

# Springboard Capstone Proposal

*L Lathrop*

*10/21/2016*

## THE PROBLEM

The average number of breakdowns leading to equine death on American racetracks equals approximately 1.8 per 1000 starts, though the same figures for synthetic and turf tracks are lower. However, even one equine death as a result of racetrack injury is one too many. This capston project, therefore, will attempt to determine the factors that lead to such events.

## THE CLIENT

The client for this project will be The Jockey Club, the Thoroughbred industry's governing body. Results from the capstone project will then be made available to industry trainers and vets so that they may classify high-risk horses and respond accordingly by not racing or exercising them.

## THE DATA

The project will rely on both public and proprietary data. The New York Racing Association (NYRA) has made data on breakdowns and deaths publicly available since 2008. This will form the core of the project. Additional data on horses *not* injured for all New York races will come from Equibase, a subsidiary of The Jockey Club. This data will include both performance (i.e., racing) and training data. Information on breeding will come from Equineline, also a subsidiary of The Jockey Club. This data may be significant because it is widely known in the industry that breeders breed horses more for speed than for stamina, though this has not always been the case.

## METHODOLOGY

Once all of the data is acquired, it will likely require significant cleaning and preprocessing. Once that is complete, the full dataset will be split into training and testing sets. The training set will be used to develop a robust classification model. This model will then be used on the testing set to determine its validity as a tool.

## DELIVERABLES

The following deliverables will result:

- \* The **code** for the model
- \* A **data journalism article** to be published in *The Blood-Horse*, which is the industry's most trusted magazine
- \* A **slide deck** that will explain the model and results
- \* A **Shiny app** that trainers and vets can use to classify their horses (This may not be complete by the end of the course.)