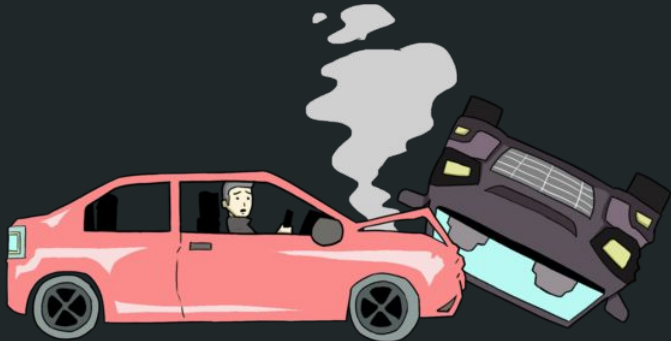# Behind The Wheel

Berk Güler
*Mustafa Barış Emektar*

The Problem

- **Exploration** of relationships between various features (such as weather conditions, time of day, and location) and their impact on accidents occured in U.S.

- **Actionable insights** for improving road safety and minimizing the consequences of accidents.

Dataset

US Accidents dataset from Kaggle
(2016 to 2023)

- Time Scope and Size
  - Accidents occurred in the 49 states of the USA.
  - 7.7 million accident records with 3.06GB of data.
- Rich Features
  - 46 different aspects such as the exact location, the weather and road conditions, when it happened.
- Severity Labels
  - Four level of accident severity (such as fatal, highly severe, moderate and minor)
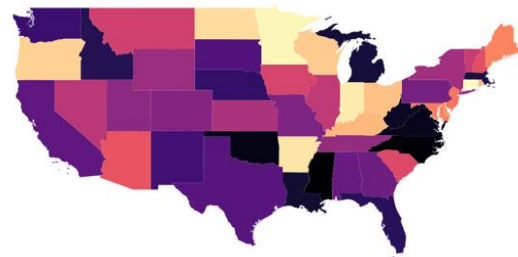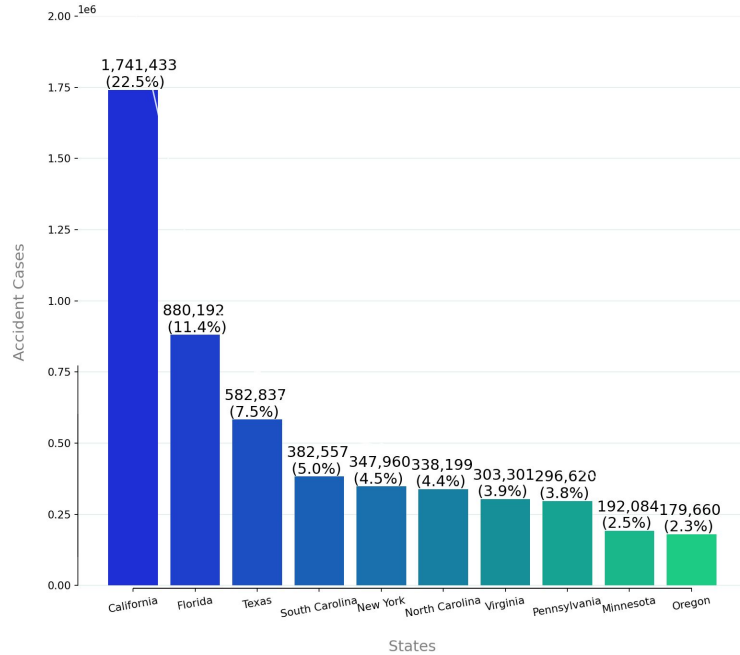

US Accident Analysis

kaggle

Solution

- Exploratory Data Analysis

- Relationship Discovery
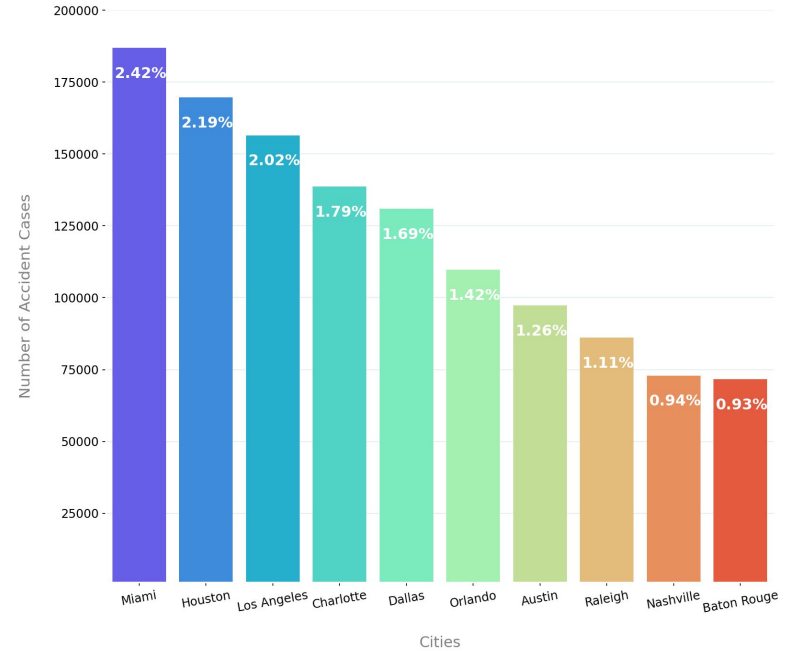
- Scaling Up

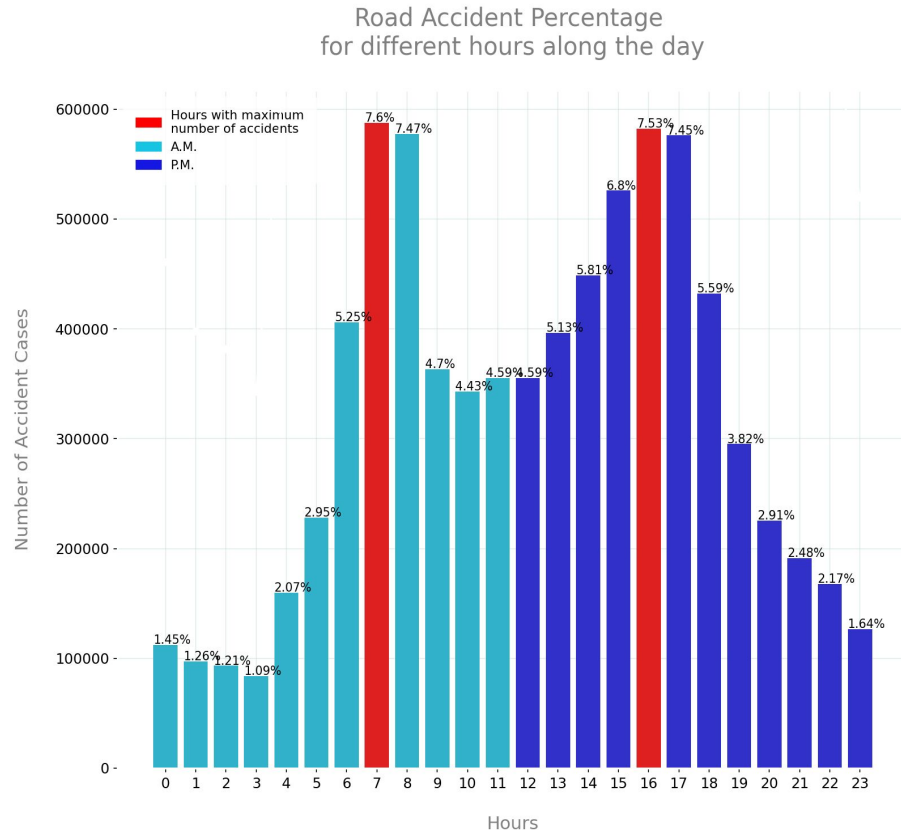- Visualization and Interpretation

Results

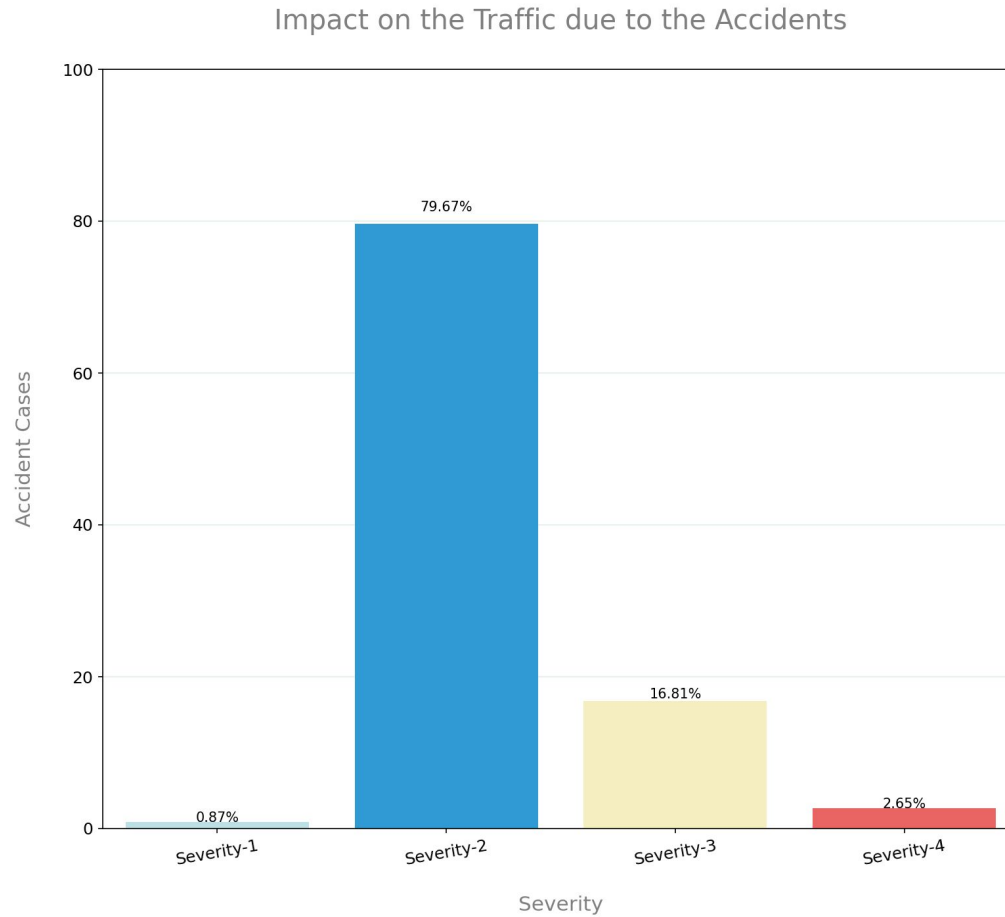## Top 10 States with most no. of Accident cases in US (2016-2023)

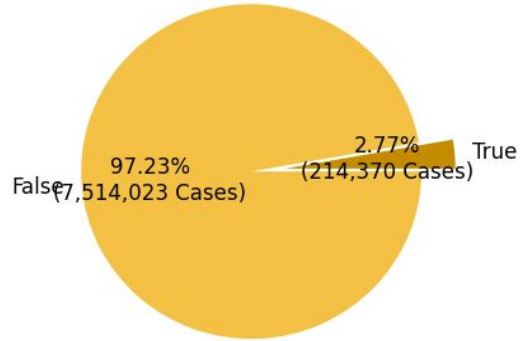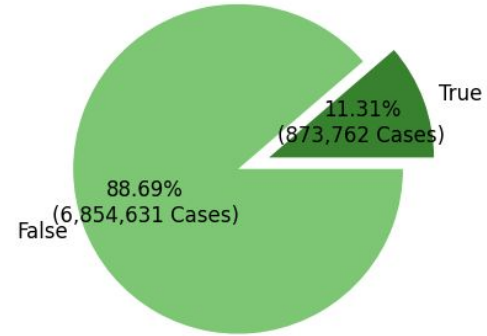## Top 10 Cities in US with most no. of Road Accident Cases (2016-2023)

Results

Road Accident Percentage
for different hours along the day

Results

Impact on the Traffic due to the Accidents

Results

Presence of Stop

2.77%
(214,370 Cases) True

97.23%
False (7,514,023 Cases)

Presence of Crossing

11.31%
(873,762 Cases) True

88.69%
(6,854,631 Cases)
False

Presence of Traffic_Signal

14.80%
(1,143,772 Cases) True

85.20%
(6,584,622 Cases)
False

Presence of Junction
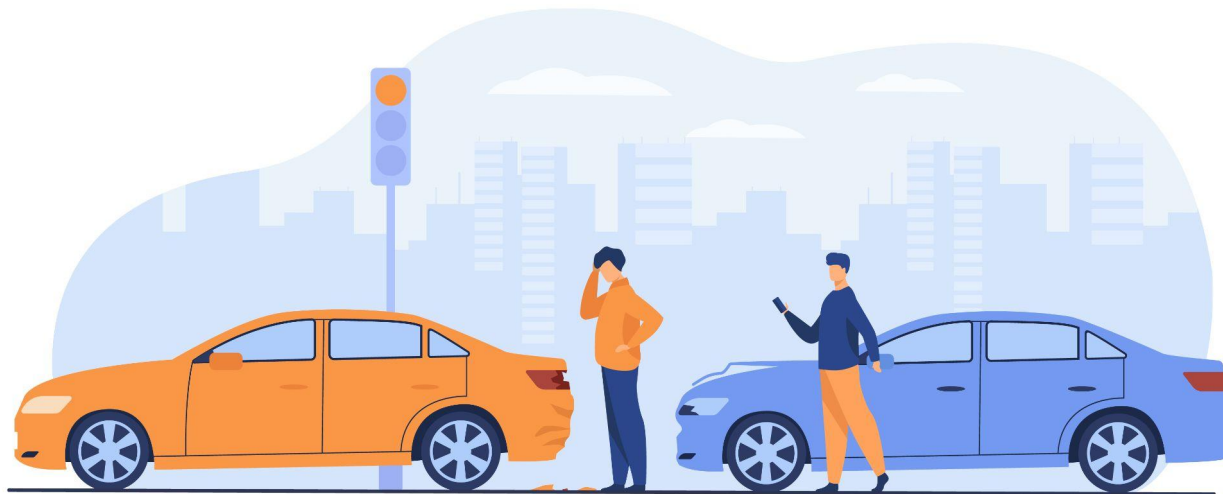
7.39%
(571,341 Cases) True

92.61%
False (7,157,052 Cases)

8
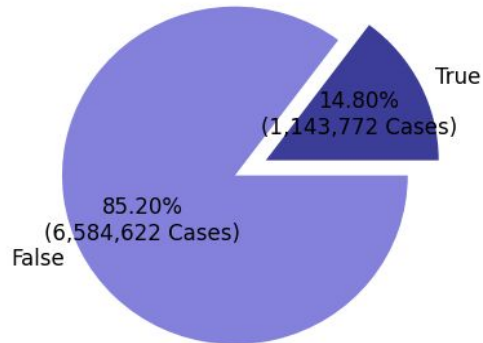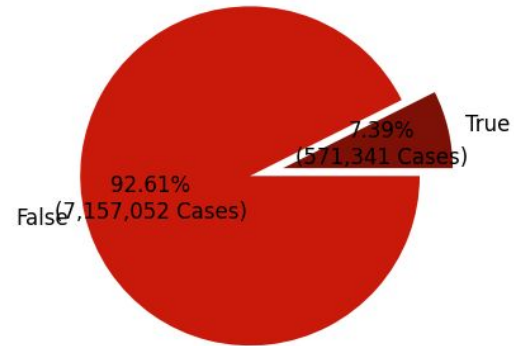
Conclusion

- Depend on multiple factors

- Other figures

- Predictions

- Apache Spark

THANK YOU

Berk Güler
Mustafa Barış Emektar