Analytics tool for E-commerce businesses

INTRODUCTION TO PROJECT

Overview

The surge in consumer demand has propelled the E-commerce sector into a period of rapid growth. Consequently, competition within the online marketplace has intensified significantly. The E-commerce industry encompasses a wide array of products, including groceries, pharmaceuticals, apparel, games, toys, hobby items, furniture, and appliances. Analyzing this industry is crucial as it provides valuable insights into product sales and profitability.

For this project, I developed a Dashboard utilizing the US Superstore Dataset within the IBM Cognos Analytics platform. The aim was to present various statistics to help retailers uncover insights such as:

- Region that accounts for a greater number of orders.
- Frequency distribution of quantity ordered.
- Percentage sales by different product class.
- Profitable products or their sub products in the last few years.
- Products that incurred losses.
- Frequently order products
- Yearly sales for various states.
- Forecasting future sales according to shipping date.
- Trend in profit/sales over time (years/months/quarters).
- No of days taken by different product class to get delivered.

With this analysis, the E-commerce business can key out various aspects of the market and trend pattern and take the required measures.

Purpose

The challenge can be addressed by conducting thorough research and competitive analysis. It's crucial to categorize competitors and consistently monitor and evaluate various aspects of their websites, including pricing, design, product offerings, and marketing strategies. There's much to be gained from studying both their strengths and weaknesses.

Dashboards empower users, particularly sellers, to actively engage in the analytics process by organizing data and presenting trends and insights visually. These data dashboards offer an unbiased perspective on performance metrics and lay a solid groundwork for further discussions.

In order to assist sellers in comprehending the true potential of their business within the market, I am embarking on the creation of an e-commerce analytics dashboard utilizing the robust capabilities of IBM Cognos Analytics.

Existing Problem

The main driver of retail growth for any business today is its E-Commerce store. However, the remarkable expansion witnessed in digital commerce over the past decade has brought forth numerous challenges for E-Commerce operations. It is anticipated that E-Commerce will encompass 17% of the industry by the end of 2022. With this burgeoning size and the escalating demand for online business, capitalizing on the digital commerce trend poses significant hurdles. Some of the challenges faced by the E-Commerce industry include:

1. **Expanding Customer Expectations** :

Retailers worldwide strive incessantly to establish their reputation and a distinct brand identity by promising exceptional experiences. In an era where the customer experience reigns supreme and tech giants like Amazon are revolutionizing the online purchasing journey with anticipatory shipping methods, meeting customer expectations becomes exceedingly difficult. Thus, competing with such giants and adapting to the ever-evolving demands of customers stands as a monumental challenge for retailers today.

2. **Agility Challenges** :

Agility refers to a business's ability to innovate, develop, and deploy digital content swiftly, and to respond to seasonal fluctuations effectively. Agility facilitates rapid digital fulfillment and is recognized as one of the most critical initiatives in the E-Commerce sector. However, numerous companies encounter difficulties in adapting or evolving according to customer needs. This is often due to inefficiencies in integrating new technologies with existing systems, consequently making market insights harder to attain.

3. **Intense Competition** :

The internet provides an equal platform and opportunity to all, thereby fostering a highly competitive environment where countless businesses offer similar products or services to the same target audience. Even niche brands must navigate intense competition to carve out their market share. The E-Commerce landscape is expected to become increasingly congested and competitive in the coming years. Startups must compete with multinational

corporations, and every enterprise is continuously innovating to gain a competitive edge in the E-Commerce space.

Proposed solution

I'm designing an Analytical Dashboard featuring column charts, pie charts, bar graphs, and more to present detailed information. This tool empowers retailers to monitor their business performance and facilitate growth. They can analyze profits and losses by region or product category. By interpreting the graphs, they can implement strategies to enhance business expansion and better serve their customers. The dashboard highlights regions with high sales, enabling them to focus on delivering quality service efficiently

EXPERIMENTAL INVESTIGATION

Preparing the Dataset

The dataset used for creating the e-commerce analytical dashboard is of a US Superstore.

Following are the steps which are to be completed to process the data and create the data module.

1. Upload the dataset to prepare data and create a data module
2. Preprocess or clean the data
3. Read the dataset
4. Analyze the dataset
5. Drop unnecessary columns
6. Change the column names
7. Remove the randomness in the columns
8. Find the missing values
9. Handle the missing values
10. Save the final data module

I have used the above mentioned process to clean and format the data and add calculated fields to the existing dataset of US superstore to add precision to our data visualizations . The calculated fields added are as follows:

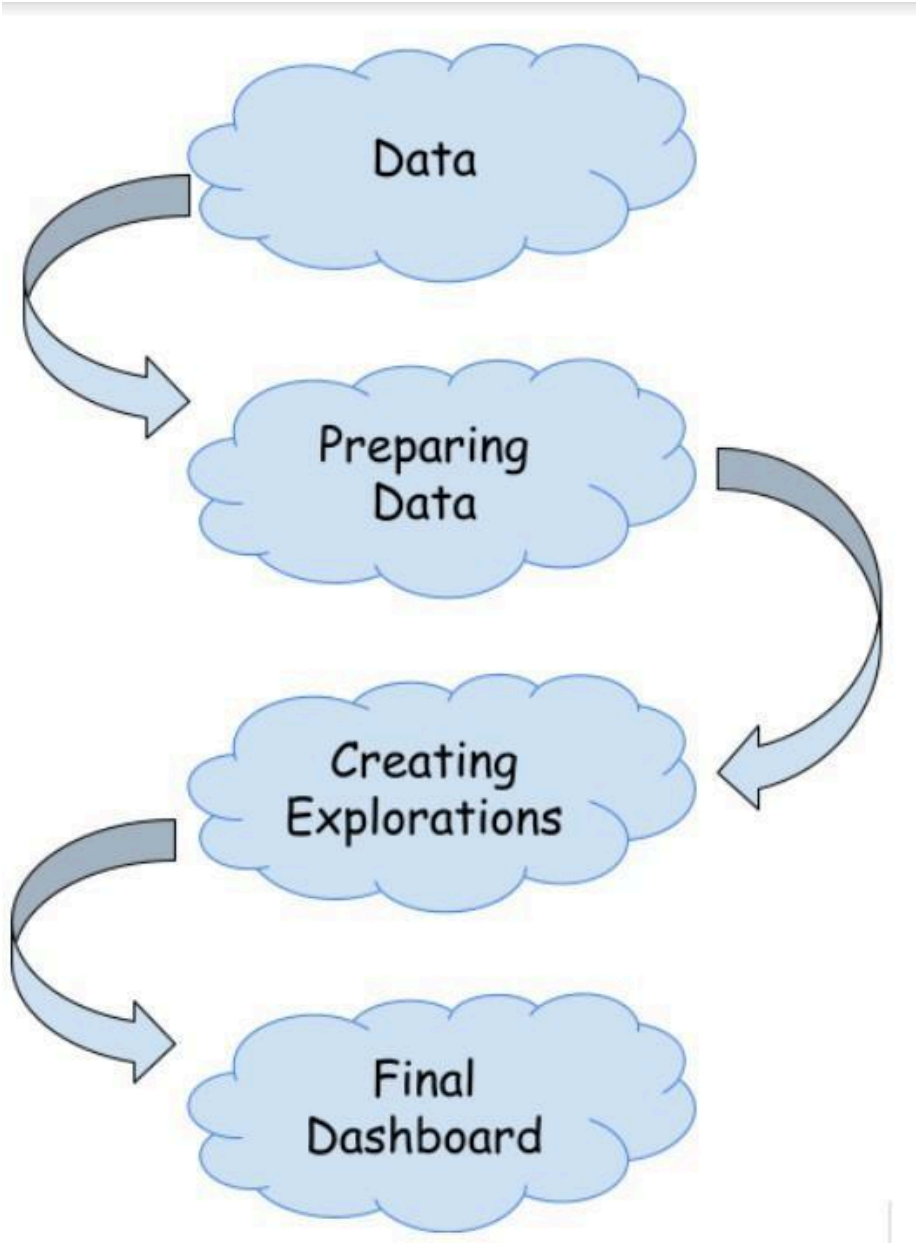
- Order_Year
- Order_Month
- Order_Quarter
- Days_to_deliver
- Shipping_year
- Shipping_quarter

Creating Explorations for Dashboard

For preparing the dashboard I have created several data explorations to analyze the dataset the several data explorations created are as follows:

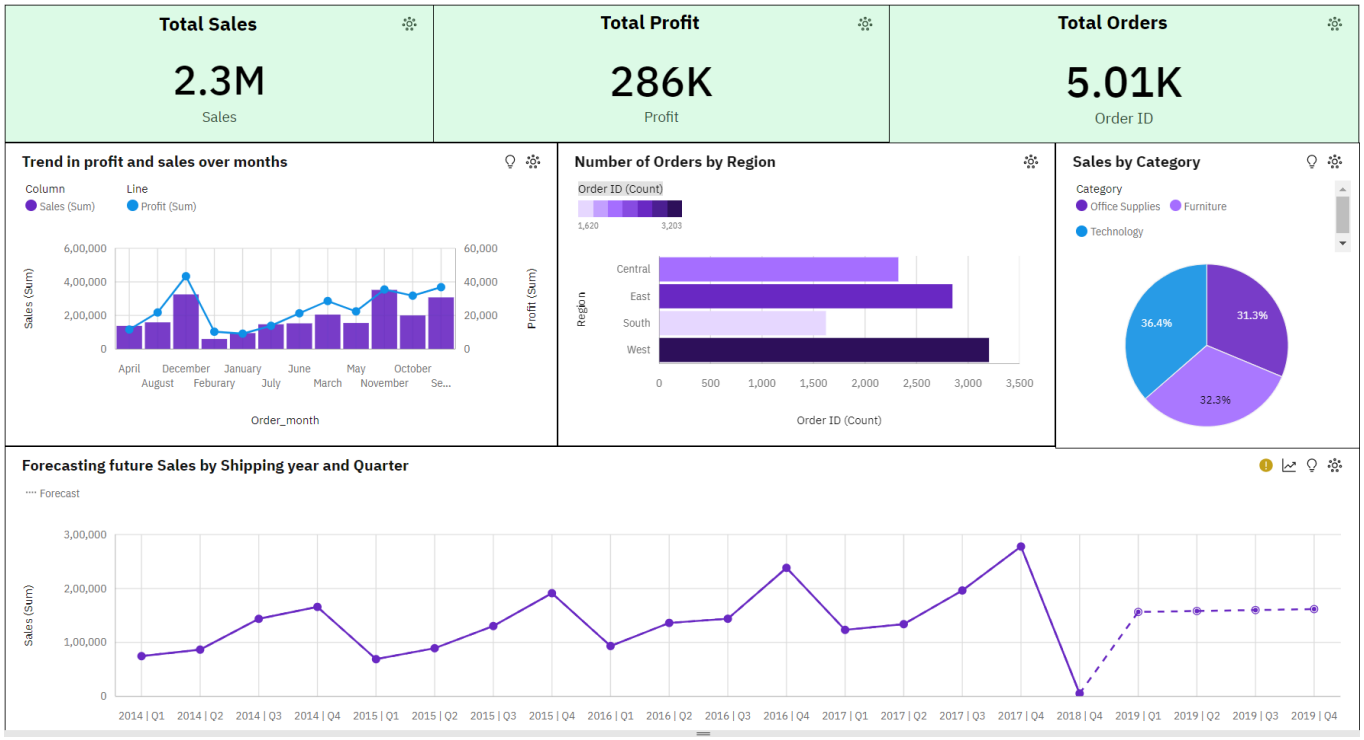
1. I have used a bar graph to display the region that accounts for the greater number of orders and the product type that was ordered greater number of times.
2. I have used a column graph to display the quantity frequency distribution of different product categories.
3. To display the % sales of different product categories I have used a pie chart.
4. I have used a stacked column graph to display both the profitable products and the products that have incurred a loss in the last few years.
5. I have used a line graph for forecasting future sales by shipping year and shipping quarter.
6. For displaying the yearly sales of various states I have again used a bar graph.
7. I have displayed the trend of both the profit and the sales over the various months using column and line graph.
8. Along with providing solutions to all the expected solution statements, I have also created a column graph showing the number of days taken by different product categories to get delivered.
9. Along with creating the above explorations I have also added some summary blocks to display the total sales ,the total profit and the total orders of the store.

FLOW CHART


Data
Preparing Data
Creating Explorations
Final Dashboard

Dashboards

Here are a few screenshots which show the final results. These are a few graphs showing different info.



OBSERVATIONS

By creating various explorations and analyzing the dataset deeply I have made the following observations by interpreting our dashboard:

1. The total sales of the store is around 2.3M , the total profit for the store is around 268K and the total orders are 5.01K.
    - City Abilene has the lowest total Sales at 1.39, followed by Elyria at 1.82.
    - New York City has the highest total Sales at over 256 thousand, followed by Los Angeles at almost 176 thousand.
    - City Philadelphia has the lowest total Profit at over minus 14 thousand, followed by Houston at nearly minus 10 thousand.
    - New York City has the highest total Profit at over 62 thousand, followed by Los Angeles at over 30 thousand.
    - City Aberdeen and Abilene have the lowest Order ID at 1.0.
    - New York City has the highest Order ID at 450.0, followed by Los Angeles at 384.0.
2. Product sub-categories that have led to huge profits in the last few years are Accessories,Appliances, Art, Binders , chairs, copiers, paper phone storage etc whereas the product sub-categories that have incurred a loss are bookcases, machines supplies and tables etc. Also, For Profit, the most significant values of Order_year are 2017 and 2016, whose respective Profit values add up to over 175 thousand, or 61.2 % of the total.
3. For displaying the yearly sales of various states I have used a bar graph. For Sales, the most significant value of Order_year is 2017, whose respective Sales values add up to almost 392 thousand, or 32.8 % of the total. For Sales, the most significant value of State is California, whose respective Sales values add up to nearly 458 thousand, or 38.3 % of the total.
4. I have observed that there is not much difference in the % sales of the various product categories that are technology , office supplies and furniture but technology with 36.4% has proved to be the product category with the highest sales percentage. Also,

- Category Furniture has the highest Total loss_percentage but is ranked #2 in Total Sales.
- Category Technology has the highest Total Sales but is ranked #2 in Total loss_percentage.

5. By quantity frequency graph I have analyzed that customers generally prefer to buy 2 or 3 quantities of products and office supplies are the most ordered products. For Quantity, the most significant value of Category is Office Supplies, whose respective Quantity values add up to almost 23 thousand, or 60.5 % of the total.

6. Looking at the trends of profit and sales I have observed that the profit has been on peak month of December and sales has been on peak in the month of November but there has been a huge dropdown in profits in the month of January.

7. **I have forecasted sales for the year 2019. Based on the current forecasting, Sales may reach almost 162 thousand(161,692 $) by Shipping_year-shipping_quarter 2019-Q4.**

8. I have displayed the trend of both the profit and the sales over the various months using column and line graphs. For Sales, the most significant values of Order_month are November, December, and September, whose respective Sales values add up to over 985 thousand, or 42.9 % of the total.

9. Along with providing solutions to all the expected solution statements, I have also created a column graph showing the number of days taken by different product categories to get delivered. For Days_to_deliever, the most significant values of Category are Office Supplies and Technology, whose respective Days_to_deliever values add up to 82, or 74.5 % of the total.

10. West (32 %), East (28.5 %), and Central (23.2 %) are the most frequently occurring categories of Region with a combined count of 8,374 items with Order ID values (83.8 % of the total) and the West region accounts for the highest no of orders with 3,203 (including repeat orders) orders and products under the office supplies category are mostly ordered.

11. Most products take around 4-5 days to deliver and office supplies take the most time to deliver and the products under technology mostly take the least time to deliver.

PROS AND CONS:

PROS:

The pros of using this Dashboard are:

1. The visual representation makes it easier to grasp various aspects such as sales, profit, and loss.

2. No longer will you waste valuable time generating reports from multiple systems. Instead, data is retrieved from a centralized source and presented as a clear visual summary.

3. By gaining deeper insights into each customer's purchasing cycle, businesses can more accurately forecast future demand using historical data. This enables better planning for demand fluctuations in the upcoming business cycle, establishing measurable goals and objectives for enhanced success.

4. Dashboards gather data from various sources and present it in easily understandable visual formats in real-time.

5. Equipped with analytics and real-time visibility into inventory stock details, sales staff are informed about available items and their locations.

CONS:

1. A flashy or cluttered design, coupled with users attempting to incorporate excessive information without comprehending constraints or considering their specific needs from the various detailed data analyses provided.

2. Difficulty arises when attaching supporting data to a dashboard, and the failure of data to refresh automatically necessitates manual intervention for both of these tasks.

3. The technologies utilized in dashboard development differ from other software solutions already utilized within organizations, which can initially pose challenges in understanding.

4. The business lacks predetermined rules and hierarchies for the utilization of dashboard metrics. As a result, each employee may employ the metrics differently, leading to a diverse array of reported data.

APPLICATIONS

There are many applications of this Dashboard in the E-commerce industry. Few of them are :

1. Assisting the company in enhancing areas like profit and sales will be facilitated.
2. Facilitating the monitoring and tracking of data will also enable the utilization of data to derive meaningful insights.

CONCLUSION

This analytical dashboard has been developed utilizing IBM Cognos Analytics. With the rapid growth of the e-commerce sector, retailers are facing intense competition. This dashboard aids in monitoring various aspects such as profits, sales, and losses, thereby providing insights into customer needs. Its visual representation simplifies business tracking, enabling retailers to effectively compete and thrive in the market.

These dashboards play a crucial role in business performance management by furnishing essential reporting and metric data. Similar to the dashboard in a vehicle, they present real-time key metrics and performance indices, facilitating decision-making and navigating the business landscape more effectively.

Recognizing the significance of dashboards in the e-commerce domain, I offer capabilities to manage, create, and share reports, track filings, and personalize content. Our dashboards empower you to:

- Build and customize a wide pot pourri of charts and graphs including

  Grouped/Stacked bar, Grouped/Stacked Line, Pie, and Map
- Schedule reports for automated delivery based on your preferred frequency, i.e., your corporate fiscal year.

Dashboards bring speed, but they also bring versatility; view the data the way that suits you best or the way that's best for sharing and storytelling to promote better agreement of IP performance throughout your organization.

Future Scope

- Integrating the dashboard with dynamic data sets to show live reports.
- Expanding the dashboard and dataset to display customer reviews for the products, say customer ratings out of 5 stars.