Introduction:

Getting from one place to another might be easy with the evolution of automobiles, but not with a street congested with cars. Spending half of your car trip sitting in a jammed traffic sucks, but what’s worse is you know that you have only 30 minutes before the flight takes off, but can’t back up or fly over the long line in front of you. No metropolis with large population density can save itself from bad traffic conditions. With the influx of job opportunities and residents from other states over the past decades, residents in Chicago have long been suffering from the crowded traffic and long commute.

Traffic congestions might be due to many things and traffic accidents can be one of the most important reasons the traffic is delayed. While it is almost impossible to eliminate accidents occurrence, having a predictive model to predict the severance of the impact it might have on the traffic flow is useful 1) for traffic incident management to allocate resources accordingly and minimize the negative impact 2) for individual drivers to plan commute accordingly, i.e. to avoid the affected routes or to factor the delay into total travel time.

In this project, I plan to study the traffic event data in Chicago over the past 5 years to develop a machine learning model that takes in important variables that might affect traffic delay like weather, incident location etc. (which will be discussed further in the Data Section) to predict the severity level (1-4) summarizing the total impact it has on traffic delay. This project also aims to present visualizations to deliver more insights on accident hotspot locations and impact of environmental stimuli.