University of Washington
BIOST 557 A Winter 2022: Applied Statistics & Experimental Design
Date: January 20th, 2023

Project Proposal

Topic: US Road Accident Analysis.
Team members:

- Aditi Kharkwal
- Akshit Miglani
- Ameya Bhamare
- Mohammad Danish Nadeem
- Tejal Kolte

Overview:
We will be statistically analysing the US road accidents dataset which records the road accidents across the US and includes various factors describing the accident and the stakeholders.
This analysis can be used by the following class of users:

1. Government to improve road conditions and manage traffic rules accordingly.
2. Applications like Uber, Lyft can manage their fares accordingly in case of an accident on the road.
3. Car drivers who are looking to avoid accidents at certain times or in certain weather conditions, be careful while driving in certain regions
4. Pedestrians to feel safe while crossing roads/walking in certain regions

Keeping these users in mind and with the attributes provided in our dataset we came up with the following questions to investigate further:

1. Is the distance of the road affected by traffic directly proportional to the severity of the accident?
2. A state wise detailed analysis on the effect of weather conditions like rainfall and wind speed on severity of accident.
3. What are the peak hours in a day during which accidents occur. Do more severe accidents occur at sunset/midnight/office leaving hours?
4. What areas in a city are prone to more severe accidents? Starting with time zones (east coast vs west coast) and then magnify it to a particular city like Seattle.
5. What is the severity of accidents near road signs?

Dataset source:
https://www.kaggle.com/datasets/sobhanmoosavi/us-accidents/versions/12?resource=download

Key attributes:

1. 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 because of the accident) and 4 indicates a significant impact on traffic (i.e., long delay).
2. Precipitation: Shows precipitation amount in inches if there is any.
3. Wind Speed: Shows wind speed (in miles per hour).
4. Start_time: Shows start time of the accident in local time zone.
5. End_time: Shows end time of the accident in local time zone. End time here refers to when the impact of accident on traffic flow was dismissed.
6. Stop: A POI annotation which indicates presence of stop sign in a nearby location.
7. Distance(mi): The length of the road extent affected by the accident.
8. City: Shows the city in address field.