ANALYTICS TOOL FOR E-COMMERCE BUSINESSES

1. INTRODUCTION

a. Overview

Data is a valuable asset in the 4G era. Wide range of analysis has to done by any organization in order to take decisions. Data in its raw form or in the form of excel sheet is very difficult to analyze and interpret.

Organizations have to process huge amounts of data which is generated each second. This data has to be processed and interpreted in order to take effective decisions. Important decisions that lead to the growth and upliftment of the organization are made by analyzing data.

With the increase in consumer demand, the E-commerce space has boomed. This also leads to an increase in fierce competition in today's online marketplace. The ecommerce industry sells a diverse product line of grocery items and merchandise products, such as food, pharmaceuticals, apparel, games and toys, hobby items, furniture and appliances. The analysis of such industry is of great importance as it gives insights for the sales and profits of various products.

b. Purpose

The main objective of the project is to develop an Analytical Dashboard where the owner/user can understand the growth/potential of the business in the market. The dataset that is used in this project is US super store data.

The Analytics tool that is being developed for the e-commerce business is Analytical dashboard. It monitors the ecommerce performance without

missing even a small detail. This dashboard gives the ecommerce team a clear overview about various aspects. It helps in tracking and analyzing the core elements in real time on day-to-day basis.

By using various metrics it is possible to potentially track various aspects of business including details of profitable products, regions, sales trends, order details, products that incurred loss. It can also be used to forecast the future sales based on the analysis of current sales. With the help of this analysis ecommerce businesses can identify various aspects of business and trends and take decisions appropriately.

2. LITERATURE SURVEY

a. Existing problem

Data is very complicated. It can be usually present in the excel sheet. In order to make sense of the data to take useful decisions, the data that is present in these excel sheets has to be interpreted. This is very tedious and it is also difficult to analyse the data and identify the pattern or trend in it.

Disadvantages of the traditional methodology for data analysis are:

- It is very difficult to make complex data analysis.

- Quick analysis of data is not possible.
- It may not be accurate.
- It limits visibility when creating complex models.
- It makes collaboration more challenging.
- It is inefficient in managing templates and data entry.
- Keeping track of multiple spreadsheets is challenging.

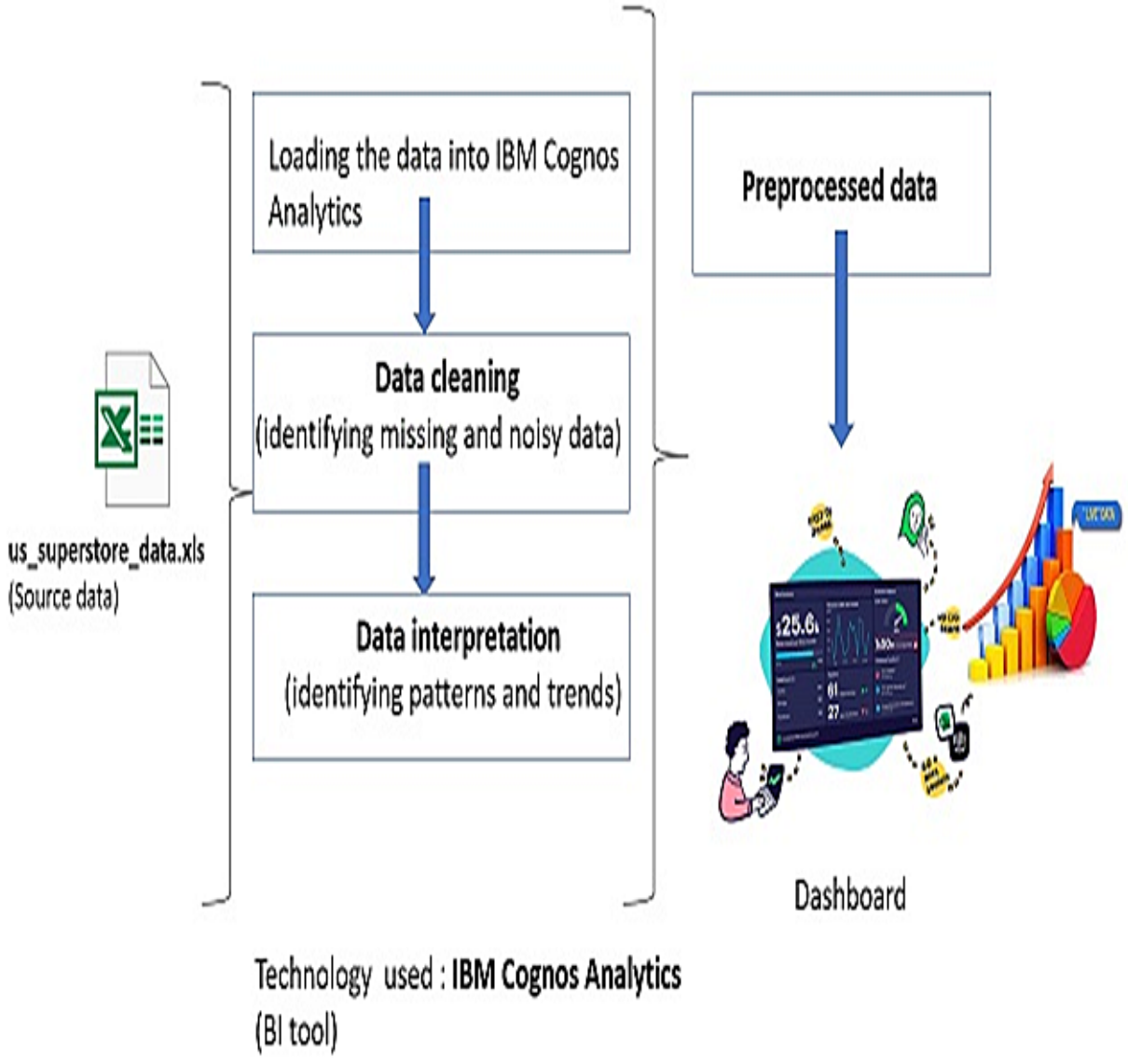
2.2 Proposed solution

Dashboard enables us to view the performance at a glance in various fields. It integrates sales and marketing performance and gives perfect insights. The real time business dashboard will save enterprise time, money and provides more clarity. It provides clear, reliable and up-to-date information to make important business decisions.

By using dashboard marketing assistant or staff doesn't need to spend time on various excel operations to get insights as dashboard provides updates automatically.

3. THEORITICAL ANALYSIS

a. Block diagram



b. Software designing

The software that is used for designing of the dashboard is IBM Cognos. IBM Cognos is a business intelligence tool that is used for reporting and analytics.

The other alternative that can be used for the project is IBM Watson Studio

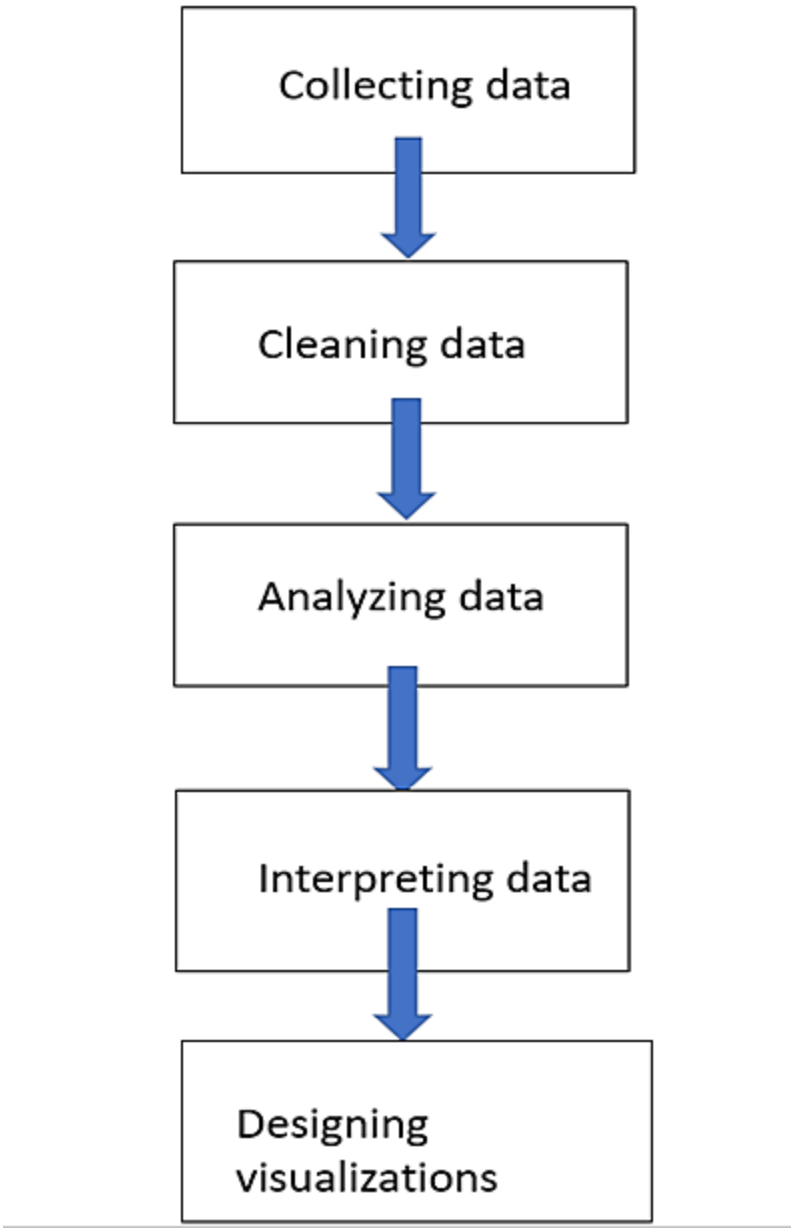
4. EXPERIMENTAL INVESTIGATIONS

	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
1	Customer ID	Customer Na	Segment	Country	City	State	Postal Code	Region	Product ID	Category	Sub-Categor	Product Nam	Sales	Quantity	Discount	Profit
2	CG-12520	Claire Gute	Consumer	United State	Henderson	Kentucky	42420	South	FUR-BO-100(	Furniture	Bookcases	Bush Somers	261.96	2	0	41.9136
3	CG-12520	Claire Gute	Consumer	United State	Henderson	Kentucky	42420	South	FUR-CH-100(	Furniture	Chairs	Hon Deluxe F	731.94	3	0	219.582
4	DV-13045	Darrin Van H	Corporate	United State	Los Angeles	California	90036	West	OFF-LA-1000	Office Suppli	Labels	Self-Adhesive	14.62	2	0	6.8714
5	SO-20335	Sean O'Donn	Consumer	United State	Fort Lauderd	Florida	33311	South	FUR-TA-1000	Furniture	Tables	Bretford CR4	957.5775	5	0.45	-383.031
6	SO-20335	Sean O'Donn	Consumer	United State	Fort Lauderd	Florida	33311	South	OFF-ST-1000(	Office Suppli	Storage	Eldon Fold 'N	22.368	2	0.2	2.5164
7	BH-11710	Brosina Hoffi	Consumer	United State	Los Angeles	California	90032	West	FUR-FU-1000	Furniture	Furnishings	Eldon Expres	48.86	7	0	14.1694
8	BH-11710	Brosina Hoffi	Consumer	United State	Los Angeles	California	90032	West	OFF-AR-1000	Office Suppli	Art	Newell 322	7.28	4	0	1.9656
9	BH-11710	Brosina Hoffi	Consumer	United State	Los Angeles	California	90032	West	TEC-PH-1000	Technology	Phones	Mitel 5320 IF	907.152	6	0.2	90.7152

- The dataset consists of US superstore data. It consists of orders in different states in the regions of South, Central, West, East.
- It also contains the details of several products with the details of the categories. Each category has a list of sub-categories in it.

- The sales details are provided in dollars along with the quantity ordered, discount, profit/loss details.
- In the dataset the details of 9.99K orders are provided. All the orders are of United States corresponding to different states in a particular region.
- The details of loss is included in the profit field with negative value
- Out of the total number of orders most of them correspond to west of United States. Details of several states in the four regions are also provided.

5. FLOWCHART

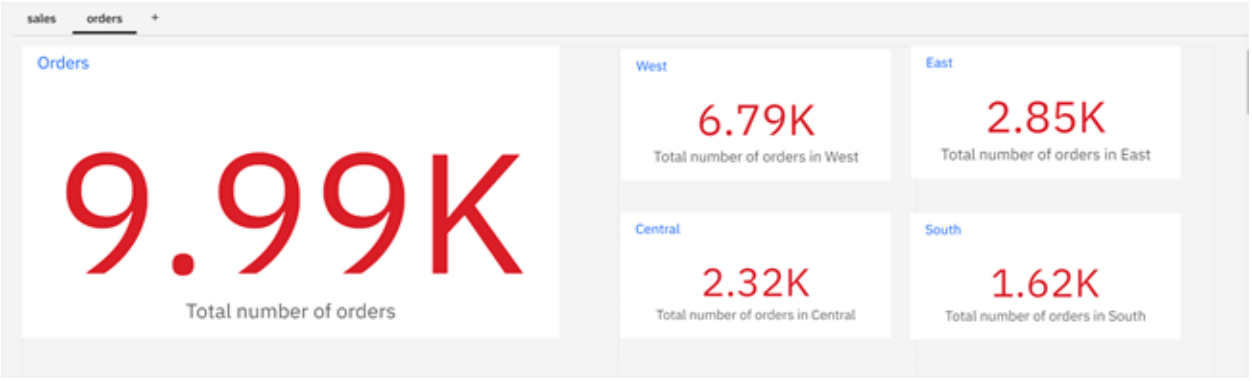

Collecting data
Cleaning data
Analyzing data
Interpreting data
Designing visualizations

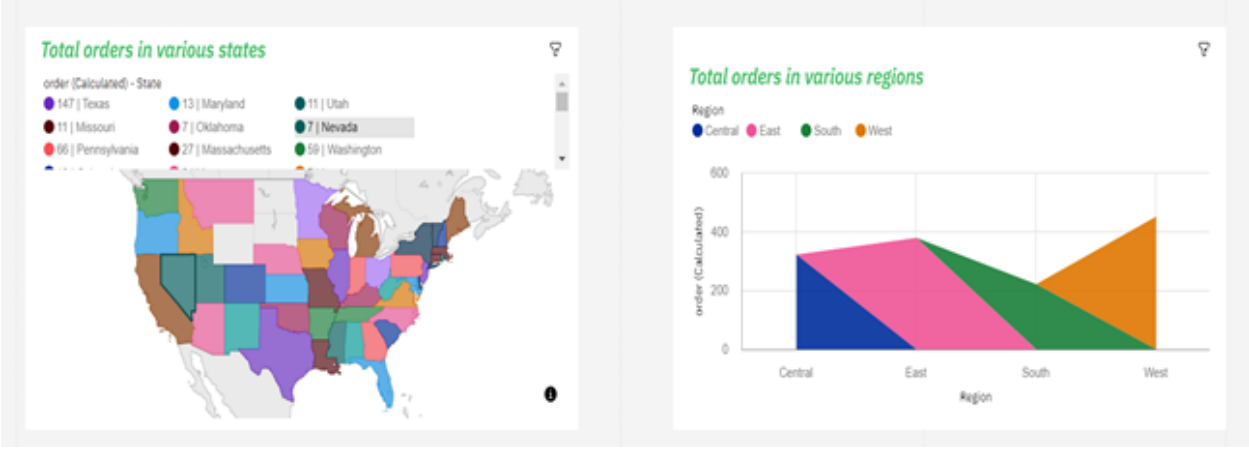
- Collection of data required for analysis from various data sources.
- Loading data into IBM Cognos Analytics.
- Performing data cleaning which involves handling missing and noisy data.
- Performing data interpretation so that useful insights can be obtained from it.
- Finally building the dashboard which provides the details for decision making.

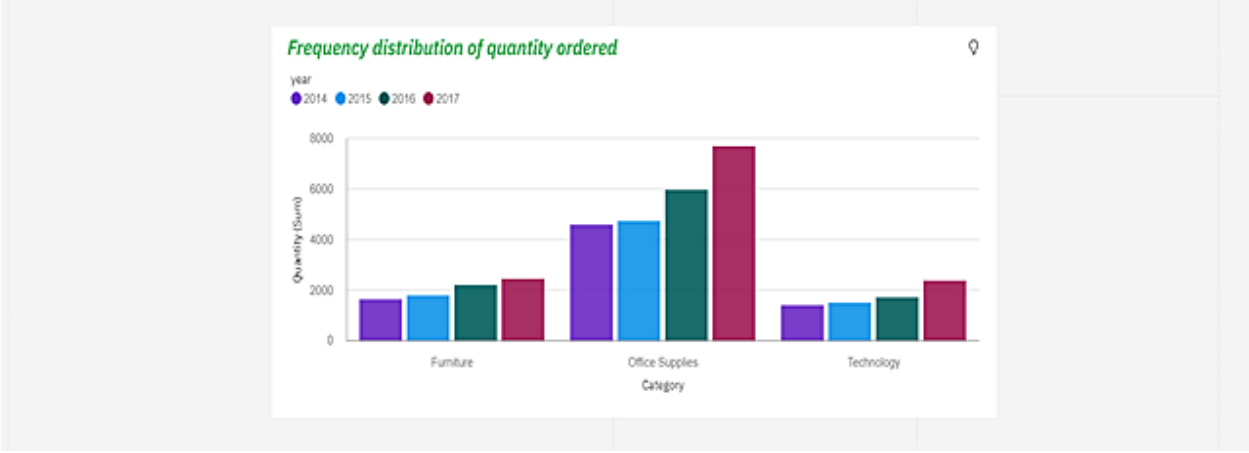
6. RESULT

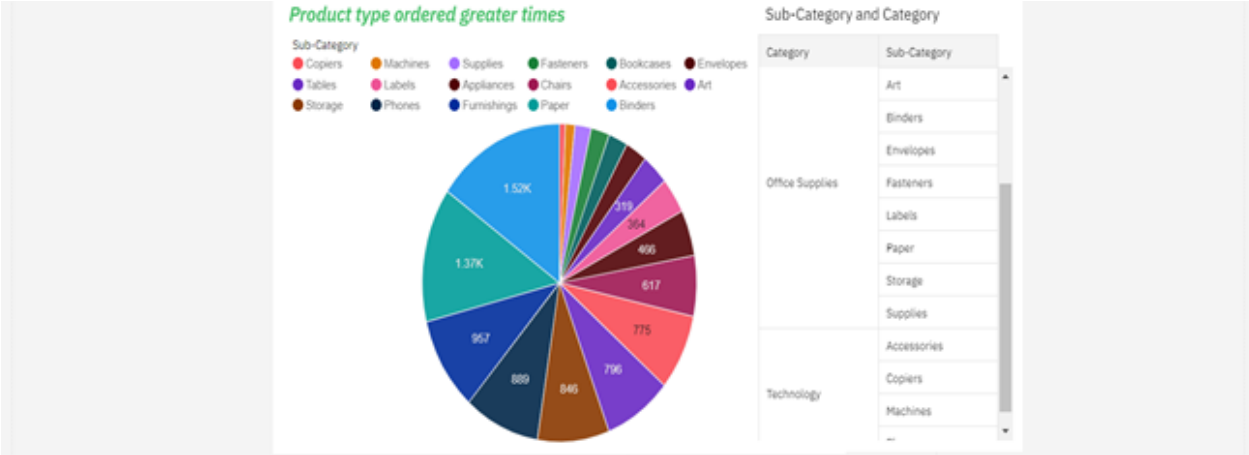
The dashboard contains two tabs. One corresponding to the sales and the other corresponding to the orders.

Orders tab:

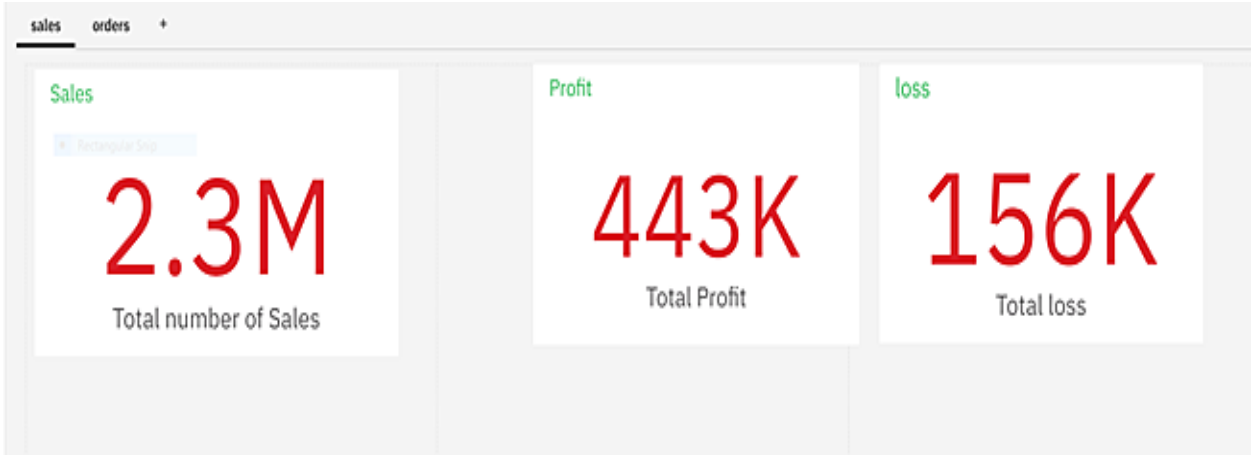



Total orders in various states
order (Calculated) - State
147 | Texas
13 | Maryland
11 | Utah
11 | Missouri
7 | Oklahoma
7 | Nevada
66 | Pennsylvania
27 | Massachusetts
59 | Washington
Total orders in various regions
Region
Central
East
South
West
600
400
200
0
order (Calculated)
Central
East
South
West
Region

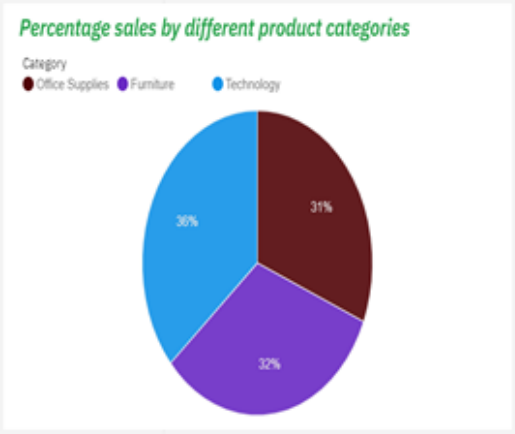

Frequency distribution of quantity ordered
year
2014
2015
2016
2017
8000
6000
4000
2000
0
Quantity (Sum)
Furniture
Office Supplies
Technology
Category

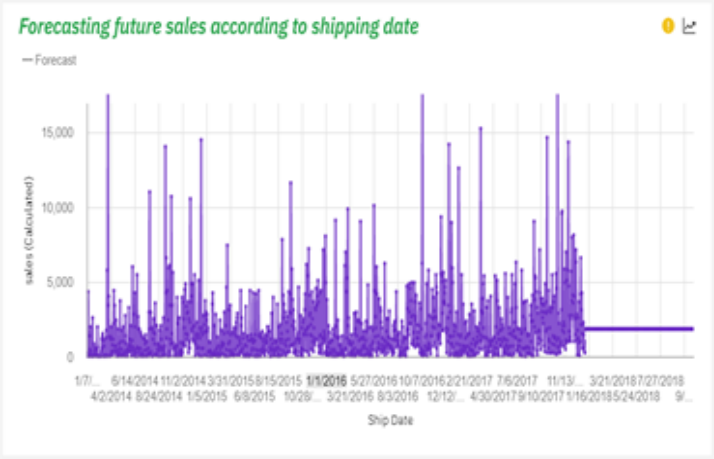

Product type ordered greater times
Sub-Category
Copiers
Machines
Supplies
Fasteners
Bookcases
Envelopes
Tables
Labels
Appliances
Chairs
Accessories
Art
Storage
Phones
Furnishings
Paper
Binders
1.52K
1.37K
957
889
846
796
775
617
496
384
319
Sub-Category and Category
Category
Sub-Category
Office Supplies
Art
Binders
Envelopes
Fasteners
Labels
Paper
Storage
Supplies
Technology
Accessories
Copiers
Machines

Sales tab:

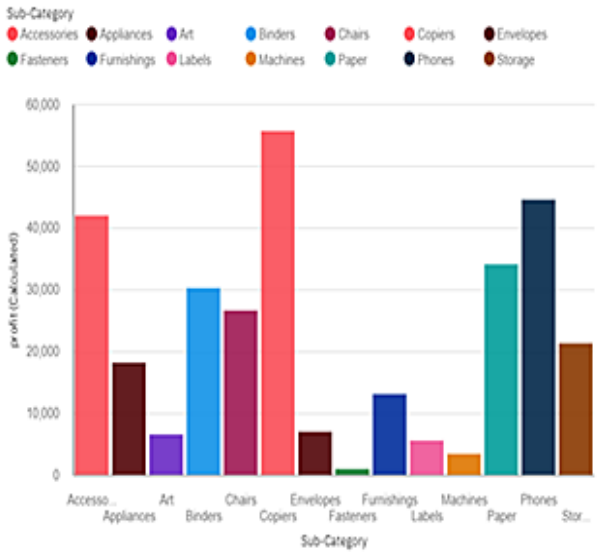

sales
orders
+
Sales
2.3M
Total number of Sales
Profit
443K
Total Profit
loss
156K
Total loss


Yearly sales for various states
Sales in Dollar (Calculated) - State
56,729.35 | Texas
3,740.02 | Utah
7,401.72 | Missouri
6,561.13 | Oklahoma
5,576.37 | Nevada
7,901.84 | Maryland
534.38 | Wyoming
46,213.76 | Washington
38,837.3 | Pennsylvania

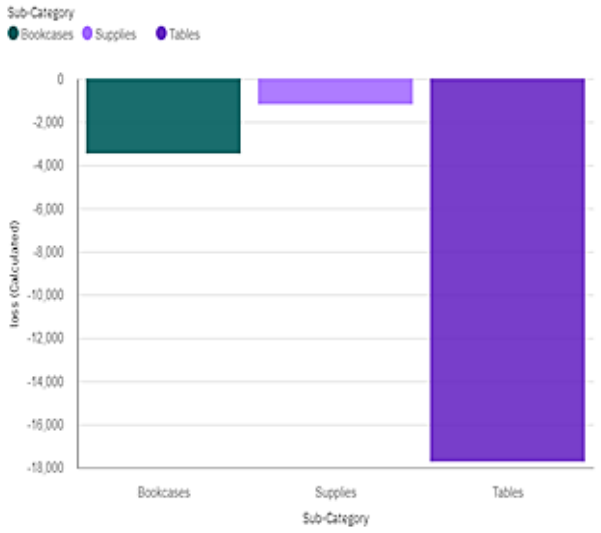

Percentage sales by different product categories
Category
Office Supplies
Furniture
Technology
31%
36%
32%

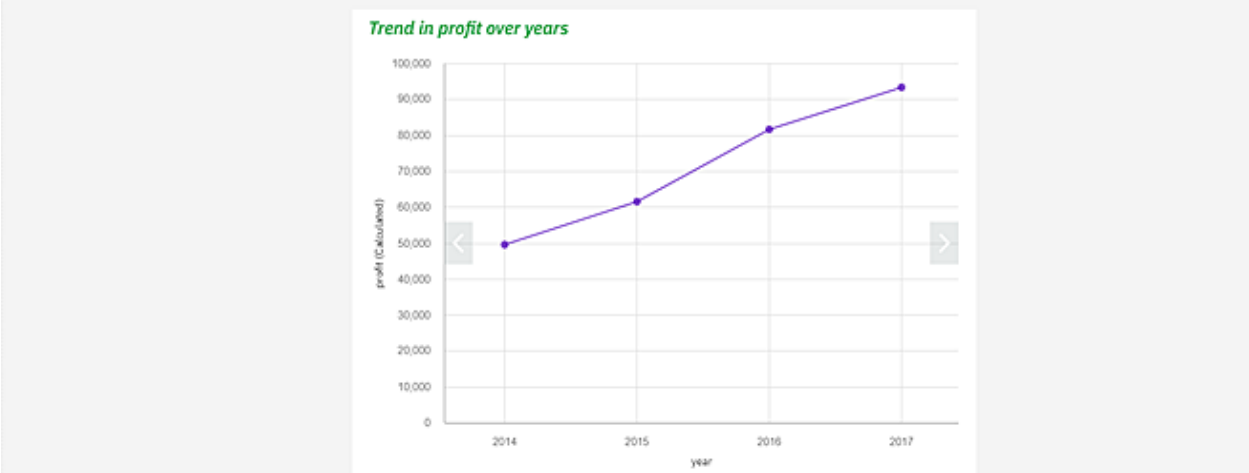

Forecasting future sales according to shipping date
Forecast
15,000
10,000
5,000
0
sales (Calculated)
Ship Date

Profitable products(profit in dollars)



Products that incurred loss(loss in dollars)





7. ADVANTAGES & DISADVANTAGES

The Analytics tool that is being developed for the e-commerce business is Analytical dashboard.

7.1 Advantages

- It monitors the ecommerce performance without missing even a small detail.
- This dashboard gives the ecommerce team a clear overview about various aspects.
- It helps in tracking and analyzing the core elements in real time on day-to-day basis.
- Using various metrics it is possible to potentially track various aspects of business including details of profitable products, regions, sales trends, order

details, products that incurred loss.

- It can also be used to forecast the future sales based on the analysis of current sales.
- It also provides the functionality of filtering, sorting and formatting of data.
- With the help of this analysis ecommerce businesses can identify various aspects of business and trends and take decisions appropriately.

7.2 Disadvantages

- Total cost of IBM Cognos is more than other tools.
- It has minimal forecasting capability.
- May not work smoothly with large datasets.

8. APPLICATIONS

Creation of dashboard using IBM Cognos Analytics can be done for analyzing the data related to e-commerce or retail.

Analytical dashboard is the absolute solution to visualize the data with ease.

Dashboard gives a huge benefit to any organization. Dynamic dashboards enable easy and flexible access to the required information to many people. Dashboard provides stunning visualizations of data in the real time. It depicts patterns, relations and other findings.

Satisfying the customer needs is very important for any ecommerce business. For this analysis of the industry in various aspects like sales and

profit is very helpful.

9. CONCLUSION

Analytical dashboards are the most suitable and stands to be the perfect solution of summarizing data. Usually, data is very large and in order to take decisions on such vast data, visualizations are generally used. The pictorial representation of data is very easy to understand and interpret. Analytical dashboards provide trends of various metrics over time, highly interactive and enable deeper analysis.

Analytical dashboards are the most suitable and stands to be the perfect solution of summarizing data. Usually, data is very large and in order to take decisions on such vast data, visualizations are generally used. The pictorial representation of data is very easy to understand and interpret. Analytical dashboards provide trends of various metrics over time, highly interactive and enable deeper analysis.

10. FUTURE SCOPE

In future dashboards with multiple, disparate data streams with high computing power can be built which can be used to handle volume of big data and artificial intelligence to make sense of it. The use of data visualization may increase both for low and high complexity analytics. Use of real time dashboards can lead to updations in few milliseconds. And for

an ecommerce portal prioritizing and solving issues at right time is vital and can be done with dashboard.

11. BIBLOGRAPHY

References:

https://www.ibm.com/docs/en/cognos-analytics/11.1.0?topic=dashboards-cognos-analytics-tutorial

https://www.ibm.com/topics/business-intelligence

https://www.idashboards.com/blog/2016/09/26/what-is-the-purpose-of-a-dashboard-2/