Cyclistic Customer Analysis

Stefanus Yudi Irwan – September 11th 2022

Google Data Analytics Portfolio

About Cyclistic

- **Cyclistic** is a bike-share program that features more than **5,800 bicycles** and **692 docking stations** in Chicago. The bikes can be unlocked from one station and returned to any other station in the system anytime.
- Cyclistic **apply flexible pricing plans :** single-ride passes, full-day passes, and annual memberships.

Single-Ride Passes
Full-Day Passes → **CASUAL RIDER**

Annual Memberships → **CYCLISTIC MEMBER**

Outline

Steps → Output

ASK	Scope of Work
PREPARE	Raw Data
PROCESS	Clean Data
ANALYZE	EDA
SHARE	Dashboard
ACT	Marketing Strategy

G & coursera

Google Data Analytics Portfolio

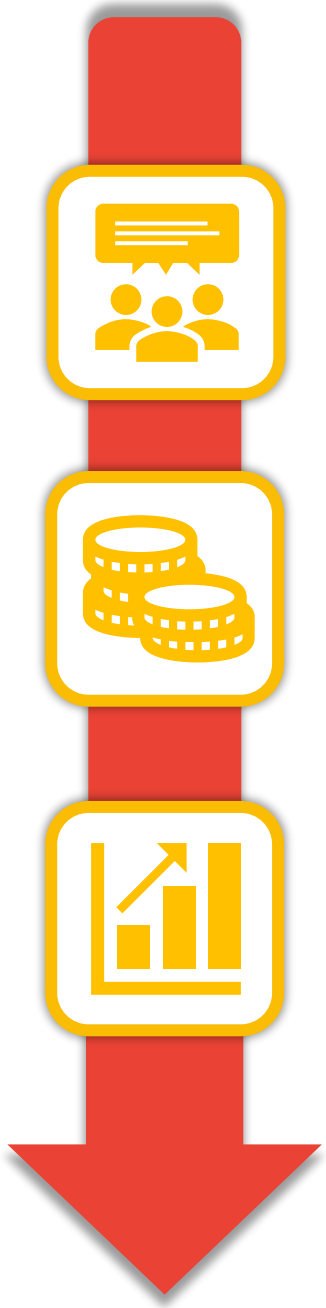
① ASK

Project Stakeholders

1. **CYCLISTIC Executive Team**
   Marketing program recommendation approval

2. **CYCLISTIC Director of Marketing**
   Responsible for bike-share program campaigns
   and initiatives development

Problems



Company want **more profit** to **expand** the business

Annual members are much **more profitable** than casual riders

Annual members will be the key to **company's future growth**.

https://unsplash.com/photos/BIu06wJQ4Cs?utm_source=unsplash&utm_medium=referral&utm_content=creditShareLink

Project Goal

"Increase the number of Annual members by at least 25% within the Q1 of the year, through conversion from casual rider to annual member"

Business Task

"**Analyze** CYCLISTIC **historical bike trip data** to identify the **behavior** of the casual rider and annual member to **enhance** the number of **annual members** through **marketing strategies** to **increase profit** and **grow** the company."

Scope Of Work Document

② PREPARE

Google Data Analytics Portfolio

Photo by Huseyn Kamaladdin:
https://www.pexels.com/photo/brown-wooden-desk-with-rolling-chair-and-shelves-near-window-667838/

Data Collection

Data Requirements.

- Bike user time data
- Bike user place data
- Bike station data
- Type of bike used
- User payment type

Scope of Data.

One year of historical data from **August 2021** till **July 2022** was used for this analysis

Data Source.

Raw data was stored within this website. The data was provided by Google Data Analytics Professional Certificate. This is an open data under this license.

Dataset Overview

1. Raw Data comprises of **12 CSV files**, represent each month bike-share historical data

2. Size of the raw data is **1,01 Giga-Byte**

3. Raw Data comprises of **5,901,463** rows, and **13** columns

Raw Data

Raw Data Composition

1. **Time Data**
   - Started_Time
   - Ended_Time
2. **Spatial Data**
   - Start_latitude
   - Start_longitude
   - End_latitude
   - End_longitude
3. **Category Data**
   - Rideable_type
   - Payment_type
4. **ID Data**
   - Ride_id
   - Start_station_id
   - Start_station_name
   - End_station_id
   - End_station_name

G & coursera
③ PROCESS
Google Data Analytics Portfolio

DATA CLEANING STEPS

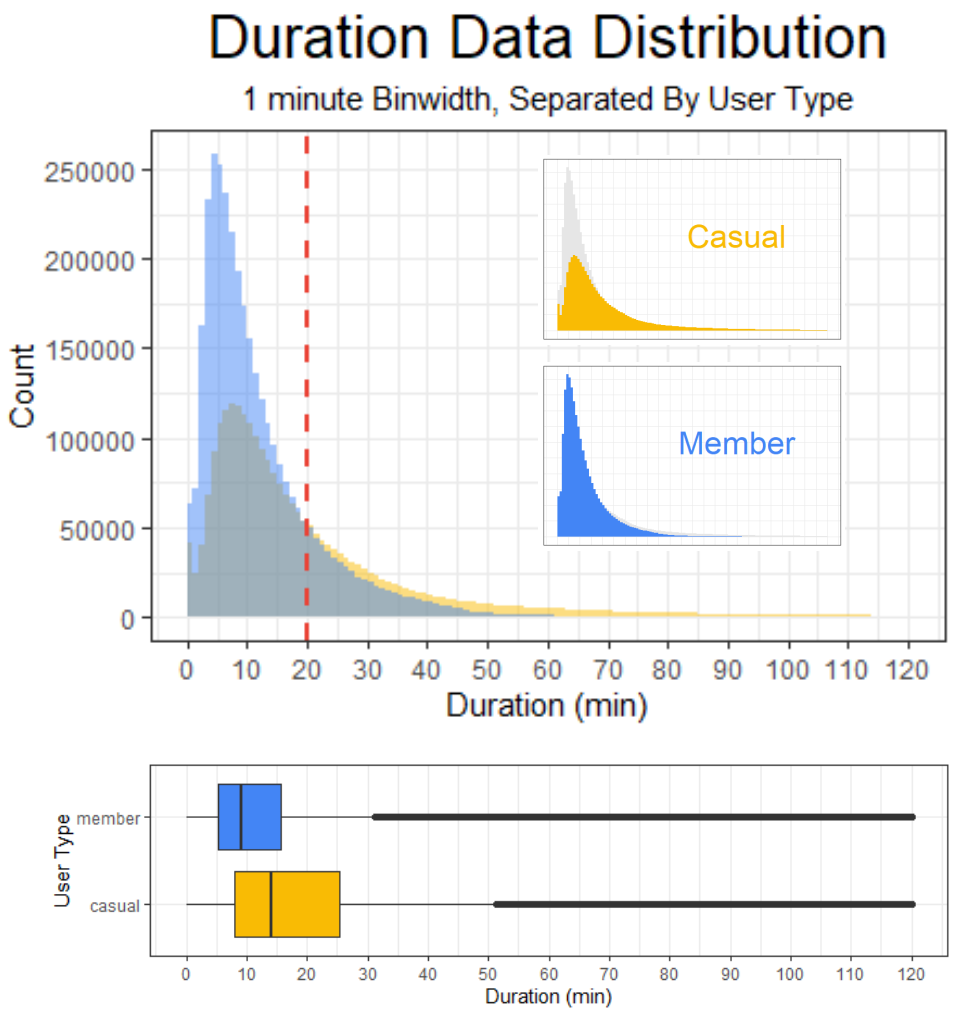
Cleaning data was done by using Microsoft Excel

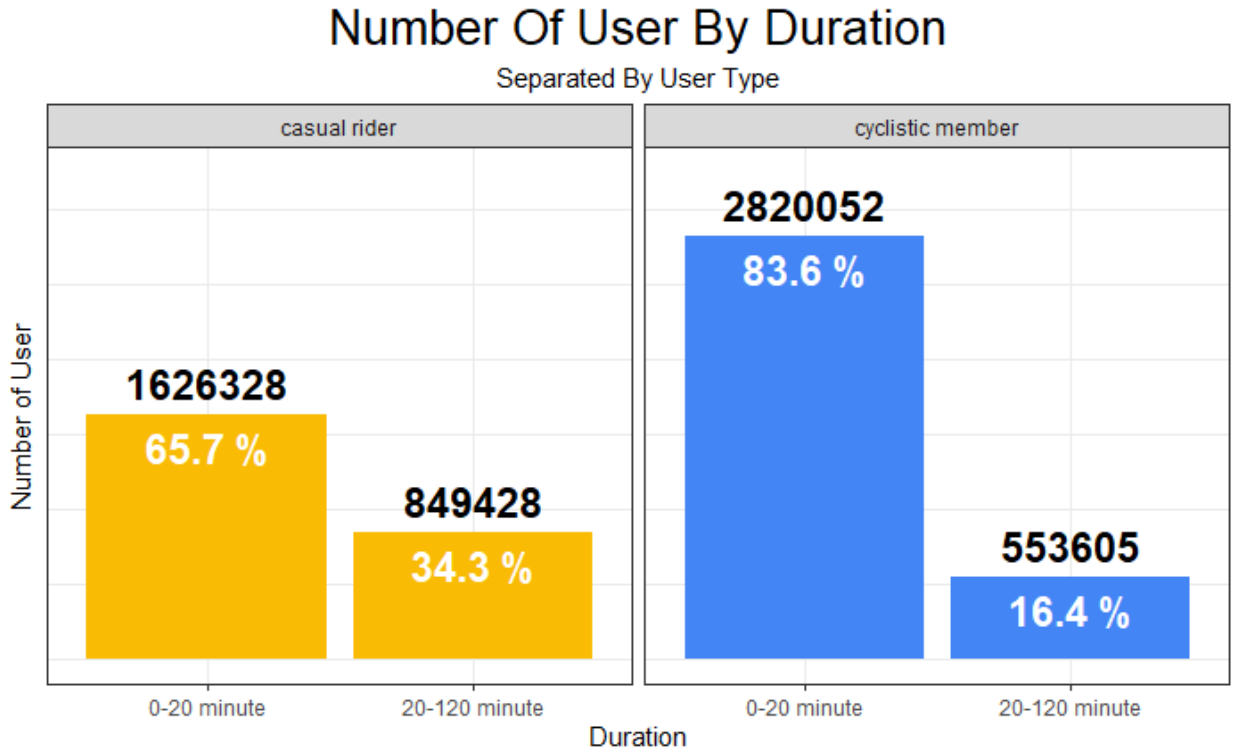
- Clean Data
- Cleaning Log

1. Import csv data to Microsoft excel
2. Adjust time data and spatial data format
3. Check data duplication
4. Check blank value
5. Check spelling for category data
6. Create column “started days name” and “duration”

G & coursera
④ ANALYZE
Google Data Analytics Portfolio

Analysis By Duration

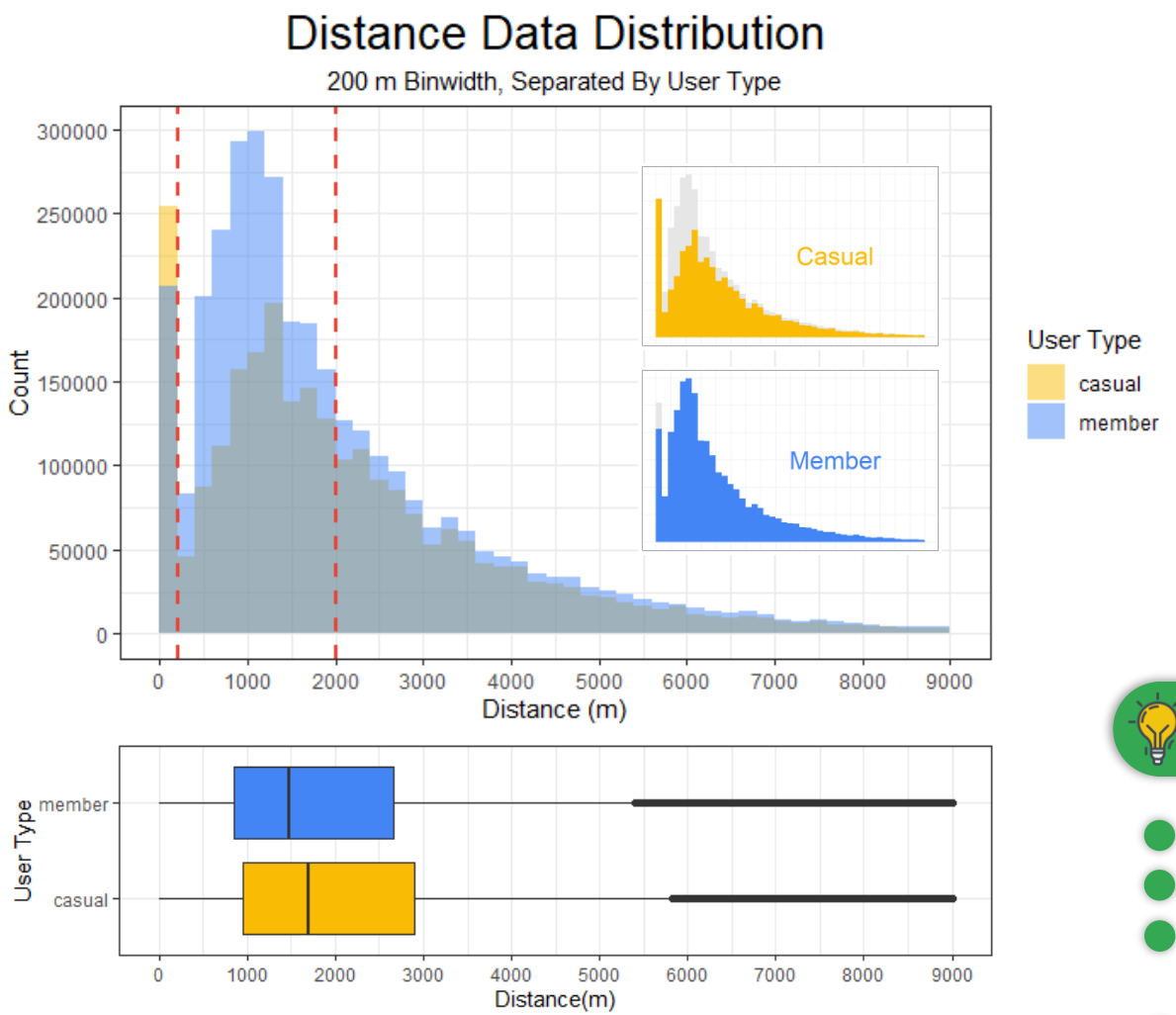


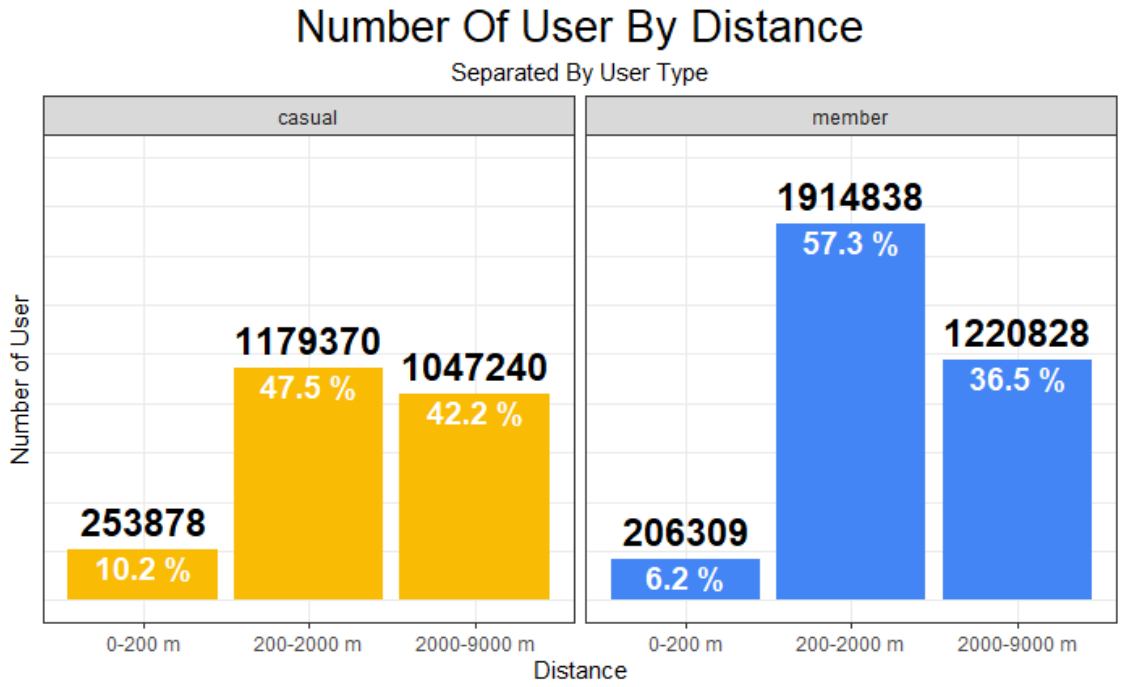


Insight.

- casual rider tend to use the bike **longer** than cyclistic member
- cyclistic member and casual rider most likely uses the bike within time **0-20 minutes.**
- more casual rider at time frame **20-120 minutes** than cyclistic member
- **casual rider** at time frame **0-20 minutes** will be **potential to be cyclistic member**

Analysis By Distance

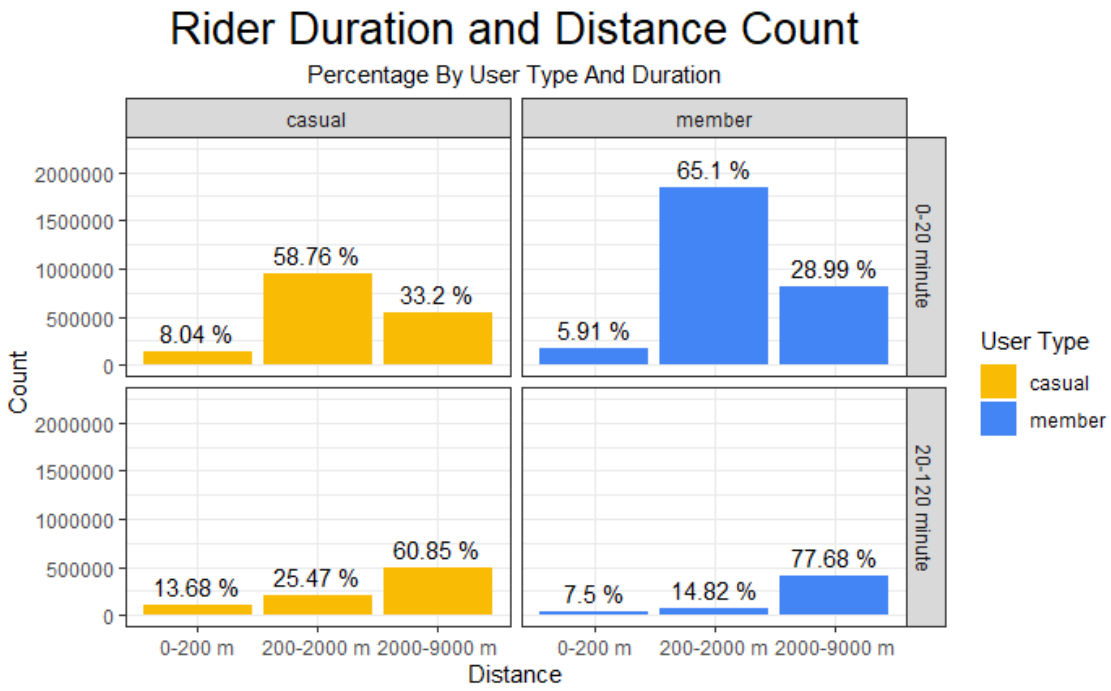


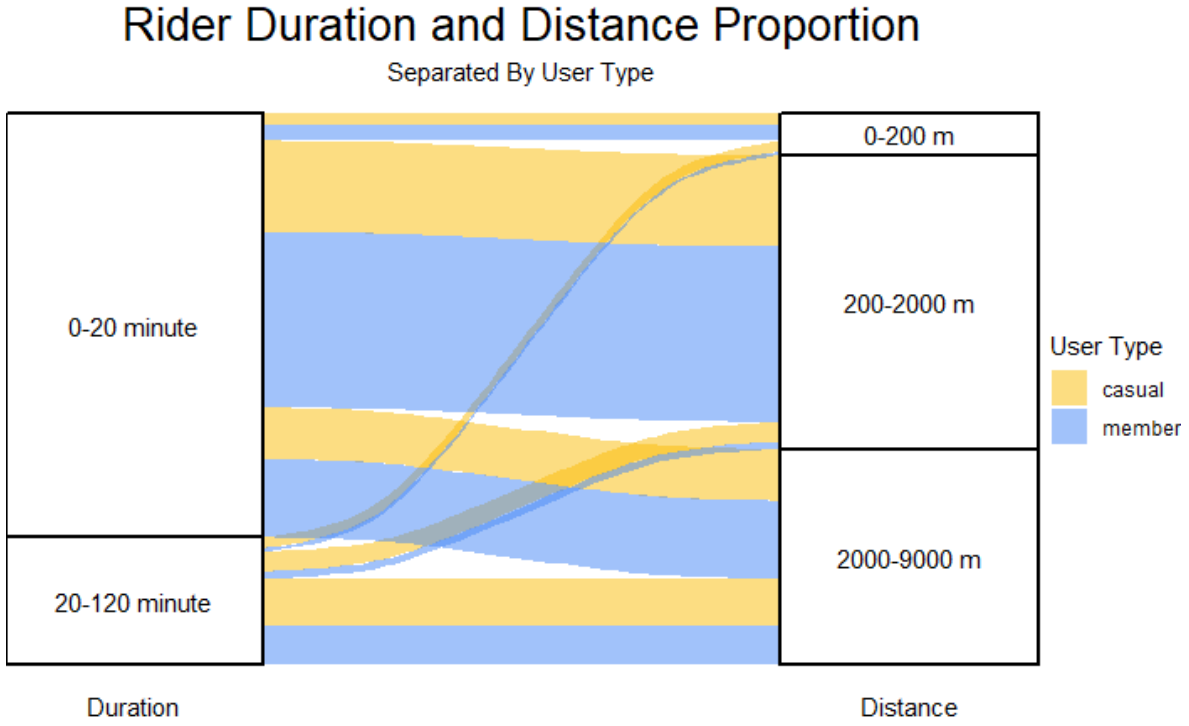


Insight.

- **most cyclistic member** use the bike for distance **200-2000 m.**
- **more casual rider** at distance **0-200 m** than cyclistic member.
- **cyclistic member** has **more difference** between **distance 200-2000m and 2000-9000m** than casual rider.
- **casual rider** at distance **200-2000 m** will be potential to be **cyclistic member.**

Analysis By Duration and Distance

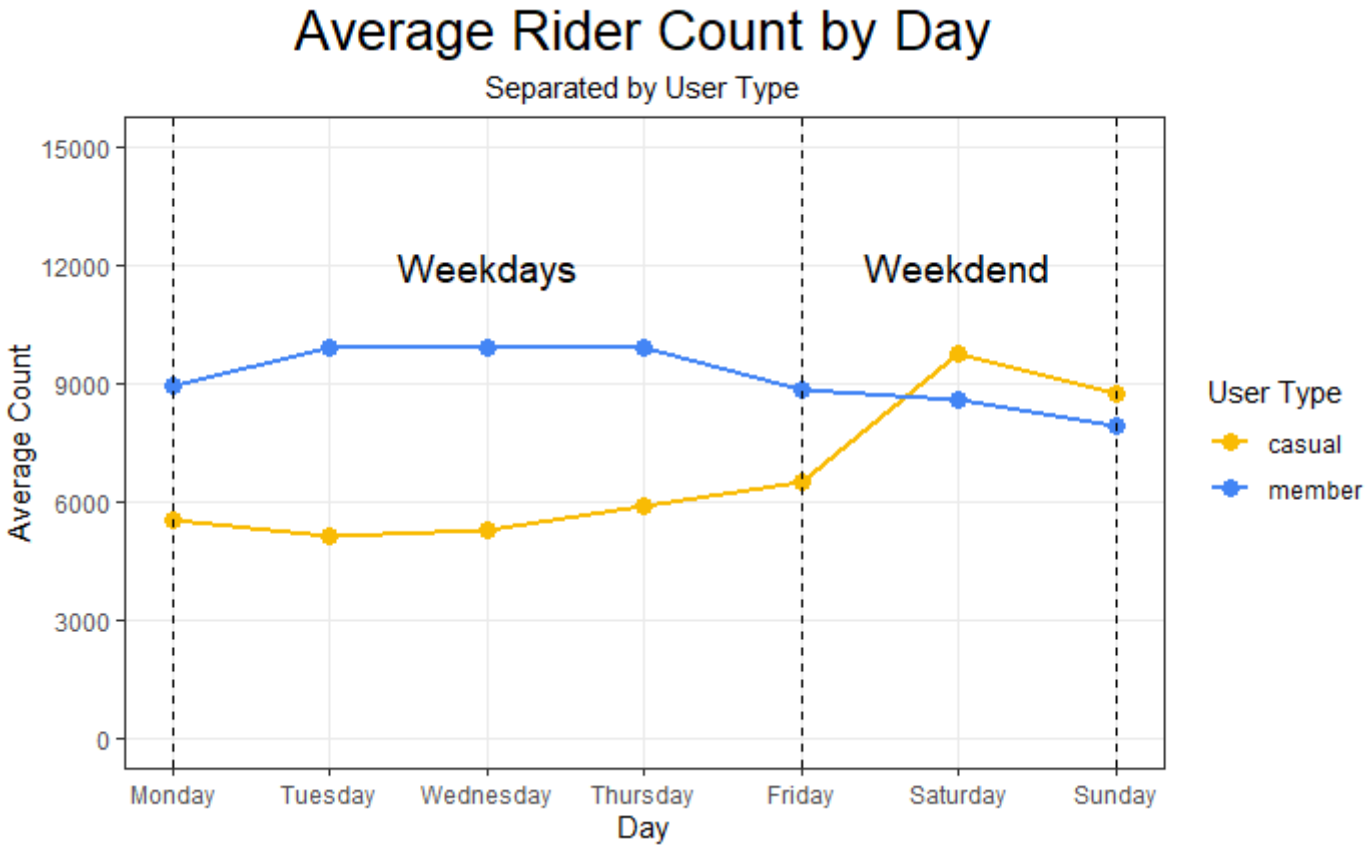




- **Cyclistic member** most likely use the bike for **duration 0-20 minute** and **distance 200-9000m**, also for **duration 20-120 minute** and **distance 2000-9000 m**
- The **difference** between the **most dominant count** and the **rest of the count** in every group is **larger** at **cyclistic member** than at casual rider
- **Casual rider** with **duration 0-20 minute** and **distance 200-9000** m **or** with **duration 20-120 minute** and **distance 2000-9000 m** will be a **good candidate** to be **cyclistic member**

Insight.

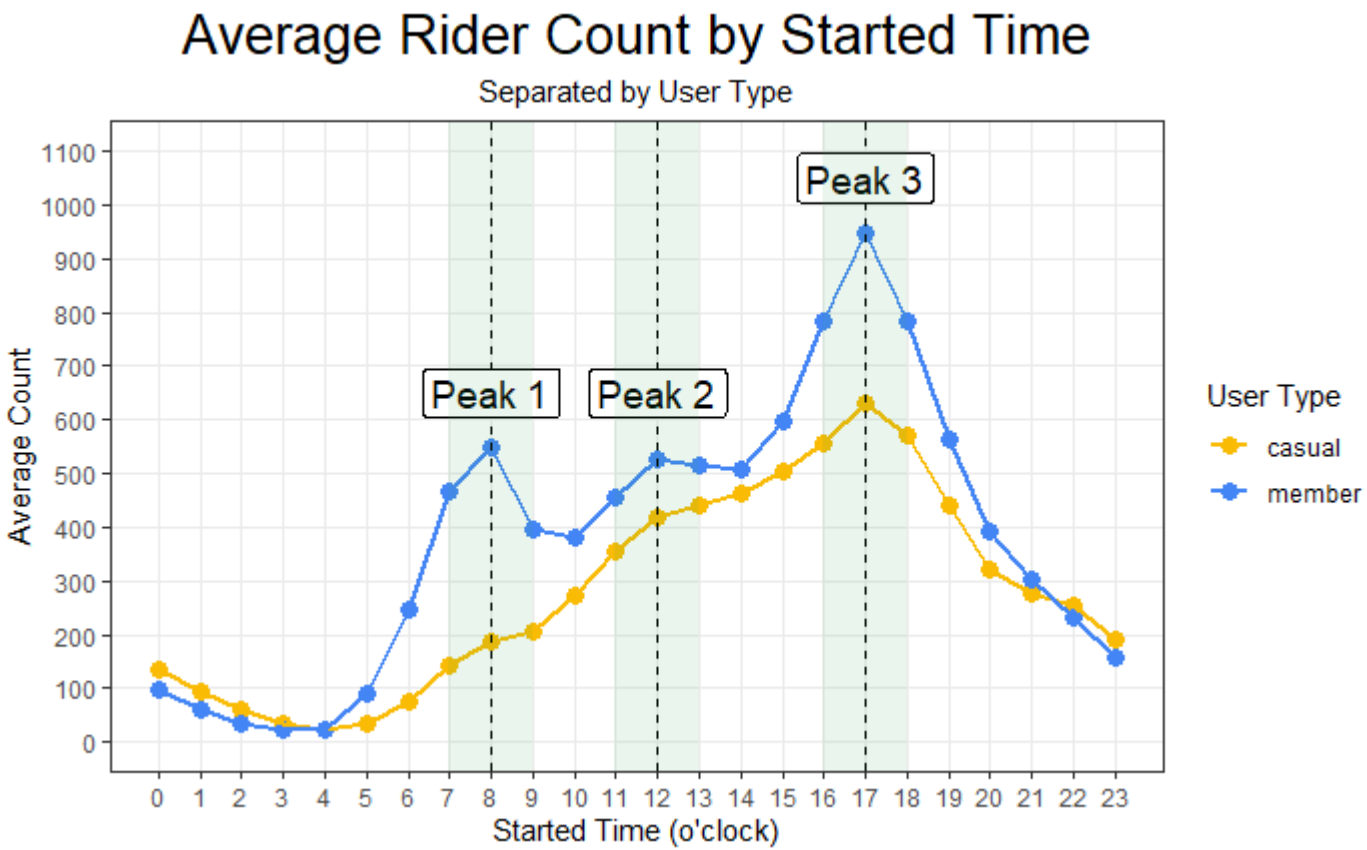
Analysis By Days



Insight.

- Average cyclistic member per day is **stable at weekdays** and **going down at weekend**.
- Average casual rider per day **raise at the weekend.**
- **Casual rider at weekdays** will be good potential to be cyclistic member
- **Weekend** is a **good time** for cyclistic member promotion activity because there are **more casual rider**

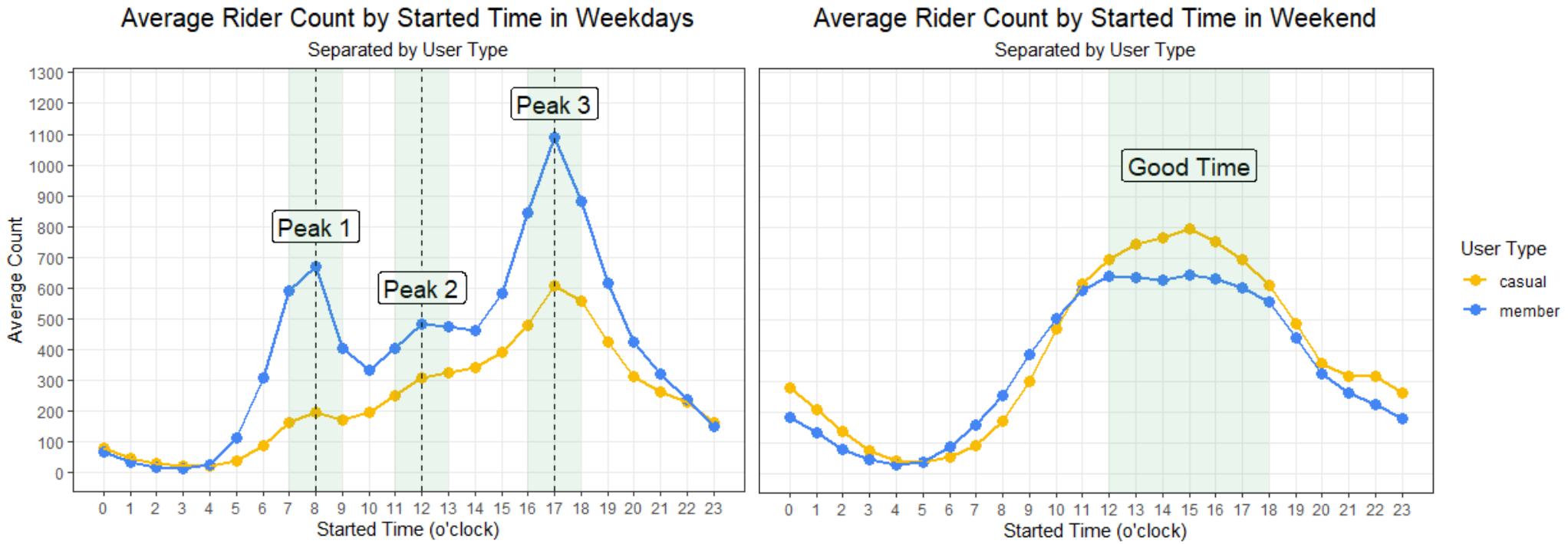
Analysis By Started Time



Insight.

- There are **3 peak time** for cyclistic member, **8**, **12**, and **17** o'clock. This is might be the time when **commuting to work**, **lunch break**, and **commuting to home.**
- Casual member only has **obvious peak** at **17 o'clock.**
- Casual member at time **7-9**, **11-13**, and **16 – 18** o'clock could be good potential candidate to be cyclistic member.

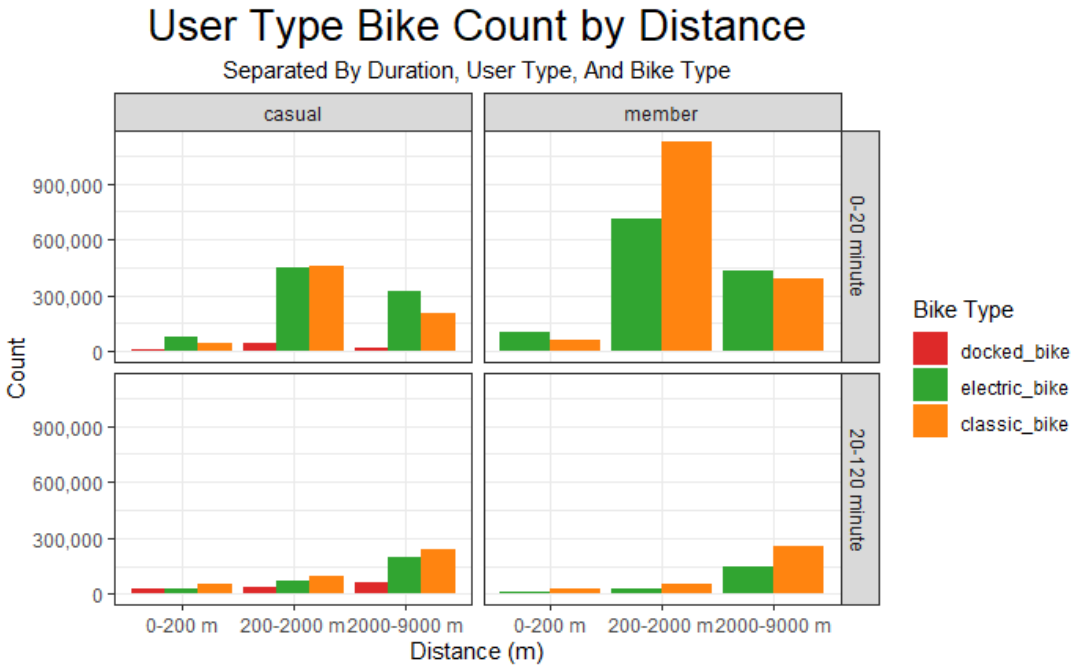
Analysis By Days and Time

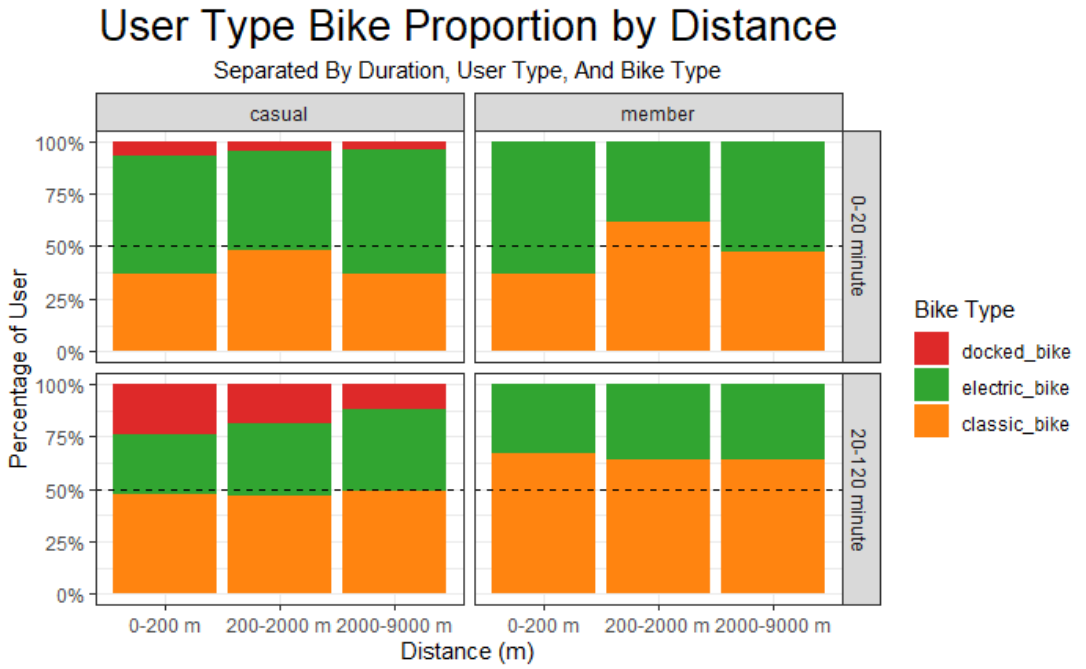


Insight.

- in weekdays there are still **3 peak time** for cyclistic member, **8**, **12**, and **17** o'clock.
- casual member at weekdays now has 2 **obvious peak** at **8** and **17 o'clock.**
- more casual rider at weekend **12 – 18 o'clock** (afternoon) this could be a good time to **promote** the **cyclistic member conversion program**
- **casual rider** at time **7-9**, **11-13**, and **16 – 18** o'clock at **weekdays** could be **good potential** candidate to be **cyclistic member**.

Analysis By Bike Type





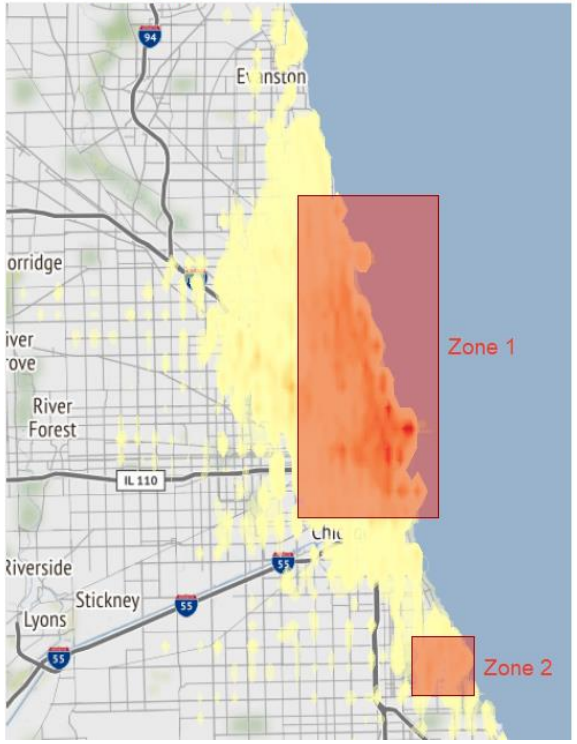
Insight.

- **Cyclistic members** prefer to use a **classic bike over an electric bike** for distances **0-200 m** and duration **0-20 minutes**, while **casual riders** are **still 50:50 for** classic bikes or electric bikes.
- For distances **2000 – 9000 m** and **duration 0-20 minutes** either cyclistic members or casual riders **prefer an electric bike to a classic bike**
- There are **no cyclistic members** who used **docked bikes**, then there is **no use** to **promote** cyclistic members to docked bike users.
- **Casual rider** with a **short duration** (0-20 minute) and **medium to long distances** (200-9000 m) who use the **classic or electric** bike is a **good group** to be cyclistic member

Analysis By Station

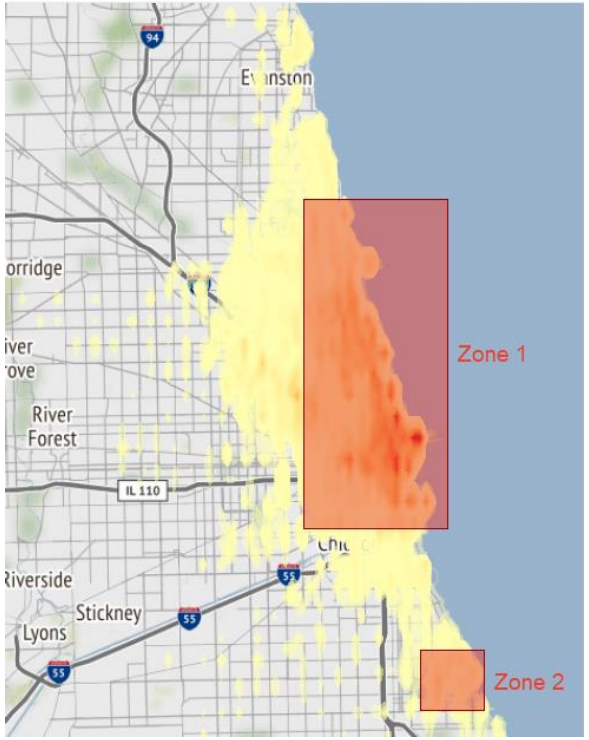
Casual Rider Start Station Density Plot

Show The Frequency of Start Station Used By Casual Member



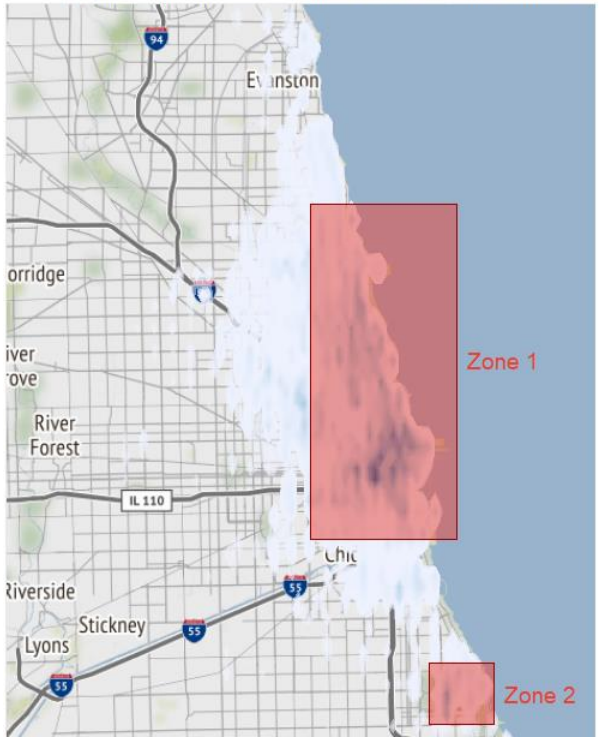
Casual Rider End Station Density Plot

Show The Frequency of End Station Used By Casual Member



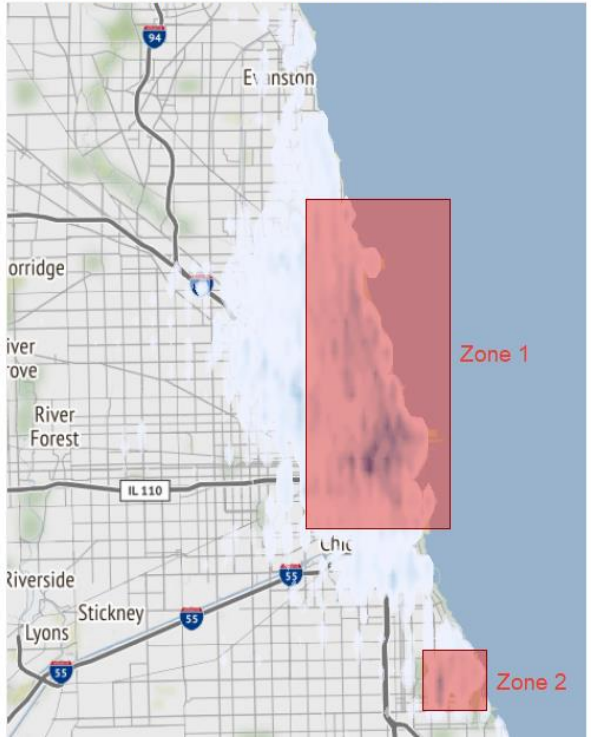
Cyclistic Member Start Station Density Plot

Show The Frequency of Start Station Used By Cyclistic Member



Cyclistic Member End Station Density Plot

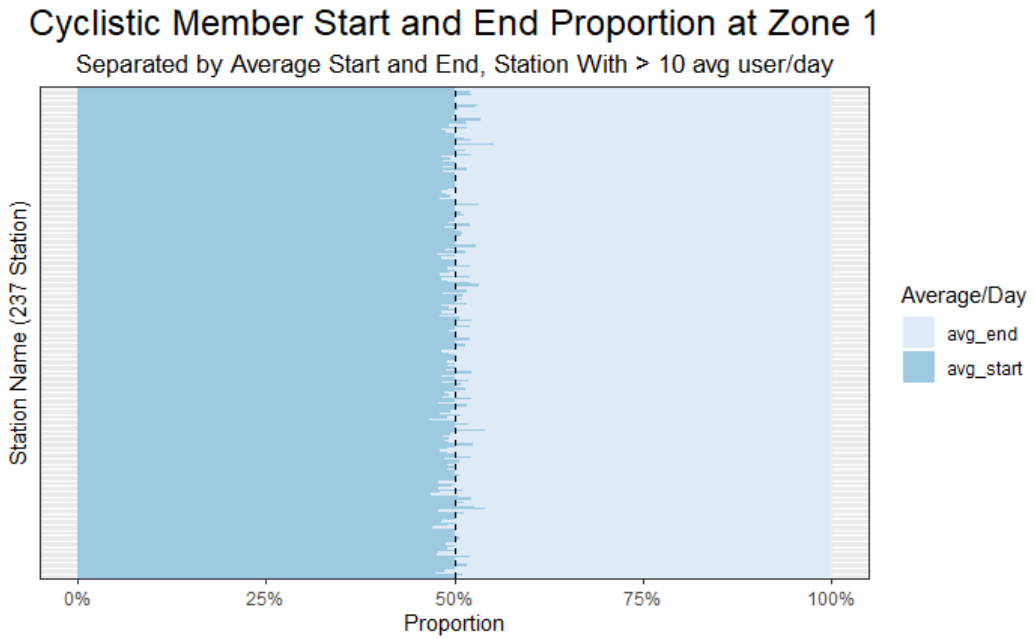
Show The Frequency of End Station Used By Cyclistic Member

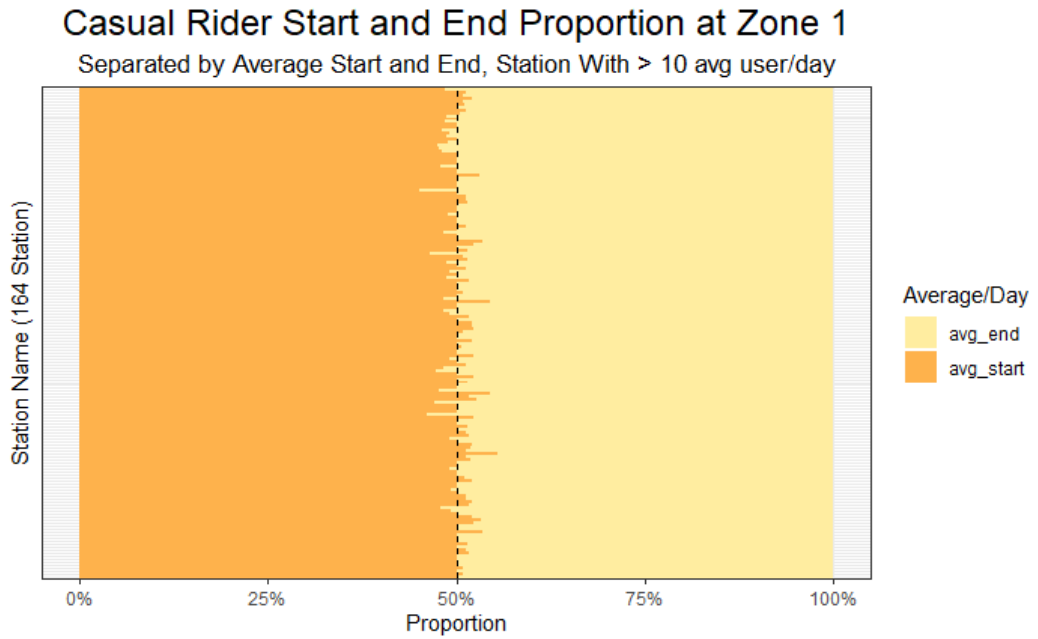


Insight.

- There are **high number of bike user** for casual rider and cyclistic member at **Zone 1 and 2**
- From this density map, we can infer that between **casual riders** and **cyclistic members** there are **similarities of the station** used
- **Zone 1** posses a good opportunity for cyclistic member marketing campaign

Analysis By Station

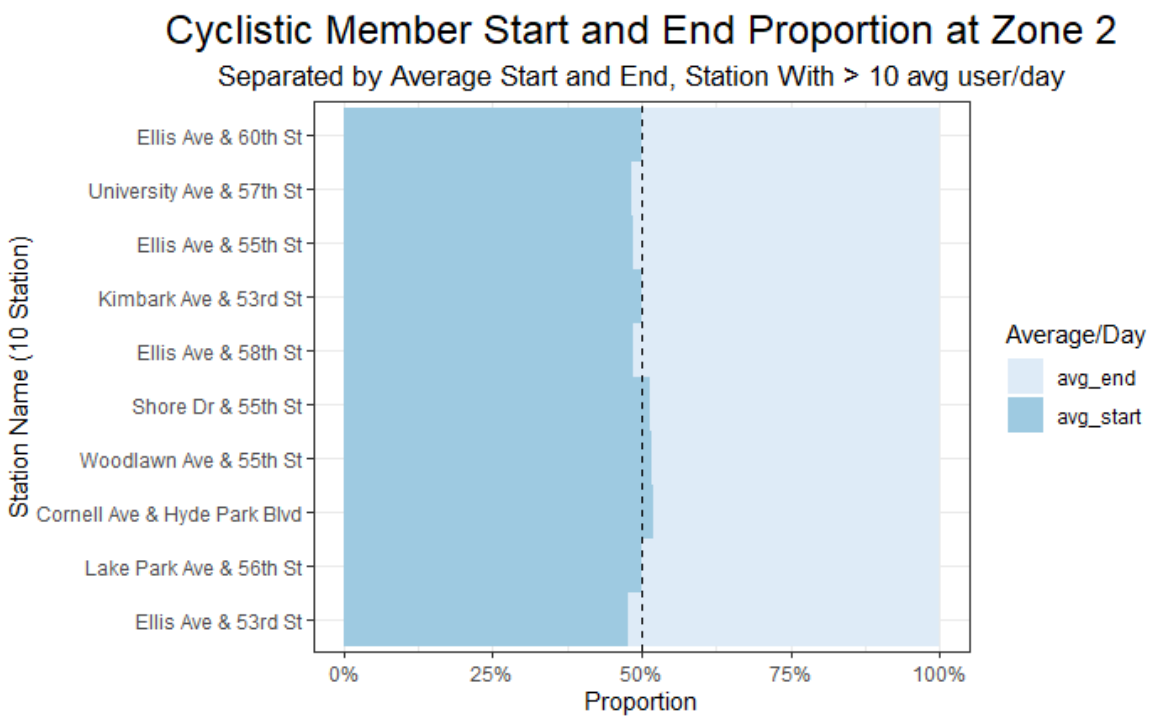


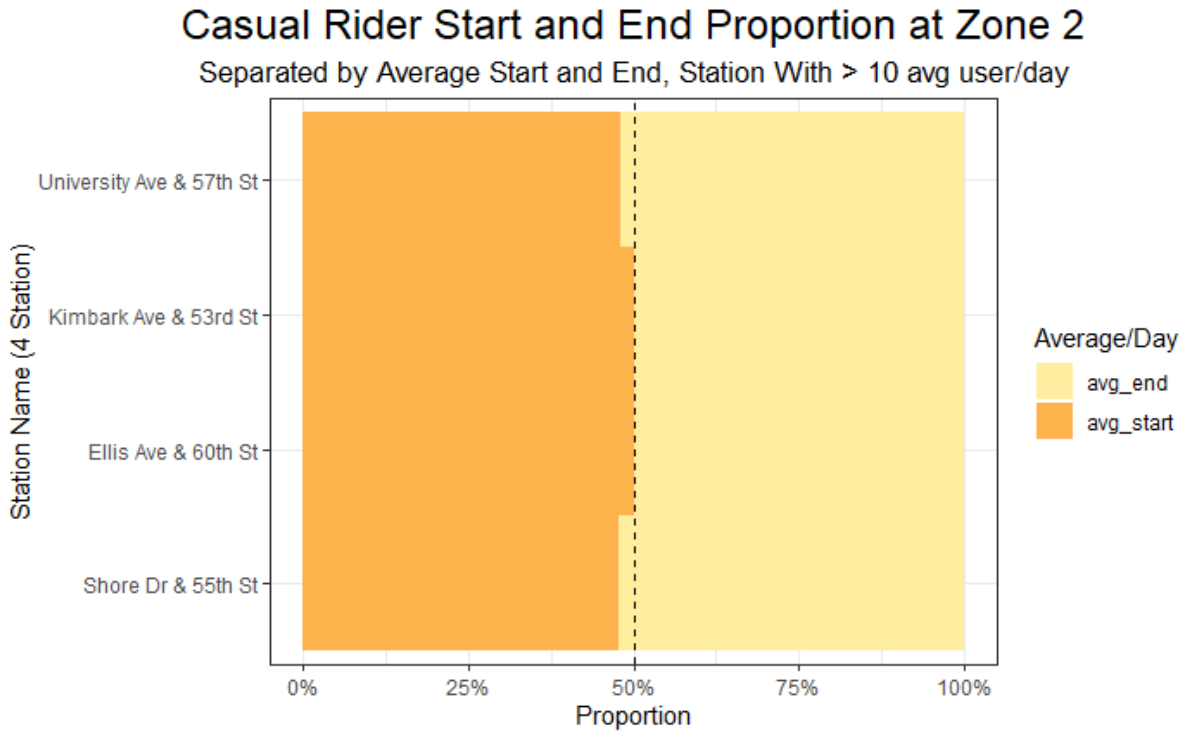


Insight.

- Station with an average user **more than 10 user/day in zone 1**, has a **50:50 proportion** between the user who **starts using the bike** and the user **who finish using the bike** in the **same station**
- In zone 1, **behaviour** of **casual riders and cyclistic member was similar**, they use the same station to start using the bike and end using the bike
- From this chart we could assume that **there are fix number of casual rider** and **cyclistic member** whom **used the bikes for commuting in zone 1**
- Casual rider use **less station** compare to cyclistic member, **164 : 237 station**, for **commuting in zone 1**

Analysis By Station



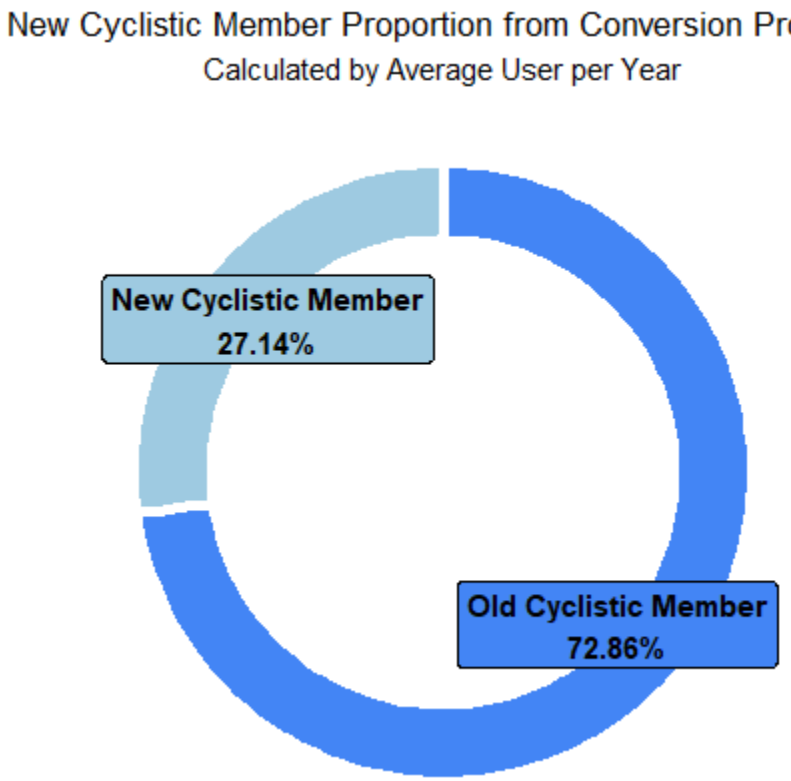


Insight.

- Station with an average user **more than 10 user/day in zone 2**, also has a **50:50 proportion** between the user who **starts using the bike** and the user **who finish using the bike** in the **same station**
- In zone 2, **behaviour** of **casual riders and cyclistic member was similar**, they use the same station to start using the bike and end using the bike
- Casual rider use **less station** compare to cyclistic member, **4:10 station**, for **commuting in zone 2**

Analysis By Station

Station Target Zone 1 (160 Station)	Station Target Zone 2 (4 Station)
Stretter Dr & Grand Ave	University Ave & 57th St
DuSable Lake Shore Dr & Monroe	Kimbark Ave & 53rd St
Millenium Park	Ellis Ave & 60th St
Michigan Ave & Oak St	Shore Dr & 55th St
DuSable Lake Shore Dr & North Blvd	
Shedd Aquarium	
Theater on the Lake	
Wells St & Concord Ln	
…	



Insight.

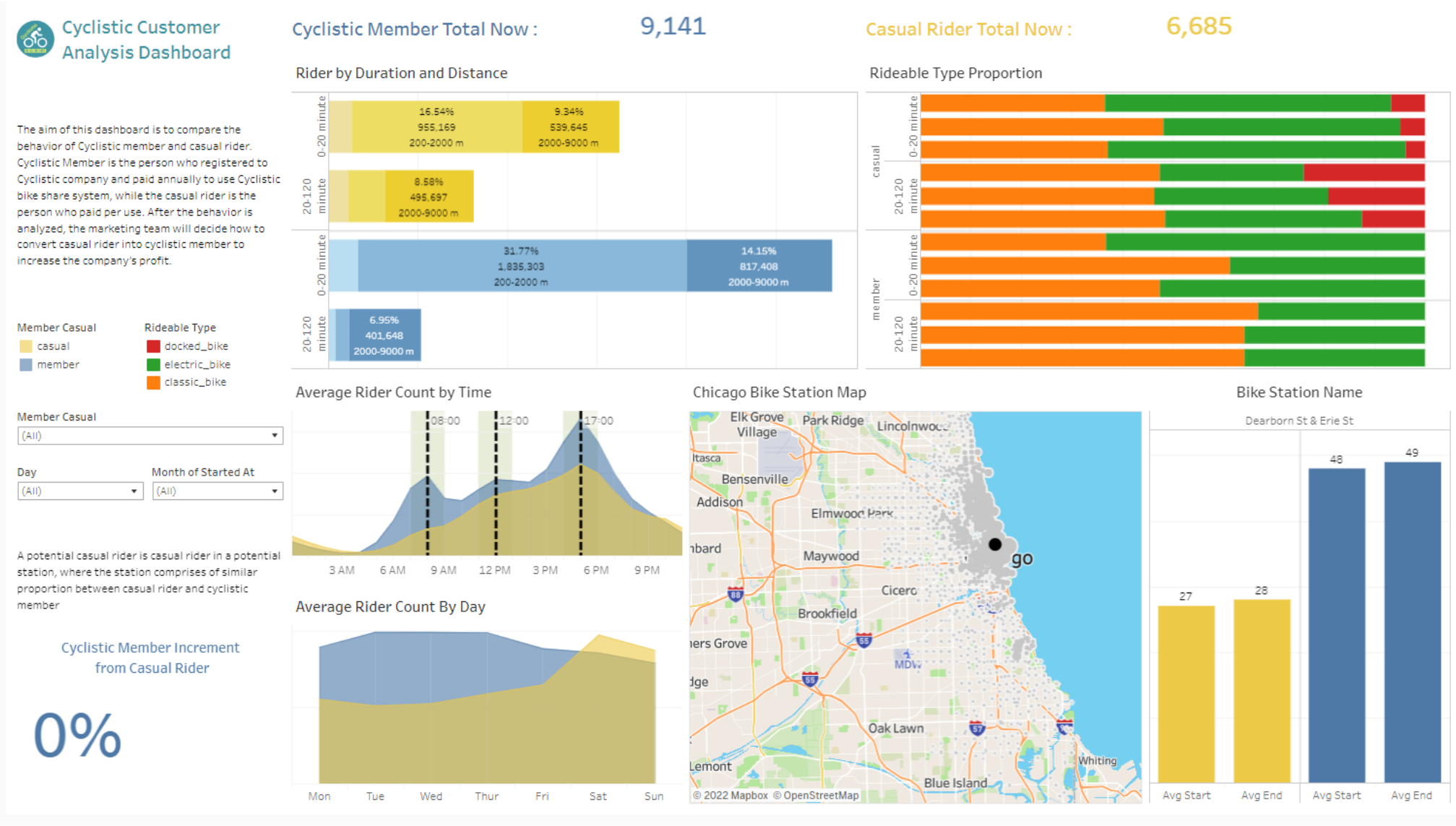
- For **efficient marketing campaign**, we could target **160 most visited station** in **zone 1** and **4 most visited station** in **zone 2** by commuted casual rider
- If we could target commuted casual rider at those station, We could get a new Cyclistic Member scheme comprises **of 27,14% new member** from casual rider and **72,86% Old Cyclistic Member.** There are **37,3% increase** in Cyclistic Member head count

Photo by fauxels: https://www.pexels.com/photo/top-view-photo-of-people-near-wooden-table-3183150/

⑤ SHARE

Google Data Analytics Portfolio

Tableau Dashboard



⑥ ACT

https://www.behr.com/colorfullybehr/best-colors-for-home-offices/

Strategy Recommendation

Stations 1

To increase the number of cyclistic members by the end of this quarter, there are **160 most-used stations in zone 1**, and **4 most-used stations in zone 2**. If we could convert all of the cyclistic members who usually use the bike for commuting at those stations, there will be a **37.3% increase in the amount of cyclistic members**.

Time 2

Run promotion program on **weekdays** when **people commute** in the morning **(7:00-9:00 A.M)**, in lunchtime **(11:00 AM - 1:00 PM)**, or in the afternoon **(4:00-6:00 PM)**. Conduct also the promotion program on **weekends** every **12:00-6:00 PM** because at this time there is **much casual rider** in Chicago.

Bike Type 3

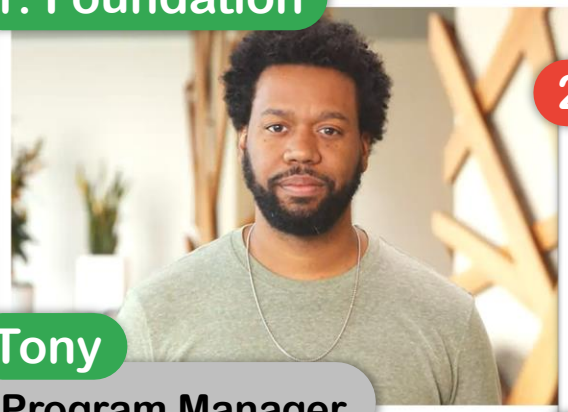
Never target **docked bike** user, just target casual riders who use **electric bikes** or **classic bikes**. First target the classic bike users, then target the electric bike users.

Duration and Distance 4

Target the casual rider who use the bike with a **duration of 0-20 minute and distance 200-2000 m** first, then the second target is the group with a **duration 0-20 minutes and distance of 2000-9000 m**, and the last target is the group with a **duration of 20-120 minutes and a distance of 2000-9000 m**.

G & coursera

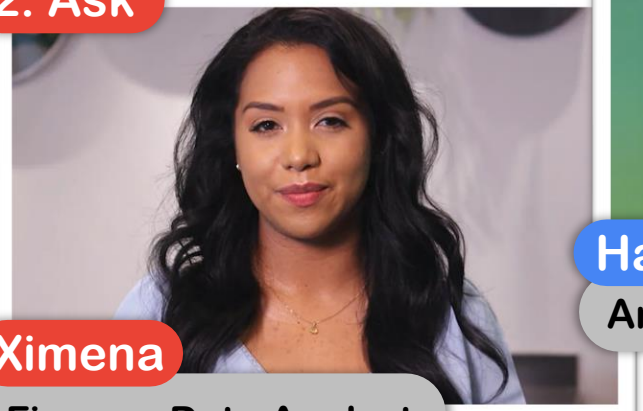
1. Foundation



Tony

Program Manager

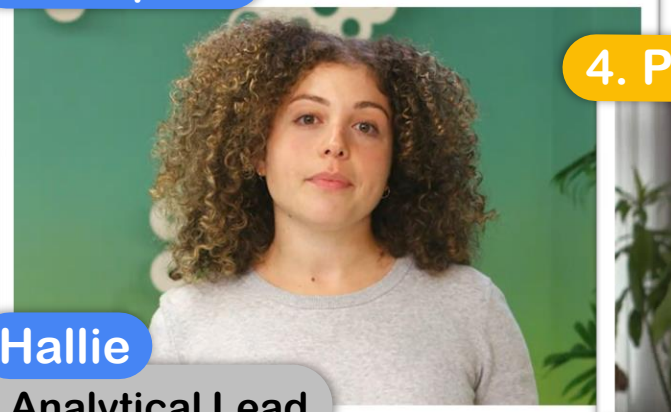
2. Ask



Ximena

Finance Data Analyst

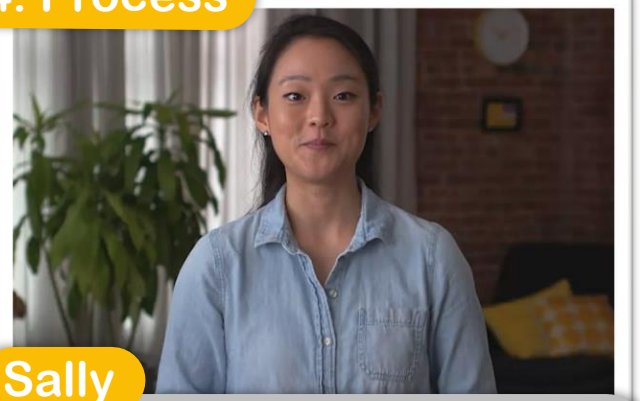
3. Prepare



Hallie

Analytical Lead

4. Process

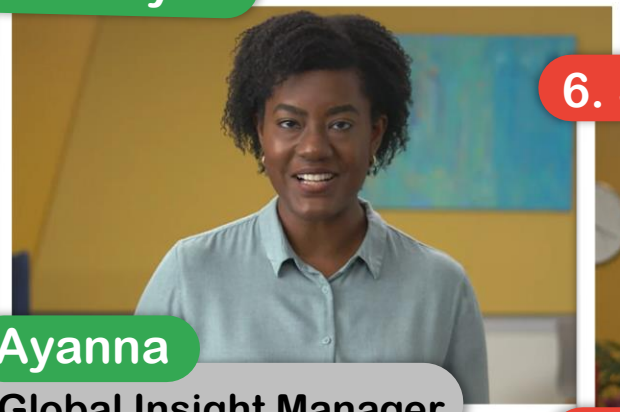


Sally

Measurement & Analytical Lead

Thank You Google

5. Analyze



Ayanna

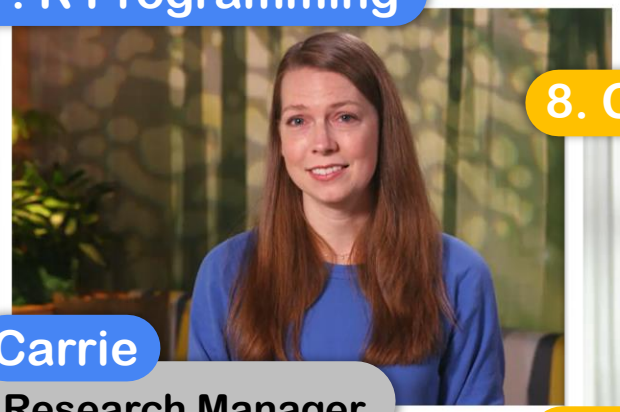
Global Insight Manager

6. Share

Kevin

Director Of Analytics

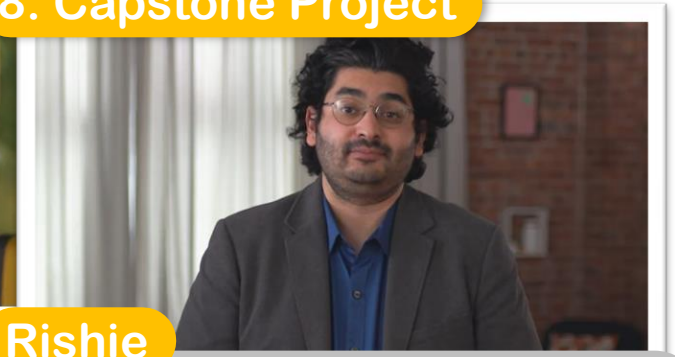
7. R Programming



Carrie

Research Manager

8. Capstone Project



Rishie

Global Analytics Skills Curriculum

Google Data Analytics Portfolio

Reach me ! for further discussion

My_Resume

yudi.stefanus22@gmail.com

https://www.linkedin.com/in/stefanusyudi22