PS4

Audrey Hopewell

2/11/2020

1   Data Sources

I am interested in gathering data from Twitter, and I am interested in the film industry/film criticism, so I think it would be interesting to collect data about movies. Since it is currently "awards season," there are plenty of people (both film critics/culture writers and lay movie fans) talking about movies on Twitter. However, this isn't typically unified by a specific hashtag, so figuring out how to filter and get the data I want would probably take some work.

2   Question 6

1. df1 is a data frame, df is a "SparkDataFrame"

2. we get an error saying "unable to find an inherited method for function 'select' for signature '"data.frame"'" because select is a Spark command that doesn't work with a normal R data frame

3. we get an error saying "unable to find an inherited method for function 'filter' for signature '"data.frame"'" because filter is a Spark command that doesn't work with a normal R data frame