Final Report 8.5.2013

Assessing the Impacts of Marketing Investments

Client: Nokia

Project Group:	Olli Rentola (Project manager)
	Tiina Kangasniemi
	Harri Mäkelin
	Vilma Virasjoki

Contents

1. INTRODUCTION **1**

2. BACKGROUND AND OBJECTIVES **2**

3. LITERATURE REVIEW **3**

3.1. MARKETING 3

3.2. MEASURING RESULTS OF MARKETING ACTIVITIES 4

3.3. METHODS FOR OPTIMIZING MARKETING STRATEGIES 4

4. DATA ANALYSIS AND PREPARATION **6**

4.1. DATA OVERVIEW 6

4.2. DATA PREPARATION 6

4.3. DATA QUALITY 7

5. METHODS AND MODELS **9**

5.1. CLUSTERING 10

5.2. PRINCIPAL COMPONENT ANALYSIS (PCA) 10

5.3. REGRESSION MODEL 11

6. RESULTS **12**

6.1. CLUSTERS 12

6.2. CLUSTER VISUALIZATION 14

6.3. MODELS 15

6.4. RETURN ON INVESTMENT (ROI) 20

6.5. MARKETING STRATEGY RECOMMENDATIONS 22

7. CONCLUSIONS **23**

8. SELF-EVALUATION **24**

8.1. SUMMARY OF THE PROJECT 24

8.2. PROJECT SCHEDULE 24

8.3. PROJECT MANAGEMENT 24

8.4. LESSONS LEARNED 24

8.5. COMMENTS ABOUT THE COURSE 24

REFERENCES **26**

1. Introduction

This project was made during spring 2013 in Aalto University School of Science for the course Mat-2.4177 Seminar on Case Studies in Operations Research. Project's client was Nokia and the aim was to find out what kinds of impacts different marketing activities have in the mobile phone field of business. Additionally, project's more far-reaching goal was to make recommendations for optimal marketing strategies for mobile phones.

Mobile industry is a relatively dynamic field in which there are several large and small players with different kinds of product ranges and brands. The fast development of new technologies and emerging innovations is a dominant feature of the industry and keeps it in constant movement. This creates instability among the players and their positions, which means possibilities for new, innovative and reactive companies to penetrate. Thus mobile industry is by nature a very competitive environment.

Moreover, mobile phones are a huge business: in 2013, the global mobile phone penetration is 96 %, as there are 6.8 billion mobile subscriptions worldwide. This inevitably means that as 100 % will soon be reached and growth rates have in fact already fallen to their lowest levels [International Telecommunication Union, 2013]. Consequently, competition for market shares will grow even more tense and the importance of high valued user experience innovations rises. While also mobile broadband subscriptions are in constant high growth [International Telecommunication Union, 2013], mobile industry has been shifting more towards smart phones.

Furthermore, consumer electronics in general are widely advertised to create brand awareness and purchases among consumers. Due to the very competitive nature of the mobile industry, mobile phones are continuously being advertised through various channels, such as printed media, television, online and outdoors. In addition, mobile phone marketing is not only done by the manufacturing companies, but also to an even larger scale by telephone operators.

The structure of this final report is as follows: First, the background of the project is explained in more detail in chapter 2 and a literature review is conducted in chapter 3. Then, the data preparation phase is described in chapter 4, while the methods and models are discussed in chapter 5. Project's results will be presented in chapter 6 and conclusions in chapter 7. Chapter 8 will furthermore discuss self-evaluation about the project in more detail.

2. Background and objectives

The guiding purpose for making this project was to repeal the saying "Half of the marketing spend is wasted, we just don't know which half." It was assumed that by analyzing statistical data from several years it could be stated more precisely how investments through various marketing channels impact final sales and brand awareness.

For the analysis, a data set containing marketing investments for selected media, sell-out data, and brand awareness survey results across multiple mobile phone brands and countries were provided by Nokia. Time spans of individual data sets varied between two and five years, but they were all collected during years 2008 and 2012.

The main objective was to find correlations between mobile phone sales and marketing investments and thus explain how different types of marketing activities contribute to growth in consumer awareness. This was done by analyzing the data statistically and by developing a mathematical model to explain sales results based on the size and character of marketing investments. In addition, the return on investment (ROI) for the marketing activities was estimated in order to assess how effective these channels are. Eventually, based on the previous analysis, marketing investment strategy recommendations were developed. However, this is largely dependent on the assumptions and interpretations made throughout the analysis.

3. Literature review

3.1. Marketing

Marketing can be defined as the activity and process of communicating the value of a product or service to customers [American marketing association, 2004]. Advertising actions are important both from the economic point of view, but also because according to the economic theory it has an effect on the customers. In short, this is because marketing is either persuasive, which means that it can change brand preferences, or it informs people about the prices and new products. Additionally, marketing may also be complementary to the offerings, which is the case when it, for example, enhances the brand status of the product or service and thus brightens up its image among some customers [Bagwell, 2007]. Furthermore, marketing can have a lasting effect on the structure of an industry and thus may be a source of competitive advantage [Doraszelski & Markovich, 2007]. As a result, companies are investing a lot of money in marketing in order to shift their demand curve upwards [Bagwell, 2007] and to build strategic capabilities.

There are many ways and channels for marketing, such as TV, printed material, online marketing and outdoor advertisements. However, it is difficult to predict how specific marketing investments relate to sales and brand awareness, which could cause inefficient investments and strategic choices. This is because advertising can be associated with many other factors in addition to sales, such as profitability and efficiency. Also, firm's sales are also affected by the past-advertisement [Bagwell, 2007]. It is thus noteworthy to realize that a static approach is rarely enough and one should consider of dynamic models, i.e. how to act through longer time periods. To overcome the problems of how to make optimal marketing decisions, different kinds of techniques have been introduced to estimate the impacts of advertising.

When it comes to marketing and sales, company's current advertising can according to studies be associated with an increase in its sales, but the effect is usually only short-term. Due to its competitive nature, increased advertising often also persuades competitors to invest in their marketing. Thus, it is challenging to determine the overall effect of marketing on primary demand. Additionally, marketing affects the brand loyalty which makes offering's demand curve less elastic. However, studies do not provide strong evidence that advertising would consistently increase brand loyalty or stabilize the market shares [Bagwell, 2007].

Structural models, which rely on economic and/or marketing theories of consumer or company behavior, have recently become more popular among marketing studies. They make it possible test behavioral theories, such as consumer demand or choices, from which they are derived and to obtain behavioral predictions of consumers: i.e. they aim both to explain and to predict. When less emphasis is put on the actual theory and data fitting is prioritized, we can talk about reduced-form models that represent the consumer's or company's historical decision rules as derived from marketing data. The resulting estimates can furthermore be used to predict future behavior of the agents, and the models can be validated by using for example time series, i.e. hold-out data [Chintagunta, Erdem, Rossi & Wedel, 2006].

In addition, it can be argued that different kinds of marketing activities should be addressed to different kinds of customers: advertising providing information on brand characteristics should primarily be aimed at new customers, while prestige creating marketing should affect both inexperienced and experienced customers [Ackerberg, 2001]. Thus, one should make the strategic choices based on the customer's or market's characteristics as well as their predicted behavior.

3.2. Measuring results of marketing activities

Effectiveness of marketing is hard to measure due to the complex and intangible nature of marketing and of its effects. To justify marketing investments marketing managers need measures and metrics to continue investing in marketing activities. In this chapter ways to measure these activities are presented.

The fundamental question of marketing performance is how to measure the effects of marketing. It is complicated to measure how different marketing activities change the way how people perceive the marketed products or services. Measurable factors can be obtained through market surveys or by following the change in the actual sales. Still, it is hard to obtain a clear ratio between the money invested and the actual performance. Due to the different nature of marketing as an investment, it is complicated to measure it with traditional measures such as return-on-investment (ROI). More sophisticated way is to use return-on-marketing-investment (ROMI), which is relatively new metric in the field of marketing. ROMI [Return on marketing investment - Wikipedia, the free encyclopedia. 2013.] describes the ratio between the incremental revenue attributable to marketing and the amount spent on marketing altogether.

ROMI metric has two forms depending on the period of measurement: short and long term ROMI. Short term ROMI measures the amount spent to marketing for every unit of revenue. Success can be measured by other variables such as change in market share. Short term ROMI measures the direct effects of marketing to the revenues, while long term ROMI takes into account the less tangible effects of marketing such as the increase in brand awareness, consideration and purchase intent. By using these different metrics the short term and long term effects can be determined.

3.3. Methods for optimizing marketing strategies

Marketing mix modeling (MMM) is a common way to find optimal marketing strategies. Marketing mix modeling takes into account multiple variables affecting the outcome and therefore helps marketing professionals to optimize their marketing activities with respect to, for example, revenues or brand awareness. This approach uses multivariate regression models to distinguish differences between performance levels of different marketing activities. MMM defines the effectiveness of each of the marketing elements in terms of contribution to the overall performance.

Another aspect of optimal advertising considers the timing and the length of campaigns. In many industries technique called “pulsing” is recognized as a useful way of advertising [Dubé, Hitsch and Manchanda, 2005]. The basic idea is that their advertising strategy is to switch systematically their advertising “on and off” at high frequency. Basically this means that there are periods of zero advertising and periods of intensive advertising instead of steady stream of advertising.

It is shown that by adjusting timing and amount of advertising one can increase the overall profits of a product life-cycle [Horsky & Simon, 1983]. Advertising heavily in the initial periods of product life cycle and decreasing the amount spent as sales grow can be beneficial for a company. This is explained by the word-to-mouth information flow through early adopters and that the information is diffused among consumers over time without significant efforts from the company itself [Horsky & Simon, 1983]. With such activities a company can cause the peak of sales to be higher and occur earlier than it would normally do.

A more straightforward way of optimizing advertising strategy is to find the optimal way to do a marketing investment in a way that maximum number are reached in a given time period. This idea could be extended so that as the result of the optimization a Pareto optimal surface would be obtained in respect to the length of the campaign, money invested and the amount of people reached.

4. Data analysis and preparation

4.1. Data overview

The basis for modeling and identification was the three datasets provided by Nokia: brand awareness data, sell-out data and media investment data. All data sets contained monthly data for multiple brand-country pairs. Before identification phase all data sets needed to be processed, unified and aggregated into a single consolidated set of data.

Brand awareness data is survey data containing results from customer surveys representing the rate of spontaneous and prompted brand awareness, brand preference, confidence intervals for each, and number of respondents. Data represents the period between January 2011 and December 2012, and it covers data from 16 countries. Perfect time series can be extracted for 15 countries and 11 brands, thus producing 165 brand-country pairs. The overall quality of the data set is relatively good, but the short time frame and low information value for less-known brands posed a real challenge for the identification of brands.

Sell-out data set contains mobile phone sales information for 9 brands in 16 countries extending from January 2008 to November 2012. These 16 countries are the same countries as in brand awareness data. The data set gives both absolute sales volume and value information and percentual value and volume share information by country for each brand-country pair. Also this data set has good quality and some missing data points are the only districting factor.

Media investment data covers advertising investments for multiple brands in 14 countries in 2010-2012. Investments are categorized by product, advertiser and advertising channel. The quality of this dataset appeared to be less good than expected and thus it demanded plenty processing. Procedures are explained in more detail in section 4.2. Data preparation.

4.2. Data preparation

Data preparation can be roughly divided into three phases: first, all data sets were processed individually in order to prepare them for aggregation. In the second phase, different types of data were aggregated into a single master table. In third phase of data preparation the data was clustered into several clusters with different types of customer profiles. This was done in order to maximize the amount of data to be used for modeling but also to enable building a more customer-oriented model. This data preparation phase is explained in more detail in section 5.1. Clustering.

Procedures for data manipulation for brand awareness data and sell-out data were relatively simple. Brand and country names were standardized and data series were led to a uniform format. This was conducted with SPSS Modeler.

Media investment data posed the largest challenge regarding the data processing. More of the data quality at this stage is discussed in chapter 4.3. As the procedures of data collecting clearly varied from country to country, a lot of manual work was required to map single media investments to specific brands. At the end, 69,6% of the monetary value of the investments had been mapped to some brand or identified as non-brand specific investment. This is presented in Figure 1. 33,7% of the monetary value was mapped for those 9 brands, for which we had received sell-out and brand awareness information (in Figure 1 “Brands in the analysis”). Investments categorized as non-brand specific were mostly made by operators. 30,4% of the monetary value of the data remained unmapped due to problems with data quality. Of this, 86,2% was investments made by operators. Mapping and sorting the data was done with Microsoft Excel.

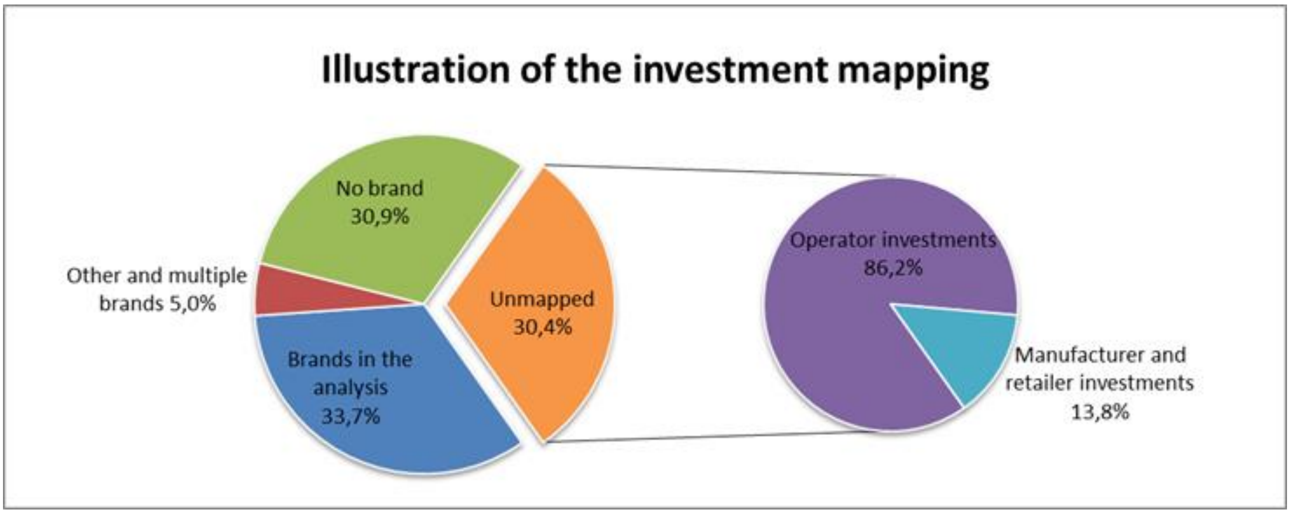

Figure 1: Media investment data division in mapping

After mapping investments for specific brands, data was aggregated on a monthly basis in order to obtain total investments per month for each brand-country pair and information how the investments were allocated between different advertising channels (TV, radio, outdoor, print, online, cinema). After aggregation, unification procedures were carried out for the data set, in order to ensure compatibility with other two datasets. If empty slots were detected within time series, they were treated as zeros and names of the countries and brands were unified. Aggregation and unification were done with SPSS Modeler.

In second phase of data modification all three data sets were aggregated into one master table which contained sell-out, investment and brand awareness information for each brand-country pair aggregated on a monthly basis. This data set is an outer join of the separate data sets. Outer join includes such country-month-brand lines for which there is only partial information available regarding investment data. These investment data points were assumed zero. As the brand awareness data had the shortest time frame, thus the data set contains fairly small amount of data extending only from January 2011 to December 2012.

4.3. Data quality

The data quality was problematic due to many reasons. When preparing the investment data for the analysis, the main quality concerns were the multiple practices in data entries, entries made in foreign languages, unknown mobile phone brands and large amount of data (altogether approx. one million data points). As low quality of data prevented automation in data manipulation, most of the processing was done by hand. Even though procedures were made with high caution, probability of some data corruption stands and this needs to be taken account when interpreting the results.

Additionally, operator marketing investments tend to promote the operator instead of a specific brand and operator marketing spend is larger than manufacturer marketing spend, thus the risk of operator investments affecting bias to the actual model is an existing risk. As practices and accuracy in recording marketing investments differ in different countries, also the reliability of the data comes to question. Because some brand-country pairs can be found for which some months seem to have significantly smaller investments than other months for the same brand-country pair, an existing risk is that only a part of the actual marketing investments is recorded and thus the input used in modeling is inaccurate. Also the definitions for the media channels are not as unambiguous as thought first; for example, whether 'outdoor' as marketing channel includes posters and public events might, and probably does, vary through countries.

Another issue that affects the data quality and the analysis is that pre-announcements, launches and pre-sales can affect the sales of a certain brand already before the actual marketing investments are made. This can for example be seen in Figure 2, where the marketing investments and the value share of brand 6 in country 11 are shown. The red line represents the value share, for which the peaks seem to happen before the peaks of the blue line, the money invested. Moreover, these peaks go one-by-one with the product release dates of the brand and country in question. Thus, one cannot explain the sales of the brand by the investments made, when the causality of the events is not reasonable in this sense in the data.

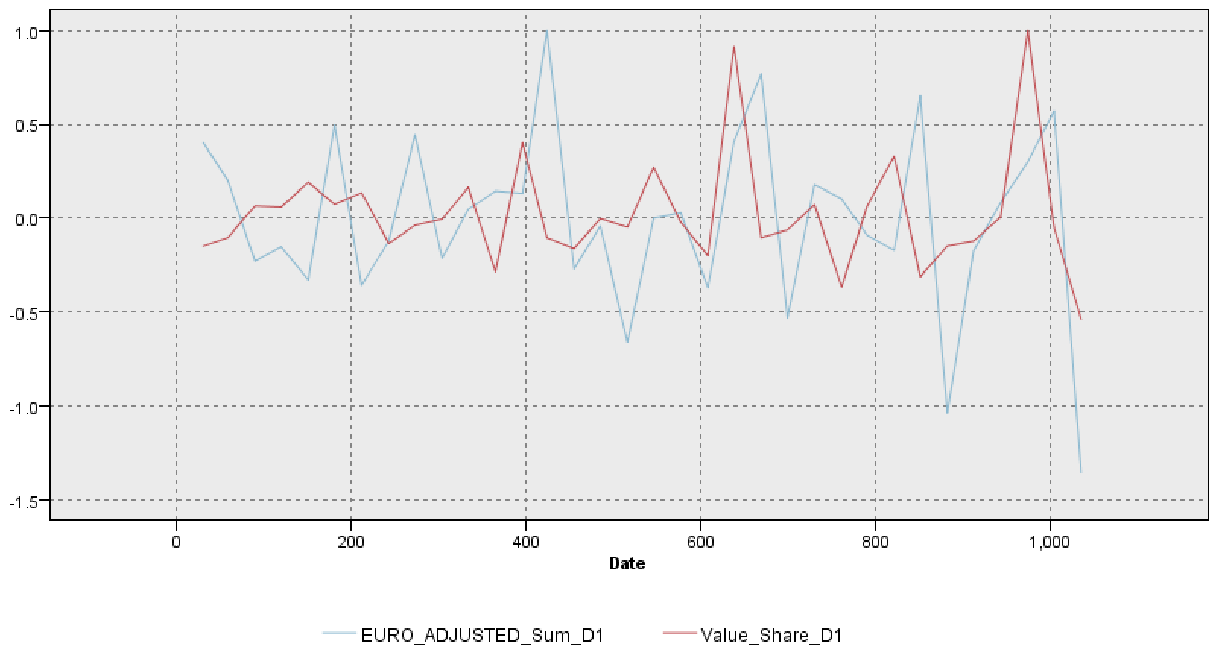

Figure 2: The marketing investments (blue) and value share (red) for brand 6 in country 11

5. Methods and models

One of the biggest challenges in the project was to find a way to create recommendations for optimizing marketing investment across several different marketing channels. More specifically, a key problem was how to draw recommendations from the estimated effectiveness of marketing investment in different marketing channels in promoting either brand awareness, preferences or, ultimately, buying decisions and market share.

To illustrate this problem, a chart displaying the relationship between total marketing investment and market share for a certain large market is provided below in Figure 3. As can be seen from the figure, the effectiveness of marketing investment varies between brands; for example, Brand 6 has attained a market share of roughly three times that of brand 7 with comparable investment share. The reason for this was assumed to lie in the differences in brand preferences in the market, i.e., more strongly preferred brands need not invest as much as less known or less preferred ones. Consequently, since it might have been difficult to fit a universal model to all of the data, it was decided that the data set be divided into segments based on brand awareness survey profiles. In other words, it was assumed that similar brand awareness and preference profiles would yield similar estimates for the effectiveness of a given marketing channel, and thus a separate model could be identified for each segment. This would also provide a way to identify the most effective marketing channels for each brand consciousness profile, on the basis of which optimal marketing channel allocation recommendations could be drawn. Therefore, a clustering algorithm was employed to segment the data into several brand consciousness profiles.

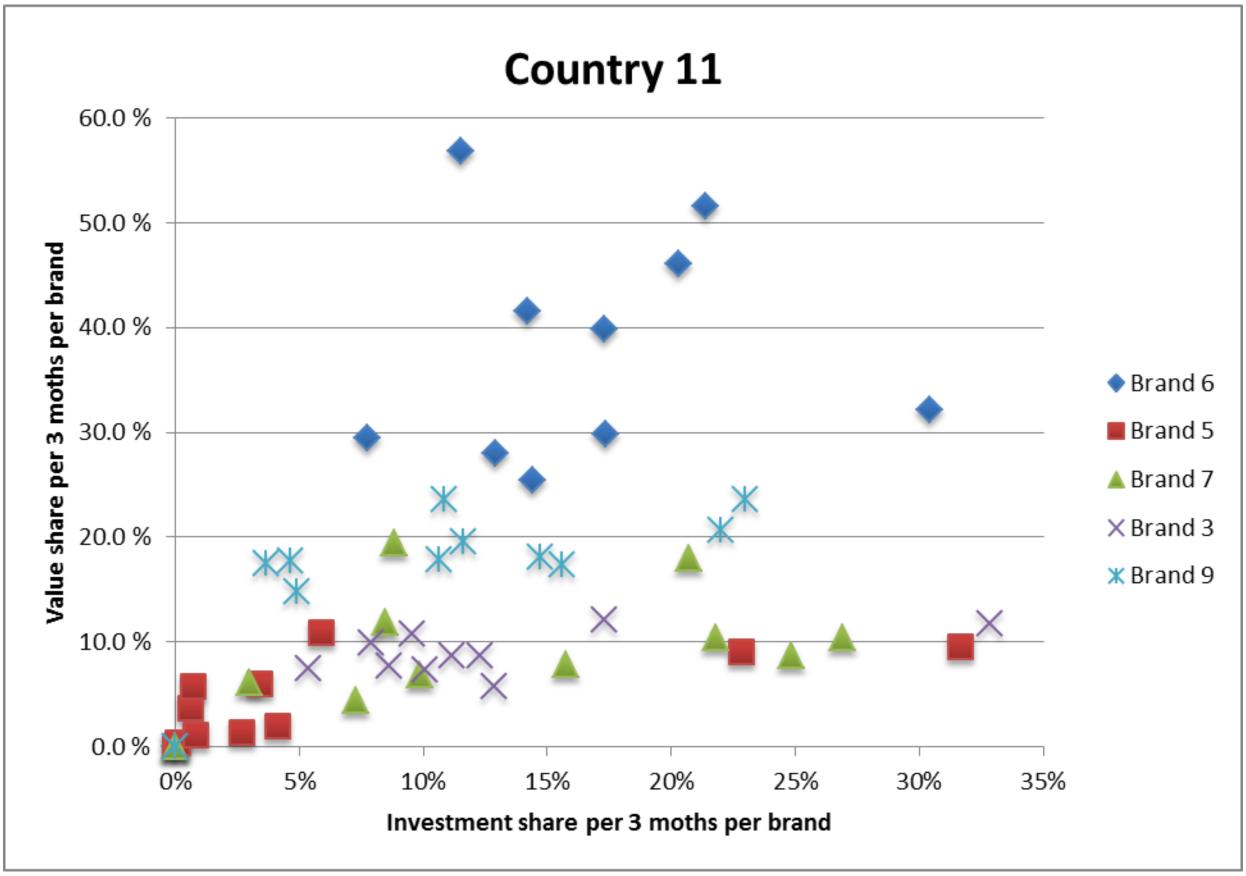

Figure 3: Example of value share (%) vs. amount invested in marketing (%)

5.1. Clustering

The basis for the modeling was made by clustering the brand awareness data into three segments. This would create so called customer profiles, which represent different kinds of customer groups with similar preferences inside them. The models can thus be made individually for each of these clusters. The decision to carry out the modeling this way was based on the hypothesis that each of the examined brands has a different kind of profile and awareness among consumers, which further affects how the marketing investments are reflected onto demand, i.e. which kinds of correlations can be found. This was already brought up in the literary review, where it was stated that different kinds of consumers (such as new and old) require different kind of marketing as well. Mainly, clustering was done due to the fact that the data set was not lengthy enough to do for example time series models or more precise models for example based on country-brand pairs. Additionally, the differences among the customers have an effect on the fact which kind of marketing strategies are optimal in each situation. Thus, so be able to create strategic recommendations for mobile phone marketing, the segments also act as a basis to separate the overall results into more precise strategies for different kinds of circumstances.

In order to create these customer profiles, brand awareness data was clustered into segments in SPSS with its K-means clustering function. The input data for the function only included the survey answers (decimals between 0 and 1 for several variables), so the data points could be put into any possible segment regardless of which brand it was recorded for. These input variables were prompted awareness, consideration, total spontaneous awareness, top of mind, preference and strength of preference.

K-means algorithm works so that the user defines how many clusters will be created, after which the algorithm divides all data points into the clusters so that their squared distance to the cluster's center is minimized. Formally, the goal is to minimize the function

$$\sum_{i=1}^{k} \sum_{x_j \subset S_i} \left| \left| x_j - c_i \right| \right|^2,$$

in which k is the number of clusters, c_i is the assigned center of each cluster, and S_i is the set of observations assigned to cluster i. At each iteration stage the algorithm then places the new cluster centers at the means of the obtained clusters:

$$c_i = \frac{1}{|S_i|} \sum_{x_j \subset S_i} x_j.$$

As a result, all the data points from the brand awareness set belong to a cluster, which now represent customer preference profiles.

5.2. Principal component analysis (PCA)

To validate and visualize the identified clusters, also a principal component analysis was made on the brand awareness survey data. Principal component analysis is a way to find those components from a data, which are the most relevant to describe it without losing any significant information. Basically, it means searching for those hyper surfaces so that the variance is maximized when the data is projected on them. The first principal component (axis) is chosen so that it has the largest possible variance in the data. The next components are again chosen so that they both are orthogonal to the preceding components and maximize the variance. Mathematically, this can be written as:

$$\max D^2(\beta^T X) = \beta_1^T \Sigma \beta_1^T = \lambda_1,$$

in which X is the data matrix, Σ is the covariance matrix of X, λ_1 is the largest eigenvalue of Σ and β_1^T is the eigenvector corresponding to λ_1. Now the first principal component is

$$y_1 = \beta_1^T X$$

The next components can hence be obtained similarly, but with the orthogonal condition. Thus, for the second component, it means

$$\max_{\substack{\beta \\ \beta'\beta=1 \\ Cov(y_1,\beta' X)=0}} \quad D^2(\beta^T X) = \beta_2^T \Sigma \beta_2^T = \lambda_2,$$

in which λ_2 is the second largest eigenvalue of Σ and β_2^T is the eigenvector corresponding to λ_2 [Mellin, 2008].

5.3. Regression model

A separate linear regression model was built for each of the three identified brand consciousness profiles. The candidate predictors were identified based on their cross correlation functions with the target variable, after which a subset of inputs was selected based on the goodness of the model as evaluated by the Akaike Information Criterion (AIC). The use of AIC reduces the risk of over-fitting the model, i.e., having too many input variables. Also, a simple check whether all coefficients were positive was applied to ensure the model was indeed feasible because it was assumed the effect of marketing investments on sales could not be negative. Forcing the coefficients to be positive was also attempted using R's penalized constrained least squares function (pcls), but it was found out that zero coefficients were too often obtained with that method. R's pcls uses quadratic penalty functions to force the solution to satisfy given linear equality and inequality constraints.

Formally, a linear regression model relates the target variable y to the predictors x_i through a linear model:

$$y = \beta_0 + \sum_i \beta_i x_i + e_i$$

in which e_i is an independent, homoscedastic normally distributed error term with mean zero.

The nature of the marketing investments did not show any need for a time-series model, which is why it was chosen to use a regression model.

6. Results

6.1. Clusters

As can be seen from Figure 4 to Figure 6, the brand awareness data was exhaustively divided into three clusters. Cluster 3 is approximately twice as big as cluster 2, which contains the least data among the clusters. The input cells of the table represent both the total distribution of the data and how the data in each cluster is divided within these. Figure 5 additionally shows how the clusters differ from each other and how their mean values and variances are located in comparison with the total distribution (the white bar). Finally, Figure 6 represents how different brands are divided into the clusters.

When further comparing the clusters with each other, cluster 1 includes observations of six brands, but more than half of its data points are covered by brands 1 and 2. Cluster 1 covers situations where the brand is not well-known and thus it also has a very weak preference. Cluster 2 on the other hand contains data from two of the most strongly preferred brands, 6 and 9. This cluster has both high awareness and high preference rates, but slightly larger variation than other clusters, especially with preference. Cluster 3 includes data from six different brands and it also has a very high awareness profile but it differs from cluster 2 by having much lower preference rate. All in all, cluster 1 includes low awareness and low preference observations, cluster 3 high and medium awareness and low preference observations and cluster 2 high awareness and high preference observations.

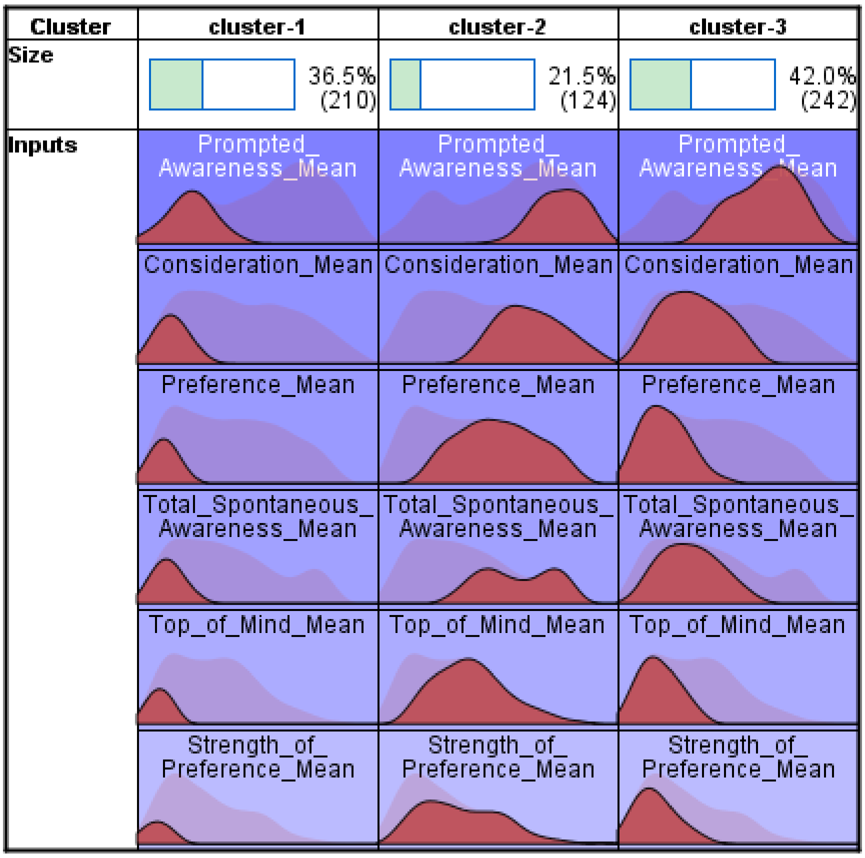

Figure 4: Brand consciousness profiles for different clusters

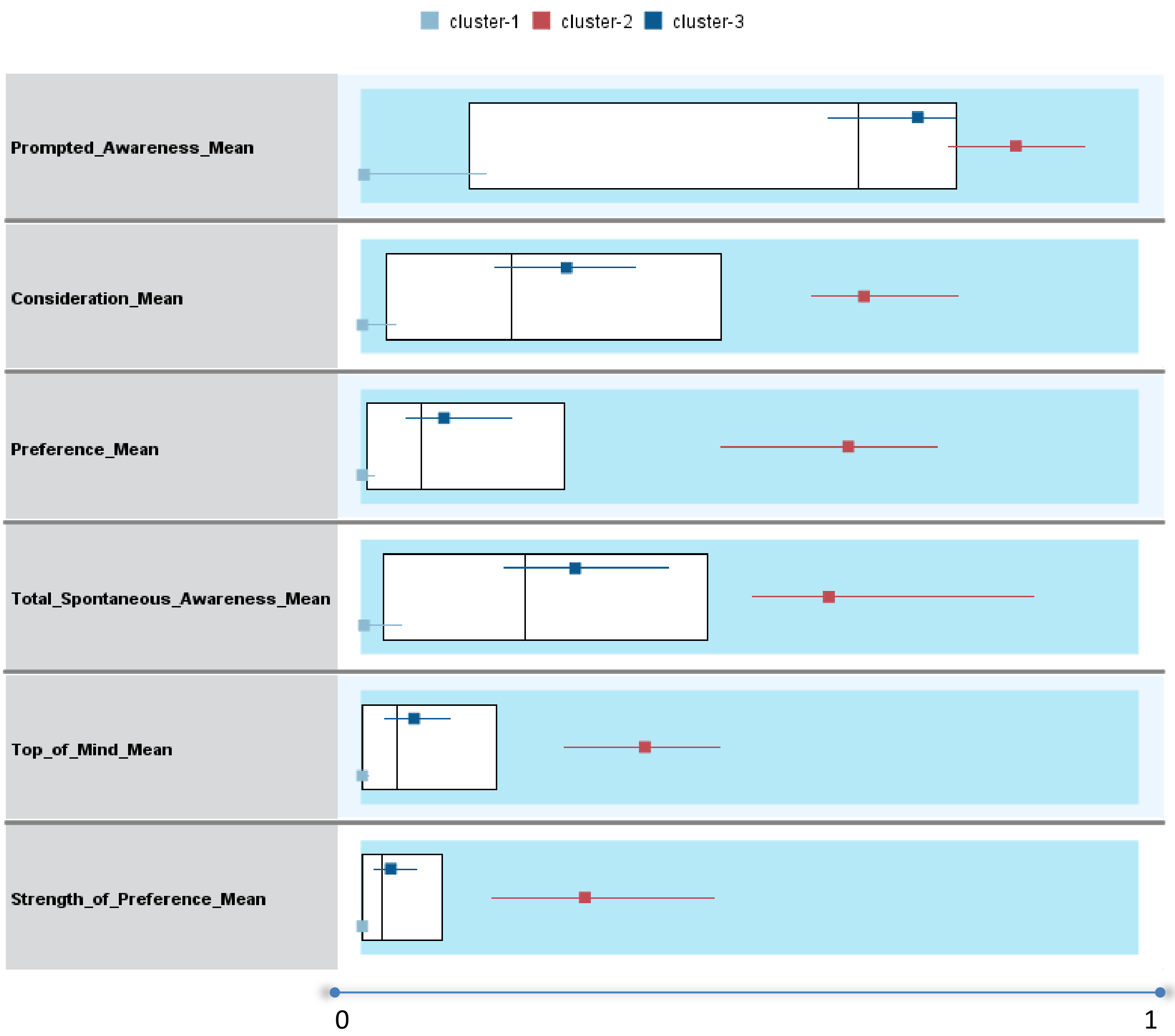

Figure 5: Average values of brand awareness survey response categories for different clusters

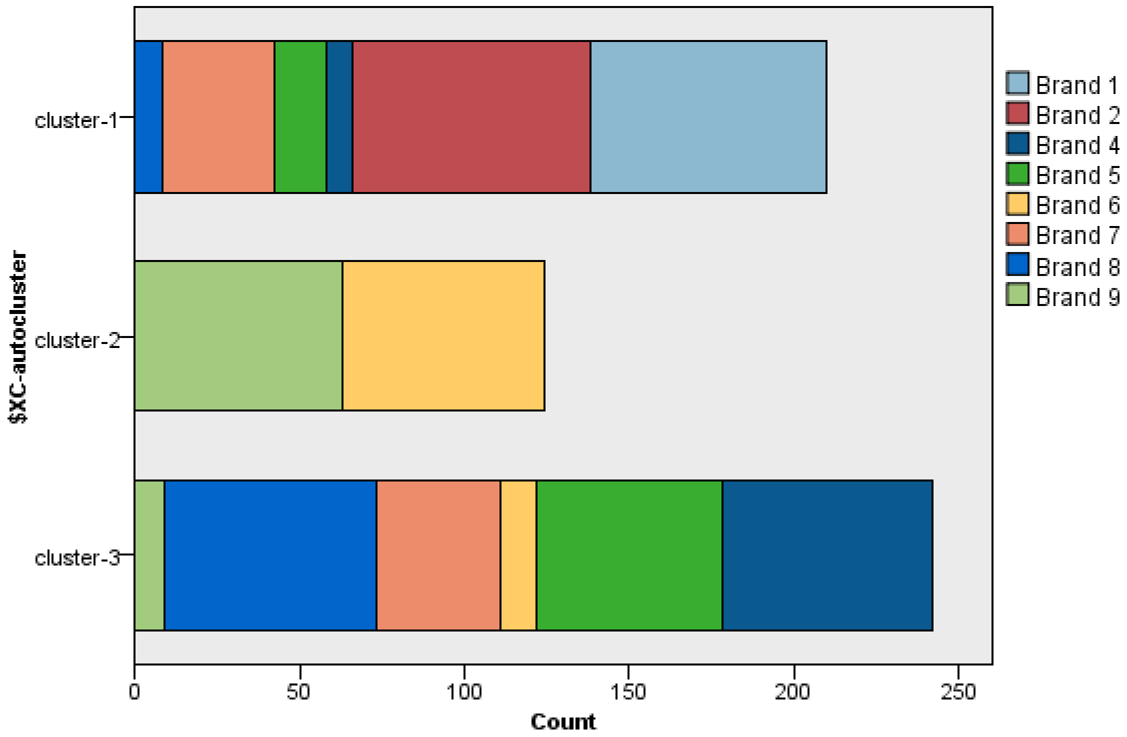

Figure 6: Distribution of brands over different clusters

6.2. Cluster visualization

To ensure the identified clusters were indeed valid and to gain a better understanding of what they looked like, a PCA was performed on the brand awareness survey data, i.e., the same variables that were used as input for the clustering algorithm. The idea was to reduce the dimension of the data so that the clusters could be visualized. Two PCA factors were generated, the loadings of which are listed below in Table 1.

Table 1: PCA factor loadings

Brand consciousness metrics	Factor	
	1	2
Top of Mind	0.926	0.011
Total Spontaneous Awareness	0.932	-0.272
Prompted Awareness	0.829	-0.465
Consideration	0.976	-0.050
Preference	0.933	0.295
Strength of Preference	0.823	0.489

The projection of the clusters into two dimensions is visualized below in Figure 6. From the upper plot it can be seen that while the clusters look reasonably well defined, cluster 2 is split into two distinct groups along the direction of Factor 1. From the lower plot it can be seen that the split clearly divides cluster 2 into two separate brands. It could therefore be argued that cluster 2 consists, in fact, of two separate clusters with both brands occupying their own cluster. However, for the purposes of the analysis, they have been kept together as they are clearly distinct from the rest of the data points and separating them would make the clusters too small to be useful for model building. Based on the factor loadings in Table 1, the higher Factor 2 scores of Brand 6 in cluster 2 imply it is somewhat more preferred but less known than Brand 9.

Figure 6: PCA cluster visualization

6.3. Models

The coefficients of the models for the different clusters are listed below in Tables 1 - 3. For cluster 1, the chosen predictors were TV, outdoor, and print marketing shares with lags 1, 0, and 0, respectively. The intercept was insignificant, which suggests absence of linear trend. Also, the print share predictor was somewhat insignificant, with its p-value close to 10%. The coefficient of determination for the model was R^2

= 21,0%, i.e., marketing effects alone cannot explain more than around one fifth of the total variation in sales. However, this was the highest coefficient of determination among the three models built for each cluster. Cluster 1 was also the cluster where the strongest correlations between the input variables and the target were observed.

Table 2: Linear regression model for cluster 1

R^2 = 21,0%	Coefficient	Significance	Share of marketing investments in cluster 1
Intercept	0,001	0,382	
TV Share of Marketing Lag 1 Diff. 1	0,126	0,000	68,141%
Outdoor Share of Marketing Diff. 1	0,569	0,001	6,563%
Print Share of Marketing Diff. 1	0,108	0,121	10,178%

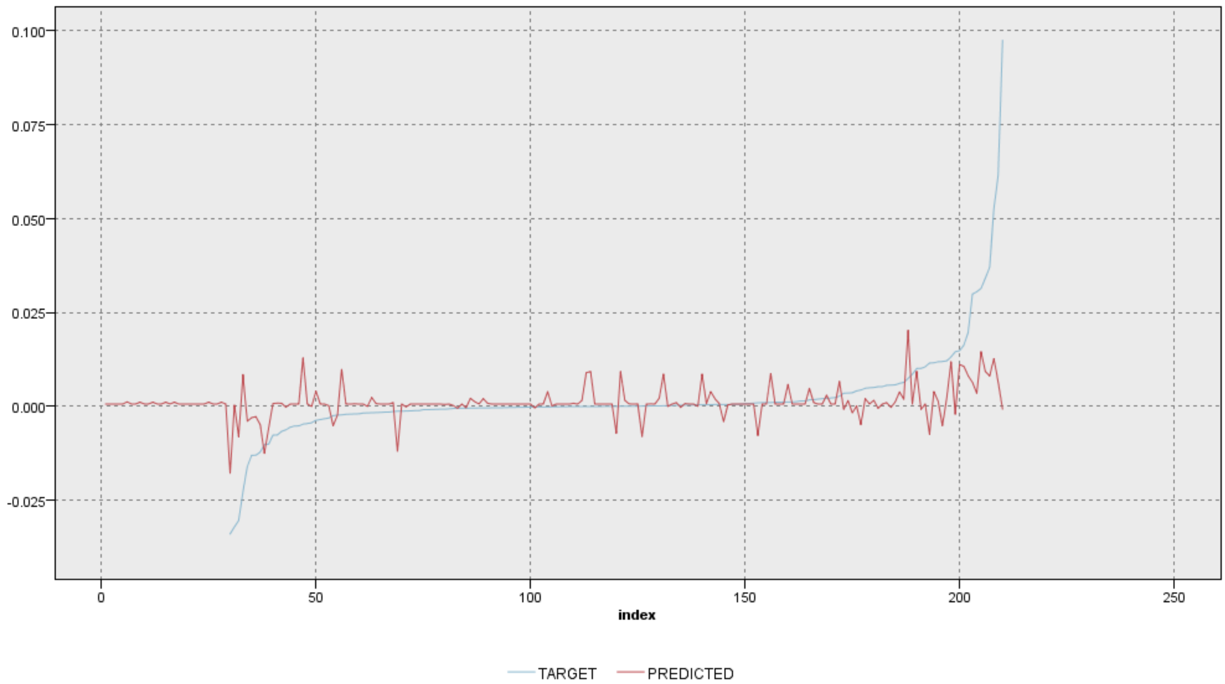

Figure 7 Predicted vs. data

The Studentized residuals of the model for Cluster 1 are plotted in a histogram chart below in Figure 8. The residual distribution is somewhat skewed to the left and does not quite follow normal distribution, which might pose questions for the validity of the model.

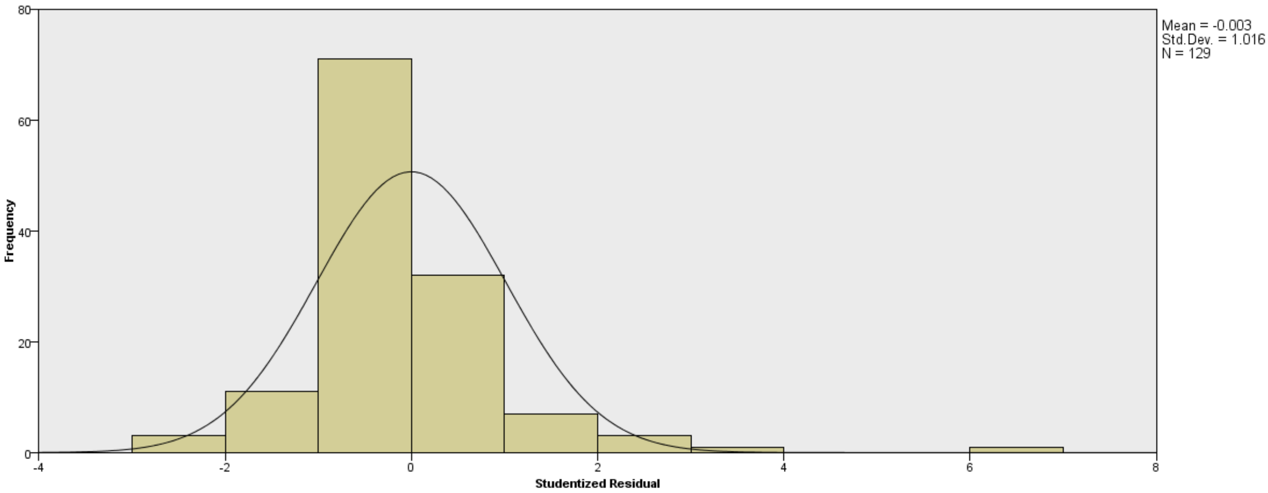

Figure 8: Cluster 1 regression residuals

Cluster 2 showed less correlation between the input variables and the target variable. The predictors that were eventually selected for the model were outdoor, cinema, and print advertising with lags of 2, 1, and 0, respectively. In this case the coefficient of determination was only R^2 = 1,6%, which was the lowest among the clusters. Also, the p-values for the predictors were much higher than for cluster 1.

Table 3: Linear regression model for cluster 2

R^2 = 1,6%	Coefficient	Significance	Share of marketing investments in cluster 2
Intercept	0,021	0,020	
Outdoor Share of Marketing Lag 2 Diff.1	0,237	0,170	5,802%
Cinema Share of Marketing Lag 1 Diff.1	0,897	0,191	0,887%
Print Share of Marketing Diff.1	0,268	0,212	8,908%

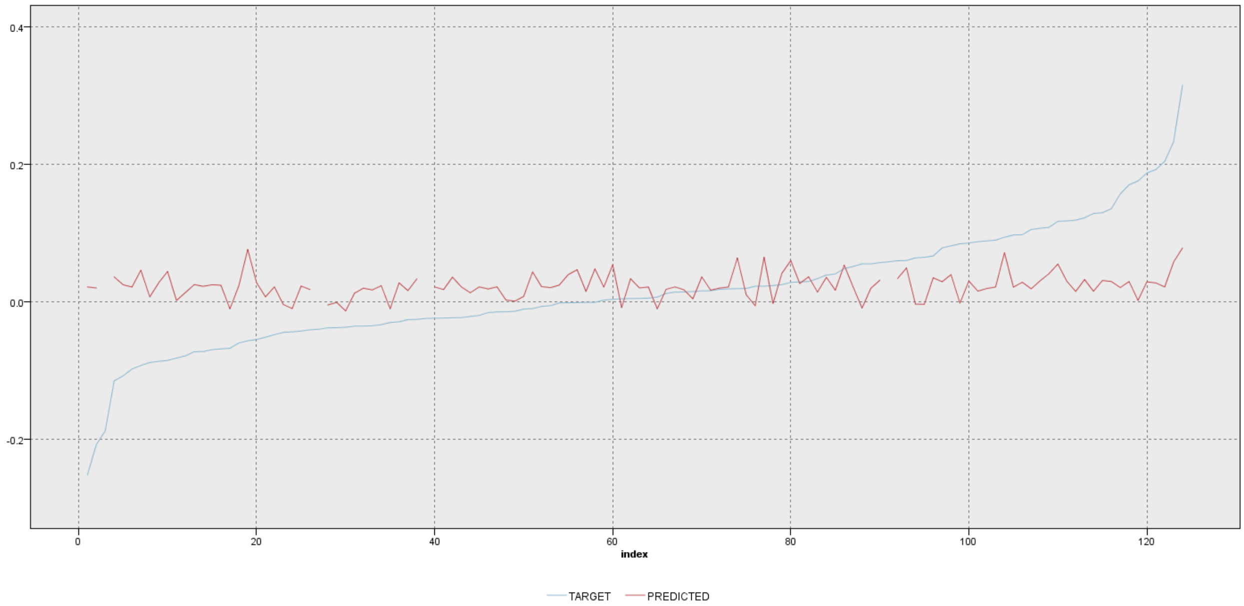

Figure 9 Predicted vs. data

The Studentized residuals of the model for cluster 2 are plotted in a histogram chart below in Figure 10. The residual distribution is yet again skewed to the left and does not follow normal distribution.

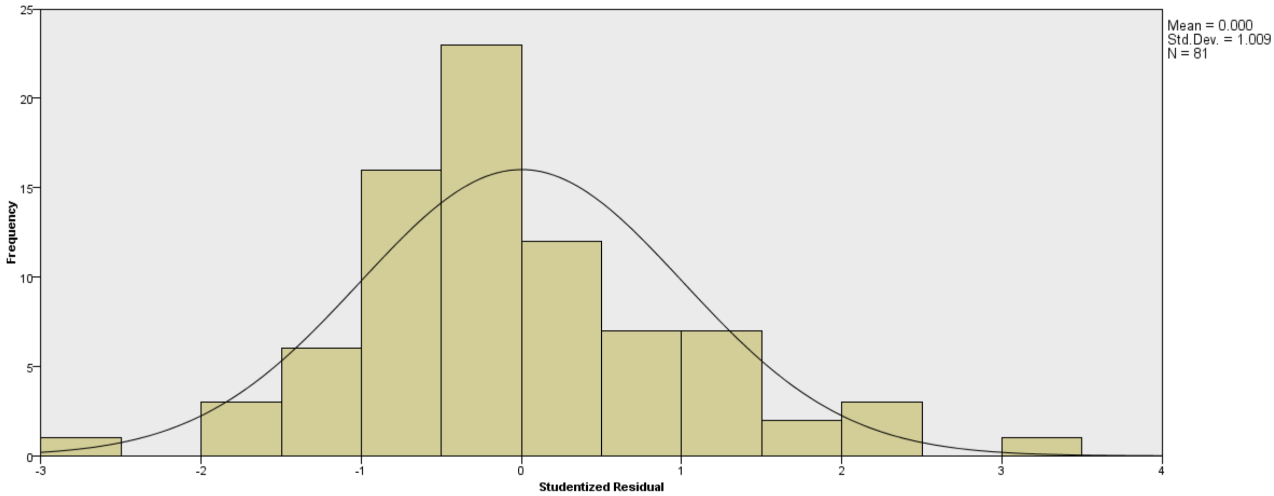

Figure 10: Cluster 2 regression residuals

Finally, for cluster 3, TV, online, cinema, and print marketing shares with lags 0, 1, 1, and 0 were selected as inputs for the model. The coefficient of determination R^2 of 8,2% was quite low. TV and online marketing seemed to be more significant predictors than cinema and print.

Table 4: Linear regression model for cluster 3

R^2 = 8,2%	Coefficient	Significance	Share of marketing investments in cluster 3
Intercept	-0,003	0,270	
TV Share of Marketing Diff.1	0,095	0,010	74,589%
Online Share of Marketing Lag 1 Diff.1	0,260	0,027	9,860%
Cinema Share of Marketing Lag 1 Diff.1	0,671	0,152	1,004%
Print Share of Marketing Diff.1	0,102	0,273	10,349%

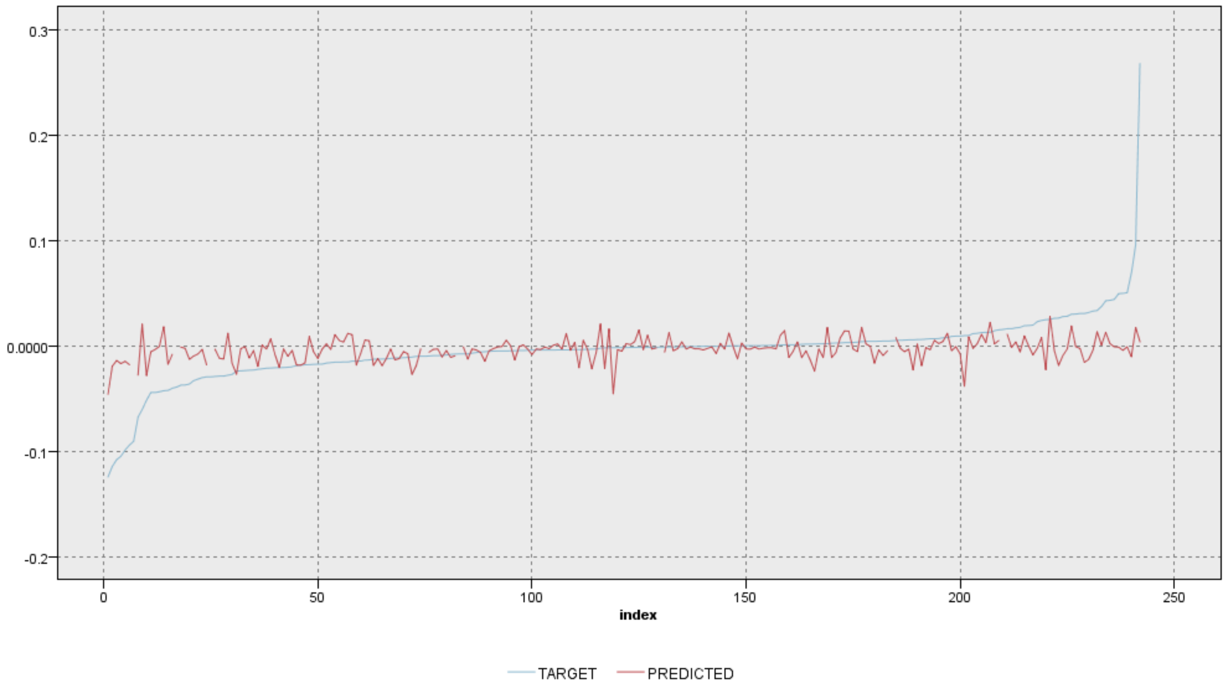

Figure 11 Predicted vs. data

The Studentized residuals of the model for Cluster 3 are plotted in a histogram chart below in Figure 12. The residual distribution has high kurtosis and does not follow normal distribution

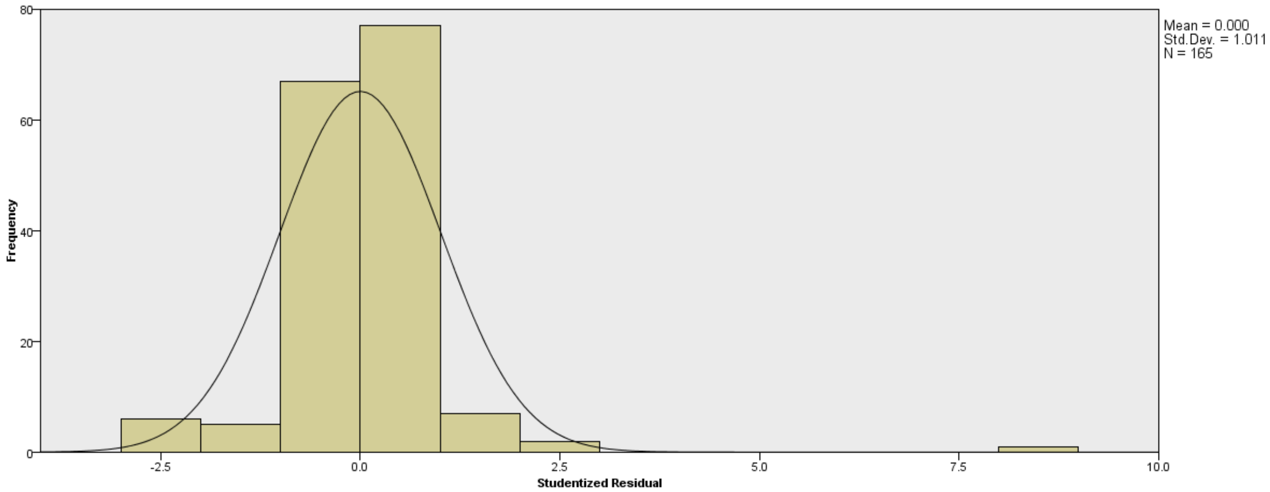

Figure 12: Cluster 3 regression residuals

Models obtained for clusters 2 and 3 do not seem to explain changes in sales well. Observations in these clusters are high awareness profiled and in the literature review it has been speculated that providing more specific information for well-known brands is more useful. Thus for example new releases or product introduction videos might explain peaks in sales better than investments made through regular marketing channels.

6.4. Return on Investment (ROI)

The measurement of return on marketing investment (ROMI) is still a difficult task. If the models are valid, an indicator for the effectiveness of different marketing channels could be easily obtained from the estimated coefficients. However, the models did not completely satisfy the necessary assumptions of linear regression.

For example, assume total investment for some brand in some market for channel X is TI_x, and market size or total sales are TS. A unit percentage increase in the marketing investment share in channel X for a certain brand would then cause an increase of $0.01 \cdot C \cdot TS$ in sales for that brand where C is the corresponding coefficient in the model. The sales-to-expenditure multiplier would then be

$$\frac{0.01 \cdot C \cdot TS}{0.01 \cdot TI_x} = C \frac{TS}{TI_x}$$

Note that the multiplier is not the same as ROMI; the calculation of ROMI would require information about brand sales margins. The multiplier for the modeled variables is listed below in Table 5.

As can be seen from the tables 5, 6 and 7, the multipliers are heavily inflated, with the exception of TV investment in cluster 3. There are several causes that might inflate the multipliers; for example, recorded total investments were very small compared to total sales, which implies not all marketing investments could be captured in the analysis – this was a major problem especially in Cluster 1. Also, marketing investment share might not be relevant for channels such as online and outdoor, where marketing investments do not really affect other competitor's marketing efforts. Using share of investment could be justified for TV, where marketing investments compete for limited air time, which might also explain the less inflated multiplier for TV in cluster 3.

Table 5: Multipliers TS/TI for marketing channels in cluster 1

Cluster 1	$C\frac{TS}{TI_x}$
TV Share of Marketing Lag 1 Diff. 1	66.79
Outdoor Share of Marketing Diff. 1	3177.40
Print Share of Marketing Diff. 1	388.01

Table 6: Multipliers TS/TI for marketing channels in cluster 2

Cluster 2	$C\frac{TS}{TI_x}$
Outdoor Share of Marketing Lag 2 Diff.1	42.32
Cinema Share of Marketing Lag 1 Diff.1	1071.04
Print Share of Marketing Diff.1	24.24

Table 7: Multipliers TS/TI for marketing channels in cluster 3

Cluster 3	$C\frac{TS}{TI_x}$
TV Share of Marketing Diff.1	1.39
Online Share of Marketing Lag 1 Diff.1	48.44
Cinema Share of Marketing Lag 1 Diff.1	1252.25
Print Share of Marketing Diff.1	17.48

The results of the analysis conform to some expectations; for example, the effect of TV advertising was stronger in cluster 1 than in cluster 3, where awareness is higher. Also, online was a significant predictor in cluster 3, which mainly consisted of higher preference brands when compared to cluster 1. Intuitively, TV and outdoor advertising should be better at raising brand awareness, whereas print and online advertising – where more specific information about the product can be conveyed to a consumer who is already considering the product – are better at driving preference and direct sales. Therefore it could also be argued that the direct effect of TV marketing on sales should be weaker than that of, e.g., online marketing. It was clear that marketing investment played a bigger role in explaining the variation in sales in Cluster 1, where brand awareness and preferences were low.

The fact that marketing investments could not explain the variation in clusters 2 was expected. Cluster 2 especially consists of brands with high preference, and consumers who strongly prefer a specific brand would be expected to not be influenced by marketing efforts. For cluster 3 average preference was slightly lower, and therefore it possibly is the reason that a slightly better coefficient of determination was obtained.

Although TV advertising seems to be less effective in driving direct or short term market share, it still has a role in raising awareness and keeping the brand in customers' minds, which would also justify the comparatively large marketing investment in TV versus other channels.

Of course, because shares of investment in a specific channel within a certain market are being compared, the ROMI for a marketing channel might be better or worse depending on how easy it is to increase that share.

6.5. Marketing strategy recommendations

Based on the clusters and the models that were made for each of them, qualitative but considerate recommendations of which kind of the marketing investments seem to be most reasonable in each case can be made. However, the results only rely on the models made in this project and the data that was available for the analysis, which means that they should not be generalized without critical examination.

According to the analysis, the model that was made for cluster 1 was the best since it had the highest coefficient of determination. In other words, the sales could be explained with the marketing investments better than in comparison with other clusters and their models. Thus, one could claim that marketing is more effective for brands that have both low brand awareness and low brand preference. For cluster 1 the significant marketing channels were TV, print and outdoor, although from these TV was the most significant.

For clusters 2 and 3 as good models were not achieved, and their models R^2 values were much lower. As speculated in section 6.4., this might be an indication that advertising of well-known and even preferred brands is not as influential over short term as it is for less known brands. However, advertising might have an effect on keeping the long term sales on a certain level. In literature review an argument was done about the importance of allocating different kinds of marketing for experienced and inexperienced customers. Especially in the case of high brand awareness, more informative advertising might be necessary. As the coefficients of determination are quite low for both models, the model could not determine which could be the best channel for advertising for high awareness and low preference situations. Hence, any of the marketing channels cannot be stated as more significant than the others and recommendations cannot be given.

All in all, it can be summed up that according to this analysis marketing has more distinct influence on short term sales when the preferences and awareness among customers are low or average than when the brand already has a high awareness rate. Additionally it is possible that new product launches are important when considering marketing for consumers with a high awareness and preference profile, and since they were not necessarily covered in our data, the models were not able to explain sales that well.

7. Conclusions

The analysis provided support for the assumption that marketing investment is more effective for brands with lower awareness and preference rates. The coefficient of determination was clearly higher for the model in cluster 1, which contained observations with both low brand preference and low brand awareness; and low for cluster 2, which contained strongly preferred brands.

While the analysis was able to uncover some insights and provide support for some of the initial assumptions and hypotheses, the models as such may be insufficient to draw firm conclusions about the impact of marketing investment to sales. The relationship between investments in different marketing channels and sales was not strong, and in many cases the investment in some of the channels was zero for long time periods, which meant there was not much data to begin with. The models did not satisfy the assumptions of linear regression, and results should therefore be taken as indicative

Regarding further lines of development, an obvious way to improve the model would be to include information about new product launches. It would not be a stretch to assume that new product launches are one if not the most significant predictor for sales, and excluding them from the analysis is a major limiting factor in terms of how accurately the impact of marketing activities can be predicted. However, compiling such a data set might take some effort, as there are several product launches for each brand over the time span of the analysis, and different countries might have different launch dates for the same product. Thus in order to be able to concentrate on building a good and extensive model, it supposedly would be beneficial to focus on one brand in a particular type of countries.

Another major problem in the analysis was the reliability of the investment data. For example, the network operator business is a whole order of magnitude larger than the handset manufacturing business, and the promotions of major operators have a significant impact on both sales and brand consciousness. However, with the available data it was hard to attribute operator marketing investments to any specific phone brands, and in many cases not enough of the total marketing spend in a certain region could be captured for analysis. Also the possibility of missing entries was extremely high. Consistent and comprehensive high-quality data is a requirement for a functioning model. Perhaps even more importantly, the phone business is becoming more and more ecosystem centric, and the actions of the ecosystem players carry more weight than those of the handset manufacturers.

8. Self-evaluation

8.1. Summary of the project

The aim of this project was to repeal the saying “Half of the marketing spend is wasted, we just don’t know which half”. As the result of the project we were not able to repeal the saying entirely. We found that in some cases the amount invested in marketing could explain some portion of the value share of a brand, but the amount explained was always relatively low. Still, interesting results in consumer behavior were found based on the preference data.

8.2. Project schedule

In the project planning stage we identified five risks. The statuses of these risks were reviewed in the beginning of each stage. In stage “Data preparation” the risk “Data quality is poor” was realized, which resulted in postponement of the next stage by two weeks. In project planning stage, this risk was identified as the most potential risk. The preventative action was to preserve much time for the stage, but we failed to estimate the workload accordingly. If more time had been available for the modeling phase, which was postponed by the data preparation, it might have been possible to come up with more complex models. However, the modeling was to a bigger extent affected by the quality of the data than mere time problems.

8.3. Project management

The project was managed in a goal oriented way which might have had a bad influence on the results of the project. From the beginning of the project we focused on finding correlations between marketing investments and market value share. Instead we should have had more open ended goals for the project: this way we could have found interesting relations, for example between marketing and the preference data, which could help to explain the effects of different marketing channels better.

8.4. Lessons learned

As a whole, the project was very eye opening and educative for the whole team. However, the most important lessons probably were not related to the contents of the analysis, but rather to the practical side of the project. Already at very early stage we realized that the data processing takes a lot of time and it needs to be done diligently in order to carry out the rest of the project in a reasonable manner. Furthermore, the data quality issues do not only prolong some stages in the project, but also pose considerable risks and interpretation matters on the final results. Also the general expectations about the results and possible hypotheses should not affect the modeling and interpretations in any way – however, staying completely objective is a difficult task.

Finally, considerable amount of the workload was cumulated to the end of the project in spite of careful scheduling in the initial project plan. Although this is a general problem with long projects, it could perhaps be avoided with more intense team work throughout the project. Also more structured communication as well as some sort of monitoring by team mates could have helped to avoid the pile up of the work load as well as decreased the iteration rounds.

8.5. Comments about the course

The concept of the course is excellent. The idea of giving much room to decide and handle the course of the project individually is good, but more counseling from the course staff should be given. It would be good to have regular meetings with the course personnel, for example once a month. The agenda of these meetings would be to get comments about the project and get consultation to the issues at hand from more experienced people. This would also create the course atmosphere a bit more intense: half a year is a good time span to work with a project, but it easily gives freedom to prioritize more urgent tasks over a long term project, which leaves too much pressure on the last weeks.

References

Ackerberg. Empirically distinguishing informative and prestige effects of advertising. RAND Journal of Economics. Vol. 32, No. 2, 2001, pp. 316-333.

American marketing association, Definition of marketing, 2004: http://www.marketingpower.com/AboutAMA/Pages/DefinitionofMarketing.aspx (Accessed 7.2.2013)

Bagwell. The Economic Analysis of Advertising. Handbook of Industrial Organization. Vol. 3. 2007, pp. 1701-1844.

Chintagunta, Erdem, Rossi and Wedel. Structural Modeling in Marketing: Review and Assessment. Marketing Science. Vol. 25, No. 6, November-December 2006, pp. 604-616.

Doraszelski and Markovich. Advertising dynamics and competitive advantage. RAND Journal of Economics. Vol. 38, No. 3, 2007, pp. 557-592.

Dubé, Hitsch and Manchanda, An Empirical Model of Advertising Dynamics. Quantitative Marketing and Economics, vol. 3, 2005, pp 107-144

Horsky and Simon, Advertising and the Diffusion of New Products, Marketing Science, Vol 2, No 1, 1983, pp. 1-17

Mellin. Pääkomponenttianalyysi, 2004: http://www.sal.tkk.fi/vanhat_sivut/Opinnot/Mat-2.112/pdf/PCOMP10.pdf (accessed 6.5.2013)

International Telecommunication Union: Facts and figures, The World in 2013: http://www.itu.int/en/ITU-D/Statistics/Documents/facts/ICTFactsFigures2013.pdf (accessed 25.4.2013)

Return on marketing investment - Wikipedia, the free encyclopedia. 2013. (accessed 25.4.2013)