

Feedback — Quiz: Exploring the Data

[Help Center](#)

Thank you. Your submission for this quiz was received.

You submitted this quiz on **Sun 11 Oct 2015 12:26 PM SGT**. You got a score of **11.00** out of **11.00**.

The purpose of this quiz is to make sure that you've downloaded the data and started working with it. Before attempting the quiz, download the data, untar it and make sure you can read it into a software system and convert the records into a usable format.

Question 1

After untaring the the dataset, how many files are there (including the documentation pdfs)?

Your Answer	Score	Explanation
<input type="radio"/> 3		
<input type="radio"/> 5		
<input checked="" type="radio"/> 7	✓ 1.00	
<input type="radio"/> 2		
Total	1.00 / 1.00	

Question 2

The data files are in what format?

Your Answer	Score	Explanation
<input type="radio"/> csv		
<input checked="" type="radio"/> json	✓ 1.00	
<input type="radio"/> .xlsx		

☐ .RData

Total

1.00 / 1.00

Question 3

How many lines of text are there in the reviews file (in orders of magnitude)?

Your Answer	Score	Explanation
<input type="radio"/> Ten million		
<input type="radio"/> One hundred thousand		
<input checked="" type="radio"/> One million	✓ 1.00	
<input type="radio"/> Ten thousand		
Total	1.00 / 1.00	

Question 4

Consider line 100 of the reviews file. "I've been going to the Grab n Eat for almost XXX years"

Your Answer	Score	Explanation
<input checked="" type="radio"/> 20	✓ 1.00	
<input type="radio"/> 2		
<input type="radio"/> 5		
<input type="radio"/> 10		
Total	1.00 / 1.00	

Question 5

What percentage of the reviews are five star reviews (rounded to the nearest percentage point)?

Your Answer**Score****Explanation**☐ 14%☒ 37%

1.00

☐ 10%☐ 30%

Total

1.00 / 1.00

Question 6

How many lines are there in the businesses file?

Your Answer**Score****Explanation**☐ Around 55 million☐ Around 15 million☐ Around 1.5 million☒ Around 60 thousand

1.00

Total

1.00 / 1.00

Question 7

Conditional on having an response for the attribute "Wi-Fi", how many businesses are reported for having free wi-fi (rounded to the nearest percentage point)?

Your Answer**Score****Explanation**☒ 40%

1.00

☐ 2%☐ 57%☐ 1%

Total

1.00 / 1.00

Question 8

How many lines are in the tip file?

Your Answer	Score	Explanation
<input type="radio"/> About 60 thousand		
<input type="radio"/> About 55 million		
<input type="radio"/> About 1.5 million		
<input checked="" type="radio"/> About 500 thousand	✓ 1.00	
Total	1.00 / 1.00	

Question 9

In the tips file on the 1,000th line, fill in the blank: "Consistently terrible _____"

Your Answer	Score	Explanation
<input checked="" type="radio"/> service	✓ 1.00	
<input type="radio"/> atmosphere		
<input type="radio"/> desserts		
<input type="radio"/> food		
Total	1.00 / 1.00	

Question 10

What is the name of the user with over 10,000 compliment votes of type "funny"?

Your Answer	Score	Explanation
<input type="radio"/> Ira		
<input checked="" type="radio"/> Brian	✓ 1.00	

☐ Jeff

☐ Roger

Total 1.00 / 1.00

Question 11

Create a 2 by 2 cross tabulation table of when a user has more than 1 fans to if the user has more than 1 compliment vote of type "funny". Treat missing values as 0 (fans or votes of that type). Pass the 2 by 2 table to `fisher.test` in R. What is the P-value for the test of independence?

Your Answer	Score	Explanation
<input type="radio"/> around 0.01		
<input type="radio"/> around 0.05		
<input type="radio"/> around 0.20		
<input checked="" type="radio"/> less than .001	✓ 1.00	
Total	1.00 / 1.00	