Analytic tools for E-commerce businesses

Table of Contents

INTRODUCTION TO PROJECT

Overview

With increase in the consumer demand, the E-commerce sector has boomed. This also led to an increase in fierce rivalry 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.

In this project we have created a Dashboard using a US SUperstore Dataset on the IBM Cognos Analytics platform . We have tried to show the following things through our project so that the retailer can discover statistics such as -

1. Region that accounts for a greater number of orders.
2. Frequency distribution of quantity ordered.
3. Percentage sales by different product class.
4. Profitable products or their sub products in the last few years.
5. Products that incurred losses.
6. Frequently order products
7. Yearly sales for various states.
8. Forecasting future sales according to shipping date.
9. Trend in profit/sales over time (years/months/quarters).
10. 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 marketand trend pattern and take the required measures.

Purpose

The challenge can be mitigated through research and competitive analysis. It is important to categorize the contenders and closely and frequently observe and assess the components of their websites, be it pricing, the design, products on display or even the marketing tools and techniques they use. There is ample to learn from the strengths as well as the weaknesses of the contention.

Users, dashboards allow **seller toSo participate and understand the analytics process by compiling data and visualising trends and occurrences**. Data dashboards provide an objective view of performance metrics and serve as an effective foundation for further dialogue.

So witht the aim of helping sellersunderstand the real potential of their business in the market we have planned to createan e-coomerce analytical dashboard using the powerful IBM cognos Analytics.

LITERATURE REVIEW

Existing Problem

An E-Commerce store is the primary retail growth engine for any business. The meteoric growth witnessed by digital commerce over the last decade has also given rise to several eCommerce challenges. eCommerce is expected to claim 17% of the industry by the end of 2022. With this increasing size and demand for online business, riding the digital commerce wave is not easy. Few challenges faced by E-commerce industry is:

1. Customers' Exploding prospect

Retailers all over the world are endlessly trying to build their reputation and a sharp brand image with the promise of a great experience. In an era where experience matters the most and tech giants like Amazon are taking the online buying process to the proverbial 'next level' with anticipatory shipping methods, it is very difficult to meet customer outlook. Thus, competing with them and fulfilling the ever-evolving customer demands is a huge challenge for retailers today.

2. Agility Challenge

Agility is the capableness of a business to introduce progress, develop and deploy digital content, react to seasonal changes, et c. rapidly. Agility drives prompt digital fulfilment, and it is identified as one of the most important initiatives in the eCommerce business. Agile shift is at the heart of digital business, and scaling is essential for making it successful.
A large number of companies find it difficult to move or change as per the needs of customers. It is commonly because they can't integrate new applied science efficiently with their existing system, and as a result of which, insight into the market gets harder.

3.Facing Competition

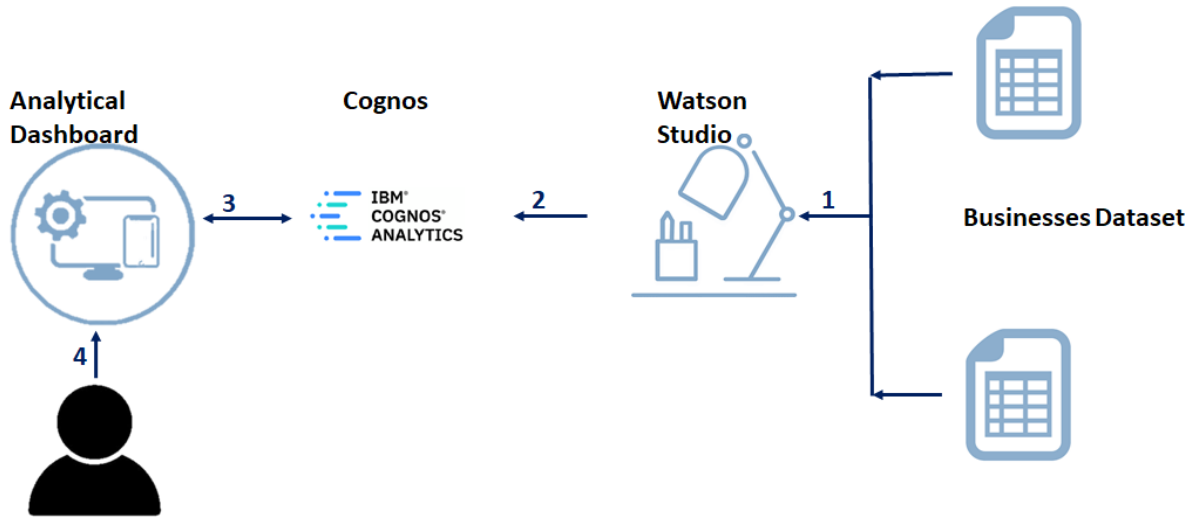
The internet offers everyone an equal platform and hence an equal opportunity. This makes the environment competitive, with possibly hundreds of other businesses offering the same products or services as you and to the same target audience. Even the most niche brands have to fight hard to carve a place for themselves. And every segment in the eCommerce space is set to get more and more congested and competitive over the years.
A home grown startup has to keep pace with a multinational giant and vice-versa. And every enterprise is continuously experimenting to come up with an advanced eCommerce business solution to gain an edge.

Proposed solution

We are creating an Analytical Dashboard which will show the details using Area graph, pie chart, bar graph, et c. This helps the retailer to keep a track of their business and help them grow. They can calculate their profit and loss in a particular region or particular product. Seeing the graph, they can adopt methods and solutions which may help them expand their business. It will also help them cater the needs of their customers. The dashboard will give an idea about the regions where the sales are high, so that they can adopt effective means to provide good service.

THEORETICAL ANALYSIS

Block Diagram



Hardware/Software Designing

The following things are required:

- 32-bit Intel® Pentium® 4 or compatible processor running at 2 GHz or greater.
- 512 MB RAM.
- Disk space: 350 MB for client components.
- A DVD-ROM drive.
- IBM cloud
- IBM Cognos Analytics

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 proccess 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. Analyse 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. Split the data into sovereign and dependent variables
11. Split the data to train and test
12. Train the machine with preprocessed data with an Appropriate Machine learning algorithm to add the calculated feilds.
13. Save the final data module

As the prediction data used for the model is classification type, we apply a logistic regression algorithm on our dataset to add the calculated fields.

Logistic regression is the suitable regression analysis to conduct when the dependent variable is binary. Like all regression analyses, logistic regression is a predictive analysis. Logistic regression is used to describe data and to explain the relationship between one dependent binary variable and one or more nominal, ordinal, interval or ratio-level independent variables.

Once the model is trained , it's ready to make accurate predications for creating the calculated fields.

We have used the above mentioned proccess to clean and format the data and add calculated fields to the existing dataset of US supperstore to add precision to aour data visualizations . The calculated feilds added are as follows:

- Order_Year
- Order_Month
- Order_Quater
- Days_to_deliver
- Shipping_year
- Shipping_quater

Creating Explorations for Dashboard

For prepaing the dashboard we have created several data explorations to analyse the dataset the several data explorations created are as follows:

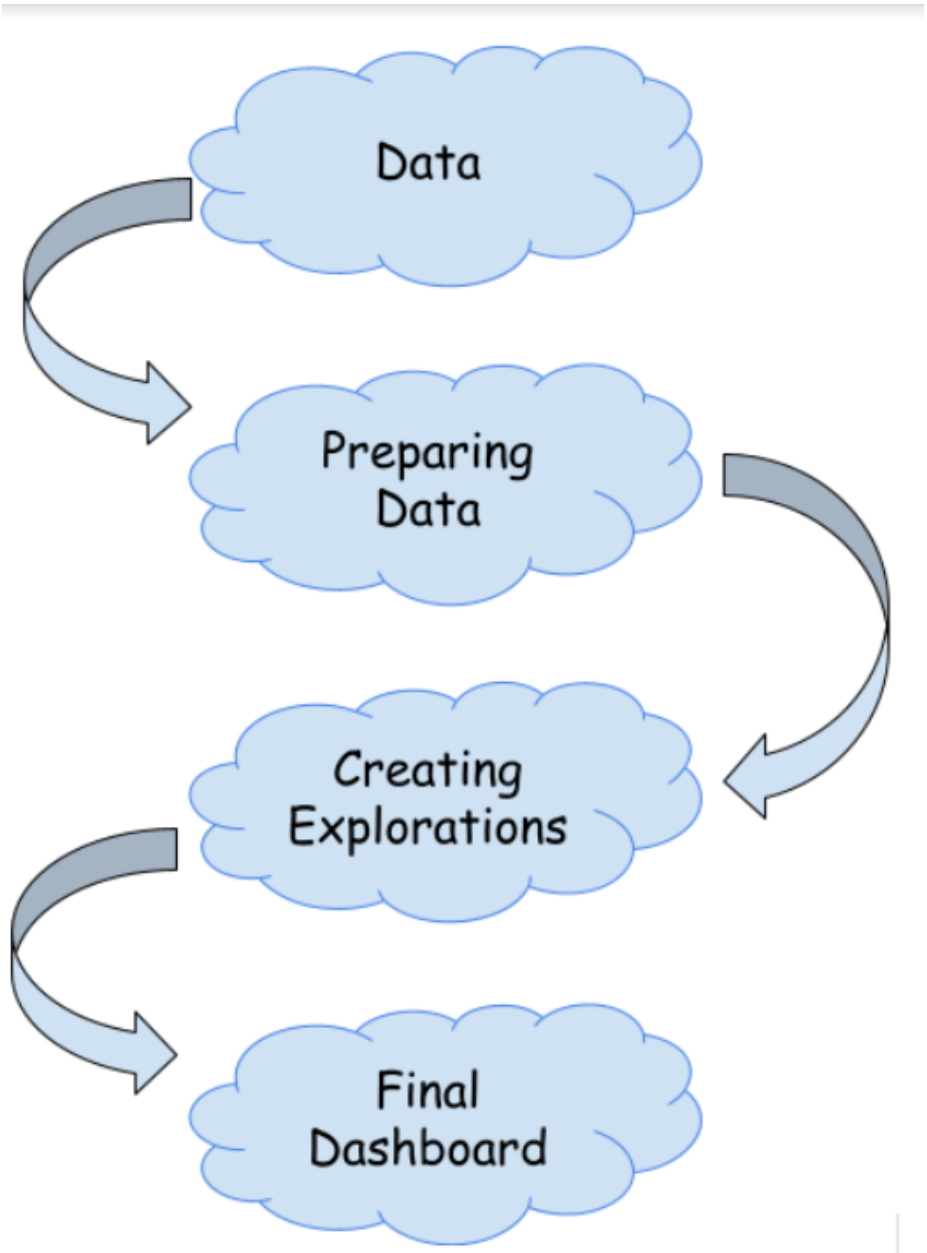
1. We have used a bar graph to display the region that accounts for greater number of orders
2. We have used a column graph to display the quantity frequency distribution of

different product categories

3. To diplay the % sales of different product categories we have used a pie chart
4. We 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. We have used a line graph for forecasting future sales by shipping year and shipping quater
6. For displaying the yearly sales of various states
7. Remove the randomness in the columns
8. Find the missing values
9. Handle the missing values
10. Split the data into sovereign and dependent variables
11. Split the data to train and test
12. Train the machine with preprocessed data with an Appropriate Machine learning algorithm to add the calculated feilds.

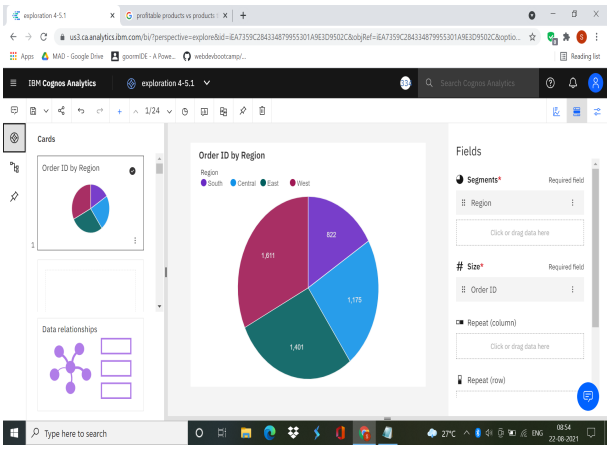
Save the final data module

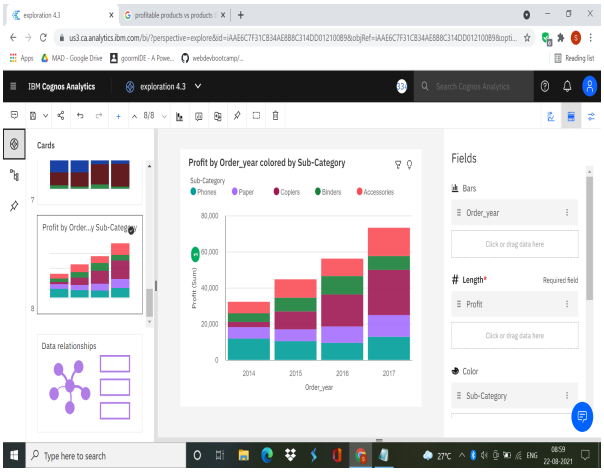
FLOW CHART

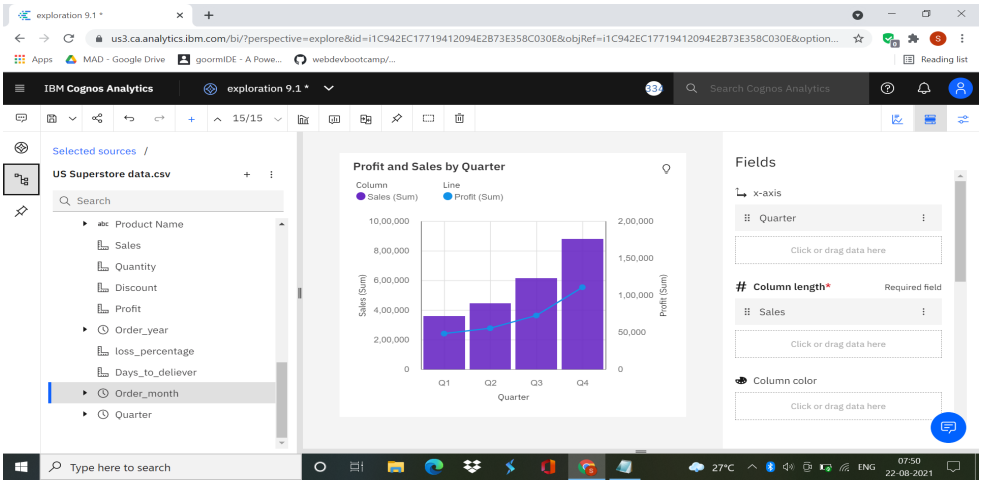


RESULT

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







PROS AND CONS:

PROS:

The pros of using this Dashboard are:

1. The graphical representation is comparatively easier to understand different things such as sales, profit and loss.
2. You are no longer wasting valuable time generating reports from multiple systems. Instead, data is drawn from a centralised source and displayed as an easy to interpret visual overview.
3. With greater insight into the buying cycle of each customer, future demand can be more accurately predicted using historic info. Businesses can more effectively plan for demand fluctuations for the next business cycle, setting measurable goals and deliverables for greater success.
4. dashboards source data from multiple areas displaying the info as easy to understand visuals in real-time.
5. With analytics and a real-time vision of inventory stock details, sales staff know what items are in stock and where they are located.

CONS:

1. Flashy or cluttered design, with users attempting to integrate too much info without understanding constraints or considering their specific needs from the range of different measurably detailed data analyses provide.
2. Trouble in attaching supporting data to a dashboard and the failure of data to refresh automatically means that both these tasks must be done manually.
3. The applied sciences used in the development of dashboards differs from other software solutions already employed in organisations and can be at first difficult to understand.
4. The business has no predetermined rules and hierarchies for how dashboard

metrics are used. This means each employee can use the metrics in different ways, resulting in a diverse set of data being reported.

APPLICATIONS

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

1. It will help the company to improve in aspects such as profit and sales.
2. It will help in monitoring and tracking data.It will also help in utilizing the data to make some sense.

CONCLUSION

This Analytical Dashboard is made using IBM Cognos Analytics. As the E-commerce industries are growing at a faster pace, the retailers are facing fierce contest. This dashboard will help them keep a track of their profit, sales, loss, et c. This will give them ideas about their customer needs. The graphical representation makes it easier to keep a track of their business. So using this Dashboard they can compete and survive in the market.

These dashboards provide critical reporting and metrics info and are integral in Business Performance Management. Much like the dashboard in your vehicle, dashboards display real-time key metrics and performance index, guiding decisions and better navigating the surrounding landscape.

We understand the imperative for dashboards in the E-commerce industry. We provide you with the ability to manage, create and share reports, monitor filings and personalise content. Our dashboards enable you to:

- Build and customise 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 organisation.

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.

Bibliography

1. https://www.youtube.com/watch?v=ASGjw8yHA2g
2. https://www.youtube.com/watch?v=qcf6a9QzkhI
3. https://www.youtube.com/watch?v=1VXO8p8yX9Y
4. https://www.youtube.com/watch?v=YvzSOCU3YME
5. https://www.kaggle.com/juhi1994/superstore-analysis

Few documents :

IBM Documentation

IBM Documentation for cognos analytics

Demonstration Video link: Google Drive link