****US-Accidents Data Analysis****

****Team Members:****

Ivy Pham

Joseph Antony Bala Vineesh Reddy Pentareddy

Poornima Pulakandam

Srujana Patil

****Team Name: Group 8****

****Introduction to Problem:****

**The US-Accidents dataset can be used for various applications such as real-time car accident prediction, to determine car accidents prone locations, accident analysis and deriving cause and effect rules to predict car accidents, and studying the impact of weather or other environmental stimuli on accident occurrence. we are working on a dataset of countrywide car accidents, covers 49 states of the USA. The dataset represents data of accidents from February 2016 to Dec 2020, using multiple APIs that provide streaming traffic incident (or event) data. We can analyze the The most recent release of the dataset can also be useful to study the impact of COVID-19 on traffic behavior and accidents.**

**US Accident Dataset has a few variables which are categorical mention about the road condition or any object that was present on the accident scene or where exactly the accident occurred and the time of day. The other variables are quantitative which mention about how the weather is impacting the accident occurence, such as wind or precipitation. Severity, as a response variable which describes the severity of the accident. Severity assigns a value from 1 to 4 to on how much impact the accident had on traffic with 4 being the most amount of interference and 1 being least amount of interference.**

****Domain Data Description:****

****Source:**** <https://www.kaggle.com/sobhanmoosavi/us-accidents>

* ****ID -** This is a unique identifier of the accident record.**
* ****Severity** - Shows the severity of the accident, a number between 1 and 4, where 1 indicates the least impact on traffic (i.e., short delay as a result of the accident) and 4 indicates a significant impact on traffic (i.e., long delay).**
* ****Start\_Time** - Shows start time of the accident in local time zone.**
* ****End\_Time** - Shows end time of the accident in local time zone.**
* ****Distance(mi)** - The length of the road extent affected by the accident.**
* ****City** - Shows the city in address field.**
* ****County** -Shows the county in address field.**
* ****State** - Shows the state in address field.**
* ****Zipcode** - Shows the zipcode in address field.**
* ****Country** - Shows the country in address field.**
* ****Timezone** - Shows timezone based on the location of the accident (eastern, central, etc.).**
* ****Weather\_Timestamp** - Shows the time-stamp of weather observation record (in local time).**
* ****Temperature(F)** - Shows the temperature (in Fahrenheit).**
* ****Wind\_Chill(F)** - Shows the wind chill (in Fahrenheit).**
* ****Humidity(%)** - Shows the humidity (in percentage).**
* ****Pressure(in)** - Shows the air pressure (in inches).**
* ****Wind\_Direction** - Shows wind direction.**
* ****Wind\_Speed(mph)** - Shows wind speed (in miles per hour).**
* ****Precipitation(in)** - Shows precipitation amount in inches, if there is any.**
* ****Weather\_Condition** - Shows the weather condition (rain, snow, thunderstorm, fog, etc.).**

****Research Questions:****

1. **Research on how weather condition impacts on the accident occurence and time affect visibility which results in the number of accidents?**
2. **Is the accident rates in different states higher due to different Weather Condition?**
3. **how wind speed and wind direction contribute to accidents?**
4. **What are the most accident prone areas in each state?**
5. **number of accidents of a state as compared to its adjacent states?**
6. **Factors that effect accident severity.**

****Future Work:****

**This Analysis can be used to predict how climatic conditions are impacting the road accidents so that we can reduce the severity in advance. We can also predict the most accident prone areas state wide.**