



BUSINESS INTELLIGENCE SOLUTION DEVELOPMENT REPORT

**MIS271: BUSINESS INTELLIGENCE &
DATA WAREHOUSING**

TABLE OF CONTENTS

<u>1. INTRODUCTION</u>	<u>3</u>
<u>2. DATA SET CLEANSING AND STANDARDISATION</u>	<u>4</u>
<u>2.1 DISCUSSION AND JUSTIFICATION</u>	<u>4</u>
<u>3. ANALYSIS AND REPORTING</u>	<u>10</u>
<u>3.1 EXCEL ANALYSIS</u>	<u>10</u>
3.1.1 PRODUCT ANALYSIS	10
3.1.2 DESCRIPTIVE ANALYSIS	13
<u>3.2 POWER BI ANALYSIS PART 1.....</u>	<u>16</u>
<u>3.2.1 POWER BI ANALYSIS PART 2.....</u>	<u>18</u>
<u>4. DASHBOARD: REGIONAL MANAGER.....</u>	<u>21</u>
<u>4.2 DASHBOARD : SALES MANAGER.....</u>	<u>25</u>
<u>5. JUSTIFICATIONS AND RECOMMENDATIONS</u>	<u>28</u>
<u>REFERENCES.....</u>	<u>32</u>

1. INTRODUCTION

The purpose of this Business Intelligence report is to decipher insights relating to sales/order trends and customer behaviour for the management team of Superstores, a retail business that specializes in selling products typically found in offices such as furniture, technology, and office supplies. The company is based in the USA and has a notable presence in many of the states within it.

Superstore has provided Group 7, a team of graduate business analysts, with a dataset containing parameters of the many different facets of the business, such as sales/order data and customer information so that the team can analyse the dataset, design BI solutions for managerial use and make a series of justifications and recommendations for the Superstores managers. There are four key sections of the report below: Cleansing and Standardizing the dataset, analysis and reporting using Microsoft Excel and Power BI, dashboard development and our justifications and recommendations for the business.

Firstly, the dataset had to be cleansed before it was used for our analysis. According to [Erhard Rahm and Hong Hai Do \(2000\)](#), the purpose of data cleaning is to detect and remove errors and inconsistencies from the dataset, so that our analysis will not lead to misleading information. The team performed this function on MS Excel before loading the data onto Power BI and significantly improved the dataset's quality provided by Superstores in the process (*refer to the data cleansing section in our report for further information*).

Next our team performed an analysis of the dataset using Microsoft Excel and Power BI, by creating visualizations and dashboards. [Dr Shekhar Pawar \(2016\)](#) states that the pros of Microsoft Excel include having an ideal interface for making data tables, contains well documented formulas and syntax and is therefore understandable by users in every industry, whereas [Nancy J \(2023\)](#) states that one of the main advantages of Power BI is that it enables more effective sharing and collaboration. Furthermore, dashboards, which are a selected group of visuals that can provide quick insight to the data, can be easily created using Power BI (*refer to the Analysis and Dashboard section below for more detail*).

Lastly, a section dedicated to key justifications and recommendations based on our analysis of the Superstores dataset has been provided so that managers of the Superstores can act on this information to improve the position of the company. This report addresses all four assessment requirements: data cleaning and standardisation, Excel-based analysis, Power BI reporting, dashboard development, and managerial recommendations.

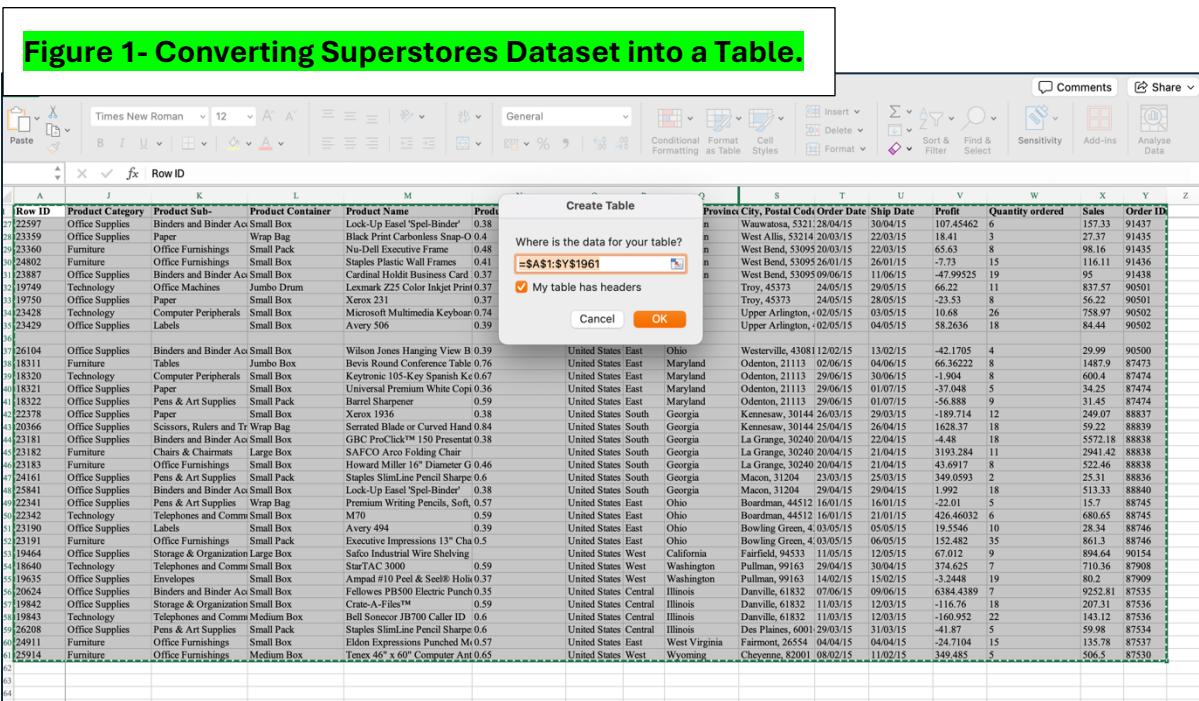
2. DATA SET CLEANSING AND STANDARDISATION

The dataset had to be cleansed on Microsoft Excel before it was used for our analysis. The purpose of data cleaning is to detect/remove errors and inconsistencies from the dataset, so that its value can be maximized, and so that our analysis will not lead to misleading information **Erhard Rahm and Hong Hai Do (2000)**. The team performed this function before loading the data onto Power BI and improved the quality of the dataset initially provided by Superstores.

2.1 DISCUSSION AND JUSTIFICATION

2.1.1 Converting data into a Table

Figure 1- Converting Superstores Dataset into a Table.



Row ID	Product Category	Product Sub-Category	Product Container	Product Name	Product	Province	City	Postal Code	Order Date	Ship Date	Profit	Quantity ordered	Sales	Order ID	
27	22597	Office Supplies	Binders and Binder Acs	Small Box	Lock-Up Easel Spel-Binder	0.38	Wauwatosa	5321	28/04/15	30/04/15	107,5462	6	157.33	91437	
28	23359	Office Supplies	Paper	Wrap Bag	Black Print Carbonless Snap-O 0.4	0.48	West Allis	5324	20/03/15	22/03/15	18.41	3	27.37	91435	
29	23360	Furniture	Office Furnishings	Small Pack	Nu-Dell Executive Frame	0.48	West Bend	53095	20/03/15	22/03/15	65.63	8	98.16	91435	
30	24040	Furniture	Office Furnishings	Small Box	Staples Plastic Wall Frames	0.37	West Bend	53095	09/06/15	20/06/15	-7.95	15	116.11	91438	
31	23887	Office Supplies	Binders and Binder Acs	Small Box	Cardinal Multi Business Card	0.37	West Bend	53095	09/06/15	11/06/15	47.99525	19	95	91438	
32	19749	Technology	Office Machines	Jumbo Drum	Lexmark Z23 Color Inkjet Print 0.37	0.37	Troy	45373	24/05/15	29/05/15	66.22	11	873.57	90501	
33	19750	Office Supplies	Paper	Small Box	Xerox 231	0.37	Troy	45373	24/05/15	28/05/15	-23.53	8	56.22	90501	
34	23428	Technology	Computer Peripherals	Small Box	Microsoft Multimedia Keybou 0.74	0.74	Upper Arlington	4021	02/05/15	03/05/15	10.68	26	758.97	90502	
35	23429	Office Supplies	Labels	Small Box	Avery 506	0.39	Upper Arlington	4021	02/05/15	04/05/15	58.2636	18	84.44	90502	
36	26104	Office Supplies	Binders and Binder Acs	Small Box	Wilson Jones Hanging View B 0.39	0.39	Westerville	43081	22/02/15	13/02/15	-42.1705	4	29.99	90500	
37	18311	Furniture	Tables	Jumbo Box	Bevis Round Conference Table 0.76	0.76	United States	East	Ohio	Westerville	43081	22/02/15	13/02/15	1487.9	87473
38	18320	Technology	Computer Peripherals	Small Box	Keytron 105-Key Spanish K 0.67	0.67	United States	East	Maryland	Odenton	21113	29/06/15	30/06/15	600.4	87474
39	18332	Office Supplies	Paper	Small Box	Universal Premium White Copy 0.57	0.57	United States	East	Maryland	Odenton	21113	29/06/15	30/06/15	342.22	87474
40	18322	Office Supplies	Pens & Art Supplies	Small Pack	Bartel Sharpener	0.59	United States	East	Maryland	Odenton	21113	29/06/15	31/07/15	-54.888	5
41	22378	Office Supplies	Paper	Small Box	Xerox 1936	0.38	United States	South	Georgia	Kennesaw	30144	26/03/15	29/03/15	-189.714	12
42	20266	Office Supplies	Scissors, Rulers and Tr Wrapping Bag	Small Box	Serrated Blade or Curved Hard 0.84	0.84	United States	South	Georgia	Kennesaw	30144	25/04/15	26/04/15	162.37	18
43	23181	Office Supplies	Binders and Binder Acs	Small Box	GBC PreClick™ 150 Presentat 0.38	0.38	United States	South	Georgia	La Grange	30240	20/04/15	22/04/15	-4.48	18
44	23182	Furniture	Chairs & Chairs	Large Box	SACFO Aroo Folding Chair	0.57	United States	South	Georgia	La Grange	30240	20/04/15	21/04/15	319.284	11
45	23183	Furniture	Office Furnishings	Small Box	Howard Miller 16" Diameter G 0.46	0.46	United States	South	Georgia	La Grange	30240	20/04/15	21/04/15	43.6917	8
46	24161	Office Supplies	Pens & Art Supplies	Small Pack	Staples SlimLine Pencil Sharpe 0.6	0.6	United States	South	Georgia	Macon	31204	23/03/15	25/03/15	349.593	2
47	25841	Office Supplies	Binders and Binder Acs	Small Box	Lock-Up Easel Spel-Binder	0.38	United States	South	Georgia	Macon	31204	29/04/15	29/04/15	1.992	18
48	22341	Office Supplies	Pens & Art Supplies	Wrap Bag	Premium Writing Pencils Soft, 0.57	0.57	United States	East	Ohio	Boardman	44512	16/01/15	16/01/15	-23.201	5
49	20252	Technology	Telephones and Comm Devices	Small Box	MTO - 1000	0.59	United States	East	Ohio	Bowling Green	45373	03/05/15	2/05/15	42.56012	6
50	23190	Office Supplies	Labels	Small Box	Avery 494	0.39	United States	West	Ohio	Bowling Green	45373	03/05/15	10/5/15	19.5546	10
51	23191	Furniture	Office Furnishings	Small Pack	Executive Impressions 13" Chair	0.5	United States	East	Ohio	Bowling Green	45373	03/05/15	10/5/15	152.482	35
52	19464	Office Supplies	Storage & Organization	Large Box	Safos Industrial Wire Shelving	0.57	United States	West	California	Fairfield	94533	11/05/15	12/05/15	67.012	9
53	18640	Technology	Telephones and Comm Medium Box	Small Box	StarTAC 3000	0.59	United States	West	Washington	Pullman	99163	29/03/15	30/04/15	374.625	7
54	19635	Office Supplies	Envelopes	Small Box	Ampad #10 Peel & Seal® Holo 0.37	0.37	United States	West	Washington	Pullman	99163	14/02/15	15/02/15	-3.2448	19
55	20624	Office Supplies	Binders and Binder Acs	Small Box	Fellowes PB500 Electric Punch 0.35	0.35	United States	Central	Illinois	Danville	61832	07/06/15	09/06/15	6384.4389	7
56	19782	Office Supplies	Storage & Organization	Small Box	Crate-A-Files™	0.59	United States	Central	Illinois	Danville	61832	11/03/15	12/03/15	-116.76	18
57	19843	Technology	Telephones and Comm Medium Box	Small Box	Bell Soncor JB700 Caller ID	0.6	United States	Central	Illinois	Danville	61832	11/03/15	12/03/15	-160.952	22
58	26208	Office Supplies	Pens & Art Supplies	Small Pack	Staples SlimLine Pencil Sharp 0.6	0.6	United States	Central	Illinois	De Plaines	60001	29/03/15	31/03/15	-41.87	5
59	24911	Furniture	Office Furnishings	Small Box	Eldon Expressions Punched M 0.57	0.57	United States	East	West Virginia	Fairmont	26554	04/04/15	04/04/15	-24.7104	15
60	25914	Furniture	Office Furnishings	Medium Box	Texex 46" x 60" Computer Amt	0.65	United States	West	Wyoming	Cheyenne	82001	08/02/15	11/02/15	349.485	5
61													306.5	87530	

According to **Muhammad Sheheryar (2023)**, table formatting in Excel makes it easier to manage large sets of data. Excel automatically applies formatting features to the table (**Figure-1**), this structured format enhances the visual clarity of the dataset, facilitating quick identification of trends, patterns, and outliers. This method streamlines the analysis and optimize consistency when imported to PowerBI.

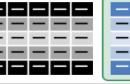
1747	19739	Medium	0	137.48	32.18	3076	Peter Hardy	Delivery Truck
1748	23816	Medium	0.07	300.97	7.18	3077	Lynne Reid	Regular Air
1749	25489	Not Specified	0.04	35.44	5.09	3078	Kate McKenna	Regular Air
1750	25490	Not Specified	0.08	3.98	0.7	3078	Kate McKenna	Regular Air
1751	5816	Medium	0.07	300.97	7.18	3079	Andrew Levine	Regular Air
1753	7489	Not Specified	0.04	35.44	5.09	3079	Andrew Levine	Regular Air
1754	7490	Not Specified	0.08	3.98	0.7	3079	Andrew Levine	Regular Air
1755	7491	Not Specified	0.01	1.76	0.7	3079	Andrew Levine	Regular Air
1756	7492	Not Specified	0.01	193.17	19.99	3079	Andrew Levine	Express Air
1757	1739	Medium	0	137.48	32.18	3079	Andrew Levine	Delivery Truck
1758	6807	Critical	0	2.21	1	3079	Andrew Levine	Express Air
1759	19756	High	0	65.99	5.99	3084	Debbie Hsu	Express Air
1760	20589	Not Specified	0.01	7.1	6.05	3084	Debbie Hsu	Regular Air
1761	20590	Not Specified	0.05	18.97	9.03	3084	Debbie Hsu	Regular Air
1762	20008	High	0.05	39.99	10.25	3086	Ted Durham	Express Air
1763	21085	Low	0.07	49.43	19.99	3089	Sandy Cannon	Regular Air
1764	20357	Critical	0.09	207.48	0.99	3095	Milton Lindsay	Regular Air

2.1.2

Removing hidden rows and columns

Figure 2- Removing hidden rows in the dataset.

Header Row First Column Filter Button
 Total Row Last Column
 Banded Rows Banded Columns



N	O	P	Q	S	T	U
0.54	United States	West	Washington	Anacortes, 98221	07/01/15	08/01/15
0.6	United States	West	California	San Gabriel, 91776	13/06/15	15/06/15
0.45	United States	East	New Jersey	Roselle, 7203	15/02/15	17/02/15
0.43	United States	Central	Minnesota	Prior Lake, 55372	12/05/15	14/05/15
0.56	United States	Central	Minnesota	Prior Lake, 55372	12/05/15	13/05/15
0.56	United States	Central	Minnesota	Prior Lake, 55372	12/05/15	13/05/15
0.36	United States	Central	Minnesota	Prior Lake, 55372	12/05/15	13/05/15
0.38	United States	East	New York	Smithtown, 11787	08/04/15	09/04/15
0.4	United States	East	New York	Smithtown, 11787	28/05/15	28/05/15
0.39	United States	East	New York	Syracuse, 13210	12/02/15	15/02/15
0.55	United States	East	New York	Syracuse, 13210	12/02/15	14/02/15
0.49	United States	West	Montana	Helena, 59601	15/05/15	16/05/15
0.39	United States	West	Montana	Missoula, 59801	21/05/15	23/05/15
0.49	United States	East	New York	New York City, 10012	15/05/15	16/05/15
0.37	United States	East	New York	New York City, 10012	21/05/15	22/05/15
0.39	United States	East	New York	New York City, 10012	21/05/15	23/05/15
0.37	United States	West	California	Laguna Niguel, 92677	28/01/15	29/01/15
0.50	United States	West	California	Laguna Niguel, 92677	28/01/15	29/01/15

Figure 3- Removing hidden columns in the dataset.

The dataset was then scanned to detect any empty columns or rows, and these were then unhidden. These blank columns/rows added no value and disrupted the dataset's integrity, and hence were removed. One hidden Row (**Figure 2**) as well as One hidden column (**Figure-3**) were found in the dataset and were then removed.

2.1.3 Deleting empty rows.

Figure 4- Deleting empty rows in the dataset.

Some rows in the dataset were not hidden but were simply empty. Similarly, these empty rows (**Figure-4**) added no value to the dataset and had to be deleted. Some of the techniques like forward fill, average analysis or Meta data referencing can be helpful in filling these empty rows but they are only effective if a certain column is empty, and since there were no existing data points empty, we decided to remove them to maintain the quality of the dataset.

2.1.4 Removing duplicate data.

Figure 5- Removing Duplicate Values from the Dataset.

Excel screenshot showing the 'Remove Duplicates' dialog box. The dialog box lists columns: Row ID, Order Priority, Discount, Unit Price, Shipping Cost, and Customer ID. The checkbox 'My list has headers' is checked. Buttons 'Cancel' and 'OK' are visible.

Identifying and removing duplicates from the dataset is an essential to maintain the quality, reliability, and accuracy of the analysis. Failure to address duplicates can cause errors in analysis. To subtract the risk, we utilised the '**Remove Duplicates**' feature of Ms Excel (Figure-5).

Our approach was to remove the duplicates based on unique keys like Customer ID or Order ID assuming they will be unique. Digging deep into the data we got to know that these unique identifiers are duplicates so we decided to remove the rows which are fully duplicated from every column. We found out that 4 rows were completely duplicated.

2.1.5 Separating city and postal codes

Figure 6- Separating City and Postal Codes using Text To Columns.

Excel screenshot showing the 'Text To Columns' wizard, Step 2 of 3. The 'Delimiters' section is selected, with 'Comma' chosen as the delimiter. A preview window shows the data with the city and postal code separated. Buttons 'Cancel', 'Back', 'Next >', and 'Finish' are visible.

In this instance we encountered a scenario that one column contained overwhelming amount of information. The column had a combination of both cities and postal code values, so we had to separate these values into different columns. To achieve this, we used an excel feature called **“Text to Columns” (Figure – 6)**. This feature helped us separate the values using commas and this improved the readability and structure of the dataset.

2.1.6 Standardization by adding Dollar signs

Figure 7- Adding Dollar signs on Excel.

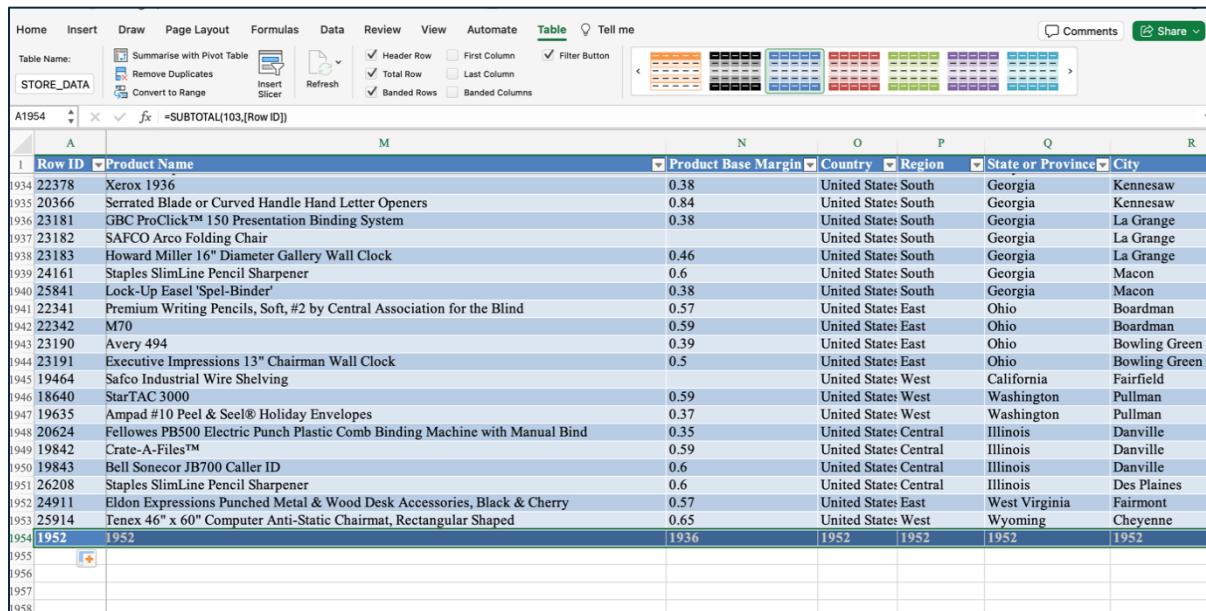
The screenshot shows a Microsoft Excel spreadsheet with data in columns V, W, X, and Y. The 'Format Cells' dialog box is open, specifically for the 'Number' category. The 'Symbol' dropdown is set to '\$ English (United States)', with the '\$' symbol highlighted. The 'Sample' cell shows '\$4.56'. The 'Decimal places' dropdown is set to 2. The 'Negative numbers' dropdown shows examples: '\$-1,234.10', '\$1,234.10', '(\$1,234.10)', and '(\$1,234.10)' in red. The 'OK' button is visible at the bottom right of the dialog box.

As some of the columns having numbers and currencies in the dataset did not have the sign assigned to them so we took the initiative to enhance the readability and scalability of the dataset. Since the store dataset was of USA, It was assumed all the transactions were done in USD so with the help of ‘Format Cells’ feature we added a dollar sign to standardize the columns and get them onto same scale (Figure – 7).

This practice aligns with presentation of the dataset for the stakeholders and such attention to detail reflects commitment towards extracting meaningful insights

2.1.7 Deleting the Product Base Margin Column

Figure 8- Deleting the Product Margin Column.



The screenshot shows a Microsoft Excel spreadsheet titled 'STORE_DATA'. The table has columns: Row ID, Product Name, Product Base Margin, Country, Region, State or Province, and City. The 'Product Base Margin' column contains 16 blank entries. The table is sorted by Row ID. The 'Product Name' column includes items like 'Xerox 1936', 'Serrated Blade or Curved Handle Hand Letter Openers', 'GBC ProClick™ 150 Presentation Binding System', and 'SAFCO Arco Folding Chair'. The 'City' column includes 'Kennesaw', 'La Grange', 'Macon', 'Boardman', 'Ohio', 'Bowling Green', 'Fairfield', 'Pullman', 'Washington', 'Illinois', 'Danville', 'Des Plaines', 'Fairmont', 'Cheyenne', and 'Wyoming'. The 'State or Province' column includes 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Georgia', 'Ohio', 'Ohio', 'Ohio', 'Ohio', 'Illinois', 'Illinois', 'Illinois', 'Illinois', 'Illinois', 'Illinois', 'Illinois', 'Illinois', 'West Virginia', and 'Wyoming'.

Row ID	Product Name	Product Base Margin	Country	Region	State or Province	City
1934	22378 Xerox 1936	0.38	United States	South	Georgia	Kennesaw
1935	20366 Serrated Blade or Curved Handle Hand Letter Openers	0.84	United States	South	Georgia	Kennesaw
1936	23181 GBC ProClick™ 150 Presentation Binding System	0.38	United States	South	Georgia	La Grange
1937	23182 SAFCO Arco Folding Chair		United States	South	Georgia	La Grange
1938	23183 Howard Miller 16" Diameter Gallery Wall Clock	0.46	United States	South	Georgia	La Grange
1939	24161 Staples SlimLine Pencil Sharpener	0.6	United States	South	Georgia	Macon
1940	25841 Lock-Up Easel 'Spel-Binder'	0.38	United States	South	Georgia	Macon
1941	22341 Premium Writing Pencils, Soft, #2 by Central Association for the Blind	0.57	United States	East	Ohio	Boardman
1942	22342 M70	0.59	United States	East	Ohio	Boardman
1943	23190 Avery 494	0.39	United States	East	Ohio	Bowling Green
1944	23191 Executive Impressions 13" Chairman Wall Clock	0.5	United States	East	Ohio	Bowling Green
1945	19464 Safo Industrial Wire Shelving		United States	West	California	Fairfield
1946	18640 StarTAC 3000	0.59	United States	West	Washington	Pullman
1947	19635 Ampad #10 Peel & Seal® Holiday Envelopes	0.37	United States	West	Washington	Pullman
1948	20624 Fellowes PB500 Electric Punch Plastic Comb Binding Machine with Manual Bind	0.35	United States	Central	Illinois	Danville
1949	19842 Crate-A-Files™	0.59	United States	Central	Illinois	Danville
1950	19843 Bell Sonecor JB700 Caller ID	0.6	United States	Central	Illinois	Danville
1951	26208 Staples SlimLine Pencil Sharpener	0.6	United States	Central	Illinois	Des Plaines
1952	24911 Eldon Expressions Punched Metal & Wood Desk Accessories, Black & Cherry	0.57	United States	East	West Virginia	Fairmont
1953	25914 Tenex 46" x 60" Computer Anti-Static Chairmat, Rectangular Shaped	0.65	United States	West	Wyoming	Cheyenne
1954	1952	1952	1936	1952	1952	1952
1955						
1956						
1957						
1958						

Our team meticulously scrutinized the dataset and identified 16 scattered blank entries within a Profit base margin column (**Figure – 8**). Furthermore, as this column was not defined in the Read Me section accompanying the dataset, its relevance and significance was unclear. Given the absence of documentation and the presence of missing values, the column was excluded from downstream analysis to prevent misinterpretation. By removing this column, we further polished the dataset before importing the dataset into Power BI.

3. ANALYSIS AND REPORTING

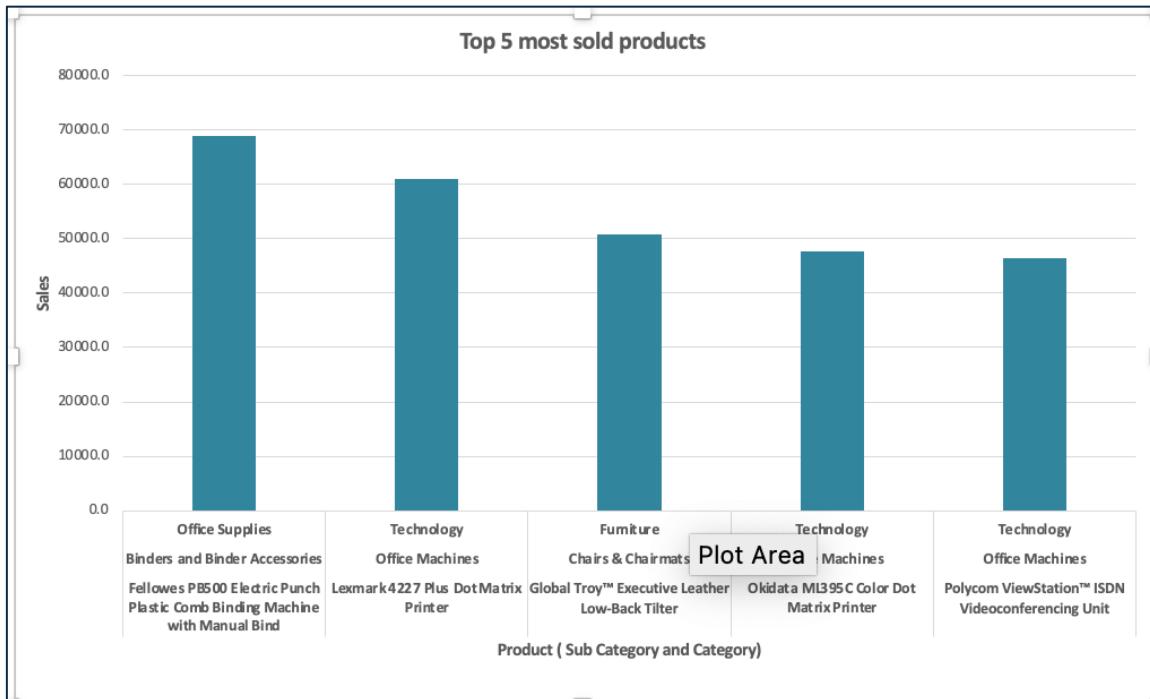
3.1 EXCEL ANALYSIS

By using Microsoft Excel, the dataset was analysed, and key findings were presented. Main analysis was done on several variables like Mode of transport, Products, Product Category, Profits etc. Our analysis provided a valuable insight into various aspects of business operations.

This section focuses on product-level performance to distinguish between high-revenue and high-volume items

3.1.1 PRODUCT ANALYSIS

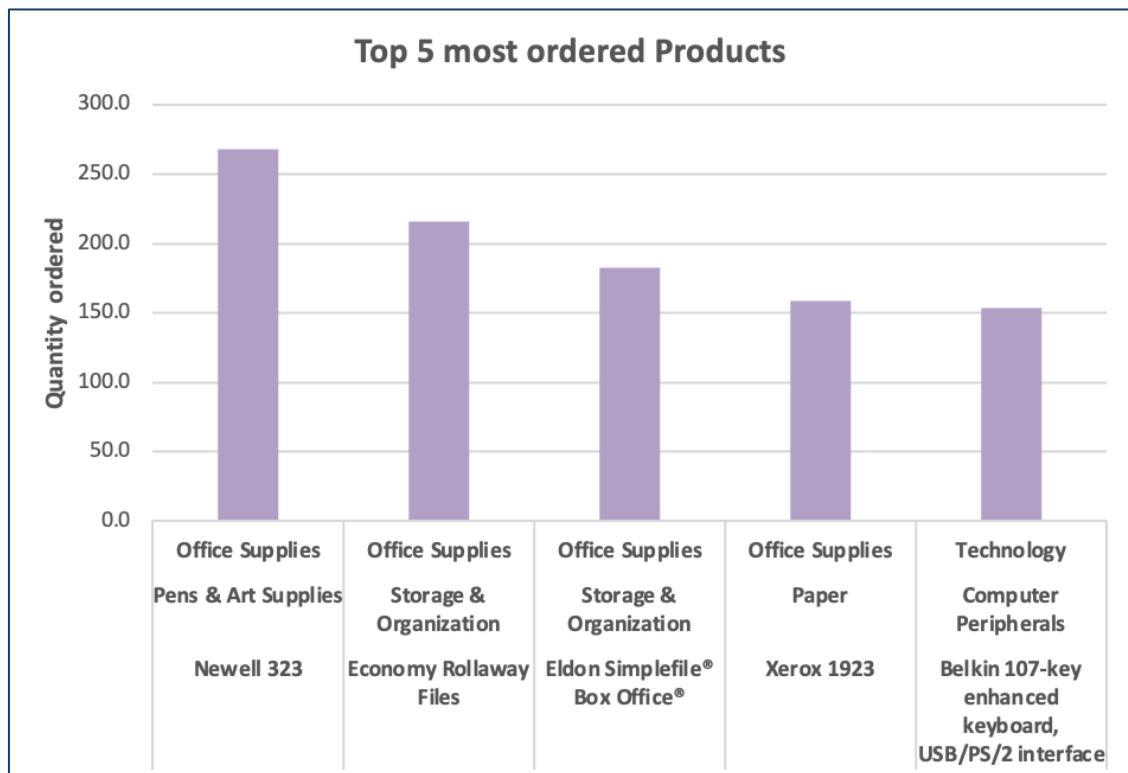
Figure 9- The Top 5 most sold products.



According to our findings, topmost (Figure 9) product that generated most sales was **Fellowes PB5 00 Electric Punch Binder** which made total sale of \$69,000 dollars followed by **Lexmark 4227 plus** which generated \$61,000 in sales and **Global Troy Leather chair** at \$50,000. Here Polycom View Station™ ISDN Video conferencing Unit is the most expensive product which is sold at \$6,783. These are the products which generated most of the sales, and in addition to being significant revenue generators, also provided valuable insights into consumer preferences and

market trends. It should be noted that though these products initiated good chunk of sales, they were not the most ordered products.

Figure 10- The Top 5 most ordered products.



The most ordered product (**Figure 10**) was **Newell 323** where 268 quantities were ordered, which are pens and sharpies stating they have a good market hold with high demand. The second most ordered item was Economy rollaway files where 216 quantities were ordered in first half of the year (Jan – June) etc. The least ordered item was Belkin 107 Key enhanced keyboard, which is a technological item. This tells us that most ordered Product category was Office supplies whereas Technological products generated the most revenue.

Figure 11- The Profit of Product Categories by Region.

Region	Furniture	Office Supplies	Technology	Total
Central	41.8	89.9	336.0	136.7
East	-4.7	192.9	315.0	179.9
South	12.3	-8.9	-121.2	-32.6
West	608.5	48.2	78.3	161.4
Total	148.1	83.6	156.6	114.8

(Figure-11) Digging deeper in the analysis, we gathered the insight that different product categories are more profitable than the other in different regions. Breaking down the findings, we discovered that the average profit of Furniture is \$148.1, and that furniture is most profitable in **Western regions** where the average profit was **\$608** and is least profitable in Eastern regions where it generated negative profit. Furniture was also not that profitable in the southern region in comparison to the Northern region. In our next product category, we can see that Office supplies was overall the least profitable category presenting a consistent challenge across all regions.

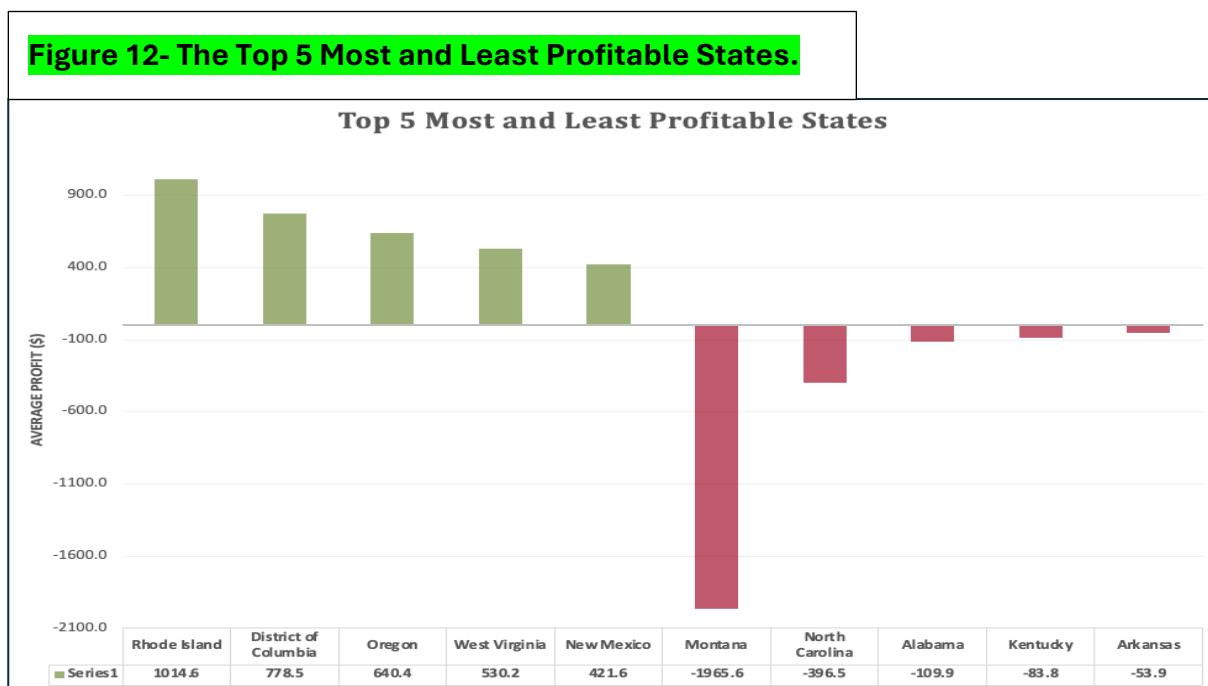
Office supplies generated the most profit in the Eastern region where the average profit was \$192.90 and is the least profitable in Southern region where it generated negative profit. Office supplies performed very well in the Central region, given that unit price of Office Supplies is the lowest among other categories.

Lastly, the most profitable items in the Central and Eastern regions were Technological products, generating an average profit of **\$336** and **\$315** respectively. Southern region saw a major tech letdown, with technological products disappointing customers and leading to negative returns of **-\$121** on average.

Even with that negative return, technological products have highest average product profits among other categories given that they are most expensive products. Our analysis on regional profitability revealed that the Eastern region is most profitable, while the Southern region lagged and experienced negative average profits.

3.1.2 DESCRIPTIVE ANALYSIS

Beyond product-level performance, state-level profitability was analysed to identify geographic patterns.



In general, the profitability of different products can vary significantly across different regions. Understanding which states are most and least profitable for the business is crucial when developing strategies to optimize the business's profits.

(Figure 12) depicts the most and least profitable states. **Rhode Island** emerged as most profitable state for the business, returning an average profit of **\$1014** per sale whereas **Montana** presented a significant challenge with major negative returns of **-\$1956** per sale.

Other profitable states were the districts of Columbia, Oregon, West Virginia and New Mexico, These states enjoyed a healthy range of average profits between **\$420 - \$790** per sale. The least profitable states besides Montana were North Carolina, Alabama, Kentucky, and Arkansas where they returned a range of negative profits dipping between **-\$52 and -\$397**.

Beyond Montana's substantial losses, Oregon was the most profitable state in the western region, whereas central and southern regions struggled with profitability which ignites the necessity of a deeper investigation into why this could be the case.

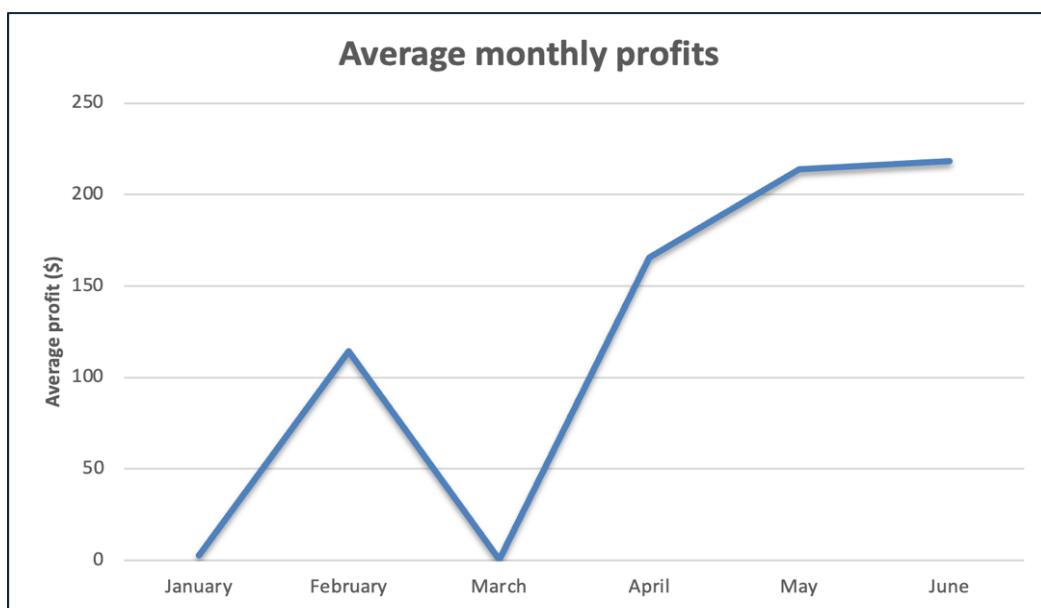
3.1.3 SEASONAL TRENDS AND PROFITS

Finally, temporal patterns were examined to understand seasonal fluctuations in sales and profit.

Figure 13- Average Monthly Sales of Superstores.



Figure 14- Average Monthly Profit of Superstores



Understanding how sales and profits fluctuate throughout the year is crucial for any business dealing with wide range of products. Even though we have data for the first half of the year, this analysis delves into monthly trends, uncovering insights into average sales by month and average monthly profits.

By exploring these patterns, we can gain valuable knowledge about consumer behaviour, profitability cycles and make informed decisions for the next half of the year.

For the first half-year of sales in **(Figure 13)**, we can see that it has a mixed trend with some peaks and trenches, which shows the unpredictable nature of the sales in the first half of the year. In the first half of the year, the average sales were \$994.

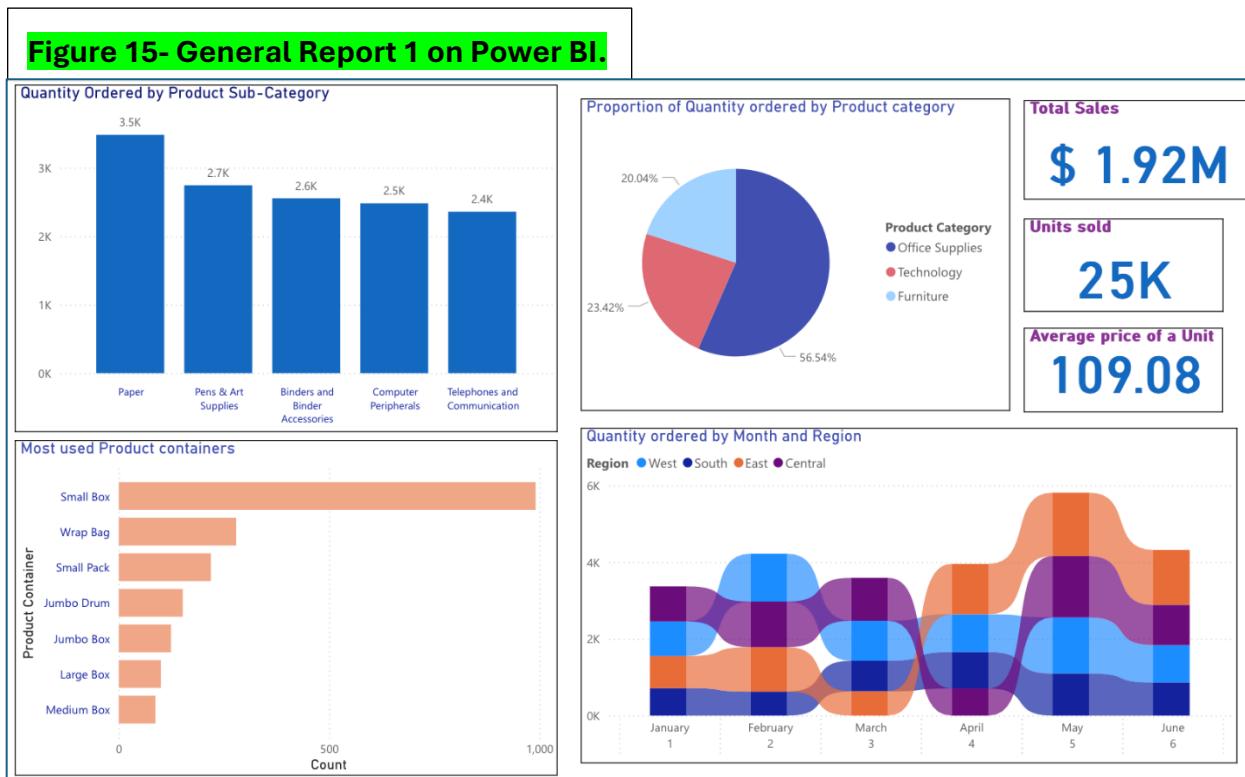
Though, the first month yielded the lowest sales (only \$743), the intriguing part is that most orders were made in January in comparison to the subsequent five months and made up a total of 370 unique orders. February presented positive contrast, boasting average sales of \$1039, however the month of March witnessed a sudden drop to \$828.

April was a strong month where it had staggering average sales of \$1214 followed by a drop and a hike again. The trends in average sales can be due to factors like price per unit and units ordered. In comparison, if we dig deeper and compare the average profit generated monthly, we can see patterns that warrant further investigations. In **(Figure 14)** we can see that despite January's lowest average sales the profits were not negative which is a good sign.

In February, high sales are accompanied by high average profit at \$114, with a drop in momentum in March, where it yielded no profits. Following this, a positive trend emerges in April with both sales and profit increasing steadily. The fascinating part lies in the last two months, despite May having lower overall sales compared to April and June, May delivered equally impressive profits. This could be merely because products with larger profit margins were ordered in this month.

In June, the key insight is that Superstores received the least orders of only 306 but generated high sales and high profits. Understanding the factors behind this success and unpacking the product demands will enable Superstores to perform consistently well throughout the entire year.

3.2 POWER BI ANALYSIS PART 1



Our first Power BI report (**Figure 15**) is designed to provide key insights on the quantity of items ordered, with respect to product categories, sub-categories and month using descriptive analysis. Using visualizations, intuitive dashboards, and in-depth analytics, we aim to provide Superstores with meaningful insights needed to make informed decisions.

The exceptional performance with regards to sales and the product quantities sold highlights the company's ability to adapt and thrive in dynamic market environments. Superstores performed well in the first half of the year, as total sales reached \$1.92 million and 25k units were sold, where average price per unit was \$109, which indicated a balanced spread in price across the products sold.

The first graph (*Quantity ordered by product sub-category*) shows that the highest ordered product sub-category is Paper with 3.5k orders, followed by Pens & Art supplies with 2.7k orders, which implies that the company's most ordered sub-categories are office supplies and stationary products. This claim is further substantiated by the **second graph** (*proportion of quantity ordered by product category*), as 56.5% of the products ordered were office supplies, which emphasises the strong demand of office-related products and the potential monopoly the company may have in this product category.

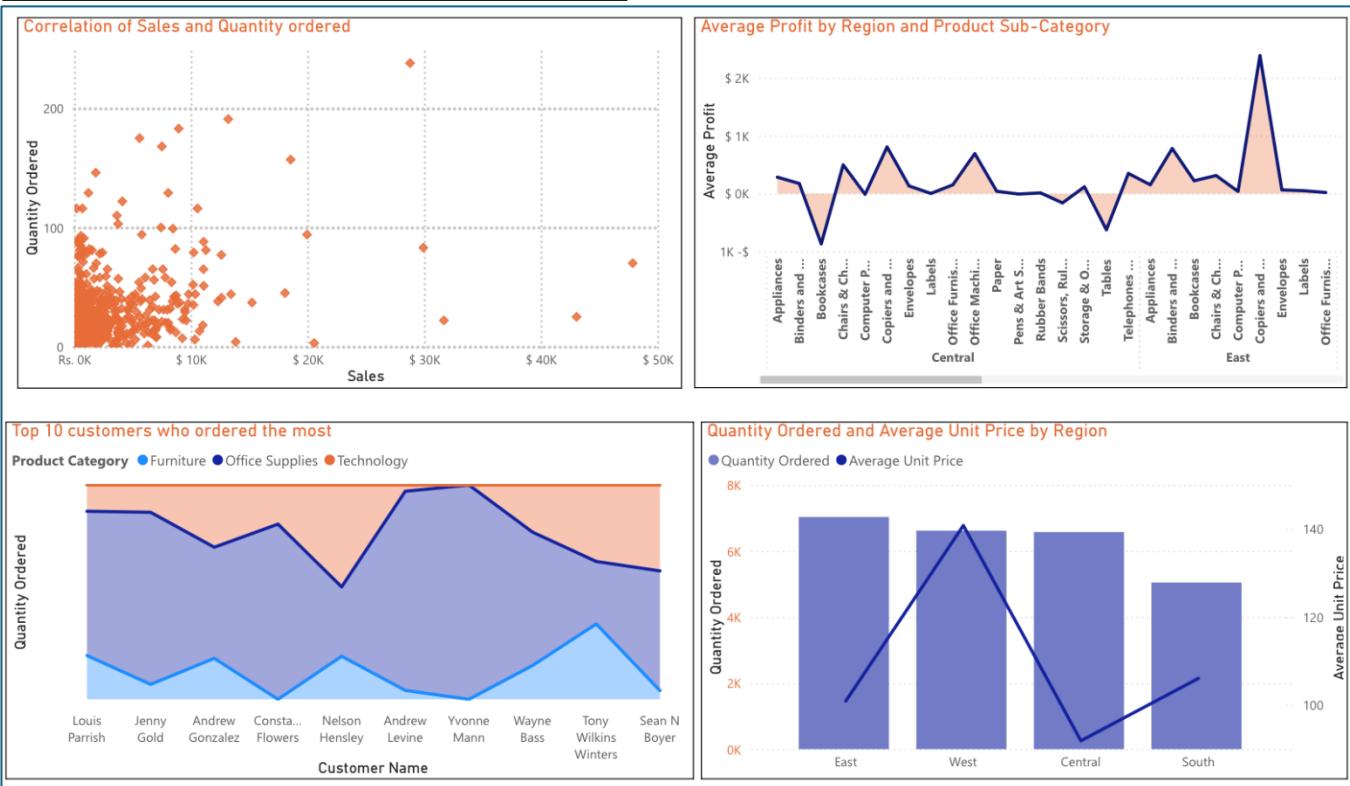
In the third graph (*Most used product containers*), we've uncovered that Superstores more commonly uses small boxes or Wrap bags for delivery, which would indicate that Superstores prioritises cost efficiency through conserving resources and contribute to environment sustainability in the process. This may lead to forming positive relationships with stakeholders and foster long-term success for the company.

Lastly, the final graph (*quantity ordered by month and region*) above showcases some unexpected fluctuations in the quantity ordered by region in the first six months. We have observed significant variations in order volumes across the North, South, East and West. In January, the quantity ordered across all 4 regions was balanced, whereas in February, this order dynamic shifted, and most orders came from western regions. In March, we can see that most orders came from the Central region, followed by West, South and East. In April, we can see a significant drop in orders, as the most orders came from East region and the least orders came from Central region.

A key insight here is that the East region's quantity ordered rises across April, May and June and maintains this momentum, whereas the central region experiences a significant drop in the quantity ordered in April, which raises the need for further enquiries.

3.2.1 POWER BI ANALYSIS PART 2

Figure 16- General Report 2 on Power BI.



In the second Power BI report (Figure: - 16), our team performed a deeper analysis to extract insights using more advanced statistical methods, by showcasing new findings by utilizing correlation and higher dimensional data analysis.

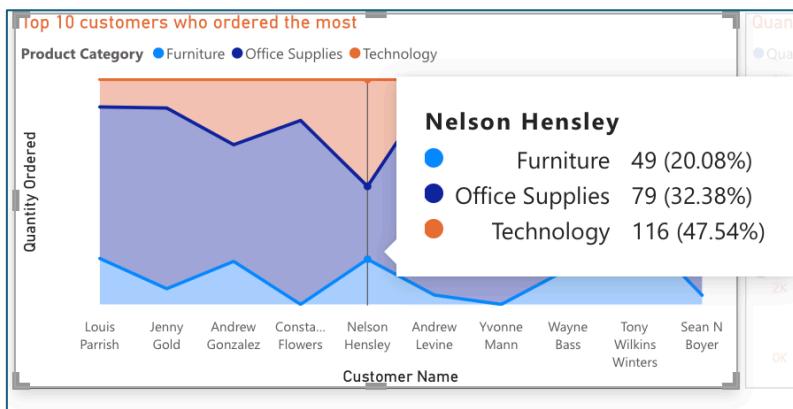
The first graph (*Correlation of sales and quantity ordered*) shows the correlation between the quantity ordered and sales, across all the products sold during the period. The **correlation coefficient** is 0.34 (*calculated in Ms Excel*), which indicates that there is a weak positive relationship, meaning that there is a slight tendency for higher quantities of products ordered to generate larger sales. This is questionable, as it suggests that aggregate correlation masks heterogeneity across product categories, indicating that category-level or price-adjusted models would provide more reliable insight. Our basic intuition suggests that higher quantities ordered would naturally lead to higher sales. This highlights the importance of controlling for unit price, discounting, and product mix when interpreting aggregate correlations.

Though another factor to look for here is that greater discounts could lead to less sales, even when large quantities of products are ordered. The factor could play a role in affecting the correlation coefficient, which may further raise the need for a deeper analysis, where the Correlation Coefficient is segmented monthly.

The second graph (*Average Profit by region and product sub – category*) provides insights on what sub-category of products are profitable in each region. In the Central region, bookcases and tables are least profitable whereas copiers, fax and office machines are most profitable. In the eastern region, we can see that bookcases and copiers & fax are most profitable whereas tables yielded negative profits in this region.

The third graph (*Top 10 customers who ordered the most*) shows the top 10 customers and their preferred product categories.

Figure: - 16-i



(Figure 16-i) shows us that customer *Nelson Hensley* ordered 49 Furniture, 79 office Supplies and 116 technological products across the period, which indicates that this customer is primary interested in Superstore's technological products. This graph can help us in determining the customer's ordering patterns to perform targeted marketing, which can eventually yield more profit.

The last graph (*Quantity ordered and average price of the unit by regions*) shows the relationship between the Quantity ordered and the average price of the unit. In the eastern region, the average unit price is \$100.93 and 7031 products were ordered, causing the total spending in the eastern region to be approx. \$710,970.83. The quantity ordered here is high despite a lower average unit price. This suggests that customers are purchasing products that are perceived to offer good value for money, leading to an increased demand of the product.

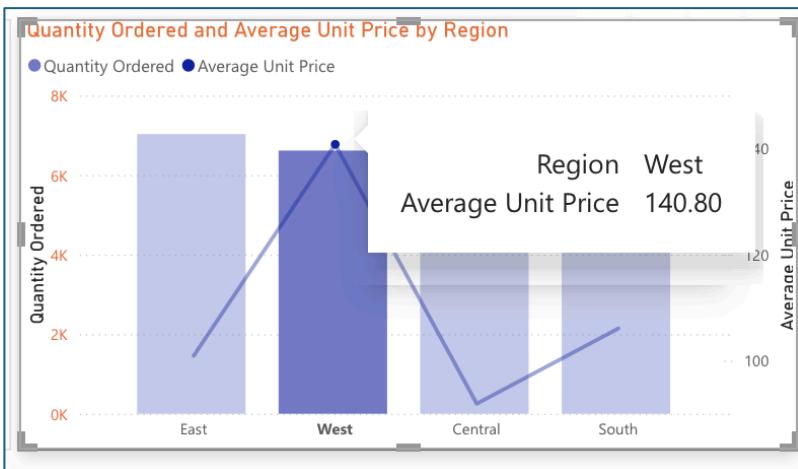


Figure: - 16- ii

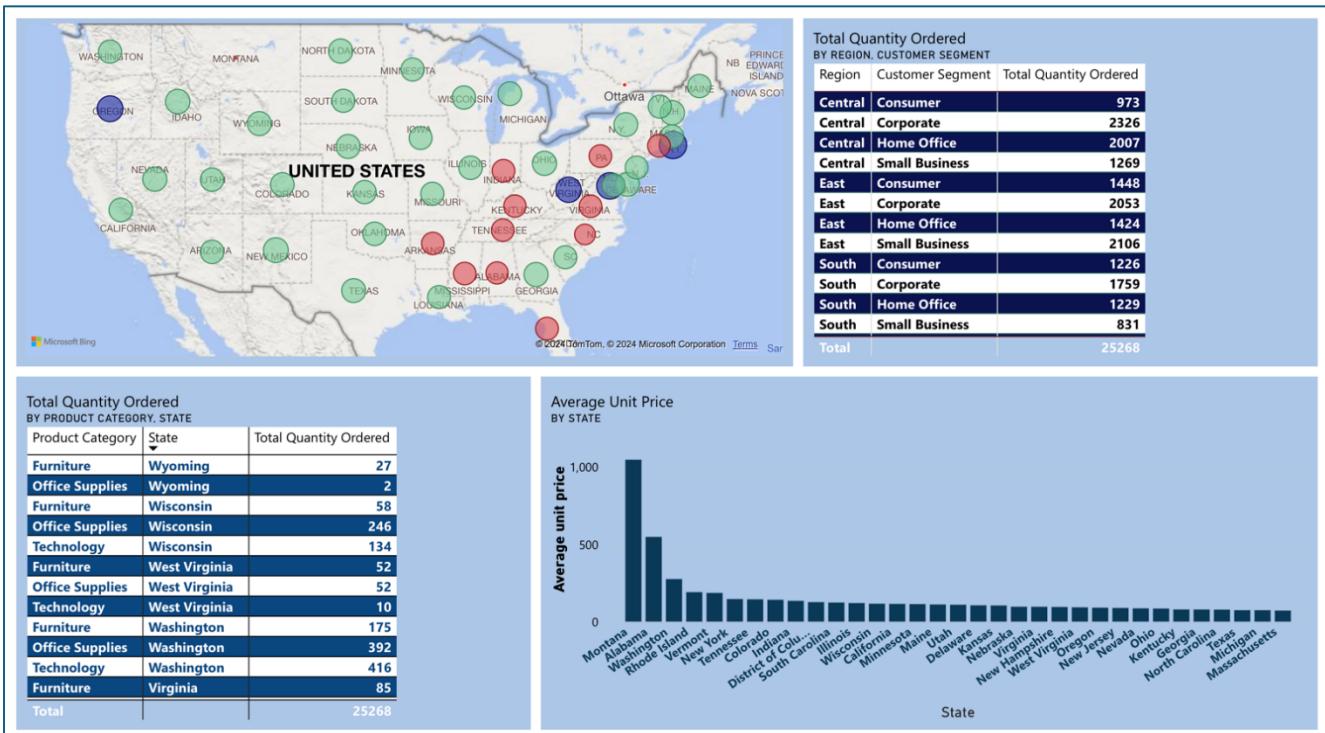
In contrast, for customers in western region (Figure:- 16- ii) the average unit price of products purchased was \$140.80 and 6617 products were ordered, causing the total spending in the western region to be approx. \$931,673.6. This could suggest that customers are willing to pay more for products with a higher perceived quality, unique features, or a good brand image. It also should be noted that on average the southern region generated the least profit.

ASSUMPTIONS :-

- The drill down features in the chart sub sections can be done for other components of the chart as well, but not all parts can be explained in the report.
- In last section, discounts have not been included to avoid financial misinterpretation.
- It is assumed that the shipping cost is incurred by the company, not the customer. Since we don't have any metric regarding customer satisfaction relating to the delivery time, we assumed all orders were successfully delivered.
- Not all our analysis in Power BI can be explained due to the word limit of the report, but the most important findings have still been included in it.

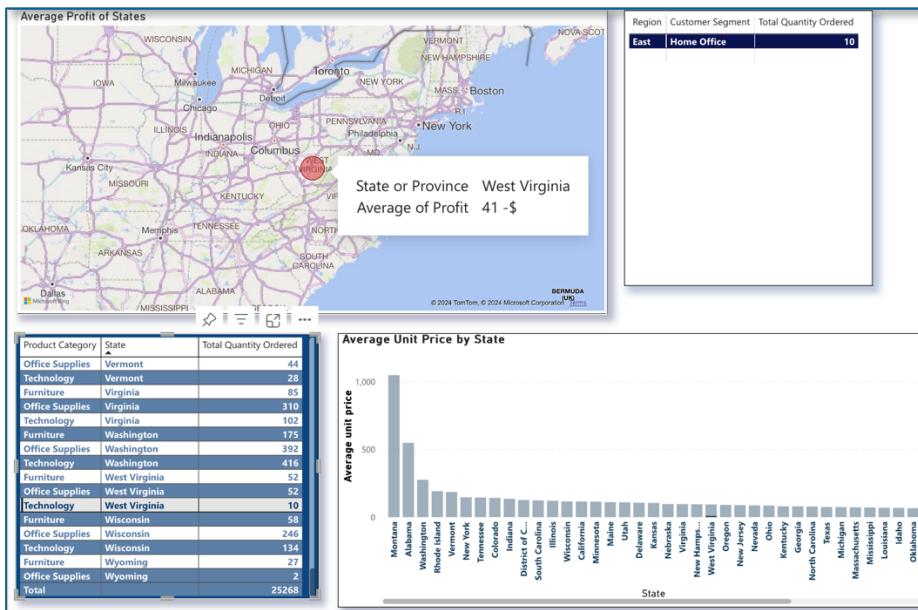
4. DASHBOARD: REGIONAL MANAGER

Figure 17- Regional dashboard



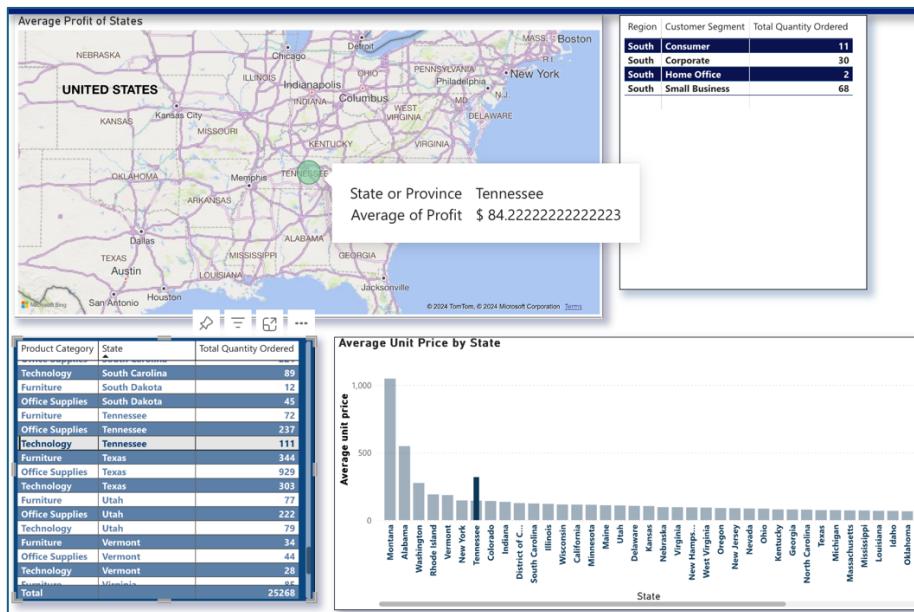
A dashboard was created that summarizes all our key findings for the Regional manager. We made four different charts including variables such as quantity ordered , profit and average unit price in different regions, allowing us to make strategic insights that can enhance Superstore's **regional performance** and help it achieve organizational goals.

Figure: - 17-i



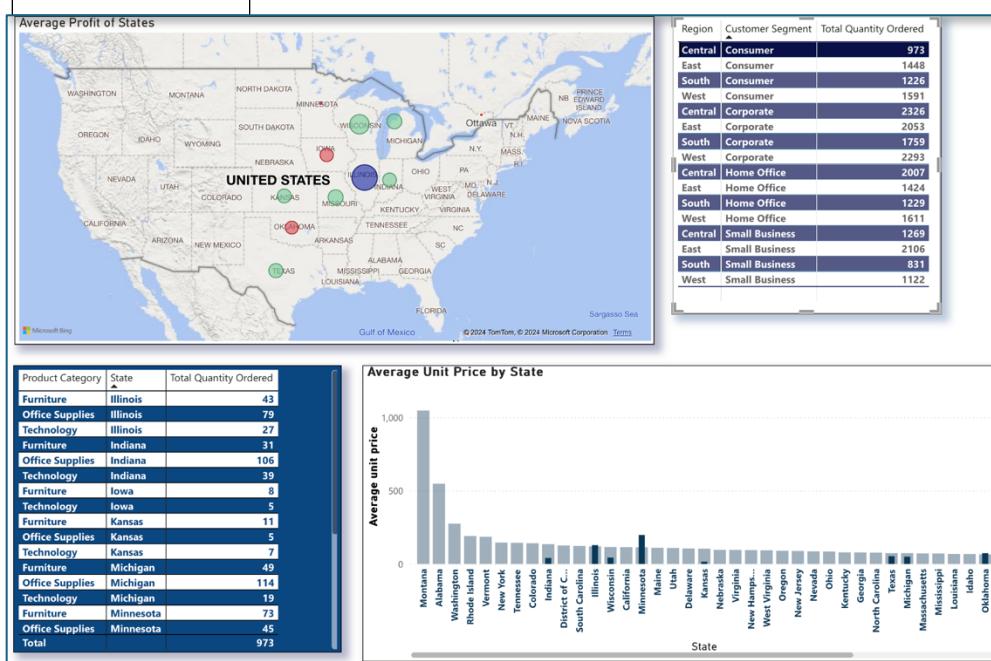
The bubble map provides us with an interactive visualization of the profit distribution of each state by representing profit levels. Through varying bubble sizes and colours, this map offers a clear and intuitive way to discern profitable regions from those with lower profitability. The information on the map can be isolated with the use of tables to give us deeper insights of the profit distribution. If we integrate the table (*Total quantity ordered by product category*) we found that some regions may be overall profitable, but some states don't perform well in certain categories. In our bubble map, West Virginia's average profit is \$530 but if we dig deeper, we can see that in (Figure: - 17-i) technological products are not profitable in West Virginia with negative profits. This insight underscores the importance of conducting granular analysis to uncover nuanced patterns and trends that may not be immediately apparent from generalised summaries. The table (*Total quantity ordered by customer segment*) also tells us that 10 orders done by Home office yielded negative profits. One of the reasons could be that technological products are expensive products that when ordered more, will eventually be more profitable.

Figure: - 17-ii



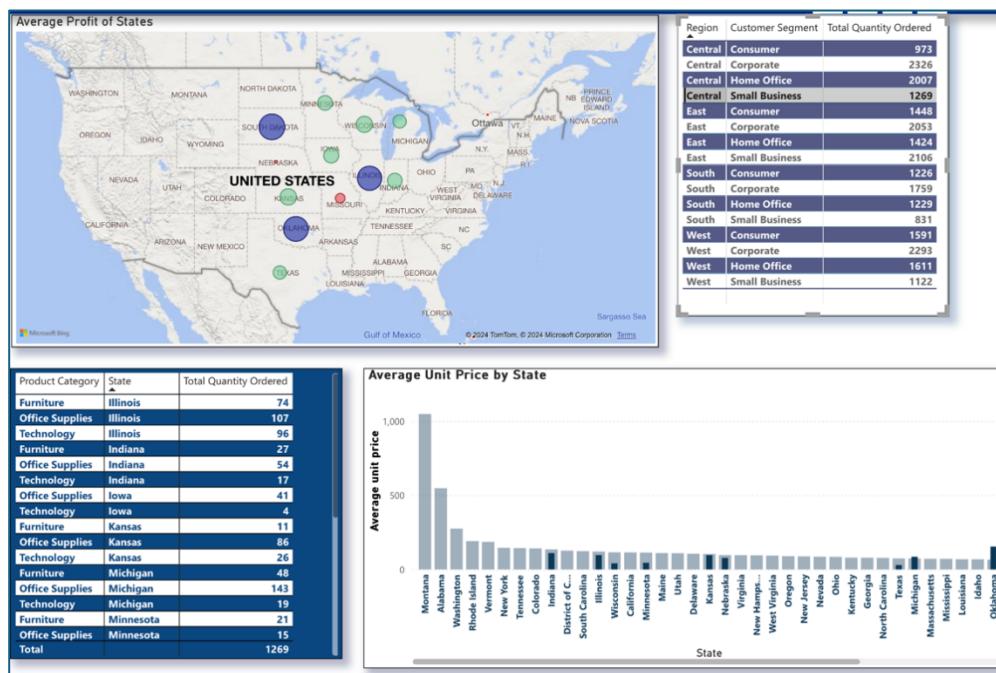
In contrast, in the state of Tennessee, average profits are shown to be - **\$39** on average which may be considered as an unprofitable, however if we look deeper into (Figure: - 17-ii) it's found that Tennessee only yielded positive returns on technological products and has an average profit of \$84 dollars, where 111 quantities were ordered in which 68 of these orders were made by small businesses. Our bottom graph also shows us the average unit price was \$318.46, whereas overall average price was \$144.37.

Figure: - 17-iii



By incorporating the table (*Total quantity ordered by customer segment*) we found that certain customer segments are more profitable in different regions. For instance, (Figure 17-iii) in the central region, the consumer customer segment that is the most profitable is the state of Illinois and the least profitable is the state of Oklahoma.

Figure: - 17-iv



In the central region, the most profitable customer segment was small businesses where most states returned positive profits (Figure: - 17-iv). By drilling down into specific product categories and customer segments, we can uncover hidden opportunities, mitigate risks, and make informed decisions that will ultimately contribute to achieving sustainable levels of profitability and success in different regions in the US.

4.2 DASHBOARD : SALES MANAGER

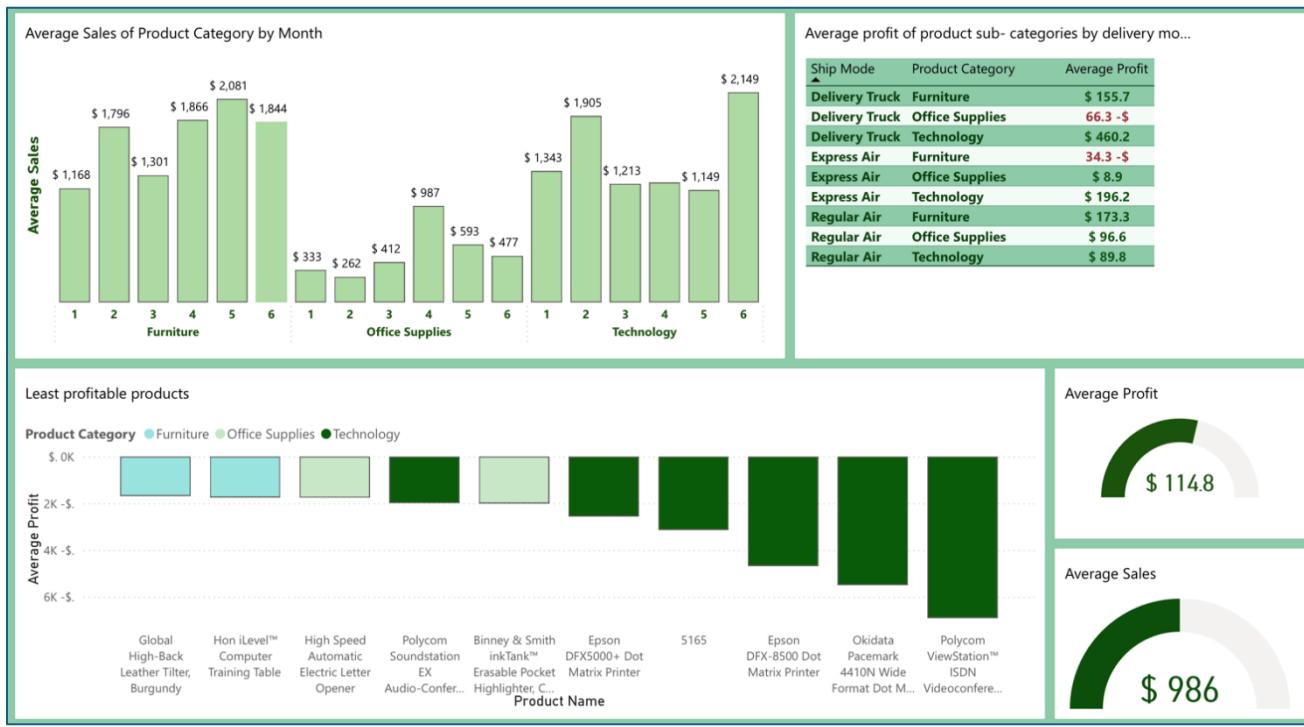


Figure 18- Sales Manager Dashboard

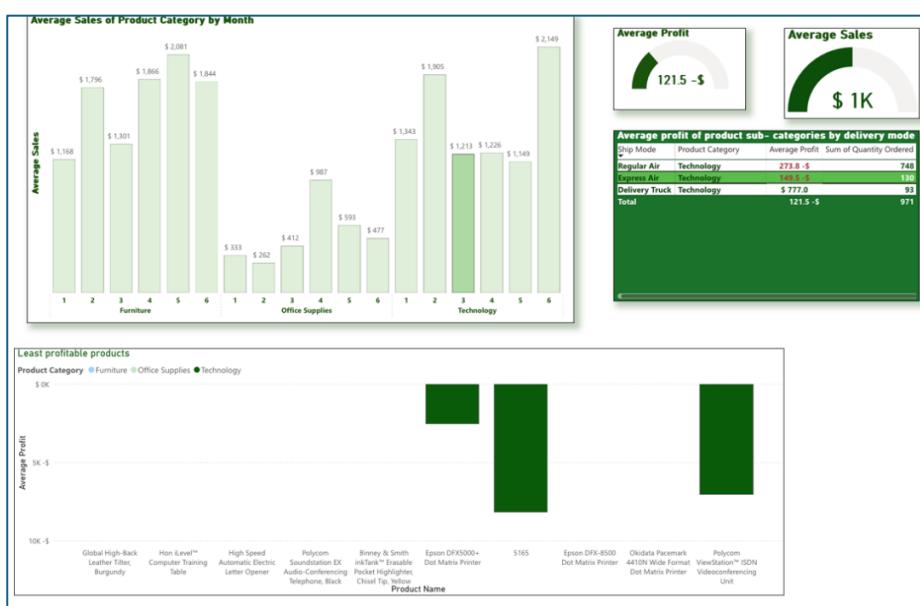
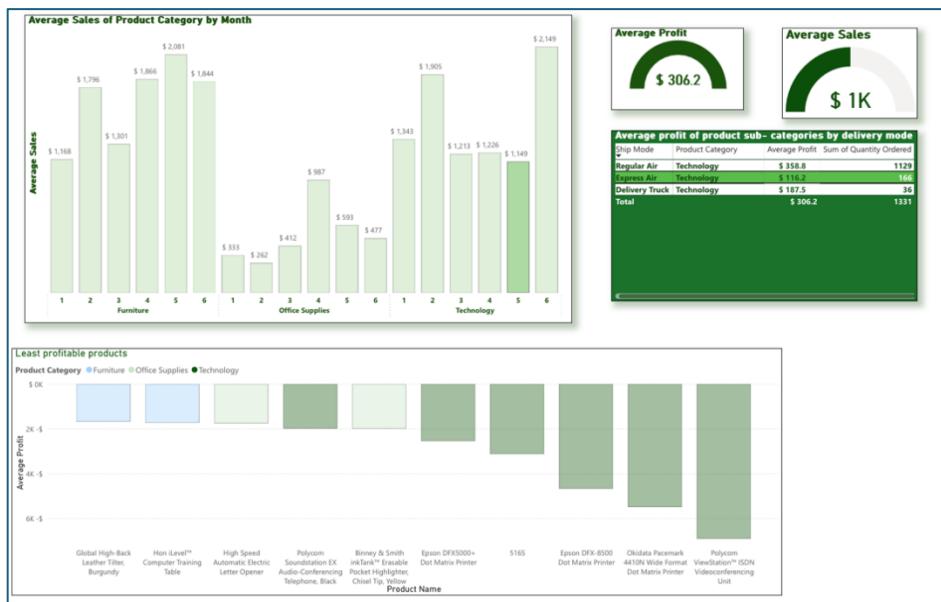


Figure: - 18-ii

(Figure- 18)- This dashboard (targeted to the sales manager) covers average profits , sale of products , mode of shipment , product- categories and sub-categories to help optimise financial performance and enable data driven decision-making within the company. Across the six-month time, Superstores achieved an average sales figure of \$986. During this period, the company managed to generate an average profit of \$114.8

Our first graph shows us the average monthly sales of different product categories. We can see that the average sales of Furniture were highest in May and lowest in January. Office supplies had an overall low average sales compared to other two categories, but that's because of its low unit prices. There was a notable hike in the sale of Office related products in April, whereas in February the average sales were the lowest (only \$262). Since technological products have higher average unit prices, June was the most profitable in terms of sales, when factoring in all product categories. After incorporating it with other graphs **(Figure: - 18)**, we found that in the month of March technological products delivered by Regular or Express air returned negative profits even though the quantity ordered was very high, 748 & 130 respectively. This could be due to higher shipping costs compared to selling price as well as extensive discounts, which resulted in low profit margins. Three of the least profitable products being sold in these months may have also influenced this outcome.



(Figure: - 18-ii)-Despite May having the lowest average sales in technological products it turned out to be the most profitable in comparison to other months. On average, May returned a profit of \$306, where the most profitable and preferred mode of delivery was regular air.

We can conclude that Delivery trucks were on average the most profitable shipping mode for technological products and Furniture, Express air was only profitable for technological products because of the high profit margins per unit, and while office supplies may be cheaper than

technological products, the consistent demand and lower shipping costs of regular air make it a viable and profitable delivery mode for this category. The top 10 least profitable products are also for your reference.

ASSUMPTIONS :-

- The dashboards target the Regional and Sales Managers. It is assumed that the Sales manager will have product knowledge, so that we can provide a deeper analysis of numbers of different products that would be meaningful.
 - The Regional Manager focuses merely on regional distributions and locations of products supply.
 - The Sales manager manages all the sales , profits of different products and sub-categories of products and is also assumed to have knowledge of all the product categories.
 - Deeper analysis can be further done in Power BI but has not been included due to word constraints of the task.
 - The months mapping has been performed in PowerBI i.e. Jan- 1 , Feb – 2 , Mar – 3 , Apr – 4 , May – 5 & Jun – 6.
-

5. JUSTIFICATIONS AND RECOMMENDATIONS

After performing our analysis on the dataset using Microsoft Excel and Power BI and creating the visualisations/dashboards above that have enabled the team to pinpoint the patterns, trends, and outliers in the dataset, we have produced five recommendations that may stimulate the growth of the business in the future.

5.1 RECOMMENDATION BASED ON EXCEL ANALYSIS :-

1. Superstores is encouraged to apply more personalised marketing strategies for their customers in the south-eastern states, to cut their losses.

It should be noted that there were numerous south-eastern states that generated a negative average profit. They were Florida, Arkansas, Mississippi, Alabama, Tennessee, Kentucky, Indiana, Virginia, North Carolina, Pennsylvania, and Connecticut. What is interesting to see is that all these states are clustered around the same geographic region, meaning that there could be common underlying reasons why the business experiences losses in this region.

A potential reason could be that many of these states have larger rural populations ([World Population Review 2020](#)), so many workers may require different tools to do their jobs, in contrast to tools required in corporate environments, which Superstores specialises in selling. Another reason could be that the general advertising pitch in these regions just does not resonate with the consumers in this area. There is also a high variation in the amount of products that were ordered across these states, and we have discovered that the correlation between the total average profit of the region with respect to the number of products ordered is quite low.

Therefore, it may be a promising idea to engage in more personalised marketing across these lower performing areas to tailor particular products to the specific needs of an individual customer, rather than apply a broader strategy across these regions ([Suman Kumar 2014](#)).

Advantages of personalised marketing are that businesses can increase customer satisfaction, engagement, and loyalty, leading to higher conversion rates and revenue, and with the right strategies and tools, Superstores can implement personalization effectively to gain their competitive advantage across these regions ([Jason Miller 2023](#)).

However, Superstores will also need to be transparent about how they collect and use customer data and to obtain consent before using it for personalization purposes, to avoid legal issues which could be our major limitation ([Jason Miller 2023](#)).

5.2 RECOMMENDATION BASED ON POWER BI REPORT :-

2. Superstores should allocate a larger portion of its budget to purchasing recyclable small box product containers.

Doing this will prevent a substantial amount of financial/material wastage from occurring, assuming that the business is already purchasing too many larger boxes for packaging ([Sustainability Victoria 2023](#)). The larger boxes are used to package bulky items, such as Furniture and Technology, but they make up much smaller portions of the product categories that are sold, which are **20.04% and 23.42% respectively**, which reduces the necessity to have massive quantities of larger boxes on hand.

Furthermore, the top three most used product containers are small boxes, Wrap Bags and Small Packs. This makes sense, as the highest proportion of quantity ordered by product category is for office supplies (**56.54%**), which are small, and hence would be shipped in smaller boxes. Finally, if Superstores were to ship their products to their customers by using sustainable packages, it would result in the following benefits.

Firstly, reducing environmental impact, since sustainable packaging materials often use less energy during production, and they can be derived from renewable sources. This aids in reducing the carbon footprint for Superstores, which in turn, will lead them to comply with any external regulatory requirements. Secondly, cost efficiency, as sustainable packages results in lighter packages being produced that are cheaper to transport and will reduce the cost of waste disposal. Lastly, an improved brand image that may cause a stronger emotional connection between the customers and Superstore's brand, which creates customer loyalty. Smaller boxes make up the most substantial portion of product containers used by the business, so the business would be able to save a lot of money, as well as obtain a more environmentally friendly image, if it were to use recyclable smaller boxes, assuming that it is not already ([Vic Ning 2023](#)).

5.3 RECOMMENDATION FOR REGIONAL MANAGER:-

3. Reward customers in the most geographically profitable regions, as well as those who have ordered the most products.

Across the six-month period, Superstores was operating across all the US states, and experienced a greater average profit in certain geographical locations compared to others. The top four best performing states were Rhode Island, Oregon, West Virginia, and the District of Colombia who all experienced an average profit above \$500 for the duration of the period. It should also be noted that the number of products that were ordered in these states was considerably low in proportion to the profit they generated, and that these states are all relatively spaced out from one another (particularly Oregon), which indicates that there are a small portion of customers that were purchasing high margin products that were ultimately responsible for driving up the average profit in these regions.

Just like how Superstores can engage in personalised marketing to cut losses, it can also inversely be used to either maintain or further stimulate the interest of the customers located in these high profit regions. Furthermore, in addition to looking for customers that generate the most profit, the business can also pinpoint which customers order the most products. As we have already deduced above, the total number of products ordered across the six-month period is **25268**. The bulk of these products were standard office supplies which made up **56.64%** of the products that were ordered, where Paper, Pens & Art supplies and Binder accessories were the top 3 most frequent products ordered.

Superstores can track down the Top 10 Customers, that our team has helped identify above, and publicly reward them with gifts (as well as tag them on social media pages) that may include large quantities of these high turnover products (among others), in order to retain their loyalty, and provide other consumers (who may follow Superstores social media pages) with an incentive to purchase more Superstore products ([Louise O Brien and Charles Jones 1995](#)).

5.4 RECOMMENDATION FOR FINANCIAL MANAGER:-

4. Superstores should review its shipping strategies and achieve product profitability by strategic pricing and minimize holding unprofitable products as inventory.

We have seen in the financial dashboard that in certain months some delivery modes are unprofitable. Superstores may want to review shipping strategies by focusing on optimizing logistics and transportation routes to reduce shipping costs, which will then increase the profit margin. As different product categories perform better in specific months, it would be possible to leverage product distribution and increase discounts, which will then lead to more quantities ordered and eventually increase sales and profits. Implementing strategies such as increasing shipping costs for prioritized products and consolidating it into a single mode of delivery for most products can potentially lead to increased profits. Furthermore, removing unprofitable products and optimizing inventory will not only reduce the number of products and increase the product quality but it will also help in focusing on the most profitable and preferred products only. By implementing these recommendations, the financial manager can drive meaningful improvements and sustain the success of superstores in the market. ([Cogsy May 18, 2022](#)).

5.5 BONUS RECOMMENDATION FOR THE FINANCIAL MANAGER

5. Superstores should allocate a month where they provide special discounts on Furniture and Technology, so that the quantity ordered comes closer to the quantity of Office Supplies ordered.

We have already determined that individually, the average sales of Furniture and Technology by month are much greater than the average sales of Office supplies by month. But despite this, even when put together, Furniture and Technology are ordered by customers less than office supplies. This could be due to a number of reasons. Superstores may already have an established reputation of being an office supply store that customers find desirable, but they may have competitors which specialise in selling Furniture and/or Technology that customers find even more desirable. A potential strategy would be to allocate a month when there is a high product turnover rate, like May, where they provide special discounts on Furniture/Technology products to increase the quantity sold. Superstores could allocate a portion of its budget into advertising this special month to attract as much attention as possible, and hopefully convince customers that the quality of the products it sells either rival/or are superior to its competitors, but also come with special discounts, that Superstores exclusively offers. Some loyal customers may stick around after the month has ended to purchase Furniture/Technology, and this way the quantity of Furniture and Technology items sold can be driven up ([Murdock 2024](#)).

REFERENCES

Dr Shekar Pawar (2016), 5 Powerful Benefits of Excel <<https://www.linkedin.com/pulse/5-powerful-benefits-excel-shekhar-pawar/>>

Is Your Inventory Killing Your Profits? [& How To Fix It]. (n.d.). Cogsy.

<https://cogsy.com/blog/is-your-inventory-killing-profits/>

Jason Miller (2023), The benefits of personalisation in Digital Marketing, <https://www.linkedin.com/pulse/benefits-personalization-digital-marketing-jason-miller/>

Leighton Murdock (2024), How can you use discounts and coupons to effectively drive sales? [https://www.linkedin.com/advice/0/how-can-you-use-discounts-coupons-effectively-drive-lw9je#:~:text=Coupons%20and%20discounts%20can%20be,of%20missing%20out%20\(FOMO\).](https://www.linkedin.com/advice/0/how-can-you-use-discounts-coupons-effectively-drive-lw9je#:~:text=Coupons%20and%20discounts%20can%20be,of%20missing%20out%20(FOMO).)

Louise O'Brien and Charles Jones (1995), Do Rewards really create loyalty? <https://hbr.org/1995/05/do-rewards-really-create-loyalty>

Muhammad Sheheryar (2023), Excel Tables: Are they worth it? Weighing Pros and Cons <https://www.linkedin.com/pulse/excel-tables-worth-weighing-pros-cons-muhammad-sheheryar/>

Marquette University (2024), Removing Hidden Data in Microsoft Excel <https://www.marquette.edu/its/help/security/hiddenexcel.shtml#:~:text=Avoid%20unintentionally%20distributing%20hidden%20information,worksheets%2C%20by%20removing%20hidden%20data>

Nancy J (2023), 5 Top Advantages of Power BI Training for Business Intelligence <https://www.linkedin.com/pulse/5-top-advantages-power-bi-training-business-intelligence-nancy-j-r9mwc/>

Rahm, Erhard & Do, Hong. (2000). Data Cleaning: Problems and Current Approaches. IEEE Data, Eng.Bull.23.3-13

https://www.researchgate.net/publication/220282831_Data_Cleaning_Problems_and_Current_Approaches

Suman Kumar Dawn (2014), Personalised Marketing: Concepts and Framework, Vol. 54, Issue. 4, (Jan-Mar 2014): 370-377., <https://www.proquest.com/docview/1516358681?pq-origsite=gscholar&fromopenview=true&sourcetype=Scholarly%20Journals>

Sustainability Victoria (2023), Accepting reusables at your business, <https://www.sustainability.vic.gov.au/recycling-and-reducing-waste/in-a-business-2/single-use-plastics-ban-centre/accepting-reusables-at-your-business>

Vaibhav Lambat (2024), How do you manage Duplicate Data Records. <https://www.linkedin.com/advice/1/how-do-you-manage-duplicate-data-records-skills-data-analysis#:~:text=Duplicate%20data%20records%20are%20a,to%20erroneous%20conclusions%20or%20decisions>

Vic Ning (2023), Top 5 Benefits of Sustainable Packaging in your Business, <https://www.linkedin.com/pulse/top-5-benefits-sustainable-packaging-your-business-vic-ning-rperc/>

World Population Review (2020), Most Rural States 2024. <https://worldpopulationreview.com/state-rankings/most-rural-states>

