States with Worst Drivers

Mohammad R Zaghian

2022-06-20

knitr::opts\_chunk$set(echo = TRUE)

Data is measuring the States with the worst drivers normalized based on the total miles driven. The three main variables are the number of crashes, how much insurance payouts are, and how much premium charges are to drivers. One main issue with the data is that it is not excluding the drivers whose licences are not issued by the state in which they had their accident.The raw form of data is in CSV format. Two of the NHTSA’s primary data systems. Fatality Analysis Reporting System captures any crash that results in death of a occupant or non occupant within the 30 days of the accident. FARS Analysts collect, translate and process police accident reports, death certificates state vehicle registration files,etc. Some specific data characteristics are modified. For example,vehicle features. General estimate system data is generated by any police report regardless of severity of the crash as long as there is damage to property damage, injury or death. GES data collectors visit 410 police jurisdictions on a weekly basis and sample about 57000 police reports annually. Trained data entry personnel translate and input about more than 90 elements into a standard format in both systems.

library(readr)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.1 ──

## ✔ ggplot2 3.3.6 ✔ purrr 0.3.4  
## ✔ tibble 3.1.7 ✔ stringr 1.4.0  
## ✔ tidyr 1.2.0 ✔ forcats 0.5.1

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()

library(knitr)

Here we use the read\_csv from readr package which works faster than normal read.csv

crashdata <-read\_csv("https://raw.githubusercontent.com/fivethirtyeight/data/master/bad-drivers/bad-drivers.csv")

## Rows: 51 Columns: 8  
## ── Column specification ────────────────────────────────────────────────────────  
## Delimiter: ","  
## chr (1): State  
## dbl (7): Number of drivers involved in fatal collisions per billion miles, P...  
##   
## ℹ Use `spec()` to retrieve the full column specification for this data.  
## ℹ Specify the column types or set `show\_col\_types = FALSE` to quiet this message.

crashdatarevised<-rename(crashdata, "Number of drivers in fatal collisions/billion miles"="Number of drivers involved in fatal collisions per billion miles","Percentage of speeing drivers in fatal collisions"="Percentage Of Drivers Involved In Fatal Collisions Who Were Speeding","Percentage Of alcohol impraired drivers in fatal collisions"="Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired","Percentage Of focused Drivers In Fatal Collisions "="Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted","Percentage Of Drivers In Fatal Collisions wihout prior accidents"="Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents")

This dataframe has 51 rows and 8 columns. The names of the columns and a brief description of each are in the table below:

name<-c("Number of drivers in fatal collisions/billion miles","Percentage of speeding drivers in fatal collisions","Percentage Of alcohol impraired drivers in fatal collisions","Percentage Of focused Drivers In Fatal Collisions","Percentage Of Drivers In Fatal Collisions wihout prior accidents","Car Insurance Premiums","Losses incurred by insurance companies for collisions per insured driver")  
description<-c("Provides the number of drivers involved in fatal collisions per billion miles","Percentage of drivers involved in fatal collisions who were speeding","Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired","Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted","Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents","Car Insurance Premiums","Losses incurred by insurance companies for collisions per insured driver")  
crash\_tbl<-data.frame(name,description)  
knitr::kable(crash\_tbl,align = "l")

| name | description |
| --- | --- |
| Number of drivers in fatal collisions/billion miles | Provides the number of drivers involved in fatal collisions per billion miles |
| Percentage of speeding drivers in fatal collisions | Percentage of drivers involved in fatal collisions who were speeding |
| Percentage Of alcohol impraired drivers in fatal collisions | Percentage Of Drivers Involved In Fatal Collisions Who Were Alcohol-Impaired |
| Percentage Of focused Drivers In Fatal Collisions | Percentage Of Drivers Involved In Fatal Collisions Who Were Not Distracted |
| Percentage Of Drivers In Fatal Collisions wihout prior accidents | Percentage Of Drivers Involved In Fatal Collisions Who Had Not Been Involved In Any Previous Accidents |
| Car Insurance Premiums | Car Insurance Premiums |
| Losses incurred by insurance companies for collisions per insured driver | Losses incurred by insurance companies for collisions per insured driver |

column\_pick3 <- select(crashdatarevised,"Number of drivers in fatal collisions/billion miles","Car Insurance Premiums ($)","Losses incurred by insurance companies for collisions per insured driver ($)")

Summarytable<-summary(column\_pick3)  
Summarytable

## Number of drivers in fatal collisions/billion miles Car Insurance Premiums ($)  
## Min. : 5.90 Min. : 642.0   
## 1st Qu.:12.75 1st Qu.: 768.4   
## Median :15.60 Median : 859.0   
## Mean :15.79 Mean : 887.0   
## 3rd Qu.:18.50 3rd Qu.:1007.9   
## Max. :23.90 Max. :1301.5   
## Losses incurred by insurance companies for collisions per insured driver ($)  
## Min. : 82.75   
## 1st Qu.:114.64   
## Median :136.05   
## Mean :134.49   
## 3rd Qu.:151.87   
## Max. :194.78

Sources

<https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/812032>

<https://fivethirtyeight.com/features/which-state-has-the-worst-drivers/>