

## **CHARAN N**

### **To Extract, Transform and Load the data:**

Step 1: Open PowerBI, click on get data from excel sheet.

Step 2: Select the necessary excel sheets and click transform.

Step 3: In the transform section, go through the data and check whether there is anything to be rectified in order to keep the data correct. It will be redirected to the power query editor.

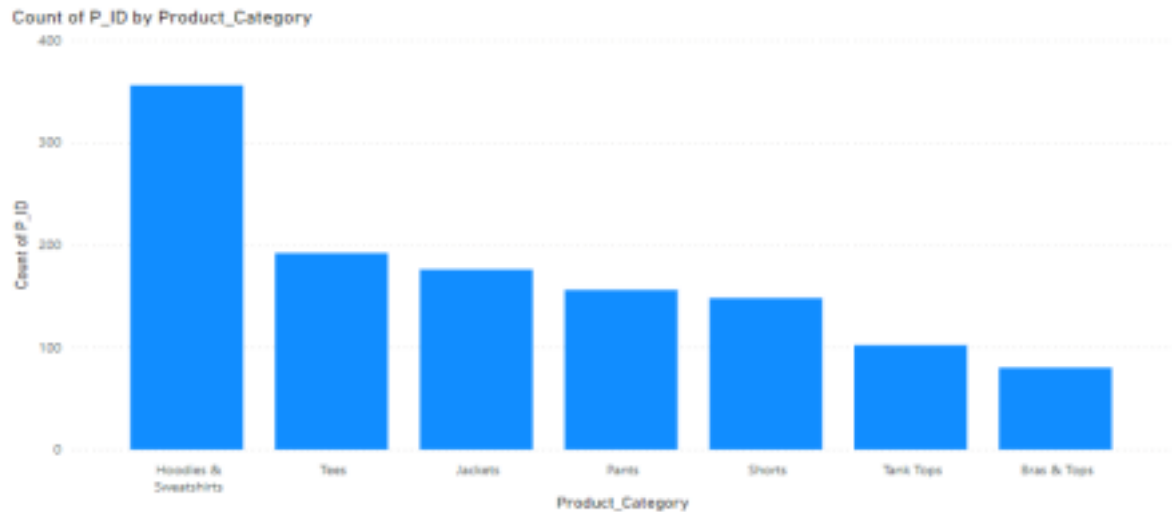
Step 4: After transforming the data, click on load. The data will be visible on the right side of the next page.

### **Data Cleaning and Transformation done in these datasets:**

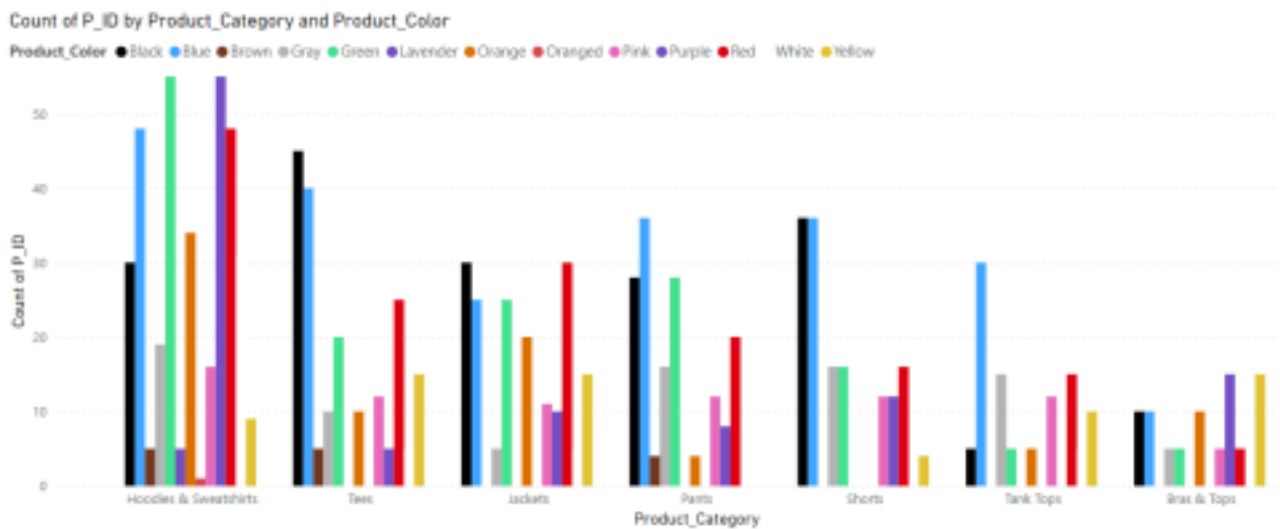
- ☐ Checked and removed all the null values in all the columns of all the tables of the datasets.
- ☐ Checked for the irrelevant rows and entries in the table.
- ☐ Checked for the wrong column names in the table. Product Cost Spelling is wrong in the Orders table. Country Spelling is wrong in retailers table. This is rectified using power query editor by right clicking the wrong column names and gave rename column.
- ☐ Checked for wrong format. The year and month format is wrong in the Orders table. To correct it, right click on the column, change format, click on year month. The format of the column will be rectified.
- ☐ These transformations are done and rectified before report process.

### **Visualizations:**

I have created a “Stacked Column Chart” for giving the information of which product is sold highest. For doing this, click on the stacked column chart and I selected “Product Category” as X axis and Count of “Product ID” as Y axis. By doing this I found “Hoodies and Sweatshirts” sold the highest. Then we can also see the order of the products that have been sold. The chart:



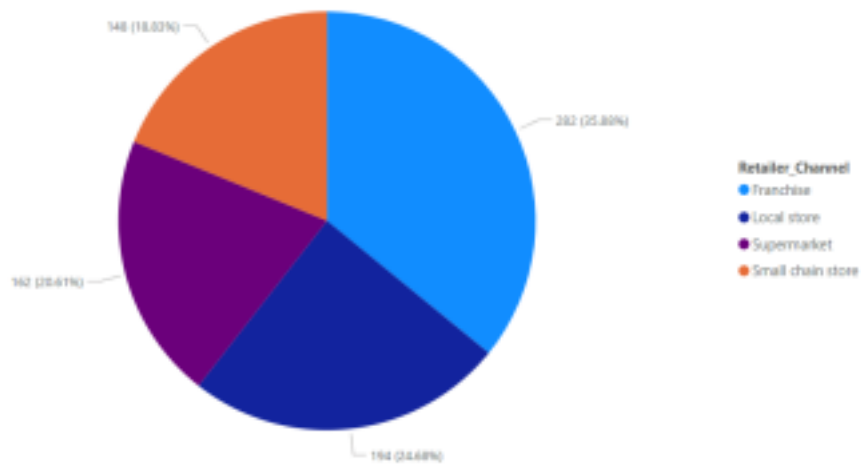
Next chart is another dimension with extra insights of previous chart by adding some attributes. This chart is made of “Clustered Column Chart”. This chart shows the number of products of different categories sold and the color of the products. This helps top understand the products sold and color of it using same chart. This chart is done by using “Product Category” as X axis and “Count of Product ID” as Y axis and “Product Color” as Legend. The chart is given below:



This chart tells the relation between the “Retailer Channels”. This shows the retailers of different Channels. Using this chart we can see the count and percentage of retailers selling in various channels. This is done by “Pie Chart” by having legend as “retailer Channel” and vales as “Count of Retailer ID”.

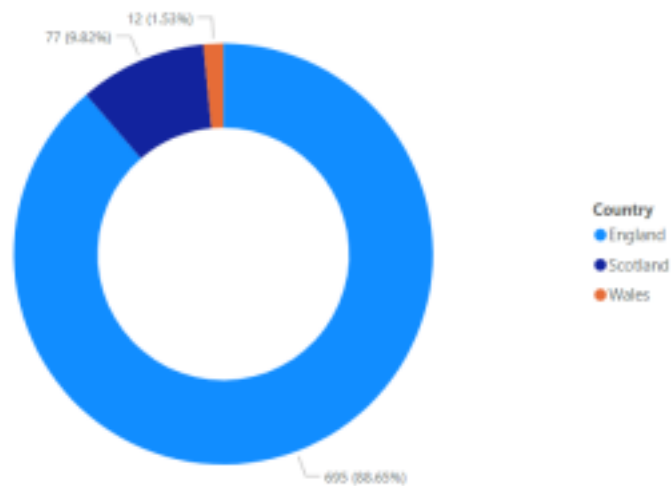
Chart:

Count of Retailer\_ID by Retailer\_Channel



Next "Donut Chart" gives the country of the retailers. This shows which country has the most number of retailers. After clicking the donut chart, I gave the legend as Country and Values as Retailer ID. The donut chart:

Count of Retailer\_ID by Country



## Created New Table:

Created new table “Retailers Type” and Connected it to “Retailers Table”

