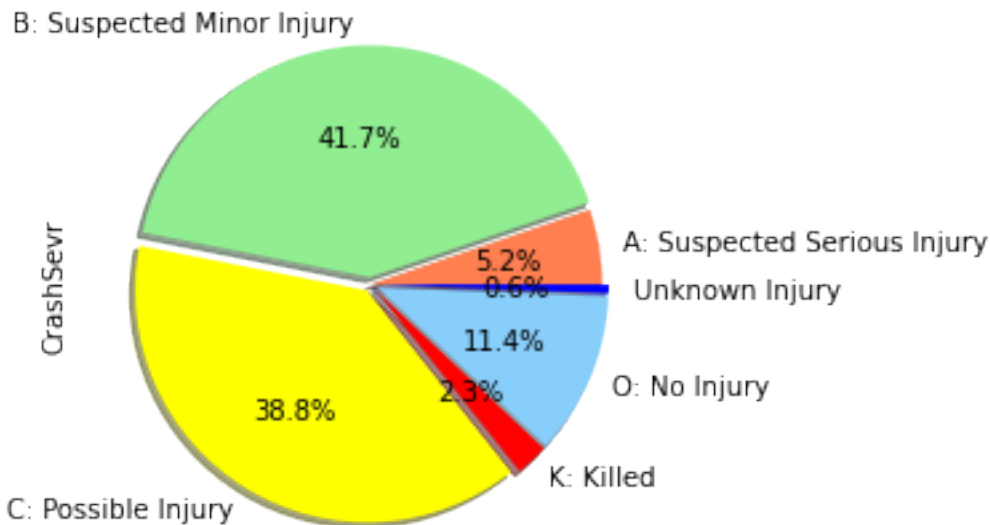


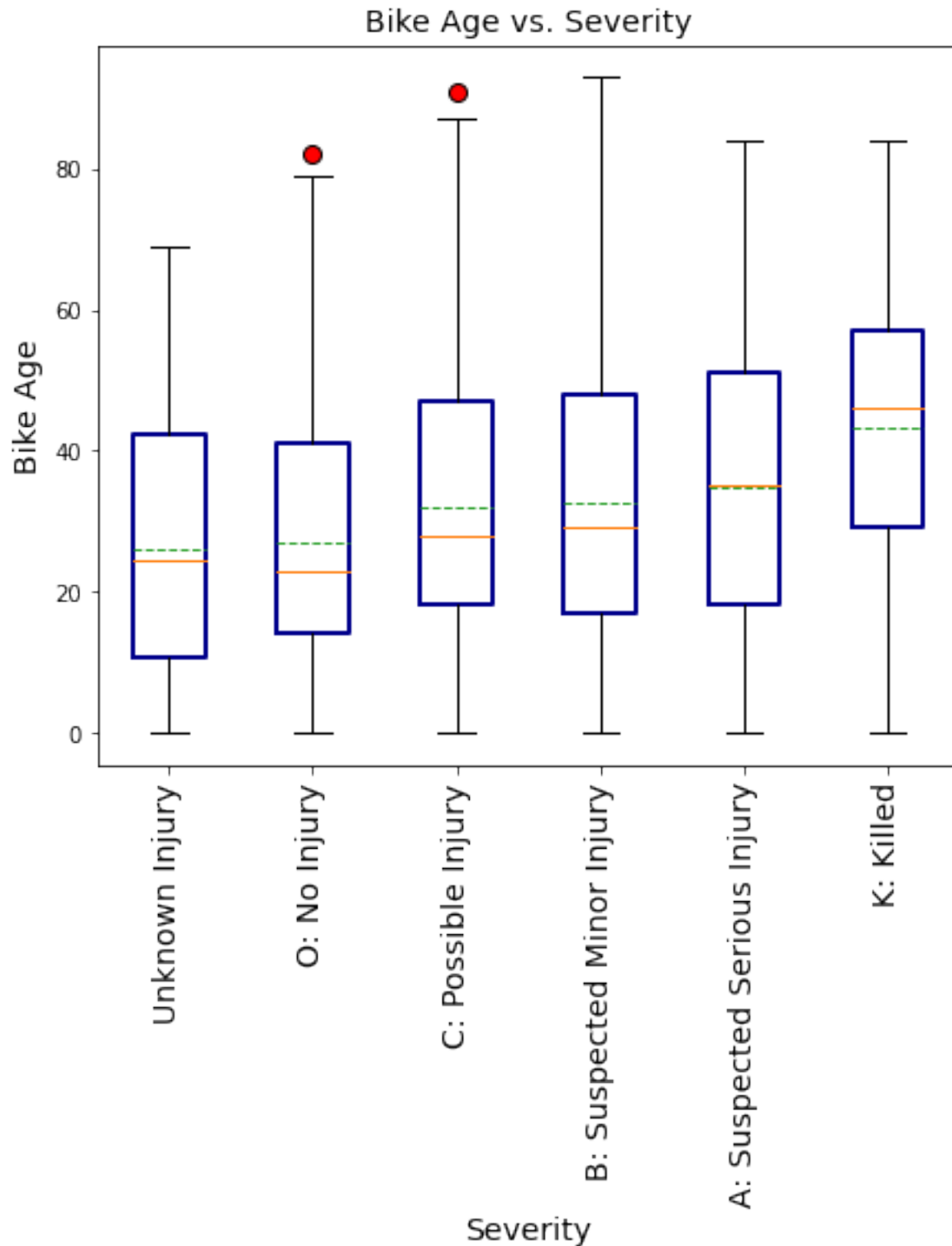
North Carolina Safety Report

For our project we wanted to find out about the safety of pedestrian and bicycle traffic in the North Carolina area. We searched the internet and found several robust datasets from the area. In fact, the datasets were so large that they took a long time to load on our computers so some cleaning and reformatting was required. We removed some of the columns that didn't contain relevant data to what we wanted to analyze using python and jupyter notebook. After doing that we picked some of the main datapoints we wanted to see visualized, such as crash severity and deaths. Included in the datasets were also the coordinates in latitude and longitude of all the reported crashes which allowed us to map the data in an interesting way using D3 and leaflet plugins, combined with JavaScript. Using a heat layer leaflet plugin it is easy to see the areas where the most crashes took place, which are around the most populous cities. Pedestrians had higher death numbers when hit by cars as compared to bicycles. MongoDB was used to help filter the data as well. We tried different JS D3 library leaflet plugins but some we had issues implementing so we went with the ones we could get to run smoothly with our data. An app was created to run our website and support the data. Here are our initial graphs with the data and some findings.

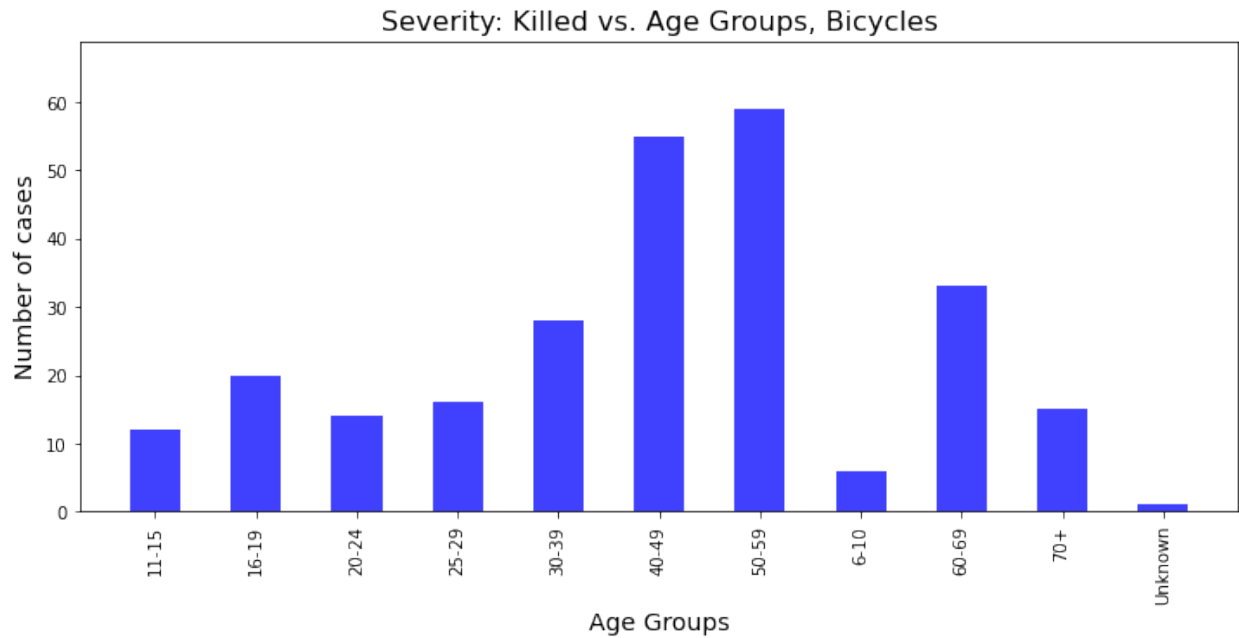
Summaries of the graphs/visualizations



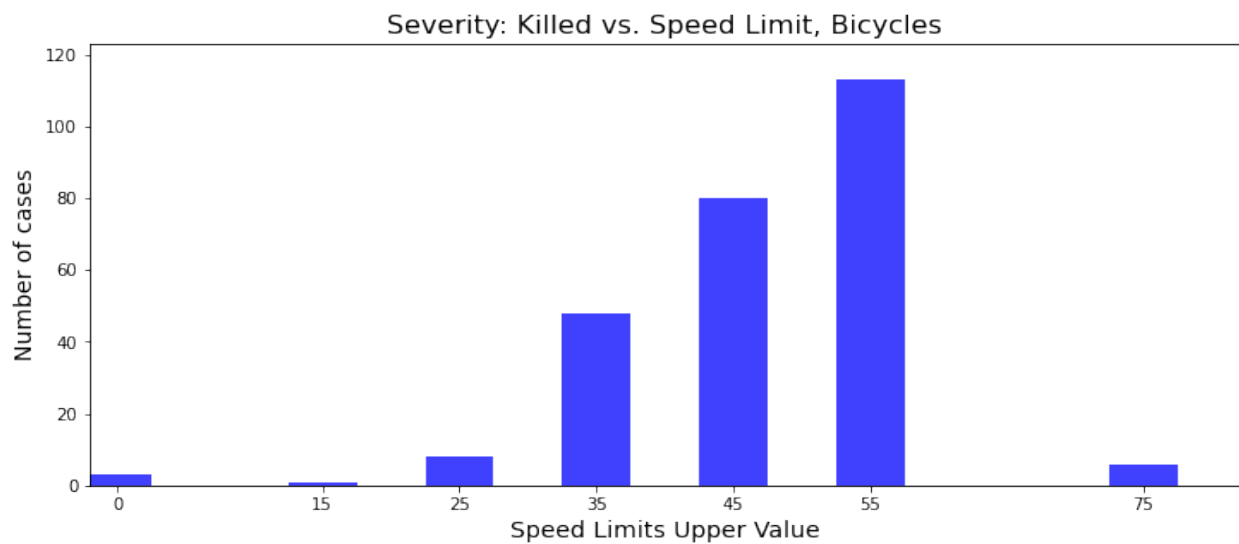
This pie chart visualizes the injury severity for the bicycle riders in the dataset. The type of injury is written next to the respective section of the pie chart.



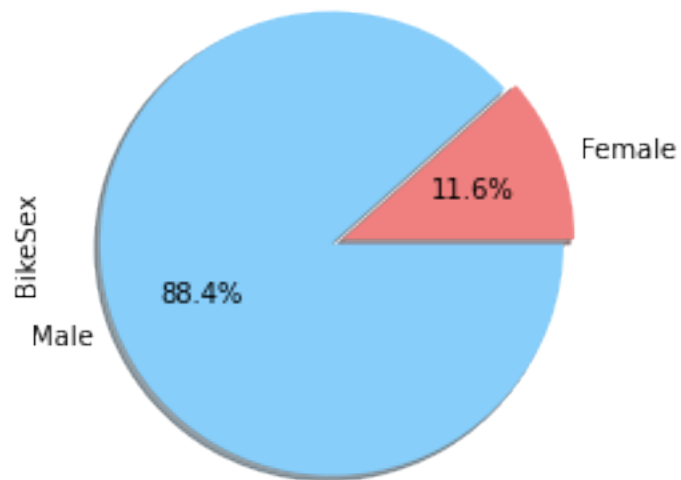
This boxplot visualizes the severity of the injuries received as compared to the ages of the bike riders. As you can see on the chart the higher the age of the rider the higher the chance you are to be seriously injured or even killed.



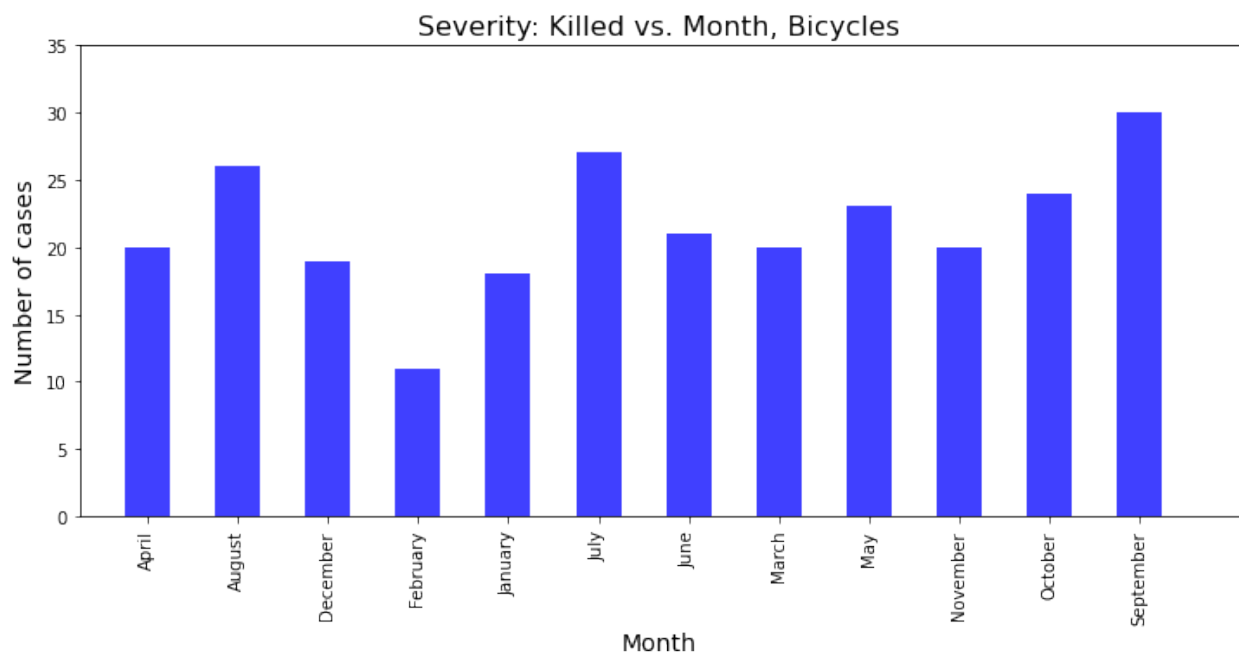
This bar chart shows the number of people killed riding bikes grouped by their respective age groups. Clearly shown is that the highest number of deaths occurred within the 40-49 age bracket and the 50-59 age bracket.



This bar chart shows the number of bicycle riders killed compared to the speed limits on the roads. The most deaths occurred on roads with a speed limit of 55.

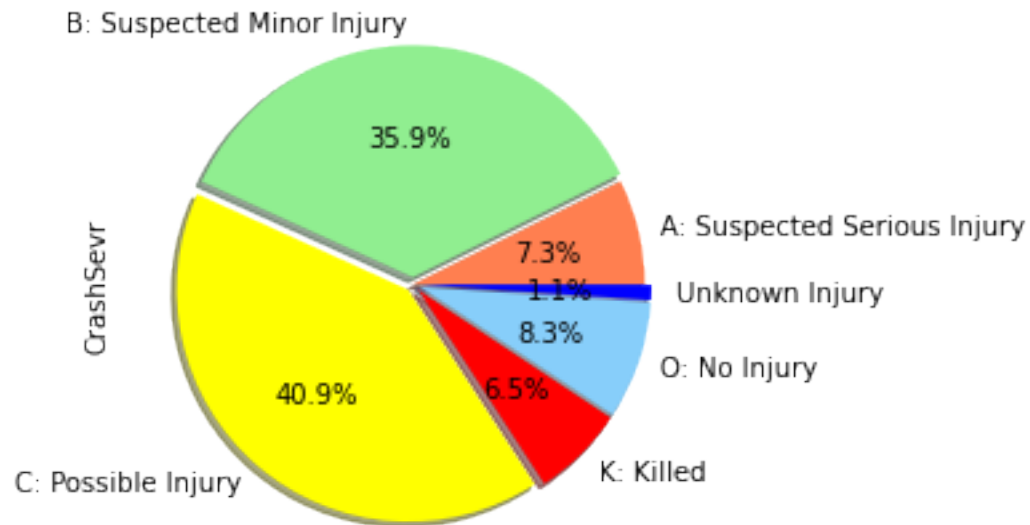


This Pie chart shows the gender breakdown of the bike riders killed in the dataset. The vast majority killed were male, as is clearly shown.

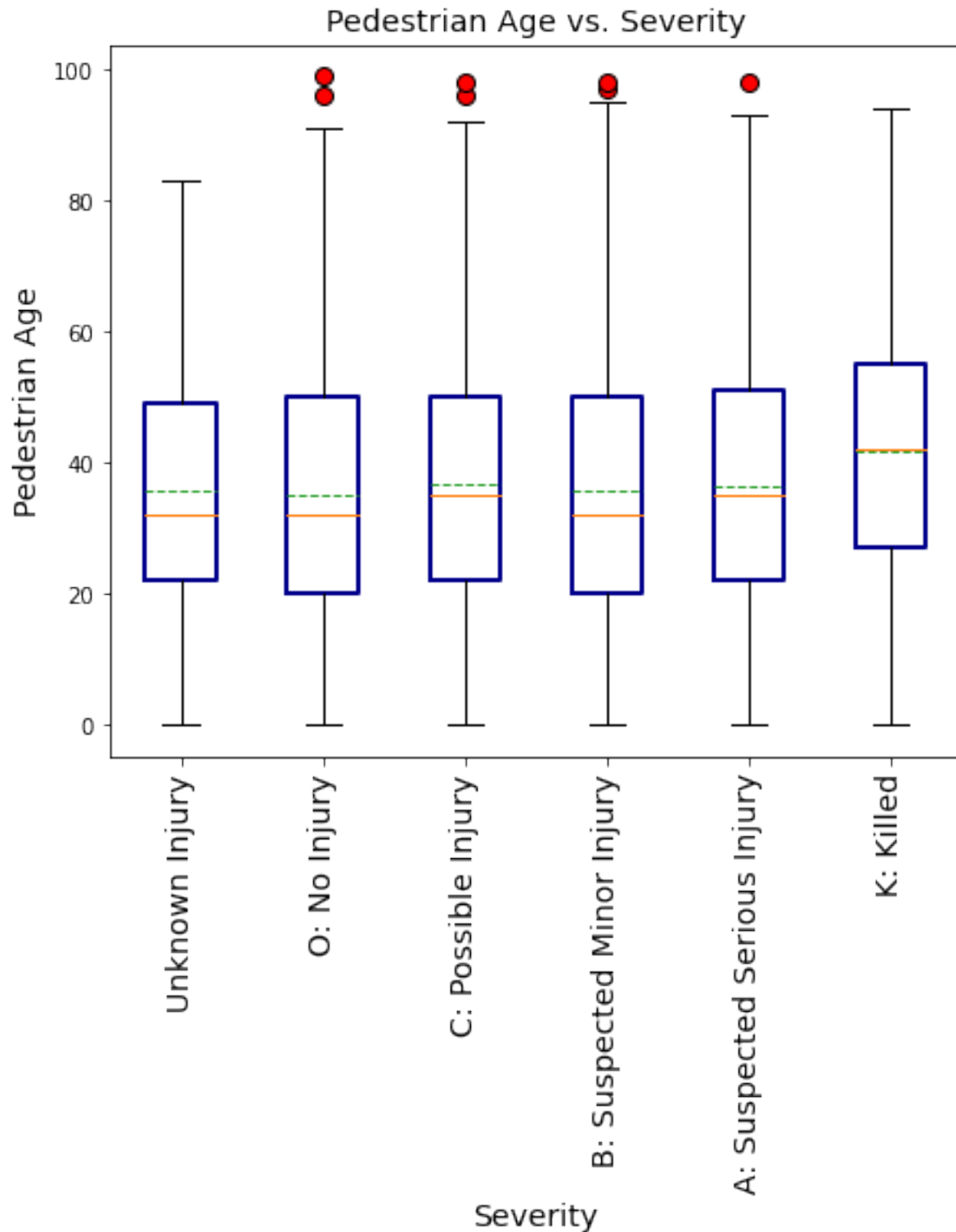


This bar chart shows the number of bike riders killed as broken down by the month of the incident. The highest numbers were from summer months with the lowest numbers in February.

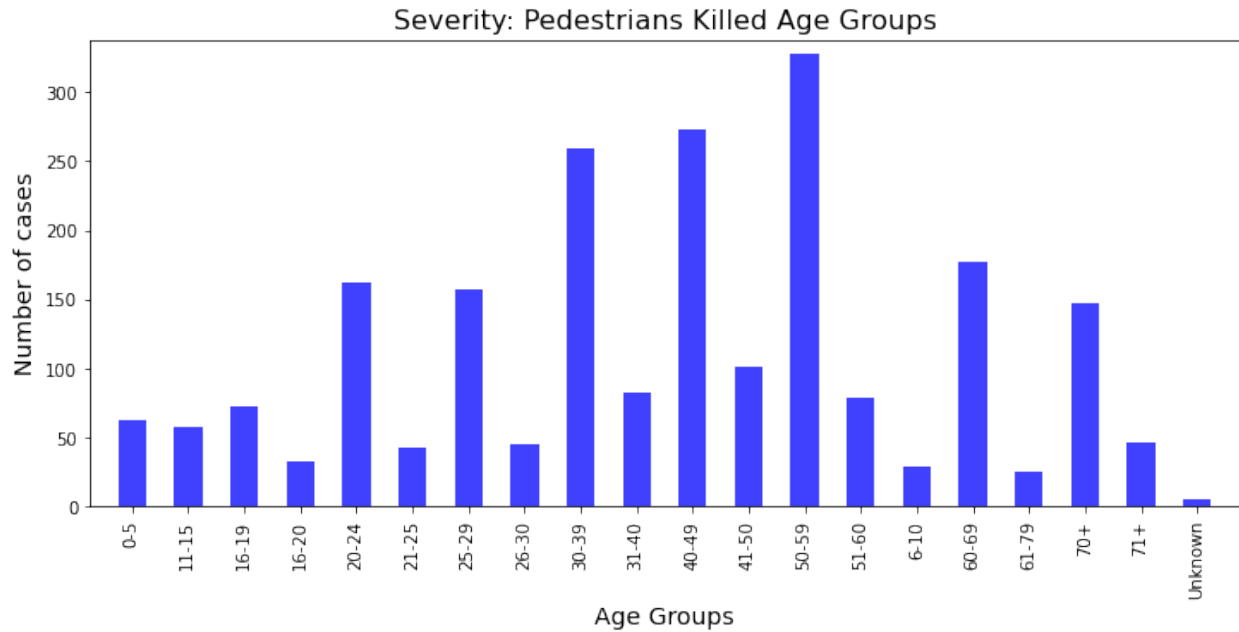
Next, we will look at the initial graphs and visualizations from the pedestrian dataset.



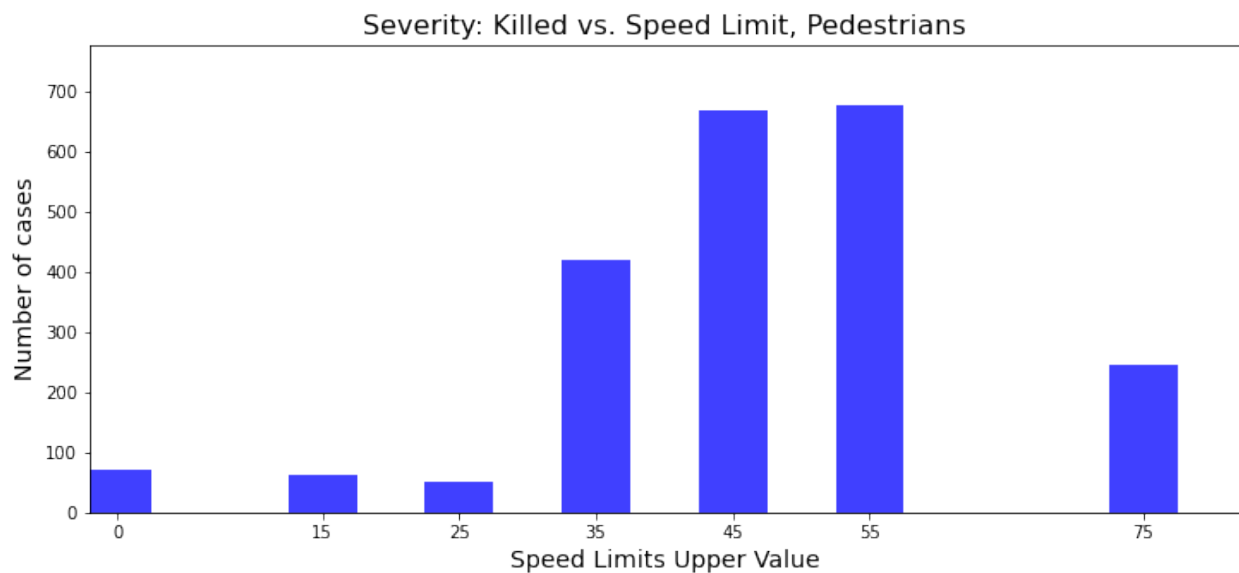
This pie chart visualizes the injury severity for the pedestrians in the dataset. The type of injury is written next to the respective section of the pie chart. The percentage of people killed walking as compared to bike riders is higher, 6.5% vs 2.3%.



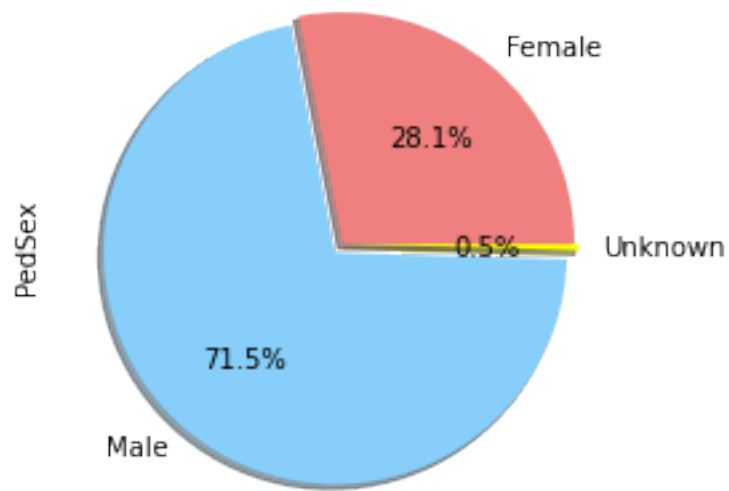
This boxplot visualizes the severity of the injuries received as compared to the ages of the pedestrians. As you can see on the chart the higher the age of the pedestrian the higher the chance you are to be seriously injured or even killed though the effect seems less than for the bike riders.



This bar chart shows the number of pedestrians killed grouped by their respective age groups. Clearly shown, again, is that the highest number of deaths occurred within the 40-49 age bracket and the 50-59 age bracket.



This bar chart shows the number of pedestrians killed compared to the speed limits on the roads. The most deaths occurred on roads with a speed limit of 55, with 45 following closely after.



This Pie chart shows the gender breakdown of the pedestrians killed in the dataset. The majority killed were male, though less than the bike riders: 71.5% vs 88.4%.

Conclusion

There were similarities in the dataset which can happen in datasets with a large number of data points. However, there were some differences when comparing the number of people killed in collisions compared to different factors. We will continue our analysis and offer more visualizations as well.