

Analyzing the Severity of the Car Accident

Zihan Ni

August 21, 2020

1. Introduction

1.1 Background

With the development of the technology and the mobility, people tend to use their personal car to travel from places to places. However, with the freedom to ride a car on the road, the car accident sometimes would occur due to a lot of reasons. More than 90 people die in car accidents everyday in the US, and about 1.35 million people die in road crashes each year worldwide. Therefore, the car accident is a huge problem that happen and potentially harm our every day life

1.2 Problem

To avoid more car accidents and out of safety concerns, the study of how to reduce the probability is conducted by using the data science. This experiment would mainly focus on how the road condition, weather, and light condition and other potential factors would cause the car accidents.

1.3 Interest

The study would be interested by every person because it matters a lot in our daily life. It also be interested by the government, fire department, and hospital for the reason that they would highly want to regulate the car accident to save more lives.

2. Data Acquisition and Cleaning

2.1 Data sources

The data of car accidents would be found in the link <https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Data-Collisions.csv>, and the Metadata form would be included under the link <https://s3.us.cloud-object-storage.appdomain.cloud/cf-courses-data/CognitiveClass/DP0701EN/version-2/Metadata.pdf>

2.2 Data Cleaning

Because this project would mainly focus in the problem that whether weather, road condition, and light condition would affect the probability of the car accident, the data cleaning would try to eliminate the row of the data which doesn't include the information of weather, road condition, and light condition.

According to the data, the weather includes the following elements: clear, raining, overcast, unknown, snowing, other, fog/smoke/smog, sleet/hail/freezing Rain, Blowing sand/dirt, sever crosswind, and partly cloudy.

The road condition would be concluded as: dry, wet, unknown, ice, snow/slush, other, standing water, sand/mud/dirt, and oil.

The light condition would be concluded as: daylight, dark-street lights on, unknown, dusk, dawn , dark-no street lights, dark-street lights off, other, dark- unknowing lighting.

2.3 Data Selection

For the data that this project would use, I put the the columns of weather, road condition, and light condition into a new data base in order to have an easy and clear view of the project.

	WEATHER	ROADCOND	LIGHTCOND
0	Overcast	Wet	Daylight
1	Raining	Wet	Dark - Street Lights On
2	Overcast	Dry	Daylight
3	Clear	Dry	Daylight
4	Raining	Wet	Daylight
5	Clear	Dry	Daylight
6	Raining	Wet	Daylight
7	Clear	Dry	Daylight
8	Clear	Dry	Daylight
9	Clear	Dry	Daylight
10	Overcast	Dry	Daylight
11	Clear	Dry	Daylight
12	Raining	Wet	Dark - Street Lights On
13	Raining	Wet	Dark - No Street Lights
14	Clear	Dry	Dark - Street Lights On
16	Overcast	Dry	Daylight
17	Overcast	Dry	Daylight
18	Clear	Dry	Daylight
19	Unknown	Dry	Unknown
20	Clear	Dry	Dark - Street Lights On
21	Clear	Dry	Daylight
22	Clear	Dry	Dark - Street Lights On

3 Methodology

3.1 Grouping the data

By using the Pandas method `value_counts`, the new data base could be easily obtained to figure out the numbers of accidents that occur in different conditions.

	value_counts
Clear	111008
Raining	33117
Overcast	27681
Unknown	15039
Snowing	901
Other	824
Fog/Smog/Smoke	569
Sleet/Hail/Freezing Rain	113
Blowing Sand/Dirt	55
Severe Crosswind	25
Partly Cloudy	5

Repeat this process for two more times to get two more data sets contain the analyze of road condition and light condition.

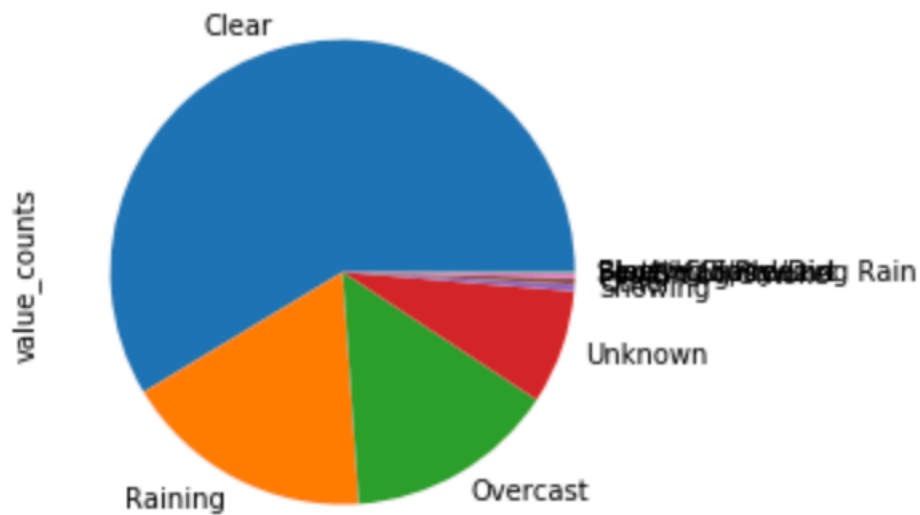
	value_counts
Daylight	116077
Dark - Street Lights On	48440
Unknown	13456
Dusk	5889
Dawn	2502
Dark - No Street Lights	1535
Dark - Street Lights Off	1192
Other	235
Dark - Unknown Lighting	11

	value_counts
Dry	124300
Wet	47417
Unknown	15031
Ice	1206
Snow/Slush	999
Other	131
Standing Water	115
Sand/Mud/Dirt	74
Oil	64

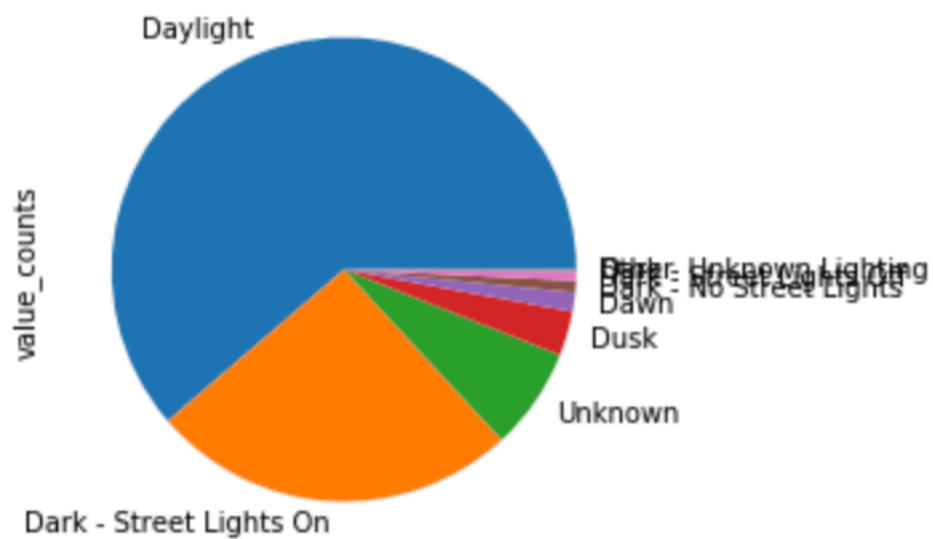
3.2 Graphing

By using the matplotlib, the pie graph is used to visualize the data.

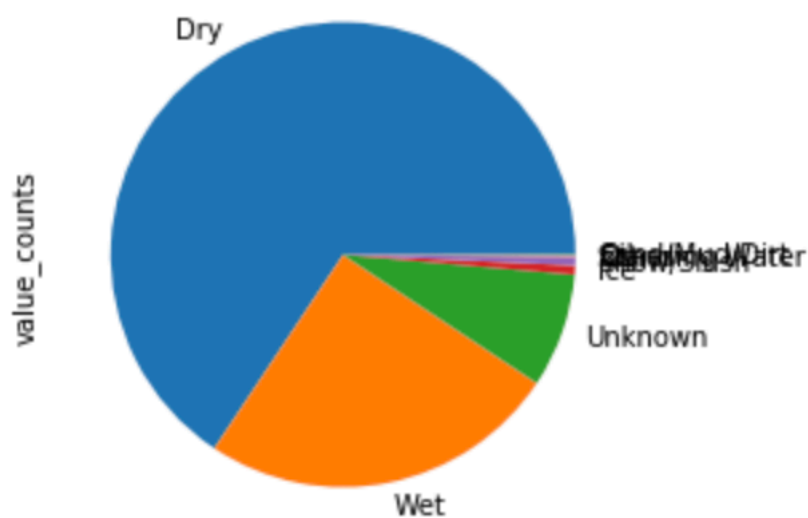
The Weather Condition for Total accidents



The Light Condition for Total Accidents



The Road Condition for Total Accidents



4. Result

4.1 Read the graph

According to the first graph, the outcome is the opposite from the common perspectives —people would normally perceive that the car accident would happen more under the special weather condition such as rainy and snowy. However, there are 111008, more than 50 percent, car accidents happen in a clear day.

The second graph indicates number of the accidents under different light conditions. For common belief, people would tend to hit the object with no lights on the road more than there is sufficient light. However, the accidents that happen under the daylight is over 50 percent in all cases.

The third graph demonstrates the road conditions. The dry road is about 124300 cases among all. It is still over 50 percent.

5. Discussion

5.1 The predictive meaning behind the graphs

According to the result of the graph, the relationship between weather, road condition, and light condition and the car accident do not fit the common belief, which the accident happens more in the rainy day, wet road, and poor light condition. The three graph all indicates that people would cause more accidents when they feel it is safe—sunny weather, dry road, enough daylight. Therefore, it is predicted that people would care less and drive faster than the time that the weather, road, and light are not suitable for driving. Moreover, people might also avoid to drive on the road when the weather, road, and light are not good, which might also cause the low rate of car accidents.

6. Conclusion.

6.1 The Conclusion and future research

From the discussion and result of the study, it is believed that bad weather, poor road and light condition is not the core factor for people to cause the car accidents. The project shows that under perfect conditions of driving, people would have a higher possibility to crash their cars and make a car accident. However, this project doesn't indicate what is the main reason for the car accidents, and why people tend to cause car accidents under great weather, perfect road condition and efficient daylight. Further study should be conducted to reveal the truth of the car accidents.

