



Car Accident Factors - Analysis

Data Analytics & Visualization



Project Information

Project Team (Group 3):

Cassidy Kratwell, Caleb O'kane, Amanda J Garza

Project Scope:

As a group we have decided to gather data from records of road accidents that occurred during January 2021 in the Kensington and Chelsea area to perform an analysis on what affects car accidents and how road safety can be improved based upon this data.

Road Accident Data Analyzed:

Junction Type: (i.e., T or staggered junction)

Junction Control: (i.e., give way or uncontrolled)

Accident Severity

Speed Limit

Urban or Rural Area

Road Types:

Number of Casualties

Questions:

What contributes the most to car accidents?

What junction incurs the least amount of accidents?

Is the accident severity affected by the speed?

Does the geographic location with road type in mind have an affect on the number of accidents that occur?

How did we go about solving these questions:

We will represent our data in bar plots, histograms, and pie charts

We completed our statistical analysis using t-tests

Project Data Sources

.csv data:

<https://www.kaggle.com/datasets/nextmillionaire/car-accident-dataset/data>

API:

Geoapify

<https://www.geoapify.com/>



Data Cleanup

- Removed 'Carriageway_Hazards'
- Changed the misspelled value in the column 'Junction_Control'

Accident_Index	307973
Accident Date	307973
Day_of_Week	307973
Junction_Control	307973
Junction_Detail	307973
Accident_Severity	307973
Latitude	307973
Light_Conditions	307973
Local_Authority_(District)	307973
Carriageway_Hazards	5424
Longitude	307973
Number_of_Casualties	307973
Number_of_Vehicles	307973
Police_Force	307973
Road_Surface_Conditions	307656
Road_Type	306439
Speed_limit	307973
Time	307956
Urban_or_Rural_Area	307973
Weather_Conditions	301916
Vehicle_Type	307973

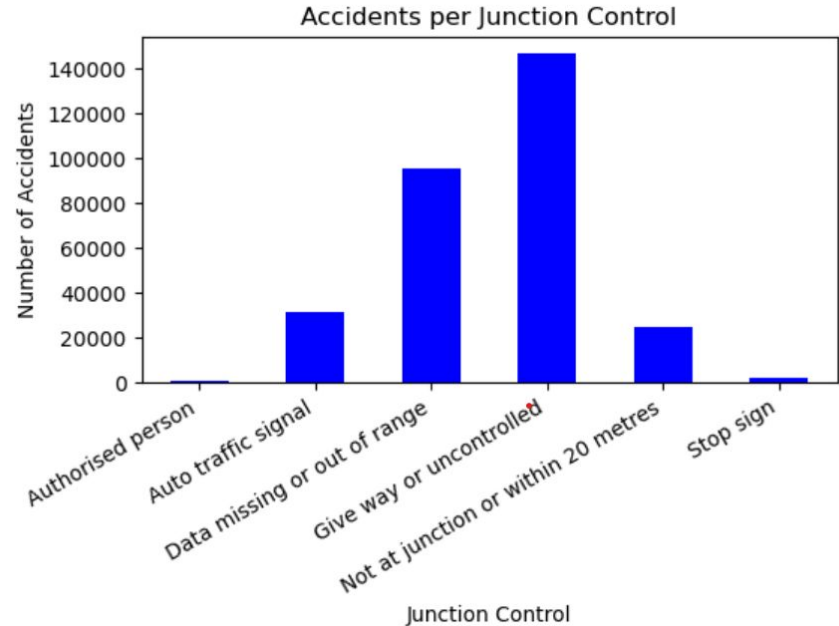
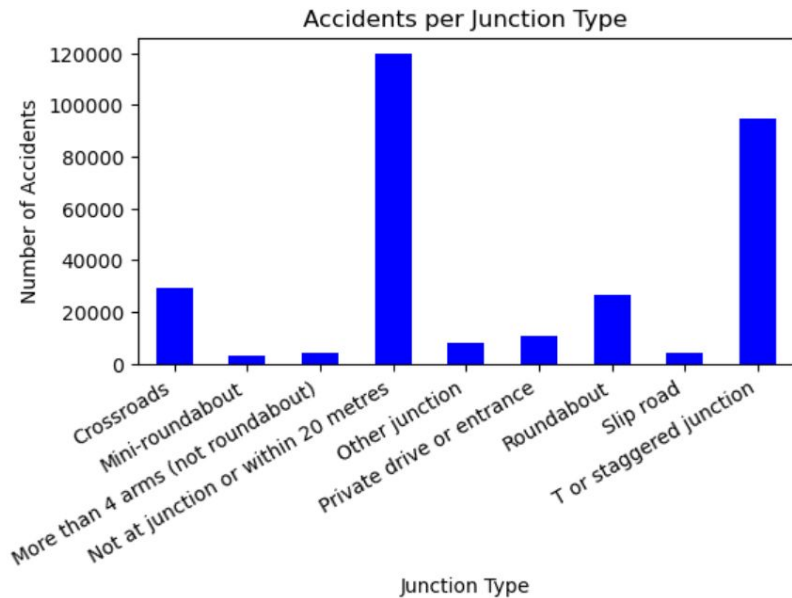
Accident_Index	300495
Accident Date	300495
Day_of_Week	300495
Junction_Control	300495
Junction_Detail	300495
Accident_Severity	300495
Latitude	300495
Light_Conditions	300495
Local_Authority_(District)	300495
Longitude	300495
Number_of_Casualties	300495
Number_of_Vehicles	300495
Police_Force	300495
Road_Surface_Conditions	300495
Road_Type	300495
Speed_limit	300495
Time	300495
Urban_or_Rural_Area	300495
Weather_Conditions	300495
Vehicle_Type	300495

Junction_Control	
Give way or uncontrolled	146515
Data missing or out of range	95427
Auto traffic signal	31471
Not at junction or within 20 metres	24889
Stop sign	1655
Authorised person	449
Auto traffic sigl	89

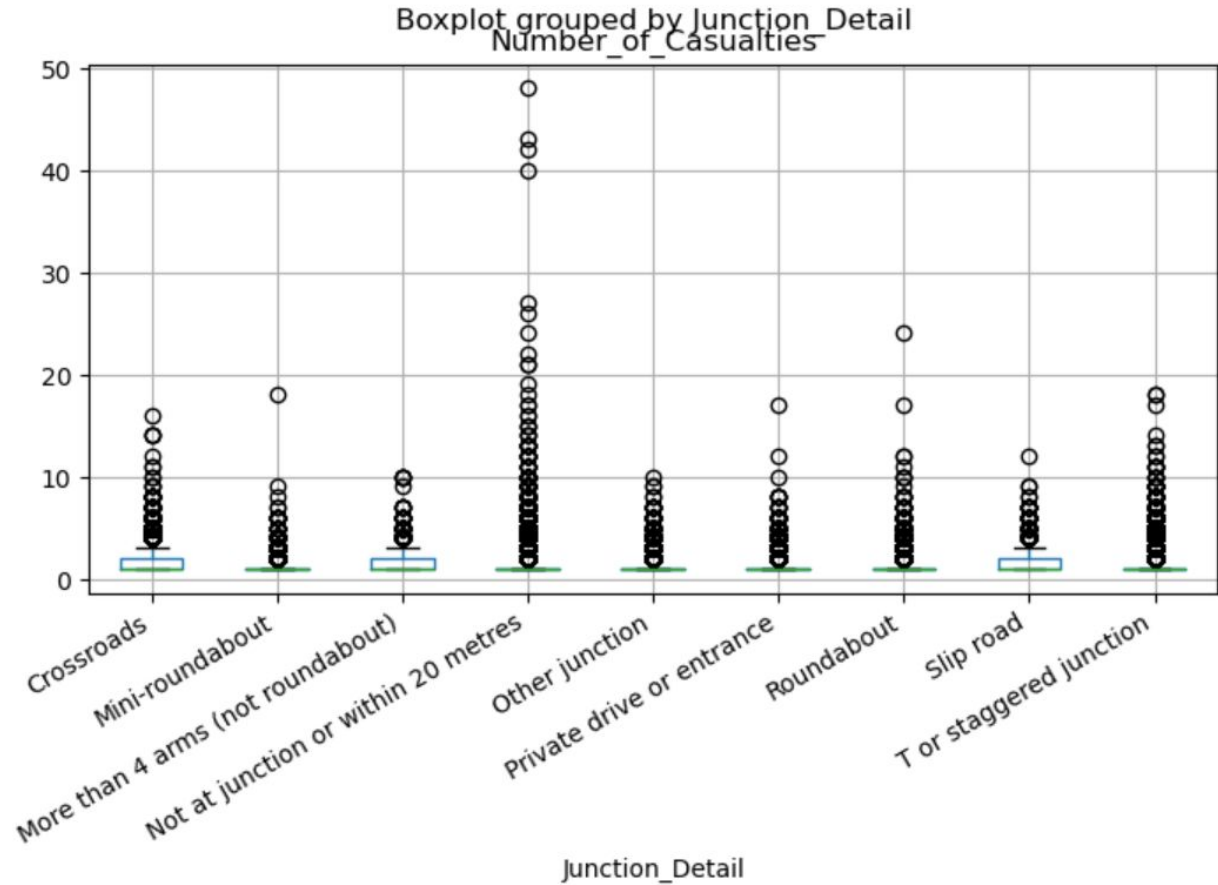


Junction_Control	
Give way or uncontrolled	146515
Data missing or out of range	95427
Auto traffic signal	31560
Not at junction or within 20 metres	24889
Stop sign	1655
Authorised person	449

What junction incurs the least amount of accidents?



Outliers



Roundabout vs T-intersection

Roundabout:

Mean: 1.287

Variance: 0.48

T-Intersection:

Mean: 1.319

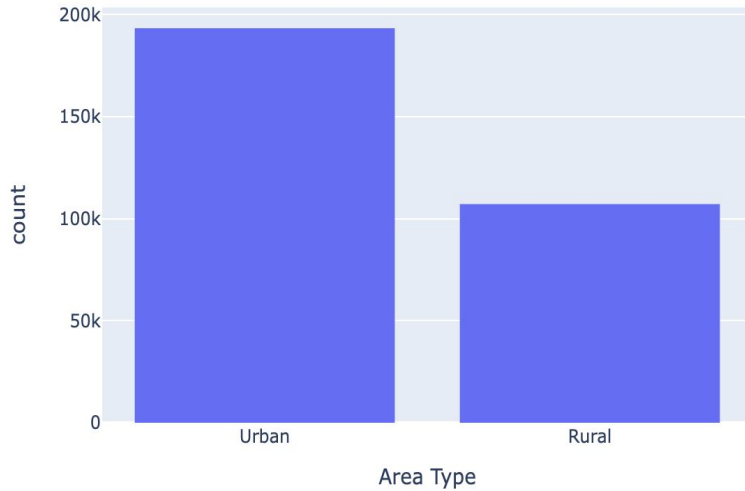
Variance: 0.54

T-test:

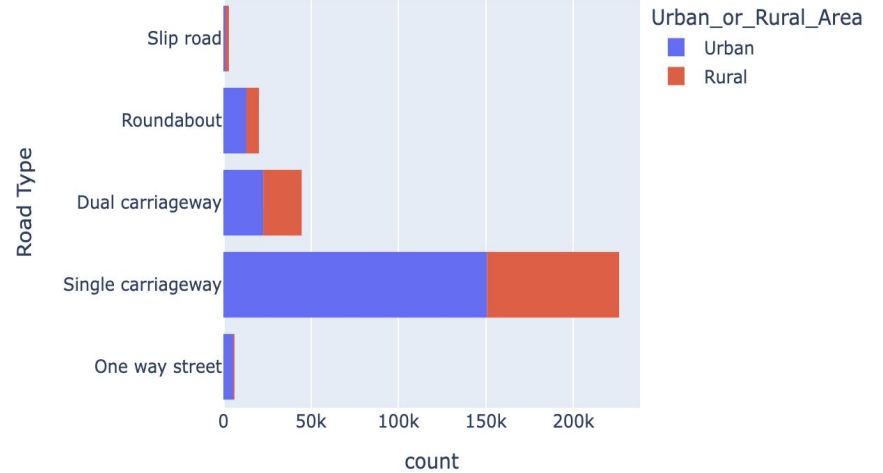
P-value: $7.8e-11$

Geographic Location vs Road Type

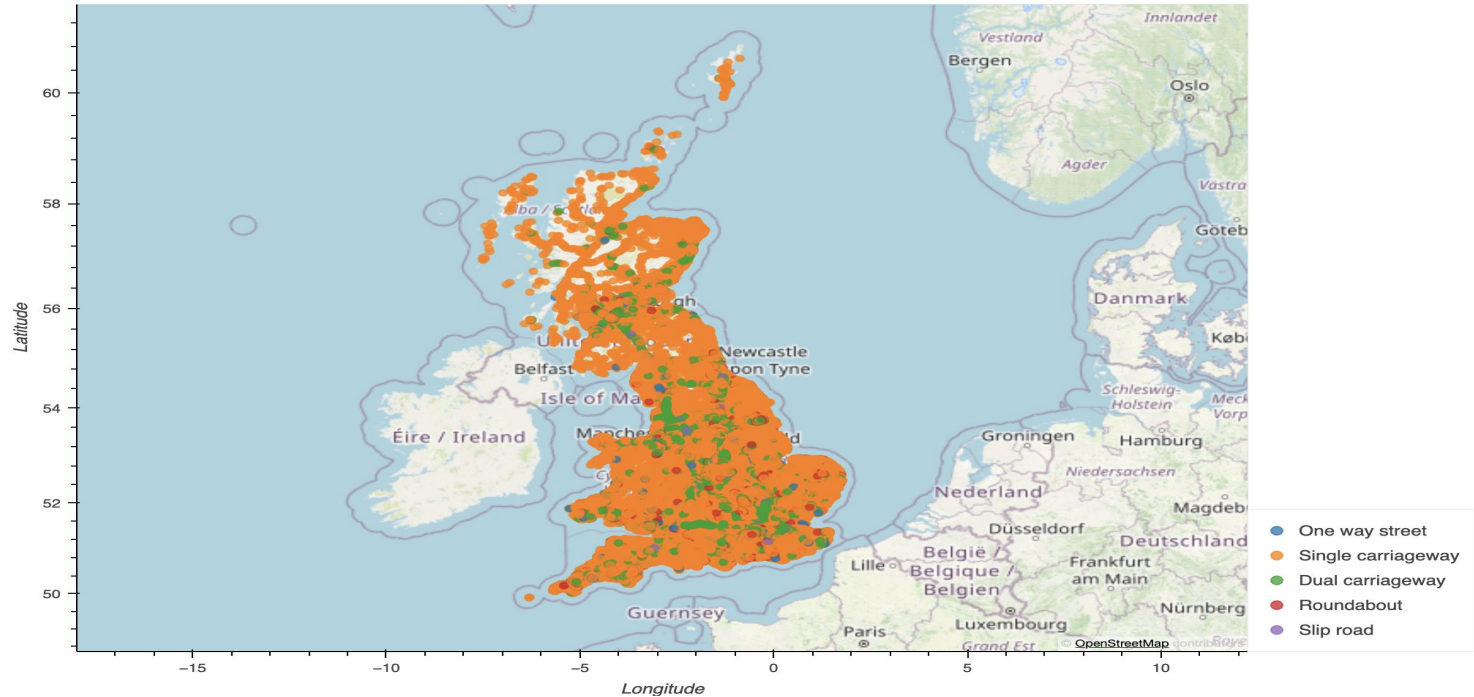
Urban vs Rural Casualties



Urban or Rural Area vs Road Type



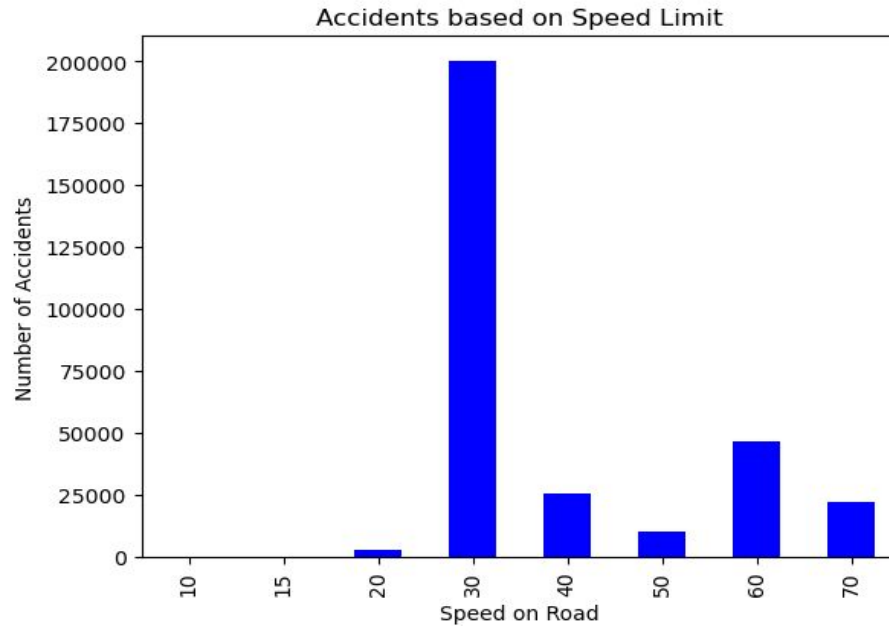
Geographic Location vs Road Type



Is the accident severity affected by the speed?

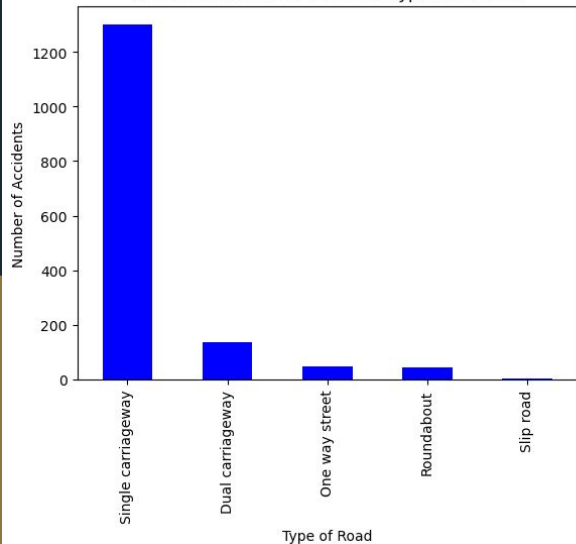


Speed Limit With The Most Car Accidents

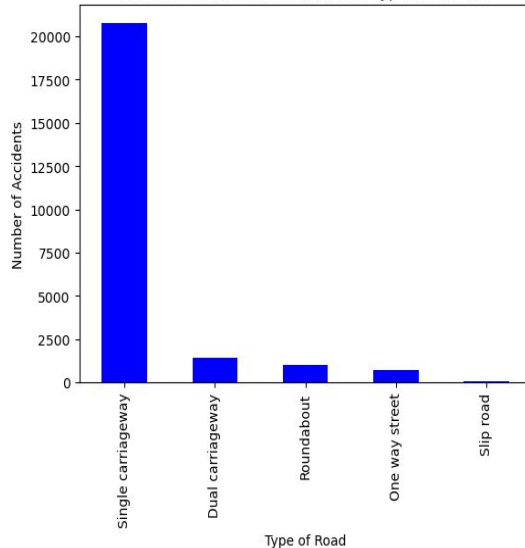


Number of Accidents Based on Road type in a 30 MPH Zone

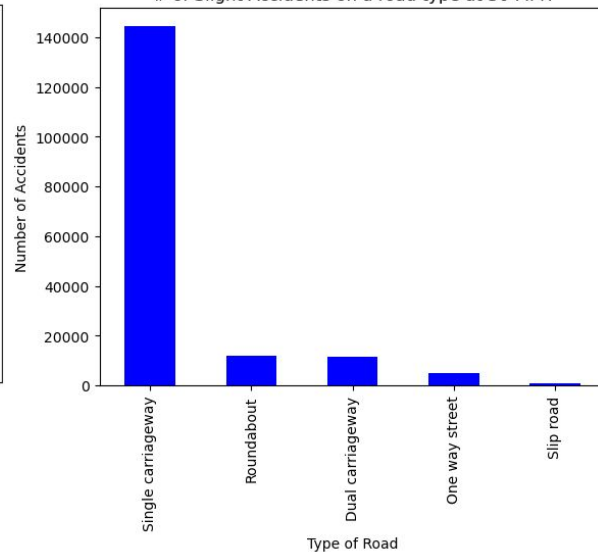
of Fatal Accidents on a road type at 30 MPH



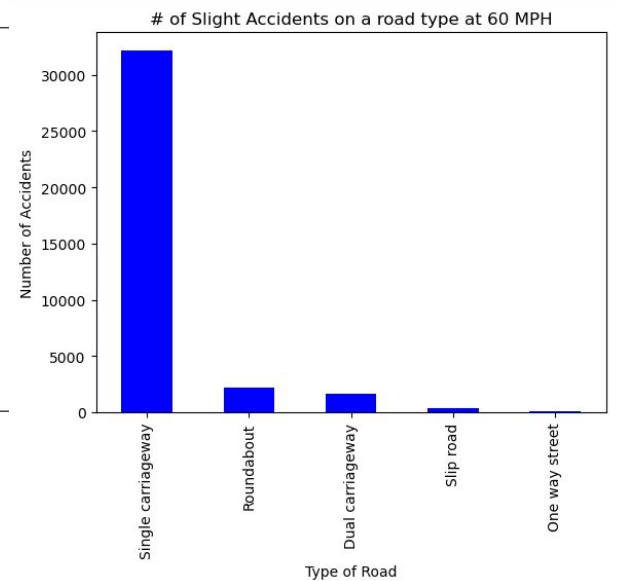
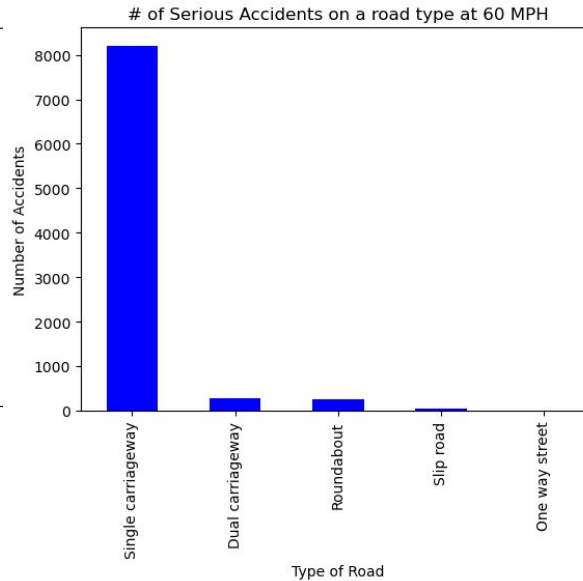
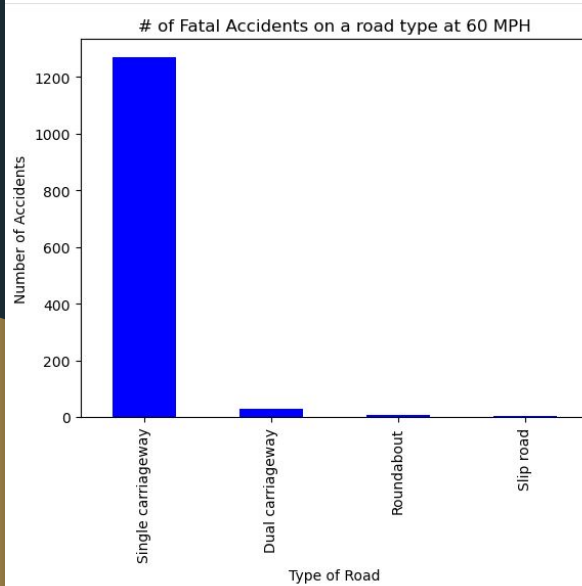
of Serious Accidents on a road type at 30 MPH



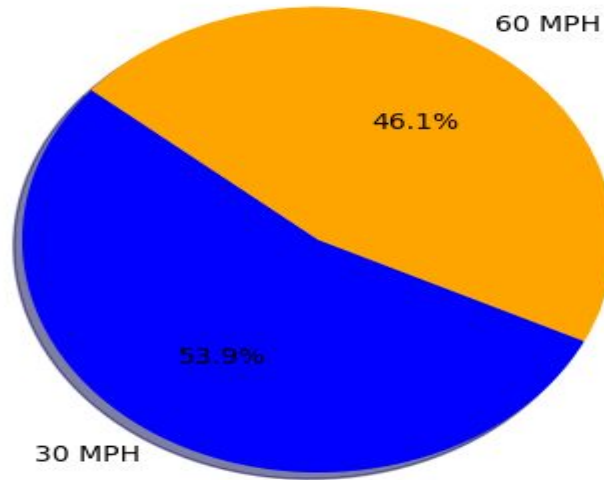
of Slight Accidents on a road type at 30 MPH



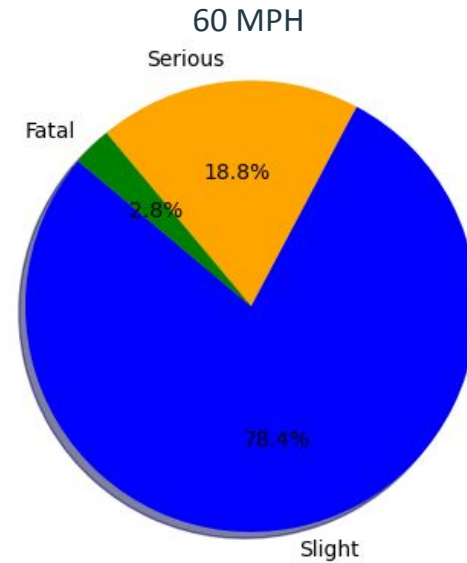
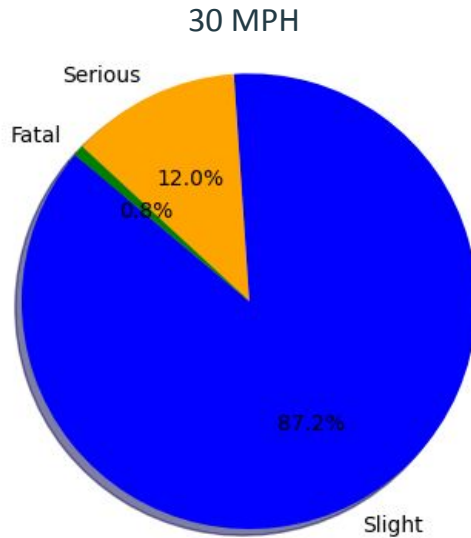
Number of Accidents Based on Road Type in a 60 MPH Zone



30 MPH vs 60 MPH Fatal Car Accidents



30 MPH vs 60 MPH Car Accident Percentages



T-test Results: Single Carriageway vs Dual Carriageway in a 30MPH zone

T-test Results: Statistic = -7.995497391534002, P-Value = 1.3866744692368699e-15, df = 14616.35828841124

Based on these results we can determine that this data is statistically significant. Dual carriageways are far more safer for drivers on the road than a single carriageway.



VS



Conclusion

- T-intersections is one of the leading causes for increased number of car accidents.
- T-Intersection vs. Roundabout
- Urban area prone to more accidents overall than rural
- Single carriageway most common accident in urban areas
- Majority of the casualties that occur due to car accidents are in urban areas rather than rural areas.
- It seems the only way we can mitigate the number of car crashes for the speed limit is to turn Single carriageways into dual carriageways.
- We would be able to decrease the number of car crashes by 1,300 for 30 MPH zones and 1,270 for 60 MPH zones.

Limitations

- The data set was limited to only two areas, Kensington and Chelsea
- Only one month of data collected
- Lack of numerical variables available
- Missing data in some of the variables
- Large amounts of outliers
- Cars may not have been going the speed limit

Questions?