Ahmed Mobarek

Demand modeling for mobile app stores

School of Electrical Engineering
Thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Engineering.

Thesis Supervisor:
D.Sc. (Tech.) Kalevi Kilkki

Thesis Instructor:
M.Sc. (Tech) Benjamin Finley

Aalto University
School of Science and Technology

AALTO UNIVERSITY
SCHOOL OF ELECTRICAL ENGINEERING

ABSTRACT OF THE
MASTER'S THESIS

Author: Ahmed Mobarek
Title: Demand modeling for mobile app stores
Date: 10.06.2014 Language: English Number of Pages: 71+9

Department of Communications and Networking
Professorship: Network Economics Code: ETA3003

Supervisor: D.Sc. (Tech.) Kalevi Kilkki
Instructor: M.Sc. (Tech) Benjamin Finley

Smartphones have reached a relatively high market share of the mobile market, creating new market opportunities. As a result, different stakeholders are investing in the mobile industry attempting to generate a higher revenue share. Hence, competition between various mobile device manufacturers has increased, as they compete for customers. These device manufacturers have created their own ecosystems, trying to lock-in their customers. These ecosystems include the application (app) stores providing services for mobile users. Currently, the two leading app stores are the Apple App Store and Google Play. Similarly, the competition exists among app developers of both stores. Therefore, it is vital to understand the user demands to design a successful app is popular in these stores.

This thesis identifies successful app categories for both app stores from the perspective of an app developer. It adopts basic descriptive analysis for the dataset provided during September and October 2013 regarding the US and Finnish markets. Furthermore, it introduces a probabilistic graphical model based on Bayesian Network, aiming to understand the dynamics of mobile app stores. The thesis defines the success indicator for each category of apps, and then compares the results of both app stores. The top successful app categories in the US market include Social Networking, Productivity, Music, Finance, Education, Sports, Entertainment, and Travel. The corresponding app categories in Finland include Social Networking, Finance, Education, Music, Productivity, Entertainment, Photos and Video, Lifestyle, Games, and News. The thesis concludes that Google Play has higher success indicators than Apple App Store both in US and Finnish markets. Additionally, the success indicator is higher for free apps compared to paid apps.

The results of this research contribute to recommendations for developers, during the development and publishing stages of an app, as well as building marketing strategies for mobile apps. Furthermore, it suggests a framework to identify successful apps in mobile app stores.

Keywords: App, Mobile app stores, Apple App Store, Google Play, App developers, Bayesian Network

Preface

This Master's Thesis has been written as a partial fulfillment for the Master of Science Degree at Aalto University, School of Electrical Engineering. The study was carried out in the Department of Communications and Networking, as part of the Network Economics research team.

I would like to thank my supervisor Kalevi Kilkki for giving me the opportunity to be part of the Network Economics research team and his guidance during this thesis project, as well as learning from his experience in the field. Additionally, I would like to express my appreciation to my instructor Benjamin Finley for his cooperative feedback and professional support while writing this thesis.

Last but not least, I thank all of my family and friends for their support during my experience studying in Finland. I am grateful to Nabil Mobarek, Sahar Abed, and Salma Mobarek who have been a huge moral support during my professional career decisions. Moreover, I highly appreciate the presence of my friends during the tough and cheerful moments of my journey. I would like to thank every person whom I have shared with special memories during my study period.

Ahmed Mobarek
10th of June 2014
Espoo, Finland

Table of Contents

List of Figures

List of Tables

Acronyms

App	Application
BN	Bayesian Network
DAG	Direct Acyclic Graph
FI	Finland
GSM	Global System for Mobile Communications
ICT	Information and Communication Technology
IT	Information Technology
US	United States

1 Introduction

The introduction of smartphones has dramatically changed the mobile industry, leading to the creation of new market opportunities. Smartphones are widely used and have penetrated the worldwide mobile market, as illustrated by the high number (61%) of US mobile phone users who had adopted a smartphone as of June 2013 (Sterling 2013). This penetration has led to a growing interest in the development of mobile application programs, more commonly referred to as "apps". Mobile apps have become a significant source of revenue for companies, such as mobile app developers, or even other companies providing a mobile app as an online tool for their customers in order to access their services. These companies can vary in size, ranging from a big corporate company developing various apps to a single person developing only one single app. In addition, these apps are published in a publically accessible market, which is competitive by nature. As a result, companies need to understand user behavior in order to increase the demand for their apps.

Development of the mobile industry has led to adoption of the term Business Ecosystem; in fact, the term refers to the business environment using a metaphor of the biological ecosystem (Moore 1993). The creation of ecosystems has opened up a new source of revenue for mobile device manufacturers in order to capture more value from their customers. This has led many mobile device manufacturers to launch app stores over the past few years, in an effort to lock-in their customers by increasing the dependency on their services. These emerging app stores reflect the importance of creating an ecosystem and understanding the behavior of different actors within the same ecosystem. The app stores are offered as services along with the smartphone software package provided by the mobile device manufacturer.

Recently, Shaughnessy (2013) has stated that nine out of ten smartphones run one of the two leading operating systems, either Android or iOS. Therefore, the two main app stores, i.e. Apple App Store and Google Play, have reshaped the whole mobile industry and successfully created profitable app store models. In July 2013, Google announced the availability of 1,000,000 apps in their store, while Apple mentioned reaching 900,000 Apple App Store apps during a conference in June 2013 (Rowinski 2013b). Google Play has a market share of 74.4%, while Apple generates 85% more revenue than Google Play (App Annie 2014). These two ecosystems have adopted different business models. Apple offers a more closed ecosystem than Google. Apple limits the access of other non-Apple devices to the store, whereas Google has introduced an open Android platform allowing users of any Android device to gain access to their store.

Studies focusing on app stores can be divided into four categories: 1) research examining Apple App Store, 2) research examining Google Play, 3) other research comparing both these app stores, and 4) researchers analyzing other app stores. When Apple App Store was first launched by Apple in 2008, many studies focused on the disruptive success of this new business model (Carare 2012; Garg & Telang 2013,Lee & Raghu 2011, Pagano & Maalej 2013, Ayalew 2011, Wang & Wang 2013, Kimbler 2010, Yamakami 2011). This thesis refers to three studies examining Google Play separately (Want 2011, Zhong & Michahelles 2013, Wang & Wang 2013). Recent research has focused on comparing both stores after the introduction of Google Play, emphasizing the difference between both business models (Kokabha 2012, Fredholm & Gunnarsson 2013, Sawant 2010, Rao & Jimenez 2011, Eaton, Elaluf-Calderwood & Sørensen 2010, Holzer & Ondrus 2011, Cuadrado & Dueñas 2012). These studies have identified different ecosystem actors to evaluate the value network chain within the platform. Additionally, other researchers have analyzed the demand for app stores using combined figures from datasets extracted from both Apple App Store and Google Play, rather than analyzing figures separately extracted from each app store (Hyrynsalmi et al. 2012, Ghose & Han 2012). Skogsberg (2013) proposes a web platform for app developers, in order to collect data about the stores and further analyze it.

Although many studies have examined Apple App Store and Google Play, only few studies have focused on modeling and comparing the demand in both platforms. Garg and Telang (2013) assume power law distribution for the demand-rank relationship of apps and provide estimations for these parameters in both Apple App Store and Google Play. These estimations are more useful to an app store provider, i.e. Apple or Google, than to an app developer. Additionally, Kim (2012) further assumes power law distribution and reports different user profiles for both platforms affecting the app demand. Kim considers the correlation between the demand and other app parameters, such as age, size, price, updates, and customer ratings. Thus, Garg and Telang (2013) have provided an abstract analysis serving the interest of the app store providers, while Kim (2012) has evaluated the user profiles in both stores, which serves both app developers and store providers.

However, app developers are typically interested in the revenue gained when launching their apps in any of the app stores. Therefore, further comparative analysis between app stores is needed in order to improve the understanding of the user behavior in both ecosystems. This understanding would enable the app developers to publish apps matching the needs of the users of these mobile app stores, thus leading to a higher probability of publishing successful apps. Hence, this would further reflect on generating more revenue by app developers.

1.1 Thesis objectives

The main objective of this thesis is to model empirically the demand for apps on both Apple App Store and Google Play. The model should enable an understanding of the dynamics of both app stores from the perspective of an app developer. This model aims to improve the understanding of the market by identifying the successful app categories for both app stores. This will be accomplished by calculating the probability of success for apps. This will be further used to propose a success indicator as a parameter of comparison between the different app categories in both app stores, i.e. Apple App Store and Google Play. In order to ensure the success of future apps, the model provides a guideline for app developers facilitating the decision making when developing apps that are most successful in terms of downloads. Hence, this will enable app developers to improve their strategies for launching successful apps. Furthermore, the thesis intends to propose a methodology for analyzing the demand of app stores, enabling its implementation on different datasets not only the one used throughout this research. The thesis will accomplish this objective by answering the following questions:

1) What is the probability of success for apps in different categories?

2) What is the difference in the probability of success between paid and free apps in different stores?

1.2 Research methods

Data from the two mobile app stores was provided by a market research company. The data reveals the ranking of apps in the Apple App Store and Google Play. It also includes download estimations for free and paid apps. The data was collected between September and October 2013; thus, it comprises two months of consecutive daily ranking information. Although there are publicly available data concerning app stores, demand information remains confidential.

This study uses several methods and tools in order to analyze data of mobile app stores. First, the data was preprocessed using Matlab in order to organize it for further analysis. Next, the analysis further utilizes basic descriptive analysis methods in order to represent graphically the processed information. Furthermore, Bayesian Network, referred to as “BN”, is used as a probabilistic graphical model to build the Success Model for the apps. This model is designed using the free version of the AgenaRisk tool, which is commonly used to simulate BN models. Further details regarding these methods and tools will be described in Chapter 3.

1.3 Thesis structure

The remainder of this thesis is organized as follows. Chapter 2 reviews the literature related to the key concepts of this analysis. Chapter 3 describes the data and the research methods used. Chapter 4 presents the results and introduces the Success Model for mobile app stores, i.e. Apple App Store and Google Play. Chapter 5 discusses the results and their significance. Finally, Chapter 6 summarizes the research findings, offers recommendations for app developers and provides suggestions for future research directions.

2 Background

This chapter introduces an overview of the basic principles, which will be followed in this study. It provides a literature review of the relevant topics that are discussed throughout previous studies. It also includes different topics that will support identifying the methodology adopted later on, to analyze the data throughout the thesis. Thus, the background chapter highlights the important findings of previous research within the scope of the thesis.

2.1 Mobile Ecosystem

The digital mobile telecommunications industry has been developing rapidly during the last 24 years since the introduction of GSM. The GSM was the first commercialized digital system, thus data services started to develop since it was first introduced. The data services enhanced developing business models within the field of telecommunications; different actors became more interested to capture the value of this growing industry. The mobile industry started adopting the recently emerging term Business Ecosystem, which is a metaphor of the biological ecosystem in business terms (Moore 1993). It resembles different business actors as living organisms and the complex relationships among these actors. The mobile ecosystem is represented by Figure 1, displaying different actors interacting within that ecosystem (ECOSYS 2004) .

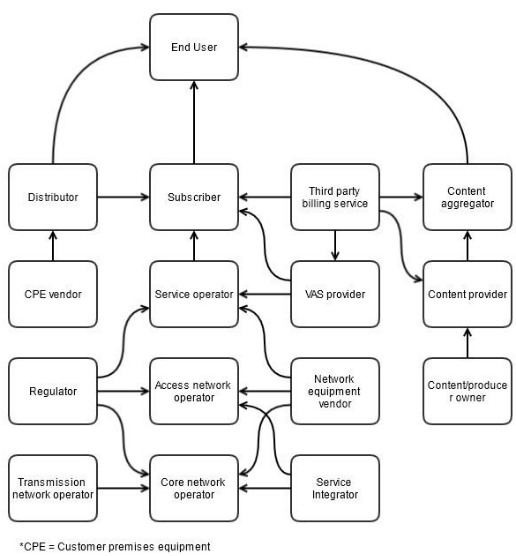


Figure 1 - Mobile Ecosystem (Ecosys 2004)

The mobile ecosystem involved different actors having various motives. Each actor in the ecosystem tries to capture as much value as possible. For example, service operators would like to provide other types of services such as being a content aggregator in order to increase their revenue margin. Mobile device manufacturers such as Apple provide the device and even act as a content network aggregator. Therefore, the mobile ecosystem involves a tough competition amongst companies aiming to capture more value of the ecosystem.

Recently, the mobile apps market has developed fast due to the need of having a platform that provides the apps to consumers. This has led to the mobile ecosystem being an ongoing topic of research. Different types of actors involved and value network within the ecosystem were examined by Tarnacha & Maitland (2006), Barnes (2002), Nyika (2010), Karvonen & Warsta (2004), Basole (2009), Buellingen & Woerter (2004). This thesis presents actors in Figure 2 that are considered of great importance for this research scope. The figure includes mobile app developer followed by mobile app store provider, operating system provider, device manufacturer, mobile network provider and finally the mobile subscriber. A mobile app developer is an actor responsible of publishing and selling the developed apps. The size of this actor varies widely, which can range from a big corporate company providing several apps to a single person providing only one app. Next, the mobile app store or platform provider is responsible for providing the platform that combines various apps within the same platform. This platform is managed by the provider, allowing easy access for the app developers to their customers. An operating system provider is the software platform provider that operates on the hardware mobile device. The mobile device manufacturer is the one who manufactures the phones. In case of Google, it does not restrict the type of device operating on the Android operating system. While Apple offers exclusive access for the iOS operating system through its own mobile devices. The following actor is the mobile network provider providing the network access, by acting as a gateway for the app store enabling exposure to the customers. Finally, the mobile subscriber who is the consumer having a need to access apps online. These actors represent the basic understanding of the mobile ecosystem, serving the research scope of this study. It provides information about the relevant actors, in order to interpret the results of this study along with the benefits for different actors of the ecosystem. Nevertheless, there are more detailed ecosystem analyses representing a wider variety of actors as discussed in previous research.

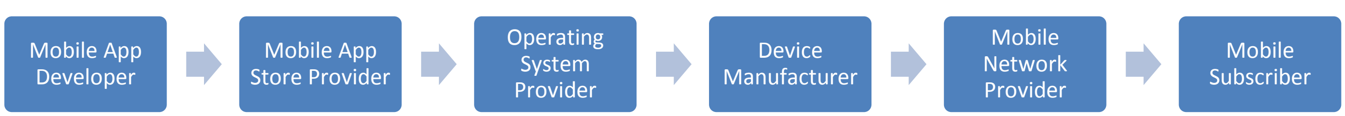


Figure 2- Mobile App Market Roles

Smartphones have become more widely popular and used by mobile users. In June 2013, the United States had 61 percent of mobile users own a smartphone (Sterling 2013). It is expected that the market of smartphones will continue to increase until the year 2017 (Dediu 2014). Smartphones run an operating system, which has become a critical part of the mobile industry, especially due to the disruptive nature of this industry. Different operating systems were analyzed from an ecosystem perspective, discussing ways in which a device manufacturer company such as Apple acts as a device manufacturer and service platform provider at the same time (Lin & Ye 2009). Each mobile operating system provider implements a different business model, trying to capture value (Kenney & Pon 2011). Additionally, these companies have lock-in strategies in order to keep the customer using their products and services. For example, Apple uses the mobile Apple App Store and the handset to capture value, while Google uses its online services as Gmail, search, and Maps. Both Apple and Google try to lock-in customers using their operating system. Apple uses its Handset for lock-in, while Google uses apps like Gmail, maps, and voice for that.

2.2 Mobile app market

Online mobile app stores have indicated a significant impact within the mobile industry, thus leading to the popularity of the smartphones. Recently, Shaughnessy (2013) mentioned that nine out of ten smartphones runs one of the two biggest operating systems either Android or iOS. Consequently, this research focuses on these two app ecosystems of iOS and Android. iOS is an operating system provided by Apple for its mobile products including iPhone, iPad, and iPod. Conversely, several device manufacturers use Android as their operating system. Even though smartphones are not the only type of mobile devices used, it changed the entire mobile market according to Want (2011) and has a promising future.

While an iOS operated smartphone comes along with Apple App Store, an Android operated smartphone can include various app stores. Nevertheless, the most popular app store for Android operated smartphones is Google Play. Despite the existence of several app stores in the market, the research scope of this thesis focuses on these two giant current market leaders in the smartphone industry. The following subsections present a comparison between Apple App Store and Google Play from a point of view of an app provider.

2.2.1 Apple App Store

When Apple first introduced the Apple App Store, it has been considered a huge transformer of the software market (Anthes 2011). The Apple App Store was launched along with the iPhone 3G in July 2008. The store is based on the iTunes platform, which was only used for music. It is currently used by different Apple devices as iPhone, iPod, iPad, and MAC PC, as well as serving global users in 123 countries. Users of Apple have access to the apps of the store by creating an Apple App Store account and registering their information, and then they can download apps from the store. In January 2013, downloads of Apple App Store apps reached 40 billion (Apple 2013). Apps of the Apple App Store are grouped based on the app category, and a special page can be found within the store presenting the top charts. The apps are either free or paid apps with the price ranging from $0.99 to $999.99. However, an app can include in-app purchases for both free and paid, where a user can further pay more in return of getting access to extra features of the app. The Apple App Store includes a Top Charts page that includes three categories of free, paid, and top grossing apps within the store. Appendix A presents screenshots of an iPhone Apple App Store for the featured and top charts. Additionally, there is a featured market in the homepage of the store displaying the best apps or even daily offers on paid apps that can be downloaded for free.

An app developer has to follow several steps in order to join Apple App Store and upload the developed apps on the store. An app provider has to pay a membership fee, and agree on the terms and conditions. Then, Apple takes the final decision of approval for joining as an app developer. Apple shares 30% of the revenue that is generated from an app, whether it is free or paid app. When a developer submits an app to the Apple App Store, Apple has the right to accept or reject the app. It has been an ongoing discussion about the reasons behind rejecting apps, usually Apple does not clarify enough when taking a decision (Jardine 2009). Recently, due to the high level of competition between both stores, Apple decided to release the first review guideline containing the basis of rejection of an app (Apple 2010).

2.2.2 Google Play

Google Play is the current market name for Google Android Market that was launched in October 2008. It operates on any smartphone supporting Android operating system. Currently, Google Play has a more market share compared with Apple in terms of number of apps (Rowinski 2013a). There are other competing app stores for Android devices such as Amazon Apple App Store, SlideMe, GetJar, Mobango, and AppsFire (Cohen 2012). Nevertheless, Google Play remains the market leader and the largest market in the Android app store market attracting more app developers. An app developer needs to pay once in order to subscribe to Google Play, and then accept the agreement terms. Additionally, the app developer needs no approval from Google Play to join the market and become identified as an app developer. Similar to Apple App Store, Google Play shares 30 percent of the revenue generated by the apps uploaded to its store. Unlike Apple, Google is not involved in the process of publishing an app on Google Play. However, Google is allowed to remove apps that do not bind to its agreement after being published. Google Play apps can also be categorized into free and paid apps, in addition to the in-app purchases available for both free and paid apps. It offers the top free, paid and grossing apps. Similar to Apple App Store Featured section, there is new free, new paid and trending apps section. A screen shot from a Samsung Mobile Phone Google Play Store is shown in Appendix B.

2.2.3 Apple App Store vs. Google Play

In January 2013, Apple apps reached 775,000 apps and the competition became tough at that time as Google Play was very close to approach the same number (Rowinski 2013a). In July 2013, Google Play announced the availability of 1,000,000 apps in their store, while Apple mentioned that Apple store reached 900,000 apps in a conference in June 2013 (Rowinski 2013b). In October 2013, Apple announced the availability of one million apps in the Apple App Store (Ingraham 2013). Figure 3 displays the historical count of number of apps in Apple App Store vs. Google Play starting from the launch of both stores to July 2012. The figure indicates the higher growth rate of apps in Google Play compared to Apple App Store. Furthermore, Figure 4 analyzes precisely the growth of Google Play apps count. In conclusion, approximately after July 2013 the app count of Google Play has exceeded the apps published in the Apple App Store. This has been followed by Apple App Store hitting the same number of apps, which is one million, later than Google Play by three months.

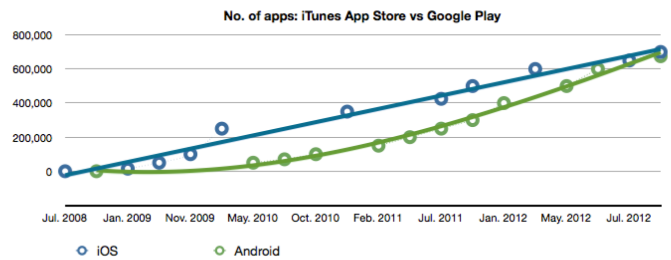


Figure 3 - App Store vs Google Play (De Vere, 2012)

The number of apps only indicates the size of the store, while cannot be used to compare the success of different app stores. Therefore, the success of the app stores can be demonstrated using the number of downloads and revenues per each store. In April 2014, App Annie (2014) announced that Google Play leads by 45 % in terms of app downloads compared to Apple App Store as shown in Figure 5. Nevertheless, Apple App Store still leads in terms of revenue by having 85% more revenue than Google Play (App Annie 2014). Even though Apple App Store is still more successful in terms of revenue, it used to earn 500% more revenue than Google Play in 2013 (Venture beat 2013). This illustrates a decline in the revenue share of Apple compared to Google Play.

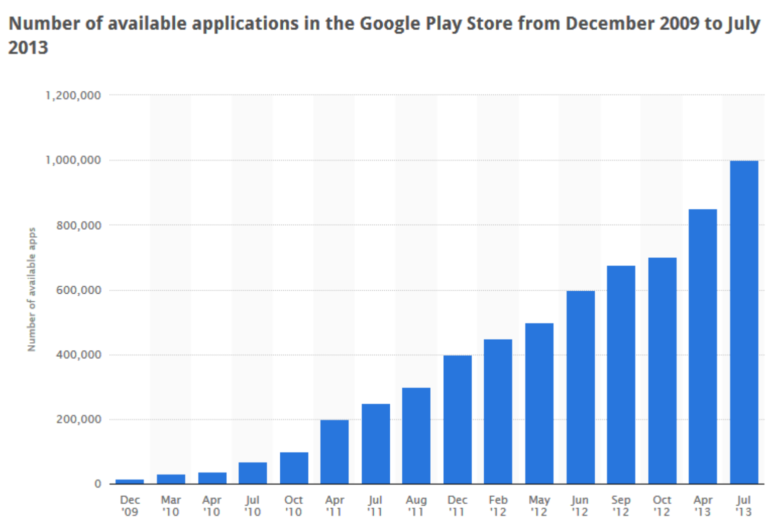


Figure 4 - Number of apps in Google Play (Statista 2014a)

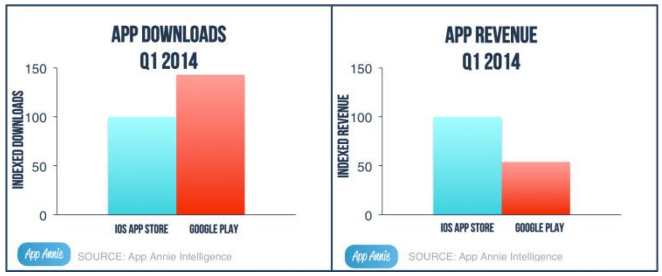


Figure 5 - App Store vs. Google Play (Downloads and Revenue) (App Annie 2014)

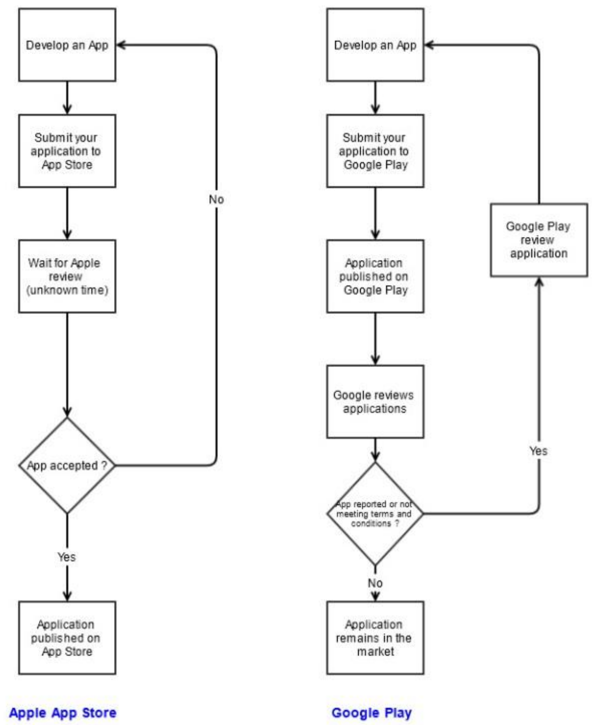


Figure 6 - Approval process in App Store and Google Play (Cuadrado & Dueñas 2012)

One major difference between both ecosystems is the approval process, which is very critical throughout the discussion of this research having a major effect on user demand in both ecosystems. The approval process comparison can be viewed in Figure 6 that was extracted from Cuadrado & Dueñas (2012) and the figure presents it in an abstract level to meet the research question requirements. The figure explains the approval process in both stores, indicating the existence of more restrictions for Apple App Store than Google Play in order to accept submitted apps. In 2009, Apple reported that 95 % of the apps get accepted within 14 days of the submission before being published (Apple 2009). Furthermore, Apple published in the same report on its website that the rejected apps represent 20% of the submitted apps. Recently, Sarno (2012) has claimed the expansion of this percentage to 30% rejected apps out of the submitted ones. On the other hand, Google Play does not face the challenge of reviewing and approving every single app before its public release in the market. However, Google Play faces a major challenge of malicious apps, it even uses an automatic anti-virus scanning system to detect and remove these apps. RiskIQ (2014) recently released in February 2014 a report indicating a 400 percent increase in the number of malicious apps in Google Play. On the other hand, it also denoted a decline in the percentage of removed malicious apps removed by Google from 60% in 2011 to 23% in 2013 (RiskIQ 2014).

As a matter of fact, these two ecosystems are based on different business models. Google Play is an open app store or market, while Apple App Store can be defined as a more controlled closed market. This has led to an interest of several researches covering the comparison both these app stores from a business model point of view. A table of comparison has been presented showing that the Apple App Store is a centralized ecosystem while Google Play is a distributed one (Rao & Jimenez 2011). Recently, companies started to favor the use of open innovation as part of their business model in the ICT industry. The question remains unclear whether Apple App Store will succeed in keeping its closed platform competing with the strong open innovation trend in the market (Yamakami 2011).

2.3 Network Effects

Network externalities theory was first developed by (Katz & Shapiro 1985, Oren & Smith 1981, Rohlfs 1974). It is defined as the value added by one user to other users of the product or service. The service becomes more valuable as the number of users increase. This theory can be easily demonstrated using different types of products or services that rely on a network of people that can be easily illustrated using a telephone service a typical example of network effects. The value of the telephone service increases as the number of subscribers of this network increases. The idea behind the service is based on networking the community together.

There is a definite difference between network effect and network externality that has been discussed and adopted by Katz & Shapiro (1994). A network effect cannot be always defined as an externality. If the network effect is internalized by the market while not having an effect on the market, this cannot be defined as an externality (Liebowitz & Margolis 1994).

Concerning the positive network effect several laws were introduced in order to help in modeling the network externality. Sarnoff's law defining the value of a broadcast network is directly proportional to the number of users. Furthermore, Metcalf's law states that the value is proportional to the square number of the users. Recently, Reed (1999) introduced Reed's law which defines the value to be proportional to twice the number of users in the network. KK-law suggested offering platforms to the customer groups, which are inexpensive to the customers (Kilkki & Kalervo 2004). All of these laws aim to understand the effect of the number of users on the success of a product or service.

The concept of the types of network effects was introduced by (Katz & Shapiro 1985, Farrell & Saloner 1985). There are direct and indirect network effects. Direct network effects, which are stronger than the other type, involve the physical effect that generates a higher value for the network. Direct network effect examples are the fax, telephone services, and recently Facebook is considered as one obvious example. The other type is the indirect one, which represents the complementary nature of different products in a network, and as a matter of fact a user adds value to the entire network.

Network effects are significant in the context of mobile apps. The apps act as services available on the online stores behaving in a similar way as those previous technologies. Successful apps would need to pass the critical mass in order to become popular and gain more customers. In fact, customers decide to join a certain phone ecosystem based on the critical mass of apps available. The concept of network effects has a significant effect on the business of the app developer. The app provider needs to understand the critical mass in order to set prices accordingly and gain more popularity for the published apps.

2.4 Long Tail vs. Superstar

Long tail is defined as the portion of the distribution having a large number of occurrences far from the head or central part of the distribution (Bingham & Spradlin 2011). The use of the long tail in business has been an area of research interest, Anderson (2014) (2016) first introduced long tail in business. The theory basically focuses on how the niche markets behave in internet based products or services. Throughout the research it has been proved that the areas of the niche market are a good source of revenue in internet based selling. This has also reflected on app usage displaying a U shape curve behavior of users, where there is high demand for most popular and niche, not for the middle apps (Verkasalo 2009). Furthermore, a mathematical model for long tail has been introduced to help in understanding and analyzing diverse long tail phenomena (Kilkki 2007). The previous long tail researchers presented analyses of different types of internet selling such as Amazon, Symbian App market, and books.

A research verifies that Google Play is not a long tail market (Zhong & Michahelles 2013). No other research focused on long tail behavior in current mobile app platforms like Google and Apple. It argues that the Google Play revenue mostly comes from the apps having high hits or ranking. This phenomenon has been defined as Superstars. This research recommends developers to focus on hit apps, so an app has to get exposed in order to be successful and attracts more revenue. Online video sales have experienced the same type of behavior and higher percentage of the revenue comes from higher ranked videos (Elberse & Oberholzer-Gee 2007). The phenomenon of Superstars was first introduced by Rosen (1981).

According to previous research, contradicting evidence has been given regarding the Long Tail behavior of digital markets. Theoretical models have been developed and empirical evidence was used to prove the validity of these models. However, during the writing of this thesis, there is no “Long Tail vs. Superstars” recent research about the current market behavior for mobile app markets. The mobile app market has changed rapidly, especially after Nokia’s decrease in market share in 2011 (Dediu 2011, Dediu 2012). The significance of the top hit apps from a perspective of an app provider is the same if both theories are taken into consideration. Therefore, this thesis provides recommendations for app providers when launching an app into any of Apple App Store and Google Play, in order to become a successful app.

2.5 Diffusion

Diffusion is defined as the process by which an innovation is communicated through certain channels over time, among members of a social system (Rogers 2003). It has been an area of interest in the field of telecommunications, especially due to the disruptive nature of the mobile telecommunications industry. The mobile industry offers digital products is considered as innovative IT products (Rogers 2003). Rogers introduced categories of people in order to describe innovation. These categories include innovators, early adopters, early majority, late majority and laggards. The above mentioned categories indicate the users who start using a new innovation sorted by time of use. The innovators for instance, are the first 2.5 percent of users starting to use the technological innovation (Rogers 2003).

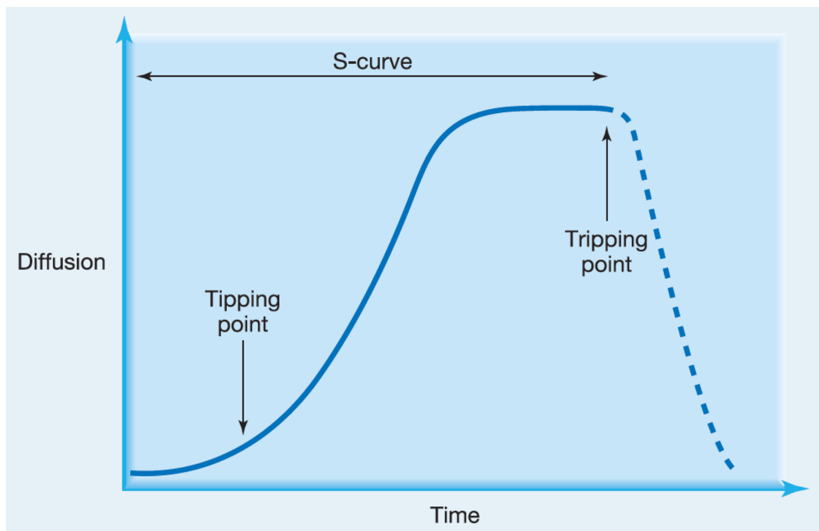


Figure 7 - The diffusion S-curve (Sood & Tellis 2005, Johnson;Scholes;& Whittington 2008)

Innovations follow the S-curve pattern of diffusion after being launched (Sood & Tellis 2005). Figure 7 shows the diffusion S-curve, and the critical points when introducing an innovation. The figure highlights two important points, because of their importance in a life cycle of a technological innovation. The first point is the tipping point, that resembles the critical mass or amount of user needed so that a successful innovation can benefit from network effects (Gladwell 2000). It is of great importance for an innovation manager to achieve this tipping point in order to attract more and more users for the innovation and make it more successful (Brown 2005). If demand drops suddenly, this is defined as the tripping point of innovation. The lifetime of an innovation depends on the nature of the product or service. In mobile app industry, the lifespan of an app is short, Nathan Ooley, President of Appmosphere Inc. mentioned that the average lifespan is around 14 months (Wolonick 2013). Therefore, it is important for an app provider to understand this curve and note the importance of the critical points. Apps are continually uploaded on the market shaping a competitive market and more effort is needed to have a Superstar app.

2.6 Clustering effect

Recently, the advances in the ICT field have led to developing different types of services. This has led ICT systems collecting enormous amount of data. Additionally, the rate of data generation has been increasing lately leading to the popularity of Big Data analysis, which refers to the process of finding meaningful information behind this massive amount of data. Additionally, organizations have realized the importance of the opportunities that lie behind this data. Even governments, universities, many others have realized the importance of making use of this Big Data. However, the return expected to be even more than the huge investments needed in such type of projects.

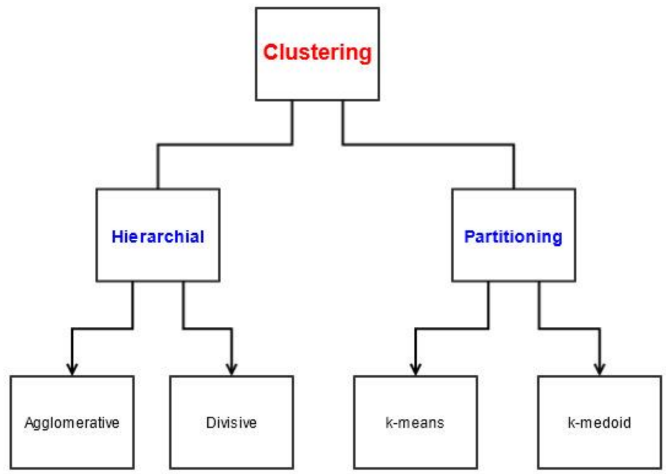


Figure 8 - Types of clustering

The increase in the amount of data generated by the current technologies creates a need for data analysis, as mentioned earlier. Clustering is one of the data mining methods used for organizing data to facilitate qualitative data analysis. Clustering is the process of grouping various objects into smaller groups having similar features. Cluster analysis was first mentioned by Driver & Kroeber (1932). Clustering algorithms are classified, as shown in Figure 8, into hierarchical and partitioning methods (Kaufman & Rousseeuw 1990, Spath. 1985). Hierarchical clustering can be divided into agglomerative and divisive clustering. Agglomerative clustering basically starts with a cluster number equal to the number of objects and then starts merging these clusters together. Divisive is the other way around, where all objects are put into one cluster, then starting to split this cluster into two main clusters. The process continues until no more clusters are required. Objects are classified to the most suitable k-clusters using the partitioning method.

Collaborative tagging is to allow internet users to manage, share and annotate online resources (Mathes 2004). The result of a complex network of users, and resources along with collaborative tag creation and management is a collection of annotations that is defined as folksonomy (Mathes 2004) . Tags are useful in a sense of categorizing the available resources (Millen, Feinberg & Kerr 2006). Several researches showed that clustering is useful in collaborative tagging (Heymann & Garcia-Molinay 2006, Begelman, Keller & Smadja. 2006). The coherent clusters of related tags are formed using clustering algorithms (Begelman, Keller & Smadja. 2006).

Petsas, Papadogiannakis & Polychronakis (2013) present a model examining four different app stores SlideMe, 1Mobile, AppChina, and Anzhi using the clustering method. It uses the affinity factor to examine elements belonging to the same category. The results confirmed the availability of clustering effect in the app stores. This means that users usually download the next app from the same category of the previous one. This explains some of the user behavior within an online app market.

2.7 Related work

This research examines previous related work regarding the analysis of mobile app stores. A list of the available research is mapped in Table 1 and is categorized with colors based on the discussed platform. It has been noticed that numerous research focused on addressing the business ecosystem behind both platforms of Apple and Google Play. Several efforts focused on the Apple App Store due to its popularity. Little research focused only on studying the Google Android market. A considerable amount of research included other app store platforms or general perspective about mobile digital stores.

Related Work	Google	Apple	Other
(Carare 2012)		y	
(Lee & Raghu 2011)		y	
(Pagano & Maalej 2013)		y	
(Ayalew 2011)		y	
(Wang & Wang 2013)		y	
(Kimbler 2010)		y	
(Yamakami 2011)		y	
(Zhong & Michahelles 2013)	y		
(Petsas, Papadogiannakis & Polychronakis 2013)	y		
(Want 2011)	y		
(Ghose & Han 2012)	y	y	
(Garg & Telang 2013)	y	y	
(Skogsberg 2013)	y	y	
(Hyrynsalmi et al. 2012)	y	y	

(Kokabha 2012)	y	y	
(Fredholm & Gunnarsson 2013)	y	y	
(Sawant 2010)	y	y	
(Rao & Jimenez 2011)	y	y	
(Eaton, Elaluf-Calderwood & Sørensen 2010)	y	y	
(Holzer & Ondrus 2011)	y	y	y
(Cuadrado & Dueñas 2012)	y	y	
(Kim 2012)	y	y	
(Lim & Bentley 2013)			y
(Liu et al. 2013)			y
(Chevalier & Goolsbee 2002)			y
(Gonçalves, Walravens & Ballon 2010)			y
(Tuunainen & Tuunanen 2011)			y
(Tuunainen, Tuunanen & Piispanen 2011)			y
(Ho & Syu 2010)			y

Table 1- Related Work Map

Previous research focusing on Apple App Store reported conclusions that can be further used throughout this thesis. These conclusions were either market analysis oriented or app user profile oriented or even app provider profile related. Market analysis from a perspective of a mobile operator has been discussed by Kimbler (2010), while taking into consideration the effect of introduction of Apple ecosystem. It concludes the difficulty for mobile operators to reach new revenue sources from the mobile app market and compete within the Apple ecosystem. This research has been followed by another one discussing the key factors behind the success of the tightly controlled business model offered by Apple ecosystem (Yamakami 2011). Furthermore, the market success of an app, which is related to the number of downloads of this app, has been affected by the positive feedback of an app (Pagano & Maalej 2013). Next, the user profiles of Apple App Store have indicated a will to pay $4.5 extra for a bestseller app than an unranked one (Carare 2012). Moreover, Lower income consumers purchase apps more than higher income ones (Ayalew 2011). Finally, an app provider profile has been researched attempting to identify the most successful profiles. It denotes that app providers offering a combination of free and paid apps have more sales than the ones offering only paid apps (Lee & Raghu 2011). Other research stressed on maintaining a high quality app in order to attract a build a good source of revenue (Wang & Wang 2013).

This thesis mapped little research analyzing Google Play separately. Google Play can be installed along with any Android operated smartphone. The concept of Android, as an open source operating system, has dramatically changed the mobile industry, specifically regarding the smartphone market. Want (2011) addresses the changes and effects of the introduction of the Android open platform. A recent study by Zhong & Michahelles (2013) has revealed that Google Play is a Superstar market, and developers should focus on having hit apps, which supports the same conclusion of a similar research done by Wang & Wang (2013) for Apple App Store. Regarding the pricing of apps, free apps should charge $0.21 per download in order to be equivalent to the average revenue of a paid app as recommended by Petsas, Papadogiannakis & Polychronakis (2013). Finally, the Android store has validated the clustering of apps based on categories, which will be further used as a base for the analysis throughout this thesis (Petsas, Papadogiannakis & Polychronakis 2013). This helps directing the research towards an objective analysis of the success of categories.

Subsequent to the emergence of Google Play, research became more directed to compare these two ecosystems. However, this thesis shows that most of the comparison between the app stores focused on the business model of the functioning ecosystem for Apple and Google. There has been detailed analysis of Google and Apple platforms from the perspective of an app provider. The research shows the value gained by the developers when joining the platform, providing a guide for the developer to differentiate between joining any of the ecosystems (Cuadrado & Dueñas 2012). Some research analyzed the complementary effects of both ecosystems, while others compared different business models of apps within the app store. Sawant (2010) shows the mobile business model of apps for different platforms from a perspective of an app provider without analyzing any details about demand behavior within the platforms itself. Kokabha (2012) focused on modeling the business ecosystem of the app stores, revealing different stakeholders in the ecosystem and the value network chain. A comparison was presented between the two business models and their success factors, stating that the motivation for joining the ecosystem is usually the most challenging part of the business model, identifying the benefits and drawbacks of managing each platform (Rao & Jimenez 2011). One of the highlighted research topics compares the value network of both Apple App Store and Google Play, while introducing different actors in the ecosystem while highlighting the difference in dynamics between both ecosystems (Eaton, Elaluf-Calderwood & Sørensen 2010). Furthermore, Holzer & Ondrus (2011) presented a comparison between different platforms from a perspective of a developer and the value gained by the developer from each platform.

Other researchers have analyzed demand for app stores using combined figures of both datasets from Apple App Store and Google Play, rather than analyzing figures separately extracted from each app store (Hyrynsalmi et al. 2012, Ghose & Han 2012). Research results reveal that multi-homing strategy is not an important critical factor affecting the sales of apps sales of an app provider (Hyrynsalmi et al. 2012). Moreover, the app demand has indicated a positive correlation for lifestyle and gaming apps, while being negative in case of multimedia and educational apps (Ghose & Han 2012). Skogsberg (2013) proposed a web platform for app developers, in order to collect data about the stores and further analyze it.

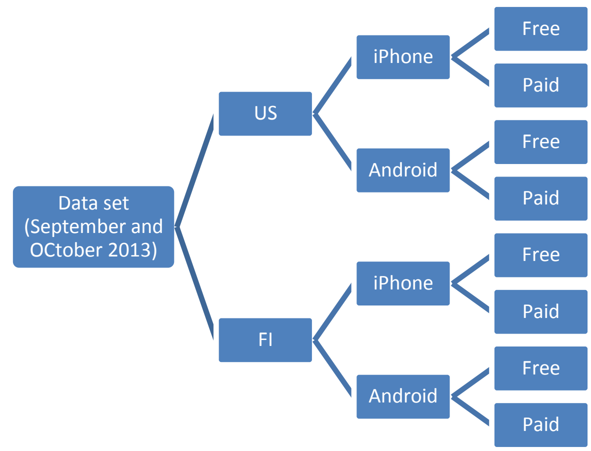
Other research aims to identify the user preference for mobile app stores. The user preferences denote a variation among different app stores or even geographical market, having a major effect on the demand of apps. Users of Google Play were found to be less likely to purchase apps, while an Apple user is expecting more benefits from the app and is more likely to buy more apps from the store (Kim 2012). Additionally, Geographic market has resembled a great influence on the demand behavior of an app, while Free-in-app has become the most common business model (Fredholm & Gunnarsson 2013). Finally, the first iPhone ranked paid app has 150 times more downloads than an app ranked as 200 in the top list (Garg & Telang 2013). While the first iPad ranked paid app has 120 times more downloads than an app ranked as 200 in the list (Garg & Telang 2013).

The last category of research mapped by this thesis examines other types of app markets or even similar mobile digital markets. In an attempt to model the diffusion of mobile digital content, Gompertz model was used to analyze the diffusion of one type of mobile digital content in the Chinese market (Chevalier & Goolsbee 2002). Furthermore, Price sensitivity of other online mobile markets shows different price elasticity between different online platforms (Lim & Bentley 2013), revealing Barnes and Noble being more prices elastic than Amazon. This has led to research aiming to identify successful apps, Entertainment has been identified as one of the main revenue sources in which users aging from 20-29 are motivated to download (Ho & Syu 2010). Additionally, Success of an app has been related to the algorithm used by the app ecosystem, simulations shows that the speed at which the content is updated affects the success of an app (Liu et al. 2013). As for the mobile industry, there has been research suggesting different possibilities for operators to gain an important role in the mobile app industry and discussing the benefits of different platforms if applied (Gonçalves, Walravens & Ballon 2010). Tuunainen & Tuunanen (2011) present a model for n-Sided markets for analyzing ICT Intensive Services Innovations. IISIn model has been applied to a comparison between Nokia OVI Store and Apple. The reasons were revealed behind the success of Apple (Tuunainen, Tuunanen & Piispanen 2011).

3 Methodology

This chapter introduces the methods and tools used by the demand analysis for mobile app stores. It provides an overview of procedures employed to reach the results of this study, while enabling the research community to make use of these results and further develop them. These tools and methods are presented as a summary to meet the aims and scope of this research.

This study includes data collected from Apple App Store and Google Play, the nature and structure of it is described later throughout this chapter. Furthermore, this chapter demonstrates the analysis tools used for examining the data for instance MATLAB and AgenaRisk. Finally, it explains Bayesian Network briefly as a tool used for building the Success Model for mobile app stores. This chapter summarizes the methods used to build the Success Model for mobile app stores, which will be clearly presented throughout the following chapters.



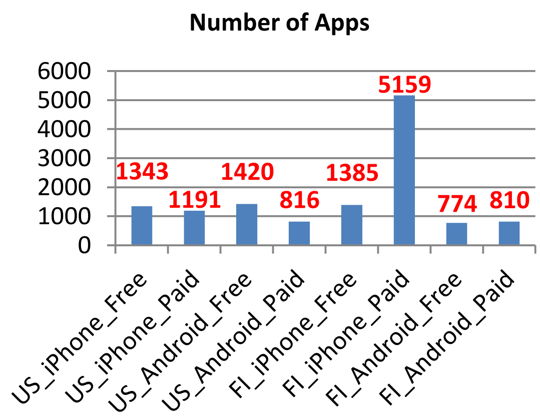


Figure 9 – Dataset map

3.1 Data Description

The dataset contains a list of top apps for iPhone Apple App Store and Google Play, and available for both United States and Finnish markets including various app types. The data has been collected for two months, September and October 2013. It must be noted that later in this thesis if the term iPhone is mentioned it refers to the Apple App Store, as well as Android referring to Google Play.

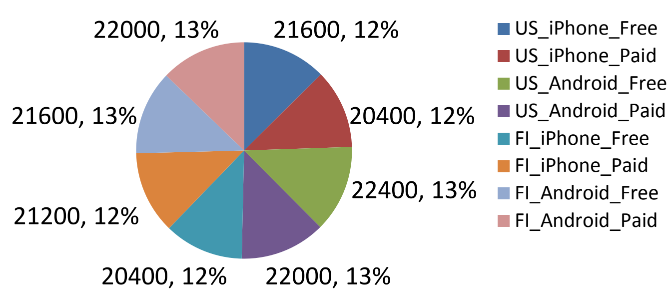


Figure 10 – Dataset records percentage map

As shown in Figure 9, apps are categorized as free or paid apps. Figure 9 also displays the number of apps mapped in each store. There is no description for Free-in-app or Paid-in-app purchases. Therefore, this research analyses the demand based only on these two app types. Additionally, Figure 10 shows the number of dataset records available. For example, free iPhone dataset records for one day are the list of top 400 apps in each store. These count as 400 dataset records, then the count of dataset records continues for the following days even if the same app is mapped among the top 400 list of the following days. Table 2 maps the categories for both stores using the following numbering for each category. This mapping is based on the information provided by the dataset for this study.

Number	Name of Category	Number	Name of Category
1	'Games'	37	'LIFESTYLE'
2	'Sports'	38	'Transportation'
3	'Photo & Video'	39	'SHOPPING'
4	'Social Networking'	40	'Cards & Casino'
5	'Music'	41	'Casual'
6	'Utilities'	42	'Shopping'
7	'Navigation'	43	'NEWS_AND_MAGAZINES'
8	'Reference'	44	'TOOLS'
9	'Entertainment'	45	'WEATHER'
10	'Productivity'	46	'RACING'
11	'ENTERTAINMENT'	47	'Media & Video'
12	'Lifestyle'	48	'Photography'
13	'SPORTS_GAMES'	49	'FINANCE'
14	'Travel'	50	'TRAVEL_AND_LOCAL'
15	'Book'	51	'MEDIA_AND_VIDEO'
16	'News'	52	'PERSONALIZATION'
17	'Health & Fitness'	53	'Sports Games'
18	'Food & Drink'	54	'Racing'
19	'Finance'	55	'News & Magazines'
20	'Catalogs'	56	'CARDS'
21	'Business'	57	'COMMUNICATION'
22	[]	58	'Weather'
23	'Social'	59	'SOCIAL'
24	'Music & Audio'	60	'HEALTH_AND_FITNESS'
25	'Tools'	61	'TRANSPORTATION'
26	'Communication'	62	'BOOKS_AND_REFERENCE'
27	'BRAIN'	63	'EDUCATION'
28	'Arcade & Action'	64	'PRODUCTIVITY'
29	'Education'	65	'SYSTEM'
30	'CASUAL'	66	'PHOTOGRAPHY'
31	'Travel & Local'	67	'LIBRARIES_AND_DEMO'
32	'Personalization'	68	'Comics'
33	'Brain & Puzzle'	69	'BUSINESS'
34	'SPORTS'	70	'Medical'
35	'MUSIC_AND_AUDIO'	71	'Libraries & Demo'
36	'ARCADE'	72	'COMICS'
		73	'UNKNOWN'

Table 2 - Original category list of the dataset

3.1.1 New mapped categories

Table 3 describes the new mapping of the categories generated to allow comparing both app stores together. This mapping takes into consideration the nature of apps mapped below for each of the categories, enabling comparison between different app stores. In addition, Figure 11 examines the availability of data at a certain day for all stores. The data is not complete for all days for each store. Therefore, it is vital to map the availability before starting to analyze the data.

Category	Number	List of Aggregated Categories
'Games'	1	Games','SPORTS_GAMES','BRAIN','Arcade & Action','CASUAL','Brain & Puzzle','ARCADE','Cards & Casino','Casual','RACING','Sports Games','Racing','CARDS'
Sports'	2	Sports','SPORTS'
'Photo & Video'	3	Photo & Video','Media & Video','Photography','MEDIA_AND_VIDEO','PHOTOGRAPHY','LIBRARIES_AND_DEMO','Libraries & Demo'
'Social Networking'	4	Social Networking','Social','Communication','COMMUNICATION','SOCIAL'
'Music'	5	Music','Music & Audio','MUSIC_AND_AUDIO'
'Utilities'	6	Utilities','Catalogs','Tools','Personalization','TOOLS','PERSONALIZATION','SYSTEM'
'Entertainment'	7	Entertainment','ENTERTAINMENT'
'Productivity'	8	Productivity','PRODUCTIVITY'
'Lifestyle'	9	Lifestyle','Food & Drink','LIFESTYLE','SHOPPING','Shopping'
'Travel'	10	Travel','Navigation','Travel & Local','Transportation','TRAVEL_AND_LOCAL','TRANSPORTATION'
'News'	11	News','NEWS_AND_MAGAZINES','News & Magazines'
'Health & Fitness'	12	Health & Fitness','HEALTH_AND_FITNESS'
'Finance'	13	Finance','FINANCE'
'Business'	14	Business','BUSINESS'
Education'	15	Education','EDUCATION'
'Weather'	16	Weather','WEATHER'
BOOKS_AND_REFERENCE'	17	BOOKS_AND_REFERENCE','Book','Comics','COMICS','Reference'
'Medical'	18	'Medical'
'UNKNOWN'	19	UNKNOWN',[]

Table 3 - New category list after mapping

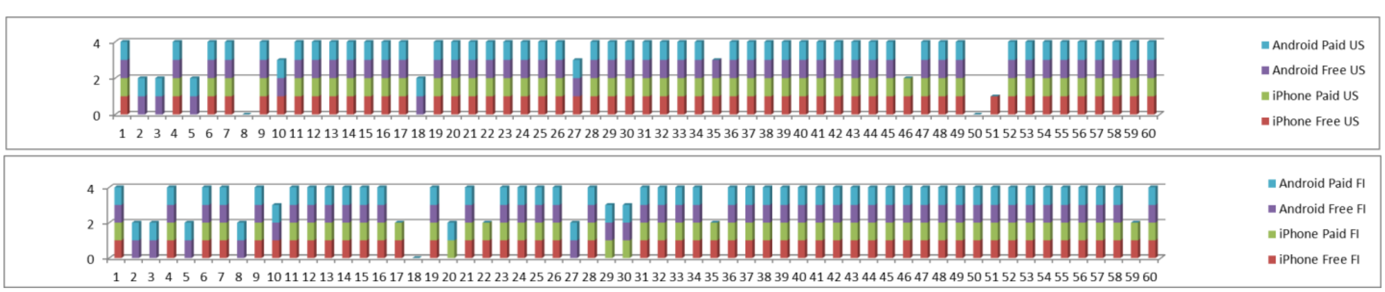


Figure 11 - Data availability per day

3.2 Data preprocessing

This thesis analyzes the data using MATLAB as a main tool. MATLAB scripts are used for preprocessing the data and organizing it. The preprocessing includes mapping the string values of the data to numbers, and mapping of the categories as mentioned earlier within this chapter. The data has some gaps in between as mentioned earlier in Figure 11; thus, if data for a certain date is missing, the rank of apps is assumed to be the same as the previous day. On the other hand, the rank is assumed to be 401 if no record has been found, while the data of that certain date is available. Likewise, the number of downloads is assumed to be the same as previous day of the data is missing; otherwise it will be equal to zero.

3.3 Steps of data analysis

Figure 12 shows the steps of data analysis followed throughout this research. The data provided is preprocessed as described earlier in this chapter using MATLAB. Furthermore, the outcome of this preprocessed data is further evaluated to represent graphically the results using different types of figures such as plots, pie charts and histograms. This evaluation identifies the probability of success of an app that can be further used as an input for the AgenaRisk tool, which will be explained later in this chapter, in order to build the Success Model. Finally, MATLAB is used to calculate the success indicator of an app using previously calculated probability of success.

The data evaluation using the MATLAB included basic descriptive statistical analysis, aiming to understand the available dataset. Descriptive statistics is used to represent quantitatively the app categories, cumulative downloads, dynamics of app rank changes, and the number of days for app survival. Furthermore, these results are used to identify the probability of success of an app within a mobile app store, which will be further used as shown in Figure 12.

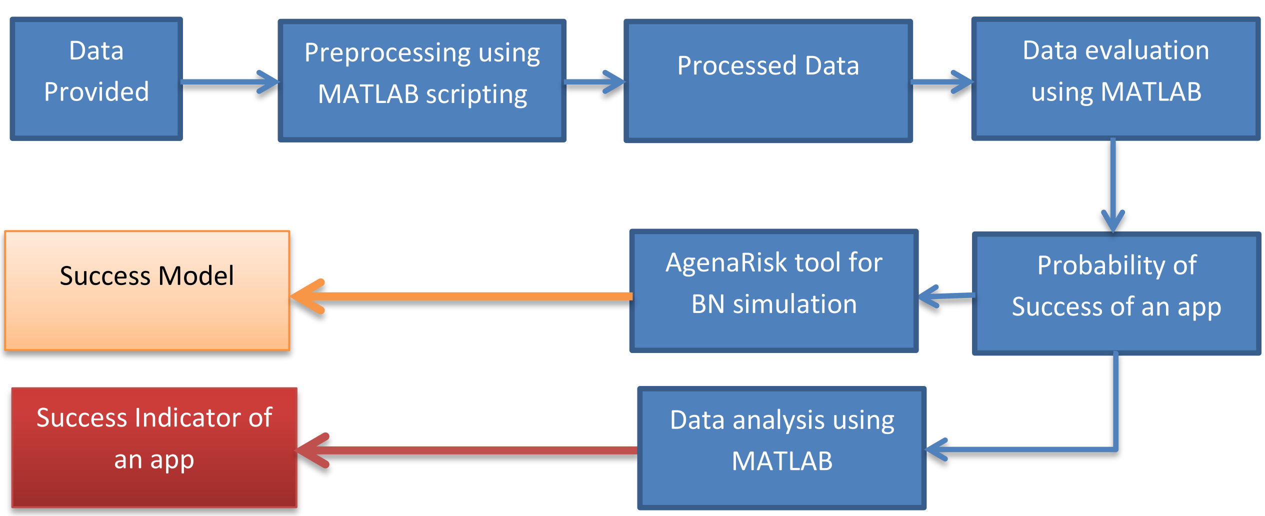


Figure 12 - Steps of data analysis

3.4 Bayes' theorem, Bayes networks, and tools

Bayes' Theorem is an important law in probability laws and statistics specifically regarding conditional probabilities. It states, as displayed in Equation 1, that the probability of a random variable X given Y is equal to the probability of Y given X multiplied by probability of Y divided by probability of X.

$$\mathbf{P(X|Y)} = \frac{\mathbf{P(Y|X) \times P(Y)}}{\mathbf{P(X)}} \quad \text{... Equation 1}$$

Bayes' Theorem reveals that P(X|Y) =P(Y|X) only if P(X) and P(Y) are equal. This is an important note to consider and realize that P(X|Y) and P(Y|X) cannot be the same unless this condition is satisfied.

<u>Simple Illustrative Example for Bayes' Theorem</u>

This example assumes that P(F) is the probability that a person joining a marathon is fit, while P(W) is the probability that a person is going to win the Marathon. The marathon has 10,000 participants, and the first 200 to reach the finish line are considered winners and earn a prize. An initial marathon test revealed that only 70% of the participants are completely fit to complete the marathon. By the end of this marathon, 185 out of the 7000 fit participants were on the top winner list. What would be the probability that the person is fit person given that this person has won in the marathon? The answer to this question is quite obvious it would be equal to $\frac{185}{200} = 0.925$. The Bayes' Theorem is applied using Equation 1 as shown below.

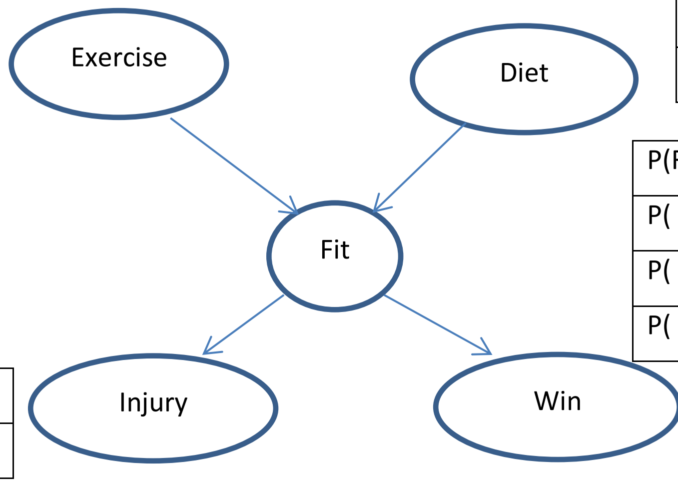
$$\mathrm{P(F|W)} = \frac{\mathrm{P(W|F) \times P(F)}}{\mathrm{P(W)}} = \frac{\frac{185}{7000} \times \frac{70}{100}}{\frac{200}{10000}} = \frac{185}{200} = 0.925$$

This example justifies the concept of the Bayes' Theorem, while showing the concept that P (W|F) is not typically equal to P (F|W). Bayesian Networks (BN) is a probabilistic graphical model, which was introduced in the late 1970s and used to represent complex systems based on direct acyclic graph (DAG). DAG is a graph with no direct cycles between its variables. BN is an analytical tool used to reveal the reasoning behind an uncertainty factor or variable, in order to predict the consequences using dependencies among its random variables (Pearl 1988). The BN presents different random variables and their conditional dependencies using the network. Conditional dependencies between the variables are based on the Bayes' Theorem mentioned above. Dependencies between variables are identified by arches, which are used to link between them. These links indicate a bidirectional flow of dependencies, although the arches are graphically pointing towards one direction only (Pearl & Russel 2001).

This thesis illustrates the BN using an example, which will be used further to examine its properties. This example addresses concepts clearly using the same example of a person joining a marathon. The BN shows the probability of having a fit runner depends on following a regular exercise and a healthy diet. Similarly, the probability of winning depends on the fitness of the runner. Additionally, the probability of having an injury depends also on the fitness. Figure 13 is the BN for this example and represents the conditional probabilities between different variables based on their dependencies.

P(E=T)	0.6
P(E=F)	0.4

P(D=T)	0.7
P(D=F)	0.3



P(F=T|E=T,D=T)	0.9
P(F=T|E=T,D=F)	0.8
P(F=T|E=F,D=T)	0.6
P(F=T|E=F,D=F)	0.4

P(I=T|F=T)	0.3
P(I=T|F=F)	0.6

P(W=T|F=T)	0.7
P(W=T|F=F)	0.4

Figure 13 - BN example

Figure 13 demonstrates a typical BN consisting of nodes representing the random variables used by the model. Each random variable can have discrete, continuous, or even hybrid values which are a mixture of both. The values of nodes in this example use discrete, precisely all of the variables are represented by Boolean values using True or False. This example shows that the probability of a runner exercising regularly and following a diet is 0.6 and 0.7 respectively. The probability of a runner is fit is calculated using the given information regarding exercise and the diet. Similarly, the probability of winning or injuries both depends on the fitness of the runner.

The previous example can be used to illustrate different representations of conditional probability within a BN. First, the relation between E, F, and W presents a causal relationship. Given evidence for the value of F, the W and E are conditionally independent. Second, the common cause relation that is revealed by the links between I, W, and F. The probability of a fit runner depends on both exercise and diet. W and I are conditionally independent given the value of F. Finally, the common effect relation can be illustrated using the BN of E, D and F, where both E and D have an effect on F. Unlike the common cause BN, E and D are conditionally independent given no information regarding F. Once F has given evidence, E and D become dependent, as the evidence is transmitted to both. It is important to understand the concept of conditional independencies, which will further affect the transmission of evidence between nodes within the BN. In conclusion, if two nodes become conditionally independent when given a value, this indicates the probabilities of these nodes will not be transmitted to each other through the intermediate node and both become independent on each other.

The joint probability distribution of Bayesian Networks is presented by the following formula:

$$P(X_1, X_2, X_3, \dots, X_n) = \prod_i P(X_i | Parents(X_i))$$

When applying this equation for joint probability distribution to example presented, the following formula is concluded:

$$P(E, D, F, I, W) = P(E) \times P(D) \times P(F|E, D) \times P(W|F) \times P(I|F)$$

If a BN is given evidence regarding a certain variable, it changes the uncertainty of another variable within the network. This can be exploited for various apps most importantly modeling for measuring uncertainty. Modeling can be first used to predict the effect of certain changes throughout the network. Finally, BN are used to find out reasoning behind a certain event, by tracking back the network and understand the reason behind its occurrence.

AgenaRisk is an analytical software tool used for predictions and risk assessment. It is based on the BN methodology and it provides an easy graphical user interface to build BN and simulating it. The tool uses a propagation algorithm for simulating the BN, while the user only interacts with the graphical user interface. It is also regarded as providing accurate results for the simulation of the BN (AgenaRisk 2004). This thesis uses the free version of AgenaRisk for the analysis. The tool is used to create nodes of the BN model and values were entered manually to each Node Probability Table (NPT) and then simulation is performed. The simulation was run based on various scenarios, given the option provided by AgenaRisk to run several scenarios. Each scenario presents results based on given evidence for a node or several nodes before running the simulation. Results include probability histograms for each node, revealing the probability of occurrence of an event.

4 Results

The previous chapter presented the methods applied to conduct this research. Thus, this chapter presents the results of the Thesis. These results are presented through statistical figures, indicating a graphical representation of the processed data subsequent to running the MATLAB scripts. They report the behavior of the mobile app market for both stores.

4.1 Categories visualization

This section introduces the category distribution of both stores for US and Finnish markets. There are 19 categories used to facilitate the comparison as described by the previous chapter. Figure 14 shows the number of apps on the top list for each category. Figure 15 indicates the percentage share of each category within the app store top list. Some figures include percentage share of a category of apps in a market, it must be noted that the percentages are calculated in reference with the market where the category is listed.

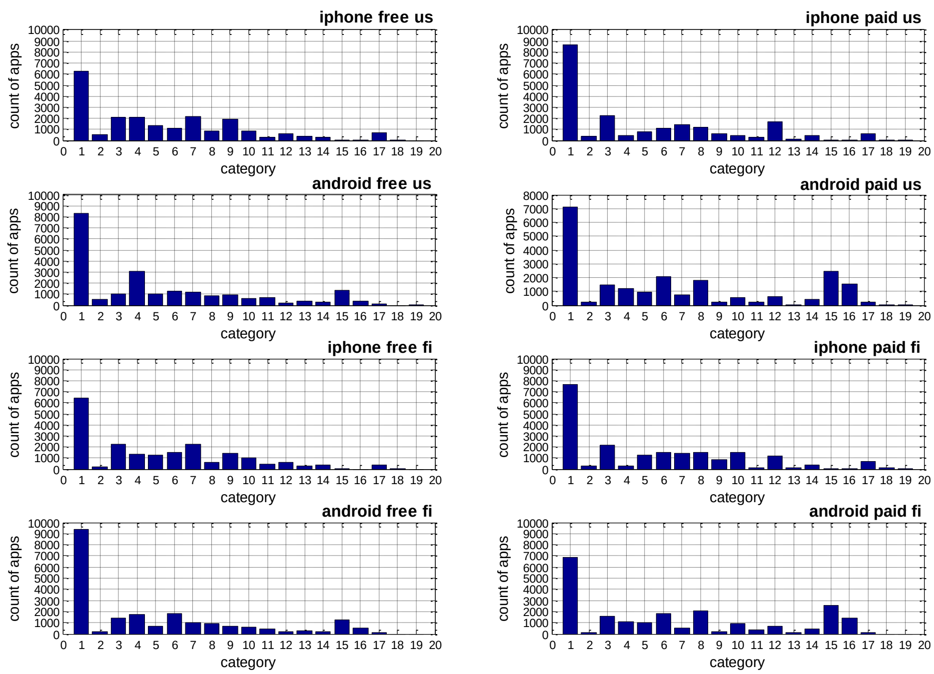


Figure 14 - Category count

4.1.1 US market

iPhone market displays Games as a leading category for the free market with 29% of the cumulative app downloads as illustrated by Figure 15, followed by Entertainment 10%, Photos & Video 10%, Social Networks 10%, and finally Lifestyle as 9%. Furthermore, Games has a higher percentage share in the iPhone paid market than the free market, in terms of percentage count in the top listed apps. Similarly, Social Networking apps are more popular in the paid store in comparison with the free store, sharing only with 2% in the iPhone paid list, while it holds 10% of the iPhone free apps. The top five listed categories in the paid iPhone market are Games 42%, Photos & Videos 11%, Health & Fitness 8%, Entertainment 7%, and Productivity 6%. This list also reveals the increase in popularity of the Health & Fitness category in the iPhone paid market.

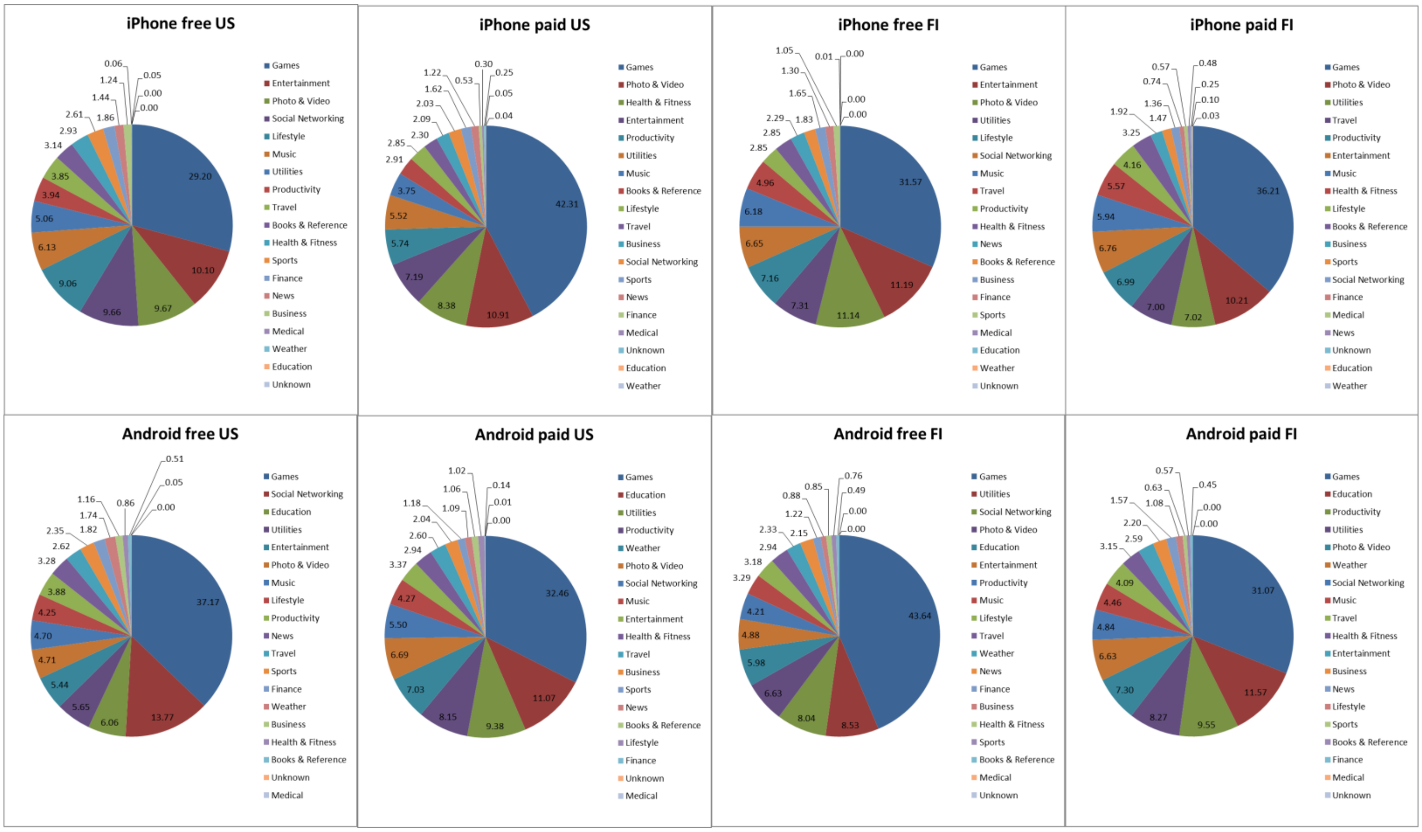


Figure 15 - Category percentage share

Android market confirms the success of Games in both free and paid markets reflected in Figure 15. It displays the free market share lead by Games 37%, followed by Social Networking 14%, Education 6%, Utilities 6%, and Entertainment 5%. Surprisingly, the Education category does not play an important role in the iPhone market, while representing 6% in free Android market, and 11% in the paid one. Furthermore, Weather apps also have higher popularity in the Android paid market compared to any other. On the other hand, some categories did not receive much of applause in the Android market. These categories include Health & Fitness, and Books & References having lower popularity in the Android market compared to the iPhone market. Additionally, Photos & Video represent higher percentage count share within the iPhone market compared to the Android. As for the Android paid market, Games are as usual on the top list with 32%, followed by Education 11%, Utilities 9%, Production 8%, and Weather 7%. By observing both app stores, Android has a different behavior in the paid market compared to the iPhone market. This can be noticed by the decrease of the Games percentage share by 5% in the Android paid store relative to the free, while the iPhone store notices an increase by 13%.

4.1.2 Finnish market

Games category still preserves its position on the top list of the iPhone Apple App Store leading by 32% in the free apps as shown in Figure 15, followed by Entertainment 11%, Photos & Video 11%, Utilities 7%, and Lifestyle 7%. Games represent higher 4% in the paid top listed apps compared to free iPhone apps. There is no major difference in the behavior of the iPhone market in Finland compared to US in terms of share of a category of apps in the top list. However, Travel apps seem to be more appealing the paid Finnish market than the US. On the other hand, categories like Health & Fitness are more appealing in the Apple App Store for the US paid market than the Finnish one. Thus, Figure 15 reveals the top paid iPhone list starting with Games 36%, Photos & Video 10%, Utilities 7%, Travel 7%, and Productivity with 7%.

Android free market shows Games 44%, Utilities 9%, Social Networking 8%, Photos & Video 7%, and Education with 6%. This explains that Games as well gain more popularity in the Android Finnish market as it was before in the US market. In contrast, Social Networking apps have gained a higher percentage in the US market than the Finnish. The Android paid top list includes Games 31%, Education 12%, Productivity 10%, Utilities 8%, and Photo & Video 7%. Travel apps gained more popularity in iPhone more than Android in the Finnish market.

4.2 Downloads visualization

This section examines the dynamics of downloads for the app stores, using figures to visualize app downloads per each category. Figure 16 compares the average number of downloads for an app for each category available on the top list. The average number of downloads is an indicator used to compare different categories between different stores and even markets. However, the pie chart of Figure 17 uses the cumulative downloads as an indicator of percentage share for each category. Both values reflect variation among different categories.

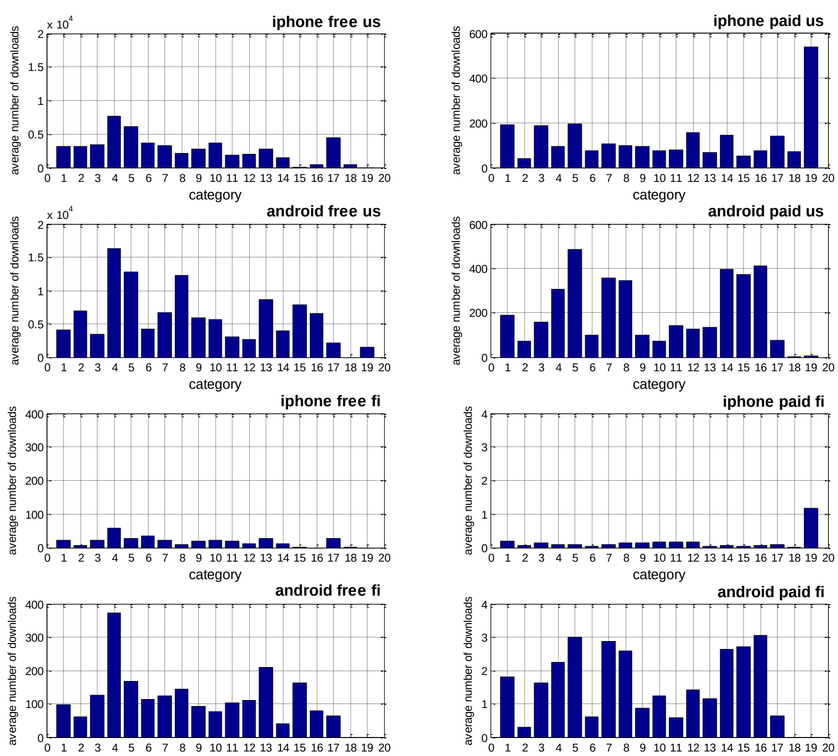


Figure 16 - Average number of downloads per category

4.2.1 US market

Downloads has always remained an important factor of the success of an app, at least in terms of visibility of an app. This visibility creates more market exposure for the app, leading to more revenue for app developers. Thus, analyzing app downloads is part of this analysis, in fact it will later have a great impact on the conclusions. Figure 17 shows the percentage of download counts per each category, which is commonly used as an indicator for an app category being successful in the reports of app stores. However, this indicator is from an app store perspective, in terms of the amount of revenue it earns from each app category. Amount of revenue gained by the app store is not included within the research scope of this thesis. It provides an analysis from an app developer perspective, as a result Figure 16 maps the average number of downloads per each category. This estimates the success of individual apps in each category based on the available dataset.

Figure 16 indicates a comparison between the average numbers of downloads in both stores. A huge gap can be easily observed between the free markets compared to the paid ones. The

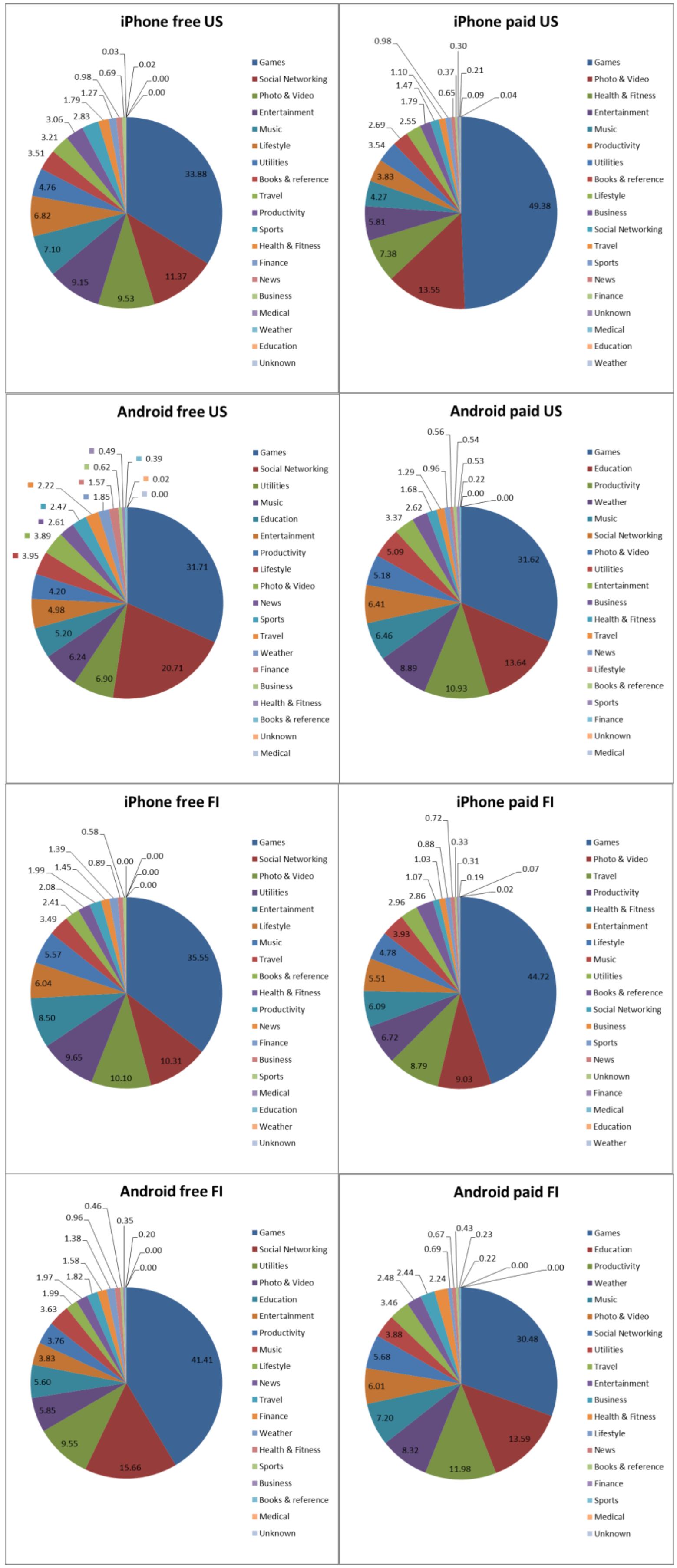


Figure 17 - Download count percentage share

free market appeals more to the app developers in terms of downloads, and seems more attractive as well as competitive. Additionally, Android market generates higher average downloads per category compared to the iPhone market. It has a considerably better performance on average for most categories. For example, category 5 Music has 6080 average downloads in iPhone free and 12800 in Android free, while generating only 194 in iPhone paid and 486 in Android paid.

This approach states if a certain category has a high download count does not typically mean that this category has a higher average downloads. For example, Music has the highest average downloads per app in iPhone free, while Games has the highest download count percentage. This indicates that the values highly depend on the number of apps available in each category. Nevertheless, the large percentage of apps per category can still indicate a large number of downloads compared to other categories. The same hypothesis applies to the Android market, where Games are still on the top download count list, but not on the top average download list as clarified by Figure 17. Education still has an important role as a category of apps having a high download count compared to other categories especially in the paid market of Android.

4.2.2 Finnish market

Figure 16 displays a huge gap between the average numbers of downloads per app in the free market relative to the paid one, which approximates to 100 times more free downloads than the paid ones. This gap between free and paid apps turns out to be larger than the one in the US market, which was only 30 times approximately. Android store performs generally better than iPhone market in Finland in terms of average number of downloads. Most of the apps in the paid iPhone store have average downloads less than one denoting a very small number. Social Networking apps have the highest average download for free Android apps, and Music has the highest average download in the paid apps.

Figure 17 shows the percentage share of apps in the Finnish market, recording no major differences between the US and Finnish markets. However, Utilities gained more popularity in the free iPhone store, while the Music gained less popularity relative to the US market. Additionally, Entertainment and Music lost some of their popularity within the paid iPhone store. Finally, Music has also received less popularity in the free Android store.

4.3 Average Rank change

4.3.1 US market

Figure 18 represents average rank change values per day for each category for iPhone free store in the US market. The rank change is an indicator of movement of an app within the top list; thus, rank change values are calculated for each app. Consequently, the average daily values are calculated for each category reflecting the dynamic behavior of apps in each category. More figures for average rank values for all US stores are available in Appendix C.

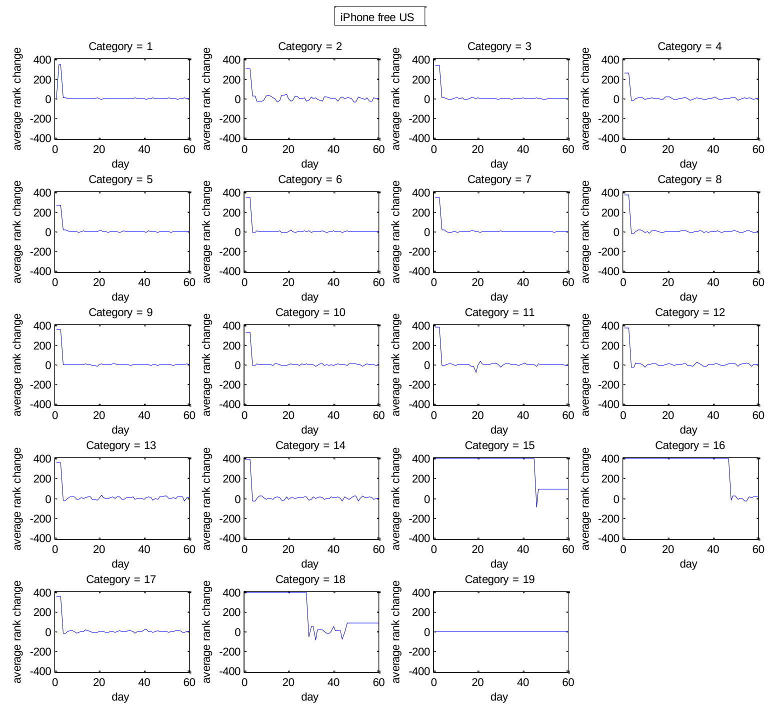


Figure 18 - Average rank change US market

Average rank change results indicates a normal distribution fitting along the 60 days as shown below for category 1 iPhone free us as a sample. The data does not fit totally a normal distribution along with excluding the outliers from calculations as shown by Figure 19, especially for the lower and middle points of the distribution as shown by the QQ plot in Figure 20 testing the validity of the approximation of the normal distribution. The normal distribution for the sample resulted in a mean of 0.098 and a variance of 10.83.

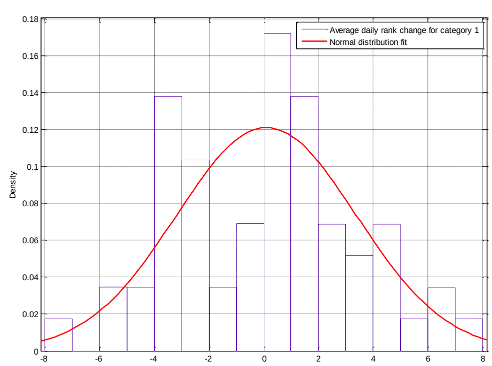


Figure 19 - Normal distribution fitting for average daily rank change of category 1 in the iPhone free US store

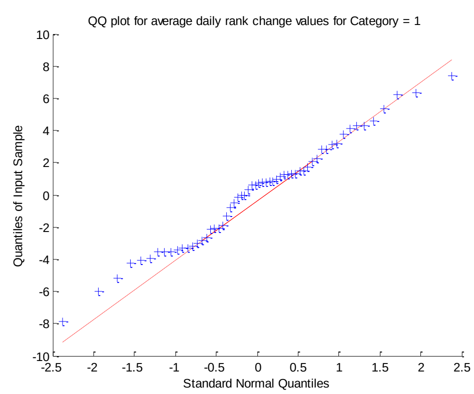


Figure 20 - QQ plot for average daily rank change of category 1 in iPhone free US store

Table 4 shows the mean values for each category within all stores subsequent to excluding the highest and lowest k values, such as $k = \frac{n \times \frac{percent}{100}}{2}$ given the percent is equal to 10. Thus, the mean values are calculated while removing 10 percent of the data regarding it as outliers for more accurate values.

This section presents interesting findings about the demand dynamics of the app stores, which has been noticed throughout the analysis of this dataset. This analysis is useful to map the dynamics of the whole app store, but is not of great value to an app developer. The mapping adopts the Ising model as an analogy for explaining the rank changes of apps in a certain category. It was developed by Ising (1925) as a mathematical model of ferromagnetism that is used in statistical mechanics. The model helps in the understanding of the behavior of complex networks, and provides general insights about the network. Newman (2003) maps different researches specifically concerning the uses for the model especially in complex social networks.

The model describes the magnetization of ferromagnetic materials like iron. These materials have the same direction magnetic dipole moments when it is magnetized. However, when exposed to high temperature, it loses its magnetization since the magnetic dipole moments cancel each other. The model defines magnetization as $M = \frac{1}{N}\sum_{i=1}^{N} S_i$. S is defined as magnetic dipole moments of atomic spins, which can have two states of +1 or -1. N is the number of ferromagnetic atoms. These atoms will have low correlation between their individual spins when exposed to infinite high temperatures, leading to positive values canceling negative ones thus zero magnetization.

This approach uses an analogy for different variables to relate this to the Ising model. The value of spins in the app stores will be ranging from -400 to +400. The N value is the number of apps. The precise values displayed in Table 4 represent the M value that was mapped by calculating the average rank change per app category in each store. This analogy examines the behavior of the whole store for each category. This behavior has resulted in average rank change

		US			
		iPhone		Android	
		Free	Paid	Free	Paid
Category	1	0.32	0.68	0.58	-0.01
	2	2.69	2.02	0.72	0.97
	3	0.53	0.42	1.46	0.017
	4	0.98	0.46	-0.04	-0.08
	5	0.97	1.10	0.68	-0.24
	6	0.035	0.32	1.63	0.13
	7	0.56	0.66	1.20	-0.06
	8	0.75	1.18	0.72	0.48
	9	0.16	0.97	0.94	0
	10	0.63	1.81	1.49	0.74
	11	1.66	2.24	0.52	-0.21
	12	1.20	1.24	0.97	0.71
	13	0.76	1.72	0.48	1.97
	14	1.05	1.79	0.17	0.52
	15	331.89	325.68	0.81	0.03
	16	328.06	321.81	0.98	-0.03
	17	0.28	0.51	2.63	0.45
	18	213.23	123.23	0	0
	19	0	0.96	325.78	328.52
Score		12/19	8/19	13/19	17/19

Table 4 - Mean average rank change US market

values in most categories to be very close to zero as shown by Table 4. Most categories have an initial value deviating from zero then start stabilizing around zero. These values have been also calculated numerically in Table 4 as mean rank values excluding the outliers. The table reveals that most of the values approximate to zero, represented by the white highlighted values lying between zero and one, since there are no decimal ranks for an app. This method aims to use this analogy as a tool of comparison between different stores. Some values highlighted in red are reported as large values, this is due to the unavailability of many apps in this category for this particular store that can be verified by Figure 14. Categories like 15, 16, and 18 in iPhone for both free and paid reported high values of M. Similarly, for category 19 in the Android market for both free and paid. Other values highlighted in green are not within the expected range but still there is no large standard of deviation from the zero mean value expected. As a result, Android store revealed 30 out of 38 values counted close to zero, compared only to 20 out of 38 values for iPhone. Thus, Android has 10 more values approximating to zero than iPhone, suggesting a lower correlation for individual behavior of apps in Android compared to iPhone. This explains when a platform is more open it results in greater diversity of apps. Thus, this diversity leads to low correlation between individual apps in the app store.

4.3.2 Finnish market

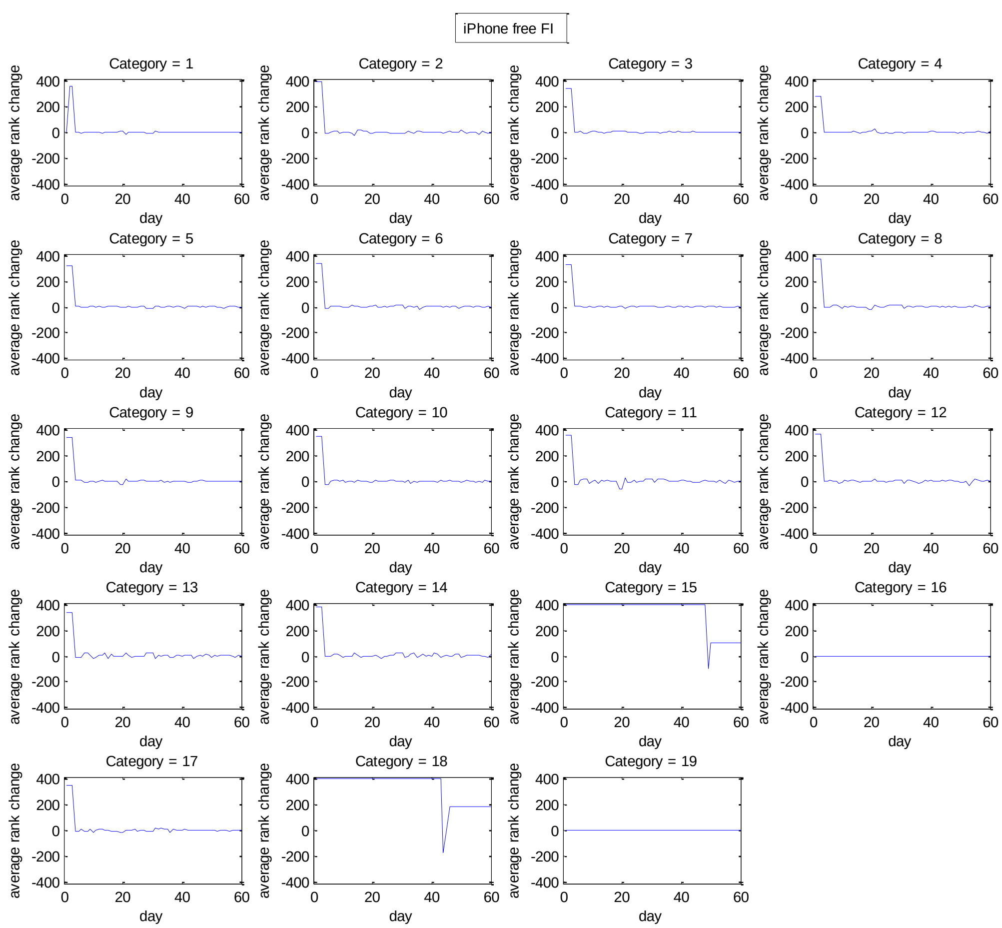


Figure 21 - Average rank change FI market

Figure 21 displays the average rank change values per day for each category for the entire iPhone free store in the Finnish market. More figures for average rank values for all Finnish stores are available in Appendix D. The average rank change behavior analysis uses an analogy of the magnetization vector, in order to describe behavior of the market. This magnetization vector values resulted in a lower correlation between the individual behaviors of apps in each category within the top list as shown in Table 5, thus reflecting higher degree of freedom for the Finnish market than the US. Some categories such as 15 and 18 reported large values highlighted in red, as a result of the low app count for these categories in the free iPhone store as shown by Figure 14. Moreover, categories 15 and 16 appear to have high values too in iPhone paid market due to low app counts. Other values are considered comparably small ones are highlighted in green, showing that they lie between 1 and 6.5 in this case. In conclusion, the Ising Model analogy confirms that the Android market is more open than the iPhone.

		FI			
		iPhone		Android	
		Free	Paid	Free	Paid
Category	1	0.18	0.16	-0.15	0.047
	2	0.29	0.95	-0.67	1.68
	3	0.67	0.53	0.67	0.032
	4	0.84	0.54	0.19	0.02
	5	0.33	0.60	0.39	0.10
	6	0.97	0.14	0.08	0.04
	7	0.70	0.40	0.21	0.35
	8	1.31	0.32	-0.13	0.33
	9	1.00	0.54	0.30	-0.18
	10	0.95	0.14	0.56	-0.18
	11	1.67	1.40	-0.12	0.16
	12	1.93	0.45	0.72	0.85
	13	1.26	1.29	-0.69	55.01
	14	1.45	0.32	0.95	-0.47
	15	351.33	305.51	0.69	0.08
	16	0	321.01	0.22	0.39
	17	-0.22	0.26	0.83	-0.52
	18	342.93	0.68	0	0
	19	0	6.51	0	0
Score		11/19	14/19	19/19	17/19

Table 5 - Mean average rank change FI market

4.4 Probability of success

The maximum days of survival values can be further used to calculate the percentage of successful apps for each category. This study assumes the criteria of success based on the percentage of apps that are above this threshold value. An app is regarded as successful when it remains on the top list for more than 15 days. Figure 22 shows the percentage of success per category for each of the stores. Table 6 sorts out the apps for each store based on the percentage of success. Table 7 combines values for each market and provides the sorted top 30 ranked categories for both app stores, based on the probability of success. Thus, this section assumes a threshold value to identify the successful app. This value is assumed as 15 days or more of consecutive availability of the app on the top list. It can be referred to as percentage of success or the probability of success of an app.

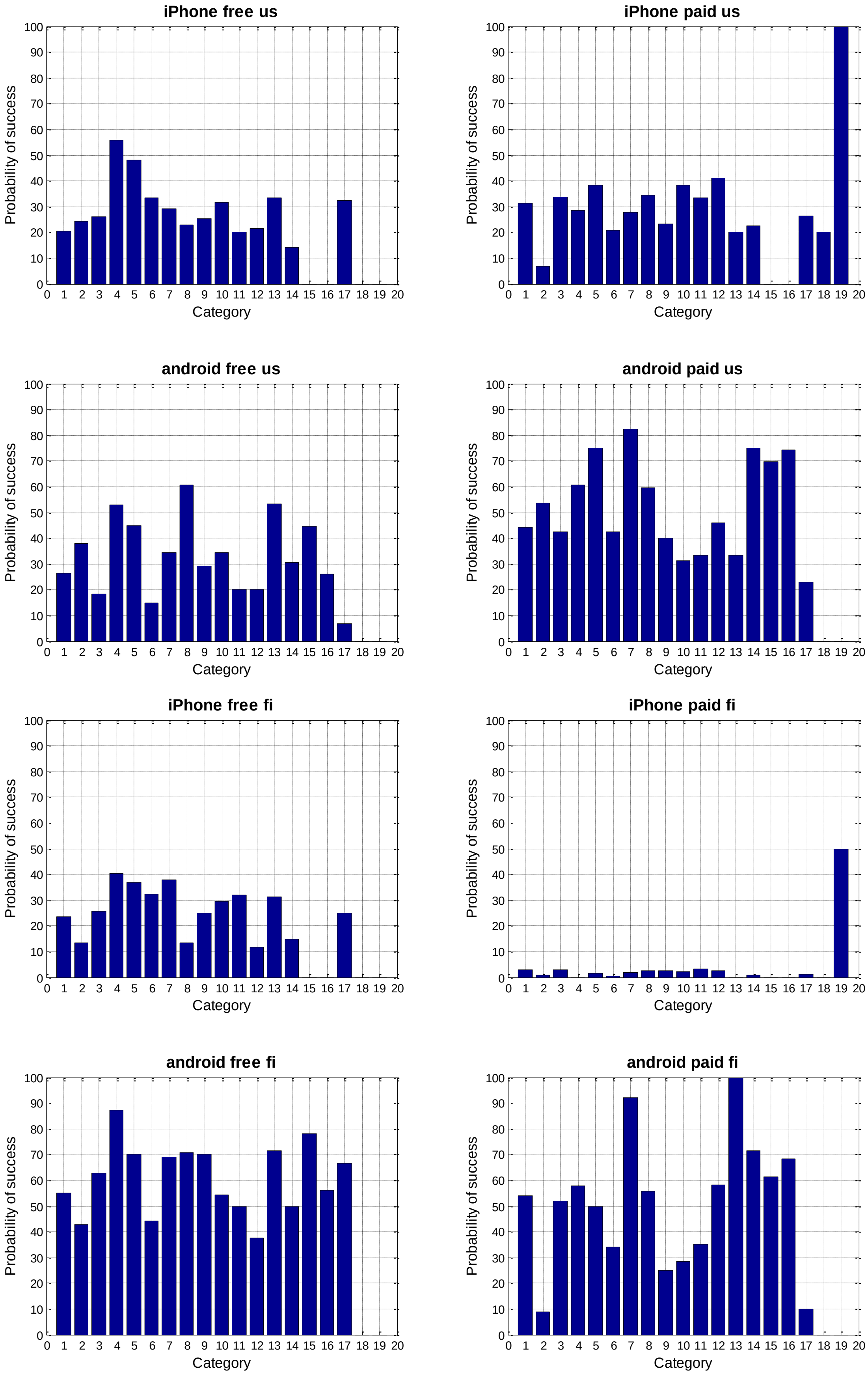


Figure 22 - Probability of success of each category

Category	iPhone Free US	Category	iPhone Paid US	Category	Android Free US	Category	Android Paid US
4	55.9	19	100	8	60.7	7	82.4
5	48.2	12	41	13	53.3	5	75
6	33.3	5	38.5	4	52.9	14	75
13	33.3	10	38.5	5	45	16	74.4
17	32.4	8	34.3	15	44.4	15	69.7
10	31.7	3	33.6	2	37.9	4	60.5
7	29.2	11	33.3	7	34.4	8	59.7
3	26	1	31.3	10	34.4	2	53.9
9	25.4	4	28.6	14	30.8	12	45.8
2	24.4	7	27.8	9	29.1	1	44.2
8	23.1	17	26.5	1	26.3	6	42.4
12	21.4	9	23.4	16	26.1	3	42.4
1	20.6	14	22.7	12	20	9	40
11	20	6	20.7	11	20	13	33.3
14	14.3	13	20	3	18.3	11	33.3
15	0	18	20	6	14.7	10	31.3
16	0	2	7	17	6.7	17	23.1
18	0	15	0	19	0	19	0
19	0	16	0	18	0	18	0
Category	**iPhone Free FI**	**Category**	**iPhone Paid FI**	**Category**	**Android Free FI**	**Category**	**Android Paid FI**
4	40.4	19	50	4	87.2	13	100
7	378	11	3.5	15	78.1	7	92.3
5	36.8	3	3	13	71.4	14	71.4
6	32.2	1	3	8	70.8	16	68.3
11	32	12	2.6	9	70	15	61.3
13	31.3	9	2.6	5	70	12	58.3
10	29.4	8	2.6	7	69	4	57.9
3	25.9	10	2.3	17	66.7	8	55.7
9	25	7	2.1	3	62.8	1	53.9
17	25	5	1.7	16	56.3	3	51.8
1	23.5	17	1.3	1	55.1	5	50
14	14.8	2	1	10	54.6	11	35.3
8	13.6	14	0.9	11	50	6	34.0
2	13.3	6	0.6	14	50	10	28.6
12	11.9	4	0	6	44.3	9	25
15	0	13	0	2	42.9	17	10
16	0	15	0	12	37.5	2	9.1
18	0	16	0	19	0	19	0
19	0	18	0	18	0	18	0

Table 6 - Ranking of apps per each store based on probability of success

Rank	Category	Percentage	Market
1	19	100	iPhone Paid US
2	7	82.4	Android Paid US
3	5	75	Android Paid US
4	14	75	Android Paid US
5	16	74.4	Android Paid US
6	15	69.7	Android Paid US
7	8	60.7	Android Free US
8	4	60.5	Android Paid US
9	8	59.7	Android Paid US
10	4	55.9	iPhone Free US
11	2	53.9	Android Paid US
12	13	53.3	Android Free US
13	4	52.9	Android Free US
14	5	48.2	iPhone Free US
15	12	45.8	Android Paid US
16	5	45	Android Free US
17	15	44.4	Android Free US
18	1	44.2	Android Paid US
19	6	42.4	Android Paid US
20	3	42.4	Android Paid US

Rank	Category	Probability	Market
1	13	100	Android Paid FI
2	7	92.3	Android Paid FI
3	4	87.2	Android Free FI
4	15	78.1	Android Free FI
5	13	71.4	Android Free FI
6	14	71.4	Android Paid FI
7	8	70.8	Android Free FI
8	5	70	Android Free FI
9	9	70	Android Free FI
10	7	69	Android Free FI
11	16	68.3	Android Paid FI
12	17	66.7	Android Free FI
13	3	62.8	Android Free FI
14	15	61.3	Android Paid FI
15	12	58.3	Android Paid FI
16	4	57.9	Android Paid FI
17	16	56.3	Android Free FI
18	8	55.7	Android Paid FI
19	1	55.1	Android Free FI
20	10	54.6	Android Free FI

Table 7 – Full ranking of categories in both US and FI markets based on the probability of success of an app category

4.4.1 US market

Figure 22 and Table 7 show a higher probability of success for Android apps compared to iPhone. The apps on the top successful list include Unknown category, which cannot be analyzed, since this value is a result of some unavailable data about the categories of these apps. Therefore, the observations focus on results of other categories. Entertainment category reports high success rates in the US market. Music has a high probability of success of 82.35% followed by Business, Weather, and Education in the Android paid market. Conversely, it is not reliable to compare the success of an app only by using the probability of success, which only indicates the number of days an app can survive on the top market list. As a result, an indicator is required to provide more accurate results to the analysis. This indicator should include the average number of downloads as mentioned earlier. It will be explained in details in the following section.

4.4.2 Finnish market

The observations of the Finnish market denote low probability of success for iPhone store apps, reaching a maximum of 50% for the Unknown category as shown in Table 7. This category is a result of dataset errors as mentioned earlier. Table 8 provides a ranking for the whole Finnish market in terms of probability of success. This ranking displays 20 apps with the largest survival probabilities including both stores. Android store apps occupied the whole list, leaving no opportunity for iPhone apps to be part of this list. Free apps recorded higher percentage share in the top 20 list than paid ones. Although the top list can provide an estimation about successful apps, simulation is needed to take into consideration all other ranked categories in all stores. This simulation will be presented by the Success Model throughout the following section. On the other hand, paid apps were on the top 2 in the list. Nevertheless, Finance had the highest probability of success, although it has a low percentage count in the top list. Therefore, the probability of success cannot be used directly as a success indicator.

4.5 Success Model

This section extends the results of the probability of success of an app to a probabilistic graphical model through a BN using the AgenaRisk software discussed earlier in the methodology chapter. The model shown by Figure 23 is the Success Model, since it represents the probability of success of an app.

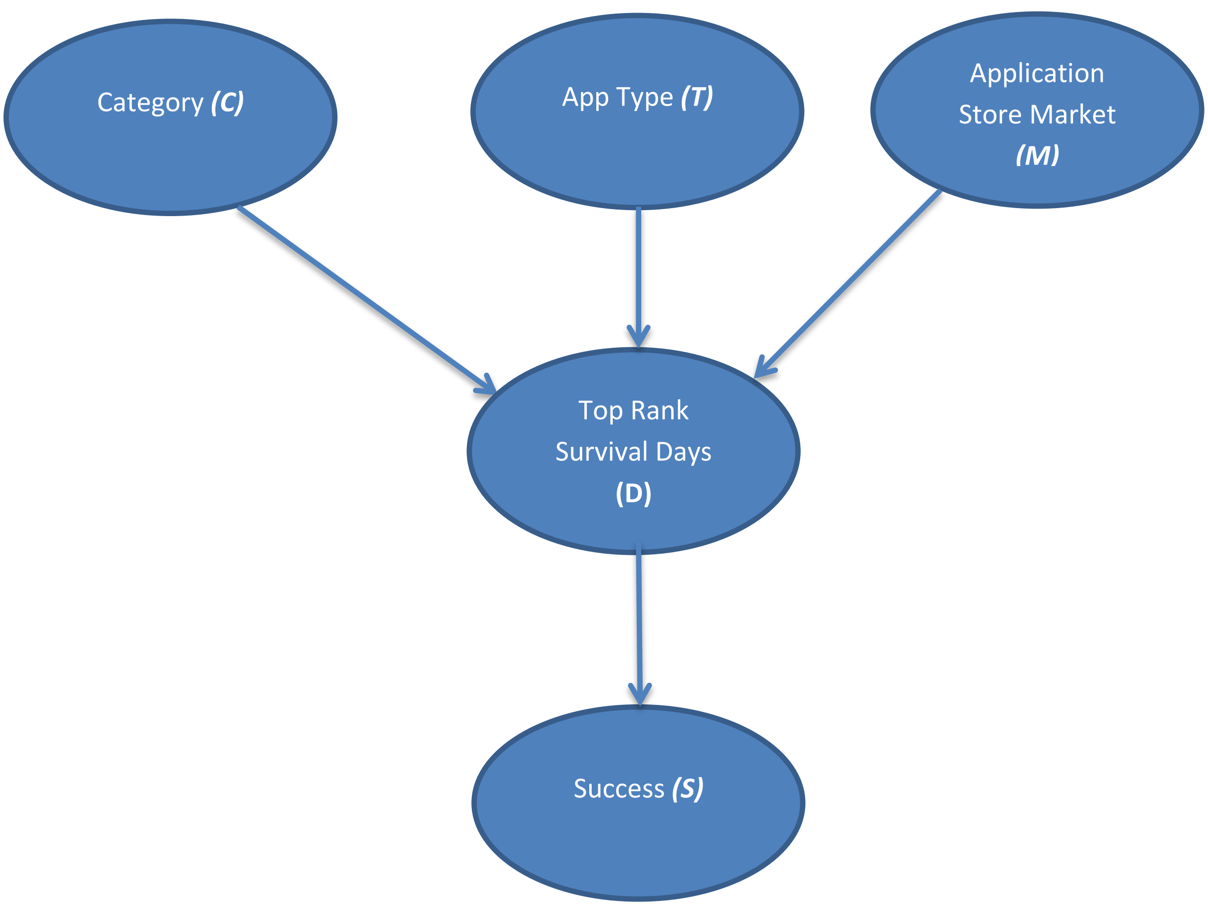


Figure 23 - Success Model

This model provides a guide for a mobile app developer to choose the market for launching an app. It provides the probability of success of an app for both stores Apple App Store and Google Play Android market. Each variable depends on the previous variable linked to it. This model assumes the free will of an app developer to choose the app category, type, and market. Thus, it is assumed that the probability to choose the market is 50% for each Apple App Store and Google Play. Furthermore, it assumes that the probability to choose free or paid App for both stores is 50%. Additionally, the model considers an app developer having an equal probability of choosing a certain category of an app.

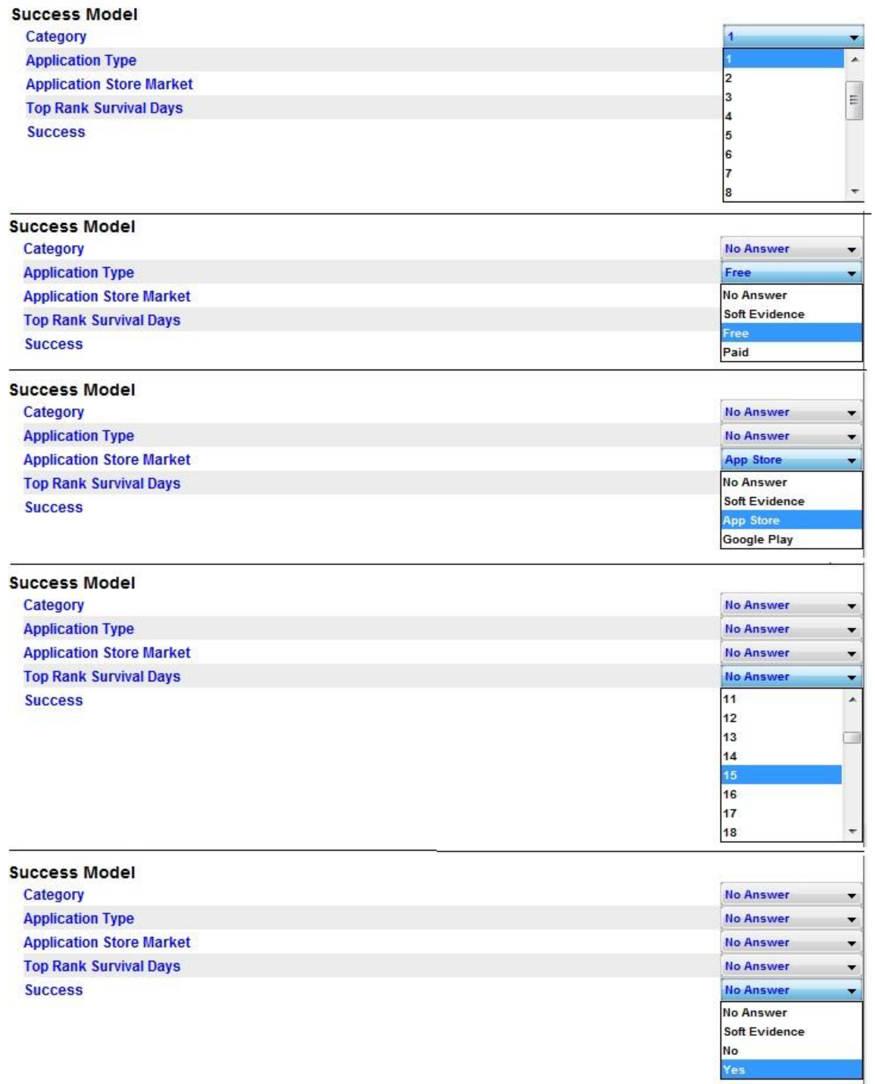


Figure 24 - Success model variables

P(m_0)=0.5 , this is the prior probability to choose iPhone Apple App Store

P(m_1)=0.5 , this is the prior probability to choose Google Android Market

P(t_0)=0.5 , this is the probability to develop free app

P(t_1)=0.5 , this is the prior probability to develop paid app

P(C_i)=$^1/_{19}$, this is the prior probability to develop an app for a certain category

P(D|C,T,M) is the probability that an app survives being listed among the top 400 for certain number of days given the category, market and the app type

P(S|D) is the probability of success of an app given the number of days

Equation 2 is the joint probability distribution function of the BN presented in Figure 23.

$$P(S, D, C, T, M) = P(S|D) \times P(D|C, T, M) \times P(T) \times P(M) \times P(C)$$

Equation 2

Figure 24 shows the variables used by the Success Model when implemented using the AgenaRisk software. The model shows the 5 different random variables mentioned earlier. The tool provides flexibility to provide any random variable given evidence. An app developer can choose between 19 different categories, as well as the app type can be chosen as free or paid. The model combines both Apple App Store and Google Play based on the available dataset. A wide range of values is provided for the maximum number of days of survival for an app. Finally, the success variable determines whether an app is successful or not based on the threshold criteria defined earlier by this chapter.

The following part of this chapter examines different scenarios using the BN model. All of these scenarios are used to explain the behavior of app stores from different angles, thus supporting conclusions concerning the behavior of apps of different app stores. It also explains different uses of the BN model and the utilization methods for an app developer.

4.5.1 US Market

Scenario 1: Simulation Example

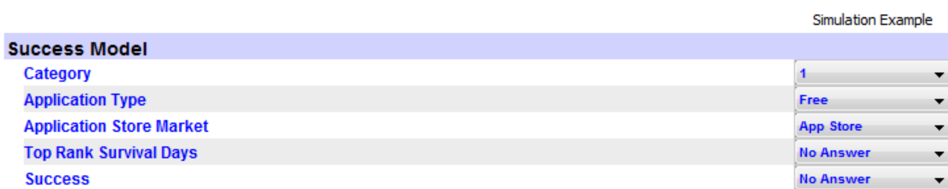


Figure 25 - Scenario 1 inputs

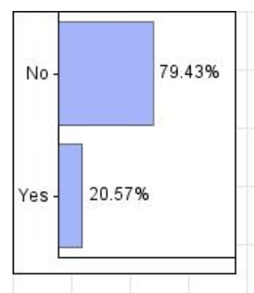


Figure 26 - App probability of success Scenario 1

The simulation of scenario 1 simply illustrates how an app developer would use the BN tool to determine the probability of success of an app by setting the inputs of the BN model as shown in Figure 25. The scenario involves category 1 along with Apple App Store and free app. Figure 26 indicates a success rate of 20.57%, as well as Figure 27 examining the distribution of number of app survival days. The survival of apps generally experiences an inverse relationship with the probability of success. However, it can be noticed that some days have larger values than the previous ones, for example day 3 having higher success rate than days 2 and 1. This study presents this scenario as an example of predictive function of a BN model.

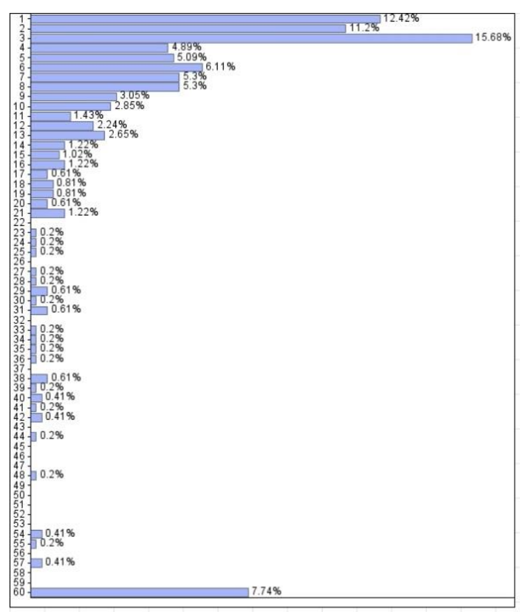


Figure 27 – Distribution of App survival days for scenario 1

Scenario 2: Free Apple App Store vs. Paid Apple App Store vs. Free Google Play vs. Paid Google Play

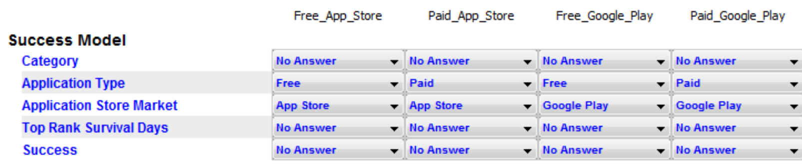


Figure 28 - Scenario 2 inputs

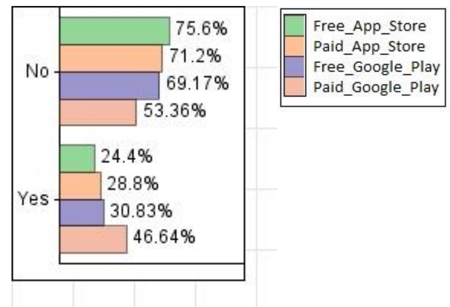


Figure 29 - App probability of success Scenario 2

This scenario introduces a comparison between different app types and stores. Figure 28 presents the inputs for the BN model, taking into consideration that app category is not given as an evidence for this scenario. Figure 29 states that 24.4% of Free Apple App Store apps were successful, 28.8% for Paid Apple App Store apps, 30.83% for Free Google Play apps, and finally 46.64% for Paid Google Play apps. Figure 30 and Figure 31 as well show the distribution of app survival days among the top list, which remains inversely proportional to the probability of success as mentioned earlier. However, there are extreme values not following this relationship such as values of the free Apple App Store for days 5, 6, 7, 9 and 15. Nevertheless, these values do not indicate an extreme deviation of the values from the expected values. However, values of free Google Play apps denote a large deviation from the expected values as shown in Figure 31. These large deviations take place at days 6, 10, 24, and 34. Moreover, paid Google Play apps indicate an increase in the percentage of apps surviving for 58 days.

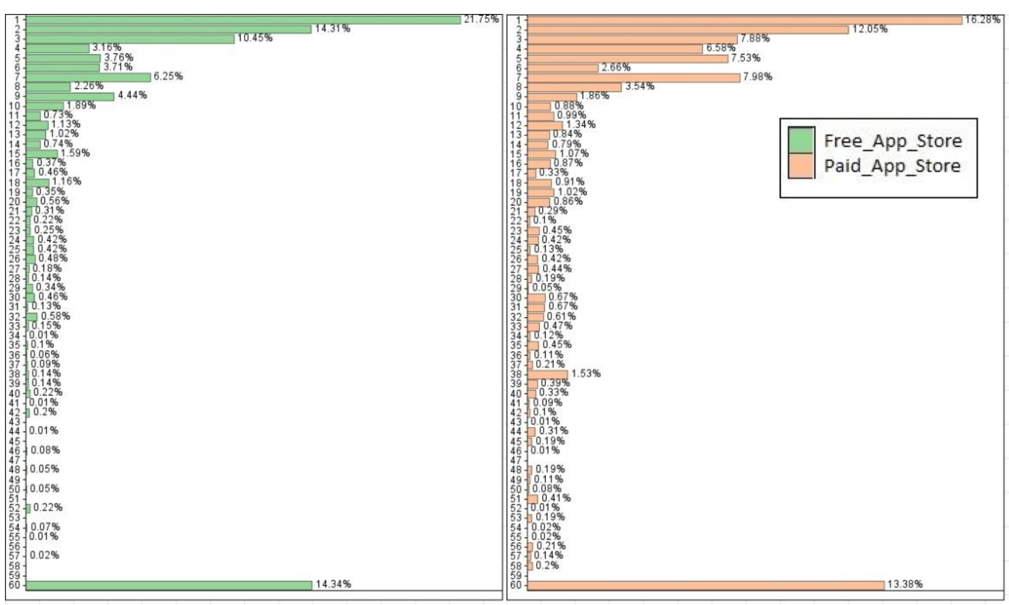


Figure 30 - Distribution of App survival days for scenario 2, Free App Store vs. Paid App Store

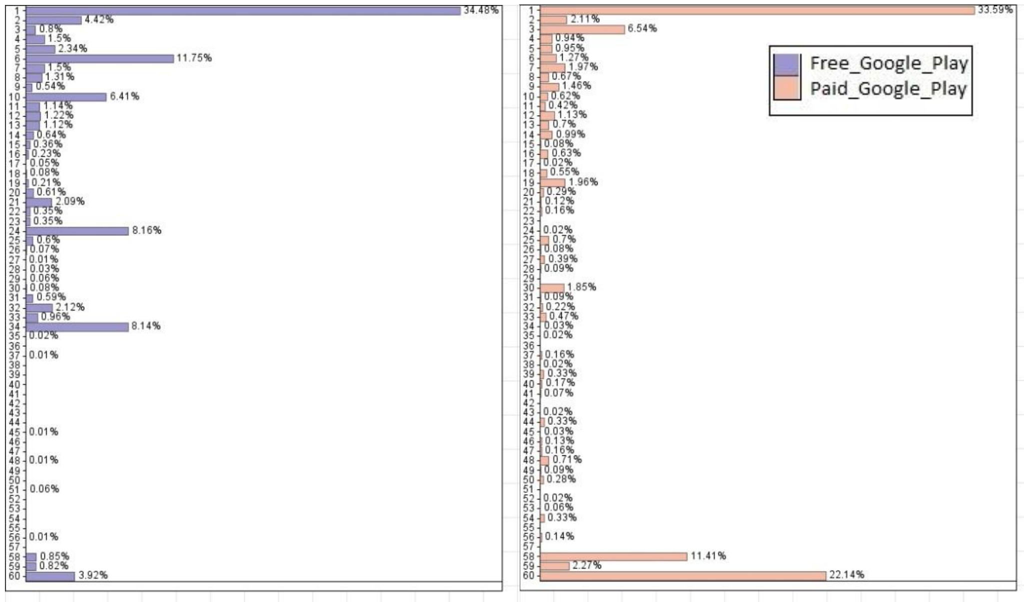


Figure 31 - Distribution of App survival days for scenario 2, Free Google Play vs. Paid Google Play

This study further investigates the reasons behind the behavior of Google Play, which does not follow the inversely proportional relationship between the number of apps and the number of survival days. The investigation includes 24 and 34 days of app survival for free Google Play apps, while not taking into consideration 6 and 10 days since these values lie below the threshold number of days assumed for a successful app.

The analysis utilizes the Success model and further exposes the distribution of app categories for these days as shown in Figure 32 and Figure 33 . Figure 32 reveals the highest percentage share belonging to category 13 Finance, followed by category 5 Music. Thus, it is vital to analyze the distribution of category 13 representing the highest percentage of 34 survival days of an app. After analyzing category 13 in Figure 34, it turned out that all of the apps in this category do not survive more than 34 days. Therefore, this explains the increase in the number of apps surviving 34 days. Moreover, category 5 has a high percentage share of apps surviving 34 days as shown in Figure 35. However, there is no further indicator for the reasons behind the behavior of these particular categories in the store.

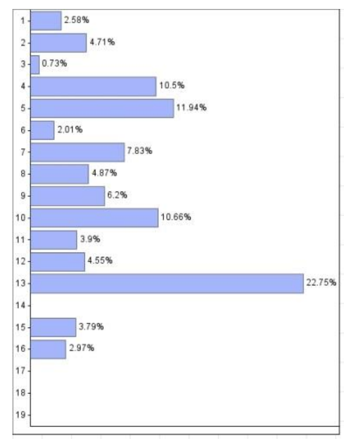


Figure 32 - Category distribution for Day 34 Google Play free

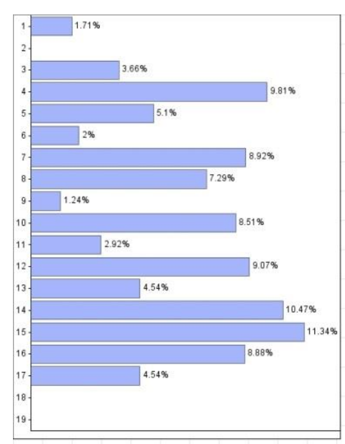


Figure 33 - Category distribution for Day 24 Google Play free

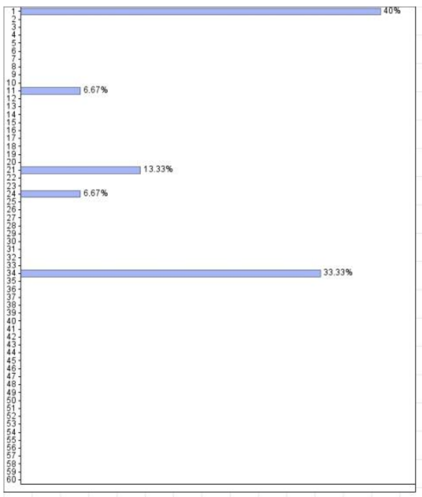


Figure 34 – Distribution of the number of app survival days for Category 13 “Finance”, Google Play Free

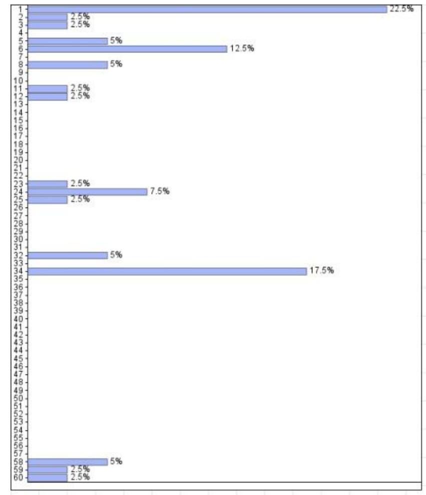


Figure 35 - Distribution of the number of app survival days for Category 5 “Music”, Google Play Free

Figure 33 displays the percentage shares of apps surviving 24 days including categories 15, 14, and 4. Figure 36, Figure 37 and Figure 38 provide the distribution of each of the categories. There is no clear explanation found behind the reasons of the behavior of these categories.

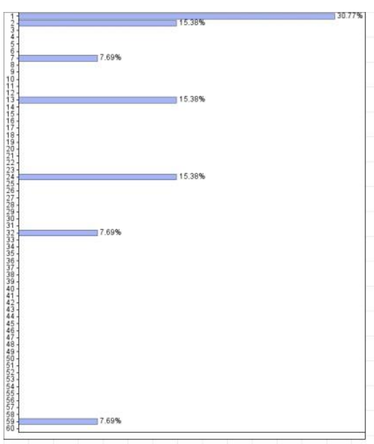


Figure 36 - Distribution of the number of app survival days for Category 14 “Business”, Google Play Free

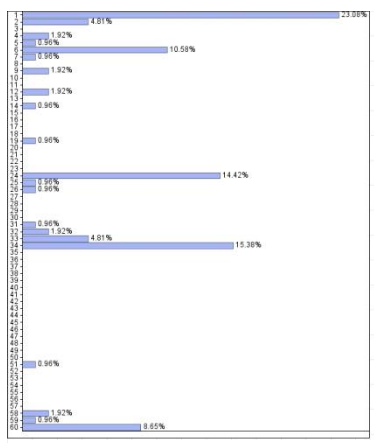


Figure 37 - Distribution of the number of app survival days for Category 4 “Social Networking”, Google Play Free

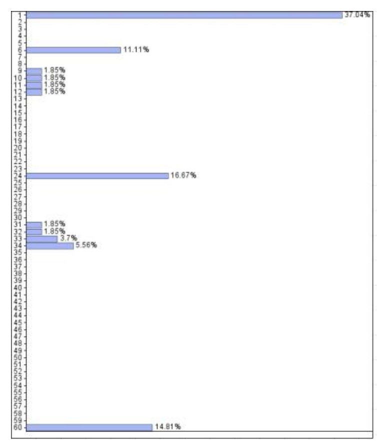


Figure 38 - Distribution of the number of app survival days for Category 15 “Education”, Google Play Free

As for the Google Play Paid store, Figure 39 shows the percentage share of each category of apps surviving 58 days. It indicates categories 5 Music, 16 Weather, and 15 Education having the highest percentage share. Music and Weather denote the largest percentage of the apps surviving 58 days, as shown in Figure 40Figure 41. Figure 42 describes Education apps survival distributions, pointing out that the second highest percentage of apps survive for 58 days in the top list.

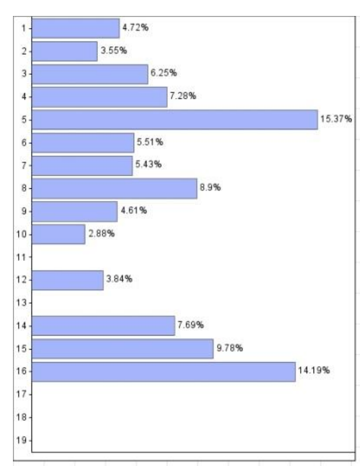


Figure 39 – Category distribution for Day 58 Google Play Paid

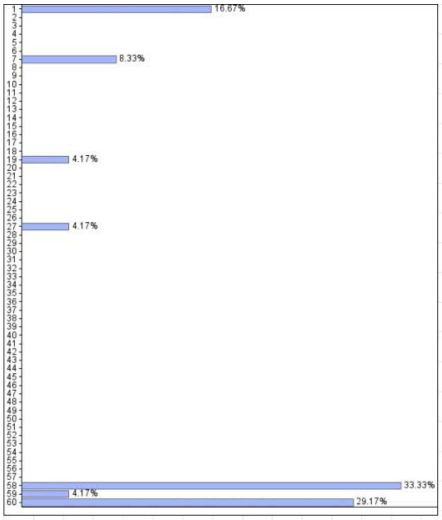


Figure 40 - Distribution of the number of app survival days for Category 5 "Music", Google Play Paid

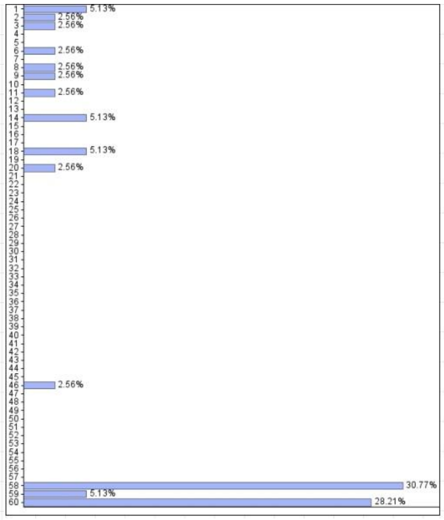


Figure 41 - Distribution of the number of app survival days for Category 16 "Weather", Google Play Paid

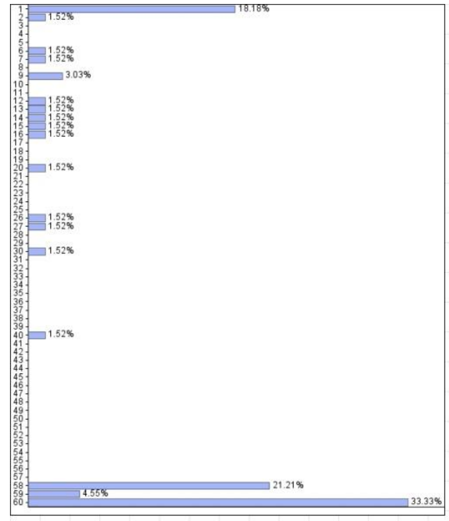


Figure 42 - Distribution of the number of app survival days for Category 15 "Education", Google Play Paid

Scenario 3: Apple App Store vs. Google Play vs. Free vs. Paid

Figure 43 - Scenario 3 inputs

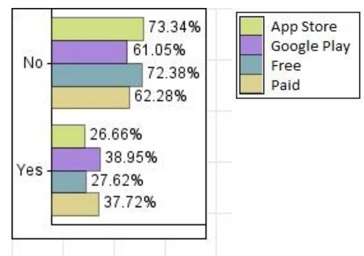


Figure 44 - App probability of success Scenario 3

BN model is utilized by this scenario in order to compare between Apple App Store, Google Play, Free, and Paid. It is significant to compare the earlier two, i.e. Apple App Store vs. Google Play, and later two, i.e. free vs. paid, as a bundle. Figure 44 shows the probability of success for each of the previously mentioned bundles. The Apple App Store achieves a probability of success of 26.66% for an app, while Google Play excels by more than 12% reaching a probability of success 38.95%. On the other hand, paid apps indicated a probability of success of 37.72% exceeding the free apps by more than 10%.

Scenario 4: Success = Yes

This study considers a very important scenario using the reasoning functionality of BN models. In this scenario, the success variable is given as input evidence "yes". Consequently, this scenario analyzes the behavior of the other random variables. Figure 45 refers to the BN calculation of the probability of an app belonging to each of

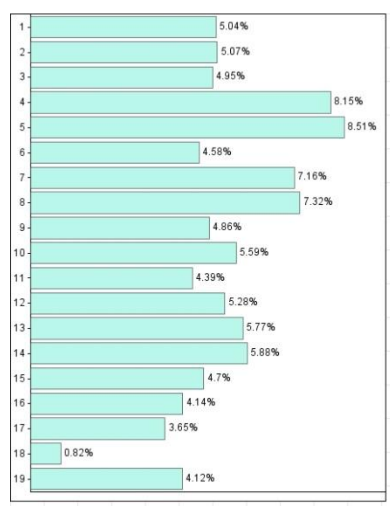


Figure 45 - Scenario 4 category success

the categories. This reveals the variation of probability of success of all app categories, led by category 5 Music followed by category 4 Social Networking, leaving the most unsuccessful category to be 18 Medical. Figure 46 presents the probability of this app belonging to Apple App Store or Google Play. The probability of an app is free or paid is also explained by Figure 47. This demonstrates higher probability of a successful app to be from Google Play and Paid. Finally, Figure 48 presents the distribution of the probability of the app survival days. The distribution starts from day 15 onwards, revealing no specific trend for the number of survival days for an app.

Figure 46 - Percentage share for each application store, scenario 4

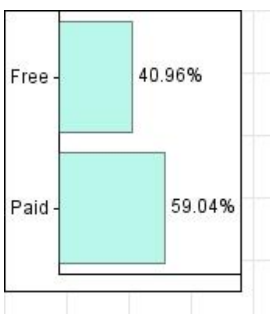


Figure 47 - Percentage share for each app type, scenario 4

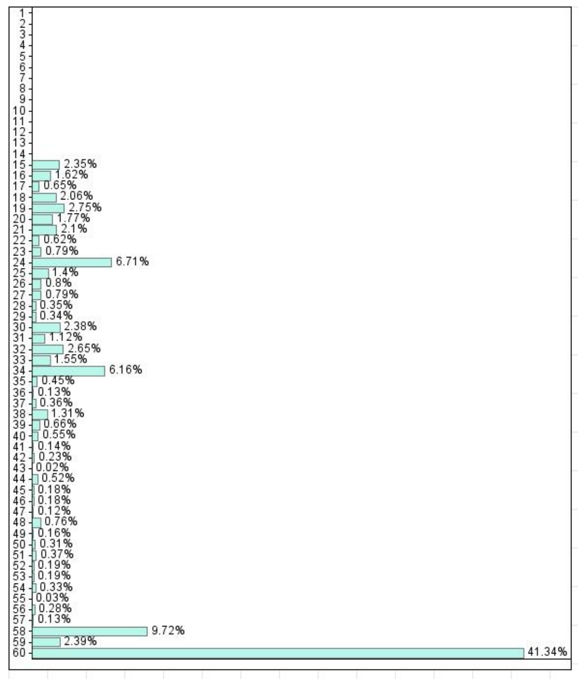


Figure 48 - Distribution of App survival days for scenario 4

Scenario 5: Success= Yes + given Category =1

This scenario can be used by an app developer whom has already developed an app and intends to publish this app. This scenario provides guidance for the app developer to choose the app store and the pricing strategy for this app based on probability of success. An app developer has to enter the inputs for the model as shown in Figure 49. Therefore, the app developer knows the category of the developed app in advance and adds this as an input to the model. The model will provide results about the app type success and the store success.

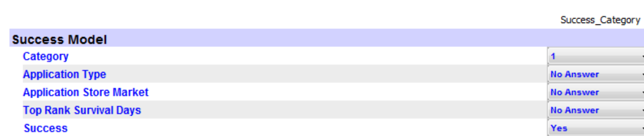


Figure 49 - Scenario 5 inputs

Figure 50 - Percentage share for each application store, scenario 5

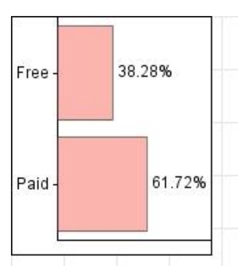


Figure 51 - Percentage share for each app type, scenario 5

Consequently, the app developer will be able to decide the pricing model for the app and the app store suitable for publishing this specific category of apps. In this case for category 1 Games, publishing the app in Google Play with a paid pricing model ensures higher probability of success for the app. Figure 50 and Figure 51 reveal the success of Google Play over app store, as well as paid apps over free ones.

Summary of scenario results

The Success model is regarded as a representation for the probability of success of an app. The model examines the dynamics of the mobile app store market using different simulation scenarios. These scenarios enable an app developer to predict the probability of success of an app for both stores, using the output results and the flexibility of changing the inputs of a scenario based on the situation. Moreover, the model uses BN, which takes into consideration the interdependence of model variables. Nevertheless, it must be noted that the probability of success does not reflect the revenues of an app. It only represents the probability of success of an app in both stores, which can be further used to calculate the success indicator. The success indicator is the result of multiplying the probability of success and the average number of downloads, which is explained later in the results of this chapter.

The model finds out a higher probability of success for Google Play relative to Apple App Store. This conclusion appears clearly in scenario 2 when paid apps in the Apple App Store succeed by 28.8%, while 24.4% for the free ones. Whereas, Google Play reports 30.83% for free apps and 46.64% for paid ones. Furthermore, scenario 3 confirms the results by Google Play leading with 38.95% followed by Apple App Store with only 26.66%. A successful app is more likely to be from Google Play with a probability of 59.37%, whereas Apple App Store is 40.63% as shown by scenario 4. The Success model further exposes the higher probability of success of paid apps in both markets over free ones. Scenario 2 explains the higher probability of success of paid apps. Paid apps appear to have 10.1% higher probability of success relative to the free ones as described by scenario 3. Similarly, scenario 4 results confirm that paid apps have 59.04% probability given the app is successful, leaving only 40.96% for the free apps.

The model uses the number of survival days as a parameter of success of an app in the store. This has resulted in the success of Google Play and paid apps in the US market over Apple App Store and free apps respectively. This explains that Google Play is less competitive than Apple App Store, indicating higher stability of apps in the top list of Google Play compared to the Apple App Store. Moreover, the higher probability of success of paid apps than free ones can reflect the higher competition in the free app store relative to the paid one. As a result, an app in the paid store is able to survive longer time among the top ranking list.

4.5.2 Finnish Market

This chapter presents two scenarios concerning the Finnish market. These scenarios can reveal the main insights of the model when applied to the Finnish market.

Scenario 1: Apple App Store vs. Google Play vs. Free v. Paid

The first scenario is exactly similar to the third scenario applied in the US market. The aim of this scenario is first to compare the probability of success of an app in Apple App Store and Google Play. Additionally, it compares the Free and Paid apps in the Finnish market. The success percentage is presented by Figure 54. It illustrates a higher probability of success for Google Play apps with 55.87% than Apple App Store with only 13.05%. Unlike the US market, free apps express higher probability of success than paid ones reaching 42.04% in the Finnish market. These results of the BN model explain the success of Google Play apps and free apps in the Finnish market. These values clarify that the Finnish paid stores along with the Apple App Store are more competitive compared to the free ones and Google Play respectively. Furthermore, these values are vital in order to predict the success indicator of a certain app category in an app store.

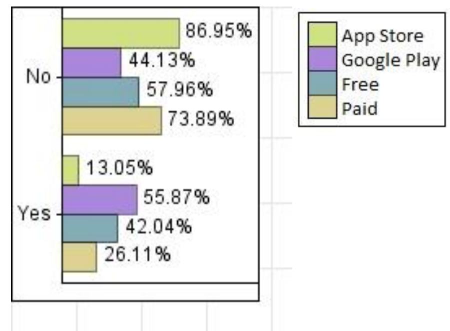


Figure 54 - App probability of success

Figure 52 - Success model scenario for Finnish market, App Store vs. Google Play

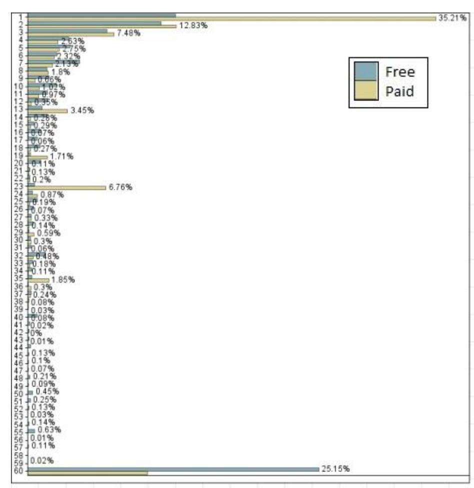


Figure 53 - Success model scenario for Finnish market, Free vs. Paid

The Finnish market shows an inversely proportional relationship between the number of survival days and the count of apps along with minor standards of deviation. However, Day 23 indicates a large variation regarding paid Google Play apps as shown in Figure 52 Figure 53. The Finnish market demonstrates a high percentage of category 13 Finance apps in day 23, followed up by category 7 Entertainment as illustrated by Figure 55. Figure 57 indicates that Finance paid apps do not survive after

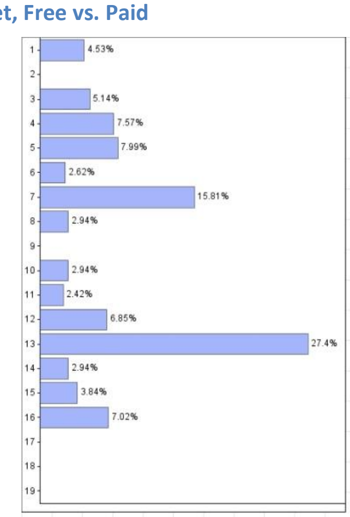


Figure 55 - Category distribution for Day 23 Google Play Paid

23 days in the Google Play store. Moreover, Entertainment apps reveal an interesting behavior with only apps surviving 1 day or 23 days or 60 or even more as shown in Figure 56.

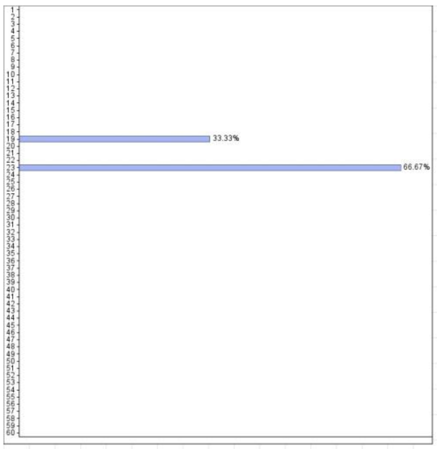


Figure 57 - Distribution of the number of app survival days for Category 13 "Finance", Google Play Paid

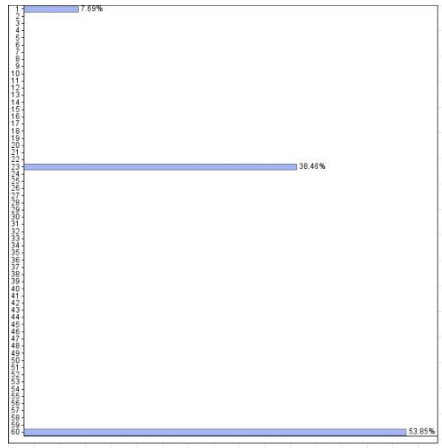


Figure 56 - Distribution of the number of app survival days for Category 7 "Entertainment", Google Play Paid

Scenario 2: Success = Yes

This scenario is initiated by giving evidence that the app is successful similar to scenario 4 in the US market. Figure 58 displays the probability distribution of each app category. This represents the leading of categories 13 Finance, 7 Entertainment, and 4 Social Networking. Although Finance category is leading in the Finnish market, this can be a result of the low count of apps available in the store.

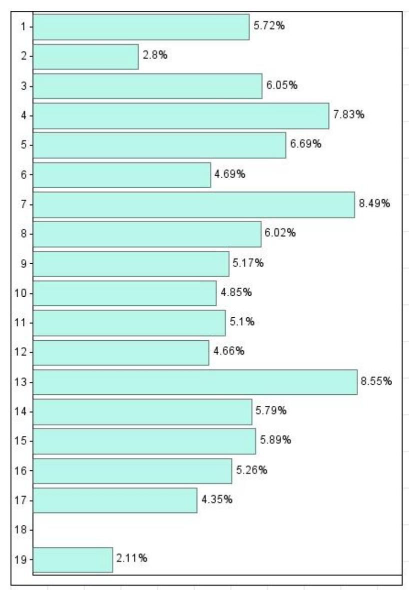


Figure 58 - Probability distribution of successful app categories Finnish market

4.6 Success Indicator

The app survival by itself cannot be used as an indicator for success. Therefore, there has to be another indicator to precisely measure the level of success of an app. This study examined the variations an average number of downloads, and the large gap between different app stores within different markets. Moreover, there is a difference between free and paid apps within the same app store. Consequently, Equation 3 calculates the success indicator that is considered as a parameter of an app success throughout this research. These values are calculated for each category and mapped in Figure 59 for each app store of the US market. The figure uses the same y axis scale, which is important to use in order to compare the success for different app stores.

$$Success\ Indicator\ =\ Probability\ of\ Success\ \times Average\ app\ downloads$$

Equation 3

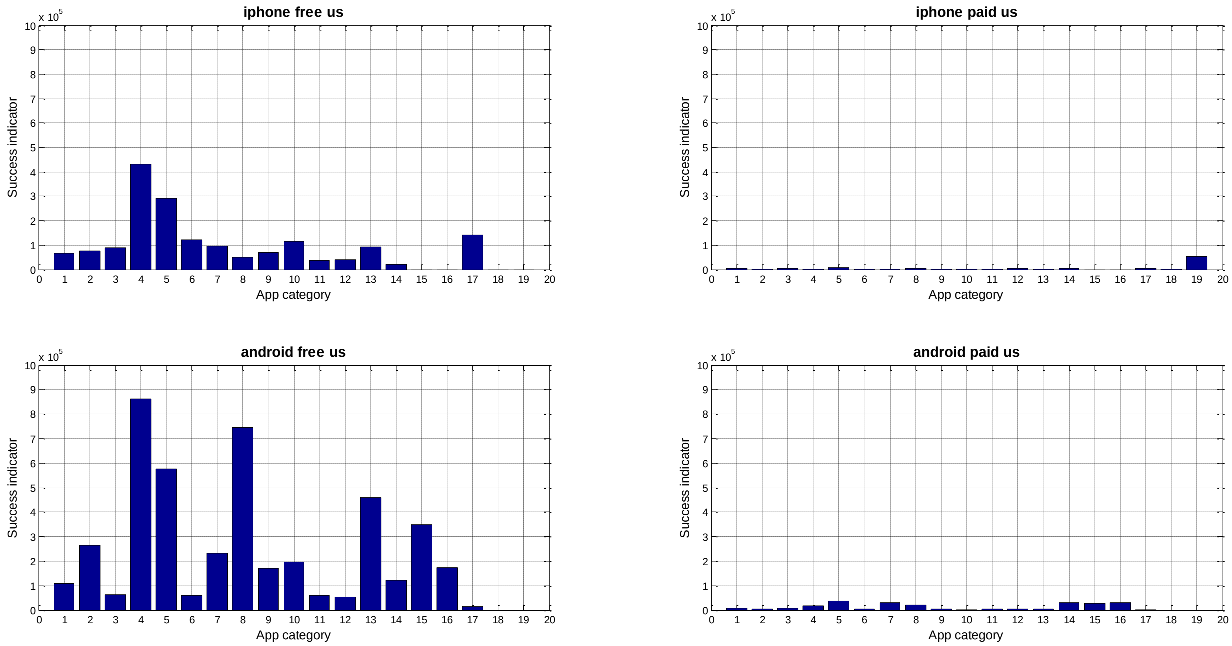


Figure 59 – Success indicator per category US market with the same y axis scale

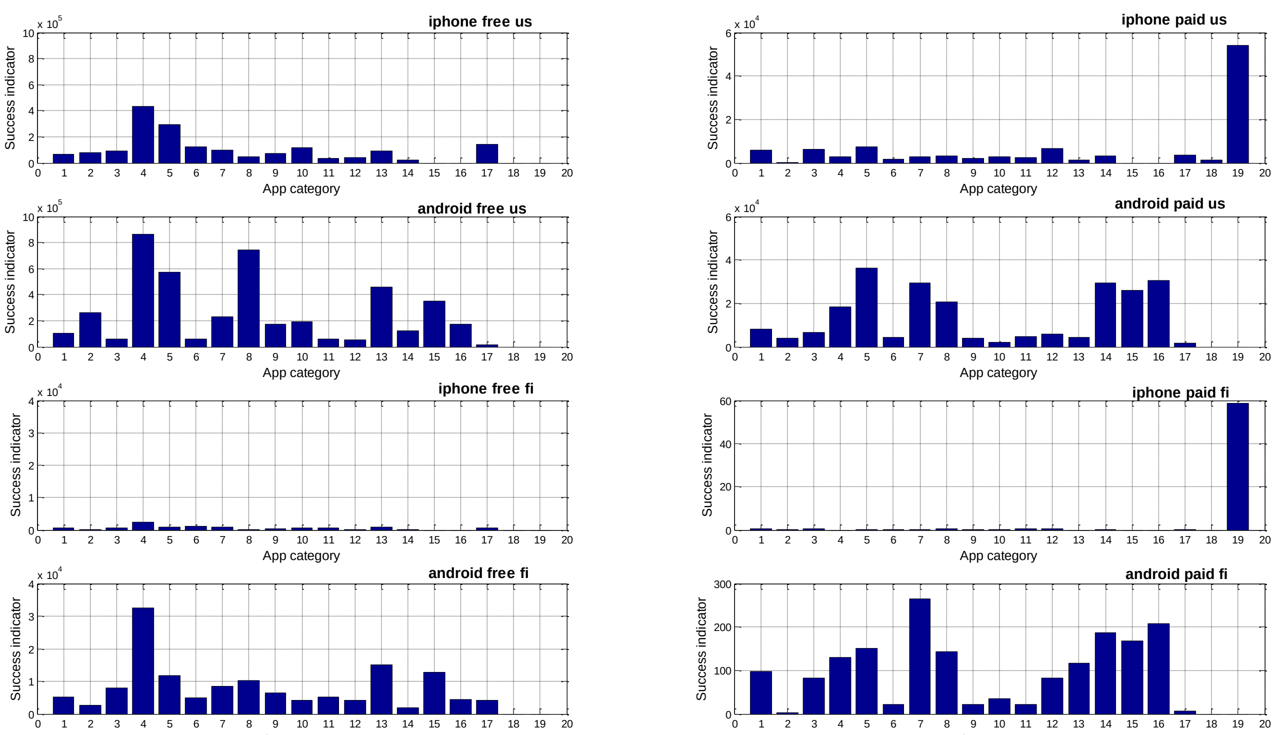


Figure 60 - Success indicator per category using different y axis scale

There has been a huge gap between values for different stores when using the same download scale. Thus, Figure 60 uses a different scale for each store in order to identify the detailed values of success. Table 9 provides a precise ranking of the top 30 successful categories for both US and Finnish markets, using the success indicator calculated by this section.

Category	iPhone Free US	Category	iPhone Paid US	Category	Android Free US	Category	Android Paid US
4	432350.5	19	54035.0	4	863419.8	5	36442.2
5	292869.5	5	7473.9	8	746016.1	16	30610.7
17	142508.2	12	6476.0	5	575187.1	14	29574.1
6	122468.6	3	6321.0	13	458805.7	7	29475.7
10	114932.7	1	6007.7	15	350626.7	15	26017.1
7	95266.8	17	3724.3	2	265214.7	8	20663.6
13	93216.3	8	3329.6	7	230628.6	4	18446.00
3	90249.9	14	3277.2	10	195648.0	1	8333.7
2	77849.9	7	2960.3	16	172206.4	3	6720.4
9	70444.3	10	2886.9	9	171328.0	12	5780.1
1	65700.6	4	2664.3	14	120540.8	11	4803.5
8	50237.3	11	2584.6	1	107343.4	13	4467.4
12	42222.5	9	2255.2	3	62675.6	6	4237.9
11	36414.2	6	1589.7	6	61174.8	9	4041.5
14	21690.4	18	1465.5	11	61134.0	2	3971.4
15	0.0	13	1320.1	12	53046.5	10	2271.8
16	0.0	2	282.7	17	14058.9	17	1728.5
18	0.0	15	0.0	19	0.0	19	0.0
19	0.0	16	0.0	18	0.0	18	0.0
Category	**iPhone Free FI**	**Category**	**iPhone Paid FI**	**Category**	**Android Free FI**	**Category**	**Android Paid FI**
4	2305.8	19	58.8	4	32480.7	7	264.9
6	1128.6	11	0.6	13	14982.5	16	208.2
5	952.5	1	0.6	15	12687.7	14	187.5
7	878.5	12	0.4	5	11780.7	15	167.1
13	857.2	3	0.4	8	10290.3	5	150.4
17	679.2	8	0.4	7	8458.7	8	143.3
10	637.0	9	0.4	3	7931.2	4	130.2
11	587.2	10	0.4	9	6450.7	13	116.1
3	561.3	7	0.2	1	5308.9	1	97.3
1	518.3	5	0.1	11	5087.1	3	83.5
9	497.4	17	0.1	6	4966.3	12	82.0
14	153.8	2	0.1	16	4496.3	10	35.4
12	132.0	14	0.1	17	4200.7	9	21.5
8	129.7	6	0.0	10	4192.7	6	21.1
2	81.2	4	0.0	12	4155.1	11	21.0
16	0.0	13	0.0	2	2586.4	17	6.4
15	0.0	16	0.0	14	2011.0	2	2.7
19	0.0	15	0.0	19	0.0	19	0.0
18	0.0	18	0.0	18	0.0	18	0.0

Table 8 - Ranking of app categories based on success indicator per store

Rank	Category	Success Indicator	Market
1	4	863419.8	Android Free US
2	8	746016.1	Android Free US
3	5	575187.1	Android Free US
4	13	458805.7	Android Free US
5	4	432350.5	iPhone Free US
6	15	350626.7	Android Free US
7	5	292869.5	iPhone Free US
8	2	265214.7	Android Free US
9	7	230628.6	Android Free US
10	10	195648.0	Android Free US
11	16	172206.4	Android Free US
12	9	171328.0	Android Free US
13	17	142508.2	iPhone Free US
14	6	122468.6	iPhone Free US
15	14	120540.8	Android Free US
16	10	114932.7	iPhone Free US
17	1	107343.4	Android Free US
18	7	95266.8	iPhone Free US
19	13	93216.3	iPhone Free US
20	3	90249.9	iPhone Free US
21	2	77849.9	iPhone Free US
22	9	70444.3	iPhone Free US
23	1	65700.6	iPhone Free US
24	3	62675.6	Android Free US
25	6	61174.8	Android Free US
26	11	61134.0	Android Free US
27	19	54035.0	iPhone Paid US
28	12	53046.5	Android Free US
29	8	50237.3	iPhone Free US
30	12	42222.5	iPhone Free US

Rank	Category	Success Indicator	Market
1	4	32480.7	Android Free FI
2	13	14982.5	Android Free FI
3	15	12687.7	Android Free FI
4	5	11780.7	Android Free FI
5	8	10290.3	Android Free FI
6	7	8458.7	Android Free FI
7	3	7931.2	Android Free FI
8	9	6450.7	Android Free FI
9	1	5308.9	Android Free FI
10	11	5087.1	Android Free FI
11	6	4966.3	Android Free FI
12	16	4496.3	Android Free FI
13	17	4200.7	Android Free FI
14	10	4192.7	Android Free FI
15	12	4155.1	Android Free FI
16	2	2586.4	Android Free FI
17	4	2305.8	iPhone Free FI
18	14	2011.0	Android Free FI
19	6	1128.6	iPhone Free FI
20	5	952.5	iPhone Free FI
21	7	878.5	iPhone Free FI
22	13	857.2	iPhone Free FI
23	17	679.2	iPhone Free FI
24	10	637.0	iPhone Free FI
25	11	587.2	iPhone Free FI
26	3	561.3	iPhone Free FI
27	1	518.3	iPhone Free FI
28	9	497.4	iPhone Free FI
29	7	264.9	Android Paid FI
30	16	208.2	Android Paid FI

Table 9 - Full ranking of app categories in both US and FI markets based on the success indicator

This success indicator is regarded as the most significant result for this study. This section demonstrates the criteria used for defining more accurate measurements for the success of an app. This indicator is a result of multiplying the probability of success and the average downloads value per category. This results in a value that is important from a perspective of an app developer, since it reflects both the revenues as a key factor and the survival of an app. This helps in identifying potential markets for an app developer when publishing their apps or even before development.

Results of the success indicator shown in Table 9 are different from the ones in Table 7, which were based on the probability of success. Therefore, the difference between both results can be easily noticed throughout the results.

4.6.1 US market

The success indicator revealed much higher success for an app in the free market compared to the paid market in addition to higher success for Android apps compared to iPhone. The list of top 30 apps in Table 9 included Unknown as the only one iPhone category. As discussed earlier, this category is a result of data errors due to unavailability of a category where the app is listed. Additionally, Games reported the highest percentage share of apps as shown in Figure 15, this contradicts the rank of Games as 17 and 23 for Android and iPhone respectively, when using the success indicator as a reference. According to the results most successful apps are Social Networking, Productivity, Music, Finance, Education, Sports, Entertainment, and Travel. These categories state high success using the success indicator proposed by this research ,which can be viewed in Table 9.

The previous analysis of Table 9 discussed the overall ranking for the US market; moreover, this part of analysis demonstrates the results of Figure 60 and Table 8, reflecting precisely the success indicator for each store. This comprehensive analysis attempts to compare different stores together, taking into consideration that overall success has been discussed earlier. The top five iPhone free successful apps included Social Networking, Music, Books & Reference, Utilities, and Travel, whereas Android free included Social Networking, Productivity, Music, Finance, and Education. This list shows different user behavior and preferences in different stores, even though users are from the same geographical market. Nevertheless, Social Networking and Music apps are appealing in both stores. The iPhone paid store top list includes Unknown, followed by Music, Health & Fitness, Photos & Video, Games, and finally Books & Reference. The top list indicates success of the Music, and Books & Reference in both free and paid iPhone stores, while users prefer to buy different app categories in paid than free ones. Users are more likely to pay for Health & Fitness, Photos & Video, and Games than other categories. This observation excludes the fact that paid apps are not as successful as free apps in the collective ranking of the US market as shown in Table 9. Additionally, the paid Android list included Music, Weather, Business, Entertainment, and Education. The only common app category with paid iPhone is the Music category. This confirms different user preferences in both markets. Education clearly indicates a large success in the Android market.

4.6.2 Finnish market

Different geographical markets experience different user behavior for both app stores, which appears in the success indicator top list in Table 9. This list indicates the top successful apps according to the calculations of this study including Social Networking, Finance, Education, Music, Productivity, Entertainment, Photos and Video, Lifestyle, Games, and News. In the Finnish market, new categories appear in the top successful list which did not appear in the US such as Photos and Video, Lifestyle, Games, and News. These results report Android having higher success in the top list than iPhone, leading with 19 apps in the top 30 app list as shown in Table 9. This table as well indicated free market as more successful than paid apps, confirming that through paid apps only occupying the last two positions in the list.

Table 8 verifies ‘Social Networking on the top five iPhone free successful app categories along with Utilities, followed by Music, Entertainment, and finally Finance. Social Networking, Finance, and Music apps are still part of the list of Android top free successful apps. Additionally, Education and Productivity have a significant popularity within the paid Android apps. In contrast, the top five paid iPhone apps have no common categories with the free ones. These successful paid apps have News, Games, Health & Fitness, Photo & Video, and Productivity. The list of the top Android apps is Entertainment, Weather, Business, Education, and Music, having no common categories with the top paid iPhone apps. Although, Music and Education apps are common top categories in both free and paid Android stores.

5 Discussion

Chapter 4 displayed an organized sequence of the results. Chapter 5 details and further analyzes the results, as well as observing the dynamics of the user demand for the mobile app stores along with discussing them with reference to other studies. This chapter contributes to direct comparison between different app stores and identifies the most successful app categories, while discussing US and Finnish markets separately. Finally, the latter part of this chapter explains the limitations of this research.

The dataset provided by the market research company includes download estimations of apps concerning both app stores. These estimations were identified by the company, by predicting download values for apps with information available about other apps. These values were employed throughout the analysis, which are not expected to be accurate. Accordingly, analyzing the absolute values will not provide reliable results. As a result, this analysis provides conclusions that are based primarily on comparing the statistical values of different stores, and the absolute values are not of the main interest concerning the scope of this research.

This study starts by analyzing the behavior of different app categories, taking into consideration the clustering effect discussed earlier by Petsas, Papadogiannakis & Polychronakis (2013). This clustering effect means that users usually download the next app from the same category as the previous one. The result of the clustering effect enables a better understanding of the dynamics of the mobile app store by mapping successful categories within the store. The thesis aims to identify the successful app categories in the mobile app stores. However, this does not change the fact that an app developer usually has a good source of revenue from a high-quality app, as mentioned by Wang and Wang (2013).

Kim (2012) reported that users of Google Play are less likely to purchase apps than Apple users. The study even mentions that the average quality of apps is higher for iPhone than Android. However, the earlier observation of Kim (2012) is valid from the perspective of an average single user of iPhone or Android. Nevertheless, this thesis shows in Table 8 and Figure 60 that Google Play apps are more successful than iPhone apps, by calculating the cumulative downloads of apps in both stores not only single user downloads. Additionally, Ghose and Han (2012) examine the demand of the South Korean market and identify a positive correlation between lifestyle and gaming apps, and a negative one in case of multimedia and educational apps. This conclusion is based on the combined datasets from both Apple App Store and Google Play. The results of this thesis support the positive correlation of lifestyle and gaming apps, as both appear on the top successful apps. However the results reveal that gaming, ranked as number 17 and 9 in US and Finland respectively as shown in Table 9, is not a leading category in the top list of apps in either of the two markets in terms of success. This supports the earlier results of positive correlation, while not indicating a huge success for gaming apps. Similarly, the lifestyle apps rank in the US and Finnish markets are 12 and 8, respectively. Music apps have shown a huge success in both Apple App Store and Google Play stores. Additionally, Education has excelled in the Android market. The results can vary due to markets in different regions, or the time of the analysis, or even it can be misleading due to the combined datasets. The combined dataset does not take into consideration different user behavior in Apple App Store and Google Play, as it assumes similar user behavior for the same category of apps in both app stores. Similarly, Entertainment was reported as one of the main revenue sources the users are motivated to download (Ho & Syu 2010). This conclusion has also been confirmed throughout this study as Entertainment ranked among the top successful categories in both US and Finnish markets.

5.1 Success Model

The Success Model is used as a tool to analyze the probability of success of an app, providing a framework for analysis and reliable numerical values for comparison. Comparing only the top ranked list in Table 7 does not provide a reliable conclusion, regarding the probability of success of an app. The Success Model provides more reliable and accurate numerical comparison while taking into consideration the interdependencies between different variables. The model indicates that Google Play has a higher probability of success than Apple App Store. Higher probability of success can be explained by a weaker competition in the store, leading to higher survival rate of the apps among the top 400 list. In other words, the movement of Google Play apps is less dynamic compared to Apple. This observation is valid for the competition among the top app list, which might not be valid for the entire store. Additionally, US market appears to have higher probability of success for paid apps compared to free ones. This can also be explained by higher competition in the free market. Unlike the US market, Finnish market reveals higher probability of success for the free apps compared to paid ones. This can be explained as being a behavior of such a small market compared to the US. Nevertheless, the probability of success is not the final indicator of a successful app. The success of mobile apps is explained precisely in the following section.

Another observation regarding the model results is the behavior of the number of survival days in each of the scenarios. All scenarios have an inversely proportional relationship between the number of surviving apps and the number of survival days. This relationship was not easily observed throughout all the values. Additionally, this behavior cannot be precisely explained using the available data. More variables are needed to explain the correlation of these values with the other factors, such as the rating of the app, price, or even the update version of an app. Nevertheless, the research examined the app categories percentage share for each of the largely deviated count of surviving apps. This has demonstrated categories such as Finance and Music having a high percentage of free apps surviving 34 days in Google Play US store. In addition to the high percentage share of Business, Social Networking and Education Google Play free apps surviving 24 days in the US top list. Unlike the free Google Play in the US, paid Google Play apps have a highest share of Music, Weather, and Education apps surviving 58 days. Whereas, for Google Play Finance and Entertainment paid apps are the most common surviving 23 days in the Finnish market.

5.2 Success of mobile apps

Results show higher success of Android apps than iPhone, which is due to the larger volume of Android devices compared to the iPhone. Any device operating using Android, i.e. smartphone or tablet, can access Google Play, while accessing Apple App Store needs an Apple device, i.e. iPhone, or iPad or iPod. This provides an advantage for Android apps once found on the top app list, therefore more exposure to larger amount of downloads. Free apps are far more successful than paid apps as shown earlier in Table 7.

5.2.1 US market

This study recommends publishing a free app of the categories Social Networking, Productivity, Music, Finance, Education, Sports, Entertainment, and Travel for Android store, and Social Networking and Music apps for the iPhone store. Social Networking and Music are the most popular apps in iPhone and Android, which could be explained by the expectations of Smartphone users. The users experience a need for instant connection with their social circles, and a device that can be used to listen to music. The behavior of the Apple App Store can be explained by the recently increased sales of iPhone compared to iPod and iPad (Statista 2014b). This indicates that users substitute their iPods by iPhone to listen to their music. Music is still a successful category in both free and paid stores, indicating being less price elastic compared to other categories. Thus, Music category does not lose its popularity among the top list of paid apps, since users still are expected to download the Music apps even if a price is charged. Moreover, Books & Reference is less price elastic compared to other categories, as it appears on both free and paid top lists for iPhone. This indicates the success of the marketing of Apple iPhone as a book reader. Additionally, iPhone users prefer to pay for apps in Health & Fitness, Photos & Video, and Games categories. The industry of games is more appealing for iPhone than Android, without even taking into consideration the remaining of the Apple App Store including iPad store, which points out the success of publishing game apps on the iPhone store. The success of iPhone Health & Fitness apps may indicate higher health awareness of the iPhone users compared to the Android users. On the other hand, different categories appeared in the Android paid list such as Weather, Business, Entertainment, and Education. Education is a successful category in both free and paid Android stores, although it does not show much success in the iPhone stores. A possible explanation is that Android devices include also tablets. Business apps for Android appear in the list of top successful apps, which can be also a result of including different types of Android device; thus, using business apps along with tablets can be very convenient to the users.

5.2.2 Finnish market

The Finnish market is a relatively small market compared to the US. Nevertheless, the behavior of the users in different stores reveals some similarities between both markets. These similarities along with differences between the markets are also discussed within this section. Table 9 recommends publishing an app in the Android free store, in order to have a successful app. The table explains the lower popularity of the iPhone apps compared to Android in the Finnish market. Social Networking is the first one on the list of most successful categories preceding Finance, Education, Music, Productivity, Entertainment, Photos and Video, Lifestyle, Games, and News.

The Finnish market indicates a success for Social Networking apps in the free market, while Music is part of the top five apps in all stores except for the paid iPhone apps. The Music apps are more price elastic in the Finnish market than the US one for the iPhone store. On the other hand, The US market has revealed the success in all stores. A possible explanation can be the popularity of Android devices in the Finnish market compared to iPhone which is not the case in the US market. Finance apps have a significant demand in the free stores of the Finnish market, while they were only popular in the Android store of the US market. Moreover, Education still remains on the top successful apps list in the Android stores of the Finnish market, as well as the US one. The reason for the success of finance apps can still be in the fact that different types of Android devices are used in Google Play. Furthermore, tablets are most probably used for educational purposes. Therefore, this increases the demand on the educational apps for Google Play. Unlike the US market, Games gained larger demand in the Android store than in the iPhone. This is related to the higher popularity of Android devices compared to iPhone within the Finnish market.

5.3 Limitations

This section points out the limitations of this analysis, and the significance of the results presented by the previous chapters. These results will be used to provide future research recommendations, taking into consideration the limitations of this study in order to improve the findings. The limitations covered will include limitations of methods used throughout the research as well as constrains with the data.

The quality of the BN model provided depends on the number of variables used in that model, affecting the accuracy of the results. Hence, a variable might be excluded from the model that might then have a major effect on the accuracy of the model. No model will be a perfect representation of real results, not to forget that all models are based on assumptions. The main tool used in this analysis is the naïve BN model, indicating building the model based on expert knowledge. However, the model can be also developed by machine learning where the data variables are added to the model and the software can automatically detect the links between variables representing the dependencies between different variables. BN model can also employ both expert knowledge and machine learning. Using the naïve BN model makes it easy to construct a simple model representing the complexity of the dynamic mobile app stores along with less need of computational power. On the other hand, correlations between some variables can be easily ignored while building and simulating the model.

The research scope did not take into consideration the correlation of various app parameters with the success of an app. This has been researched earlier by Kim (2012) who studied the correlation between the demand and other parameters such as age, size, price, updates, and customer ratings. Likewise, positive feedback has shown a major effect on the user downloads of an app (Pagano & Maalej 2013).

Another major limitation of the data is the unavailability of the price information for each app. As a result, the analysis was based on the count of app downloads assuming equal revenue per download for all apps. This is unrealistic when it is applied in practice. Nevertheless, this still provides a good indicator of success of an app. While the data only included free and paid apps, this study does not take into consideration the Free-in-app and Paid-in-app apps. Distimo (2013) reports that free-in-apps share the majority of revenue in Apple App Store and Google Play. In-app purchases even accounted for 79% of the revenue generated by Apple App Store in March 2014 (Distimo, 2014). Recently, Free-in-app has become the most common business model (Fredholm & Gunnarsson 2013). Moreover, the primary goal of this thesis is to compare Apple App Store and Google Play, although the data provided for Apple App Store only included the iPhone market, thus limiting the results to iPhone as a representative of the Apple App Store. This limitation ignores the iPad Apple App Store, thus providing less number of downloads for the Apple App Store. Additionally, tablets are included in the Google Play while not included in the iPhone Apple App Store. These different devices introduce various user preferences based on the type of device.

The dataset provided included data about the categories of the apps, for each app one category was listed. However, some apps were not listed in a certain category due to some data errors. In practice, an app can be listed under two categories in Apple App Store. In this case, an app provider defines one of the categories as the primary one. Therefore, this analysis only included the primary category of the apps into consideration. In addition, the time-frame presented in this study is two months, which is not a large time-frame. The results and models presented in this thesis can be further extended and applied on larger datasets. Larger time-frame will enable the validation of the model, while taking into consideration that the app store market is dynamic and changes quickly. Thus, demand behavior needs to be validated using updated datasets.

The previously mentioned time-frame limitation restricts the identification of the time of app introduction into the app store. The lack of this information opens up the question of defining the precise phase of the app lifecycle, which is expected to have a direct effect on the performance of the app. For instance, an app might have been introduced two months earlier, and it might have been traced surviving only two days during the beginning of this dataset. This situation remains unchanged at any time-frame for any dataset, even if the time-frame was extended, there will remain apps that were introduced before the time-frame of the dataset. On the other hand, the results of analyzing the behavior of app categories generate more accurate results than the ones based specifically on the behavior of each app. Nevertheless, additional app information, such as the dates when the app was published as well as the dates of updated versions, will help excluding the apps with unidentified lifecycle.

6 Conclusion

This chapter summarizes the key findings and recommendations for future research in the field. Chapter 5 discussed user preferences for the mobile application (app) stores. Users of various app stores have shown different preferences, even for the same geographical market. These preferences should drive an app developer while designing new apps, to meet these needs, thus maximizing the revenue. Accordingly, the conclusion presents the key findings related to user behavior concerning both app stores.

6.1 Key Findings

This research focused on the app developer perspective. It examined differences in ranking results between the average download per app for each category compared to the cumulative percentage of downloads for all apps in each category. For instance, although Games appear on the top of the list of categories with a high percentage share in terms of cumulative downloads, it is still not one of the top successful apps in terms of probability of success defined by this research. This stems from the probability of success being calculated on the basis of the average success per app for each category, not as a cumulative parameter for the entire category. The cumulative parameter is usually of interest to the app store provider, not to app developers. Hence, the study provides guidelines helping developers predicting the market demand. It also uses an analogy of the Ising Model for approximating the behavior of apps within the app stores. The analogy examines low correlation between the individual behaviors of Google Play apps and indicates that it is a more open platform compared to Apple App Store, as a result of less control over the market.

This thesis assumes that a successful app is one that remains 15 consecutive days on the top list, guaranteeing that the demand for the app is high enough to keep it on the list. Thus, this thesis presents the Success Model for apps in both Google Play and Apple App Store. The model utilizes Bayesian Network, a type of probabilistic graphical models, to convey the probability of success of an app in both stores. This model presents simulations that have been developed using the dataset available and provides a useful tool for app developers. The results of the simulations have shown higher probability of success for a Google Play app than an Apple App Store app in the US market. Additionally, it shows that paid apps have higher probability of success than free ones in the US market. Although the Finnish market, similar to the US, revealed higher probability of success for Google Play apps than the Apple App Store, it also yielded opposite results of higher probability of success for free apps than paid apps. This may result from the small market and low popularity of Apple products in Finland. Paid Google Play apps in the US market have shown high probability of success for Entertainment, Music, and Business categories compared to other apps in the US market. The Finnish market shows different user demand with high probability of success for Finance and Entertainment listed as paid, and Social Networking apps listed as free.

The probability of success of an app does not necessarily correspond to high revenue for an app. Consequently, this research further defines a success indicator for mobile apps. This success indicator multiplies the probability of success of an app and the average expected number of downloads for the app. Both values depend on the app store, app type and the category where the app is listed. The top successful apps in the US market included Social Networking, Productivity, Music, Finance, Education, Sports, Entertainment, and Travel. The top successful apps in Finland included Social Networking, Finance, Education, Music, Productivity, Entertainment, Photos and Video, Lifestyle, Games, and News. Google Play brings higher success indicators than Apple App Store both in US and Finnish markets. Additionally, the success indicator is higher regarding free over paid apps.

This conclusion contributes to recommendations for app developers when developing and publishing an app, as well as building marketing strategies for a mobile app. Furthermore, it suggests a framework to identify successful apps in mobile app stores. This framework can be further adopted and implemented on a different scale by various stakeholders.

6.2 Future Research

This thesis provides a guideline for mobile app developers benefits for taking strategic decisions regarding investments in mobile apps, and points out the benefit of app developer. It also provides a framework of analysis in order to examine similar types of datasets. Although this research is based on a dataset collected over two months, the average lifespan of an app is around 14 months (Wolonick 2013). Hence, it is recommended that the time-frame of the dataset be extended, for example a dataset of two years, enabling various analytical approaches. Additionally, this study analyzed the success of an app using the number of downloads, assuming that all apps have a similar price. It would be more precise to introduce the actual price values of apps.

The BN was used as an analysis framework, offering a dynamic environment by understanding uncertainties regarding the success of an app. The framework is very useful specifically for the reasoning and predictive functionalities. The results have demonstrated success of this method regarding the scope of the analysis, which can be further extended by allowing the BN tool to automatically determine the correlations between variables. This can be achieved with various tools available in the market. Moreover, other variables can be taken into consideration, such as app age, size, update, version and price, which will help consider the correlations between these variables and the success of an app. Finally, the same analysis might provide different results based on the chosen geographical markets chosen. Therefore, comparing results in different markets is essential.

7 References

AgenaRisk 2004, 'Agena's Bayesian Network Technology. Agena White Paper W0904/01, Version 01.00'.

Anderson, C 2004, 'The Long Tail', *Wired*, vol 12, no. 10.

Anderson, C 2006, 'The Long Tail', *London: Random House*.

Anthes, G 2011, 'Invasion of the Mobile Apps', *Communications of the acm*, vol 54, no. 9, pp. 16-18.

App Annie 2014, '*App Annie Index – Market Q1 2014: Revenue Soars in the United States and China',* Available: http://blog.appannie.com/app-annie-index-market-q1-2014/, Last accessed 30 April 2014

Apple 2009, '*Apple Answers the FCC's Questions',* Available: http://www.apple.com/hotnews/apple-answers-fcc-questions/, Last accessed 30 April 2014.

Apple 2010, '*App Store Review Guidelines',* Available: http://stadium.weblogsinc.com/engadget/files/app-store-guidelines.pdf, Last accessed 30 April 2014.

Apple 2013, '*App Store Tops 40 Billion Downloads with Almost Half in 2012',* Available: http://www.apple.com/pr/library/2013/01/07App-Store-Tops-40-Billion-Downloads-with-Almost-Half-in-2012.html, Last accessed 30 April 2014.

Ayalew, R 2011, 'Consumer behavior in Apple's App Store', Master of Science Thesis, Human Computer Interaction, Uppsala University.

Barnes, SJ 2002, 'The Mobile Commerce Value Chain: Analysis and Future Developments.', *International Journal of Information Management*, vol 22, no. 2, pp. 91-108.

Basole, RC 2009, 'Visualization of Interfirm Relations in a Converging Mobile Ecosystem', *Journal of Information Technology advance online publication*.

Begelman, G, Keller, P & Smadja., F 2006, 'Automated Tag Clustering: Improving search and exploration in the tag space', *The Collaborative Web TaggingWorkshop at WWW*.

Bingham, A & Spradlin, D 2011, *The Long Tail of Expertise*, Pearson Education.

Brown, S 2005, 'The tripping point', *Marketing Research*, vol 17, no. 1, pp. 8-13.

Buellingen, F & Woerter, M 2004, 'Development perspectives, firm strategies and apps in mobile commerce', *Journal of Business Research*, vol 57, no. 12, pp. 1402-1408.

Carare, O 2012, 'The impact of bestseller rank on demand: Evidence from the App market', *International Economic Review*, vol 53, no. 3, pp. 717-742.

Chevalier, J & Goolsbee, A 2002, 'Measuring prices and price competition online:Amazon and Barnes and Noble', *National Bureau of Economic Research*.

Cohen, K. 2012, '*To Google Play or Not to Google Play? That is the Andriod App Store Question',* Available: http://marketingland.com/to-google-play-or-not-to-google-play-that-is-the-app-store-question-25689, Last accessed 30 April 2014.

Cuadrado, F & Dueñas, JC 2012, 'Mobile Application Stores:Success Factors, Existing Approaches,and Future Developments', *Communications Magazine, IEEE*, vol 50, no. 11, pp. 160-167.

De Vere, K. 2012, '*Tracking growth: the iTunes app store vs Google Play'.* Available: http://www.insidemobileapps.com/2012/09/26/tracking-growth-the-itunes-app-store-vs-google-play/, Last accessed 30 April 2014.

Dediu, H. 2011, '*Nokia vs. Android',* Available: http://www.asymco.com/2011/08/22/nokia-vs-Android/, Last accessed 30 April 2014.

Dediu, H. 2012, '*The phone market in 2012: a tale of two disruptions',* Available: http://www.asymco.com/2012/05/03/the-phone-market-in-2012-a-tale-of-two-disruptions/, Last accessed 30 April 2014.

Dediu, H. 2014, '*When will smartphones saturate?',* Available: http://www.asymco.com/2014/01/07/when-will-smartphones-saturate/, Last accessed 30 April 2014.

Distimo 2013, ' *Free Apps Generate the Most Revenue, Google Play Grows.',* Available: http://www.technobuffalo.com/2013/12/28/free-apps-generate-the-most-revenue-google-play-grows/, Last accessed 30 April 2014.

Distimo 2014, ' *Distimo: In-app purchases accounted for 79% of iOS revenue in January Read more: Distimo: In-app purchases accounted for 79% of iOS revenue in January',* Available: http://www.fiercedeveloper.com/story/distimo-app-purchases-accounted-79-ios-revenue-january/2014-03-24, Last accessed 30 April 2014.

Driver, HE & Kroeber, AL 1932, 'Quantitative expression of cultural relationships', *University of Caltfornia Publications in American Archeology and Ethnology*, pp. 2111-256.

Eaton, BD, Elaluf-Calderwood, SM & Sørensen, C 2010, 'A Methodology for Analysing Business Model Dynamics for Mobile Services using Control Points and Triggers', *Intelligence in Next Generation Networks (ICIN)*, Berlin.

Ecosys 2004, 'Business models in telecommunications', First results on the economics of the new service provisioning in 3G and beyond. The CELTIC initiative, technoeconomics of integrated communication systems and services

Elberse, A & Oberholzer-Gee, F 2007, 'Superstars and Underdogs:An Examination of the Long Tail Phenomenon in Video Sales', *ASSA Conference Paper, Harvard Business School Working Paper No. 07–015*, Boston, MA.

Farrell, J & Saloner, G 1985, 'Standardization, Compatibility, and Innovation', *The RAND Journal of Economics*, vol 16, no. 1, pp. 70-83.

Fredholm, M & Gunnarsson, M 2013, 'Mobile Apps Industry', Master of Science Thesis, Business Administration, School of Economics and Management, Lund University, Lund.

Garg, R & Telang, R 2013, 'Inferring App demand from publicly available data', *MIS Quarterly*, vol 37, no. 4, pp. 1253-1264.

Ghose, A & Han, SP 2012, 'Estimating demand for mobile apps in the new mobile economy', *Thirty Third International Conference on Information Systems*, Orlando.

Gladwell, M 2000, *The Tipping Point*.

Gonçalves, V, Walravens, N & Ballon, P 2010, '"How about an App Store?" Enablers and Constraints in Platform Strategies for Mobile Network Operators', *Mobile Business and 2010 Ninth Global Mobility Roundtable (ICMB-GMR)*, Athens.

Heymann, P & Garcia-Molinay, H 2006, 'Collaborative Creation of Communal Hierarchical Taxonomies in Social Tagging Systems', Computer Science Department, Stanford University, InfoLab Technical Report 2006-10.

Holzer, A & Ondrus, J 2011, 'Mobile application market: A developer's perspective', *Telematics and Informatics*, vol 28, no. 1, pp. 22-31.

Ho, H-Y & Syu, L-Y 2010, 'Uses and Gratifications of Mobile Application Users', *International Conference on Electronics and Information Engineering (ICEIE 2010)*, Kyoto.

Hyrynsalmi, S, Makila, T, Jarvi, A, Suominen, A, Seppanen, & Knuutila, T 2012, 'App Store, Marketplace, Play! An Analysis of Multi-Homing in mobile software ecosystems', *IWSECO*.

Ingraham, N . 2013, '*Apple announces 1 million apps in the App Store, more than 1 billion songs played on iTunes radio',* Available: http://www.theverge.com/2013/10/22/4866302/apple-announces-1-million-apps-in-the-app-store, Last accessed 30 April 2014.

Ising, E. 1925, 'A contribution to the theory of ferromagnetism.', Z. Phys 31.1, pp. 253-258.

Jardine, M. 2009, '*A Well-Timed Letter of Rejection',* Available: http://tapbots.com/blog/app-store/a-well-timed-letter-of-rejection, Last accessed 30 April 2014.

Johnson, G, Scholes, K & Whittington, R 2008, *Exploring Corporate Strategy, EIGHTH EDITION*, Harlow, pp. 333

Karvonen, J & Warsta, J 2004, 'Mobile Multimedia Services Development- Value Chain Perspective', *3rd international conference on Mobile and ubiquitous multimedia*.

Katz, ML & Shapiro, C 1985, 'Network Externalities, Competition, and Compatibility', *American Economic Association*, vol 75, no. 3, pp. 424-440.

Katz, ML & Shapiro, C 1994, 'Systems Competition and Network Effects', *The Journal of Economic Perspectives*, vol 8, no. 2, pp. 93-115.

Kaufman, L & Rousseeuw, PJ 1990, *Finding Groups in Data: an Introduction to Cluster Analysis*, John Wiley & Sons.

Kenney, & Pon, B 2011, 'Structuring the Smartphone Industry: Is the Mobile Internet OS Platform the Key?', *Journal of Industry, Competition and Trade*, vol 11, no. 3, pp. 239-261.

Kilkki, K 2007, 'A practical model for analyzing long tails.', *First Monday*, vol 12, no. 5.

Kilkki, K & Kalervo, M 2004, 'KK-law for Group Forming Services', *International Symposium on Services and Local Access*, Edinburgh.

Kim, MJ 2012, 'The Interdependence between Smartphones and Apps: The Role of Platforms', *working paper*, Available: http://www.econ.umn.edu/~kimx1423/_content/JMP_MinjungKim.pdf , Last accessed 30 April 2014.

Kimbler, K 2010, 'App Store Strategies for Service Providers', *Intelligence in Next Generation Networks (ICIN)*, Berlin.

Kokabha, MR 2012, 'Mobile Application Ecosystems from the Application Developers' Perspectives: the Cases of App Store and Android Market', Master of Science Thesis, International Business, School of Economics, Aalto University, Helsinki.

Lee, G & Raghu, TS 2011, 'Product Portfolio and Mobile Apps Success: Evidence from the App market', *AMCIS*, Detroit, Michigan.

Liebowitz, SJ & Margolis, SE 1994, 'Network Externality: An Uncommon Tragedy', *The Journal of Economic Perspectives*, vol 8, no. 2, pp. 133-150.

Lim, SL & Bentley, PJ 2013, 'Investigating App Store Ranking Algorithms using a Simulation of Mobile App Ecosystems', Software Systems Research Centre, Bournemouth University,United Kingdom and Department of Computer Science, University College London, United Kingdom.

Lin, F & Ye, W 2009, 'Operating System Battle in the Ecosystem of Smartphone Industry', *International Symposium on Information Engineering and Electronic Commerce*.

Liu, AX, Wang, Y, Chen, X & Jiang, X 2013, 'Understanding the diffusion of mobile digital content: a growth curve modelling approach', *Information Systems and e-Business Management*, pp. 1-20.

Mathes, A 2004, 'Folksonomies-Cooperative Classification', *Computer Mediated Communication - LIS590CMC (Doctoral Seminar),Graduate School of Library and Information Science,University of Illinois Urbana-Champaign, December.*

Millen, D, Feinberg, J & Kerr, B 2006, 'Social bookmarking in the enterprise', *Proceedings of the Special Interest Group on Computer-Human Interaction conference on Human Factors in computing systems*.

Moore, JF 1993, 'Predators and Prey: A New Ecology of Competition', *Harvard business review*, vol 71, pp. 75-83.

Newman, MEJ 2003, 'The Structure and Function of Complex Networks', *2003 Society for Industrial and Applied Mathematics*, vol 45, no. 2, pp. 167-256.

Nyika, J 2010, 'The Mobile Applications Ecosystem: A Primer', *White paper. Triton Federal Solutions, [Online]* Available: http://www.ppc.com/assets/pdf/white-papers/Mobile-Apps-Ecosystem.pdf, Last accessed 30 April 2014

Oren, SS & Smith, SA 1981, 'Critical Mass and Tariff Structure in Electronic Communications Markets', *The Bell Journal of Economics*, vol 12, no. 2, pp. 467-487.

Pagano, D & Maalej, W 2013, 'User Feedback in the AppStore: An Empirical Study', *Requirements Engineering Conference (RE)*, Rio de Janeiro.

Pearl, J 1988, *Probabilistic Reasoning in Intelligent Systems*, San Francisco, pp. 1-28.

Pearl, J & Russel, S 2001, 'Bayesian networks', *Handbook of Brain Theory and Neural Networks*, vol Report (R-277), pp. 157-160.

Petsas, T, Papadogiannakis, A & Polychronakis, M 2013, 'Rise of the Planet of the Apps: A Systematic Study of the Mobile App Ecosystem', *The 2013 conference on Internet measurement conference*, New York.

Rao, B & Jimenez, B 2011, 'A Comparative Analysis of Digital Innovation Ecosystems', *Technology Management in the Energy Smart World (PICMET)*, Portland, OR.

Reed, D 1999, '"That Sneaky Exponential—Beyond Metcalfe's" , Available: http://www.contextmag.com/archives/199903/DigitalStrategyReedsLaw.asp', *Context*, Last accessed 30 April 2014

RiskIQ 2014, '*RiskIQ Reports Malicious Mobile Apps in Google Play Have Spiked Nearly 400 Percent',* Available: http://www.riskiq.com/company/press-releases/riskiq-reports-malicious-mobile-apps-google-play-have-spiked-nearly-400, Last accessed 30 April 2014.

Rogers, EM 2003, *Diffusion of Innovations. (5th ed.)*, Free Press, New York.

Rohlfs, J 1974, 'A Theory of Interdependent Demand for a Communications Service', *The Bell Journal of Economics and Management Science*, vol 5, no. 1, pp. 16-37.

Rosen, S 1981, 'The economics of superstars', *American Economic Review*, vol 71, p. 845–58.

Rowinski, D. 2013a, '*Google Play Will Beat App Store To 1,000,000 Apps',* Available: http://readwrite.com/2013/01/08/google-play-to-hit-1-million-apps-before-apple-app-store#awesm=~otvLC454uYVyn2, Last accessed 30 April 2014.

Rowinski, D. 2013b, '*Google Play Hits One Million Android Apps',* Available: http://readwrite.com/2013/07/24/google-play-hits-one-million-Android-apps#awesm=~otvNCeD1utKrGB, Last accessed 30 April 2014.

Sarno, D. 2012, '*Apple Says It Gets 26,000 Submissions of Apps Every Week',* Available: http://www.hispanicbusiness.com/2012/3/20/apple_says_it_gets_26000_submissions.htm, Last accessed 30 April 2014.

Sawant, S 2010, 'Mobile marketing strategies for B2C companies', Master of Science Thesis, System Design and Management, MIT - Massachusetts Institute of Technology.

Shaughnessy, H. 2013, '*Can These Competitors Break Apple and Google's Stranglehold On The Mobile OS?',* Available: http://www.forbes.com/sites/haydnshaughnessy/2013/04/09/ios-vs-Android-can-competitors-break-apple-and-googles-stranglehold/, Last accessed 30 April 2014.

Shvachko, K, Kuang, H, Radia, S & Chansler, R 2010, 'The Hadoop Distributed File System', *2010 IEEE 26th Symposium on Mass Storage Systems and Technologies (MSST)*.

Skogsberg, P 2013, 'Quantitative indicators of a successful mobile application', Master of Science Thesis, Communication Systems, School of Information and Communication Technology, KTH Royal Institute of Technology, Stockholm, Sweden.

Sood, A & Tellis, G 2005, 'Technological evolution and', *Journal of Marketing*, vol 69, no. 3, pp. 152-168.

Spath, H 1985, *Cluster Dissection and Analysis: Theory, FORTRAN programs, Examples*, Ellis Horwood Ltd.

Statista 2014a, ‘Number *of available applications in the Google Play Store from December 2009 to July 2013',* Available: http://www.statista.com/statistics/266210/number-of-available-apps-in-the-google-play-store/, Last accessed 30 April 2014.

Statista 2014b, '*iPhone, iPad and iPod sales from 1st quarter 2006 to 2nd quarter 2014 (in million units)',* Available: http://www.statista.com/statistics/253725/iphone-ipad-and-ipod-sales-comparison/, Last accessed 30 April 2014.

Sterling, G. 2013, '*Pew: 61 Percent In US Now Have Smartphones',* Available: http://marketingland.com/pew-61-percent-in-us-now-have-smartphones-46966. Last accessed 30 April 2014

Tarnacha, A & Maitland, C 2006, 'Standards Competition and the Role of Intermediationi n the Supply of Mobile Applications', *34th Annual Research Conference on Communication, Information and Internet Policy (TPRC)*.

Tuunainen, VK & Tuunanen, T 2011, 'IISIn - A Model for Analyzing ICT Intensive Service Innovations in n-Sided', *System Sciences (HICSS), 2011 44th Hawaii International Conference*, Kauai, HI.

Tuunainen, VK, Tuunanen, T & Piispanen, J 2011, 'Mobile Service Platforms, Comparing Nokia OVI and App Store with the IISI^n model', *ICMB 2011, The 10th International Conference on Mobile Business*, Como.

Venture beat 2013, '*Comparing Apples and Googles: The App Store vs. Google Play (infographic)',* Available: http://venturebeat.com/2013/07/17/comparing-apples-and-googles-the-app-store-vs-google-play-infographic/, Last accessed 30 April 2014.

Verkasalo, H 2009, 'Open Mobile Platforms Modeling the Long-Tail of Application Usage', *Fourth International Conference on Internet and Web Apps and Services*.

Wang, L & Wang, H 2013, 'Marketing Strategy for Mobile apps on iOS', Master of Science Thesis, International Business School, Jonkoping University.

Want, R 2011, 'Android: Changing the Mobile Landscape', *Pervasive Computing, IEEE*, vol 10, no. 1, pp. 4-7.

Wolonick, J. 2013,' *Where is the booming app market going?',* Available: http://www.usatoday.com/story/tech/2013/03/07/booming-app-market-minyanville/1970245/, Last accessed 30 April 2014.

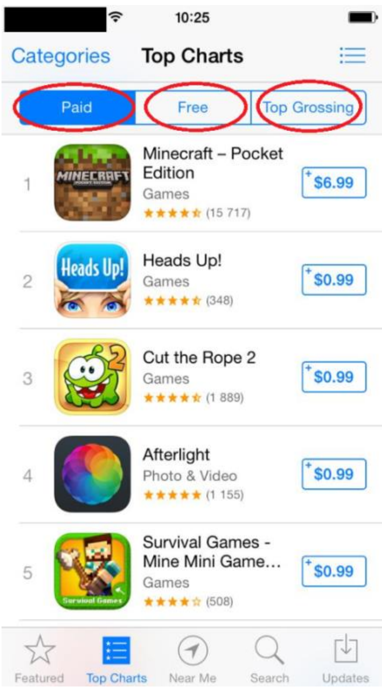
Yamakami, T 2011, 'A Three-dimension Analysis of Driving Factors for Mobile Application Stores: Implications of Open Mobile Business Engineering', *2011 Workshops of International Conference on Advanced Information Networking and Apps*, Biopolis.

Zheng, Z, Zhu, J & Lyu, MR 2013, 'Service-generated Big Data and Big Data-as-a-Service: An Overview', *IEEE International Congress on Big Data*.

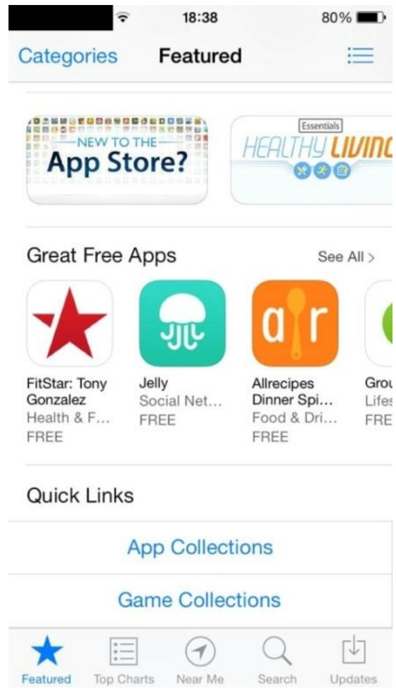
Zhong, N & Michahelles, F 2013, 'Google Play Is Not A Long Tail Market', *Proceedings of the 28th Annual ACM Symposium on Applied Computing*.

8 Appendix

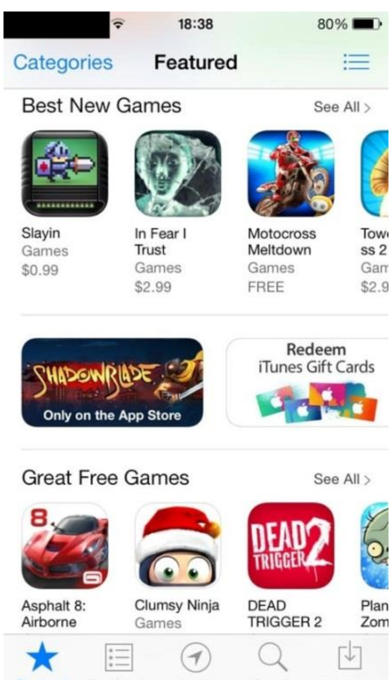
Appendix A – Apple App Store Top charts and Featured page



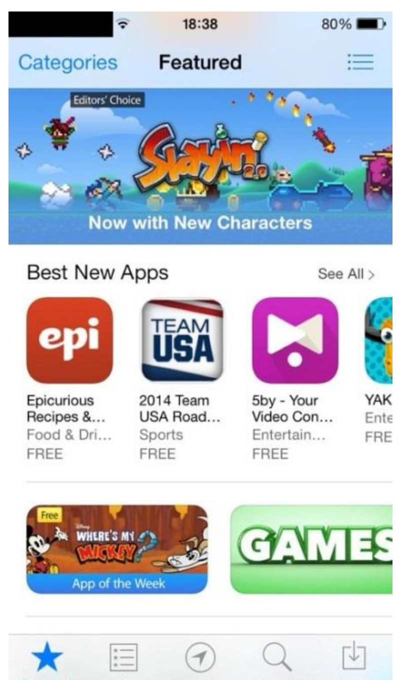
Top Charts App Store



Featured Page (1) App Store

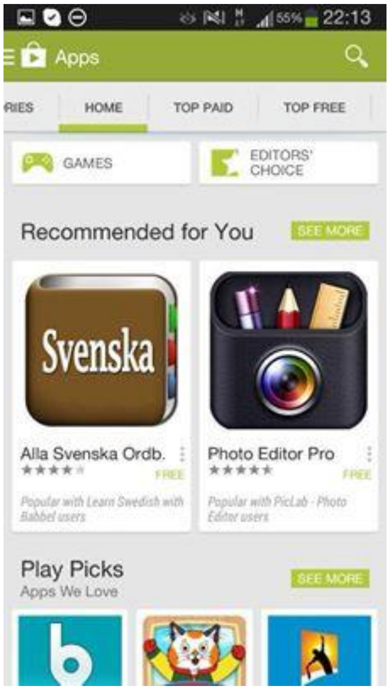


Featured Page (2) App Store

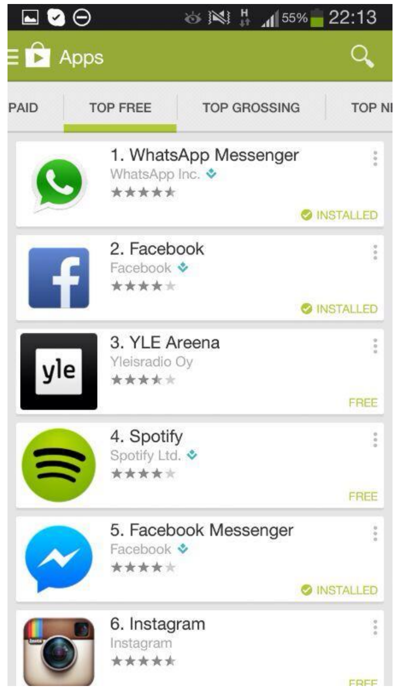


Featured Page (3) App Store

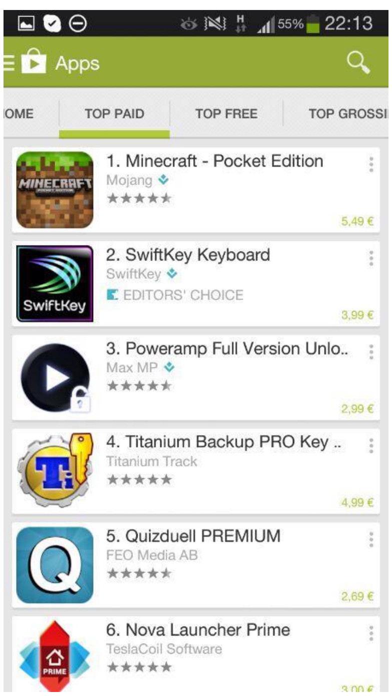
Appendix B – Google Play Home page and top application lists



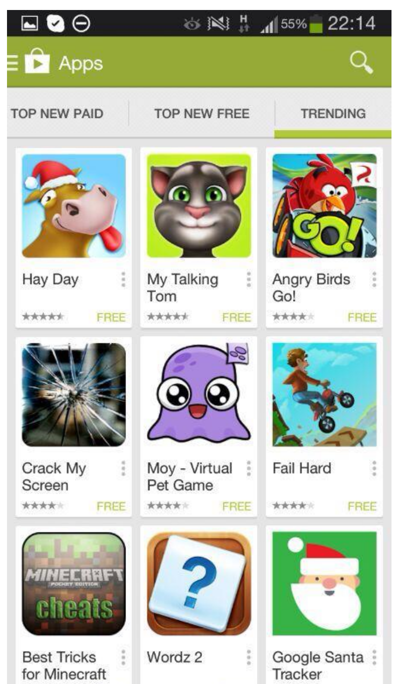
Home Page Google Play



Top Free Google Play

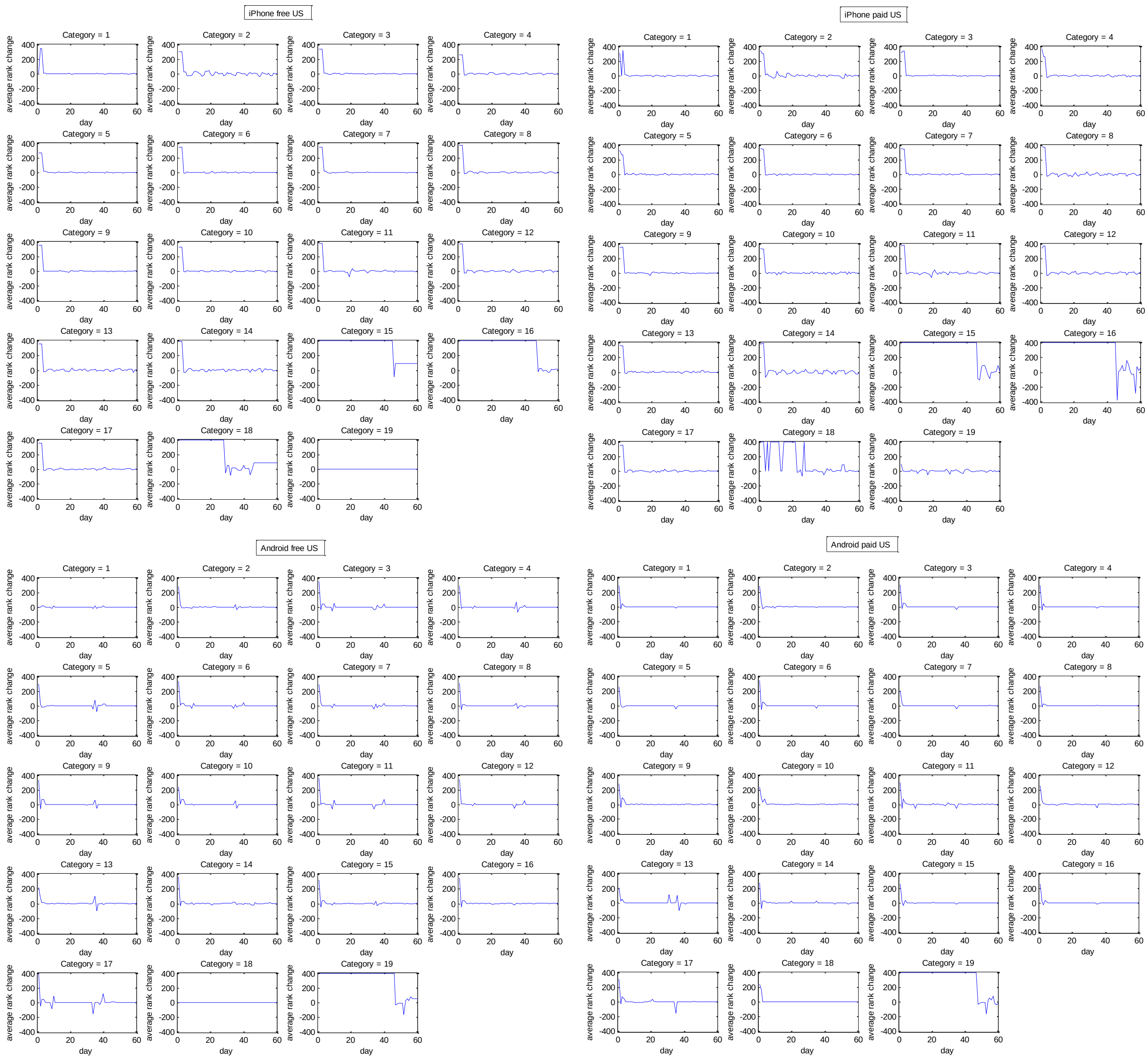


Top Paid Google Play



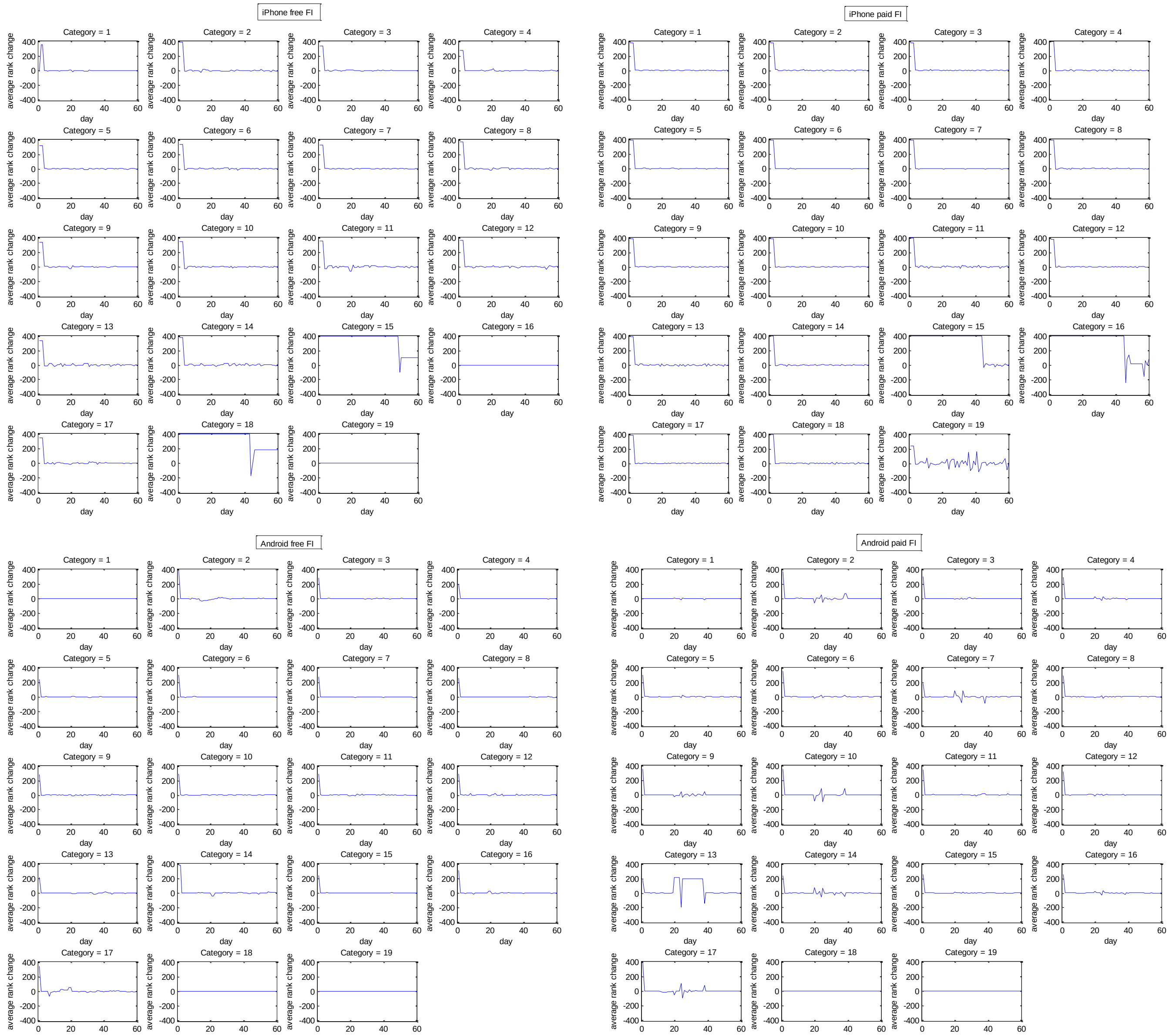
Trending Google Play

Appendix C – Average rank change per category for each app store in the US market



Average rank change US market

Appendix D - Average rank change per category for each app store in the Finnish market



Average rank change FI market