<INT217> PROJECT REPORT

(Project Semester August-December 2022)

“ECOMMERCE SALES ANALYSIS”

Submitted by

‘PRATYAY DAS’

Registration No: 12209742

P26J:M.Sc. (Statistics and Data Analytics)

Under the Guidance of

Dr. Prakash Singh Tanwar(26199)

Associate Professor

Discipline of CSE/IT

Lovely School of Chemical Engineering & Sciences.

Lovely Professional University, Phagwara

CERTIFICATE

This is to certify that ‘Pratyay Das’ bearing Registration no. ‘12209742’ has completed INT217 project titled, **“ECOMMERCE SALES ANALYSIS”** under my guidance and supervision. To the best of my knowledge, the present work is the result of his/her original development, effort and study.

Signature and Name of the Supervisor
Designation of the Supervisor
School of Chemical Engineering And Sciences.
Lovely Professional University
Phagwara, Punjab.

Date: 27th November 2022.

DECLARATION

I, Pratyay Das student of LOVELY PROFESSIONAL UNIVERSITY (M.Sc. Statistics and Data Analytics) 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: 27th November 2022 Signature : Pratyay Das

Registration No. 12209742 Name of the student: Pratyay Das

ACKNOWLEDGEMENT

I would like to thank following people who helped me in completing my dissertation report.

First of all, I would like to appreciate and thanks my guide ……for giving me the opportunity to study and work under his expert guidance, encouragement and support through out the semester.

Secondly, I would like to thank my parents, for giving me the knowledge I asked and for the encouragement. Without their wishes this project wouldn't have been successful .I would also like to thank my friends for their support.

I am grateful to THE **LOVELY PROFESSIONAL UNIVERSITY** for giving me this opportunity.

CONTENTS:

OUR FINAL DASHBOARD

INTRODUCTION

Consumers are getting more comfortable every year with purchasing online or through their mobile phones. The sales based on online purchase are generally referred to as Ecommerce sales. Ecommerce stands for electronic commerce and refers to a digital platform and a business models where you can purchase products online. Every time you purchase a product online , you're participating in ecommerce economy.

In our given ecommerce sales analysis dashboard, we did a brief analysis for a given ecommerce sales dataset for years 2011, 2012, 2013, 2014 for the segments consumer, corporate and home office in the region Central, East, South and West.

So, in this project first of all we collected our dataset and then organised it according to the requirements and then with some formatting and formulas we constructed the pivot charts and made an interactive dashboard.

Using the dash board we visualized the various aspects of our objectives based on analysis using tables and different kinds of charts.

OBJECTIVES / SCOPES OF ANALYSIS:

An excel dashboard is an one-pager (mostly, but not always necessary) that helps managers and business leaders in tracking KPIs or metrics and take decision based on it. It contains charts , tables and views that are backed by data.

So in this project, I also tried to make an interactive dashboard with some objectives.

The objectives are as follows:

1. Sales And profit analysis.
2. Category-wise profit.
3. Category-wise sales share percentage.
4. Sales by State display with Map chart.
5. Top 5 subcategories by sales

SOURCE OF DATASET:

The dataset of Ecommerce sales is collected from GOOGLE SHEETS.

WEBSITE:

https://docs.google.com/spreadsheets/d/1L6aBX0uNlzKiJb7JHdkNUile18s9CI4r/edit#gid=1589100670

ETL PROCESS:

In computing, extract, transform, load (ETL) is a process to prepare data for analysis, especially in data warehousing. Data extraction involves extracting data from homogenous or heterogenous sources, while data transformation processes data by transforming them into a proper storage format/structure for the purposes for querying and analysis; finally, data loading describes the insertion of data into the final target location such as 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 used together, and finally delivers data in a presentation-ready format so that application developers and end users can make decisions.

It is a data integration process that combines data from multiple data sources into a single, consistent data source that is loaded into a data warehouse or other target systems.

a) **Extract** : During data extraction, raw data is copied or exported from source locations to a staging area.

b) **Transform** : In the staging area, the raw data undergoes data processing. Here, the data is transformed and consolidated for it's intended analytical use case. This phase can involve the following tasks:

i. Filtering, cleansing, de-duplicating, validating and authenticating the data.
ii. Performing calculations, transactions or summarizations based on raw data. This can include changing row and column headers for consistency, converting currencies or other other units of measurement, editing text strings and more.
iii. Conducting audits to ensure data quality and compliance.
iv. Removing, encrypting, or protecting data governed by industry or governmental regulators.
v. Formatting the data into tables or joined tables to match the schema of the target data warehouse.

c) **Load** : in this step, the transformed data is moved from the staging area into a target data warehouse. Typically, this involves an initial loading of all data, followed by periodic loading of incremental data changes and, less often, full refreshes to erase and replace data in the warehouse. For most organizations that use ETL, the process is automated, well-defined, continuous and batch-driven. Typically, ETL takes place during off-hours when traffic on the source systems and data warehouse is at it's lowest.

	Row ID	Order ID	Year	Order Date	Ship Date	Ship Mode	Customer ID	Customer Na	Segment	Country	City	St
9958		9957 US-2011-143287	2011	11-11-2011	17-11-2011	Standard Class	KN-16705	Kristina Nunn	Home Office	United States	New Rochelle	Ne
9959		9958 US-2011-143287	2011	11-11-2011	17-11-2011	Standard Class	KN-16705	Kristina Nunn	Home Office	United States	New Rochelle	Ne
9960		9959 US-2011-143287	2011	11-11-2011	17-11-2011	Standard Class	KN-16705	Kristina Nunn	Home Office	United States	New Rochelle	Ne
9961		9960 CA-2014-137421	2014	07-10-2014	12-10-2014	Standard Class	AJ-10945	Ashley Jarboe	Consumer	United States	Chandler	Ar
9962		9961 CA-2014-141446	2014	17-09-2014	19-09-2014	Second Class	CL-12700	Craig Leslie	Home Office	United States	Florence	Ke
9963		9962 CA-2012-168088	2012	19-03-2012	22-03-2012	First Class	CM-12655	Corinna Mitche	Home Office	United States	Houston	Te
9964		9963 CA-2012-168088	2012	19-03-2012	22-03-2012	First Class	CM-12655	Corinna Mitche	Home Office	United States	Houston	Te
9965		9964 CA-2012-143700	2012	26-07-2012	26-07-2012	Same Day	AS-10240	Alan Shonely	Consumer	United States	Philadelphia	Pe
9966		9965 CA-2013-146374	2013	06-12-2013	11-12-2013	Second Class	HE-14800	Harold Engle	Corporate	United States	Newark	De
9967		9966 CA-2013-146374	2013	06-12-2013	11-12-2013	Second Class	HE-14800	Harold Engle	Corporate	United States	Newark	De
9968		9967 CA-2013-146374	2013	06-12-2013	11-12-2013	Second Class	HE-14800	Harold Engle	Corporate	United States	Newark	De
9969		9968 CA-2014-153871	2014	12-12-2014	18-12-2014	Standard Class	RB-19435	Richard Bierne	Consumer	United States	Plainfield	Ne
9970		9969 CA-2014-153871	2014	12-12-2014	18-12-2014	Standard Class	RB-19435	Richard Bierne	Consumer	United States	Plainfield	Ne
9971		9970 CA-2014-153871	2014	12-12-2014	18-12-2014	Standard Class	RB-19435	Richard Bierne	Consumer	United States	Plainfield	Ne
9972		9971 CA-2012-103772	2012	28-06-2012	02-07-2012	Standard Class	MP-17470	Mark Packer	Home Office	United States	Smyrna	Ge
9973		9972 CA-2012-103772	2012	28-06-2012	02-07-2012	Standard Class	MP-17470	Mark Packer	Home Office	United States	Smyrna	Ge
9974		9973 CA-2013-130225	2013	12-09-2013	18-09-2013	Standard Class	RC-19960	Ryan Crowe	Consumer	United States	Houston	Te
9975		9974 US-2013-103674	2013	07-12-2013	11-12-2013	Standard Class	AP-10720	Anne Pryor	Home Office	United States	Los Angeles	Ca
9976		9975 US-2013-103674	2013	07-12-2013	11-12-2013	Standard Class	AP-10720	Anne Pryor	Home Office	United States	Los Angeles	Ca
9977		9976 US-2013-103674	2013	07-12-2013	11-12-2013	Standard Class	AP-10720	Anne Pryor	Home Office	United States	Los Angeles	Ca
9978		9977 US-2013-103674	2013	07-12-2013	11-12-2013	Standard Class	AP-10720	Anne Pryor	Home Office	United States	Los Angeles	Ca
9979		9978 US-2013-103674	2013	07-12-2013	11-12-2013	Standard Class	AP-10720	Anne Pryor	Home Office	United States	Los Angeles	Ca

THIS IS MY DATASET FOR ECOMMERCE SALES.

ANALYSIS ON THE DATASET

OBJECTIVE 1 : SALES AND PROFIT ANALYSIS (COMBO CHART)

This objective gives a brief analysis for the sales and profit of the products being sold. It is mandatory for any ecommerce platform to keep the track of the sales and profit earned to make crucial decisions, business leaders and managers can make decision based on the analysis.

We use pivot tables , pivot charts and dashboard visualization to do so.

Firstly we make a pivot table. The purpose of pivot table is to summarise the data.

We select the whole data set with **Ctrl+A**, then we go to insert and then click Pivot table, a dialogue box appears ‘Pivot table from table or range’, there we select ‘New Worksheet’ and then click ‘OK’.

A new sheet opens up with an optimizable ‘Pivot table fields’.

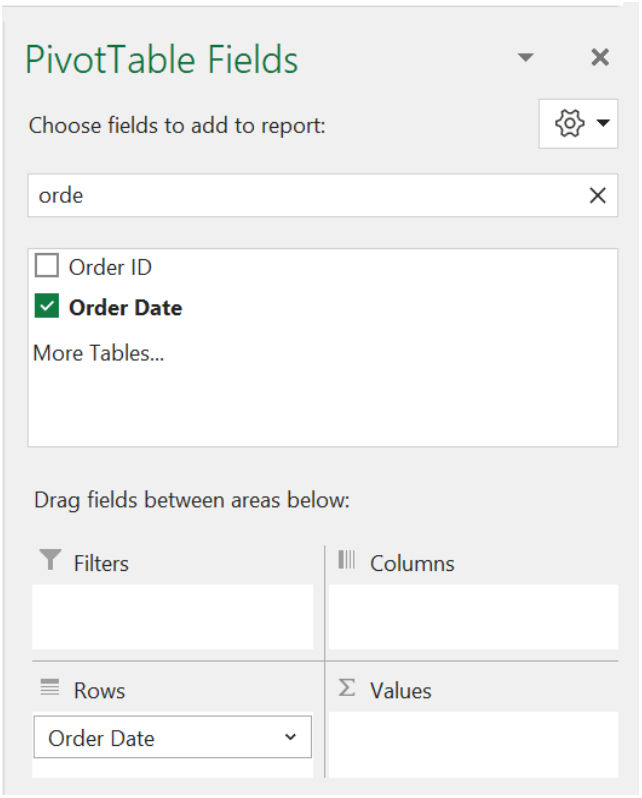


FIGURE 1.1

there itself we select ‘order date’, by selecting or dragging down ‘order date’, we add it to the ‘Rows’ box of “ Drag fields between areas below” and we obtain the following column.

Row Labels
Jan
Feb
Mar
Apr
May
Jun
Jul
Aug
Sep
Oct
Nov
Dec
Grand Total

FIGURE 1.2

Now we select ‘Profit’ AND ‘Sales’in a similar way and add it to the ‘values’ box.

By following up the above steps we are able to make a pivot table for the sum of sales and sum of profit for all the months and get the grand total of sum of sales and profits.

Row Labels	Sum of Sales	Sum of Profit
Jan	95365.9376	9199.27
Feb	60172.6494	10288.8987
Mar	199252.984	26806.5737
Apr	141851.5674	13437.8068
May	156122.2867	22308.5128
Jun	147082.6113	20899.3345
Jul	149580.838	13535.383
Aug	159589.451	21896.7154
Sep	309770.0997	37293.1925
Oct	197115.2017	31469.6638
Nov	349120.074	35825.5366
Dec	332177.1595	43436.1339
Grand Total	**2297200.86**	**286397.0217**

FIGURE 1.3

Since the values in the table are not in the proper manner with many numbers after the decimal places. So we try to format the numbers by adding Dollar ($) and converting it to thousands for better display and understanding , this would also be helpful to us when we make pivot chart out of the pivot table which we later put in the dashboard.

To do that, we select the data press **Ctrl+1** and the ‘format cells’ dialogue box appears and then we use custom formatting.

Where in the 'Type’ box we write ($0.00, “k”) and hence we a get a clean pivot table .

Row Labels	Sum of Sales	Sum of Profit
Jan	$3.08k	$0.43k
Feb	$0.67k	$0.23k
Mar	$3.24k	-$0.83k
Apr	$9.90k	$2.56k
May	$2.31k	$0.61k
Jun	$3.57k	-$0.02k
Jul	$4.15k	$0.51k
Aug	$3.27k	$0.39k
Sep	$3.00k	$0.36k
Oct	$6.13k	$0.67k
Nov	$3.44k	$0.29k
Dec	$10.53k	$2.65k
Grand Total	**53295.289**	**7848.9109**

FIGURE 1.4

Now we select a cell from the table, then click on ‘Pivot Table Analyze’ and then ‘Pivot Chart’.

Insert chart dialogue box appears and then we select combo.

We choose the type and axis for our data series and ‘OK’ and we get our combo chart which we further wish to use in our dashboard. Later we will add slicer to make the dashboard interactive. As this is a combo-chart so we will have three axes in total, one for sales, on for profits and the other for months for each year from 2011 to 2014. We will do some more required formatting for the chart after we add it to our respective dashboard.

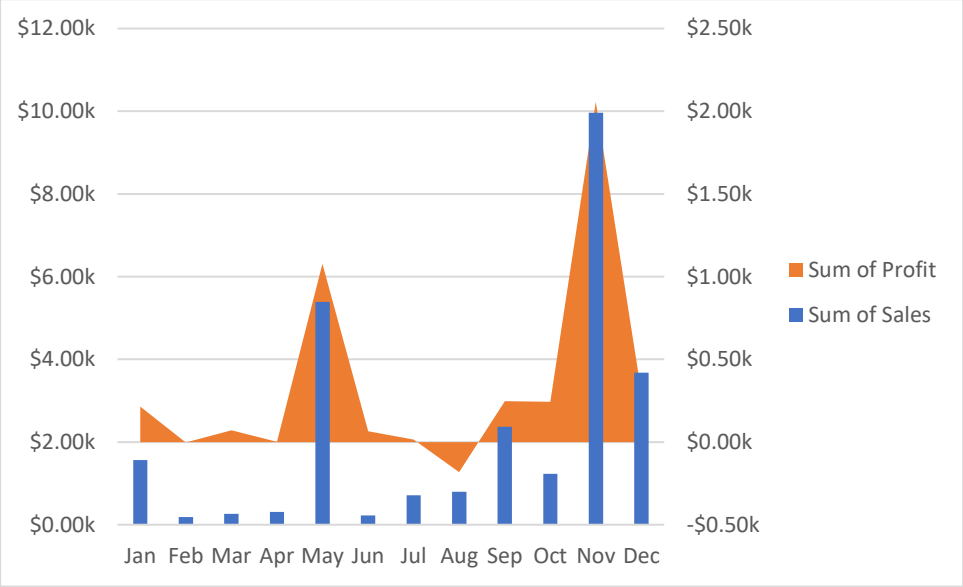


FIGURE 1.5

OBJECTIVE 2 : CATEGORY-WISE PROFIT ANALYSIS (WATERFALL CHART)

As we can see there are various categories and sub-categories in our given dataset of ecommerce sales. Thus it is an important aspect to make category-wise profit analysis.

Product profitability helps to determine the profit of the product we sell in the market.

We need to evaluate the profit while we already have total revenue because the revenue we get is pointless if we get no profit on the products that we sell.

We will do the same with the help of pivot table and chart. Here we will develop a waterfall chart to visualize our objective.

We will move to the sheet containing our dataset, we select the whole dataset with **Ctrl+A** , we go to insert and click pivot table, a dialogue box appears ‘Pivot table from table or range’, there we select ‘New Worksheet’ and then click ‘OK’.

A new sheet opens up with an optimizable ‘Pivot table fields’.

In the ‘PivotTable fields’, in the ‘Rows’ box we put ‘categories’ and in the ‘values’ box we put ‘profit’

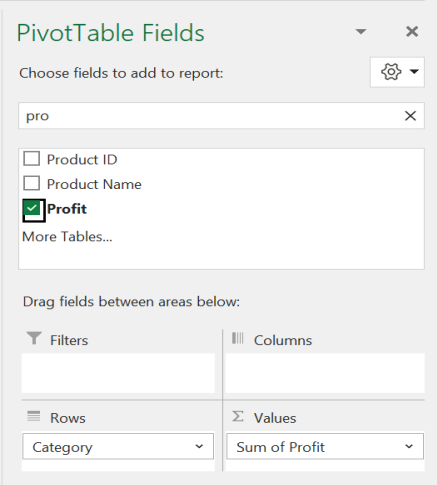


FIGURE 2.1

We obtain our pivot table:

Row Labels	Sum of Profit
Furniture	-131.1135
Office Supplies	1246.2762
Technology	2982.3425
Grand Total	**4097.5052**

FIGURE 2.2

but we wish to make a waterfall chart for the analysis and to make such kind of chart, we need to copy the data outside the pivot table and create another table for reference:

CATEGORY	PROFIT
Furniture	-$0.13k
Office Supplies	$1.25k
Technology	$2.98k
Grand Total	$4.10k

FIGURE 2.3

Here we can observe that the values in the ‘Profit’ column being formatted , we did this by following up these steps:

i. Ctrl+1
ii. We move to conditional formatting.
iii. In the ‘type box’ we type ($0.00, “k”)
iv. Press ‘OK’ .

We use the second table , we select the table, then we go to insert and select waterfall chart and thus we have our required chart:

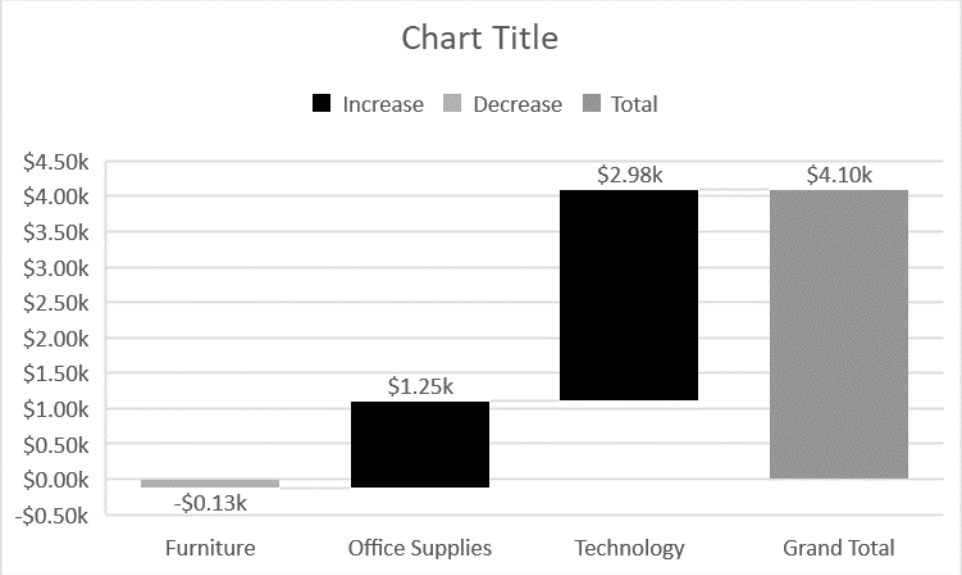


FIGURE 2.4

OBJECTIVE 3: <u>CATEGORY-WISE SALES SHARE PERCENTAGE (PIE-CHART)</u>

We have three different categories of products which are furniture, office supplies and technology. In this objective we use pivot table and pie chart to visualize category-wise sales share percentage , which is an important analysis to derive inferences.

Whereas sales per share is a ratio that computes the total revenue earned per share over a designated period, whether quarterly, semi annually, or trailing twelve months.

We do so by repeating the same steps as earlier :

1. Select the whole table with **CTRL+A.**
2. Go to insert and click pivot table.
3. ‘Pivot table from table or range’ dialogue box appears.
4. We select new worksheet
5. Press ok.

We are moved to a new sheet .

In the pivotable fields , we put ‘category’ in the row box and in the values box we take ‘sales’ to get the table:

Row Labels	Sum of Sales
Furniture	741999.7953
Office Supplies	719047.032
Technology	836154.033
Grand Total	**2297200.86**

FIGURE 3.1

There is requirement for formatting , we follow the following steps:

1. Right click on any cell from the sales columns.
2. Select ‘Show value as’
3. Then click ‘% of Grand Total’.
4. We get all the sale column values as percentage.

Row Labels	Sum of Sales
Furniture	28.77%
Office Supplies	33.03%
Technology	38.20%
Grand Total	**100.00%**

FIGURE 3.2

Next part is to make pie chart using the following table.

1) We select the table
2) Click pivot table analyze.
3) Select Pie.
4) We go for the doughnut shape
5) And Ok.

Following the steps we obtain our pie chart which requires some formatting.

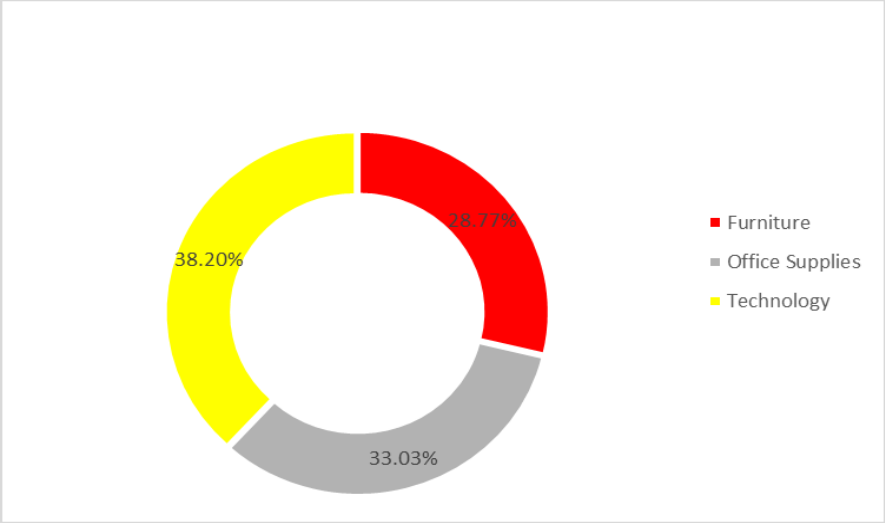


FIGURE 3.3

We added the data labels and we formatted the data series by increasing the doughnut hole size and increasing the doughnut explosion size.

OBJECTIVE 4 : SALES BY STATE (MAP CHART)

This is the simplest objective in terms of understanding where we try to display the sales for different years in different regions of different categories of products.

We use pivot table and map chart to do so.

To make the pivot table we will follow the same steps we have been following so far.

1. Go to insert and click pivot table.
2. 'Pivot table from table or range' dialogue box appears.
3. We select new worksheet
4. Press ok.

In this case we will take states and sales in the pivot table fields as per as our requirement.

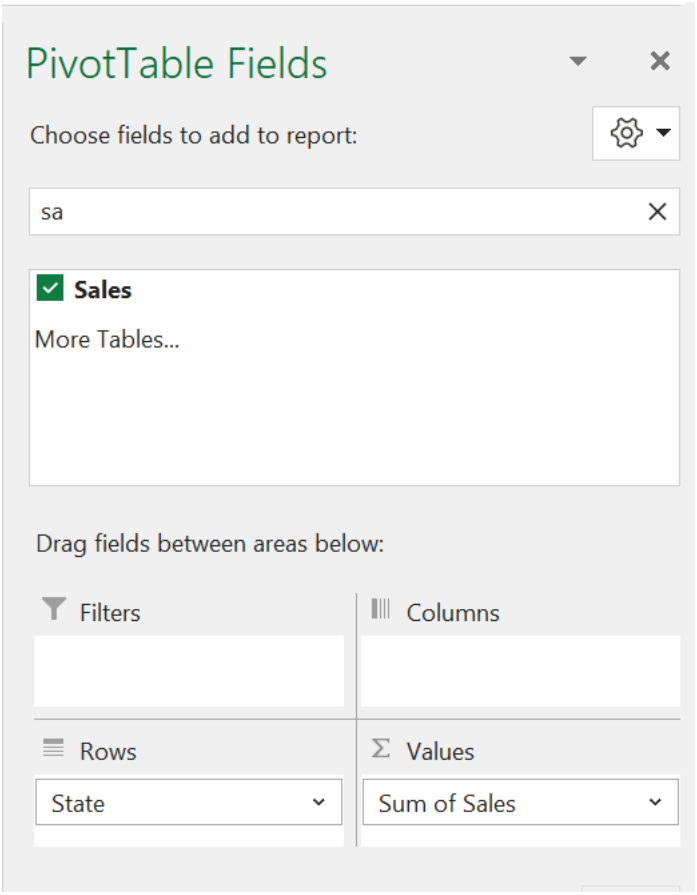


FIGURE 4.1

We obtain our pivot table , but we cannot obtain a map chart using a pivot table thus we are required to make a separate table and we just do not copy the matter from the pivot table to another table.

What we do here is make two separate columns one for states and the other for sum of sales,

We wish to build the table using formulas so we type '=' and the select a cell from the pivot table and press OK , this will get our values copied to our new table .

Our next step is to format the table, with the help of formatting we convert the sales values in terms of Dollars by using custom formatting from the format cells dialogue box which appears upon pressing Ctrl + 1

After that we select the table :

Insert > Maps > Filled Map

We obtain our map chart:

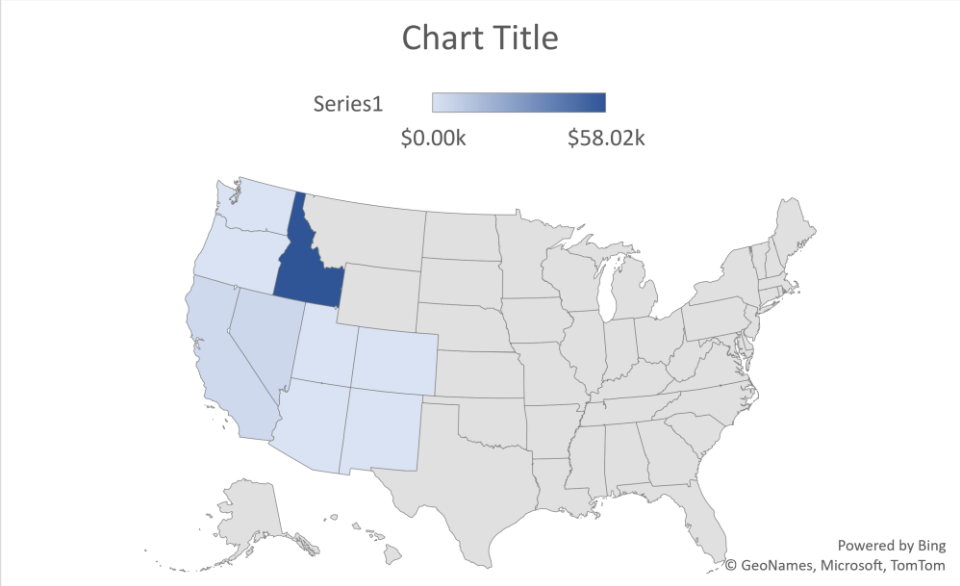


FIGURE 4.2

This is a dynamic chart, so it will contribute in developing an interacting dashboard.

OBJECTIVE **5: TOP 5 SUB-CATEGORIES BY SALES**

In our given dataset we have various sub-categories of the products , sub-categories helps us to understand the sales pattern and trend more accurately.

There are five different sub-categories that are present in our dataset namely Accessories, binders, phones, tables and copiers.

By using bar chart we will visualize our objective.

At first we create a pivot table using the following steps:

1. Go to insert and click pivot table.
2. ‘Pivot table from table or range’ dialogue box appears.
3. We select new worksheet
4. Press ok.

In the pivotTable field we put ‘sub-category’ in the row box and ‘sales’ in the values box.

Now in the table itself at Row labels we can find a drop down arrow, we click it and then click value filters and then ‘top 10’.

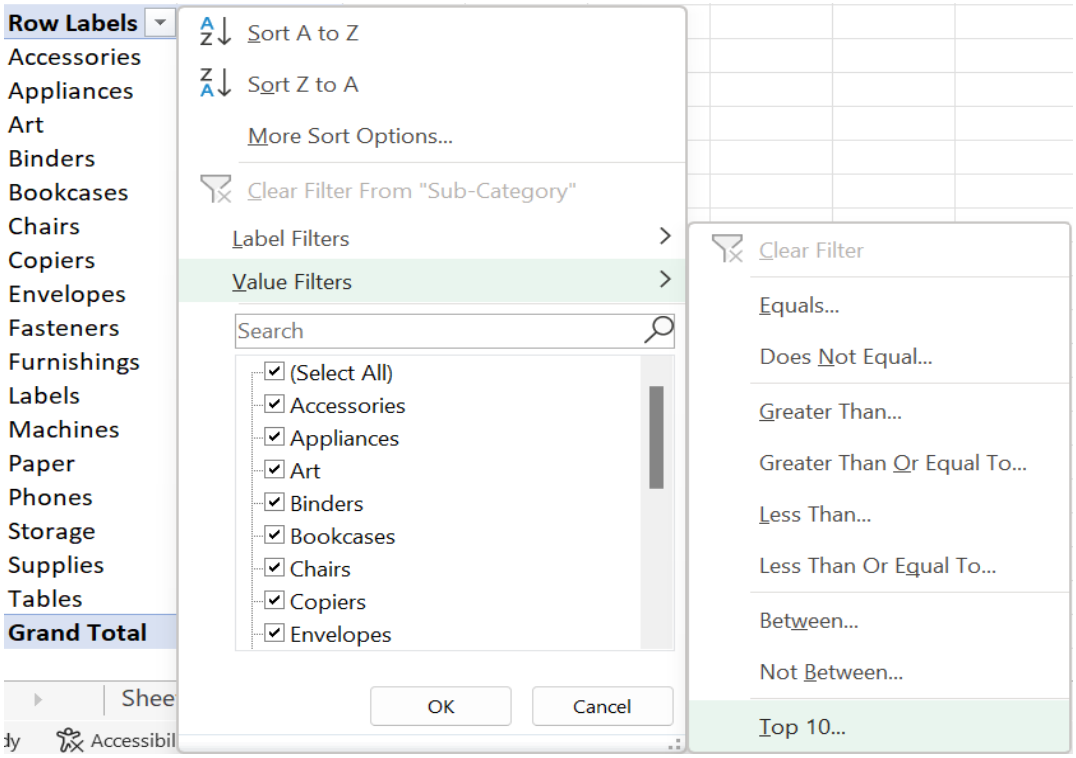


FIGURE 5.2

A top 10 sub-category filter (Sub-category) dialogue box appears, we reduce the top 10 value to 5 and press ‘OK’.

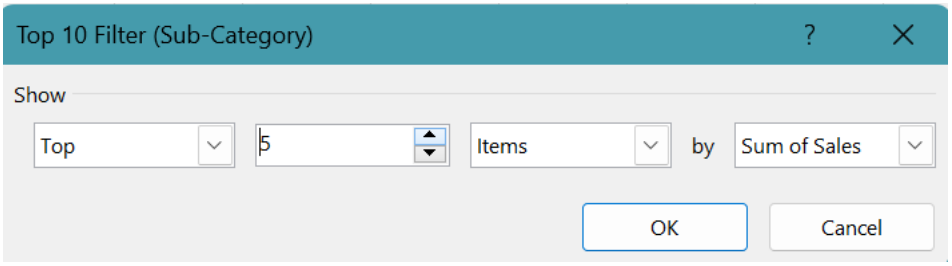


FIGURE 5.3

And thus we obtain our required pivot table:

Row Labels	Sum of Sales
Binders	203412.733
Chairs	328449.103
Phones	330007.054
Storage	223843.608
Tables	206965.532
Grand Total	**1292678.03**

FIGURE 5.4

Now we create a bar chart using the pivot table. We do it in the following steps:

1. Select any cell from the Pivot table
2. Click pivot table analyze.
3. Click pivot chart
4. From the insert chart dialogue box we select Bar.

And thus we obtain out required chart for the objective:

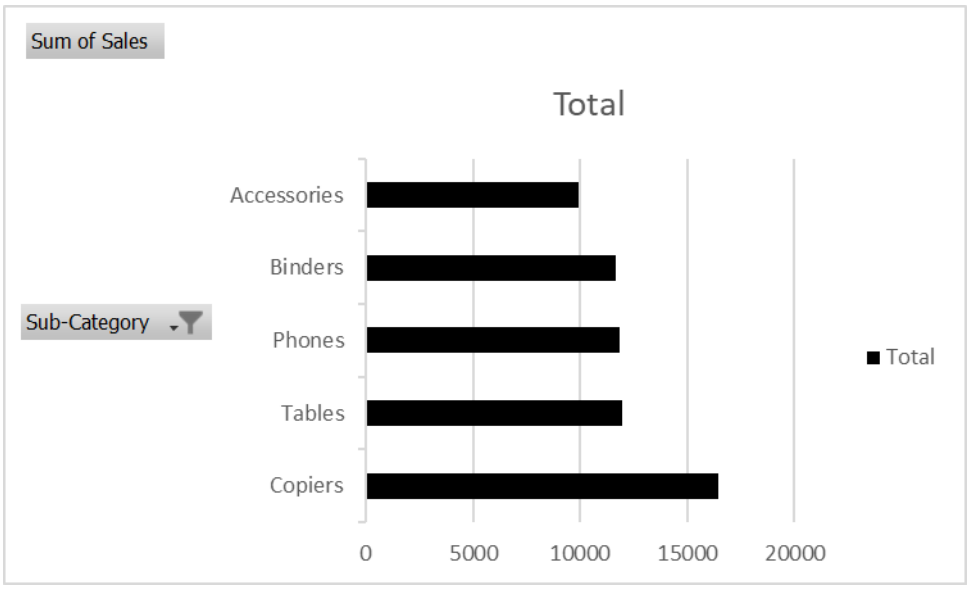
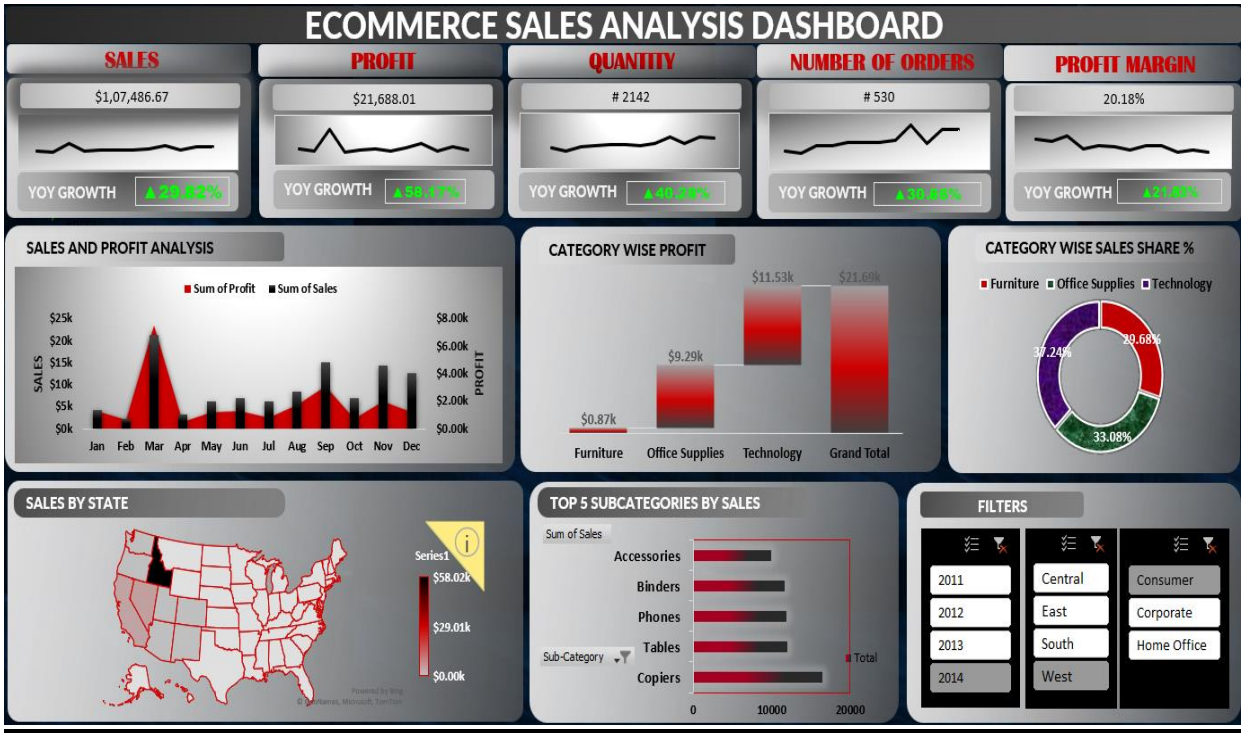


FIGURE 5.5

In our dashboard we also visualised Year on Year growth using line charts and some required formulas.

**Year on year growth = (Sales for the year 2014 / Sales for the year 2015) – 1

OUR FINAL DASHBOARD



REFERENCES:

1) DATASET FROM GOOGLE.
2) HOW TO BUILD INTERACTIVE DASHBOARD (YOUTUBE)
3) LINKEDIN FOR EXCEL FORMULAS
4) Excel Dashboards and Reports - Michael Alexander, John Walkenbach - Google Books
5) https://databox.com/ecommerce-analytics-guide

BIBLIOGRAPHY

[1] **M. &. W. J. Alexander, Excel Dashboards and Reports, 2013.**

[2] **M. Alexander, Excel 2019 Bible, Wiley, 2019.**

[3] **G. Harvey, Excel 2019 All-in-one For Dummies, For Dummies, 2019.**

[4] **M. GIrvin, Slaying Excel Dragons, Holy Macro: Books.**

[5] **M. Humphrey, Excel For Biginners, Independently Published.**

[6] **W. Winston, Microsoft Excel 2019 Data Analysis and Business Modeling, USA: Microsoft Press, 2019.**

[7] **B. Jelen and Alexander, Michael, Microsoft Excel 2019 Pivot Table Data Crunching, USA: Microsoft Press, 2019.**

[8] **R. Bovey, Professional Excel Development, Addison-Wesley Professional; 2nd edition (6 May 2009), 2009.**

[9] **B. Jelen, Excel 2016 VBA and Macros, Que Publishing, 2016.**

[10] **P. McFedries, Microsoft Excel 2019 Formulas and Functions, USA: Microsoft Press, 2019.**

[11] **M. Alexander, Excel VBA Programming for Dummies, USA: For Dummies, 2018.**