Predicting Fatal Accidents by Location

By: Edwin Johnson

1. Introduction

1.1 Background

Accidents are one the highest probability causes of death, comparable to fatalities caused by firearms in the United States. It is therefore reasonable and expected that governments and individuals take the necessary precautions and make investments to infrastructure in order to prevent accidental deaths. To optimize for this reduction, it becomes necessary to understand the features that cause fatalities and to explore possible reasons or correlated factors that make any accident a fatal one.

1.2 Problem

This project aims to find whether the location of the fatal accidents and their nearby venues affect whether the accident will become a fatal one.

1.3 Interest

Government officials would be interested as previously stated to decrease or even eliminate fatalities. If an accident is clustered around a certain area it would be in that government officials' best interest to act or improve infrastructure at that location. Individuals as well would be interested in obtaining the knowledge to avoid being another statistic.

2. Data

The datasets this project will use is from Kaggle dataset Killed or Seriously Injured (KSI) Toronto and the Foursquare API for location and venue details.

The Kaggle dataset has the location in latitude and longitude, date, time, types of accidents such as pedestrian, vehicle, or if alcohol was involved.

These features in addition to the nearest venues in 1 mile of the location of the accident would be used to make predictions as to where fatalities are clustered and the types of the accidents.