

TABLE OF CONTENTS

OVERVIEW	2
Project Background & Motivation	2
Project Objectives	3
DATA INTEGRATION AND FILTERING	3
Data Collection.....	3
Data Summary	3
Delivery Data	3
Promotion Data.....	4
Outlets Data	5
Challenges.....	5
DATA CLEANING	6
1. Redundant Columns – Delivery Data	6
2. Data Noises.....	6
DATA PREPARATION	7
1. New Columns Created.....	7
2. Outliers identification for Reorder Frequency	8
3. Discontinued Ingredients	9
EXPLORATORY DATA ANALYSIS	10
REVISED SCOPE OF WORK	19
REVISED METHODOLOGY	21
Cluster Analysis	21
Monte Carlo Model	21
REVISED WORK PLAN	22
BIBILOGRAPHY	23

OVERVIEW

Project Background & Motivation

KOI, one of the most well-known bubble tea brands in Singapore was founded in 2006. It was first launched in 2 locations around Taiwan’s Taichung area with focus on product development and innovation. Their first overseas store was subsequently opened in Singapore in 2007. This was part of the effort for international brand expansion. Upon opening their first outlet in 2007, it became the hype among the younger generation and attracted a long queue during that period. Till date, there are a total of 46 outlets widespread in Singapore, 34 outlets in China, and presence in Taiwan, Macau, Hong Kong, Japan, Indonesia, Vietnam, Cambodia, Thailand, Myanmar and Malaysia.

With the increasing number of outlets in Singapore coupled with the extensive menu of drinks available at KOI, they are facing logistics issues in their daily operation. Employees from each of its outlet are required to manually order their required stock for next-day delivery. As such, Koi are facing problems such as their employees might under or over order for the next day.

To decide on the amount of required stock for the next-day delivery, employees are required to estimate the next day sales at the same time consider the shop storage limits. Additionally, the employees should be aware of any special promotions that will be held or any events that is happening around their outlet as it might affect their next-day sales. Overall, it has led to a steep learning curve for the new employees.

Lastly, KOI often hold promotional events that help to boost their sales revenue and brand reputation. Currently, there are significant difficulty in determining the effectiveness of such promotional campaigns. Furthermore, it is important for KOI to hold such promotional campaigns at the right time to maximize their sales profit.

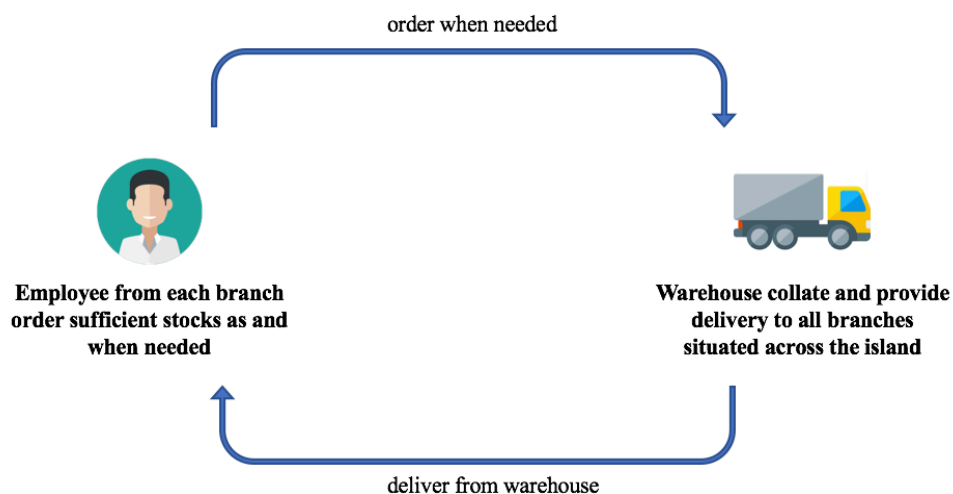


Figure 1: KOI’s re-order delivery system

Project Objectives

By utilizing the data from their current database, we wish to discover meaningful and informative insights which will allow KOI employees to make better decision in doing stock ordering, and thus remove the steep learning curves for their new employees. Further on, we wish to explore different factors that could impact KOI sales (i.e. demographics of people living in the area, schools around the area). To achieve the above mentioned, we will perform an in-depth analysis on the data collected with the following objectives in mind:

1. Reorder Optimal Stock Quantity

- To allocate the optimal restock amount. How much should employee order to ensure that they will not run out of the ingredients needed?
- To analyze the time series trend. How much should the employee order in this current period?

DATA INTEGRATION AND FILTERING

Data Collection

To facilitate our analysis, KOI has kindly provided our team with a data of 47 outlets with 2 years (from Jan 2016 - Dec 2017) worth of data collected. The client wish to focus on the latest business fiscal year, henceforth we will be provided with the latest data obtained. The type of data obtained are summarize in the table below.

Data Summary

File Name	Description
Delivery Data	Information regarding the delivery ordered per branch over a period of two years
Promotion Data	Information regarding the promotional campaigns proposed by KOI over the period of two years.
Outlets Data	Information regarding 47 different branches of KOI

Delivery Data

A row in this table represents a specific delivery ordered by a branch. The detailed description of the main columns in this table are as follow:

Column name	Description
Delivery Date	The date where the order is being delivered to the store.
Created Date	The data where the order is being placed by the staff.
Time	The time where the order is being placed by the staff.
Order Num	The unique id of the order.
Type	The different type of order being placed.
Store	The store outlet name and code that placed the order.
Product Num	The unique id of the product being ordered.
Product	The name of the product being ordered.
UOM	The unit of measures of the delivery order type. For instance: bottles, box, bag and roll.
Qty	The amount of product being placed.

New Columns Created	Description
Created Date	Date format for the date that the order is being placed
Date Diff	The date difference between this current order and the next created order
Actual Date Diff	The actual date difference between this current order and the next created order in an outlet
Product Category	The type of item delivered including bubble tea ingredients, non-ingredients and miscellaneous
Include?	To identify products with time series trends and thus to be included in our model building

Promotion Data

A row in this table represent a specific promotional campaign held in a period. The detailed description of the main columns in this table are as follow:

Column name	Description
-------------	-------------

Promotion Period	The timeframe of promotions within the year
Main Menu Board	The product which is promoted via the stated column channel (Main Menu Board)
Table Menu	The product which is promoted via the stated column channel (Table Menu)
A4	The product which is promoted via the stated column channel (A4)

Outlets Data

A row in this table represents the KOI branch outlet information. The detailed description of the main columns in this table are as follow:

Column name	Description
Code	The unique code of the branches town. For instance, Ang Mo Kio (AMK), Paya Lebar (PYL).
Code2	The unique code of the branches in the town.
Outlets	The specific location of the branches.
Address	The address information of the branches.

Challenges

As a growing SME in Singapore without relevant IT expertise, KOI's database design was outsourced to an external IT firm. However, it has been observed that the database design is not constructed in an efficient manner. There were several columns in the dataset that are considered redundant, and in the delivery order, there is no significant segregation between bubble tea ingredients, non-ingredients such as cups, straw and seal as well as other miscellaneous such as coins and cleaning supplies.

Furthermore, our team have identified couple of ingredients that have been discontinued from their business, however the ingredient still exists in the dataset. Hence, to overcome this, we have implemented the following measures:

- Upon the conclusion of our EDA, for those ingredients which do not have significant time series across all outlets, we have created a new column called "Include?" to indicate whether that row of data should be included in our cleaned dataset.

- There is no specific product category for all the ingredients given. Examples include: coins, cleaning supplies, bubble tea ingredients such as tapioca pearl and coffee bean, non-ingredients such as cups, straws, plastic bags and seal. Henceforth, we have created a column named “Product Category” to group similar products together.

DATA CLEANING

Before exploring the data, our team had taken a brief look at the data provided. We observed that there are some noises in the data, which we have proceeded to clean it prior to our exploratory data analysis. Listed below are the steps of data cleaning undertaken:

1. Redundant Columns – Delivery Data

Our group have identified number of redundant columns which we deemed not important to our analysis, and have subsequently removed it. The highlighted table below illustrate the columns removed:

Delivery Data:

Column name	Description
Time	The time where the order is being placed by the staff.
Type	The different type of order being placed.

2. Data Noises

We have identified several overseas branches order in our dataset, and outlets that have been closed. We deemed these as noises and insignificant to our analysis and hence removed it from our dataset. Additionally, as our data analysis are only focusing on two years’ analysis – 2016 and 2017. We deemed 2015 and 2018 data as noises and decided to remove them. The list below identifies the outlets that we have removed from our dataset:

Overseas Branches	<ul style="list-style-type: none"> • KOI TE (Thailand) • Cambodia Karanak KOI Café • Jakarta KOI • KOI Myanmar • Vietnam – Cong TY TNHH KOI NHANH
--------------------------	--

Closed Outlets

- Esplanade (May 2017)

DATA PREPARATION

Moving on to data preparation, we have undertaken the following steps to prepare our data for the Exploratory Data Analysis.

1. New Columns Created

As identified previously in data summary, we have identified several new columns created to aid in our analysis. Following which, we will explain the underlying reasons behind our columns creation in the table below:

Delivery Data:

New Columns Created	Reason for creating the column
Created Date	Reason for creating a new column created date was due to the date format that KOI provided. The date created provided include date time format which we do not need.
Date Diff	Data difference illustrates what is the difference in number of days for this outlet current order against the previous order. This allow us to understand how frequent an outlet does their reordering process also known as “reorder frequency”
Actual Date Diff	Actual date difference helps to standardize the difference in number of days for this outlet current order against the previous order across for all products in a single order (i.e. date diff column above will only reflect the first row of the order, whereas actual date diff will bring down the same date diff reorder days for all products in the same order)
Product Category	Reason for creating a Product Category column was because there was no significant segregation for the items. Henceforth, we have created 3 product categories to segregate the products. Ingredients refer to bubble tea ingredients, non-ingredients refer to non-bubble tea ingredients but are crucial to business operation and lastly miscellaneous for others. For illustration

	<p>purpose, examples on how we segregate the products are as follow:</p> <ul style="list-style-type: none"> • Ingredients → Pearl, Milk Powder, Jasmine Tea etc. • Non-Ingredients → Cups, Straw, Cup Seal, Plastic Bags etc. • Miscellaneous → Coins, Cleaning Supplies, Uniform etc.
Include?	<p>Reason for creating the include column was for our future model building. As there are some products identified with non-significant time-series trends, include will allow us to segregate products that have time-series trends and products without.</p> <p>Those with time-series trend, it will be marked “Yes”. Those without will be marked with “No”</p>

Outlet Info:

New Columns Created	Reason for creating the column
Segment	Reason for creating a new column “Segment” was because we wish to split the outlets with similar reorder frequency into groups to compare and make necessary business recommendations to our sponsor. We have split the outlets into 3 segments – Segment 1 for Reordering Frequency of ≤ 5 days, Segment 2 for Reordering Frequency of ≤ 9 days, and Segment 3 for Reordering Frequency with outliers of ≥ 10 days.
Region	Reason for creating a new column “Region” was to split our outlets into different constituency region in Singapore to enable us to provide better business recommendation regarding delivery and reorder frequency to our sponsor

2. Outliers identification for Reorder Frequency

Whilst undergoing through our preliminary exploratory data analysis, we identified outliers for reorder frequency. It is as illustrated in the chart below:

Store with ≥ 7 order date different

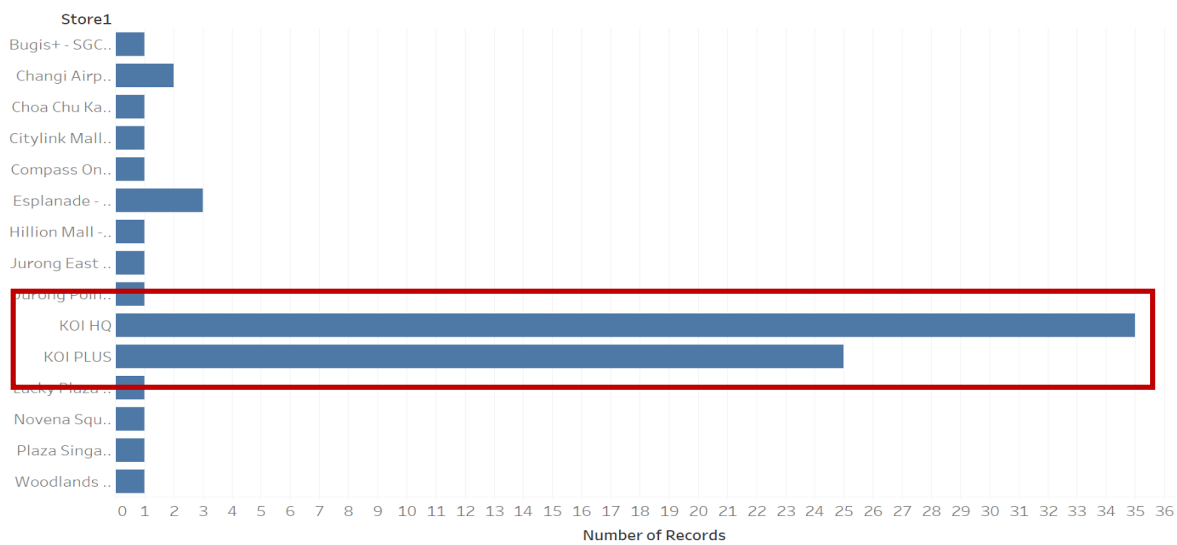


Figure 1: Outliers for Reorder Frequency

From Figure 1 above, we could clearly identify KOI HQ and KOI PLUS as the outlier for reorder frequency. Our team have done a deeper dive, and decided to remove these outlets from our analysis. For KOI HQ as it is not a customer facing branch, we have deemed it insignificant for our analysis. Secondly, as consulted with our sponsor, we have confirmed that KOI PLUS relies on KOI Singapura for its goods ordering. Furthermore, our results also show that the restock products of KOI Plus are mainly non-perishable goods such as cleaning supplies, cups and seal. Henceforth, the reorder frequency for KOI PLUS tends to stretch for a long period before their next order.

3. Discontinued Ingredients

Whilst undergoing through our preliminary exploratory data analysis, we have also identified several ingredients which have been discontinued from use. The list below illustrate ingredients that have been discontinued, and thus we have removed it from our dataset as it is considered insignificant for our analysis.

Discontinued Ingredients:

- Mango Syrup
- Mango Sticker
- Passion Fruit Syrup
- Passion Fruit Sticker
- Coffee Mate 1KG

EXPLORATORY DATA ANALYSIS

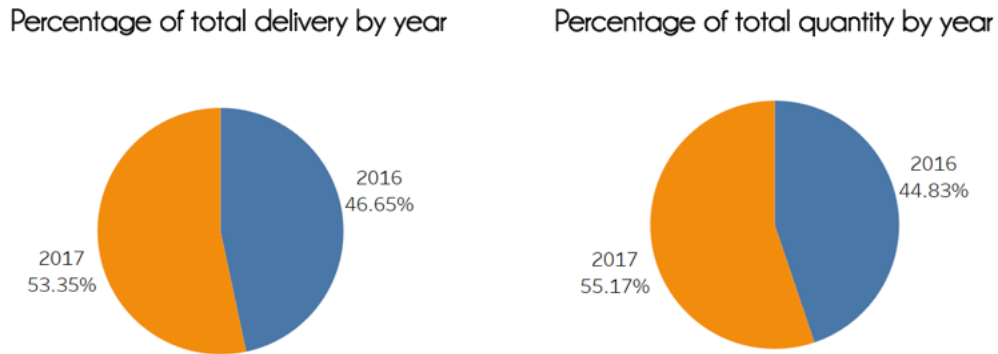


Figure 2: 2016/2017 Delivery Data Analysis

As shown in Figure 2, we can identify that the percentage of total delivery in 2016 stood at 46.65% and 53.35% in 2017. In comparison to the percentage of total quantity ordered, 2016 stood at 44.83% and 2017 stood at 55.17% increment. Upon further investigation in our EDA, we realized that it is mainly due to the new outlets open in 2016 and 2017, which made 2016 data collectively lesser than 2017 dataset. There are 4 new outlets opened in 2016, namely Woodlands, Seng Kang and The Arcade. Woodlands outlet opened in May 2016, Seng Kang in September 2016, Arcade in November 2016 and Jurong Point in December 2016. These 4 outlets do not provide full year dataset for the year 2016. Furthermore in 2017, there are 7 new KOI outlets opening across the island. They are Harbourfront, Jurong East, Raffles City, Hillion Mall, CCK, Novena Square and Tampines Mall. Henceforth, these months of data contribute to the increment in 2017 data.

We have further highlighted the contributions of new outlets by delivery and quantity in Figure 3 below:

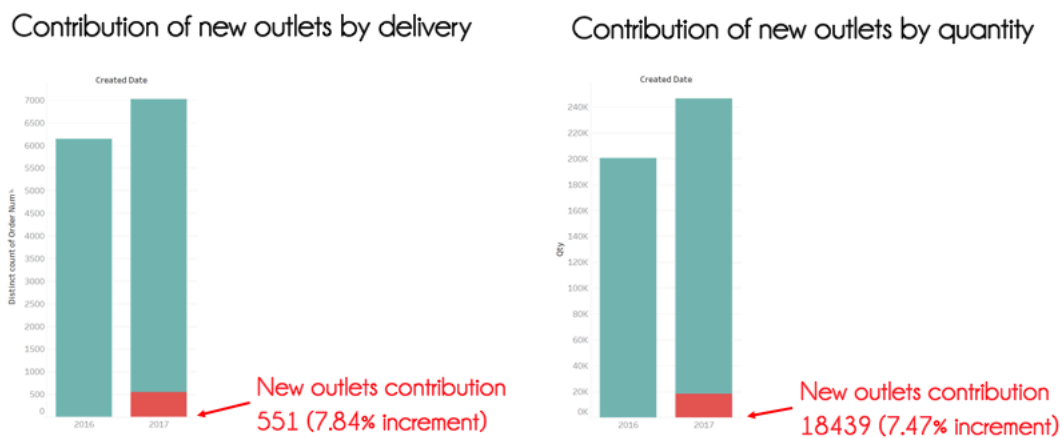


Figure 3: Contributions of new outlets by delivery and quantity

As shown in Figure 3, new outlets contributed to 551 (7.84%) increment of delivery counts and 18,439 (7.47%) of quantity counts. While the remaining contribution can be broken down into contributions by various existing outlets as shown in Figure 4 below.

Contribution of existing outlets by delivery

Contribution of existing outlets by quantity

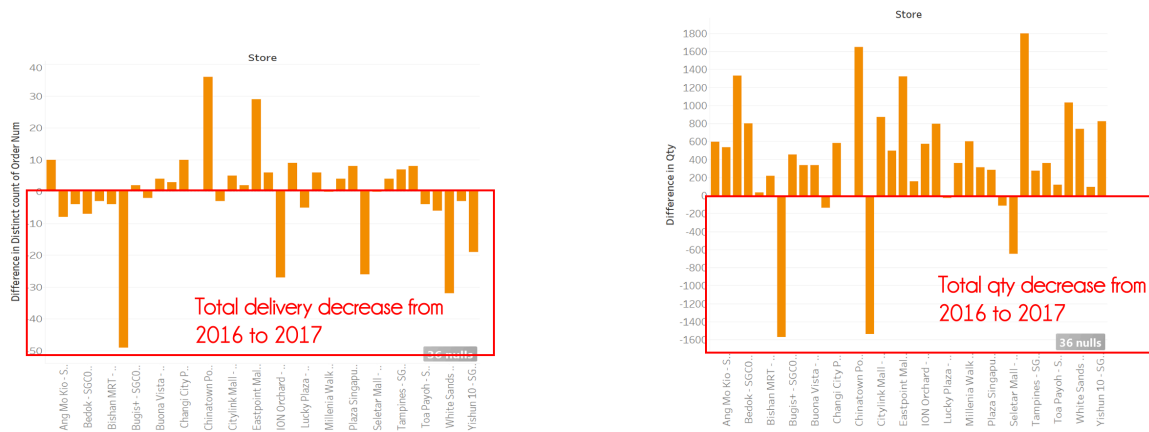


Figure 4: Contributions of existing outlets breakdown by individual outlets

As shown in Figure 4, we can identify that there are several outlets that had shown total delivery decrement from 2016 to 2017 (i.e. 2017 do less deliveries than 2016). Examples of outlets decrease from 2017 as compared to 2016 are Ang Mo Kio, Bishan and ION Orchard. However, in comparison to the chart on the right regarding contribution of new outlets by quantity, it displayed otherwise. Henceforth we can conclude that although the number of delivery decrease from 2017 as compared to 2016, it does not justify that the order decrease. With that we have decided to focus on the chart on the right which shows contribution of new outlets by quantity.

As such, we have decided to focus on the two outlets with significant decrement recorded. Figure 5 below illustrate the trend analysis on the quantity ordered for Boon Lay MRT outlet and City Square Mall outlet.

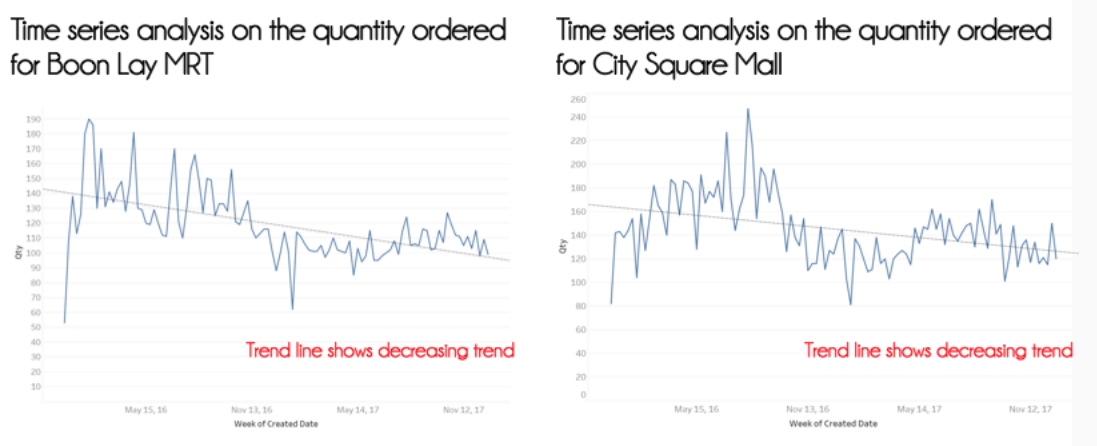


Figure 5: Time series analysis on the quantity ordered for outlets

From Figure 5 above, we could identify that the time series trend for the two outlets which shown significant quantity decrement. Henceforth, we have noted and will highlight to our sponsors regarding the outlets business performance for their consideration.

Moving on to our next analysis, we have constructed Figure 6 below to highlight the top 15 outlets on delivery ordered and quantity ordered.

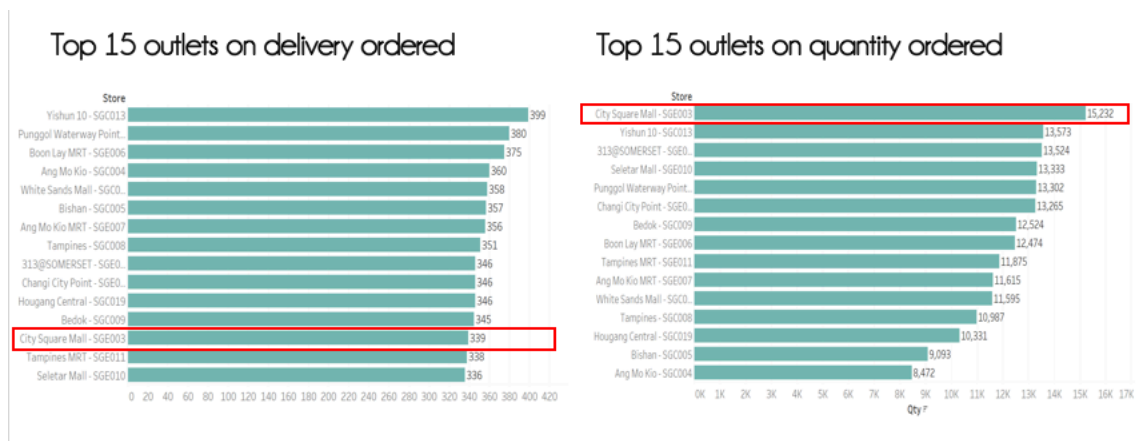


Figure 6: Top 15 outlets on delivery ordered and quantity ordered

From Figure 6 above, we have identified “City Square Mall” as the abnormally with low number of deliveries recorded (rank 13) but high quantity ordered (rank 1). We have drilled down to understand the underlying reason behind it

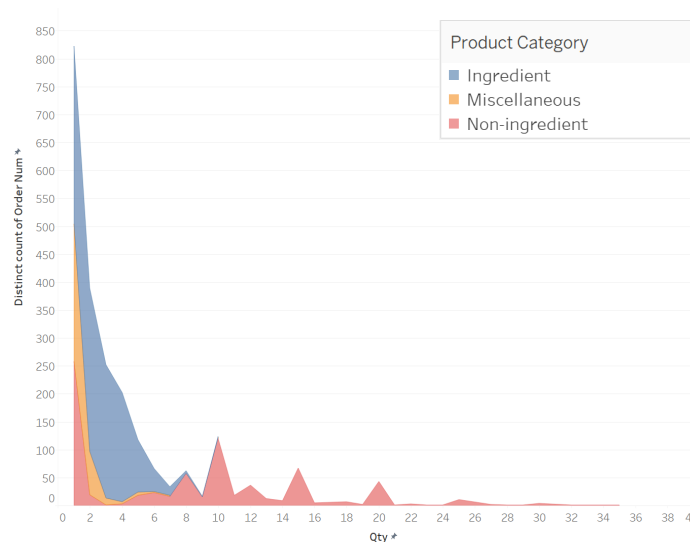
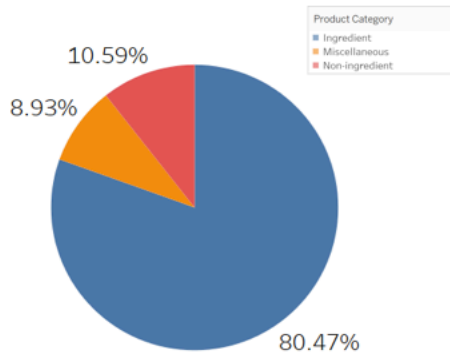


Figure 7: City Square Mall Ordering Trends filter by Product Category

From Figure 7 above, we can see that “Non-Ingredients” (red area) makes up the bulk of high quantity order, whereas Miscellaneous and Ingredients trends relatively normal. We want to understand the underlying reason behind it, henceforth drilled down to make comparison against the average percentage across all outlets as shown in Figure 8 below:

Composition of Average Quantity Ordered by Product Category



Composition of Quantity Ordered by Product Category for City Square Mall

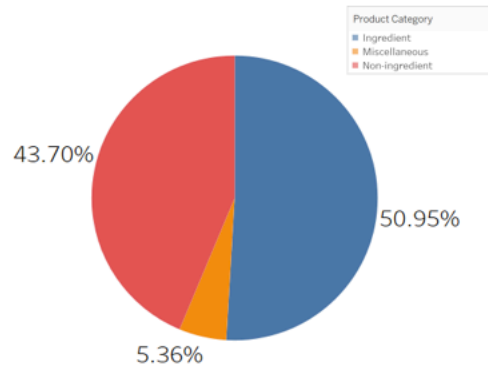
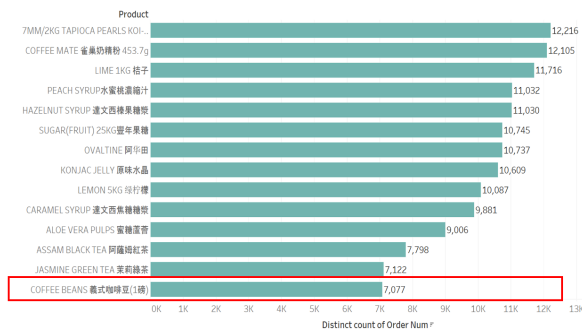


Figure 8: Comparison of Product Category Composition of Average and City Square Mall

From Figure 8 above, we can see that on average (chart on the right), Ingredients makes up to be about 80.47% of the quantity ordered, Non-Ingredient makes up about 10.59% and Miscellaneous makes up about 8.93% across all outlets. Whereas in comparison in City Square Mall, Ingredients makes up to be about 50.95% and Non-Ingredients makes up to be about 43.70% of the quantity ordered. There is a big contrast between City Square Mall and the average amount ordered, thus a point to note and highlight to our sponsor.

Moving on to our next analysis, we have constructed Figure 9 below to identify top products on delivery and quantity ordered.

Top 15 Products on Delivery Ordered



Top 15 Products on Quantity Ordered

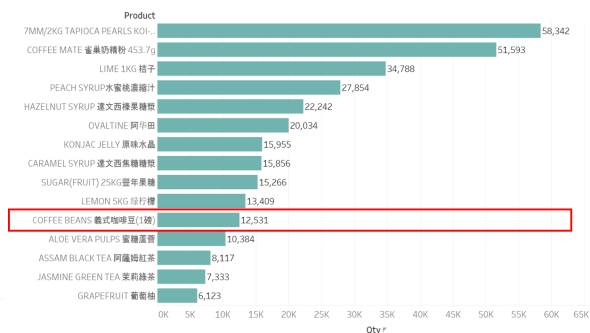


Figure 9: Top 15 Products on Delivery and Quantity Ordered

From Figure 9 above, we can identify the top 15 products on delivery and quantity ordered. For instance, Coffee Beans had 7077 delivery orders across all outlets and 12,531 quantity ordered counts. We want to identify what are the ratio between the delivery and quantity (i.e. how many quantity ordered per delivery). Henceforth we have created Figure 10 below to illustrate the ratios:

Top Products Ordered by Ratio

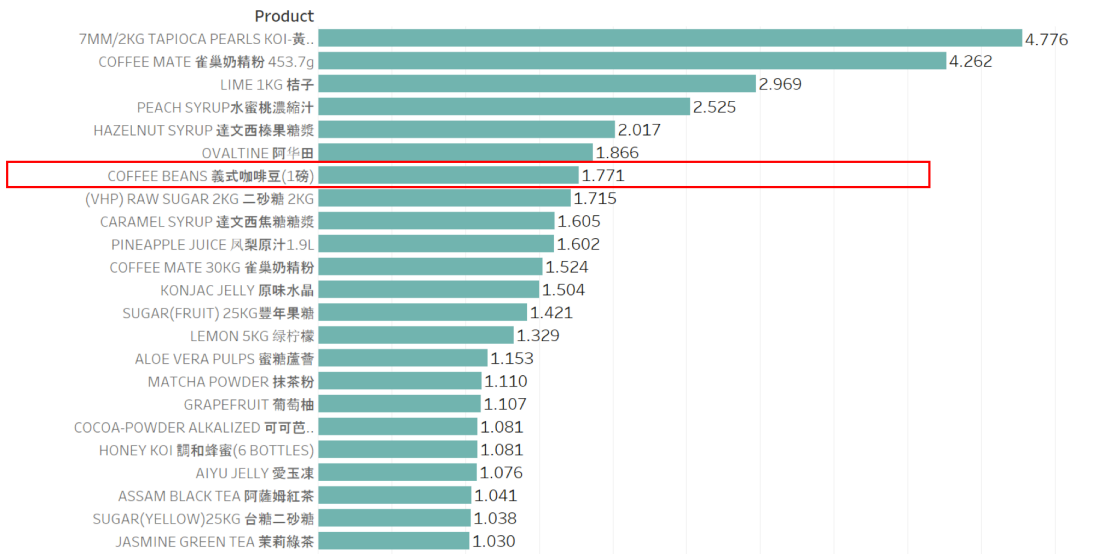


Figure 10: Top 15 Products Ordered by Ratio

From Figure 10 above, we can see that coffee beans stood at 1.771 per delivery. Whereas the top spot remains for Tapioca Pearls at 4.776. This means that for every delivery order, there is an average of 1.771 coffee beans and 4.776 of Tapioca Pearls ordered. From this ratio, it'll allow us to gauge what are the average amount per delivery that we can propose in our prediction in the later part of this project.

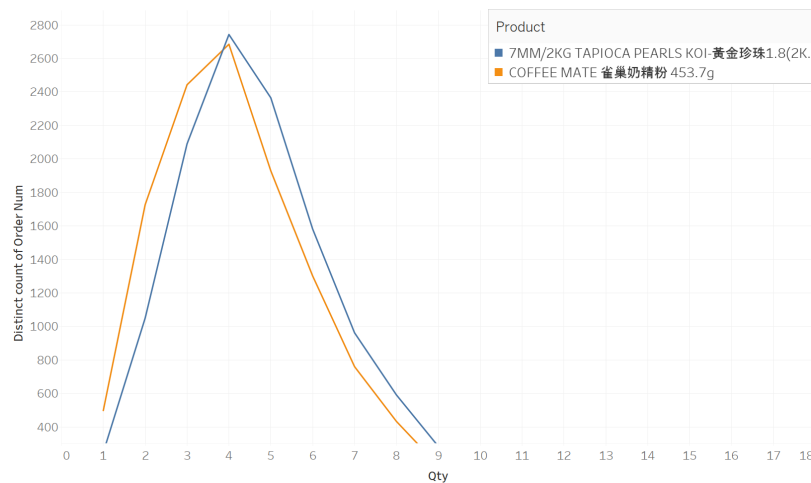


Figure 11: Number of Deliveries for Each Quantity Ordered

Furthermore, from Figure 10 above, we identify that the top 2 products look similar in terms of deliveries, quantity ordered as well as ratio at 4.776 for Tapioca Pearls and 4.262 for Coffee Mate. Henceforth, we have drilled down to look in-depth into their order patterns. As shown in Figure 11 and discover that they have similar order patterns. Following which, we have drilldown into Figure 12 below on a weekly basis.

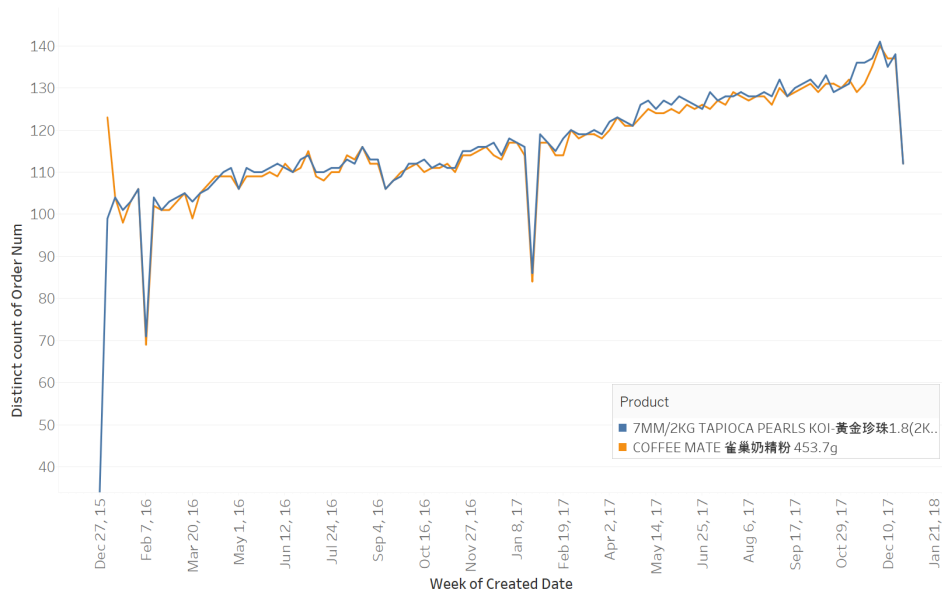


Figure 12: Time Series Trend for the Number of Deliveries per week

From Figure 12 above, we can see that both Tapioca Pearls and Coffee Mate follow a very close pattern on a weekly basis. Henceforth this is a point to note for our future model development as well as business recommendations for our sponsor on the market basket analysis products for future reorder purpose.

Moving on to our next analysis, we are exploring into Reordering Frequency. We have implemented a new column called “Segment” to compare the reordering frequency patterns between outlets and make comparison between various outlets in the same segment.

Segment 1:

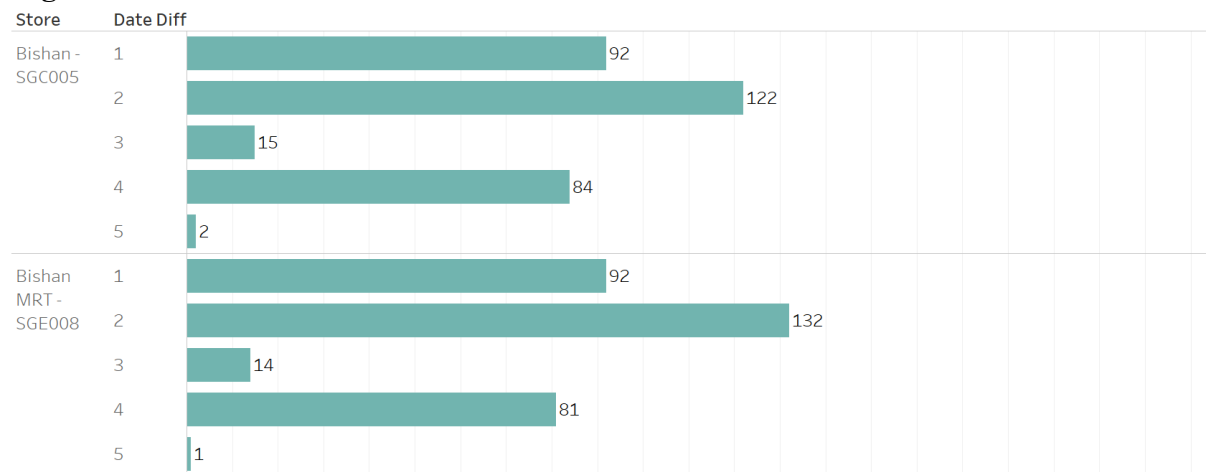


Figure 13: Segment 1 Outlets Reordering Frequency

Since the above examples in the Segment 1 have similar reordering frequency, this enable us to make further evaluation and business recommendations to our sponsor about delivery.

Segment 2:

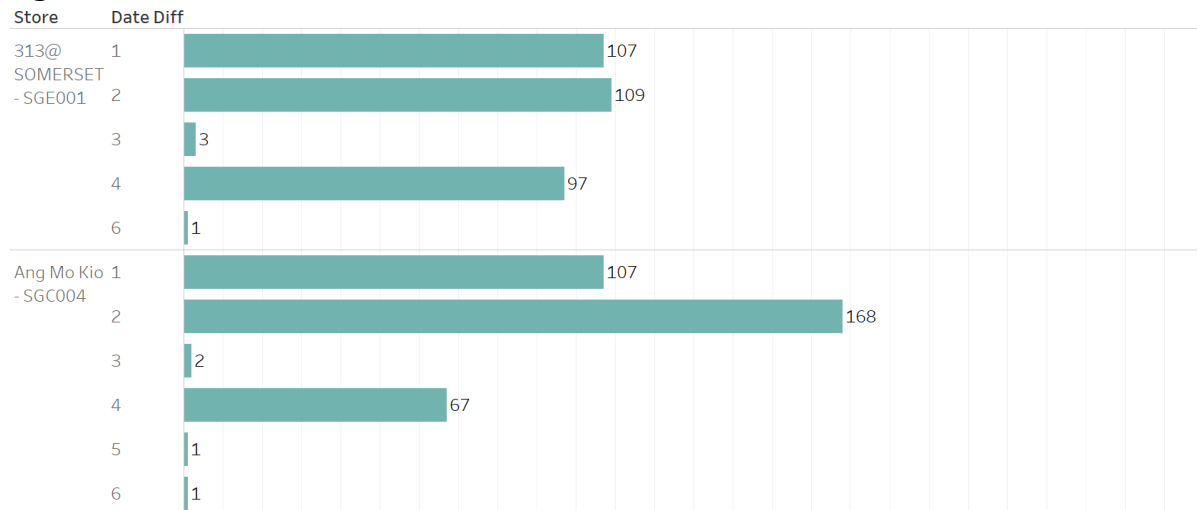


Figure 14: Segment 2 Outlets Reordering Frequency

Similarly, since the above outlets in Segment 2 have similar reordering frequency, this enable us to make further evaluation and business recommendations to our sponsor about delivery.

Segment 3:

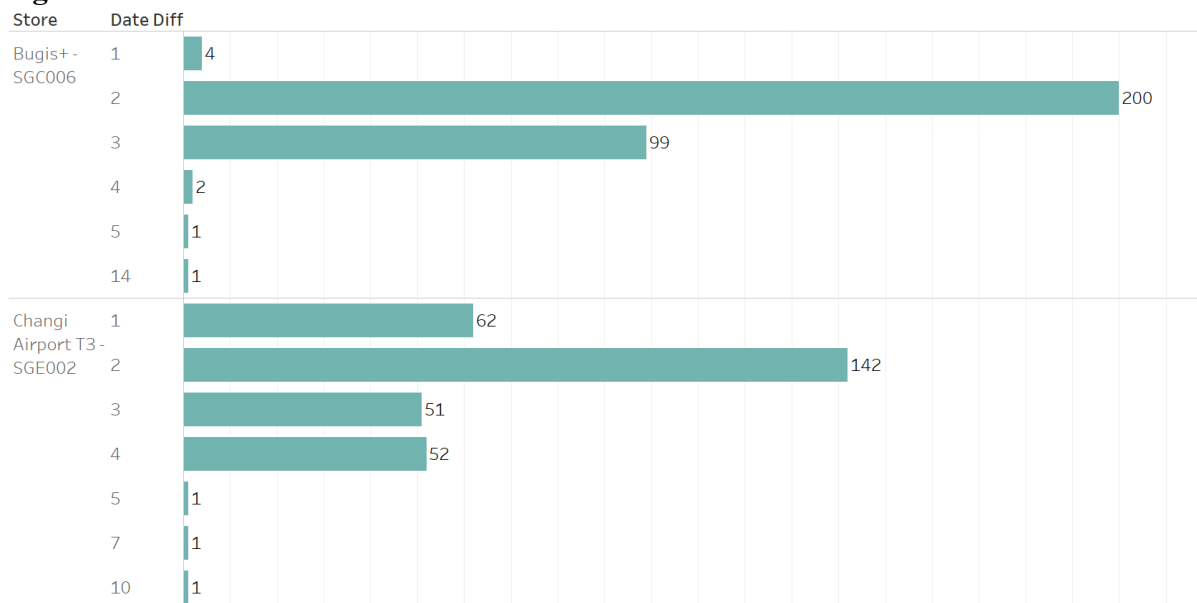


Figure 15: Segment 3 Outlets Reordering Frequency

Regarding outlets in Segment 3, it consists of the outliers and henceforth there is a need for us to drilldown to understand the underlying reasons behind these outliers.

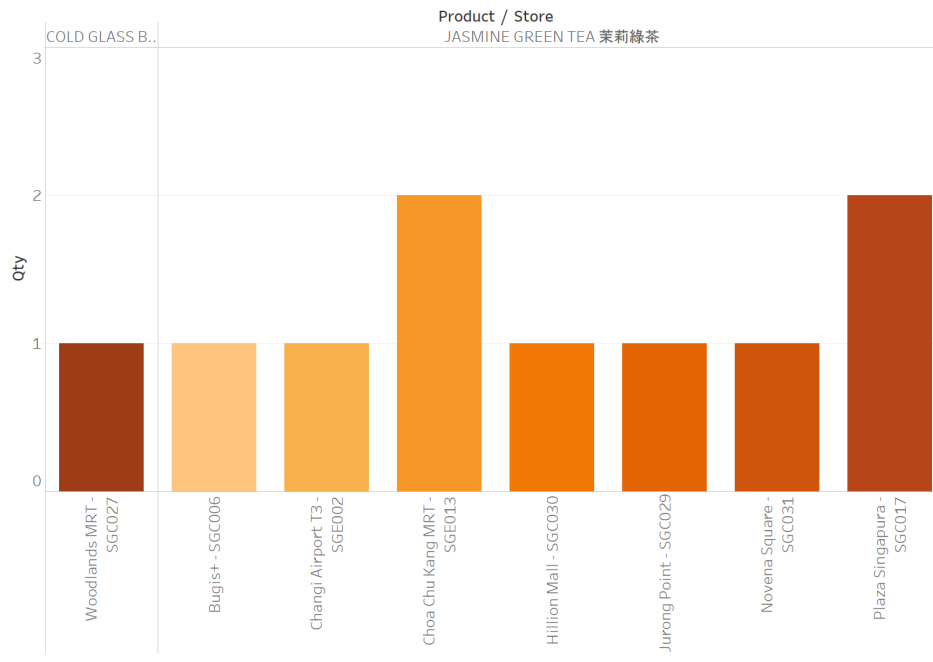


Figure 16: Products ordered with abnormal date difference of ≥ 9 days

From Figure 16 above, we have identified the products that causes the outliers. Jasmine Green Tea and Cold Glass Bottle are the two products that contributed to the outliers with Jasmine Green Tea of a higher concern since it occurs in 7 outlets across all outlets.

Subsequently, we have considered order frequency composition for monthly basis in 2016 to identify what months causes outliers. This is as shown in Figure 17 below:

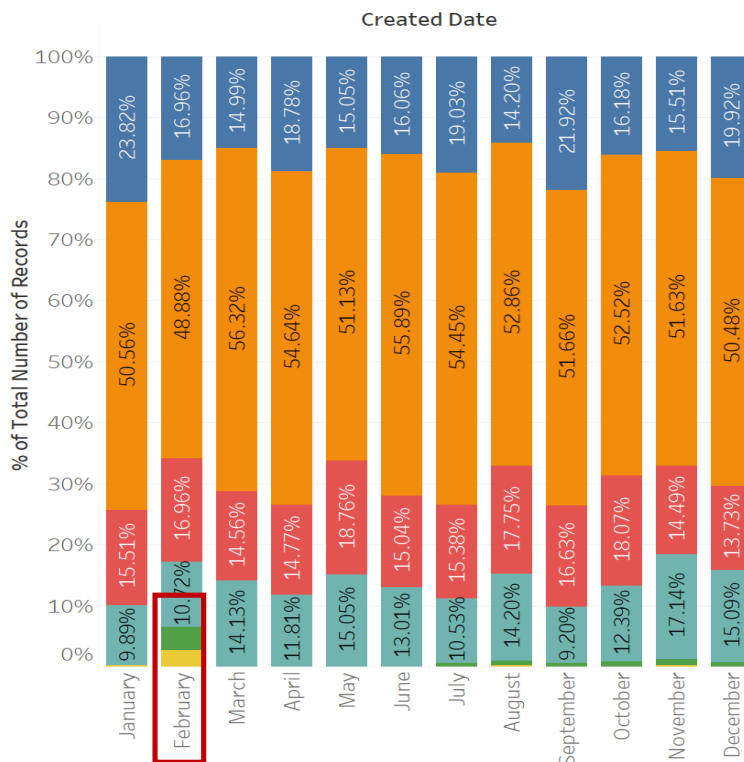


Figure 17: Order Frequency Composition for Monthly in 2016

To further understand the breakdown of reorder frequency by monthly basis, our team have constructed in Figure 17. The chart shows that in months of Feb, there are some abnormally as indicated in the red box. We have further drilled down into the month of Feb, and it is as shown in Figure 18 below.

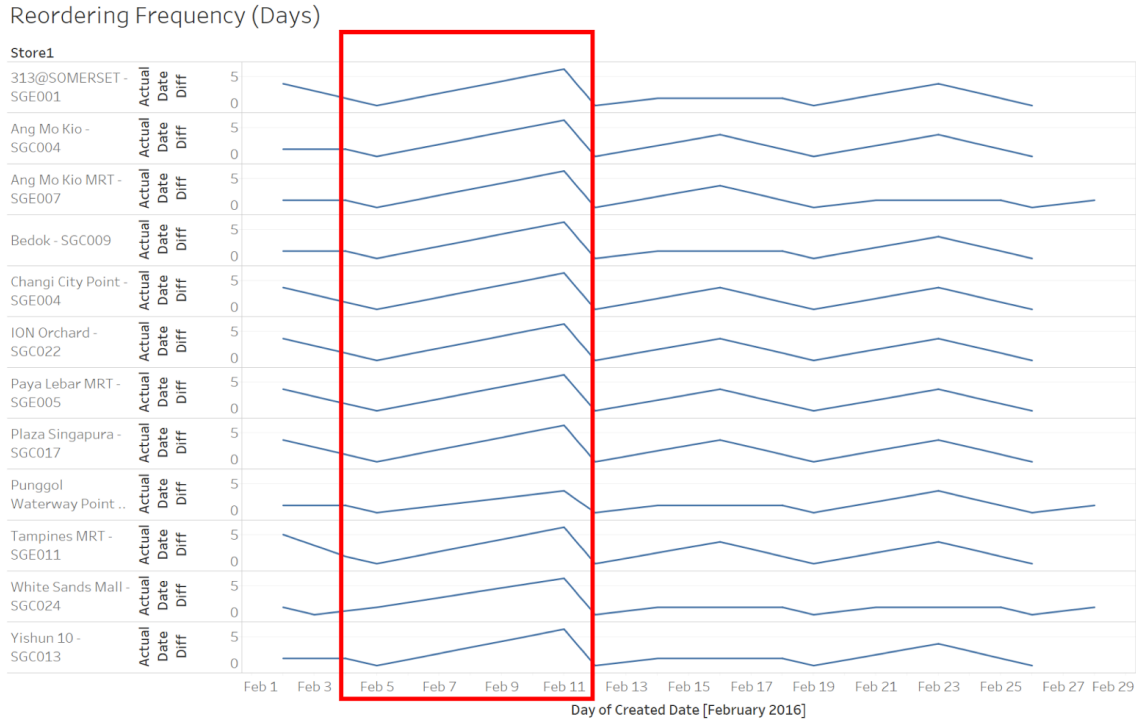


Figure 18: Order Frequency drilled-down into February

Upon drilled-down to understand the underlying reason behind outlier of order frequency over 6 days, we realized that there is a similar trend across all the branches including 313 Somerset, Ang Mo Kio, Ang Mo Kio MRT, Bedok, Changi City Point, ION Orchard, Paya Lebar MRT and Plaza Singapura. We concluded that the long order frequency of 6 days between 5th Feb 2016 to 11th Feb 2016 was due to the Chinese New Year long holiday.

REVISED SCOPE OF WORK

Our scope of work mainly remains the same as stated in our project proposal. However, with the comments from sponsor and supervisor, our team have decided to revise some of the initial scope as specified in our project proposal.

From the feedback, we got from Professor Kam, we will be focusing on the logistic analysis that our sponsor is the most interested in, instead of encompassing the effectiveness of promotion and optimal location analysis based on target population living in the area. The aim of our prediction analysis is to reduce the steep learning curve for new employees. We will be providing a custom set of recommendations for each outlet to re-order.

The table below will better illustrate our updated work scope:

Completed ✓	Removed ∅	In-Progress and New Tasks +
<ol style="list-style-type: none"> 1. Data cleaning on all 5 tables by identifying unrelated data to our analysis (i.e. oversea outlet delivery record which is insignificant in Singapore context) 2. Identify the most popular product across all outlets in year 2016 and 2017 3. Identify the frequency of delivery across two years 4. Identify the most popular branch across all outlets basing on the reorder quantity accumulated over 2 years 5. Identify the reordering frequency in days across all outlets 6. Identify outliers of reorder frequency (days) 7. Identify composition and reason behind the outliers in reorder frequency (days) 	<ol style="list-style-type: none"> 1. Evaluate the effectiveness of the promotional campaigns held 2. Location analysis which is best for outlet opening based on the age group and demographics living in the area 	<ol style="list-style-type: none"> 1. Remove outliers for model building 2. Prepare dataset for model building 3. Clustering for time-series similarities with wrapping method 4. Exponential Smoothing using the weighted moving average to conduct prediction 5. Train and test the model using 70% training and 30% test model 6. Monte Carlo model prediction to have another view on the optimal reorder quantity of each product 7. Comparison of the two models to evaluate which is more effective 8. Written report, final presentation, poster,

<p>8. Deep dive of outliers by months to understand which month contribute to the highest no. of outliers</p> <p>9. Identify underlying reason behind the outliers by outlet</p> <p>10. Identify time series trend of each product across different outlets</p> <p>11. Identify products without time series trend and label it as "No" in created field "Include?"</p>		<p>insights and recommendations for KOI</p>
---	--	---

REVISED METHODOLOGY

As the client are mainly interested in the logistic part of the business, we will narrow down our analysis to focus on the logistic of the business.

Cluster Analysis

Firstly, cluster analysis will be carried out to determine the existence of time-series similarities trend using the wrapping method among KOI's products for each individual outlet. Next, we will be forming time-series exponential smoothing using the weighted-moving average to conduct prediction. Subsequently, we will also be conducting training and test model validation to ensure that our model is not over-fitted. Thereafter, we hope to translate the results obtained to a set of recommended quantities for each product in all the outlet to help KOI reduce its steep learning curve for new employees.

Monte Carlo Model

As consulted from Professor Kam, he has recommended our team to conduct Monte Carlo simulation to stimulate our prediction analysis. Monte Carlo is a computerized mathematical technique that allows people to account for risk in quantitative analysis and decision making. Monte Carlo perform risk analysis by building models of possible results by substituting a range of values for any factor that have uncertainty. It then calculates result repeatedly, each time using different random values. Monte Carlo will then produce distributions of possible outcome values. In our scenario, we will be putting in quantity as our range of values that have uncertainty, basing on each product in the individual outlet. Furthermore, we will split to weekly basis to standardize the period. Thereafter, we will be comparing the results with the results produced from cluster analysis and evaluate the better model for our end analysis. This result will then be submitted to our sponsor to aid them in better restock decision analysis.

REVISED WORK PLAN

S/N	Task		Assigned to	Week -1	Week 0	Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9	Week 10	Week 11	Week 12	Week 13	Week 14	Week 15	
				25th Dec - 31st Dec	1st Jan - 7th Jan	8th Jan - 14th Jan	15th Jan - 21st Jan	22nd Jan - 28th Jan	29th Jan - 4th Feb	5th Feb - 11th Feb	12th Feb - 18th Feb	19th Feb - 25th Feb	26th Feb - 4th Mar	5th Mar - 11th Mar	12th Mar - 18th Mar	19th Mar - 25th Mar	26th Mar - 1st Apr	2nd Apr - 8th Apr	9th Apr - 15th Apr	16th Apr - 22nd Apr	
1	Gathering Requirements	Confirm Project Sponsor	All																		
2		Gather Requirements	All																		
3	Data Gathering	Gather Data from Sponsor	Ryan																		
4	Project Proposal	Refining Project Scope	All																		
5		Proposal Preparation	Angie, Shi Jia																		
6		Wiki Page Preparation	Ryan, Shi Jia																		
7	Research	Literature Review	Angie, Shi Jia																		
8	Proposal Deadline - 14th January 2017 (Week 1)																				
9	Data Exploratory	Data Preparation	All																		
10		Data Cleaning	All																		
11		Data Exploration	All																		
12	Data Exploratory	Generate Findings	All																		
13	Project Revision	Review Progress and Findings with Sponsor	All																		
14	Interim	Interim Report Preparation	Angie, Shi Jia																		
15		Interim Presentation Preparation	Angie, Ryan																		
16		Interim Wiki Update	Ryan																		
17	Interim Practicum Presentation - 26th Feb to 4th March (Week 8)																				
18	Model Building	Data Normalisation	All																		
19		Quantitative Forecasting Analysis	All																		
20		Correlation Analysis	All																		
21		Model Validation	All																		
22	Insights and Recommendations	Generate Visualizations from Analysis Results	All																		
23		Generating Insights	All																		
24		Formulate Recommendations	All																		
25	Project Review	Review Progress and Findings with Sponsor	All																		
26		Align Final Deliverables with Sponsor Requirement	All																		
27	Final Preparation	Abstract Paper Preparation	Shi Jia																		
28		Full Paper Preparation	Shi Jia																		
29		Final Wiki Update	Ryan																		
30		Final Presentation Preparation	Angie																		
31		Final Paper Preparation	Angie, Ryan																		
32		Poster Submission Preparation	Shi Jia, Ryan																		
33	Abstract Paper Submission - 1st April 2018 (Week 12)																				
34	Full Paper Submission - 8th April 2018 (Week 13)																				
35	Undergraduate Conference on Data Analytics Conference Day - 14th April to 15th April (Week 14)																				
36	Analytics Practicum Final Submission - 22nd April 2017 (Week 15)																				

Legend	
Planned	
Actual	
Milestone	

BIBLIOGRAPHY

Forecasting methods. Retrieved January 14, 2018, from <http://mech.at.ua/Forecasting.pdf>

Forecasting data and methods. Retrieved January 14, 2018, from <https://www.otexts.org/fpp/1/4>

Correlation (Pearson, Kendall, Spearman). Retrieved January 14, 2018, from <http://www.statisticssolutions.com/correlation-pearson-kendall-spearman/>

One Map SG School Query. Retrieved January 14, 2018, from <https://www.onemap.sg/main/v2/schoolquery>

Correlation (Pearson, Kendall, Spearman). Retrieved January 14, 2018, from <http://www.statisticssolutions.com/correlation-pearson-kendall-spearman/>

Data Analysis. Retrieved January 14, 2018, from <http://learntech.uwe.ac.uk/da/Default.aspx?pageid=1442>

What is a correlation? And data analysis tools. (2017, July 28). Retrieved January 14, 2018, from <https://www.certs.es/en/blog/correlation-and-data-analysis-tools>

Monte Carlo Simulation. Retrieved February 12, 2018 from http://www.palisade.com/risk/monte_carlo_simulation.asp