Traffic Flow in Inclement Weather and Factors Affecting the Road Safety

Section 1: introduction

In this project, I will work with a dataset on the Road Accidents. The dataset can be downloaded from here: https://www.kaggle.com/search?q=road+accidents.

Weather causes a variety of impacts on the transportation system. While severe winter storms, hurricanes, or floodings can result in major stoppages or evacuations of transportation systems and cost millions of dollars, day-to-day weather events such as rain, fog, snow, and freezing rain can have a serious impact on the mobility and safety of the transportation system users. These weather events can result in increased fuel consumption, delay, number of accidents, and significantly impact the performance of the transportation system. The overall goal of the project work undertaken in this project is to develop a better understanding of the impacts of weather on traffic flow.

Vehicle design and performance have two potential effects on safety: first, those aimed at reducing the risk of crashing; second, those aimed at reducing the consequences when accidents do occur. Vehicle mass, size and speed have long been important concerns for manufacturers and for safety analysts. While we may not be able to change the weather, we can have a better understanding regarding how certain weather events can affect road conditions. This knowledge allows us to be better prepared to stay safe while driving. However, rain, snow, falling temperature, high wind, and fog have a great effect on the road surface and driving conditions.

Comparing casual severity on weather condition and road surface (wet/damp) instead of dry road and counting how many of the accidents were on a wet/damp vs dry surface and the type of vehicle is the goal of this project.

The analysis also produced more detailed, multidimensional analysis results that can be found in below report. The summaries and diagrams show how various parameters are influenced over a range of weather and traffic conditions.

Justify why is it important/useful to solve the problem?

Adverse weather conditions can impact arterial mobility by affecting driver behavior and traffic signal operations. This paper will describe several studies that have been conducted on various road facilities in different climates to document weather impacts on arterial traffic flow. A few successful strategies to modify traffic signal timing in response to poor road weather conditions will be introduced. A discussion of weather-related parameters in simulation models and the benefits of weather-responsive signal timing are also discussed. The premise is that the better the community understands traffic flow under varying weather conditions the better able it is to manage road networks and improve highway operations.

What does the analysis/model building tell you?

Weather causes a variety of impacts on transportation systems during and after weather events . These impacts can be short- or long-term, and direct or indirect. However, the linkages betwe en inclement weather conditions and traffic flow in existing analysis tools remain tenuous. The model is contingent upon identifying causality severity in several condition road such as 1st Roa d Class, Road Surface, Lighting Conditions, Weather Conditions, Casualty Class, Casualty Severit y and display causality severity after each car accident.

What are your recommendations?

Weather events such as rain, snow, sleet, fog, high winds, and flooding reduce roadway capacity. These events can cause slick pavement, lower traffic speeds, increase speed variability, affect traffic volume, increase delay, escalate crash risk, disrupt access to roads (e.g., lane obstruction, pavement buckling) and damage road infrastructure (e.g., traffic control devices). My recommendation would be to Comparing casual severity on weather condition and road surface (wet/damp) instead of dry road and counting how many of the accidents were on a wet/damp vs dry surface.

Milestone 1

Adverse weather conditions can impact arterial mobility by affecting driver behavior and traffic signal operations. Weather causes a variety of impacts on the transportation system. In milestone 1, I will describe the original idea for the analysis/model building. Comparing casual severity on weather condition and road surface (wet/damp) instead of dry road and counting how many of the accidents were on a wet/damp vs dry surface and the type of vehicle is the clearly identified issues in this milestone.

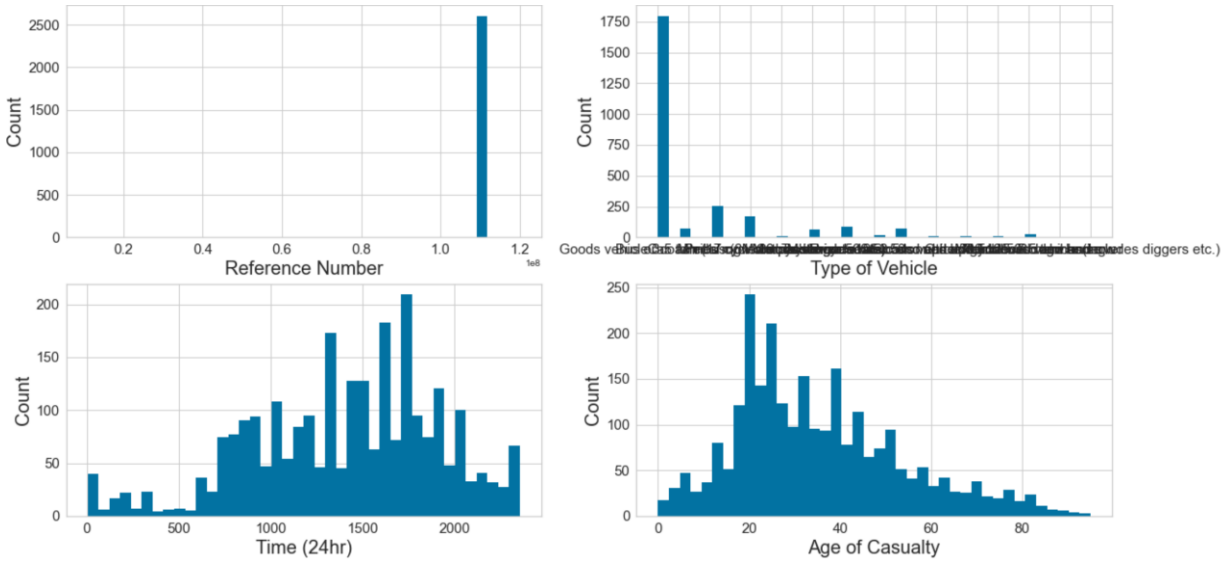
This section will show two sets of bar chart which displayed the categorical and numerical data to get the better description of weather condition on wet/damp road surface instead of dry road.

Bar Chart 1: The bar chart represents the Reference Number which are 1.1

Bar Chart 2: The bar chart represents the Type of Vehicle which most Type of Vehicle accident are car

Bar Chart 3: The bar chart represents the Time(24hrs) which most Time(24hrs) accident happened in the evening between 12:00 to 18:00 o'clock

Bar Chart 4: The bar chart represents the Age of Casualty which most Age of Casualty accident reported between 20 to 30 years old

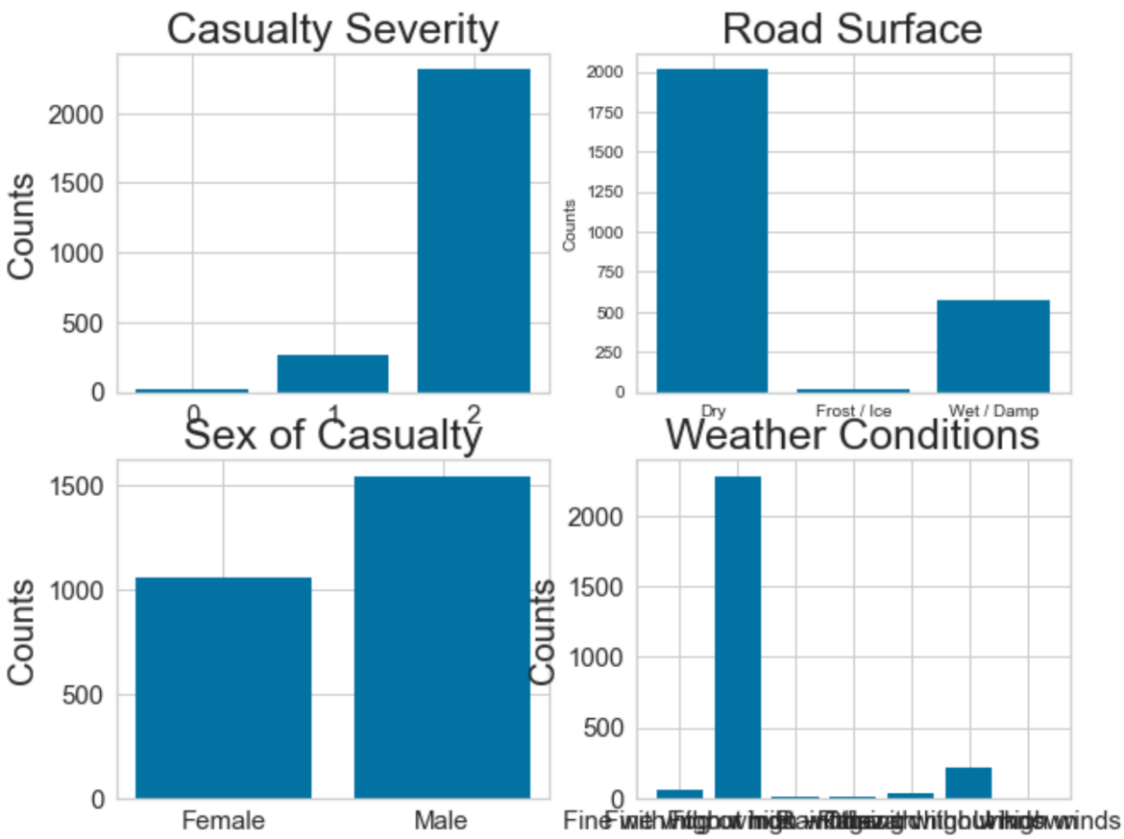


Bar Chart 1: The bar chart represents the most passengers on the Casualty Severity data set got Slight

Bar Chart 2: The bar chart represents the most accident happened on the dry Road Surface

Bar Chart 3: The bar chart represents the slightest passengers were male in Sex of Casualty bar chart.

Bar Chart 4: The bar chart represents the most accident recorded in Fine without high winds on 'weather conditions' situation



Milestone 2: Dimensionality & Feature Reduction and Feature Engineering

In the previous milestones I have created an idea, located the data, and have started the analysis. In this section I have moved the data preparation of the project. The following steps have done for the model building and evaluation phase.

- Dropped unnecessary data
- Filled in missing data
- Replaced fare price with its log; verified the data is more "normal"
- Created dummy variables for the categorical data
- Split data into training and test sets

```
#create a new dataframe without 'Number of Vehicles' column
X_train.drop('Number of Vehicles', axis = 1, inplace = True)
X_test.drop('Number of Vehicles', axis = 1, inplace = True)
```

```
#create dummy variables for the categorical columns
X_train = X_train.replace({'Pclass': {1: '1st', 2: '2nd', 3: '3rd'}})
X_test = X_test.replace({'Pclass': {1: '1st', 2: '2nd', 3: '3rd'}})
#get dummy variables for the categorical columns dropping the first dummy variable
#we drop the first dummy variable, so we get one fewer than the number of categories
#this prevents collinearity between the dummy variables
X_train = pd.get_dummies(X_train, drop_first = True)
X_test = pd.get_dummies(X_test, drop_first = True)
```

```
#check dimensions of the training and test feature matrices
print(X_train.shape)
print(X_test.shape)
#view the training features
X_train.head()
```

```
(2082, 36)
(521, 34)
```

	Time (24hr)	Age of Casualty	log_Number_of_Vehicles	1st Road Class_A(M)	1st Road Class_B	1st Road Class_Motorway	1st Road Class_Unclassified	Road Surface_Frost / Ice	Road Surface_Wet / Damp	Lighting Conditions_Darkness: street lighting unknown
0	830.0	31.0	1.098612	0	0	0	1	0	0	0
1	805.0	32.0	1.609438	0	0	1	0	0	0	0
2	1825.0	19.0	1.098612	0	0	0	0	0	0	0
3	1845.0	24.0	1.098612	0	0	0	0	0	1	0
4	1805.0	16.0	1.098612	0	0	0	1	0	0	0

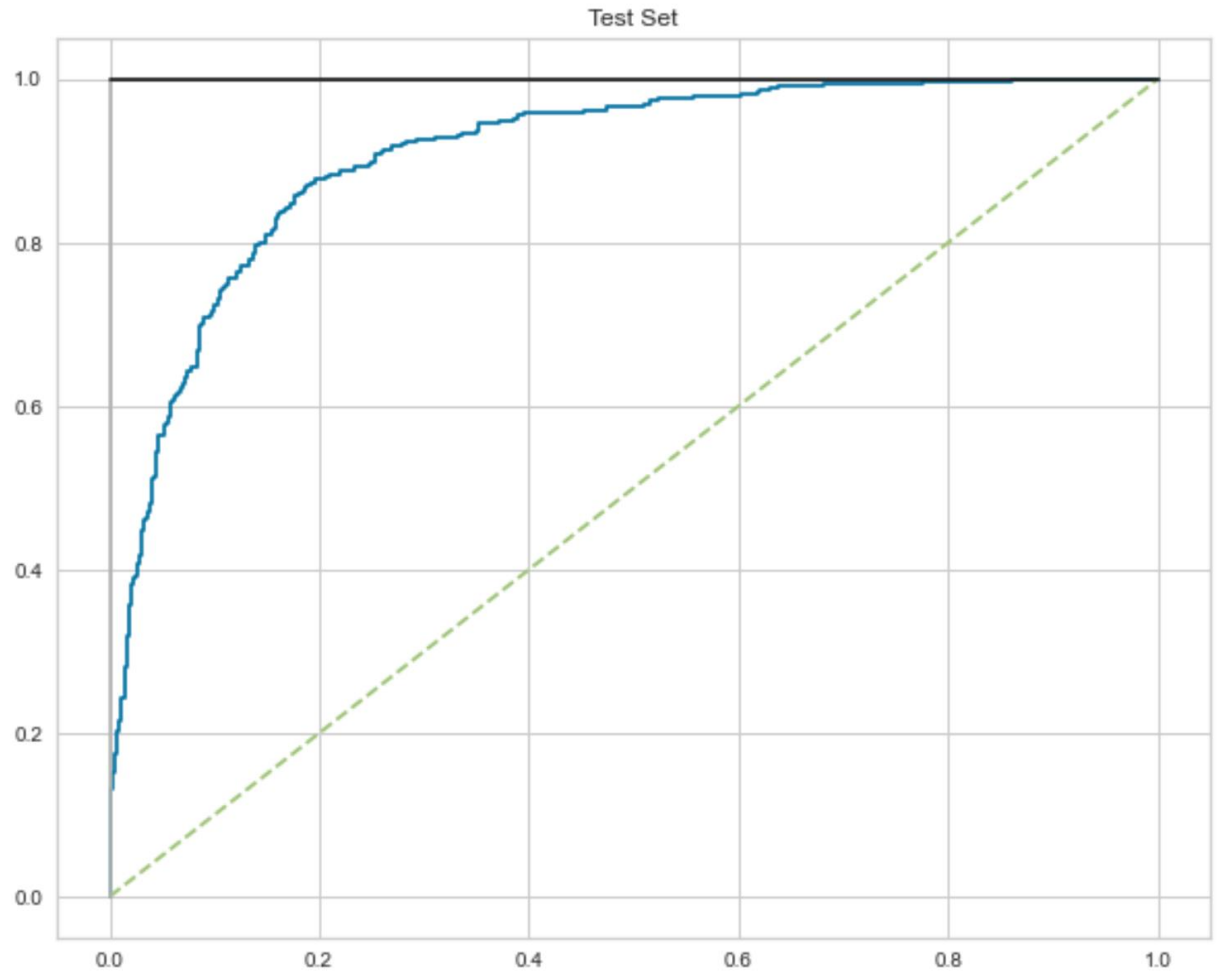
5 rows × 36 columns

Milestone 3: Model Selection & Evaluation

For my model, I used logistic regression. Logistic regression is a classification algorithm which is how I structured my dataset. Usually when the class distribution is unbalanced, accuracy is considered a poor choice as it gives high scores to models which just predict the most frequent class. Because I have highly imbalanced classes, i.e., almost all of the target values are "Slight". However, accuracy for our training dataset with tuning is 89.34%.

The ability to have high values on Precision and Recall is always desired but, it's difficult to get that for this data sets. Depending on the type of application we need to either increase Precision or Recall.

```
# Plot ROC curve
plt.title("Test Set")
plt.plot(false_positive_rate, true_positive_rate)
plt.plot([0, 1], Ls='--')
plt.plot([0, 0], [1, 0], c=".7"), plt.plot([1, 1], c=".1")
#plt.ylable("True Positive Rte")
#plt.xlable("False Positive Rte")
plt.show()
```



Conclusions

- Road Surface plays an important role in predicting in Casualty Severity
- Accuracy for our training dataset with tuning is 89.34%89.34% accuracy
- Road Surface, Lighting Conditions, Weather Conditions, and Casualty Class are useful in predicting Casualty Severity.
- Because of categorical features KNN model and metric makes sense for my problem