BSCMS2001 Business Data Management

Capstone Project Proposal

Name: Durga Shree N

Roll No.: 21f2000610

Estimating Uncertainty in Supermarket Sales During Festive and Non-festive Seasons and Proposals for Customer Satisfaction Improvement

Executive summary

The purpose of this project is to help a supermarket deal with a few of its challenges by providing useful insights from everyday sales and purchase data. This origination, a relatively new one in the locality, is looking for strategies to gain popularity and trust from the public in par with other old stores. This project aims to summarize the data and point out a chuck of products that require alteration in service strategies. It is also expected to derive insights on the uniformity of the quantity of products sold and quantify the deviations through variance. Determining the elasticity of quantity of products sold during festive and non-festive seasons supports the shopkeepers with deciding on the right quantity of stocking and the timing for stocking. It is understood from the interaction with the shopkeeper that re-stocking has been done on rough estimations and informal sale pattern observations which has led to understocking and overstocking at times. It's the peak time for the store to not lose customers and start gaining better profits as well and make the right choice about dealers for supplies. Hence, quantifiable outcomes for amending the service strategy and maintenance are expected to be developed from this work.

Background of the organization

Name of the firm: SLV Super Market

Location: No.1, Kaliamman Kovil St, Denkanikottai, Tamil Nadu 635107

Services Provided:

Denkanikotta is a town and the numerous retail shops that run on the main road provide service by picking out groceries as the customers ask for and packing them. None of these grocery stores generate a digitalized bill with GST nor have a price tag on every product with

their QR code. This supermarket is first of its kind in the locality to allow customers to put products into their carts by looking around and bill at the end. Actual price and discounted price are printed on the products to improve the turnover. They initiated the door delivery service during covid and it is still in practice now.

External view of the supermarket

Internal view of the supermarket

Scope of the supermarket:

Items sold at the supermarket:

- Dairy products (milk, curd, paneer, cheese, butter, etc.)
- Toiletries (Shampoos, soaps, toothpaste, bushes, etc.)
- Storage containers (Baskets, boxes, cups, etc.)

- Packed food items (Jam, honey, ketchup, etc.)
- Kitchen supplies (cooking handles, frying pans, etc.)
- Cooking essentials (cereals, pulses, oils, etc.)
- Household supplies (Detergent, disinfectants, etc.)
- Snacks (Biscuits, chocolates, etc.)
- Beverages (coffee and tea powders, flavoured juices, etc.)
- Skincare (Moisturizers, razors, facewash, etc.)
- Baby care products (Diapers, baby powders, cerelac, etc.)

Founders and Organization Structure:

The supermarket is being run by two brothers and the eldest one is a retired Indian Army soldier, Nikhil, whose native is Denkanikotta. The supermarket was opened in 2020. The two brothers take care of billing, stocking, placing orders, etc while a salesboy is hired to help customers navigate through the store.

Problem statement

Objectives:

- To help the shopkeepers identify the fast-moving quantity/size of packets of a product.
- To estimate the variance in sales during normal days and festive seasons.
- To propose strategies for improving customer satisfaction and reducing dead capital on products while ensuring sufficient stocks.
- To determine how early an order for a product should be placed.
- To summarize the value of transactions and type of products made with different dealers.

Background of the problem

Context to the problem:

The supermarket outsources the products from different dealers. Hence, supply happens throughout the week based on different procedures the dealers follow. Some orders for supplies are placed through phone calls, some by visiting the wholesale supplier, and others by contract with those dealers who visit the stores to get orders. Each supplier has a different mode of service and the duration for shipment and delivery changes.

Especially the cooking essentials are purchased in bulk and the supermarket packs them into packets of 50gm/100gm/250gm/500gm/1000gm based on the type of product.

Cooking essentials are generally ordered 1.5 times the normal order capacity during festive seasons.

The issue to be addressed:

The shopkeepers generally make the same number of packets of all quantities. This leaves packets of a certain size to be sold out quickly while others just remain in stock for a long time. Re-packing them also consumes extra manpower and time.

The owner roughly observes a pattern of sales and does re-stocks. The stocks are increased by a certain amount during the festive season. This most often is creating overstocking or understocking. The owners are not informed of quantified uncertainty during normal days and its deviation during the festive season.

Problem-solving approach

1. Methods to solve the problem

Data Pre-processing

The data that the organization might provide could have details of the customer, billing number and other details. It is essential to remove the data irrelevant to the problem statement. The null values are to be corrected and the data should be formatted into tables that can support a structured form for analysis.

Pivot tables

These can be used to organize and summarize the raw data into the required organized format. Tabulating in terms of quantity sold per week/month could give more information than learning about each day as each day will be a very magnified picture for a supermarket dataset. Accumulating the quantities as per requirement is essential before initiating any method to analyze and pivot tables can do the arrangement excellently.

Variance

The quantity of a product sold each day varies. The deviation can increase drastically during the festive season or remain the same. These differences can be captured in terms of *variance*. When the variance of quantity purchased is analysed for a specific duration and if deviations

in variance is seen based on the presence and absence of festive season in that particular period, then those products contributing to the difference in variance can be called elastic in terms of purchase quantity. It is essential to term them as elastic or inelastic in terms of quantity because the products generally purchased from a supermarket are everyday essentials and not luxuries. The increase in purchase quantity could be a resultant of family gatherings, rituals, etc.

Python logic

As mentioned in the problem statement based on the product the quantity purchased varies (Customers prefer 1Kg of rice packet than 1/4Kg and 50gm of pepper powder than 200gm), such inferences can be drawn by automating on every product as it is easier than estimating manually. A function can be defined to identify and separate those products whose purchase quantity varies based on the size of the packet above a threshold. The shopkeepers can plan on concentrating the packet sizing based on observed pattern in data. Dataframes can be used for processing this data.

2. Intended dataset

Sales data for a span of time: Sales data of months which include a festive period and data of purchases performed in the month. The sales details required:

- Product name and quantity of packet
- Price
- Quantity of each product sold in a day
- Date of sale

Purchase data for the same span of time: The purchase details required:

- Product name and quantity of packet
- Distributor
- Quantity of purchase
- Date of purchase

3. Tools used for analysis

Pivot table (Microsoft Excel)

Pivot tables can be used to summarize the whole data and also filters can be applied on attributes to visualize the summary based on only specific values or ranges of an attribute.

Since the data that the supermarket will be able to give is on the everyday sales and purchase, we need pivot tables to organize and summarize data in terms of products being sold or sales for a specific duration (a week/ month).

Pandas dataframe (Python)

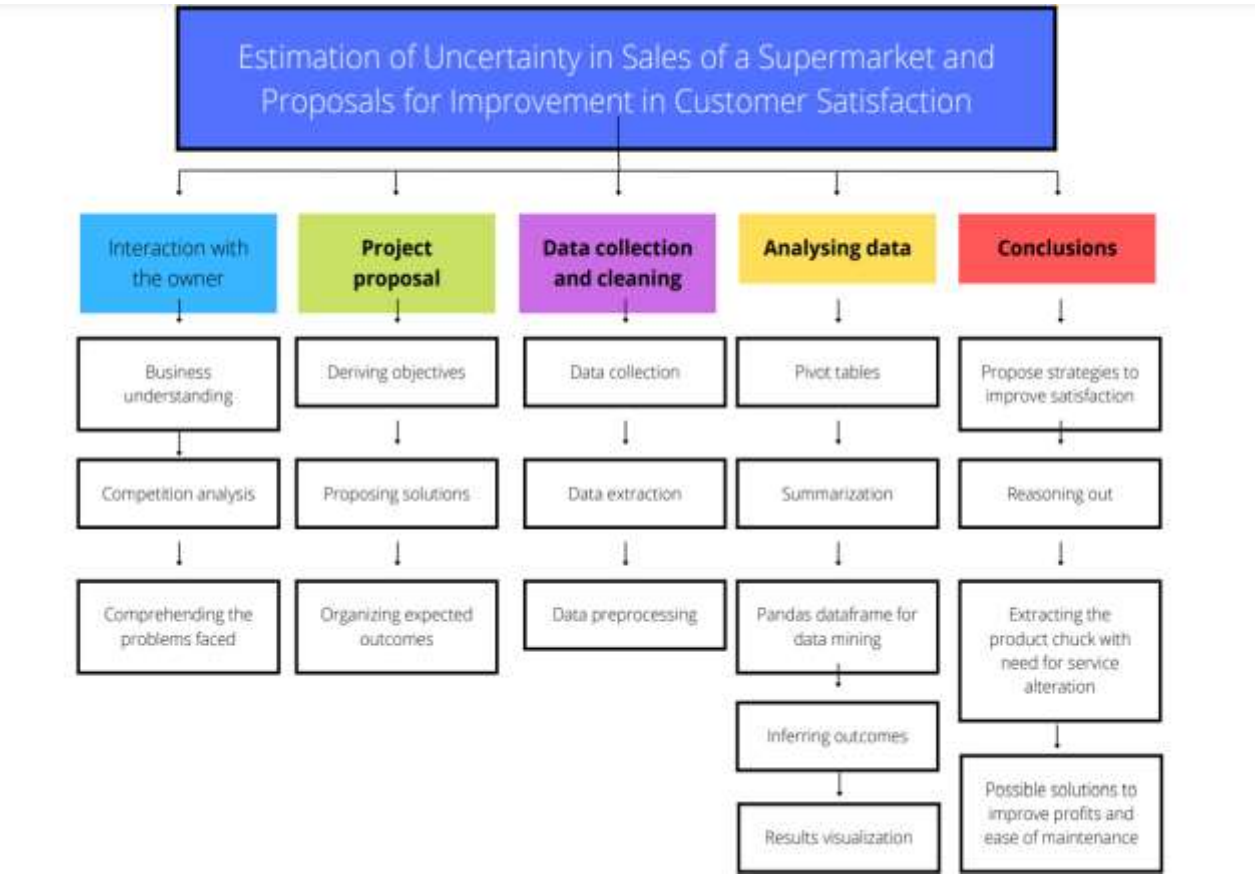
It provides with numerous built-in functions to access and automate processing of tabular data with python logic. This can significantly help in extracting products with analogous features based on thresholds in their features.

Matplotlib (Python)

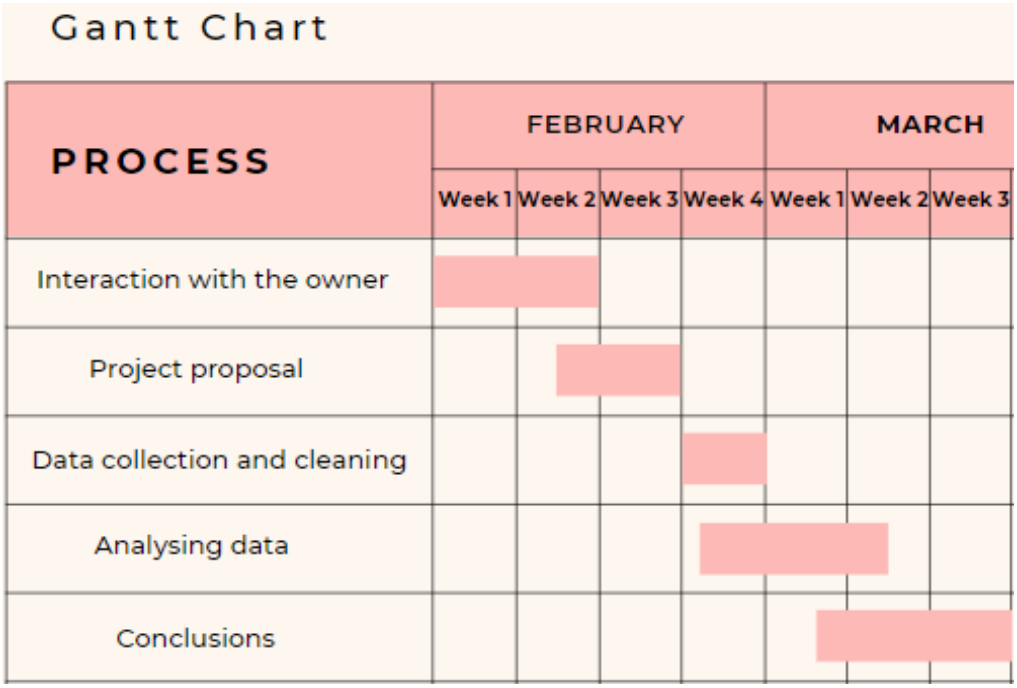
This is a library in Python used for creating charts that are both static and interative. This supports easier visualization by passing the required dataframes used for analysis. These visualizations will play a crucial role in decision-making.

Expected timeline

Work Breakdown Structure



Gantt Chart



Expected outcome

A customer who has come to purchase 200gms of cashew would be fine with purchasing 2 100gms packets but the satisfaction of getting the quantity of their requirement readily will be more than trying to collect packets to satisfy their requirement. If a large number of customers are purchasing 200gms of cashew when compared to 100gms, the organization should concentrate on increasing packets of the requirement rather than making packets of all sizes in equal numbers. *The shopkeeper will be able to prune their internal process of packing using insights from the analysis. The outcome from the analysis will be a specific mention to those products which are preferred in specific quantity packets.*

Losing customers, turnover and profit when there is no stock is essentially as bad as rotten groceries due to heavy stocking. *The analysis of selling rates of products will provide an estimation of how early a product has to be placed order for in a season to avoid overstocking and understocking.*

The supermarket has dealers supplying different types of products. *The project aims to summarize the level, frequency and significance of dealers based on transactions.*