The dataset is an evaluation of all reported NYPD vehicle report data. Through visualization, we can learn some neat things about our data!

Data Story 1:

In the 10 Axis Parallel Coordinates Display, we select with brushing reported instances where cyclists have been injured. We can also select reported instances where pedestrians have been injured, and we can see that the two seem to be mutually exclusive. From this we can ascertain certain conclusions. The first conclusion we may initially draw is that cyclists simply are not "good" at injuring pedestrians. Perhaps they are able to dodge the pedestrians very well, or, the more likely scenario, they usually go on the roads. The other conclusion one may draw is a more thorough one. Since these injuries are NYPD reported data, we can induce that a human being must have called the police for that specific instance. It is very probable that in any injury involving a cyclist and a pedestrian either it would be too trivial to call the police, meaning that any damage done to either party is of minimal value.

Data Story 2:

In the 10 Axis Parallel Coordinates Display, we select with brushing reported instances where cyclists have been injured. We then can select or simply view that the time frame in which these instances have been reported was between the hours of 10 AM to 8 PM (note the transformations were done from a 24 hour time scale multiplied by .1). We can once again come to multiple conclusions for this. The first conclusion one may come to is that these hours are the most dangerous hours for cyclists. However, it is important to note that correlation does not imply causation. Therefore the more enlightened conclusion to come by is that peak cyclist traffic occurs at 10 AM to 8 PM. This also makes sense as we can apply some domain knowledge and further back this theory up by stating that the sun is usually out at these hours in NYC.

Data Story 3:

In the 10 Axis Parallel Coordinates Display, we select with brushing reported instances where persons have been injured. We then select instances to further filter, where pedestrians have been injured. We see that 0 cyclists have been injured. We can then deselect our pedestrian injured filter, then select instances where cyclists have been injured while still maintaining our persons injured filter. We notice how these two groups are mutually exclusive. We can come to a conclusion that if an accident occurs where a person (who is in the car) is injured, it is either with a cyclist OR with a pedestrian. We can notice a similar pattern as with our first data story, it is very likely that cyclists will be in the road, while pedestrians will be on the sidewalk. Therefore we can make a stronger conclusion: cyclists do not ride in sidewalks. Using domain knowledge, this makes complete sense as it is illegal for cyclists to ride on sidewalks in NYC. However it is neat that we can show this with data!