PS4 Yuhas

Michael Yuhas

February 13, 2018

1 Data Scraping

- Biotechnology firm financial data

- Book data on Arthurian Legend potentially from Gutenberg

- international bankruptcy data

- music audio data and pattern analysis

2 Exercise Answers

6.4: The class of df1 is a data frame and the class of df is a Spark program data frame. 6.5.b: The function cannot run because the data frame is not formatted in a way that allows the spark package function to be used. 6.6.b: The function cannot run for the same reason as previously stated.