# 



# School of InfoComm Technology

**Data Discovery & Visualisation**

Diploma in Data Science

April 2022 Semester

**ASSIGNMENT 1**

**(Individual Assignment)**

**Submission Deadline:**

**12 June 2022 (Sunday), 12 Noon**

|  |  |  |
| --- | --- | --- |
| **Student Name** | **:** | Jyoshika Barathimogan |
| **Student Number** | **:** | S10222388 |
| **Video submission link** | **:** |  |

**Penalty for late submission:**

10% of the marks will be deducted every calendar day after the deadline.

**NO** submission will be accepted after 18 June 2022 (Saturday), 12:00 Noon

Table of Contents

[School of InfoComm Technology 1](file:///C:\Users\jyosh\Desktop\IT%20-%20DDV\ASSIGNMENT%201\assignment%201%20report%20done.docx#_Toc105968185)

[Introduction 3](#_Toc105968186)

[Project Objective (exploratory questions) 3](#_Toc105968187)

[Data preparation and Data analysing 5](#_Toc105968188)

[Exploratory Data Analysis and Visualization 6](#_Toc105968189)

[Dashboard 18](#_Toc105968190)

[Recording 22](#_Toc105968191)

# Introduction

The goal of my assignment is to design and implement dashboards that helps to answer exploratory question that will be useful for the online electronic retailer management. Using the given data set, I have created visualization and dashboards with data storytelling. Moreover, by cleaning and preparing the data set have helped me to make a better dashboard.

# Project Objective (exploratory questions)

Customer analysis

* Which customers generate the most profit?

Background: helps the management to see what customer generate the most profit and help the company by anticipating and satisfying the customer demand to help improve their profits.

* What is the most desired volume of orders customer placed?

Background: helps the management to know what the number of quantity that customer order. It helps the company to better plan their budgets

* Which months customer placed the most orders?

Background : it helps to determine whether the month is most popular by customer orders. By knowing what month are popular, it helps the company to better allocate its resources and budgets and plan ahead.

Shipment analysis

* Which shipment mode that customer prefer?

Background : helps to determine what shipment mode is popular and hence helps the company to plans their shipping better.

* What shipment mode affects turnaround time?

Background: Helps to determine the order status based on the shipment mode. It helps to determine what shipment mode is early or late and helps the management team to better forecast the volume of shipment based on the order date.

* Does shipment mode affect the order value?

Background: To see whether there is a relationship between the shipment mode and the order value. Hence the management can better consider the shipment fee.

Product analysis

* Which products generate the most sales?

Background: looking at the sales generated by the products category , it helps the company to better plan their budgets and discounts given.

* Which products generate the most profits?

Background: looking at the profit by products category helps management to better understand whether there are some products are doing well and not so well. This helps the company to better allocate their resources and their stocks.

* Which products have the highest profit ratio?

Background: looking at the profit ratio by products category. Then the company can look further into the reasons and determine how to help improve their profit-ratio. Factors that can affects are discount given and the type of product sub-category.

Trend analysis

* Which months have the highest number of sales and it corresponding profits made?
* Which months have the highest number of orders placed?

Background: it helps to determine whether the month affects the sales and profit. By knowing what month are popular, it helps the company to better allocate its resources and budgets and plan ahead.

Regions analysis

* Which regions have the highest sales ?

Background: the managers will understand which regions are doing well and not so well. Then they can dig deeper to understand what exactly is casing the loses also dig deeper into the council. Also replicate some of those best practices across other regions.

* Which products in regions are doing well based on their profit percentile?

Background: Helps to see what products in regions are doing well based on their profit percentile. Then they can dig deeper to see what exactly is losing the decrease in profit.

* What is the highest order value by regions?
* Which regions have the highest profit ration?

Background : helps to determine what regions are progressing better. So the company can better plan ahead to help improve the sales and profit and shows what regions are promising.

Country analysis

* Which country have the highest sales?
* Which country have the highest profit?

Background : the manager will understand which countries thar are doing well and not so well. Then they can dig deeper to understand what exactly is casing the loses also dig deeper into the council. Also replicate some of those best practices across other countries. There is also a factor of population in the country that will affect the sales and profit.

Price analysis

* What is the minimum order value placed?
* What is the average order value placed?

Background: it helps the management to predict the budgets and also work on their discounting to helps improve their numbers.

* Do discounts affect the profit made?

Background: to find out the problem area and find out why. It helps to determine which regions are profitable whether discounting is the key reason behind the loses. However, it might be also necessary for that particular region and maybe to penetrate the market or maybe given more market share.

Market analysis

* Which markets have the highest profits?
* Which markets have the highest sales?

Background: look at all the 6 markets to understand what markets are performing well. Dig deeper and find out what is causing it for example budget allocation to markets. Also replicate some of those best practices across other markets.

# Data preparation and Data analysing

The data given is clean but there are some nulls that I will ignore when doing my visualisation and dashboards.

I have separated data into few important segments in the data set given to better help with my dashboards.

The order segments contain order ID and Order date. Order ID helps to determine the count of order places based on the filter such as month or year. While Order date can helps as a filter for many dashboards.

The shipment segments contain Ship mode and ship Date. Ship mode can help to determine the popular ship mode customer prefer. And using Ship date and order date , we can determine the order status or turnaround time.

The customer segments contain customer ID , Customer Name and Segment. Customer ID can help to determine the count of customer placing orders. Customer name can be useless to determine what customer generates the highest profit and sales for better targeting. The segment helps to determine whether is it a consumer or corporate.

The map segments contain latitude and longitude that helps to create map visualization.

The Location segments contains City, State and regions , country and ISO3.However I feel that only some data are important such as regions and countries as it shows a bigger picture. I believe that company are not really interested in drilling down to individual cities and states as a overall pictures can shows much more. Hence, I have ignored states, city and IS03.

Regions and country can help show many data such as profit by country/regions.

The product segments contain product ID and Product Name which I am not going to use for my data as it shows no context to the viewers. As a manager viewing the dashboard, I cannot tell straight away the key concepts as the naming of these data are not efficient in conveying the meeting.

Therefore, I used the category segment that contains sub-category and category. This data can help shows many visualizations such as the products with highest sales and profit. This information is valuable for the company to make better decision.

The data segments contain Quantity , sales, profit, and discount. These data are very crucial for the company to determine their success and progression.

# Exploratory Data Analysis and Visualization

Based on the exploratory question I will be designing visualization.

Dashboard 01 Visualization 1

Graphical user interface, text, application

Description automatically generated

**Background:** This visualization shows the profit ratio by 12 different regions. It contains a filter based on the month and year of the order Date and the 2 different types of segments. The difference in colour shows the 3 different categories.

**Purpose:** The purpose of this visualization aims to shows what regions are doing well based on their profit ratio and how does consumer and corporate affects the numbers. This helps to determine what regions are successful based on category and overall performance. The category indicated what regions are successful hence help the company to plan to sell more based on the category to increase their profits.

Dashboard 01 Visualization 02

Map

Description automatically generated

**Background:** This visualization shows the profit ratio by countries. The colour on the countries indicates the profits and the profit ratio.

**Purpose:** The purpose of this visualization aims to shows what countries are doing well based on their profit ratio and profit. The red areas indicated that the profit value is negative and hence showing that this country helps improvement. While the darker the purple is, the better the country is performing. This helps the company to allocate the necessary resources to helps the countries that are in the red zones and even replicate best practices.

Moreover, it shows what counties have the highest/ lowest profit by just looking at the colour

Dashboard 01 Visualization 03

Chart, bar chart

Description automatically generated

**Background:** This visualization shows the profit ratio and profit based on the sub-category. The colour indicated the profit ratio while the bar graph is plotted against profit.

**Purpose:** The purpose of this visualization aims to shows what sub-category are doing better. This helps to understand what sub-category is generating the greatest number of profits and its profit-ratio. The darker the purple colour indicates the pro-ratio is high. Hence this graph shows a positive relationship between profit and profit-ratio (the higher the profit , the greater the profit-ratio) . This helps the management to determine what sub-category products are successful and helps with better planning.

Dashboard 02 Visualization 01

Chart, line chart

Description automatically generated

**Background:** This visualization shows sales trend based on the range of month chosen. This line graph contains filters such as by category , segment, and range of dates.

**Purpose:** The purpose of this visualization aims to shows the monthly sales trend to determine what month have the highest sales generated. Moreover, it graphs also shows sales based on category and segment that is useful for the management. As it determines what category generates the highest sales and the sales generated by consumer or corporate. This help the company to forecast sales and plan ahead.

Dashboard 02 Visualization 02

Graphical user interface, application

Description automatically generated

**Background:** This visualization shows the sales generated by 13 different regions. The purple bar indicated the regions which have the highest sales generated. This line graph contains filters such as by category , segment, and range of dates.

**Purpose:** The purpose of this visualization aims to shows sales generated by each region. It helps the management to determine the region that generates the highest sales in one look.

It is useful as it shows promising regions and hence help company to plan better.

Dashboard 02 Visualization 03

Chart, bar chart, histogram

Description automatically generated

**Background:** This visualization shows the sales generated by sub-category. The purple bar indicated the sub-category which have the highest sales generated. This line graph contains filters such as by category , segment, and range of dates.

**Purpose:** The purpose of this visualization aims to shows sales generated by each sub-category. It helps the management to determine the sub-category that generates the highest sales in one look. It is useful as it shows promising sub-category and hence help company to plan better.

Dashboard 02 Visualization 04

Map

Description automatically generated

**Background:** This visualization shows the sales generated by country. The size of the circle indicated the sales generated ( the bigger the size of the circle, the greater the sales generated). This map contains filters such as by category , segment, and range of dates.

**Purpose:** The purpose of this visualization aims to shows sales generated by countries. It helps the management to determine which country that generates the highest sales in one look. It is useful as it shows promising country and hence help company to plan better.

Dashboard 03 Visualization 01

Chart, bar chart

Description automatically generated

**Background:** This visualization shows the sales generated by each category based on the month chosen. The colour indicated the 3 different category and the greater the area of colour, the greater the sales generated. This bar graph contains filters such as by year and month.

**Purpose:** The purpose of this visualization aims to shows sales generated by each category. It helps the management to determine which days in the month sales are high and low. This can help to show the number of sales based on company promotions and discounts given. For example, May 5 the company have 5.5 sales. Hence the management can look into may 5th to determine the sales generated. This graph is meant for in dept analysis on sales based on days of the month.

Dashboard 03 Visualization 02

Graphical user interface, text, application, email

Description automatically generated

**Background:** This visualization displays the region with the highest number of sales. This visualization graph contains filters such as by year and month.

**Purpose:** The purpose of this visualization aims to shows highest sales generated by the regions. Hence shows the data of the manager to take note.

Dashboard 04 Visualization 01

Chart, bar chart

Description automatically generated

**Background:** This visualization shows the profit , sales and discounts based on the sub-category. The greener the colour the higher the profit generated, while the darker the colour red, the profit value lesser than 0. This visualization has a filter based on the country selected in another visualization.

**Purpose:** The purpose of this visualization aims to shows how discount affects the profits made by the company. By one look, the manager can determine sub-category that are not doing well and determine if discounts are the issue. For example, lenses have negative profits generated and discounts tends to play a key role as the discount given is generally high. Thus, the manager can made better decision when assigning discounts to the sub-category based on the country.

Dashboard 04 Visualization 02

Map

Description automatically generated

**Background:** This visualization shows the profit based on the country in the map. The greener the colour the higher the profit generated, while the darker the colour red, the profit value lesser than 0. This visualization is also used as a filter.

**Purpose:** The purpose of this visualization aims to shows what country have the highest profit and what countries need help. It is important for the manager, as it shows what area needs improvement and hence, they can dig deeper to help improve their overall profit. Also replicate some of those best practices across other countries.

Dashboard 05 Visualization 01

Chart, scatter chart

Description automatically generated

**Background:** This visualization shows the profit percentile based on the sub-category on different regions. The four different colours are different category. For example. the green colour (leading) shows sub-category products in a region that have high profit percentile for 2019 and 2020.

**Purpose:** The purpose of this visualization aims to shows which sub-category products in a region progression based on profit. The green colour indicates promising sub-category products in a region while the blue colour indicates more work have to done to help increase the profit of sub-category products in a region. In one look, the manger knows what areas are performing well and not so well.

Dashboard 06 Visualization 01

Chart

Description automatically generated

**Background:** This visualization shows the number of orders in a period of 2019 and 2020. this shows a trend. This visualization contains a filter such as category and segment.

**Purpose:** The purpose of this visualization aims to show the numbers of orders placed in a period of 2019 and 2020. This shows a fluctuating trend of orders placed. But when seen by year, the number of orders placed increased. Hence the company can take advantage of the trend and check what cause the increase such as discount or seasons to better plan for the coming years.

Dashboard 06 Visualization 02

**Chart, treemap chart

Description automatically generated**

**Background:** This visualization shows the number of order value by product sub-category. The darker the blue colour, the highest he average sales. This visualization contains a filter such as category and segment.

**Purpose:** The purpose of this visualization aims to show the average sales by each sub-category. Thus, by looking at it, the company can better plan for the sub-category that have the largest size as it is the only that generates the highest average sales.

Dashboard 06 Visualization 03

A screenshot of a computer

Description automatically generated with medium confidence

**Background:** This visualization shows the average sales based on the shipment mode. The 4 colours represent 4 different ship mode. This visualization contains a filter such as category and segment.

**Purpose:** The purpose of this visualization aims to see whether there is a relationship between shipment mode and the average sales generated. Hence helps the company to plan their shipping mode and increase efficiency or shipping charge.

Dashboard 07 Visualization 01

Chart, treemap chart

Description automatically generated

**Background:** This visualization shows the order status / turnaround time. the darker the colour, the greater the number of order placed. It uses a calculation field to determine the order status. (IF [Ship Date] = [Order Date] = 0 THEN "Early"

ELSEIF [Ship Date] - [Order Date] > 0 AND [Ship Date] - [Order Date] <= 3 THEN "On-Time"

ELSE "Late") . Also, this contains a filter based on the year.

**Purpose:** The purpose of this visualization aims to understand the order status and how efficiency each shipment mode it. For example, only express shipping mode is used to send orders that are placed on the same day as the shipping date. This helps the manager to understand the shipping mode usage and what orders are sent early. Late or on-time and improve based on the data.

Dashboard 07 Visualization 02

**Chart

Description automatically generated**

**Background:** This visualization shows the number of order placed under 4 types of shipping mode based on 12 month in a year. The 4 colour represents the 4 types of shipping mode. Also, this contains a filter based on the year.

**Purpose:** The purpose of this visualization aims shows what shipping mode is popular with the customer based on the number of orders placed. Moreover, it shows what months have the highest orders placed. For example, the popular shipping mode is standard.

Visualization 01

Chart, bar chart

Description automatically generated

**Background:** This visualization shows the top 10 customer that generates the highest profits. This bar graph contains a filter by category.

**Purpose:** The purpose of this visualization aims shows what customer generates the highest profit and also based on category. This helps the manager to better target the top 10 customer and also increasing the profit in the future. The category helps to company to show what to target for what customer.

Visualization 02

**Chart, bar chart, histogram

Description automatically generated**

**Background:** This visualization customer order frequency. This graph contains a filter based on month and year.

**Purpose:** The purpose of this visualization aims shows the distinct count of orders by customer gives the number of orders each customer made. This determines the number of orders that a customer is most likely to make. It shows in a form of the probability graph, and this can help to better plan for discounts that is given based on order count of customer.

Dashboard 08 Visualization 01

A picture containing chart

Description automatically generated

**Background:** This visualization shows the number of sales based on the markets. This line graph shows a trend of sales throughout 2019 to 2022.

**Purpose:** The purpose of this visualization aims shows trend is sales based on the different markets. It helps to determine what markets are doing well and what market is lacking behind. It helps the manager to better plan budget or allocate resources or discounts to help some markets grow.

Dashboard 08 Visualization 02

**Graphical user interface, application

Description automatically generated**

**Background:** This visualization shows the number of orders based on the markets on different category.

**Purpose:** The purpose of this visualization aims shows what category have the highest number of orders placed based on the markets.

Dashboard 08 Visualization 03

**Text

Description automatically generated**

**Background:** This visualization shows top 5 sub-category that generates the highest number of profits based on the market. It shows the sub-category name and its profit generated.

**Purpose:** The purpose of this visualization aims shows the top 5 sub-category that is generating the highest profit based on the markets. It helps the manager to oversee all the performance of market and what sub-category is popular among the markets.

# Dashboard

Dashboard 01

**Chart

Description automatically generated with medium confidence**

This dashboard shows the profit-ratio based on Regions, Country, and Sub-category. Hence shows what regions generates the highest profit and profit-ratio and also it is applicable for countries as well as sub-category. The user can use the filter by order period for more analysis and also use to segment filter to determine how consumer or corporate affects.

Dashboard 02

Graphical user interface, application

Description automatically generated

This dashboard shows the sales by monthly sales trend, sub-category, regions and by country.

This displays the sales , profit , quantity, and discount for the user to take note. Moreover, users can use the filter on category to determine hat areas are doing and not so well also use to segment filter to determine how consumer or corporate affects.

Dashboard 03

**Graphical user interface

Description automatically generated**

This dashboard shows the sales by days of the month for in dept analysis. Also lets the users know the top region that generated the most sales

Dashboard 04

A picture containing graphical user interface

Description automatically generated

This dashboard shows the sales, discount, and profit. But for this, there is more importance to the profit and how discounts affect it. Also let the user know what countries are doing well and not so well.

Dashboard 05

**Chart, scatter chart

Description automatically generated**

This dashboard shows the profit percentile from 2019 to 2020. User can immediately know the progress of each area.

Dashboard 06

**A screenshot of a computer

Description automatically generated with medium confidence**

This dashboard shows the Order value based on sub-category, regions. Moreover it also shows the minimum and maximum order value to helps the company o predict their budget and discount. The user can use the filter by category for more analysis and also use to segment filter to determine how consumer or corporate affects.

Dashboard 07

**Chart

Description automatically generated**

This dashboard shows shipping mode based on order count and the order status. This helps the company to better plan their shipping . the users can use the order period filter for more analysis on the data.

Dashboard 08

Graphical user interface, application

Description automatically generated

To interact with this dashboard, double click on the market until the market name is changed to black while the rest is changed to grey.

This dashboard shows the profit, sales and quantity based on the chosen market. The company can analysis each market individually with trend analysis on sales. Moreover it also shows the top 5 sub-category with the highest sales for each market.

# Recording

