

# Cyclistic

- Anar Seyf (anar.seyf@gmail.com)
  - October 2021
  - Capstone project | Google Data Analytics course #8 | Coursera
- 

## Goal

The objective of the report is to show differences between **Members** (those who purchased an annual membership) and **Casual** riders (those who pay for single rides or daily passes). Every part of the analysis shows (**TODO - synonym**) those categories side-by-side.

---

## Data

The **main data source** of the report is a 12-month window (October 2020 to September 2021, inclusive) from the dataset provided by Divvy ([link TODO](#)). (See data cleaning notes in the Appendix.) The data consists of over 5 million unique ride records with these columns:

- unique ride ID;
- bicycle type (classic, docked, electric);
- ride start and end time;
- start and end docking station (ID and name);
- start and end latitude/longitude;
- status (member or casual)

The **secondary data source** is the Chicago daily weather dataset from NOAA ([link TODO](#)) for the same time period.

The 5 million individual records are condensed into hourly, daily, weekly, and monthly **aggregates**, grouped by status (member vs casual).

Throughout the report, more emphasis is given to *ride count*, as the main indicator of activity volume, than to ride *ride duration*.

---

## Analysis

### 1. Members ride more often, casual users ride for longer.

An average ride is about **28 minutes** long for casual users, **14 minutes** long for members. Members make an average of **7,479** rides per day, casual users **6362** rides per day. (The number of distinct riders is unavailable in the data, so we can only refer to the overall population.)

Ride volume is lowest in February and highest in July-August. Casual ride volume exceeds that of Members in the summer months only.

(The month of February shows an abnormal peak in ride duration. It may be partially attributable to low ride volume and thus higher variance in the data, but is otherwise not readily explained here.)

Table 1: Average ride duration (minutes)

Status	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Casual	24	24	22	21	33	27	28	29	28	27	25	24
Member	13	13	12	13	18	13	14	14	14	14	14	13

Table 2: Average daily rides (count)

Status	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep
Casual	4,573	2,888	956	575	354	2,677	4,490	8,159	12,147	14,044	13,131	11,962
Member	7,676	5,614	3,220	2,502	1,380	4,592	6,582	8,705	11,754	12,058	12,431	12,863

## 2. Members ride all week, casual riders prefer weekends.

Members' usage remains nearly flat, falling on Sunday. Casual users do most riding starting on Friday and through the weekend, peaking on Saturday.

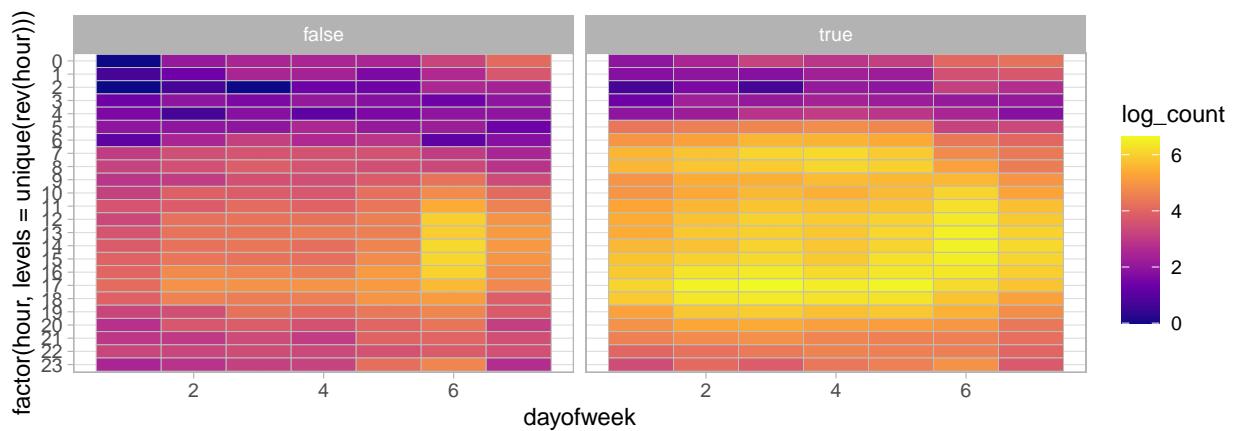
Table 3: Average daily duration (minutes)

Status	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Casual	28	25	24	24	26	30	32
Member	13	13	13	13	13	15	16

Table 4: Average daily rides (count)

Status	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Casual	5,047	4,763	4,871	5,104	6,421	9,923	8,429
Member	7,107	7,701	8,002	7,807	7,666	7,550	6,513

## By weekday



## Hourly

**TODO** — duration by weekday.

### 3. Seasonal and Weather

**TODO** — eliminate 0 rain/snow circles.

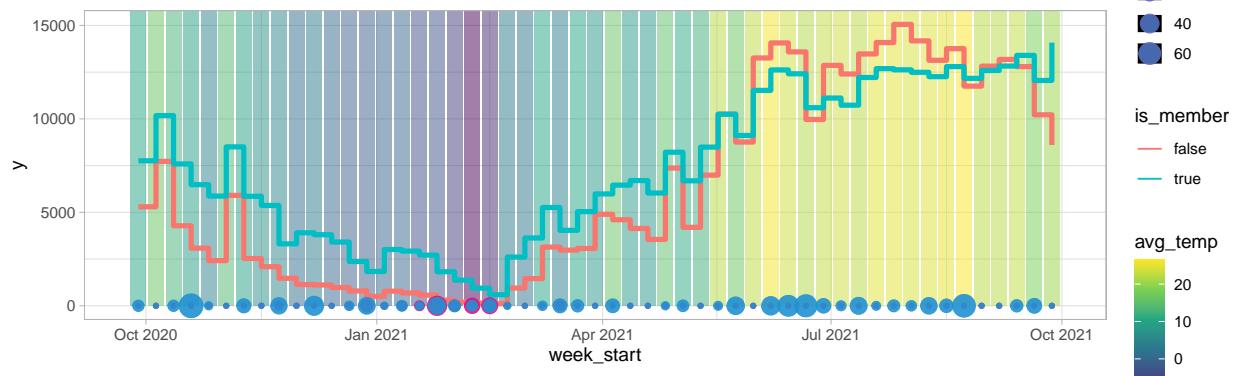


Table 5: Correlation: weather and ride count

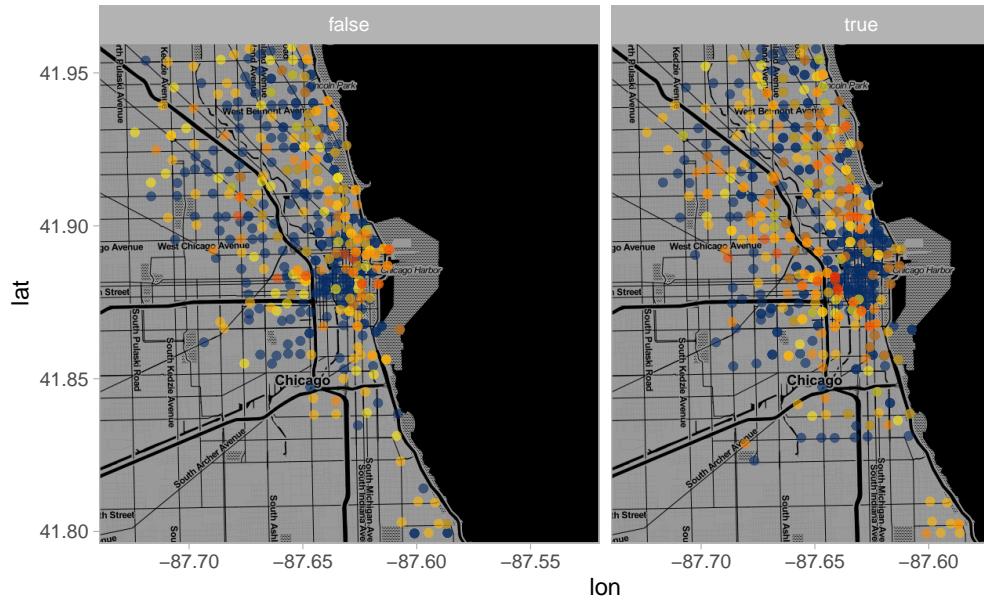
Status	Temperature	Rain	Snow	Wind
Casual	0.81	0.00	-0.20	-0.20
Member	0.91	-0.01	-0.29	-0.22

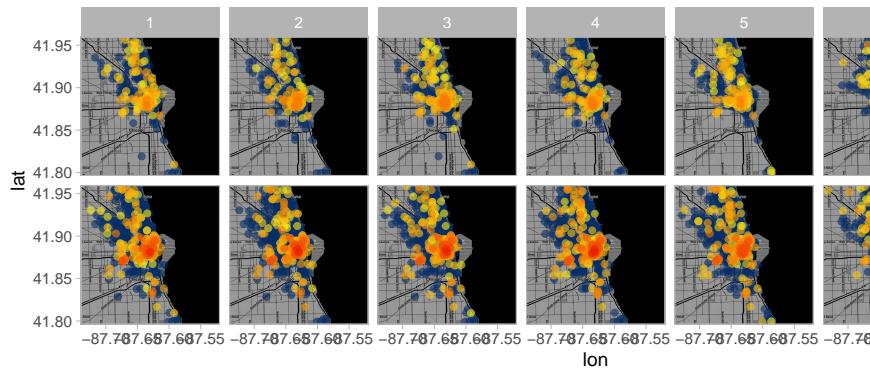
### 4. Geographical

#### Stations

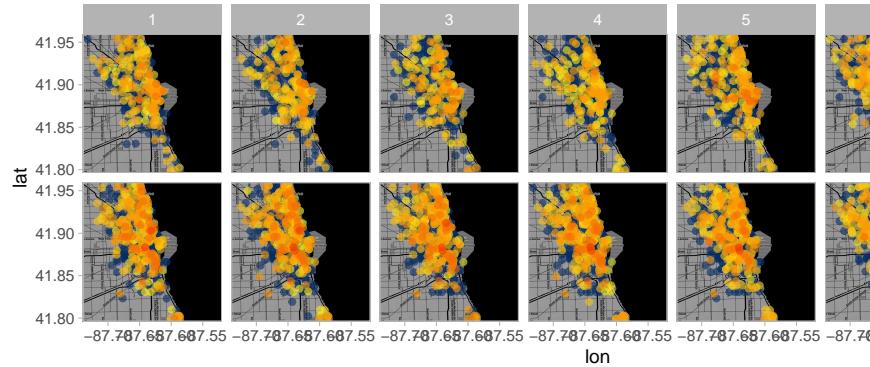
1288 stations. Following: **casual vs members**, **8am vs 4pm**, **Monday-Friday**.

Each dot is a bike docking station. Stations highlighted **yellow-red** have more arrivals than departures for the given hour, suggesting an influx of bike traffic at that location.





**8am, Monday-Sunday, casual vs members**



**4pm, Monday-Sunday, casual vs members**

## Appendix

### A. Data cleanup

### B. Links

Data sources (incl. Google Maps, Stamen).

### C. Full data (weekly?)

Monthly by weekday

**Weekly Counts and Average Duration** Full year, weekly.

Monthly by bike type

**Temperature**

**Rain**

**Wind**