PREDICTING THE SEVERITY OF ACCIDENTS IN SEATTLE

Predicting the severity of accidents in Seattle

- The World Health Organization (WHO), states that "there were 1.35 million road traffic deaths globally as at 2016, with millions more sustaining serious injuries and living with long-term adverse health consequences
- The United States recorded six million automobile collisions every year. With estimated recorded fatalities of nearly 40,000 annually. at 2013, the reported road traffic fatalities in the United States stood at 32,719 (71% M, 29% F)

Data acquisition and cleaning

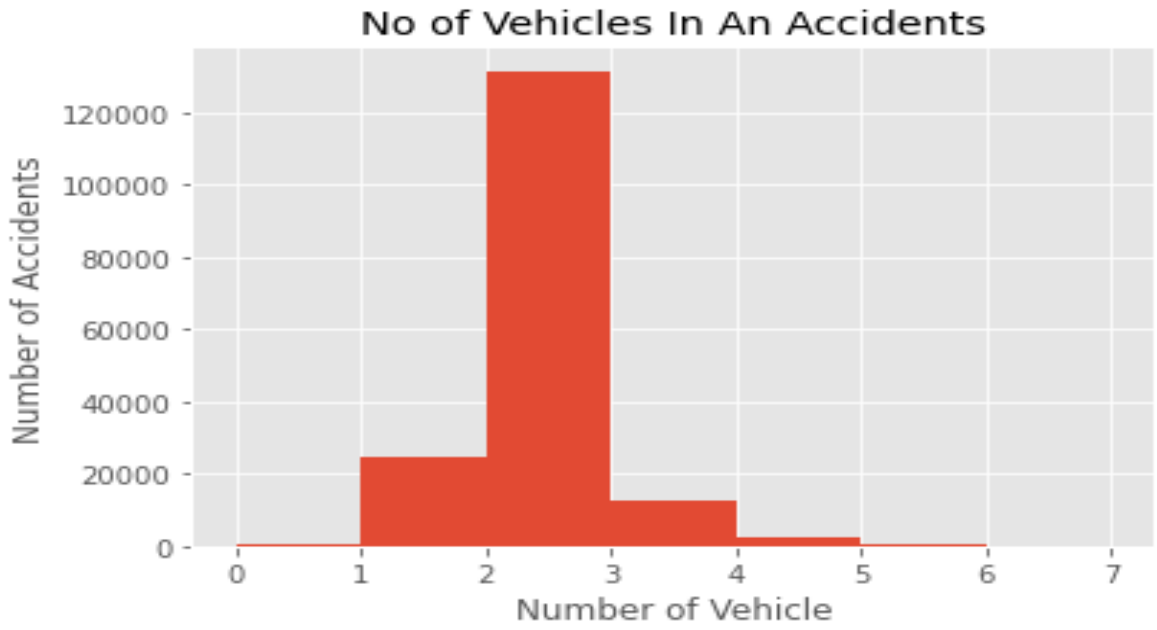
- The collision for Seattle data was obtained from the shared link in the Coursera website. It cover a period of over 15 years
- It contained 194,673 rows and 38 columns with various data attribute
- Irrelevant data, missing data and columns containing unknown data and blank data were identified and dropped from the dataset
- Speed was dropped from the dataset because it contained 185,340 missing data
- The final dataset contained 172,242 rows and 6 columns comprising of ACCTYPE, WEATHER, SEVERITYCODE, PERSONCOUNT, VEHICLECOUNT and ROADCOND.

Number of vehicles in an accident:

```
[49]: bins = np.arange(df.VEHCOUNT.min(),8,1)
      plt.hist(df.VEHCOUNT, bins = bins)

      plt.title('No of Vehicles In An Accidents')
      plt.ylabel('Number of Accidents')
      plt.xlabel('Number of Vehicle')
```

```
[49]: Text(0.5, 0, 'Number of Vehicle')
```

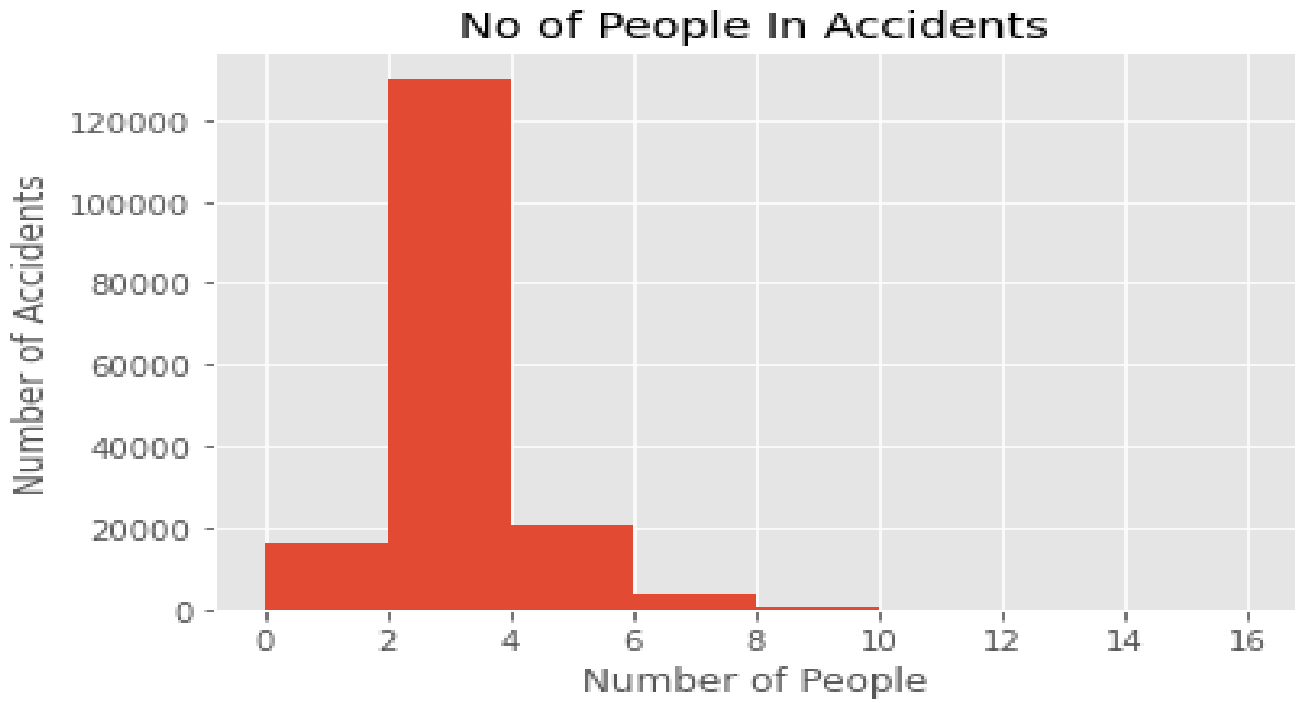


Number of persons involved:

```
[57]: bins = np.arange(df.PERSONCOUNT.min(),17,2)
      plt.hist(df.PERSONCOUNT, bins = bins)

      plt.title('No of People In Accidents')
      plt.ylabel('Number of Accidents')
      plt.xlabel('Number of People')
```

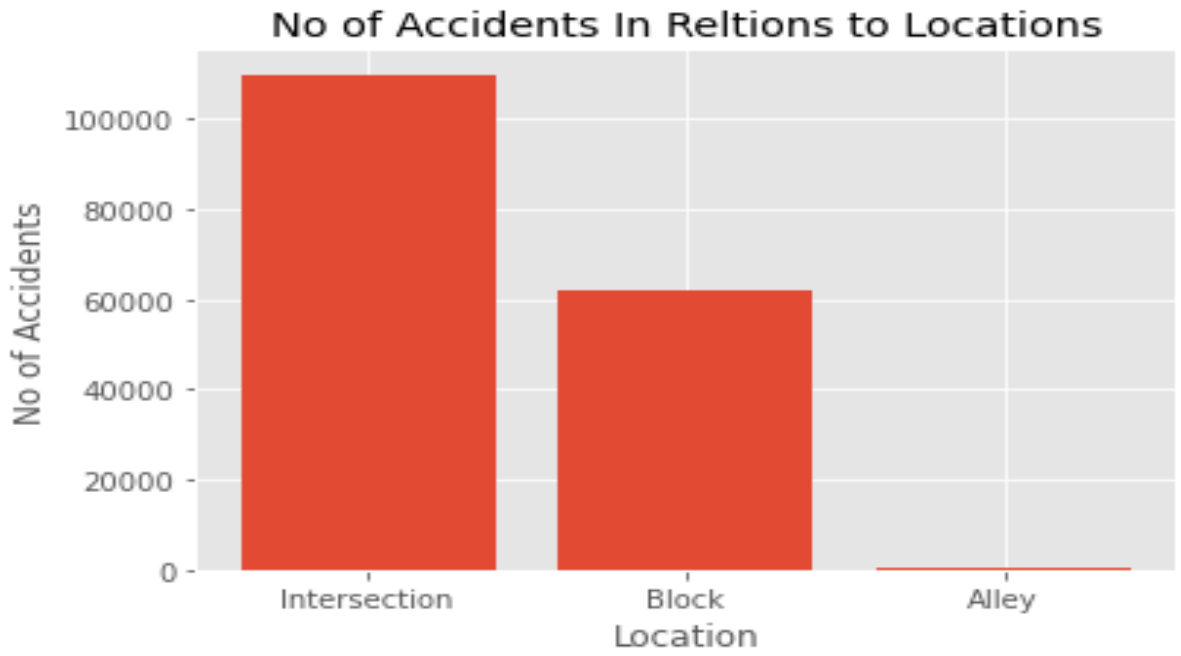
```
[57]: Text(0.5, 0, 'Number of People')
```



ACCTYPE

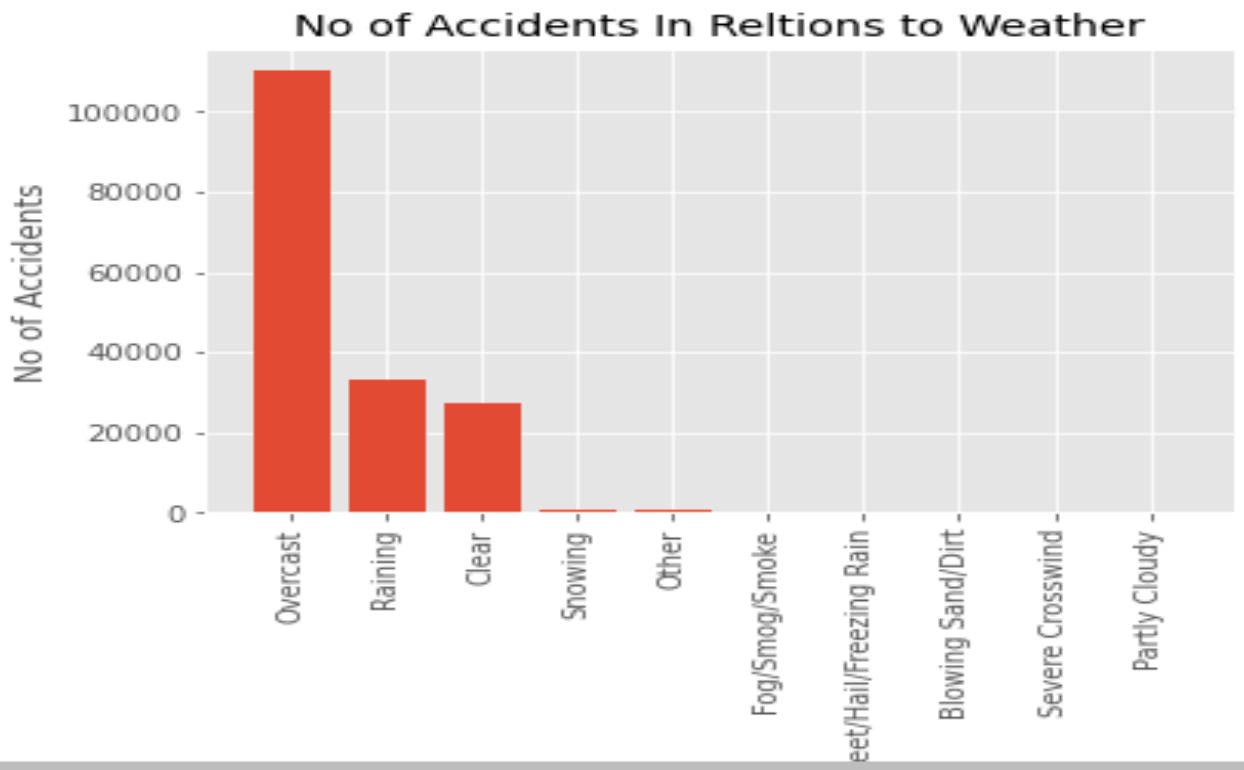
```
[43]: X = df.ADDRTYPE.unique()
      Data = df.ADDRTYPE.value_counts()
      plt.bar(X, height=Data)
      plt.xlabel('Location')
      plt.ylabel('No of Accidents')
      plt.title('No of Accidents In Reltions to Locations')
```

```
[43]: Text(0.5, 1.0, 'No of Accidents In Reltions to Locations')
```



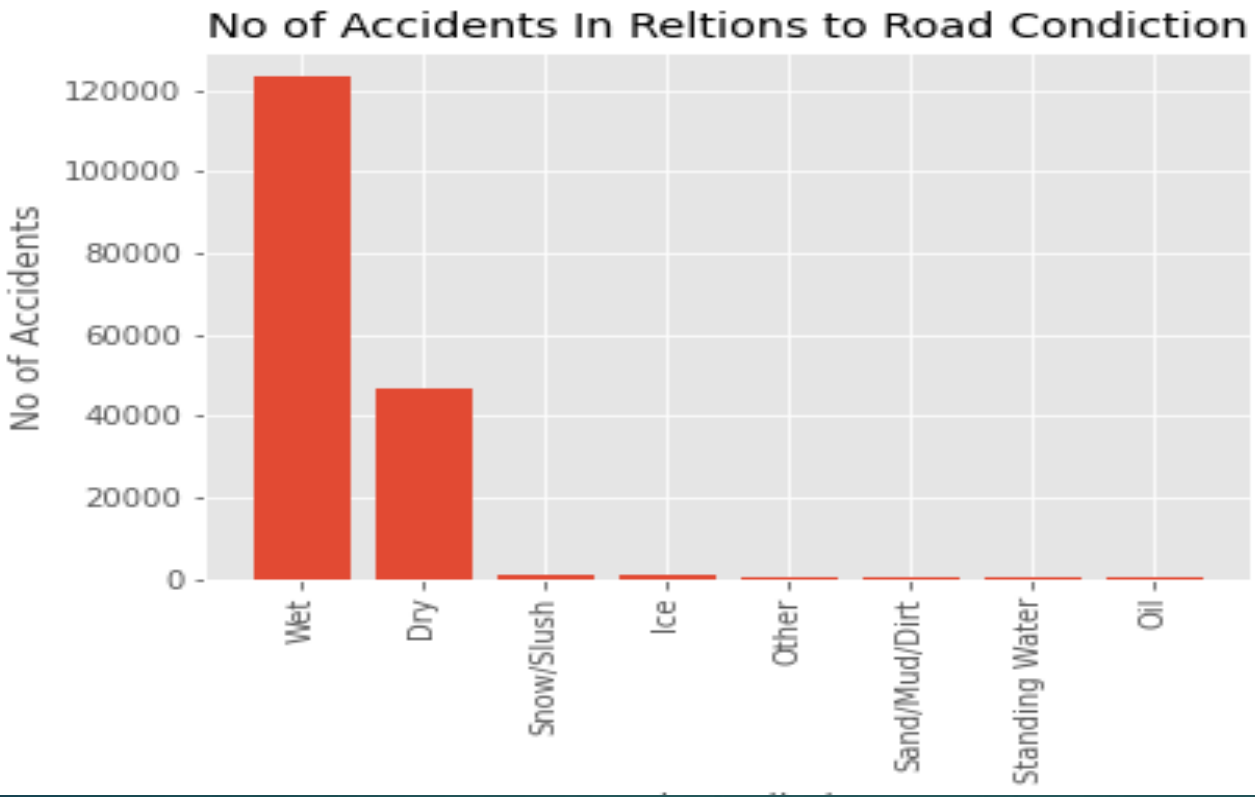
Weather:

```
Data = df.WEATHER.value_counts()
plt.bar(X, height=Data)
plt.xlabel('Weather')
plt.ylabel('No of Accidents')
plt.title('No of Accidents In Reltions to Weather')
plt.xticks(rotation= 90)
plt.show()
```

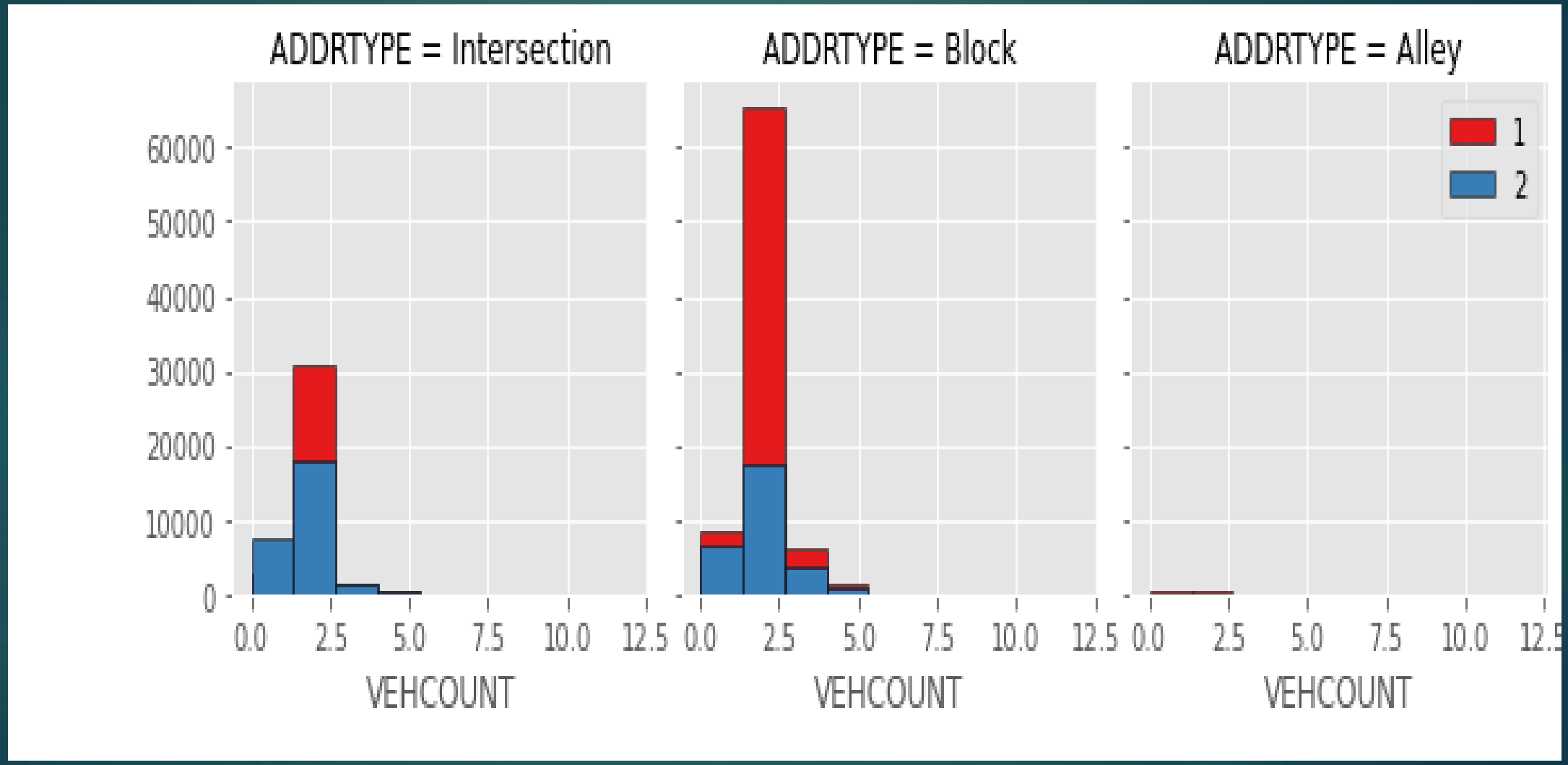


Road Condition

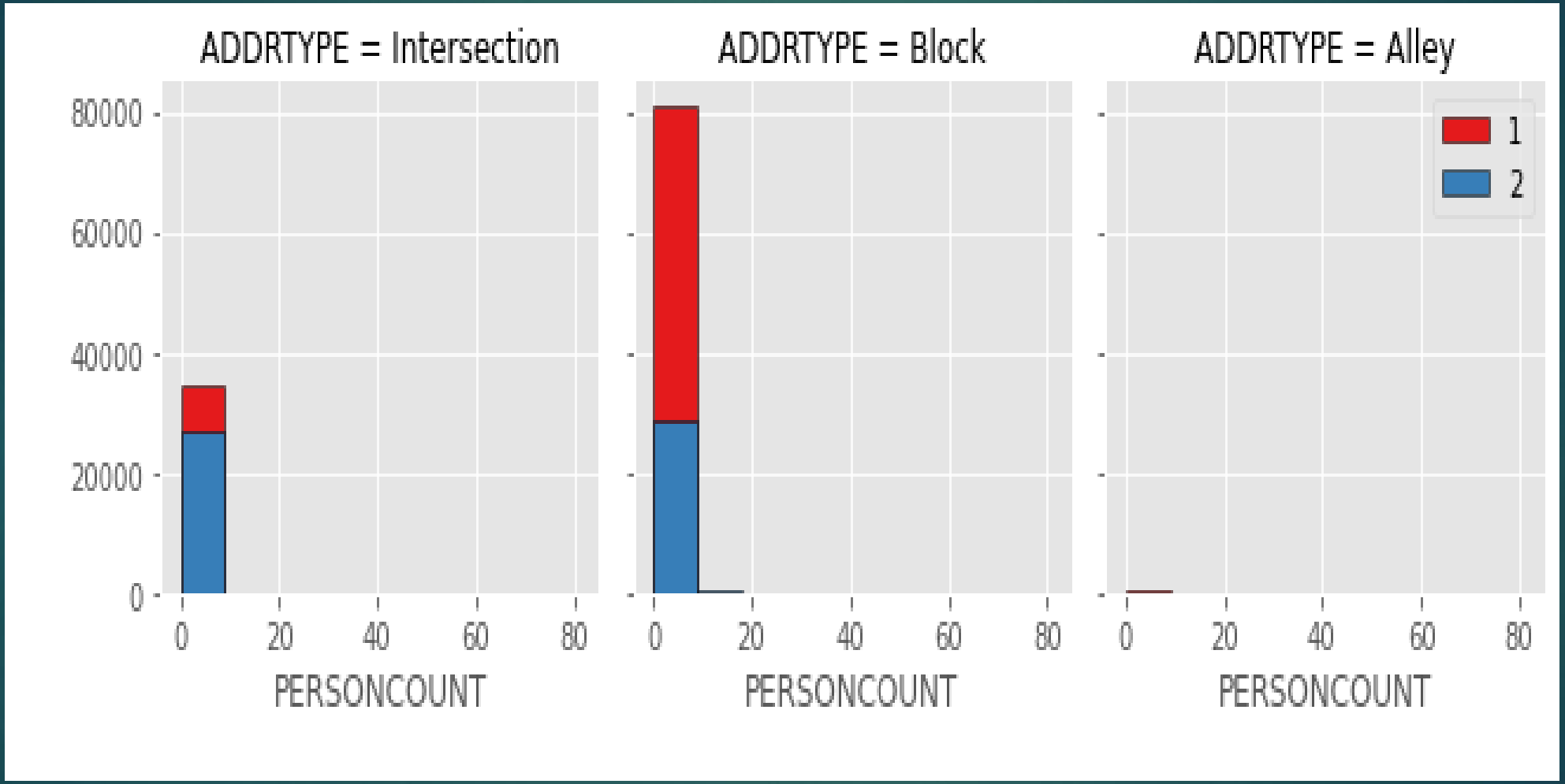
```
Data = df.ROADCOND.value_counts()
plt.bar(X, height=Data)
plt.xlabel('Road Condiction')
plt.ylabel('No of Accidents')
plt.title('No of Accidents In Reltions to Road Condiction')
plt.xticks(rotation= 90)
plt.show()
```



Accident Severity



Person Count Severity



Conclusion

- Despite this analysis, there needs to be a closer inspection of certain other variables for insights. There is a considerable amount of property loss, however a lot of these accidents are minor and avoidable according to the analysis. These findings could be helpful to the Seattle Police Department in enforcing some new measures to prevent future accidents.