# BUSINESS INTELLIGENCE PROJECT

## Chosen Dataset: Dashi Foods

**Link**: https://ibacity-my.sharepoint.com/personal/abeeratariq_iba_edu_pk/_layouts/15/onedrive.aspx?ga=1&id=%2Fpersonal%2Fabeeratariq%5Fiba%5Fedu%5Fpk%2FDocuments%2FBI%20Project%20Datasets%20for%20Students%2FDashi%20Foods

## Group Members:

- Muhammad Sumair 22995
- Khansa Junaid 23074
- Saad Ahmed Siddiqui 23030

## About the Dataset:

This dataset pertains to Dashi Foods, a local Pakistani company focused on creating and introducing new food products that offer both convenience and great taste. The dataset comprises multiple sheets with a database-like structure, with various ID columns linking the sheets together. Sheets are as follows:

- **Customer:** This sheet contains information about the customers like whether they are active or not, where do they live etc. The data in this sheet is completely dimensional.
- **Items:** This sheet is like the Customer sheet; this sheet was also rich in Dimensions. Moreover, this sheet contains the details of the items like item name, its category etc.
- **CustomerWise:** This sheet contains the details of each transaction of a customer. It shows the items the customer is buying, their quantity, price etc.
- **CustomerLedger:** This sheet contains all the information regarding the balances of each customer. It gives an idea of how the transactions are being endured and through what means.
- **Purchase:** This sheet contains the data regarding the purchase of each item, the category and quantity of the product ordered by the customers
- **Datadump:** This sheet contains the data of both the customers, and the products. It can be used for customer profiling or their analysis because it contains the brands of the products that each customer is buying.

## Data Cleaning & EDA:

We checked all six sheets individually to clean the data efficiently. Firstly, we check the missing values in each of the columns. There were very few missing values, so to treat them we replaced the values with unknown because we could not drop the entire columns as the missing values were less than 70%. For example, in NTN number we replaced the missing value with “unknown”

and similarly in rest of the columns with similar data. We then checked the columns that are potentially the primary key in each sheet. We checked whether they are unique or not so that we are aware that the data is entered correctly.

Furthermore, we have made a MSNO matrix, MSNO bar to check any missing values that were left, and the heatmap to check the correlation between the columns for the analysis. After this, we checked the inconsistencies in the columns in all the sheets. Very few data was inconsistent like in the column “Mobile Number”, “NTN”, and “Contact Number”; there were dashes in these columns where there was no data, so we replaced those dashes with the null values and then replaced it with “unknown” in order to make it consistent with all the other columns in different sheets because we have used “unknown” in the instances where there were missing values, to make it look consistent.

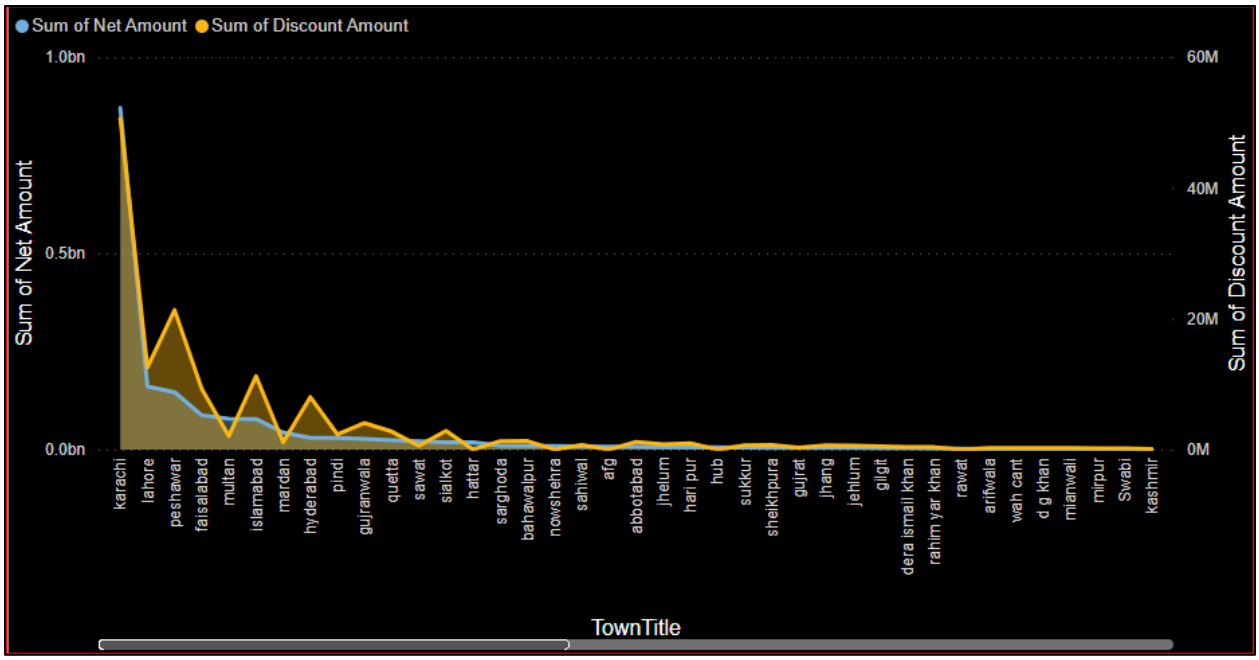
We then also performed univariate analysis of the data, that is we created histograms and box plots for anomaly and outlier detection by performing Tietjen Moore test. Lastly, we performed bi-variate analysis of data in which we performed Chi Squared test to check for dependencies.

## Problem Statements:

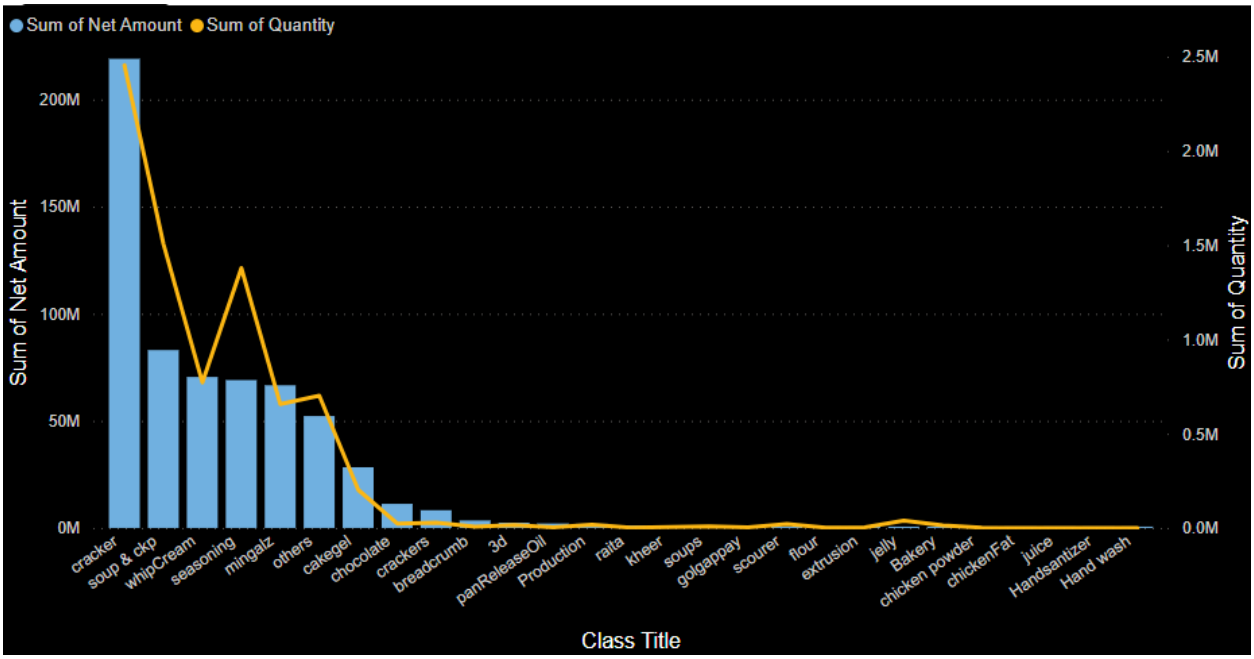
1. **Understanding the Sales Pattern of Dashi Foods:** In this problem statement we are analyzing sales of Dashi Food with respect to different factors. We are observing which towns have the highest sales, which distributor is making the highest sales, and what brands are being sold the highest and most importantly which months have the highest sales. Furthermore, we have also analyzed which items are selling the most. Analysis of sales pattern in this way can help Dashi Foods fill their inventory and the towns where they should be more focused, because understanding the sales pattern can help them grow in revenue as well as gaining customer loyalty. Customer loyalty can be gained by focusing and innovating the items that are sold the most, they can improve the prices etc. Understanding the sales pattern or its trend can be very useful for any brand; thus, we are analyzing this for Dashi Foods.
2. **Demonstrate consumer purchasing preferences and their behavior in relation to Dashi Foods' online store:** In this problem statement we are analyzing the customers of Dashi Foods. This is an important tool for any brand or company, which is to understand its customers and their preferences, because knowing customers preferences and their behaviors the company can play with the pricing, discounts, and the demand and supply of the products accordingly. In analyzing this problem statement, we can also analyze what type of customers Dashi Foods has and in which city or town what item has the highest popularity. We can also see which of our customers are registered and what type of transactions do they prefer so that the company can follow that transaction type of the customers to make it easy for them, to attract more customers.

## Charts:

1. **Problem Statement 1:**

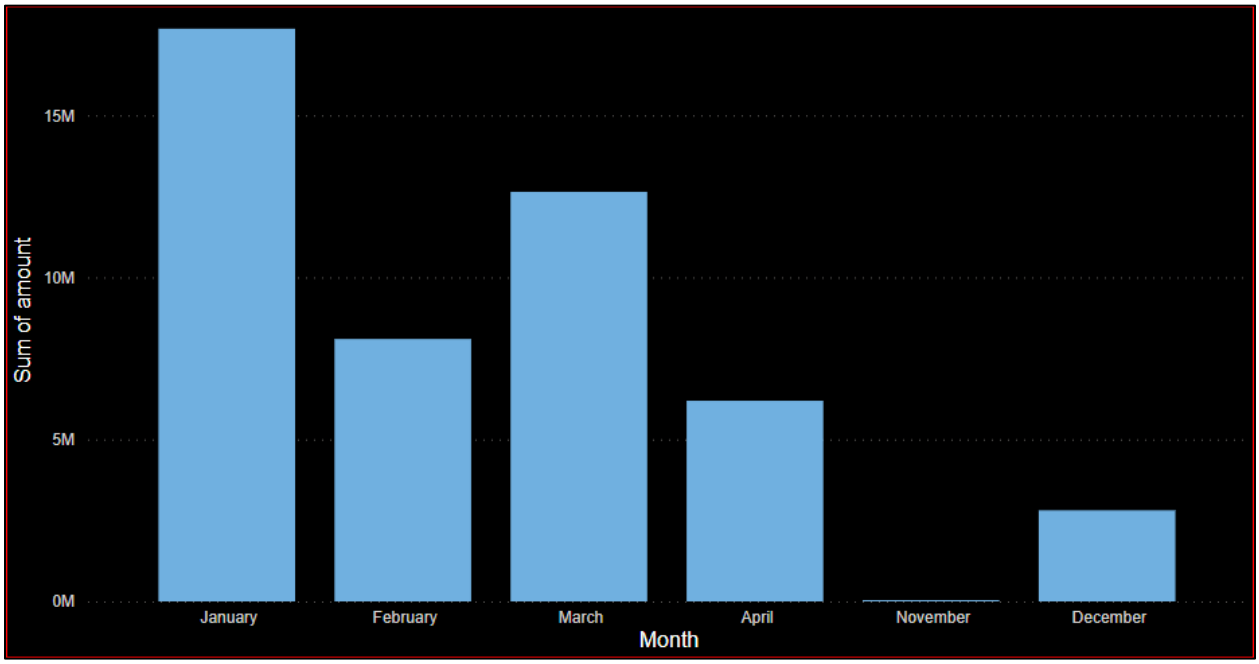



This is an area chart which shows the total net amount or sales and the total discount amount with respect to the towns or cities. The blue area line represents the net amount, and the yellow area line represents the discount amount. We can see that the highest sales and highest discount offered is in Karachi.

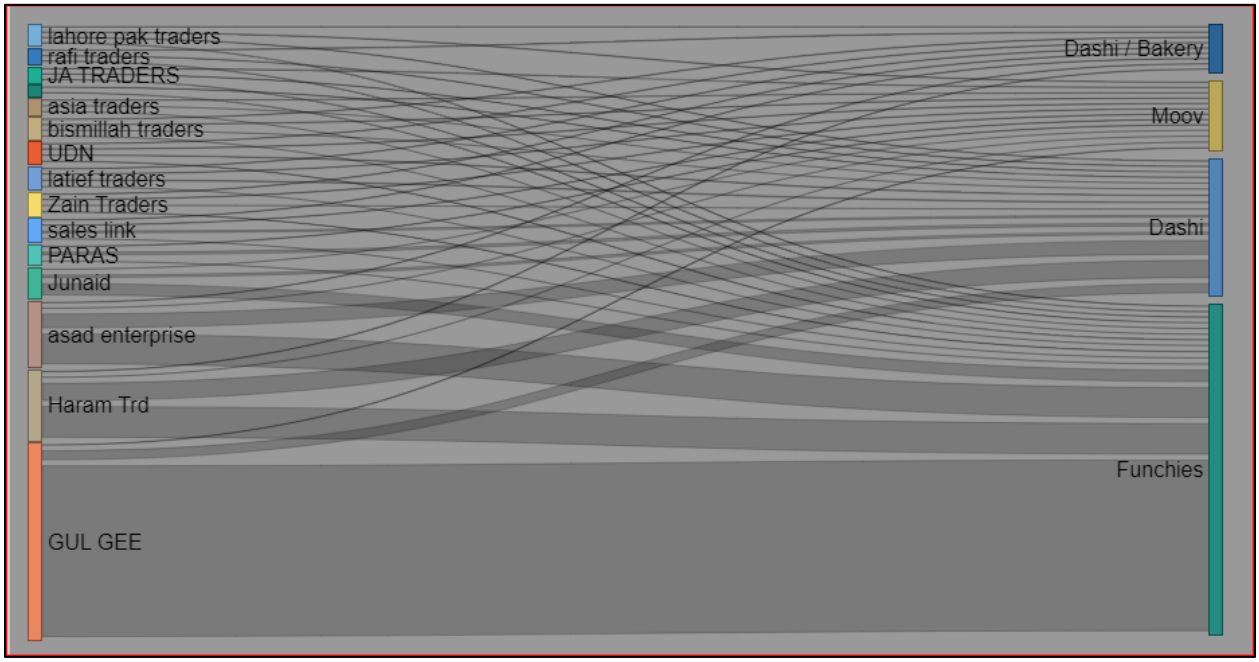


This is a bar and line chart which shows the total sales with respect to each class. Each class consists of items for example in this chart we can see that crackers make up to the highest selling Class in terms of revenue. Which means that items in the Class “Crackers” combine up to be a major part of sales. Moreover, the yellow line represents the total quantity, which is also highest for the Class “Crackers” and the lowest for “Hand Wash”.


This is a word cloud chart which shows different items according to the times they have been bought. The size of the word which is the item name represents the times it has been bought. We can see that the size of the word “Chicken” is the biggest hence it has been bought regularly. Similarly, we can see Powder, Seasoning, Papad, Crackers and Mingalz in bigger sizes. So we can deduce that these are the products which are sold on a regular basis as they are highlighted in this chart.



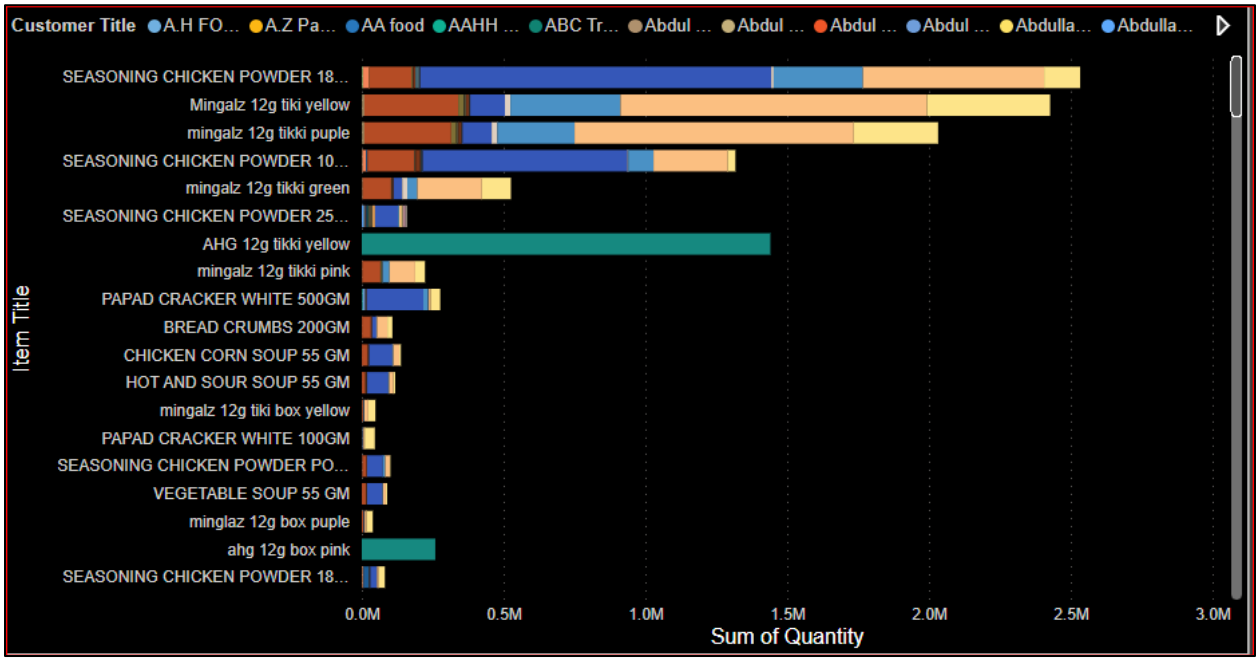

This is a simple bar chart which shows the net amount or sales in each month. We can see that January has the highest sales while November has the lowest sales. The company can see the trend of sales in each month with this chart.



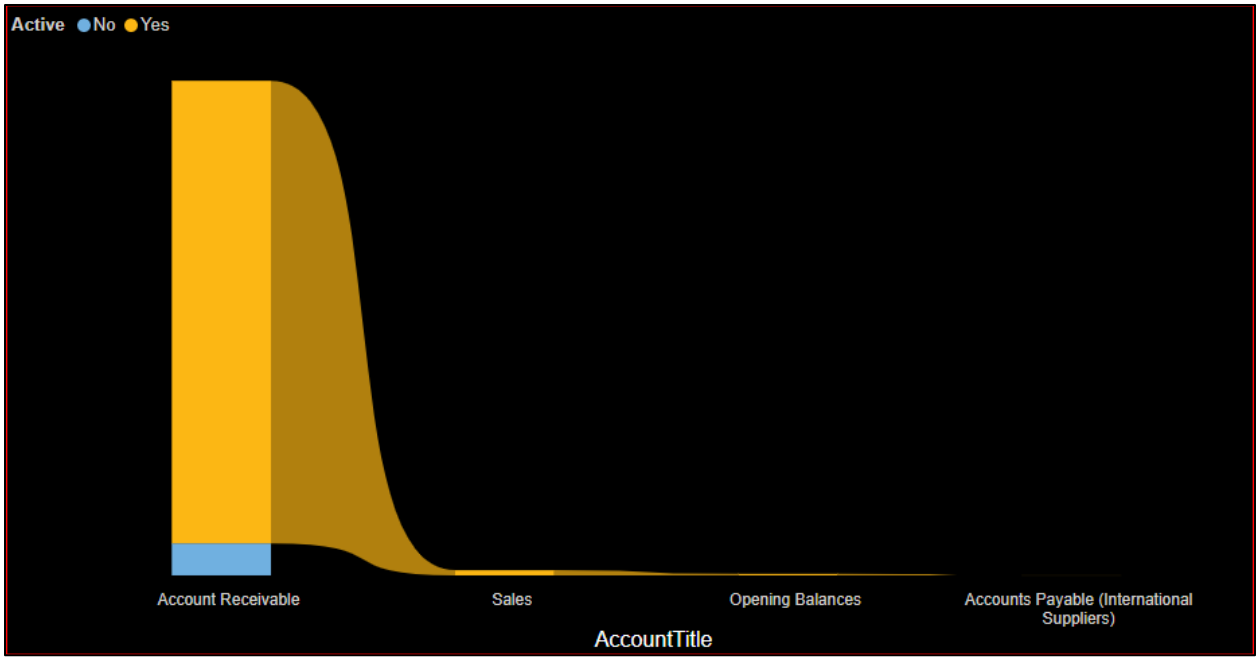

This is a Sankey chart which shows distributors name on the left side and the brand names on the right side. The size or weight is defined by the quantity sold. We can see that the highest selling products are from the brand Funchies which is by the distributor Gul Ghee. Similarly other

distributors also combine and make up the full quantity of Funchies, and the brand which is second highest in terms of quantity is Dashi and furthermore. So, we can analyze which distributor is selling the highest products of which brand.

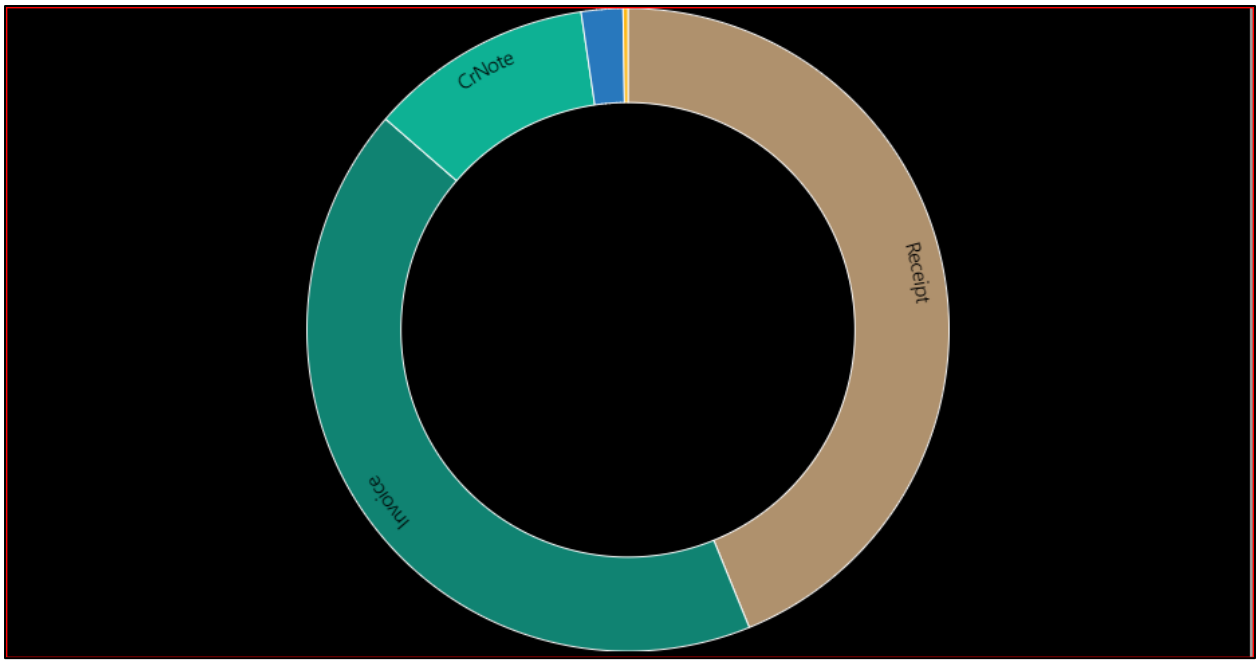
## 2. Problem Statement 2:



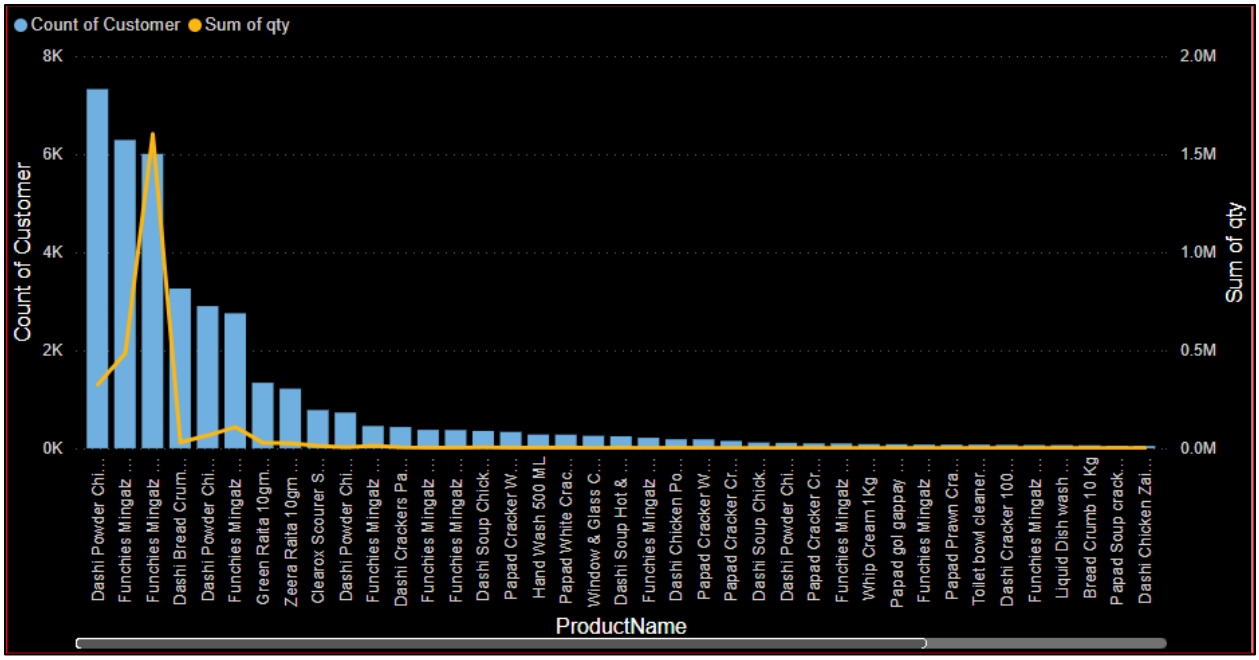

This is a horizontal bar chart which shows each item according to the total quantity that was bought by the customers. The colors represent the customers. So, in this chart we can see that “Seasoning Chicken Powder” is bought the most by the customers with over 2.5 million kg. Furthermore, we can also observe that it was mostly bought by AA food which holds a major part of buying “Seasoning Chicken Powder”. Similarly, if we scroll down, we will see that there are some items which are never bought by the customers.



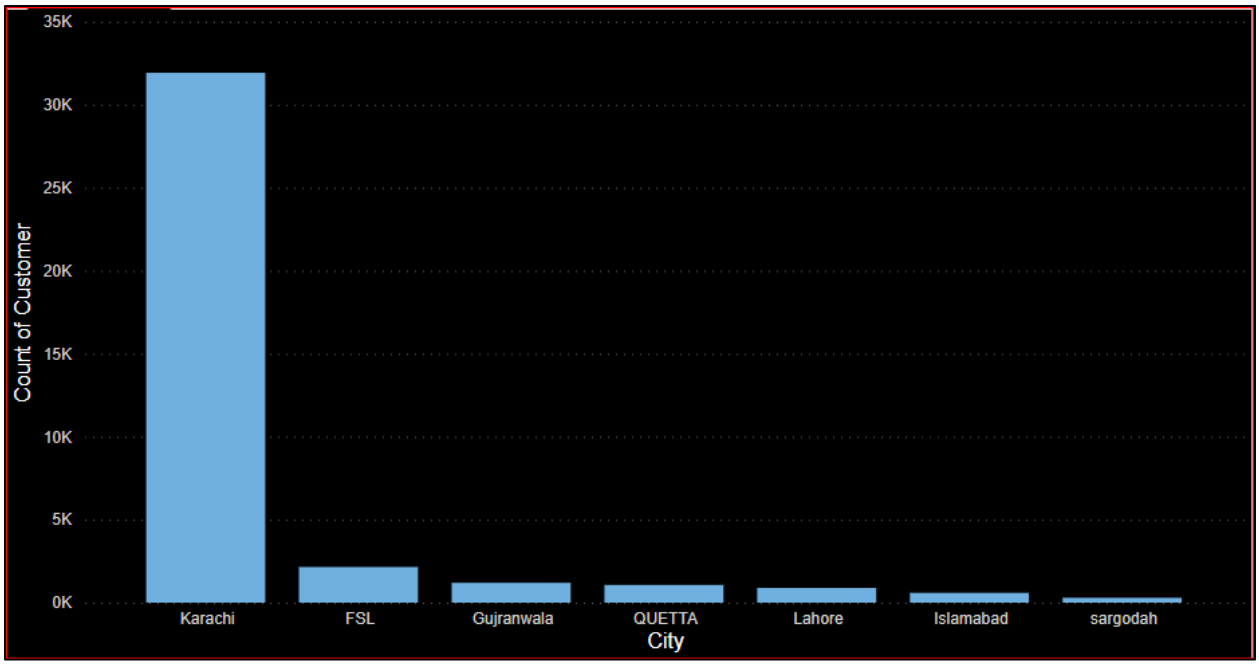

This is a ribbon chart which shows different Account Titles. The colors in this chart represent whether the account is active or not. We can see that there are almost all the accounts that are Account Receivables and majority of them are active which is represented by yellow color and very few are inactive. Moreover, there are no Account Payables as we can see in the chart.



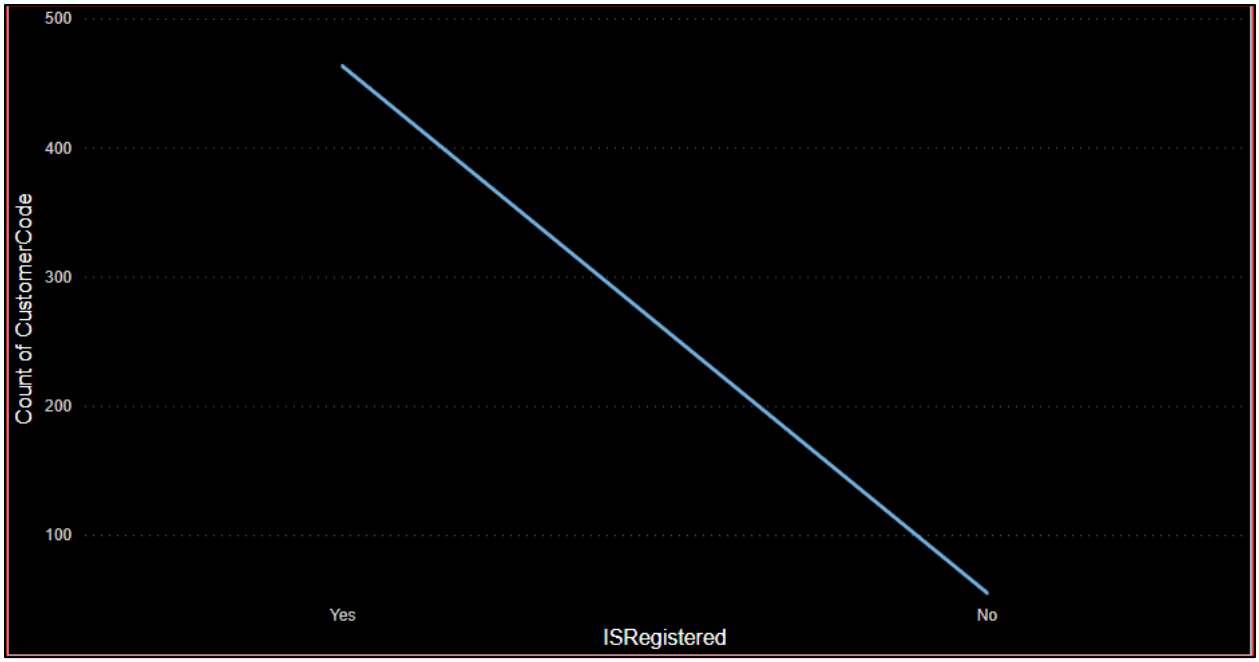

This is a sunburst chart which shows how many transactions are done on which transaction type. This is to analyze which transaction type is preferred by the customers. We can see that most transactions are on Receipt, which means that customers prefer it more. And there are almost 71 transactions on Invoice Return, which is the lowest.



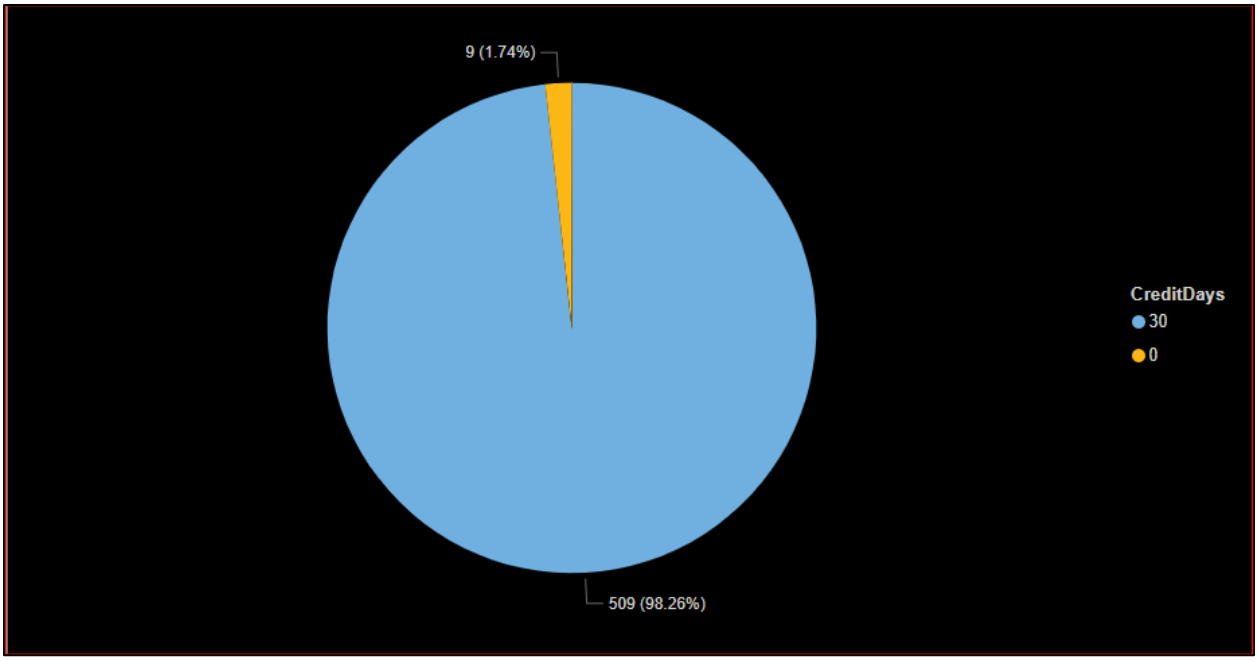

This is a bar and line chart showing which customers are buying which product. The line shows the total quantity. We can see that the customer buys Dashi Powder the most as it has the highest quantity, while Funchies Mingalz has the highest number of quantities bought. We can analyze that there might be a few customers that are traders or distributions which are buying huge quantity of Funchies.




This is a simple bar chart which shows which city has the highest number of customers. We can see that Karachi has the highest number of customers while Sargodah has the lowest.

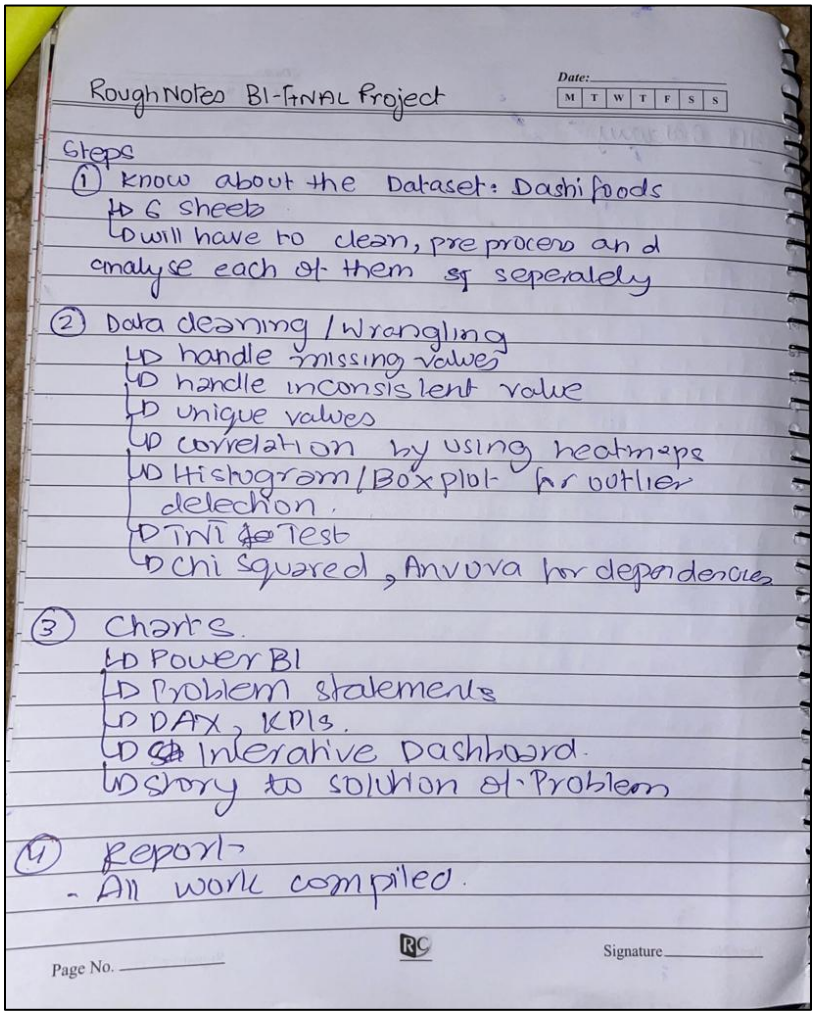



This is a simple line chart showing how many customers are registered and how many are not because the company does have NTN of all the customers. In this, we can see that most of the company's customers are registered.




This is a pie chart which represents what percentage of customers prefer the credit of 30 days and what percentage of customers are not on credit terms. We can see most customers are using a credit service of 30 days.

## Dashboards on Paper:



Rough Notes BI-FINAL Project

Date: M T W T F S S

Steps

(1) Know about the Dataset: Dashi foods
- 6 sheets
  - will have to clean, preprocess and analyse each of them seperately

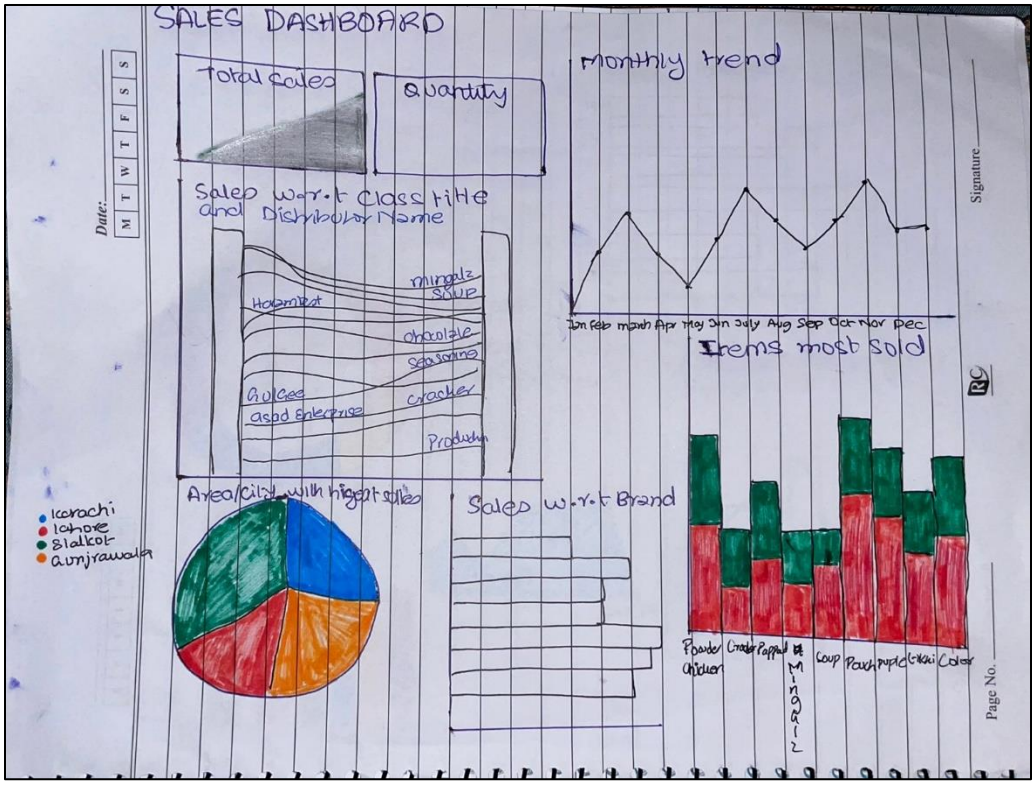
(2) Data cleaning / Wrangling
- handle missing values
- handle inconsistent value
- unique values
- correlation by using heatmaps
- Histogram / Boxplot for outlier detection.
- TWI Test
- Chi Squared, Anvova for dependencies

(3) Charts.
- Power BI
- Problem statements
- DAX, KPIs.
- Interative Dashboard.
- story to solution of Problem

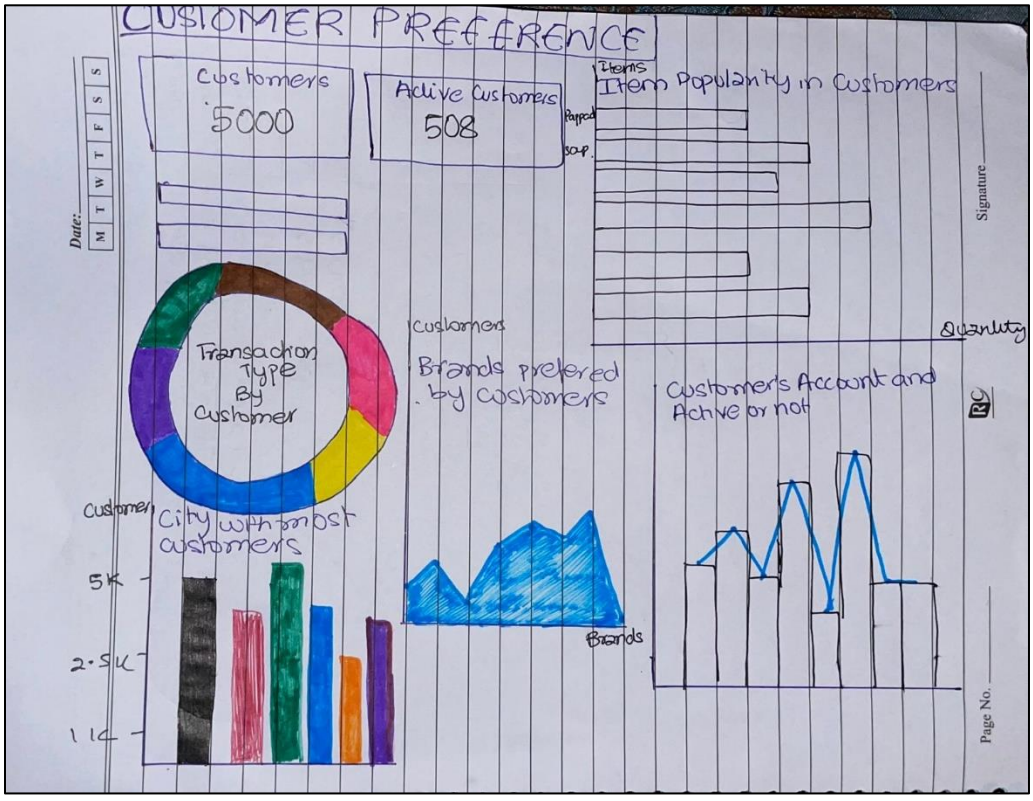
(4) Report-
- All work compiled.

Page No. _____ Signature _____

Rough Notes



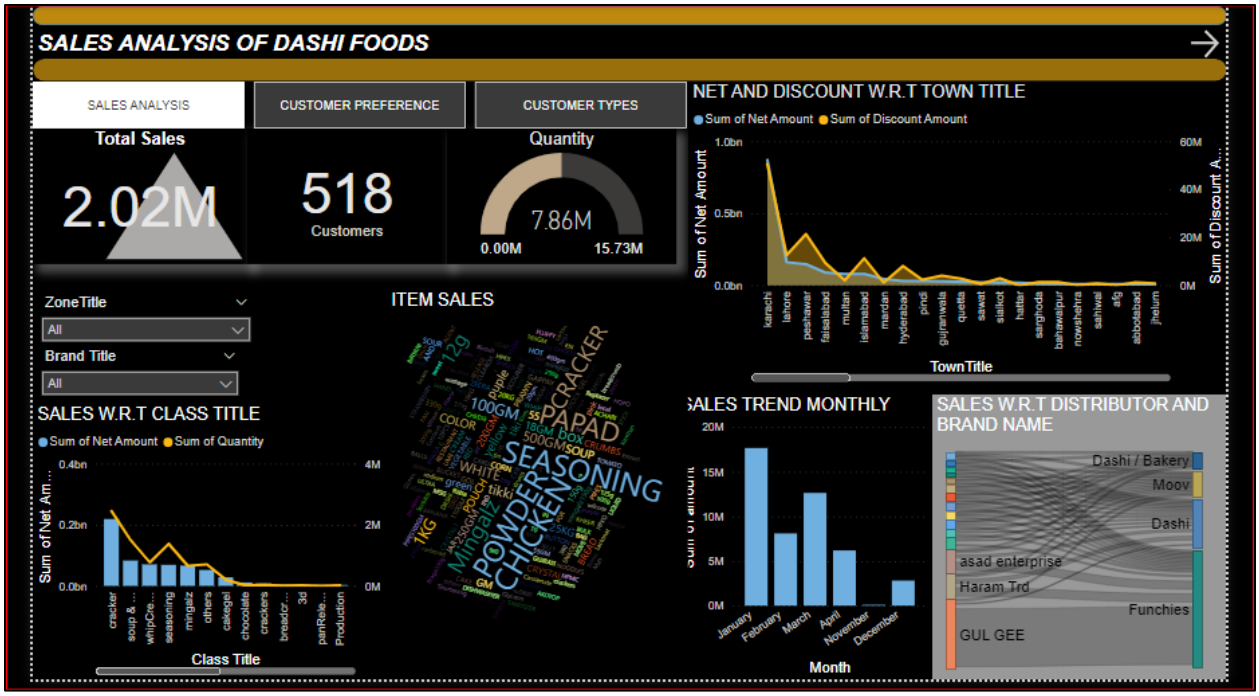

**Problem Statement 1**




**Problem Statement 2**

## Dashboards:

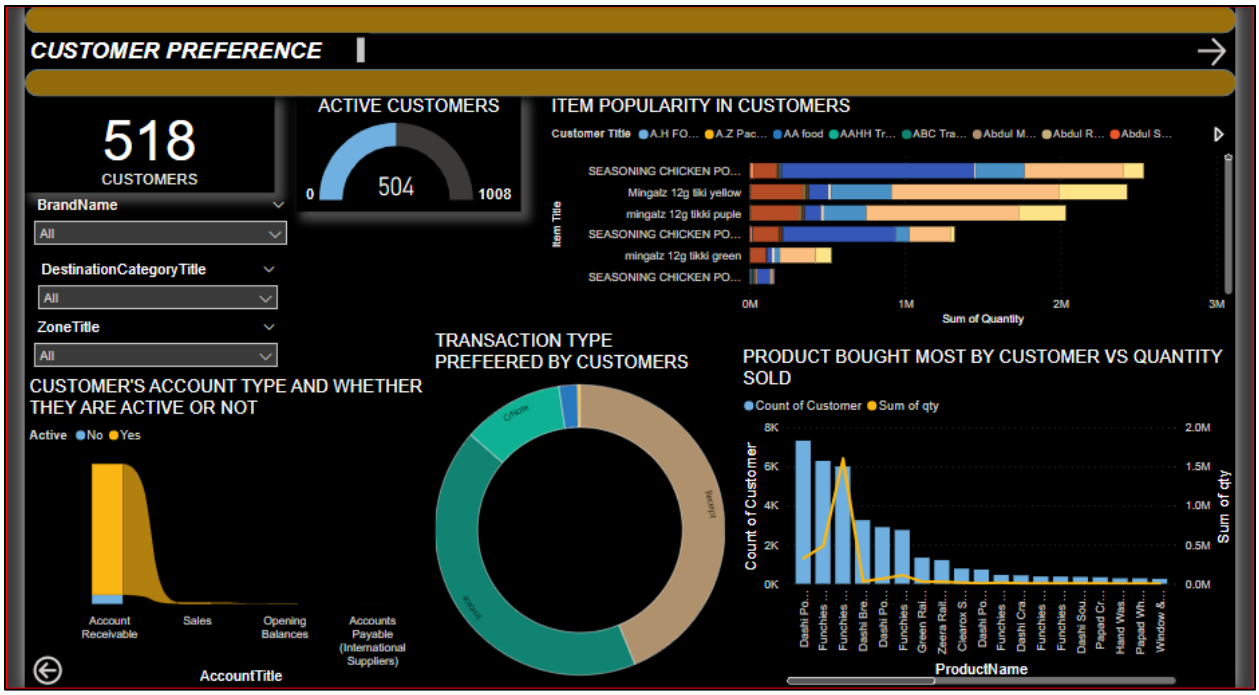
- **Problem Statement 1:**

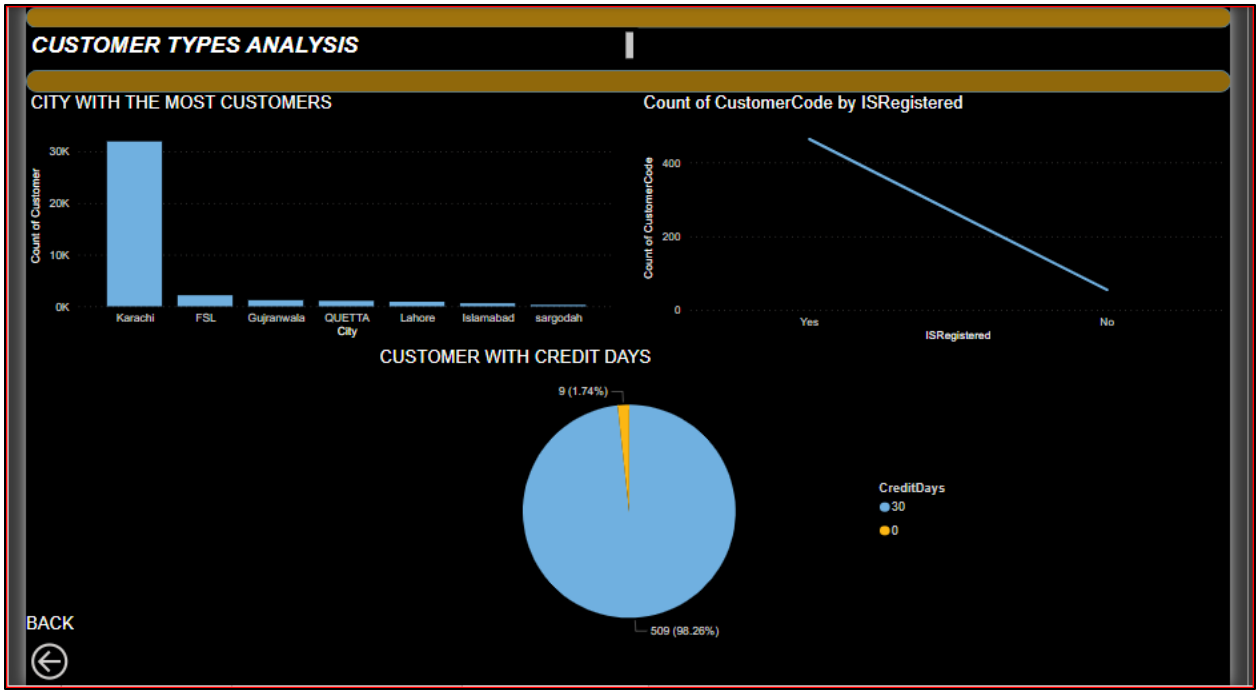



This is our dashboard for the first problem statement which is about understanding the sales patterns. We have used DAX and KPIs to calculate total sales, total quantity, and the total customers the company has. Furthermore, we have used drop down menus for Zone and Brand title which can be used for analysis of sales pattern in different zones and of different brands for example we want to see the sales pattern of finished goods in Sindh, we can use the drop-down menus for that. Moreover, all these charts can be helpful in understanding the sales because from these charts in the dashboard we can see what items are sold the most, which distributor is selling what brands the most, what are the months when the sale is the highest, which city has the highest sale and discounts etc.

By looking at these charts, the company can change their prices and offer discounts accordingly. Additionally, it would be easier for the company to fill in the stock of items according to the charts that have been made from the data available because these charts can help the company grow in future by providing the past data.

- **Problem Statement 2:**




This is our dashboard for the second problem statement which is about the company's customer preferences and their behavior. We have divided our dashboard into two stories where the first one represents the customer preferences and the second one represents types of customers. We have made DAX to calculate the total number of customers and the number of customers that are active. Furthermore, we have used the drop-down menus for Brand Name, Destination

Category, and the Zone which can be used for the analysis of customer preferences. This is so because with the help of drop-down menus we can see where most customers are attracted to specific brands and what type of customers they are. For example, the company wishes to check how many customers that have restaurants in Punjab are buying the products of Dashi Bakery.

This dashboard can help the company understand their customers better, because the company can see what are the items and products that are most popular amongst their customers, what type of transactions are preferred by their customers, what type of accounts are preferred by the customers and whether most of the customers are registered or not. Secondly, the company can also analyze which city has the highest customers of Dashi Foods and how many customers are registered and have a valid NTN number, and what percentage of the customers are more comfortable with and without the credit terms. In this way, the company can offer their services to the customers according to their preferences so that more and more customers are attracted to Dashi Foods and help them incur profits.

## Solutions:

- **Problem Statement 1 - Understanding the Sales Pattern of Dashi Foods:**
  1) We can see that Karachi has the highest number of sales of all the cities, so the company should cut off on little discounts to the customers there and offer greater discounts in the towns where there is lower sale so that more customers from that town are attracted. This is so, because the town like Karachi has a greater purchasing parity, so the customers there will still buy the products in less discounts, but the towns with lower purchasing power will tend to buy more if they see inexpensive or cheaper items comparatively.
  2) We can analyze that January and March have the highest and the second highest sales. So, the company should offer schemes and special discounts in other months so that the customers are attracted more and start to buy in other months as well. In this way, Dashi Food will manage to have good business the whole year.
  3) The company should offer discounts to the distributors who are buying more. And the company should offer some schemes or seasonal offers on the brands that are bought more by the distributors. Secondly, incentives should be given to small distributors to encourage them to expand.
  4) The high selling items should be fully stocked in each of the towns. Since the company knows which product is selling the most in which town, they should not let the stock of the popular products or items finish off. In this way, people will be able to buy the products instantly without any hesitation of stock finishing up.

- **Problem Statement 2 - Demonstrate consumer purchasing preferences and their behavior in relation to Dashi Foods' online store:**
  1) Since the company will be aware of the fact that which items are bought the most and the least by the customers, the company can start to lower the prices of the items that are not sold. Furthermore, the company can also start to offer discounts or schemes like “Buy a certain item and get the other item”. In this way the items that are not sold or are sold at a very slow rate will get more attention from the customers.
  2) The company should deal with the customers in the form of receipts as we can see that most of the customers are comfortable in dealing with receipts and invoices. In this way, customers will be satisfied, and they will not hesitate to order more, because they would know that Dashi Foods keeps customer service its utmost priority.
  3) We can see that there are very few accounts that are not active; so, the company should offer some incentives like discounts, vouchers, and gifts to the customers that are inactive as a welcome back present. There is a good chance that those customers will start to buy again from Dashi Foods because almost every customer is attracted by gifts, vouchers, and discounts.
  4) Since we know by the analysis that most of the customers of Dashi Foods are from Karachi, the company should start some campaigns and advertisements in other cities where there are very less or no customers, so that the people are encouraged to buy from Dashi Foods.
  5) There are very few customers that are not registered; so, the company should encourage them to get themselves registered. This can be done by telling them the benefits of being registered and how the tax charges will be reduced once they are registered, because taxes for unregistered companies are greater as compared to the companies that are registered.
  6) We can see that most of the customers are comfortable on credit terms with the company. So, the company should offer their customers some leverage for credit terms, and on the side, Dashi Food should offer some discounts to the customers that are not on credit terms. In this way, most of the customers will tend to pay early or without any credit basis to attain the discounts. And this would also be beneficial for the company as well.

## Team Members Contributions:

- **Muhammad Sumair:** Worked on identifying the problem statements, ideas for charts about what should each chart be about, writing the report, and identifying solutions to each problem statement.
- **Khansa Junaid:** Worked on identifying the problem statements, making the charts for both the problem statements, gathering background data knowledge file cleaning and the EDA.
- **Saad Ahmed Siddiqui:** Worked on identifying the problem statements, understanding the problem statements, and identifying solutions to each problem statement.