**Contents**

**Title** 1

**Abstract** 1

**Introduction** 2

Primary Analysis Objective2

**Materials and Methods** 3

Data Source 3

Statistical Analysis 3

Model Assumptions 4

Primary Objective Analysis 4

Model Selection 10

Goodness of Fit 16

**Result 17**

**Title**

To determine the influence of factors affecting the market share of food product.

**Abstract**

For a food product manufacturing firm, It is important to account influencing factors which directly affects the success of product and market share associated to it. Analysis of sales data collected over a period of time, can play pivotal role in enhancing the product and this can also lead to increase in sales as well as profit.

The objective of this analysis is to determine the influencing factors (independent Variables) for a better market share of food product. For this study, Data collected by National Database (Nielsen) over period of 36 consecutive months (From Sep,1999 to Aug,2002) is used as data set.

The Average monthly market share of product (in percentage) is monitored against the input features from data set. Statistical analysis carried out on given data to better understand the influence of factors. Input features from the data set are average monthly price of product in dollars (Price), Gross Nielsen Rating points(gnrpoints), Discount Price(discount), Package promotion(promotion), Month and year.

**Introduction**

A food product can only be successful if the manufacturing firm knows exactly about the influencing factors associated to the market share. Better Understanding of demand and supply towards target audience could lead to product domination in the market. Also, A comprehensive analysis of product data could suggest the better marketing strategies for increasing the market share. For this study, Data collected by National Database (Nielsen) is used. This data has been collected for 36 consecutive months starting from Sept,1999 to end of Aug,2002. The factors taken in account for this study are Market Share, Identification no. for each month, average monthly price of product in dollars(Price), Gross Nielsen Rating points(gnrpoints) which refers to an index of the amount of advertisement exposure that the product received, Discount price(discount) which refers to presence or absence of discount price(1 if discount,0 otherwise), Package promotion(promotion) which refers to presence or absence of package promotion during period(1 if promotion,0 otherwise), Months and year (1999-2002).

In this study, each input feature from the data set has been analyzed against the market share with the help several statistical methods to determine its influence over market share. Finally, a best model was selected on the basis of R^2(proportion of variation in the outcome that is explained by the predictor variables) and CP (A parameter that penalizes inclusion of addition variables to model) for combination of input features.

**Primary Analysis Objectives**

To investigate the linear association between market share (in percentage) and factors accounted in the study for influencing the market share of the food product.

**Materials and Methods**

**Data Source**

The is collected from 1999 to 2002 from a national database (Nielsen) for 36 consecutive months. The data consists of factors influencing the market share of a particular product produced by a large packaged food manufacturer. Each record in the data contains 6 variables i.e. Month (January to December), Year (1999 to 2002), Package Promotion(1 if package promotion is present, 0 otherwise), Discount Price (1 if discount is present, 0 otherwise), Gross Nielson rating points (An index of the amount of advertising exposure that the product received) and Price (Average monthly price of product in dollars) which collectively affect the market share of the product.

**Statistical Analysis**

We have got the data in excel (.xlsx) format which consisted of 36 records. We are going to use multiple linear regression to find the correlation between the market share and the rest of the factors and we have analyzed the data keeping this goal in mind. The data doesn’t have any null values in the preliminary analyses. We have performed linear regression for each of the predictor variable with the target to confirm the individual relationship between them. We have also changed the type of predictor variable Month from character to numerical value so we can fit the model easily.

**Primary Objective Analysis**

We are going to explore the individual predictors and the impact that they can have over the response variable before selecting the best model. It will help us in determining the importance of each factor and also detecting outliers or skewness in data so that we can transform the data accordingly. After exploring the dataset, we will check for linearity between the factors and the target variable individually.