

# Marketing Analysis for an Elevator Company by Using Predictive Analysis

<sup>1</sup>Praveesh.k

*Data Science and Business Systems*  
*SRM Institute of Science and Technology*  
Chennai, India  
pk3801@srmist.edu.in

Dr.M.Sangeetha

*Data Science and Business Systems*  
*SRM Institute of Science and Technology*  
Chennai, India  
sangeetk@srmist.edu.in

**Abstract**—This report summarizes a study of marketing analysis and future price prediction, which margins between the years 2021 to 2022. This provides a cross-sectional prediction of the profit in an elevator company's income. From this study it is seen that based on the increase or decrease in the number of clients/buyers we can provide plans for risk management strategies and technical cost factors which may help significantly for the improvement in the economic stability and profit of the company. It helps in the prediction of the categorizing clients/buyers within the group of long-term investors and short-term investors. Insurance is one of the important features that provides benefits for the client as well as the company. Factors associated with long-term insurance are identified by using exploratory analysis and data preprocessing

**Index Terms**—component, formatting, style, styling, insert

## I. INTRODUCTION

In these present days data analysis has become a part of daily life. This period where there is rapid globalization and development in the sector of technology, the competition within the world of business has become high. In this increasingly competitive business world, the emergence of new companies has become a challenge for the companies that have been around for a long time. Hence these old companies have the possibility of losing those old clients to the new companies. This project uses the data that has been collected by an elevator company pertaining to their products materials which are being used by their elevators. For this project of data analysis, two sets of data values are obtained, which are from one of the company that manufactures the raw materials and the other from the company which uses these raw materials for their elevators. By applying the methods of data preprocessing, finding the missing values, exploratory analysis and in regression these two sets of data are compared. The result which is obtained by the application of these methods of data analysis are R<sup>2</sup> score, OLS Regression Results which is used to obtain the detailed version of regression format, standard errors, kurtosis, prob[JB], and cond No. By the application of data analysis, the companies have the advantage of reducing errors. The cost of production and maintaining their products quality

## II. LITERATURE SURVEY

Factors Related with Promoting Edges at Illinois Grain Lifts, 1982-83: An Exhibition Investigation Sara Helen R. Thompson and Stanley M. Dziura, Jr.

This paper summarizes an investigation of promoting edges in 1982 and 1983 at a cross part of grain lifts in Illinois. Factors related with promoting edges are distinguished by

relapse examination. It is shown that the promoting edge at Illinois grain lifts is extremely delicate to gamble with the board methodologies and specialized cost factors some of which might suggest huge scope economies in grain marketing. Lower marketing edges are related with more prominent limit usage. Expanded limit use might be accomplished by expanding the volume of grain promoted in Illinois, or by lessening the quantity of lifts that product grain. Energy-productive lifts and elevators in Europe: An examination of energy proficiency possibilities and strategy measures

Lifts and elevators stand out enough to be noticed from an energy proficiency point of view previously. To close this hole, this paper examines energy proficiency possibilities and recommends strategy measures for the European lift and elevator market. As a precondition, the electrical energy interest of lifts and elevators in the European market is dissected, in view of master assessments and the consequences of a checking effort. The assessed current interest is contrasted with situations where the execution of most ideal that anyone could hope to find innovation for the applicable gear is expected and the investment funds possibilities thereof are determined. The outcomes demonstrate that impressive specialized proficiency possibilities exist for lifts (over 60

## III. PROPOSED METHODOLOGY

### A. Exploratory Analysis

Exploratory analysis is a scientific methodology which perceives the general examples in the information. The examples incorporate anomalies and elements of the information that may be startling.

The first move toward quite a while investigation is EDA. Comprehension of where anomalies happen and how the factors are connected will assist one in the plan of factual examinations that will with delivering ideal results. If there should arise an occurrence of natural checking information, lo-cales are likely impacted by different stressors. Subsequently, before one endeavours to relate stressor or factor to natural reaction factors doing investigation of stressor relationship as the underlying step is vital. EDA can give bits of knowledge into up-and-comer causes that ought to have been remembered for a causal evaluation.

### B. Scatterplots

Data on connections between sets of factors can be acquired from Scatterplots and Relationship coefficient. Notwithstanding, while dissecting various factors, essential techniques for multivariate representation can give more prominent bits

C. Histogram basically

A histogram basically sums up the circulation of the information by setting perceptions into spans (also called classes or receptacles) and by including the quantity of perceptions in every stretch. The y-hub can address the quantity of perceptions, percent of aggregate, part of aggregate (or likelihood), or thickness (in which the level of the bar duplicated by the width of the span compares to the overall recurrence of the stretch). In view of how the stretches are characterized will decide the presence of the histogram.

D. Boxplot

A smaller rundown of the conveyance of a variable is given by Box plot. A standard boxplot comprises of (1) 25th and 75th percentiles, (2) an even line on the crate at the middle, and (3) vertical lines (stubbles) drawn from each end (quartile) to the outrageous worth. In the event of gentle variety of the standard boxplot, hairs are the ones that stretch out to a ranged distance from the pivot, and the exceptions past the range are distinguished. The estimation of range (S) is:  $S = 1.5 \times (75\text{th percentile} - 25\text{th percentile})$

E. Combined Appropriation Capabilities (CDF)

The combined circulation capability CDF is a capability  $F(X)$  which is the likelihood that the perceptions of a variable are not more prominent than a predefined esteem. The opposite CDF which is likewise often utilized, is the likelihood that the perceptions are more prominent than a predetermined worth. In developing the CDF, loads (e.g., consideration probabilities from a likelihood configuration) are utilized. By utilizing this the likelihood that the worth of the variable in the factual populace is under a predefined worth can be assessed. In the event that forequivalent weighting of perceptions, the CDF applies just to the noticed qualities. A quantile (Q) plot also called likelihood plot, which is a graphical method for portrayal for contrasting a variable with a specific, hypothetical dissemination or to contrast it with the dispersion of another variable. One normal utilization of the Q plot is to check whether a variable is similarly conveyed.

F. Scatterplots

Scatterplots are graphical portrayal of coordinated information plotted with one variable on the even pivot and the other variable plotted on the upward hub. Information are typically plotted in the diagram with the even hub comprising of proportions of a powerful boundary (free factor) and the upward pivot comprising of proportions of a trait that might answer the compelling boundary (subordinate variable). Scatterplots are one of the valuable initial phase in any examination since they help in the representation of connections and recognize potential issues (e.g., exceptions) which can impact resulting factual investigations.

G. Co-relationship Examination

Estimating the covariance of two erratic factors in a matched informational index is known as Relationship examination. The relationship coefficient of two factors X and Y are the ones used to communicate covariance. The relationship coefficient is a unitless number which goes from -1 to +1. The normalized level of co-relationship among X and Y is the size of the connection coefficient. This is the heading of the affiliation, which can be either sure or negative.

H. linear regression

The Straight Relapse is an AI calculation in light of directed learning. It plays out a relapse task. Relapse models an objective expectation esteem in light of free factors. It is generally utilized for figuring out the connection among QTY and Quality feedback full cost with the given data we are able to identify the intercept and slope of the given data.

I. OLS Regression Results

Common Least Squares (OLS) is the most popular of the relapse methods. It is likewise a beginning stage for all spatial relapse investigations. It gives a worldwide model of the variable or interaction you are attempting to comprehend or foresee; it makes a solitary relapse condition to address that cycle. There are various assets to assist you with studying both OLS relapse and Geologically Weighted Relapse. Begin with Relapse investigation rudiments. Then, work through the Relapse Investigation instructional exercise. This point will cover the consequences of your examination to assist you with grasping the result and diagnostics of OLS.

IV. RESULTS AND DISCUSSION

A. Training performance

The data obtained as a result of this project contains information about the raw materials quantity and quality at their manufacturing site and the data of their quantity and quality at their site of delivery which is the elevator manufacturing company. By using an Excel sheet we are able to identify the total value count, material missing, and the grand total.

Count of Claim No	Column Labels		
Row Labels	SI-Delivery ctr / warehousing / Picking	SI-Transp	Grand Total
Material Missing	2	42	44
Material Damages/Retrieval		143	143
Material Wrong	11	1	12
Grand Total	14	186	200
Row Labels	Sum of Quality feedback Full Cost		
SI-Delivery ctr / warehousing / Picking		13586.89	
SI-Transport		1402194.4	
Grand Total		1318081.29	

Fig.1.Excel total

```
In [1]: import numpy as np
import pandas as pd
import seaborn as sns
import matplotlib.pyplot as plt
from sklearn.model_selection import train_test_split
import statsmodels.api as sm
from scipy.stats import norm
from sklearn.linear_model import LinearRegression
from sklearn import preprocessing
import math

In [2]: df=pd.read_excel('OC file OR tracker 2019.xlsx')
```

EXPLORATORY ANALYSIS

```
In [3]: df.head()
```

```
Out[3]:
```

MMAC	FL	Return	Comercio	Presenta	Presenta	Feedback	Quality	COST	Quality	Remarks
	Comments	Request	Action	Action	Action1	Score	Index	Number	Index	
003030	...	NA	NA	NA	NA	NA	3	NA	02135	WIP/Complete

Fig.2.Listofpackagesimported

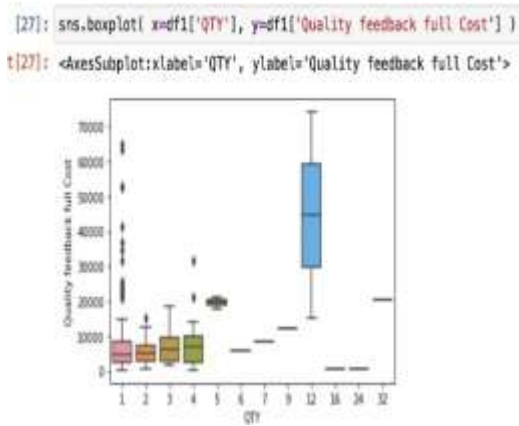


Fig.3.Boxplot

V. CONCLUSION

From our above investigation we can reason that we have figured out how to furnish the clients with the data with respect to which section of raw material stop their resources in to or buy and when and how much returns they can anticipate on selling their finished products. It also helps clients to minimize their errors and to maintain the standard of their products. For this, we have utilized a Different Straight Relapse model

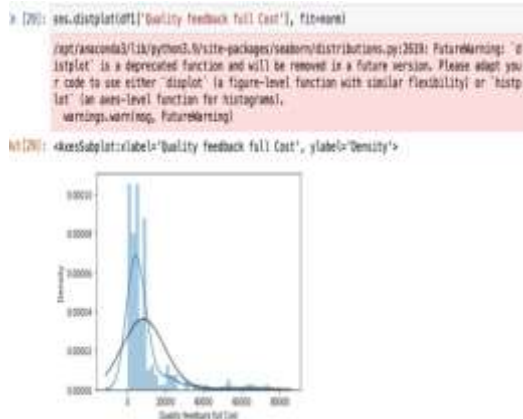


Fig.4.Sea bordistplot

```
In [41]: from sklearn.metrics import r2_score
```

```
In [42]: print(r2_score(y,y_pred))
```

0.016355802195719637

Fig.5.r2score

The given data work show that the 'r2score' of the given data. The r2 scores show the correlation of the data. Lower the values show the better correlation. For example, if the value is 0.5 or 0.7, how the data is better correlation.

```
> [43]: sm.OLS(x,y).fit()
```

```
> [44]: print(ols_summary())
```

OLS Regression Results				
Dep. Variable:	0	R-squared (uncentered):	0.112	0.112
Model:	OLS	Adj. R-squared (uncentered):	0.117	0.118
Method:	Least Squares	F-statistic:	1.57e+09	10.18
Date:	Tue, 02 May 2023	Prob (F-statistic):	1.57e+09	<0.001
Time:	01:20:58	Log-Likelihood:	1060.	1060.
No. Observations:	200	AIC:	1060.	1060.
Df. Residuals:	199	BIC:	1060.	1060.
Df. Model:	1			
Covariance Type:	nonrobust			
	coef	std err	t	P> t
x1	0.0001	1.70e-05	5.371	0.000
				10.425
				0.971
				1.700
				15100.900
				0.00
				1.00

Fig.6.OLS

In this given dataset the by applying OLS model we have obtain R-squared value and obtain the standard errors in the given data.

REFERENCES

- [1] Basu Swastha Irawan, Modern Marketing Management. 2nd Edt. Second Printing. Liberty. Yogyakarta, 2005.
- [2] M. Armstrong, and S. Taylor, S., "Armstrong's Handbook of Human Resource. Management Practice," Kogan Page, 2014.
- [3] H. Sjahruddin, and S. Akbar, "Impact of Promotion, Price, and Product Differentiation in Improving Purchase," 2020.
- [4] Sitharthan, R., Vimal, S., Verma, A., Karthikeyan, M., Dhanabalan, S. S., Prabakaran, N., ... & Eswaran, T. (2023). Smart microgrid with the internet of things for adequate energy management and analysis. Computers and Electrical Engineering, 106, 108556.
- [5] Tjiptono, Fandy, "Marketing Strategy," Third Edition., Yogyakarta CV. Andi Offset, 2008.
- [6] D.C. Luvizon, D. Picard, and H. Tabia, "Multi-task deep learning for real-time 3d human pose estimation and action recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 43, no. 8, pp. 2752-2764, 2020.
- [7] Cao Guoen, Zhang Xin, Deng Ronget. al, "Design of Home-Based Care Interactive Product System Based on SET Analysis Method," Mechanical Design, Dec 2014.
- [8] Cai Wen, "Product Design of Electronic Medicine-Taking Equipment Based on User-Oriented," Mechanical Design, Sept 2015.
- [9] Naz, Saeeda, "Zoning features and 2DLSTM for Urdu text-line recognition," Procedia Computer Science, vol. 96, pp. 16-22, 2016.
- [10] K. Khan, "Urdu text classification using decision trees. High-Capacity Optical Networks and Enabling Emerging Technologies (HONET)," In: 2015 12th International Conference on IEEE, 2015.

- [11] Liao, Minghui. TextBoxes: A Fast Text Detector with a Single DeepNeuralNetwork.AAAI.2017.
- [12] Kanwal, Kehkashan, Assistive Glove for Pakistani Sign Lan-guage Translation. In: Multi-Topic Conference (INMIC), 2014 IEEE17thInternational.IEEE,2014.
- [13] Moshika, A., Thirumaran, M., Natarajan, B., Andal, K., Sambasivam, G., &Manoharan, R. (2021).Vulnerability assessment in heterogeneous web environment using probabilistic arithmetic automata. IEEE Access, 9, 74659-74673.