

# PS4 McGuire

Joel McGuire

February 2018

## 1 Data Sources

Twitter, specifically the text from a random sample of all tweets during a given period.

## 2 Other Questions

6.4 What is the class of each? We get a `SparkDataFrame` and a `data.frame`.

6.(5.b / 6.b) What happens when you use `filter` or `select` in a `SparkR` session on a traditional dataframe? It says there's no inherited method for a `data.frame`. R doesn't know what to do because the method is specifically built into the base language of `SparkR` for `SparkDataFrames`. `SparkR` appears to be built built on the same root package as `dplyr`, `plyr`.