restrictions on data sharing, it is not uncommon for posts to never be crawled or indexed by Google at all. Unlike Facebook, which hides data from Google, or Twitter, which directs Google not to follow most of its links, Google+ data is immediately and fully accessible to the company that built it.

Google+ posts pass link equity. Pages and posts on Google+ not only accumulate PageRank, but also, as links to posts are followed, they pass link equity on as well. When you share a link on Google+, the anchor text becomes the title of the page you are sharing. Some important things to remember about followed links within Google+ are:

- Only "shared" links (the links that show up beneath a post) are followed. Any external links within the post body itself are not followed, so these do not pass any link equity.
- For obvious reasons, uploaded images do not pass external link equity. Some people like to upload a screenshot of a page and then link to it in the body of the post. While a good image may increase post popularity and click-through rate, these posts do not pass link equity.
- Certain links in your Google+ "About" page are also followed and pass link equity.

Google+ is optimized for semantic relevance. Unlike Facebook or Twitter, each post you make in Google+ has most of the characteristics of a full-blown blog posting.

- Each post has its own URL.
- The first 45-50 characters of the post appear in the title tag.
- Just like a blog post, entries can be long and complex in order to explore a subject deeply. Various correlation studies have shown a strong relationship between longer pages and higher rankings.
- If a post is reshared, it can accumulate internal links from the Google+ platform, all with relevant anchor text.

Because of these factors, each post has the potential to send strong semantic signals to Google's search algorithm. This not only helps the post itself to rank in Google's search results, but potentially sends relevancy signals to a URL shared by the post.

4.1.5 Backlinks

For many years, links formed the absolute basis for search engine rankings, for SEO's, and for the analysis of ranking factors. This was also the reason for the highly tactical manipulations in this sector over a long period. These times have largely passed. Most of webmasters are also convinced that links will continue to lose relevance in the age of semantic contexts and machine learning with a user focus. For search engines, what it really matters is the best and most relevant content

#29. Root Citation – Citation Flow

A metric designed to predict how influential a link in a site might be, by considering the links pointing to it. *It does not bother about the quality of links*. If there are more domains pointing to a blog post, then the more influential it is. The perfect example is porn sites. They have a very high CF but these links are often not qualitative. A website with a lot of links pointing to it will be influential and will get a good citation flow. It is good to know that if TF increase, then CF should also increase. However, if CF increase, there is no proof that TF will increase too. So, if a website with high CF flow points to your website, your CF should get a boost. But if websites with high CF but low TF link to your website, it will negatively impact your TF.

#30. Root Trust – Trust flow

A metric designed to decide how trustworthy the link is. Based on the QUALITY of backlinks pointing to the site. If there are authoritative, trustworthy backlinks to a site, then the greater is the trust flow. A good trust flow is harder to get than a citation flow and so the CF will always get higher than the TF and not the contrary because there are a lot of links but not all of them are relevant. Moreover, even if

you are developing a qualitative linking strategy, there are always directories or other sorts of backlinks that are not very qualitative pointing to your website.

If a site has high trust flow, then it means that it has a high-quality backlink profile. High-quality backlinks also boost Google rankings**.** High trust flow is the clear sign of having high-quality content to Google and other search engines. If there are more backlinks to a site that carry the trust flow, then that means that the site is getting decent organic traffic.

#31. External Backlinks

Backlinks are links that are directed towards a specific website. The number of backlinks is an indication of the popularity or importance of that website. Backlinks are important for SEO because some search engines, especially Google, will give more credit to websites that have a respectable number of quality backlinks, and consider those websites more relevant than others in their results pages for a search query. When search engines calculate the relevance of a site to a keyword, they consider the number of quality inbound links to that site. A search engine considers the content of the sites to determine the quality of a link. When the external links come from sites with content related to your site, these links are considered more relevant to your site. If external links are found on sites with unrelated content, they are considered less relevant. The higher the relevance of external links, the greater their quality. Search engines look for natural links which are built slowly over time. While it is fairly easy to manipulate links on a web page, trying to achieve a higher ranking, it is a lot harder to influence a search engine with external backlinks from other websites. This is also a reason why backlinks factor is so highly into a search engine's algorithm. Search engine's criteria for quality external links has gotten even tougher during the last years, as many webmasters tried to achieve higher rankings by using deceptive techniques, such as hidden links or automatically generated pages which provide inbound links to websites. These pages are called link farms and not only they are disregarded by search engines, but also google may penalize your site because it understands that the link is manipulative.

#32. Referring Domains

Different types of back links are weighted differently in SEO. Some of the most valuable types of links for link building are educational and government sites. These are highly valuable they tend to have higher click through rates for traffic to your site. The logic being that users are more likely to follow a link if it is recommended by a highly trusted site as it provides respected information. Although there is no proof that Google treats educational and government links any differently, they are much more trusted so will have higher Flow Metric scores suggesting they are more influential. The domain information on the summary page of Site Explorer gives you a quick summary of how many back links your site has from government and educational sites as well as how many domains link to your site.

Not only can the source of the backlink be important on how influential it is but also the type of link that it is. When a link is deleted from a site and re-crawled by Google that link will not count anymore. It is important to know how many of each type of link that your site has as search engines give different weights to different types of links. Therefore, if the majority of your links are redirects they will be less influential on your site rankings.

#33. URL Citation

This is similar with #29. Root Citation – Citation Flow with the only difference that counts only for a certain URL

#34. URL Trust

This is similar with #30. Root Trust – Trust flow with the only difference that counts only for a certain URL

#35. URL External Backlinks

This is similar with #31. External Backlinks with the only difference that counts only for a certain URL

#36. URL Referring Domains

This value is similar with #32. Referring Domains. The only difference is that counts only unique domains and that count only for a certain URL.

4.1.6 Mobile

The days of building websites targeted solely at desktop or laptop environments are over. Users can access websites from a variety of internet-enabled devices. What is more, the popularity of smartphones and cheaper data packages from network providers have driven a sharp rise in mobile web usage. Mobile devices also offer some fantastic features typically not available on the desktop – functionality such as clicking a hyperlink in your website to call your phone number, or adding your contact details to an address book and location-aware content for your visitors. With the growing importance of mobile devices and the diversity of access it brings, it's vital for website owners, designers and developers to think smarter and broader about how to enable visitors to engage with their sites.

#37. Mobile Pagespeed

This value is similar with #21. Desktop Pagespeed. The main difference is that for this ranking factor it counts the performance of a website only for mobile devices, while in the Desktop Pagespeed factor it was counting the performance of desktop computers.

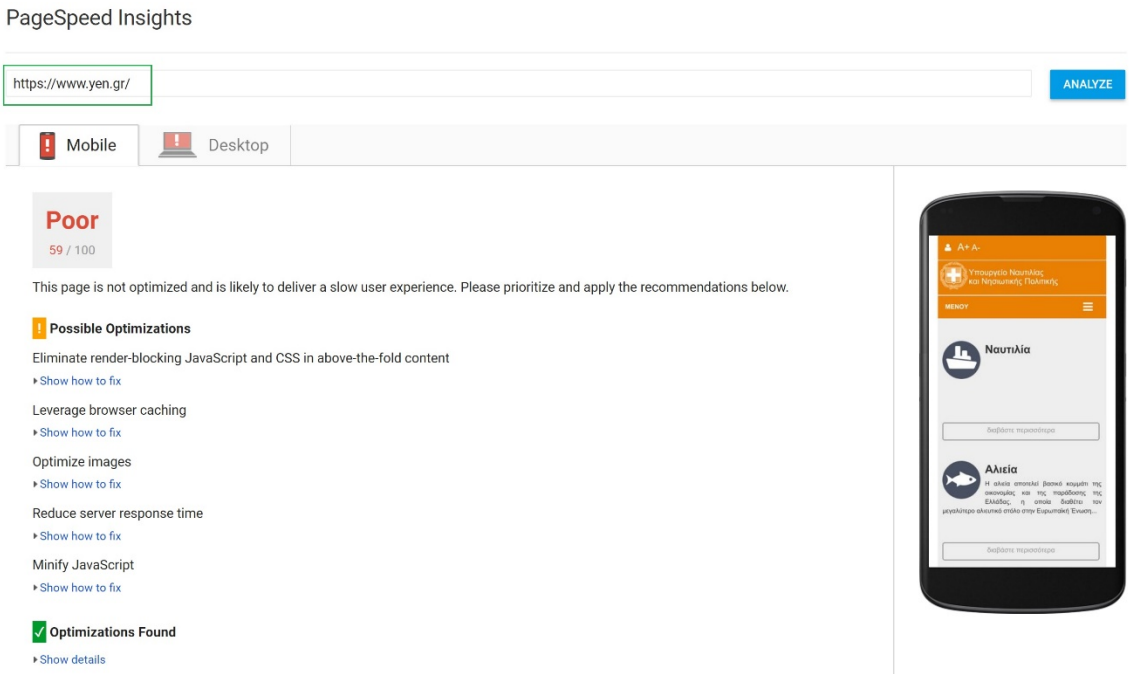


Figure 4.1.6.1: Checking the performance of www.yen.gr for the mobile version according to Pagespeed insights.

Gary Illyes from Google have stated that Google will be updating the page speed ranking factor to specifically look at the page speed of mobile pages when it comes to the mobile-friendly algorithm. Until today, many of the ranking signals Google uses for mobile rankings are based on desktop web pages, not mobile web pages. As a result, if a website has a really fast desktop web page, but the mobile version is really slow, it currently doesn't hurt the mobile rankings. When Google updates their mobile-friendly algorithm, they are planning to add mobile-specific page speed, as a factor and not rely on the desktop version [CITATION Sle16 \l 1033].

#38. Mobile UX

The Mobile App UX Principles report is a conversion optimization framework tailored to "Smartphone" mobile apps. It defines the key considerations when assessing apps, in order to identify how to improve the user experience (UX), optimize conversion and measure performance. It covers the whole customer journey,

conversion points at each stage and usability. UX is evaluated from 4 different categories: Adopt, use, transact and return [CITATION Gri15 \l 1033].

- **Adopt.** The focus at this stage is to remove all roadblocks to usage of your mobile application. Applications have to present their content as quickly as possible. First impressions count and a splash screen gives to users a quick image of the application. Users should never wait so it has to be fast. Tips / help should be used in the context of what user is doing. The home screen needs to provide the user functionality to complete their priority tasks and only primary navigation and content should be visible by default, with secondary content hidden - but available via tap or swipe - off-screen. One of the main points of difference between mobile apps and mobile sites, is that apps enable a user to be persistently logged-in, to benefit from a high level of convenience and personalization. So, applications should request sing-up but only the at the right moment, when it absolutely need it [CITATION Gri15 \l 1033].
- **Use.** At this stage, it is needed simple conversion decisions. Enable users to quickly search for what they want and consider products and services. An excellent search facility will help users find what they want quickly and easily, in order to satisfy their needs and drive conversion. Product screens are also important as it is where users make key conversion decisions, e.g. add to basket, add to wish list, locate store, call now, etc. Strong product screens enable users to quickly transact, save for later, and share the items they have to make decisions on[CITATION Gri15 \l 1033].
- **Transact.** The focus here is to ensure first time users progress through each checkout stage with minimal effort, with reassuring messages at each stage, and convert without hesitation. At basket stage: direct users towards checkout or continue shopping, confirm items and costs, and enable users to edit items. At personal and address details stage: reassure users to progress to payment quickly and provide a single-screen checkout for ultimate convenience. At payment stage: employ convenient data capture methods such as scan cards, express payment etc. [CITATION Gri15 \l 1033]

- **Return.** The focus at this stage is to give users reasons to return in order to retain customers and encourage member loyalty. Apps are the ideal touchpoint for customers and members to self-serve and manage accounts and transactions anytime and anywhere. Give customers control and the level of convenience that only apps afford. Send notifications to users alerting them about highly relevant, timely and personal events, content, or messages [CITATION Gri15 \l 1033].

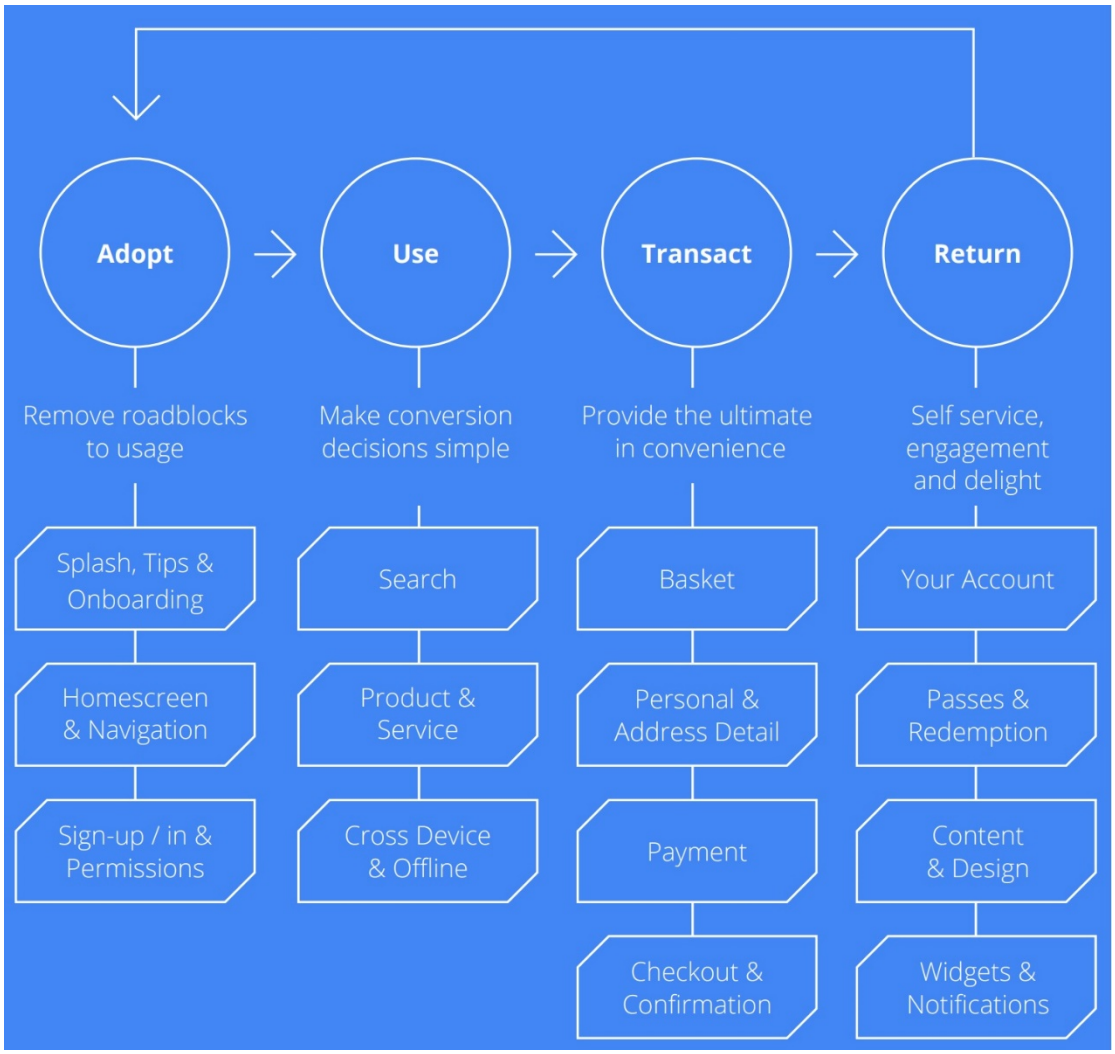


Figure 4.1.6.2: Mobile usability Hygiene (source: Mobile App UX principles)

#39. Mobile Compatibility

Websites need to have a clear policy and strategy in place for being accessible to as many internet-enabled devices as possible. The less website contains special types of

web content, such as Flash, Silverlight or Java, the more mobile devices can access the content.

- *Mobile compatible* site**.** A mobile compatible website is an HTML based website that does not contain Flash. It means that a website is viewable on a Smartphone or Tablet, but it is not optimized for these devices. These websites have been designed for web browsers on PCs with much larger screens and navigation using a mouse. By entering in a mobile compatible website by a smartphone or tablet device, everyone could easily notice a few things, such as that the user has to scroll left, right, up or down and figure out how the site "works", media may load slower, text is tiny so visitors have to zoom in to read it and the links are small and difficult to click with the fingertip. In a website like the one described above it is likely that visitors will abandon it very quickly. Poor experiences do not convert customers.
- *Mobile optimized* site**.** A mobile optimized website is a website that is designed specifically for a Smartphone or Tablet, not a desktop computer or laptop. A mobile optimized website does not require zoom to read text. On a mobile optimized site, the navigation is built for efficiency, the images and media are optimized for quick loading and the content is succinct for maximum effectiveness. Additionally, mobile-only functionality includes tap-to-call, tap-to-email and map functionality.

A hotel's optimized website for mobile devices not only is much easier to navigate and to be used compared to the experience of a compatible website also it will have less bounce rates, higher engagement levels and improved conversions.

The image 4.1.6.3 presents the two forms of the same website in order to exactly understand the main differences between the two cases and to illustrate the importance of having an optimized website for mobile devices.

Figure 4.1.6.3: How the same website is represented in the optimized and in the compatible form

Starting at 21 April 2015, Google Search has expanded its use of mobile compatibility as a ranking signal. This change has affected mobile searches in all languages worldwide and has a significant impact on Google Search results. Users find it easier to get relevant, high quality search results optimized for their devices.

#40. Font size legibility

One factor that website designers often overlook when creating websites is the readability of their content. This is probably due to the fact that content readability is not accounted in Google's search ranking algorithm, at least not that anyone knows of or could confirm. But factors like content readability is one of the many indirect ranking factors that actually matters and has a big impact on SEO.

Content readability is the level of ease to understand a written text. The scientific way of measuring readability is based on factors such as:

- Speed of perception
- Perceptibility at a distance
- Perceptibility in peripheral vision
- Visibility
- Reflex blink technique
- Rate of work (e.g., reading speed)
- Eye movements
- Fatigue in reading

Basically, if a written text is easy to read and understand, then it has good readability. There are also plenty of algorithms that are designed to measure content readability and that is how robots or search engines measure them. There are many reasons why the content should be readable, even though search engines do not take the score of content's readability into their search ranking algorithm.

This is because many of search engine's ranking algorithm is based on human behavior on a page. By improving the website's content readability, automatically improve their behavior on page. Their page on time, exit rate, bounce rate and social signals will all improve and through all these, search engines understand that visitors like the webpage content. It is a fact that many users are closing a page after reading the first few sentences because the text is just too hard to read. People read stuff online to get the information they want in the fastest and simplest way possible. So, designers have to make all these information available to them as easily as possible. An article with a good content readability will make the readers stay longer on page because they will probably read the whole text. Consequently, this will improve the bounce and exit rates which is perfect for SEO.

Focusing especially to the font, a website needs to have a font that is easily to read. Novelty fonts are suitable for headers and sub headers as they are short and it distinguish them from the body text. However, the body text should be in standard sans-serif fonts because they are easier to read on the screen. Arial and Helvetica are

always a good choice for body text. Serif fonts are more suitable for printed mediums such as newspaper.

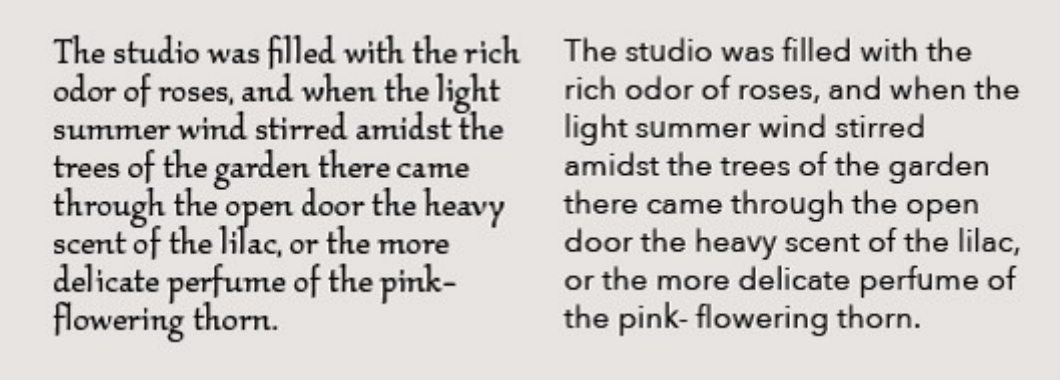


Figure 4.1.6.4**:** The text on the right side is using sans-serif fonts. It is obvious that sans-serif fonts are a lot easier to be read on screen.

The font size should not be too big neither too small. Fonts that are too small will be hard on the eyes while fonts that are too big will take up too much space. The size of text is controlled by the media queries and is declared for different sized screens. Different text size is chosen for mobile users and different text size for desktops. With this technic, every font can be readable on any device regardless of its size.

#41. Mobile Friendly

Mobile is changing the world. Today, everyone has smartphones with them, constantly communicating and looking for information. In many countries, the number of smartphones has surpassed the number of personal computers and so, having a mobile-friendly website has become a critical part of having an online presence. Search Console's Mobile-Friendly Test Tool is a quick, easy way to test whether a page on your site is mobile-friendly.

The Mobile Friendly test tool is very easy to use, the user has just to type in the full URL of the web page that he wants to test (https://search.google.com/test/mobile-friendly). Any redirects implemented by the page will be followed by the test. The test typically takes less than a minute to run. Test results include a screenshot of how the page looks to Google on a mobile device, as well as a list of any mobile

usability problems that it finds. Mobile usability problems are issues that can affect a user that visits the page on a mobile (small screen) device, including small font sizes (which are hard to read on a small screen) and use of Flash (which is not supported by most mobile devices). If for some reason the tool is not able to access the page, it will display an error describing the problem. Access problems include network connectivity issues or the site being down. External resources such as image, CSS, or script files included by a page may be blocked to Googlebot by a robots.txt file. In this case, the test results page will show a "blocked resources" warning. Depending on the blocked resource, this could have a large effect on the page. For example, a blocked large image could make a page appear to be mobile-friendly when it is not, or a blocked CSS file could result in incorrect font styles being applied (for example, too small for a device). This affects both the mobile usability score and Google's ability to crawl your page. For this reason, it is important resources not to be blocked to Googlebot by robots.txt [CITATION Goo17 \l 1033].

The Mobile-Friendly Test tool can identify the following usability errors:

- Flash usage. Most mobile browsers do not render flash-based content. Therefore, mobile visitors will not be able to use a page that relies on flash in order to display content, animations, or navigation. It is recommended designing the page animations using modern web technologies.
- Viewport not configured. Because visitors to your site use a variety of devices with varying screen sizes (from large desktop monitors, to tablets and small smartphones) your pages should specify a viewport using the meta viewport tag. This tag tells browsers how to adjust the page’s dimension and scaling to suit the device.
- Fixed-width viewport. This report shows those pages with a viewport set to a fixed width. Some web developers define the viewport to a fixed pixel size in order to adjust a non-responsive page to suit common mobile screen sizes. To fix this error, adopt a responsive design for your site’s pages, and set the viewport to match the device’s width and scale accordingly.
- Content not sized to viewport. This report indicates pages where horizontal scrolling is necessary to see words and images on the page. This happens

when pages use absolute values in CSS declarations, or use images designed to look best at a specific browser width (such as 980px). To fix this error, make sure the pages use relative width and position values for CSS elements, and make sure images can scale as well.

- Small font size. This report identifies pages where the font size for the page is too small to be legible and would require mobile visitors to "pinch to zoom" in order to read. After specifying a viewport for your web pages, set your font sizes to scale properly within the viewport.
- Touch elements too close. This report shows the URLs for sites where touch elements, such as buttons and navigational links, are so close to each other that a mobile user cannot easily tap a desired element with their finger without also tapping a neighboring element. To fix these errors, make sure to correctly size and space buttons and navigational links to be suitable for your mobile visitors.

4.1.7 Alternative sources

#42. Moz Rank

MozRank quantifies link popularity and is Moz's version of Google's classic PageRank algorithm. Pages earn MozRank based on the other pages on the web that link to them and the MozRank of those linking pages. The higher the MozRank of the linking pages, the higher the MozRank of the page receiving those links. In this way, it reflects a type of raw link equity for any given webpage on the Internet.

Similar to the way that Google's original PageRank is calculated, MozRank is based on a logarithmic scale between 0 and 10. Thus, it's much easier to improve from a MozRank of 3 to 4 than it is to improve from 8 to 9. The "average" web page actually has a MozRank that is only a tiny fraction (~0.05). Thus, if you have a MozRank of 1 or 2, you're way more "important" than the average. In other words, "average" on

the web can be a bit of a misnomer as the vast majority of pages you'll find ranking in Google or on webpages you visit in a normal surfing session are likely to be MozRank 2, 3, 4 or higher. But because the web is so vast, there are many unimportant, barely linked-to pages on billions of websites that drag down the overall average.

There's no specific "good" or "bad" MozRank score, but higher generally means more and more important links point to the page, subdomain, or root domain [CITATION Moz17 \l 1033].

#43. Woorank

WooRank is a digital SEO solution that provides automatic website reviews as well as personalized tips for boosting traffic, leads & sales.

It covers: Mobile rendering, SEO, Social Media data, backlinks, SERP ranking, usability & servers optimization. A crucial tool for Web agencies, Designers, Webmasters, SEO experts, and Digital Marketers, WooRank provides efficient services to over 40,000 companies. These companies rely on this tool to analyze their websites and receive easy-to-understand, actionable insights on which actions they need to optimize. WooRank also provides a dynamic grade on a 100-point scale that represents your overall Internet Marketing effectiveness at a given time [CITATION Woo17 \l 1033].

Chapter 5: Greek ferry routes websites case study

In this chapter, it is going to be presented the results of the case study, which was made for the Greek ferry routes websites. After reviewing and analyzing 43 different factors in the previous chapter, now it will be tested and checked if all these factors

really correlate with the google search results machine and how and in what degree they can affect a google search.

The subject of the survey which was made has to do with the Greek tourism and more specifically concerns ferry tickets and tourist websites that offer the opportunity to book and reserve a ticket on line.

5.1 Defining the specifications and the parameters of the survey

In order, the survey results to be as reliable as possible, we had to define and follow specific parameters so that the keywords on which the survey will be based meet specific requirements. The main parameters were about localization, survey subject and keyword research.

- Localization. The language of the keywords which was used, was decided to be the Greek one and it was also used the Greek version of Google search machine platform (google.gr). The reason for the usage of non-English language is to protect results from potential unidentified factors with global impact.
- Survey subject. The survey's subject is the Greek ferry routes and the tourist websites that provide ferry routes information, availability of tickets, prices of the tickets and on-line reservation of ferry tickets. The reason for choosing to study Greek ferry routes is due to the fact that Greece is made up of too many islands and have been officially recorder more than 6,000 islands. 200 of them have permanent residents and so Greek ferry companies operate on these islands for the whole year. Ferry routes websites have extreme traffic during summer period from vacation travelers ,while for the rest of the year the recorded traffic is in a medium level due to residents. It becomes clear enough that for social and economic factors, shipping companies should give importance to their websites, if they want to achieve better ranking in a google search inquiry.
- Keywords. After studying result from Google Trends website and Google Adwords database it was found that the related keyword with the most searches for ferry routes was a 2 words phrase. The first word was

«δρομολόγια» which means routes and the second word which was following was changing and had to do with different islands destinations. The survey was decided to be limited at Cycladic destinations and the islands which were included in the survey had to meet specific criteria. The first criterion has to do with the alternative ways of transportation to and from the island and for this reason all the islands that had an airport were excluded from the survey. The other main factor was about the permanent residents. So, all the islands without official recorded permanent residents were also excluded.

5.2 Islands meeting the requirements

After filtering all the parameters of the survey, we concluded at 16 different islands that fulfil the specific conditions. So the final keywords for our study are:

- Δρομολόγια Άνδρος (Andros Routes)
- Δρομολόγια Φολέγανδρος (Folegandros Routes)
- Δρομολόγια Ίος (Ios Routes)
- Δρομολόγια Κέα (Kea Routes)
- Δρομολόγια Κουφονήσι (Koufonisi Routes)
- Δρομολόγια Κύθνος (Kythnos Routes)
- Δρομολόγια Σέριφος (Serifos Routes)
- Δρομολόγια Σίφνος (Sifnos Routes)
- Δρομολόγια Τήνος (Tinos Routes)
- Δρομολόγια Αμοργός (Amorgos Routes)
- Δρομολόγια Ανάφη (Anafi Routes)
- Δρομολόγια Δονούσα (Donousa Routes)
- Δρομολόγια Ηρακλειά (Iraklia Routes)
- Δρομολόγια Κίμωλος (Kimolos Routes)
- Δρομολόγια Σχοινούσα (Shinousa Routes)
- Δρομολόγια Σίκινος (Sikinos Routes)

Figure 5.2.1: Map of Kyklades islands

5.3 Methodology of survey

Data collection: The selected keywords had been searched at google.gr one by one, from the same PC, with the same IP and at the same Day (June 12th, 2016). During the search, for every google inquiry made, it was recorder the 20 first results.

Factors selection: The factors which were tested are 43 (the 43 factors that were presented and explained in chapter 4), which can someone find easily and free through third party sources, freeware applications, or browser add-ons. The 43 factors were divided in following categories: Technical, User Experience, Content, Social Networks, Backlinks, Mobile and Alternative Sources.

5.4 Survey results

At this section, it is presented the results of the survey made. The graphs show the average values of all the results for the 16 searches made.

#1. Title tag

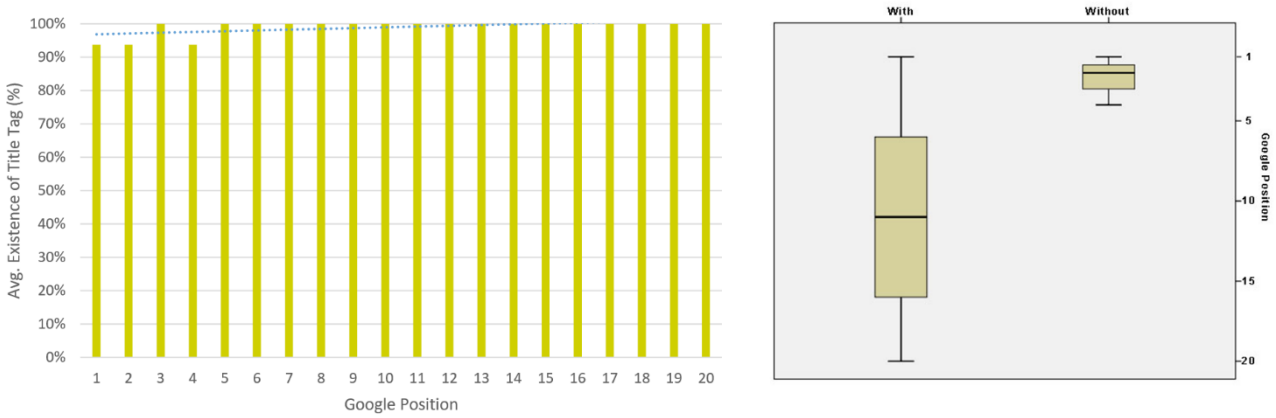


Figure 5.4.1: Average search results relative with title tag factor

According to the graph, no correlation exists about the google ranking of the 20 first results regarding with the factor title tag. There is no valuable statistical results due to the fact that almost all websites uses title tag.

#2. Meta Description

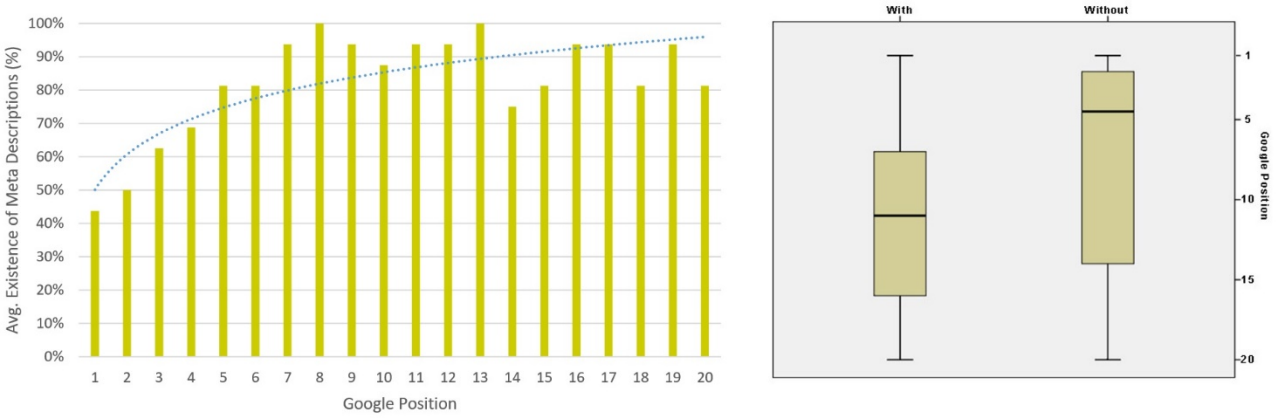


Figure 5.4.2: Average search results relative with meta description factor

Despite the fact that meta description is an important factor for the google ranking improvement, in the survey appears to have opposite effects for the websites using it. Almost 50% of the high ranked websites do not use any meta description. With this data, **the correlation value is 0.45**

#3. H1

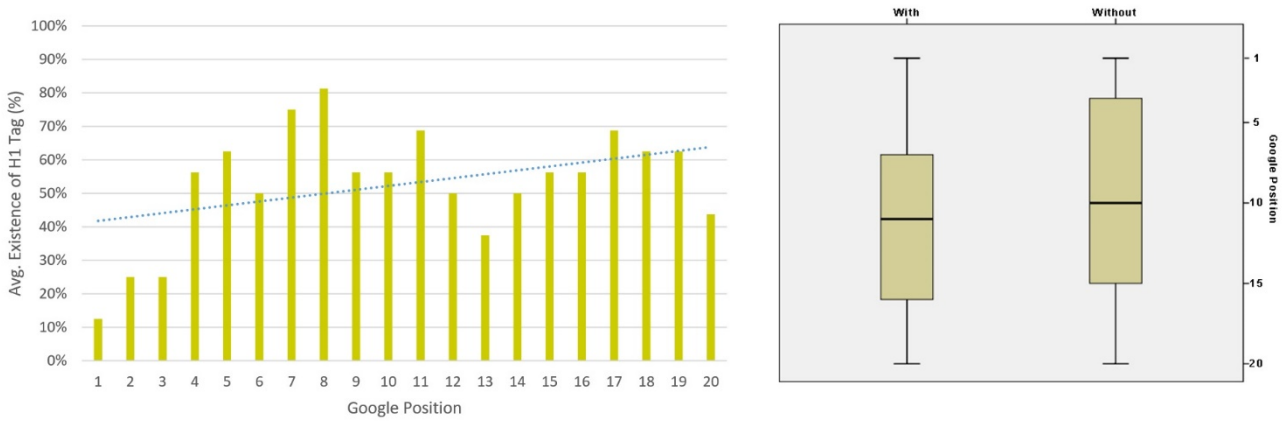


Figure 5.4.3: Average search results relative with H1 factor

According to the graph, the existence of H1 factor does not seems to affect the Google search rankings. Almost half of the average top 20 websites does not use H1 tags.

#4. H2

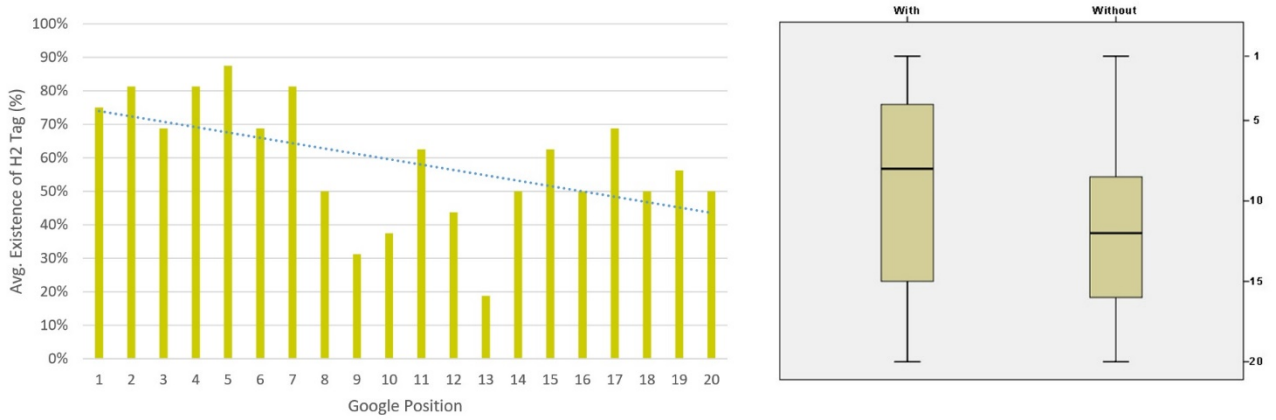


Figure 5.4.4: Average search results relative with H2 factor

According to the data of the survey, there is an important correlation between the h2 tag existence and Google ranking position. More specifically, websites using the factor occupy higher ranking positions than those who do not. **The correlation value here is -0.55.**

#5. Alt Attribute

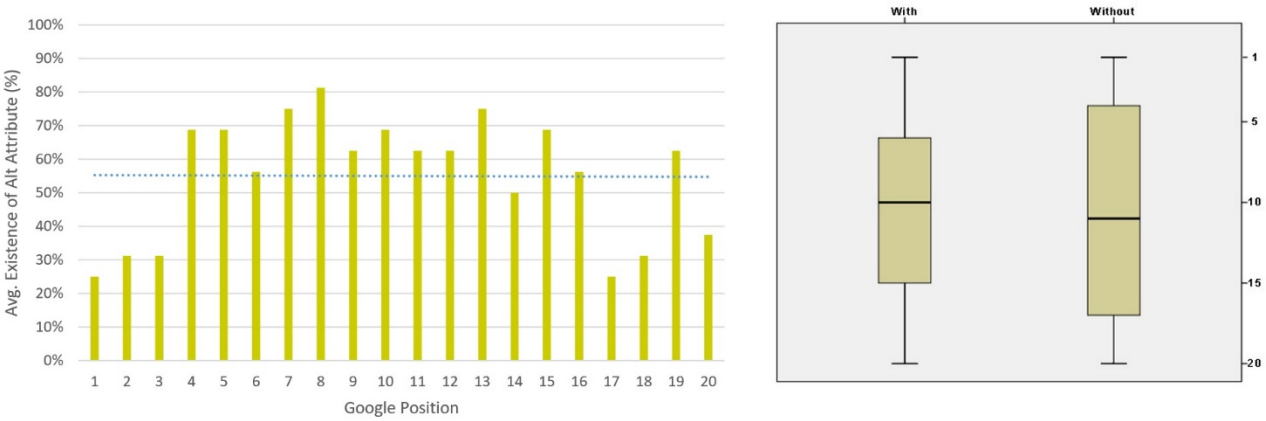


Figure 5.4.5: Average search results relative with alt attribute factor

The majority of the websites in the first three places do not use alt attribute. The factor is used in an average rate of 55%. So, there is not any significant result for the regular usage of Alt Attribute for the Greek Ferry Route Websites.

#6. Broken Links

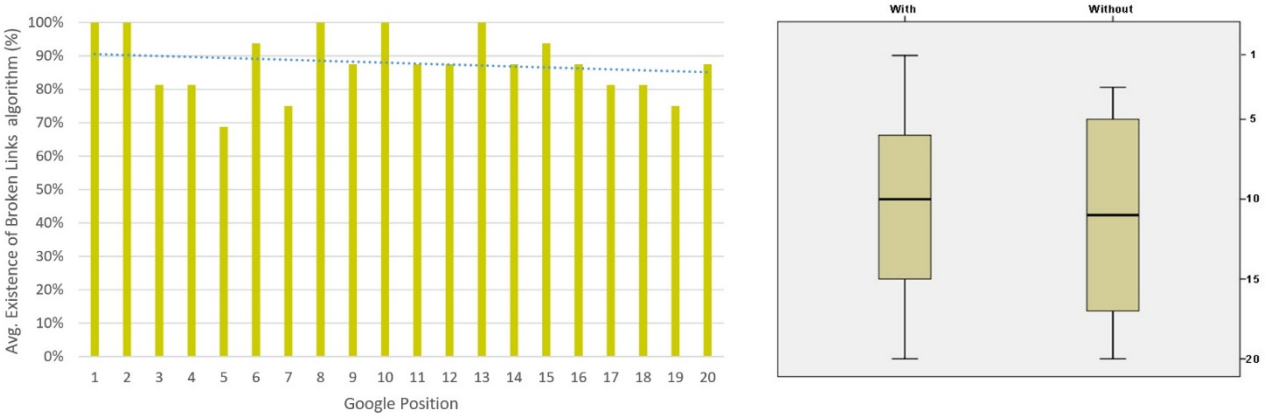


Figure 5.4.6: Average search results relative with broken links factor

All the websites which were appearing in the two first places have no broken links. As a general observation, the less broken links exists, the higher a website appears in google ranking position, even though there is no significant correlation.

#7. WWW Resolve

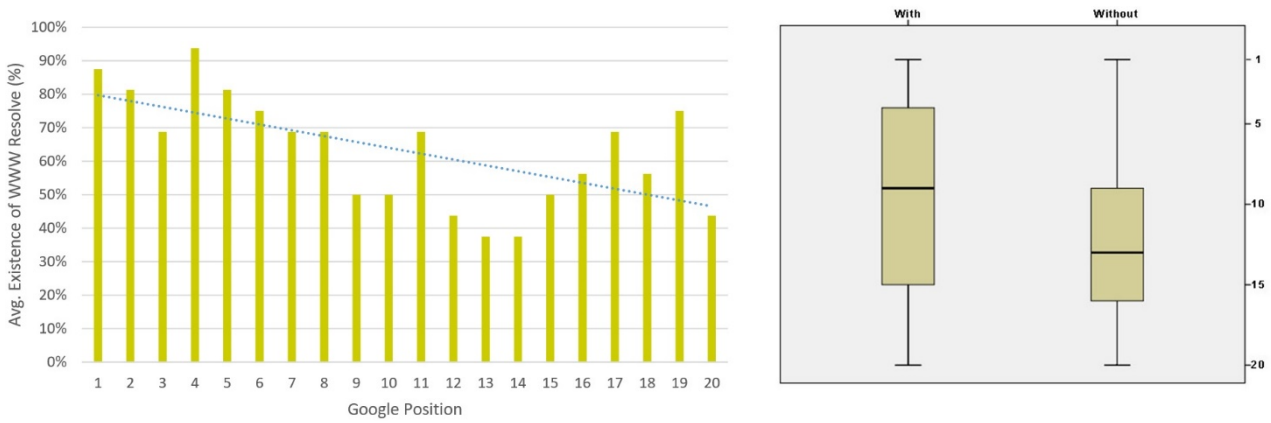


Figure 5.4.7: Average search results relative with WWW resolve factor

This graph shows how WWW resolve factor greatly affects Google rankings. The higher we move in ranking position, lesser the existence of www resolve. It is clear enough that duplicate content **has a high correlation value of -0,59.**

#8. IP Canonicalization

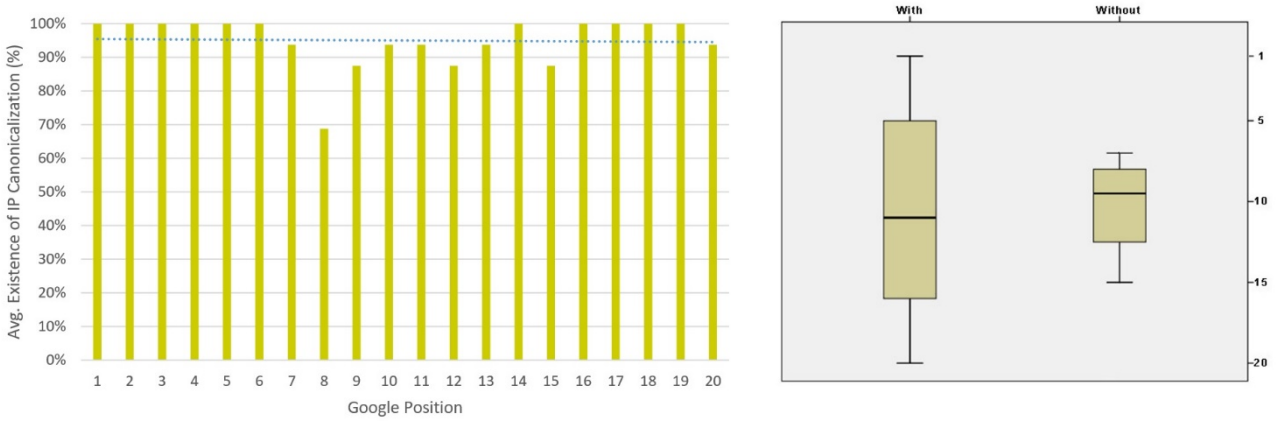


Figure 5.4.8: Average search results relative with IP canonicalization factor

The survey results show that most of the Greek Ferry Route Websites use IP Canonicalization algorithm. On the other hand, the websites that use it less are appearing between the positions 6 and 13. To conclude no significant correlation exist here.

#9. Robots.txt

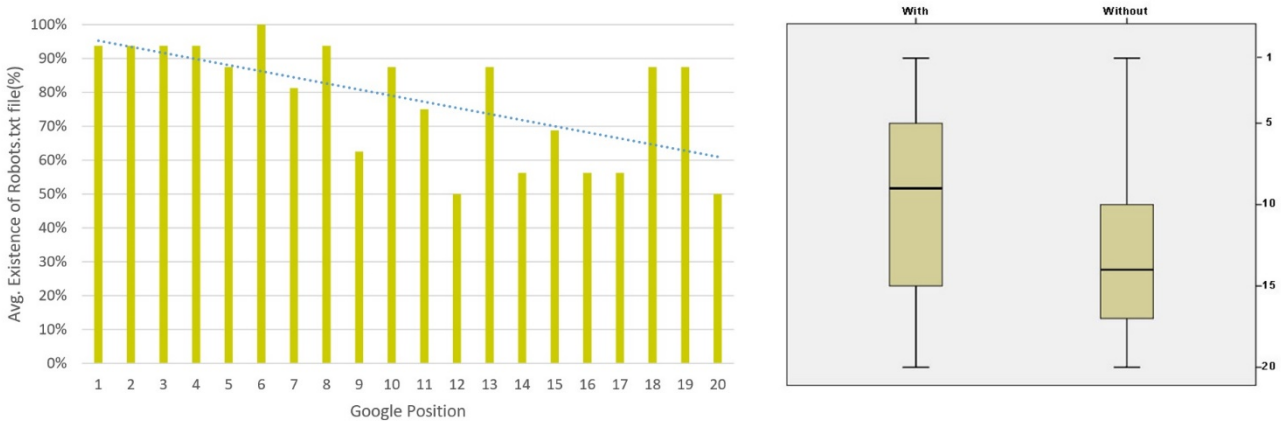


Figure 5.4.9: Average search results relative with robots.txt factor

According to the graph, robots.txt. factor has a great effect at the Google search rankings. Particularly, the higher we move in ranking position, the less existence of Robots.txt files become. Statistically **the correlation value of this factor is -0,52**

#10. XML Sitemap

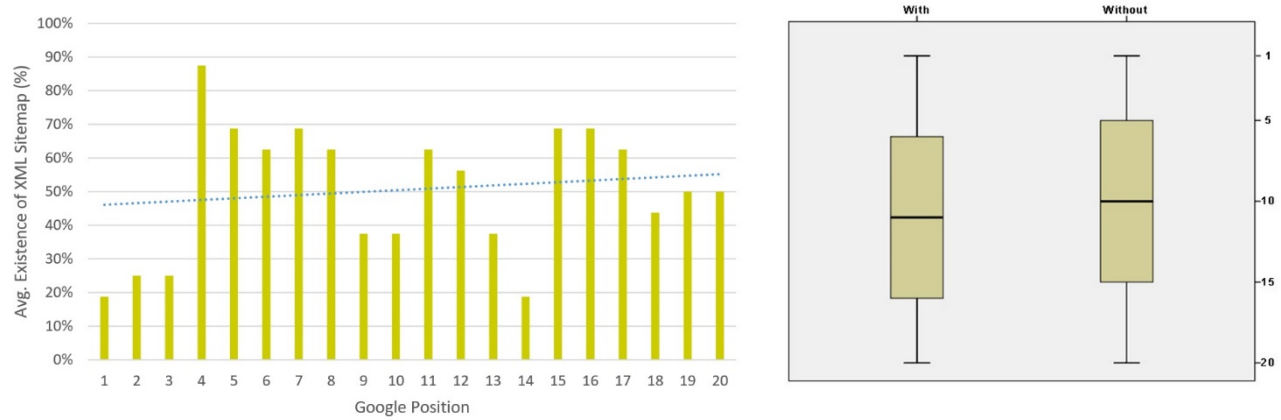


Figure 5.4.10: Average search results relative with XML sitemap factor

The results regarding the XML sitemap factor show that there is not a specific allocation in order to extract a safe conclusion. It seems that it does not affect the Google search rankings and it does not exist any statistical significance.

#11. URL Rewrite

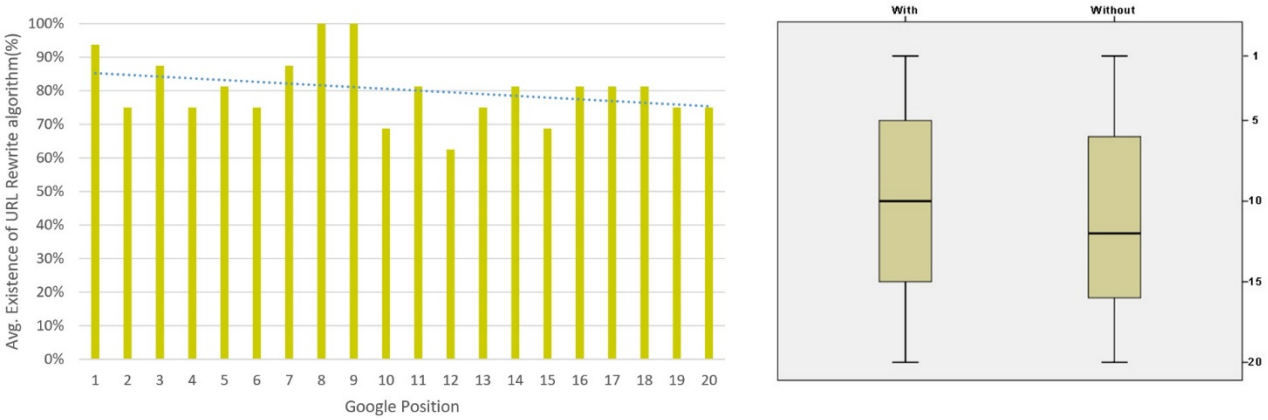


Figure 5.4.11: Average search results relative with URL rewrite factor

Analyzing the graphs data, it seems that there is a correlation with this factor and the Google search rankings. Websites from the tenth place and below do not give necessary importance at URL rewrite factor, while in the first ten places websites seems to take care of it more. In general, from this factor appears to be some correlation but without statistical significance.

#12. Underscores in the URLs

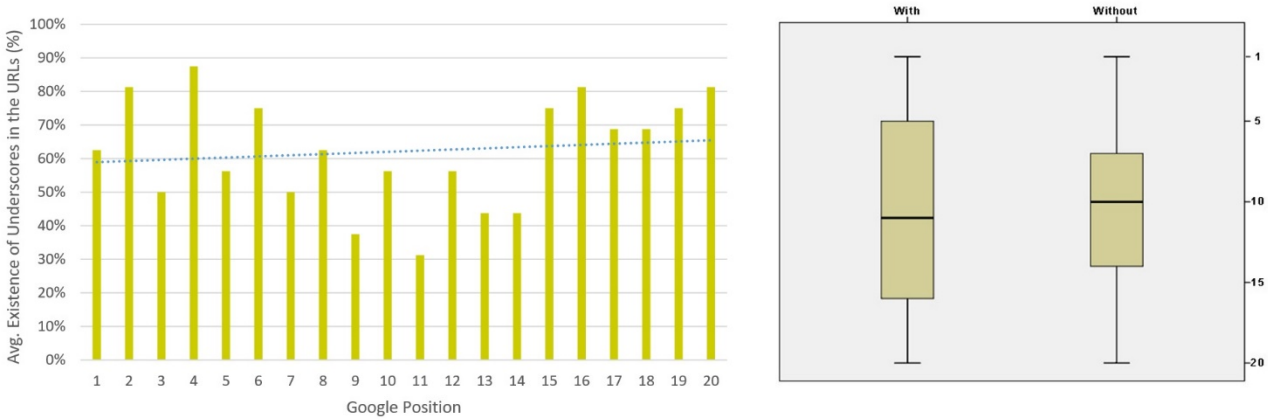


Figure 5.4.12: Average search results relative with underscores in the URLs factor

The results regarding the underscores in the URLs factor show that there is not a specific allocation in order to extract a safe conclusion. It seems that it does not affect the Google search rankings in our case study and it does not exist any statistical significance.

#13. Blocking Factors

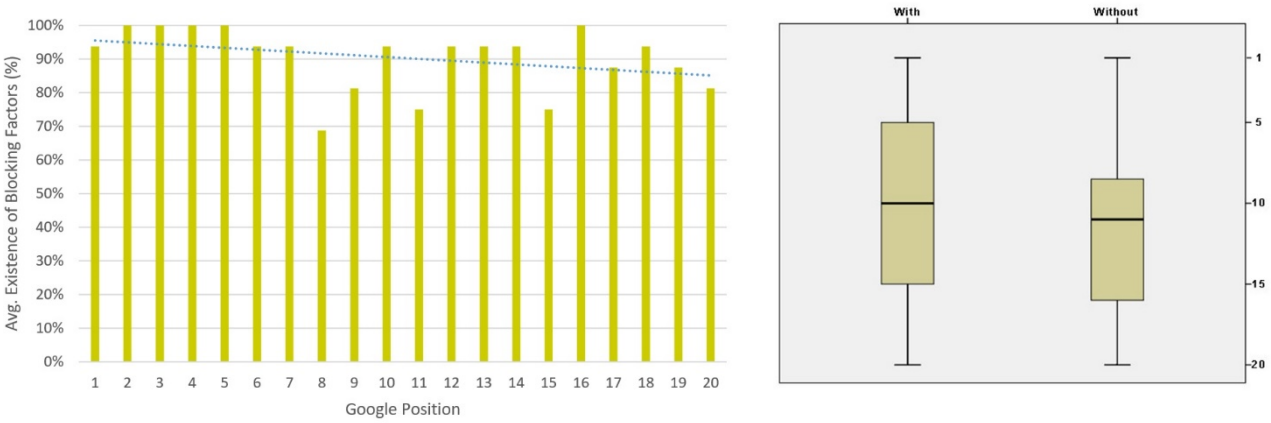


Figure 5.4.13: Average search results relative with blocking factor

Analyzing the graphs data, it seems that there is a correlation with this factor and the Google search rankings. The websites from 8th position till 20th are paying less attention to blocking factors, with the exception of 16th position. Even though this correlation is without statistical significance.

#14. Custom 404 Page

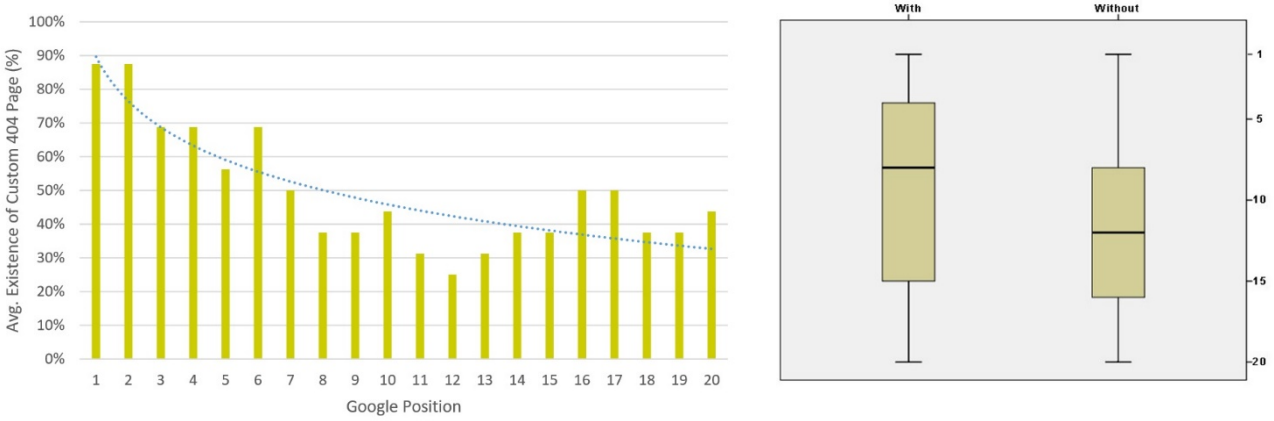


Figure 5.4.14: Average search results relative with custom 404 page factor

According to the graph, custom 404-page factor has a great effect at the Google search rankings. Particularly, the higher we move in ranking position, the more the websites use custom webpages in order to reduce customer frustration and link them to other pages on the same site. Statistically **the correlation value of this factor is -0,64.**

#15. Language

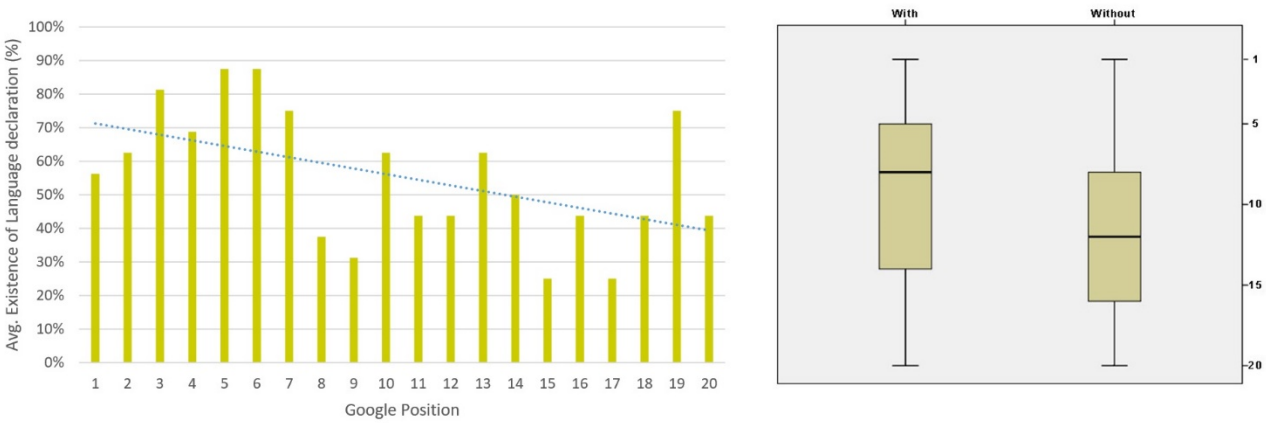


Figure 5.4.15: Average search results relative with language factor

According to the survey more than the half of the Greek ferry root websites just let Google search engine to detect their language. However, the rest of the websites which declare their language in HTML seems to gain highest Google ranking positions. Statistically **the correlation value of this factor is -0,52.**

#16. Structured Data Markup

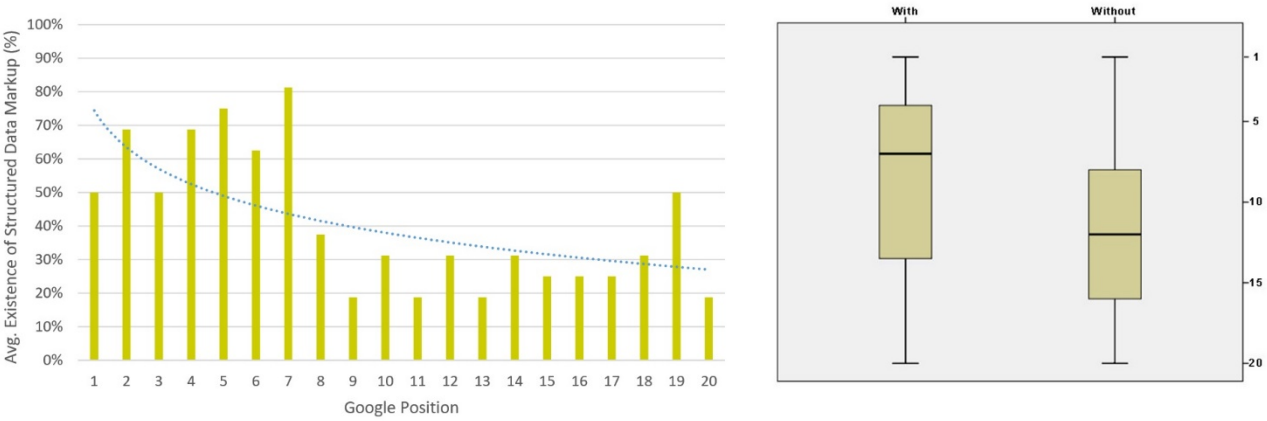


Figure 5.4.16: Average search results relative with structured data markup factor

Analyzing the graphs data, it is easily noticed that the majority of the websites do not contain structured data markup in most of cases. On the other hand, It seems that there is a strong correlation with the Google rankings as the few sites using it are be placed in the first top positions. Statistically **the correlation value of this factor is -0,62.**

#17. SSL Secure

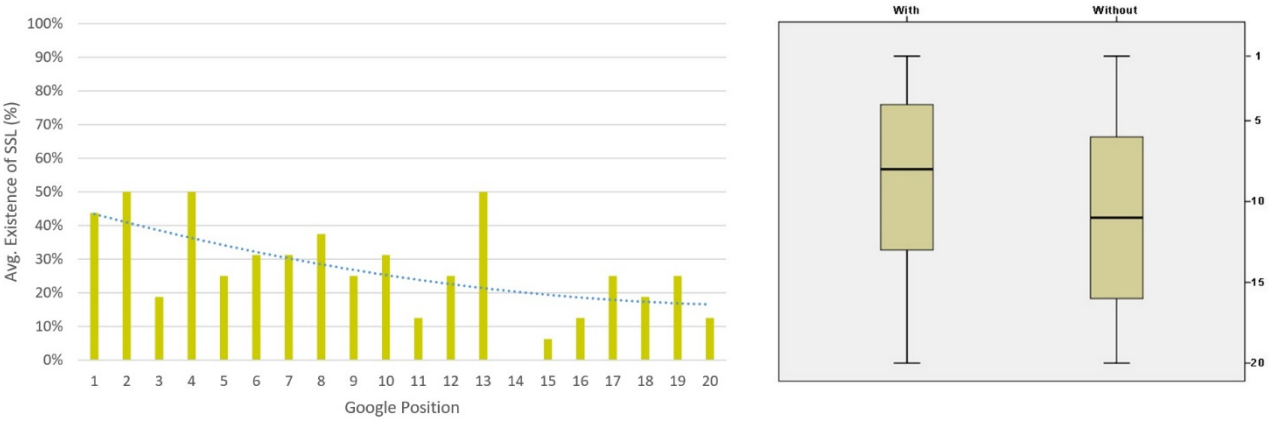


Figure 5.4.17: Average search results relative with SSL secure factor

The survey results show that most of the Greek Ferry Route Websites do not give the necessary importance to SSL secure factor. However, it seems that there is a strong correlation between SSL Certificate existence and Google ranking position as safety is one of the most important factors. Statistically **the correlation value of this factor is -0,55.**

#18. Page Size (Mb)

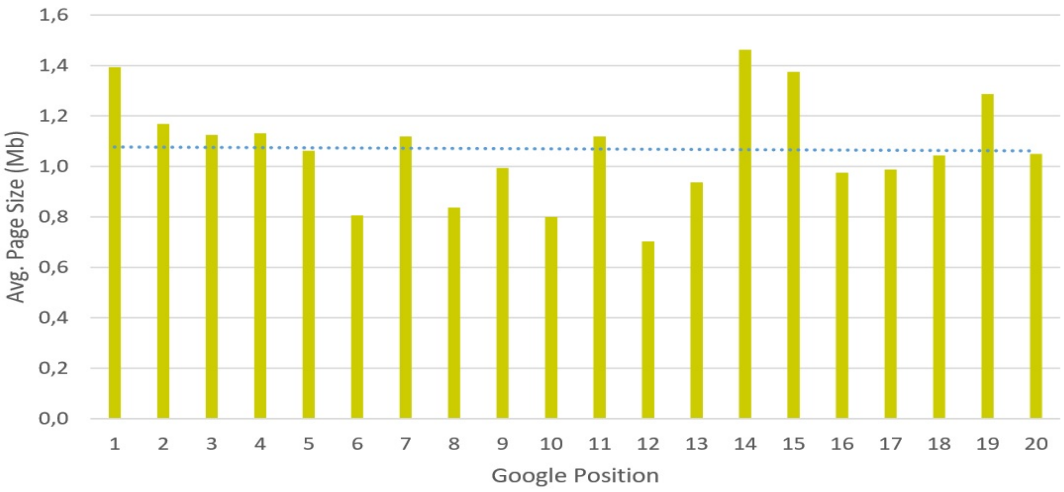


Figure 5.4.18: Average search results relative with Page size factor

The survey results show that most of the Greek Ferry Route Websites have the same page size with small deviations between them. As a result, no safe conclusion can

be extracted that this factor really affects the Google search rankings in our case study and does not exist any statistical significance.

#19. Google Analytics

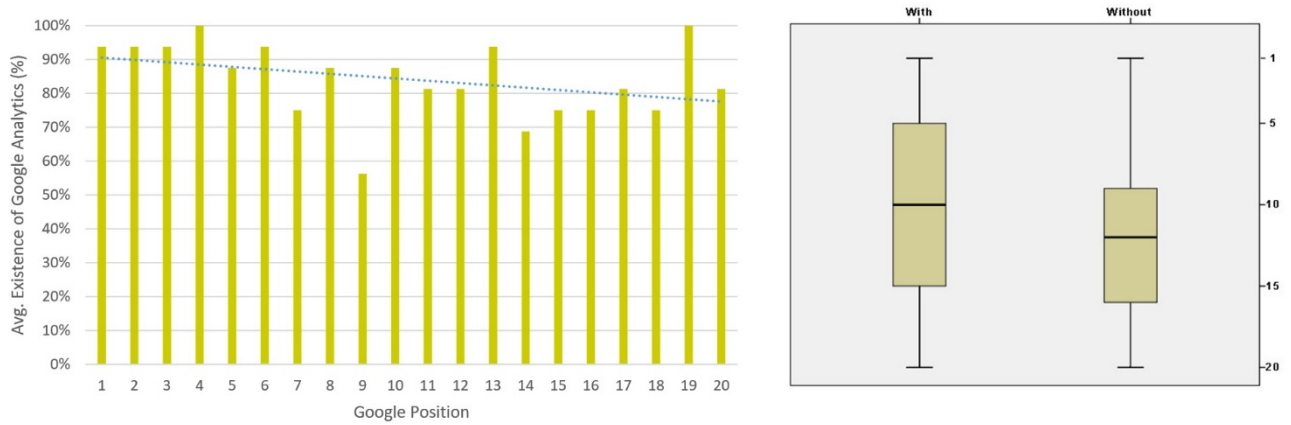


Figure 5.4.19: Average search results relative with Google analytics factor

According to the survey results most of the websites have installed Google's analytics. As a factor, seems to have a light impact in the Google rankings but from a statistical point of view it does not have any significant correlation.

#20. Directory Browsing

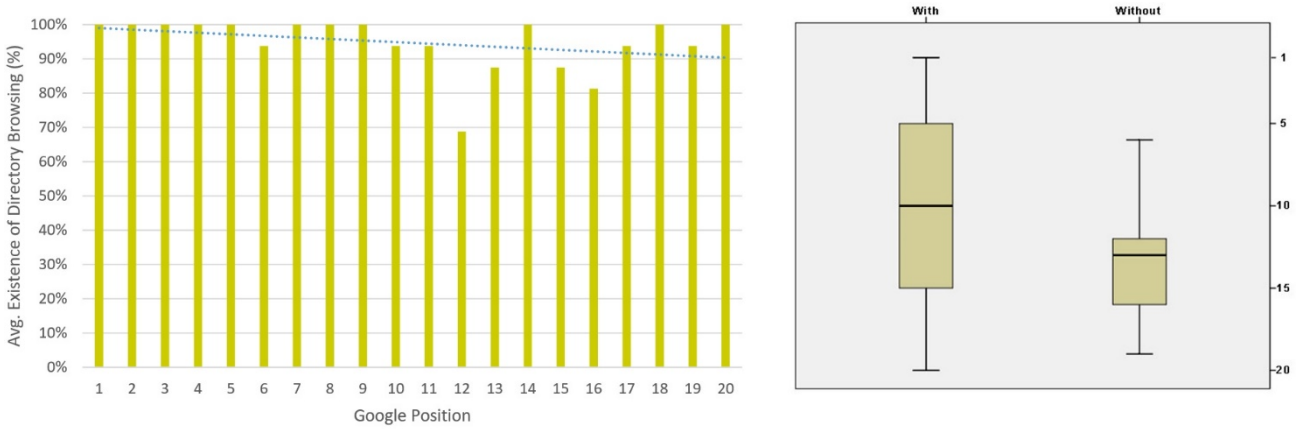


Figure 5.4.20: Average search results relative with directory browsing factor

According to the survey results most of the Greek ferry route websites use directories. As a factor, seems to have a light impact in the Google rankings but from a statistical point of view it does not have any significant correlation.

#21. Desktop Pagespeed

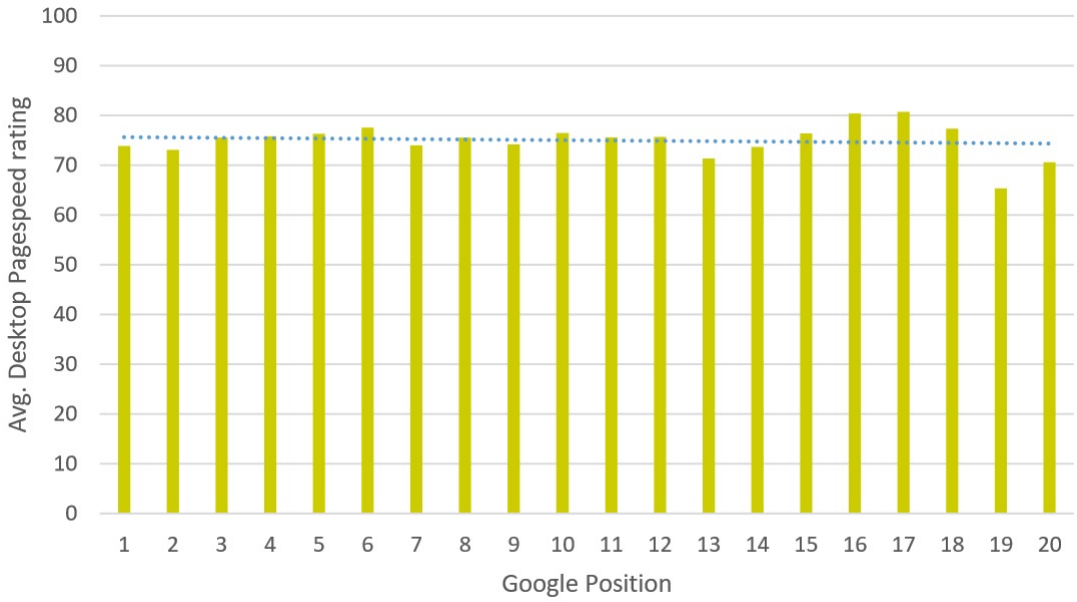


Figure 5.4.21: Average search results relative with desktop pagespeed factor

According to the survey results most of the Greek ferry route websites (with an average percentage of 75%) consult PageSpeed Insights. For this reason, no safe conclusion can be extracted, that this factor really has an impact at the Google search rankings and in our case study does not exist any statistical significant result.

#22. Load Time (sec)

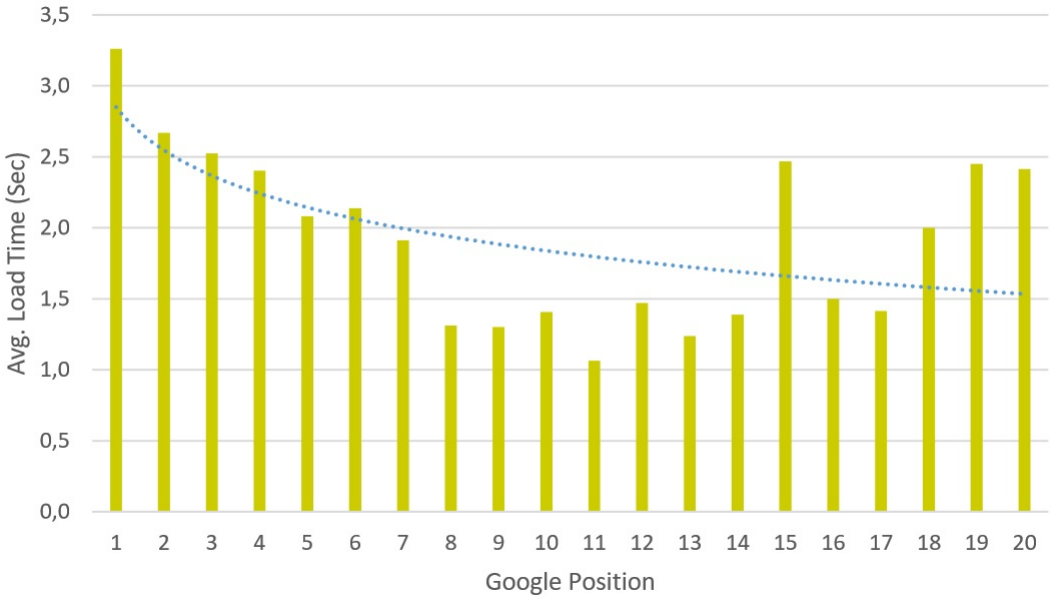


Figure 5.4.22: Average search results relative with load time (sec) factor

Analyzing the graph partially (until the 14 position), it is noticed that the load time factor is heavily affecting the Google search rankings. The difference between the half first places (until position 7) and the last half places (position 8 to 14) is enormous and clearly the load time of a website has a great impact in the search results. However, three particular positions (15, 19 and 20) is opposed to the rest of the research, and so does not exist a statistical significant result.

#23. Ping

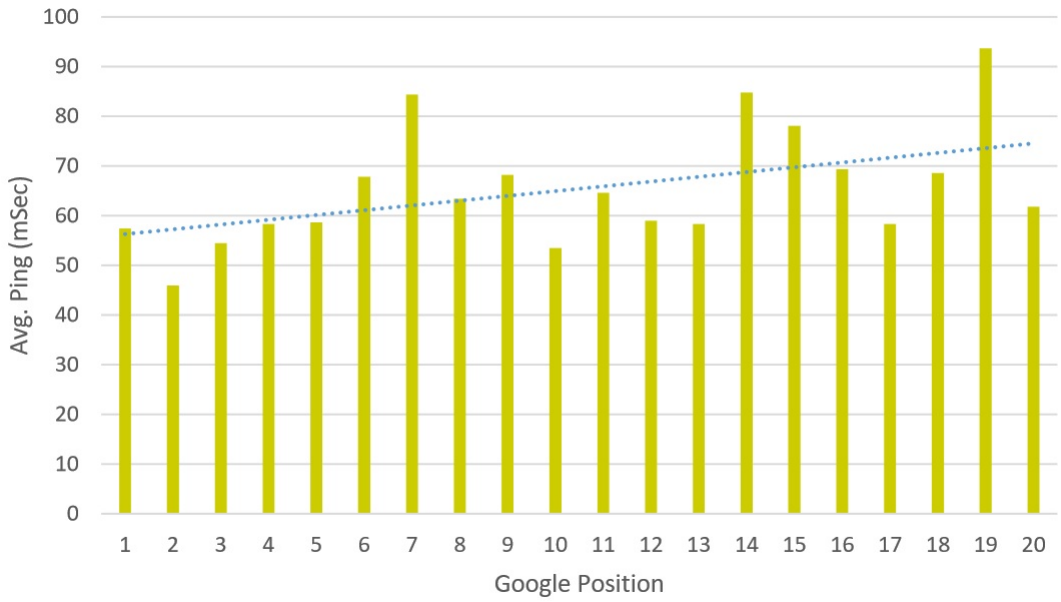


Figure 5.4.23: Average search results relative with ping factor

According to the survey results the Greek ferry root websites with smaller ping time are placed in the top positions. Smaller ping times are achieved in servers geographically close to users. As a result, websites which are hosted in Greece or even in Europe have better ranking position comparatively with others in far distances. Statistically **the correlation value of this factor is 0,39.**

#24. Keyword in domain

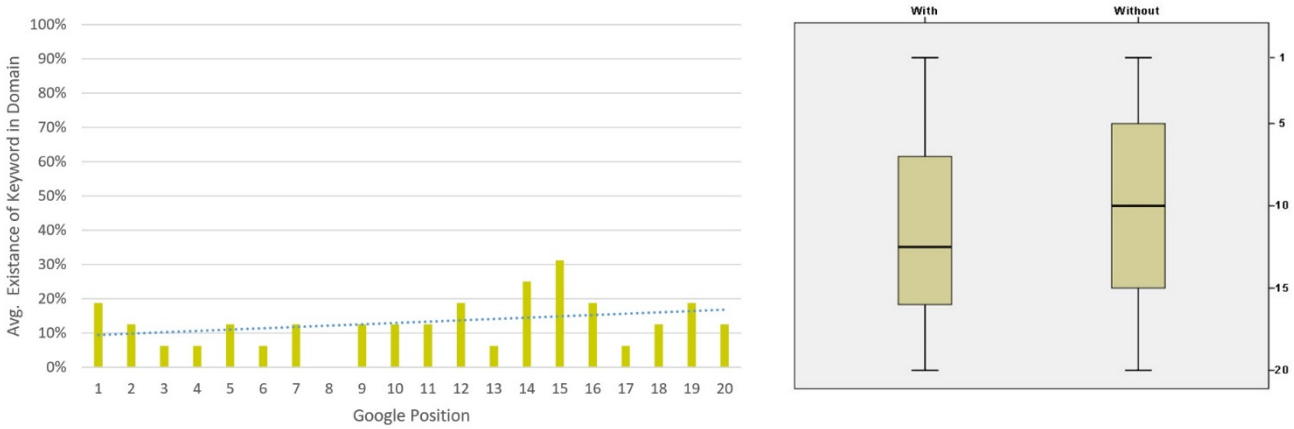


Figure 5.4.24: Average search results relative with keyword in domain factor

The survey results show that the minority of the Greek ferry route websites have a keyword in their domain. As a result, no safe conclusion can be extracted that this factor really affects the Google search rankings and in our case study and it does not exist any statistical significance.

#25. Keyword in Sub – Directory

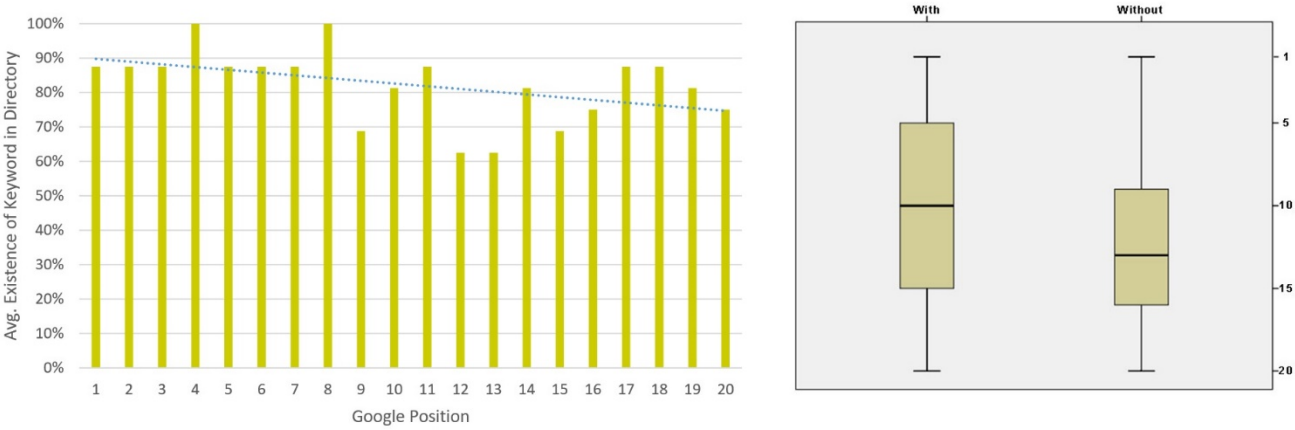


Figure 5.4.25: Average search results relative with keyword in sub – directory factor

Analyzing the graph, it is noticed that the websites that are placed in the top 8 positions of the ranking, are using a keyword in the directory in the same percentage with very small deviations. In the remaining positions, the percentage is obviously lower and from this it becomes clear that there is a strong correlation about this factor. Statistically **the correlation value of this factor is 0,50.**

#26. Domain Age

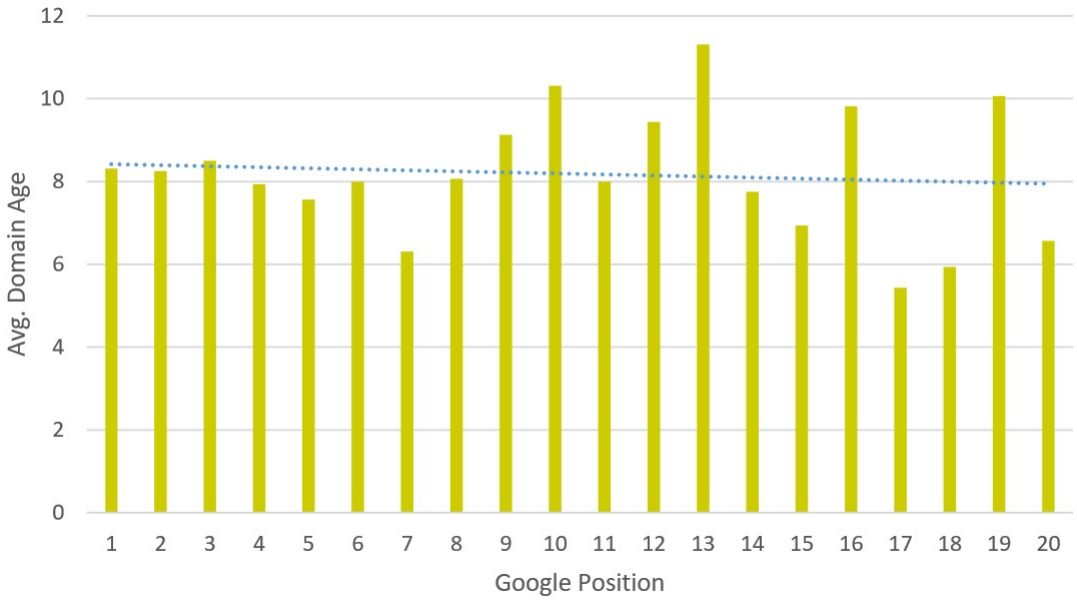


Figure 5.4.26: Average search results relative with keyword in domain age factor

The survey results show that most of the Greek Ferry Route Websites (until 16th position) have the same domain age with small deviations between them. As a result, no safe conclusion can be extracted that this factor really affects the Google search rankings and does not exist any statistical significance.

#27. Facebook activity

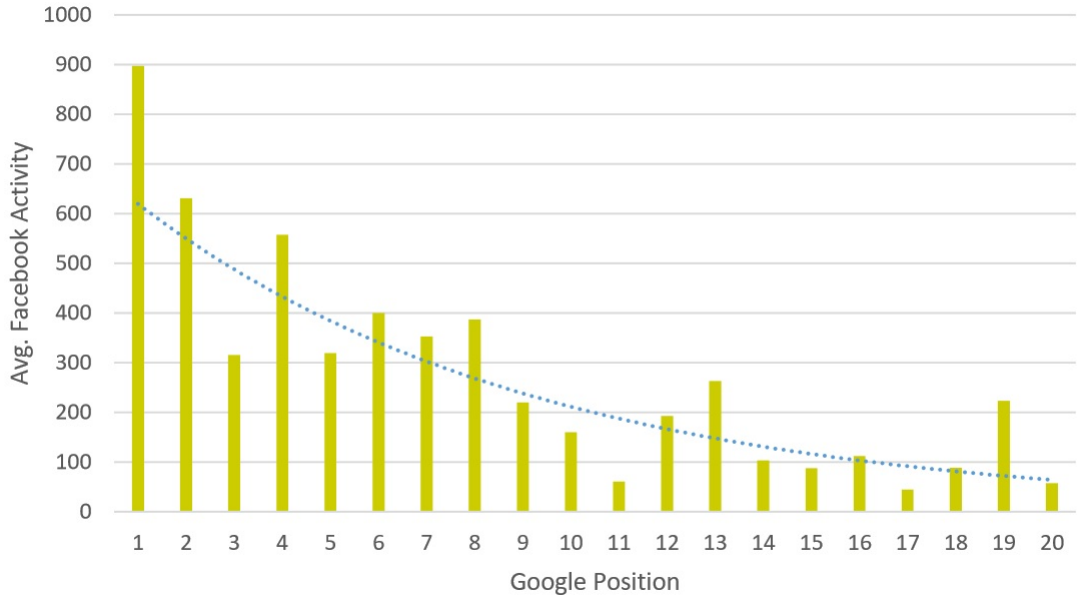


Figure 5.4.27: Average search results relative with Facebook activity factor

According to the graph, Facebook activity has a great effect at the Google search rankings. Particularly, the higher we move in ranking position, the more the websites use Facebook. An interesting element is the fact that all pages are using Facebook. Statistically **the correlation value of this factor is -0,65.**

#28. Google+ activity

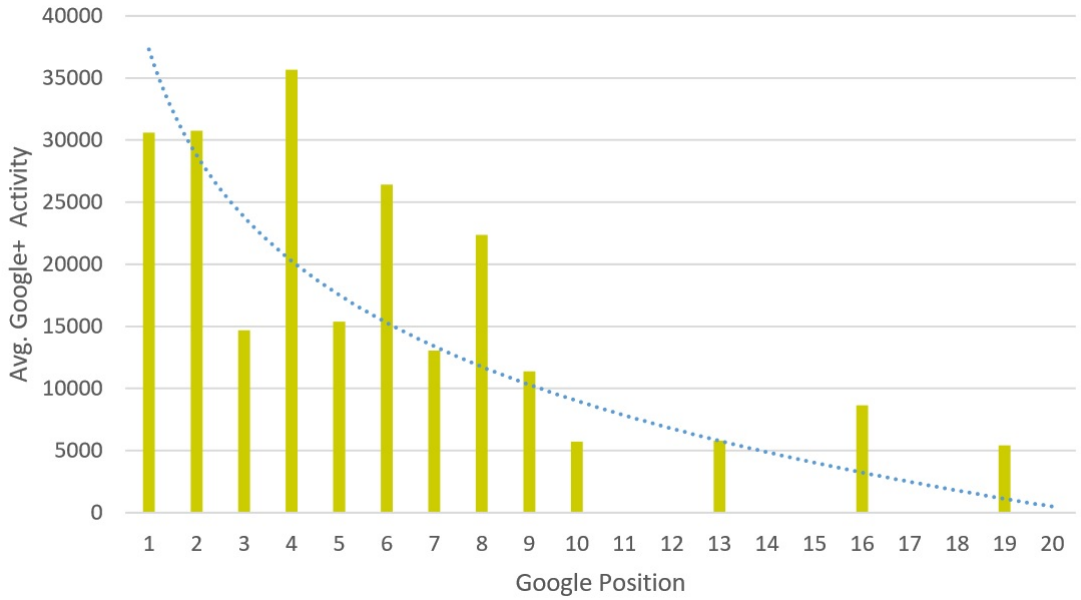


Figure 5.4.28: Average search results relative with Google+ activity factor

According to the graph, Google+ activity has a great effect at the Google search rankings. Particularly, the higher we move in ranking position, the more the websites use Google+. In contrast with Facebook, in some particular positions none website uses Google+. Statistically **the correlation value of this factor is -0,65.**

#29. Root Citation – Citation Flow

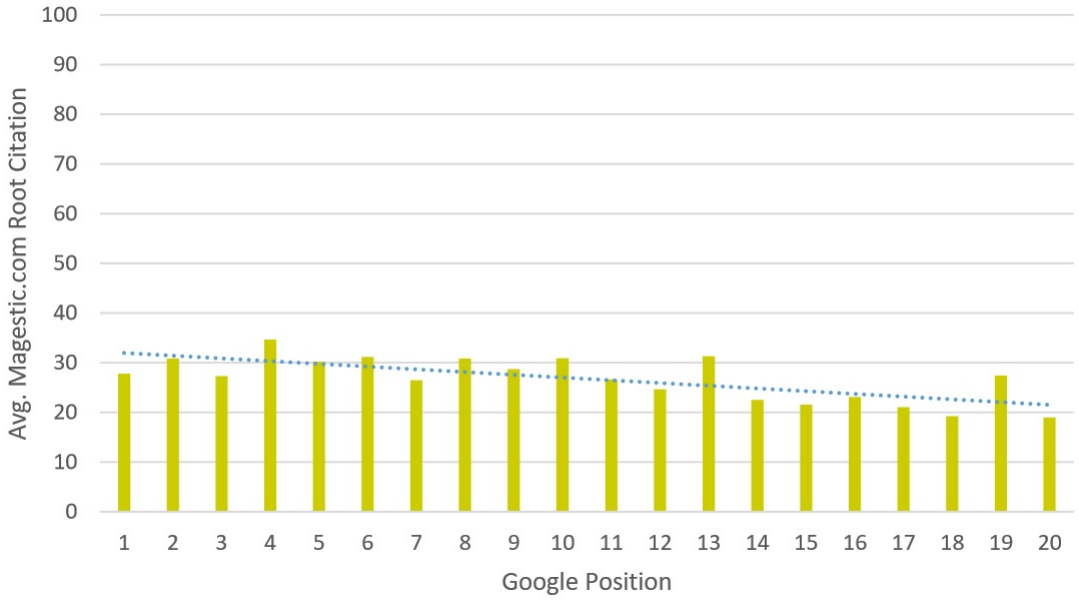


Figure 5.4.29: Average search results relative with root citation factor

According to the graph, root citation has a big effect at the Google search rankings. Particularly, the higher we move in ranking position, the more the websites have other websites pointing on them. Statistically **the correlation value of this factor is -0,50.**

#30. Root Trust – Trust flow

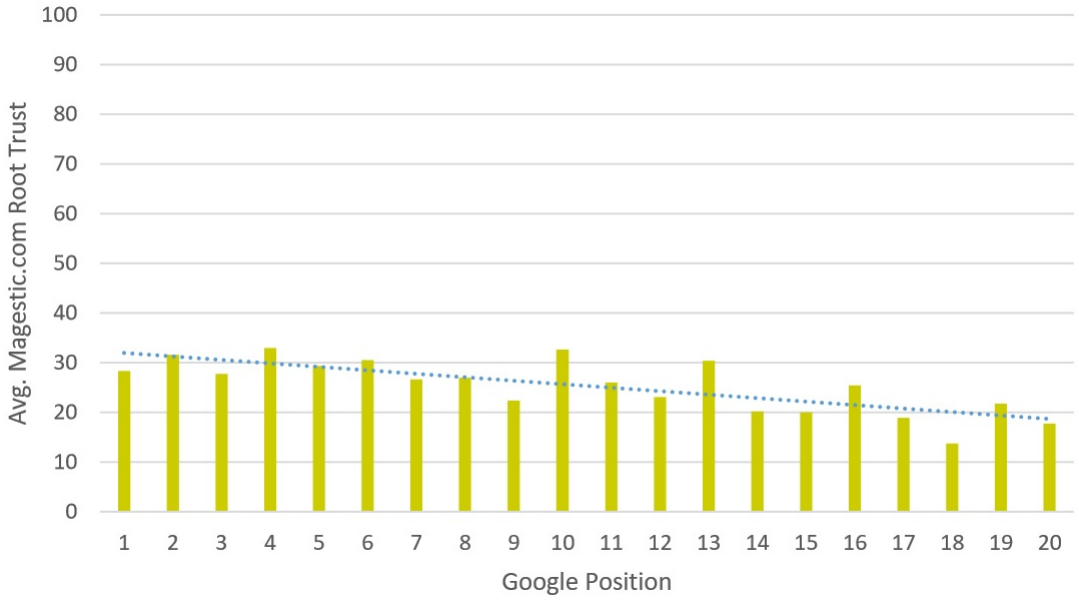


Figure 5.4.30: Average search results relative with root trust factor

According to the graph, root trust has a big effect at the Google search rankings. Particularly, the higher we move in ranking position, the more the websites have quality backlinks pointing on them. Statistically **the correlation value of this factor is -0,61.**

#31. External Backlinks

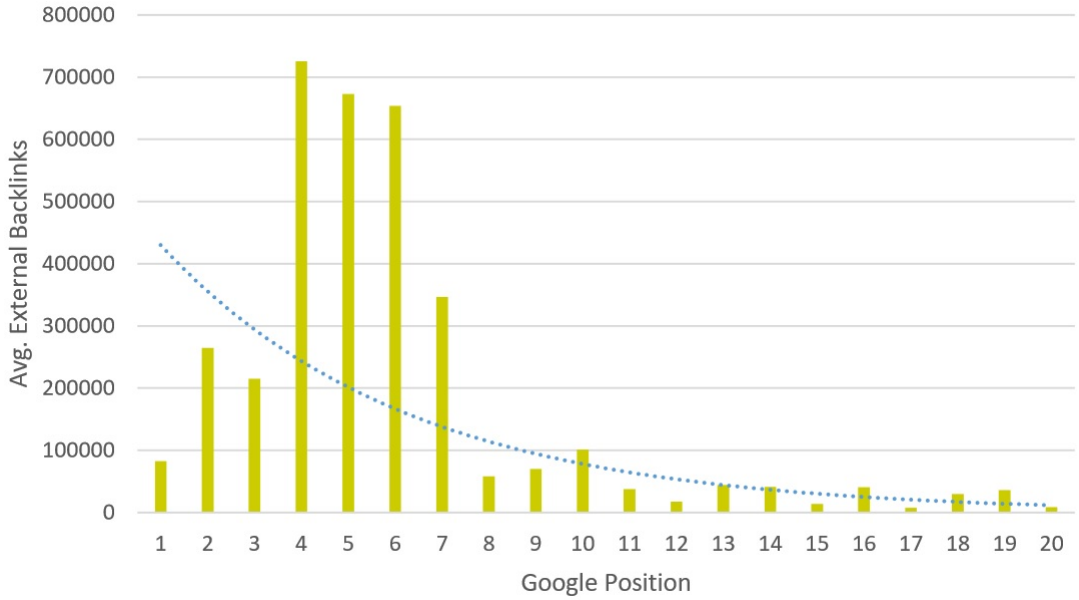


Figure 5.4.31: Average search results relative with external backlinks factor

According to the graph, external backlinks have enormous effect at the Google search rankings. The survey's distribution about the Greek ferry roots is a little complicated and it is not clear enough how it affects the final results. In the three first places, there are websites with less external backlinks comparatively with the places 4,5,6,7. In any case we can say with certainty that they have a decisive role in improving the position of a website. Statistically **the correlation value of this factor is -0,65.**

#32. Referring Domains

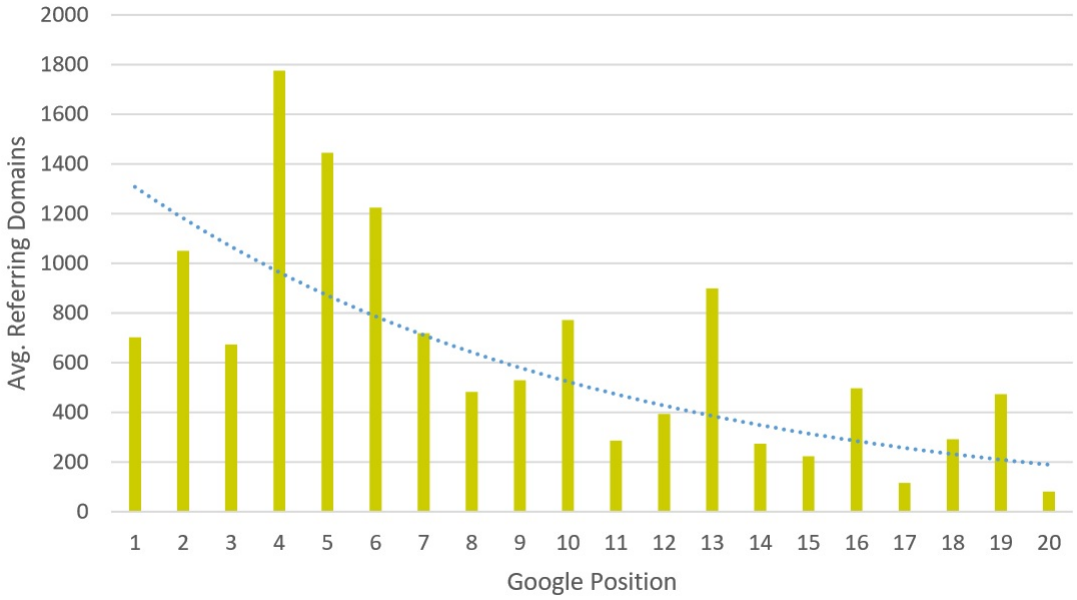


Figure 5.4.32: Average search results relative with referring domains factor

According to the graph, referring domains have a big effect at the Google search rankings. In this factor applies the same as the previous one (#31. External Backlinks). In the first three places, there are less referring domains than the places 4,5,6 but in general it can improve the final Google ranking. Statistically **the correlation value of this factor is -0,56.**

#33. URL Citation

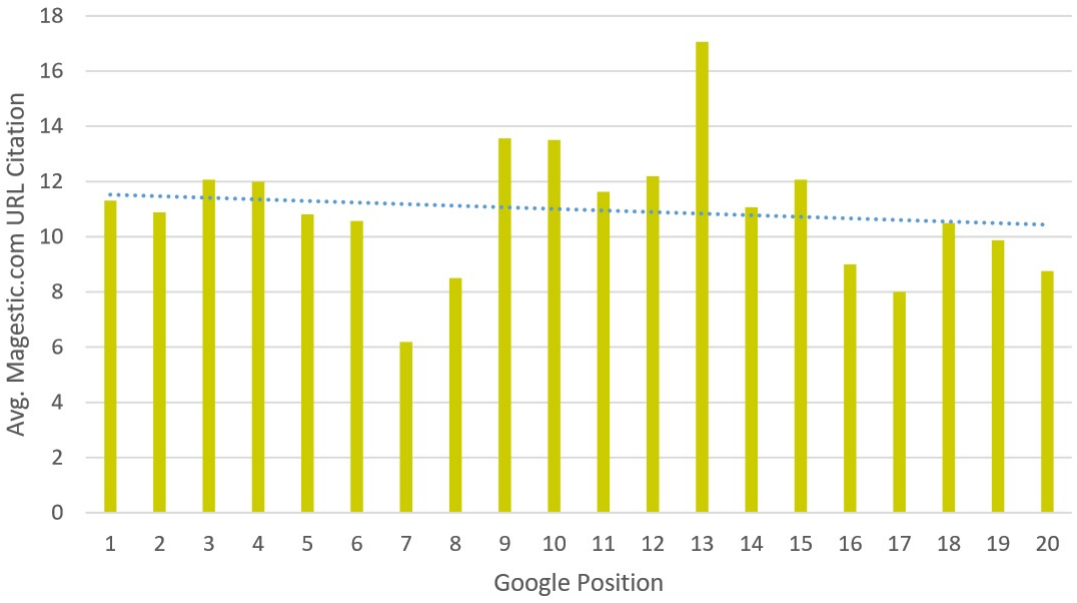


Figure 5.4.33: Average search results relative with URL citation factor

The results regarding the URL citation factor show that there is not a specific allocation in order to extract a safe conclusion. It seems that it does not affect the Google search rankings in our case study and it does not exist any statistical significance.

#34. URL Trust

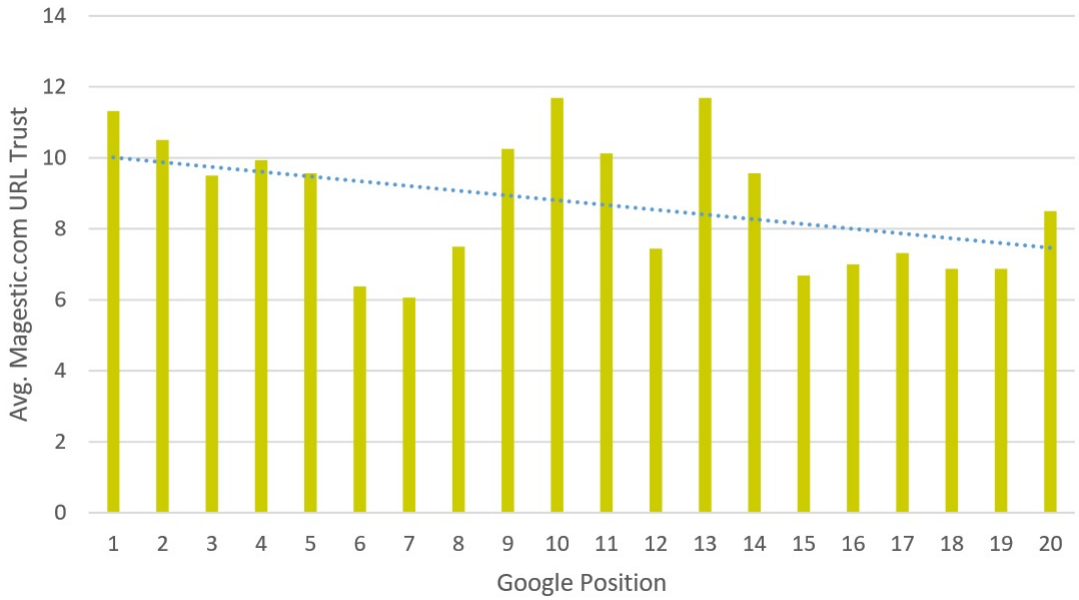


Figure 5.4.34: Average search results relative with URL trust factor

The results regarding the URL trust factor show that there is not a specific allocation in order to extract a safe conclusion. It seems that it does not affect the Google search rankings in our case study and it does not exist any statistical significance.

#35. URL External Backlinks

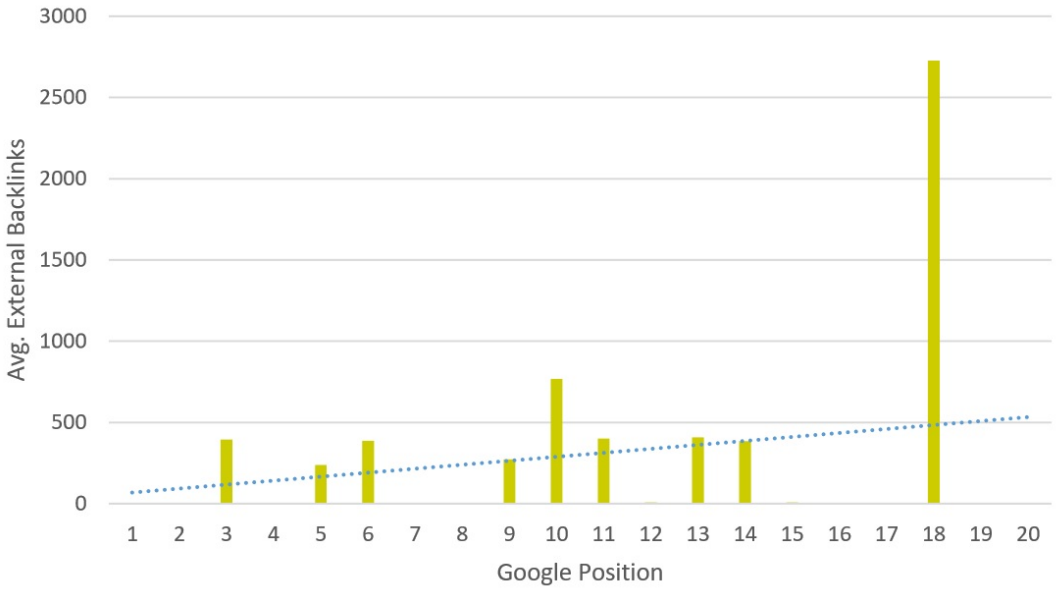


Figure 5.4.35: Average search results relative with URL external backlinks factor

The results regarding the URL external backlinks factor show that there is not a specific allocation in order to extract a safe conclusion. Half of the ranking positions do not even have URL external backlinks and the websites placed in the 18th position have more than 2.500. It seems that it does not affect the Google search rankings in our case study and it does not exist any statistical significance.

#36. URL Referring domains

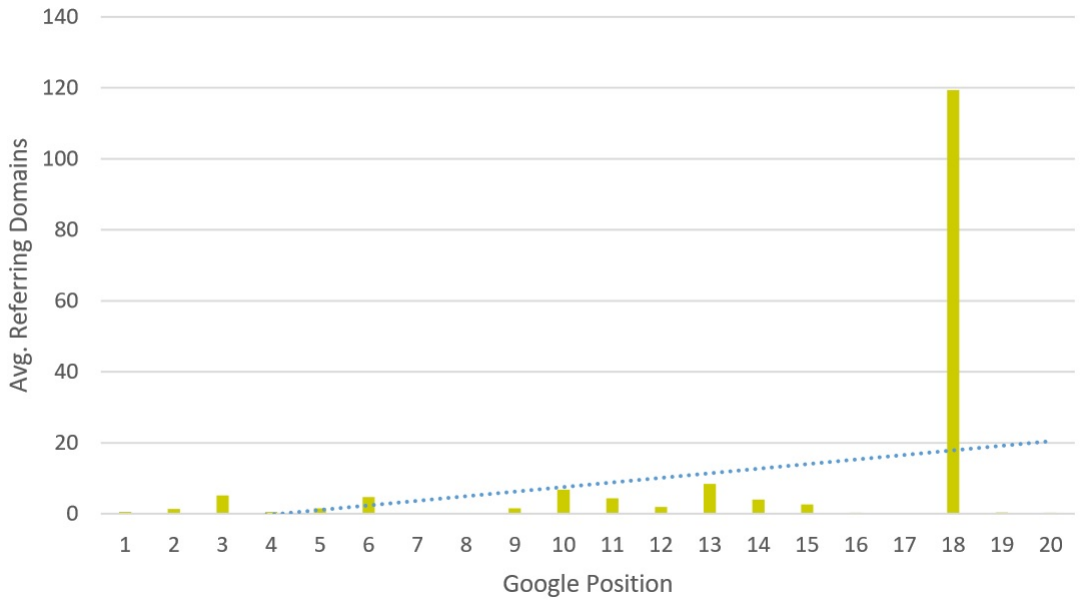


Figure 5.4.36: Average search results relative with URL referring domains factor

The results regarding the URL referring domains factor show that there is not a specific allocation in order to extract a safe conclusion. Half of the ranking positions do not even have URL external backlinks and the websites placed in the 18th position have more than 100. It seems that it does not affect the Google search rankings in our case study and it does not exist any statistical significance.

#37. Mobile Pagespeed

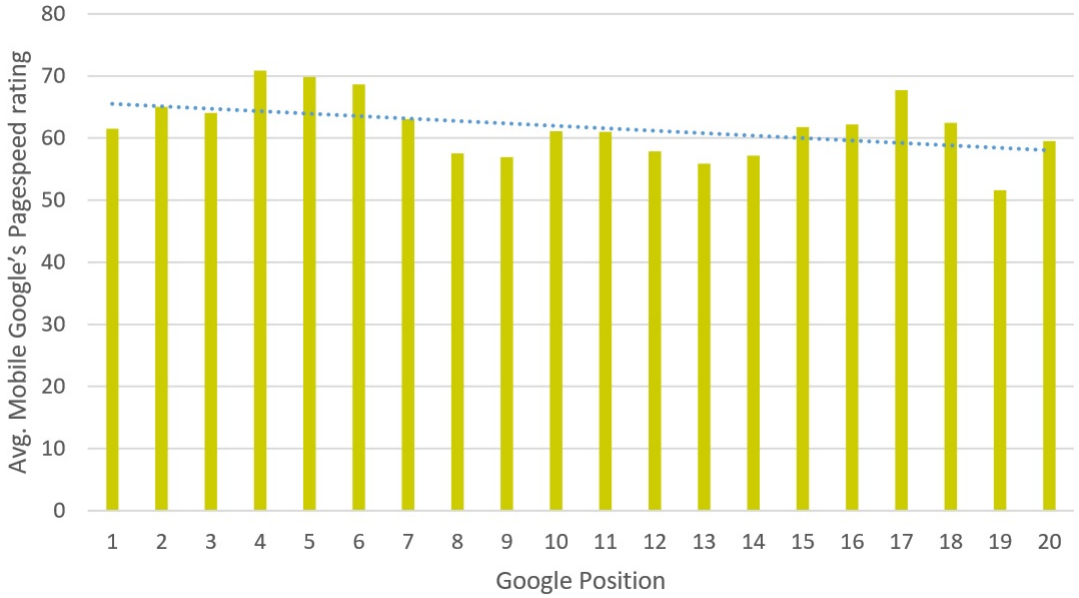


Figure 5.4.37: Average search results relative with mobile pagespeed factor

The survey results show that most of the Greek Ferry Route Websites have the same mobile pagespeed with small deviations between them. As a result, no safe conclusion can be extracted that this factor really affects the Google search rankings and does not exist any statistical significance.

#38. Mobile UX

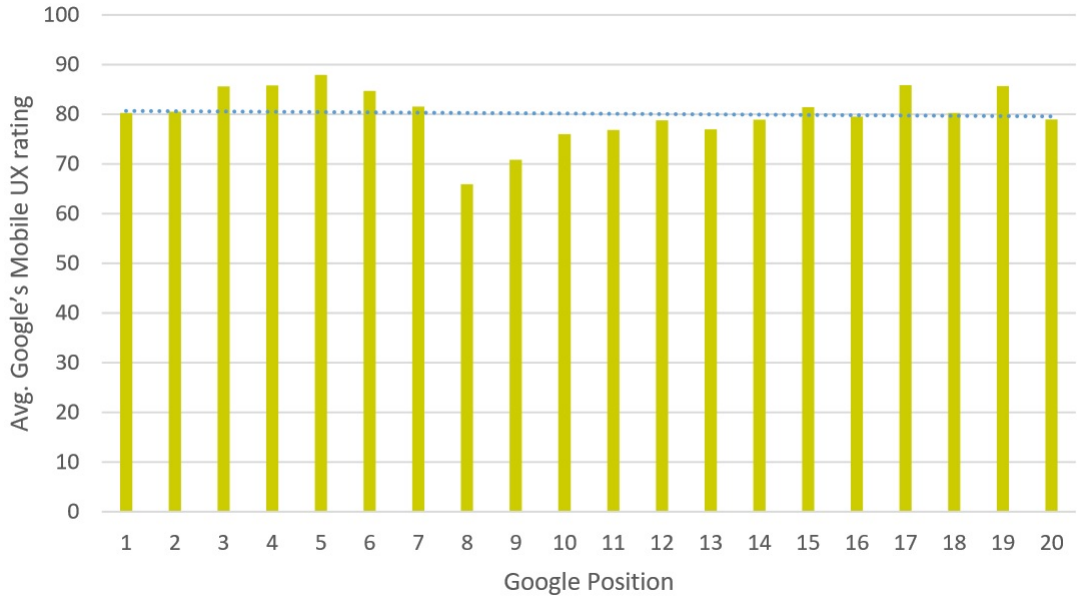


Figure 5.4.38: Average search results relative with mobile UX factor

The survey results show that most of the Greek Ferry Route Websites have the same mobile UX with very small deviations between them. As a result, no safe conclusion can be extracted that this factor really affects the Google search rankings and does not exist any statistical significance.

#39. Mobile Compatibility

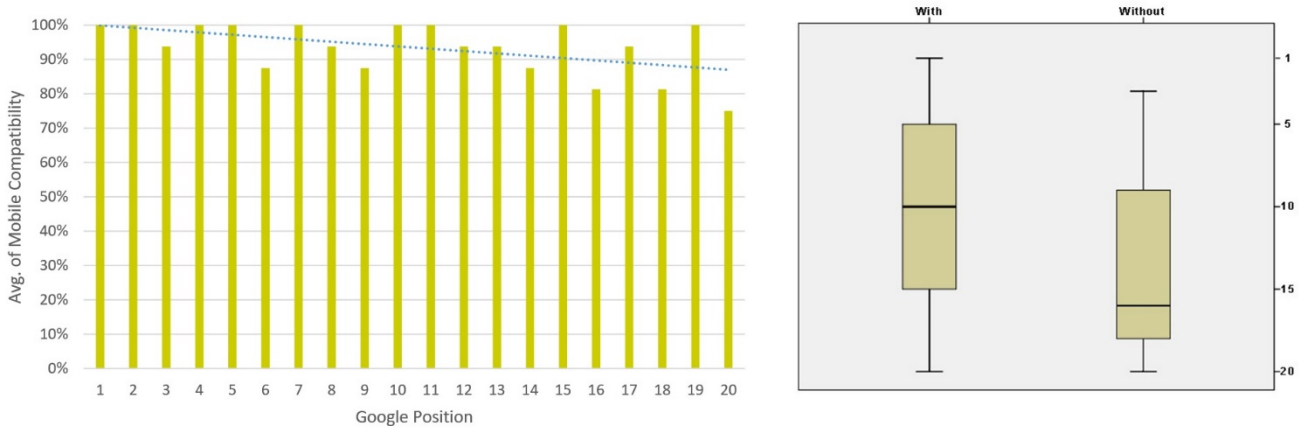


Figure 5.4.39: Average search results relative with mobile compatibility factor

According to the graph, mobile compatibility has a big effect at the Google search rankings. With a few exceptions, the higher we move in ranking position, the more the websites are mobile compatible. Statistically **the correlation value of this factor is -0,49.**

#40. Font size legibility

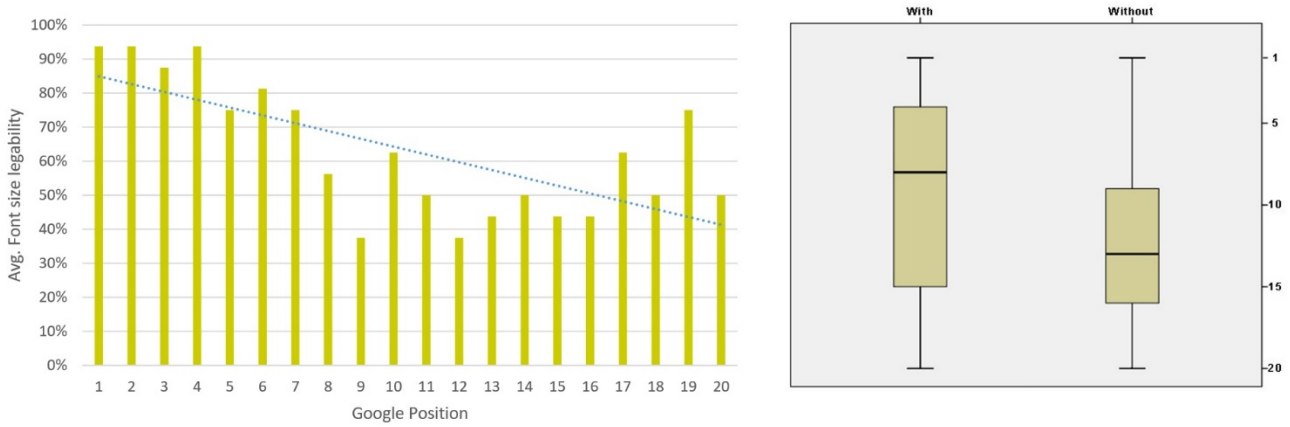


Figure 5.4.40: Average search results relative with font size legibility factor

According to the data of the survey, there is an important correlation between the font size legibility and Google ranking position. More specifically, websites high font size legibility occupy higher ranking positions than those who do not. **The correlation value here is -0.66.**

#41. Mobile Friendly

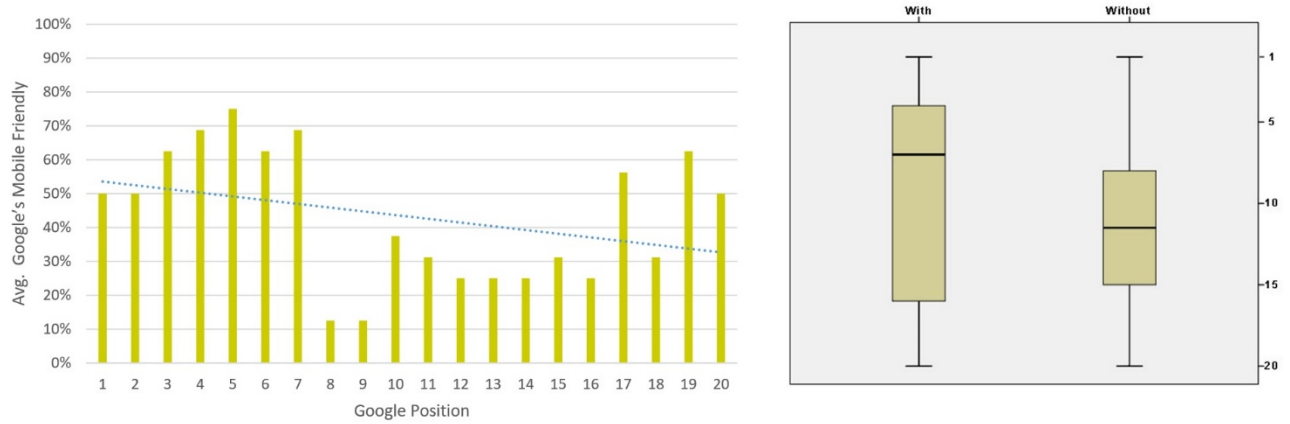


Figure 5.4.41: Average search results relative with mobile friendly factor

Analyzing the graph partially (until the 16 position), it is noticed that the mobile friendly factor is heavily affecting the Google search rankings. The difference between the half first places (until position 7) and the last half places (position 8 to 16) is enormous and clearly the percentage of mobile friendliness of a website has a great impact in the search results. However, three particular positions (17, 19 and 20) is opposed to the rest of the research, and so does not exist a statistical significant result.

#42. Moz Rank

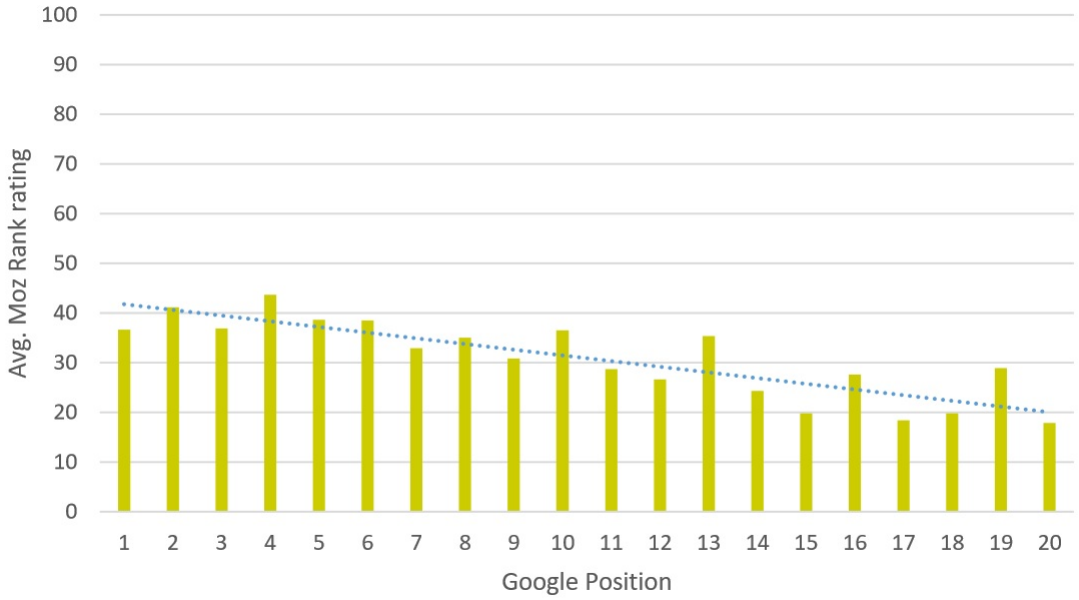


Figure 5.4.42: Average search results relative with Moz rank factor

According to the graph, Moz rank has a big effect at the Google search rankings. Particularly, the higher we move in Google ranking position, the higher is the corresponding Moz ranking of the websites. Statistically **the correlation value of this factor is -0,69.**

#43. Woorank

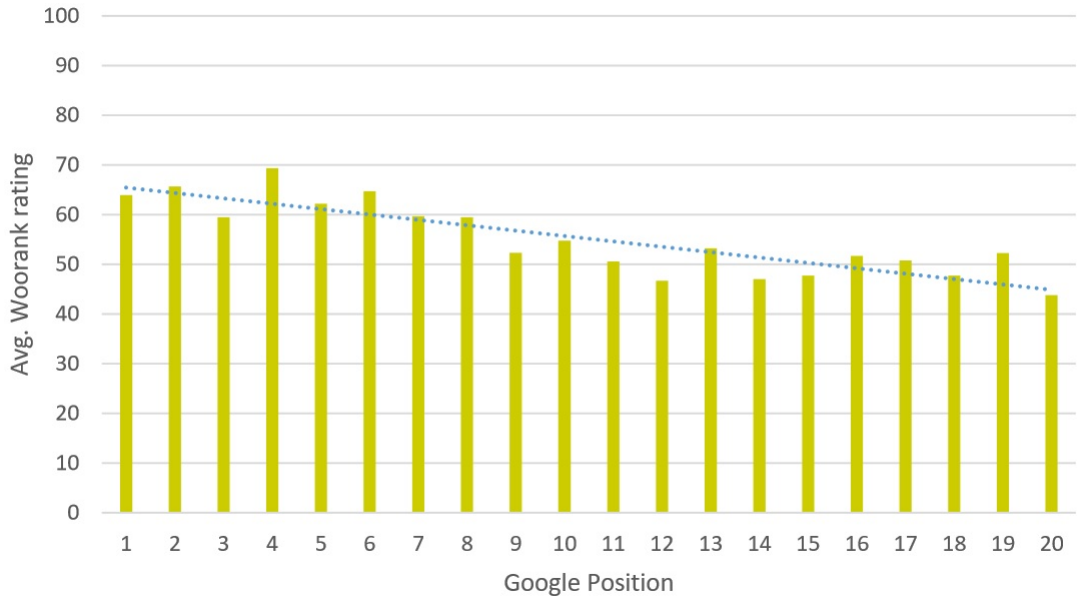


Figure 5.4.43: Average search results relative with Woorank factor

According to the graph, Woorank has a big effect at the Google search rankings. Particularly, the higher we move in Google ranking position, the higher is the corresponding Woo ranking of the websites. Statistically **the correlation value of this factor is -0,65.**

5.5 Conclusion

After analyzing all the 43 different factors the survey's results lead us to the following conclusions. We can divide the factors into 3 categories depending on the degree that affect the results of a search in Google machine. There are factors that affect positively to a large extent, factors that do not affect at all and have no impact,

5.5.1 Table Factor Categories

Positive effect		**No effect**	**Negative effect**		**No correlation**
Factor Name	Correlation Value	Factor name	Factor name	Correlation Value	Factor name
#42. Moz Rank	-0.69	#3. H1	#2. Meta Description	0.45	#1. Title tag
#27. Facebook activity	-0.67	#5. Alt Attribute	#23. Ping	0.39	#18. Page Size (Mb)
#40. Font size legibility	-0.66	#6. Broken Links			#21. Desktop Pagespeed
#28. Google+ activity	-0.65	#8. IP Canonicalization			#24. Keyword in domain
#31. External Backlinks	-0.65	#10. XML Sitemap			#26. Domain Age
#43. Woorank	-0.65	#12. Underscores in the URLs			#36. URL Referring domains
#14. Custom 404 Page	-0.64	#22. Load Time (sec)			#37. Mobile Pagespeed
#16. Structured Data Markup	-0.62	#33. URL Citation			#38. Mobile UX
#30. Root Trust – Trust flow	-0.61	#34. URL Trust			
#7. WWW Resolve	-0.59	#35. URL External Backlinks			
#32. Referring Domains	-0.56	#41. Mobile Friendly			
#4. H2	-0.55				
#17. SSL Secure	-0.55				
#9. Robots.txt	-0.52				
#15. Language	-0.52				
#25. Keyword in Sub – Directory	-0.50				
#29. Root Citation – Citation Flow	-0.50				
#39. Mobile Compatibility	-0.49				
#11. URL Rewrite	<0.49				
#13. Blocking Factors	<0.49				
#19. Google Analytics	<0.49				
#20. Directory Browsing	<0.49				

and factors that negatively affect the final ranking in a Google search. Finally, there is one more category which differs from the other 3, as the survey's results could not give us a correlation or all the websites that was participated in the survey had the same characteristics and so the results were the same for all. The table 5.5.1 summarizes all the factors, divided each one into the corresponding category with the degree to which they affect the results of a Google search.

Table 5.5.1: The four factor categories as they emerged from the survey made

According to the table 5.5.1 it is easily understood that half of the factors (22 factors) can positively influence Google rankings. On the other hand, only 2 factors have negative effects, while 11 factors had no effect on the survey made. Finally, there are 8 factors that did lead us to a safe conclusion as the available data were the same for every website participated in the survey and could not be used.

The most important factors, in which the webmasters have to give more importance if they want to improve a website's position, is the Moz rank, the Facebook activity, the font size legibility, the Google+ activity, the external backlinks and the Woorank.

References

airtickets.gr, 2017. *About us.* [Online]
Available at: http://www.airtickets.gr/about-us
[Accessed 18 September 2017].

Allen, R., 2017. *Search Engine Statistics 2017.* [Online]
Available at: http://www.smartinsights.com/search-engine-marketing/search-engine-statistics/
[Accessed 2017 September 18].

Amadeus, 2007. *Service Fees and Commission Cuts,* s.l.: Amadeus.

Beal, V., 2016. *The Difference Between the Internet and World Wide Web.* [Online]
Available at: http://www.webopedia.com/DidYouKnow/Internet/Web_vs_Internet.asp
[Accessed 22 December 2016].

Cisco, 2017. *The Zettabyte Era: Trends and Analysis ,* s.l.: Cisco public.

Dellaert, B. G., 2014. *How Tolerable is Delay?Consumers' Evaluations of Internet Web Sites after Waiting.* [Online]
Available at:
https://www.researchgate.net/publication/4783290_How_Tolerable_Is_Delay_Consumers%27_Evaluations_of_Internet_Web_Sites_After_Waiting
[Accessed 18 September 2017].

eBizMBA, 2017. *Top 15 Most Popular Search Engines.* [Online]
Available at: http://www.ebizmba.com/articles/search-engines
[Accessed 18 September 2017].

Enge, E., 2013. *Does Facebook Activity Impact SEO?.* [Online]
Available at: https://www.stonetemple.com/does-facebook-activity-impact-seo/
[Accessed 18 September 2017].

Everts, T., 2015. *Page bloat: The average web page size is more than 2MB.* [Online]
Available at: https://www.soasta.com/blog/page-bloat-average-web-page-2-mb/
[Accessed 18 September 2017].

forthcrs, 2017. *Company.* [Online]
Available at: http://www.forthcrs.com/company/
[Accessed 18 September 2017].

forthcrs, 2017. *Products.* [Online]
Available at: http://www.forthcrs.com/products/
[Accessed 18 September 2017].

Google, 2010. *Using site speed in web search ranking.* [Online]
Available at: https://webmasters.googleblog.com/2010/04/using-site-speed-in-web-search-ranking.html
[Accessed 18 September 2017].

Google, 2017. *Mobile-Friendly Test Tool.* [Online]
Available at: https://support.google.com/webmasters/answer/6352293
[Accessed 18 September 2017].

Griffiths, S., 2015. *Mobile App UX Principles,* s.l.: Google.

Helander, M. G. & Khalid, H. M., 2000. Modeling the customer in electronic commerce. *Applied Ergonomics .*

Hellenic Chamber of Hotels, 2016. *HOTEL RESIDENCES OF GREECE,* s.l.: Hellenic Chamber of Hotels.

Hellenic Statistical Authority, 2016. *Press release,* s.l.: Hellenic Statistical Authority.

Hellenic Statistical Authority, 2016. *Survey on the Use of Information and Communications Technologies by Households and Individuals,* s.l.: Hellenic Statistical Authority.

illinois.edu, 2002. *The first search engine, Archie.* [Online]
Available at: https://chip.web.ischool.illinois.edu/people/projects/timeline/1990archie.htm
[Accessed 18 September 2017].

McGee, M., 2015. *Google's Gary Illyes: HTTPS May Break Ties Between Two Equal Search Results.* [Online]
Available at: http://searchengineland.com/googles-gary-illyes-https-may-break-ties-between-two-equal-search-results-230691
[Accessed 18 September 2017].

MozRank, 2017. *What is MozRank?.* [Online]
Available at: https://moz.com/learn/seo/mozrank
[Accessed 18 September 2017].

netmarketshare, 2017. *Market Share Reports.* [Online]
Available at: http://www.netmarketshare.com/
[Accessed 2017 September 18].

Peppard, J. & Butler, P., 1998. CONSUMER PURCHASING ON THE INTERNET: PROCESSES AND PROSPECTS. *European Management Journal.*

PwC, 2016. *IAB internet advertising revenue report ,* s.l.: Interactive Advertising Bureau (IAB) .

rapidwebseo, n.d. *Matt Cutts: Does Domain Age Really Matter?.* [Online]
Available at: http://rapidwebseo.com/matt-cutts-does-domain-age-really-matter.php
[Accessed 18 September 2017].

S. Shunmuga Krishnan, R. K. S., 2012. *Video Stream Quality Impacts Viewer Behavior: Inferring Causality Using Quasi-Experimental Designs,* s.l.: s.n.

Schwartz, B., 2013. *Google: Site Speed Penalty Coming To Mobile Web Sites.* [Online]
Available at: http://searchengineland.com/google-mobile-site-speed-162977
[Accessed 18 September 2017].

Seymour, T., Frantsvog, D. & Kumar, S., 2011. History Of Search Engines. *International Journal of Management & Information Systems ,* Volume 15.

Slegg, J., 2016. *Next Mobile Friendly Update Includes Page Speed; Coming in Months.* [Online]
Available at: http://www.thesempost.com/next-mobile-friendly-update-includes-page-speed-coming-in-months/
[Accessed 18 Septeber 2017].

statcounter, 2017. *Desktop, Tablet & Console Search Engine Market Share in Greece.* [Online]

Available at: http://gs.statcounter.com/search-engine-market-share/desktop-tablet-console/greece
[Accessed 18 September 2017].

thomascook.com, 2017. *Thomas Cook History.* [Online]
Available at: https://www.thomascook.com/thomas-cook-history/
[Accessed 18 September 2017].

Toothman, J., n.d. *What's the difference between the Internet and the World Wide Web?.* [Online]
Available at: http://computer.howstuffworks.com/internet/basics/internet-versus-world-wide-web1.htm

University System of Georgia, n.d. *Internet Search Services.* [Online]
Available at: http://www.usg.edu/galileo/skills/unit07/internet07_07.phtml
[Accessed 2017 September 18].

usabilitynet, n.d. *Cost justifying Usability.* [Online]
Available at: http://www.usabilitynet.org/management/c_cost.htm
[Accessed 18 September 2017].

van den Bosch, A., Bogers, T. & de Kunder, M., 2016. Estimating search engine index size variability: a 9-year longitudinal study. *Scientometrics*, 9 February.

visitgreece.gr, 2017. *Greek Islands.* [Online]
Available at: http://www.visitgreece.gr/el/greek_islands
[Accessed 18 September 2017].

WooRank, 2017. *So What Is WooRank?.* [Online]
Available at: https://www.woorank.com/en/p/about
[Accessed 18 September 2017].