Project Description:	
Project Name: Do changes in your socioeconomic status impact your likelihood of being involved in a pedestrian – motor vehicle accident.	
Project Type: Data Analysis	**Start Date:** July 9, 2019
Author of this document: Jacob Geeves	**End Date:** August 27, 2019

Project Summary:

Every year pedestrians and transit users continue to be involved in motor vehicle collisions. Though the city continues to employ multiple methods to improve road safety for these users, the idea that inequality and safety are tied to the socioeconomic status of an area is prevalent. The purpose of this project is to study the relationship between a person's socioeconomic status and their likelihood to be involved in a pedestrian – motor vehicle accident. This information can then be leveraged in determining locations where road safety improvements will have the biggest impact in lowering this socioeconomic inequality in the city of Toronto.

Literature Review:

There are a number of papers and studies that have drawn the link between socioeconomic status and the likelihood of being involved in a vehicular accident. However, these studies have often taken a much broader or much narrower view of these impacts.

These studies have often either taken a much broader approach to the impact of socioeconomics on vehicular accidents by looking at all accidents or have focused on a much narrower subset of data while pulling additional variables such as subset of pedestrian accidents been much more focused taking into account the prevalence of road features in each socioeconomic differed based on socioeconomics of the area and its impact on child safety. Two examples of such studies include:

Chen HY, Ivers RQ, Martiniuk ALC, et al, Socioeconomic status and risk of car crash injury, independent of place of residence and driving exposure: results from the DRIVE Study Journal of Epidemiology & Community Health 2010;64:998-1003.

Rothman L, Cloutier M, Manaugh K, et al, Spatial distribution of roadway environment features related to child pedestrian safety by census tract income in Toronto, Canada Injury Prevention Published Online First: 01 April 2019. doi: 10.1136/injuryprev-2018-043125

Data Sources:

We will be using 2 primary data sources for this review:
- The KSI Data set available from the Toronto Police Service
  - This database contains all traffic data related to persons that have been killed or seriously injured in a traffic collision (KSI) for the last 10 years. (2008 – 2018)
  - We will be specifically focusing on Pedestrian and TTC/Municipal vehicle related accidents
  - http://data.torontopolice.on.ca/datasets/ksi
- 2016 Canadian Census data as pulled from SimplyAnalytics.com
  - Specifically, we will be extracting the average income for each burrow in Toronto

Methods:

This project fill focus on the use of three methods while performing the data analysis:
- Geospatial analysis:
  - We will be looking to map the accidents pulled from the KSI dataset
  - We will also be looking to map the average income as pulled from the census data.
- Clustering:
  - We will be looking to use clustering to determine if likelihood of an accident decreases based on income and if so, what are the optimal ranges.
- Correlation:
  - We will be looking to correlate accident frequency to income range
  - We will be looking to identify how that frequency changes year over year to see if any improvement is noticeable
  - We will be looking for correlation between income, accident rate and time of day/year to determine if these factors play a major role in the accident rate.

Limitations/Constraints:

The project has the following constraints:
- We will be limiting the study to a single census year. Meaning average income will be a static variable and we will infer a certain economic stability of an area over the 10-year timeframe in the KSI data set.
- Over the 10-year timeframe in the KSI data set, we do not know what changes or safety measures may have been implemented in a given area to help improve safety.

- We will be pulling data only for Toronto area.
- We will be limiting our review to pedestrian and TTC accidents and assuming all individuals involved in the accidents are from the surrounding communities.

The project has the following limitations:
- Income by area is limited to a single variable as we will only be using one year of census data.
- The income of individuals involved in an accident is not included in the KSI data set.
- There is no way to confirm if the individuals involved in an accident lives in the area where the accident occurred
- There is no way to confirm the rate at which people from a given community may walk in the community. This may impact our data if we assume people of lower socioeconomic status are more likely to walk or take transit.

Project Milestones:

- Project Proposal – July 15, 2019
- Sprint #1 – Data Collection – July 29, 2019
- Sprint # 2 – August 12, 2019
- Sprint #3 – August 20, 2019
- Final Project Submission – August 27, 2019