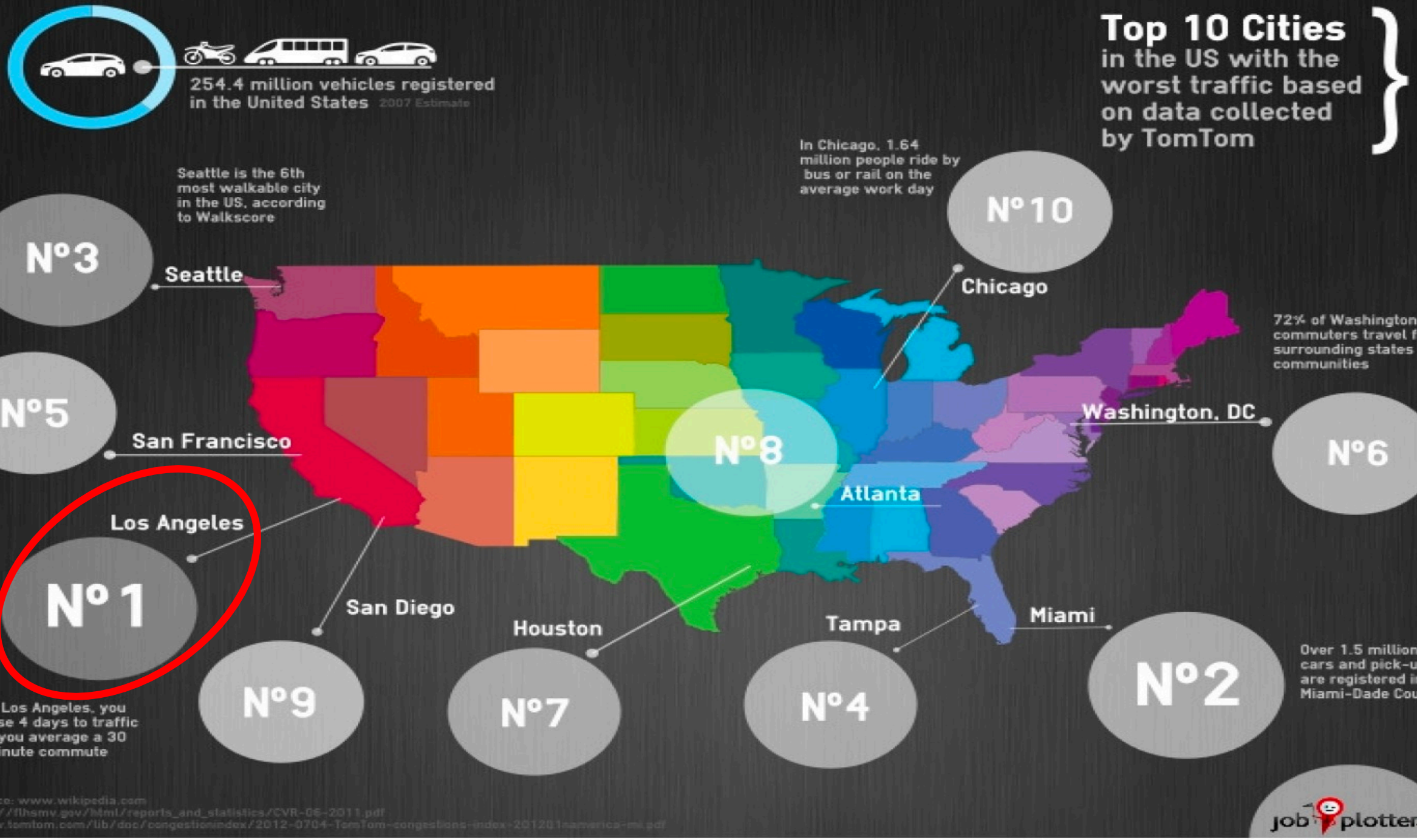
Los Angeles Traffic Collision Analysis

Group: Jon Kaplan, Tyler Marsh, John Fields

IDENTIFYING PATTERNS AND CORRELATIONS IN ORDER TO HELP REDUCE CRASHES AND POTENTIALLY SAVE LIVES

JT


254.4 million vehicles registered in the United States 2007 Estimate
Top 10 Cities in the US with the worst traffic based on data collected by TomTom
In Chicago, 1.64 million people ride by bus or rail on the average work day
Nº10
Seattle is the 6th most walkable city in the US, according to Walkscore
Nº3
Seattle
Chicago
72% of Washington, DC commuters travel from surrounding states and communities
Nº5
San Francisco
Washington, DC
Nº6
Nº8
Atlanta
Los Angeles
Nº1
San Diego
Houston
Tampa
Miami
Nº9
Nº7
Nº4
Nº2
Over 1.5 million cars and pick-ups are registered in Miami-Dade County
In Los Angeles, you lose 4 days to traffic if you average a 30 minute commute
Source: www.wikipedia.com
http://flhsmv.gov/html/reports_and_statistics/CVR-06-2011.pdf
www.tomtom.com/lib/doc/congestionindex/2012-0704-TomTom-congestions-index-20120130america-mi.pdf
jobplotter.com
JT

Los Angeles Traffic Collisions Dataset

Data Sources: Kaggle, Weather, Income

This dataset includes **90,881 rows and 22 feature variables**. Each row corresponds to a traffic collision, and includes these variables:

Time Occurred	Council Districts
Area Name	Median Income
Victim Age	Date
Victim Sex	Weekday
Victim Descent	Longitude
Premise Description	Latitude
Address	HighTemp
Cross Street	LowTemp
Location	Precipitation
Zip Codes	Sunrise Time
Median Income	Sunset Time

Sample Data

Time Occurred	Area Name	Victim Age	Victim Sex	Victim Descent	Premise Description	Address	Cross Street	Location	Zip Codes	Council Districts	Median Income	Date	year	month	weekday	hours	longitude	latitude
13	Devonshire	74	F	A	STREET	NORDHOFF ST	LOUISE AV	{'latitude': '34.2359', 'human_address': '{"ad...	18514	2	55024.39	2017-12-31	2017	12	6	13	-118.5113	34.2359

Los Angeles Traffic Collisions Dataset

Data Questions We Explored:

What are the most dangerous intersections?

What are the most common collision areas in Los Angeles?

What are the best/worst times of the day for accidents? Best/worst month?

What is the demographic makeup of victims in collisions?

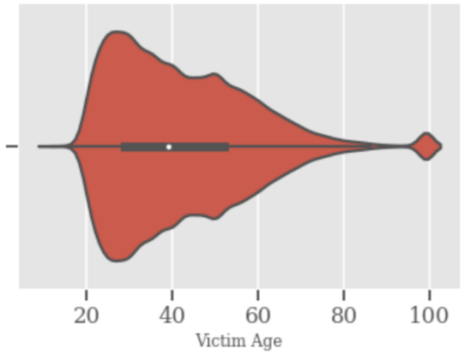
What is the relationship between income and collision victims?

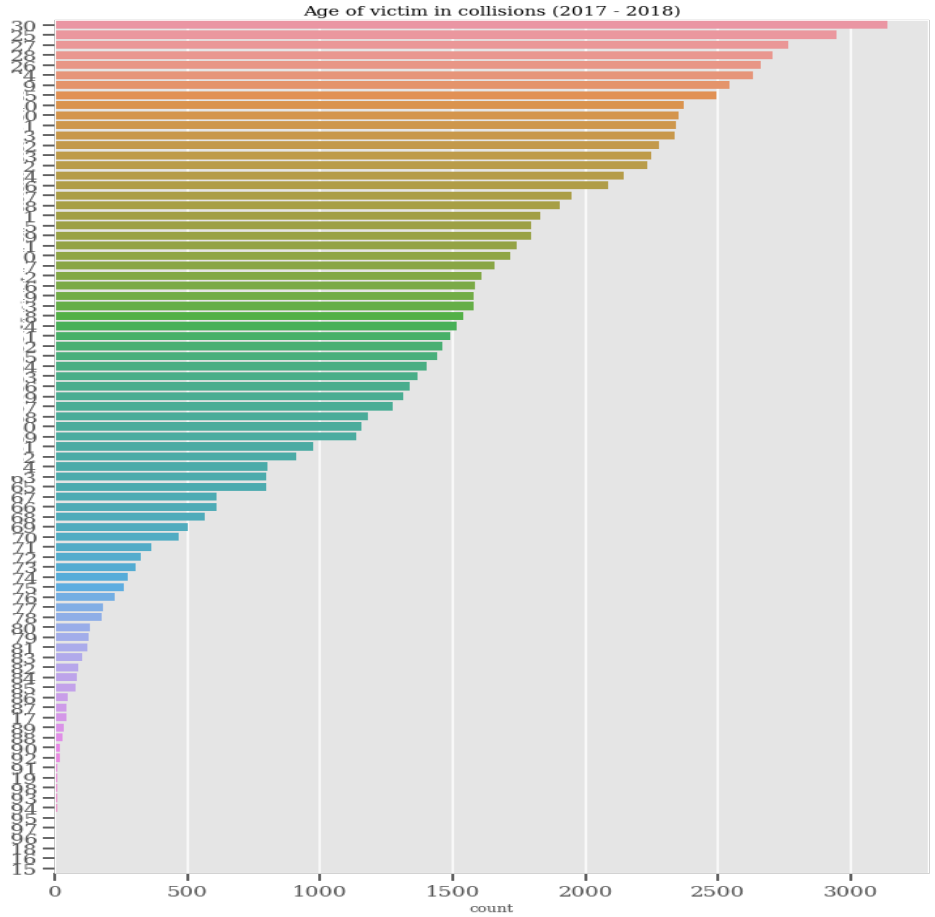
Do certain temperatures or weather play a factor?

Does the amount of daylight play a factor?

Los Angeles Traffic Analysis: Demographics

Age





Gender

There are substantially more men involved in collisions compared to women:

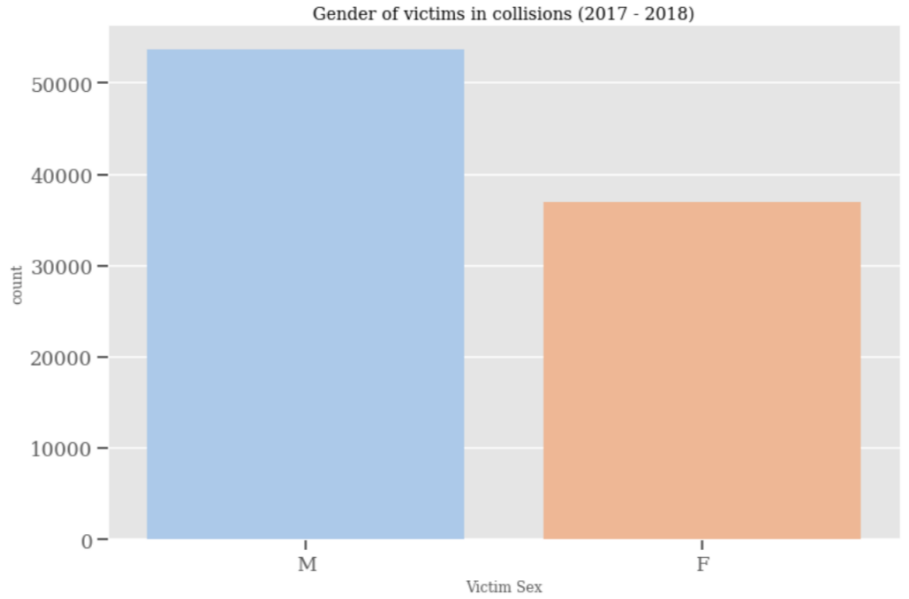


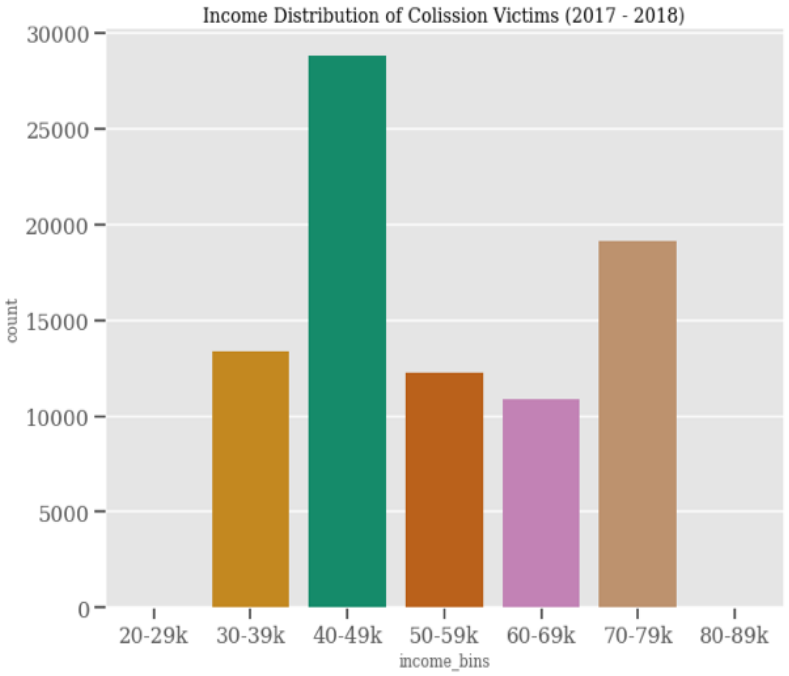
FIGURE 14

Los Angeles Traffic Analysis: Demographics

Median Income

Victim Income Breakdown		
Income Range	Count	Percentage
30-39k	13410	15.86%
40-49k	28816	34.08%
50-59k	12266	14.51%
60-69k	10900	12.89%
70-79k	19166	22.66%

TABLE 3 - VICTIM INCOME BREAKDOWN



Ethnicities

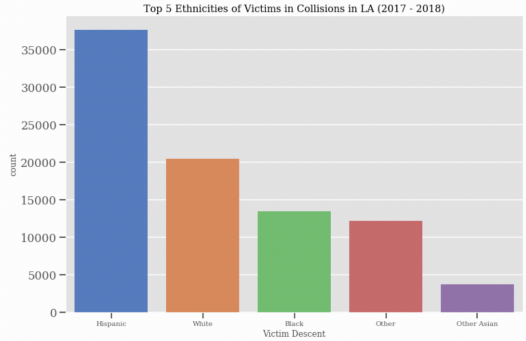
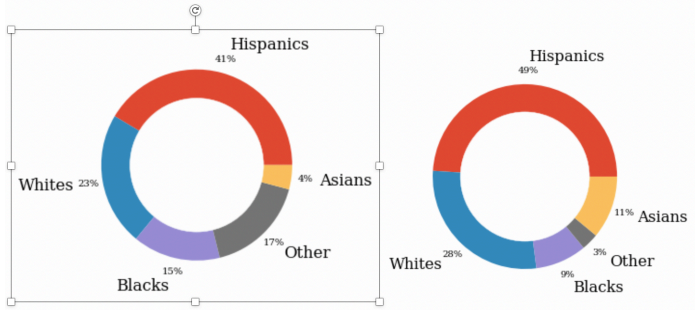


FIGURE 8



LA population source: https://censusreporter.org/profiles/16000US0644000-los-angeles-ca/

About the program(s)

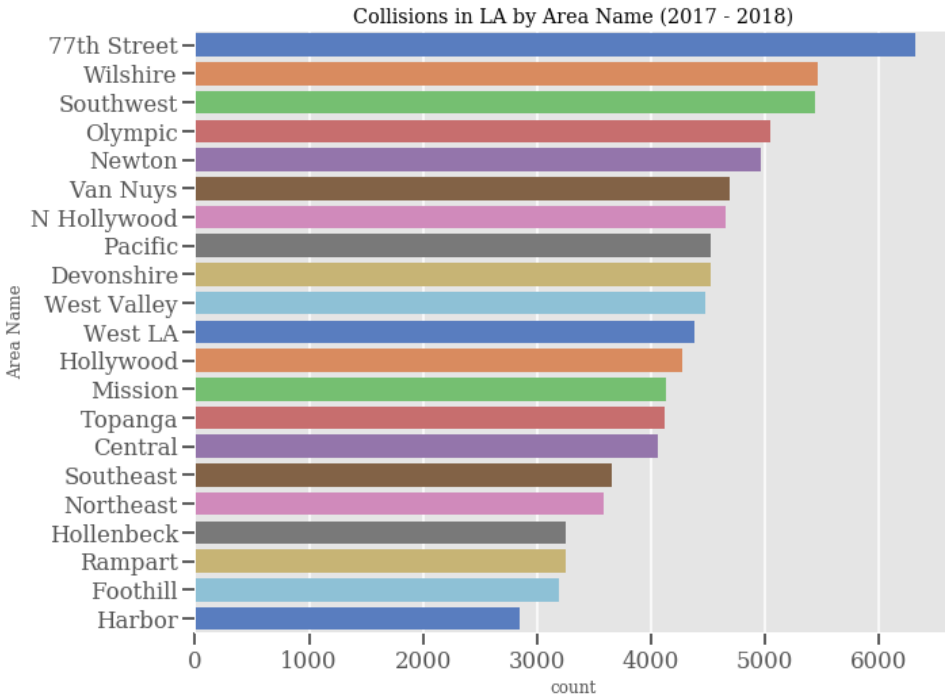
Scraping Weather

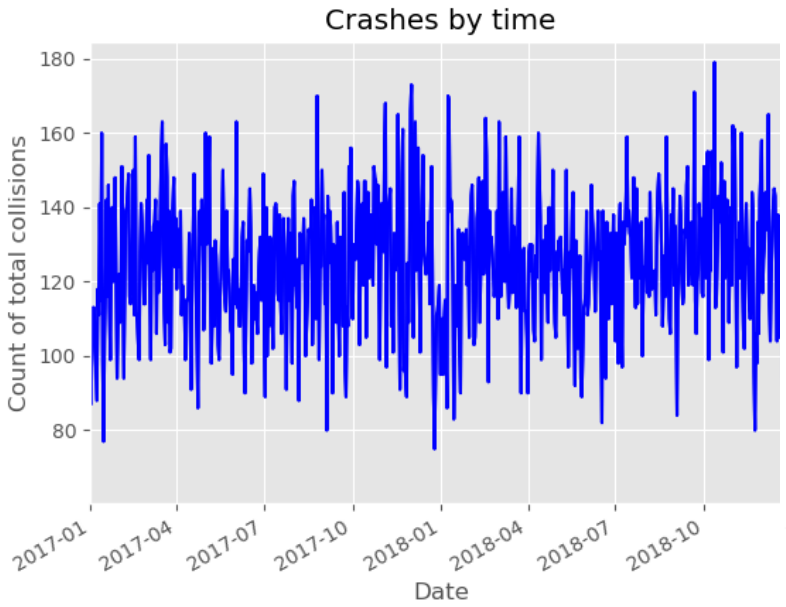
```
def get_dates(year, months):...

def empty_df():...

def insert_data(base_url, the_xpath, dates):...

dates17 = get_dates(2017, 12) #make dates list

#create container for data

df17 = empty_df() #make data frame

#get the data into a list of lists

data_points = insert_data(base_url1, the_xpath1, dates17)

df17 = df17.append(pd.DataFrame(data_points, columns=df17.columns))

#get working directory

my_cwd = os.getcwd()

df17.to_csv(my_cwd,'/2017weather.csv', index=False)
```

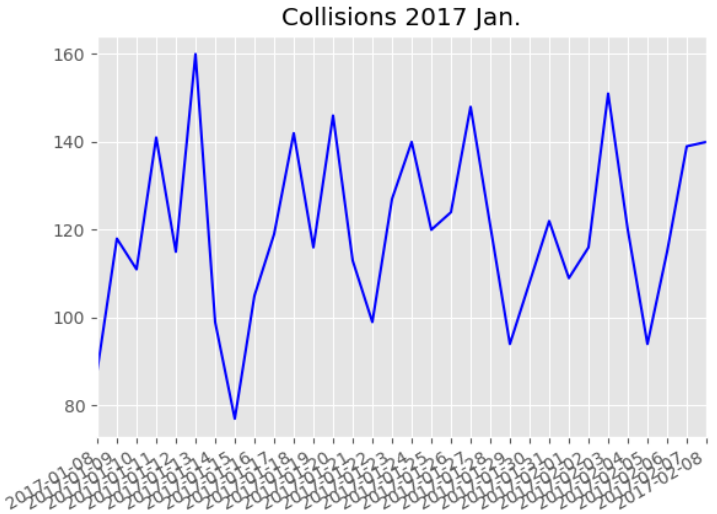
Adding Robust Modular Code

```
def vc_fun(df, column_name, *args):

vc_fun(sunny_df, 'VictimSex')

def get_coords(df):...

def map_LAcoords(tuple_list, size,
desired_file_name, type):...

coordSunny = get_coords(sunny_df)

#make the map save it
map_LAcoords(coordSunny, 13,
'SunnyCrash', 'heat')
```

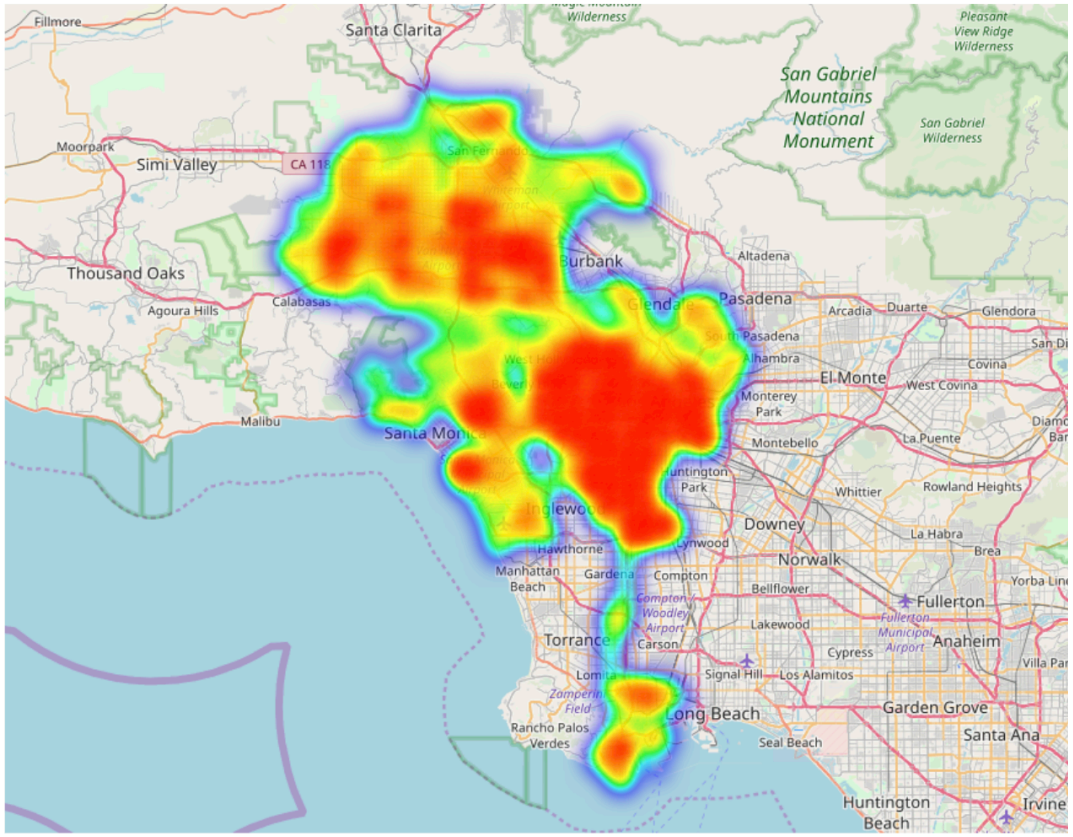
Los Angeles Traffic Analysis: Location/Time







Los Angeles Traffic Analysis: Time of Day



HEAT MAP OF CRASHES BETWEEN SUNRISE AND SUNSET

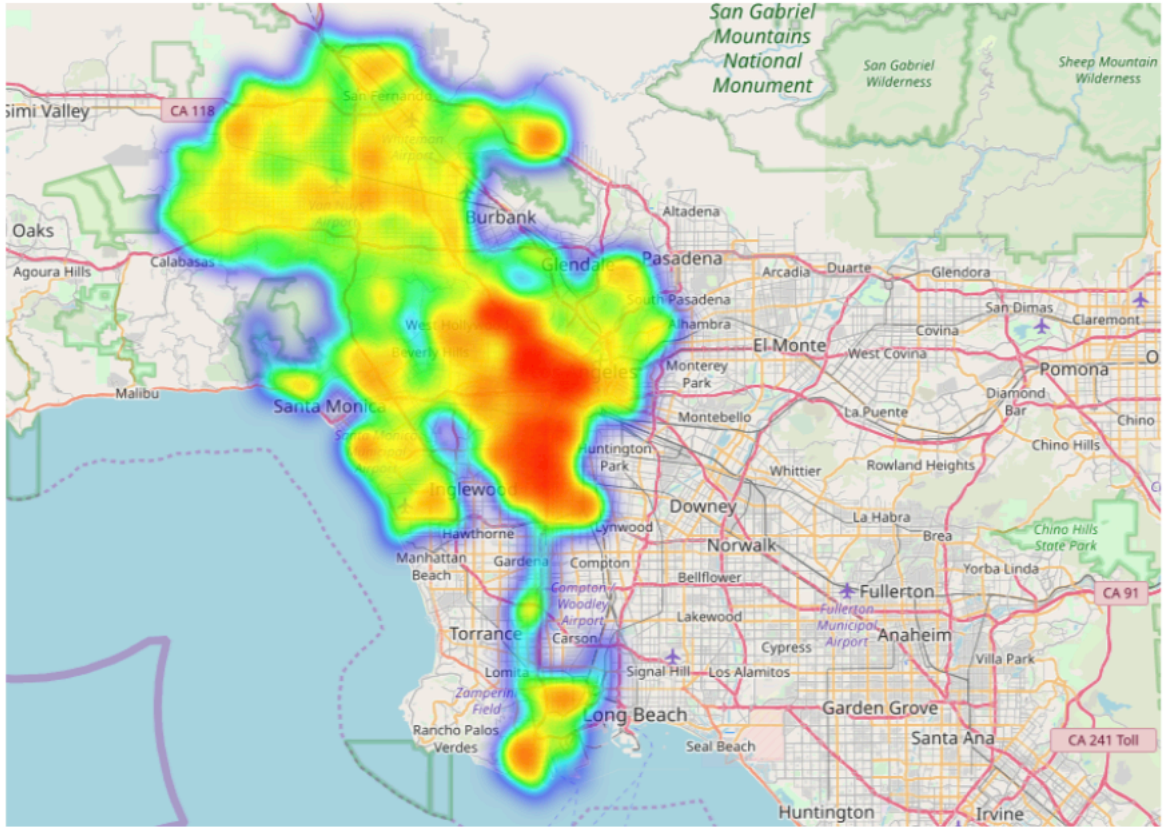
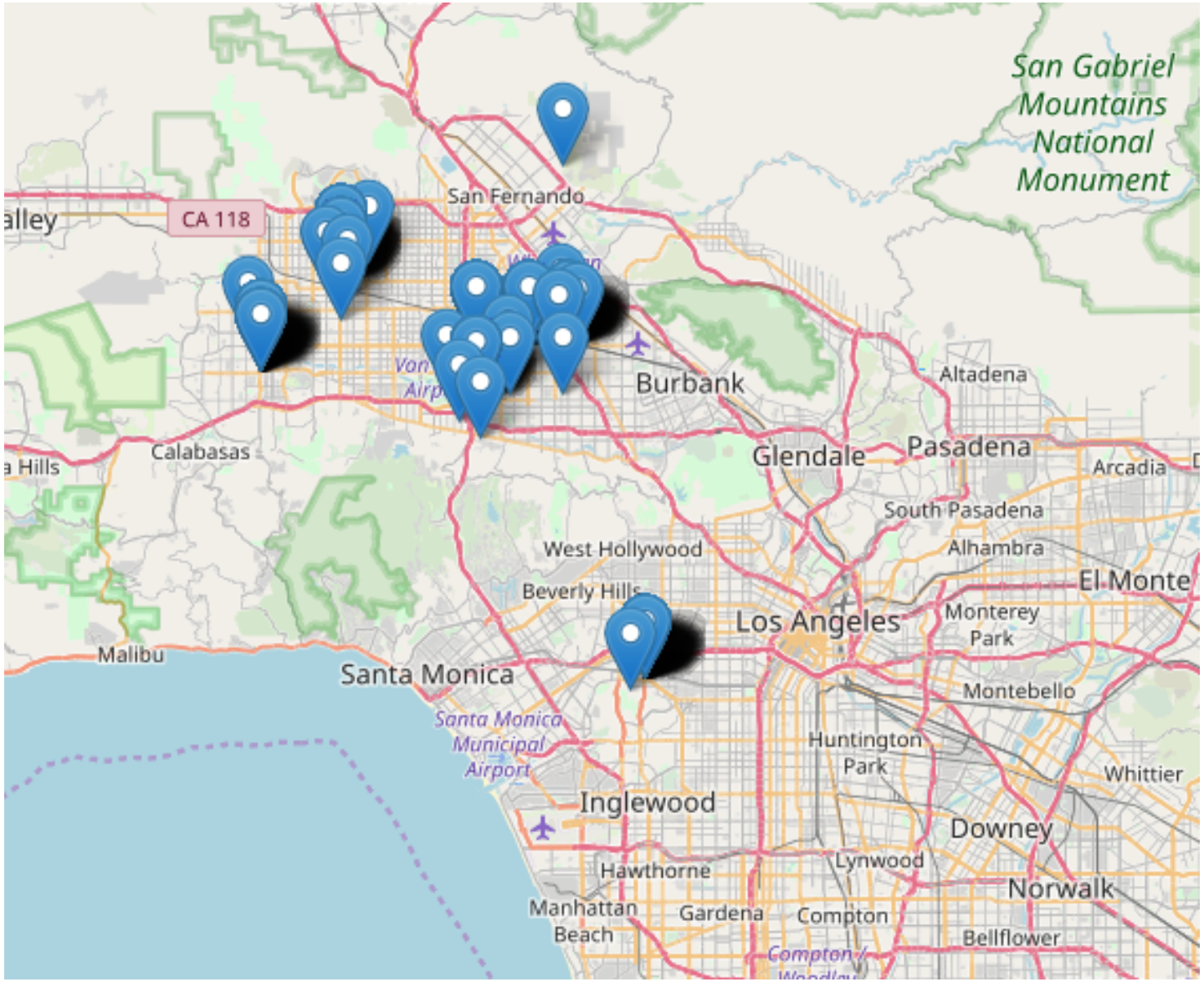


FIGURE 7, HEAT MAP OF CRASHES BEFORE SUNRISE AND AFTER SUNSET

Most collisions by Address and Cross street

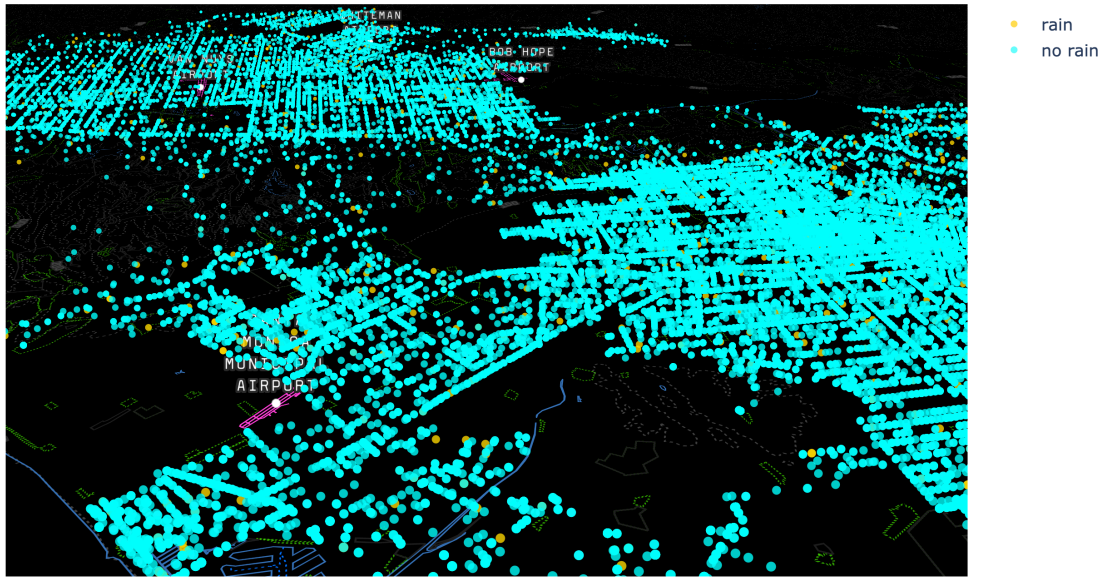
Over 400 Collisions occurred in these area. We can see how the collisions were in the concentrated area of crashes to the north of LA in the crashes during day light.



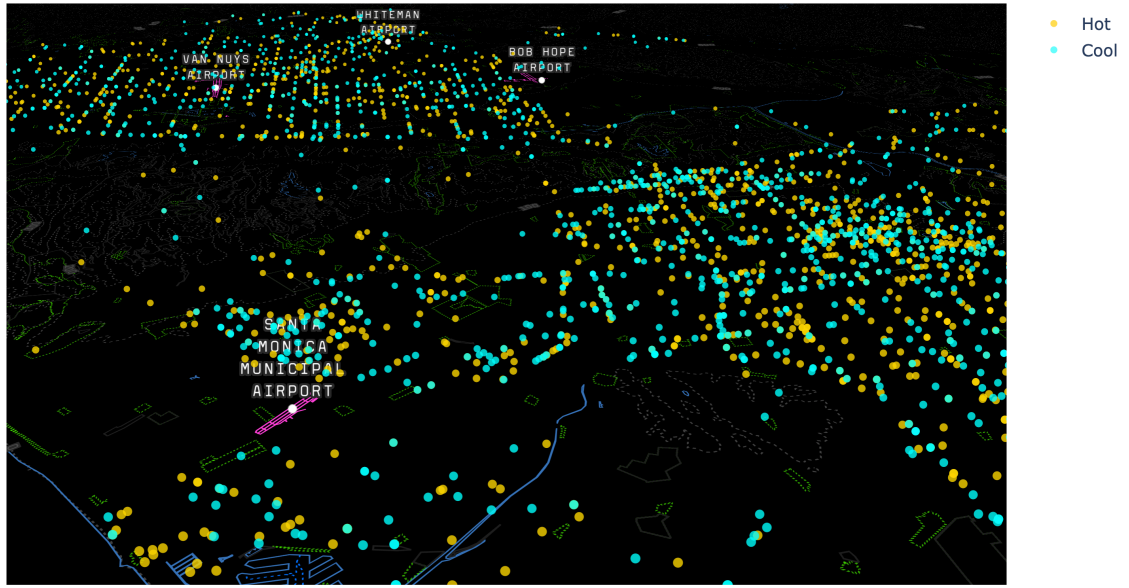
Tyler

Los Angeles Traffic Analysis: Weather

LA Accidents 2017-2018 with/without precipitation



LA Accidents 2017-2018 on Hot and Cool Days



Demo of zooming and panning in MapBox/Plotly

Los Angeles Traffic
Conclusions and Recommendations

The results of this study provide additional insights that can be used by city planners, government officials and citizens to better understand the Los Angeles traffic conditions in 2017 and 2018.

We recommend the use of this information in the following areas:

1. Additional studies of dangerous intersections and locations.

2. Warnings on interactive road signs during rush hour when it is raining.

3. Utilize this data and interactive maps for planning sessions with subject matter experts and citizens when starting new projects which impact roadways.

4. Public service campaign to educate the public (with targeted messages to men).

Visualization Packages

- MapBox/Plot.ly
  - https://plot.ly/python/scattermapbox/
- Folium
  - https://pypi.org/project/folium/
- MatPlotLib
  - https://matplotlib.org/