Optimizing the stocks of stationery firm

BUSINESS DATA MANAGEMENT - CAPSTONE PROJECT

Report By
Gagneet Kaur
Roll Number : 21f1006856

Table of Contents

Executive Summary

Sarvjot Communications is a stationery shop which offers all kinds of stationery items required for school or college students. The business problems that could be identified from the discussion with the owner are as follows:

- the problem of reaching a large customer base.
- overstocking or stock outs at different times
- generation of less revenue

The approach followed to do this capstone project is as follows:

Step 1: Define your goals

Step 2: Decide how to measure goals

Step 3: Collect your data

Step 4: Analyze your data

Step 5: Visualize and interpret results

In the project proposal first two steps were completed and the goal chosen was to optimize the stocks of the selected stationary firm. Then the data was collected and pre-processed by entering the data into excel workbook manually.

The raw data collected from the Stationery Firm is described in a detailed manner under the Metadata heading. The data is further explained using descriptive statistics which helped in breaking down the data and infer the story that data can tell.

There are 2 sheets of different category data that has been collected to solve the business problem faced by the Stationery Firm. Sales data tells us about the weekly sales of the products sold by the firm. Revenue data tells us about the total sales amount and the leftover revenue after rent of the shop is paid.

Pictorial representation of data has also been done for visualization and better understanding of the data collected.

With all the useful information, further proceedings can be done to analyze the data using the methods taught in the Business Data Management course to solve the business problem of stockouts and overstocking.

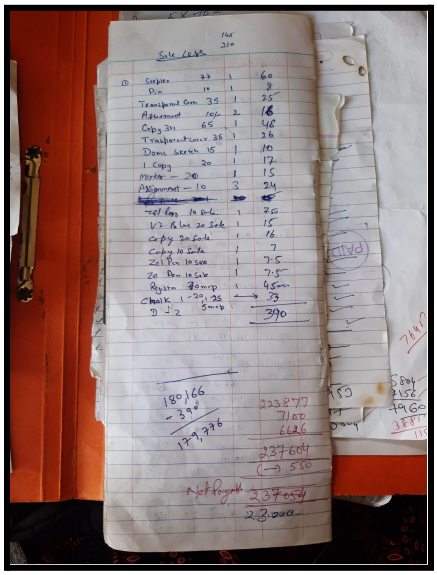
Proof of Originality

Shop Board

Shop Location

Owner of the shop

Primary Data



Video Link : [Interaction with the shop owner]

Metadata

In the workbook, there are 3 worksheets namely:

- **Sales Data :** This sheet holds the data of all the items of the shop, their rate and weekly sale of each item.
- **Revenue Data :** This sheet holds the revenue data across all the three months date-wise
- **Revenue Data (Weekly) :** This sheet holds the revenue data generated weekly

1. Sales data worksheet

Name of the Item	Rate of the Item (in Rupees)	Sale in Week-1 (Dec 5,'22 - Dec 11,'22)	Sale in Week-2 (Dec 12,'22 - Dec 18,'22)	Sale in Week-3 (Dec 19,'22 - Dec 25,'22)	Sale in Week-4 (Dec 26,'22 - Jan 1,'23)	Sale in Week-5 (Jan 2,'23- Jan 8,'23)

Sale in Week-6 (Jan 9,'23- Jan 15,'23)	Sale in Week-7 (Jan 16,'23- Jan 22,'23)	Sale in Week-8 (Jan 23,'23- Jan 29,'23)	Sale in Week-9 (Jan 30,'23- Feb 5,'23)	Sale in Week-10 (Feb 6,'23- Feb 12,'23)	Sale in Week-11 (Feb 13,'23- Feb 19,'23)	Sale in Week-12 (Feb 20,'23- Feb 26,'23)

The metadata of the above mentioned column headers are explained as follows:

- **Name of the Item :** This column contains the names of the items sold by the stationery firm.
- **Rate of the Item (in rupees) :** This column indicates the price of the item.
- **Sale in Week-1 (date) :** The column header has date according to the weeks. This column indicates the quantity of a particular item sold in one week. There are a total 12 columns for the quantity of products sold in weeks 1 to 12. Week-1 basically refers to the week starting from Dec 5, 2022.

2. Revenue data worksheet

Date	Total Sale Amount	Rent	Account	Cash Balance

The metadata of the above mentioned column headers are explained as follows:

- **Date :** This column includes the date from the first week of December to the last week of February.
- **Total Sales Amount :** This column indicates the total amount of revenue generated by the sales of products in a day.
- **Rent :** This column tells the amount of rent to be paid for the shop.
- **Account :** This attribute refers to the amount that was paid online
- **Cash Balance :** This column refers to the amount that was paid by cash.

3. Revenue data (Weekly) worksheet

Week	Total Sale Amount	Rent	Account	Cash Balance

The metadata of the above mentioned column headers are explained as follows:

- **Week :** This column refers to the week number.
- **Total Sales Amount :** This column indicates the total amount of revenue generated by the sales of products in a week.
- **Rent :** This column tells the amount of rent to be paid for the shop.
- **Account :** This attribute refers to the amount that was paid online
- **Cash Balance :** This column refers to the amount that was paid by cash.

* Link to the Project Data : [BDM Project Data]

Descriptive Statistics

Descriptive Statistics				
Name of the Item	TOTAL SALE	MAX SALE	MIN SALE	Revenue Generated (Rate*Total Sale)
Spiral Binding Notebook	30	4	0	₹ 3,300.00
A4 Size Notebook	56	7	2	₹ 5,040.00
King Size Notebook	60	7	3	₹ 3,600.00
General Long Register	47	5	2	₹ 2,820.00
Test Notebook	81	15	1	₹ 810.00
Pen Type-1	400	50	20	₹ 10,000.00
Pen Type-2	176	25	5	₹ 11,440.00
Pencil Pack	108	16	3	₹ 7,560.00
Sketches	77	11	2	₹ 1,155.00
Wax Colours	60	8	3	₹ 600.00
Crayons (cylindrical box)	80	11	2	₹ 4,800.00
Fevicol	124	18	5	₹ 1,240.00
Charts	73	10	2	₹ 730.00
Notebook Cover	31	4	1	₹ 930.00
Pen-Pencil	80	9	3	₹ 1,600.00
Eraser	118	15	4	₹ 590.00
Painting Colours	14	3	0	₹ 280.00
A4-Sheet Bundles	30	3	1	₹ 8,100.00

These statistics were calculated using Excel formulas like MAX, MIN and SUM. These formulas were applied to each item and generated descriptive statistics for the entire dataset. These statistics helped to understand the distribution of sales for each item and identify patterns and trends.

The following are the insights from the above table:

- The most sold item is ‘Pen Type-1’ that is the pen whose amount ranges from Rs.10 to Rs.25.

- The least sold item is ‘Painting Colors’.
- It can be seen that there were weeks when NO ‘Spiral Binding Notebook’ and ‘Painting Colors’ could be sold.
- In the above table, the column ‘Revenue Generated’ shows the total amount of revenue generated by a particular item in the span of all 12 weeks. It is calculated by multiplying the rate of the product with its total sale. It can be seen that the most revenue generating product is ‘Pen Type-2’ which gave Rs.11440 with its sale in a 12 weeks span.

With the use of pictorial representations such as bar charts and pie chart:

- It was found that in Week-10 the maximum number of products were sold generating the least amount of revenue. Week-10 could be considered as the outlier of the dataset.
- The most amount of revenue was generated in Week-12 around Feb 20 to Feb 26
- Week-5 saw the least number of product sales.
- ‘Pen-type 1’ had 24% of its share in the number of products sold in the span of 12 weeks\

Detailed Explanation of Analysis Process/Method

❖ Introduction

- The analysis of the data that has been collected is done through time-series analysis technique because the data has been collected over the span of 12 weeks from the first week of December to the last week of February.

- There are four main types of analyzing data and the following written approach is adopted for analyzing the data collected:
  - ➔ Descriptive analysis
  - ➔ Diagnostic analysis
  - ➔ Predictive analysis
  - ➔ Prescriptive analysis

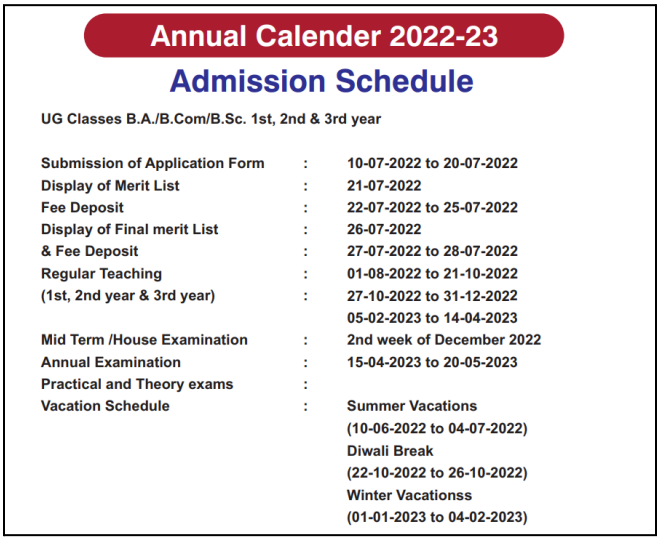
- Descriptive analysis identifies what has already happened. Basic descriptive analysis of the data had already been presented.

❖ Objective

- The main objective of this project is to identify the sales of the stocks based on a particular time period.

- Other than collecting the raw data an effort to understand the customer base was done.

- While discussing with the shop owner it was concluded that 80% of the customers of the shop are students and the rest of the 20% are office people and general working professionals. Focus was to get to know the timelines that affect the larger customer base, i.e, Students.
- This is the annual calendar of the only college in our town. It was taken from the website of the college (gcp.ac.in) to get an idea of the timely needs of students according to their institute calendar.



Annual Calender 2022-23

Admission Schedule

UG Classes B.A./B.Com/B.Sc. 1st, 2nd & 3rd year

Submission of Application Form	:	10-07-2022 to 20-07-2022
Display of Merit List	:	21-07-2022
Fee Deposit	:	22-07-2022 to 25-07-2022
Display of Final merit List	:	26-07-2022
& Fee Deposit	:	27-07-2022 to 28-07-2022
Regular Teaching	:	01-08-2022 to 21-10-2022
(1st, 2nd year & 3rd year)	:	27-10-2022 to 31-12-2022
		05-02-2023 to 14-04-2023
Mid Term /House Examination	:	2nd week of December 2022
Annual Examination	:	15-04-2023 to 20-05-2023
Practical and Theory exams	:	
Vacation Schedule	:	Summer Vacations
		(10-06-2022 to 04-07-2022)
		Diwali Break
		(22-10-2022 to 26-10-2022)
		Winter Vacationss
		(01-01-2023 to 04-02-2023)

Figure 1

- This is the annual calendar of one of the schools in our town. This helped to get information about the general timings of the regular school days, exams and holiday months of all of the schools in the town because the timeline followed is almost the same for every school here.

2. ANNUAL ACADEMIC CALENDAR:

UNIT TEST I	JUNE
TERM I	SEPTEMBER
UNIT TEST II	DECEMBER
ANNUAL EXAM (TERM II)	MARCH

Figure 2

UT I	1st Week of June
T I	September 15 to 30
UT II	Last week of November
T II	March 1 to 20

Figure 3

- These calendars were used to make a general timeline for the stationery firm to keep stocks accordingly.

❖ Pre-Processing of the Data

- Getting the data wasn't an easy task because there was not any sales data written on paper or any digital source. The discussion with the owner was the only means to write the data

manually into excel. The data was entered into the spreadsheet by getting the sales figure from the shop owner orally.

- The data had only the Name of the Product, their rate and the quantity sold weekly.

- The 'Paste Transpose' function of excel was very handy as it helped to get the Week-Wise Sales data by transposing the rows and columns of the primary data. This helped to calculate the sales per-week much more easily.

Weekly Sales	
Week	TOTAL WEEKLY SALES
1	105
2	122
3	116
4	107
5	64
6	136
7	160
8	160
9	161
10	183
11	179
12	152

Table 1

- The revenue data was on papers and had to be entered into excel for further data processing. There was a relation between the 'Total Sale Amount' and the rest of the three variables, i.e., 'Rent', 'Account' and 'Cash Balance'. The formula in the 'figure-4' was used to calculate the 'Cash Balance' from the rest of the three variables. Cash balance basically was the amount of money that was paid through cash by customers.

Date	Day	Total Sale Amount	Rent	Account	Cash Balance
05-12-2022	Monday	₹ 1,520.00	₹ 300.00	₹ 60.00	=C2-D2-E2

Figure 4

- The analysis was to be done on a weekly basis. Revenue Data was given on the daily basis which was further cleaned and processed and converted into weekly revenue data by using formulas and pivot tables.

- Using the formula shown in 'figure-5', the week of the particular date was identified.

Date	Day	Total Sale Amount	Rent	Account	Cash Balance	Week
05-12-2022	Monday	₹ 1,520.00	₹ 300.00	₹ 60.00	₹ 1,160.00	=YEAR(A2)&"-"&TEXT(WEEKNUM(A2,2),"00")
06-12-2022	Tuesday	₹ 3,450.00	₹ 300.00	₹ 1,790.00	₹ 1,360.00	2022-50

Figure 5

- Pivot table was further used to transform the daily data into the weekly data. Since using the above formula, the last week of December and first week of January were treated separately, they had to be merged together and then a final table was formed. The table is shown below.

Week	Total Sale Amount	Rent	Account	Cash Balance
1	10375	2100	3005	5270
2	10006	2100	2816	5090
3	12191	2100	6491	3600
4	6365	1500	1795	3070
5	8498	2100	2978	3420
6	7417	1800	1997	3620
7	8532	2100	3232	3200
8	8316	2100	2766	3450
9	6977	1800	1677	3500
10	5976	1500	966	3510
11	8336	2100	2126	4110
12	14044	2100	2454	9490

Table 2

❖ Descriptive Analysis

- With the help of Excel, different formulas could be used to get the Summary Statistics of the data which was extremely helpful in the analysis process.
- Further to understand the data precisely and easily, pictorial representations were used.
- To know the share of each item in the sales, a pie chart was used. It was filtered to Top 5 products with respect to their sales.
- Another pie chart was generated to check the best 4 products wrt their share in the revenue.
- Line charts helped a lot to see the trend across the days of the week.
- Column charts were extremely useful in showing data changes over the period of time and for illustrating comparisons among items.

- The descriptive statistics were calculated using Excel formulas like MAX, MIN and SUM. These formulas were applied to each item and generated descriptive statistics for the entire dataset.

Results and Findings

The insights drawn from the analysis of the data are the following:

➢ To find out the top 4 products who had the most sales in the span of 12 weeks, SORT and FILTER function was used to create the 'Table-3'.

Best 4 Products and Others (wrt Sale)	
Name of the Item	TOTAL SALE
Pen Type-1	400
Pen Type-2	176
Fevicol	124
Eraser	118
Others	827

Table-3

- Visualizing data gives more insights. Thus the following chart was generated out of the above table.

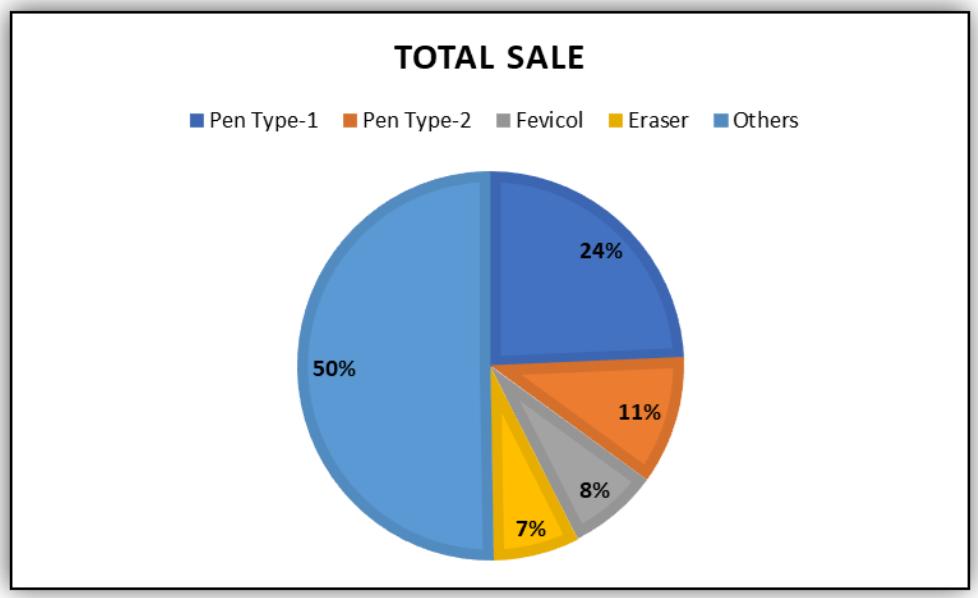


CHART-1 – PIE CHART (Product Sales Share)

- Pie charts are used to represent the proportional data or relative data in a single chart.

- Chart-1 basically shows the 'Share' of Quantity of a product sold across all the twelve weeks.

- From the above chart it can be clearly seen that ‘PEN TYPE 1’ is the most sold product. This chart helped us to identify the most favorable item.

➢ On the basis of revenue, the best 4 products were filtered out. The same can be seen in ‘Table-4’

Best 4 Products and Others (wrt Revenue)	
Items	**Revenue Generated**
Pen Type-2	11440
Pen Type-1	10000
A4-Sheet Bundles	8100
Pencil Pack	7560
Others	27495

Table-4

- To find out which product contributed to the most share of the total revenue the following chart was generated out of the ‘Table-3’.

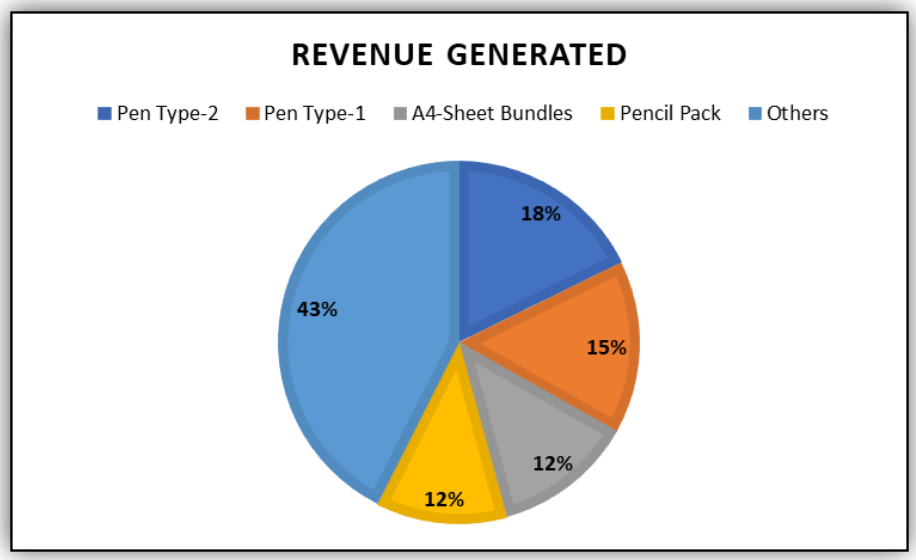


CHART-2 – PIE CHART (Share of Revenue)

- Chart-2 represents the ‘Share’ of Revenue generated by the best four products and the rest of the products across all the twelve weeks.

- From the above chart it can be clearly seen that ‘PEN TYPE 2’ contributed the most to the revenue generated in 3 months . This chart helped us to identify the most profitable product.

- From ‘Chart-1’ and ‘Chart-2’ it can be concluded that ‘PEN TYPE-1’ and ‘PEN TYPE-2’ dominate the overall sales and revenue generation of the shop.

➢ Weekly Revenue Generation can also help us in a better way to check the profitability across the weeks. The following chart - 'CHART-4' was prepared for the same.

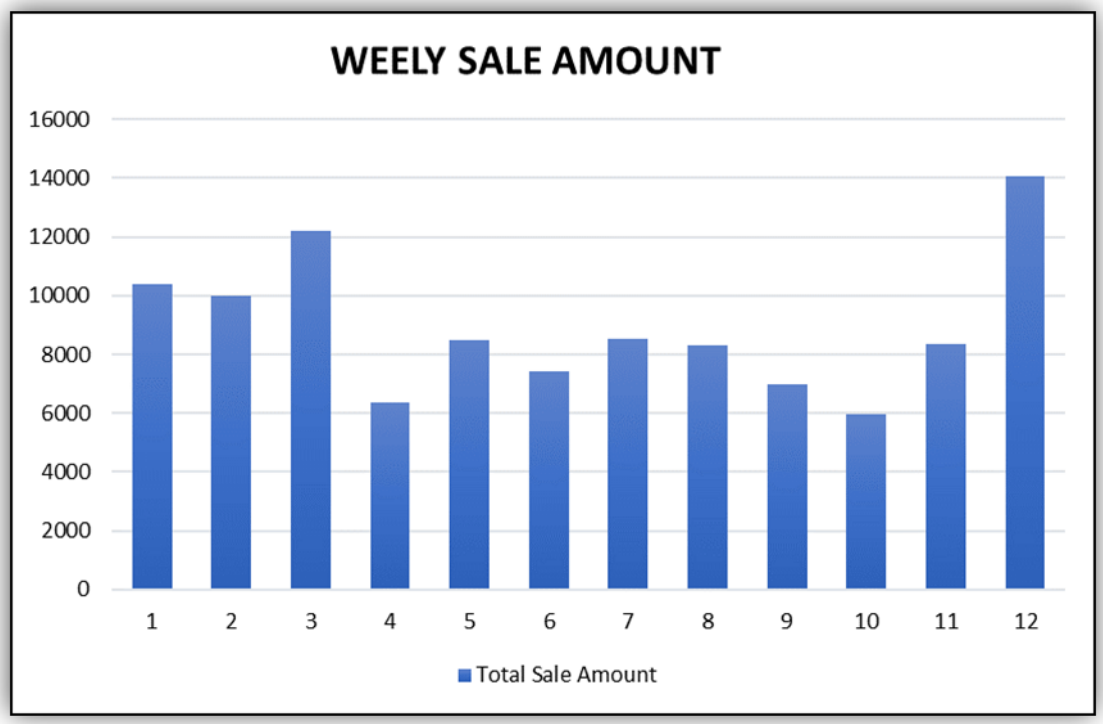


CHART-4 – COLUMN CHART (Weekly Revenue of all products altogether)

- Chart-4 is the column representation of the Weekly Revenue generated from the sales of products.
- Column charts work well in showing data changes over a period of time by displaying the comparisons among subjects on an overall chart. They are often used to show data comparisons in a visual way.
- Here it can be seen that Week-10 generated the least amount of revenue.
- Week-12 generated the most amount of revenue.
- Although Week-5 saw the least sales as represented by 'CHART-3' but from 'CHART-4' we can conclude that with regards to revenue generation Week-5 did a lot better than a lot of weeks. This means that high price items were sold during the 5th week, i.e, the first week of January.