

Data Visualisation: Empowering Business with Effective Insights

K.R. Manikandan
Data Analyst



Internship Programme

11 Sep 2024

INTRODUCTION

This is a virtual internship Programme case study with the company, TATA. This program is hosted through the site called "Forage". This internship helps to learn data basics such as data cleaning, modelling, visualization and storytelling.

PROJECT SCENARIO

An online retail store has hired a consultant to review their data and provide insights that would be valuable to the CEO and CMO of the business. The management wants to analyse what the major contributing factors are to the revenue so they can strategically plan for next year. Suggest metrics that would enable them to make the decision on expansion.

Project description:

The tasks divided in to 4 phases.

TASK_1

Create each four questions with reasons for choosing it, related to the business solutions in the CMO and CEO point of view for future strategy planning of marketing, expansion of the business in future. This task contain 8 questions.

TASK_2

CEO and CMO explained the situations of business task . Find out the charts and graphs suitable visualizations for it. This task contain 5 questions.

TASK_3

CEO and CMO explained the situations of business task . Create suitable visualizations for it use the tools Tableau or Power BI . This task contains 4 questions.

TASK_4

Analytics, visualization, dashboards as back up , presentation of Reports got from data insights and data driven decisions.

Input data _ Dataset Summary

dataset in excel/csv format

InvoiceNo column has invoice number of each purchase made by customers

StockCode has the products stock code.

Description has the products description and information.

InvoiceDate has dates from when the purchase happened.

UnitPrice has the price of the product.

CustomerID tells us who purchased the product in numerical

Country is where the customer resides in.

DATA PREPARATION & DATA PREPROCESSING

- ✓ Excel to clean the dataset with the following steps:
- ✓ Removed Rows in the Quantity column showing negative numbers or below 1
- ✓ Removed Rows in the UnitPrice column that are below \$0
- ✓ Removed Rows with blanks

*****Added Revenue Column from multiplying UnitPrice with Quantity**



TASK_1

Concerns the CMO may have

1. **How does the highest purchased product , perform in each region ?**

Percentage of most sold product regionwise identification helpful to CMO strategy and target the more demanded products inventory and changing the marketing strategy of the less purchased products regionswise. The trends will indicate the purchasing habits of the customers in the particular region.

2. **Which product purchased repeatedly ? what is the frequency between next purchase ?**

Identification of repeated purchase of product by the customers. How long it takes for reordering the same product. Helpful to CMO plan the marketing strategy, make sure the availability of the products at particular seasons.

3. **Which region the most repeated purchase happened ?**

This is very important to find out the patterns , which region the purchasing happened repeatedly. Then the CMO can improve the customer relationship and award programmes and referral programmes to encourage to customers in particular regions. Also improve the strategy of offers to the other region customers that way increase the sales in other regions

4. **Which product has purchased more,how many times they purchased ?**

This is very important to CMO to analyse the trends of product which purchased repeatedly. It helps to ensure the product availability at needed time. Understanding the demand of the product helpful the define the future strategy planning of the production and sales.

Concerns the CEO may have

1. **Which region is producing the most profit, and which region is producing the lowest?**

The CEO is interested in the viewing revenue by the regions, to assess which regions are generating the highest revenue and which regions are generating lower revenue.

As it is based on the fundamental source of income for the business.

- **Revenue analysis** is focused on earnings and how to increase it. Using the data and analysis.
- The CEO will be able to decide on how to further generate revenue in the regions that are already generating the most revenue. For the regions that are not generating enough revenue.
- The CEO will then study the reasons why there is a lack of sales in those regions and try to improve the products and make them more suitable for those regions.

2 Did customers based have significant impact on revenue ?

Overall, revenue customers have a significant impact on a company's revenue growth. Understanding their characteristics, behavior, and preferences is crucial for developing effective strategies and optimizing revenue.

3 What is the quarterly and monthly sales and revenue breakdown ?

This metric is one of the key indicators used by management, investors, and external analysts to gauge a company's performance and its effectiveness in generating revenue from operations.

4 which month has the highest revenue and lowest revenue ?

A monthly trend of revenue will provide the CEO with insights on how the revenue is fluctuating each month.

This will enable the CEO to analyze ,

- ❖ How the internal changes inside the company have had impact on the sales.
- ❖ How a new product launch has led to an increase in revenue during the month
- ❖ How the introduction of a new region has led to an increase in revenue for the online store.
- ❖ If there have been any delays internally that would have caused a potential decrease.

Such analysis are vital for the senior management as it would enable them to plan and try to make the customer experience as smooth as possible.

TASK_ 2

Agenda of task_2

CEO and CMO to provide visuals on the metrics that they wish to analyse for the online retail store. gather the requirements and provide them with the type of visual that would be best suited to the scenario. The senior management wants to understand how their business is performing and what areas are the key strengths of the company. They are also focused on identifying opportunities that would lead to growth and generate more revenue in the future.

Requirements

Multiple visualisation requests from the CEO , need to provide the visual which would explain the data and insights in the simplest possible manner. The visual should adequately convey the information that trying to present.

Questions 1/ 5

The CEO of the retail store is interested to view the time series of the revenue data for the entire year. The CEO is interested in viewing the seasonal trends and wants to dig deeper into why these trends occur. This analysis will be helpful for the CEO to forecast for the next year.

1 Which visual would most likely help the CEO analyse the data?

A line chart

is used to represent repeated measurements taken over regular time intervals. Time is always displayed on the horizontal axis and values on the vertical axis. The line chart would enable the CEO to see important changes in the data, like seasonality or cyclic behaviour, which will provide a better understanding of the revenue and help forecast better.

Questions 2/ 5

The CMO is interested in viewing the top 10 countries which are generating the highest revenue. Additionally, as a subcomponent, they would also like to see which products are contributing to the total revenue being generated by each country. Which visual would enable the

2 CMO to view the revenue for each country and the breakdown by products on a single chart?

Stacked bar chart

The chart allows users to compare subcomponent pieces across different categories. The height or length of the bars will represent the total revenue generated by each country.

Each bar will be divided into the products sold, where the major portion will be allocated to the products that are generating more revenue. This will allow the CMO to view the total revenue data as well as the revenue for each individual product.

Questions 3/ 5

The CEO of the online retail store wants to see how much average revenue is generated by each country. They are interested in viewing the following metrics on the visual: Minimum value First quartile value Median value Third quartile value Maximum value

3 Which chart would you create to show the above metrics for the average revenue generated by each country?

Box plot

The CEO's request would best be fulfilled using a box plot. The box plot would show the distribution of data based on a five number summary ("minimum", first quartile, median, third quartile, and "maximum"). Boxplots are used to graphically demonstrate the locality, spread and skewness groups of numerical data. By using the boxplot, we can easily spot outliers and the distribution of the plot.

Questions 4/5

The CMO of the online retail store wants to view the information on the top 10 customers by revenue. They are interested in a visual that shows the greatest revenue-generating customer at the start and gradually declines to the lower revenue-generating customers. The CMO wants to target the higher revenue-generating customers and ensure that they remain satisfied with their products.

4 Which visual would help the CMO understand the data on revenue generated by the top 10 customers?

COLUMN CHART

Column charts are used to display comparisons between different items. Alternatively, you can view a comparison of items over time. In this case, the column chart would display the top 10 customers who are contributing the most to revenue. The first bar on the left would be the longest one and would display information for the customer who bought the most goods from the online store. The revenue contributed by each customer would gradually decline as the bars move to the right.

Questions 5/5

The CEO is looking to gain insights on the demand for their products. They want to look at all countries and see which regions have the greatest demand for their products. Once the CEO gets an idea of the regions that have high demand, they will initiate an expansion strategy which will allow the company to target these areas and generate more business from these regions. He wants to view the entire data on a single view without the need to scroll or hover over the data points to identify the demand.

5 Which chart would be most useful to provide the CEO information on the demand in each region?

MAP CHART

A map chart would be the best option for visuals here as it will allow the CEO to view the demand information on a single view. The map chart will have all the countries on a single map and the demand will be displayed by highlighting the area of each country. Dark highlights would mean that the countries have high demand for products whereas the countries that are highlighted light colour are the ones that have low demand. Geographical data is best visualised using map charts as they are very easy to view and the underlying values are also easy to analyse.

TASK_3

Agenda of task_3

Create the visuals around four of the questions that the CEO and CMO have requested. can use either Tableau or Power BI to create the visuals .

Data cleanup

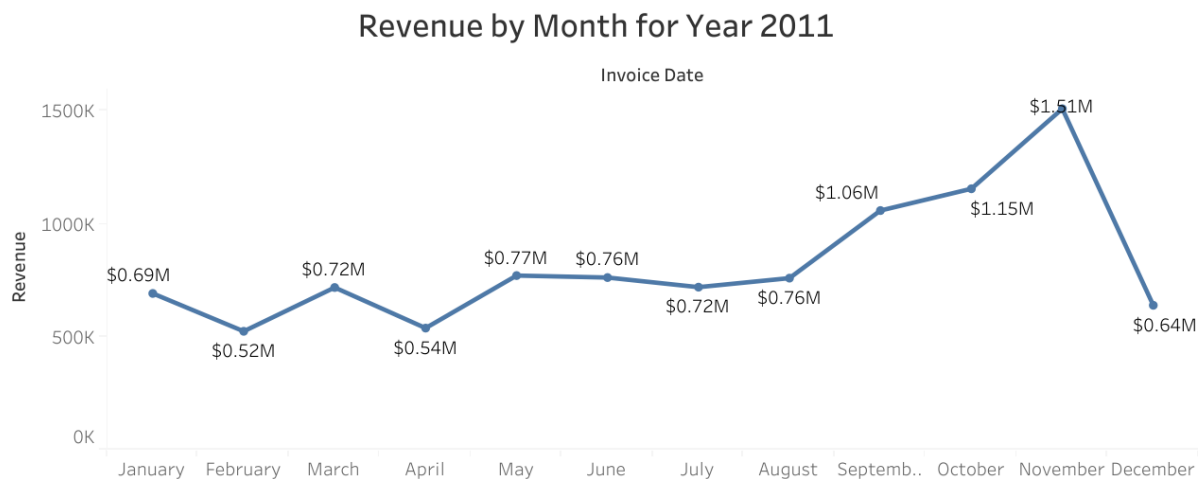
data contains some returns to the store which are provided in negative quantities and there are unit prices which were input in error

- Create a check that the quantity should not be below 1 unit
- Create a check that the Unit price should not be below \$0

use conditional formulas where the logic would state that if the conditions are met then the tool should exclude the data from analysis. Also use data transformation methods to get rid of the bad data.

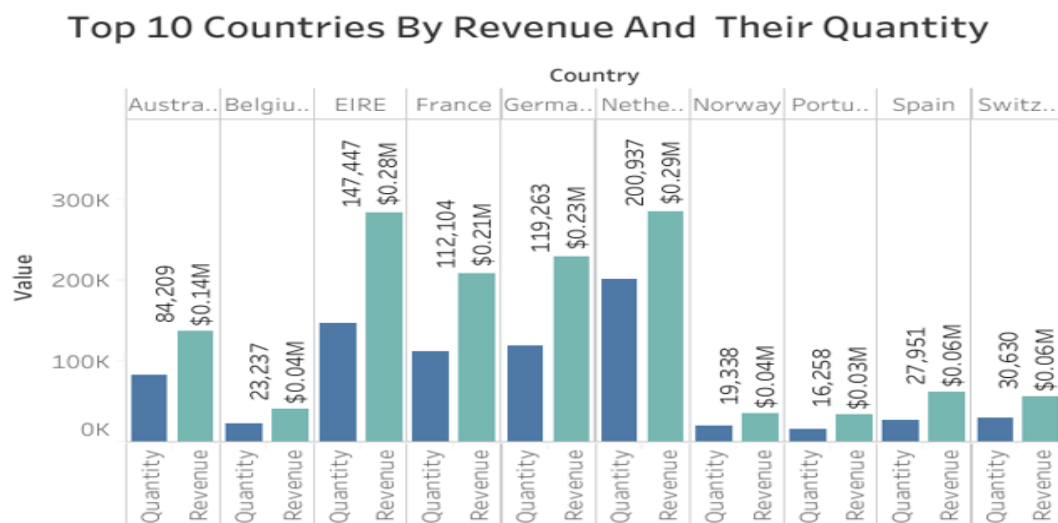
Question 1

The CEO of the retail store is interested to view the time series of the revenue data for the year 2011 only. He would like to view granular data by looking into revenue for each month. The CEO is interested in viewing the seasonal trends and wants to dig deeper into why these trends occur. This analysis will be helpful for the CEO to forecast for the next year.



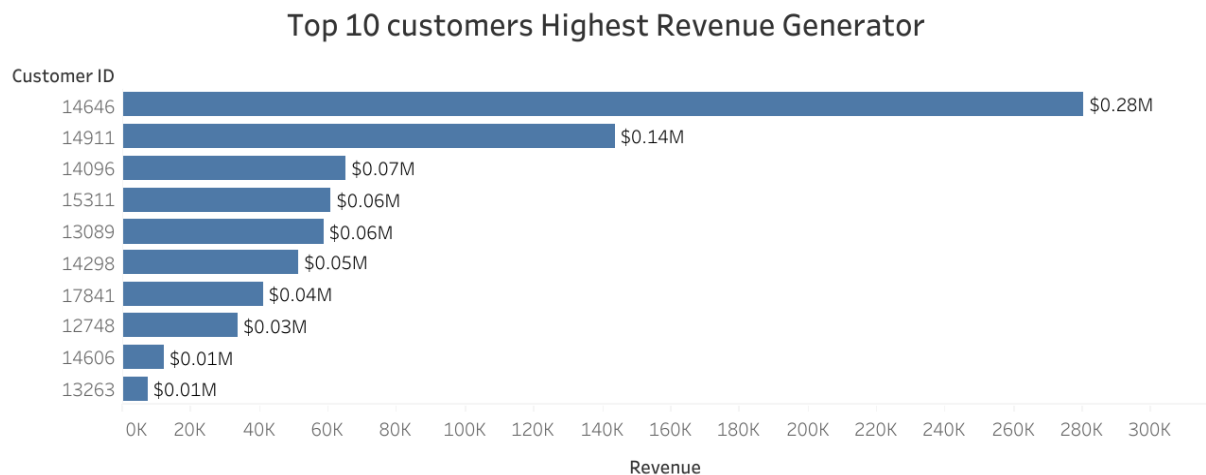
Question 2

The CMO is interested in viewing the top 10 countries which are generating the highest revenue. Additionally, the CMO is also interested in viewing the quantity sold along with the revenue generated. The CMO does not want to have the United Kingdom in this visual.



Question 3

The CMO of the online retail store wants to view the information on the top 10 customers by revenue. He is interested in a visual that shows the greatest revenue generating customer at the start and gradually declines to the lower revenue generating customers. The CMO wants to target the higher revenue generating customers and ensure that they remain satisfied with their products.



Question 4

The CEO is looking to gain insights on the demand for their products. He wants to look at all countries and see which regions have the greatest demand for their products. Once the CEO gets an idea of the regions that have high demand, he will initiate an expansion strategy which will allow the company to target these areas and generate more business from these regions. He wants to view the entire data on a single view without the need to scroll or hover over the data points to identify the demand. There is no need to show data for the United Kingdom as the CEO is more interested in viewing the countries that have expansion opportunities.



Final Report

DATA ANALYSIS INTERNSHIP PROJECT

Data Visualisation: Empowering Business with Effective Insights

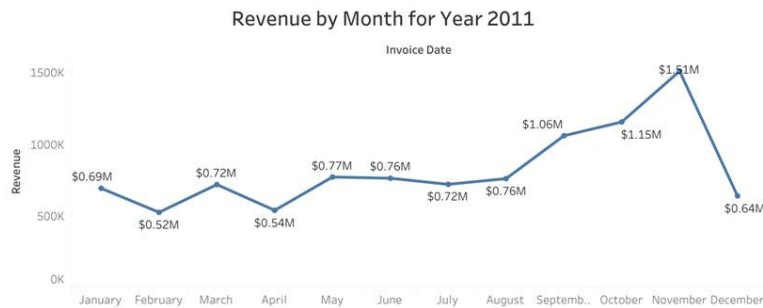


Prepared by K.R.MANIKANDAN
Submitted on 11th Sep 2024

DASHBOARD MARKET ANALYSIS



Time series Analysis

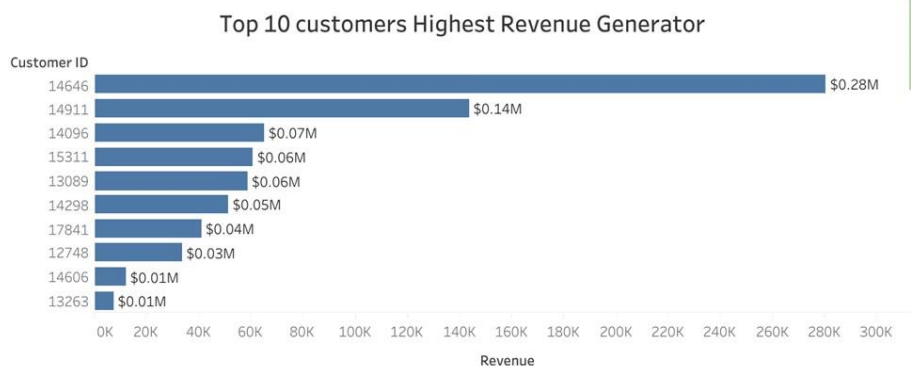


Highest Monthly Revenue is November \$ 1.51M

Lowest Monthly Revenue is February \$ 0.52M

CEO wants to know the Revenue of 2011 for each month, seasonal trends identified for future forecasting . Time series analysis done. In the month of August to November Revenue increasing November is highest. February is lowest.

Customer Analysis

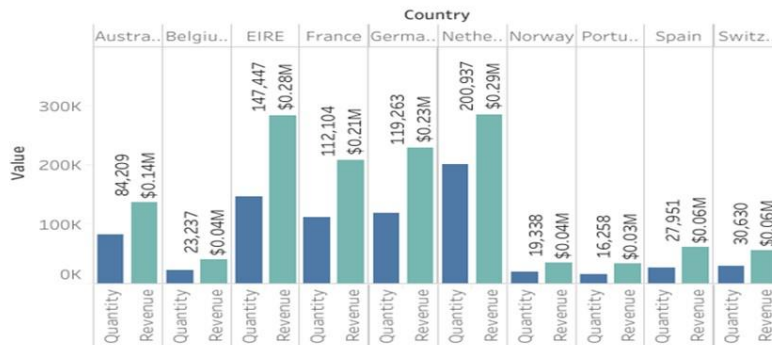


Highest Revenue generated customer ID is 14646 \$ 0.28 M

CMO wants to know the Revenue of year 2011 (exclude UK since having potential market), top 10 highest revenue generated customers to prepare the customer satisfaction strategy for keep retain the customers.

Revenue and quantity Analysis

Top 10 Countries By Revenue And Their Quantity



Highest Revenue generated country is Netherlands
\$ 0.29 M

Lowest Revenue generated country is Portugal
\$ 0.03 M

CMO wants to know the Revenue of year 2011, generated by the top 10 countries with quantity and exclude potential country UK from this analysis. Revenue and quantity compared by Bar charts Netherlands is highest revenue generated country with quantity. Portugal generated the least revenue.

Product Vs Regional Analysis

Demanded products regionwise



Highest Revenue generated customer ID is 14646
\$ 0.28 M



CEO wants to know the Revenue of year 2011 regionalwise per product (exclude UK since having potential market), to plan the strategy for expand the business. Asia, African countries business need to be expandable. Keep the momentum of sales by customer satisfaction and more products launch in European and US market.

