Cyclistic: Strategies on Increasing Annual Membership

Huy Nguyen

01

Introduction

Context

01 Cyclistic makes bike-share more inclusive to people with disabilities and riders who can't use a standard two-wheeled bike

02 Majority of riders opt for traditional bikes; about 8% of riders use the assistive options

03 Users are more likely to ride for leisure, but about 30% use them to commute to work each day.

04 Until now, Cyclistic's marketing strategy relied on building general awareness and appealing to broad consumer segments.

Business Objectives

Moreno believes that maximizing the number of annual members will be key to future growth.

Data: from the past 12 months. The data has been made available by Motivate International Inc

Task: see how annual members and casual riders use Cyclistic bikes differently

Recommendations:
Short-distance discounts
Docked bikes
Poverty areas

02

Data

Data Overview

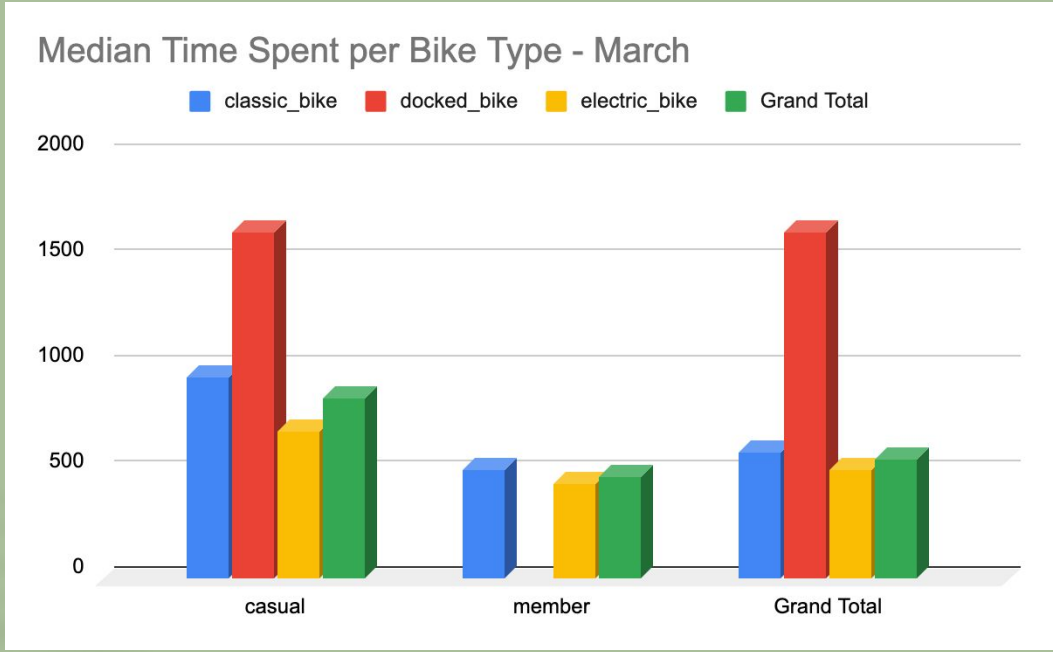
1. There are 12 data sets from December 2021 to November 2022.
2. There are 13 columns for each data. However, we are only interested in the following:
    - rideable_type: the bike type (classic,docked,electric).
    - start_lat & start_lng: starting latitude and longitude of each trip.
    - end_lat & end_lng: ending latitude and longitude of each trip.
    - started_at: time and date that the trip started.
    - ended_at: time and date that the trip ended.
    - member_casual: whether the person is a casual or a member. This is our response variable.
3. There are several issues with some of the data:
    - The dates and time are not in the correct class. We will convert it to POSIXlt.
    - The start and end dates are sometimes not in correct chronological order.
    - For rideable_type and member_casual, there are missing values.
    - Some data is too large for Excel to handle, so we will mostly use Tableau and R. Excel will be used for February and August.

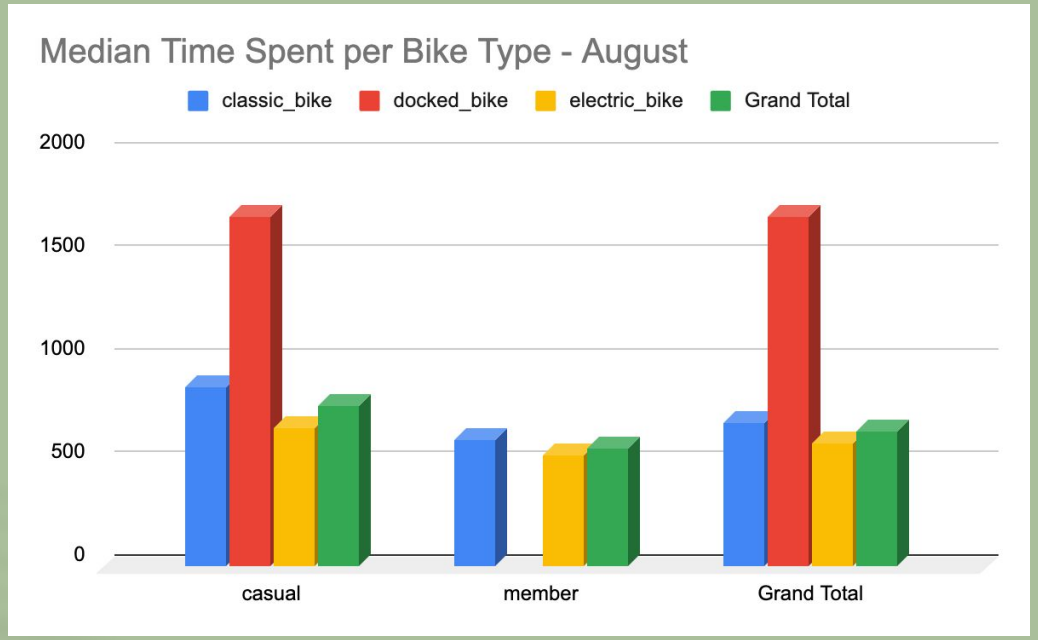
Data Cleaning and Transformation

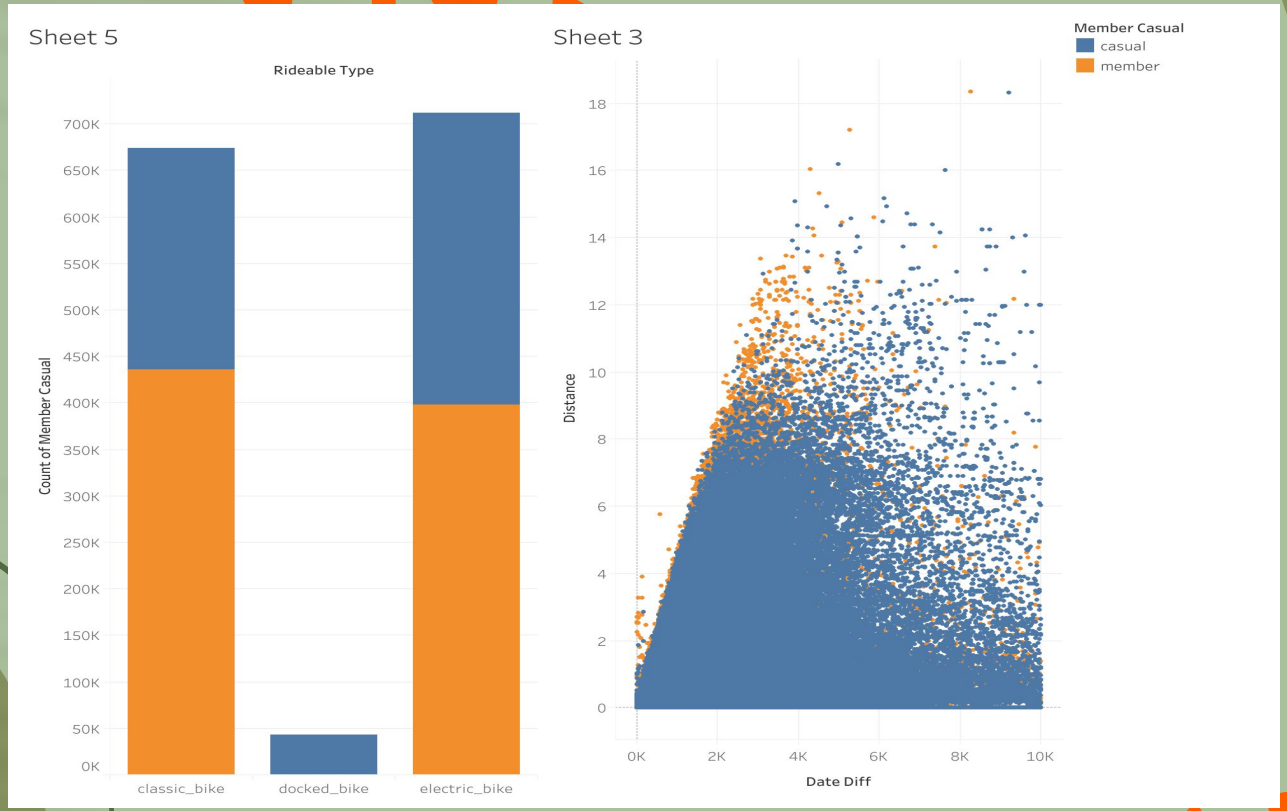
- We will filter in R to get only the aforementioned columns.
- Then, we convert started_at and ended_at to the right format. We make sure these columns have no NAs. Now, we take ended_at - started_at to have the time travelled. We call this date_diff. We can now remove started_at and ended_at. We will also filter to have date_diff strictly positive.
- We remove NAs for rideable_type and member_casual, then convert them to type factor in R.
- Now, given that each month can have very different number of data points, we will row bind the data with minimum (January), median (October), maximum (June) to have dataframe new_data with more than 1 million rows.
- We will use this data in Tableau. There, we will use haversine formula to calculate the distance travelled for each trip. We will also combine latitude and longitude columns to have our starting coordinates and ending coordinates.
- For the data of March and August separately, we will pivot through Excel. This is to have more data points and show general trends that we will discuss later.

03

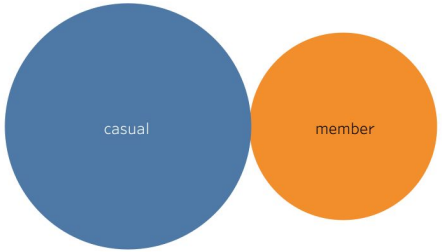
Analysis


Median Time Spent per Bike Type - March
classic_bike
docked_bike
electric_bike
Grand Total
2000
1500
1000
500
0
casual
member
Grand Total


Median Time Spent per Bike Type - August
classic_bike
docked_bike
electric_bike
Grand Total
2000
1500
1000
500
0
casual
member
Grand Total

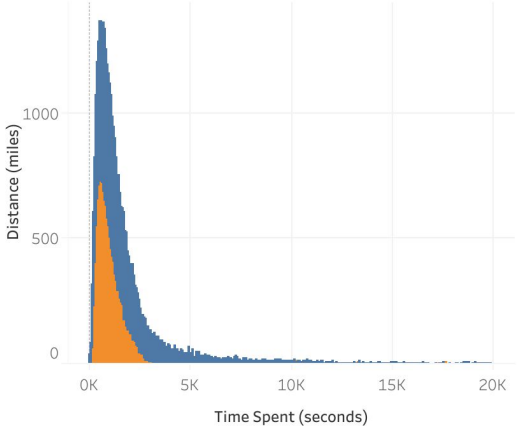

Sheet 5
Rideable Type
Count of Member Casual
700K
650K
600K
550K
500K
450K
400K
350K
300K
250K
200K
150K
100K
50K
0K
classic_bike
docked_bike
electric_bike
Sheet 3
Member Casual
casual
member
Distance
18
16
14
12
10
8
6
4
2
0
0K
2K
4K
6K
8K
10K
Date Diff

Average time Vs. Membership

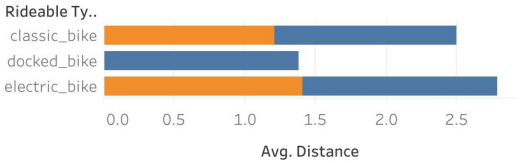


Distance by membership

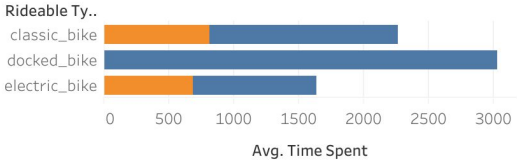
Time spent vs. Distance



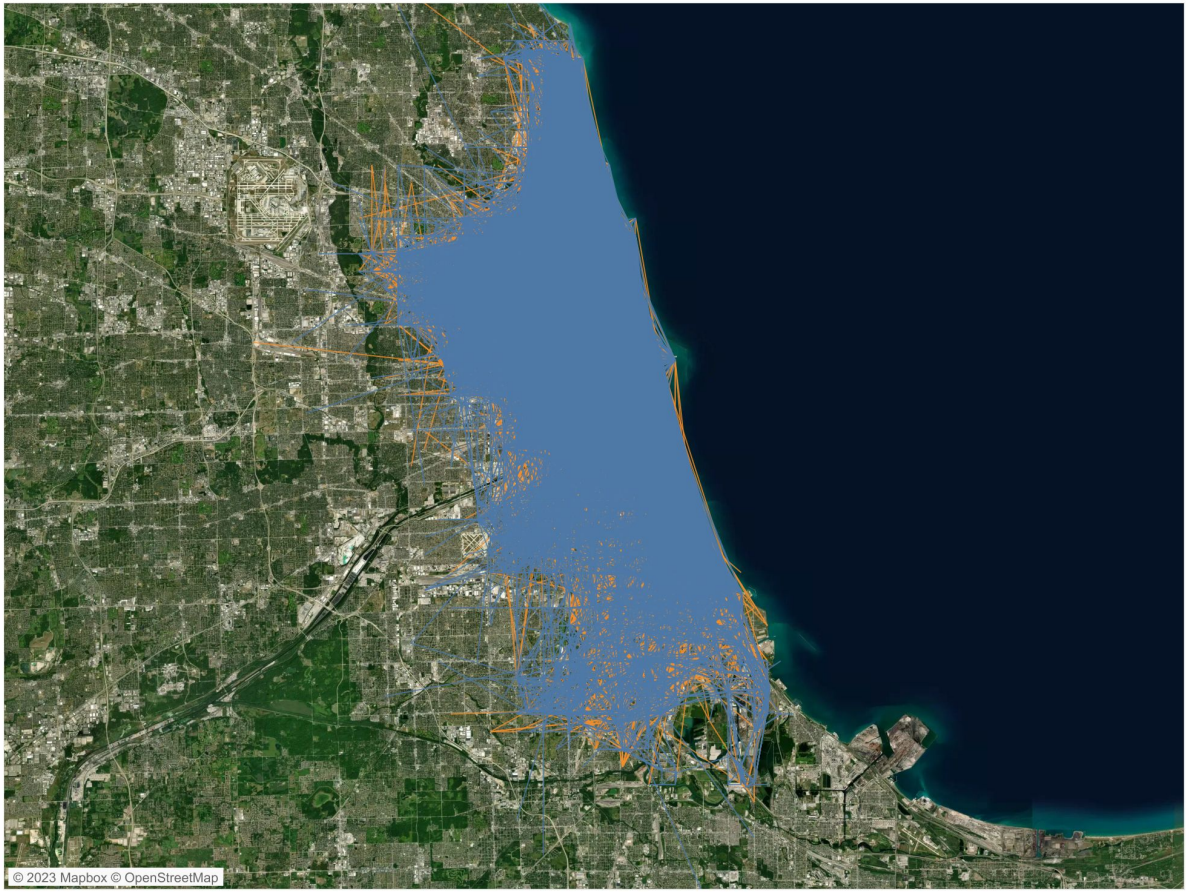
Average distance by bike type

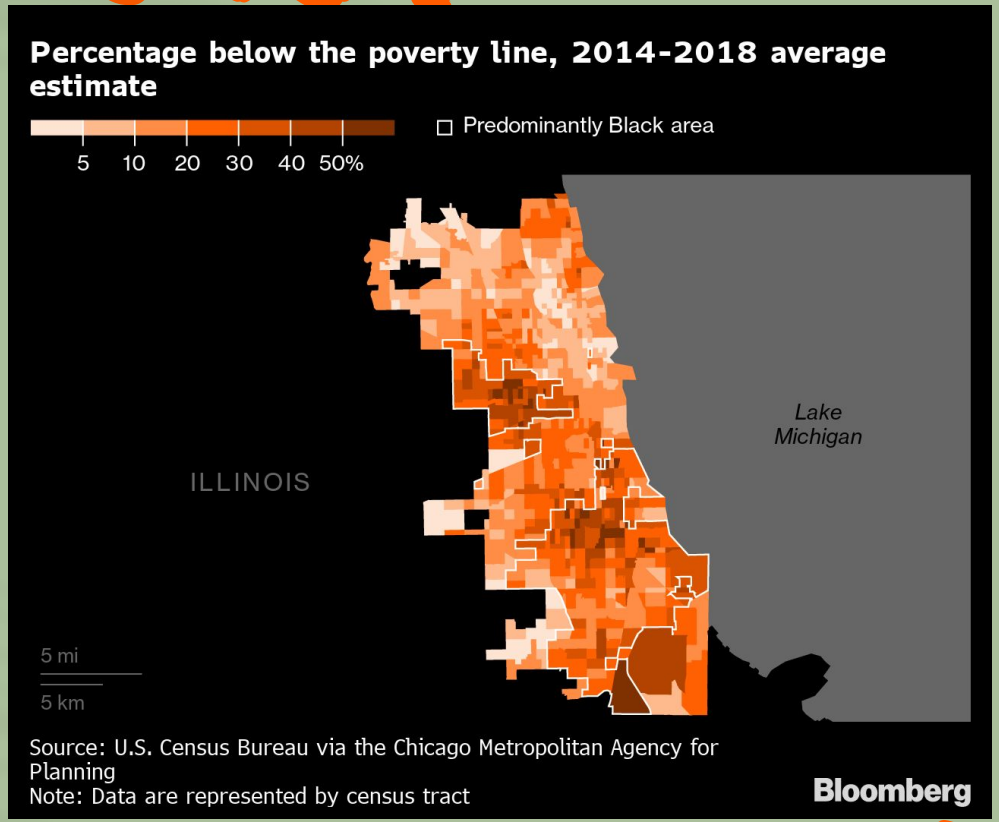


Average time spent by bike type



Paths of Bicycles in Greater Chicago




Percentage below the poverty line, 2014-2018 average estimate
5
10
20
30
40
50%
Predominantly Black area
ILLINOIS
Lake Michigan
5 mi
5 km
Source: U.S. Census Bureau via the Chicago Metropolitan Agency for Planning
Note: Data are represented by census tract
Bloomberg

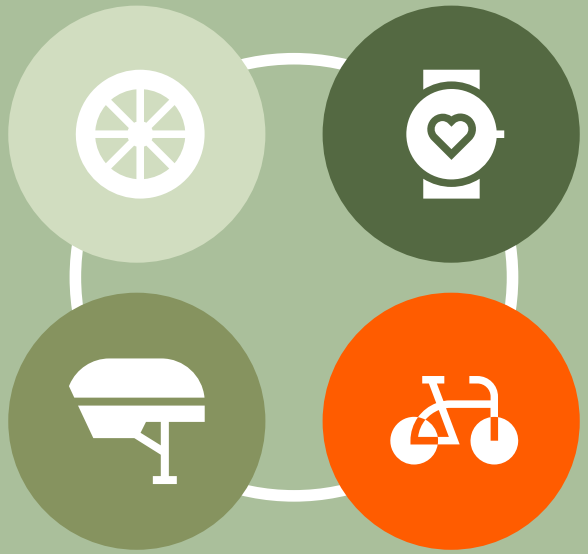
Findings

- There are no member for docked bikes, which also have the least counts.
- Across all bike types, members tend have less travelled time.
- Most of the very large travel time and distance are made by casual. In fact, one travelled all the way from/to Montreal, Canada!
- At around 2000 seconds (approx. 30 minutes) to 4000 seconds (approx 1 hour), many members do travel at long distance.
- The distribution of distance for member and casual are very similar. The box plot shows that most people (75%) do not travel over 3 miles.
- When looking at the paths travelled for member and casual. We see that in low-income areas (South Chicago) there seem to be more paths taken by members.
- Still, the majority of the trips by both members and casuals travelled east of Des Plaines river, from Evanston to the Illinois-Ohio border.

Logistic Model

Testing Accuracy:
64%

All Factors are
Significant



30% Test
70% Train
Data: new_data

Full Model

04

Conclusion Recommendation

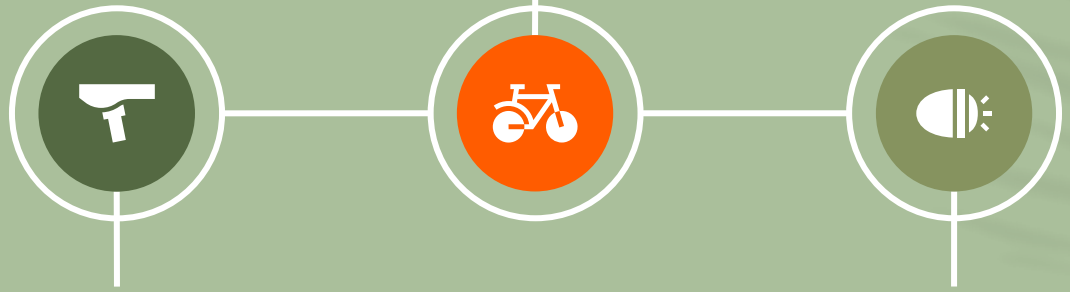
Conclusion

Members tend to spend less time on each trip even if they may travel more distance.

People from low socioeconomic areas may be more likely to be members

Docked bikes have no members.

Recommendations



Membership for docked bikes, more docked bikes in downtown Chicago

More bikes in low-income areas, discounted membership

Targeted membership advertisement for trips with short travel time

Thanks

CREDITS: This presentation template was created by **Slidesgo**, including icons by **Flaticon**, and infographics & images by **Freepik**