

# Accident severity

Seattle Police Department  
Car Accident

# Introduction

In recent year ther is certain increase in accident. This cause serious damage in property , parts of body and even life. So to understand and find the real cause ,we have to work on certain data and find pattern to resolve the problem or to minimise it.

By this we able to solve problem like like driving suggestion ,carefull analysis of fatal accident, depending on the weather and road conditions on any given day. This also contains their employers, insurance firms, emergency and health care personal.

# DataSet

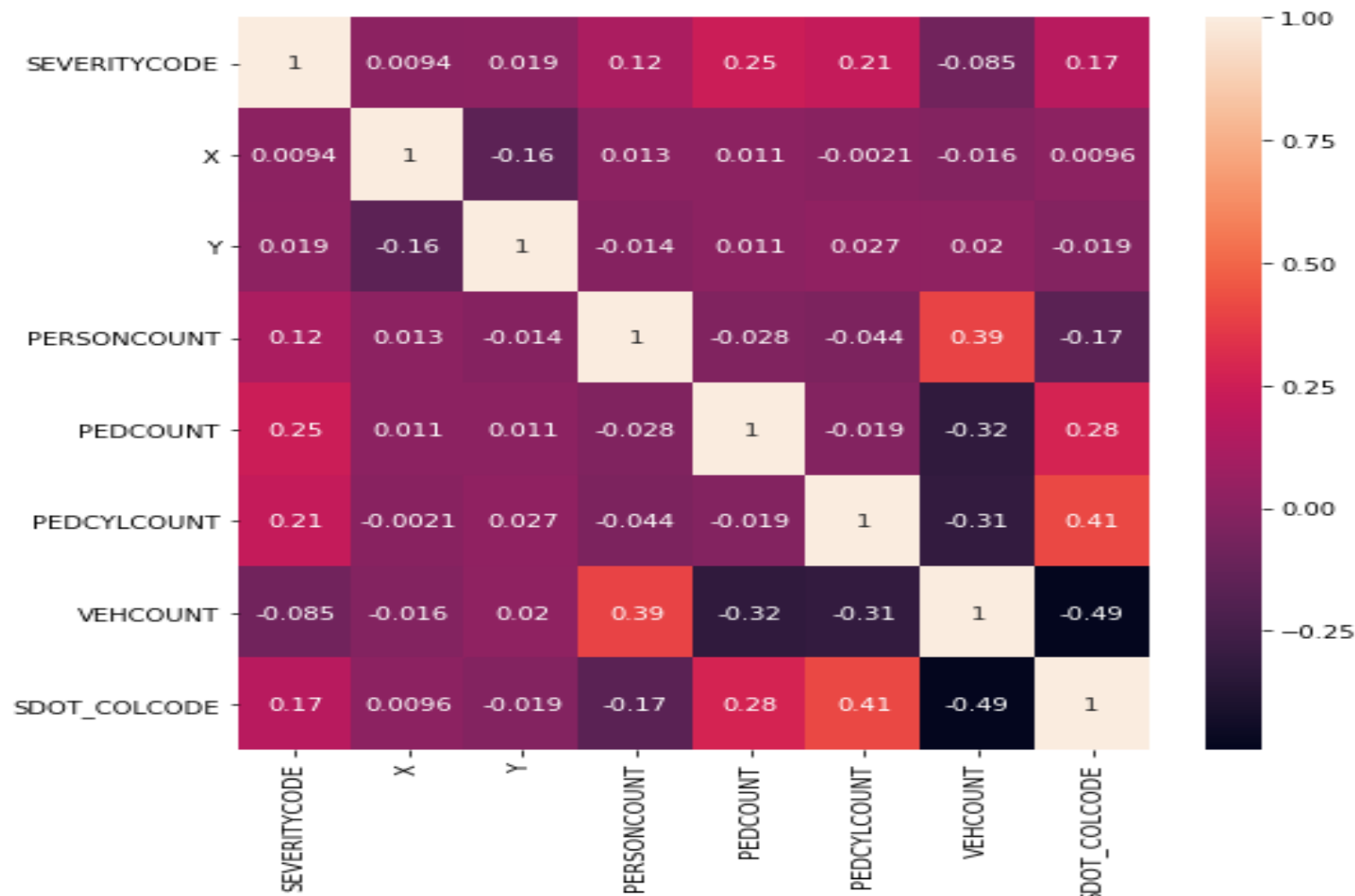
This is an standard data set published by the Seattle Police Department, with over 194673 observations with 37 attributes collected over the last 15 years. By this huge data we have to make a effective model to prevent future accident and reduce severity, so it can be use by people for getting security and also use by companies to build a reilable system .

# Methodology

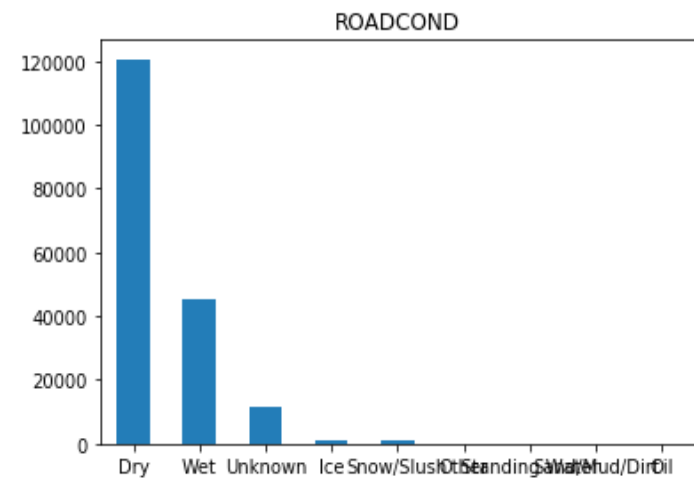
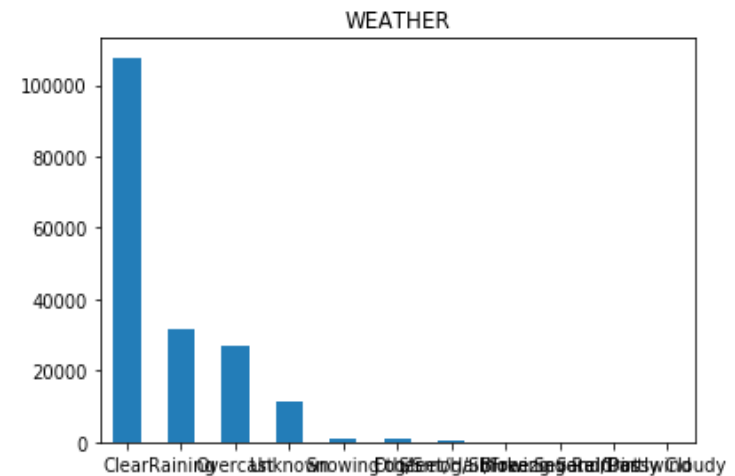
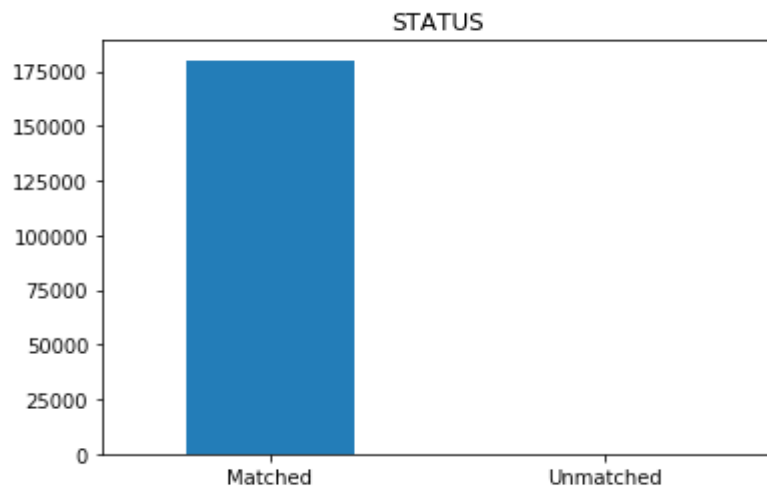
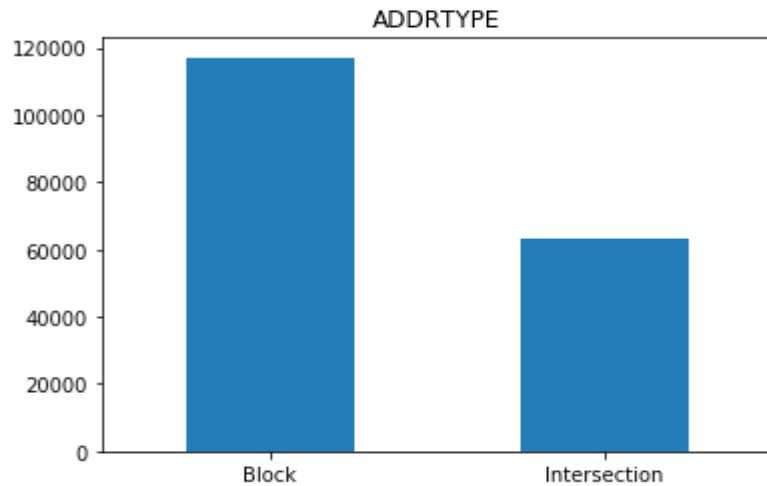
In this project python is used for easily availability of functionality, coding is performed on IBM watson jupyter notebook. In python data analysis is easy to perform and python also contain sufficient library for data transformation like Pandas, Numpy, Matplotlib, and Seaborn .Sequence have to follow =

- Data understanding , Data collection , Data visualization , Data Transformation , Data model and Evaluation .

# Visualisation (heat map)



# Visualization



# Model and Code

- Import modules and run Random forest

```
y=pd.DataFrame()  
y=x['SEVERITYCODE']  
x=x.drop(['SEVERITYCODE'], axis=1)
```

```
from sklearn.ensemble import RandomForestClassifier  
from sklearn.model_selection import cross_val_score,train_test_split  
from sklearn.svm import SVC  
from sklearn.pipeline import make_pipeline  
from sklearn.preprocessing import StandardScaler  
from sklearn.metrics import classification_report, confusion_matrix, accuracy_score,mean_absolute_error
```

```
x_train, x_test, y_train, y_test = train_test_split(x, y, test_size=0.25, random_state=0)
```

```
tree=RandomForestClassifier(n_estimators=100)  
tree.fit(x_train, y_train)  
y_pred = tree.predict(x_test)
```

# Result

By the graph analysis and after filling missing values we are ready to apply the ML model , These models are use for prediction new events ,while trains with previously happened event that is why ML models are appropriate .I am using random forest because it take several decision and came to right conclusion . Ramdom forest is taken from sklearn liberary with 100 estimators.



# Conclusion

Although this analysis has given us some good insight, there needs to be a closer inspection of certain other variables. It seems like a lot of these accidents are minor and avoidable. Having said that there is still a considerable amount of loss of property and these findings can be helpful to the Seattle PD in enforcing some new measures to prevent future accidents.

# Thank You

- Created by Siddharth Kajle