

PS4_Li

Donald Li

February 2018

1 Data Scraping

I would be interested in taking a look at data gathered from Twitter. Mainly in trying to find the relationship between peoples attitudes on environmentalism and see if they have a relationship between to the weather in their city. It's easy to be green when it is convenient but what about when it is 102F? Depending on the outcome the data could be used to find the optimal times to introduce legislation in areas where views on global warming are conflicting with science.

2 Answers the HW

```
4 > class(df1) "data.frame"
   > class(df) "SparkDataFrame" attr(,"package") "SparkR"
5
```

I think the sparkR uses a different function than select to print data from columns.

6 It is probably because df1 is classified as a "data.frame" as opposed to df which is SparkR