GOOGLE CAPSTONE PROJECT: CHICAGO BIKE-SHARE

Business Task

- Analysis of Existing Usage of Cyclistic Bike-share Service
- Strategy for Conversion of Casual Riders to Annual Memberships
- Evaluate Use of Digital Media as a Marketing Channel

Background

Cyclistic is a bike-share company in Chicago with 5,824 bicycles and 692 docking stations. Cyclistic sets itself apart by also offering reclining bikes, hand tricycles, and cargo bikes, making bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike (about 8% of riders). The majority of riders opt for traditional bikes. Cyclistic users are more likely to ride for leisure, but about 30% use them to commute to work each day. Customers have 3 options for using the service:

1. Annual pass
2. Single-ride
3. Full-day pass

These first category of users are Cyclistic members and the others are referred to as casual riders. Currently Cyclistic relies upon building general awareness of the brand and appealing to a broad range of consumer segments.

My role is as a junior analyst in the market analysis team wants to understand how casual riders and annual members use Cyclistic bikes differently. From these insights, my team will design a new marketing strategy to convert casual riders into annual members. But first, Cyclistic executives must approve the recommendations, so they must be backed up with compelling data insights and professional data visualizations. Stakeholders include senior executives, the market analysis team and the Director of Marketing, Lily Moreno, believes the company's future success depends on maximizing the number of annual memberships.

Preparation of Data

Data sources used:

- **Cyclistic Rides dataset 8/2021-7/2022:** Bike ride data from Cyclistic with locations, membership details, start times/end times and start locations/end locations.

- **Chicago Crime Dataset 8/2022:** Data on crimes recorded in Chicago in the first 8 months of 2022
- **ZORI Chicago Multifamily Rents:** Data provided by zillow.com on average rents achieved in Chicago districts sorted by Postcode

Cleaning & Manipulation of Data

The data is obtained from a verified source, Motivate International (licence here). The data is contained in multiple zipped data files. After glimpsing individual files in MS Excel, it was clear that, due to being very large data files, it would be more convenient to organize the data in R-studio than in MS Excel. After reviewing the structure of the data using the STR, head, colnames functions, the principal issues with the data were blank cells, unclear column headings and date format. As the data was anonymized for any sensitive or personal data, it was deemed not necessary to use additional password protection as the new csv file is stored on a secure laptop. The data was transformed and consolidated in one .csv file for analysis and presentation in Tableau. Additional columns created for weekdays and duration of ride (per instruction) and additional column for months and seasons, blank cells were filtered and removed.

As the premise of the analysis included a geographical component, additional external data was obtained for Chicago apartment rents and crime statistics. This data is publicly available and obtained from ZORI and the City of Chicago. As with the core data, the data was cleaned and transformed to be useful for analysis.

Cyclistic Ride Data

```
library("tidyverse")
library(readr)
library(devtools)
library(lubridate)
library(janitor)
library(ggplot2)
library(data.table)
library(dplyr)
library(hms)


distinct(cyclistic_trips)
cyclistic_trips$date <- as.Date(cyclistic_trips$started_at)
cyclistic_trips$weekday <- format(as.Date(cyclistic_trips$date), "%A")
cyclistic_trips$month <- format(as.Date(cyclistic_trips$date), "%b_%y")
cyclistic_trips$year <- format(cyclistic_trips$date, "%Y")
```

```
cyclistic_trips$time <- format(cyclistic_trips$started_at, format = "H%:M%")
cyclistic_trips$time <- as_hms(cyclistic_trips$started_at)
cyclistic_trips$hour <- hour(cyclistic_trips$time)

cyclistic_trips <- cyclistic_trips %>% mutate(time_of_day =
                                    case_when(hour == "0" ~ "night",
                                              hour == "1" ~ "night",
                                              hour == "2" ~ "night",
                                              hour == "3" ~ "night",
                                              hour == "4" ~ "night",
                                              hour == "5" ~ "night",
                                              hour == "6" ~ "night",
                                              hour == "7" ~ "morning",
                                              hour == "8" ~ "morning",
                                              hour == "9" ~ "morning",
                                              hour == "10" ~ "morning",
                                              hour == "11" ~ "midday",
                                              hour == "12" ~ "midday",
                                              hour == "13" ~ "midday",
                                              hour == "14" ~ "midday",
                                              hour == "15" ~ "midday",
                                              hour == "16" ~ "evening",
                                              hour == "17" ~ "evening",
                                              hour == "18" ~ "evening",
                                              hour == "19" ~ "evening",
                                              hour == "20" ~ "night",
                                              hour == "21" ~ "night",
                                              hour == "22" ~ "night",
                                              hour == "23" ~ "night",
                                              hour == "24" ~ "night")
)
```

```
cyclistic_trips$ride_duration <- round(cyclistic_trips$ride_duration, digits = 1)


head(cyclistic_trips)
str(cyclistic_trips)
colnames(cyclistic_trips)
summarise(cyclistic_trips)


cyclistic_trips <- cyclistic_trips %>%
  select(-c(start_station_id, end_station_id, ride_id))
cyclistic_trips <- cyclistic_trips %>%
  select(-c(season))
  str(cyclistic_trips)
write.csv(cyclistic_trips, file = "cyclistic_trips.csv", row.names = FALSE)
```

Real Estate Data

```
# Data sourced from zillow.com at https://www.zillow.com/research/data/ #


install.packages("dplyr")
library(dplyr)


Zillow_Chicago_multifamily <- Zip_ZORI_AllHomesPlusMultifamily_Smoothed
Zillow_Chicago_multifamily <- Zillow_Chicago_multifamily %>%
  filter(Zillow_Chicago_multifamily$MsaName == "Chicago, IL")


head(Zillow_Chicago_multifamily)


write.csv(Zillow_Chicago_multifamily, "chicago_rental_data.csv", row.names =
FALSE)
```

Crime Data

```
# Chicago crimes datasets can be found at https://data.cityofchicago.org/Public-Safety/Crimes-2022/9hwr-2zxp/data #


library(dplyr)

library(tidyverse)


colnames(Crimes_2022)


Crimes_2022 %>%

select(-c("ID", "IUCR", "Case Number", "Block", "Ward", "District"))


head(Crimes_2022)

write.csv(Crimes_2022, file = "chicago_crime_2022_data.csv", row.names = FALSE)
```

Analysis

Membership Dynamics

- **Seasonality:** The service is very seasonal with far less use by casual users during the famously cold Chicago winter months. While work commutes remain at a greatly diminished level during winter months, casual ridership collapses at other means of transport are preferred and few tourists are present.
- **Strong Initial Adoption by young IT professionals in North River district**: Heat maps demonstrate that the membership service achieved greatest successful marketing of membership among a young demographic working predominantly in a few tech companies in the Northern River District. It should also be considered that these customers have relatively short bike rides (<15mins) and value the convenience of the service. Cyclistic could drive profitability by increasing fees for the service.
- **Social Discrimination against disadvantaged neighborhoods?** There are few members outside River Norther and the wealthy northern suburbs, and little to suggest that there is adoption by disenfranchised socioeconomic groups living outside of the Downtown and Northern suburbs. Different marketing strategies may be employed after further study of their travel habits and transport preferences. The company would need to adapt its marketing to avoid being accused of racism and social discrimination
- **Limited Potential for Tourist Market:** The relative light transport between traditional tourist destinations such as the northern beaches and Daley Park suggests that the service is not being adopted to a meaningful extent by tourists.

Member Acquisition Strategy

- The client acquisition strategy should address 3 categories of potential members:
    - **Corporate Membership as Employee incentives:** Chicago is a growing tech hub so this is a large addressable market and the offices in River North are the most expensive in the city. The strong adoption by wealthy IT service-sector professionals of the Cyclistic service in River North suggests that the persona of the brand resonates with values shared by this valuable demographic. Many of these companies would be happy to include Cyclistic among their employee incentives and, since some have thousands of employees, can drive huge increases in membership e.g. sell memberships at discounts based on 10's/100's of memberships
    - **Campaign to Convert Casual Riders to Annual Memberships:** There is a strong rationale for forecasting further growth among this group and an attractive ROE on marketing compared to tourist segment or other socio-economic groups. The campaign should focus on the convenience of the service e.g. point-to-point travel, and attractiveness of brand identity rather than offering a value proposition compared to other forms of transport. With greater potential for weekend use in summer months, the campaign will get better ROE in the spring months.
    - **Other Segments should be de-prioritized.** Cyclistic should focus on maximizing market share in their core demographic. The ROE on other segments will be lower than the River North IT community.

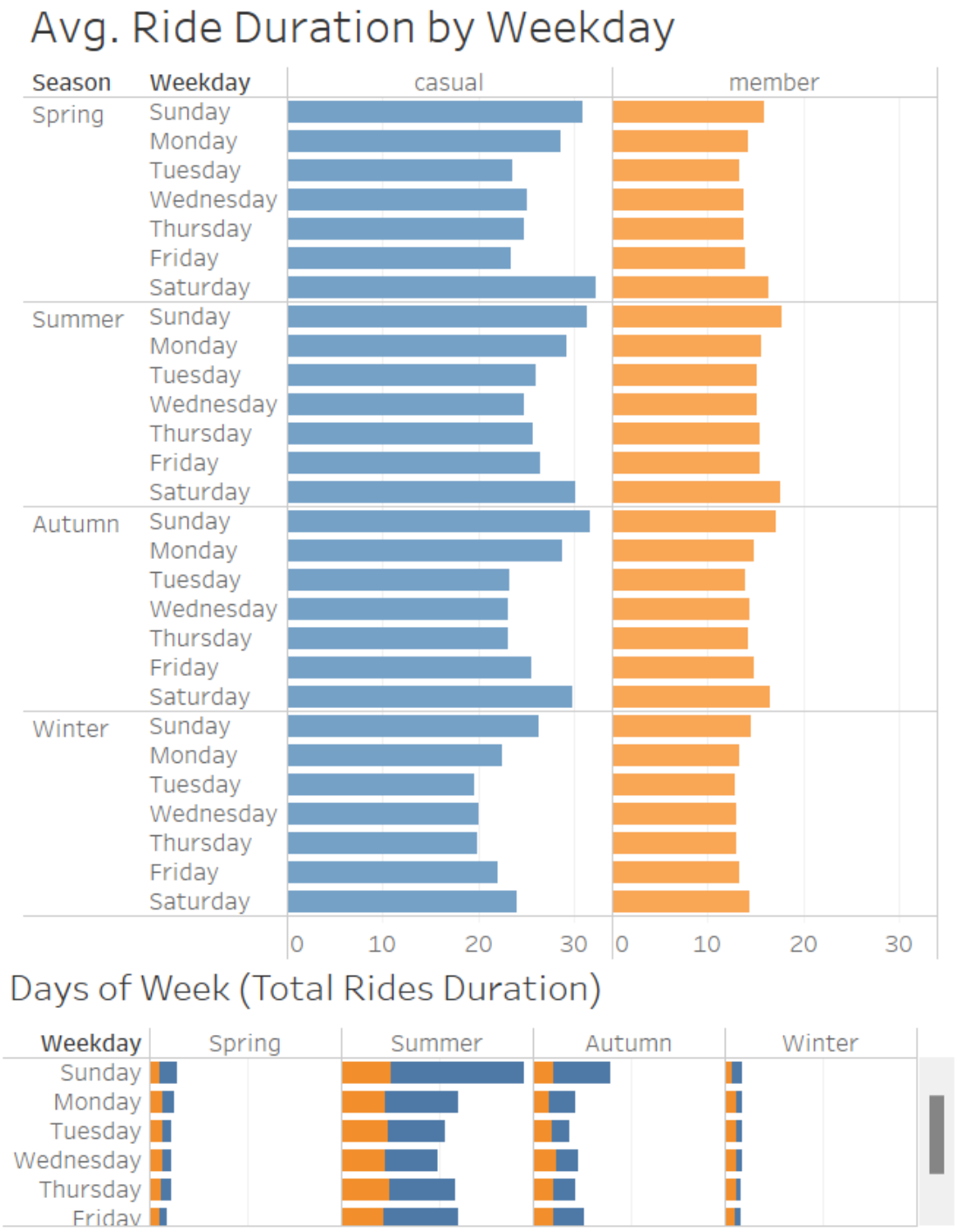
Application of Digital Media

- **Branding**: Online media such as Instagram and TicTok offer effective platforms for branding to the target customer segment. Facebook and linkedin are appropriate digital platforms for connecting and reinforcing brand identiy with corporate customers.
- **Customer Segments & Product Positioning**: Analysis indicates that the highest returns on equity for marketing will be expanding the membership base amongh young professionals in the River Norther district. Because of the concentrated geographical area, as well as digital media, effective off-line media will also be an important part of building brand awareness and interest of Cyclistic.

Visualizations Supporting Key Findings

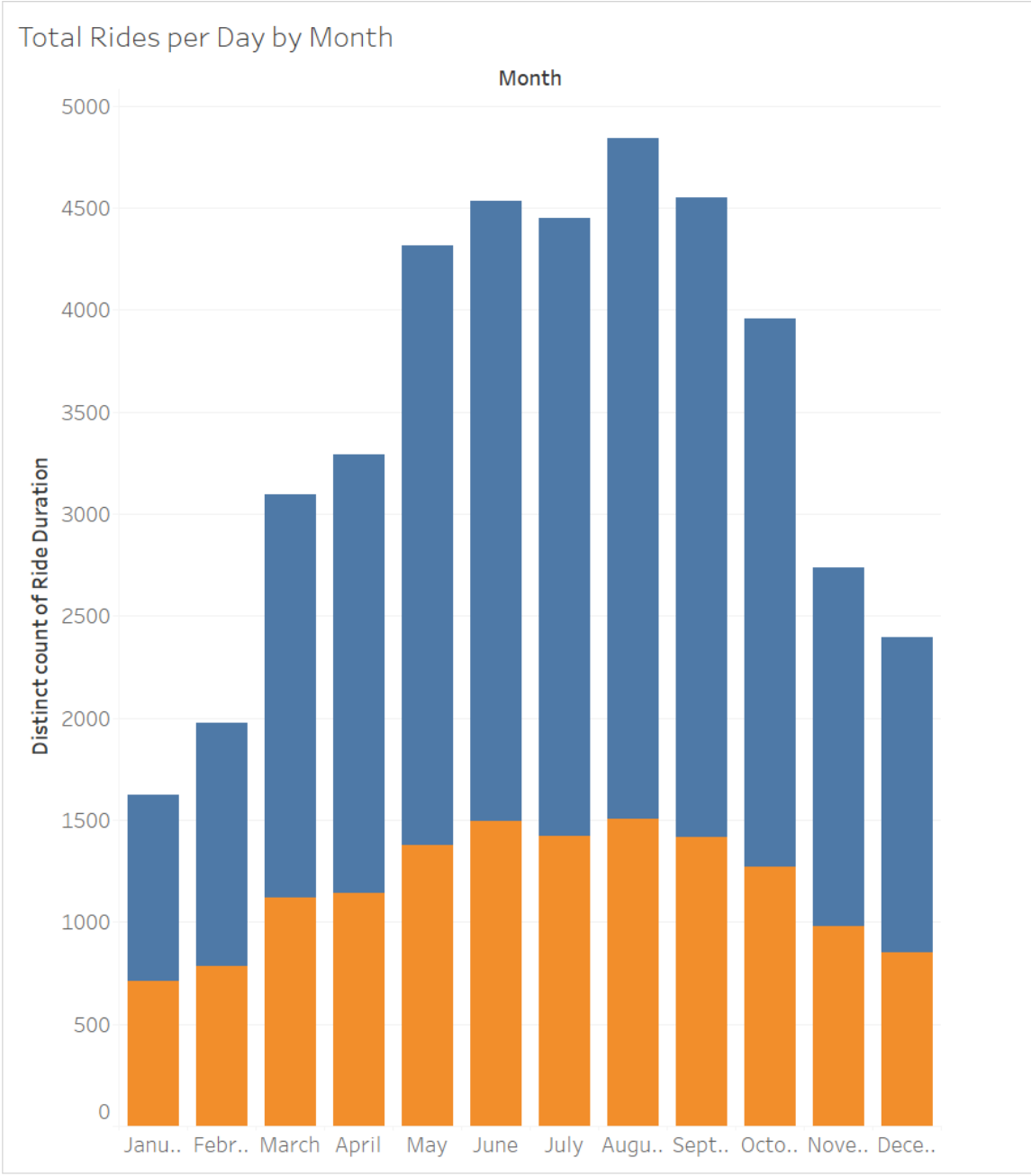
Tableau was used for the analysis and presentation of data for sharing. Interactive slides can be accessed here

Ride Dynamics

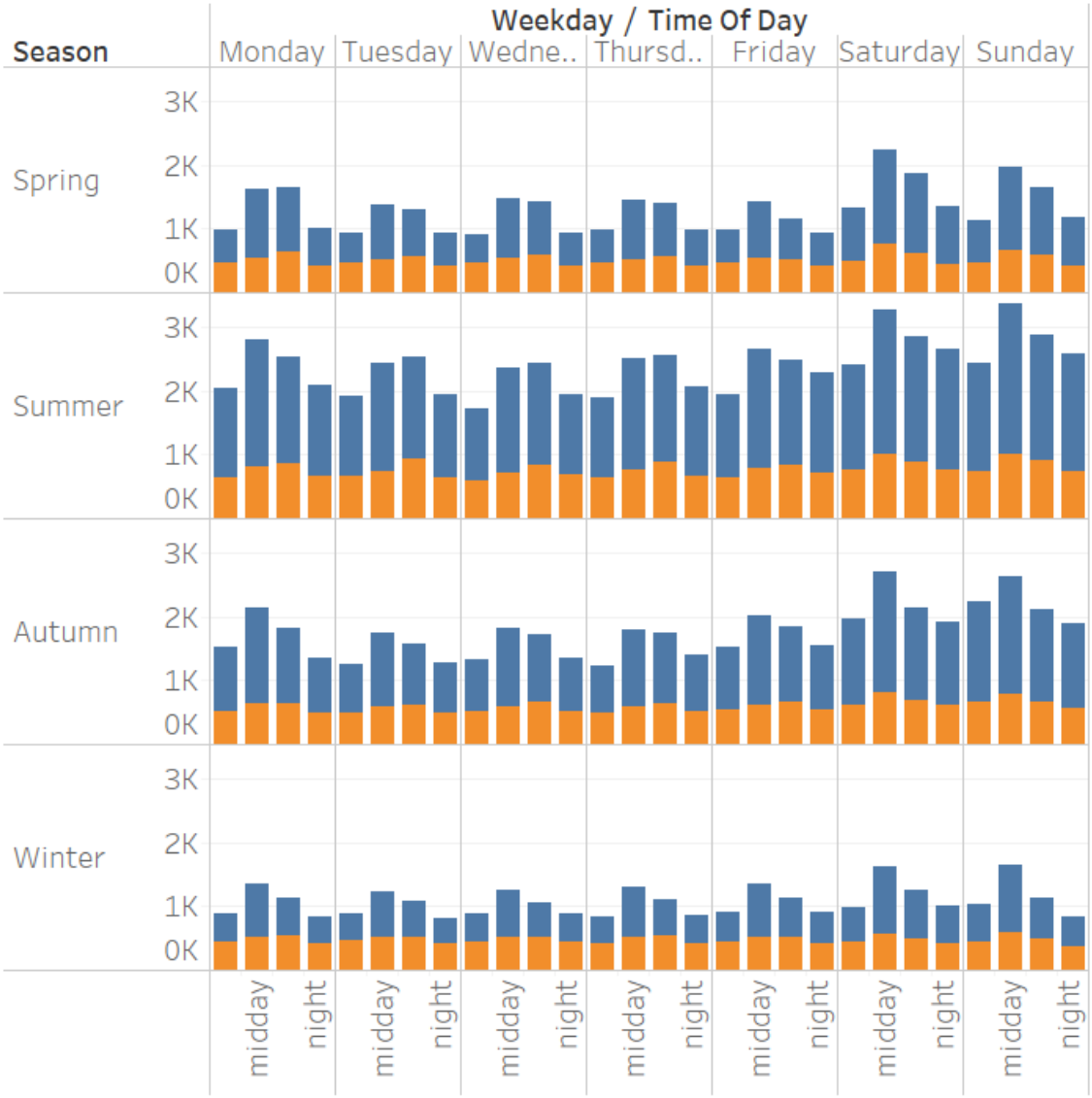


Members are very consistent users of the service, with a lot of users active on 10-15min rides in morning and evening, presumably on commutes to and from the office. Casual riders use the

service for longer rides (almost x2 times the members average) and both groups are more active during warmer months. The relative disproportionate correlation in increase between the casual riders and member use in summer months suggests that some tourists use the service as casual riders.



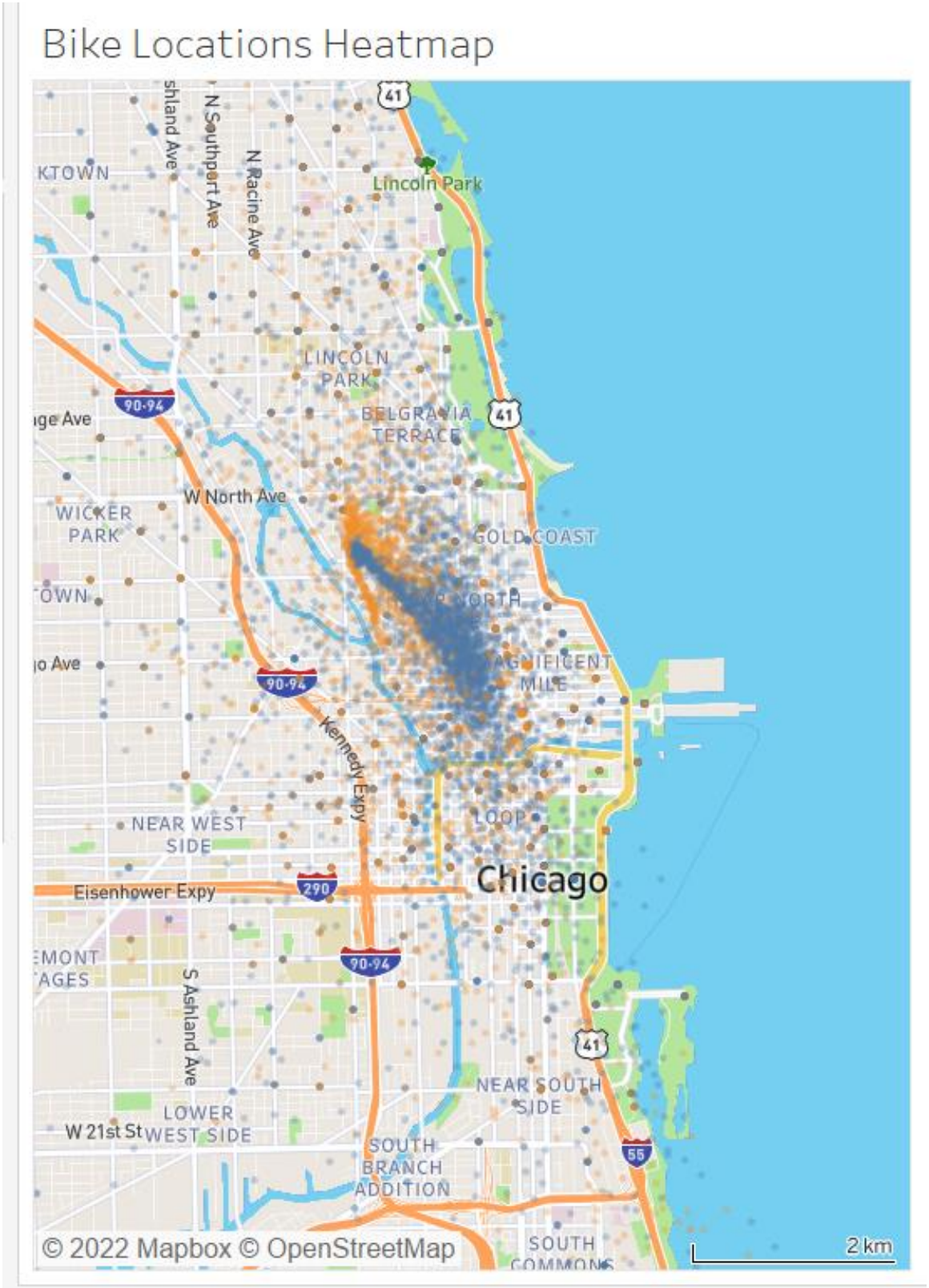
Rides by Time, Day, Season



Rides per Day by Time, Day, Month

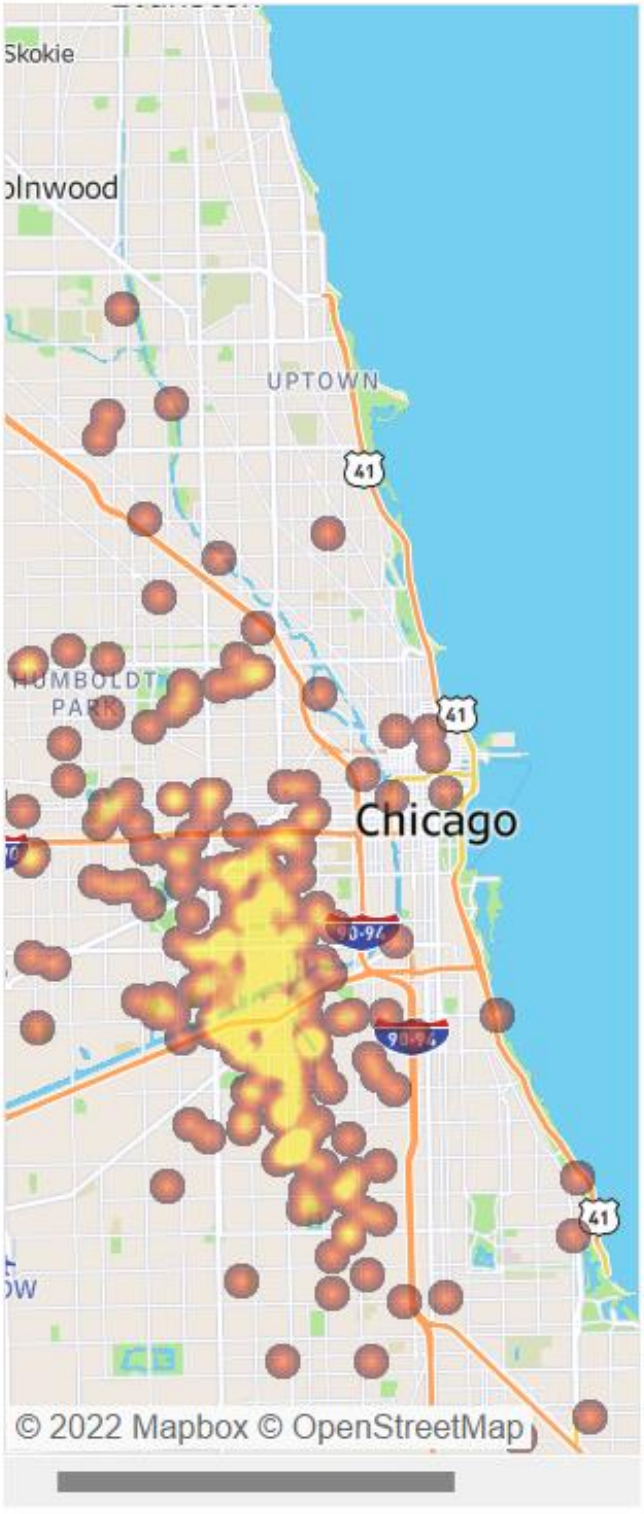
Geography

Usage is highly concentrated in the River North area of Chicago. The district is the location for the most expensive offices and apartments in the city and the lowest crime rate.



Rent 1,000 3,500

Crime Map



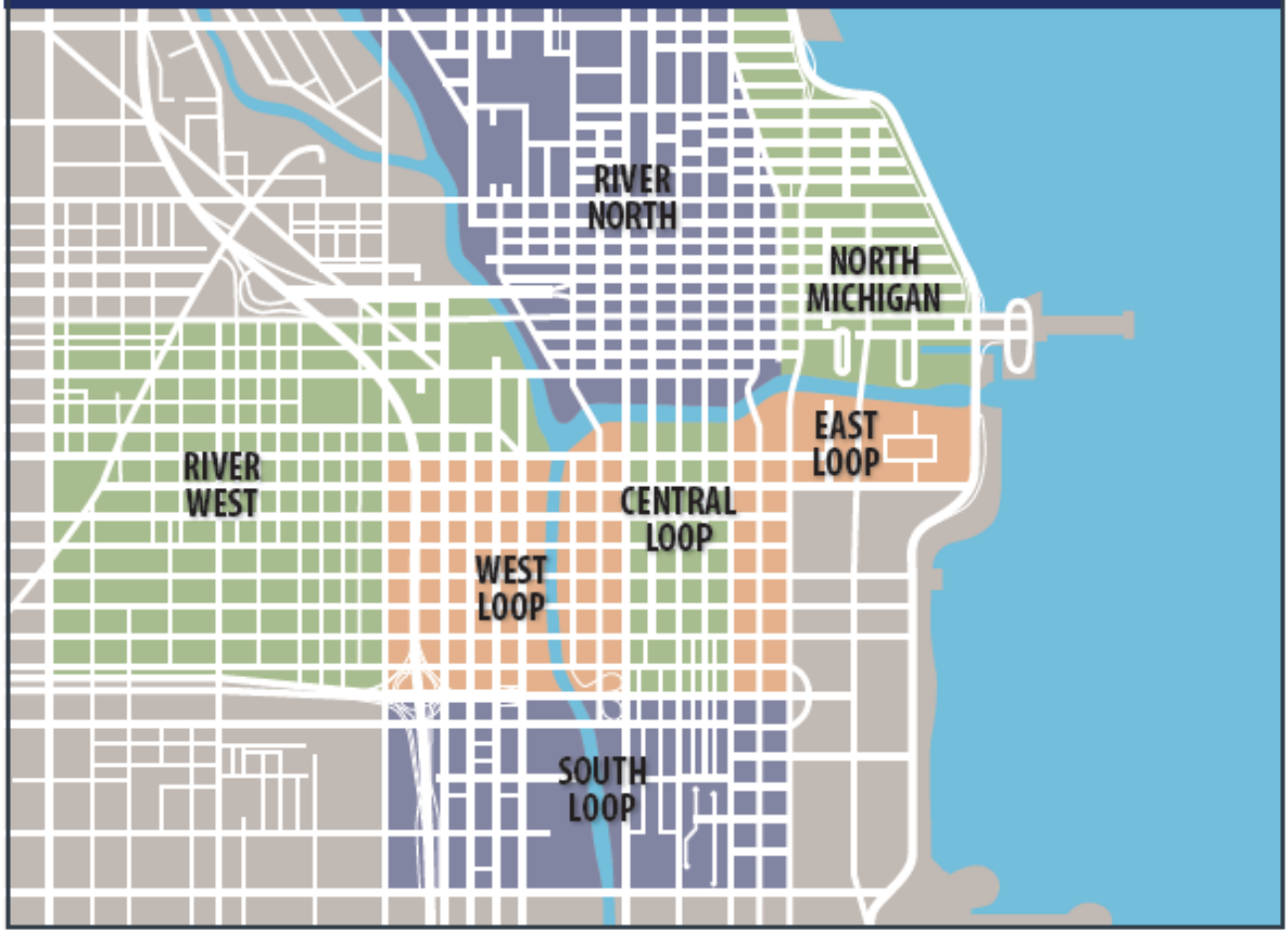
Avg. Rents by Postcode

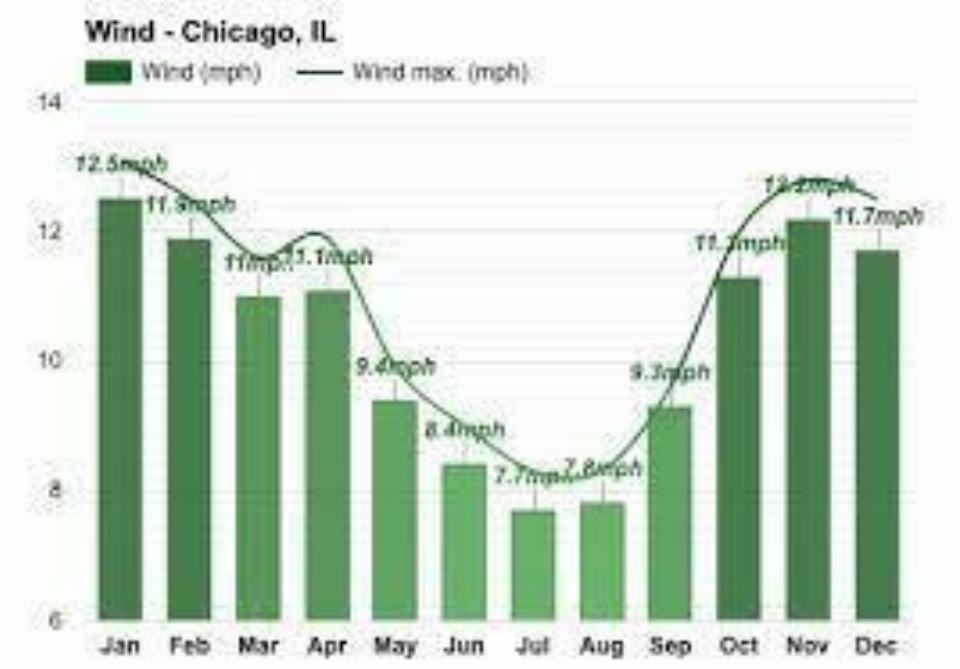
Chicago CBD Submarket Snapshot | Q1

A look at basic metrics in each downtown submarket

Submarket	Total RBA (SF)	Direct Available Space (SF)	Sublet Available Space (SF)	Total Available (SF)	Vacancy Rate			Leasing Activity (SF)	Net Absorption (SF)	Average Asking Rent
					Overall Vacancy	Class A	Class B			
CBD	177,331,557	25,016,225	3,168,131	28,184,356	10.9%	11.7%	10.6%	4,300,500	440,654	$27.30
West Loop	56,012,481	9,384,172	972,028	10,356,200	12.0%	13.0%	10.1%	2,587,327	214,221	$29.56
Central Loop	47,525,675	6,099,807	1,017,894	7,117,701	10.8%	12.6%	8.9%	699,915	21,246	$24.66
East Loop	27,480,707	3,858,298	628,919	4,487,217	13.1%	11.0%	21.1%	336,097	172,987	$23.18
River North	18,650,369	2,082,146	357,972	2,440,118	8.2%	5.6%	10.4%	211,657	-70,311	$31.99
North Michigan Ave	16,971,658	1,413,890	114,165	1,528,055	7.5%	9.7%	5.4%	245,812	36,442	$24.79
River West	7,222,166	1,867,749	67,468	1,935,217	10.6%	14.5%	9.7%	214,593	75,429	$31.98
South Loop	3,468,501	310,163	9,685	319,848	8.6%	N/A	12.5%	5,099	-9,360	$18.70

CBD Submarket Map





In view of the acutely cold weather in the Great Lakes region during the winter and spring months, there are weather constraints on the attractiveness of bike share for over half of the year. Corporate memberships are annual which make them more attractive from a client acquisition cost perspective and profitability perspective i.e. much less marketing cost.

Additional Date and Research

Acquiring casual riders and members during the summer months can be achieved by summer memberships and weekend passes but the marketing team might want to avoid confusing the message since adoption by a highly attractive client group has already been demonstrated and should be targeted as a priority. While campaigns are developed and run for this segment, further data analysis on the other market segments could help to define future marketing strategies for other client segments.