INTRODUCTION TO DATA MANAGEMENT

(Project Semester August-December 2022)

SALES ANALYSIS OF SUPERSTORE

Submitted by
Yendluri Pavan Ram Chandar
Registration No.12003141

Programme and Section: CSE(Honour) and KM082

Course Code: INT 217

Under the Guidance of
Baljinder Kaur
UID:27952

Discipline of CSE/IT
Lovely School of Computer Science & Engineering
Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that Yendluri Pavan Ram Chandar bearing Registration no. 12003141 has completed INT217 project titled, **"SALES ANALYSIS OF SUPERSTORE"** under my guidance and supervision. To the best of my knowledge, the present work is the result of his original development, effort and study.

Baljinder Kaur

School of Computer Science and Engineering
Lovely professional University
Phagwara, Panjab.

Date:10-11-2022

DECLARATION

I, Yendluri Pavan ram Chandar student of B.Tech CSE under CSE/IT Discipline at, Lovely Professional University, Punjab, hereby declare that all the information furnished in this project report is based on my own intensive work and is genuine.

Date:10-11-2022

Registration No.: 12003141   Yendluri Pavan Ram Chandar

Acknowledgement

I would like to express my special thanks of gratitude to my teacher Baljinder kaur who gave me the golden opportunity to do this wonderful project on the topic Data Analysison Sales which also helped me in doing a lot of Research and I came to know about so manynew things.

Secondly, I would also like to thank my friends who helped me a lot in finalizing this project within the limited time frame.

Table of Content

INTRODUCTION

Data Analysis is a process of inspecting, cleansing, transforming, and modeling data to discover useful information, inform conclusions, and support decision-making. Data analysis has multiple facets and approaches, encompassing diverse techniques under a variety of names, while being used in different business, science, and social science domains.

The Analytics team of a Super Store anywhere in the world would want to design a Sales and Performance dashboard to analyze the sales based on various product categories and other factors which have a role to play in the running of the store. The store managing head, or the owner wants to add user control for product category, so users can select a category and can see the trend month-wise and product-wise accordingly.

The Analytics team would also want to analyze various other things like how many days the store takes to ship the product, how many times a customer orders a product, how much time is there between the customer's first and second order, etc.

The Super Store's database keeps track of the following data fields:

- Order ID – Id of the order created by the customer.

- Order Date and Ship Date – Date when the item was ordered and the date when the item was shipped to the person.

- Shipping Mode – Mode of shipping

- Product Category – Category of the product

- Product – The product that was ordered

- Sales – Quantity of the items ordered

- Discount – Discount value on the product

- Profit – How much profit was earned on that product

- Order Priority – Priority of the order determining delivery

- Customer Name – Name of the person who ordered the item.

- Customer ID – Unique ID for each Customer

- Region – Region where the sales was made.

- Order Month – Month when the order took place.

SCOPE OF ANALYSIS

The super store wants to see and analyze the sales trend month-wise and product-wise , work upon the lagging segments, and outperform employees accordingly. The Analytics team also wants to create analyze the database in depth to help the super store grow exponentially. The Analytics team wishes to answer the following objectives: -

Sales, Quantity and Profit of each product category

1. Segment Distribution of each product category

2. Sales and Profit throughout months of a sales of each product category

3. Regional Sales of each product category

4. Overall Sales Trend throughout months of a sales year

5. Distribution of Order Priority

6. Customer Ordering Trend

7. Analysis of work Performance of Regional Managers

8. Comparison of sales and profit product category wise

9. Analyzing Shipping priority share.

Aim of this project is to answer the above objectives in the form of visualization by creating a dashboard to convey the answers effectively and efficiently.

ETL PROCESS

In computing, extract, transform, load (ETL) is a process in database usage to prepare datafor

analysis, especially in data warehousing. Data extraction involves extracting data from homogeneous or heterogeneous sources, while data transformation processes data by transforming them into a proper storage format/structure for the purposes of querying and analysis; finally, data loading describes the insertion of data into the final target database suchas an operational data store, a data mart, or a data warehouse. A properly designed ETL system extracts data from the source systems, enforces data quality and consistency standards, conforms data so that separate sources can be used together, and finally delivers data in a presentation-ready format so that application developers can build applications and end users can make decisions.

Precisely, ETL is defined as a process that extracts the data from different RDBMS source systems, then transforms the data (like applying calculations, concatenations, etc.) and finally loads the data into the Data Warehouse system. ETL stands for Extract, Transform and Load.

Before ETL, the dataset looked like this. **This data is taken from Kaggle.**

Day	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID	Category	Sub-Category	Product Name
8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-10001798	Furniture	Bookcases	Bush Somerset Collection Bookcase
8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Cl
12	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-10000240	Office Supplies	Labels	Self-Adhesive Address Labels for Typewrit
11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular T
11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-10000760	Office Supplies	Storage	Eldon Fold 'N Roll Cart System
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-10001487	Furniture	Furnishings	Eldon Expressions Wood and Plastic Desk A
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-10002833	Office Supplies	Art	Newell 322
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002275	Technology	Phones	Mitel 5320 IP Phone VoIP phone
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-10003910	Office Supplies	Binders	DXL Angle-View Binders with Locking Rings
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-10002892	Office Supplies	Appliances	Belkin F5C206VTEL 6 Outlet Surge
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-10001539	Furniture	Tables	Chromcraft Rectangular Conference Tables
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002033	Technology	Phones	Konftel 250 Conference phone - Charcoal b
15	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina	28027	South	OFF-PA-10002365	Office Supplies	Paper	Xerox 1967
5	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washington	98103	West	OFF-BI-10003656	Office Supplies	Binders	Fellowes PB200 Plastic Comb Binding Mac
22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-10002311	Office Supplies	Appliances	Holmes Replacement Filter for HEPA Air Cl
22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-10000756	Office Supplies	Binders	Storex DuraTech Recycled Plastic Frosted I
11	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wisconsin	53711	Central	OFF-ST-10004186	Office Supplies	Storage	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H
13	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah	84084	West	OFF-ST-10000107	Office Supplies	Storage	Fellowes Super Stor/Drawer
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-AR-10003056	Office Supplies	Art	Newell 341
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	TEC-PH-10001949	Technology	Phones	Cisco SPA 501G IP Phone
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-BI-10002215	Office Supplies	Binders	Wilson Jones Hanging View Binder, White,
9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AR-10000246	Office Supplies	Art	Newell 318
9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AP-10001492	Office Supplies	Appliances	Acco Six-Outlet Power Strip, 4' Cord Lengtl
16	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-CH-10002774	Furniture	Chairs	Global Deluxe Stacking Chair, Gray
25	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah	84057	West	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular T
16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	OFF-BI-10001634	Office Supplies	Binders	Wilson Jones Active Use Binders
16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	TEC-AC-10003027	Technology	Accessories	Imation 8GB Mini TravelDrive USB 2.0 Flas
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-BO-10004834	Furniture	Bookcases	Riverside Palais Royal Lawyers Bookcase, F
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10000474	Office Supplies	Binders	Avery Recycled Flexi-View Covers for Bind
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-FU-10004848	Furniture	Furnishings	Howard Miller 13-3/4" Diameter Brushed
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-EN-10001509	Office Supplies	Envelopes	Poly String Tie Envelopes
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-AR-10004042	Office Supplies	Art	BOSTON Model 1800 Electric Pencil Sharp
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10001525	Office Supplies	Binders	Acco Pressboard Covers with Storage Hool

Through the process of ETL, we are going to clean the dataset and bring all the entities to their proper data format.

Step 1: Removing the blank cells from the dataset.

For this, select the whole dataset. Go to Find and Select in the Home tab of excel. Select Goto Special from the drop-down menu and then tick the blank option. All the blank cells willbe selected. Then go to Delete option in the home tab again and select Delete Rows from the drop-down menu. This will remove any rows with blank cells.

Day	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID	Category	Sub-Category	Product Name
8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-10001798	Furniture	Bookcases	Bush Somerset Collection Bookcase
8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-10000454	Furniture	Chairs	Hon Deluxe Fabric Upholstered Stacking Cl
12	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-10000240	Office Supplies	Labels	Self-Adhesive Address Labels for Typewrit
11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular T
11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-10000760	Office Supplies	Storage	Eldon Fold 'N Roll Cart System
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-10001487	Furniture	Furnishings	Eldon Expressions Wood and Plastic Desk A
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-10002833	Office Supplies	Art	Newell 322
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002275	Technology	Phones	Mitel 5320 IP Phone VoIP phone
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-10003910	Office Supplies	Binders	DXL Angle-View Binders with Locking Ring:
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-10002892	Office Supplies	Appliances	Belkin F5C206VTEL 6 Outlet Surge
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-10001539	Furniture	Tables	Chromcraft Rectangular Conference Tables
9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002033	Technology	Phones	Konftel 250 Conference phone - Charcoal b
15	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina	28027	South	OFF-PA-10002365	Office Supplies	Paper	Xerox 1967
5	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washington	98103	West	OFF-BI-10003656	Office Supplies	Binders	Fellowes PB200 Plastic Comb Binding Mac
22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-10002311	Office Supplies	Appliances	Holmes Replacement Filter for HEPA Air Cl
22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-10000756	Office Supplies	Binders	Storex DuraTech Recycled Plastic Frosted I
11	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wisconsin	53711	Central	OFF-ST-10004186	Office Supplies	Storage	Stur-D-Stor Shelving, Vertical 5-Shelf: 72"H
13	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah	84084	West	OFF-ST-10000107	Office Supplies	Storage	Fellowes Super Stor/Drawer
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-AR-10003056	Office Supplies	Art	Newell 341
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	TEC-PH-10001949	Technology	Phones	Cisco SPA 501G IP Phone
27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-BI-10002215	Office Supplies	Binders	Wilson Jones Hanging View Binder, White,
9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AR-10000246	Office Supplies	Art	Newell 318
9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AP-10001492	Office Supplies	Appliances	Acco Six-Outlet Power Strip, 4' Cord Length
16	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-CH-10002774	Furniture	Chairs	Global Deluxe Stacking Chair, Gray
25	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah	84057	West	FUR-TA-10000577	Furniture	Tables	Bretford CR4500 Series Slim Rectangular T
16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	OFF-BI-10001634	Office Supplies	Binders	Wilson Jones Active Use Binders
16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	TEC-AC-10003027	Technology	Accessories	Imation 8GB Mini TravelDrive USB 2.0 Flas
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-BO-10004834	Furniture	Bookcases	Riverside Palais Royal Lawyers Bookcase, F
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10000474	Office Supplies	Binders	Avery Recycled Flexi-View Covers for Bind
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-FU-10004848	Furniture	Furnishings	Howard Miller 13-3/4" Diameter Brushed
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-EN-10001509	Office Supplies	Envelopes	Poly String Tie Envelopes
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-AR-10004042	Office Supplies	Art	BOSTON Model 1800 Electric Pencil Sharpe
17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10001525	Office Supplies	Binders	Acco Pressboard Covers with Storage Hool

Step 2: Removing columns which are not properly defined or not crucial to our analysis.

For this we will columns which are redundant like the column with just the index numbers. For this we will select that particular column and then go to delete option in the home tag andthen select Delete Columns from the drop-down menu.

Row ID	Order ID	Order Date	Year	Month	Day	Ship Mode	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID
1	CA-2016-152156	08-11-2016	2016	November	8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-BO-10001798
2	CA-2016-152156	08-11-2016	2016	November	8	Second Class	CG-12520	Claire Gute	Consumer	United States	Henderson	Kentucky	42420	South	FUR-CH-10000454
3	CA-2016-138688	12-06-2016	2016	June	12	Second Class	DV-13045	Darrin Van Huff	Corporate	United States	Los Angeles	California	90036	West	OFF-LA-10000240
4	US-2015-108966	11-10-2015	2015	October	11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	FUR-TA-10000577
5	US-2015-108966	11-10-2015	2015	October	11	Standard Class	SO-20335	Sean O'Donnell	Consumer	United States	Fort Lauderdale	Florida	33311	South	OFF-ST-10000760
6	CA-2015-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-FU-10001487
7	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AR-10002833
8	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002275
9	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-BI-10003910
10	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	OFF-AP-10002892
11	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	FUR-TA-10001539
12	CA-2014-115812	09-06-2014	2014	June	9	Standard Class	BH-11710	Brosina Hoffman	Consumer	United States	Los Angeles	California	90032	West	TEC-PH-10002033
13	CA-2017-114412	15-04-2017	2017	April	15	Standard Class	AA-10480	Andrew Allen	Consumer	United States	Concord	North Carolina	28027	South	OFF-PA-10002365
14	CA-2016-161389	05-12-2016	2016	December	5	Standard Class	IM-15070	Irene Maddox	Consumer	United States	Seattle	Washington	98103	West	OFF-BI-10003656
15	US-2015-118983	22-11-2015	2015	November	22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-AP-10002311
16	US-2015-118983	22-11-2015	2015	November	22	Standard Class	HP-14815	Harold Pawlan	Home Office	United States	Fort Worth	Texas	76106	Central	OFF-BI-10000756
17	CA-2014-105893	11-11-2014	2014	November	11	Standard Class	PK-19075	Pete Kriz	Consumer	United States	Madison	Wisconsin	53711	Central	OFF-ST-10004186
18	CA-2014-167164	13-05-2014	2014	May	13	Second Class	AG-10270	Alejandro Grove	Consumer	United States	West Jordan	Utah	84084	West	OFF-ST-10000107
19	CA-2014-143336	27-08-2014	2014	August	27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-AR-10003056
20	CA-2014-143336	27-08-2014	2014	August	27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	TEC-PH-10001949
21	CA-2014-143336	27-08-2014	2014	August	27	Second Class	ZD-21925	Zuschuss Donatelli	Consumer	United States	San Francisco	California	94109	West	OFF-BI-10002215
22	CA-2016-137330	09-12-2016	2016	December	9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AR-10000246
23	CA-2016-137330	09-12-2016	2016	December	9	Standard Class	KB-16585	Ken Black	Corporate	United States	Fremont	Nebraska	68025	Central	OFF-AP-10001492
24	US-2017-156909	16-07-2017	2017	July	16	Second Class	SF-20065	Sandra Flanagan	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-CH-10002774
25	CA-2016-106320	25-09-2015	2015	September	25	Standard Class	EB-13870	Emily Burns	Consumer	United States	Orem	Utah	84057	West	FUR-TA-10000577
26	CA-2016-121755	16-01-2016	2016	January	16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	OFF-BI-10001634
27	CA-2016-121755	16-01-2016	2016	January	16	Second Class	EH-13945	Eric Hoffmann	Consumer	United States	Los Angeles	California	90049	West	TEC-AC-10003027
28	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-BO-10004834
29	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10000474
30	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	FUR-FU-10004848
31	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-EN-10001509
32	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-AR-10004042
33	US-2015-150630	17-09-2015	2015	September	17	Standard Class	TB-21520	Tracy Blumstein	Consumer	United States	Philadelphia	Pennsylvania	19140	East	OFF-BI-10001525

Step 3: Giving proper and appropriate column names.

The dataset does not have proper columns so our next step would be to giver proper column names to the columns wherever required.

ANALYSIS OF DATASET

z

1. Monthly Sales and profit of each categoryDescription:

By knowing about sales and profit over month we can know about the months which are more profitable for sales and hence customize our advertisement plan to increase the sales even more. After finding out the sales and profit we visualize the result with the help of a stacked bar graph.
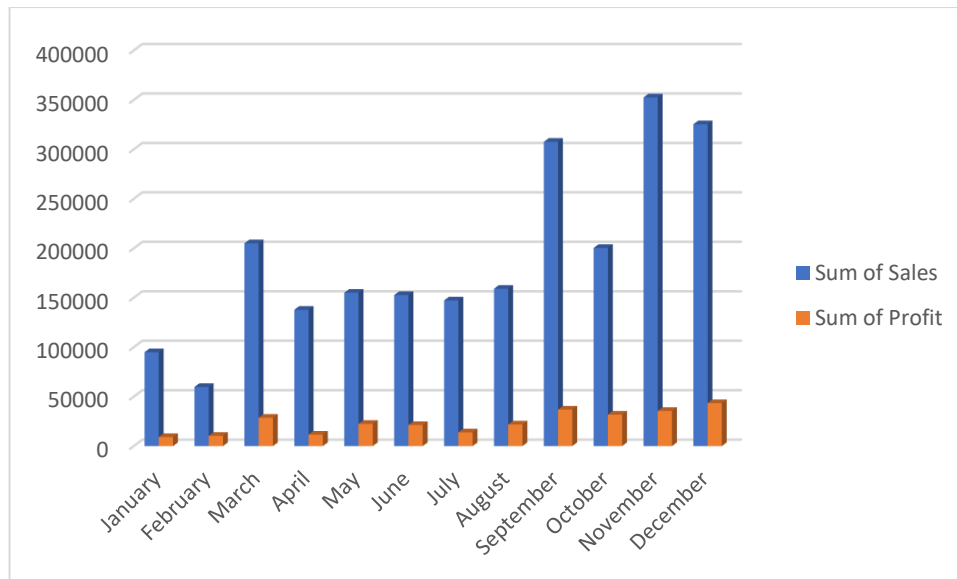
Specific function and requirements

We have to create a pivot table. No specific functions are used. We then put the priority c and count of their respective sales in the columns of the pivot table.

Results:

Row Labels 🔽	Sum of Sales	Sum of Profit
January	94924.8356	9134.4461
February	59751.2514	10294.6107
March	205005.4888	28594.6872
April	137762.1286	11587.4363
May	155028.8117	22411.3078
June	152718.6793	21285.7954
July	147238.097	13832.6648
August	159044.063	21776.9384
September	307649.9457	36857.4753
October	200322.9847	31784.0413
November	352461.071	35468.4265
December	325293.5035	43369.1919
Grand Total	2297200.86	286397.0217

Visualization:

The results are then visualized in the form of a stacked bar graph for both profit andsales

X

2.Segment Distribution of each product category:

Description:

By knowing which segment of sales has the greatest number of sales and which has least we can identify factors which affect the sales and thereby improve our strategy of making sales.

Specific function and requirements

We must create a pivot table. No specific functions are used. We then put the priority c and count of their respective sales in the columns of the pivot table.

Results:

Row Labels ▾	Count of Sales
Furniture	2121
Office Supplies	6026
Technology	1847
Grand Total	9994

Visualization:

We will use a pie chart to visualize the distribution.
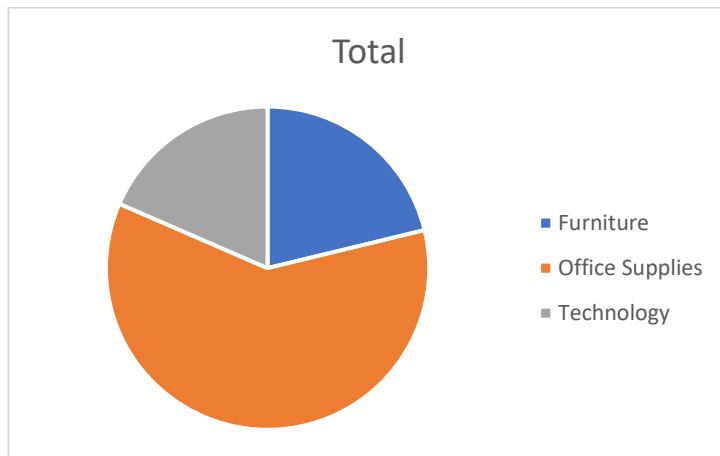
The chart of the total no of the segments in the superstore .

The percentage of the sales done on each category.
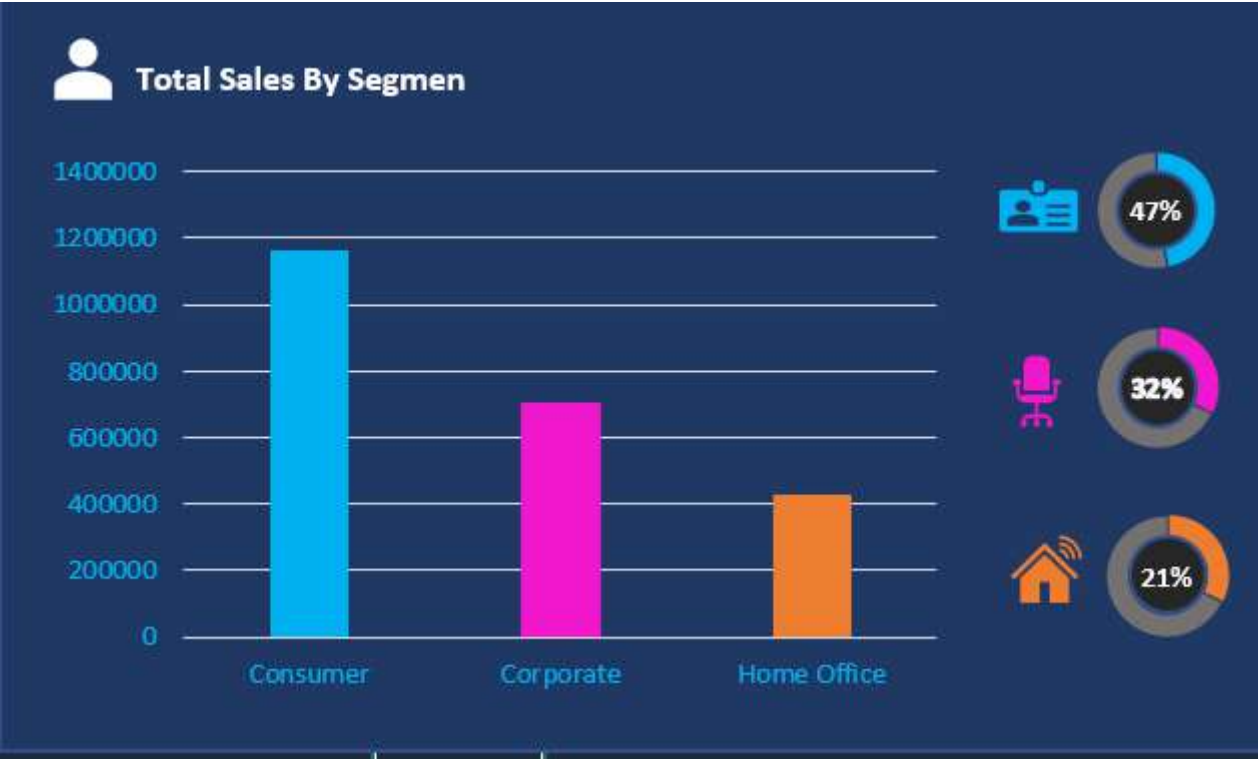
The total sales done the different segment and the highest sales done on the consumer segment and the second highest sales done on the corporate segment and the lowest sales done on the home office segment and on the right side of the graph and the percentage of the sales done on the each category compared to the total sales.
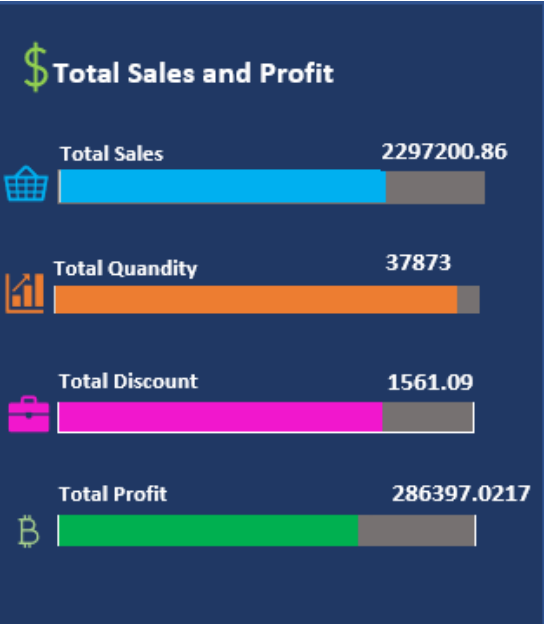
This graph show the total sales , quantity, discount, profit

Profit Sales by Region

Central	39706.3625 14%	East	91522.78 32%
South	46749.4303 16%	West	108418.4489 38%
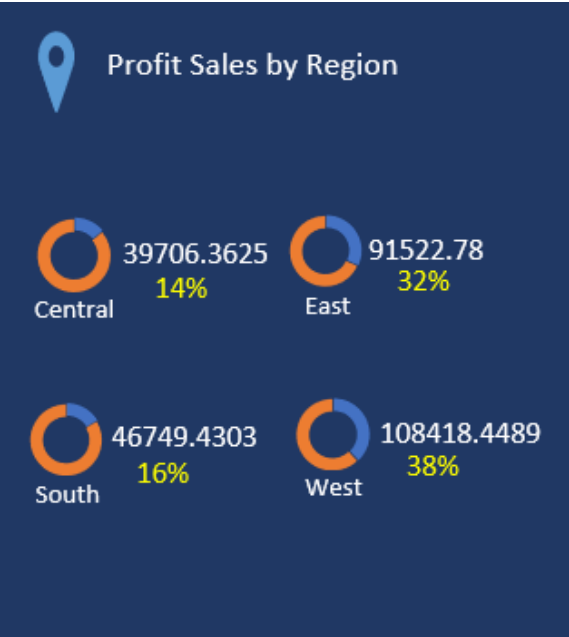
These graphs represents the profit persentage on the particular region compared to total other regions and the total profit on the particular region and the persentage on the profit on the particular region.

Sum of Ship Mode (Sales)

2297200.86
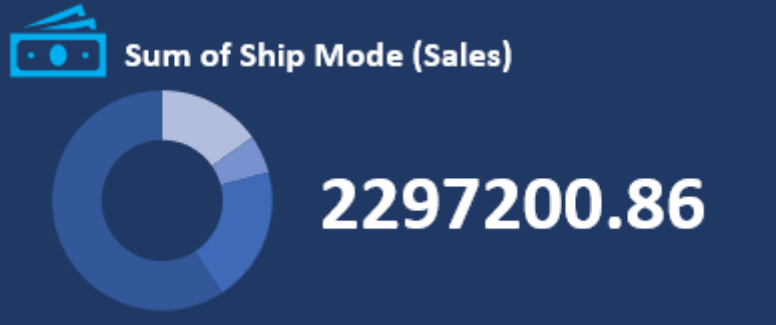
This graph is based on the saled done on the different shiping modes and the total sales is shown on the right side of the graph.


Sum of Ship Mode (Quantity)

37873

This graph is based on the shipping mode based on the category means how huch quandity is on the different shipping modes .
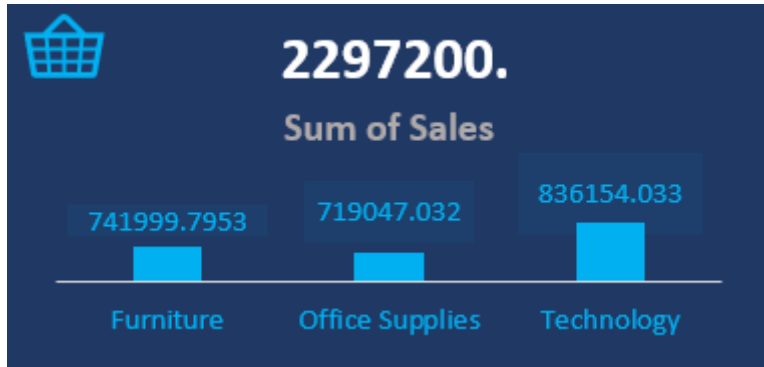
This graph based on the amount of sales based on the category (Furniture,Office Supplies,Technology).the quandity is high on the Technology category and the second highest sales is on the Office Supplies category and the lowest sales is on the Furniture sales category.


This graph based on the amount of quantity based on the category (Furniture,Office Supplies,Technology).the quandity is high on the Office Supplies and the second highest quantity is on the Furniture quantity category and the lowest quantity is on the Technology category.
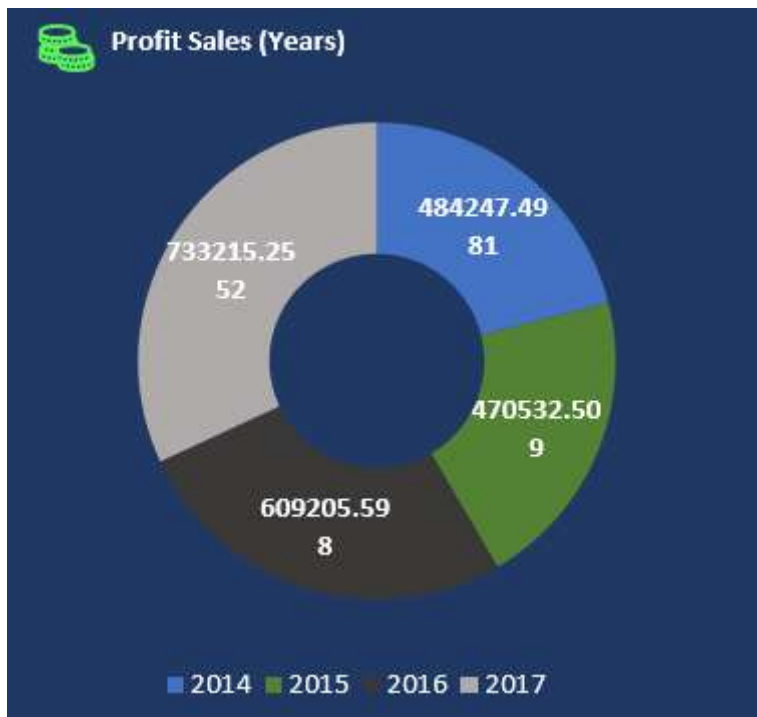
Tatal profit analysis based on years .profit is high on the year 2017 and second highest on the year 2016 and third highest on the year 2015 and lowest sales on the year on 2014.

Final view of Dashboard;
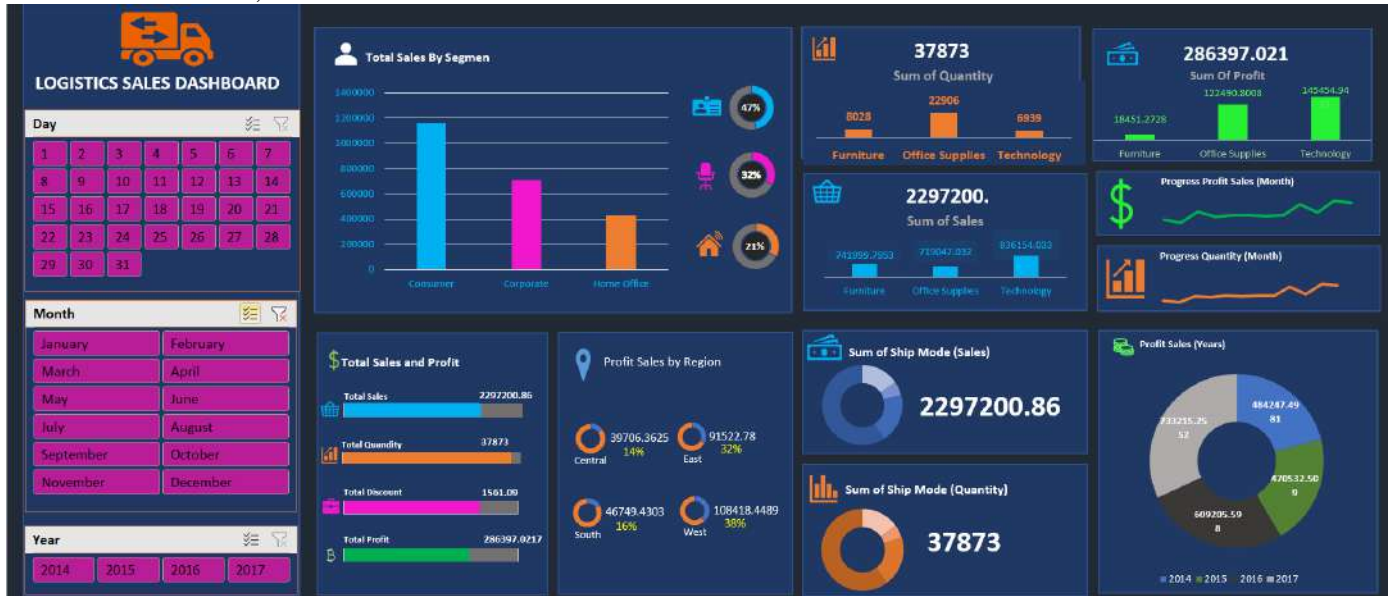
Bibliography

- ➢ YouTube

- ➢ Analytics Vidhya

- ➢ Kaggle

- ➢ Used MS Word 2016 Information from Google.com

- ➢ Learnt some things from Tutorial Point YouTube channel