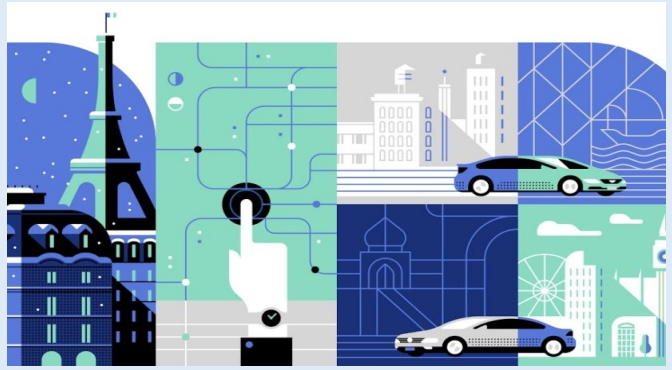
Uber Data Analysis Project

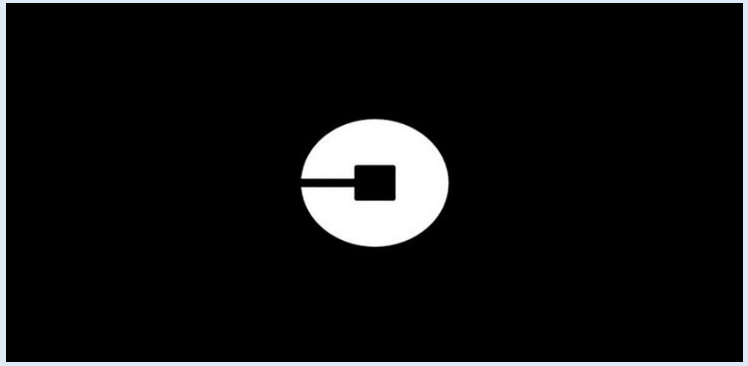
By

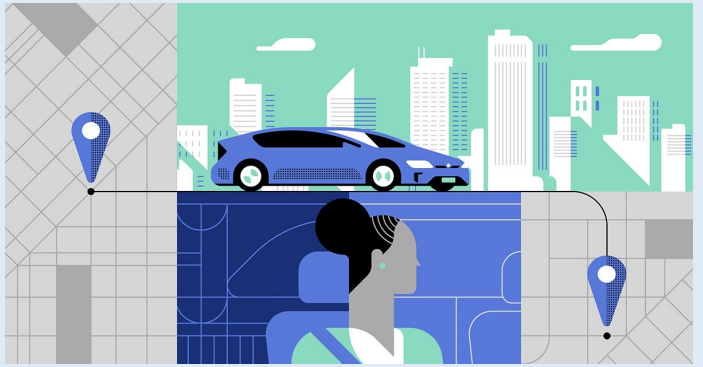
Juan Moctezuma

What is Uber?

- Platform that allows the user (passenger) to request a ride or a service to the closest Uber partner/driver, but the driver cannot select the user or their respective destination.
- The Uber mobile app determines the cost of fare once the trip is completed.
- There are several types of service that the rider can select. Each city or state might have different services.
- San Diego, CA has the following: UberASSIST, UberBLACK, UberESPAÑOL, UberLUX, UberPOOL, UberSELECT, UberSUV, UberUCSD, UberX and UberXL.







What was the project about?

- Launching a 2-question survey from May to December 2017 while working as an Uber Partner/Driver regarding:
  - Where the passenger(s) is/are originally from.
  - What do they like the most about San Diego, CA.
- Recording and compiling the passenger's opinion in an Excel file.
- Understanding the attributes that make San Diego an appealing location for many individuals.
- Combining the passengers' information recorded on the Uber Partner's mobile app with data collected during trips.
- Creating visuals displaying relevant data not only from the survey but from other observed patterns as well!

What happened after 2017 was over?

- All data was reviewed and then pivot tables, charts and graphs were created out of the passengers' information dataset.
- Passengers' responses were labeled as Not Available (N/A), Other (O), People (P) and Weather (W).
- Extra tables and graphs containing climate-related data were created. The reason being was that most passengers voted for weather as San Diego's best quality!
- Since other patterns were reflected on the data, I decided to include extra information, such as the total number of services per category performed during 2017.

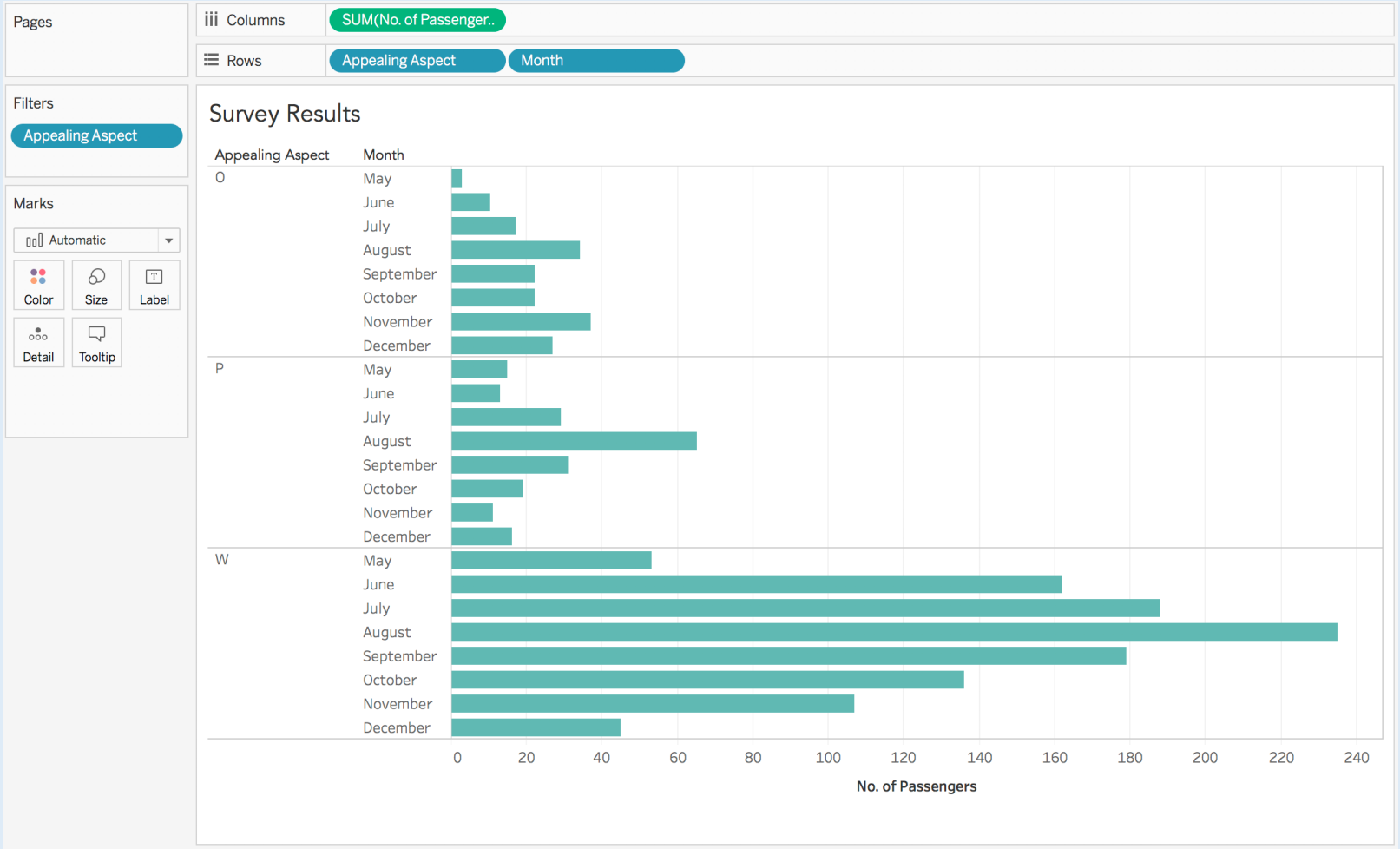
Survey Results

- The pivot table shows that during this 8-month period, 51.2% of the passengers voted for weather. The result came from the following computation: 1105 ÷ 2159 x 100%. Therefore, most riders loved San Diego's weather!
- The number of votes represent the counted number of places of origin of passengers as given by them.

For all categories:	No. of Passengers
N/A	683
Other	172
People	199
Weather	1105
Total No. of Passengers:	**2159**

Survey Results (Continued)

- The following image shows the same outcome, however this chart was produced by Tableau Public using the same compiled data.



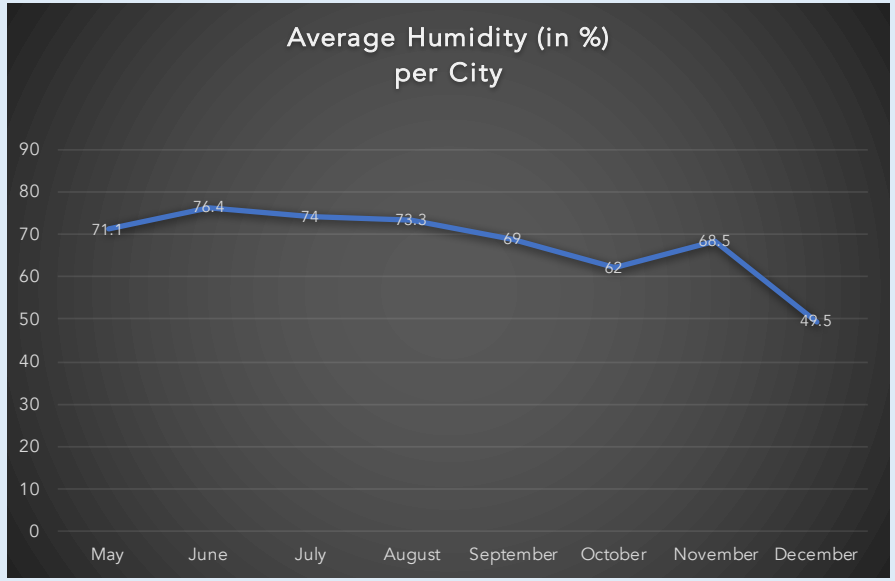
Survey Results (Continued)

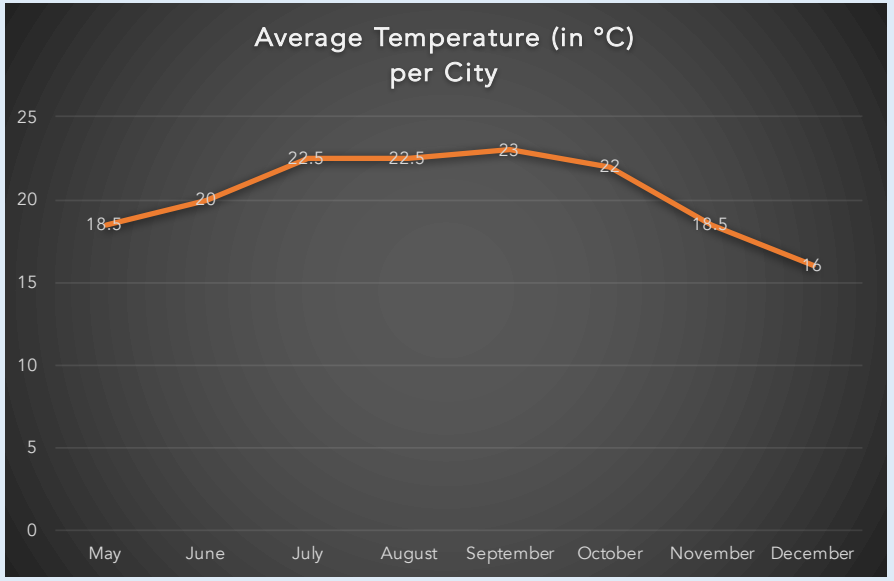
- The pivot table below contains a small portion/number of voters listed with their respective places of origin. Where each place of origin equals one vote. That image only contains a small section of the voters (places with 19 or more votes). The complete table is located on the Survey Results section from the Excel worksheet titled Uber Data Analysis Project.
- The number of votes represent the counted number of places of origin of passengers as given by them.
- Creating separate tables for each type of territory would be unpractical because some rides have people from different cities, states/provinces or countries, during the same trip.

Individuals from the following places voted for:	No. of Passengers
⊟ Weather (Total)	1105
San Diego, CA	68
Texas, US	28
California, US	28
San Francisco, CA	25
Chicago, IL	24
New York, US	20
New York City, NY	19

Weather Results

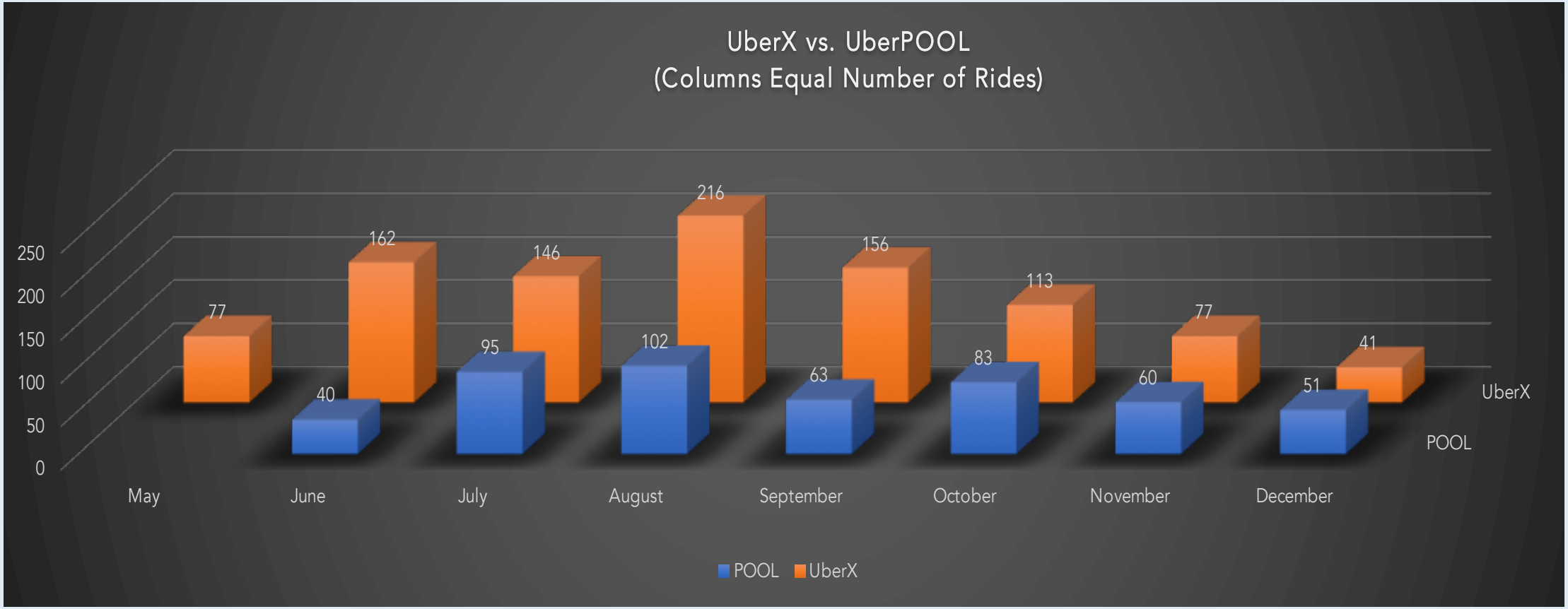
- San Diego, CA has a cold semi-arid (BSk) climate according to Köppen-Geiger classification. Its average temperature was 20°C or 68°F, humidity levels were not excessive and had barely any precipitation.
- Most of the cities, states and countries being compared to San Diego, CA, were not classified as cold semi-arid. In fact, the majority of the places individuals voted for (meaning where they originate from) have harsher conditions.
- The tables below contain the weather results from San Diego, CA during 2017.





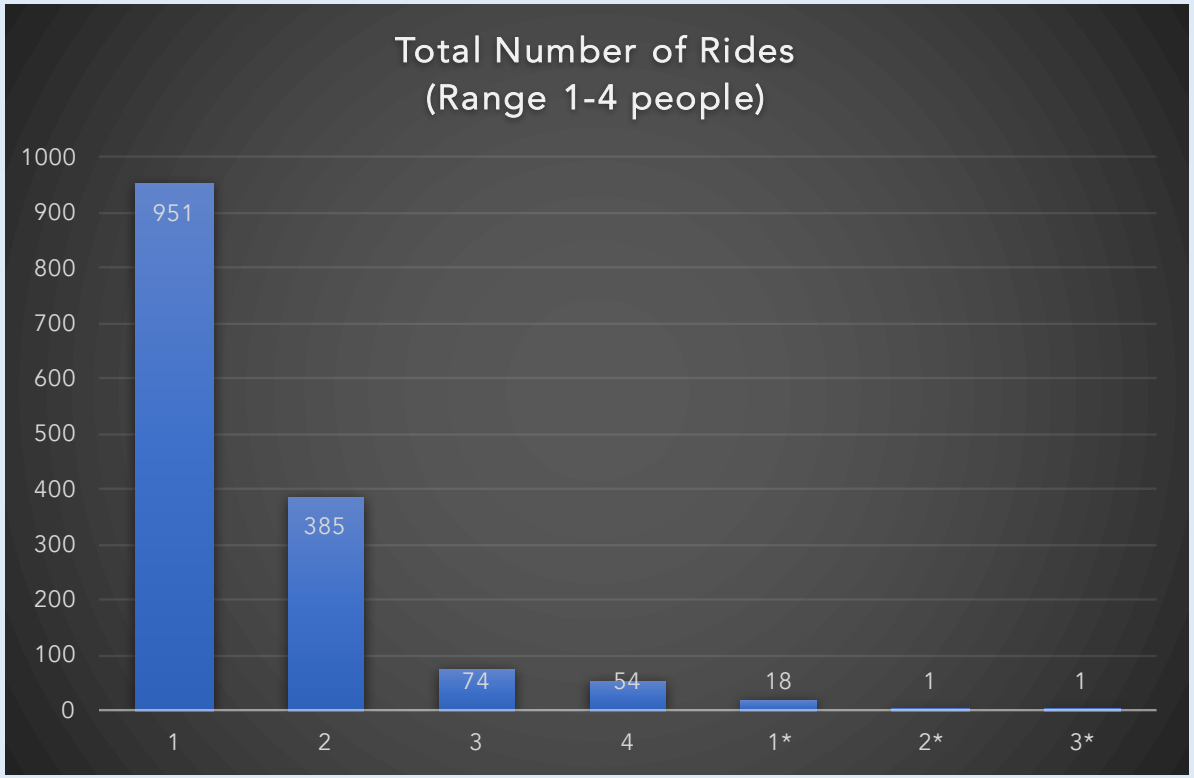
What else can we observe?

- There were more UberX requests than UberPool rides. Only December had more Pool rides.
- There were zero UberPool rides in May.



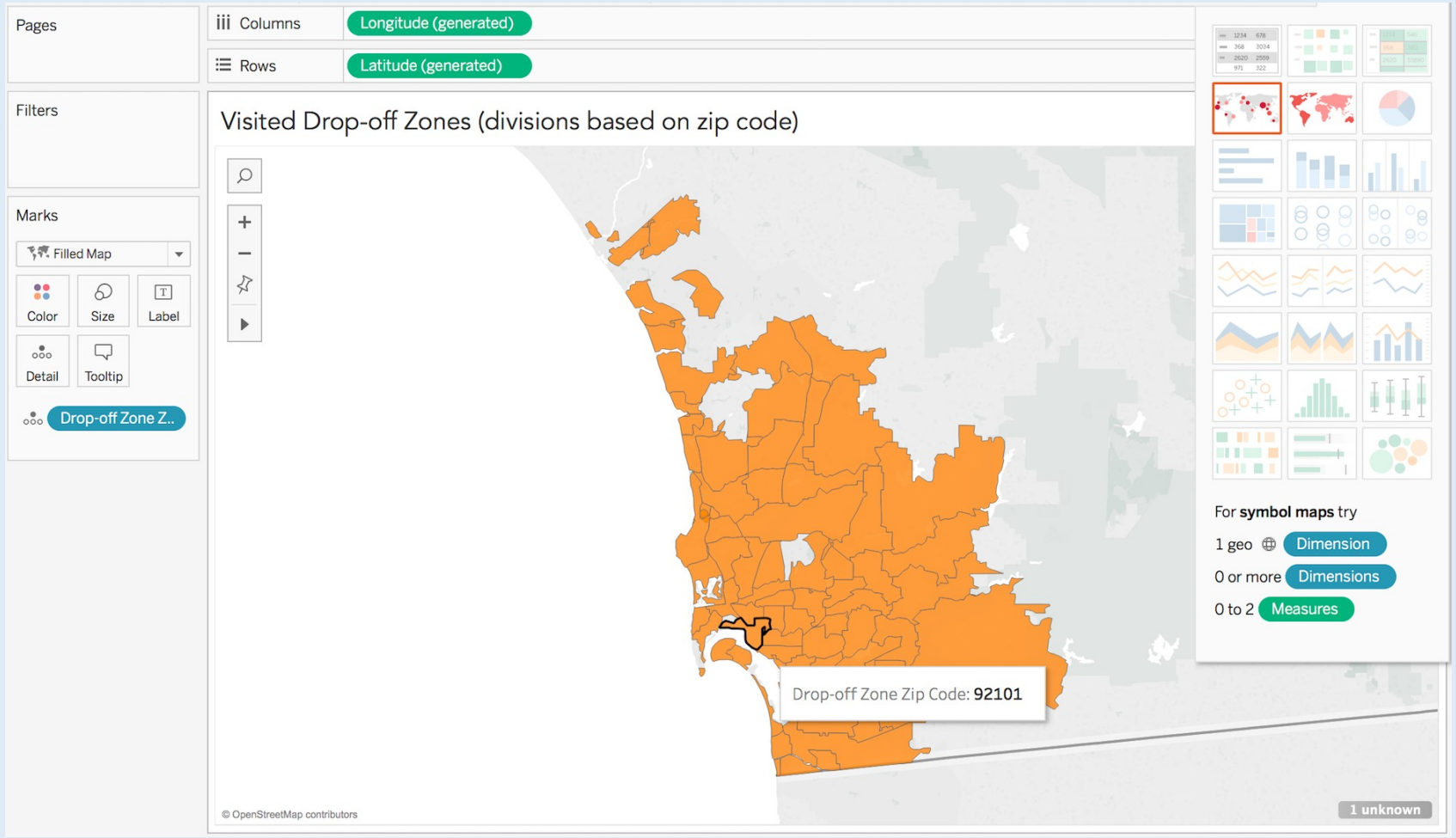
What else can we observe? (Continued)

- There were much more rides with a single passenger than rides with 2 to 4 people. The asterisk represents repeated passengers. Which means that were only 18 rides with individuals that had been aboard my vehicle more than one time.



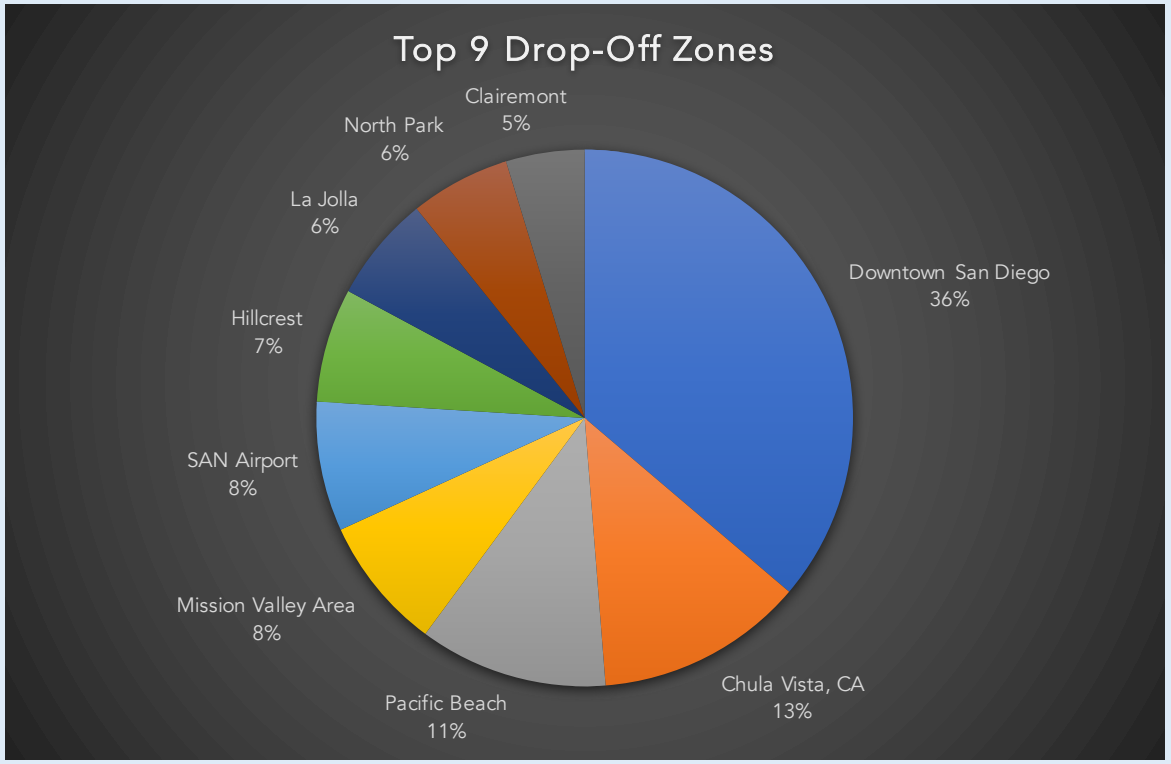
What else can we observe? (Continued)

- The orange sections represent the visited cities and districts from San Diego County.
- The “1 unknown” on the bottom right corner represents a district in Tijuana, Mexico.



One last observation!

- The most popular drop-off addresses were located in Downtown San Diego!
- The percentages on the table below are NOT the total percentages representing the entire dataset's drop-off locations. The percentages correspond to the total number of rides that involved the top 9 areas ONLY.
- Unfortunately listing all neighborhoods or nearby cities wouldn't fit the pie chart.



How can we use this in a real world application? Why should we care?

- The passenger information table could be set as an example of data analysis being applied in order to encourage further studies or research that seeks a deeper understanding of what ideal weather conditions for human beings are.
- The top list of cities contained in the Excel's workbook could become targets for an advertising campaign that promotes San Diego's tourism, specially during both peak summer and winter season. Perhaps, San Diego could be portrayed as an ideal site for a weekend getaway!

THANKS FOR WATCHING!

References

- Climate-Data.org. (2017). Retrieved from https://en.climate-data.org.
- Rubel, F. (2017). *World Maps of Köppen-Geiger climate Classification*. Retrieved from http://koeppen-geiger.vu-wien.ac.at.
- Uber Profile. (2017). *Uber Profile*. Retrieved from https://www.partners.uber.com.
- Weather Underground. (2017). Retrieved from https://www.wunderground.com.