R Project Report

Nikita Belbase

4/28/2020

## Road Accidents Study in the UK

### Packages used

library(ggplot2)  
library(knitr)  
library(shiny)  
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

### Summary

Weather is one of the important factors that impact road accidents among other factors like driving under influence, mobile phone use, speeding, etc. In Britain, extreme weather conditions contribute a total of 22% to the total number of road accident victims. A majority of fatal accidents occur on icy and wet pavements and during heavy rainfall. However, this study suggests that among the total number of accidents that were observed, 81.73% of accidents happened in fine weather conditions. This report is conducted to know if the change in climate increases/decreases road injuries.

This study can be useful to educate the drivers on how to operate under certain weather conditions. It can also be used by vehicle manufactures to safeguard the vehicles by add extra safety measures.

According to the Department for Transport of the United Kingdom, in 2012 the number of casualties was comparatively lower than the previous years. As 2012 experienced heavy rainfall, the number of road users especially motorcyclists and cyclists reduced. This analysis of the report manifests that when the weather conditions are not favorable, people do not tend to go out. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/358190/rrcgb2013-03.pdf>

“Nevertheless, when inclement conditions are experienced, the general inclination is to assume that accident numbers will increase by the very presence of that particular hazard. However, the relationship between adverse weather and road accidents is far from straightforward” <https://rmets.onlinelibrary.wiley.com/doi/pdf/10.1017/S1350482799001139>

This research aims to find:  
• The kind of weather that contributes to the highest number of road accidents casualties.  
• The average number of casualties affected by the variation in weather.  
• The count of road accidents grouped by their severity.  
• The number of vehicles involved in road accidents due to different types of weather conditions.  
The researcher’s opinion towards this study was following the general notion towards accidents caused by hazardous weather i.e. severe weather conditions will cause a greater number of road accidents as well as casualties. The data for this study was sourced from Kaggle website, linked [here](https://www.kaggle.com/akshay4/road-accidents-incidence)

The information available on the Kaggle website was studied to get hold of a basic understanding of the dataset. In addition to that, the support file that was downloaded was also studied to analyze the data further. For example, under column accident\_severity, ‘1’ stands for Fatal, ‘2’ stands for Serious, and ‘3’ stands for Slight. Microsoft Excel was used to apply filters and view the data.

This study depicts that the highest number of road accidents (233216) occurred in fine weather conditions with no high winds followed by raining with no high winds (31375). Whereas, the least number of accidents (357) took place while it was snowing with high winds.

The mean number of causalities caused by the variation in weather is 1.61. The median that defines a value that lies at the midpoint of the frequency distribution of the assigned value is 1.

The study shows that out of 285331 accidents that were researched, 3648 were Fatal and 242477 accidents were classified as Slight.

severityFreq <- group\_by(UKAccidents, accident\_severity)  
knitr::kable(summarise(severityFreq, n()))

|  |  |
| --- | --- |
| accident\_severity | n() |
| 1 | 3648 |
| 2 | 39206 |
| 3 | 242477 |

The number of vehicles that were involved in the road accidents caused by change in weather state are calculated. Serious number of accidents (492823) were taken place while the climate was fine with no high winds while during snowfall and high winds, least number of vehicles (820) were counted.

freq <- group\_by(UKAccidents, weather\_conditions)  
knitr::kable(summarise(freq, n()))

|  |  |
| --- | --- |
| weather\_conditions | n() |
| Fine + high winds | 4716 |
| Fine no high winds | 233216 |
| Fog or mist | 1404 |
| Other | 3887 |
| Raining + high winds | 4627 |
| Raining no high winds | 31375 |
| Snowing + high winds | 357 |
| Snowing no high winds | 1050 |
| Unknown | 4699 |

### Introduction

The purpose of the research is to know the impact of weather change on road accidents. The report is not only limited to disclose the frequency of road accidents but it also calculates the number of vehicles affected in the road accidents. Moreover, the report also aims to classify the severity caused by the road accident. The object of this project is to apprise the reader about the importance of safe driving.

### Literature Review

This study has taken 9 different kind of weather in account for the analysis of the role of various weather in road accidents. Usually people think that there is more risk of getting into accident when the weather is hazardous. In a research conducted in The Great Britain in 2018, 7% of the accident occurred due to weather whereas 40% of accidents took place because the driver failed to look properly, followed by speeding (21%), carelessness (16%), and maneuver (13%). <https://www.statista.com/statistics/323079/contributing-factors-leading-to-road-accidents-in-great-britain-uk/>

### Theory

The null and alternative hypothesis of the report is stated as: H0 (Null Hypothesis): Hazardous weather conditions contribute to an increase in road accidents. H1 (Alternative Hypothesis): Hazardous weather conditions do not contribute to an increase in road accidents.

### Data

Road Accidents Incidence data was originated from <https://data.gov.uk/> and sourced from [Kaggle](https://www.kaggle.com/akshay4/road-accidents-incidence). The dataset in Kaggle was cleaned and merged therefore further flittering was not required.

### Methodology

The number of road accidents and the weather data related to those accidents was available in the same dataset that was sourced from Kaggle. So, using Microsoft Excel, the data was viewed to read and understand the data. Further details available on the Kaggle website (linked above) was also studied.

After having a thorough understanding of the data, the data was imported to R studio. In R studio, a subset of the data was chosen that contained only four columns using the following command:

head(UKAccidents)

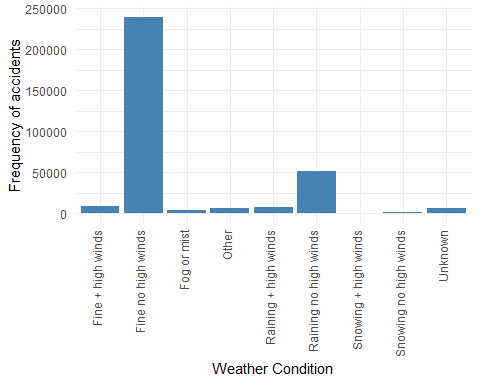
## weather\_conditions number\_of\_casualties number\_of\_vehicles accident\_severity  
## 1 1 1 2 3  
## 2 1 1 2 3  
## 3 1 1 2 3  
## 4 1 1 1 3  
## 5 1 1 2 3  
## 6 1 1 2 3

Then, the values for the columns weather\_conditions and accident\_severity were inserted using the following command.

A bar graph to visualize the number of accidents caused by different weather conditions was plotted as follows:

plot <- ggplot(data.frame(UKAccidents),aes(x= weather\_conditions, y= number\_of\_casualties))+   
 geom\_bar(stat = "identity", fill= "steelblue") +  
 theme\_minimal() +  
 ylim(0,240000) +  
 xlab("Weather Condition")+  
 ylab("Frequency of accidents")+  
 theme(axis.text.x=element\_text(angle=90 ,hjust=1,vjust=0.1))  
plot

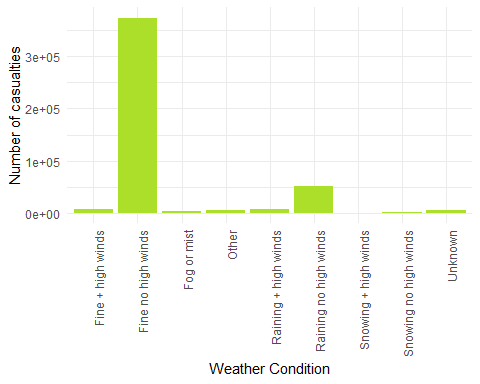
## Warning: Removed 81817 rows containing missing values (geom\_bar).



From the above graph, we can conclude that the highest number of accidents occur during fine weather conditions with no high wind.

To visualize the relationship between the number of casualties of road accidents and the weather condition, the following piece of code was used:

sumOfCasualties <- aggregate(UKAccidents$number\_of\_casualties, by=list(Category=UKAccidents$weather\_conditions), FUN=sum)  
  
#the graph is plotted to show the number of casualties due to weather conditions   
plot <- ggplot(data.frame(sumOfCasualties),aes(x= Category, y= x))+   
 geom\_bar(stat = "identity", fill= "#ABDF2A") +  
 xlab("Weather Condition")+  
 ylab("Number of casualties")+  
 theme\_minimal() +  
 ylim(0,375000) +  
 theme(axis.text.x=element\_text(angle=90 ,hjust=1,vjust=1))  
plot



Descriptive statistics was calculated as follows:

mean(UKAccidents$number\_of\_casualties, na.rm = TRUE)

## [1] 1.613239

median(UKAccidents$number\_of\_casualties, na.rm = TRUE)

## [1] 1

Anova was used to determine the p-value:

fit <- aov(UKAccidents$number\_of\_casualties ~ UKAccidents$weather\_conditions, data=UKAccidents)  
summary(fit)

## Df Sum Sq Mean Sq F value Pr(>F)   
## UKAccidents$weather\_conditions 8 2753 344.2 161.1 <2e-16 \*\*\*  
## Residuals 285322 609471 2.1   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

The number of accidents based of their severity was determined as follows:

severityFreq <- group\_by(UKAccidents, accident\_severity)  
summarise(severityFreq, n())

## # A tibble: 3 x 2  
## accident\_severity `n()`  
## <chr> <int>  
## 1 Fatal 3648  
## 2 Serious 39206  
## 3 Slight 242477

### Results

fit <- aov(UKAccidents$number\_of\_casualties ~ UKAccidents$weather\_conditions, data=UKAccidents)  
summary(fit)

## Df Sum Sq Mean Sq F value Pr(>F)   
## UKAccidents$weather\_conditions 8 2753 344.2 161.1 <2e-16 \*\*\*  
## Residuals 285322 609471 2.1   
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Since the p-value, calculated using Anova, is 2.35e-10 which is less than 0.5 at 95% confidence interval, the null hypothesis (H0) is rejected. Also, it can be stated that the findings of this research are significant.

### Implications

This study’s results have some implications for road safety and researchers. While going through different research conducted on the same topic, it is shown that there are less accidents when the weather is normal compared to hazardous weather like snow, rain, fog and high wind. The reason behind this is that people are not prompt to go out when the weather is unfavorable due to which, number of accidents are lower.

### Conclusion

It can be concluded that adverse weather conditions are not responsible for an increase in accidents.

### References

(*Reported Road Casualties in Great Britain: 2018 Annual Report* [2018](#ref-govuk2019)) (*How Bad Weather Increases Driving Accidents*, [n.d.](#ref-lawgroup2019)) Babbar ([2017](#ref-kaggle2017)) Department ([n.d.](#ref-statista2018))

Babbar, Akshay. 2017. *Road Accidents Incidence*. <https://www.kaggle.com/akshay4/road-accidents-incidence>.

Department, Statista Research. n.d. *Distribution of Contributing Factors Leading to Road Accidents in Great Britain in 2018*. <https://www.statista.com/statistics/323079/contributing-factors-leading-to-road-accidents-in-great-britain-uk/>.

*How Bad Weather Increases Driving Accidents*. n.d. <https://www.blumenshinelawgroup.com/how-bad-weather-increases-driving-accident/>.

*Reported Road Casualties in Great Britain: 2018 Annual Report*. 2018. <https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/834585/reported-road-casualties-annual-report-2018.pdf>.