Analysis of Traffic Fatality Records

This manuscript (permalink) was automatically generated from uiceds/project-team-front-row@f40c8ad on October 24, 2024.

Authors

- **Justin Rebholz** ✉
  · jrebholz12
  Department of Civil and Environmental Engineering, University of Illinois

- **Cameron Kimber** ✉
  · ckimber2
  Department of Civil and Environmental Engineering, University of Illinois

- **Hannah Daggett** ✉
  · hed2
  Department of Civil and Environmental Engineering, University of Illinois

- **Riley Kelch** ✉
  · rileykelch
  Department of Civil and Environmental Engineering, University of Illinois

✉ — Correspondence possible via GitHub Issues or email to Justin Rebholz <rebholz4@illinois.edu>, Cameron Kimber <ckimber2@illinois.edu>, Hannah Daggett <hed2@illinois.edu>, Riley Kelch <rjkelch2@illinois.edu>.

Abstract

Description

The dataset that will be used for this project is the Fatality Analysis Reporting System created by the National Highway Safety Administration. The data will be obtained from the NHTSA's FARS database, which is publicly accessible.The FARS dataset is available in the CSV format. The specific dataset that our project will be focused on is labeled "accidents" and includes 32K+ instances and 52 columns. The columns descriptions are described in the Fatality Analysis Reporting System (FARS) Analytical User's Manual 1975-2015. Columns are described in the below table:

Column	Description
STATE	This data element identifies the state in which the crash occurred. The codes are from the General Services Administration's (GSA) publication of worldwide Geographic Location Codes (GLC).
ST_CASE	This data element is the unique case number assigned to each crash. It appears on each data file and is used to merge information from the data files together.
VE_TOTAL	This data element is the number of contact motor vehicles that the officer reported on the PAR as a unit involved in the crash.

Link: https://www.kaggle.com/datasets/nhtsa/2015-traffic-fatalities

Plan and Proposal

Using the FARS dataset we aim to understand the trends in traffic fatalities in a given year and what factors are affecting those trends. We will also look at how the different variables play a role in the severity of the accident and identify geographic regions that are more prone to accidents. The trends in traffic fatalities found through this project can be used to inform policy makers and ultimately decrease the number of traffic fatalities.

Exploratory Data Analysis

Description

The dataset that will be used for this project is the Fatality Analysis Reporting System created by the National Highway Safety Administration. The data will be obtained from the NHTSA's FARS database, which is publicly accessible.The FARS dataset is available in the CSV format. The specific dataset that our project will be focused on is labeled "accidents" and includes 32K+ instances and 52 columns. The columns descriptions are described in the Fatality Analysis Reporting System (FARS) Analytical User's Manual 1975-2015. Columns are described in the below table:

Column Abbreviation	**Description**
State	This data element identifies the state in which the crash occurred. The codes are from the General Services Administration's (GSA) publication of worldwide Geographic Location Codes (GLC).
ST_CASE	This data element is the unique case number assigned to each crash. It appears on each data file and is used to merge information from the data files together.
VE_TOTAL	This data element is the number of contact motor vehicles that the officer reported on the PAR as a unit involved in the crash.

Link: https://www.kaggle.com/datasets/nhtsa/2015-traffic-fatalities

Plan and Proposal

Using the FARS dataset we aim to understand the trends in traffic fatalities in a given year and what factors are affecting those trends. We will also look at how the different variables play a role in the severity of the accident and identify geographic regions that are more prone to accidents. The trends in traffic fatalities found through this project can be used to inform policy makers and ultimately decrease the number of traffic fatalities.