# Capstone Project 1 Data Story

We would like to explore the relationships between the weather and the time/average speed of a taxi trip.

The first thing we look at is the speed vs trip distance of the taxi trips, colored by the levels of precipitation. Darker dots represent taxi trips during times of higher levels of precipitation. From this scatter plot, we can see that in general the darker dots are concentrated near the bottom of the range, which indicates that maybe higher levels of precipitation is related to lower average speeds. We can do the same plot but this time for the temperatures, with darker dots representing higher temperatures. Here we can see that pattern isn't as clear, but still most of the taxi trips during low temperatures have low speeds and lower distances.
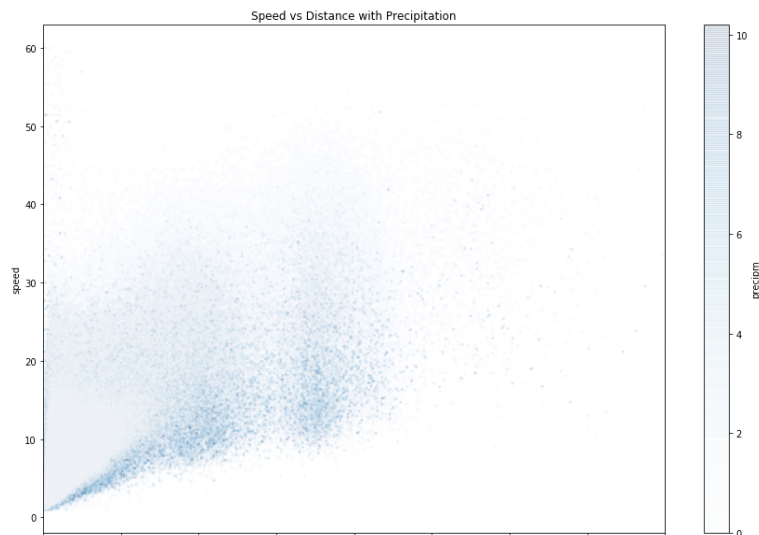
Next we can look at the distributions of speeds for the different weather conditions (rain, snow, clear). First we can plot the distributions against each other on a single plot so that they can be compared. From this, we can see that the distribution of speeds for the rainy trips is generally smaller than the distributions of speeds for the snow and clear trips. We can continue this exploratory analysis for different ranges of trip distances, as chosen using the 25, 50, and 75 percentiles of trip distances. From these plots, we can see that in general the speeds for the rainy days are generally lower than the speeds for the snow and clear days. One strange thing of note is for the trip distances in the 75 to 100 percentile, the rain remains slower but the speeds for trips in snow are faster. One possible explanation for this could be that when it snows less people go on trips of longer distances, so the roads may be less congested. For the snow days the relationships between the weather and the speed remain unclear. It does appear that the rain affects the speed across all trip distances.

Speed Distribution for Distances Between 0 and 188.9 miles
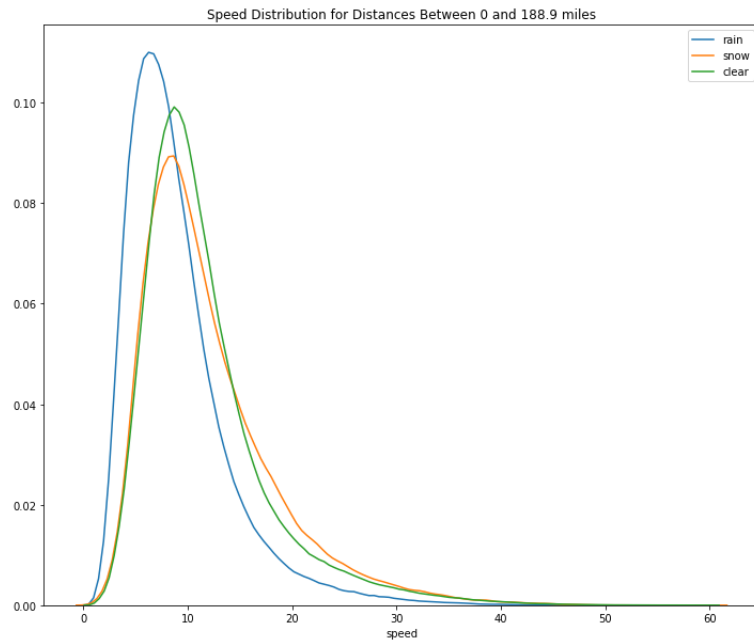
Using our suspicion that in general there are less cars on the roads during the snow days making the roads less congested on longer trips (generally using the highways), we can take a look at the distribution of trip distances for rain, snow and clear trips. From this we can see a big difference in that when it snows, the trip distances are generally smaller than when it rains or when it's clear.

Distribution of Trip Distances
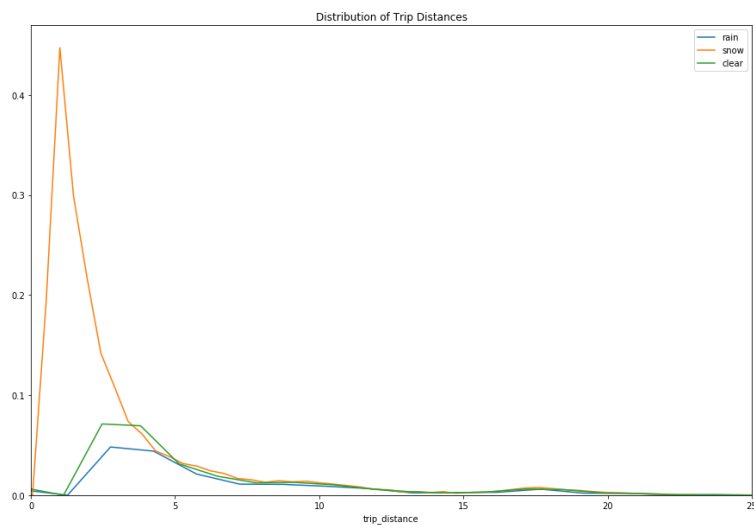
Next, we can maybe examine the relationship between the time of day and the speed of the taxi trip. First let's look at the number of taxi trips that happen per hour. We can plot them grouped by the weather condition. From this we can see that in general the number of taxi trips peaks at around 8am and 6pm, and is generally higher during the day leading up to 10pm and lower late night / early morning leading to 5am. Finally we can look at the average speed of a taxi trip for each hour of the day. Here we can see that the average speed peaks at 5am and is much

higher late night / early mornings. It is also much lower during the day leading up to around 6 or 7 pm. This suggests that there is a strong relationship between the hour of the day and the speed of the taxi trip.
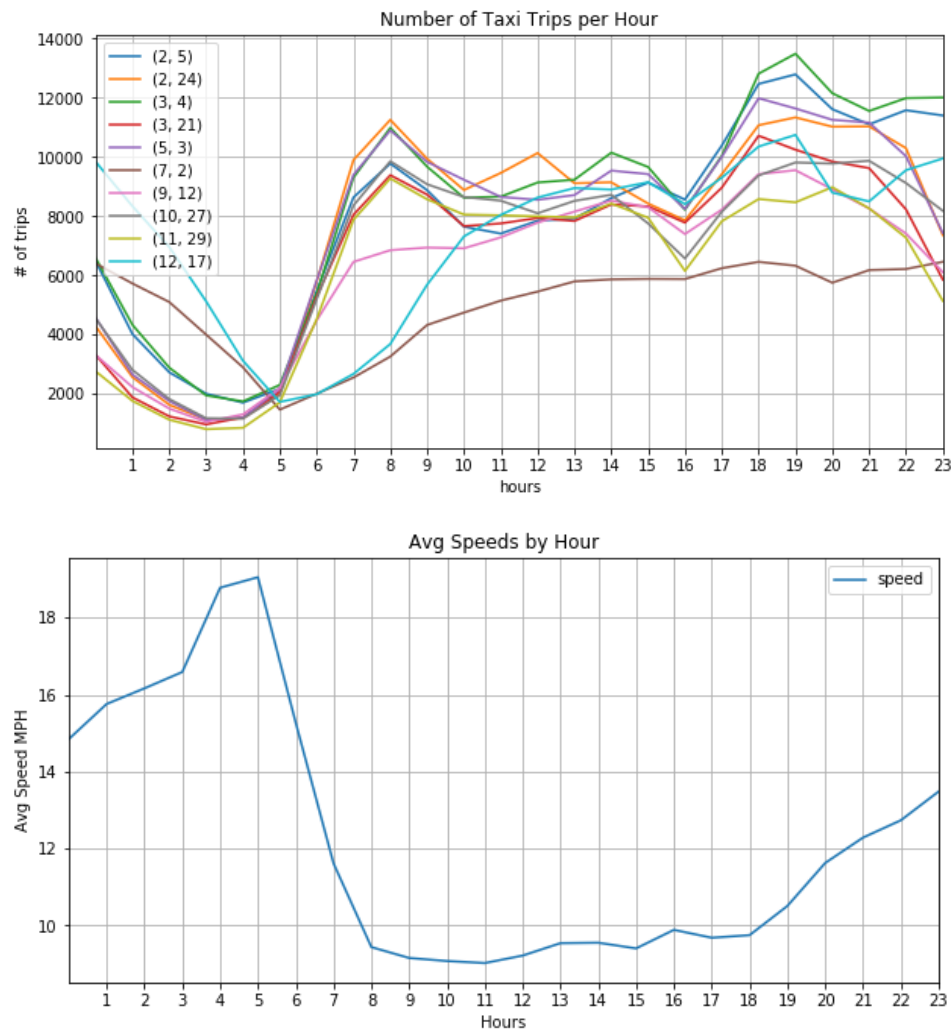
In conclusion, from this analysis we can hypothesize that higher precipitations generally slow down taxi trip speeds, as does the hour of the day. We can also hypothesize that days in which it rains will have slower travel speeds than clear and snow days.