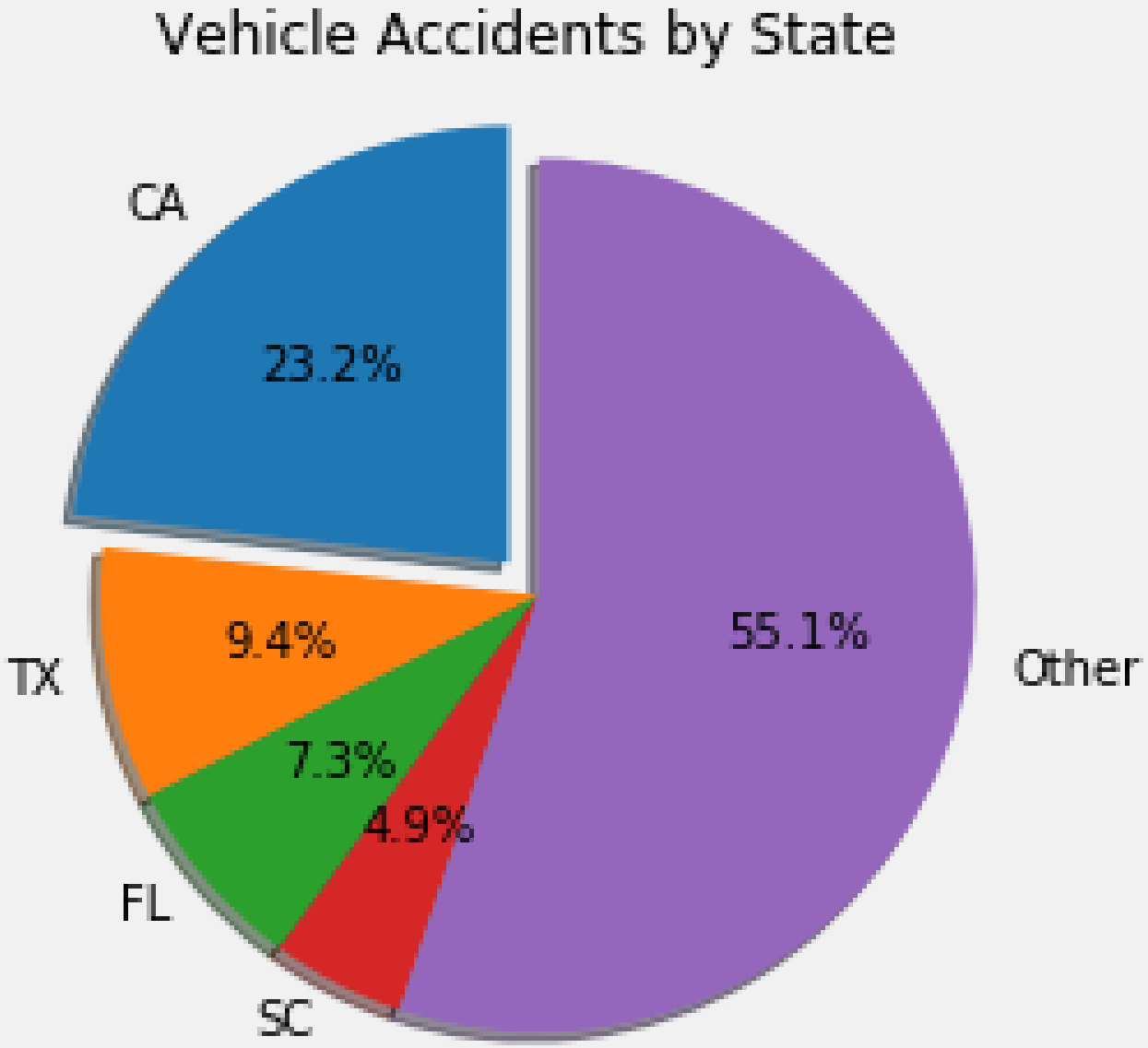
# UNDERSTANDING FACTORS ATTRIBUTING TO TRAFFIC ACCIDENTS IN CALIFORNIA

Team 6

John O'Boyle, Derek Yang, Isaias Larios, Kaixin Lin

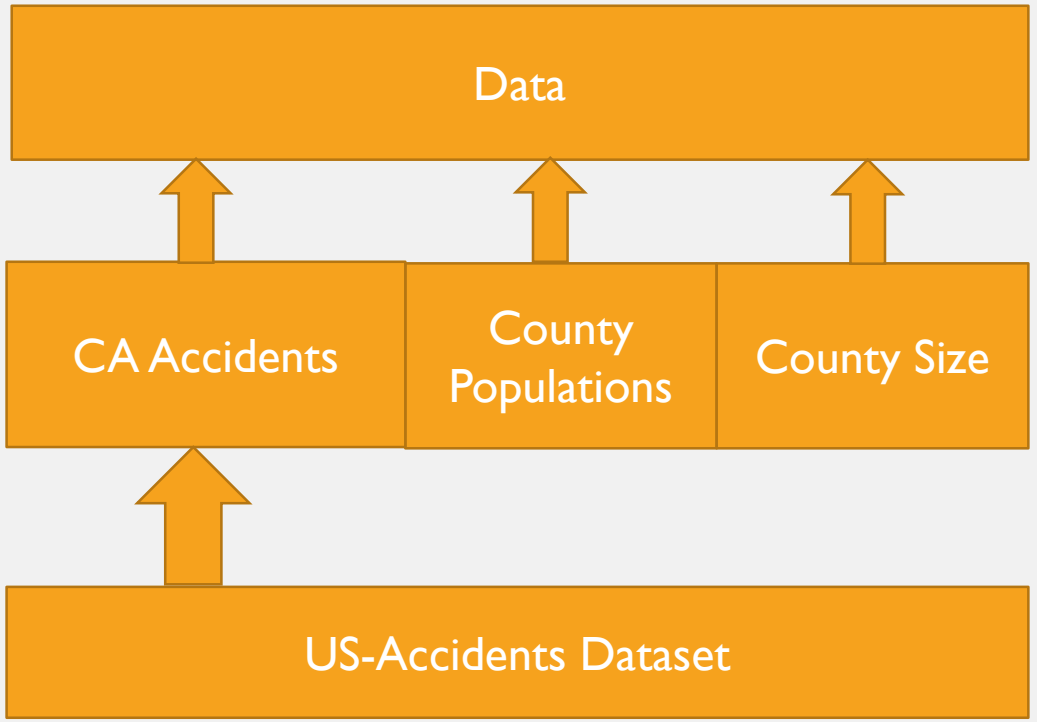
# MOTIVATION AND OBJECTIVE

- The state of California has the most licensed drivers in the United States.
- We also account for a large portion of vehicle accidents in the United States.
- Analysis of accident data would provide insight into the factors that lead to vehicle accidents and their severity in the most populous state in America



# US-ACCIDENTS OPEN DATASET

- US-Accidents open dataset[1] ( https://smoosavi.org/datasets/us_accidents ) contains data on 3.5 million traffic accidents collected from February 2016 to June 2020 for the Contiguous United States.
- California County size information https://www.counties.org/pod/square-mileage-county
- 2020 California County Population https://worldpopulationreview.com/us-counties/states/ca




[1] Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, and Rajiv Ramnath. "A Countrywide Traffic Accident Dataset.", arXiv preprint arXiv:1906.05409 (2019).
[1] Moosavi, Sobhan, Mohammad Hossein Samavatian, Srinivasan Parthasarathy, Radu Teodorescu, and Rajiv Ramnath. "Accident Risk Prediction based on Heterogeneous Sparse Data: New Dataset and Insights." In proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, ACM, 2019.

# FACTORS TO ANALYZE

**Weather**

Time of day and traffic patterns

Traffic hotspots and high-risk locations

# FACTORS TO ANALYZE

Weather

**Time of day and traffic patterns**

Traffic hotspots and high-risk locations

# FACTORS TO ANALYZE

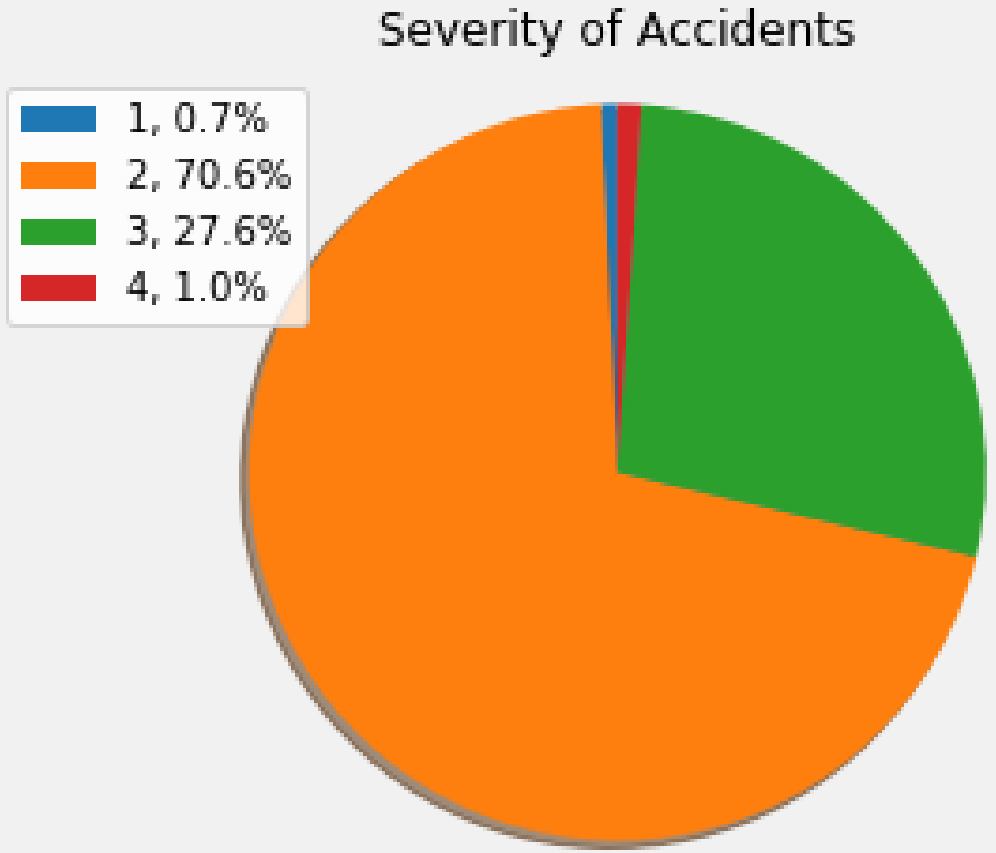
Weather

Time of day and traffic patterns

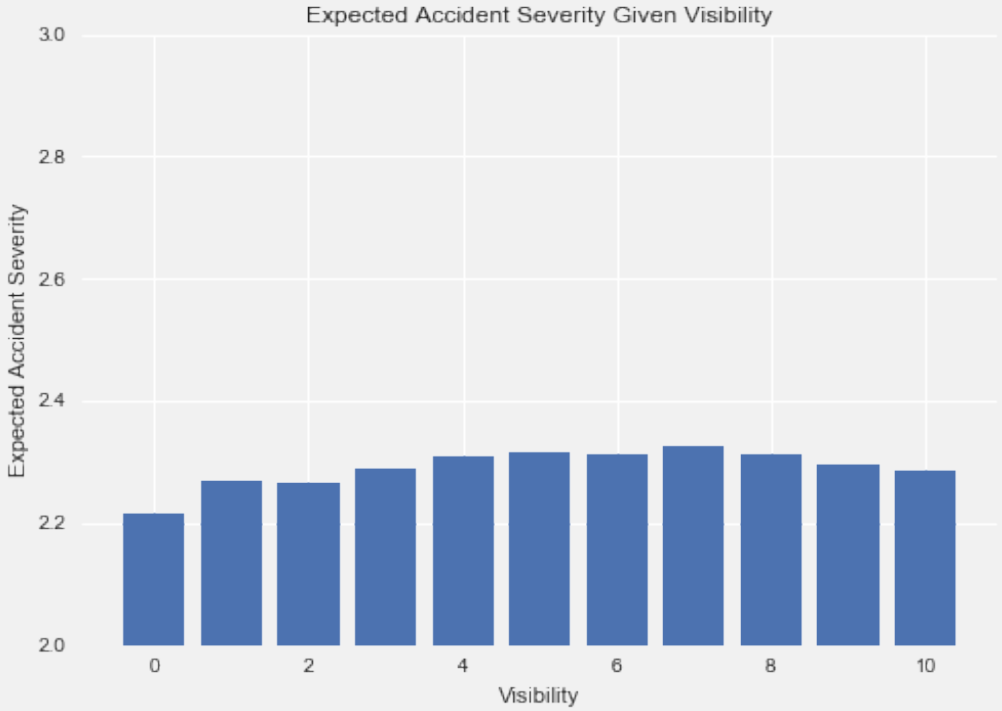
**Traffic hotspots and high-risk locations**

# SEVERITY

- Severity of an accident is measured by the impact of the accident on the surrounding traffic.
- A 1 indicates low traffic impact while a 4 can be large collisions which impact traffic for a longer period.

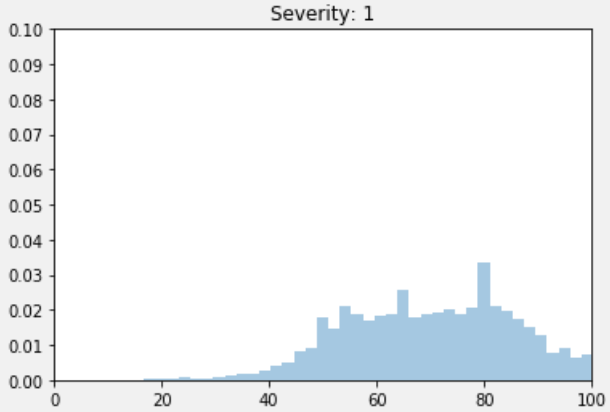


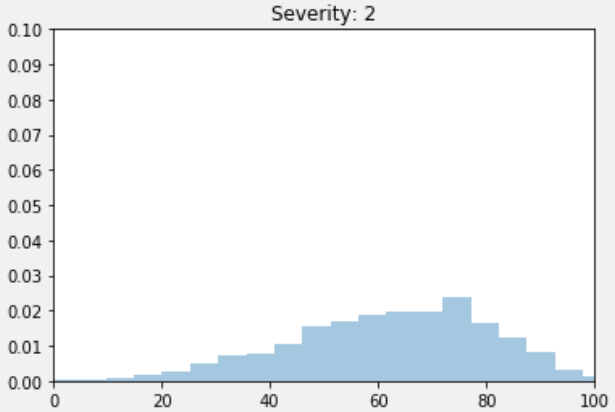
# VISIBILITY AFFECTING ACCIDENT SEVERITY

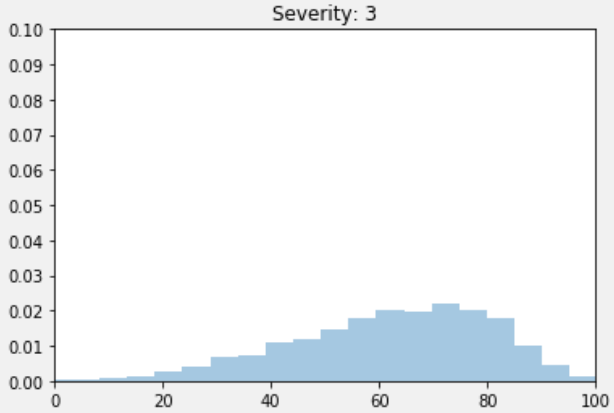


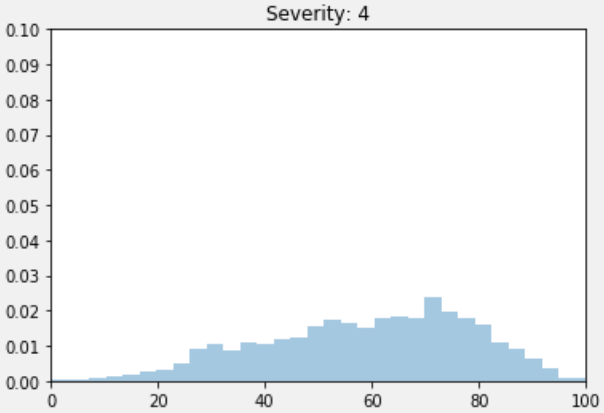

| Accident Severity | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Expected Visibility: E(Vis) | 9.5 | 9.21 | 9.22 | 9.14 |
| Variance V(Vis) | 2.24 | 3.99 | 3.74 | 4.45 |
| Sample Size | 197 | 149,242 | 105,986 | 261 |

# EFFECT OF TEMPERATURE ON ACCIDENT SEVERITY



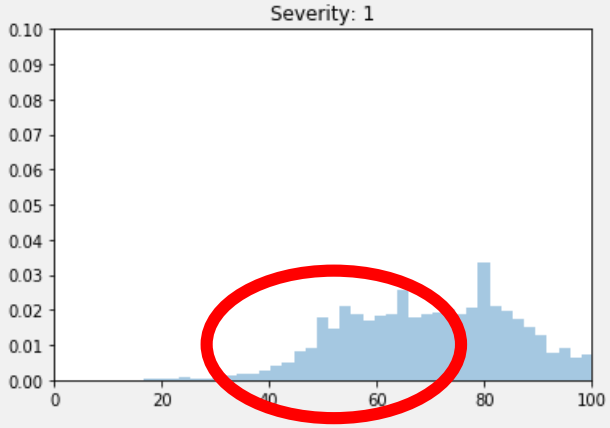



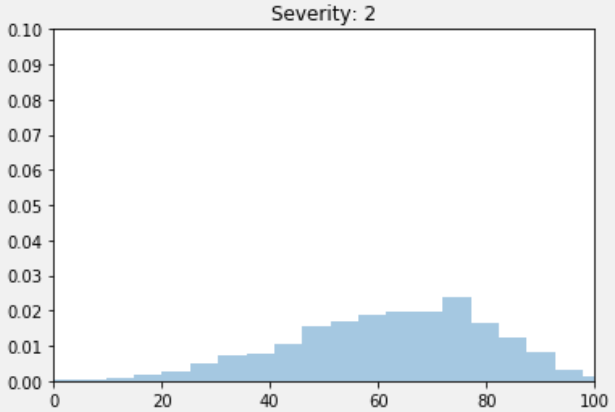



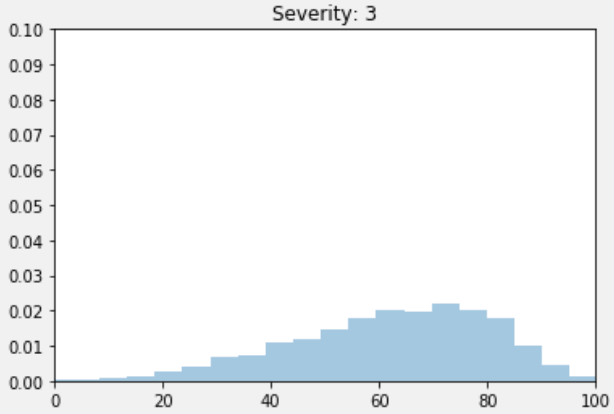



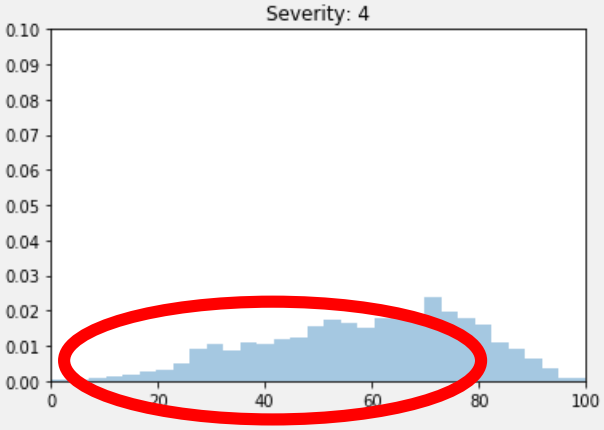

| Severity | Temperature(F) mean |
|---|---|
| 1 | 70.741 |
| 2 | 61.994 |
| 3 | 61.859 |
| 4 | 59.021 |

# EFFECT OF TEMPERATURE ON ACCIDENT SEVERITY



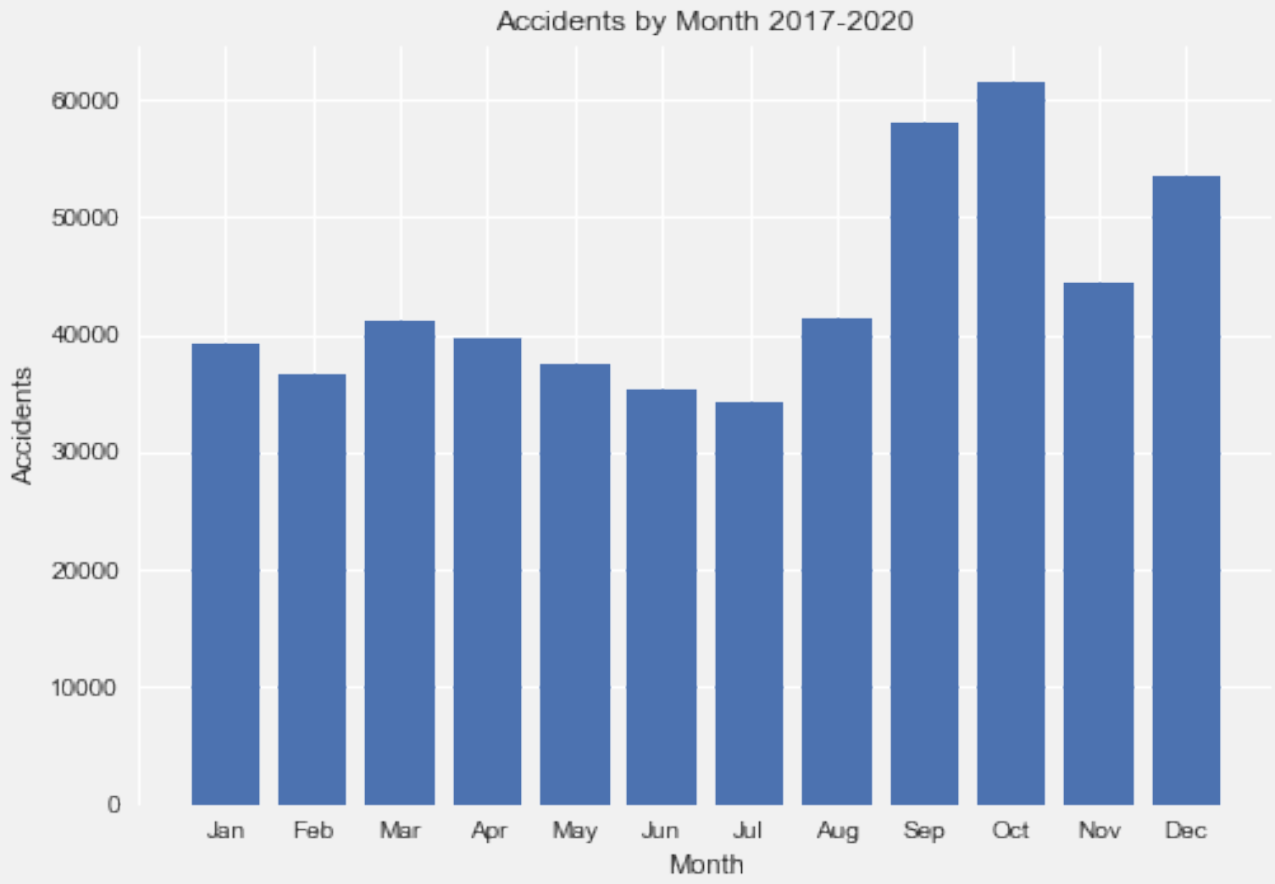




| Severity | Temperature(F) mean |
|---|---|
| 1 | 70.741 |
| 2 | 61.994 |
| 3 | 61.859 |
| 4 | 59.021 |

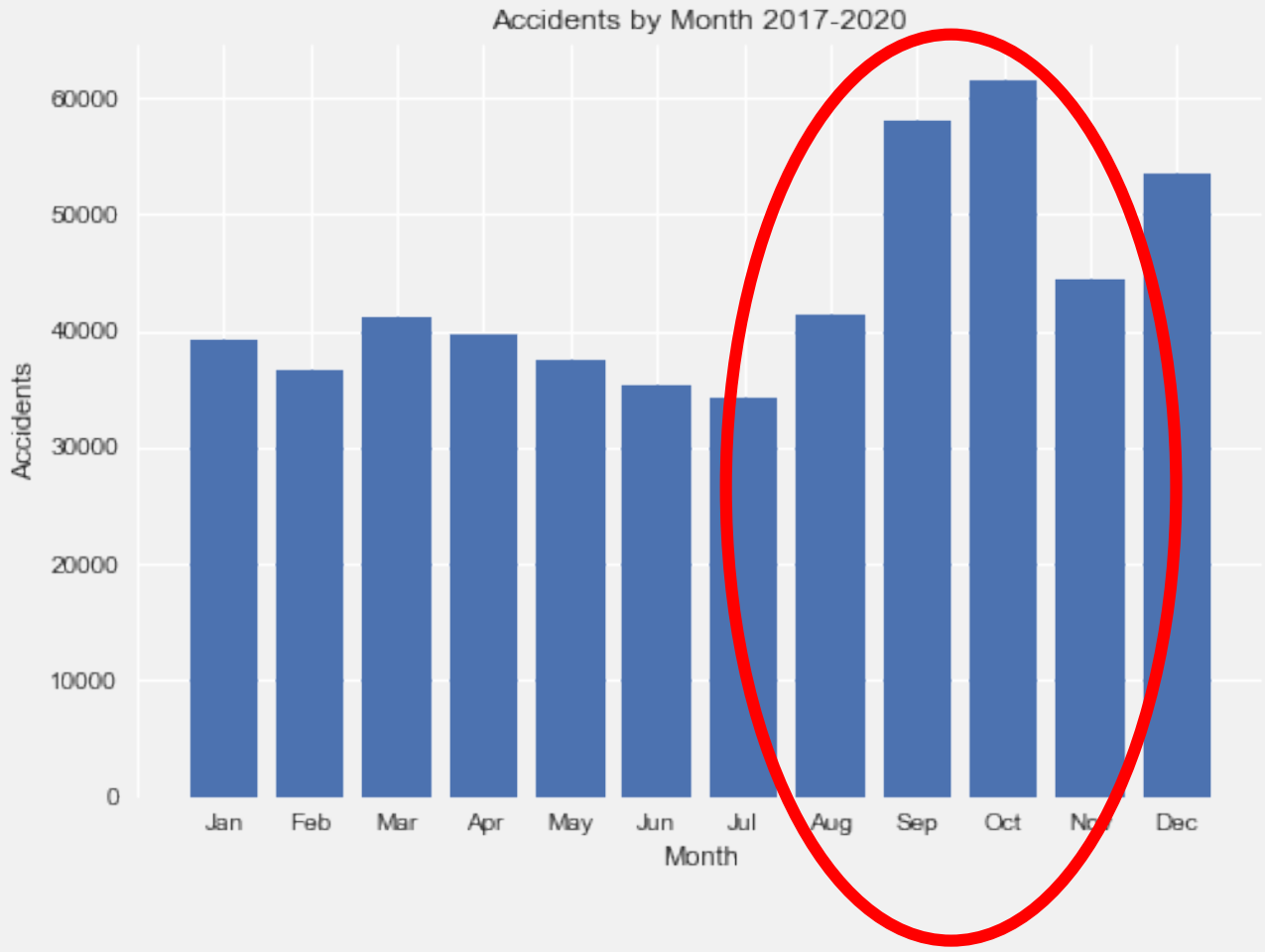
# EFFECT OF MONTH ON NUMBER OF ACCIDENTS

- Starting August, the number of accidents begin to increase.
- Factors that can contribute are:
  1)Younger drivers tend to drive more after summer
  2) Longer commutes due to holidays that occur at end of year



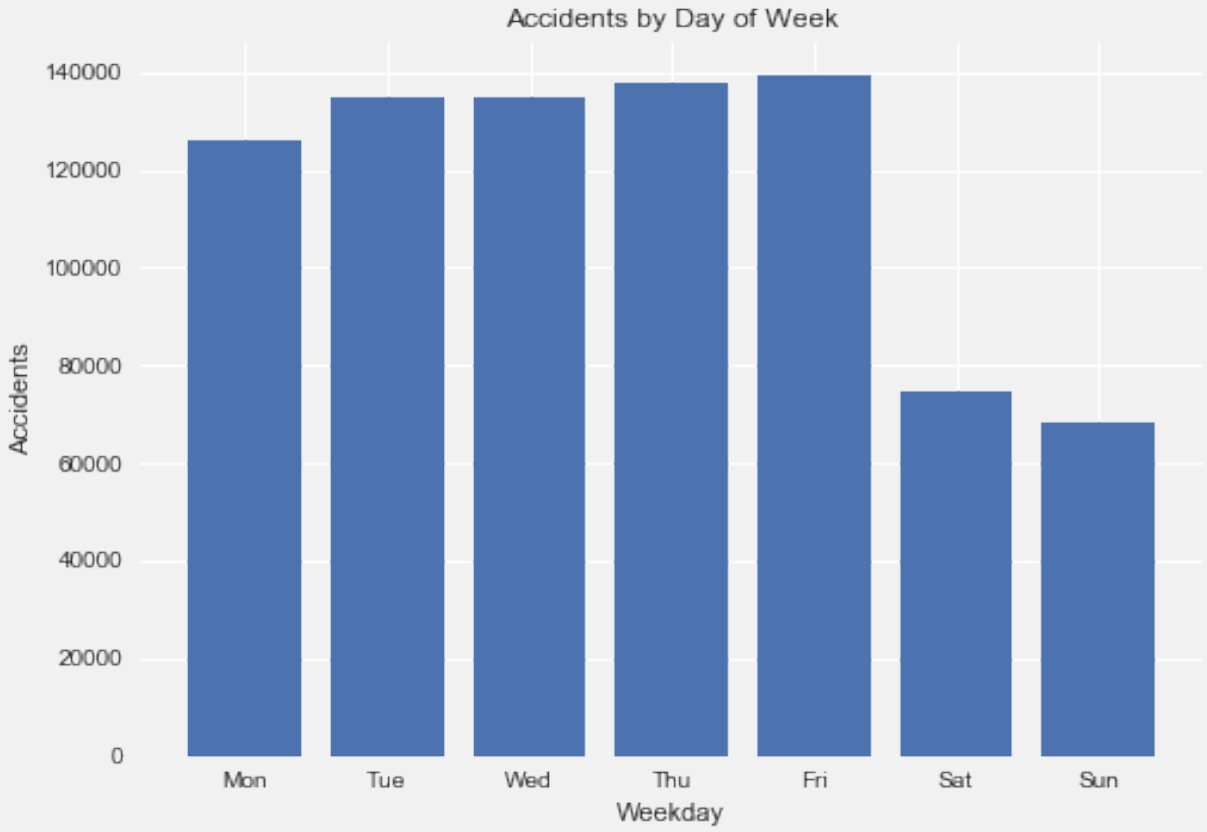
# EFFECT OF MONTH ON ACCIDENT OCCURANCES

- Starting August, the number of accidents begin to increase.
- Factors that can contribute are:
  1)Younger drivers tend to drive more after summer
  2) Longer commutes due to holidays that occur at end of year



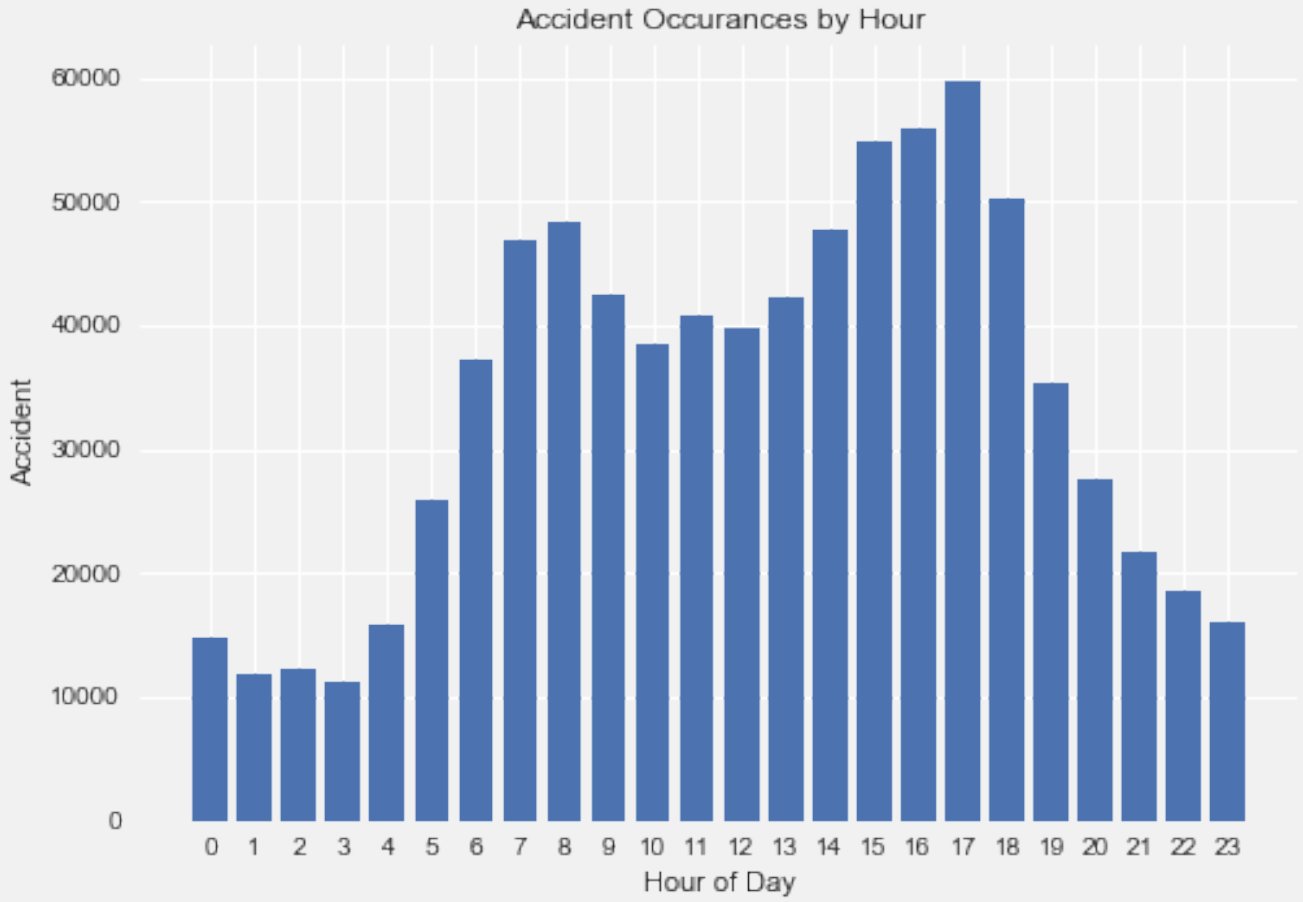
# EFFECT OF DAY OF THE WEEK ON ACCIDENT OCCURRENCES

- Despite previous beliefs, the weekend exhibits a drop in the number of accidents.
- The work week (Mon-Friday) have a consistent number accidents which could be contributed to accidents occurring while driving to or from work.



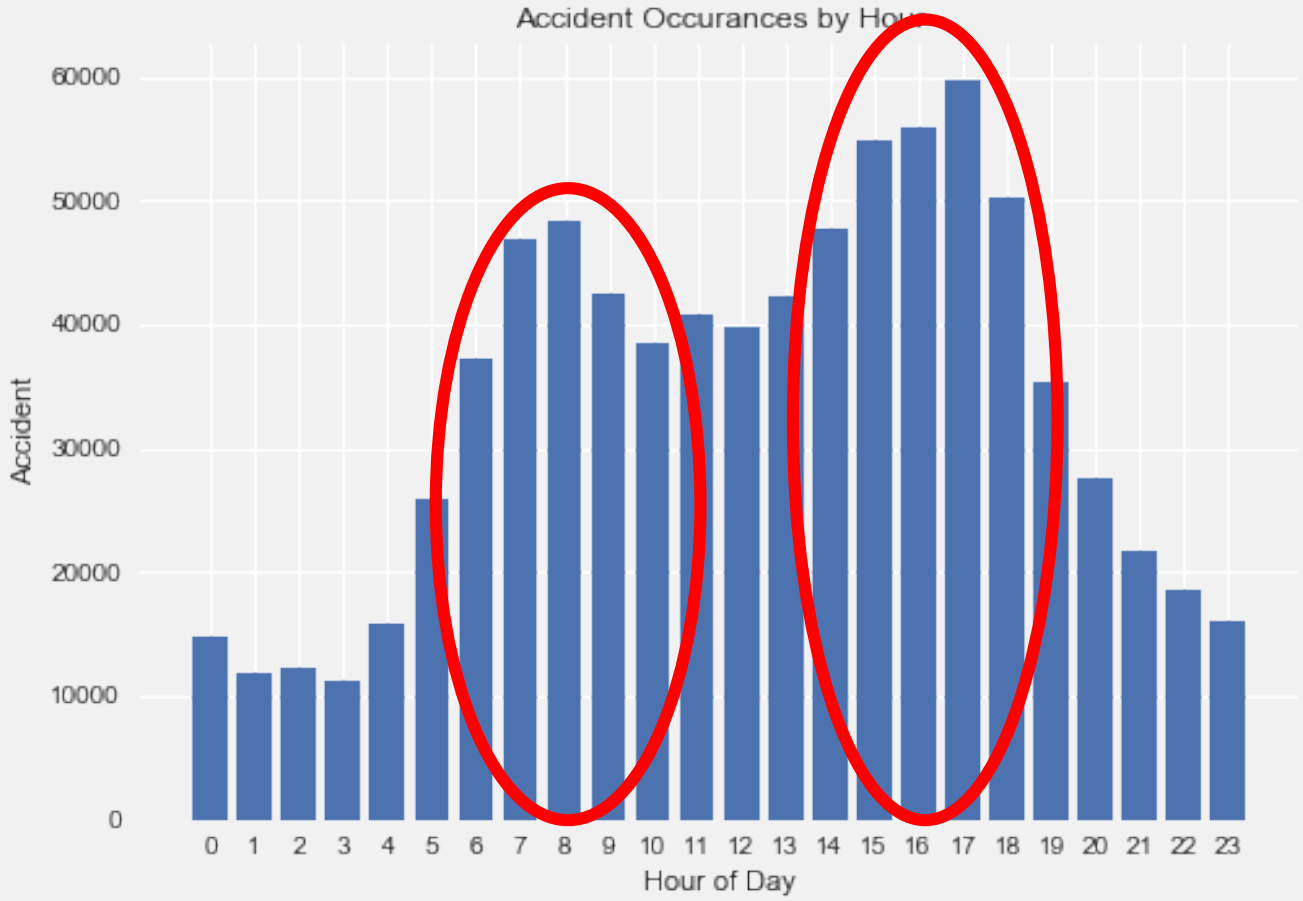
# EFFECT OF HOUR OF DAY ON ACCIDENT OCCURRENCES

- Accidents are quite low early in the morning or late at night but exhibit two distinct peaks.
- Between 8am-10am and 3-6pm we see the most accidents occurring.
- These times are related to “rush hour” when drivers are heading to and from work.
- Supports conclusion from accidents by day of week that accidents are more likely to happen on daily work commutes.



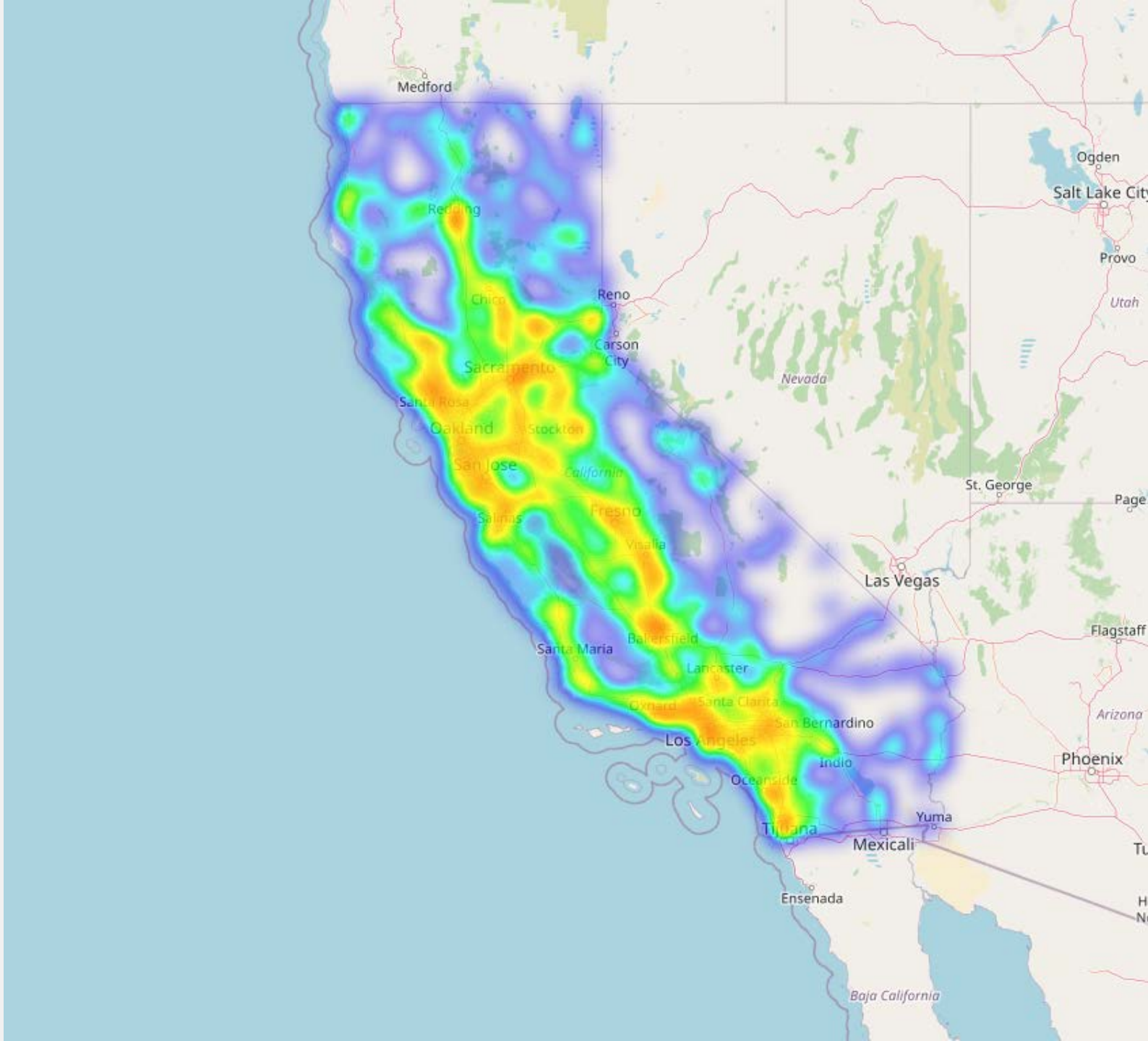
# EFFECT OF HOUR OF DAY ON ACCIDENT OCCURRENCES

- Accidents are quite low early in the morning or late at night but exhibit two distinct peaks.
- Between 8am-10am and 3-6pm we see the most accidents occurring.
- These times are related to “rush hour” when drivers are heading to and from work.
- Supports conclusion from accidents by day of week that accidents are more likely to happen on daily work commutes.



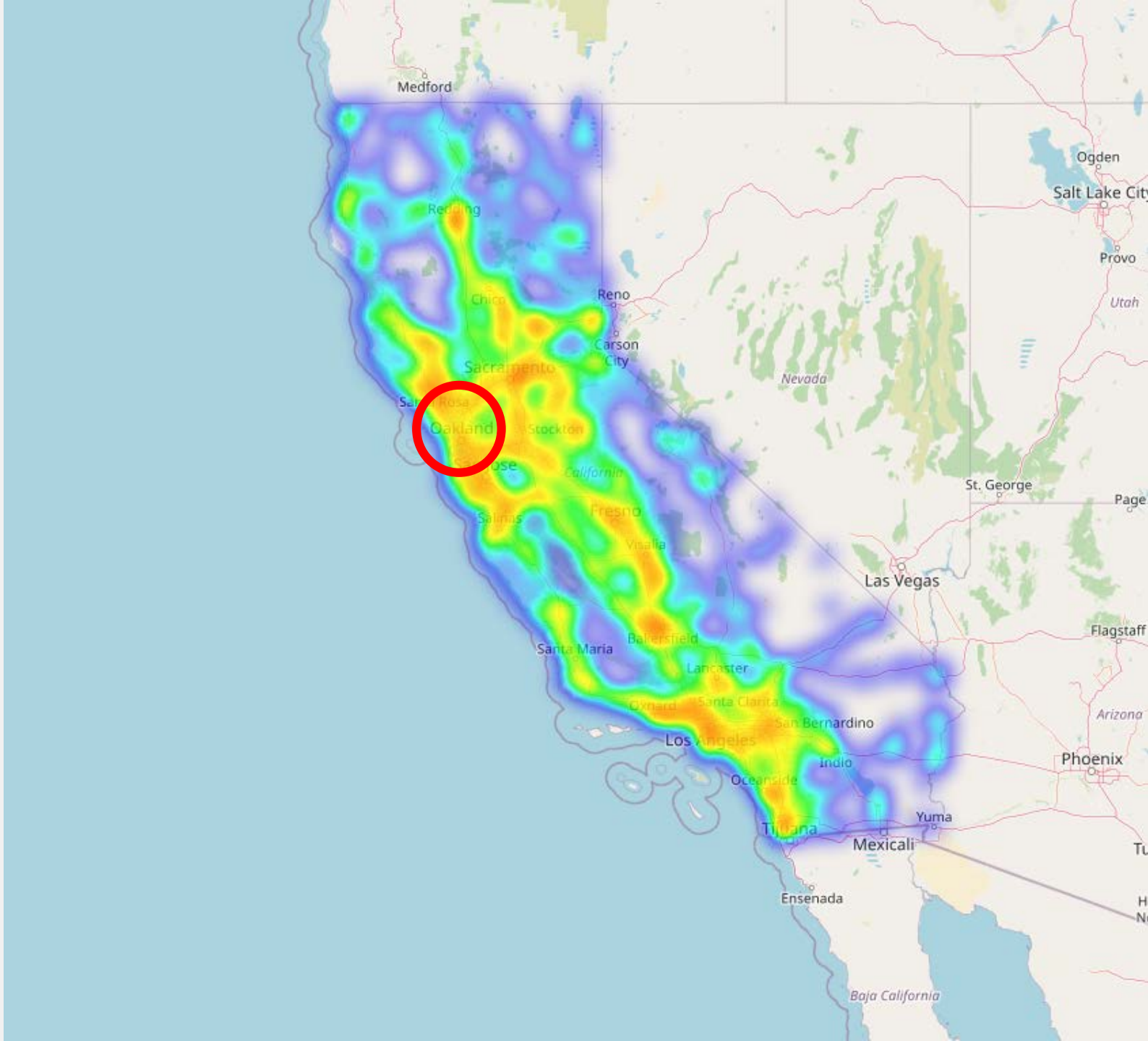
# CALIFORNIA ACCIDENT HEATMAP

- We can see that areas that people generally believe to be dangerous areas to drive in are indeed hotspots.



# CALIFORNIA ACCIDENT HEATMAP

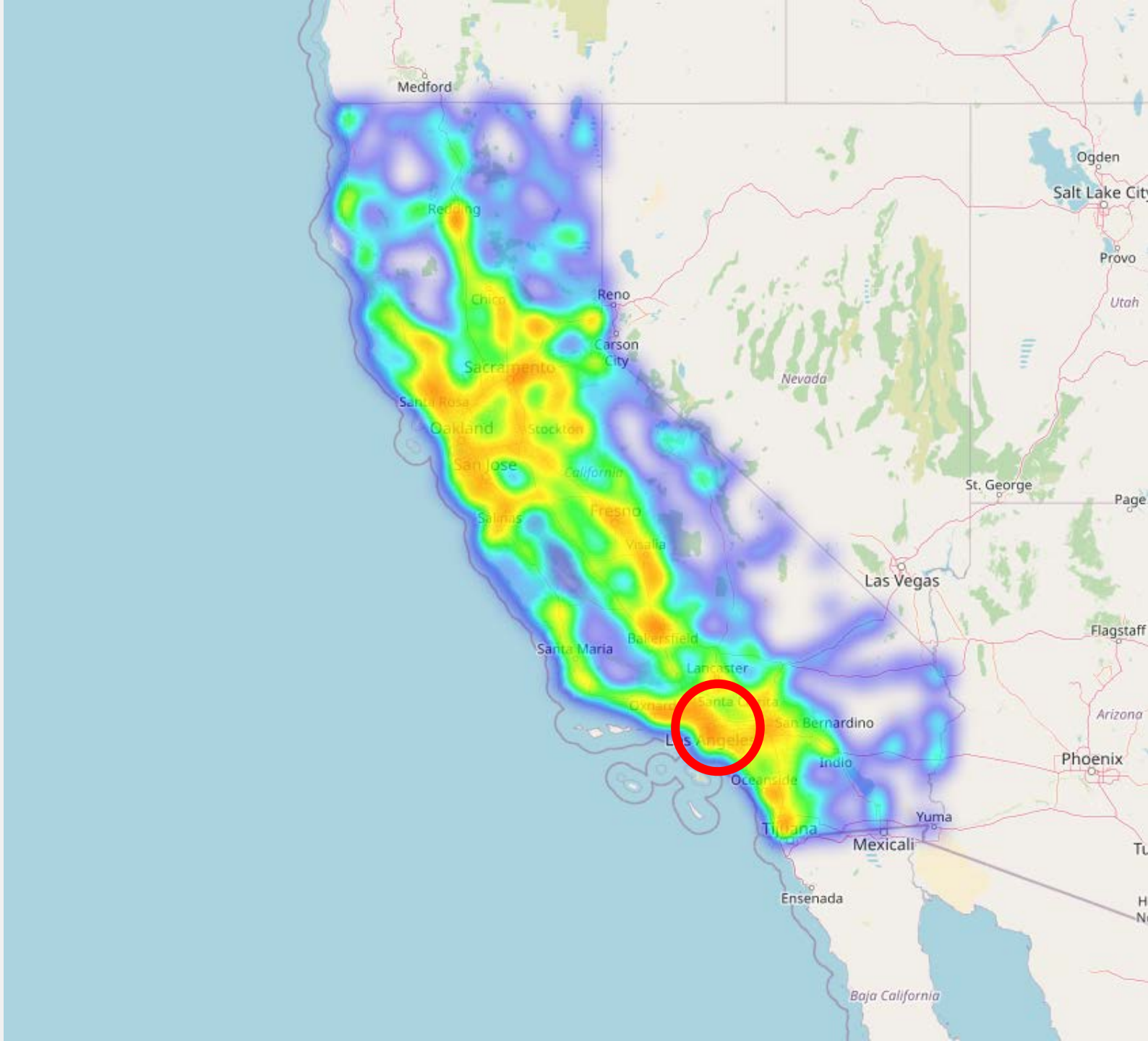
- We can see that areas that people generally believe to be dangerous areas to drive in are indeed hotspots.
- **San Francisco**



# CALIFORNIA ACCIDENT HEATMAP

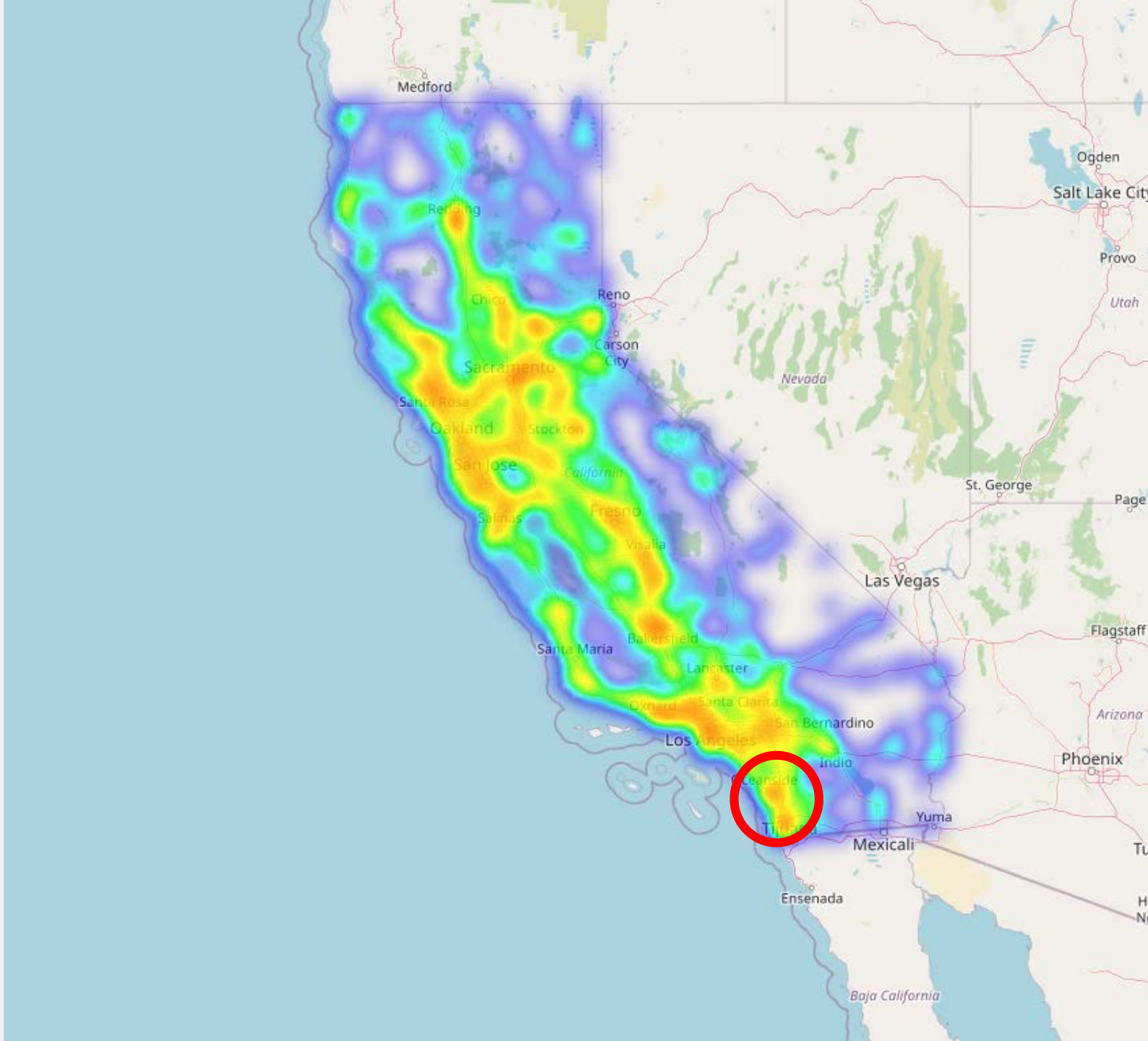
- We can see that areas that people generally believe to be dangerous areas to drive in are indeed hotspots.

- San Francisco, **Los Angeles**



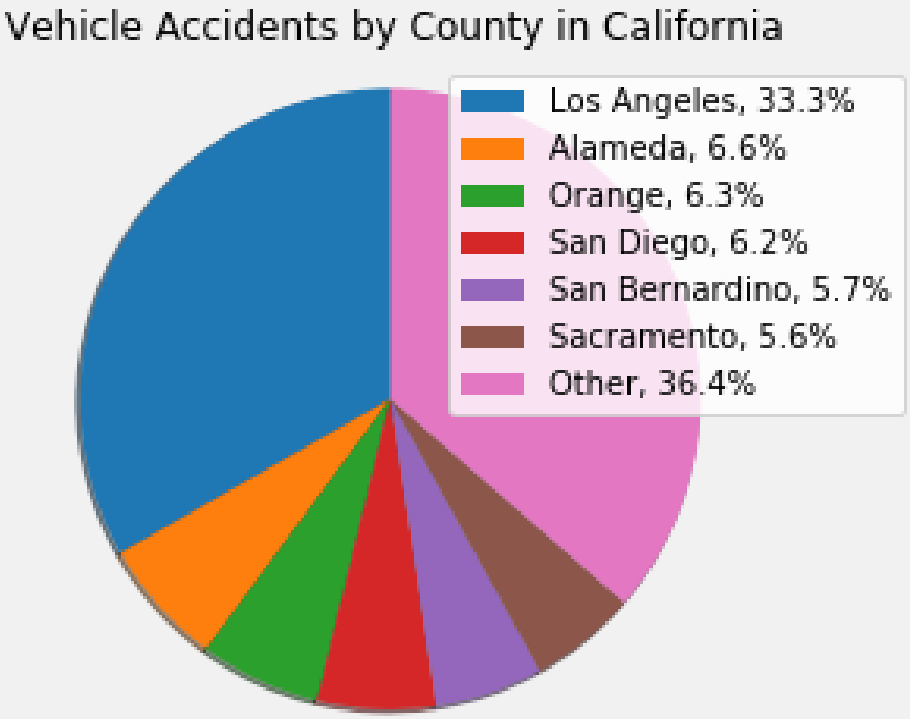
# CALIFORNIA ACCIDENT HEATMAP

- We can see that areas that people generally believe to be dangerous areas to drive in are indeed hotspots.
- San Francisco, Los Angeles, **San Diego**



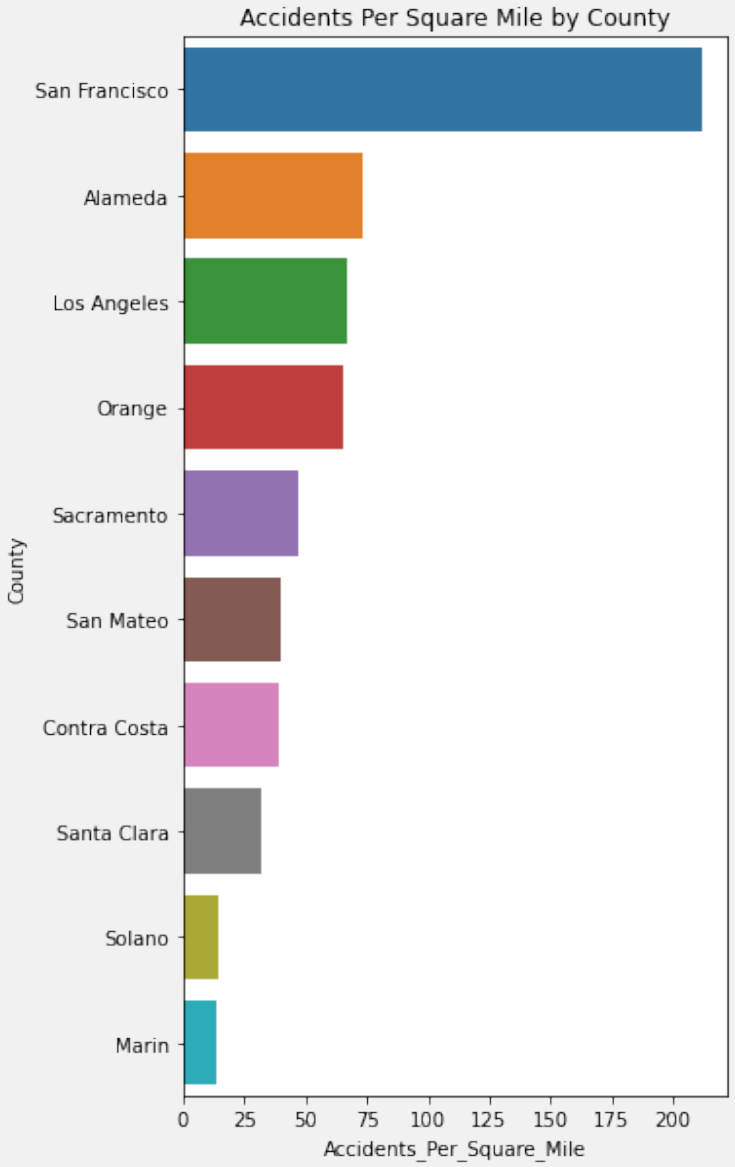
# COUNTIES WITH THE MOST ACCIDENTS

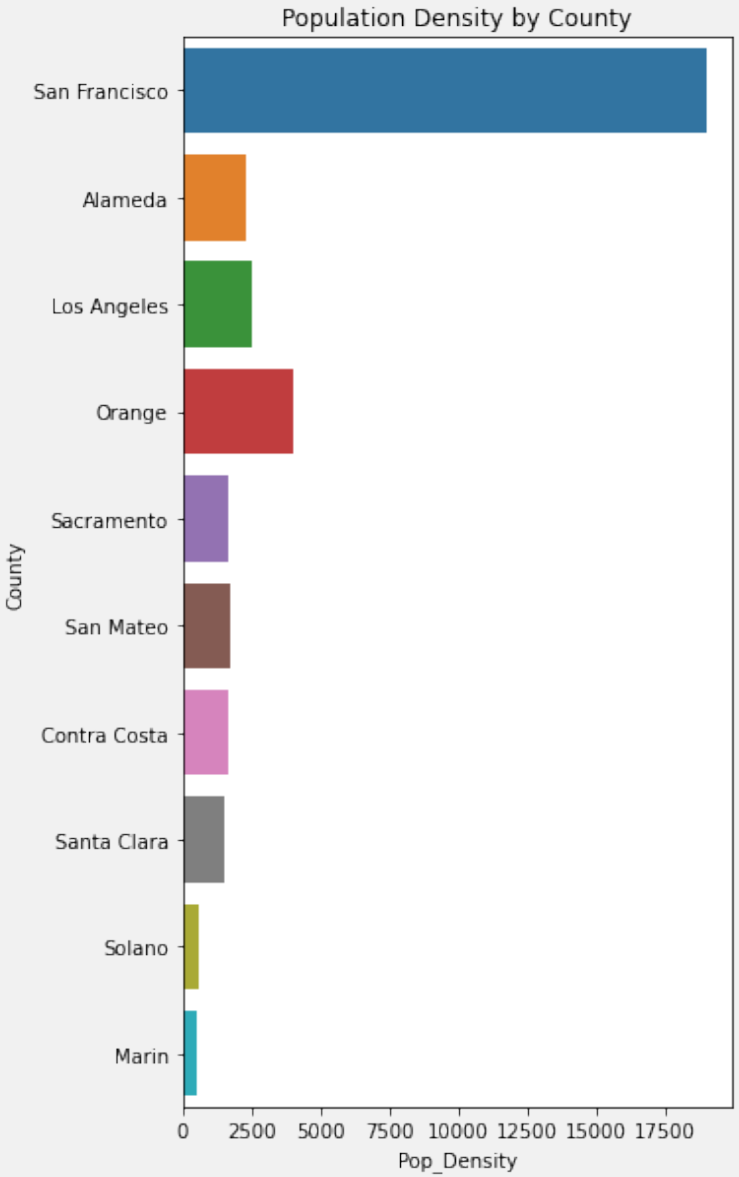
- The top 6 counties out of the 58 in California comprise nearly 63.6% of accidents in the state.
- However, these top 6 counties are also some of the most populous counties in California.
- Other possible ways to measure?



# ACCIDENTS PER SQUARE MILE

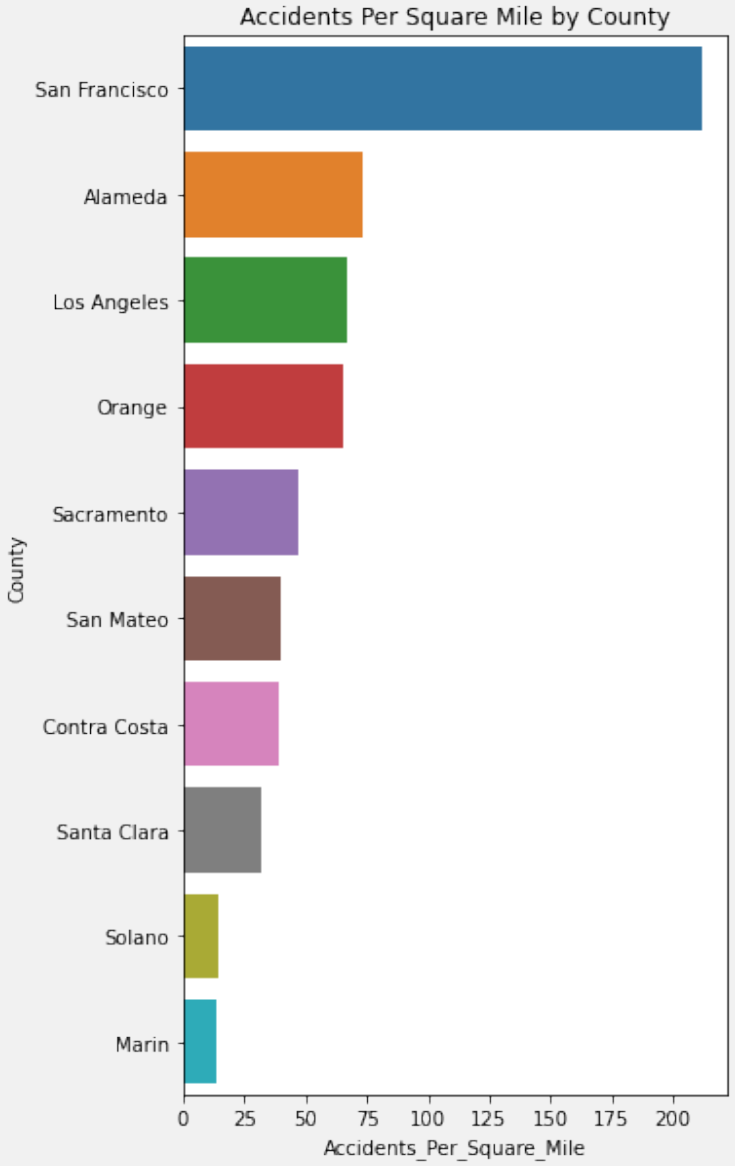
- There is a clear correlation between accidents per square mile and population density
- San Diego and Sacramento accounted for large number of accidents but are not listed in the top 12 counties when looking at per square mile.
- Orange county has a higher population than LA or Alameda yet lower accidents per square mile

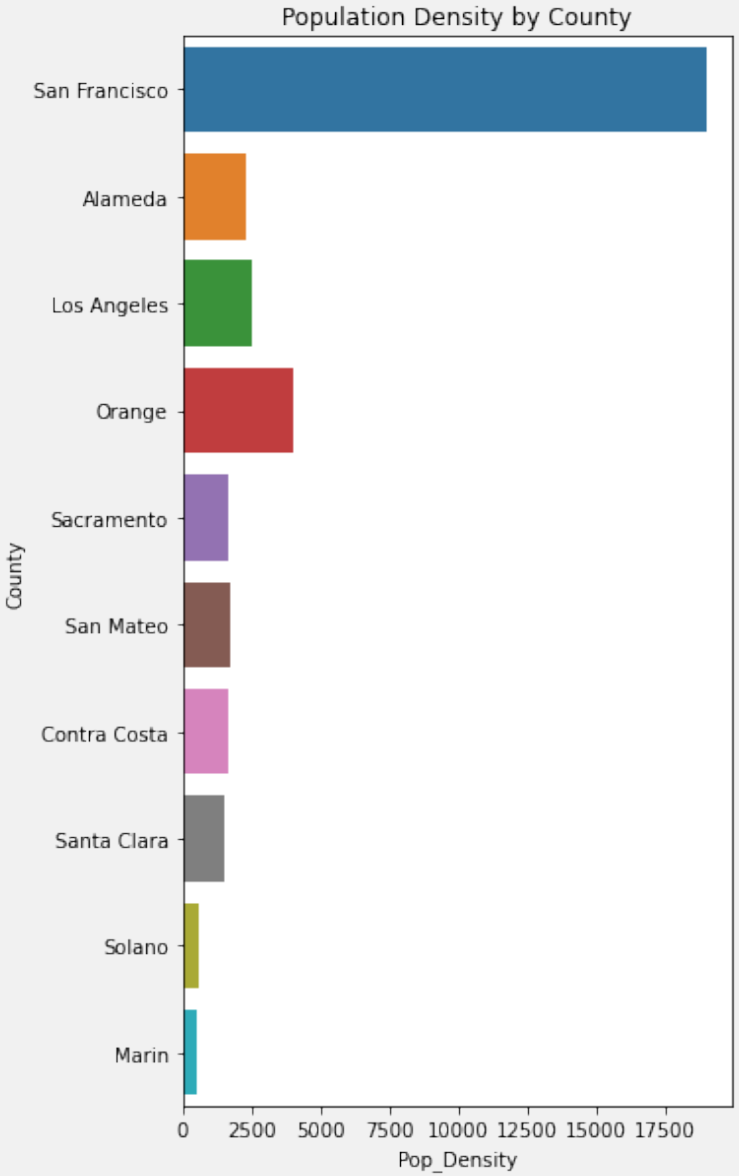


# ACCIDENTS PER SQUARE MILE

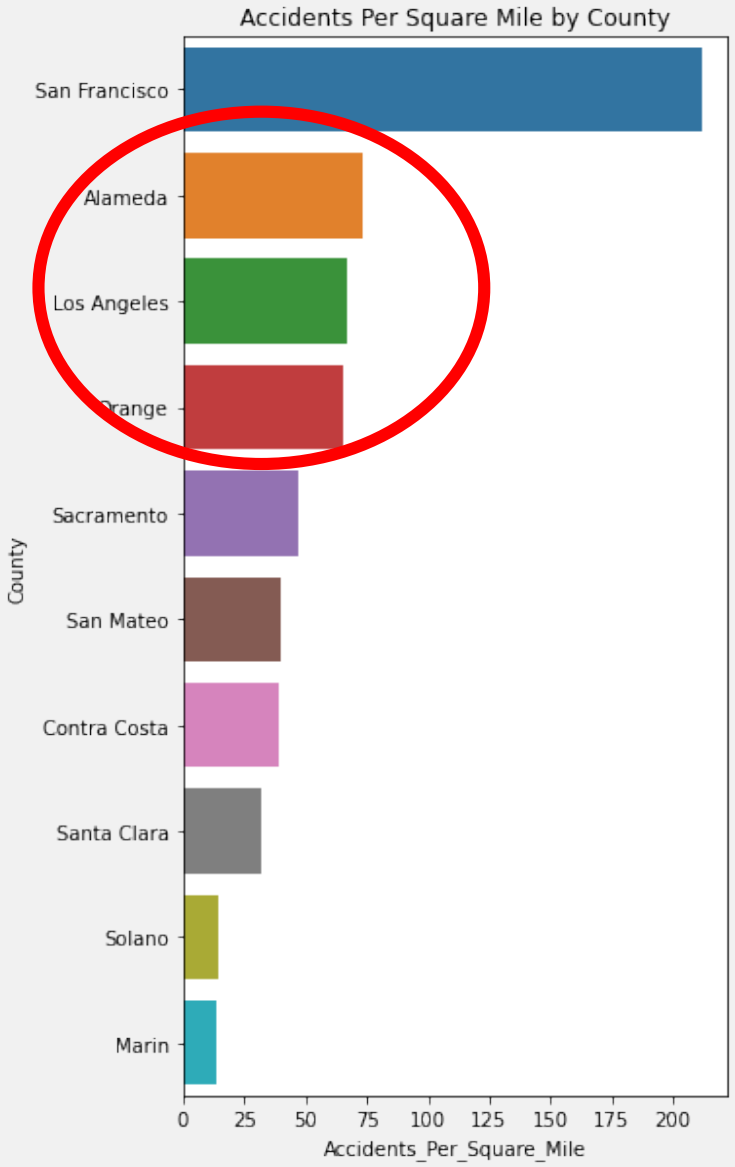
- There is a clear correlation between accidents per square mile and population density
- **San Diego and Sacramento accounted for large number of accidents but are not listed in the top 12 counties when looking at per square mile.**
- Orange county has a higher population than LA or Alameda yet lower accidents per square mile

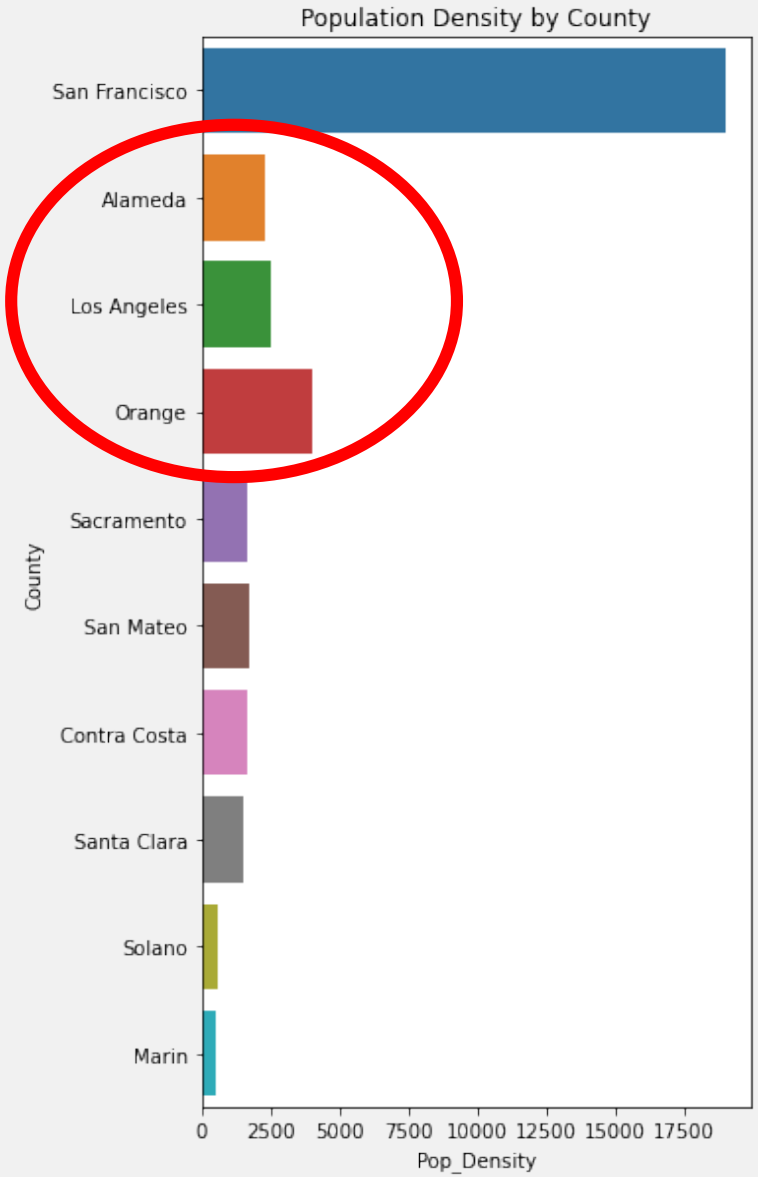


# ACCIDENTS PER SQUARE MILE

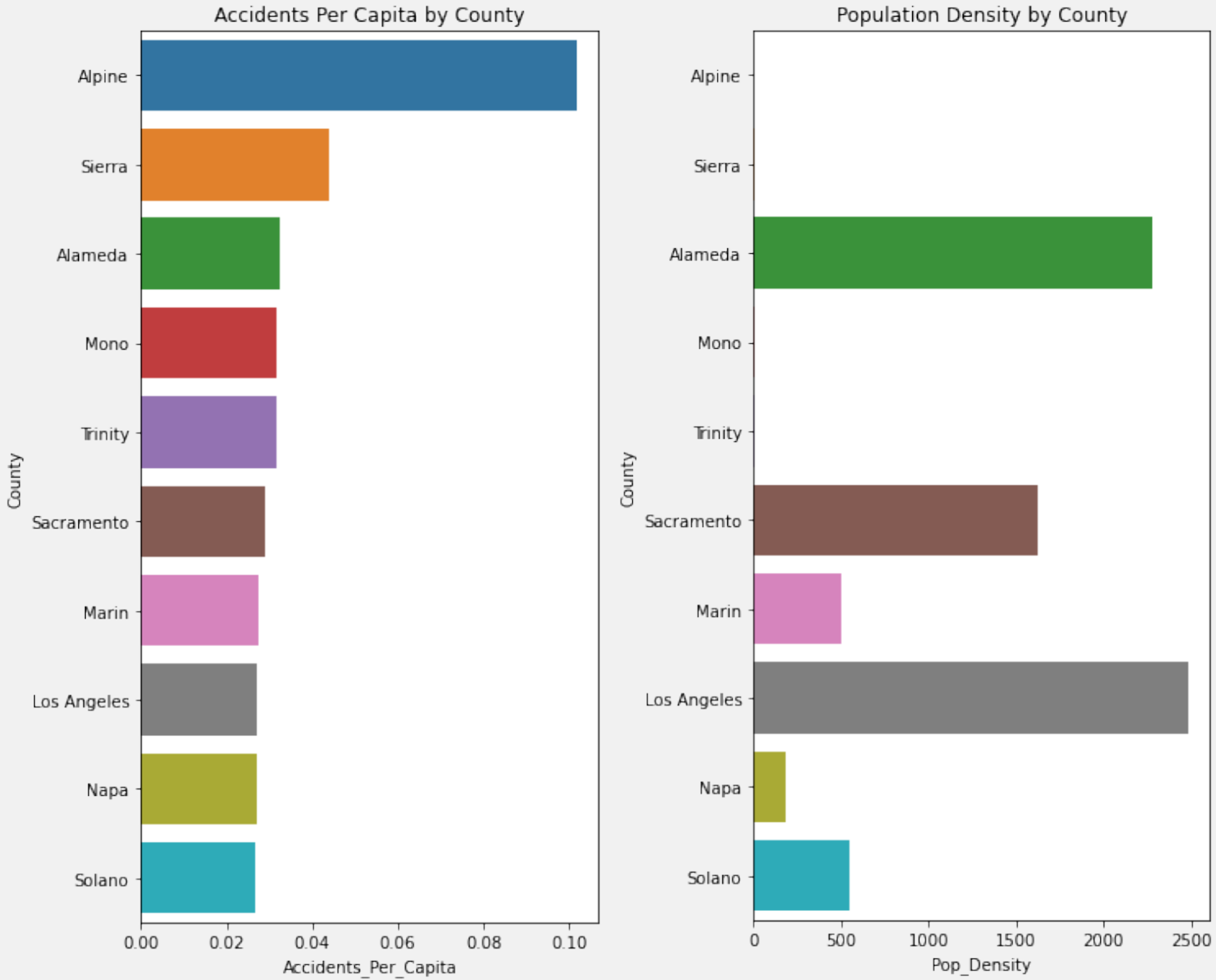
- There is a clear correlation between accidents per square mile and population density
- San Diego and Sacramento accounted for large number of accidents but are not listed in the top 12 counties when looking at per square mile.
- **Orange county has a higher population than LA or Alameda yet lower accidents per square mile.**



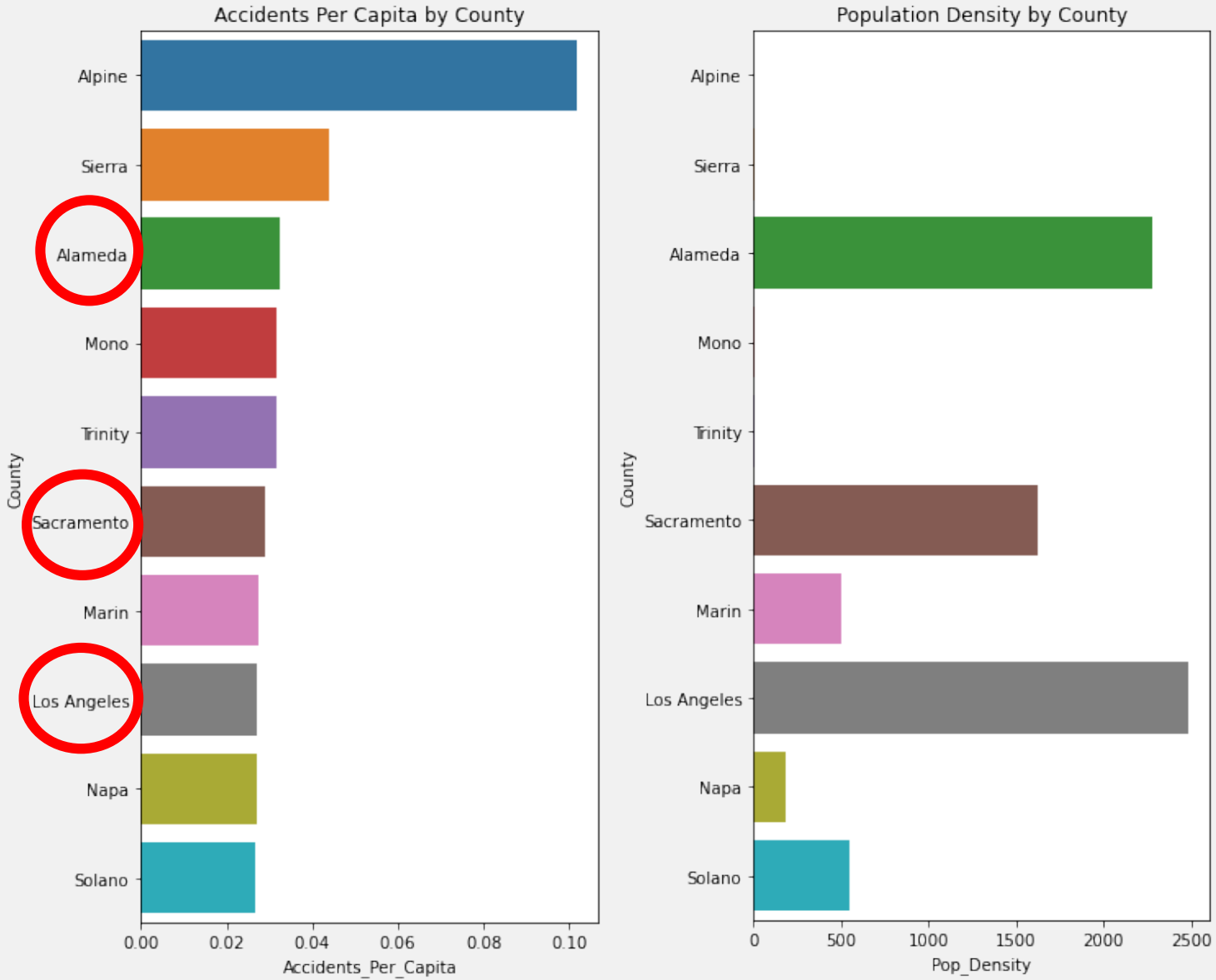
# ACCIDENTS PER CAPITA

- Using a per capita approach, we see that only three counties with high population appear.
- Counties like Alpine are listed at the top of this graph, this is because they are mountainous locations (i.e.: dangerous roads) and could have a high amount of traffic passing through.



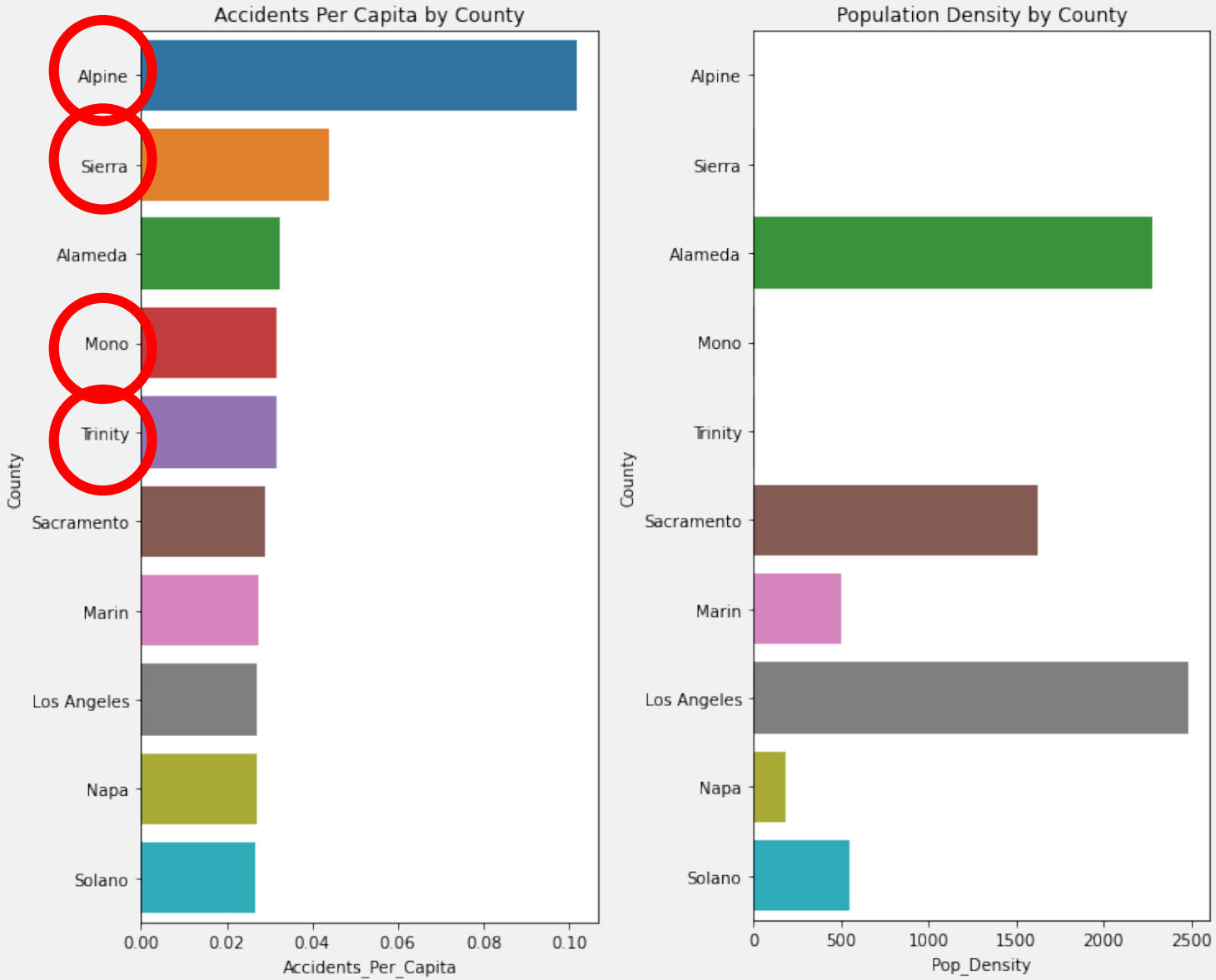
# ACCIDENTS PER CAPITA

- **Using a per capita approach, we see that only three counties with high populations appear.**
- Counties like Alpine are listed at the top of this graph, this is because they are mountainous locations (i.e.: dangerous roads) and could have a high amount of traffic passing through when compared to their population.

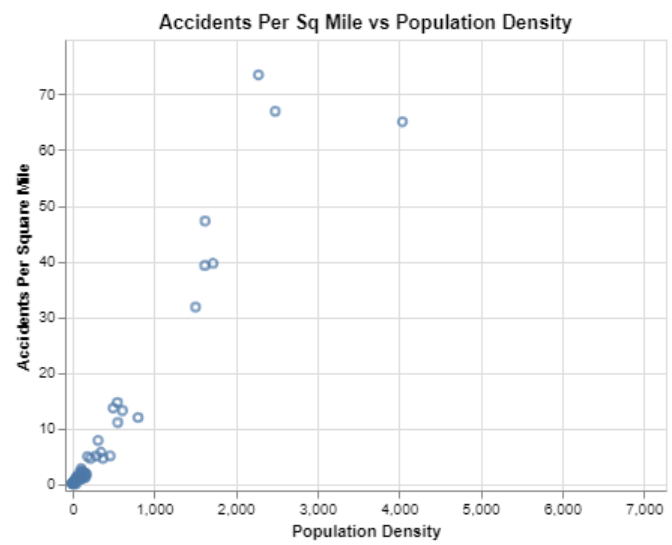


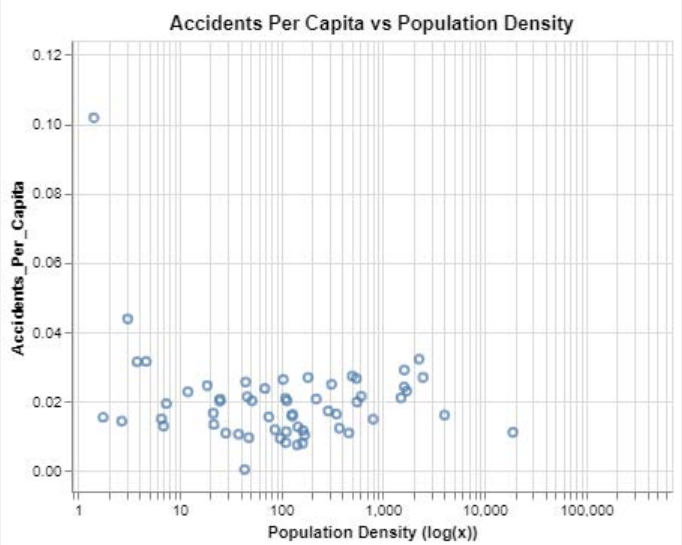
# ACCIDENTS PER CAPITA

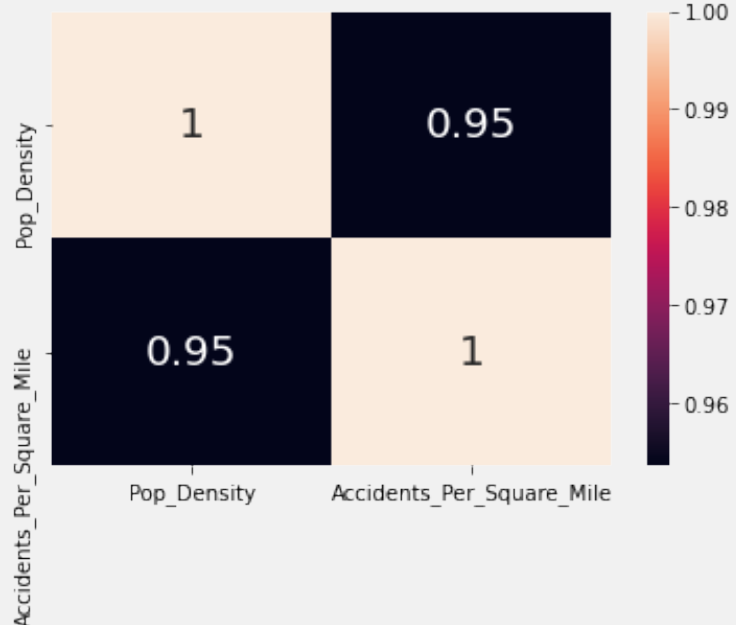
- Using a per capita approach, we see that only three counties with high population appear.
- **Counties like Alpine are listed at the top of this graph, this is because they are mountainous locations (i.e.: dangerous roads) and could have a high amount of traffic passing through.**

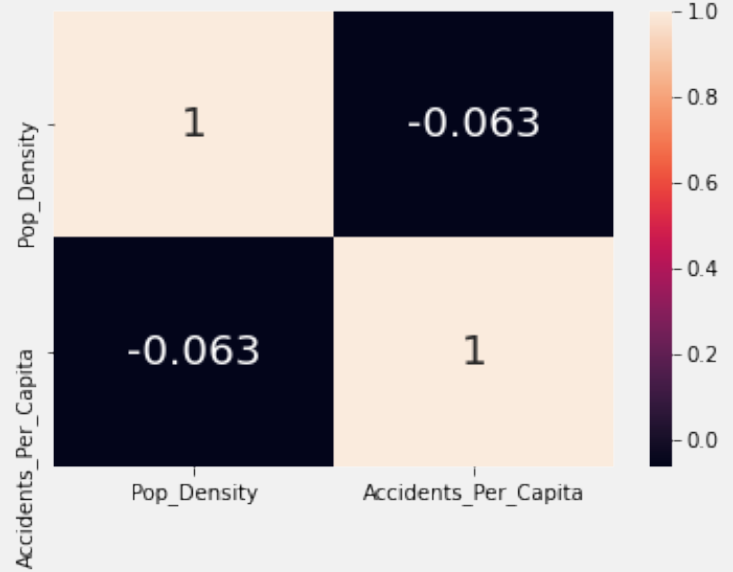


# DOES POPULATION DENSITY AFFECT ACCIDENT LIKELIHOOD?

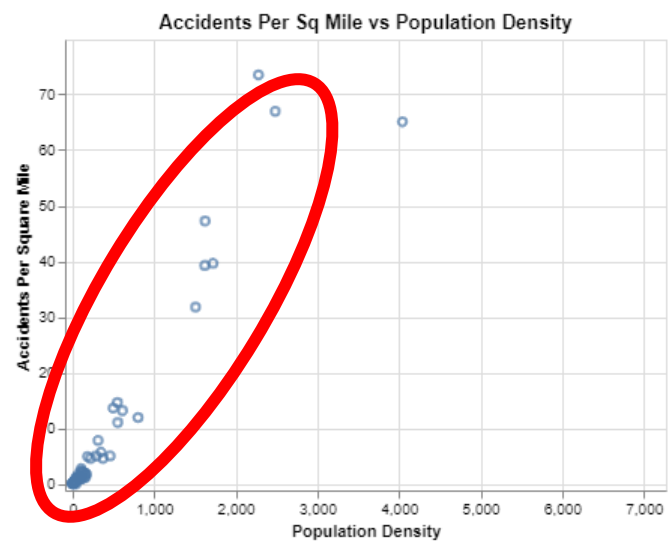


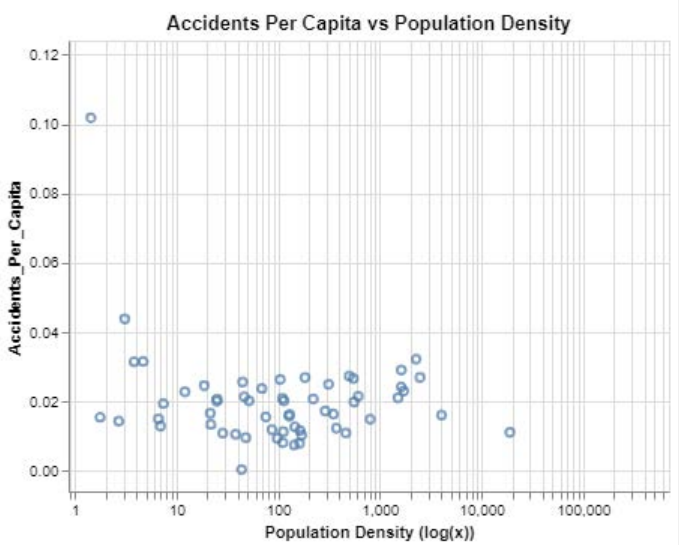



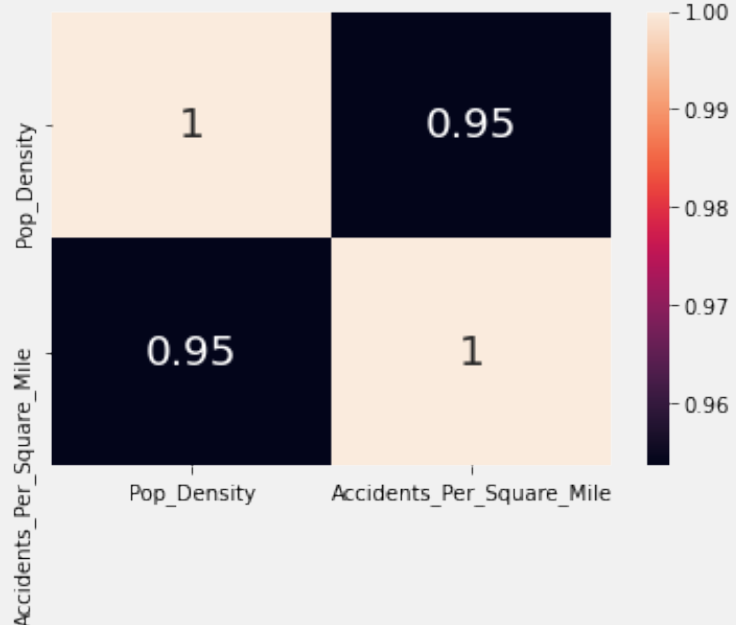



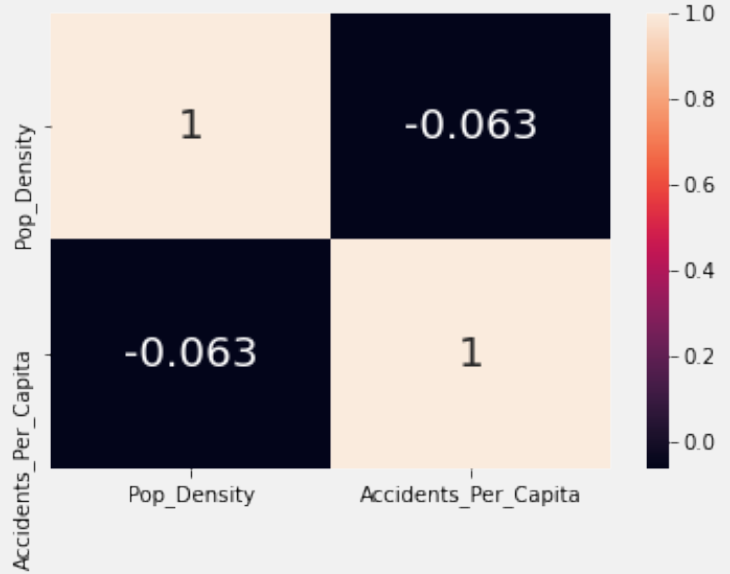



# DOES POPULATION DENSITY AFFECT ACCIDENT LIKELIHOOD?

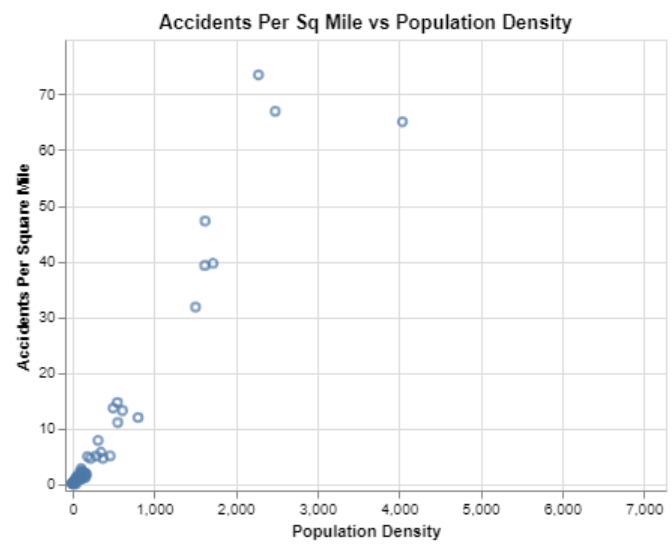


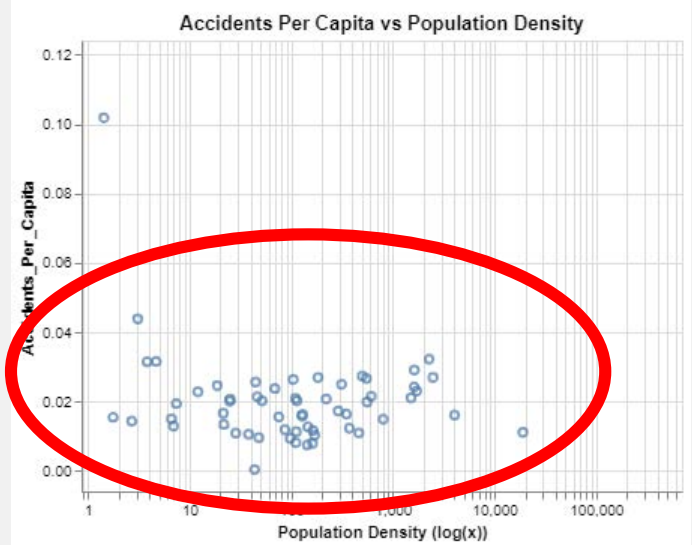



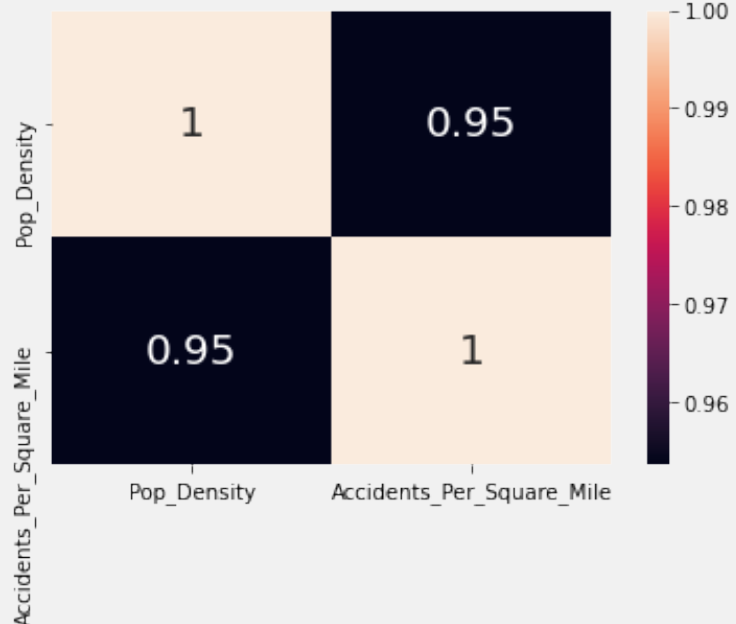



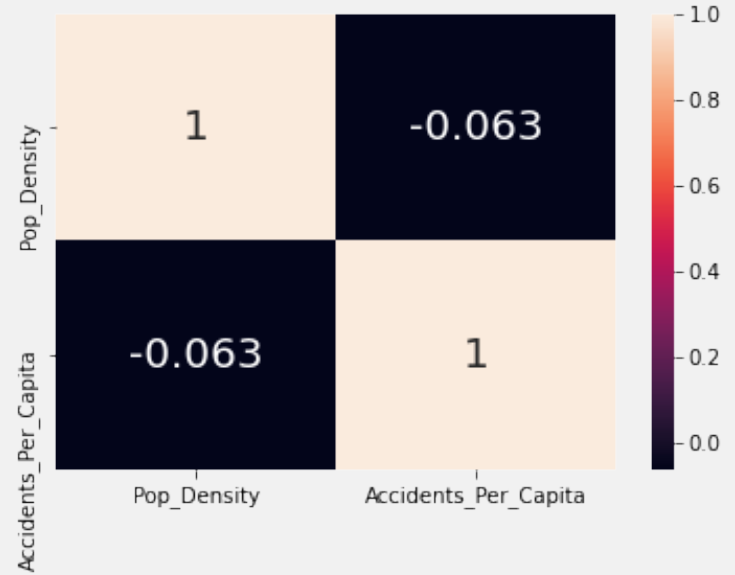



# DOES POPULATION DENSITY AFFECT ACCIDENT LIKELIHOOD?



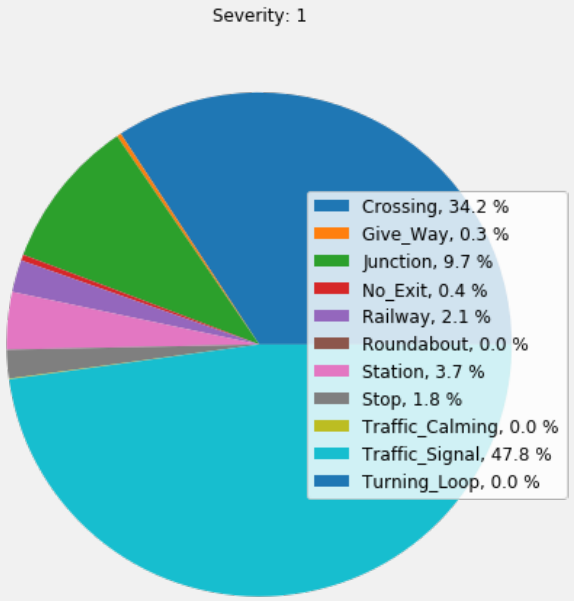



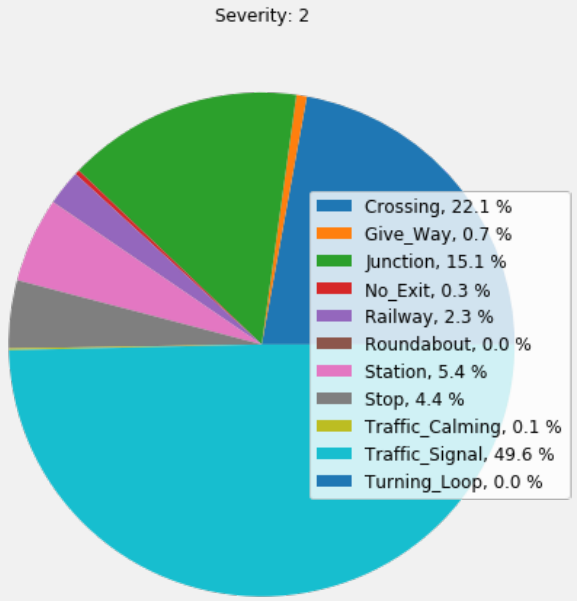



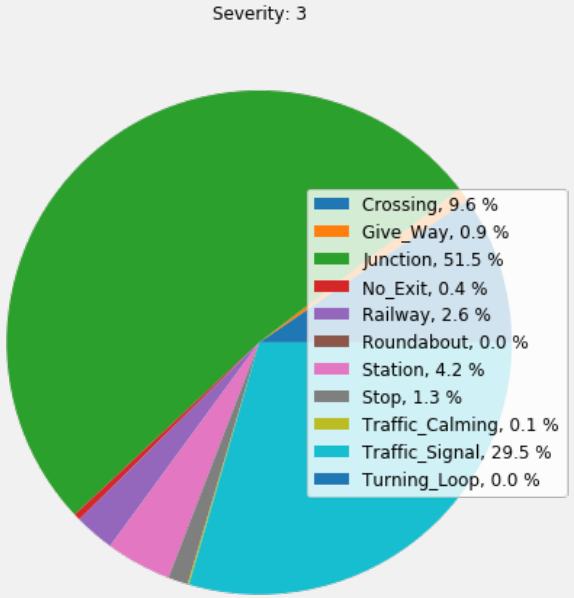



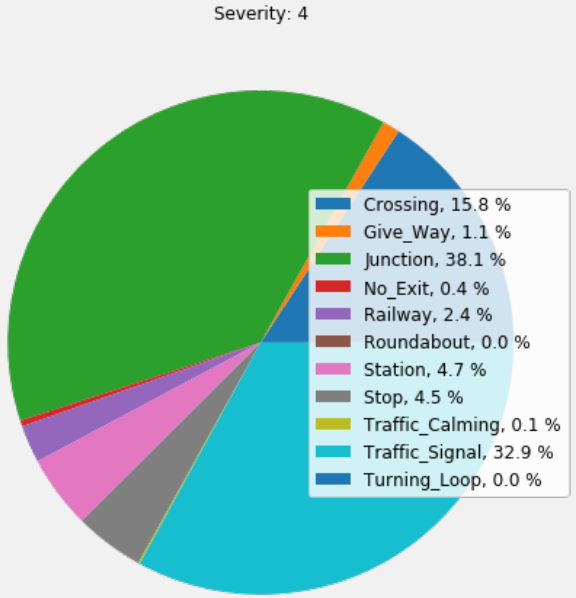
# HIGH-RISK LOCATIONS

- Crossings, traffic signals, and junctions are the majority contributors in accident location.
- Traffic signals and crossings comprise more of the total at lower severities.
- Junctions comprise the majority of accidents at higher severities.



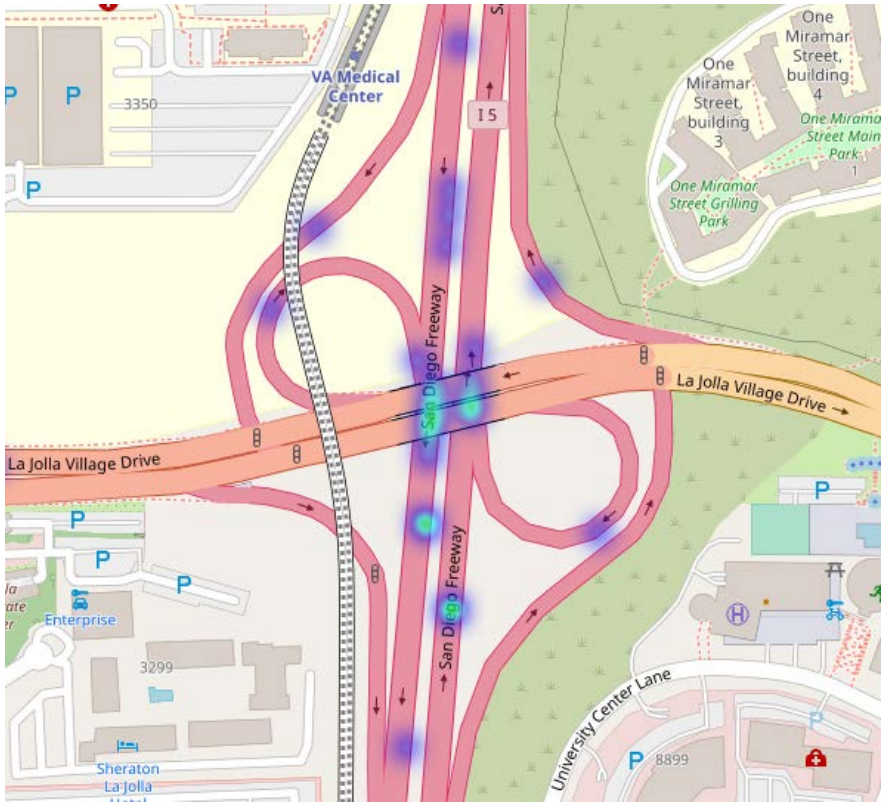
# CONCENTRATED ACCIDENT AREAS

- Connection of the Oakland Bay Bridge and the West Side of Yerba Buena Island had the most accidents (in all of CA) within 100 yards.
- On any given day there is approximately a 45% chance an accident will occur within 100 yards of that location

# CONCENTRATED ACCIDENT AREAS (LOCAL)

San Diego County

City of La Jolla



# CONCLUSIONS

**No Effect or Reduced Risk**

Visibility

**Increased Risk**

# CONCLUSIONS

**No Effect or Reduced Risk**

- Visibility was not found to affect the severity of an accident

Temperature

**Increased Risk**

# CONCLUSIONS

**No Effect or Reduced Risk**

- Visibility was not found to affect the severity of an accident

Month

**Increased Risk**

- Lower Temperatures were associated to increased severity

# CONCLUSIONS

**No Effect or Reduced Risk**

- Visibility was not found to affect the severity of an accident

Day of Week
Hour of Day

**Increased Risk**

- Lower Temperatures were associated to increased severity
- Months towards the end of the year (August-December) were found to be more likely

# CONCLUSIONS

## County Location

### No Effect or Reduced Risk

- Visibility was not found to affect the severity of an accident

### Increased Risk

- Lower Temperatures were associated to increased severity
- Months towards the end of the year (August-December) were found to be more likely
- The daily work commute is found to have the most accident occurrences

# CONCLUSIONS

## High-risk Locations

### No Effect or Reduced Risk

- Visibility was not found to affect the severity of an accident
- Higher populated counties such as LA were not found to increase likelihood of an accident per capita

### Increased Risk

- Lower Temperatures were associated to increased severity
- Months towards the end of the year (August-December) were found to be more likely
- The daily work commute is found to have the most accident occurrences
- Mountainous Regions are found to have a very high per capita rate of accidents

# THANK YOU

Questions or Comments?

**No Effect or Reduced Risk**

- Visibility was not found to affect the severity of an accident
- Higher populated counties such as LA were not found to increase likelihood of an accident per capita.
- Crosswalks and traffic signals area associated with lower severity accidents

**Increased Risk**

- Lower Temperatures were associated to increased severity
- Months towards the end of the year (August-December) were found to be more likely
- The daily work commute is found to have the most accident occurrences
- Mountainous Regions are found to have a very high per capita rate of accidents
- Junctions are hotspots for high severity accidents.