# Tracking the Danger Zone

## The Dangers of Horse Racing in New York State

## Data Description

Within the state of New York, any incident, death, or breakdown of any horse at any facility under the Gaming Commission’s jurisdiction is reported to the Presiding Judge or Stewart at that facility which then provides the details of the incident to the Commission’s main office in Schenectady, New York. The verified data is entered into a database that contains 13 components that detail over 4,000 incidents at New York State equine racetracks between 2009 and 2019.

## Compelling Story

As an equestrian and horse owner, I believe that injuries and deaths happen on equine racetracks more often than they should. I am interesting in understanding if certain division of racing, certain tracks or other factors influence the high number of incidents.

## Audience

The audience would be the New York State Gaming Commission, horse owners, jockeys and gamblers with concerns around animal cruelty. In addition, anyone with a love of horses may be interested in this information.

## Text

Review text from other strong posters to arrive on what to write

### Sources

Add sources for improvements on tracks, trainers by starts

## Plots

1. What are the most reoccurring incidents happening on racetracks?  
   Single Dimensional: Frequency of Incidents by Incident Type (barplot – Plot1)  
   \*\* Make this focal point \*\*  
   \*\* sort by largest number decreasing \*\*
2. Are the incidents different for the two racing divisions: harness or thoroughbred?  
   Multi-Dimensional: Incident Type by Division (barplot – Plot4)  
   *needs a color change, maybe different type of plot?*   
   \*\* Maybe do each type of incident and then color code by division? \*\*  
   \*\* maybe stacked bar taking up entire bar)
3. Do specific tracks encounter a higher number of incidents?  
   Single Dimensional: Frequency of Incidents by Type (piechart – Plot2)  
   *needs to be normalized – set up to be a percentage – check Lab6 (red/white stacked plot)*Multi-Dimensional: Incident Type by Track (Plot5) – check this and if important
4. Does weather play a role in the types of incidents?  
   *Needs to be created – plot type TBD.   
   Working on function for historical weather by zipcode of track*
5. Are there any commonalities between the incidents across division and/or track?  
   WordCloud on Descriptions
6. Does track location in New York State have any bearing on accidents and deaths?  
   Map of NY
7. Death over years in general?  
   (I think I have this … maybe a step graph).
8. Are their certain trainers that have more incidents than others?  
   *Plot to be created*
9. What is the trend in the number of incidents over time?  
   *Completed with PowerBI (only death and by division and by track) – Need to recreate in R  
   maybe clean up –* can I use PowerBI dataset?  
   Check Lab6 – region by type over years.