Course Title: Software Project V

Teacher Information:

Fariha Jahan

Lecturer, Department of Computer Science & Engineering

Daffodil International University (DIU)

Dhaka, Bangladesh

Course code: CSE-317

Group Name

Little Giant's

Group Members

Nusrat Jahan Marjana

ID: 191-15-12999

Azmain Iqutedar Anik

ID: 191-15-12988

Md.Shakib Uddin

ID: 191-15-12995

Tahrima Tazim

ID:191-15-12998

Shahriar Sagor

ID: 191-15-12996

Applying Data Visualization on Shop marketing strategy: An empirical case of coffee shop industry in New York

Abstraction: This project is based on **Data visualization.** It makes **data** more natural and cleaner. To visualize the data, creator find a good dataset and push this in **Power BI**. Then they preprocessed the data set, create Relationship model and some charts such as, Stacked bar chart, stacked column chart, Stacked Area chart, 100% Stacked bar chart, Clustered Bar Chart, Map, Matrix, Donut chart, Pie Chart, Treemap, Gauge card and Slicer. **Coffee shop sample data**"- dataset is used in this project. The result of this dataset quit satisfying. But pastry inventory is a failed project for them.

Introduction: **Data visualization** is the graphical representation of processed data or information. It is the easiest way to understand Big data or information in short time. By using visual elements like charts, graphs, cards and maps, **data visualization** tools provide an approachable way to see and understand results, trends, outliers, and patterns in **data**. This is a report which based on Data visualization. It gives us a clear idea of what the information means by giving it visual context through maps or graphs. This makes the data more natural and easier to understand. Regular **visualization** helps to focus human mind on what they want. One of the problems many people face when aiming for goals is that they lose focus, becoming distracted by the goings on of life. When anyone **visualize** on a regular basis, it helps them to reset goal and it focuses their mind.

In this project. Datasets are collected from **Kaggle**, name of dataset "**Coffee shop sample data**". Kaggle allows user to find and publish dataset, it can still be a great learning tool for beginners.

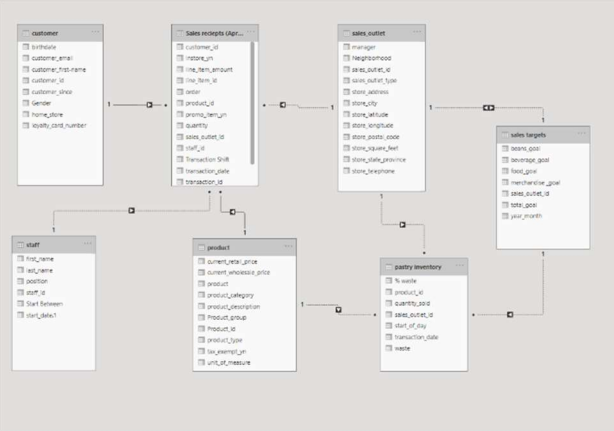
Here, **Power BI** desktop used as data visualization tool. Now, the Goal is to connect dataset into power BI and Visualize the data. After insert selected dataset, user Pre-preprocessed the data. Then set relations between entities and create different types of charts, graphs and. map for visualize the dataset. The result of this dataset quit satisfying. But pastry inventory was not carried good result or profit for them. They Face a big lose or wastage in this category of product.

Methodology: In this project, user visualize the dataset "**Coffee shop sample data**". This sample data module contains representative retail data from a fictional coffee chain. (Source: IBM). Here are Csv files of this dataset:

- **Sales Receipts** (Columns:15, Rows-49,894)
- **Pastry Inventory** (Columns:7, Rows-307)
- **Sales Targets** (Columns:7, Rows-8)
- **Customer** (Columns:8, Rows-2246)
- **Product** (Columns:10, Rows-88)
- **Sales Outlet** (Columns:12, Rows-9)
- **Staff** (Columns:6, Rows-55)

To enter new dataset in Power BI desktop, Click on **Get data** option under Insert tab. Here user can find some common data

sources, such as Excel, CSV, SQL Server etc. From those users can select any option and click on **connect.** Then select Dataset file from Pc which can find from any data site. In this project user input the selected dataset. After input any dataset, user must have Pre-preprocessed the data. Data Pre-preprocessing includes data cleaning, data integration, data transformation and data **reduction**. Before data visualization user try to prepare a Clean data set which has no Error or null Value. This all process is called Data Pre-preprocessing. After this take a look in relational model. Here Entities are connected to each other with their relationship.



After set the relational model properly, user can go in dashboard/report option. Here user make different type of charts for visualize the dataset.

Results & Discussion:

Results are the presentation of data and hence findings or investigations. In here user create 5 Dashboard/Report for visualization.

- **Report-1:** This report is based on Pastry Inventory. There are 3 type of visuals: Stacked Column Chart, Clustered Bar Chart and Slicer.

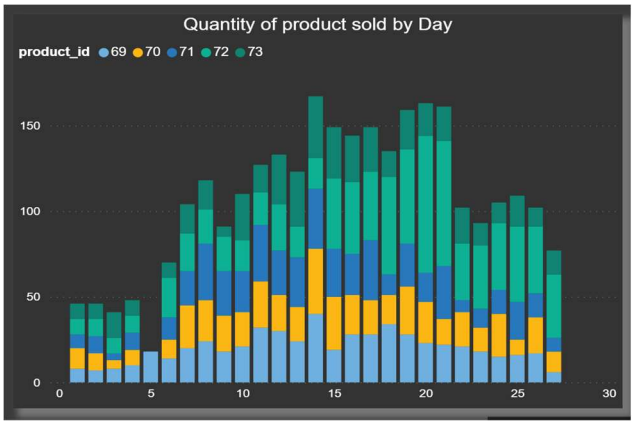


Fig 1: Stacked column Chart on Quantity of product sold by Day

In this column chart (Fig-1), 3 fields are used, transaction date **(**X axis), product id(legend) and Quantity of product sold(values).

Here they describe the amount of sold pastries in days. There are 5 kind of pastry, which can carry 5 number (69,70,71,72,73). By using the Slicer, it is easy to understand every single pastry's sold amount and selling progress.

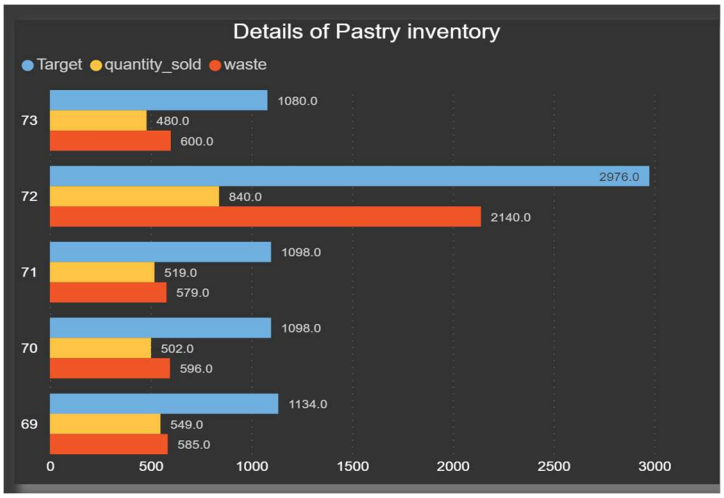


Fig 2: Clustered Bar Chart on details of pastry inventory

In this bar chart (Fig-2), 3 fields are used product id (X axis), target, quantity sold and waste (Values).

Here they try to take the full details about pastry inventory (In April 2019). There are 5 kind of pastry, in every pastry selling target failed in this inventory. 50-80% Product has wasted. By using the Slicer, it is easy to understand for every single pastry's sold amount, selling amount and wastage.

- **Report-2:** This report is based on Sells and Outlets. There are 4 type of visuals: Map, Clustered column Chart, Stacked Area Chart and Slicer.

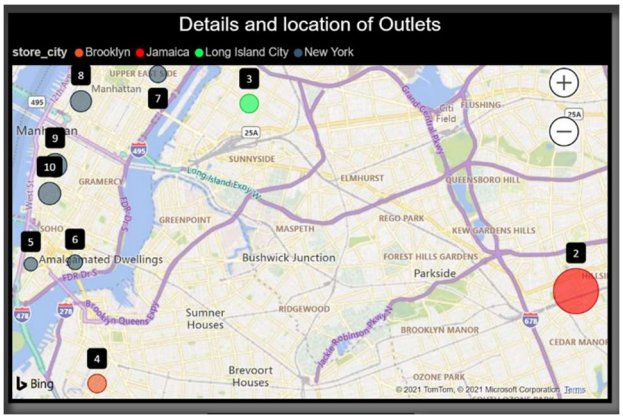


Fig 3: Map on details and location of outlets

From this map (Fig-3), there are 4 cities where 8 retails outlet and 1 ware house. Here full details are available about outlets. So, anyone can go there easily.

Fig 4: Clustered column Chart on sells goal in outlets

In this column chart (Fig-4), 2 fields are used, Seal's outlet Id **(**X axis) and total goal(values). Here Outlet 5 which is in New York city, has biggest selling goal than any other outlets in 2021.

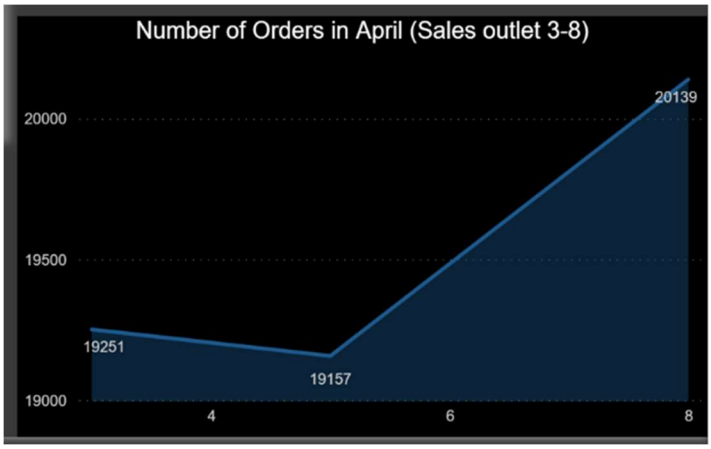


Fig 5: Stacked Area Chart on number of orders in April

In this area chart (Fig-5), 2 fields are used, Seal's outlet Id **(**X axis) and order (values). This is taken from Sales receipt of April 2019 (Outlet 3-8). Here outlet 8 completed 20,139 order which is the best amount in this chart. But Outlet 5 Completed less order than other outlets. It carried a biggest lose for this outlet. For that in 2021, they take a biggest selling goal.

- **Report-3:** This report is based on Staff and Customer. There are 2 type of visuals: 100% Stacked Bar Chart, Stacked column Chart.

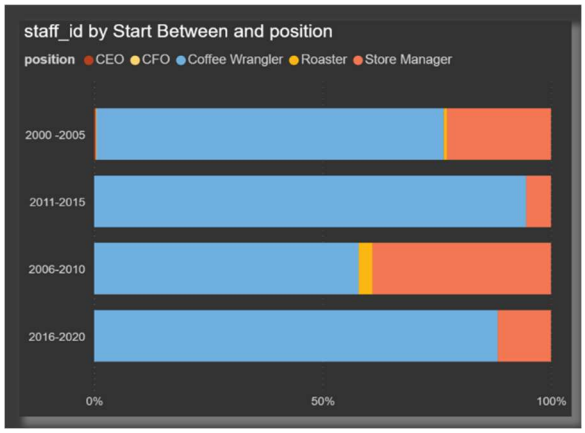


Fig 6: 100% Stacked Bar Chart on staff id by Start Between and position

In this bar chart (Fig-6), 3 fields are used, Start between **(**X axis), position(legend)and Staff Id(values). Here most of the staff are coffee wrangler and 2011-2015 they joined most. CEO and CFO of joined in the bigging of the years and then others are joined recent years.

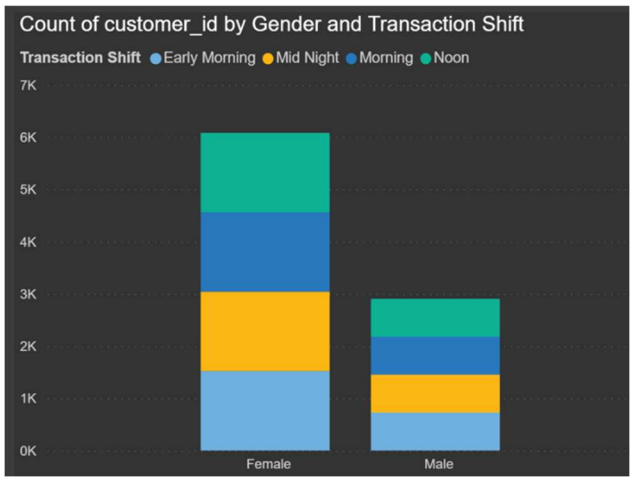


Fig 7: Stacked Column Chart on Count of Customer id by gender and transaction shift

In this Column chart (Fig-7), 3 fields are used, Gender **(**X axis), Transition shift(legend)and Customer Id(values). Most of the female Customer visited in the Coffee shops and almost in every shift they came there.

- **Report-4:** This report is based on Product category. There are 4 type of visuals: Treemap, Donut chart, Pie chart and Slicer.

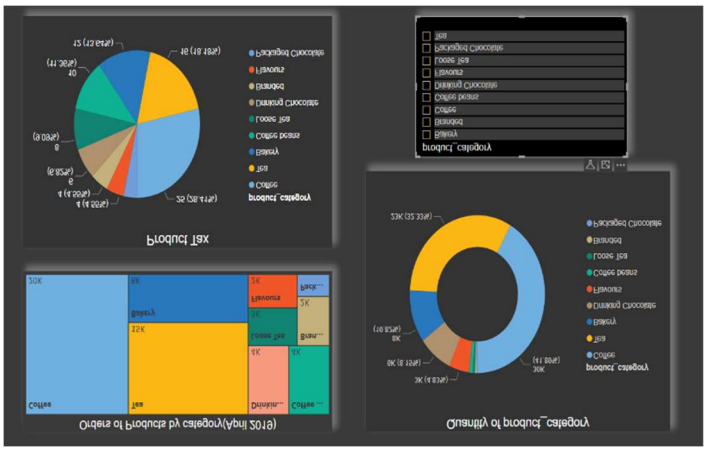
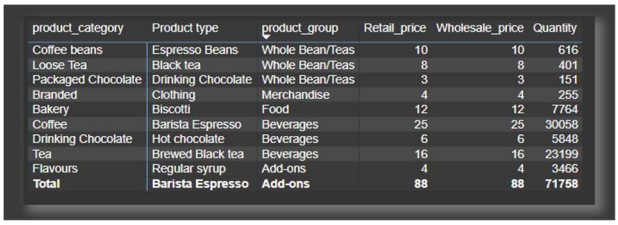


Fig 8: Treemap, Donut chart, Pie chart and Slicer on Order, Quantity and Current retail price of product

In this Charts, they try to describe all the possible way of product categories for visualization. Coffee is the most Countable category in that data. Using slicer, give them the best result for every category in this visualization.

- **Report-5:** This report is based on All Product and Target order. There are 4 type of visuals: Matrix, Gauge, Card and Slicer.



product_category	Product type	product_group	Retail_price	Wholesale_price	Quantity
Coffee beans	Espresso Beans	Whole Bean/Teas	10	10	616
Loose Tea	Black tea	Whole Bean/Teas	8	8	401
Packaged Chocolate	Drinking Chocolate	Whole Bean/Teas	3	3	151
Branded	Clothing	Merchandise	4	4	255
Bakery	Biscotti	Food	12	12	7764
Coffee	Barista Espresso	Beverages	25	25	30058
Drinking Chocolate	Hot chocolate	Beverages	6	6	5848
Tea	Brewed Black tea	Beverages	16	16	23199
Flavours	Regular syrup	Add-ons	4	4	3466
Total	**Barista Espresso**	**Add-ons**	**88**	**88**	**71758**

Fig 9: Matrix on Product category

In this Matrix, they describe some details about product categories for visualization. Product type, group, retail price, whole sell price and quantity.

Fig 10: Card on Quality Sold of pastry

In here, they describe the Quantity of sold pastry in April 2019.This is a good amount which they sold in only one month but they face some wastage also.

Fig 11: Gauge on Order product from sales receipt

In April,2019, They took a target and that is 40000 Order they will must complete in this month. They successfully complete the target and around 19000 order overloaded. They got a good amount of profit in that month.

Conclusion:

Data visualization helps people organize, understand, and **use data** to its fullest potential. This might include pie charts, bar graphs or any other visual aid that helps its audience find the answers they need. In this project, they use a medium dataset of Coffee shop which is in good position but their past pastry inventory is failed and they get some lose. So, in this focus part, they have to give more and more attention. In 2021 they take some target if these are completed successfully then they will get a great profit in this year.