

## **UNDERSTANDING THE DATASET**

**TEAM ID : PNT2022TMID05235**

### **Context**

Shopping online is currently the need of the hour. Because of this COVID, it's not easy to walk in a store randomly and buy anything you want. I this I am trying to understand a few things like

### **Customers Analysis**

Profile the customers based on their frequency of purchase - calculate frequency of purchase for each customer

Do the high frequent customers are contributing more revenue

Are they also profitable - what is the profit margin across the buckets

Which customer segment is most profitable in each year.

How the customers are distributed across the countries- -

### **Product Analysis**

Which country has top sales?

Which are the top 5 profit-making product types on a yearly basis

How is the product price varying with sales - Is there any increase in sales with the decrease in price at a day level

What is the average delivery time across the counties - bar plot

I will keep updating the analysis.

### **Content**

Once you download the file the rows you see are the details of the order done online by people across the globe in the time frame 1-jan-2011 to 31-dec-2014. There are no missing values in the majority of columns except postal code, you can drop it if not required.