Our project will be an investigation into thoroughbred racing, to determine if a correlation exists from the horse lineage, betting odds, and payouts. The winning ancestry of a thoroughbred is one of the key aspects used to determine the odds of a horse winning a given race. The opening odds are based on this information, as well as recent race outcomes. As more wagers are placed on a particular horse, their odds decrease because the house never loses. We will examine past race history looking for indicators that horses with lower odds of victory, but good genes, have a better chance of large payout.

The first and only source we have for our project right now comes from the website kaggle.com, which has a file which contains horse racing records. It holds about 13,000 horses and 30,000 different individual matches between the horses, their markets, and their runners. The attributes for the markets and horses include their strong wins and wins in general, information regarding their last twenty starts and their days since last run. For the horses themselves, some of the information they hold is what age they are and how much money they have won. We wanted to ask if there are some traits which can help predict winners, so those attributes will help with that.

In order to determine the best payout probability for horse racing, we are comparing data sets against the official betting odds. In addition to reverse-engineering the method of odds generated by the derby, we are seeing if there is a correlation between siring genealogy vs. the win ratio. The main purpose of this study would help people place optimal bets at derbies, but the data analysis should also show the optimal genealogy for a winning race horse, which would be used to help raise and train these steeds.

A dynamic pie chart will enable us to switch different views on optimal payout probabilities based on rider vs. horse, weather vs. horse, house odds vs. horse. We will also have a graph on siring genealogy vs wins; wins based on generation.