



Quiz 1: Getting Started

Quiz, 6 questions

1
point

1. The `en_US.blogs.txt` file is how many megabytes?
- ☒ 200
 - ☐ 250
 - ☐ 100
 - ☐ 150

1
point

2. The `en_US.twitter.txt` has how many lines of text?
- ☐ Around 2 hundred thousand
 - ☒ Over 2 million
 - ☐ Around 1 million
 - ☐ Around 5 hundred thousand

1
point

3. What is the length of the longest line seen in any of the three `en_US` data sets?
- ☐ Over 11 thousand in the blogs data set
 - ☐ Over 40 thousand in the news data set
 - ☐ Over 11 thousand in the news data set
 - ☒ Over 40 thousand in the blogs data set

1
point

4. In the `en_US` twitter data set, if you divide the number of lines where the word "love" (all lowercase) occurs by the number of lines the word "hate" (all lowercase) occurs, about what do you get?
- ☐ 0.25
 - ☐ 2
 - ☐ 0.5
 - ☒ 4

1
point

5. The one tweet in the `en_US` twitter data set that matches the word "biostats" says what?
- ☐ They just enrolled in a biostat program
 - ☐ They need biostats help on their project
 - ☐ It's a tweet about Jeff Leek from one of his students in class
 - ☒ They haven't studied for their biostats exam

1
point

6. How many tweets have the exact characters "A computer once beat me at chess, but it was no match for me at kickboxing". (I.e. the line matches those characters exactly.)
- ☐ 2
 - ☐ 0
 - ☐ 1
 - ☒ 3

☐ I, **Vijay Ramanujam**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.
[Learn more about Coursera's Honor Code](#)

Submit Quiz