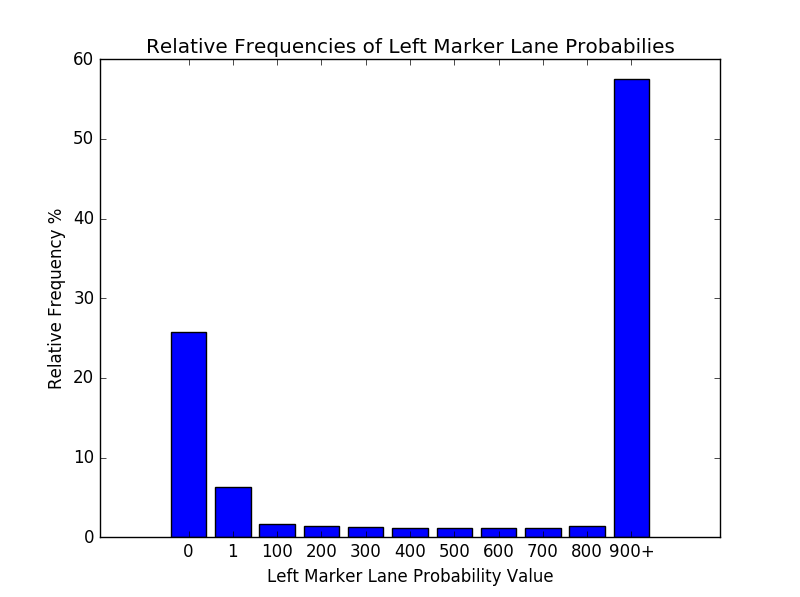
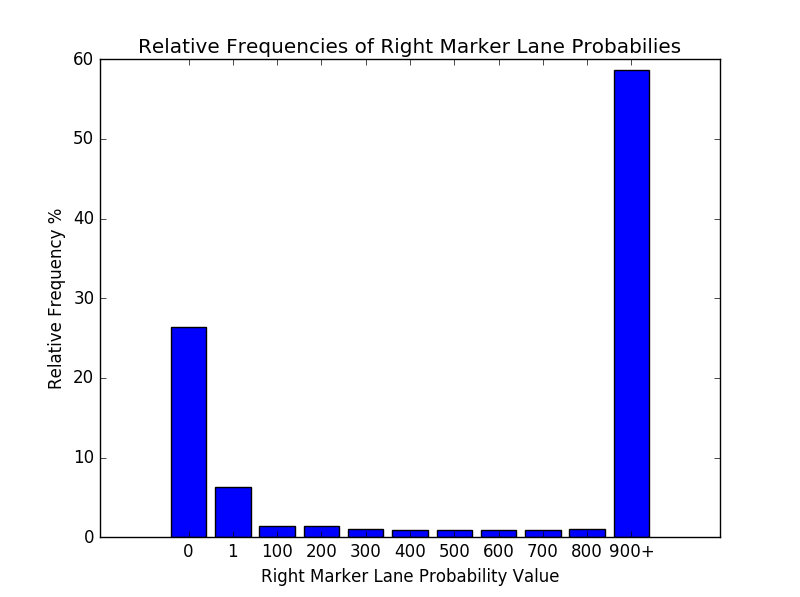
SHRP 2 contains multiple variables that measure lane positioning by the machine vision of a front facing camera mounted on the vehicle. Two variables measure confidence the machine vision is providing accurate lane positioning data, left marker probability and right marker probability. For insight as to how much of the lane positioning data is accurate, and at what levels of camera probability lane positioning data is accurate, a number of procedures were taken.

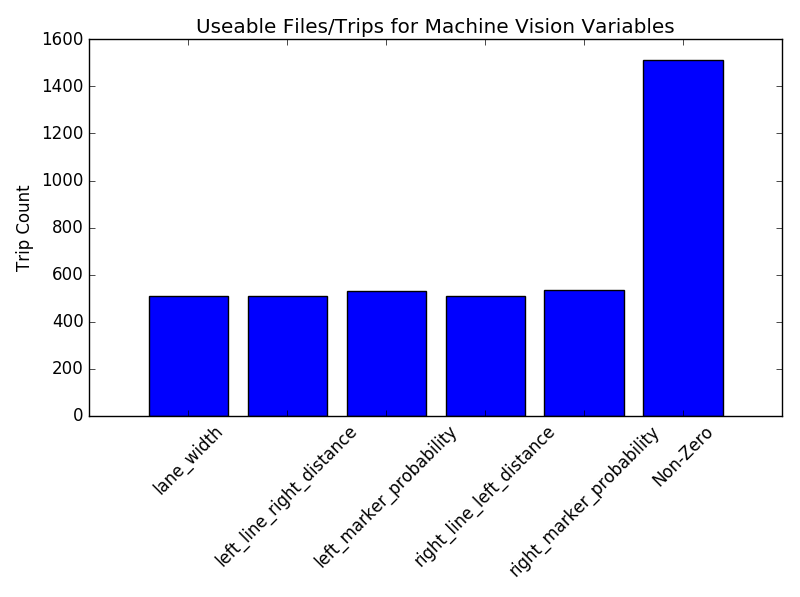
The first step in the procedure was to analyze the relative frequencies of lane probabilities occurring throughout all trips in the data set. This step was carried out by first concatenating all trips contained in the data set. Ranges of camera probability values were then binned and displayed in a histogram. The first bin measures the relative frequencies of zeros occurring throughout the dataset. Camera probability values are then binned in ranges of 100, 1024 being the maximum camera value possible. The results are displayed below:



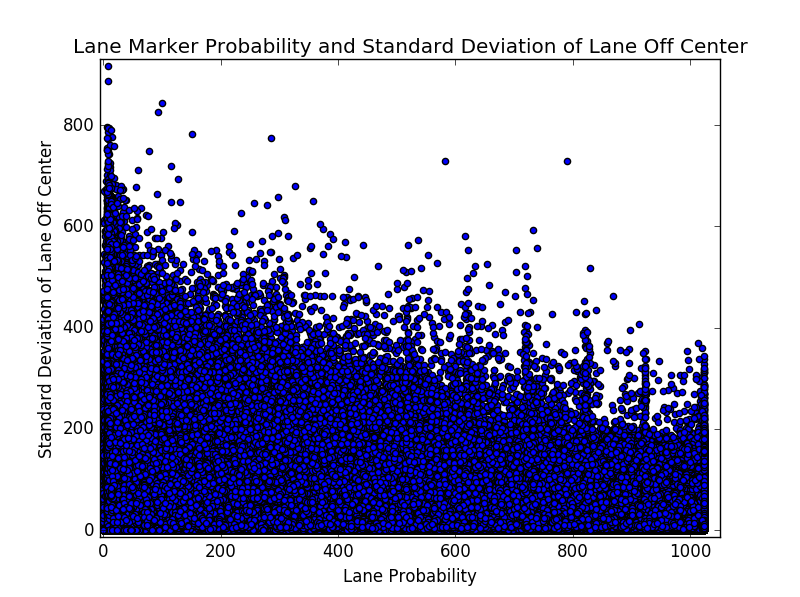


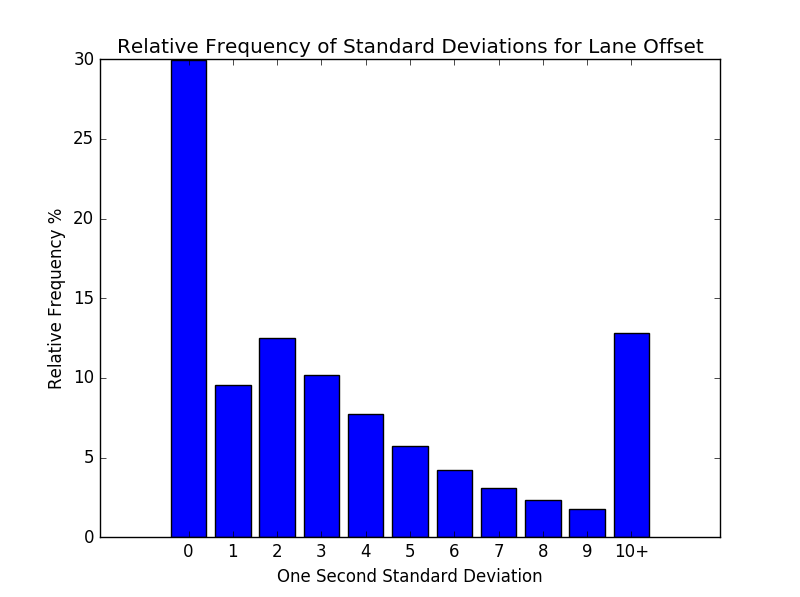
Note the frequency of 0’s occurring throughout the data set and the frequency of camera probability values falling between 900 and 1024. Considering these, it can be concluded 26% of the data sets lane positioning data is completely unusable. Of the trips with useable lane positioning data, camera probability was near the max (900-1024 range) 58% of the time.

To asses if 0’s occurring in the data set were isolated only to certain trips, a running count was taken across the data set. If all values in the machine vision columns were 0, 1 was added to the count for each machine vision variable. The results were as follows:

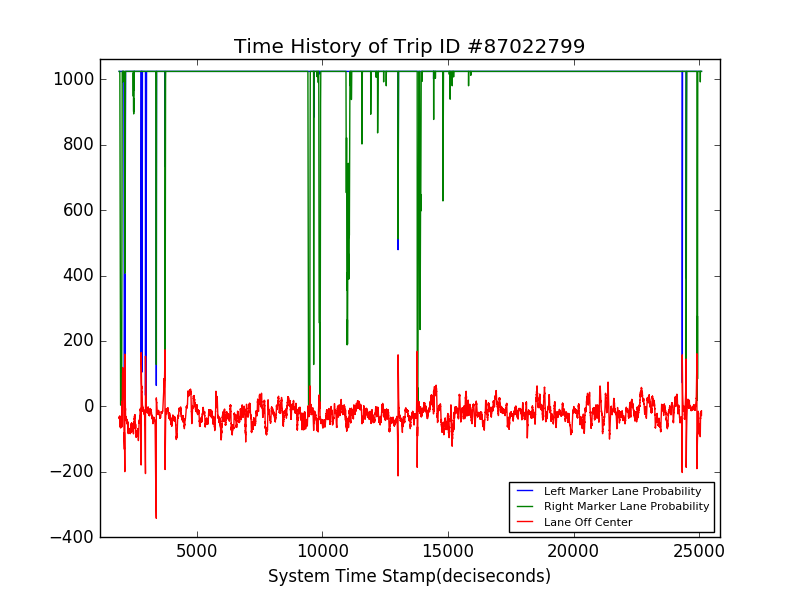


The next step in the procedure was to relate the variance of lane positioning variables to their corresponding levels of camera probability. It was hypothesized that at lower levels of camera probability there would be higher variance in lane positioning. To test this idea, the larger average of the two camera probabilities contained in a one second time window were plotted with the one second standard deviation of lane off center. The downward sloping relationship between camera probability and standard deviation of lane off center can be interpreted from the scatter plot below:

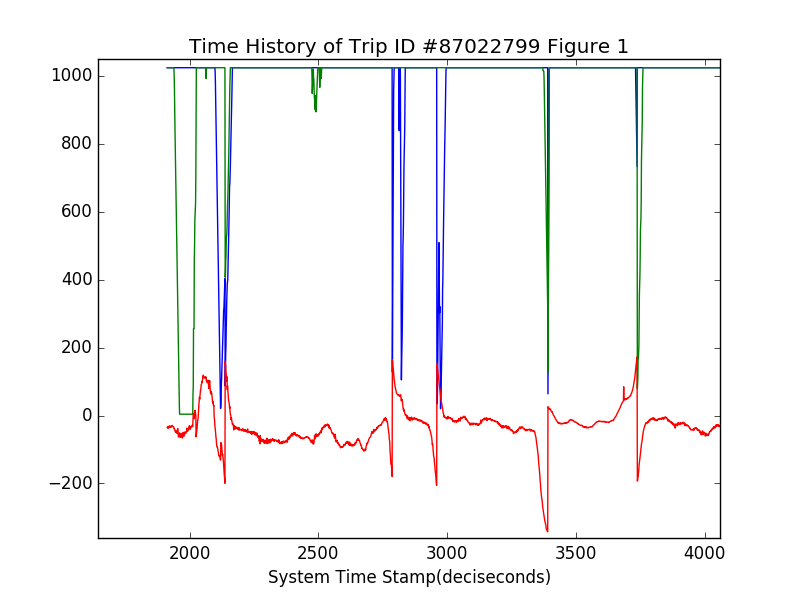


The relative frequencies of one second standard deviations plotted in the chart above can be visualized below: 

The next part of the process was to observe the relationship between camera confidence and the variance of lane positioning variables within individual trips. Trips were first ranked by their average left lane probability. The timeseries plot below is for the best trip in the dataset ranked by average left lane probability:



Note that readings for lane off center become considerably more volatile with corresponding drops in camera confidence:

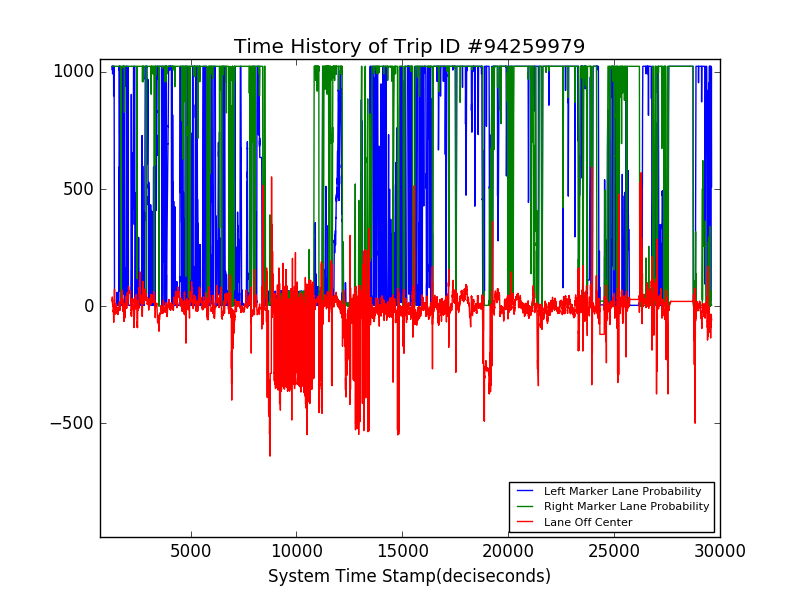


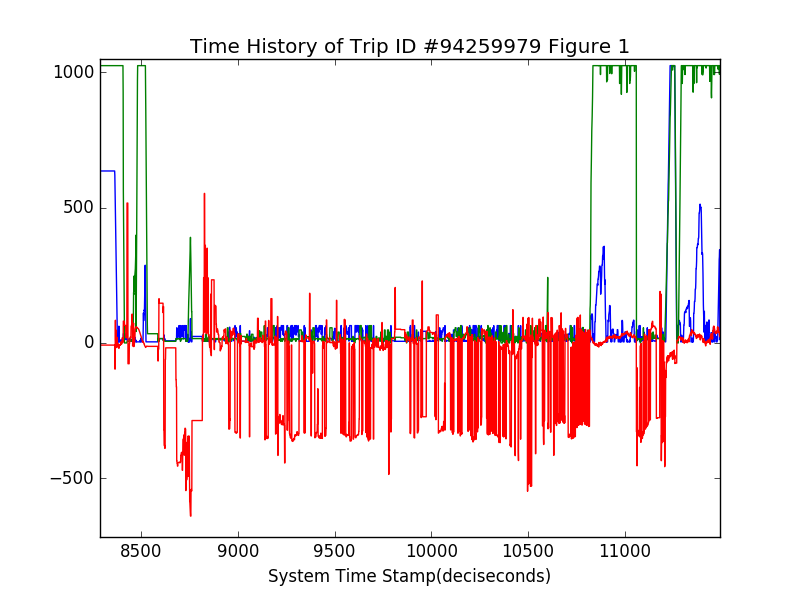
The volatility in this trip between 2,000 and 2,500 deci seconds can be explained by a lane change:

The volatility in this trip between 2,500 and 3,500 deci seconds can be explained by the driver entering an exchange ramp to a sepearte multiple lane highway. Note that readings from the left and right marker fall near 0 in the time sereis as the driver enters the on ramp.



Next is a timeseries plot for a trip with an average left marker probability of 509, or roughly 50% of the max. Observe the increase in the volatility of lane off center when both lane probabilities drop near 0:





An observation one can make from the trip with average probability 50% of the max is that the resolution looks considerably more blurry. The image also appears to have been taken from a camera with a wider angle than the previous: