School of Production Engineering and Management

Technical University of Crete
Chania, Greece

Evaluation of different customer experience metrics in a GameTech company

In partial fulfillment of the requirements for the Degree Master of Science in Technology and Innovation Management

Submitted by: **ELENI CHATZIMITSOU**

Supervisor: Prof. Evangelos Grigoroudis

Members of the Committee: Prof. Nikolaos Matsatsinis

Assoc. Prof. Stelios Tsafarakis

March 2021

Acknowledgments

I wish to express my sincere appreciation and gratitude to my supervisor, Professor Evangelos Grigoroudis, for his convinced guidance, trust and encouragement to continue when the road got tough.

I would also like to thank Professor Stelios Tsafarakis for his support at the first stages of this thesis. I am also thankful for the support I received by all team members of Stoiximan and especially by George Kourakos and Niovi Kalantzi for their trust and support.

Last but not least, the biggest “thank you” goes to my two lovely daughters for their patience for the moments I was too busy studying and writing this thesis.

Abstract

The purpose of this thesis is to examine the predictive ability of different customer experience metrics for firm's growth. Customer experience, incorporating the whole customer journey along with the satisfaction and loyalty components, is arguably a factor driving firm growth.

The Net Promoter Score (NPS) metric, the overall Satisfaction and the Expectations fulfilment have been tested in this thesis against firm revenue and other metrics building up firm growth.

The NPS metric is widely adopted with more than two thirds of Fortune 1000 companies using the specific metric, claiming to be correlated with company's revenue growth. Yet, a significant number of academics and professional argue about its actual global correlation with firm growth, suggesting among other the "overall satisfaction" and loyalty as better metrics for firm growth prediction. Noticeably, different metrics may apply better to different industries, where especially the gambling industry involves complicated purchase decision making, often driven by satisfaction, hedonism and addiction.

Upon reviewing relevant literature, a research was conducted based on Stoiximan, a GameTech company engaged in online gambling, users raw data and answers related to the NPS, overall satisfaction and expectations fulfillment. According to the research findings, the NPS metric does not work well as a firm growth predictor, confirming the literature findings. Yet, this indicates that firm growth may be more accurately predicted via implementing a combination of the NPS metric along with a different metric in the online gambling industry.

Table of Contents

List of figures

List of tables

1. Introduction

This thesis addresses the topic of predicting a firm's growth by measuring customers' loyalty and satisfaction. Upon reviewing the relative literature, a case study for Stoiximan, a Kaizen Gaming (formerly named GML Interactive Ltd) owned firm is examined. Not only the topic, but also the industry examined render this thesis valuable.

Consumer behavior has proven to be a hot topic both in academia and in businesses over the past few years, with the focus shifting towards offering personalized experiences, rather than just services. Consequently, more resources are spent nowadays towards analyzing customer satisfaction and designing a holistic customer journey in order to build up loyal customers pools and thrive in an extremely competitive environment.

As per industry, the Gambling industry has been thriving over the past couple of decades with the insertion of online Gambling firms operating worldwide and offering sports gambling and casino options. The online gambling industry has been disrupting the physical gambling and casino industry, raising a total revenue of more than $100 billion in Europe and the US annually. (European Gaming and Betting Association, 2020), (Global Market Insights, 2020)

1.1. Purpose of this study

The purpose of this study is to explore the predictive ability of different customer experience metrics for company's growth. A research will be conducted in a sports gaming & betting firm. More specifically, this study will focus on analyzing the correlation of three customer experience metrics to company's revenue growth. In particular, the customer feedback metrics suggested for this research are the following: Net Promoter Score (NPS), the Overall Satisfaction and the Expectancy confirmation / disconfirmation question.

NPS is widely adopted with more than two thirds of Fortune 1000 companies using the metric, claiming to be correlated with company's revenue growth. Yet, the number of those who dispute it and argue that an "overall satisfaction" or a “loyalty” indicator is a better metric for growth is increasing. Therefore, initially in this study, a literature review on NPS as an indicator of measuring a company's growth by presenting the arguments of its supporters and of those who dispute it, will be conducted.

The research will examine the correlation of the responses of the Sportsbook and Casino customers from each question and their actual behavior three months after. Seven dimensions are chosen as depended variables (e.g. average revenue per customer, average amount per bet, number of bets placed by the player).

The results of this study will focus on evaluating the value and reliability of customer-based metrics in predicting revenue growth.

1.2. The problem

Loyalty and consumer feedback metrics (CFMs) as Morgan and Rego (2006) name them, are used both in the academic research and in the real business world in order to predict future purchase behavior and companies’ revenue growth. (Morgan & Rego, 2006)

A widely applied and embraced by the business world metric for evaluating a firm’s growth is the Net Promoter Score, abbreviated NPS, since it was introduced in a Harvard Business Review article back in 2003. The NPS focuses on measuring customer loyalty as a measure of the likelihood a customer might suggest a firm to another consumer. (Reichheld, 2003)

Other metrics focus solely on customer satisfaction or retention and less often used as firm growth predictors. Yet, lately there has been an increasing debate around the NPS metric superiority over other customer satisfaction, loyalty or retention metrics as a firm growth and revenue predictor.

Given that revenue and firm growth predictors do not dictate a firm’s strategic planning, the metrics used as predictors do not affect the firm growth or revenue

unilaterally, and therefore one should keep in mind that the outcome of this thesis cannot be interpreted as a means of formulating strategic planning.

1.3. Research goals

The main research goal of this thesis is to identify which consumer feedback metric or combinations of metrics a sports gaming and betting firm should monitor in order to be able to predict its revenue growth and increase its business performance.

In order to meet the research goal, primary data collected by a real game tech users have been used for evaluating their experience, namely their satisfaction and loyalty. The data have been evaluated under the NPS method. Upon evaluating user's experience, NPS results have been examined per their linear correlation to the game tech firm fiscal indices used for evaluating revenue and growth.

1.4. Thesis structure

This thesis consists of a total of five chapters.

The first chapter is an introduction to the thesis, where the thesis purpose is presented, the problem examined is introduced and the research goals and thesis structure are provided.

The literature review findings on the field of customer experience, satisfaction and loyalty and their connection with firm growth are provided and discussed in the second chapter of this thesis.

The third chapter consists of the online gambling industry review, where various gambling products and different gambling markets are reviewed, concluding in the Greek gambling market.

Upon presenting the gambling firm analyzed, the research structure results are presented in the fourth chapter. The results for every customer experience, satisfaction and loyalty dimension examined, including all the component variables, are presented and discussed, while correlations and linear regression results are provided as well towards responding the research questions.

The conclusions drawn are provided in the fifth chapter, including limitations faced and suggestions for future research.

2. Literature Review

The literature review in the field of customer experience finding is presented and discussed in this chapter. Customer experience, satisfaction and loyalty are thoroughly discussed before providing literature review findings on different customer experience metrics. Before proceeding with the literature review findings, it is worth pointing out that they have been utilized in order to design questionnaires distributed to the sports gaming and betting firm users.

2.1. Customer experience

Customer experience has been a hot topic over the past years in various management fields, without a unanimously accepted definition being provided by literature. Yet, many businesses are mobilized towards improving their customers' experience, as a means of boosting profitability and improving organizational performance in total. In that context, many businesses adopt a customer-centric strategy in order to gain and maintain a sustainable competitive advantage.

Elaborating the latter, creating and maintaining a competitive advantage is a far more difficult task than ever before, as firms operate in highly competitive environments, where various tools and methods are put in use in order to make the difference and succeed.

The traditional means of competitive advantage creation are arguably obsolete and incumbent firms are arguably inherent competitive advantage holders. Nowadays, creating competitive advantage stems from hearing the customers' voice and making all the necessary amendments to offer them unique and utmost experiences. (Pine & Gilmore, 1999)

There are numerous examples of firms bringing customer experience to the business forefront, such as Uber, Airbnb, Amazon and others. These firms offer customers a communication channel through which their voice is heard and through customers' evaluation and recommendations, the experiences offered are tailored to the constantly changing consumer needs. Under that context, and especially under the

accelerating digital transformation of firms and markets, customers seek personalized experiences, offering them increased ease and safety.

According to Abbott (1955), experiences are created through customers' interaction with physical objects, while when it comes to services provision, experiences are driven by the value customers receive when interacting with a firm. Therefore, customer experience may prove to be rather difficult to interpret when comparing services and products provision. When people buy services, they are delivered with value, whose evaluation is subject to how each different customer perceives value. (Abbott, 1955)

Moreover, purchasing services creates a different set of emotions to customers, who are the inner drivers of customer experience. Thus, customer experiences are heavily personalized, as different customers experience different feelings when interacting with the same business. Moreover, customers build up a relation when interacting repeatedly with a firm, which, upon positive experiences, may lead to emotional engagement. (Grigoroudis & Siskos, 2010)

A successful customer experience may be described as unique and unforgettable, although in different business fields experiences may be judged upon more robust criteria, such as service time, waiting, service outcome etc. In that context, various tools and methods have been developed for measuring and evaluating customer experience, which may not apply horizontally to every market. (Krassadaki & Matsatsinis, 2015)

Essentially, the customer experience refers to all the experiences, thoughts and emotions a customer has for a firm. Customers are not connected to businesses only when interacting for purchasing a product or some services. Instead, customers and businesses cultivate a deeper relationship, where experiences are driven by consumer needs, which drive them towards a business. When referring to customers taking a purchase decision, this is based on their needs. From this point on and on, coming back to the same business for catering for their needs, is driven by positive customer experiences. This course of logic leads to cultivating commitment and emotional engagement.

Customer-centric firms are devoted to understanding their customer needs and making all the necessary amendments in order to cater for them in the way customers have optimum experiences, where this is a long term plan rather than a one shot. (Grigoroudis & Siskos, 2010), (Johnston & Kong, 2011)

When it comes to online gambling, things change significantly. Online gamblers experiences creation is far more complicated, as a different set of emotions drive purchase decisions. In the online gambling landscape, customers may experience different emotions one after the other, upon winning or losing stakes, without these experiences being the protagonist in taking a purchase decision. (Teichert, et al., 2017)

2.2. Customer satisfaction and loyalty

A necessary element in managing customer experience and engagement is the ability to measure and monitor customer reactions, attitudes, and perceptions. Customer satisfaction has been researched since the 70s. It refers to a measure of how products and services offered by a firm meet customers' expectations. Under a marketing metrics perspective, customer satisfaction is defined as "the number of customers, or the percentage of total customers, whose reported experience with a firm, its products, or its services exceeds specified satisfaction goals". (Grigoroudis & Siskos, 2010), (Farris, et al., 2010)

The satisfaction level may derive by comparing a firm's delivered performance with the customers' expectations. According to the literature, customer satisfaction can be measured in various ways. The most common type of question customers are asked is "How satisfied were you from …?". (Bolton, 1999)

Moreover, other simple and more complex effects caused or driven by customer satisfaction, such as the impact on customers' emotions have drawn researchers' attention. Some common emotions researched upon include happiness, trust, indulgence and frustration. These emotions served in that specific order proposed by the emotional value hierarchy. Customers feeling happy or trust would definitely recommend their experience to other people, while customers feeling indulged or frustrated would never recommend their experience to other people. (Laros & Steenkamp, 2005), (Grigoroudis & Siskos, 2010)

Yet, that is not the typical case for the Gambling Industry, since according to other researchers, there are some concerning asymmetries between satisfaction and firm performance. Issues such as customers' habit often overwhelm customers' satisfaction when it comes to predicting firm performance. The real issue behind the vagueness of customer satisfaction and firm performance in the gambling industry, including the online gambling industry, is that customers develop a feeling of hedonism, which may overcome their sensation of satisfaction when it comes to spending more money on gambling activities. In that aspect, research indicates a "hedonic experience factor" which should be taken into consideration when predicting the firm performance via customer satisfaction is raised. (Said, et al., 2003), (Grigoroudis & Siskos, 2004), (Back & Lee, 2015), (Alba & Williams, 2013), (Io, 2016)

The latter will be taken into consideration when interpreting the case study analysis results. A quick heads up for the reader includes a recommendation of expanding currently used customer feedback metrics by inserting some controls regarding customers' pleasure, in order to capture the whole picture of the mechanisms underlying between customer experience and firm performance.

Customer satisfaction summarizes to customer expectations towards a product, a service or a brand. Since customers' expectations sit behind customers' satisfaction, these can be divided in three broad clusters;

- product or service performance, as perceived by customers
- implied costs and needed effort for purchasing a product or experiencing a service
- social approval related expectations

Therefore, customers initiate a three-folded evaluation process where their expectations are weighted against their prior experience and any relevant information available, for instance from word of mouth. The more the information customers know of, the easier and the more accurately they will evaluate their expectations from a product or service. (Cardozo, 1965)

Literature indicates that customers' satisfaction and customers experience are negatively related to the evaluation effort and the level of expectations, while the more positively a product or service meets customers' expectations, the more likely it is to retain these customers. Moreover, customers are highly prejudiced, from any prior relative experience, and their degree of satisfaction is significantly affected by, in case of extremely poor or extremely positive prior experiences. Finally, in terms of reviewing the customer satisfaction evaluation process, literature indicates positive relation with the whole experience a customer is living when buying a product or enjoying a service. (Cardozo, 1965), (Dixon, et al., 2010), (Ezenwafor, et al., 2020), (Grigoroudis & Siskos, 2010), (Gao & Lai, 2015)

The aformentioned theory applies to the gambling industry as well, according to gambling-specialized papers. (Said, 2002), (Jeon & Hyun, 2013)

Overall, customer satisfaction is a standard metric used in the field of marketing, perceived as a leading driving force of customer experience and as mentioned above, proven to be correlated with firm performance. Yet, the link between customer satisfaction and customer loyalty is not bidirectional: a loyal customer is definitely a satisfied customer, but the opposite is not always true. In simple words, satisfaction is a necessary but not a sufficient condition for loyalty. (Grigoroudis & Siskos, 2010)

A key difference between customer satisfaction and loyalty can be spotted in their conceptual framework. On the one hand, customer satisfaction is a metric, while on the other hand loyalty is a broader concept, hard to be limited down to a metric. According to literature, customer loyalty has to do with consumer behavior and consumer attitude, where loyalty mandates satisfaction, and drives customer retention and customer promoting an experience to other people. (Grigoroudis & Siskos, 2004)

Loyal customers tend to prefer a specific brand over other competitive brands usually based on different loyalty attitudes. Loyalty attitudes on their turn depend on different combinations of emotions, opinions and knowledge about a brand. As implied above, loyalty may not always come along with loyal attitudes or the other

way round. In fact, loyalty may be rational, behavioral or emotional. (Bilgihan, et al., 2016), (Balakrishnan & Griffiths, 2018)

Under rational loyalty, customers are attracted to special offers, discounts and generally premium pricing. Therefore, rational loyalty refers to the ration behind being loyal to a product or service brand which comes at a lower cost. In that case, rational loyalty is rather not rock solid, as customers may easily switch to a competitive brand offering more premium pricing and fulfilling the same needs. Literature indicates that this kind of loyalty does apply to the online gambling industry.

Under behavioral loyalty, customers prefer a specific brand repeatedly mainly driven by comfort and convenience reasons, such as proximity, ease of use etc. Likewise, with rational loyalty, behavioral loyalty is not rock solid, since customers may switch to a competitive brand in case they are offered with more premium comfort. According to literature, this kind of loyalty applies to the online gambling industry as well.

Finally, under emotional loyalty, customers' behavior and attitudes are driven by the inner feeling a brand creates. Emotional or attitudinal loyalty comes when customers feel appreciated by a firm and is the most rock solid kind of loyalty, also applying as the most rock solid kind of loyalty in the online gambling industry. This is the main reason, most of the resources spent on customer relations management are directed to creating, safeguarding and enhancing customers emotional loyalty. (Grigoroudis & Siskos, 2004), (Bilgihan, et al., 2016), (Balakrishnan & Griffiths, 2018)

In this context, customer satisfaction is arguably sufficient when it comes to managing customers' relations and building the strategic planning upon them. On the contrary, customer loyalty is a sustainable competitive advantage foundation and has been recognized so in terms of both academia and real businesses in the field of strategic marketing. Conclusively, firm performance may be more well-tied to customer engagement, which includes satisfaction as a loyalty dependent. (Fleming & Asplund, 2007), (Brodie, et al., 2011), (Vivek, et al., 2012)

2.3. Net Promoter Score

Over the past few years, the aspects of customer satisfaction, retention and loyalty have gained increased interest over both the academia and real businesses. Firms have been always investing on customer satisfaction and loyalty, yet only lately have these investments been modeled and justified on facts and figures.

Traditionally, customer satisfaction and especially customer loyalty have been perceived as multidimensional concepts, with the overwhelming complexity of satisfaction and loyalty models not facilitating the selection of a proper model per business case.

The Net Promoter Score model was introduced back in 2003 by Reichheld Fred, marking the beginning of a new era for evaluating customer loyalty. The NPS model was introduced as an innovative metric, calculated by surveying customers regarding their willingness to suggest their experience to someone in their social circle.

Shortly after the NPS model introduction, academics suggested its strong correlation with firms' growth, increasing the model's value rapidly and reshaping how customer loyalty is integrated into firms' strategies. (Reichheld, 2003), (Grigoroudis & Siskos, 2010)

The NPS model is revised in this paragraph in order to facilitate the examination of the NPS model in the research part of this thesis.

According to the literature, the word-of-mouth plays an incredibly significant role in creating momentum for products, services, brands or firms. The NPS model sits on the impact of the word-of-mouth on a firm's market dynamics.

Despite the effect of the word-of-mouth on sales might resemble simple, it proves to be a rather complex concept according to the literature. In fact, the presumed linkage between the word-of-mouth and firms' market dynamics does not apply to every industry and every market.

Some researchers suggested a pivotal contribution of customer loyalty and word-of-mouth in retaining and expanding consumers' pools. Moreover, different promotion

strategies and customer retention strategies have a different impact on sales growth. (Danaher & Rust, 1996), (Rust, et al., 2000), (Grigoroudis & Siskos, 2010)

Definitely, various other factors, such as economic ones, industry expansion, innovation and change integration etc., in addition to customer loyalty, drive firms' growth. The NPS underlying concept is that customers enjoying sufficiently great experiences to lead them to recommend their experiences to another person can reflect a firm's potential sales and revenue growth.

The NPS model measures how willing customers are to recommend an experience to others. Obviously, customers willing to recommend an experience to others are most likely satisfied. As mentioned above, according to Reichheld (2003), although there is a strong evidence about a significant positive correlation between NPS results and firms' revenue, the model cannot work as a panacea. In industries or markets where the NPS model results do not drive growth or cannot be used as a sole growth predictor, NPS results are still linked with revenue and growth, yet with a less strong correlation.

All in all, the NPS results do not prejudice firms' growth but a firm growing always comes along with high NPS scores.

Moreover, these findings directed researchers and professionals into an entirely different approach to customer loyalty surveys, breaking down prior complex surveys into simpler ones revolving around customers' recommendation willingness. (Grigoroudis & Siskos, 2004)

2.4. Consumer experience feedback metrics across industries

According to literature, understanding the mechanisms behind Customer Feedback Metrics (CFMs), consumer behavior is a prerequisite for examining these phenomena and facets impact on firm performance and growth. Among CFMs, various researchers have proven a strong relation between customer satisfaction and performance, while on contrary other CFMs correlation with firm performance is yet to be proven. (Keiningham, et al., 2005) (Gupta & Zeithaml, 2006), (Gupta, et al., 2006), (O'Sullivan & McCallig, 2012), (Hanssens, et al., 2014)

Marketing research has highlighted various metrics for measuring customer feedback, where CFMs refer mostly to indices measuring not only customer satisfaction but also the effect of satisfaction and loyalty to their willingness to share their experiences with other consumers. (Grigoroudis & Siskos, 2010)

CFMs have become a hot topic for customer relations management. They apply mostly to marketing and customer relations management as their effect is rather cumulative and these metrics can be assessed more accurately and more fruitfully in the long run. (Gupta & Zeithaml, 2006)

Upon literature review, two different logics are identified in terms of examining various CFMs.

The first logic, proposed by Bolton et al. (2004), focuses on the time horizon of various metrics. CFMs are divided into metrics evaluating past and present customers' feedback and into metrics evaluating future customers' feedback. In this context, forward looking metrics are more powerful when it comes to predict future firm performance, while past and present metrics may serve as measures for evaluating mechanisms shaping past and present firm performance based on past and present customers' experiences.

As mentioned above, the NPS metric is a forward looking metric, as it examines the probability that current users may propose their experience with a brand to future users. (Reichheld, 2003)

On the other hand, a backward looking metric is the Customer Effort Score, shortly CES, metric, which measures customers' level of effort put in past experiences in order to receive the service they desired. (Dixon, et al., 2010)

In the same paper by Dixon et al. (2010), the CES metric is proposed as a more accurate predictor of future firm growth than the NPS metric or other customer satisfaction metrics, as it can reveal the causality behind prior firm performance and customer experience.

Finally, when it comes to customer satisfaction metrics, given the fact that the level of customer expectation fulfillment is measured, they tend to be more useful for

evaluating the current status of firm performance and its linkage with customer satisfaction and cannot serve well as a future firm performance predictor as not only do they focus on contemporary issues but also other crucial retention driving factors such as loyalty and the holistic customer experience and journey are not taken into consideration. (de Haan, et al., 2015)

The second course of logic for assessing CFMs sits on focusing outliers and the proportion of outliers to the total sample size. Outliers refer to customers providing either extremely positive or extremely negative answers. According to Morgan and Rego (2006), customers providing extremely positive or negative answers may be clustered in two boxes, named two top boxes, where the proportion of these two boxes to the overall answers scale is measured. This proportion is proposed as a rather effective measure of predicting customers future purchase decisions, therefore it can be used as a rather effective firm performance predictor. (Morgan & Rego, 2006)

Another paper “translated” the two top box theory to the original NPS metric by distinguishing between very positive, mediocre and very negative customer satisfaction responses. (Reichheld, 2003)

This transformed NPS metric may prove to be rather effective and in fact more effective than the original NPS metric, which as mentioned above is under severe criticism, as according to more recent literature, customers are driven by extreme positive or negative experiences more intensely than by mediocre experiences with brands. This effect of extreme experiences inserts some severe non-linearity between customer satisfaction and firm performance. (Van Doorn & Verhoef, 2008), (Verhoef, et al., 2009), (Henderson, et al., 2014)

Moreover, according to other researchers, customers in the services industry tend to give extremely high scores when their satisfaction or loyalty are measured, which means that metrics such as the NPS may prove to be rather poor when it comes to predicting firm performance due to such outlying values. (Cronin, et al., 2000)

Yet, the NPS metric is mostly used on its original version, scoring customers satisfaction on a sale from 0 to 10, where each scale has the same weight. In case the

NPS metric is transformed, attributing more weight on extremely low and mostly on extremely high scores, it should expectedly produce more accurate results when predicting the firm performance is raised.

In general, combining two or more CFMs in order to build a custom satisfaction metric can result in more powerful methods for assessing customers satisfaction and may result in more effective firm performance prediction based on customer satisfaction.

2.5. Net promoter score and firm growth

As mentioned above, different metrics apply to different industries and market setups. In fact, customer loyalty expands much further from repeated purchases, since customers preferring a firm repeatedly are not necessarily loyal to that firm. Only the other way round works, since as mentioned above loyal customers do proceed to repeated purchases from a specific firm.

Furthermore, the repetitiveness of purchases is due to other factors, such as the level of a consumer need to cover, the buying power, extenuating circumstances etc. Pure loyalty undoubtedly favors profitability. Hence, loyal customers are not necessarily more profitable for firms, but the increased pool of loyal customers leads to increased profitability, and therefore increased revenue, as the retention costs drop. (de Haan, et al., 2015)

Moreover, pure customer loyalty is a main top line growth driver. Failure to retain customers will eventually lead to decreased revenue and decelerated growth. Taking the other factors affecting the repetitiveness of purchases into consideration, Reichheld suggests that loyal customers will increase their purchases as their buying power increases, regardless of needs to meet.

The tendency of loyal customers to attract new customers comes at no cost for firms and proves to lead to increased profitability in mainstream and mature markets. That stands as in mature and mainstream markets the difficulty in attracting and retaining customers is reflected on increased marketing costs.

All in all, literature implies that the effectiveness of the NPS model as a firm's growth predictor is not for granted, but works for most market setups. (Reichheld, 2003), (Schulman & Sargeant, 2013)

The more complex the buying decision process, the less likely it is for the NPS model to predict firm's growth accurately. Some typical industries where the NPS model works as an accurate firm growth predictor is the FMCG[1] sector, where the purchase decision is much simpler than the one in the gambling industry, even more when it comes to online gambling. (Korneta, 2018), (Hayes, 2017), (Grigoroudis, et al., 2008), (Mecredy, et al., 2018), (van Doorn & Leeflang, 2013)

[1] *Fast-moving consumer goods such as food and hygienic products*

3. Online gambling industry

The online gambling industry is briefly discussed in this chapter in order to present the industry size, trends and dynamics. The analysis provided is focusing from the global, to the European and finally to the Greek online gambling industry.

3.1. Introduction

The gambling industry has always been a revenue generator on a global scale. However, over the past decade, the gambling industry undergone a severe transformation with the emergence of the online gambling industry. The key driver for the online gambling industry has been the various technological advances of the past decade, which facilitated gambling online from any place and at any time.

The online gambling industry comes along with significant advantages over the traditional offline gambling industry.

Initially, consumers are offered with increased ease of sports or casino gambling using either a computer or a smartphone or any other mobile device, along with the spread of online gambling advertisements around the world, which boosts the industry dynamics.

Consumers now have access to desktop and mobile web pages and applications for sports and casino betting. At the same time, a large network of affiliate firms are engaged into promoting the online gambling agencies.

As mentioned below, the European market is dominating the global online gambling market, where a bunch of gambling agencies consolidated under a few online gambling groups are operating. Some key Groups operating globally are Betfair, William Hill and Paddy Power.

The Asia Pacific market and the North America market are anticipated to record the highest growth rate upon adopting new legislation which started allowing for online gambling. It is worth noticing that these areas are densely populated, while the proportion of young citizens, who use to gamble more heavily, is increased if compared to the European market. (NCRG, 2013)

3.2. Sportsbook

Online sports betting involves live and fixed odds betting. Therefore, players can choose over betting on live events or on fixed events before a sports event has started. Players are offered with various sports options as well as with various options within each sport or the ability to combine bets from different sports.

Odds are offered in various types, but three odds types are prevailing: (Investopedia, 2020)

- decimal or European type
- fractional or British type
- moneyline or American type

Online betting firms generate revenue from a predefined rake, which varies from firm to firm. Rakes define a constant percentage of revenue gained as a percentage of the total stakes placed. Under the rake logic, a player placing bets on a binary option cannot reach a breakeven point with any bets combination. The percent difference from the break-even point equals the rake percentage. Rakes may vary upon players' VIP levels, meaning that players with higher VIP level, judged upon their betting activity, may take advantage of lower rakes, therefore may take advantage of higher earnings. More specifically, the firm allocates users on six VIP levels;

- Negative
- Bronze
- Silver
- Gold
- Platinum
- Diamond

Moreover, some online gambling firms, such as Betfair offer exchange betting platforms, where players compete one another, with the exchange platform making profit out of predefined rakes as well. This means that a player eager to place a bet on a binary option will be allowed to place that bet only if another bettor is eager to place an equivalent bet on the same binary option alternative outcome.

As mentioned in the previous chapter, the purchase decision in online gambling is a much more complicated process than in other industries, as online gambling is addictive and players may keep up gambling despite losing money or being dissatisfied. (NCRG, 2013), (D'Astous & Gaspero, 2013)

3.3. Casino

Apart from sports betting, online gambling firms offer online casino platforms, where players have access to more or less the same range of offline casino games. Firms generate profit from a predefined rake, which differs from game to game, while lately some online gambling firms have launched online platforms where players compete one another, e.g., in poker tables.

Moreover, a latest trend followed by most of the online gambling firms is offering live casino deals, especially for the most popular casino games, such as blackjack and roulette. In live casino platforms, players place bets live, while the dealer is a human interacting with players via video and voice.

Casino betting comes second among the online gambling activities both globally and on the European market, as mentioned below. Moreover, players registering on an online gambling firm can have access both to sports betting and to casino betting under the same account and managing the same capital. (NCRG, 2013), (D'Astous & Gaspero, 2013)

3.4. European - Global online gambling industry

Over the past decade online gambling emerged as a technological advance. Lately, online gambling is becoming more and more an online activity, as players are offered with various gambling options via all possible devices (PCs, smartphones etc.). Consumers can gamble online more or less wherever they are across the world.

The global online gambling industry recorded a gross revenue of $58.9 billion in 2019 with an expected compound annual growth rate of roughly 16.5% until 2026. The offline gambling industry recorded a gross revenue of $450 billion in 2019. The industry is driven by continuous technological advances such as artificial intelligence and machine learning, virtual reality and cloud computing, which drive the market

towards offering consumers with increasingly more gambling options at more tempting odds, resulting from more accurate odds calculations, as well as with more friendly and interactive online gambling environments.

All these technological advances have shaped a highly dynamic industry, which is larger in Europe than in the US. However, this gap is expected to be bridged in the mid run, as the European market is growing seemingly slower than the US market. More specifically, the European market is projected to grow at a 14% CAGR until 2026, while the US market is projected to grow at a 20% CAGR until 2026. (Global Market Insights, 2020)

Figure 3-1: US and North America Online Gambling Market facts and figures (Global Market Insights, 2020)

In 2018, the European online gambling industry recorder a gross revenue of €22.2 billion or 23.2% of the total gambling industry. The offline gambling industry recorded an astonishing gross revenue of €73.5 billion. Yet, the online gambling industry is growing 10% faster than the offline gambling industry in Europe. (European Gambling and Betting Association, 2020)

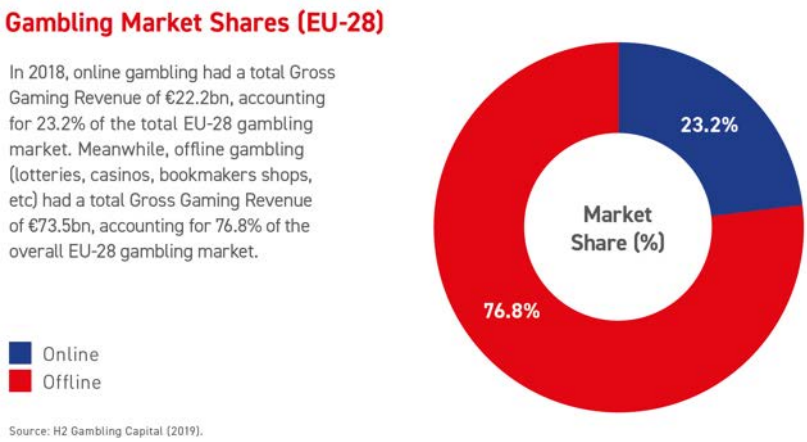


Figure 3-2: European Gambling Market Shares (European Gambling and Betting Association, 2020)

The European online gambling industry is projected to grow to roughly €30 billion by 2022, without incorporating the impact of the Covid19 pandemic, which has driven a large portion of the offline gambling activity towards online gambling. (European Gambling and Betting Association, 2020)

The Covid19 pandemic has accelerated the online gambling along with the accelerated digital transformation of everyday life. Consumers are offered with more secure digital payment options, while the emergence of digital currencies is expected to boost the online gambling industry dynamics.

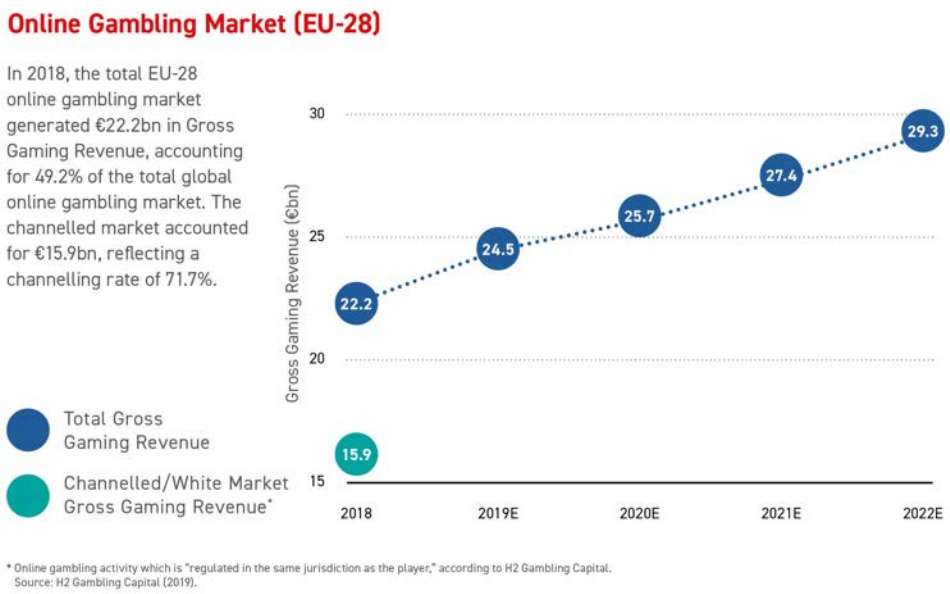


Figure 3-3: Online gambling European Market (European Gambling and Betting Association, 2020) (European Gambling and Betting Association, 2020)

All in all, the European online gambling industry is the global market leader, with a rough 50% of the global industry shares.

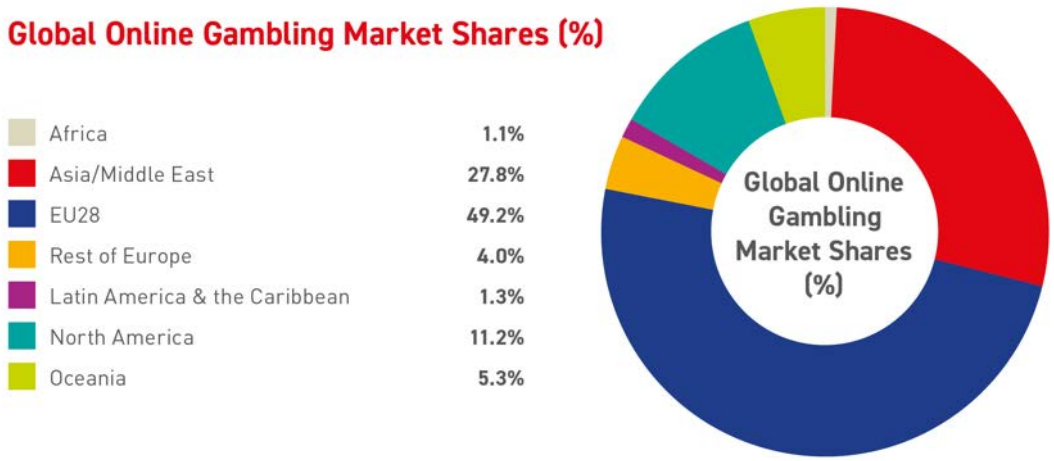


Figure 3-4: Global Online Gambling Market Shares (European Gambling and Betting Association, 2020)

Upon legislation reforms in the US, online sports gambling has been legitimized in some US states, the North American market shares are projected to grow significantly.

When it comes to online gambling activities popularity, sports betting has the lion's share with a 42.5% over casino betting with 32.4% in the European market.

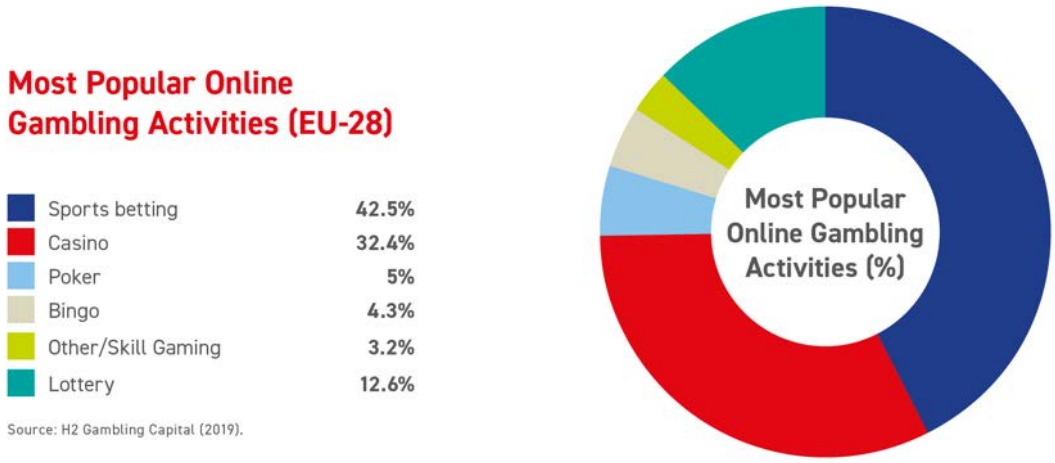


Figure 3-5: Most popular online gambling activities (European Gambling and Betting Association, 2020)

3.5. Greek online gambling industry

The Greek online gambling industry was irregulated a few years ago and numerous firms operated in the country without paying taxes and without operating under an explicit regulatory framework.

In 2019, the Greek government voted for the national online gambling regulations, following European Commissions instruction and closing a decade of lawless operations. The first attempt to adopt a regulatory framework was marked back in 2011, resulting in many small online gambling operating in Greek domains ceasing their operations. On April 1st, the 4002/2011 law was sent to the European Commission for approval.

This led to new licenses being attributed to new firms, marking the beginning of a new era for the Greek gambling and online gambling market. The priorly monopolized by OPAP market, now faces harsh online competition, but as mentioned below, OPAP is the major and controlling shareholder of Stoiximan, a leader in the Greek online gambling market.

The Greek online gambling market is superintended by the Hellenic Gaming Commission (HGC) and the lately adopted licensing regime includes sports betting

permits at a €3 million cost and a casino and poker permits at a €2 million cost. Both permits types have a duration of seven years, at the end of which the permits must be renewed at their initial costs.

The online gaming revenue is bound to a 35% tax, while the maximum casino stake is limited to €2 million. Moreover, online gambling firms may promote their operations on their own.

In 2019, the turnover of the Greek gambling market reached €16 billion, marking a 14.7% annual growth, while the gross online gambling gross revenue reached €2.23 billion, marking a 5.25% annual growth.

As far as the online and offline equilibrium is concerned, the online gambling turnover reached €8.5 billion, or 53.3% of the total Greek online gambling market, marking an average 20% annual growth. The 2019 online gambling net revenue reached €437.3 million, standing for 19.6% of the total (offline and online) market net revenue and marking an 11% year on year growth.

Finally, when it comes to online gambling activities, sports betting accounted for 73% of the total online gambling revenue, while casino betting accounted for 23,9% and poker accounted only for 2,6%. (Calvin Ayre, 2020)

The offline casino revenue accounted for 11% of the total (offline and online) casino activities, as a result of the struggling land based casinos. Moreover, 75% of the online casino revenue stemmed from slot machines. (Calvin Ayre, 2020)

The Greek lottery and betting operator (OPAP) accounted for more than 61% of the Greek gambling 2019 revenue, while third party lottery products accounted for only 7.8% of the Greek gambling market. (Calvin Ayre, 2020)

4. Case study

The firm analyzed is Stoiximan, a Kaizen Gaming (formerly named GML Interactive Ltd) owned company. Kaizen Gaming is the leading GameTech firm in the Online Gaming and Betting industry in Greece and Cyprus, operating the Stoiximan subsidiary. The company offers online sports betting as well an online casino.

Kaizen Gaming was founded back in 2014, and it constantly expands its operations not only in Greece and Cyprus, but also in the rest of Europe and lately in Brazil, under the Betano brand. Stoiximan is a market leader in Greece and Cyprus, while Betano is leading in Romania and among the leaders in the German and Portuguese markets. Currently, the Kaizen Gaming Group has a workforce of more than 800 worldwide.

Kaizen Gaming is present with two brands, Stoiximan and Betano, while it is worth mentioning that OPAP is a Kaizen Gaming major shareholder and a Stoiximan major and controlling shareholder. As of November 2020, OPAP holds 84.99% of the Stoiximan subsidiary and a 36.75% of the Betano subsidiary. (OPAP, 2020)

Aiming towards better understanding of the firm analyzed, a comprehensive group's description via their Linkedin profile is provided below; (Kaizen Gaming (Stoiximan/Betano), 2020)

"Kaizen Gaming is the leading GameTech company in Greece and one of the fastest-growing in Europe. Being International with a local approach, the company is currently operating in 6 countries, with the Stoiximan brand in Greece and Cyprus, and with its international brand, Betano in Germany, Romania, Portugal, and recently in Brazil. Our aim is to leverage cutting-edge Technology in order to provide the optimum experience to those who trust us for their entertainment. People at the core of everything we do, our team of 700+ talented and enthusiastic people fuels our international expansion with their passion, maintaining an "eyes-on-the-customer" approach and a unique OneTeam spirit. Continuous improvement is what we strive for, from professional development to team-bonding activities, while being efficient and making things happen is what our team's minds are set on." (Kaizen Gaming (Stoiximan/Betano), 2020)

“Responsibility has been in our DNA right from the start of our operations, back in 2013. Our wide product offering is adapting to our customers' needs while ensuring that it remains a solely recreational activity. To this end, besides our extensive sports sponsorships program across markets that vary from popular teams to Olympic Champions, we also deploy a social responsibility program that focuses on Technology, Safetainment, and Sports.” (Kaizen Gaming (Stoiximan/Betano), 2020)

4.1. Data collection and measures

Data has been collected by distributing questionnaires to a group of customers over three different periods. Selecting to collect different data in different periods of time was intentional in order to capture the impact of customers’ experience on their loyalty and analyze better the impact of customer loyalty on the firms’ revenue over a bigger time span.

4.2. Descriptive statistics

The survey’s descriptive statistics are reviewed and discussed in this paragraph. This aims to build up a better knowledge about the research sample before proceeding with the correlations and regressions parts.

The paragraph is divided into three distinct sections (demographics, satisfaction and loyalty, financial performance metrics).

4.2.1. Demographics

As per gender, an overwhelming 96% of the users are males, while as per age group 54% of the users are between 35 and 54 years old, followed by a 34% users between 25 and 34 years old. (Figure 4.1 and Figure 4.2)

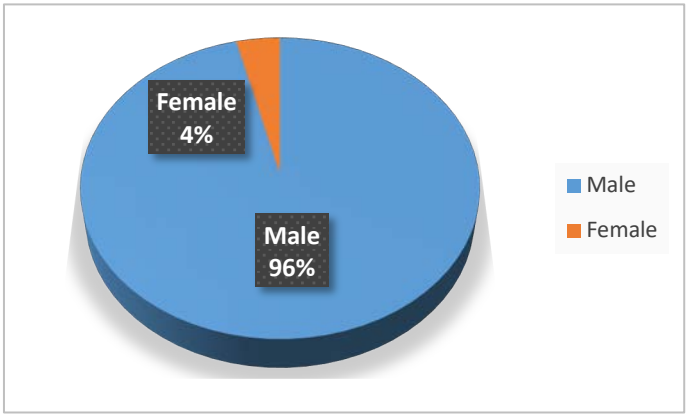


Figure 4.1: Gender

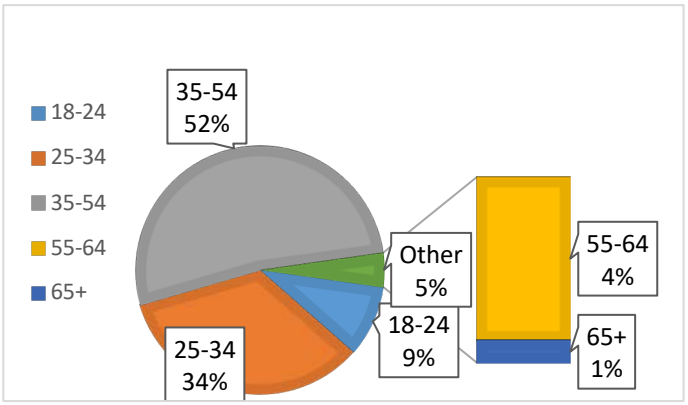


Figure 4.2: Age group

As far as the users VIP level is concerned, the bronze members prevail standing for 60.8% followed by Silver members standing for 10.3%. 9.2% of the respondents stand for those holding a negative VIP level, 9.1% for the Gold level and only 9.5% hold at least a Platinum level. (Figure 4.3)

The firm's users are clustered into seven different VIP level ranks, according to their total stakes placed. Users gain VIP point upon placing bets and mandate reaching and maintaining a specific number of points in order to reach and maintain their VIP level.

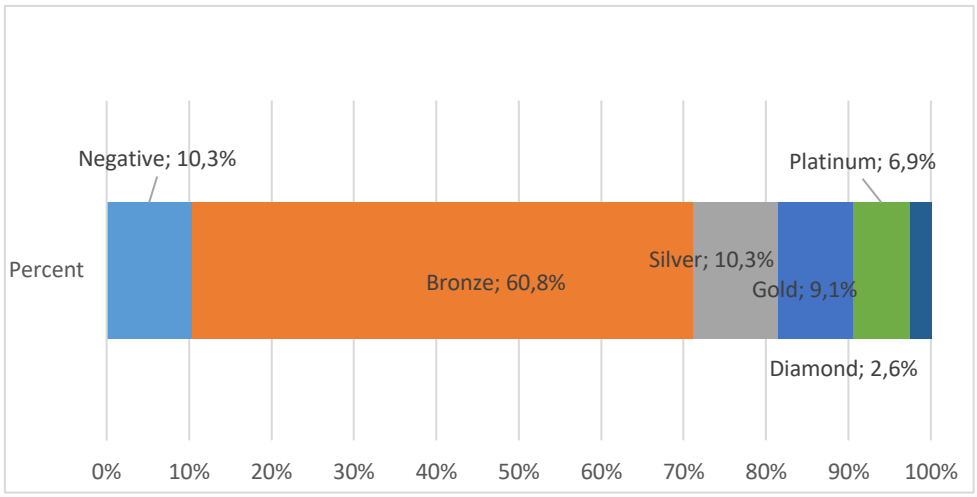


Figure 4.3: VIP level

4.2.2. Customer satisfaction and loyalty

It is worth reminding that respondents answering the NPS questions with 1 to 6 are detractors, 7 and 8 are passives and 9 and 10 are promoters. Upon evaluating the data collected, the NPS equals the difference of the Promoters minus Detractors percentages:

$$NPS = \%\ of\ promoters - \%\ of\ detractors = 70.4\% - 11\% = 59.4\%$$

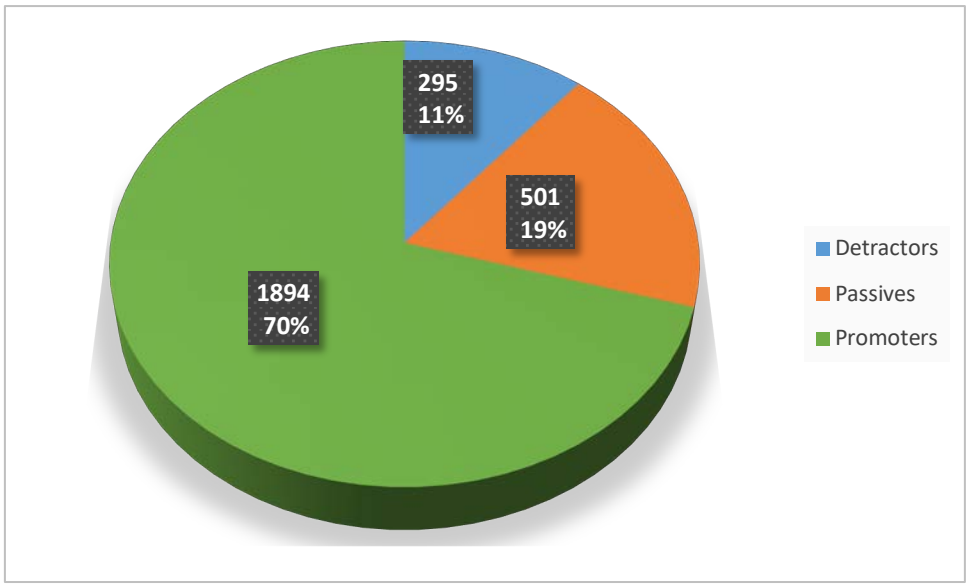


Figure 4.4: REC categories

The NPS of 59.4% is good news, therefore at this point the author expects these good news reflection on firm growth. Moreover, the positive NPS may be considered as a result of the customer satisfaction orientation of Stoiximan. All in all, the satisfied customers prevail over dissatisfied ones or more precisely those willing to recommend Stoiximan to others are more than those who are not willing to do so.

It is worth noticing that passive customers are not part of the above mentioned equation, since the NPS is not affected by passive customers. More specifically, 70.4% of the users are promoters, 19% of the customers are passives and 11% of the customers are detractors.

Finally, digging further in the NPS numeric scores it is worth noticing that 1532 users out of a total of 2690 respondents or 56.95% of them are totally willing to recommend the Stoiximan services to others, answering a 10 on the 1 to 10 NPS Likert scale. Correspondingly, 82.9% of the respondents scored at least 8 on the 1 to 10 NPS Likert scale.

Conclusively, the mean NPS was measured equal to 8.73 with 2,131 standard deviation. (Figure 4.5)

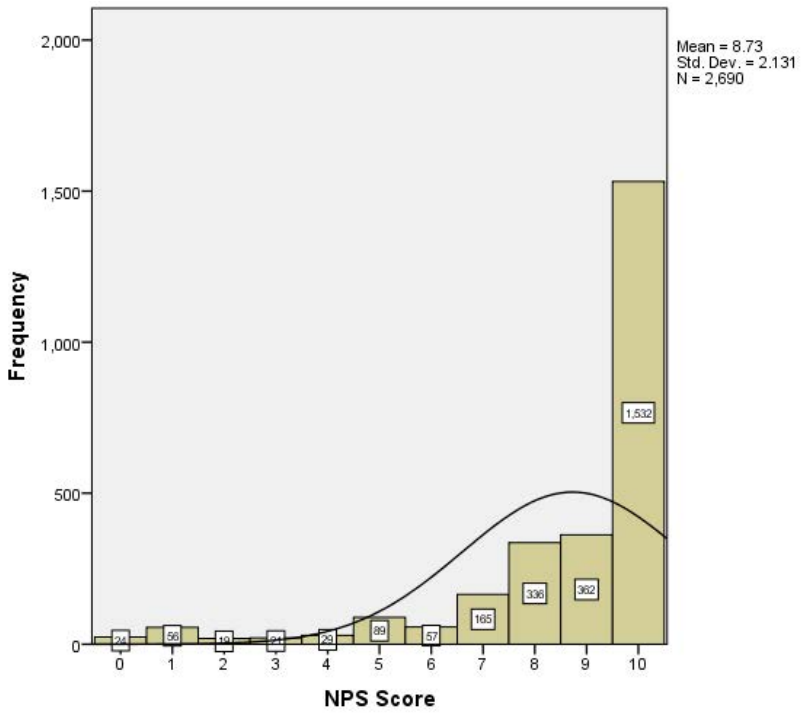


Figure 4.5: NPS histogram

As far as the customer overall satisfaction is concerned, 21% of them are totally satisfied, followed by a 44.7% of very satisfied customers. Cumulatively, 97.1% of the customers are generally satisfied. (Figure 4.6)

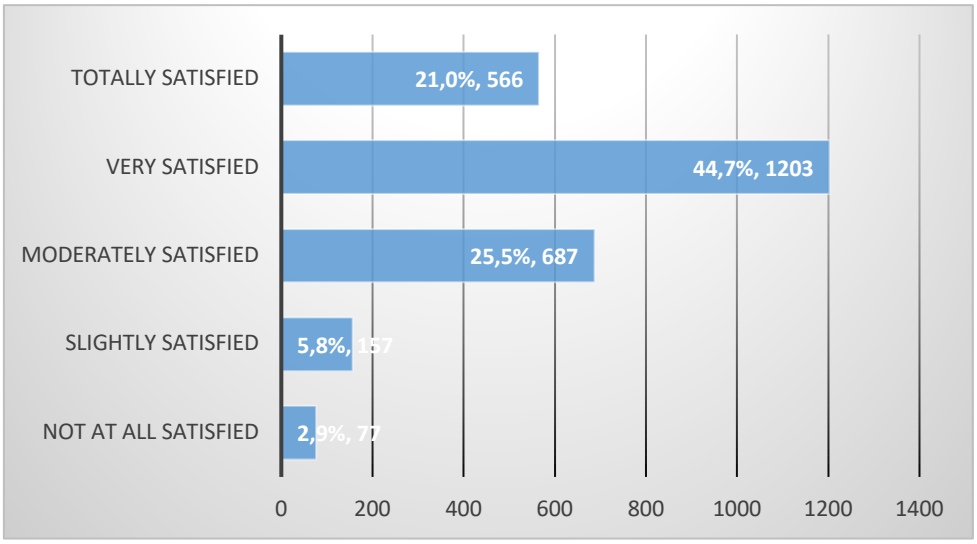


Figure 4.6: Overall customer satisfaction (relative frequency %, frequency)

As far as the customers; expectations confirmation is concerned, users whose expectations are about as expected confirmed prevail, standing for 32.3% of the sample, while 88.9% of the users feel that their expectations are generally confirmed or exceeded. (Figure 4.7)

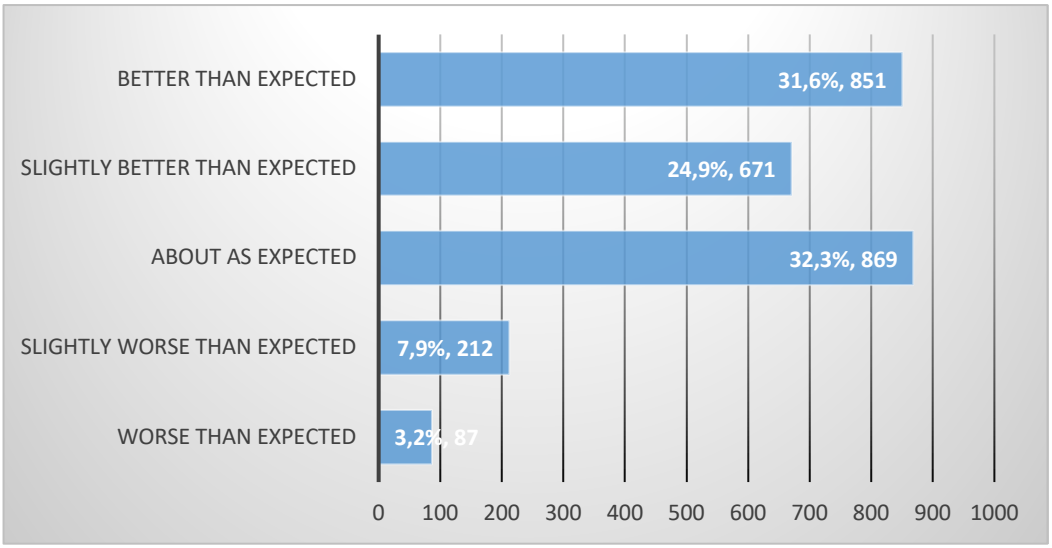


Figure 4.7: Customers expectations confirmation (relative frequency %, frequency)

4.2.3. Financial performance metrics

The average revenue per user mean was measured equal to 302.36€ with a 1,386.53€ standard deviation. The users' revenue in sport bets outweighs the casino bets mean with 239.36€ over 63€. The huge revenue dispersion is worth noticing, since the users total average revenue range from -11,883.98€ to 53,423.95€, adding up to a 65,307.93€ range. Yet, this dispersion is mostly due to outliers, while in all cases the average revenue ranges from -1.000€ to +1.000€. Examining the average revenue per user and per bet type, the sport bet range is much bigger than the casino bet range with a 69,152.93€ over 11,779.91€ respectively.

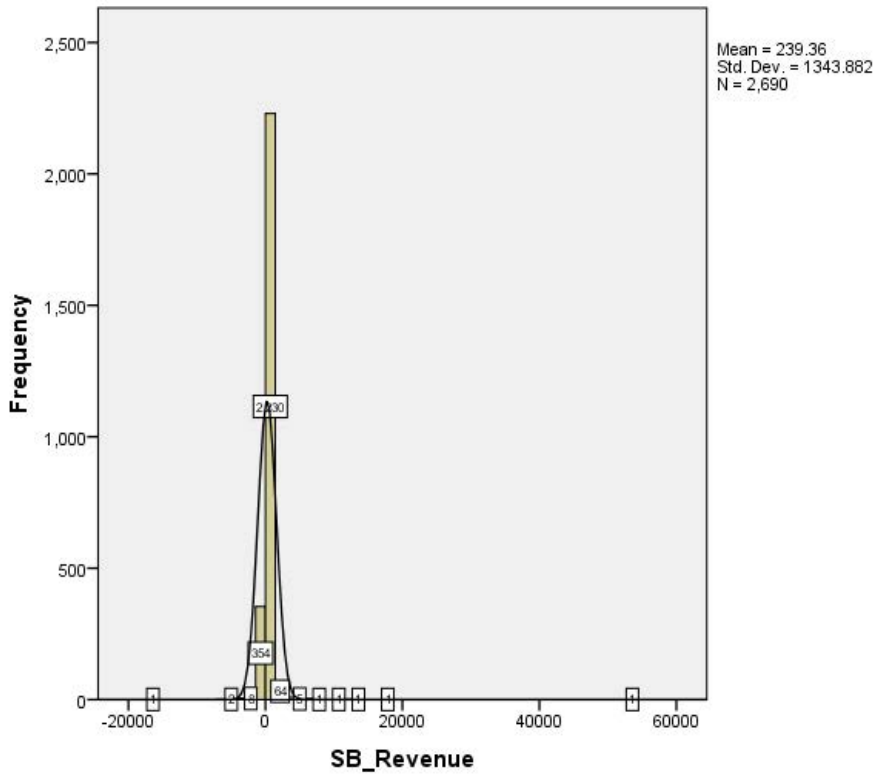


Figure 4.8: Sport bet average revenue per user histogram (ND)

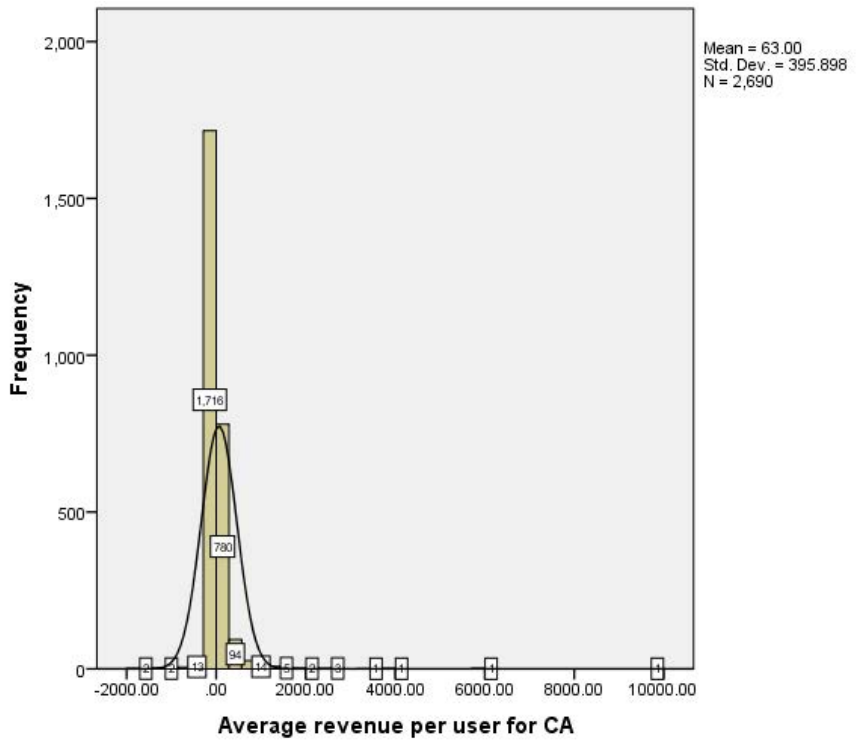


Figure 4.9: Casino bet average revenue per user histogram (ND)

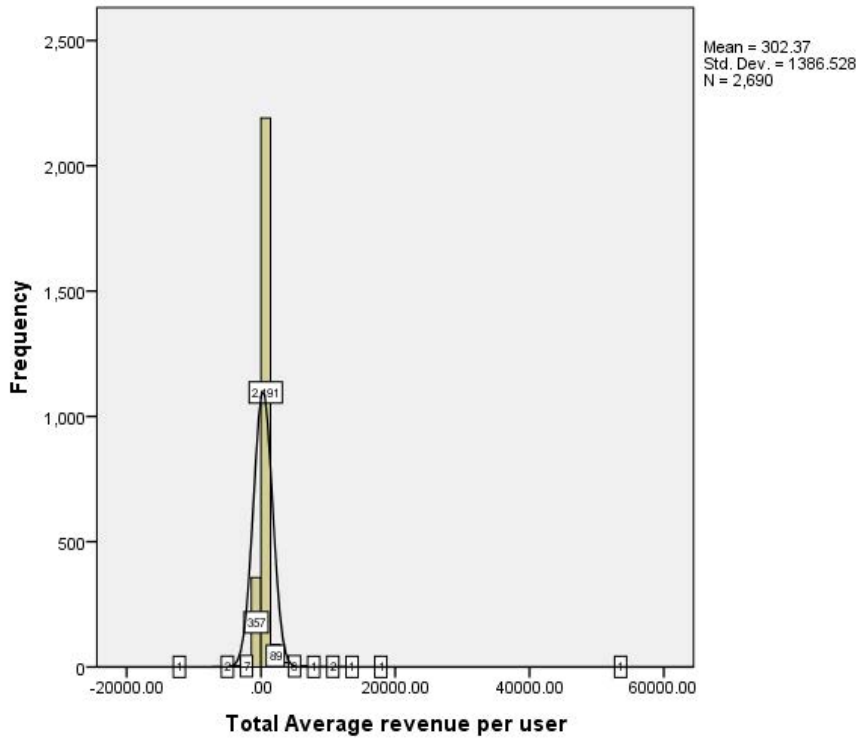


Figure 4.10: Total average revenue per user histogram (ND)

The mean turnover per user on sport bet equals 2,072.84€ with a 9,439.14€ standard deviation, while the mean turnover on casino bet equals 1,489.39€ with an 8,288.61€ standard deviation. The sport bet turnover is larger and more volatile than the casino bet one.

Both sport and casino bet turnovers resemble a normal distribution, but are definitely not normally distributed, since the curve is highly leptokurtic The distributions displacement to the left is due to the fact that turnovers can be only positive.

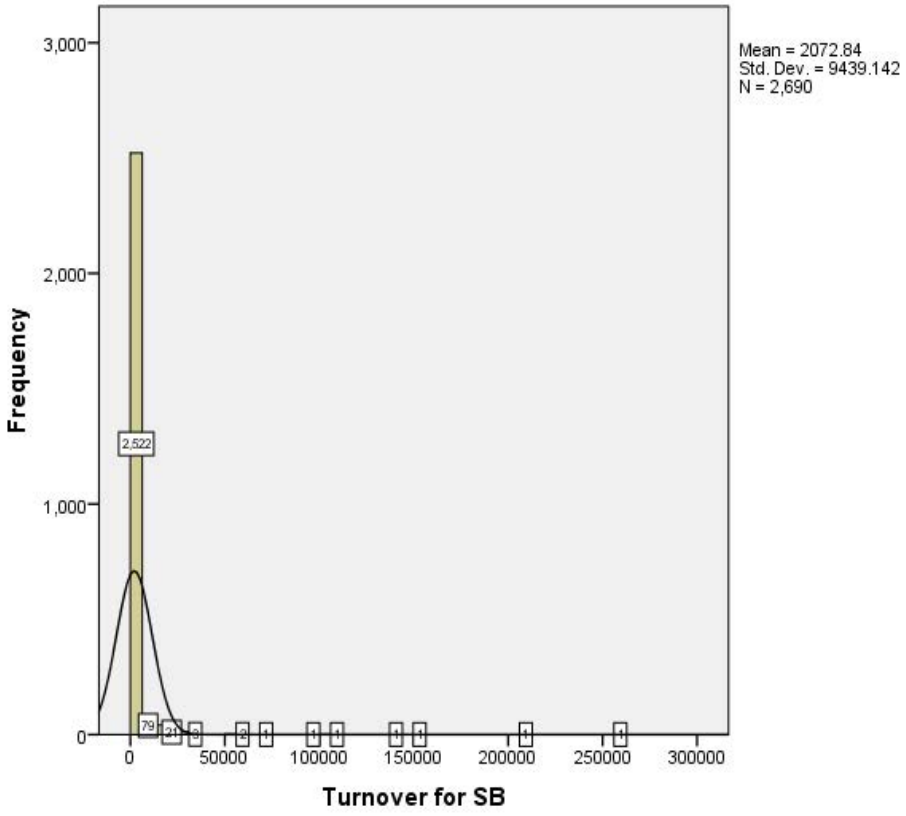


Figure 4.11: Sport bet turnover histogram

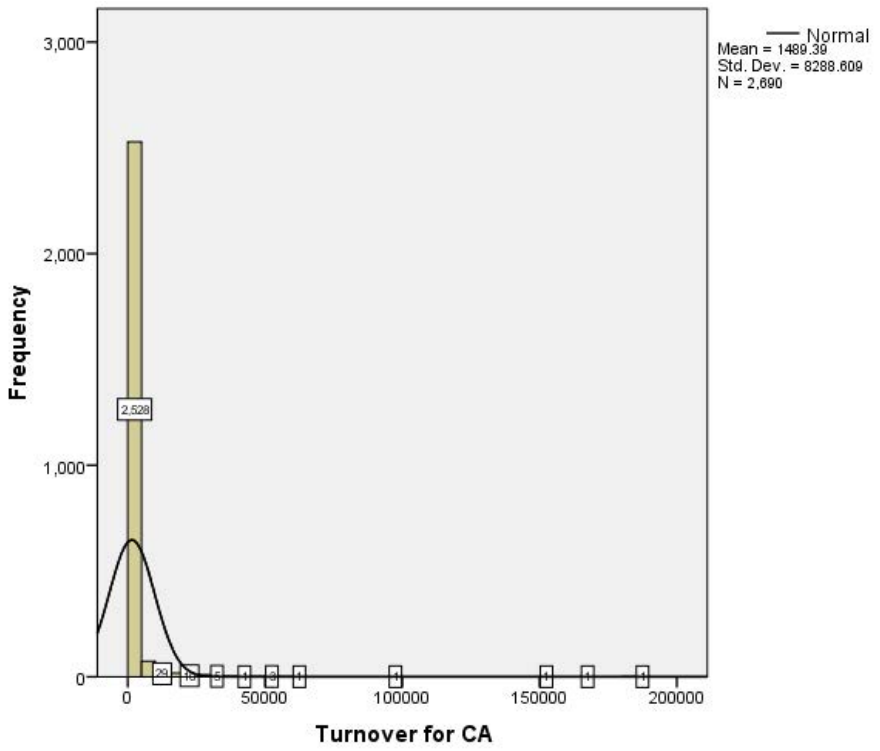


Figure 4.12: Casino bet turnover histogram

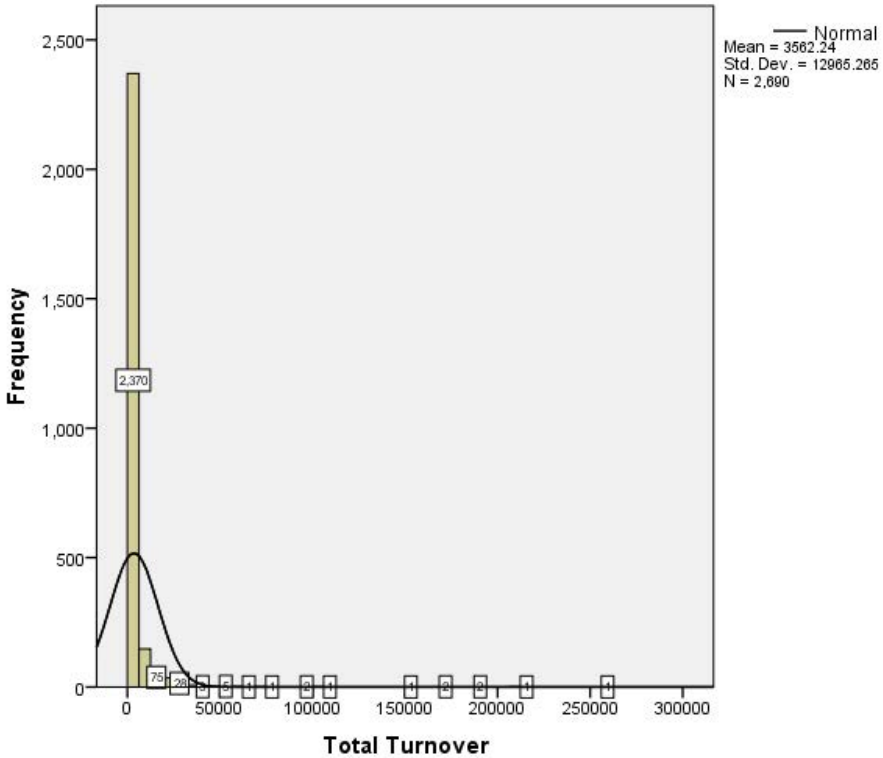


Figure 4.13: Total turnover histogram (ND)

As far as the users' average stake per bet is concerned, users place higher and more volatile bets on sports rather than on the casino, with a 21.99€ and a 2.74€ mean and a 337.56€ and a 31.4€ standard deviation, respectively. Similarly, with the turnover, the average stake per bet distributions are displaced to the left due to the positive stakes restriction. (

Figure 4.14 and Figure 4.15)

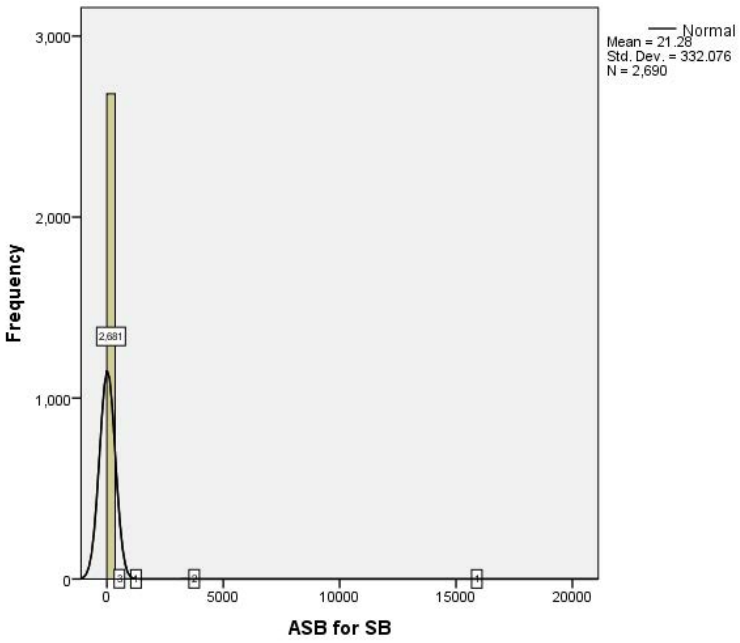


Figure 4.14: Average stake per sport bet histogram

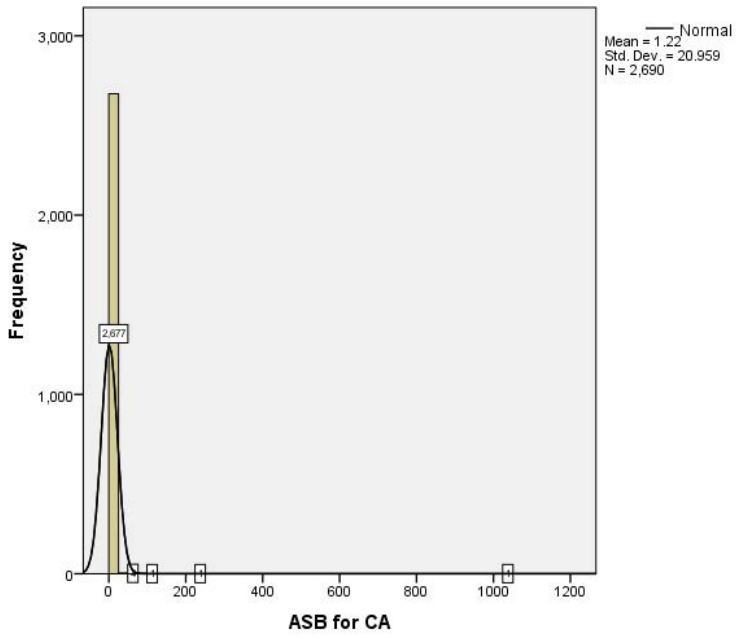


Figure 4.15: Average stake per casino bet histogram

The mean number of bets for casino equals 1,974.47 with a 6,502.76 standard deviation while the mean number of sport bets equals 182.5 with a 302.41 standard deviation. All in all, upon reviewing the last three metrics, users settle more, smaller and making less turnover bet on casino than on sports.

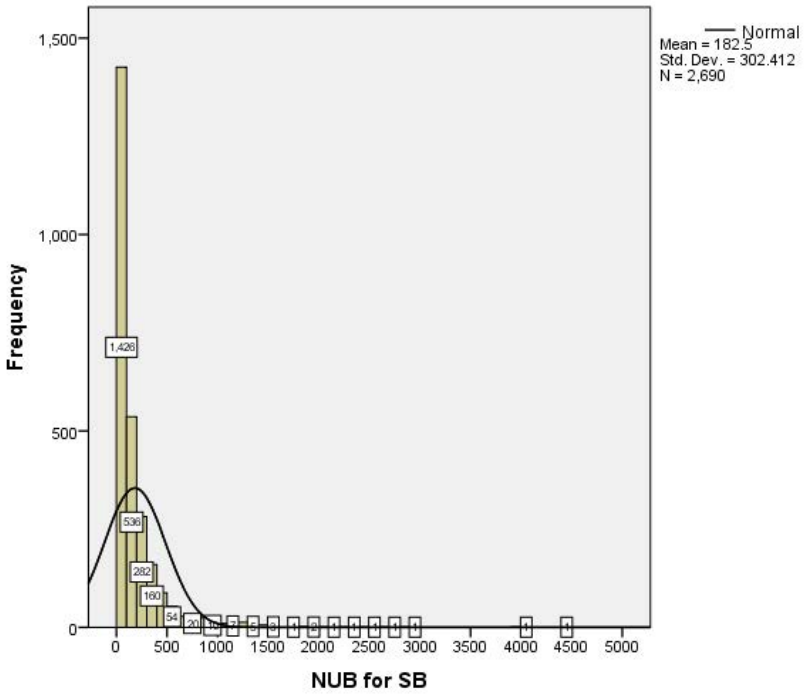


Figure 4.-16: Number of sport bets (ND)

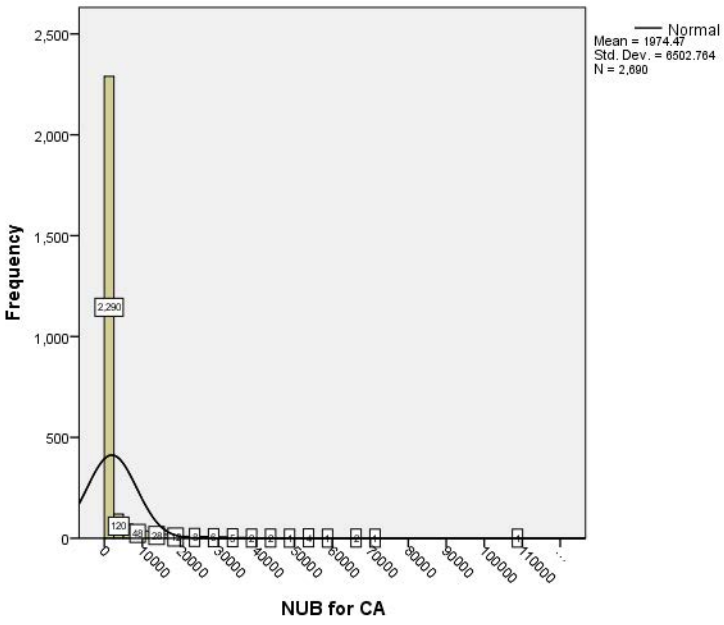


Figure 4.17: Number of casino bets (ND)

Both the number of casino and sport bets distributions have high positive skewness and high kyrtosis. The skewness reflects the fact that the number of bets is a positive number always.

Moreover, users were active for 18 days on sport bets on average, while they were active for 8.2 days on the casino on average during the 29 days data collection period. Both variables standard deviations are close to 8 days.

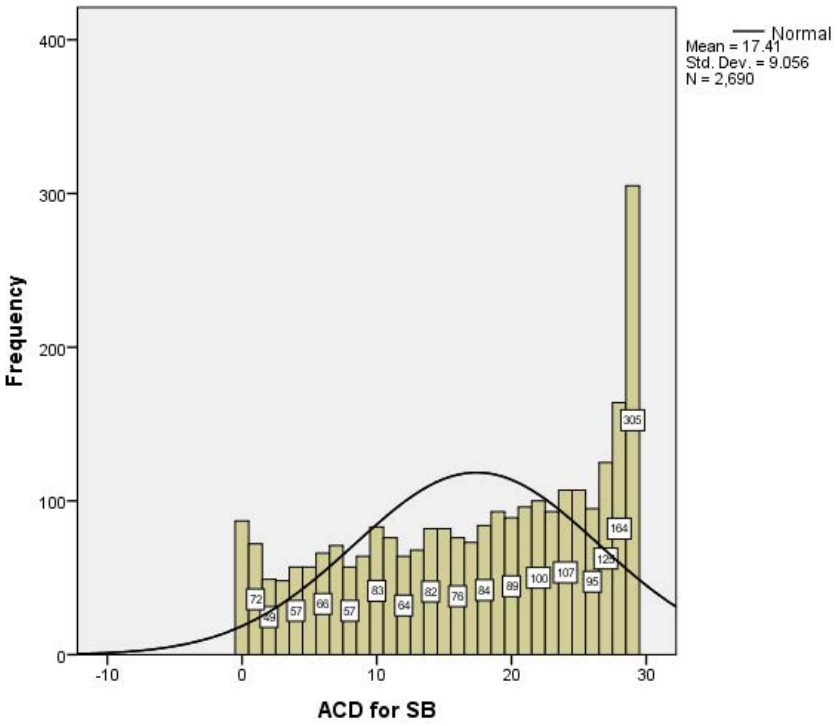


Figure 4.18: Active days on sport bets (ND)

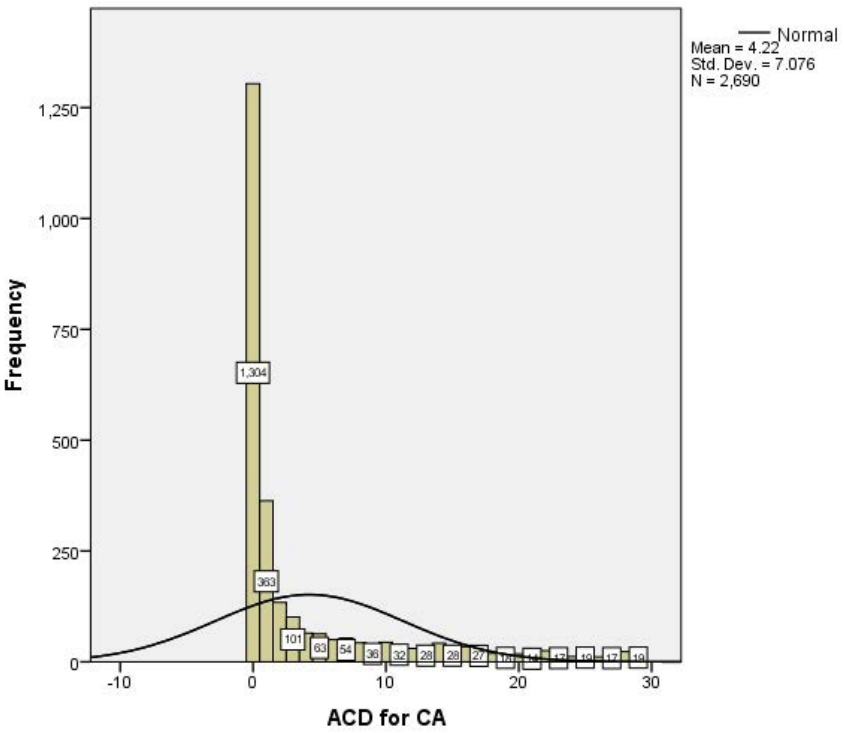


Figure 4.19: Active days on casino (ND)

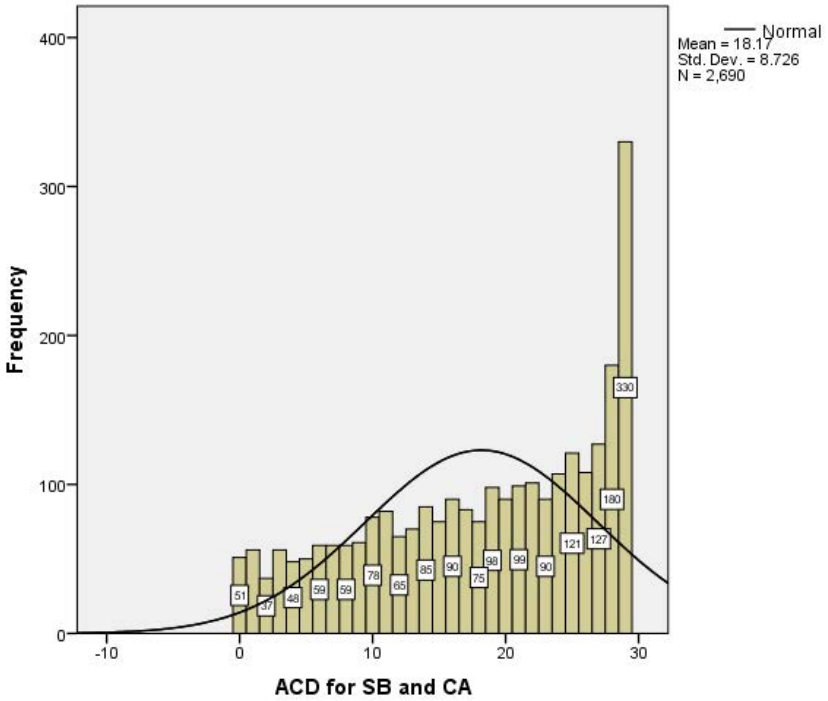


Figure 4.20: Active days on total (ND)

Active days on sport bets, on casino and on total are distributed with a small right tail due to the finite data collection period duration and highly leptokurtic.

The mean deposits per user equal 25.35€ with a 36.84€ standard deviation.

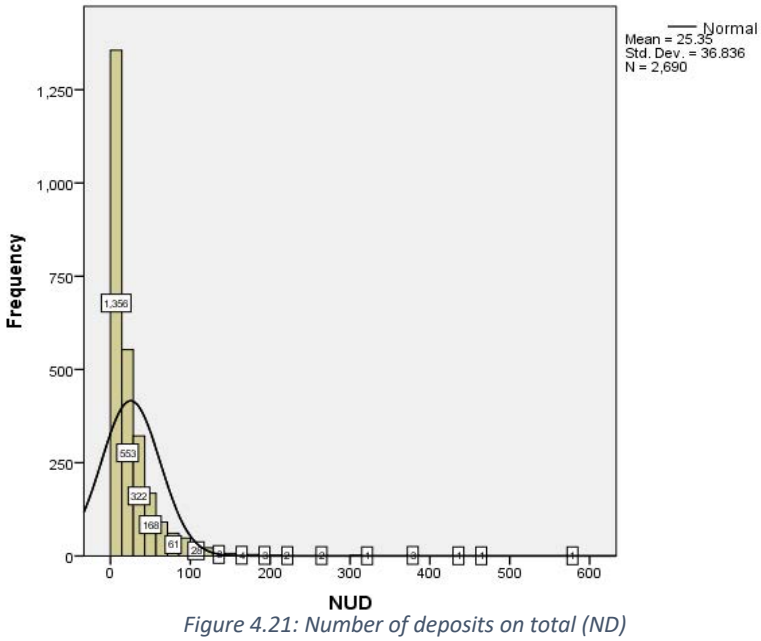


Figure 4.21: Number of deposits on total (ND)

As far as the amount deposited per user, during the period examined, is concerned, the mean deposit was 649.87€ with a standard deviation of 1765.78€.

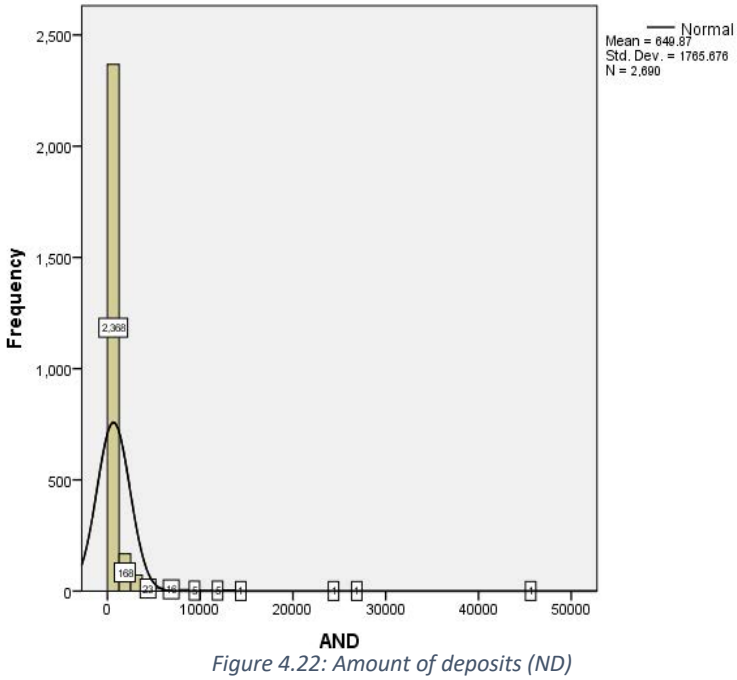


Figure 4.22: Amount of deposits (ND)

The descriptive statistics are summarized in Table 4.1:

METRIC	ABBR.	SPORT BET	CASINO BET	TOTAL
AVERAGE REVENUE PER USER	ARPU	239.36 €	63.00 €	302.36 €
TURNOVER	TUR	2,072.84 €	1,489.39 €	3,562.24 €
AVERAGE STAKE PER BET	ASB	21.99 €	2.74 €	-
NUMBER OF BETS	NUB	182.5	1,947.47	-
ACTIVE DAYS	ACD	18	8.2	18.17
NUMBER OF DEPOSITS	NUD	-	-	25.35
AMOUNT OF DEPOSITS	AND	-	-	649.87 €

Table 4.1: Variables overview

4.3. Correlation analysis

The data to be analyzed consist of 3 categorical ordinal independent variables (REC, SAT, EXP), 7 scale dependent variables (ARPU, TUR, ABS, NUB, ACD, NUD, AND) and 2 categorical ordinal demographic variables (Age group, VIP level) and 1 categorical nominal demographic variable (Gender). In order to identify correlations between the depended variables and independent variables, two methods took place in the analysis procedure.

First, the Spearman correlation was performed under the assumption that all the variables are ordinal variables (even the scale ones). The Spearman's correlation coefficient (rho) and p-value showed statistically significant correlations between the variables under examination.

Then, the Kruskal-Wallis test took place in order to reveal statistically significant differences on a continuous dependent variable by a categorical independent variable. By using these analysis methods, I examined the correlation strength between the variables and also, the significant differences between the created groups. So, we have all the information needed to focus on the parameters that affect each variable and combine them to a model.

Before providing the correlations results, the methods used for examining correlations are discussed further below.

The Spearman's rank-order correlation is the nonparametric version of the Pearson product-moment correlation. Spearman's correlation coefficient, (ρ, also signified by

rho) measures the strength and direction of association between two ranked variables. The Spearman correlation coefficient, rho, can take values from +1 to -1. A rho of +1 indicates a perfect positive association of ranks, a rho of zero indicates no association between ranks and a rho of -1 indicates a perfect negative association of ranks. The closer the rho is to zero, the weaker the association between the ranks.

The general form of a null hypothesis for a Spearman correlation is:

H0: There is no [monotonic] association between the two variables.

The Kruskal-Wallis test is a nonparametric (distribution free) test, and it is used when the assumptions of one-way ANOVA are not met (e.g. Kruskal–Wallis test does not assume a normal distribution of the residuals). Both the Kruskal-Wallis test and one-way ANOVA assess for significant differences on a continuous dependent variable by a categorical independent variable (with two or more groups). In the ANOVA, we assume that the dependent variable is normally distributed and there is an approximately equal variance on the scores across groups. However, when using the Kruskal-Wallis Test, we do not have to make any of these assumptions. Therefore, the Kruskal-Wallis test can be used for both continuous and ordinal-level dependent variables.

Null hypothesis H0: Null hypothesis assumes that the samples (groups) are from identical populations or otherwise that there are no systematic or consistent differences among the treatments being compared. That means that the medians of all groups are equal.

Alternative hypothesis H1: Alternative hypothesis assumes that at least one of the samples (groups) comes from a different population than the others or otherwise that at least one population median of one group is different from the population median of at least one other group.

The distribution of the Kruskal-Wallis test statistic approximates a chi-square distribution, with k-1 degrees of freedom, if the number of observations in each group equals 5 or more.

If the calculated value of the Kruskal-Wallis test is less than the critical chi-square value, then the null hypothesis cannot be rejected. On the other hand, ifthe calculated value of Kruskal-Wallis test is larger than the critical chi-square value, then we can reject the null hypothesis and say that at least one of the samples comes from a different population.

Fundamental assumptions:

- We assume that the samples drawn from the population are random
- We also assume that the observations are independent of each other
- The measurement scale for the dependent variable should be at least ordinal

A significant Kruskal–Wallis test indicates that at least one sample stochastically dominates one other sample. The test does not identify where this stochastic dominance occurs or for how many pairs of groups stochastic dominance obtains. For analyzing the specific sample pairs for stochastic dominance, pairwise Mann-Whitney tests without Bonferroni correction has been used.

The Mann-Whitney U test is used to compare differences between two independent groups when the dependent variable is either ordinal or continuous, but not normally distributed. The Mann-Whitney test can be viewed as an alternative to the independent-measures t test. The test uses the data from two separate samples to test for a significant difference between two treatments or two populations. The null hypothesis for the Mann-Whitney test simply states if there is a systematic or consistent difference between the two treatments (populations) being compared.

The correlations are provided into four parts, namely users profile variables, REC categories, EXP categories and SAT categories.

4.3.1. Summary

The categorical variables correlation tests results are summarized in the table 4.2 below;

Table 4.2: Spearman correlation tests for all categorical and demographic variables

Spearman's rho		Categories of REC	Expectations Confirmation/ Disconfirmation	Overall Satisfaction	Gender	Age Group	Customer's VIP Level
Categories of REC	Correlation Coefficient	1.000	.414**	.400**	.022	.063**	.032
	Sig. (2-tailed)	.	.000	.000	.256	.001	.092
	N	2690	2690	2690	2690	2690	2690
Expectations Confirmation/ Disconfirmation	Correlation Coefficient	.414**	1.000	.588**	-.052**	.128**	.048*
	Sig. (2-tailed)	.000	.	.000	.007	.000	.012
	N	2690	2690	2690	2690	2690	2690
Overall Satisfaction	Correlation Coefficient	.400**	.588**	1.000	-.041*	.152**	.068**
	Sig. (2-tailed)	.000	.000	.	.032	.000	.000
	N	2690	2690	2690	2690	2690	2690
Gender	Correlation Coefficient	.022	-.052**	-.041*	1.000	-.002	.022
	Sig. (2-tailed)	.256	.007	.032	.	.929	.258
	N	2690	2690	2690	2690	2690	2690
Age Group	Correlation Coefficient	.063**	.128**	.152**	-.002	1.000	.074**
	Sig. (2-tailed)	.001	.000	.000	.929	.	.000
	N	2690	2690	2690	2690	2690	2690
Customer's VIP Level	Correlation Coefficient	.032	.048*	.068**	.022	.074**	1.000
	Sig. (2-tailed)	.092	.012	.000	.258	.000	.
	N	2690	2690	2690	2690	2690	2690

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

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

In particular, based on the results presented in the table above, Gender is statistically significantly correlated with EXP and was found independent with the REC categories and SAT variable.

Statistically significant correlations according to Spearman's correlation coefficient rho were found between:

- REC and EXP (p-value=0.000). This is a positive relationship (rho=0.414) which means that as EXP increases REC increases too.
- REC and SAT (p-value=0.000). This is a positive relationship (rho=0.4) which means that as SAT increases REC increases too.
- REC and AGE Group (p-value=0.001). This is a positive relationship (rho=0.063) which means that as AGE Group increases REC increases too.
- EXP and SAT and (p-value=0.000). This is a positive relationship (rho=0.588) which means that as EXP increases SAT increases too.
- EXP and GENDER (p-value=0.007). This relationship is being presented by the figures 4.23 to 4.24. The expectations of men are fulfilled more than those of women.
- EXP and AGE Group (p-value=0.000). This is a positive relationship (rho=0.128) which means that as AGE Group increases EXP increases too.
- EXP and VIP Level (p-value=0.012). This is a positive relationship (rho=0.048) which means that as VIP Level increases EXP increases too.
- SAT and GENDER (p-value=0.032). This relationship is being presented in the Figures 4.25 and 4.26 below. Men are more satisfied.
- SAT and AGE Group (p-value=0.000). This is a positive relationship (rho=0.152) which means that as SAT increases AGE Group increases too.
- SAT and VIP Level (p-value=0.000). This is a positive relationship (rho=0.068) which means that as SAT increases VIP Level increases too.
- Age Group and VIP Level (p-value=0.000). This is a positive relationship (rho=0.074) which means that as Age Group increases VIP Level increases too.

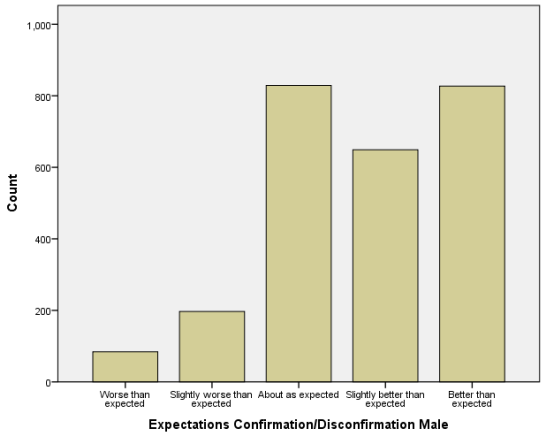


Figure 4.23: EXP in male users.

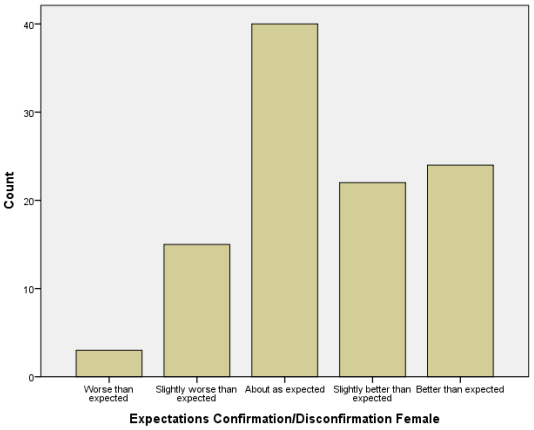


Figure 4.24: EXP in female users

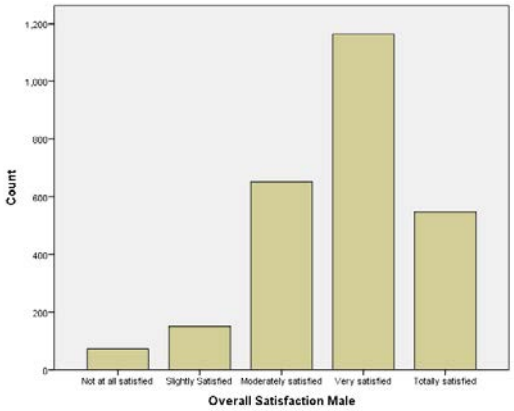


Figure 4.25: SAT in male users

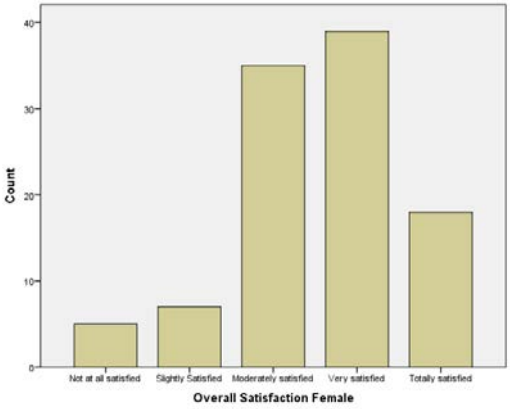


Figure 4.26: SAT in female users

4.3.2. Profile variables

The profile variables include turnover, average stake per bet / round, active days, deposits amount, revenue, number of bets (sports/ casino) and number of deposits. Their assessed intervals and their descriptive statistics are given in Appendix D.

The summarized correlation test results are provided in the table below. The statistically significant correlations are marked green, while cases of scale data treated as ordinal data are marked red. The scale data treated as ordinal data do have an ordinal essence.

Table 4.3: Spearman and Chi-square correlation results (profile variables)

	Gender (nominal)	Age (ordinal)	VIP level (ordinal)
Avg revenue per user SB (scale)		Spearman rho=0.076 (p=0.000)	Spearman rho= 0.505 (p=0.000)
Avg revenue per user CA (scale)		Spearman rho=-0.061 (p=0.002)	Spearman rho= 0.216 (p=0.000)
Avg revenue per user total (scale)		Spearman rho= 0.071 (p=0.000)	Spearman rho= 0.590 (p=0.000)
Avg revenue per user SB (ordinal)	Chi-Square test df=4, p-value=0.542	Spearman rho= 0.073 (p=0.000)	Spearman rho= 0.032 (p=0.096)
Avg revenue per user CA (ordinal)	Chi-Square test df=4, p-value=0.195	Spearman rho= 0.030 (p=0.126)	Spearman rho= .0004 (p=0.830)
Avg revenue per user total (ordinal)	Chi-Square test df=4, p-value=0.867	Spearman rho= 0.068 (p=0.000)	Spearman rho= 0.030(p=0.114)
Turnover SB (scale)		Spearman rho= 0.092 (p=0.000)	Spearman rho= 0.453 (p=0.000)
Turnover CA (scale)		Spearman rho=-0.094 (p=0.000)	Spearman rho= 0.220 (p=0.000)
Turnover total (scale)		Spearman rho=0.080 (p=0.000)	Spearman

	Gender (nominal)	Age (ordinal)	VIP level (ordinal)
			rho= 0.538 (p=0.000)
Turnover SB (ordinal)	Chi-Square test df=4, p-value=0.001	Spearman rho= 0.085 (p=0.000)	Spearman rho= 0.443 (p=0.000)
Turnover CA (ordinal)	Chi-Square test df=4, p-value=0.025	Spearman rho=-0.021 (p=0.274)	Spearman rho= 261 (p=0.000)
Turnover total (ordinal)	Chi-Square test df=4, p-value=0.35	Spearman rho= 0.075 (p=0.000)	Spearman rho= 0.525 (p=0.000)
Average stake per bet SB (scale)		Spearman rho=-0.043 (p=0.026)	Spearman rho= 0.313 (p=0.000)
Average stake per bet CA (scale)		Spearman rho= -0.134 (p=0.000)	Spearman rho= 0.194 (p=0.000)
Average stake per bet SB (ordinal)	Chi-Square test df=4, p-value=0.086	Spearman rho=-0.009 (p=0.630)	Spearman rho= 0.033 (p=0.084)
Average stake per bet CA (ordinal)	Chi-Square test df=4, p-value=0.888	Spearman rho=-0.003 (p=0.862)	Spearman rho= 0.012 (p=0.544)
Number of bets SB (scale)		Spearman rho= 0.167 (p=0.000)	Spearman rho= 0.321 (p=0.000)

	Gender (nominal)	Age (ordinal)	VIP level (ordinal)
Number of bets CA (scale)		Spearman rho=-0.116 (p=0.000)	Spearman rho= 0.185 (p=0.000)
Number of bets SB (ordinal)	Chi-Square test df=4, p-value=0.00	Spearman rho= 0.157 (p=0.000)	Spearman rho= 0.019 (p=0.325)
Number of bets CA (ordinal)	Chi-Square test df=4, p-value=0.021	Spearman rho=-0.085 (p=0.000)	Spearman rho=-0.038 (p=0.050)
Active days SB (scale)		Spearman rho= 0.156 (p=0.000)	Spearman rho= 0.215 (p=0.000)
Active days CA (scale)		Spearman rho=-0.124 (p=0.000)	Spearman rho= 0.177 (p=0.000)
Active days total (scale)		Spearman rho= 0.151 (p=0.000)	Spearman rho= 0.249 (p=0.000)
Active days SB (ordinal)	Chi-Square test df=4, p-value=0.000	Spearman rho= 0.146 (p=0.000)	Spearman rho= 0.027 (p=0.159)
Active days CA (ordinal)	Chi-Square test df=4, p-value=0.003	Spearman rho=-0.037 (p=0.052)	Spearman rho=-0.026 (p=0.177)
Active days total (ordinal)	Chi-Square test df=4, p-value=0.000	Spearman rho= 0.138 (p=0.000)	Spearman rho= 0.024 (p=0.205)

	Gender (nominal)	Age (ordinal)	VIP level (ordinal)
Number of deposits (scale)		Spearman rho=-0.029 (p=0.131)	Spearman rho= 0.472 (p=0.000)
Number of deposits (ordinal)	Chi-Square test df=4, p-value=0.084	Spearman rho=-0.026 (p=0.185)	Spearman rho= 0.021 (p=0.266)
Amount of deposits (scale)		Spearman rho= 0.027 (p=0.158)	Spearman rho= 0.595 (p=0.000)
Amount of deposits (ordinal)	Chi-Square test df=4, p-value=0.794	Spearman rho= 0.029 (p=0.135)	Spearman rho= 0.035 (p=0.069)
REC (scale)		Spearman rho= 0.026 (p=0.173)	Spearman rho= 0.051 (p=0.008)
REC (ordinal)	Chi-Square test df=2, p-value=0.361	Spearman rho= 0.063 (p=0.001)	Spearman rho= 0.032 (p=0.092)
EXP (ordinal)	Chi-Square test df=4, p-value=0.035	Spearman rho= 0.128 (p=0.000)	Spearman rho= 0.048 (p=0.012)
SAT (ordinal)	Chi-Square test df=4, p-value=0.175	Spearman rho= 0.152 (p=0.000)	Spearman rho= 0.068 (p=0.000)

A statistically significant correlation was found between ARPU and all three demographic variables, except for the ARPU_CA and age pair.

A statistically significant correlation was found between TUR and all three demographic variables, except for the TUR_CA and age pair.

A statistically significant correlation was found between ASB and all three demographic variables, except for the ASB_CA and age pair.

A statistically significant correlation was found between NUB and all three demographic variables, except for the NUB_SB and VIP level pair.

A statistically significant correlation was found between ACD and both age and gender, except for the ACD_CA and age pair. The ACD variable is not correlated with the VIP level.

A statistically significant correlation was found between NUD and both age and gender. The NUD variable is not correlated with the VIP level.

A statistically significant correlation was found between AND and both age and gender. The AND variable is not correlated with the VIP level.

4.3.3. Net Promoter Score (REC)

The three REC categories (detractors, passives and promoters) were tested for their potential correlation with the seven dependent variables. The Spearman and Kruskal-Wallis tests results are provided below, while results from Mann-Whitney tests are provided in Appendix E, where the correlations are examined between different REC scales.

The Kolmogorov-Smirnov and Shapiro-Wilks normality test results are mentioned, and the supporting tables are provided in Appendix F, while it is worth mentioning that all variables were found normally distributed and with normally distributed residuals.

4.3.3.1. REC and ARPU

The REC categories are not correlated with the ARPU variable according to the Spearman's tests results.

Table 4.4: REC and ARPU Spearman test

Spearman's rho		Categories of REC	Average revenue per user for SB	Average revenue per user for CA	Total Average revenue per user
Categories of REC	Correlation Coefficient	1.000	.025	-.029	.029
	Sig. (2-tailed)	.	.197	.129	.135
	N	2690	2690	2690	2690

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

Table 4.5: Kruskal-Wallis test ranks ARPU_SB and REC

	Categories of REC	N	Mean Rank
Average revenue per user for SB	Detractors	295	1330.36
	Passives	501	1303.41
	Promoters	1894	1358.99
	Total	2690	

	Average revenue per user for SB
Chi-Square	2.155
df	2
Asymp. Sig.	.340

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.34>0.05, therefore the three REC categories ARPU_SB medians are equal.

Table 4.6: Kruskal-Wallis test ranks ARPU_CA and REC

	Categories of REC	N	Mean Rank
Average revenue per user for CA	Detractors	295	1411.20
	Passives	501	1351.75
	Promoters	1894	1333.61
	Total	2690	

	Average revenue per user for CA
Chi-Square	3.127
df	2
Asymp. Sig.	.209

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.209>0.05, therefore the three REC categories ARPU_CA medians are equal.

Table 4.7: Kruskal-Wallis test ranks ARPU_CA and REC

	Categories of REC	N	Mean Rank
Total Average revenue per user	Detractors	295	1326.87
	Passives	501	1297.62
	Promoters	1894	1361.07
	Total	2690	

Table 4.8: Kruskal-Wallis test statistics ARPU_TOTAL and REC

	Total Average revenue per user
Chi-Square	2.834
df	2
Asymp. Sig.	.242

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.242>0.05 (Table 4.8), therefore the three REC categories ARPU_TOTAL medians are equal.

4.3.3.2. REC and TUR

The REC categories are not correlated with the TUR variable according to both tests results.

Table 4.9: REC and TUR Spearman test

Spearman's rho		Categories of REC	Turnover for SB	Turnover for CA	Total Turnover
Categories of REC	Correlation Coefficient	1.000	.020	-.026	.025
	Sig. (2-tailed)	.	.312	.177	.191
	N	2690	2690	2690	2690

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

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

Table 4.10: Kruskal-Wallis test ranks TUR_SB and REC

	Categories of REC	N	Mean Rank
Turnover for SB	Detractors	295	1285.07
	Passives	501	1350.60
	Promoters	1894	1353.56
	Total	2690	

	Turnover for SB
Chi-Square	2.012
df	2
Asymp. Sig.	.366

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.366>0.05, therefore the three REC categories TUR_SB medians are equal.

Table 4.11: Kruskal-Wallis test ranks TUR_CA and REC

Ranks

	Categories of REC	N	Mean Rank
Turnover for CA	Detractors	295	1475.86
	Passives	501	1294.57
	Promoters	1894	1338.67
	Total	2690	

	Turnover for CA
Chi-Square	12.815
df	2
Asymp. Sig.	.002

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.02<0.05, therefore the three REC categories and TUR_CA medians are not equal. A statistically significant correlation between the three REC categories and the TUR_CA is detected.

Table 4.12: Kruskal-Wallis test ranks TUR_SB and REC

	Categories of REC	N	Mean Rank
Total Turnover	Detractors	295	1336.28
	Passives	501	1298.05
	Promoters	1894	1359.49
	Total	2690	

	Total Turnover
Chi-Square	2.526
df	2
Asymp. Sig.	.283

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.215>0.05, therefore the three REC categories TUR_TOTAL medians are equal.

4.3.3.3. REC and ASB

The REC categories are not correlated with the ASB variable according to both tests results.

Table 4.13: REC and ASB Spearman test

Correlations				
Spearman's rho		Categories of REC	Average stake per bet_SB	Average stake per bet_CA
Categories of REC	Correlation Coefficient	1.000	.009	-.025
	Sig. (2-tailed)	.	.643	.199
	N	2690	2690	2690

Table 4.14: Kruskal-Wallis test ranks ASB_SB and REC

	Categories of REC	N	Mean Rank
Average stake per bet_SB	Detractors	295	1331.04
	Passives	501	1337.47
	Promoters	1894	1349.88
	Total	2690	

Table 4.15: Kruskal-Wallis test statistics ASB_SB and REC

	Average stake per bet_SB
Chi-Square	.216
df	2
Asymp. Sig.	.898

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.898>0.05, therefore the three REC categories ASB_SB medians are equal.

Table 4.16: Kruskal-Wallis test ranks ASB_CA and REC

	Categories of REC	N	Mean Rank
Average stake per bet_CA	Detractors	295	1477.65
	Passives	501	1290.68
	Promoters	1894	1339.42
	Total	2690	

	Average stake per bet_CA
Chi-Square	13.469
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the three REC categories and ASB_CA medians are not equal. A statistically significant correlation between the three REC categories and the ASB_CA is detected.

4.3.3.4. REC and NUB

The REC categories are not correlated with the NUB variable according to the Spearman's test results.

Table 4.17: REC and NUB Spearman test

Spearman's rho		Categories of REC	Number of bets for SB	Number of Rounds for CA
Categories of REC	Correlation Coefficient	1.000	.019	-.036
	Sig. (2-tailed)	.	.333	.062
	N	2690	2690	2690

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

Table 4.18: Kruskal-Wallis test ranks NUB and REC

	Categories of REC	N	Mean Rank
Number of bets for SB	Detractors	295	1268.97
	Passives	501	1365.07
	Promoters	1894	1352.24
	Total	2690	

	Number of bets for SB
Chi-Square	3.325
df	2
Asymp. Sig.	.190

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.19>0.05, therefore the number of bets in the three REC categories medians are equal.

Table 4.19: Kruskal-Wallis test ranks NUB_CA and REC

	Categories of REC	N	Mean Rank
Number of Rounds for CA	Detractors	295	1493.73
	Passives	501	1302.53
	Promoters	1894	1333.78
	Total	2690	

	Number of Rounds for CA
Chi-Square	14.344
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the three REC categories and NUB_CA medians are not equal. A statistically significant correlation between the three REC categories and the NUB_CA is detected.

4.3.3.5. REC and ACD

The REC categories are not correlated with the ACD variable according to the Spearman's tests results, except for the REC categories and ACD_CA pairs. The corresponding significance factor was found 0.039<0.05, therefore the relative H0 is rejected.

Table 4.20: REC and ACD Spearman test

Spearman's rho		Categories of REC	Active Days for SB	Active Days for CA	Active Days for both SB and CA
Categories of REC	Correlation Coefficient	1.000	.036	-.040*	.034
	Sig. (2-tailed)	.	.062	.039	.081
	N	2690	2690	2690	2690

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

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

Table 4.21: Kruskal-Wallis test ranks ACT_SB and REC

	Categories of REC	N	Mean Rank
Active Days for SB	Detractors	295	1185.99
	Passives	501	1392.69
	Promoters	1894	1357.86
	Total	2690	

	Active Days for SB
Chi-Square	14.810
df	2
Asymp. Sig.	.001

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the three REC categories and ACD_SB medians are not equal. A statistically significant correlation between the three REC categories and the ACD_SB is detected.

Table 4.22: Kruskal-Wallis test statistics ACD_CA and REC

	Active Days for CA
Chi-Square	12.915
df	2
Asymp. Sig.	.002

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.02<0.05, therefore the three REC categories and ACD_SB medians are not equal. A statistically significant correlation between the three REC categories and the ACD_CA is detected.

Table 4.23: Kruskal-Wallis test ranks ACD_SB and ACD_CA and REC

	Categories of REC	N	Mean Rank
Active Days for both SB and CA	Detractors	295	1206.42
	Passives	501	1381.68
	Promoters	1894	1357.59
	Total	2690	

	Active Days for both SB and CA
Chi-Square	11.039
df	2
Asymp. Sig.	.004

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.004<0.05, therefore the three REC categories and ACD_SB and ACD_CA medians are not equal. A statistically significant correlation between the three REC categories and the ACD_SB and ACD_CA is detected.

4.3.3.6. REC and deposits variables (NUD and AND)

The REC categories are not correlated with the NUD variable according to both tests results.

Table 4.24: REC and NUD Spearman test

			Categories of REC	Number of deposits
Spearman's rho	Categories of REC	Correlation Coefficient	1.000	.014
		Sig. (2-tailed)	.	.478
		N	2690	2690
	Number of deposits	Correlation Coefficient	.014	1.000
		Sig. (2-tailed)	.478	.
		N	2690	2690

Table 4.25: Kruskal-Wallis test ranks NUD and REC

	Categories of REC	N	Mean Rank
Number of deposits	Detractors	295	1369.12
	Passives	501	1297.31
	Promoters	1894	1354.57
	Total	2690	

Table 4.26: Kruskal-Wallis test statistics NUD and REC

	Number of deposits
Chi-Square	2.462
df	2
Asymp. Sig.	.292

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.292>0.05, therefore the number of deposits in the three REC categories medians are equal.

The REC categories are not correlated with the AND variable according to the Spearman's test results.

Table 4.27: REC and AND Spearman test

			Categories of REC	Amount of deposits
Spearman's rho	Categories of REC	Correlation Coefficient	1.000	.023
		Sig. (2-tailed)	.	.238
		N	2690	2690
	Amount of deposits	Correlation Coefficient	.023	1.000
		Sig. (2-tailed)	.238	.
		N	2690	2690

Table 4.28: Kruskal-Wallis test ranks AND and REC

	Categories of REC	N	Mean Rank
Amount of deposits	Detractors	295	1343.21
	Passives	501	1297.93
	Promoters	1894	1358.44
	Total	2690	

Table 4.29: Kruskal-Wallis test statistics AND and REC

	Amount of deposits
Chi-Square	2.409
df	2
Asymp. Sig.	.300

a. Kruskal Wallis Test

b. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.3>0.05, therefore the users deposited amount medians in the three REC categories are equal.

4.3.3.7. Conclusions

- REC and EXP (p-value = 0.000). This is a positive relationship (rho=0.414) which means that **as EXP increases REC increases too.**
- REC and SAT (p-value = 0.000). This is a positive relationship (rho=0.4) which means that **as SAT increases REC increases too.**
- REC and AGE Group (p-value = 0.001). This is a positive relationship (rho=0.063) which means that **as AGE Group increases REC increases too.**

4.3.4. Expectations

The five EXP categories (Worse than expected, Slightly worse than expected, About as expected, Slightly better than expected, Better than expected) were tested for their potential correlation with the seven dependent variables. The Spearman and Kruskal-Wallis results are provided and discussed below, while results from Mann-Whitney tests are provided in Appendix E, where the correlations are examined between different EXP scales.

4.3.4.1. EXP and ARPU

Table 4.30: EXP and ARPU Spearman test

Spearman's rho		Expectations Confirmation/Disconfirmation	Average revenue per user for SB	Average revenue per user for CA	Total Average revenue per user
Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.028	-.032	.036
	Sig. (2-tailed)	.	.149	.097	.063
	N	2690	2690	2690	2690

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

Table 4.31: Kruskal-Wallis test ranks ARPU and EXP

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Average revenue per user for SB	Worse than expected	87	1306.28
	Slightly worse than expected	212	1314.94
	About as expected	869	1331.83
	Slightly better than expected	671	1337.60
	Better than expected	851	1377.32
	Total	2690	
Average revenue per user for CA	Worse than expected	87	1488.76
	Slightly worse than expected	212	1400.41
	About as expected	869	1369.23
	Slightly better than expected	671	1270.85
	Better than expected	851	1351.80
	Total	2690	
Total Average revenue per user	Worse than expected	87	1306.01
	Slightly worse than expected	212	1313.62
	About as expected	869	1330.76
	Slightly better than expected	671	1317.19
	Better than expected	851	1394.85
	Total	2690	

Table 4.32: Kruskal-Wallis test statistics ARPU and EXP

	Average revenue per user for SB	Average revenue per user for CA	Total Average revenue per user
Chi-Square	2.317	13.400	5.222
df	4	4	4
Asymp. Sig.	.678	.009	.265

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected for the ARPU_CA as p-value=0.009<0.05, therefore the ARPU_CA median is not equal among the five EXP scales.

All in all, the ARPU_CA medians are statistically significantly different for users allocated in all possible different EXP levels.

4.3.4.2. EXP and TUR

The EXP levels are not correlated with the TUR variable according to the Spearman's tests results, apart from the EXP and TUR_SB pair. As shown below, the p value=0.008<0.05, which indicates a positive relationship. Given the rho 0.051 value, this relationship is weak.

Table 4.33: EXP and TUR Spearman test

Correlations

Spearman's rho		Expectations Confirmation/Disconfirmation	Turnover for SB	Turnover for CA	Total Turnover
Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.051**	-.035	.032
	Sig. (2-tailed)	.	.008	.067	.097
	N	2690	2690	2690	2690

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

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

Table 4.34: Kruskal-Wallis test ranks EXP and TUR

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Turnover for SB	Worse than expected	87	1357.93
	Slightly worse than expected	212	1230.00
	About as expected	869	1321.96
	Slightly better than expected	671	1346.20
	Better than expected	851	1396.49

Turnover for CA	Total	2690	
	Worse than expected	87	1425.98
	Slightly worse than expected	212	1466.98
	About as expected	869	1344.90
	Slightly better than expected	671	1309.83
	Better than expected	851	1335.75
	Total	2690	
Total Turnover	Worse than expected	87	1359.43
	Slightly worse than expected	212	1303.18
	About as expected	869	1325.34
	Slightly better than expected	671	1331.60
	Better than expected	851	1386.17
	Total	2690	

Table 4.35: Kruskal-Wallis test statistics EXP and TUR

	Turnover for SB	Turnover for CA	Total Turnover
Chi-Square	9.178	9.264	3.791
df	4	4	4
Asymp. Sig.	.057	.055	.435

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is not rejected for the ARPU_CA as all the p-values are higher than 0.05, which leads to the assumptions that the five TUR groups' medians are equal among the five EXP scales.

4.3.4.3. EXP and ASB

The EXP levels are not correlated with the ASB variable according to the Spearman's tests results.

Table 4.36: EXP and ASB Spearman test

Spearman's rho		Expectations Confirmation/Disconfirmation	Average stake per bet_SB	Average stake per bet_CA
Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	-.004	-.035
	Sig. (2-tailed)	.	.849	.070
	N	2690	2690	2690

The assumption for normal distribution of the residuals has been checked and all the results is that the null hypothesis is rejected by both tests Kolmogorov-Smirnov and Shapiro-Wilk. So, the cases of normality distributed residuals are rejected, and Kruskal- Wallis tests has been used.

Table 4.37: Kruskal-Wallis test ranks EXP and ASB

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Average stake per bet_SB	Worse than expected	87	1471.23
	Slightly worse than expected	212	1338.52
	About as expected	869	1345.48
	Slightly better than expected	671	1315.12
	Better than expected	851	1358.36
	Total	2690	
Average stake per bet_CA	Worse than expected	87	1456.47
	Slightly worse than expected	212	1466.58
	About as expected	869	1339.20
	Slightly better than expected	671	1312.75
	Better than expected	851	1336.24
	Total	2690	

Table 4.38: Kruskal-Wallis test statistics EXP and ASB

	Average stake per bet_SB	Average stake per bet_CA
Chi-Square	3.557	10.023
df	4	4
Asymp. Sig.	.469	.040

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected for the ASB_CA as p-value=0.04<0.05, therefore the ASB_CA medians are not equal among the five EXP scales.

Conclusively, the correlations results indicate that the ASB_CA medians are statistically significantly different among three out of five EXP scales combinations.

4.3.4.4. EXP and NUB

The EXP levels are significantly correlated with the NUB variable according to the Spearman's tests results with p-values of 0.003 and 0.008 respectively for NUB_SB and NUB_CA. More specifically, there is a positive correlation (rho=0.057) between the EXP scales and NUB_SB and a negative correlation (rho=-0.051) between the EXP scales and NUB_SB

The assumption for normal distribution of the residuals has been checked and all the results is that the null hypothesis is rejected by both tests Kolmogorov-Smirnov and Shapiro-Wilk. So, the cases of normality distributed residuals are rejected, and Kruskal- Wallis tests has been used.

Table 4.39: EXP and NUB Spearman test

Spearman's rho		Expectations Confirmation/Disconfirmation	Number of bets for SB	Number of Rounds for CA
Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.057**	-.051**
	Sig. (2-tailed)	.	.003	.008
	N	2690	2690	2690

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

Table 4.40: Kruskal-Wallis test ranks EXP and NUB

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Number of bets for SB	Worse than expected	87	1268.67
	Slightly worse than expected	212	1206.00
	About as expected	869	1325.77
	Slightly better than expected	671	1371.29
	Better than expected	851	1387.92
	Total	2690	
Number of Rounds for CA	Worse than expected	87	1436.23
	Slightly worse than expected	212	1496.35
	About as expected	869	1355.12
	Slightly better than expected	671	1303.10
	Better than expected	851	1322.25
	Total	2690	

Table 4.41: Kruskal-Wallis test statistics EXP and NUB

	Number of bets for SB	Number of Rounds for CA
Chi-Square	11.529	13.633
df	4	4
Asymp. Sig.	.021	.009

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected both for the NUB_SB and for the NUB_CA with p-value=0.021 and 0.009<0.05, therefore the NUB_SB and NUB_CA medians are not equal among the five EXP scales.

Conclusively, the NUB_SB and NUB_CA medians are found to have the same statistically significant differences with 3 out of 5 EXP scale pairs.

4.3.4.5. EXP and ACD

The EXP levels are significantly correlated with the ACD variables according to the Spearman's tests results with p-values of 0.00, 0.01 and 0.012 respectively for ACD_SB, ACD_CA and ACD both. More specifically, there is a positive correlation (rho=0.068 and rho=0.063) between the EXP scales and ACD_SB and ACD both and a negative correlation (rho=-0.048) between the EXP scales and ACD_CA.

Table 4.42: EXP and ACD Spearman test

Correlations

Spearman's rho		Expectations Confirmation/Disconfirmation	Active Days for SB	Active Days for CA	Active Days for both SB and CA
Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.068**	-.048*	.063**
	Sig. (2-tailed)	.	.000	.012	.001
	N	2690	2690	2690	2690

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

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

The assumption for normal distribution of the residuals has been checked and all the results is that the null hypothesis is rejected by both tests Kolmogorov-Smirnov and Shapiro-Wilk. So, the cases of normality distributed residuals are rejected, and Kruskal- Wallis tests has been used.

Table 4.43: Kruskal-Wallis test ranks EXP and ACD

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Active Days for SB	Worse than expected	87	1168.59
	Slightly worse than expected	212	1184.97
	About as expected	869	1326.86
	Slightly better than expected	671	1393.05
	Better than expected	851	1385.12
	Total	2690	
Active Days for CA	Worse than expected	87	1414.98
	Slightly worse than expected	212	1501.67
	About as expected	869	1349.96
	Slightly better than expected	671	1311.10
	Better than expected	851	1322.06
	Total	2690	
Active Days for both SB and CA	Worse than expected	87	1161.74
	Slightly worse than expected	212	1213.82
	About as expected	869	1325.83
	Slightly better than expected	671	1386.64
	Better than expected	851	1384.74
	Total	2690	

Table 4.44: Kruskal-Wallis test statistics EXP and ACD

	Active Days for SB	Active Days for CA	Active Days for both SB and CA
Chi-Square	18.848	12.890	15.624
df	4	4	4
Asymp. Sig.	.001	.012	.004

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected for all three ACD variables with p-values of 0.001, 0.012 and 0.004 for the ACD_SB, ACD_CA and ACD both variables, therefore all the ACD medians are not equal among the five EXP scales.

According to tables this section tables, the following correlations were identified;

- There is a statistically significant difference between the 'ACD_SB' medians for the Expectations Confirmation/Disconfirmation categories '1-4', '1-5', '2-3', '2-4' and '2-5'
- There is a statistically significant difference between the 'ACD_CA' medians for the Expectations Confirmation/Disconfirmation categories '2-3', '2-4' and '2-5'
- There is a statistically significant difference between the 'ACD_Both_SB and CA' medians for the Expectations Confirmation/Disconfirmation categories '1-4', '1-5', '2-4' and '2-5'

4.3.4.6. EXP and NUD

The EXP levels are not correlated with the NUD variable according to the Spearman's tests results.

Table 4.45: EXP and NUD Spearman test

			Expectations Confirmation/Disconfirmation	Number of deposits
Spearman's rho	Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.013
		Sig. (2-tailed)	.	.506
		N	2690	2690

The assumption for normal distribution of the residuals has been checked and all the results is that the null hypothesis is rejected by both tests Kolmogorov-Smirnov and Shapiro-Wilk. So, the cases of normality distributed residuals are rejected, and Kruskal-Wallis tests has been used.

Table 4.46: Kruskal-Wallis test ranks EXP and NUD

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Number of deposits	Worse than expected	87	1302.04
	Slightly worse than expected	212	1360.54
	About as expected	869	1327.79
	Slightly better than expected	671	1356.10
	Better than expected	851	1355.93
	Total	2690	

Table 4.47: Kruskal-Wallis test statistics EXP and NUD

	Number of deposits
Chi-Square	1.083
df	4
Asymp. Sig.	.897

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is not rejected for the NUD as the p-value is higher than 0.05, which leads to the assumption that the number of deposits median is equal among the five EXP scales.

4.3.4.7. EXP and AND

The EXP levels are not correlated with the AND variable according to the Spearman's tests results.

Table 4.48: EXP and AND Spearman test

			Expectations Confirmation/Disconfirmation	Amount of deposits
Spearman's rho	Expectations Confirmation/Disconfirmation	Correlation Coefficient	1.000	.022
		Sig. (2-tailed)	.	.265
		N	2690	2690

The assumption for normal distribution of the residuals has been checked and all the results is that the null hypothesis is rejected by both tests Kolmogorov-Smirnov and Shapiro-Wilk. So, the cases of normality distributed residuals are rejected, and Kruskal-Wallis tests has been used.

Table 4.49: Kruskal-Wallis test ranks EXP and AND

	Expectations Confirmation/Disconfirmation	N	Mean Rank
Amount of deposits	Worse than expected	87	1389.57
	Slightly worse than expected	212	1287.95
	About as expected	869	1336.44
	Slightly better than expected	671	1336.44
	Better than expected	851	1371.74
	Total	2690	

Table 4.50: Kruskal-Wallis test statistics EXP and AND

	Amount of deposits
Chi-Square	2.626
df	4
Asymp. Sig.	.622

a. Kruskal Wallis Test

b. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is not rejected for the AND variable as the p-value is higher than 0.05, which leads to the assumption that the amount of deposits median is equal among the five EXP scales.

4.3.4.8. Conclusions

- EXP and SAT (p-value = 0.000). This is a positive relationship (rho=0.588) which means that **as SAT increases EXP increases too**
- EXP and GENDER (p-value = 0.035) This is a strong relationship (rho=0.588).
- EXP and AGE Group (p-value = 0.000). This is a positive relationship (rho=0.128) which means that **as AGE Group increases EXP increases too**
- EXP and VIP Level (p-value = 0.015). This is a negative relationship (rho=-0.047) which means that **as VIP Level decreases, EXP increases. This reflects users growing expectations as they are engaged more in the services offered.**

4.3.5. Satisfaction

The five SAT categories (Not at all satisfied, Slightly Satisfied, Moderately satisfied, Very satisfied, Totally satisfied) were tested for their potential correlation with the seven dependent variables. The Spearman and Kruskal Wallis tests results are provided and discussed below, while results from Mann-Whitney tests are provided in Appendix E, where the correlations are examined between different SAT scales.

4.3.5.1. SAT and ARPU

The five SAT categories are not correlated with the ARPU_CA variable according to the Spearman's tests results. The five SAT categories are correlated with the ARPU_SB and the ARPU_CA variables according to the Spearman's tests results. The corresponding significance factors were found 0.001<0.05 and 0.002<0.005 therefore the respective null hypotheses are rejected.

Table 4.51: SAT and ARPU Spearman test

Spearman's rho		Overall Satisfaction	Average revenue per user for SB	Average revenue per user for CA	Total Average revenue per user
Overall Satisfaction	Correlation Coefficient	1.000	.063**	-.029	.061**
	Sig. (2-tailed)	.	.001	.129	.002
	N	2690	2690	2690	2690

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

Table 4.52: Kruskal-Wallis test ranks ARPU and SAT

	Overall Satisfaction	N	Mean Rank
Average revenue per user for SB	Not at all satisfied	77	1285.49
	Slightly Satisfied	157	1241.48
	Moderately satisfied	687	1296.63
	Very satisfied	1203	1358.09
	Totally satisfied	566	1415.08
	Total	2690	
Average revenue per user for CA	Not at all satisfied	77	1447.14
	Slightly Satisfied	157	1402.64
	Moderately satisfied	687	1384.16
	Very satisfied	1203	1301.56
	Totally satisfied	566	1362.28
	Total	2690	
Total Average revenue per user	Not at all satisfied	77	1308.94
	Slightly Satisfied	157	1270.31
	Moderately satisfied	687	1303.07
	Very satisfied	1203	1339.59
	Totally satisfied	566	1435.39
	Total	2690	

Table 4.53: Kruskal-Wallis test statistics ARPU_SB and REC

	Average revenue per user for SB	Average revenue per user for CA	Total Average revenue per user
Chi-Square	10.855	9.651	11.345
df	4	4	4
Asymp. Sig.	.028	.047	.023

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the all three ARPU variables as all p-values < 0.05, therefore all three ARPU medians are not equal among the five SAT scales. Therefore, there is statistically significant difference between SAT and ARPU (Figures 4.27, 4.28 and 4.29).

Figure 4.23: Matrix plot SAT and ARPU_SB

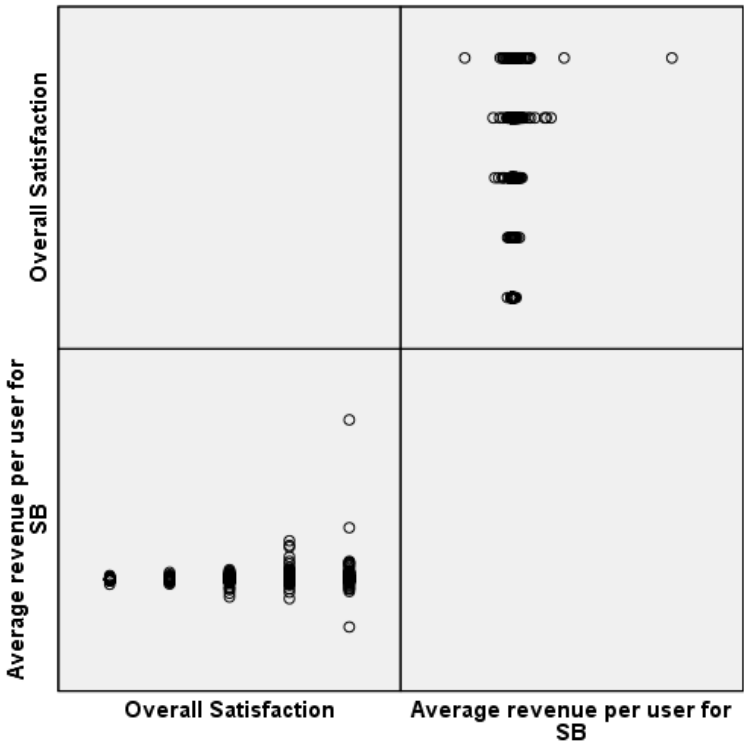


Figure 4.24: Matrix plot SAT and ARPU_CA

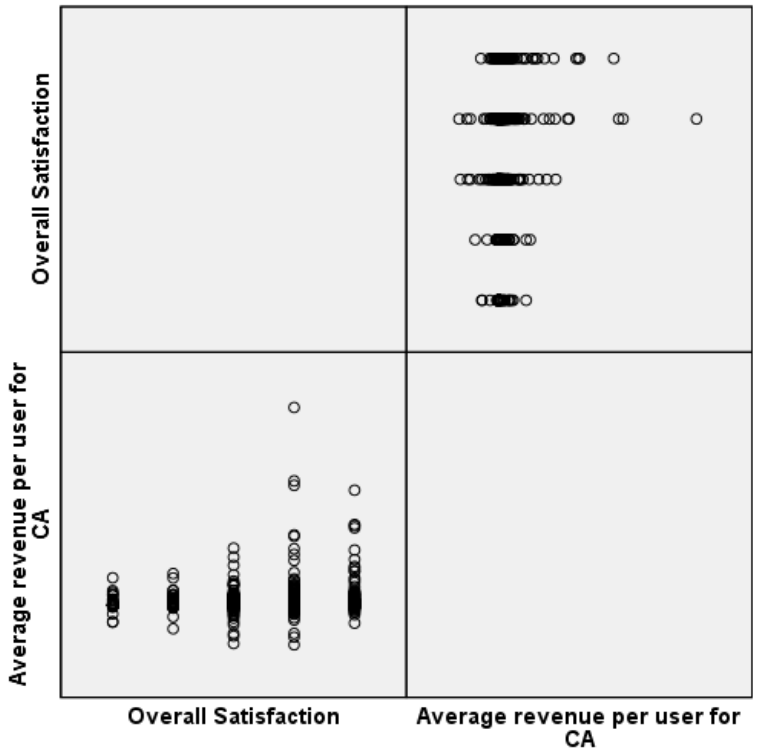
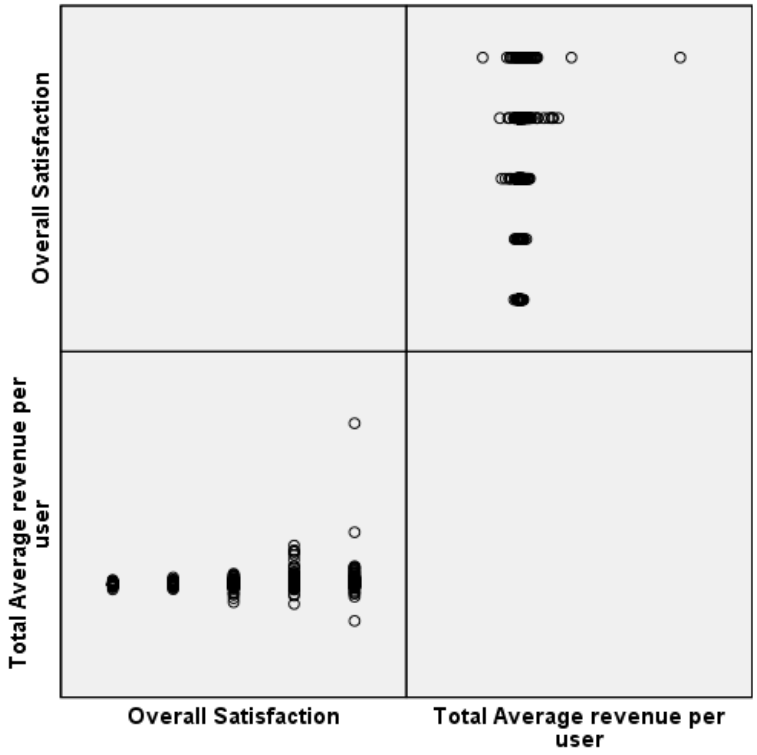


Figure 4.25: Matrix plot SAT and ARPU_TOT



4.3.5.2. SAT and TUR

The five SAT categories are correlated with all three TUR variables according to the Spearman's tests results. The corresponding significance factors were found 0.000<0.05, 0,001<0,005 and 0,009<0,05 therefore the respective null hypotheses are rejected. The correlation between TUR_CA and the SAT categories is negative.

Table 4.54: SAT and TUR Spearman test

Correlations

Spearman's rho		Overall Satisfaction	Turnover for SB	Turnover for CA	Total Turnover
Overall Satisfaction	Correlation Coefficient	1.000	.095**	-.063**	.050**
	Sig. (2-tailed)	.	.000	.001	.009
	N	2690	2690	2690	2690

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

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

Table 4.55: Kruskal-Wallis test ranks TUR and SAT

	Overall Satisfaction	N	Mean Rank
Turnover for SB	Not at all satisfied	77	1268.45
	Slightly Satisfied	157	1254.03
	Moderately satisfied	687	1243.50
	Very satisfied	1203	1373.18
	Totally satisfied	566	1446.33
	Total	2690	
Turnover for CA	Not at all satisfied	77	1457.85
	Slightly Satisfied	157	1548.04
	Moderately satisfied	687	1403.33
	Very satisfied	1203	1272.71
	Totally satisfied	566	1358.54
	Total	2690	

Total Turnover	Not at all satisfied	77	1334.19
	Slightly Satisfied	157	1373.11
	Moderately satisfied	687	1284.46
	Very satisfied	1203	1338.10
	Totally satisfied	566	1429.19
	Total	2690	

Table 4.56: Kruskal-Wallis test statistics TUR and SAT

	Turnover for SB	Turnover for CA	Total Turnover
Chi-Square	25.853	32.393	11.140
df	4	4	4
Asymp. Sig.	.000	.000	.025

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the all three TUR variables with the five SAT categories as all p-values < 0.05, therefore all three TUR medians are not equal among the five SAT categories. Therefore, there is statistically significant difference between SAT and TUR.

Figure 4.26: Matrix plot SAT and TUR_SB

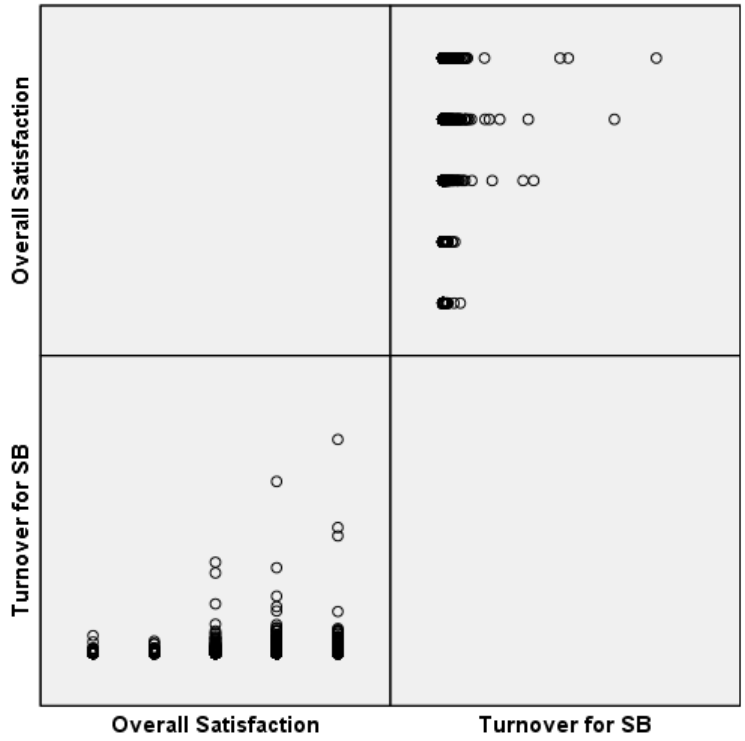


Figure 4.27: Matrix plot SAT and ARPU_CA

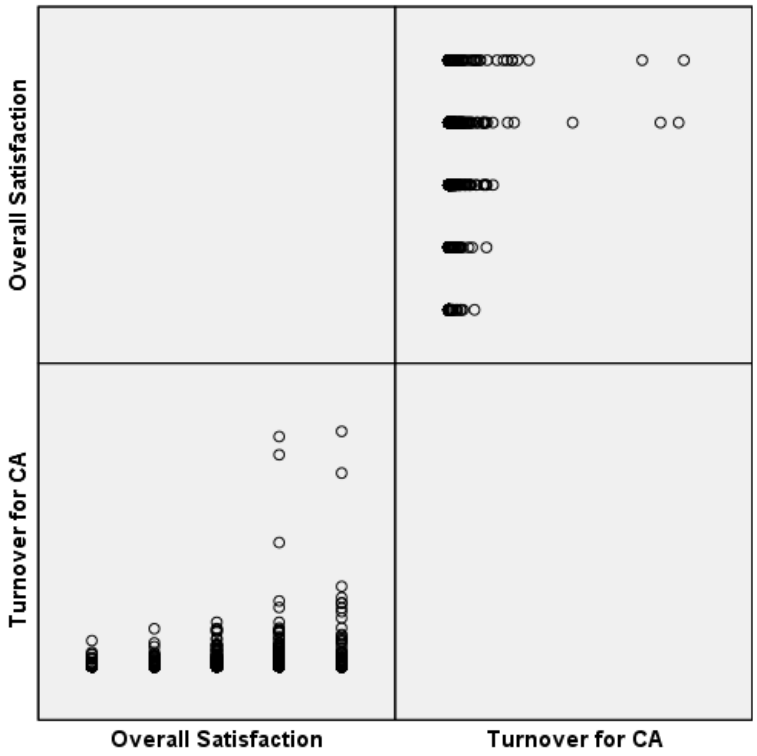
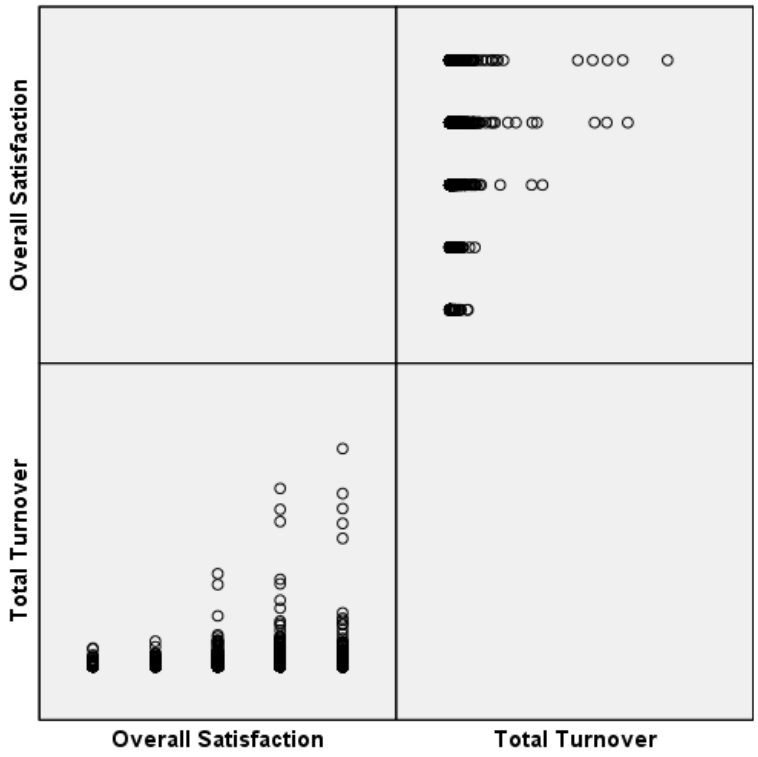


Figure 4.28: Matrix plot SAT and ARPU_TOT



4.3.5.3. SAT and ASB

The five SAT categories are correlated with both ASB variables according to the Spearman's tests results. The corresponding significance factors were found 0.000<0.05, 0.004<0,005 and 0.013<0.05 therefore the respective null hypotheses are rejected. The correlation between ASB_CA and the SAT categories is negative.

Table 4.57: SAT and TUR Spearman test

Spearman's rho		Overall Satisfaction	Average stake per bet_SB	Average stake per bet_CA
Overall Satisfaction	Correlation Coefficient	1.000	.056**	-.048*
	Sig. (2-tailed)	.	.004	.013
	N	2690	2690	2690

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

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

Table 4.58: Kruskal-Wallis test ranks ASB and SAT

	Overall Satisfaction	N	Mean Rank
Average stake per bet_SB	Not at all satisfied	77	1382.32
	Slightly Satisfied	157	1315.44
	Moderately satisfied	687	1272.57
	Very satisfied	1203	1355.15
	Totally satisfied	566	1416.84
	Total	2690	
Average stake per bet_CA	Not at all satisfied	77	1450.81
	Slightly Satisfied	157	1526.40
	Moderately satisfied	687	1387.98
	Very satisfied	1203	1279.75
	Totally satisfied	566	1369.18
	Total	2690	

Table 4.59: Kruskal-Wallis test statistics TUR and SAT

	Average stake per bet_SB	Average stake per bet_CA
Chi-Square	11.427	25.528
df	4	4
Asymp. Sig.	.022	.000

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the both ASB_SB and ASB_CA variables with the five SAT categories as both p-values < 0.05, therefore both ASB medians are not equal among the five SAT categories. Therefore, there is statistically significant difference between SAT and ASB.

Figure 4.29: Matrix plot SAT and ASB_SB

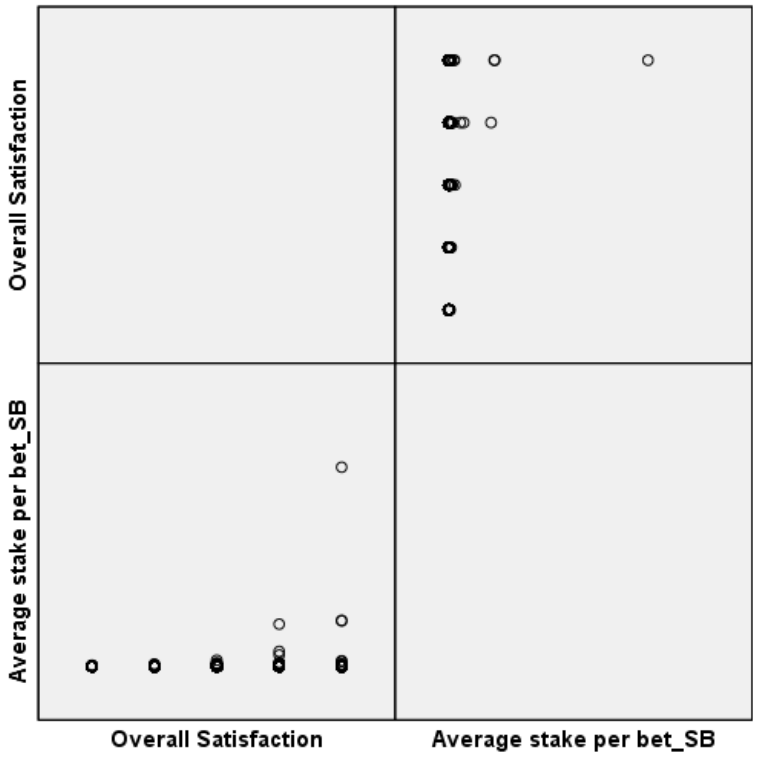
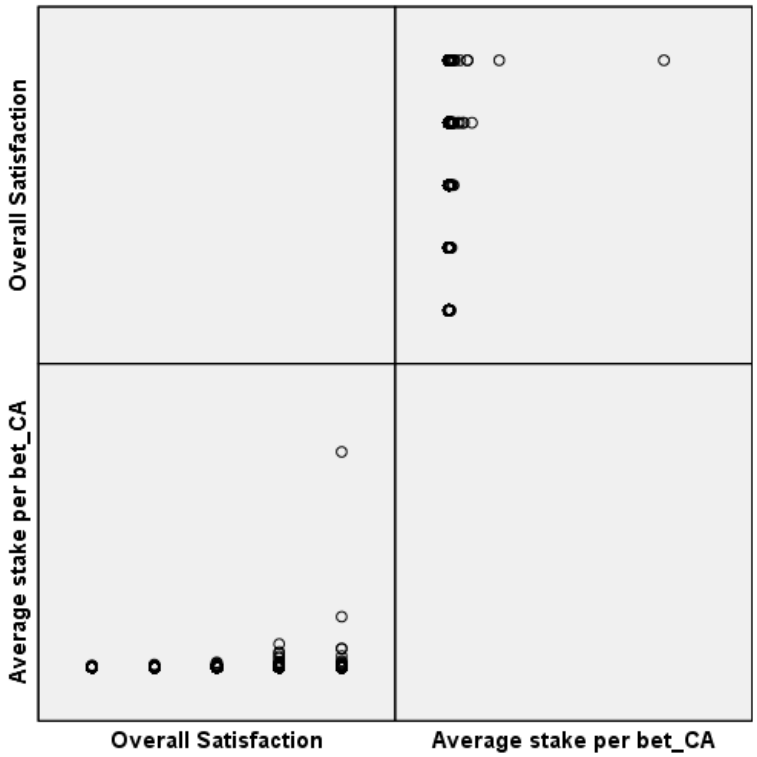


Figure 4.30: Matrix plot SAT and ASB_CA



4.3.5.4. SAT and NUB

The five SAT categories are correlated with both NUB variables according to the Spearman's tests results. The corresponding significance factors were both found 0.000<0.05, therefore the respective null hypotheses are rejected. The correlation between NUB_CA and the SAT categories is negative.

Table 4.60: SAT and NUB Spearman test

Spearman's rho		Overall Satisfaction	Number of bets for SB	Number of Rounds for CA
Overall Satisfaction	Correlation Coefficient	1.000	.068**	-.093**
	Sig. (2-tailed)	.	.000	.000
	N	2690	2690	2690

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

Table 4.61: Kruskal-Wallis test ranks NUB and SAT

	Overall Satisfaction	N	Mean Rank
Number of bets for SB	Not at all satisfied	77	1194.82
	Slightly Satisfied	157	1274.56
	Moderately satisfied	687	1272.28
	Very satisfied	1203	1383.41
	Totally satisfied	566	1393.98
	Total	2690	
Number of Rounds for CA	Not at all satisfied	77	1493.90
	Slightly Satisfied	157	1568.75
	Moderately satisfied	687	1435.86
	Very satisfied	1203	1264.72
	Totally satisfied	566	1325.40
	Total	2690	

Table 4.62: Kruskal-Wallis test statistics TUR and SAT

	Number of bets for SB	Number of Rounds for CA
Chi-Square	15.386	43.421
df	4	4
Asymp. Sig.	.004	.000

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the both NUB_SB and NUB_CA variables with the five SAT categories as both p-values < 0.05, therefore both NUB medians are not equal among the five SAT categories. Therefore, there is statistically significant difference between SAT and NUB.

Figure 4.31: Matrix plot SAT and NUB_SB

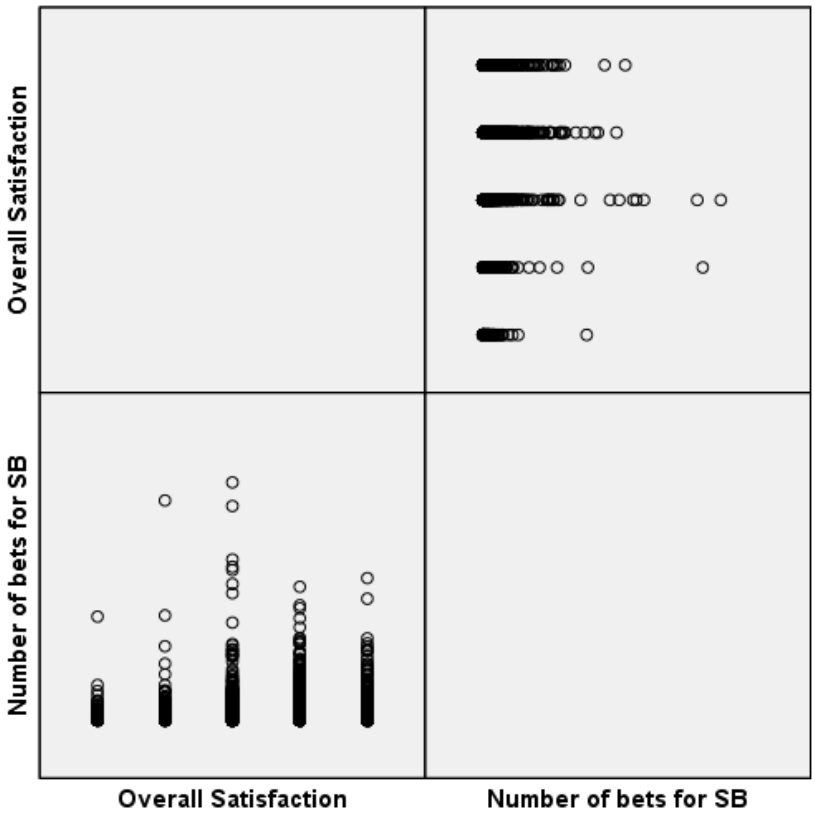
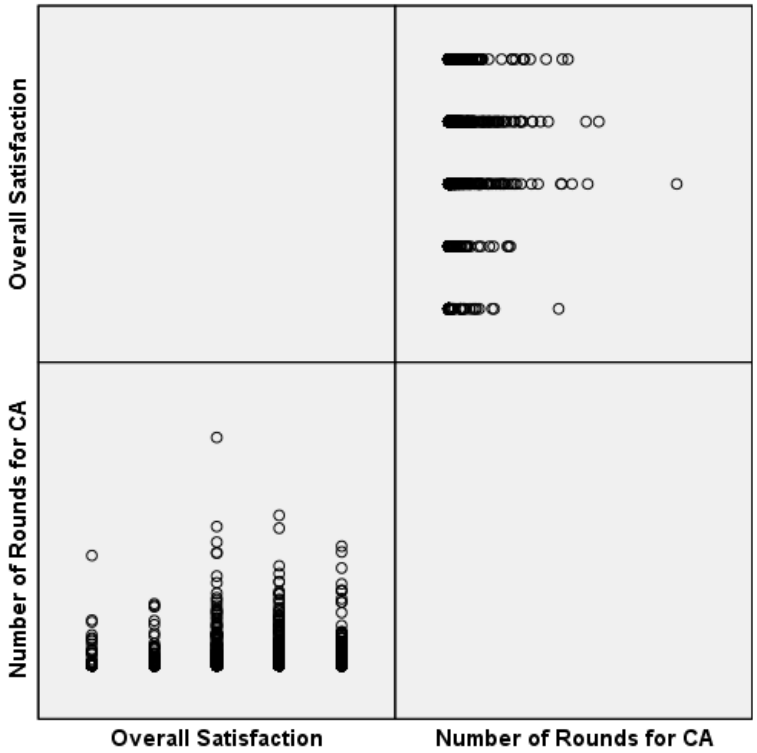


Figure 4.32: Matrix plot SAT and NUB_CA



4.3.5.5. SAT and ACD

The five SAT categories are correlated with all three ACD variables according to the Spearman's tests results. The corresponding significance factors were all found 0.000<0.05, therefore the respective null hypotheses are rejected. The correlation between ACD_CA and the SAT categories is negative.

Table 4.63: SAT and ACD Spearman test

Correlations

Spearman's rho		Overall Satisfaction	Active Days for SB	Active Days for CA	Active Days for both SB and CA
Overall Satisfaction	Correlation Coefficient	1.000	.103**	-.089**	.094**
	Sig. (2-tailed)	.	.000	.000	.000
	N	2690	2690	2690	2690

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

Table 4.64: Kruskal-Wallis test ranks ASB and SAT

	Overall Satisfaction	N	Mean Rank
Active Days for SB	Not at all satisfied	77	1056.94
	Slightly Satisfied	157	1183.06
	Moderately satisfied	687	1250.87
	Very satisfied	1203	1413.07
	Totally satisfied	566	1401.06
	Total	2690	
Active Days for CA	Not at all satisfied	77	1487.82
	Slightly Satisfied	157	1554.39
	Moderately satisfied	687	1434.26
	Very satisfied	1203	1267.08
	Totally satisfied	566	1327.13
	Total	2690	
Active Days for both SB and CA	Not at all satisfied	77	1072.06
	Slightly Satisfied	157	1195.12
	Moderately satisfied	687	1264.06
	Very satisfied	1203	1403.98
	Totally satisfied	566	1398.97
	Total	2690	

Table 4.65: Kruskal-Wallis test statistics TUR and SAT

	Active Days for SB	Active Days for CA	Active Days for both SB and CA
Chi-Square	39.796	40.180	32.583
df	4	4	4
Asymp. Sig.	.000	.000	.000

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the all three ACD variables with the five SAT categories as all three p-values < 0.05, therefore all three ACD medians are not equal among the five

SAT categories. Therefore, there is statistically significant difference between SAT and ACD.

4.3.5.6. SAT and NUD

The five SAT categories are not correlated with the NUD variable as the significance equals 0.143>0.05, therefore the respective null hypothesis cannot be rejected.

Table 4.66: SAT and ACD Spearman test

			Overall Satisfaction	Number of deposits
Spearman's rho	Overall Satisfaction	Correlation Coefficient	1.000	.028
		Sig. (2-tailed)	.	.143
		N	2690	2690
	Number of deposits	Correlation Coefficient	.028	1.000
		Sig. (2-tailed)	.143	.
		N	2690	2690

Table 4.67: Kruskal-Wallis test ranks NUD and SAT

	Overall Satisfaction	N	Mean Rank
Number of deposits	Not at all satisfied	77	1294.10
	Slightly Satisfied	157	1335.72
	Moderately satisfied	687	1343.96
	Very satisfied	1203	1318.88
	Totally satisfied	566	1413.66
	Total	2690	

Table 4.68: Kruskal-Wallis test statistics NUD and SAT

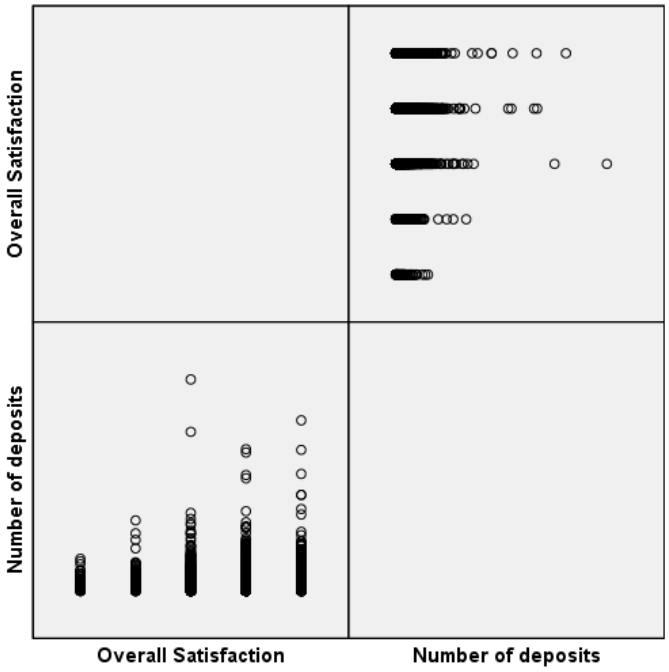
	Number of deposits
Chi-Square	6.144
df	4
Asymp. Sig.	.189

a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 cannot be rejected as the p-value equals 0.189>0.05, therefore the NUD medians are equal among the five SAT categories. Therefore, there is no statistically significant difference between SAT and BUD.

Figure 4.33: Matrix plot SAT and NUD



4.3.5.7. SAT and AND

The five SAT categories are positively correlated with the AND variable according to the Spearman's tests results as the significance factor was found 0.006<0.05, therefore the respective null hypothesis is rejected.

Table 4.69: SAT and AND Spearman test

			Overall Satisfaction	Amount of deposits
Spearman's rho	Overall Satisfaction	Correlation Coefficient	1.000	.053**
		Sig. (2-tailed)	.	.006
		N	2690	2690

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

Table 4.70: Kruskal-Wallis test ranks ASB and SAT

	Overall Satisfaction	N	Mean Rank
Amount of deposits	Not at all satisfied	77	1326.04
	Slightly Satisfied	157	1308.87
	Moderately satisfied	687	1306.08
	Very satisfied	1203	1331.02
	Totally satisfied	566	1436.92
	Total	2690	

Table 4.71: Kruskal-Wallis test statistics TUR and SAT

	Amount of deposits
Chi-Square	10.432
df	4
Asymp. Sig.	.034

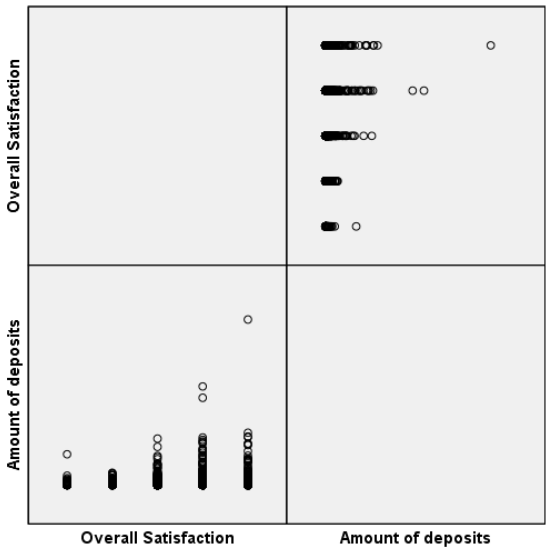
a. Kruskal Wallis Test

b. Grouping Variable: Overall Satisfaction

The H0 is rejected for the AND variable with the five SAT categories as the p-value equals 0.034< 0.05, therefore the AND median is not equal among the five SAT categories.

Therefore, there is statistically significant difference between SAT and AND.

Figure 4.34: Matrix plot SAT and AND



4.3.6. Conclusions

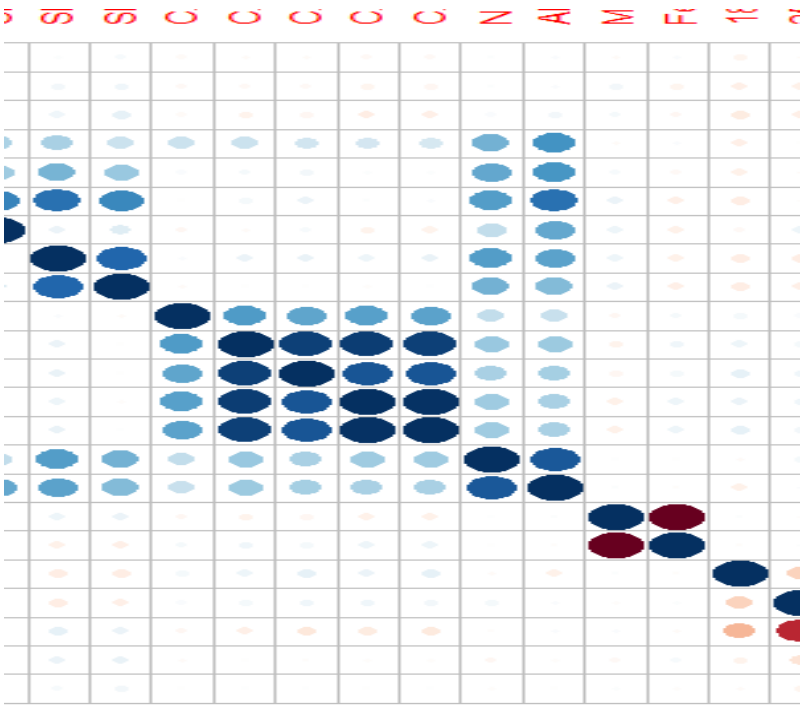
The 4.3 paragraphs brief conclusions are provided in Table 4.72 below and further elaborated in this paragraph:

Table 4.72: Correlations conclusions

	REC	SAT	EXP
ARPU	-	SB, CA, Total (medians)* SB: + Total: +	CA (medians)
TUR	CA (medians)	SB, CA, Total (medians) SB: + CA: - Total: +	SB: +
ABS	CA (medians)	SB, CA (medians) SB: + CA: -	CA (medians)
NUB	-	SB, CA (medians) SB: + CA: -	Medians and SB: +, CA: -
ACD	SB, CA, Both (medians) CA: -	SB, CA, Both (medians) SB: + CA: - Both: +	Medians and SB: +, CA: - Both: +
NUD	-	-	-
AND	-	+ And medians	-
SAT	+		+
EXP	+	+	
AGE	+	+	+
GENDER	-	Sig. Correlation	Sig. Correlation
VIP Level	-	-	-

The results worth discussing are provided below per categorical ordinal variable, while all the correlation analysis results are presented in the following graph:

Figure 4.35: Correlation matrix



The REC variable was found positively correlated with SAT, EXP and AGE group and negatively correlated with ACD_CA. Therefore, as REC increases, SAT, EXT and Age group increase too, while ACD_CA is decreasing.

Moreover, the variables TUR_CA, ABS_CA, NUB_CA, ACD_SB, ACD_CA and ACD (both SB and CA) are correlated with the three REC categories.

The Turnover Casino in the Detractors (TUR_CA) group is statistically significantly higher than both in the Passives and the Promoters groups.

The Average Stake per Bet for casino (ASPB_CA) in the Detractors group is statistically significantly higher than both in the Passives and the Promoters groups.

The Number of Rounds for CA (NUB_CA) in the Detractors group is statistically significantly higher than in the Passives group and Number of Rounds for CA in the Detractors group is statistically significantly higher than in the Promoters group.

Active Days for SB (ACD_SB) in the Detractors group is statistically significantly lower than in the Passives group and Active Days for SB in the Detractors group is statistically significantly lower than in the Promoters group.

Active Days for CA (ACD_CA) in the Detractors group is statistically significantly higher than in the Passives group and Active Days for CA in the Detractors group is statistically significantly higher than in the Promoters group.

Active Days for both in the Detractors group is statistically significantly lower than in the Passives group and Active Days for both in the Detractors group is statistically significantly lower than in the Promoters group.

No other variable was found to be correlated with the REC variable.

As far as the EXP variable is concerned it is positively correlated with TUR_SB, NUB_SB, ACD_SB, ACD_both, EXP, AGE group and VIP Level, negatively correlated with NUB_CA and ACD_CA and correlated with GENDER.

That means that as EXP increases, TUR_SB, NUB_SB, ACD_SB, ACD_both, EXP, AGE group and VIP Level increases too, while NUB_CA and ACD_CA are decreasing.

Also, the variables ARPU_CA, ABS_CA, NUB_CA, NUB_SB, ACD_SB, ACD_CA and ACD_both have significant differences between the 5 EXP categories so, we can conclude that they are also correlated with EXP.

As far as the SAT categories are concerned, they are positively correlated with ARPU_SB, ARPU_total, TUR_SB, TUR_Total, ABS_SB, NUB_SB, ACD_SB, ACD_both, AND, SAT, AGE group and VIP Level. They are negatively correlated with TUR_CA, ABS_CA, NUB_CA and ACD_CA and correlated with GENDER.

That means that as SAT increases, ARPU_SB, ARPU_total, TUR_SB, TUR_Total, ABS_SB, NUB_SB, ACD_SB, ACD_both, AND, EXP, AGE group and VIP Level, increases too, while TUR_CA, ABS_CA, NUB_CA and ACD_CA are decreasing.

Also, the variables ARPU_SB_CA_Total, TUR_SB_CA_Total ABS_SB_CA, NUB_SB_CA, ACD_SB_CA_both have significant differences between the 5 SAT categories so, we can conclude that they are also correlated with SAT.

All the other variables in the analysis have no impact in the SAT categories.

Concluding, the results so far show that there are more correlations for the SAT variable than the others. This could be a sign of more sensitivity for this variable. The strength of the relationship between the independent variables (REC, SAT, EXP) and the dependent ones will be given by the regression model results. So, based on the strength of the models that will be created for each variable, we will have a measure to compare their ability to forecast future behaviors. Also, NUD variable does not correlate with any of the three independent variables and the AND variable correlates only with the SAT scale.

4.4. Regression analysis

Each categorical variable (Male, Female, VIP & AgeGroup[2]) was converted to numerical (dummy variables) using one-hot encoding. As a result, the Gender variable was split into *Gender_Male* and *Gender_Female* and the Age Group into *AgeGroup1, AgeGroup2, AgeGroup3, AgeGroup4, AgeGrou5*, according to the 5 different age groups (18-24, 25-34, 35-54, 55-64, 65+).

For each response variable (*CA_ARPU, CA_TUR, CA_ASB, CA_NUB, CA_ACD*, *SB_ARPU, SB_TUR, SB_ASB, SB_NUB, SB_ACD)*, the exploratory variables REC_cat, SAT, EXP, VIP_Level, AND_euro, NUD, Gender (*Gender_Male* and *Gender_Female*) and AgeGroup (*AgeGroup1, AgeGroup2, AgeGroup3, AgeGroup4, AgeGrou5*) were used. Additionally, for each selected variable, the remaining response variables were treated as predictors (exploratory variables). For example, when selecting *CA_ARPU* as a response variable, apart from the aforementioned exploratory variables, *CA_TUR, CA_ASB, CA_NUB, CA_ACD* were also considered as exploratory in the model. In total, the model consists of 17 exploratory variables.

[2] Here, the variables VIP level and AgeGroup were treated as a categorical, nominal variables

4.5 CASINO Regressions Analysis Results

The metrics used to evaluate the contribution of each subset of selected variables to the model are the RMSE (Root Mean Square Error), the R^2 (R squared), MAE (Mean Absolute Error), and the standard deviation of the resamples. Under this thesis context, the behavior of the RMSE and R^2 is examined in order to assess the impact of the exploratory variables and the performance of our model. Each variable (or subset of variables) that is selected is presented by the model in descending order of importance (impact to the model performance).

4.5.1 Casino Average Revenue per User

This regression analysis treated the Casino Average Revenue per User as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Casino Average Revenue per User predictors.

Recursive feature selection

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	253.7	0.5995	71.06	119.9	0.1789	20.20
8	213.2	0.6778	66.13	105.7	0.2910	20.32
16	244.4	0.6116	66.41	111.9	0.1903	20.32
17	246.1	0.6083	66.94	112.7	0.1886	20.55

Table 4.73: CA_ARPU model

The top 5 variables (out of 8): CA_TUR, VIP_Level, CA_ASB, AND_euro, CA_NUB

The CA_ARPU is selected as a response variable.

We observe that 8 out of 17 variables perform best in terms of R^2 and RMSE, with values 0.6778 and 213.2 respectively. These variables achieved the highest R^2 and lowest RMSE with respect to the other subsets. The top 5 of such variables are CA_TUR, VIP_Level, CA_ASB, AND_euro, CA_NUB.

However, the value of R^2 (=0.6778) is sufficiently large. This indicates that the independent variables are explaining a lot of the variation of the dependent variable (CA_ARPU). Adding more (non-correlated) independent variables to our model that somehow relate to our dependent variable (context) could further increase the model's performance.

The value of R^2 (=0.6778) here is medium high enough and indicates a medium high fit of the model with the data.

4.5.2 Casino Turnover

This regression analysis treated the Casino Turnover as the dependent variable and a ll the other variables in the casino bets and demographics dataset as independent va riables seeking to find the variables that act as Casino Turnover predictors.

Recursive feature selection

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	4173	0.8068	694.3	2950	0.1319	321.5
8	4301	0.8008	734.7	3036	0.1087	326.2
16	4273	0.7998	714.5	3023	0.1207	321.3
17	4301	0.7996	728.4	3050	0.1202	327.5

Table 4.74: CA_TUR model

The top 4 variables (out of 4): CA_ASB, CA_ARPU, CA_NUB, AND_euro

The CA_TUR was treated as a response variable. We observe that 4 variables perform best in terms of R^2 and RMSE, with values 0.8068 and 4173 respectively. These variables achieved the highest R^2 and lowest RMSE with respect to the other subsets. These variables are CA_ASB, CA_ARPU, CA_NUB, AND_euro.

The value of R^2 (=0.8068) here is very high and indicates a very good fit of the model with the data.

4.5.3 Casino Average Stake per Bet

This regression analysis treated the Casino Average Stake per Bet as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Casino Average Stake per Bet predictors.

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	12.39	0.5705	1.489	19.72	0.1080	1.493
8	12.13	0.6665	1.402	19.77	0.1860	1.510
16	12.91	0.5485	1.449	19.57	0.1446	1.502
17	12.91	0.5075	1.471	19.52	0.1708	1.496

Table 4.75: CA_ASB model

The top 5 variables (out of 8): CA_TUR, CA_NUB, NUD, CA_ARPU, CA_ACD

The CA_ASB was treated as a response variable.

We observe that only 5 variable performs best in terms of R^2 and RMSE, with values 0.6665 and 12.13 respectively. These are CA_TUR, CA_NUB, NUD, CA_ARPU, CA_ACD. Such results indicate that the average stake per bet (CA_ASB) depends mainly on the total amount of money placed by the player in Casino rounds (CA_TUR) the number of rounds played (CA_NUB) and the number of times that the customer deposited money in his account (NUD). This is something that is intuitively reasonable.

The value of R^2 (=0.6665) is relatively high, indicating that more exploratory variables are needed for the model to fit better.

4.5.4 Casino number of bets

This regression analysis treated the Casino number of bets as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Casino number of bets predictors.

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD
1	6022	0.2189	2166.4	1669	0.18580	563.9
2	1783	0.9382	355.2	1056	0.05034	130.4
3	2303	0.8927	578.1	1053	0.05961	160.9
4	2641	0.8575	690.8	1033	0.06404	166.0
5	2893	0.8284	781.8	1059	0.07264	178.4
6	2575	0.8639	662.0	1059	0.06519	162.1
9	2669	0.8535	691.6	1032	0.06428	159.8

Table 4.76: CA_NUB model

The top 2 variables (out of 2): CA_ASB, CA_TUR

The CA_NUB was treated as a response variable.

We observe that 2 variables perform best in terms of R^2 and RMSE, with values 0.9382 and 1783 respectively. These variables are CA_ASB, CA_TUR. Such result indicates that the number of rounds played by a Casino player (CA_NUB) depends on the average stake per bet (CA_ASB) as well as the total amount of money placed by the player in Casino rounds (CA_TUR).

The value of R^2 (=0.9382) is quite high indicating a particularly good fit of the model with the actual data.

4.5.5 Casino active days

This regression analysis treated the Casino active days as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Casino active days predictors.

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD
1	3.648	0.7451	1.760	0.3156	0.04516	0.1408
2	3.208	0.8001	1.587	0.2877	0.03657	0.1313
3	3.077	0.8169	1.535	0.2691	0.03099	0.1249
4	3.051	0.8199	1.540	0.2649	0.02998	0.1297
5	3.027	0.8226	1.536	0.2452	0.02774	0.1175
6	3.048	0.8199	1.507	0.2646	0.03008	0.1217
11	3.041	0.8208	1.490	0.2616	0.03000	0.1234

Table 4.77: CA_ACD model

The top 5 variables (out of 5): CA_NUB, CA_TUR, AND_euro, CA_ASB, CA_ARPU

The CA_ACD was treated as a response variable.

We observe that 5 variables perform best in terms of R^2 and RMSE, with values 0.8226 and 3.027 respectively. These variables are CA_NUB, CA_TUR, AND_euro, CA_ASB, CA_ARPU. Such results indicate that the number of active days for a Casino player (CA_NUB) mainly depends on the aforementioned variables, each of which are presented with descending order of importance.

The value of R^2 (=0.8226) is very high and indicates a very good fit of the model with the data.

4.5.6 Number of deposits

This regression analysis treated the Number of deposits as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Number of deposits predictors.

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	24.18	0.5203	13.03	5.384	0.08773	1.408
8	24.26	0.5162	12.69	5.255	0.08563	1.346
16	24.62	0.5036	12.70	5.004	0.08279	1.333
17	24.53	0.5069	12.69	4.923	0.08072	1.303

Table 4.78: NUD model

The top 4 variables (out of 4): AND_euro, CA_ACD, CA_ASB, VIP_Level

The AND_euro was treated as a response variable.

We observe that 4 variables perform best in terms of R^2 and RMSE, with values 0.5203 and 24.18 respectively. These variables are AND_euro, CA_ACD, CA_ASB, VIP_Level.

The value of R^2 (=0.5203) is medium high, indicating a medium high fit of the model with the data. Additional exploratory variables need to be supplied to the model (SB variables)

4.5.7 Amount of deposits

This regression analysis treated the Amount of deposits as the dependent variable and all the other variables in the casino bets and demographics dataset as independent variables seeking to find the variables that act as Amount of deposits predictors.

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
1	1.347	0.4809	483.5	699.9	0.1455	104.3
2	1.293	0.5271	449.6	732.8	0.1721	105.1
3	1.284	0.5315	452.3	753.7	0.1849	110.8
4	1.290	0.5270	463.7	759.4	0.1857	108.0
5	1.300	0.5232	475.9	761.0	0.1821	101.0
6	1.302	0.5274	457.5	732.8	0.1788	102.7
9	1.336	0.5101	466.2	736.0	0.1776	106.4

Table 4.79: AND_euro model

The top 3 variables (out of 3): VIP_Level, CA_TUR, CA_ARPU

The AND_euro was treated as a response variable.

We observe that 3 variables perform best in terms of R^2 and RMSE, with values 0.5315 and 1284 respectively. These variables are VIP_Level,CA_TUR, CA_ARPU. Such results indicate that the total amount that the customer deposited in his account (AND_euro) depends mainly the Customer VIP Level he/she is categorized (VIP_Level), the total amount of money placed by the player in Casino rounds (CA_TUR) and the average revenue per Casino user (CA_ARPU).

The value of R^2 (=0.5315) is medium high, indicating a medium high fit of the model with the data. Additional exploratory variables need to be supplied to the model (SB variables).

4.6 Casino Model fitting

As aforementioned, since there is no significant correlation among the dependent variables and the other exploratory variables (SAT, EXP, REC), neither normality regarding our data, we embarked on a non-parametric approach. Apart from Random

Forest Regression, four (4) different models were tested: Generalized Linear Model (GLM), Stochastic Gradient Boosting (GBM), Bagged Decision Tree (Bagged Tree) and Multivariate Adaptive Regression Spline (MARS). In order to train the models, 75% of the total dataset was used as train set. The remaining 25% was used a test set. In addition, 10-fold cross validation was repeated 5 times. For training each model, only the variables that have the greatest impact to each response variable were considered (according to the RFE selection).

Casino variables

For variables including outliers, the extreme values were removed.

CA_ARPU: 4 observations above the value of "5.000"

CA_TUR: 16 observations above the value of "31.000"

CA_ASB: 17 observations above the value of "20"

CA_NUB: 10 observations above the value of "50.000"

The results with reference to each models' performance are shown below.

Mtry[3]: Number of variables that is randomly collected to be sampled at each split time.

4.6.1 Casino Average Revenue per User Modelling

The results of the model fitting for the 5 independent variables identified as strong Casino Average Revenue per User predictors are provided in this section.

Selected exploratory features from RFE (via RF): CA_TUR, VIP_Level, CA_ASB, AND_euro, CA_NUB

[3] *Model variable that shows how many variables the model uses to have the lowest possible error (RMSE)*

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	1.219.113	1.698.341	2.130.464	2.478.880
2	Bagged Tree	2.359.134	3.328.671	6.553.620	7.738.312
3	MARS	2.451.542	2.720.573	7.538.710	8.111.099
4	GBM	3.950.414	6.330.802	10.079.898	13.457.220
5	GLM	2.618.605	2.996.753	7.733.263	8.670.863

Table 4.80: CA_ARPU Modelling results

Best Model: Random Forest (RF)

(A)

The final value used for the model was mtry = 5.

Rsquared = 0.7245891

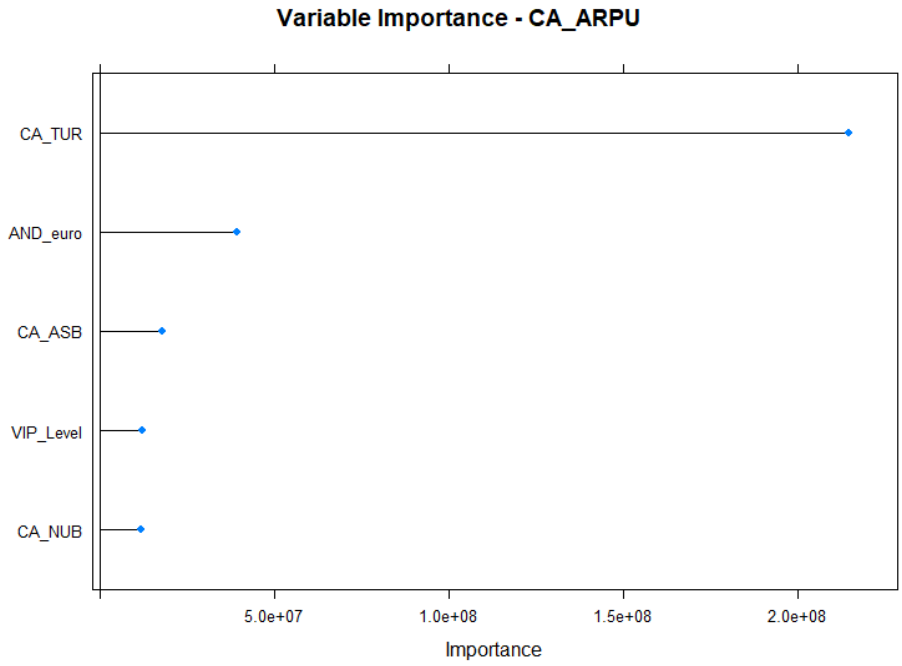


Figure 4.36: ARPU_CA importance

The importance of each variable to the model is illustrated where the response variable is CA_ARP. The optimal "mtry" of the model equals to 5 indicating that all the 5 variables contribute to the model, with variable CA_TUR contributing the most.

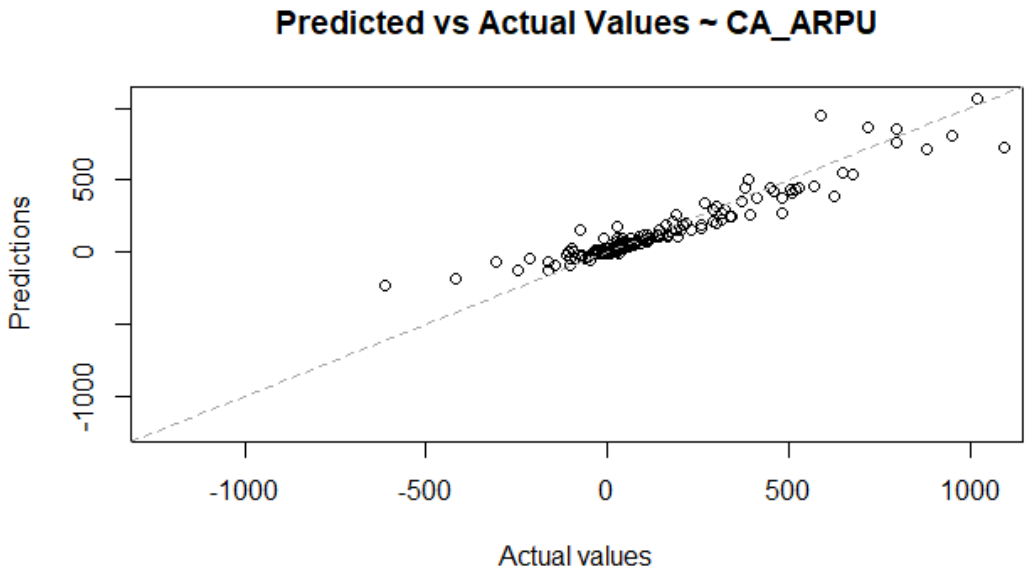


Figure 4.37: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the (in absolute value) prediction error (MAE) on the test set is 24.78880 ~ 25 euro.

Above is shown the model's (RF) predictive performance. In relation to the Average Revenue per User, we observe that the model did not capture with accuracy large values (possible outliers). This is a reason that increases RMSE's sensitivity. Nevertheless, the overall fit with the data (train and test) is good.

4.6.2 Casino Turnover Modelling

The results of the model fitting for the 4 independent variables identified as strong Casino Turnover predictors are provided in this section.

Selected exploratory features from RFE (via RF): CA_ASB, CA_ARPU, CA_NUB, AND_euro

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	3.062.913	3.812.741	1.129.881	1.249.328
2	Bagged Tree	5.837.608	7.155.717	6.752.028	8.615.704
3	MARS	7.851.310	11.460.337	7.890.107	8.107.570
4	GBM	9.386.221	9.191.333	25.782.502	32.298.379
5	GLM	6.809.153	9.072.036	14.901.649	17.213.252

Table 4.81: CA_TUR_Modelling results

Best Model: Random Forest (RF)

(A)

The final value used for the model was mtry = 2.

Rsquared = 0.8556966

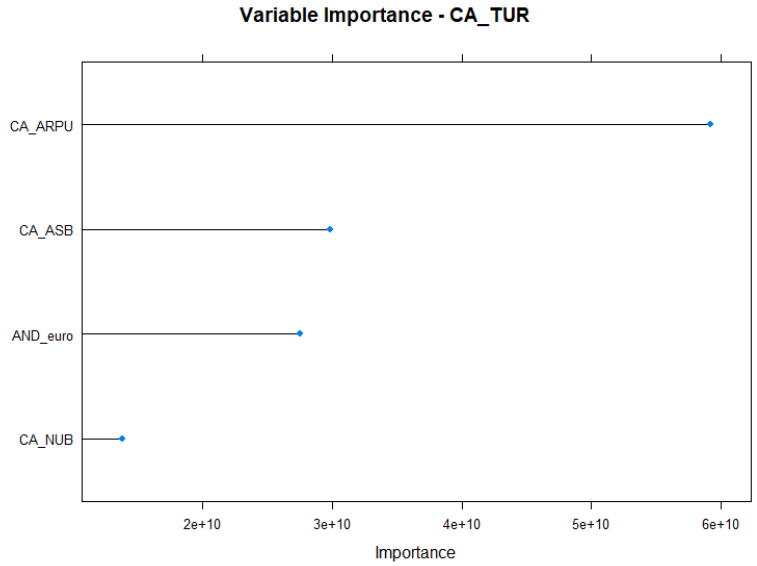


Figure 4.38: TUR_CA importance

Here, the value of "mtry" that results to the minimum RMSE is 2. As it is observed from the variable Importance plot, the 2 variables that contribute the most (explain) to the model are CA_ARPU and CA_ASB.

(B)

Performance on test set – Random Forest

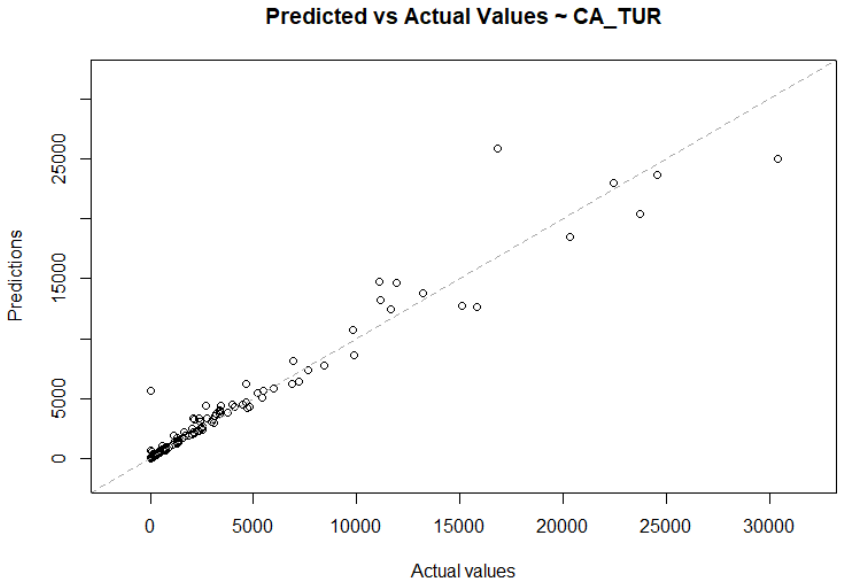


Figure 4.39: Performance on test set – Random Forest

Random Forest (RF), on the test set, has the best performance with respect to other models selected in terms of RMSE and MAE. Bagged Tree performs best regarding the train data, however, the difference with respect to the RF performance it is not significant. Thus, we choose the RF as the most suitable model for prediction.

4.6.3 Casino Average Stake per bet Modelling

The results of the model fitting for the 5 independent variables identified as strong Casino Average Stake per bet predictors are provided in this section.

Selected exploratory features from RFE (via RF): CA_TUR, CA_NUB, NUD, CA_ARPU, CA_ACD

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	929.786	1.047.168	0.2412352	0.199688
2	Bagged Tree	2.046.959	3.949.714	15.045.548	3.190.086
3	MARS	2.503.133	2.651.105	65.602.450	8.068.459
4	GBM	2.399.443	4.732.762	14.218.191	3.354.387
5	GLM	2.208.598	4.263.795	26.153.964	4.250.938

Table 4.82: CA_ASB_Modelling rersults

Best Model: Random Forest (RF)

(A)

The final value used for the model was mtry = 2.

Rsquared = 0.7326957

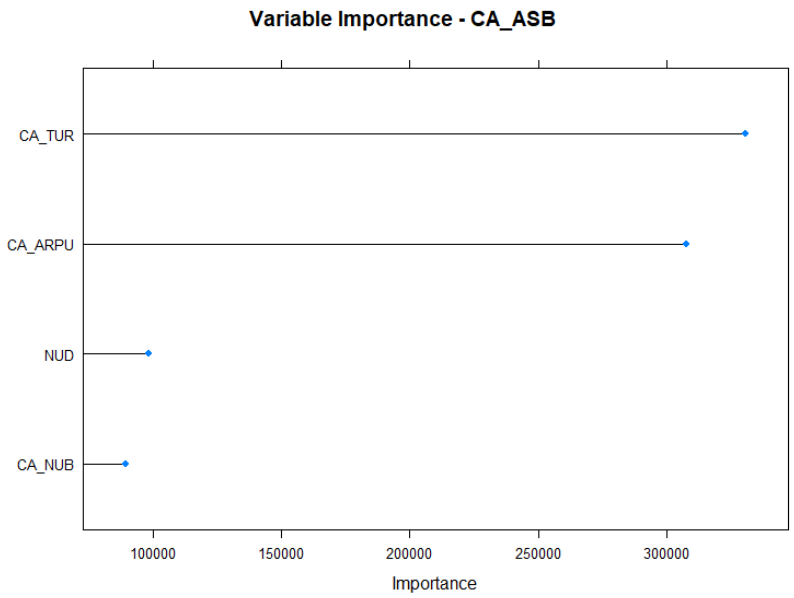


Figure 4.40: ASB_CA importance

The value of "mtry" that results to the minimum RMSE is 2. As it is observed from the Variable Importance plot, the 2 variables that contribute the most (explain) to the model are CA_TUR and CA_ARPU. Variables NUD and CA_NUB appear to contribute the least to the model.

(B)

Performance on test set – Random Forest

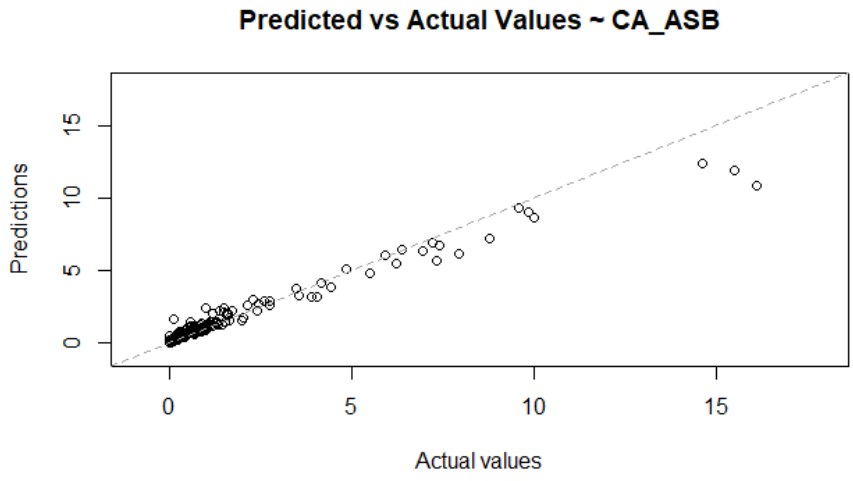


Figure 4.41: Performance on test set – Random Forest

In terms of RMSE (as well as MAE), apart from Random Forest, Bagged Tree and GLM have also a good performance. According to the MAE, on the test set, the average prediction error (Random Forest) is 1.599688.

4.6.4 Casino number of bets Modelling

The results of the model fitting for the 2 independent variables identified as strong Casino number of bets predictors are provided in this section.

Selected exploratory features from RFE (via RF): CA_ASB, CA_TUR

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	9.659.683	1.177.059	1.210.576	1.265.849
2	Bagged Tree	21.676.949	3.221.426	6.717.448	7.125.793
3	MARS	44.248.452	4.788.470	15.054.627	13.338.344
4	GBM	78.384.620	8.638.583	30.042.625	29.002.726
5	GLM	60.760.963	6.574.101	26.052.475	24.582.175

Table 4.83: CA_NUB_Modelling results

Best Model: Random Forest (RF)

A)

Tuning parameter 'mtry' was held constant at a value of 2.

Rsquared = 0.9512655

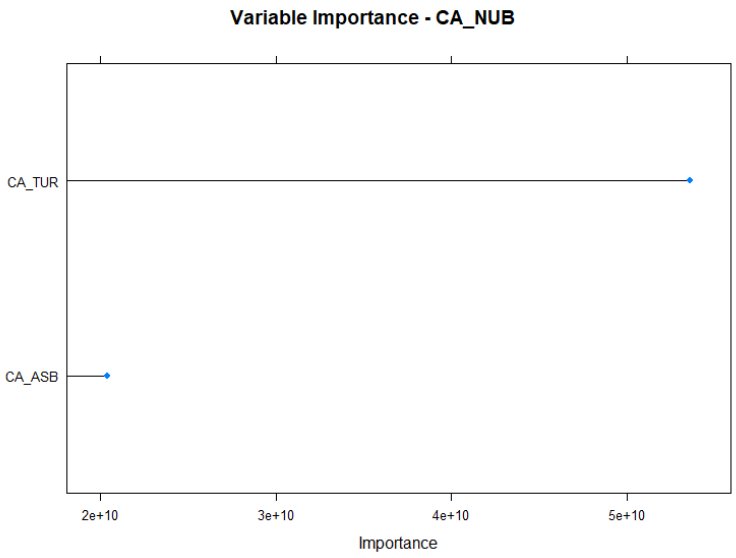


Figure 4.42: NUB_CA importance

The value of "mtry" that results to the minimum RMSE is 2. As it is observed from the Variable Importance plot, the 2 variables differ in terms of their exploratory impact to the model. Variable CA_TUR explains most of the variance in relation to the response variable CA_NUB, contrary to the variable CA_ASB.

(B)

Performance on test set – Random Forest

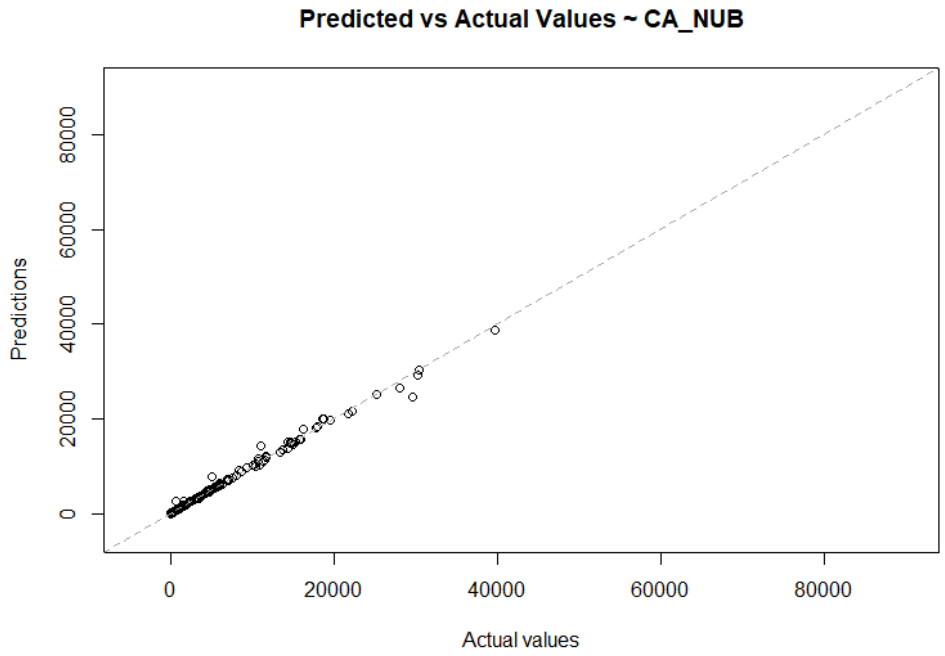


Figure 4.43: Performance on test set – Random Forest

Above is shown the model's (RF) predictive performance. In relation to the Number of Rounds played by a customer, we observe that the model captures the unseen data with high accuracy. This signifies that the variable CA_NUB is well explained by the model and in particular by the exploratory variables CA_ASB, CA_TUR.

4.6.5 Casino active days Modelling

The results of the model fitting for the 5 independent variables identified as strong Casino active days predictors are provided in this section.

Selected exploratory features from RFE (via RF): CA_NUB, CA_TUR, AND_euro, CA_ASB, CA_ARPU

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	1.332.688	1.309.148	0.6640569	0.6465337
2	Bagged Tree	2.901.741	2.773.907	16.541.827	16.146.937
3	MARS	6.208.457	6.098.042	40.600.085	39.878.719
4	GBM	5.326.633	5.106.697	26.787.607	25.088.226
5	GLM	5.060.612	4.892.797	35.829.129	34.746.550

Table 4.84: CA_ACD_Modelling results

Best Model: Random Forest (RF)

(A)

The final value used for the model was mtry = 3.

Rsquared = 0.8284512

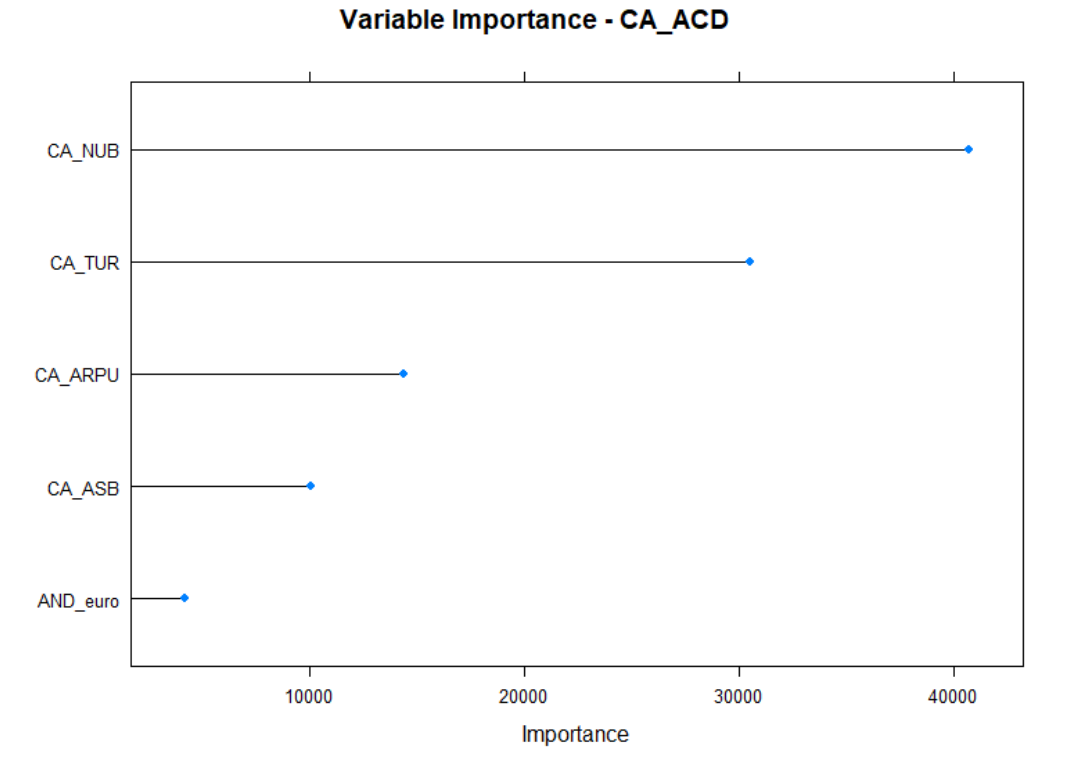


Figure 4.44: ACD_CA importance

The value of "mtry" that results to the minimum RMSE is 3. As it is observed from the Variable Importance plot, the 3 variables that contribute the most (explain) to the model are CA_NUB, CA_TUR and CA_ARPU.

(B)

Performance on test set – Random Forest

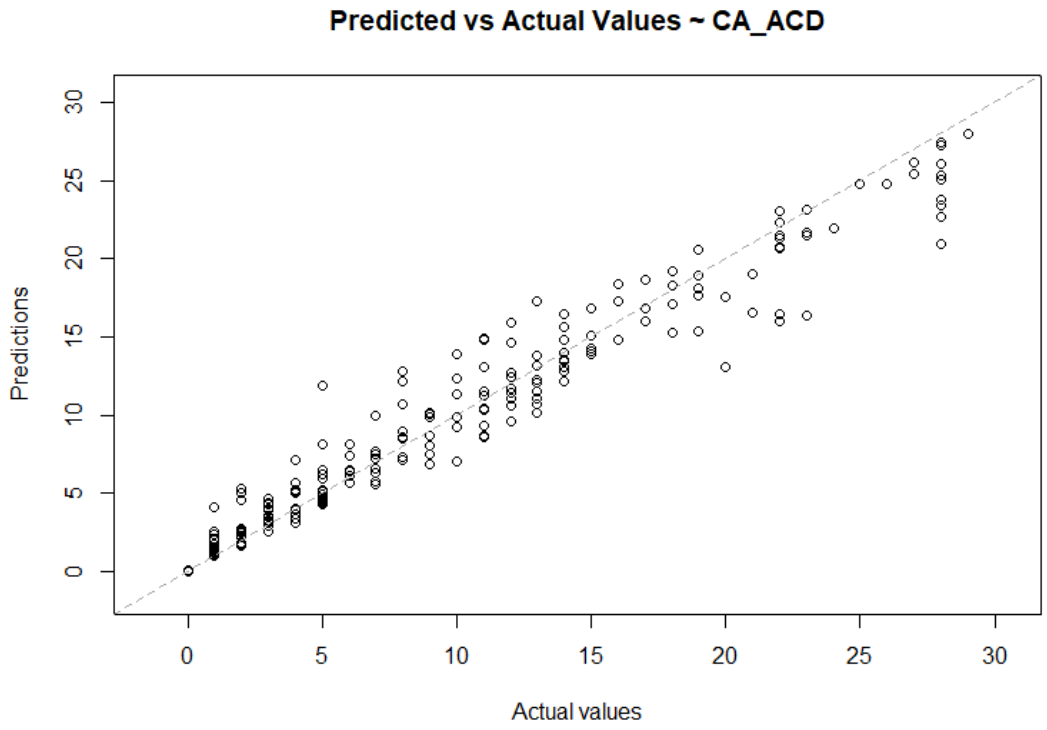


Figure 4.45: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the prediction error (MAE) on the test set is 0.6465337 ~ 1 day.

4.7 SPORTSBOOK Regressions Analysis Results

4.7.1 Sports bets average revenue per user

This regression analysis treated the Sports bets average revenue per user as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Sports bets average revenue per user predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	1.157	0.6044	269.0	1.118	0.2206	115.0
8	1.167	0.6313	255.9	1.113	0.2405	114.0
16	1.158	0.6043	247.5	1.126	0.2606	116.2
17	1.138	0.6452	243.7	1.077	0.2639	107.4

Table 4.85: SB_ARPU model

The top 5 variables (out of 17): VIP_Level, SB_NUB, NUD, AND_euro, SB_ASB

The SB_ARPU was treated as a response variable.

We observe that 17 variables perform best in terms of R^2 and RMSE, with values 0.6452 and 1138 respectively. These variables achieved the highest R^2 and lowest RMSE with respect to the other subsets. The top 5 of such variables are VIP_Level, SB_NUB, NUD, AND_euro, SB_ASB.

4.7.2 Sports bets turnover

This regression analysis treated the Sports bets turnover as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Sports bets turnover predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	5.137	0.8062	906.0	4.161	0.1057	474.6
8	5.322	0.7815	965.9	4.155	0.1149	442.5
16	5.322	0.7860	936.6	4.140	0.1133	443.5
17	5.304	0.7830	948.4	4.092	0.1146	439.3

Table 4.86: SB_TUR model

The top 4 variables (out of 4): SB_ASB, AND_euro, SB_ARPU, SB_NUB

The SB_TUR was treated as a response variable.

We observe that 4 variables perform best in terms of R2 and RMSE, with values 0.8062 and 5137 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. These variables are SB_ASB, AND_euro, SB_ARPU, SB_NUB.

4.7.3 Sports bets average stake per bet

This regression analysis treated the Sports bets average stake per bet as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Sports bets average stake per bet predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Recursive feature selection

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	189.8	0.5387	20.73	310.7	0.2916	22.51
8	189.0	0.5979	19.09	308.7	0.3035	22.69
16	187.5	0.6015	18.71	307.8	0.2978	22.60
17	186.9	0.6004	18.76	307.6	0.2929	22.53

Table 4.87: SB_ASB model

The top 5 variables (out of 17): SB_TUR, SB_ARPU, AND_euro, VIP_Level, SB_NUB

Here the SB_ASB is selected as a response variable.

We observe that 17 variables perform best in terms of R2 and RMSE, with values 0.6004 and 186.9 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. The top 5 of such variables SB_TUR, SB_ARPU, AND_euro, VIP_Level, SB_NUB.

4.7.4 Sports bets number of bets

This regression analysis treated the Sports bets number of bets as the dependent variable and all the other variables in the sportsbook and demographics dataset as

independent variables seeking to find the variables that act as Sports bets number of bets predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	170.3	0.7003	62.74	49.21	0.09503	8.267
8	177.5	0.6617	68.24	50.92	0.10754	8.746
16	171.6	0.6858	64.75	50.46	0.10403	8.407
17	173.8	0.6780	66.41	50.14	0.10237	8.494

Table 4.88: SB_NUB model

The top 4 variables (out of 4): SB_ASB, SB_TUR, SB_ACD, SB_ARPU

Here the SB_NUB is selected as a response variable.

We observe that 4 variables perform best in terms of R2 and RMSE, with values 0.7003 and 170.3 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. These variables are SB_ASB, SB_TUR, SB_ACD, SB_ARPU.

4.7.5 Sports bets active days

This regression analysis treated the Sports bets active days as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Sports bets active days predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	4.752	0.7249	3.627	0.2566	0.03414	0.2147
8	4.644	0.7382	3.578	0.2501	0.03291	0.2014
16	4.627	0.7400	3.546	0.2482	0.03216	0.1984
17	4.636	0.7390	3.560	0.2499	0.03267	0.2005

Table 4.89: SB_ACD model

The top 5 variables (out of 16): SB_NUB, NUD, SB_ASB, SB_TUR, AND_euro

Here the SB_ACD is selected as a response variable.

We observe that 16 variables perform best in terms of R2 and RMSE, with values 0.7400 and 4.627 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. The top 5 of such variables are SB_NUB, NUD, SB_ASB, SB_TUR, AND_euro.

4.7.6 Number of deposits

This regression analysis treated the Number of deposits as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Number of deposits predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	24.19	0.5601	11.47	5.200	0.07300	11.341
8	23.47	0.5860	11.11	4.483	0.06354	10.193
16	23.18	0.5955	10.93	4.279	0.07163	0.9603
19	22.96	0.6341	10.80	4.170	0.07468	0.9402

Table 4.90: NUD model

The top 5 variables (out of 19): AND_euro, SB_ASB, VIP_Level, TOTAL_ACD, TOTAL_ARPU

The NUD was treated as a response variable.

We observe that 19 variables perform best in terms of R2 and RMSE, with values 0.6341 and 22.96 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. The top 5 of such variables are AND_euro, SB_ASB, VIP_Level, TOTAL_ACD, TOTAL_ARPU.

4.7.7 Amount of deposits

This regression analysis treated the Amount of deposits as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Amount of deposits predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Outer resampling method: Cross-Validated (10 fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	803.5	0.7515	239.5	332.8	0.09030	44.70
8	781.3	0.7692	223.2	329.8	0.09304	42.12
16	782.8	0.7685	221.9	334.3	0.09617	43.06
19	783.3	0.7693	222.2	336.6	0.09674	43.55

Table 4.91: AND_euro model

The top 3 variables (out of 8): TOTAL_TUR, NUD, TOTAL_ARPU

The AND_euro was treated as a response variable.

We observe that 8 variables perform best in terms of R2 and RMSE, with values 0.7692 and 781.3respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. The top 3 of such variables are TOTAL_TUR, NUD, TOTAL_ARPU.

4.7.8 Total average revenue per user

This regression analysis treated the Total average revenue per user as the dependent variable and all the other variables in the sportsbook and demographics dataset as independent variables seeking to find the variables that act as Total average revenue per user predictors.

Recursive feature elimination (RFE) via Random forest (RF)

Recursive feature selection

Outer resampling method: Cross-Validated (10-fold, repeated 5 times)

Resampling performance over subset size:

Variables	RMSE	Rsquared	MAE	RMSESD	RsquaredSD	MAESD Selected
4	750.1	0.4685	231.1	370.6	0.2471	51.28
8	743.9	0.5003	229.3	385.3	0.2579	51.02
16	710.1	0.5350	222.2	382.9	0.2668	50.35
19	693.2	0.5462	220.1	383.6	0.2717	50.49

Table 4.92: TOTAL_ARPU model

The top 5 variables (out of 19): VIP_Level, AND_euro, NUD, CA_NUB, TOTAL_ACD

The TOTAL_ARPU was treated as a response variable.

We observe that 19 variables perform best in terms of R2 and RMSE, with values 0.5462 and 693.2 respectively. These variables achieved the highest R2 and lowest RMSE with respect to the other subsets. The top 5 of such variables are VIP_Level, AND_euro, NUD, CA_NUB, TOTAL_ACD.

4.8 Sportsbook model fitting

Sport bet variables

For variables including outliers, the extreme values were removed.

SB_ARPU: 13 observations above the value of “5.000”

SB_TUR: 42 observations above the value of “20.000”

SB_ASB: 47 observations above the value of “100”

And_EURO: 3 observations above the value of “20.000”

TOTAL_ARPU: 34 observations above the value of “3.000”

Mtry: Number of variables that is randomly collected to be sampled at each split time.

4.8.1 Sports bets average revenue per user Modelling

The results of the model fitting for the 5 independent variables identified as strong Sports bets average revenue per user predictors are provided in this section.

Selected exploratory features from RFE (via RF): VIP_Level, SB_NUB, NUD, AND_euro, SB_ASB

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	2.699.546	2.976.807	6.722.967	7.236.589
2	Bagged Tree	4.211.531	4.316.533	12.160.786	12.301.732
3	MARS	5.299.828	5.329.135	22.730.950	23.839.985
4	GBM	16.214.088	16.726.695	108.380.740	119.401.080
5	GLM	6.478.572	7.405.277	30.961.158	29.574.461

Table 4.93: SB_ARPU_Modelling results

The final value used for the model was mtry = 3.

Rsquared = 0.7245891

(A)

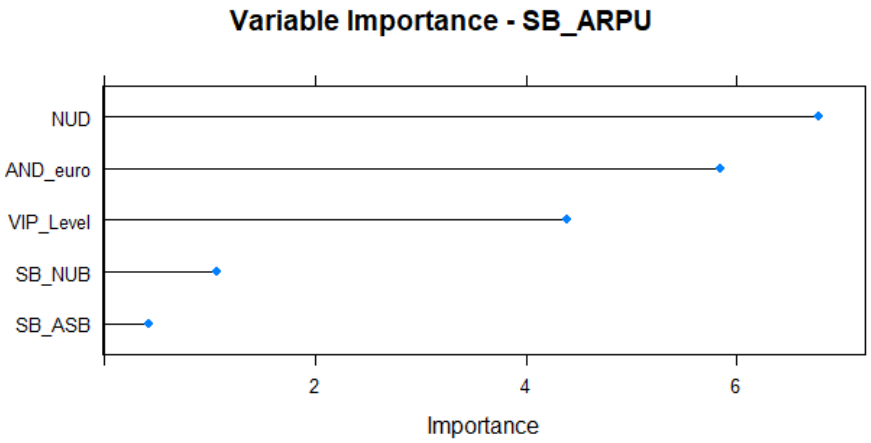


Figure 4.46: ARPU_SB importance

The 'mtry' value with the lowest RMSE is 3. As it is observed from the Variable Importance plot, the 3 variables that contribute the most (explain) to the model are NUD, AND_euro, and VIP_level. This indicates that the gross revenue earned by the company can mainly be explained by the number of times that the customer deposited money in his/her account (NUD), the total amount of money deposited (AND_euro) and the customer VIP level (VIP_level).

(B)

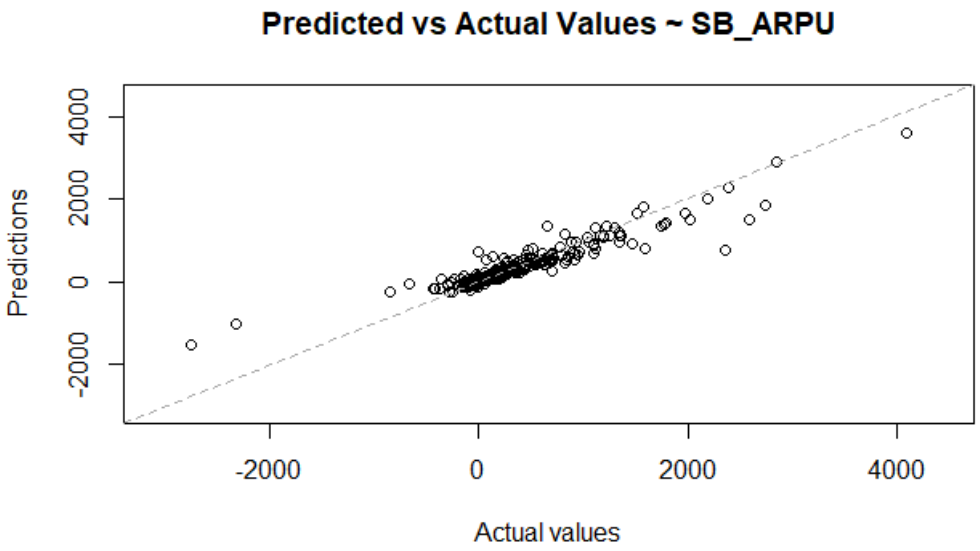


Figure 4.47: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the (in absolute value) prediction error (MAE) on the test set is 72.36589 ~ 70 euro.

4.8.2 Sports bets turnover Modelling

The results of the model fitting for the 4 independent variables identified as strong Sports bets turnover predictors are provided in this section.

Selected exploratory features from RFE (via RF): SB_ASB, AND_euro, SB_ARPU, SB_NUB

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	3.112.083	3.835.737	3.645.667	3.781.211
2	Bagged Tree	5.711.651	5.756.585	4.438.870	4.852.164
3	MARS	5.735.077	5.773.626	3.831.960	4.277.409
4	GBM	3.167.242	4.536.862	8.264.550	6.109.594
5	GLM	4.171.456	5.309.903	8.359.073	8.676.947

Table 4.94: SB_TUR_Modelling results

The final value used for the model was mtry = 4.

Rsquared = 0.8381638

(A)

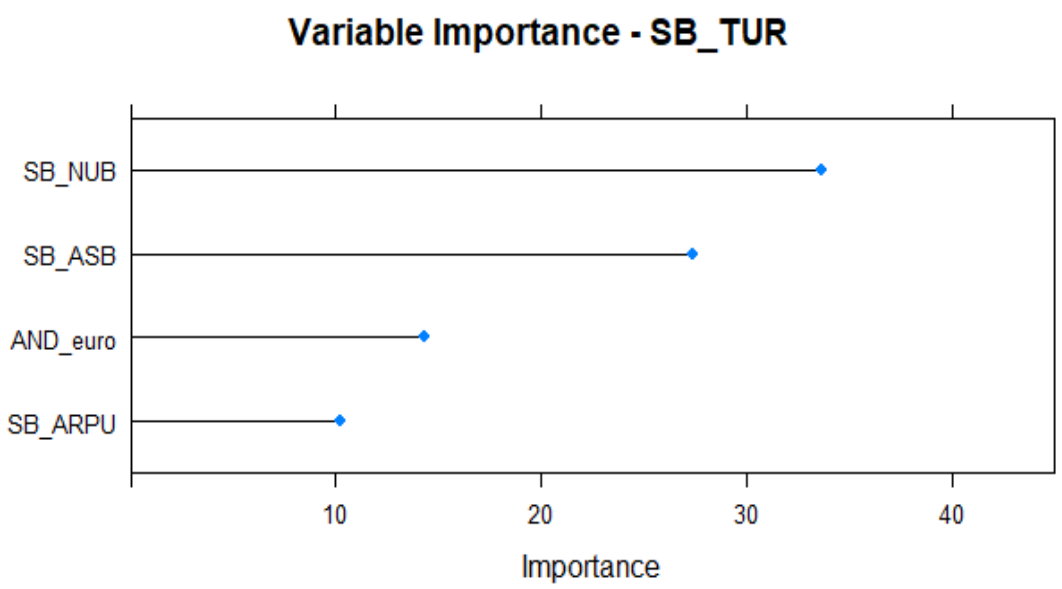


Figure 4.48: TUR_SB importance

The 'mtry' value with th elowest RMSE is 4. As it is observed from the Variable Importance plot, the 4 variables that contribute the most (explain) to the model are SB_NUB, SB_ASB, AND_euro and SB_ARPU. This indicates that the total amount of money placed by a player in Sportsbook bets is particularly related to the number of bets played (SB_NUB) and the amount that has placed on average on every bet (SB_ASB).

(B)

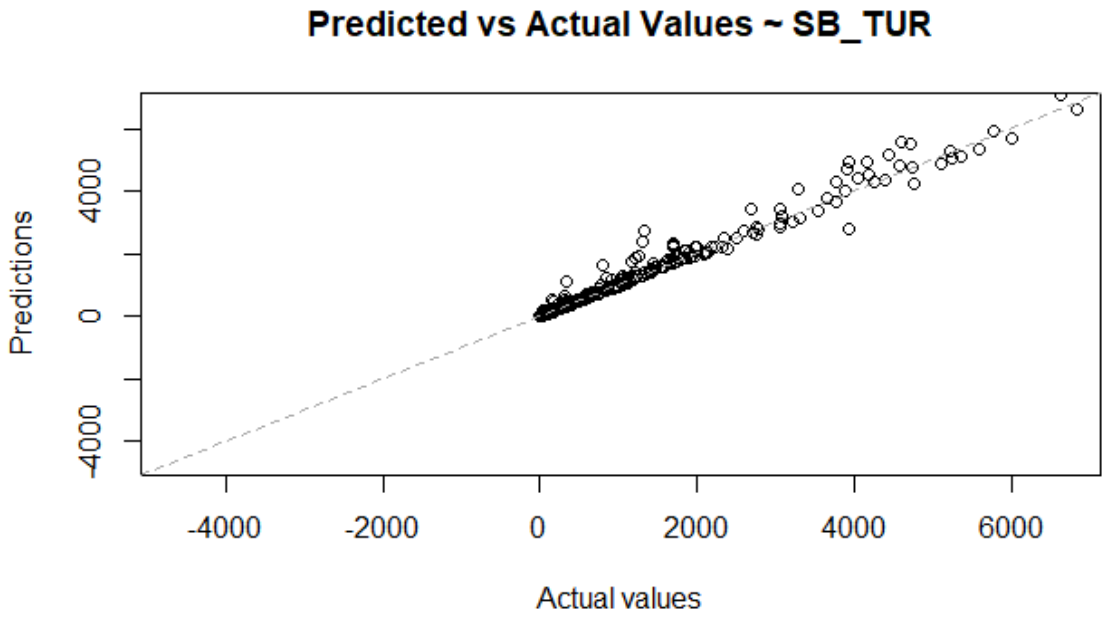


Figure 4.49: Performance on test set – Random Forest

In terms of RMSE, all models (except for MARS and GBM) perform well, with Random Forest (RF) having the lowest value in both sets.

4.8.3 Sports bets average stake per bet Modelling

The results of the model fitting for the 5 independent variables identified as strong Sports bets average stake per bet predictors are provided in this section.

Selected exploratory features from RFE (via RF): SB_TUR, SB_ARPU, AND_euro, VIP_Level, SB_NUB

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	682.244	808.657	0.975708	113.473
2	Bagged Tree	771.069	891.399	2.302.849	443.344
3	MARS	3.633.727	4.351.506	56.112.199	6.298.239
4	GBM	1.646.691	2.789.267	26.538.471	3.604.395
5	GLM	2.144.709	3.112.237	39.485.360	4.293.050

Table 4.95: SB_ASB_Modelling results

(A)

The final value used for the model was mtry = 4.

Rsquared = 0.7784627

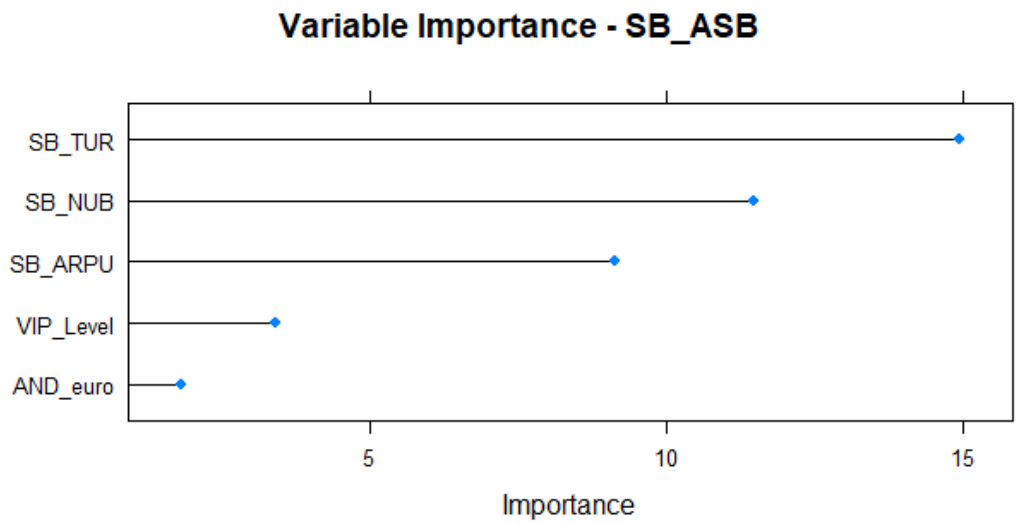


Figure 4.50: ASB_SB importance

The 'mtry' value with the lowest RMSE is 4. As it is observed from the Variable Importance plot, the 4 variables that contribute the most (explain) to the model are SB_TUR, SB_NUB, SB_APRU and VIP_Level. Such variables play an important role in explaining the variance of the model that is interrelated with the response variable SB_ASB.

(B)

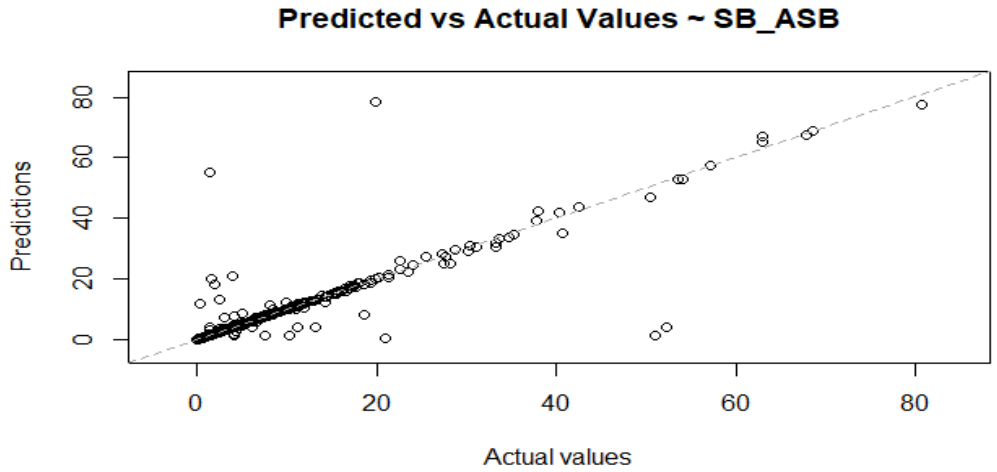


Figure 4.51: Performance on test set – Random Forest

In terms of RMSE (as well as MAE), apart from Random Forest, Bagged Tree and GLM have also a good performance. According to the MAE, on the test set, the average prediction error (Random Forest) is 1.13473 .

4.8.4 Sports bets number of bets Modelling

The results of the model fitting for the 4 independent variables identified as strong Sports bets number of bets predictors are provided in this section.

Selected exploratory features from RFE (via RF): SB_ASB, SB_TUR, SB_ACD, SB_ARPU

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	6.017.441	7.056.429	1.114.689	9.875.981
2	Bagged Tree	15.865.865	16.885.268	7.131.328	66.661.157
3	MARS	27.909.907	34.351.711	11.858.073	185.427.568
4	GBM	47.954.624	55.565.053	21.084.838	297.550.191
5	GLM	28.106.953	31.686.209	13.459.537	135.863.564

Table 4.96: SB_NUB_Modelling results

(A)

The final value used for the model was mtry = 4.

Rsquared = 0.7710673

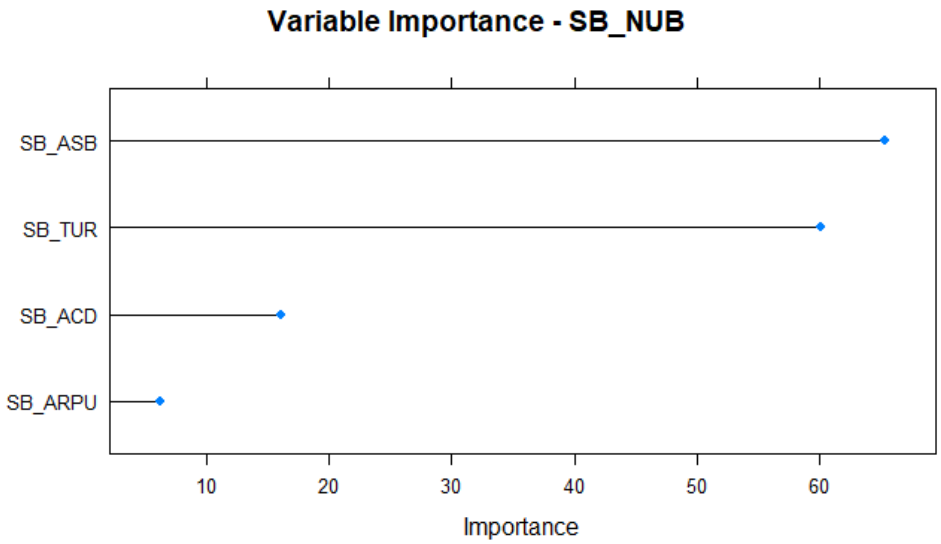


Figure 4.52: NUB_CA importance

The 'mtry' value with the lowest RMSE is 4. As it is observed from the Variable Importance plot, the variables that contribute (explain) to the model are SB_ASB, SB_TUR, SB_ACD, SB_ARPU. However, SB_ASB, SB_TUR appear to have the greatest share in terms of contribution/impact to the exploratory power of the model.

(B)

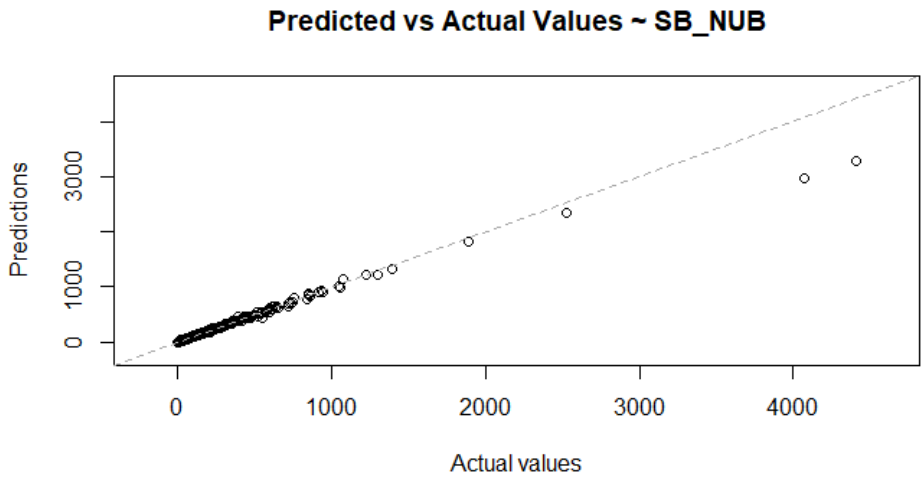


Figure 4.53: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest outperforms the other models in the train test as well as the test set.

4.8.5 Sports bets active days modelling

The results of the model fitting for the 5 independent variables identified as strong Sports bets active days modelling predictors are provided in this section.

Selected exploratory features from RFE (via RF): `SB_NUB, NUD, SB_ASB, SB_TUR, AND_euro`

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	2.148.141	2.260.818	1.604.103	1.689.574
2	Bagged Tree	4.844.291	5.126.365	3.830.814	4.014.331
3	MARS	10.049.177	10.294.881	8.142.360	8.320.714
4	GBM	7.031.943	7.045.108	5.421.819	5.392.086
5	GLM	7.880.013	7.963.770	6.596.243	6.727.942

Table 4.97: SB_ACD_modelling

(A)

The final value used for the model was mtry = 2.

Rsquared = 0.881094

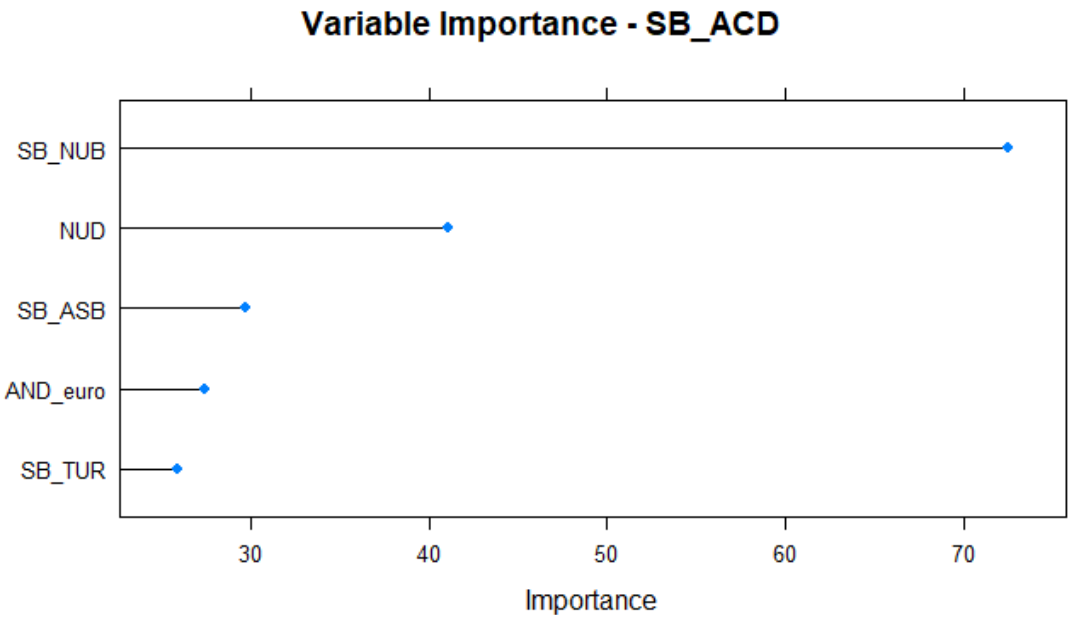


Figure 4.54: ACD_SB importance

The 'mtry' value with the lowest RMSE is 4. As it is observed from the Variable Importance plot, the variables that contribute (explain) to the model are SB_NUB and

NUD. According to the model results, this is an indication that the number of days within the month that a customer has played at least one is related mainly on the number of bets played by a customer for Sportsbook bets and the number of times he/she deposited money in their account.

(B)

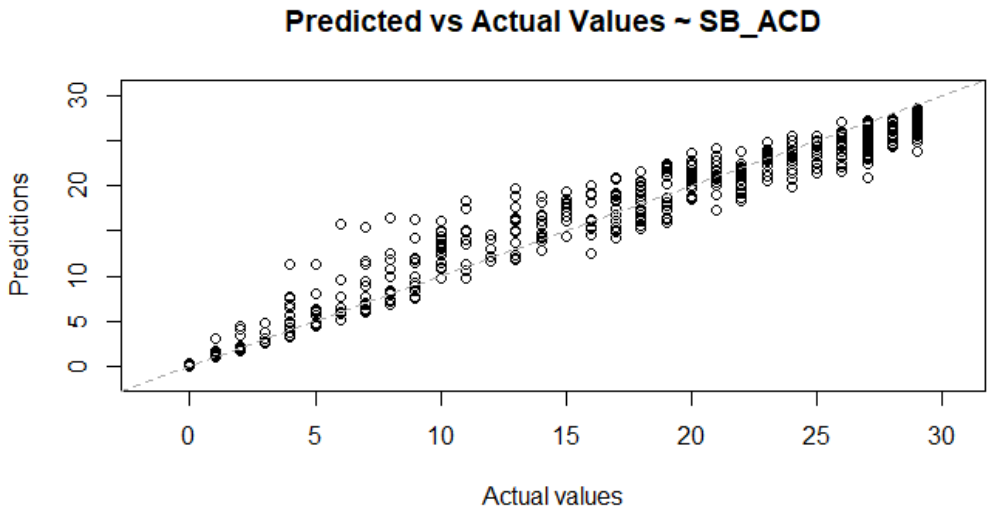


Figure 4.55: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the prediction error (MAE) on the test set is 1.689574 ~ 2 days.

It is worth mentioning that the variables **TOTAL_ACD**, **TOTAL_ARPU** and **TOTAL_TUR** were considered as additional exploratory variables for the model. All of them, represent the sum of the corresponding Sport Bet (SB) and Casino (CA) variables. In addition, correlated variables such as CA_ACD,SB_ACD,CA_ARPU,SB_ARPU,CA_TUR and SB_TUR were removed from the model before training.

4.8.6 Number of deposits modelling

The results of the model fitting for the 5 independent variables identified as strong Number of deposits predictors are provided in this section.

Selected exploratory features from RFE (via RF): `AND_euro`, `SB_ASB`, `VIP_Level`, `TOTAL_ACD`, `TOTAL_ARPU`

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	13.443.533	12.808.026	4.516.759	5.234.972
2	Bagged Tree	16.819.742	15.300.430	10.879.370	10.831.464
3	MARS	48.278.217	46.814.528	33.921.563	31.888.628
4	GBM	37.238.160	36.890.043	23.478.371	22.895.573
5	GLM	25.541.749	21.600.150	14.880.034	13.583.716

Table 4.98: NUD_modelling results

(A)

The final value used for the model was mtry = 3.

Rsquared = 0.7838293

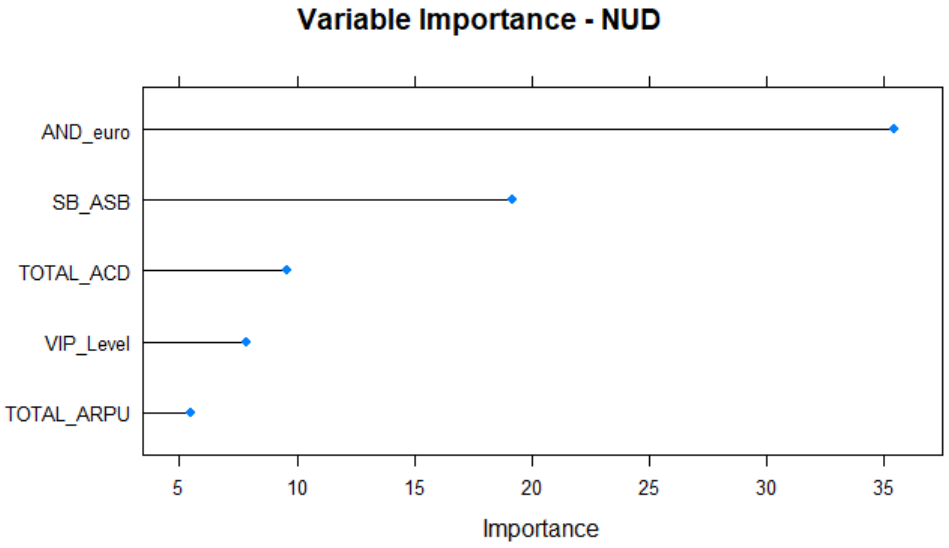


Figure 4.56: NUB_SB importance

The value of "mtry" that results to the minimum RMSE is 3. As it is observed from the Variable Importance plot, the variables that contribute (explain) to the model are AND_euro, SB_ASB, TOTAL_ACD. The variable AND_euro appears to have the greatest share in terms of contribution/impact to the exploratory power of the model.

(B)

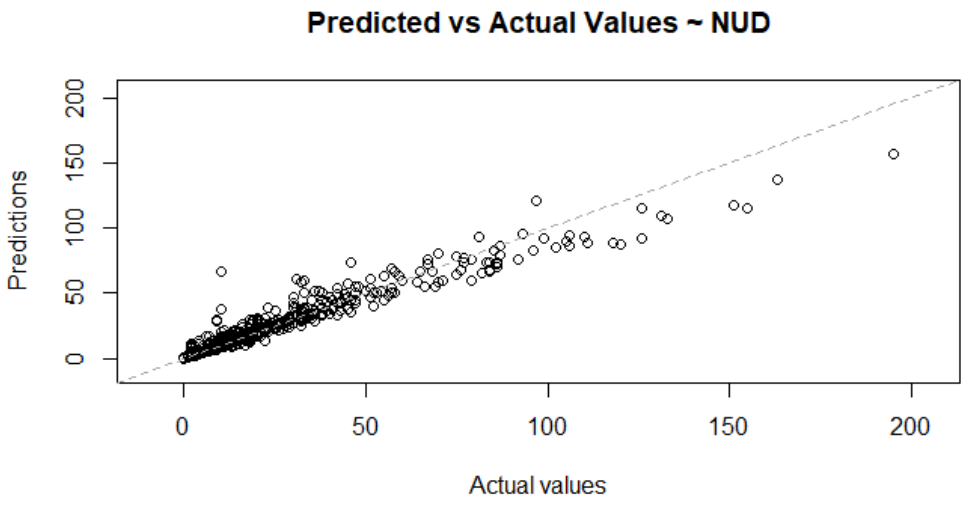


Figure 4.57: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the prediction error (MAE) on the test set is 5.234972 ~ 5 times.

4.8.7 Amount of deposits modelling

The results of the model fitting for the 3 independent variables identified as strong Amount of deposits predictors are provided in this section.

Selected exploratory features from RFE (via RF): TOTAL_TUR, NUD, TOTAL_ARPU

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	4.006.901	3.523.593	1.116.351	8.926.448
2	Bagged Tree	7.708.396	6.364.654	2.976.006	25.165.438
3	MARS	33.137.238	28.376.469	8.096.358	66.671.715
4	GBM	16.394.926	12.634.157	6.002.105	52.579.594
5	GLM	9.222.125	5.100.702	2.959.718	22.895.887

Table 4.99: AND_euro_modelling results

(A)

The final value used for the model was mtry = 2.

Rsquared = 0.8465739

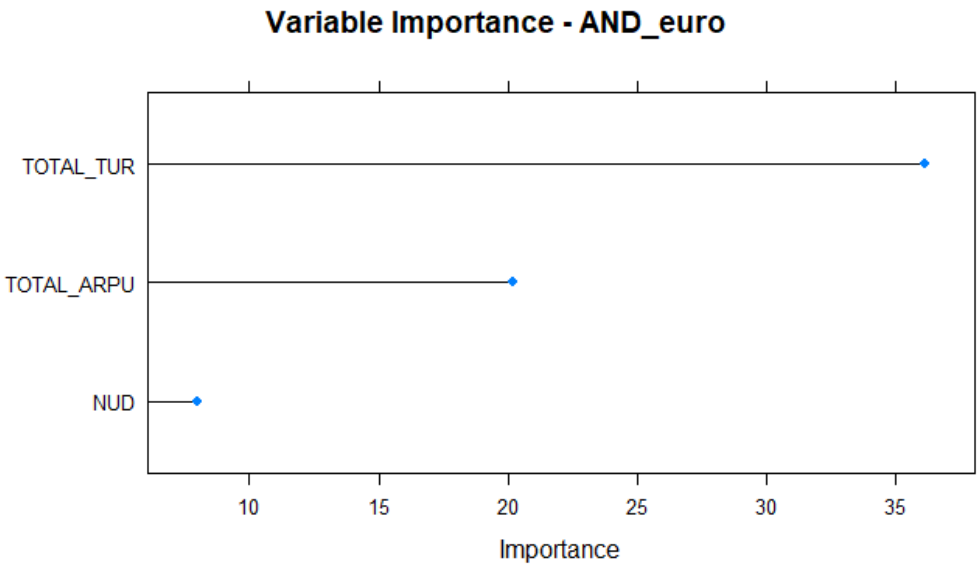


Figure 4.58: AND_SB importance

The 'mtry' value with the lowest RMSE is 2. As it is observed from the Variable Importance plot, the 2 variables that contribute the most (explain) to the model are TOTAL_TUR, TOTAL_ARPU. Such variables play an important role in explaining the variance of the model that is interrelated with the response variable AND_euro.

(B)

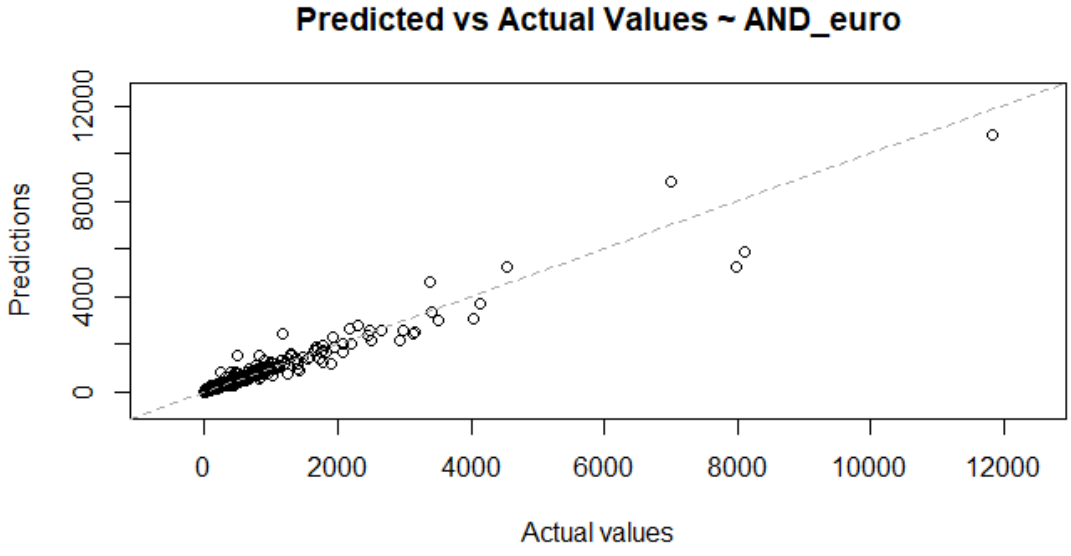


Figure 4-59: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the (in absolute value) prediction error (MAE) on the test set is 89.26448 ~ 89 euro.

4.8.8 Total average revenue per user modelling

The results of the model fitting for the 5 independent variables identified as strong Total average revenue per user predictors are provided in this section.

Selected exploratory features from RFE (via RF): `VIP_Level, AND_euro, NUD, CA_NUB, TOTAL_ACD`

	predictor	RMSE.train	RMSE.test	MAE.train	MAE.test
1	Random Forest	2.675.566	2.993.465	522.257	743.694
2	Bagged Tree	5.978.231	8.663.045	1.447.558	1.853.244
3	MARS	19.219.364	29.063.169	6.341.698	7.968.754
4	GLM	8.132.338	10.569.790	2.549.141	2.867.155
5	GBM	11.362.063	12.409.475	4.635.795	4.616.378

Table 4.100: TOTAL_ARPU_modelling results

(A)

The final value used for the model was mtry = 2.

Rsquared = 0.7382643

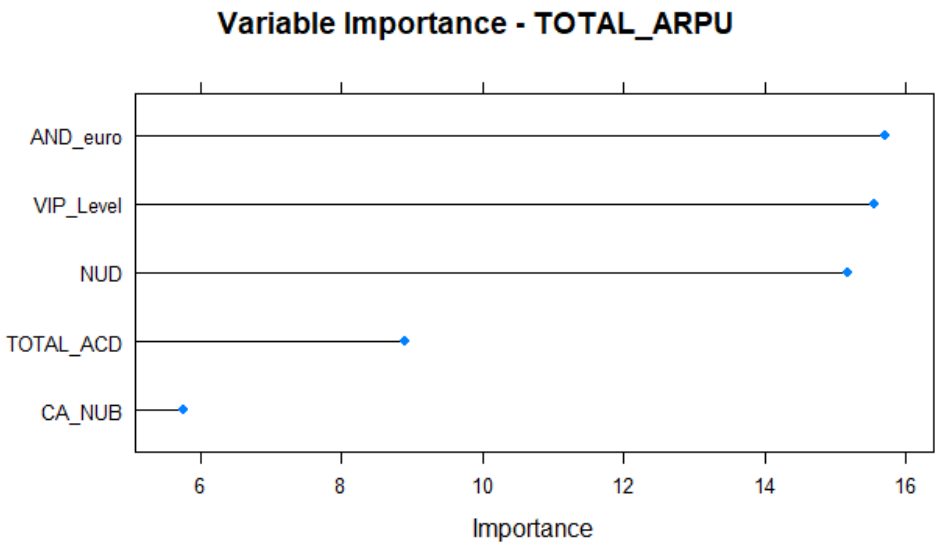


Figure 4.60: ARPU_TOT importance

The 'mtry' value with th elowest RMSE is 2. As it is observed from the Variable Importance plot, the 2 variables that contribute the most (explain) to the model are AND_euro, and VIP_Level. Also, the variable NUD seems that contributes almost equally to the model. Such variables play an important role in explaining the response variable TOTAL_ARPU.

(B)

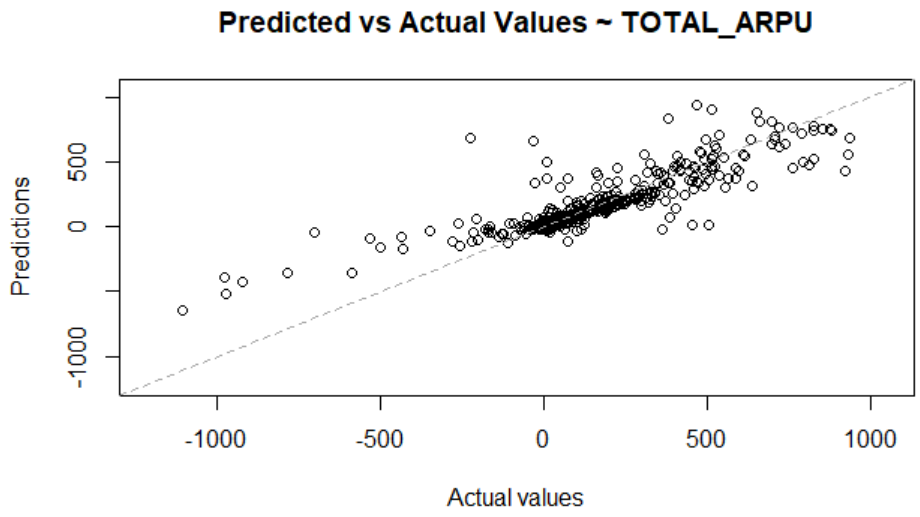


Figure 4.61: Performance on test set – Random Forest

In terms of RMSE and MAE, Random Forest performs best, and, on average, the (in absolute value) prediction error (MAE) on the test set is 74.3694 ~ 75 euro.

4.9 Regressions Analysis Conclusions

The variance of the response variable SB_ARPU is mostly explained by the variables NUD, AND_euro and VIP_Level. On the other hand, the largest amount of variance of the dependent variable CA_ARPU is solely explained by the exploratory variable CA_TUR. This indicates that the company revenue from Casino users is mainly depended on the total amount of money placed by the player in Casino rounds, whereas the company revenue from Sportsbook players rely on the number of times that the customer deposited money in his/her account, the total amount of money deposited and his/her VIP_Level.

The above can be interpreted as follows: In casino games, from the player perspective, there is lack of strategy since "randomness" is more present. Thus, the casino revenue

is mainly explained by the casino rounds (the player bet for each round). Contrarily, regarding sport games, there is more likely to exist a strategy profile for each player. In this study, such profile consists of information relevant to the number of times that the customer deposited money in his/her account, the total amount of money deposited and his/her VIP_Level.

The variable TOTAL_ARPU, the total company revenue, mostly depends on 5 variables. The AND_euro, the VIP_Level, NUD, TOTAL_ACD and NUB. However, the most important are the total amount that the customer deposited in his/her account (AND_euro), the VIP level (VIP_Level) and the number of times that the customer deposited money in his/her account (NUD)

Conclusively, the REC, EXP and SAT variables did not lead to well-fitted models, indicating that it is not possible to accurately predict the financial performance metrics used for forecasting company growth.

5. Conclusions

5.1. Summary of findings and discussion

Various FCMs can be used for measuring customers' experience, such as satisfaction, confirmation/disconfirmation of expectations or loyalty. The NPS metric has been proposed by various researchers as a robust method of measuring loyalty. NPS loyalty is measured as the customers' willingness to promote their experience to others, representing a metric of the word-of-mouth promotion intensity. Moreover, various researchers suggest using the NPS metric for forecasting firm performance or firm growth. According to literature, the NPS metric may prove to be rather accurate for most firm and market setups, yet when it comes to the tertiary sector and especially to the gambling industry, the NPS metric is arguably a good firm performance forecasting means. The purchase decision mechanism in the services sector and especially in the gambling industry is more complex, as the purchase decision sits on more and more vague feelings such as pleasure and hedonism. Literature indicates other metrics as alternatives or supplements to the NPS metric, when it comes to forecasting firm performance in the services sector and in the gambling industry, which does not necessarily cancel out utilizing the NPS metric.

According to the survey results, 96% of the respondents are males, while the 35-54 years old age group represents 54% of the sample. 60.8% of the sample are bronze members, while only 10.3% are silver members.

The sample NPS was measured 59.4%, where 70.4% are promoters, 19% are passives and only 11% are detractors. A cumulative 97.1% of the respondents are generally satisfied, while a cumulative 88.9% of the respondents' expectations are generally fulfilled.

All variables were tested for correlation. The results are presented below in Figure 5.1.

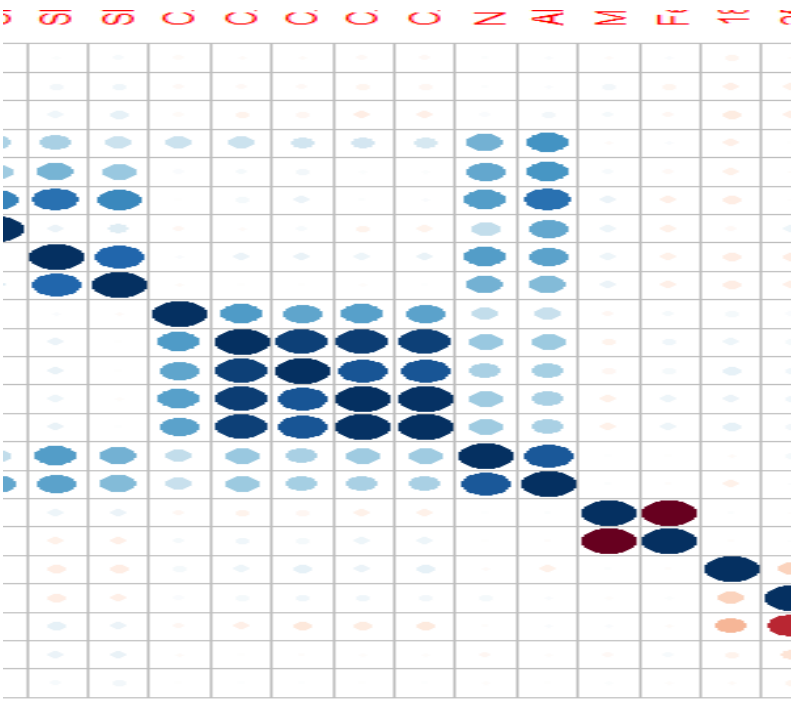


Figure 5.1: Correlations analysis summary

The REC variable was found positively correlated with SAT, EXP and AGE group and negatively correlated with ACD_CA. Moreover, the variables TUR_CA, ABS_CA, NUB_CA, ACD_SB, ACD_CA and ACD (both SB and CA) are correlated with the three REC categories.

The gambling intensity for the casino users is not correlated with the NPS metrics. More specifically, the Turnover, the Average Stake per Bet and the Number of Rounds for casino users in the Detractors group is statistically significantly higher compared to both the Passives and the Promoters groups.

The Active Days for SB in the Detractors group is statistically significantly lower than in both the Passives and the Promoters group. On the contrary, the Active Days for CA in the Detractors group is statistically significantly higher than in both the Passives and the Promoters group. The SB effect is higher, leading to lower ACD_TOTAL in Detractors than in Passives and Promoters.

Regarding the EXP variable, it is positively correlated with TUR_SB, NUB_SB, ACD_SB, ACD_both, EXP, AGE group and VIP Level, negatively correlated with NUB_CA and ACD_CA and correlated with GENDER. Also, the variables ARPU_CA, ABS_CA, NUB_CA, NUB_SB, ACD_SB, ACD_CA and ACD_both have significant differences between the 5

EXP categories so, we can conclude that they are also correlated with EXP. All in all, casino users tend to gamble more rounds and for more days despite not getting their expectations fulfilled. When it comes to sports book, only the case of NUB is correlated with the level of expectations fulfillment.

As far as the SAT categories are concerned, they are positively correlated with ARPU_SB, ARPU_total, TUR_SB, TUR_Total, ASB_SB, NUB_SB, ACD_SB, ACD_both, AND, SAT, AGE group and VIP Level and negatively correlated with TUR_CA, ABS_CA, NUB_CA and ACD_CA and correlated with GENDER. Moreover, the variables ARPU_SB_CA_Total, TUR_SB_CA_Total ABS_SB_CA, NUB_SB_CA, ACD_SB_CA_both have significant differences between the 5 SAT categories so, we can conclude that they are also correlated with SAT. All in all, the more satisfied the users, the higher the total average revenue per user, the total turnover, the active days and the total average stake per bet for sports book and casino users.

Among three metrics, overall satisfaction is better correlated with the financial performance metrics.

As far as predicting firm performance by using any of the three metrics used, the firm generates most revenue from casino users due to the total amount of money gambled in casino rounds, while when it comes to the sports book, the firm generates most revenue due to the number and amount of deposits and due to the users VIP level.

These results indicate that casino users follow a more random gambling strategy, therefore the more the rounds the more the revenue, while when it comes to sports book players, they seem to follow a more strategic gambling strategy.

Among the above mentioned variables, the total amount of deposits explains the biggest part of the total average revenue per user.

As far as predicting firm performance, neither of the REC, EXP and SAT variables lead to well-fitted prediction models, indicating that it is not possible to predict the financial performance using any of these metrics.

Upon combining the survey findings with the literature review findings, it is suggested that the utilized metrics, Net Promoter Score, Overall Satisfaction and Expectations

Fulfillment, seem incapable of predicting company growth, which can be explained by the customers complex feelings driving their purchase decisions on either the sports book or the casino platforms. According to literature, inserting other supplementary metrics, such as customers' pleasure metrics may lead to more valuable results, as pleasure is a significant component of firm performance in the services and in the gambling industry. All in all, there are more complex underlying mechanisms between customer experience and firm performance in the gambling industry. Moreover, studying alternative or supplementary metrics such as the CES metric is indicated as a go option by the literature.

5.2. Limitations and suggestions for further research

As far as limitations are concerned, the current study was conducted in three different stages, as the NPS, SAT and EXP related questions were provided to participants in this order for 20 days, 20 days and 7 days, respectively. This adds up to the survey biasness, as customers may have either experienced either high losses or high profits, which may have altered their original answers.

Moreover, the firm is already applying the NPS metric for forecasting growth, which did not leave space for studying a second supplementary metric such as the CES metric, which according to literature may apply better to the services industry, including the gambling industry.

A latter limitation identified was the complexity of the data analysis, which led to time planning deviations, as the research topic proved to be rather more complex than it was initially estimated. This has led to producing a complex analysis report, which is limited to the essentials despite its volume.

As far as future research is considered, it is suggested studying different FCMs simultaneously or repeating the same study for larger periods of time may, expecting that either choice may have led to different results, as the customer responses may have been biased by either sudden events or by the short periods during they were collected.

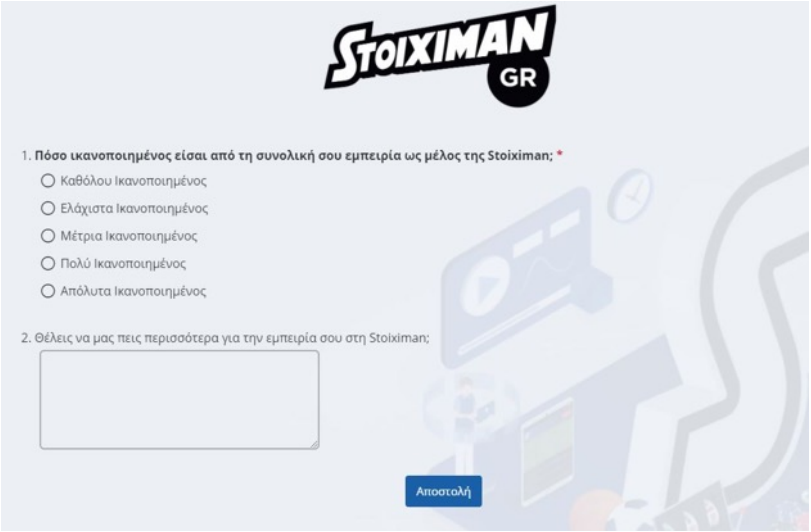
APPENDIX A: SURVEY QUESTIONS

1) **"How likely are you to recommend Stoiximan to a friend?"**

On a Likert scale question from 0-10 (NSP metric)

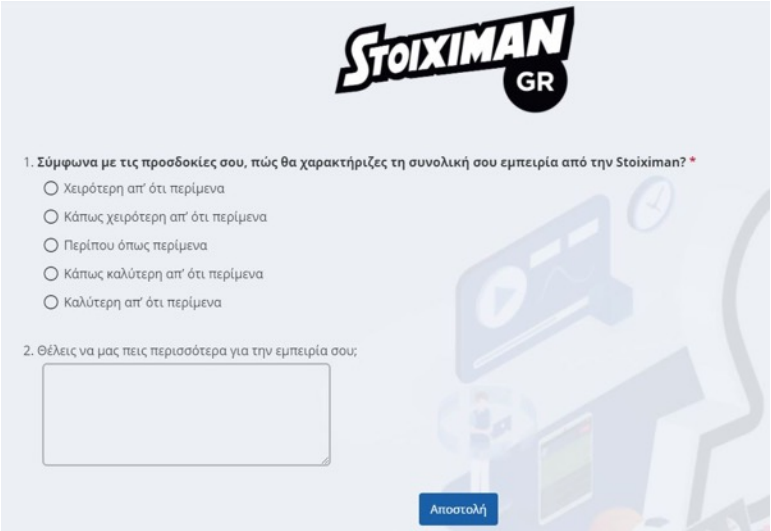
2) "How would you rate your overall satisfaction with Stoiximan?"

5. Very satisfied
4. Satisfied
3. Neither satisfied or dissatisfied
2. Dissatisfied
1.Very dissatisfied



3) **"Compared to your expectations, what is your overall (visiting/ playing) experience from Stoiximan?"**

5. Better than expected
4. Somehow better than expected
3. More or less as expected
2. Somehow worse than expected
1. Worse than expected



<u>Demographics and other</u>

1) Gender
Male
Female
2) Age group
18-24
25-34
25-54
55-64
65+
3) VIP level
Negative
Bronze
Silver
Negative VIP
Gold
Platinum
Diamond

The survey periods were:

- *1st Question **24/10-14/11***
- *2nd Question **19/12-07/01***
- *3rd Question **21/01-27/01***

The depended variables data were collected in February (1-29/2/2020).

APPENDIX B: DATA DICTIONARY

#	Variable	Abbreviations	Description
1	Gender	Gender	Gender
2	Age Group	Age	Age Group
3	NPS Score: How likely are you to recommend Stoiximan to a friend?	REC / REC_cat	The Net Promoter Score is an index that measures the willingness of customers to recommend a company's products or services to others
4	Compared to your expectations, what is your overall (visiting/ playing) experience from Stoiximan?	EXP	
5	How would you rate your overall satisfaction with Stoiximan?	SAT	
6	Customer_Created		Date that customer created the account in our site
7	Customer_Monthly_VIPLevel_Name	VIP_level	This variable segments our customers based on their value
8	SB_Revenue	ARPU_SB	Sportsbook Gross Gaming Revenue
9	SB_Turnover	TUR_SB	Total amount of money placed by the player in Sportsbook bets
10	SB_Avg_Stake_per_Bet	ASB_SB	The amount that a customer has placed on average on every bet
11	SB_Number_of_Bets_Placed	NUB_SB	Number of bets played by a customer
12	SB_Active_Days	ACD_SB	Number of days within the month that a customer has placed at least one bet
13	CA_Revenue	ARPU_CA	Casino Gross Gaming Revenue
14	CA_Turnover	TUR_CA	Total amount of money placed by the player in Casino rounds
15	CA_Avg_Stake_per_Round	ARPU_CA	The amount that a customer has placed on average on every casino round
16	CA_Rounds	ASB_CA	Number of casino rounds played by a customer
17	CA_Active_Days	ACD_CA	Number of days within the month that a customer has played at least one casino round

#	Variable	Abbreviations	Description
18	Active_Days_Both_SB_and_CA	ACD_Both_SB_CA	Number of days within the month that a customer has at least one casino round or/and a placed bet
19	Total Revenue	ARPU_TOTAL	The total sum of the sportsbook & casino gross gaming revenue per player
20	Total Turnover	TUR_TOTAL	The sum of the total amount placed by the players in sportsbook bets and in Casino rounds.
21	Number_of_Deposits	NUB	Number of times that the customer deposited money in his account
22	Deposits_Amount_euro	AND	Total amount that the customer deposited in his account

APPENDIX C: DATA TABLES

Categorical Variables

Gender

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Male	2586	96.1	96.1	96.1
	Female	104	3.9	3.9	100.0
	Total	2690	100.0	100.0	

Age Group

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	18-24	250	9.3	9.3	9.3
	25-34	915	34.0	34.0	43.3
	35-54	1408	52.3	52.3	95.7
	55-64	103	3.8	3.8	99.5
	65+	14	.5	.5	100.0
	Total	2690	100.0	100.0	

Customer's VIP Level

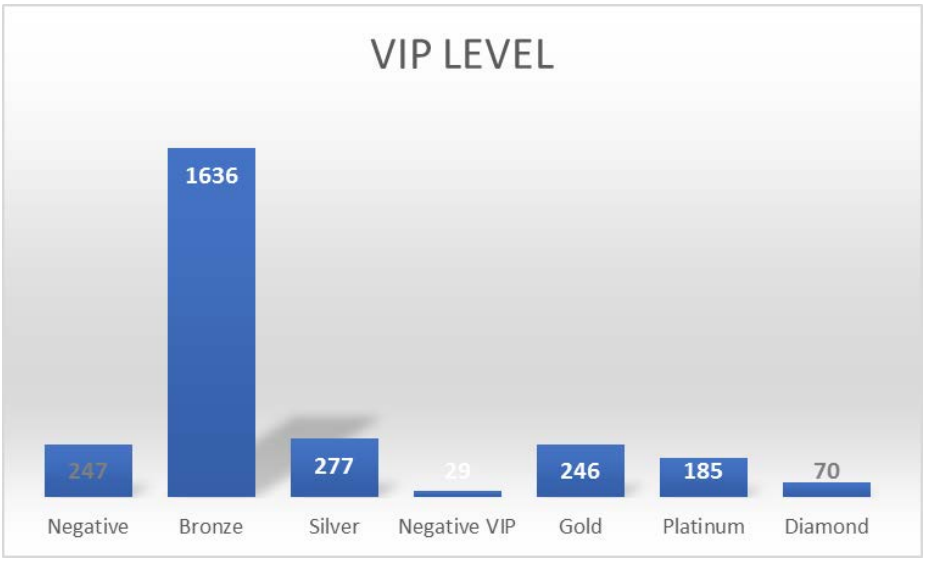
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Negative	247	9.2	9.2	9.2
	Bronze	1636	60.8	60.8	70.0
	Silver	277	10.3	10.3	80.3
	Negative VIP	29	1.1	1.1	81.4
	Gold	246	9.1	9.1	90.5
	Platinum	185	6.9	6.9	97.4
	Diamond	70	2.6	2.6	100.0
	Total	2690	100.0	100.0	

Categories of REC

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Detractors	295	11.0	11.0	11.0
	Passives	501	18.6	18.6	29.6
	Promoters	1894	70.4	70.4	100.0
	Total	2690	100.0	100.0	

Overall Satisfaction

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Not at all satisfied	77	2.9	2.9	2.9
	Slightly Satisfied	157	5.8	5.8	8.7
	Moderately satisfied	687	25.5	25.5	34.2
	Very satisfied	1203	44.7	44.7	79.0
	Totally satisfied	566	21.0	21.0	100.0
	Total	2690	100.0	100.0	



Expectations Confirmation/Disconfirmation

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	Worse than expected	87	3.2	3.2	3.2
	Slightly worse than expected	212	7.9	7.9	11.1
	About as expected	869	32.3	32.3	43.4
	Slightly better than expected	671	24.9	24.9	68.4
	Better than expected	851	31.6	31.6	100.0
	Total	2690	100.0	100.0	

NPS Score (REC)

Descriptive Statistics

	N	Range	Minimum	Maximum	Mean	Std. Deviation	Variance
NPS Score	2690	10	0	10	8.73	2.131	4.540
Valid N (listwise)	2690						

One-Sample Kolmogorov-Smirnov Test

		NPS Score
N		2690
Normal Parameters[a,b]	Mean	8.73
	Std. Deviation	2.131
Most Extreme Differences	Absolute	.294
	Positive	.275
	Negative	-.294
Kolmogorov-Smirnov Z		15.254
Asymp. Sig. (2-tailed)		.000

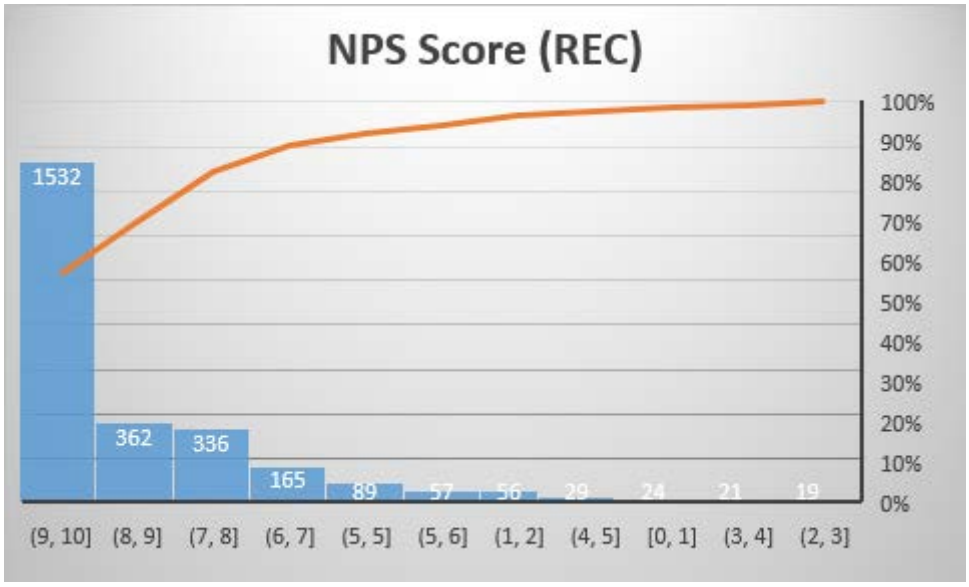
a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Pareto Chart with Cumulative Curve

The Pareto chart plots the distribution of the data in descending order of frequency with a cumulative curve on a secondary axis as a percentage of the total.



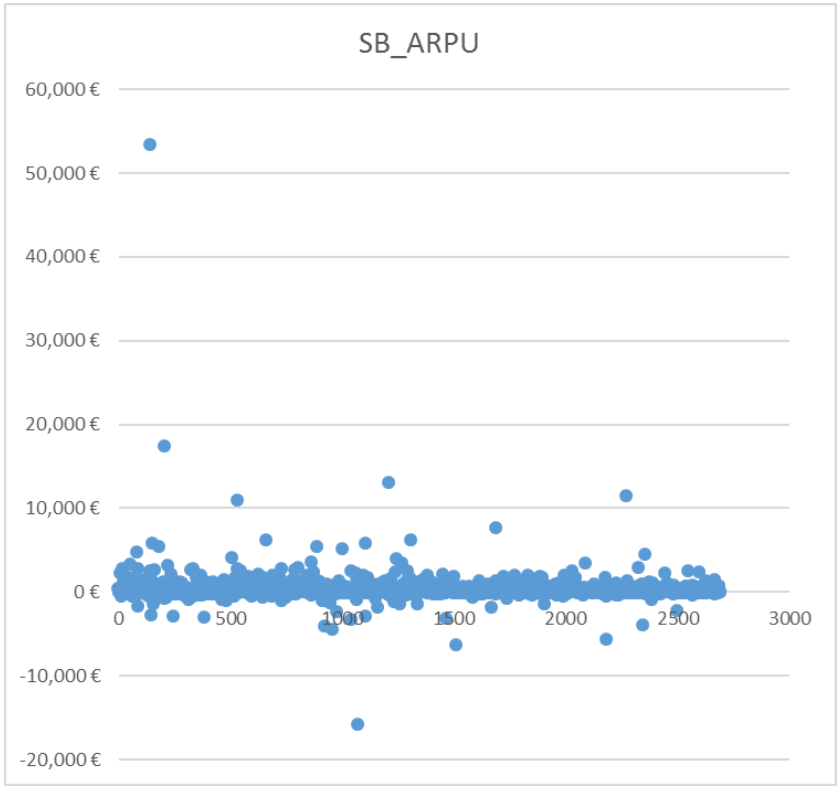
Continuous Variables

Average Revenue Per User (ARPU)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Average revenue per user for SB	2607	-15,728.98	53,423.95	239.36	1,343.88
Average revenue per user for CA	1194	-1,958.91	9,821.20	63.00	395.90
Total Average revenue per user	2640	-11,883.98	53,423.95	302.36	1,386.53
Valid N (listwise)	2690				

Dot plot (sports bet)



One-Sample Kolmogorov-Smirnov Test

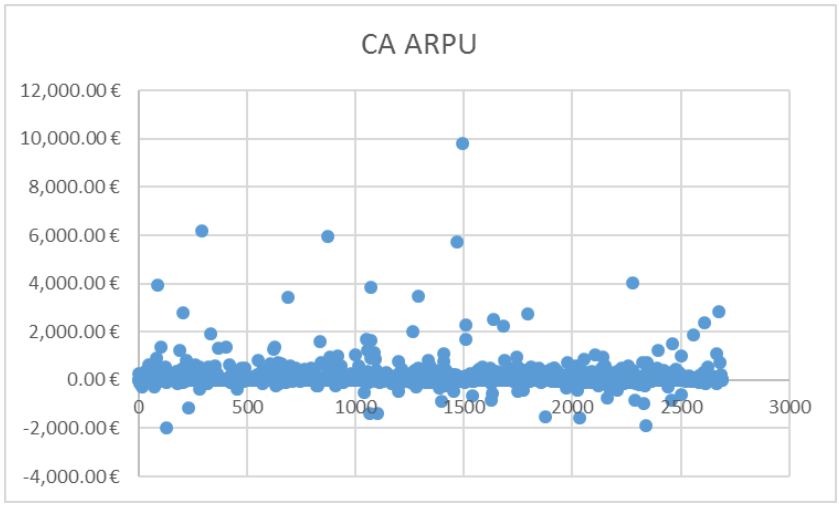
		Average revenue per user for SB
N		2690
Normal Parameters[a,b]	Mean	239.3629
	Std. Deviation	1343.88168
Most Extreme Differences	Absolute	.341
	Positive	.298
	Negative	-.341
Kolmogorov-Smirnov Z		17.710
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Dot plot (casino bet)



One-Sample Kolmogorov-Smirnov Test

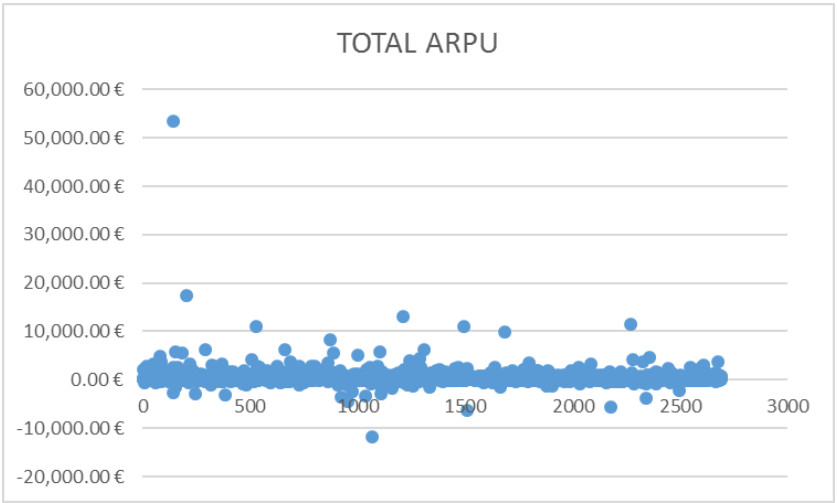
		Average revenue per user for CA
N		2690
Normal Parameters[a,b]	Mean	63.0023
	Std. Deviation	395.89757
Most Extreme Differences	Absolute	.365
	Positive	.349
	Negative	-.365
Kolmogorov-Smirnov Z		18.930
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one..

Dot Plot (total)



One-Sample Kolmogorov-Smirnov Test

		Total Average revenue per user
N		2690
Normal Parameters[a,b]	Mean	302.3652
	Std. Deviation	1386.52819
Most Extreme Differences	Absolute	.323
	Positive	.287
	Negative	-.323
Kolmogorov-Smirnov Z		16.732
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Turnover (TUR)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Turnover for SB	2607	.00	26,1756.7	2,072.84	9,439.14
Turnover for CA	1194	.00	18,8375.5	1,489.39	8,288.61
Total Turnover	2640	.00	261,756.7	3,562.24	12,965.26
Valid N (listwise)	2690				

One-Sample Kolmogorov-Smirnov Test

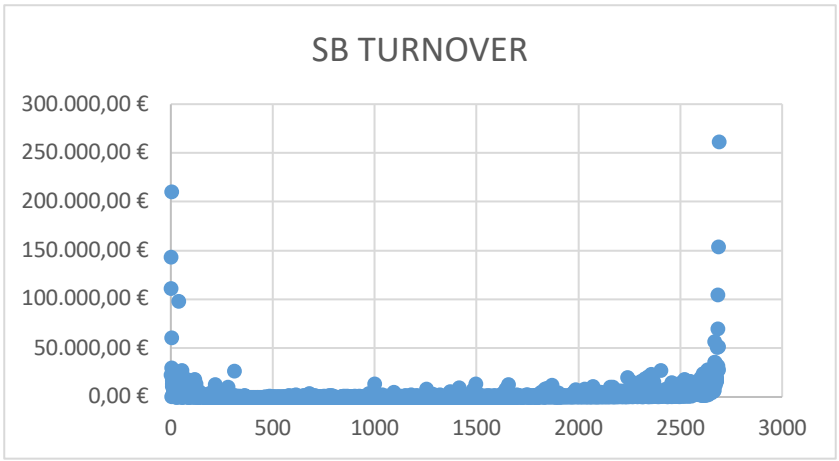
		Turnover for SB	Turnover for CA	Total Turnover
N		2690	2690	2690
Normal Parameters[a,b]	Mean	2072.8419	1489.3946	3562.2365
	Std. Deviation	9439.14159	8288.60935	12965.26500
Most Extreme Differences	Absolute	.413	.429	.392
	Positive	.344	.366	.315
	Negative	-.413	-.429	-.392
Kolmogorov-Smirnov Z		21.425	22.234	20.318
Asymp. Sig. (2-tailed)		.000	.000	.000

a. Test distribution is Normal.

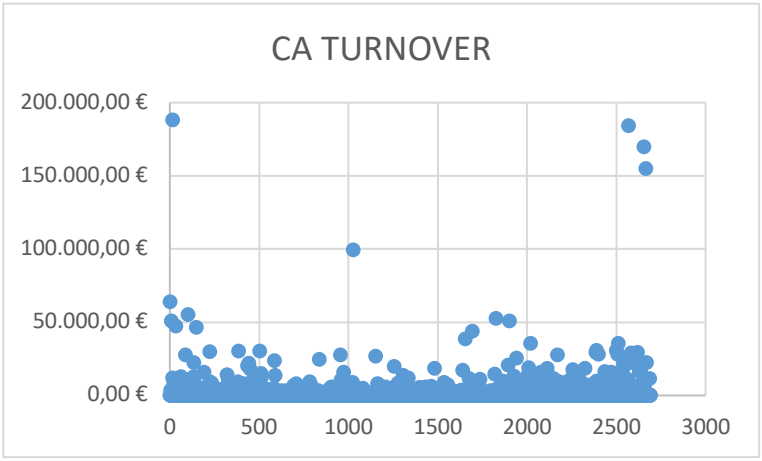
b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

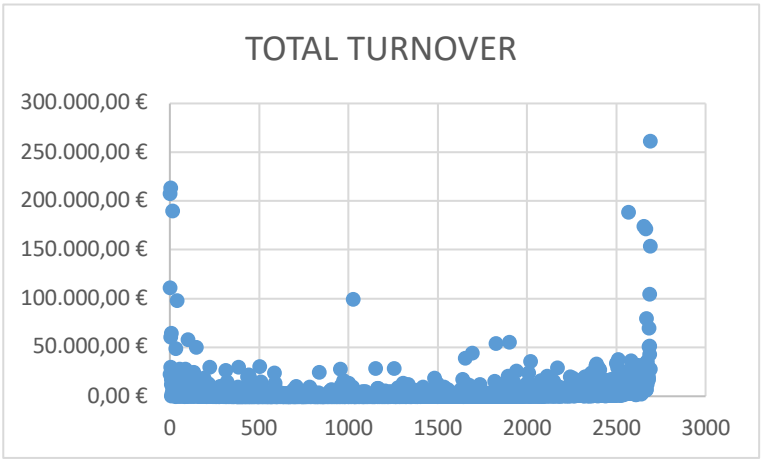
Dot Plot (sports bet)



Dot Plot (casino bets)



Dot Plot (total)



Average stake per bet (ASB)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Average stake per bet_SB	2603	.10	15,962.56	21.99	337.56
Average stake per bet_CA	1194	.00	1,034.27	2.74	31.4
Valid N (listwise)	2690				

One-Sample Kolmogorov-Smirnov Test

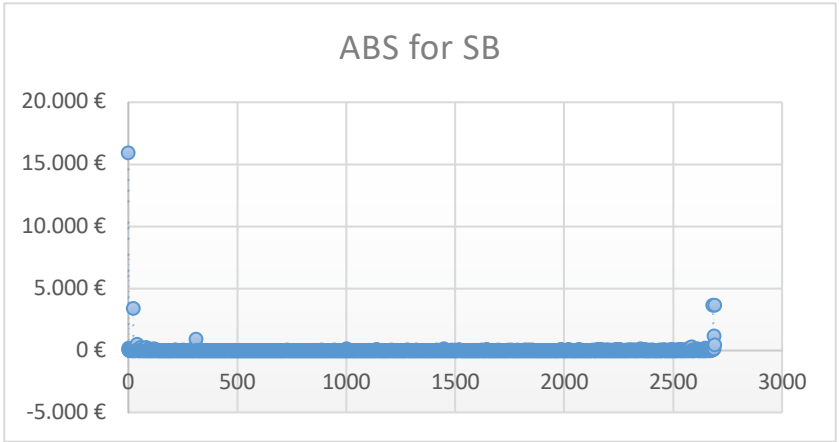
		Average stake per bet_SB	Average stake per bet_CA
N		2690	2690
Normal Parameters[a,b]	Mean	21.2790	1.2179
	Std. Deviation	332.07622	20.95924
Most Extreme Differences	Absolute	.474	.477
	Positive	.426	.427
	Negative	-.474	-.477
Kolmogorov-Smirnov Z		24.608	24.731
Asymp. Sig. (2-tailed)		.000	.000

a. Test distribution is Normal.

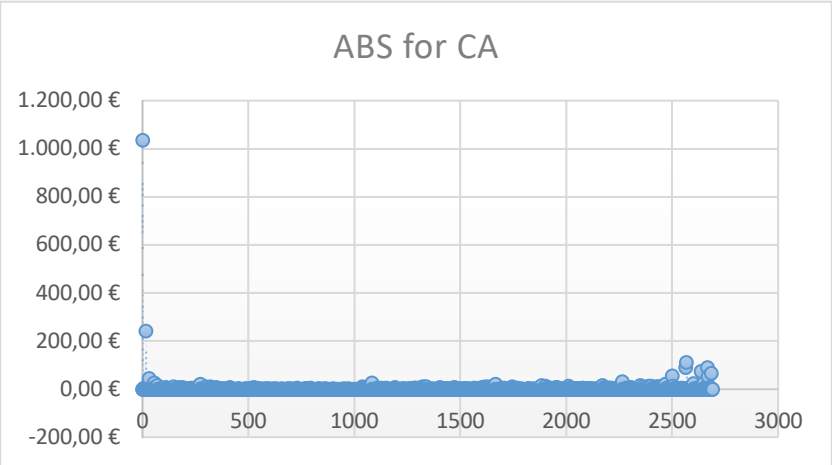
b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Dot Plot (sports bet -> smooth lines observed)



Dot Plot (casino bet -> smooth lines observed)



Number of Bets (NUB)

Descriptive Statistics

	N	Minimum	Maximum	Median	Mean	Std. Deviation
Number of bets for SB	2690	0	4,409	90	182.50	302.412
Number of Rounds for CA	2690	0	109,583	4	1,974.47	6,502.764
Valid N (listwise)	2690					

One-Sample Kolmogorov-Smirnov Test

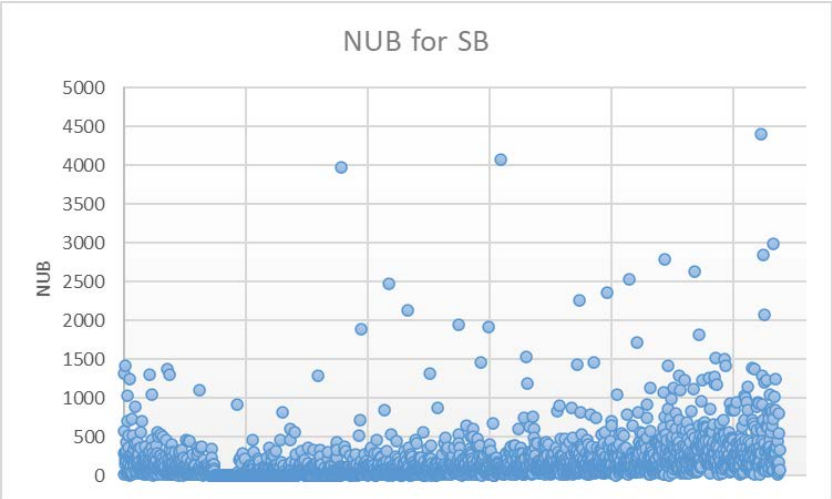
		Number of bets for SB	Number of Rounds for CA
N		2690	2690
Normal Parameters[a,b]	Mean	182.50	1974.47
	Std. Deviation	302.412	6502.764
Most Extreme Differences	Absolute	.273	.381
	Positive	.209	.355
	Negative	-.273	-.381
Kolmogorov-Smirnov Z		14.164	19.745
Asymp. Sig. (2-tailed)		.000	.000

a. Test distribution is Normal.

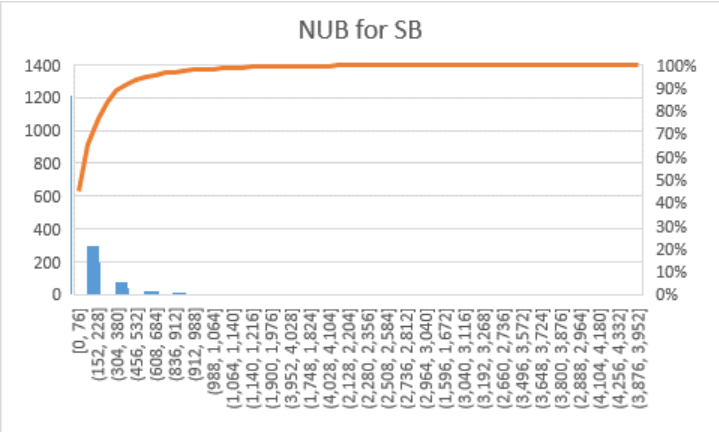
b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal

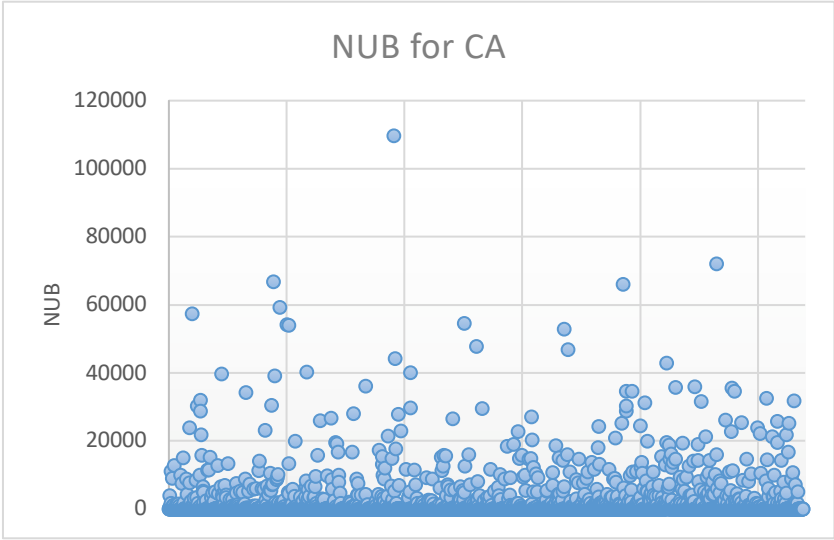
Dot Plot (sports bet -> smooth lines observed)



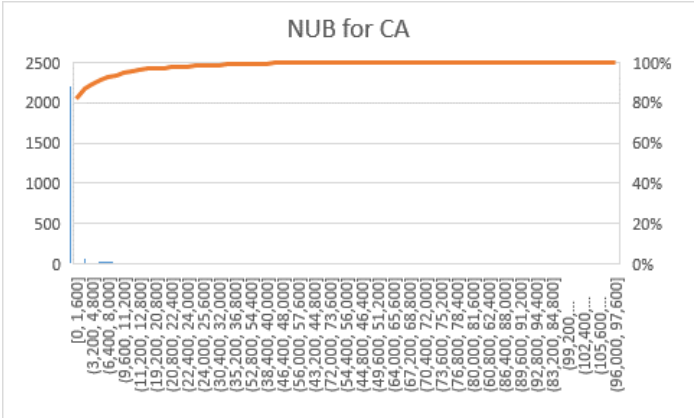
Pareto Chart with Cumulative Curve (sport bet)



Dot Plot (casino bet -> smooth lines observed)



Pareto Chart with Cumulative Curve (casino bet)



Active Days (ACD)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Active Days for SB	2603	1	29	18	8.618
Active Days for CA	1386	1	29	8.2	8.038
Active Days for both SB and CA	2639	1	29	18.53	8.432
Valid N (listwise)	2690				

One-Sample Kolmogorov-Smirnov Test

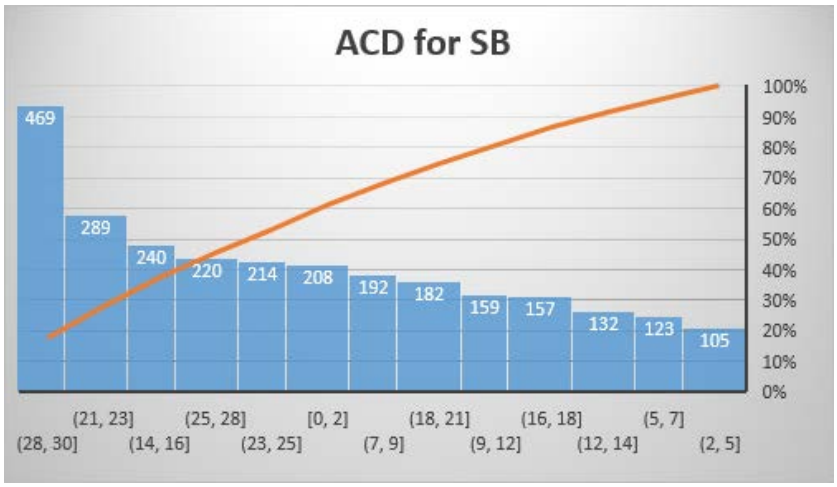
		Active Days for SB	Active Days for CA	Active Days for both SB and CA
N		2690	2690	2690
Normal Parameters[a,b]	Mean	17.41	4.22	18.17
	Std. Deviation	9.056	7.076	8.726
Most Extreme Differences	Absolute	.102	.295	.110
	Positive	.100	.295	.107
	Negative	-.102	-.275	-.110
Kolmogorov-Smirnov Z		5.298	15.321	5.681
Asymp. Sig. (2-tailed)		.000	.000	.000

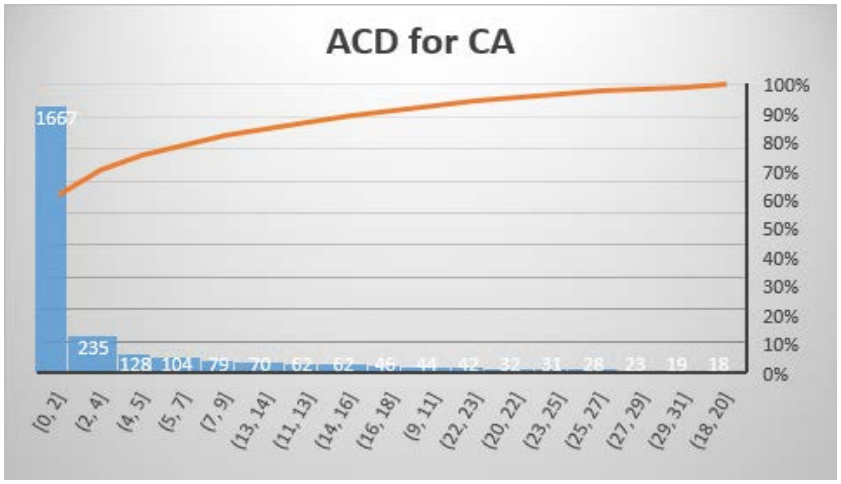
a. Test distribution is Normal.

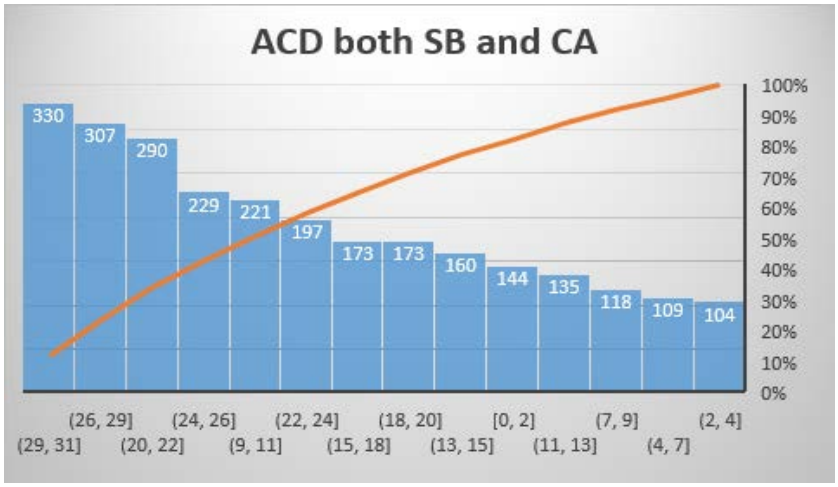
b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Pareto Charts with Cumulative Curve







Number of Deposits (NUD)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Number of deposits	2690	0	573	25.35	36.836
Valid N (listwise)	2690				

One-Sample Kolmogorov-Smirnov Test

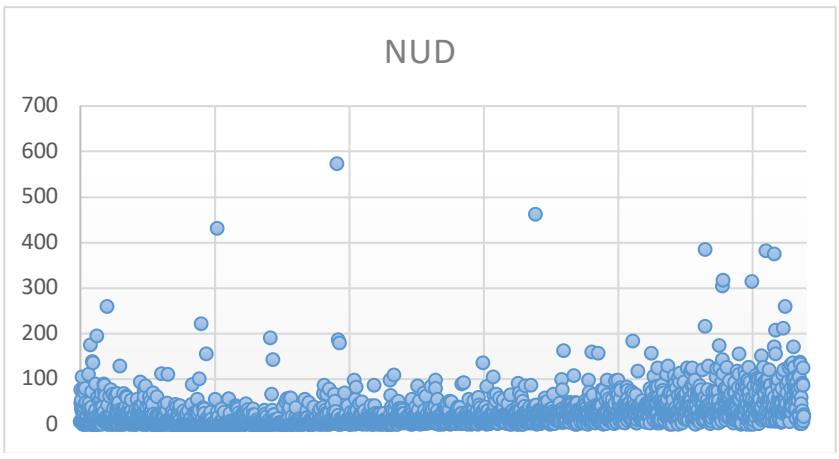
		Number of deposits
N		2690
Normal Parameters[a,b]	Mean	25.35
	Std. Deviation	36.836
Most Extreme Differences	Absolute	.246
	Positive	.183
	Negative	-.246
Kolmogorov-Smirnov Z		12.740
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Dot Plot (-> smooth lines observed)



Amount of Deposits_euro (AND)

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
Amount of deposits	2690	.00	45724.00	649.8687	1765.67642
Valid N (listwise)	2690				

One-Sample Kolmogorov-Smirnov Test

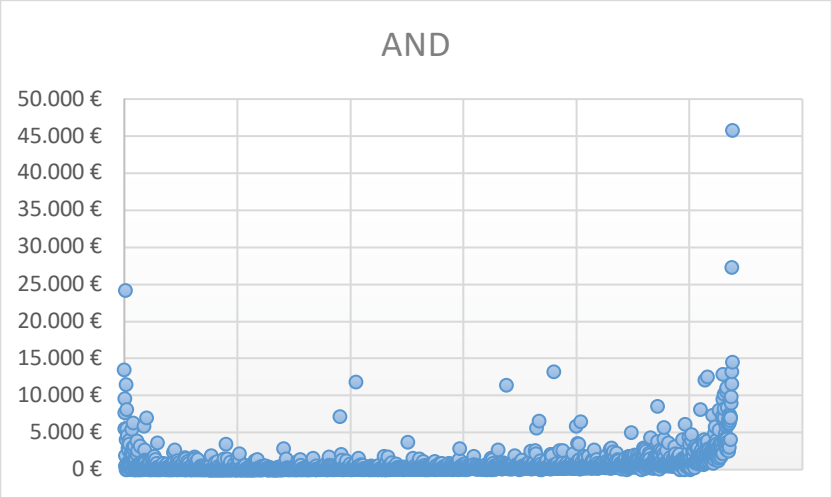
		Amount of deposits
N		2690
Normal Parameters[a,b]	Mean	649.8687
	Std. Deviation	1765.67642
Most Extreme Differences	Absolute	.356
	Positive	.288
	Negative	-.356
Kolmogorov-Smirnov Z		18.486
Asymp. Sig. (2-tailed)		.000

a. Test distribution is Normal.

b. Calculated from data.

The assumption for normal distribution of the data has been tested by a Kolmogorov-Smirnov Z test and the result is that the null hypothesis is rejected (p-value=0.000) and the data distribution is not the normal one.

Dot Plot (-> smooth lines observed)



APPENDIX D: PROFILE VARIABLES

Turnover[€]	**Revenue[€]**
1) 0-100	1) -15.000 -100 euro
2) 101-500	2) 100 - 500
3) 501 - 1000	3) 501 - 1000
4) 1001 - 5.000	4) 1001 - 5.000
5) >5.000	5) >5.000
Average stake per bet/round	**Number of bets (sports/casino)**
0 -10 euro	0 -20
11 - 20	21 - 50
21 - 30	51 - 100
31 - 50	101 - 200
>51	> 201
Active Days	**Number of deposits**
0-3	0-7 --> 906
4-7	8-12 --> 354
8-12	13-20 --> 376
13-20	21-50 --> 682
>21	> 51 --> 371
Deposits Amount[€]	
0-50 --> 653	
51-200 --> 756	
201-500 --> 563	
501-1000 --> 328	
>1001 --> 391	

APPENDIX E: MANN-WHITNEY TESTS

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the TUR_CA combination. Initially the Mann-Whitney test confirms there is at least one significant correlation between the possible REC categories and TUR_CA combinations, as the p-value=0.00<<<0.05 in the following table;

Table E.1: Mann-Whitney U test statistics TUR_CA and REC

	Turnover for CA
Mann-Whitney U	63,803.5
Wilcoxon W	189,554.5
Z	-3.506
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Categories of REC

Table E.2: Mann-Whitney U test ranks TUR_CA and Detractors – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Turnover for CA	Detractors	295	1,191.14	351386.00
	Promoters	1894	1,080.03	2045569.00
	Total	2189		

	Turnover for CA
Mann-Whitney U	251,004.0
Wilcoxon W	2,045,569.0
Z	-3.074
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.02<0.05, therefore the Detractors TUR_CA median is statistically significantly different (higher) than the Promoters median. The issue of consumer behavior in the gambling industry and the incapability of the NPS metric to predict firm growth has been raised in the literature review section. More specifically,

the case of users getting a hedonism sensation despite losing was mentioned. The lowest turnover by promoters than detractors can be explained by users going “tilt” and continuing gambling despite losing.

Table E.3: Mann-Whitney U test ranks TUR_CA and Passives – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Turnover for CA	Passives	501	1167.21	584774.50
	Promoters	1894	1206.14	2284435.50
	Total	2395		

	Turnover for CA
Mann-Whitney U	459023.5
Wilcoxon W	584774.5
Z	-1.241
Asymp. Sig. (2-tailed)	.215

a. Grouping Variable: Categories of REC

The H0 cannot rejected as p-value=0.215>0.05, therefore the TUR_CA promoters median is statistically significantly equal to passives median.

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the ASB_CA combination.

Table E.4: Mann-Whitney U test ranks ASB_CA and REC and Detractors – Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Detractors	295	433.81	127972.50
	Passives	501	377.71	189233.50
	Total	796		

	Average stake per bet_CA
Mann-Whitney U	63482.500
Wilcoxon W	189233.500
Z	-3.617
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.00<<<0.05, therefore the Detractors ASB_CA median is statistically significantly different (higher) than the Passives median. This adds up to the issue raised in the previous set of variables correlation, regarding detractors' higher turnover that was justified by users gambling more heavily on the casino despite being moderately satisfied.

Table E.5: Mann-Whitney U test ranks ASB_CA and Detractors – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Detractors	295	1191.84	351593.50
	Promoters	1894	1079.92	2045361.50
	Total	2189		

	Average stake per bet_CA
Mann-Whitney U	250796.500
Wilcoxon W	2045361.500
Z	-3.096
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.02<0.05, therefore the Detractors ASB_CA median is statistically significantly different (higher) than the Promoters median. This adds up to the issue raised in the previous set of variables correlation, regarding detractors' higher turnover that was justified by users gambling more heavily on the casino despite being dissatisfied, in other words most likely losing money.

Table E.6: Mann-Whitney U test ranks ASB_CA and Passives – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Passives	501	1163.97	583146.50
	Promoters	1894	1207.00	2286063.50
	Total	2395		

	Average stake per bet_CA
Mann-Whitney U	457395.500
Wilcoxon W	583146.500
Z	-1.372
Asymp. Sig. (2-tailed)	.170

a. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.17>0.05, therefore the Passives ASB_CA median is statistically significantly equal with the Promoters median.

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the NUB_CA combination.

Table E.7: Mann-Whitney U test ranks NUB_CA and REC and Detractors – Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Detractors	295	434.63	128216.00
	Passives	501	377.23	188990.00
	Total	796		

	Number of Rounds for CA[4]
Mann-Whitney U	63239.000
Wilcoxon W	188990.000
Z	-3.581
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.00<<<0.05, therefore the Detractors NUB_CA median is statistically significantly different (higher) than the Passives median. This adds up to the issues raised in the previous set of variables correlation, regarding detractors'

[4] *The terminology for casino is rounds instead of bets*

higher number of rounds and turnover that was justified by users gambling more heavily on the casino despite being moderately satisfied.

Table E.8: Mann-Whitney U test ranks ASB_CA and Detractors – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Detractors	295	1207.10	356093.50
	Promoters	1894	1077.54	2040861.50
	Total	2189		

	Number of Rounds for CA[5]
Mann-Whitney U	246296.500
Wilcoxon W	2040861.500
Z	-3.471
Asymp. Sig. (2-tailed)	.001

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the Detractors NUB_CA median is statistically significantly different (higher) than the Promoters median. This adds up to the issues raised in the previous set of variables correlation, regarding detractors' higher number of rounds and turnover that was justified by users gambling more heavily on the casino despite being dissatisfied, in other words most likely losing money.

[5] *The terminology for casino is rounds instead of bets*

Table E.9: Mann-Whitney U test ranks ASB_CA and Passives – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Passives	501	1176.31	589331.00
	Promoters	1894	1203.74	2279879.00
	Total	2395		

	Number of Rounds for CA
Mann-Whitney U	463580.000
Wilcoxon W	589331.000
Z	-.843
Asymp. Sig. (2-tailed)	.399

a. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.399>0.05, therefore the Passives NUB_CA median is statistically significantly equal with the Promoters median.

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the ACD_SB combination.

Table E.10: Mann-Whitney U test ranks ACD_SB and REC and Detractors – Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for SB	Detractors	295	359.84	106152.00
	Passives	501	421.27	211054.00
	Total	796		

	Active Days for SB
Mann-Whitney U	62492.000
Wilcoxon W	106152.000
Z	-3.644
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.00<<<0.05, therefore the Detractors ACD_SB median is statistically significantly different (higher) than the Passives median. This adds up to the issue raised in the previous set of variables correlation, regarding detractors' higher active days that was justified by users gambling more heavily on the casino despite being moderately satisfied.

Table E.11: Mann-Whitney U test ranks ACD_SB and Detractors – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for SB	Detractors	295	974.15	287374.00
	Promoters	1894	1113.82	2109581.00
	Total	2189		

Table E.12: Mann-Whitney U test statistics ACD_CA and Detractors – Promoters combination

	Active Days for SB
Mann-Whitney U	243714.000
Wilcoxon W	287374.000
Z	-3.535
Asymp. Sig. (2-tailed)	.000

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.000<<<0.05, therefore the Detractors ACD_SB median is statistically significantly different (higher) than the Promoters median. This adds up to the issue raised in the previous set of variables correlation, regarding detractors' higher active days that was justified by users gambling more heavily on the casino despite being dissatisfied, in other words most likely losing money.

Table E.13: Mann-Whitney U test ranks ACD_SB and Passives – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for SB	Passives	501	1222.42	612434.00
	Promoters	1894	1191.54	2256776.00
	Total	2395		

Table E.14: Mann-Whitney U test statistics ACD_CA and Passives – Promoters combination

	Active Days for SB
Mann-Whitney U	462211.000
Wilcoxon W	2256776.000
Z	-.890
Asymp. Sig. (2-tailed)	.373

a. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.373>0.05, therefore the Passives ACD_SB median is statistically significantly equal with the Promoters median.

Table E.15: Mann-Whitney U test ranks ACD_SB Detractors, Passives and Promoters

	Categories of REC	N	Mean Rank
Active Days for CA	Detractors	295	1488.45
	Passives	501	1314.62
	Promoters	1894	1331.40
	Total	2690	

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the ACD_CA combination.

Table E.16: Mann-Whitney U test ranks ACD_SB Detractors and Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for CA	Detractors	295	430.73	127066.00
	Passives	501	379.52	190140.00
	Total	796		

Table E.17: Mann-Whitney test statistics ACD_CA and REC

	Active Days for CA
Mann-Whitney U	64389.000
Wilcoxon W	190140.000
Z	-3.200
Asymp. Sig. (2-tailed)	.001

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the Detractors and Passives REC categories and ACD_CA medians are not equal. A statistically significant correlation between the Detractors and Passives REC categories and the ACD_CA is detected.

Table E.18: Mann-Whitney U test ranks ACD_CA Detractors and Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for CA	Detractors	295	1205.72	355686.50
	Promoters	1894	1077.76	2041268.50
	Total	2189		

Table E.19: Mann-Whitney test statistics ACD_CA and Detractors and Promoters combination

	Active Days for CA
Mann-Whitney U	246703.500
Wilcoxon W	2041268.500
Z	-3.433
Asymp. Sig. (2-tailed)	.001

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.01<0.05, therefore the Detractors and Promoters REC categories and ACD_CA medians are not equal. A statistically significant correlation between the Detractors and Promoters REC categories and the ACD_CA is detected.

Table E.20: Mann-Whitney U test ranks ACD_SB Passives and Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for CA	Passives	501	1186.10	594236.00
	Promoters	1894	1201.15	2274974.00
	Total	2395		

Table E.21: Mann-Whitney test statistics ACD_CA and Passives and Promoters combination

Test Statistics[a]

	Active Days for CA
Mann-Whitney U	468485.000
Wilcoxon W	594236.000
Z	-.463
Asymp. Sig. (2-tailed)	.643

a. Grouping Variable: Categories of REC

The H0 cannot be rejected as p-value=0.643>0.05, therefore the Passives and Promoters REC categories and ACD_SB medians are statistically significantly equal.

The Mann-Whitney test was deployed to identify the statistically significant difference in the REC categories medians with the ACD_SB and ACD_CA combination.

Table E.22: Mann-Whitney U test ranks ACD_SB and REC and Detractors – Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for both SB and CA	Detractors	295	365.83	107921.00
	Passives	501	417.73	209285.00
	Total	796		

Table E.23: Mann-Whitney U test statistics ACD_SB and REC and Detractors – Passives combination

	Active Days for both SB and CA
Mann-Whitney U	64261.000
Wilcoxon W	107921.000
Z	-3.080
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.002<<<0.05, therefore the Detractors ACD_SB and ACD_CA median is statistically significantly different (higher) than the Passives median.

Table E.24: Mann-Whitney U test ranks ACD_SB and Detractors – Promoters combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for both SB and CA	Detractors	295	988.59	291633.00
	Promoters	1894	1111.57	2105322.00
	Total	2189		

Table E.25: Mann-Whitney U test statistics ASB_CA and Detractors – Promoters combination

	Active Days for both SB and CA
Mann-Whitney U	247973.000
Wilcoxon W	291633.000
Z	-3.113
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.002<<<0.05, therefore the Detractors ACD_SB and ACD_CA median is statistically significantly different (higher) than the Promoters median.

Table E.26: Mann-Whitney U test ranks ACD_SB and Detractors and Passives combination

	Categories of REC	N	Mean Rank	Sum of Ranks
Active Days for both SB and CA	Detractors	295	365.83	107921.00
	Passives	501	417.73	209285.00
	Total	796		

Table E.27: Mann-Whitney U test ranks ACD_SB Detractors and Promoters combination

	Active Days for both SB and CA
Mann-Whitney U	64261.000
Wilcoxon W	107921.000
Z	-3.080
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Categories of REC

The H0 is rejected as p-value=0.02<0.05, therefore the Detractors and Passives REC categories and ACD_SB and ACD_CA medians are not equal. A statistically significant correlation between the Detractors and Passives REC categories and the ACD_SB and ACD_CA is detected.

Upon identifying that there is a statistically significant difference among the five scaled expectations medians, the Mann-Whitney test was deployed to identify the statistically significant difference in specific pairs of ARPU_CA medians.

Table E.28: Mann-Whitney U test ranks ARPU and EXP - Worse than expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average revenue per user for CA	Worse than expected	87	433.04	37674.50
	Slightly better than expected	671	372.56	249986.50
	Total	758		

	Average revenue per user for CA
Mann-Whitney U	24530.500
Wilcoxon W	249986.500
Z	-2.701
Asymp. Sig. (2-tailed)	.007

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.007<0.05, therefore the ARPU_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.29: Mann-Whitney U test ranks ARPU and EXP - Slightly worse than expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average revenue per user for CA	Slightly worse than expected	212	473.92	100471.50
	Slightly better than expected	671	431.91	289814.50
	Total	883		

	Average revenue per user for CA
Mann-Whitney U	64358.500
Wilcoxon W	289814.500
Z	-2.302
Asymp. Sig. (2-tailed)	.021

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.021<0.05, therefore the ARPU_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.30: Mann-Whitney U test ranks ARPU and EXP - About as expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average revenue per user for CA	About as expected	869	795.03	690878.00
	Slightly better than expected	671	738.74	495692.00
	Total	1540		

	Average revenue per user for CA
Mann-Whitney U	270236.000
Wilcoxon W	495692.000
Z	-2.729
Asymp. Sig. (2-tailed)	.006

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.006<0.05, therefore the ARPU_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.31: Mann-Whitney U test ranks ARPU and EXP - Slightly better than expected and Better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average revenue per user for CA	Slightly better than expected	671	735.64	493616.50
	Better than expected	851	781.89	665386.50
	Total	1522		

	Average revenue per user for CA
Mann-Whitney U	268160.500
Wilcoxon W	493616.500
Z	-2.265
Asymp. Sig. (2-tailed)	.024

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.024<0.05, therefore the ARPU_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

The Mann-Whitney test was deployed to identify the statistically significant difference in the ASB_CA medians.

Table E.32: Mann-Whitney U test ranks ASB_CA and EXP - Slightly better than expected and About as expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Slightly worse than expected	212	582.73	123538.50
	About as expected	869	530.82	461282.50
	Total	1081		

Table E.33: Mann-Whitney U test statistics ASB_CA and EXP- Slightly better than expected and About as expected pair

	Average stake per bet_CA
Mann-Whitney U	83267.500
Wilcoxon W	461282.500
Z	-2.359
Asymp. Sig. (2-tailed)	.018

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.018<0.05, therefore the ASB_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.34: Mann-Whitney U test ranks ASB_CA and EXP - Slightly worse than expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Slightly worse than expected	212	479.90	101739.50
	Slightly better than expected	671	430.02	288546.50
	Total	883		

Table E.35: Mann-Whitney U test statistics ASB_CA and EXP - Slightly worse than expected and Slightly better than expected pair

	Average stake per bet_CA
Mann-Whitney U	63090.500
Wilcoxon W	288546.500
Z	-2.730
Asymp. Sig. (2-tailed)	.006

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.006<0.05, therefore the ASB_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.36: Mann-Whitney U test ranks ASB and EXP - Slightly worse than expected and Better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Average stake per bet_CA	Slightly worse than expected	212	573.43	121568.00
	Better than expected	851	521.68	443948.00
	Total	1063		

	Average stake per bet_CA
Mann-Whitney U	81422.000
Wilcoxon W	443948.000
Z	-2.395
Asymp. Sig. (2-tailed)	.017

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.017<0.05, therefore the ASB_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

The Mann-Whitney test was deployed to identify the statistically significant difference in the NUB medians.

Table E.37: Mann-Whitney U test ranks NUB_SB and EXP - Slightly better than expected and About as expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of bets for SB	Slightly worse than expected	212	502.25	106476.50
	About as expected	869	550.45	478344.50
	Total	1081		

	Number of bets for SB
Mann-Whitney U	83898.500
Wilcoxon W	106476.500
Z	-2.016
Asymp. Sig. (2-tailed)	.044

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.044<0.05, therefore the NUB_SB median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.38: Mann-Whitney U test ranks NUB_SB and EXP - Slightly worse than expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of bets for SB	Slightly worse than expected	212	399.85	84768.50
	Slightly better than expected	671	455.32	305517.50
	Total	883		

	Number of bets for SB
Mann-Whitney U	62190.500
Wilcoxon W	84768.500
Z	-2.760
Asymp. Sig. (2-tailed)	.006

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.006<0.05, therefore the NUB_SB median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.39: Mann-Whitney U test ranks NUB_SB and EXP - Slightly worse than expected and Better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of bets for SB	Slightly worse than expected	212	475.18	100737.50
	Better than expected	851	546.16	464778.50
	Total	1063		

	Number of bets for SB
Mann-Whitney U	78159.500
Wilcoxon W	100737.500
Z	-3.012
Asymp. Sig. (2-tailed)	.003

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.003<0.05, therefore the NUB_SB median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.40: Mann-Whitney U test ranks NUB_CA and EXP - Slightly worse than expected and About as expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Slightly worse than expected	212	586.56	124350.00
	About as expected	869	529.89	460471.00
	Total	1081		

	Number of Rounds for CA
Mann-Whitney U	82456.000
Wilcoxon W	460471.000
Z	-2.492
Asymp. Sig. (2-tailed)	.013

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.013<0.05, therefore the NUB_SB median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.41: Mann-Whitney U test ranks NUB_CA and EXP - Slightly worse than expected and Slightly better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Slightly worse than expected	212	490.43	103972.00
	Slightly better than expected	671	426.70	286314.00
	Total	883		

Table E.42: Mann-Whitney U test statistics NUB_CA and EXP - Slightly worse than expected and Slightly better than expected pair

Test Statistics[a]

	Number of Rounds for CA
Mann-Whitney U	60858.000
Wilcoxon W	286314.000
Z	-3.364
Asymp. Sig. (2-tailed)	.001

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.001<0.05, therefore the NUB_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

Table E.43: Mann-Whitney U test ranks NUB_CA and EXP - Slightly worse than expected and Better than expected pair

	Expectations Confirmation/Disconfirmation	N	Mean Rank	Sum of Ranks
Number of Rounds for CA	Slightly worse than expected	212	586.74	124388.50
	Better than expected	851	518.36	441127.50
	Total	1063		

	Number of Rounds for CA
Mann-Whitney U	78601.500
Wilcoxon W	441127.500
Z	-3.074
Asymp. Sig. (2-tailed)	.002

a. Grouping Variable: Expectations Confirmation/Disconfirmation

The H0 is rejected as p-value=0.002<0.05, therefore the NUB_CA median for users with low EXP is statistically significantly different (higher) than the one for users with higher EXP.

APPENDIX F: TESTS OF NORMALITY

Tests of Normality

Average revenue per user for SB	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.339	2690	.000	.206	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

Tests of Normality

Average revenue per user for CA	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.365	2690	.000	.244	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

Tests of Normality

Total Average revenue per user	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.321	2690	.000	.246	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

Tests of Normality

Turnover for SB	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.407	2690	.000	.173	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

Turnover for CA	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.425	2690	.000	.159	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

Total Turnover	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.385	2690	.000	.239	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used. The Kruskal-Wallis test is about checking the equality of medians and not the equality of means.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.471	2690	.000	.025	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.473	2690	.000	.027	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used. The Kruskal-Wallis test is about checking the equality of medians and not the equality of means.

Tests of Normality

Number of Bets for SB	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.272	2690	.000	.529	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

Number of Rounds for CA	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.378	2690	.000	.340	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

Tests of Normality

Active Days for SB	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.097	2690	.000	.930	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

Active Days for CA	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.290	2690	.000	.668	2690	.000

a. Lilliefors Significance Correction

Tests of Normality

Active Days for both SB and CA	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.104	2690	.000	.928	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected and the Kruskal-Wallis test has been used.

Tests of Normality

Number of deposits	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.240	2690	.000	.602	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

Tests of Normality

Amount of deposits	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Standardized Residual	.349	2690	.000	.335	2690	.000

a. Lilliefors Significance Correction

The null hypothesis is rejected by both tests since p-value = 0.000 <0.05 in all cases. So, the case of normality distributed residuals is rejected, and the Kruskal-Wallis test has been used.

References

Abbott, L., 1955. *Quality and Competition: An Essay in Economic Theory.* 1 ed. New York: Columbia University Press.

Alba, J. W. & Williams, E. F., 2013. Pleasure principles: A review of research on hedonic consumption. *Journal of Consumer Psychology,* 23(1), pp. 2 - 18.

Back, K.-J. & Lee, C.-K., 2015. Determining the Attributes of Casino Customer Satisfaction: Applying Impact-Range Performance and Asymmetry Analyses. *Journal of Travel & Tourism Marketing,* 32(6), pp. 747 - 760.

Balakrishnan, J. & Griffiths, M. D., 2018. Loyalty towards online games, gaming addiction, and purchase intention towards online mobile in-game features. *Computers in Human Behavior,* Volume 87, pp. 238 - 246.

Bilgihan, A., Madanoglu, M. & Ricci, P., 2016. Service attributes as drivers of behavioral loyalty in casinos: The mediating effect of attitudinal loyalty. *Journal of Retailing and Consumer Services,* Volume 31, pp. 14 - 21.

Bolton, R. N., 1999. A Dynamic Model of the Duration of the Customer's Relationship with a Continuous Service Provider: The Role of Satisfaction. *Journal of Marketing Research,* 17(1), pp. 171 - 186.

Brodie, R., Hollebeek, L., Jurić, B. & Ilić, A., 2011. Customer Engagement: Conceptual Domain, Fundamental Propositions, and Implications for Research. *Journal of Service Research,* 14(3), pp. 252 - 271.

Calvin Ayre, 2020. *Greece's online gambling plans get EC nod, applications flood in.* [Online]
Available at: https://calvinayre.com/2020/04/02/business/greece-online-gambling-plans-european-commission-okay/
[Accessed 20 October 2020].

Cardozo, R., 1965. An Experimental Study of Customer Effort, Expectation, and Satisfaction. *Journal of Marketing Research,* 2(3), pp. 244 - 249.

Cronin, J. J., Brady, M. K. & Hult, G. T., 2000. Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing,* 76(2), pp. 193 - 218.

Danaher, P. & Rust, R., 1996. Determining the optimal return on investment for an advertising campaign. *European Journal of Operational Research,* 95(3), pp. 511 - 521.

D'Astous, A. & Gaspero, M. D., 2013. Explaining the performance of online sports bettors. *International Gambling Studies,* pp. 371-387.

de Haan, E., Verhoef, P. C. & Wiesel, T., 2015. The predictive ability of different customer feedback metrics for retention. *International Journal of Research in Marketing ,* Volume 32, pp. 195 - 206.

de Haan, E., Verhoef, P. C. & Wiesel, T., 2015. The predictive ability of different customer feedback metrics for retention. *International Journal of Research in Marketing,* Volume 32, pp. 195-206.

Dixon, M., Freeman, K. & Toman, N., 2010. Stop trying to delight your customers. *Harvard Business Review,* 88(7-8), pp. 116 - 122.

Dixon, M., Freeman, K. & Toman, N., 2010. *Stop Trying to Delight Your Customers,* s.l.: Harvard Business Review.

European Gambling and Betting Association, 2020. *European Gambling and Betting Association.* [Online] Available at: https://www.egba.eu/eu-market/ [Accessed 11 October 2020].

European Gaming and Betting Association, 2020. *Gambling is becoming more and more an online activity.* [Online] Available at: https://www.egba.eu/eu-market/ [Accessed 12 September 2020].

Ezenwafor, E. C., Ayodele, A. A. & Ejiroghene, P. A., 2020. Self Service Technology and Customer Satisfaction in the Nigerian Online Sport Betting Industry. *European Journal of Business and Management ,* 12(8), pp. 1-5.

Farris, P. W., Bendle, N. T., Pfeifer, P. E. & Reibstein, D. J., 2010. *Farris, Paul W.; Neil T. Bendle; Phillip E. Pfeifer; David J. Reibstein (2010). Marketing Metrics: The Definitive Guide to Measuring Marketing Performance.* 1 ed. New Jersey: Pearson Education Inc..

Fleming, J. H. & Asplund, J., 2007. *Human Sigma: Managing the Employee-Customer Encounter.* 1st ed. New York: Gallup Press.

Gao, B. W. & Lai, I. K., 2015. The effects of transaction-specific satisfactions and integrated satisfaction on customer loyalty. *International Journal of Hospitality Management,* Volume 44, pp. 38 - 47.

Global Market Insights, 2020. *Online Gambling Market.* [Online] Available at: https://www.gminsights.com/industry-analysis/online-gambling-market#:~:text=Online%20Gambling%20Market%20size%20surpassed,learning%2C%20will%20drive%20market%20growth.
[Accessed 12 October 2020].

Global Market Insights, 2020. *Online Gambling Market by Type (Betting [Boxing & UFC, eSports, Football, Formula1, Horse Racing, Tennis], Casino [Baccarat, Blackjack, Roulette, Slots], Poker, Lottery, Bingo), Device Type (Desktop, Mobile), Regional Outlook, Price Trends, Competitive Ma,* s.l.: Global Market Insights.

Grigoroudis, E. et al., 2008. The assessment of user-perceived web quality: Application of a satisfaction benchmarking approach. *European Journal of Operational Research,* 187(3), pp. 1346-1357.

Grigoroudis, E. & Siskos, Y., 2004. A Survey of Customer Satisfaction Barometers: Some Results from the Transportation-Communications Sector. *European Journal of Operational Research,* 152(2), pp. 334-353.

Grigoroudis, E. & Siskos, Y., 2010. *Customer Satisfaction Evaluation: Methods for Measuring and Implementing Service Quality.* 1 ed. New York: Springer.

Gupta, S. et al., 2006. Modeling Customer Lifetime Value. *Journal of Service Research,* 9(2), pp. 139 - 155.

Gupta, S. & Zeithaml, V., 2006. Customer metrics and their impact on financial performance. *Marketing Science,* 25(6), pp. 718 - 739.

Hanssens, D. M. et al., 2014. Consumer attitude metrics for guiding marketing mix decisions. *Marketing Science,* 33(4), p. 534 – 550.

Hayes, B., 2017. *Stop Listening to the Net Promoter Score (NPS) Dogma and Follow the Evidence.* [Online]
Available at: https://customerthink.com/stop-listening-to-the-net-promoter-score-nps-dogma-and-follow-the-evidence/
[Accessed 11 October 2020].

Henderson, C. M., Steinhoff, L. & Palmatier, R. W., 2014. Consequences of customer engagement: how customer engagement alters the effects of habit-, dependence-, and relationship-based intrinsic loyalty. *Marketing Science Institute Working Papers Series,* 14(121), pp. 1 - 39.

Investopedia, 2020. *The Math Behind Betting Odds & Gambling.* [Online]
Available at: https://www.investopedia.com/articles/dictionary/042215/understand-math-behind-betting-odds-gambling.asp
[Accessed 15 December 2020].

Io, M.-U., 2016. Exploring the impact of hedonic activities on casino-hotel visitors' positive emotions and satisfaction. *Journal of Hospitality and Tourism Management,* Volume 26, pp. 27 - 35.

Jeon, S. M. & Hyun, S. S., 2013. Examining the influence of casino attributes on baby boomers' satisfaction and loyalty in the casino industry. *Current Issues in Tourism,* 16(4), pp. 343 - 368.

Johnston, R. & Kong, X., 2011. The customer experience: a road-map for improvement. *Managing Service Quality,* 21(1), pp. 5-24.

Kaizen Gaming (Stoiximan/Betano), 2020. *Linkedin profile Kaizen Gaming (Stoiximan/Betano).* [Online]

Available at: https://www.linkedin.com/company/kaizen-gaming/about/ [Accessed 2020 October 2020].

Keiningham, T. L., Perkins-Munn, T., Aksoy, L. & Estrin, D., 2005. Does customer satisfaction lead to profitability?: The mediating role of share-of-wallet. *Managing Service Quality,* 15(2), pp. 172 - 181.

Korneta, P., 2018. Net promoter score, growth, and profitability of transportation companies. *International Journal of Management and Economics,* pp. 1 - 13.

Krassadaki, E. & Matsatsinis, N., 2015. A multi-criteria and statistical framework for measuring and analysing customers' experience. *International Journal of Decision Support Systems,* 2015(1), pp. 18-41.

Laros, F. J. M. & Steenkamp, J.-B., 2005. Emotions in consumer behavior: A hierarchical approach. *Journal of Business Research,* Volume 10, pp. 1437 - 1445.

Mecredy, P., Wright, M. J. & Feetham, P., 2018. Are promoters valuable customers? An application of the net promoter scale to predict future customer spend. *Australasian Marketing Journal,* Volume 26, pp. 3-9.

Morgan, N. & Rego, L., 2006. The value of different customer satisfaction and loyalty metrics in predicting business performance. *Marketing Science,* 25(5), pp. 426 - 439.

NCRG, 2013. *Internet Gambling: An Emerging Field of Research,* Washington DC: National Center for Responsible Gaming.

O'Sullivan, D. & McCallig, J., 2012. Customer satisfaction, earnings and firm value. *European Journal of Marketing,* 46(6), pp. 827 - 843.

OPAP, 2020. *Stoiximan Investment Update.* [Online] Available at: https://investors.opap.gr/en/results-and-news/news/regulatory-announcements/2020/19-11-2020
[Accessed 27 November 2020].

Pine, B. J. & Gilmore, J. H., 1999. The Experience Economy. *Harvard Business Review*, 12 10.

Reichheld, F. F., 2003. The one number you need to grow. *Harvard Business Review,* December, 81(12), pp. 46 - 54 & 124.

Reichheld, F. F., 2003. The One Number You Need to Grow. *Harvard Business Review*, December, pp. 46 - 52.

Rust, R., Zeithaml, V. A. & Lemon, K. N., 2000. Driving customer equity. *The Free Press*, p. 46.

Said, L., 2002. Customer retention in online games market. *Information Management & Marketing Seminar Series*, 22 November.

Said, L. R., Mizerski, D. & Lam, D., 2003. *Comparing the Effect of Habit in the Online Game Play of Australian and Indonesian Gamers.* Adelaide, Proceeding of the ANZMAC Conference, pp. 2526 - 2532.

Schulman, K. & Sargeant, A., 2013. Measuring donor loyalty: key reasons why Net Promoter Score (NPS) is not the way. *International Journal of Nonprofit and Voluntary Sector Marketing,* 18(1-6), pp. 1-6.

Teichert, T., Gainsbury, S. & Mühlbach, C., 2017. Positioning of online gambling and gaming products from a consumer perspective: A blurring of perceived boundaries. *Computers in Human Behavior.*

van Doorn, J. & Leeflang, P. S. H. T. M., 2013. Satisfaction as a Predictor of Future Performance: A Replication. *International Journal of Research in Marketing,* 30(3), pp. 314-318.

Van Doorn, J. & Verhoef, P. C., 2008. Critical incidents and the impact of satisfaction on customer share. *Journal of Marketing,* 72(4), pp. 123 - 142.

Verhoef, P. C. et al., 2009. Customer Experience Creation: Determinants, Dynamics and Management Strategies. *Journal of Retailing,* 85(1), pp. 31 - 41.

Vivek, S. D., Beatty, S. E. & Morgan, R. M., 2012. Customer Engagement: Exploring Customer Relationships Beyond Purchase. *Journal of Marketing Theory and Practice,* 20(2), pp. 122 - 146.