# 2004-NPcrestHori_B-W

# School of InfoComm Technology

**Data Exploration & Analysis Assignment**

Diploma in DS

April 2022 Semester

**ASSIGNMENT 2**

(40% of DEA Module)

11th July 2022 – 7th August 2022

**Submission Deadline:**

**Presentation: 7th August 2022 (Sunday), 11:59PM**

**Report and files: 7th August 2022 (Sunday), 11:59PM**

|  |  |  |
| --- | --- | --- |
| **Tutorial Group** | **:** | **T03** |
| **Student Name** | **:** | **Brian Low De Quan** |
| **Student Number** | **:** | **S10228107D** |

**Penalty for late submission:**

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

**NO** submission will be accepted after 14th August 2022 (Sunday), 11:59PM.

# **Table of Content**

Contents

[1](#_Toc110720497)

[1](#_Toc110720498)

[School of InfoComm Technology 1](file:///C:\Users\Brian\Desktop\Poly%20Folder\Year%202%20Semester%201\DEA\Assignment%202\Brian%20Low%20De%20Quan_S10228107D.docx#_Toc110720499)

[**1.** **Table of Content** 2](#_Toc110720500)

[**2.** **Abstract/Overview** 3](#_Toc110720501)

[**3.** **Dashboards** 4](#_Toc110720502)

[ **Question 1** 5](#_Toc110720503)

[ **Question 2** 6](#_Toc110720504)

[ **Question 3** 7](#_Toc110720505)

[ **Question 4** 8](#_Toc110720506)

[ **KPI Dashboard** 9](#_Toc110720507)

[**4.** **Further Analysis** 11](#_Toc110720508)

[ **Univariate** 11](#_Toc110720509)

[ **Bivariate** 12](#_Toc110720510)

[ **Multivariate** 13](#_Toc110720511)

[**5.** **Summary** 14](#_Toc110720512)

# **Abstract/Overview**

Our company, which is founded in 2020, called FoodFresh has established their brands which consists of BunFresh, KopiFresh, NoodleFresh, RamenFresh and CakeFresh. For the past 2 years, the company have branches off to all part of Singapore and looking into expanding their business operation even further.

As a result, the CEO of FoodFresh set up a business analytics team to provide business intelligence reports and analysis to aid in company growth and provide the business with the right tracks of sales inquiries that can aid them in identifying the key performance that influences their sales regardless of good or bad.

We were given the task of identifying trends and contributing to the creation of recommendations for improving business performance using the critical analysis we discovered in our dashboards. Developing important questions can also help managers make informed marketing decisions. My report's goal is to offer management advice and critical analysis to help the company achieve its goals.

# **Dashboards**

To better decide what my company needs out of all the visuals I have made to support the company growth and to deliver the most insights for the company expansion, I completed my four questions individually before developing my KPI dashboard.

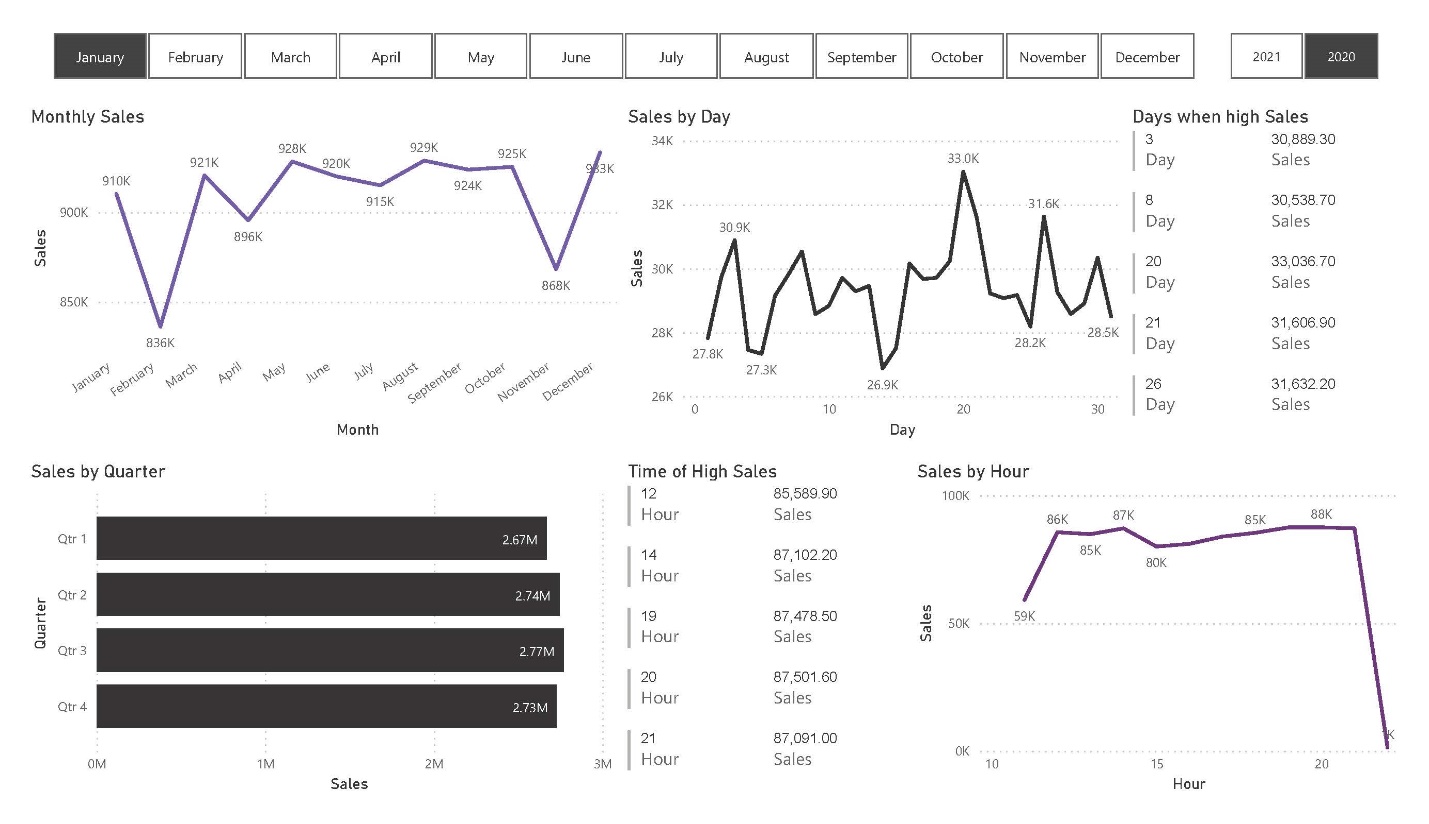
First off, since the dataset simply gave pricing and quantity of each transaction, which cannot be read unless combined to build a sales feature, I have also generated features to make my visualizations more beautiful and understandable. Examples of these features include sales features.

While exploring the 5 datasets, I have found out that Customers dataset only includes Members registered into the company while Transaction dataset include transactions from members and non-members, which will showcase blanks while creating visuals. But I will be showing my steps in making them into feature that can be easily comprehend in the next part.

The default feature customer ID does not provide much insight, so I added a “Membership” feature and divided customer ID into two sections: 0 for Non-Members and any number other than 0s for Members. This is simpler to handle than the raw data.

The hourly feature I generated last is taken from the time features. Hourly allows me to determine precise sales for each day depending on the hours and may determine the number of people who visited the stand on specific days.

## **Question 1**

****

**Question: “Which period has the most Sales?”**

**Rationale:**

The purpose of this questionnaire is to identify if there is any seasonality in sales as this will help to optimise management of inventories, staffing, and timing of appropriate marketing campaigns.

**Address business question:**

To assist me organize the graphics into a particular month and year, I first developed slicers to represent each year and month. It displays monthly sales, responsive to the date slicer choice, starting from the top left corner.

To show the sales for each specific quarter for each year, I have constructed a bar chart. This will enable the business to identify the high-sale seasons. To identify the peak sales for days and hours of each day, I have generated line charts for the hours and days. I have designed a multi-row chart to provide the organization a better understanding of which days and the time of day have the most sales to assist them in identifying peak sales.

## **Question 2**

****

**Question: “Who are the Top 10 High Spending Customer?”**

**Rationale:**

Grouping ‘High Spending’ consumers to identify the most appropriate target for FreshFood products and thus, to efficiently and effectively use marketing activities to engage them and to keep them as FreshFood customers.

**Address business question:**

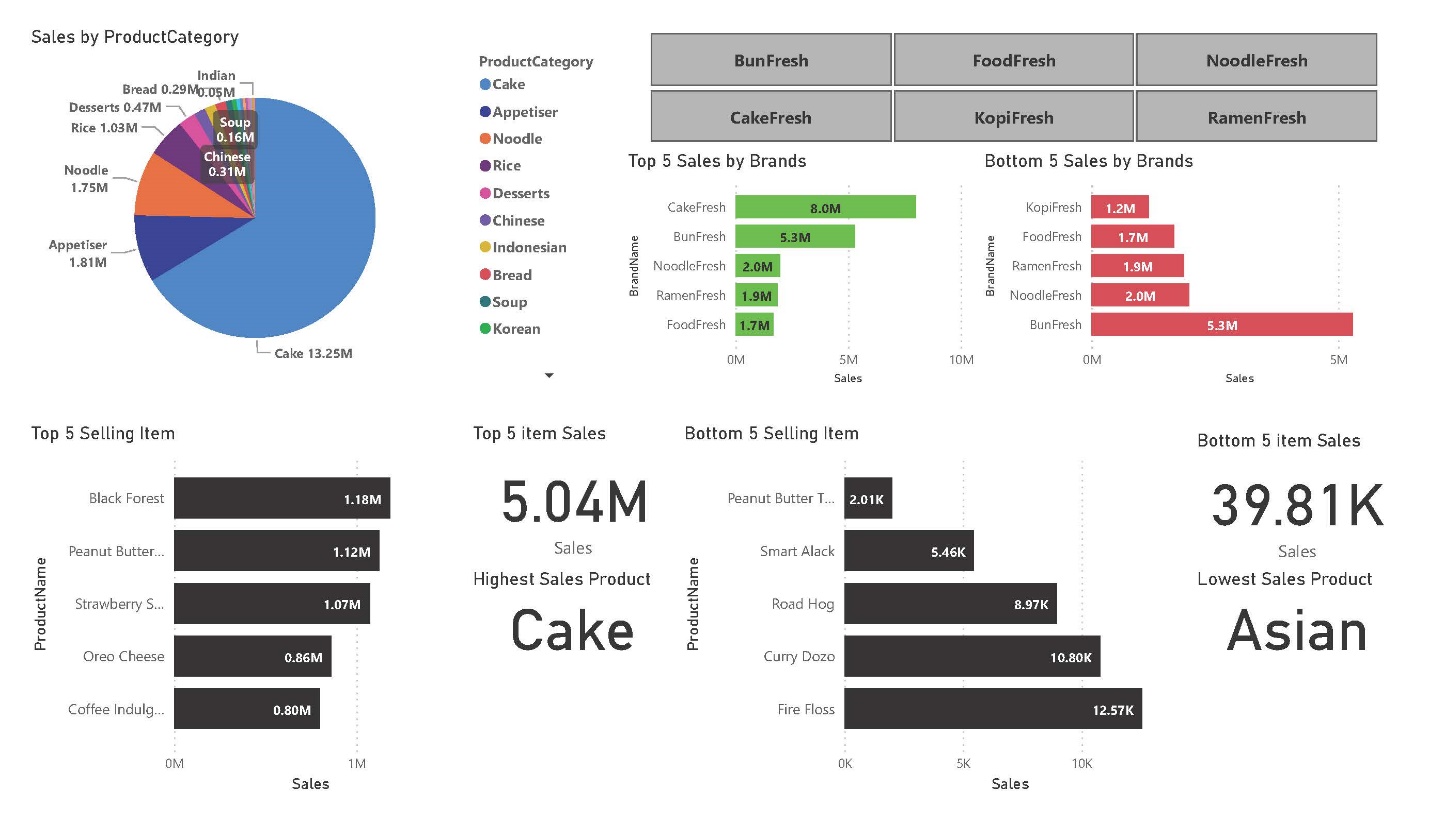
To categorize the dashboard depending on months and years, I constructed a slicer. I made a bar chart to display the top 10 customers, which might change depending on the month, to illustrate the top 10 spending customers in response to the query. This was created with the idea of only applying to customers that are members which is registered into the customer dataset. To learn more about them, I also made tree maps to determine the ages of the top 10 customers, allowing me to identify commonalities and the age range that our items appeal to.

To identify a trend, I also made a line chart from membership sales showing that customers who applied for membership typically spend more than those who did not.

A cluster bar chart showing the sales broken down by age group is displayed after that to determine which age groups our items are most popular with and create a wider selection of those products.

To uncover the general trends in the popularity of different age groups and develop new items that would appeal to these age groups and draw in more customers overall, I have also constructed a scatter plot to display the sales by age.

## **Question 3**

****

**Question: “What are the top & bottom 5 selling items/brands?”**

**Rationale:**

It is generally a good business practice for a company to pay more attention to its high selling products, to make some effort to sustain those low selling products and decide if it is still sustainable/profitable for the business to continue low selling products.

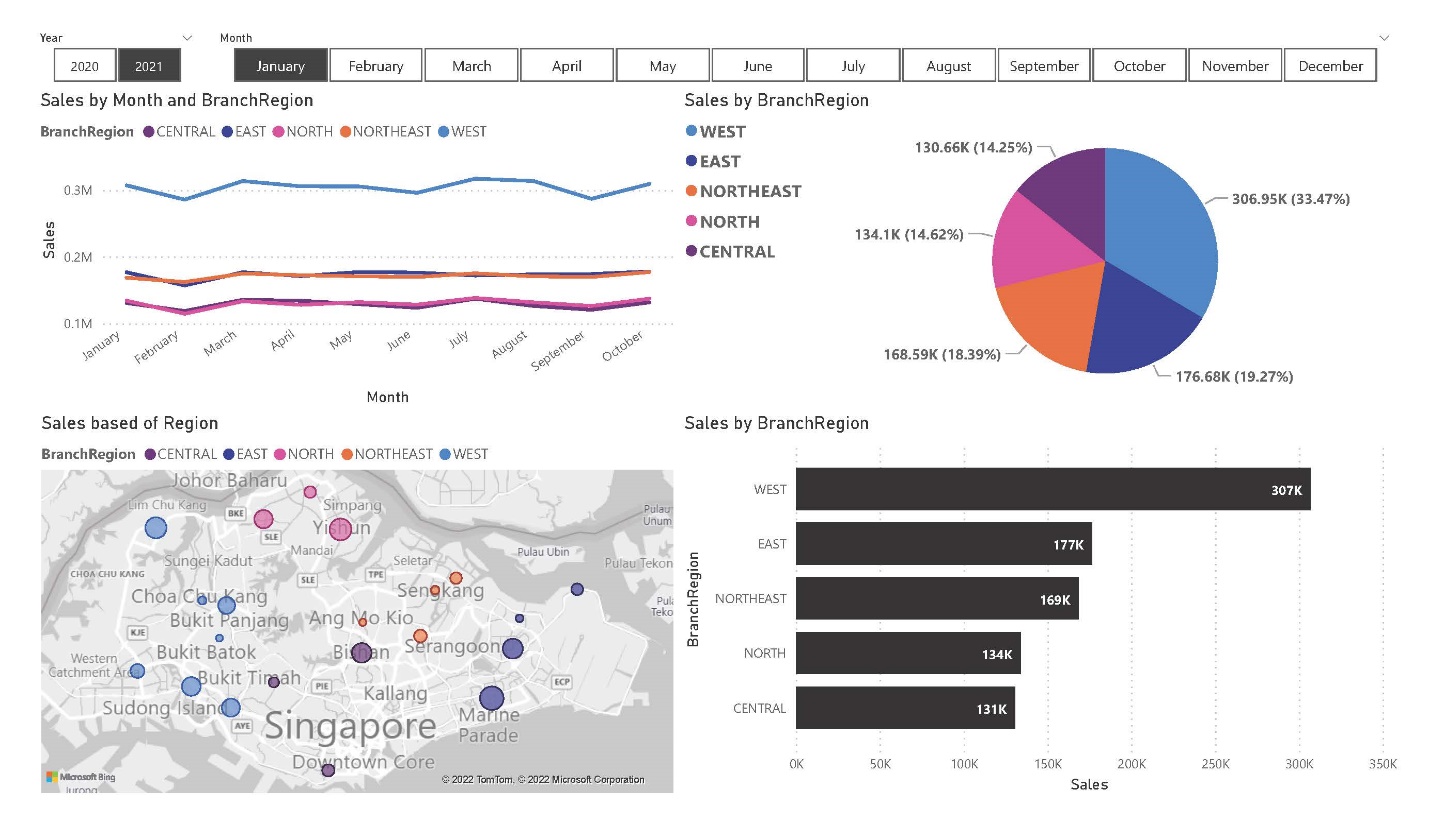
**Address business question:**

For our organization to view sales for each individual brand and determine its top and bottom goods as well as its sales, I designed a slicer to indicate each brand name.

To display the popularity of products and the amount of money they brought in for various brands, I made a pie chart. In addition, we created the top 5 and bottom 5 sales of a brand name to determine which brand is more popular with our consumers.

I have prepared additional bar charts for the company to do in-depth analyses of the top 5 products sold for each distinct brand so that the company may choose whether keeping the lowest-selling products is advantageous or introducing new products depending on the highest-selling products.

## **Question 4**

****

**Question: “Regions that has the most sales?”**

**Rationale:**

It is impossible to overstate the significance of location for business success. A company's location puts it in a good position to draw consumer base to find out popularity in that region. Since there are always areas in a city that have a reputation, whether it be a reputation for modest living or a reputation for rich extravagance, a business's location also aids in the creation of a brand and image.

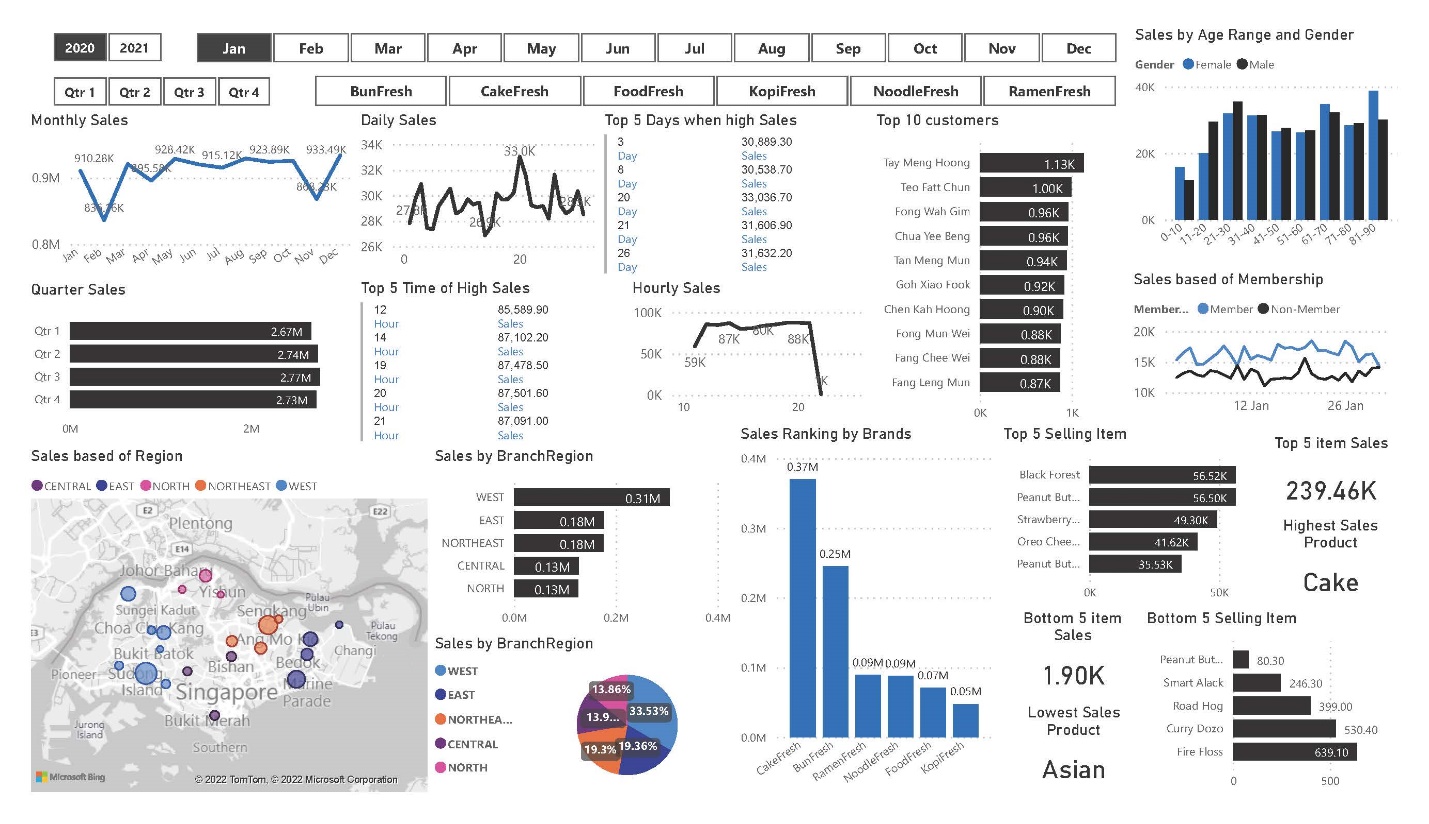
**Address business question:**

I have made two slicers to represent the years and months to respond to the query. Create regional sales by region each month to highlight the peak sales for a given region. For the company to know which region has the biggest sales, I designed a pie chart to show the proportion of sales.

I included a map visual with sales as the size of the bubble, followed by legends to distinguish the region with colours, to learn more about the towns inside the region that produce the sales.

Finally, I produced a bar chart to display the individual sales by region so that the company could determine where to grow its business in 2023. This allowed me to determine the precise values of sales in each region.

## **KPI Dashboard**



Finally, I will combine the results of all four questions into a single Key Performance Indicators dashboard for the organization to observe. I have made several slicers to divide the images into categories based on years, months, quarters, and brand names so that I can browse through my KPI Dashboard.

We will determine the seasonality of sales using visuals that begin at top left and display the monthly sales of each individual year, the quarterly sales of each specific year, then the daily sales of each month chosen from a slicer and the hourly sales of each day. With this, we may change the number of employees for each month to reduce labour costs, and with daily sales, we can identify the days when many customers frequent the store and perhaps run promotions to boost sales on those days. Finally, hourly sales help to control inventory so that there are always goods in stock throughout the day.

Moving forward, we'll be more attentive to our customers by displaying visuals of the top 10 people who have bought things from our stall over the past few years, months, and quarters. Finally, we can compare the difference in sales between members and non-members to have a better gauge of when is the right opportunity to create activities to draw in more potential customer and possibly increase the number of members that subscribe to Food Fresh. The age range of customers that have purchased our item that is classified between males and females so we can create products that can attract new customer in specific age range.

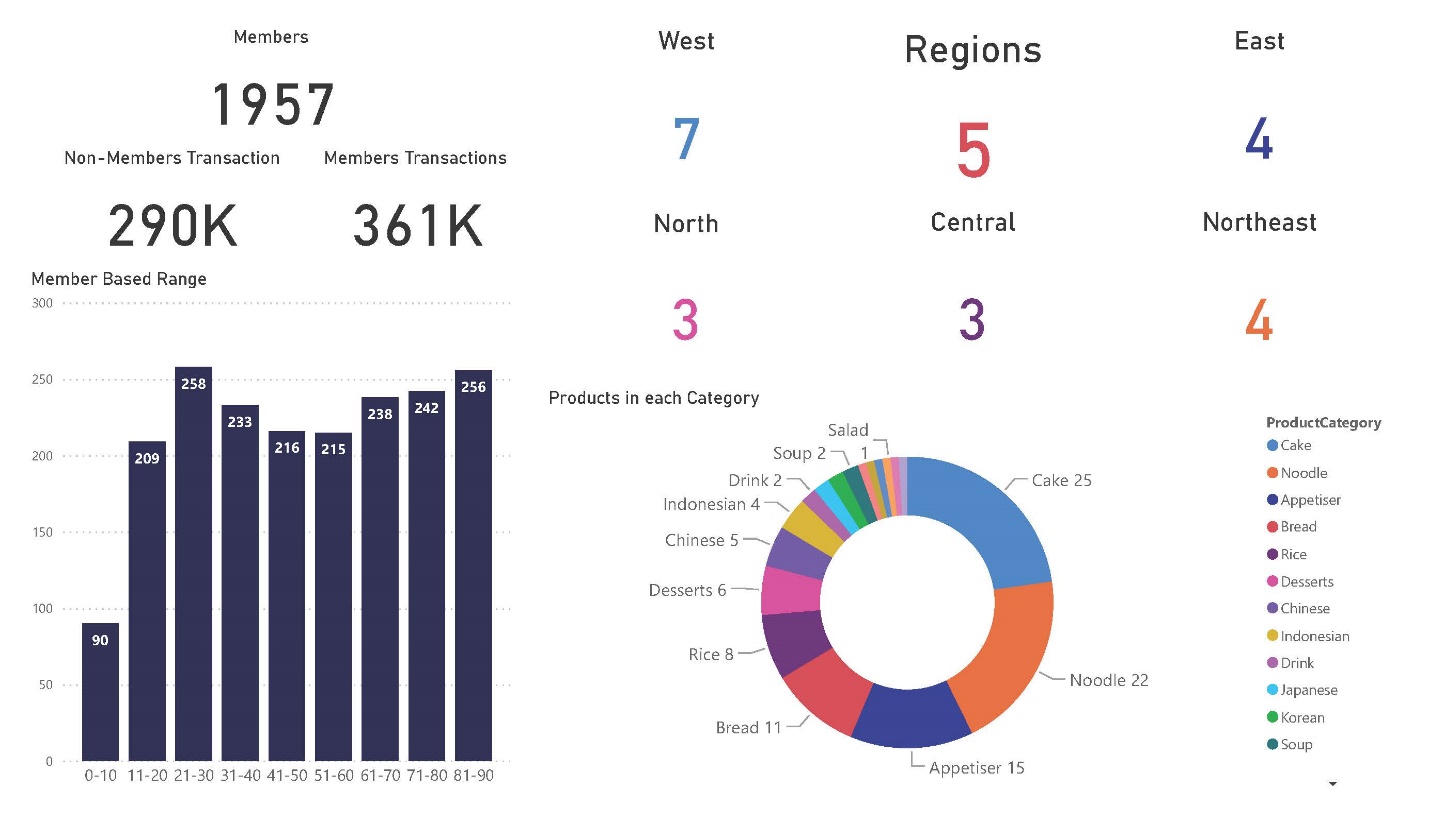
To answer the third question, we ranked each of the company's six brands to determine which was the most well-liked and to generate ideas for potential revisions that could boost sales for the brands at the bottom of the list. While displaying the ranking of products with lowest sales and making a critical determination of whether keeping them as 2023 products would be a wise decision, more in-depth analysis into the specific products ranking that have the highest sales may lead to the creation of new varieties of products to attract new customers.

To answer the final question, we've used maps of Singapore and the West, Northeast, East, North, and Central regions of the country, with the size of the bubbles representing the number of sales in each town. Following the pie chart that displays the specific percentage of the total sales in each region, there is a bar graph that compares the total sales of each region. This might aid the business in deciding whether adding more branches there will significantly increase sales in 2023 or closing branches that would not generate as much revenue as those in other regions to reduce annual costs.

# **Further Analysis**

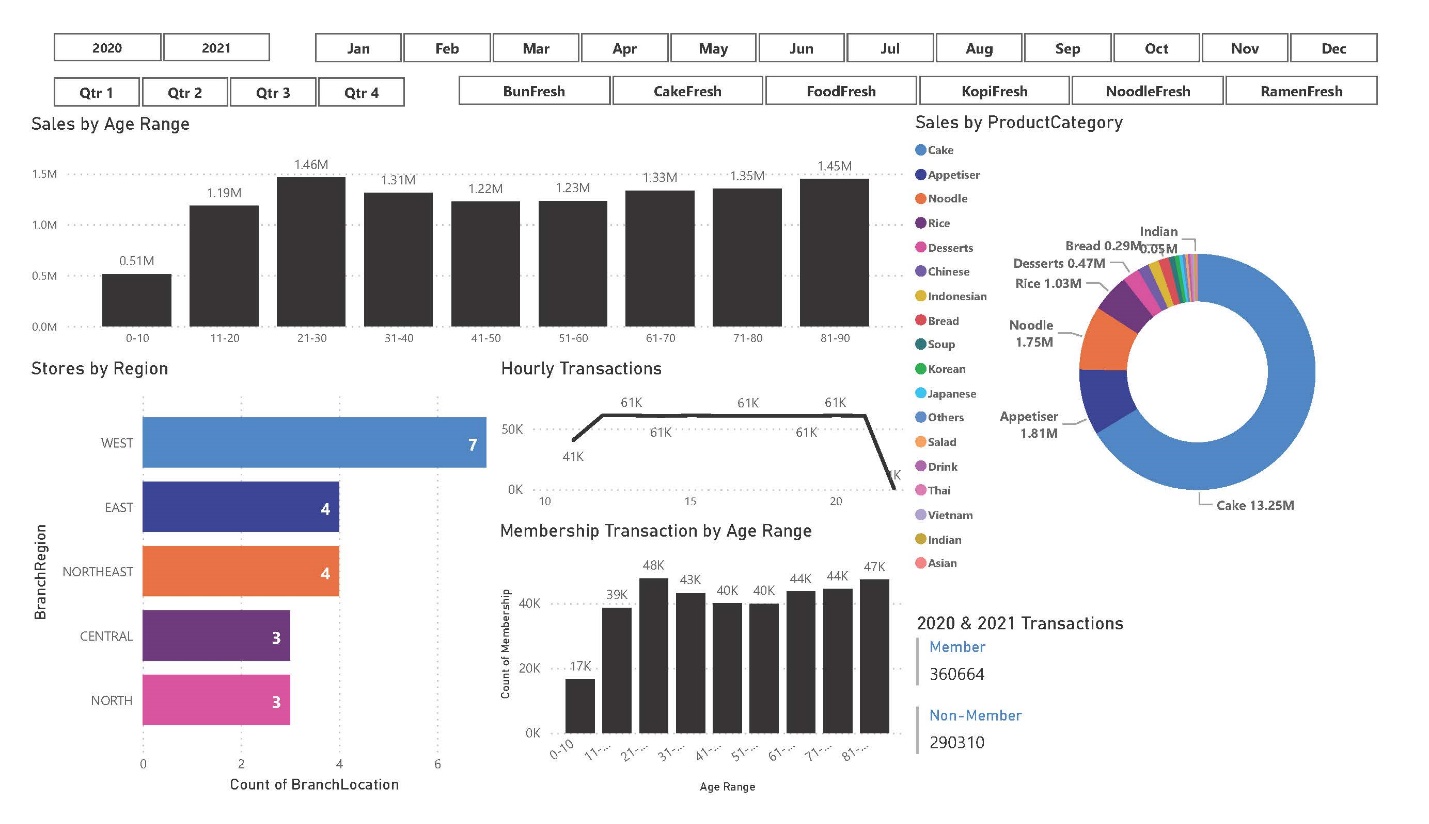
While creating meaningful visuals to help build my KPI Dashboard I have found interesting data which can help company make better decision for the upcoming 2023 sales. The data have been consolidated into dashboard representing univariate, bivariate and multivariate analysis.

## **Univariate**



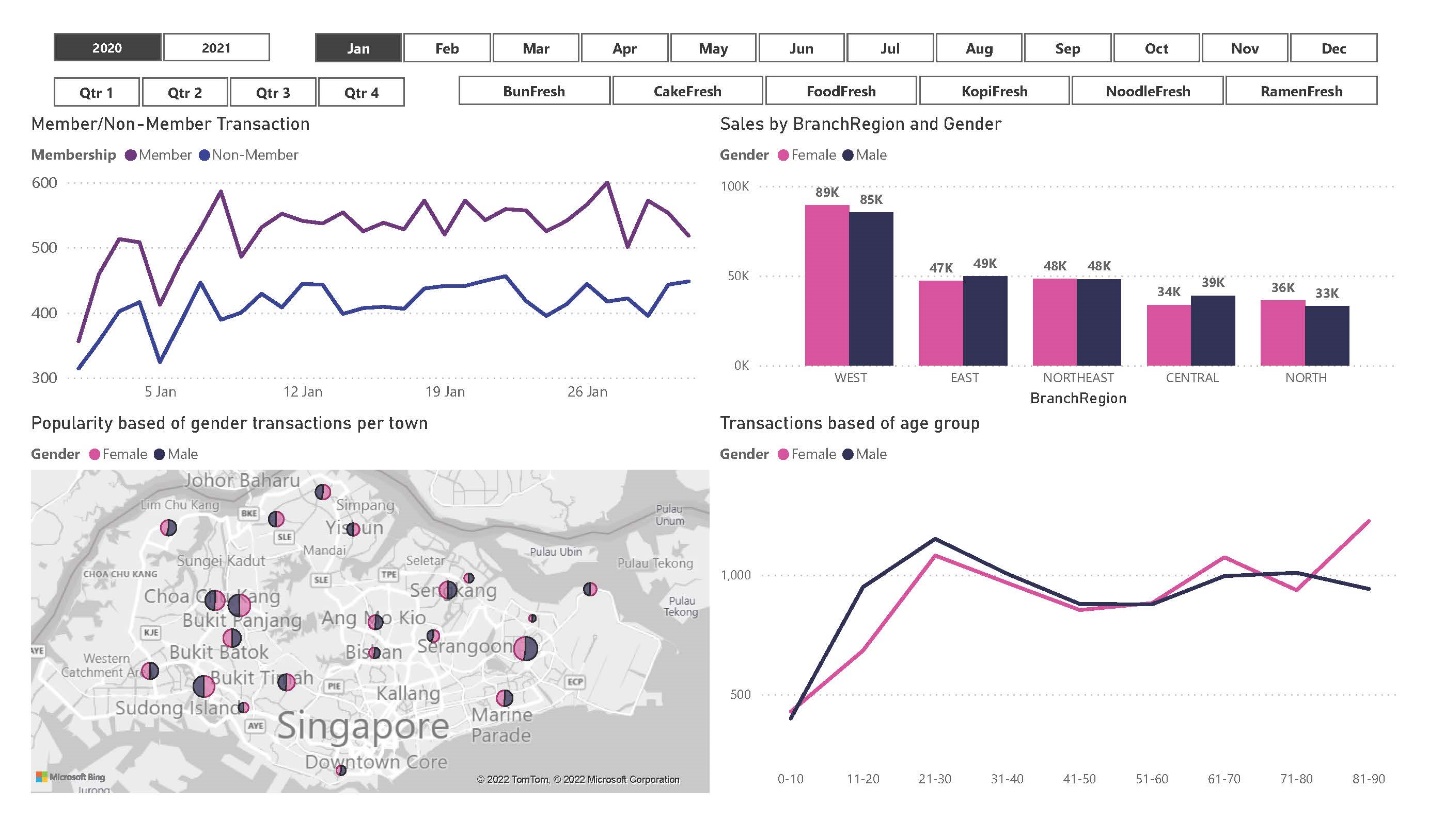
I saw several features that needed much more in-depth research while trying to understand the significance behind each feature, therefore I combined them into a single dashboard that is classified as a univariate analysis. I've come to the realization that the customer database provided to me only contains customers who are members, whereas the data provided for the transactions database only includes non-members. As a result, I've discovered that there are 1957 members overall, of whom 361K transactions came from them and the remaining 290K from non-member transactions. Additionally, I've classified them before counting the number of stores in order to learn more about how many branches have been open in each region. I created a bar chart to determine the number of members in each group in order to satisfy them with the appropriate promotion. By knowing more about our Members' age range, we can keep their support for the following year 2023. Last but not least, in order to avoid overcrowding our menus with dishes while also introducing new dishes that are well-liked by our clients based on data from a donut chart I made, I have calculated the number of products created since 2020.

## **Bivariate**

****

The dashboard above was created with bivariate analysis in mind to help the organization choose the following objective for 2023. When I was trying to determine the age range that brings in the most consumers, I discovered that adults between the ages of 21 and 30 and senior citizens between the ages of 81 and 90 make up most of our customers. With this information in mind, we can create healthy meals for adults because they are frequently busy at work during the day and may require a boost during their lunch break. This is demonstrated in the line graph I created above, which shows the increase in sales transactions from hour 11 to hour 12, indicating that customers normally arrive at this hour to have lunch. Since cakes accounted for the bulk of sales, we can assume that cakes will be a much more viable alternative for us to draw more potential customers from that age bracket. As for the elderly, we can make products that are simpler for them to bite down so they can enjoy their tea breaks. After that, I made a new chart to show how many members there are in each age group. As you can see, most members are between the ages of 21 and 30 and 81 to 90. Members have made up 360,664 of the total sales transactions during the last two years, while the remaining 290,310 clients have not registered to become members. Accordingly, 650,974 transactions were made in total during the two years, with sales by non-members accounting for 50% of those transactions. We can certainly organize activities to increase the number of members in 2023 to increase revenue for our company.

## **Multivariate**

****

Through multivariate analysis, I can thoroughly examine the transactions done by members and non-members throughout the course of the two years of company and identify moments when a significant portion of members buy our items and times when a significant portion of unregistered consumers buy them. With this information, the business can decide whether to create promotions to entice members to purchase more goods and customers to apply for memberships by offering them free vouchers. Following that, by categorizing regions into female and male sales, I can infer that most customers came from the west side, with a higher female consumer than a male consumer. With a map visual to help prove my point, I show the size of each town's customer base classified as female or male. Finally, using the multi-line chart, I can see that most transactions come from people between the ages of 21 and 30, as well as those between the ages of 61 and 70, which will help the company make better decisions when dealing with specific age groups.

# **Summary**

To summarize my entire report, my dashboard will be able to help company make better decisions when deciding their road ahead for 2023. Such as seasonality in sales as this will help to optimise management of inventories, staffing, and timing of appropriate marketing campaigns. Follow by identifying the most appropriate target for FreshFood products and thus, to efficiently and effectively use marketing activities to engage them and to keep them as FreshFood customers. In addition to paying more attention to its high selling products, to make some effort to sustain those low selling products and decide if it is still sustainable/profitable for the business to continue low selling products. Finally, able to draw consumer base to find out popularity in that region. Since there are always areas in a city that have a reputation, whether it be a reputation for modest living or a reputation for rich extravagance, a business's location also aids in the creation of a brand and image.

With further analysis, we have more knowledge about the data we are dealing with, which can help company make better decision for the upcoming 2023 sales. Finding out the amount of members I have throughout the 2 years of operating a business, products that keeps our sales going, how age groups affects our sales and by understanding more about the customers we are selling our products with, we will be able to produce much higher sales depending on the methods we used during the specific time of the year or the regions our customers stayed in.