Quiz 1 Topics 1-4_

Due Jun 16 at 7pm   **Points** 100   **Questions** 20

Available Jun 2 at 8:33pm - Jun 16 at 7:46pm 14 days   **Time Limit** None

Instructions

The Dataset used:

College Dataset Data Dictionary.pptx 📄

college Dataset.csv 📄

Attempt History

	Attempt	Time	Score
LATEST	**Attempt 1**	38 minutes	95 out of 100

Score for this quiz: **95** out of 100
Submitted Jun 16 at 10:13am
This attempt took 38 minutes.

Question 1	5 / 5 pts

Exploratory data analysis relies on the data to determine how the analysis proceeds.

Correct!

- ⦿ True

- ◯ False

Question 2

5 / 5 pts

Linear regressions and hypothesis testing never done as part of exploratory data analysis.

○ True

Correct!

◉ False

Question 3

5 / 5 pts

Most data scientists get their data through broad Google searches.

○ True

Correct!

◉ False

Question 4

5 / 5 pts

The first step in conducting research is to find data that is interesting.

○ True

Correct!

⊙ False

Question 5   5 / 5 pts

The research statement "This study proposes to conduct research about science programs in secondary schools in the United States" is a bad research statement because:

○ It is too broad.

○ Overall, the statement is not well defined.

Correct!   ⊙ All the answers given.

○ It is probably not realistic in terms of scope.

Question 6   5 / 5 pts

Which question is NOT a question you should use to evaluate your research project?

○ Can I produce an insightful answer?

Correct!   ⊙
          Is it unique, i.e. can I make sure others will not be able to duplicate it?

○ What is the cost?

○ Will I be able to explain the results?

Question 7   5 / 5 pts

Given the following problem statement/research question, which of the
following is <u>NOT</u> a sub-problem of the original statement/question?

*The purpose of this study will be to analyze the motets of Palestrina
written between 1575 and 1580, discover their distinctive contrapuntal
characteristics, and to analyze the same characteristics in the motets of
William Byrd written between 1592 and 1597 (to determine what a
comparative study of these two analyses may reveal).*

○ To analyze the motets of Palestrina written between 1575 and 1580 and
discover their distinctive contrapuntal characteristics.

○ None of the answers given.

Correct!

⦿ To compare the motets of Palestrina written between 1575 and 1580
with those written by today's composers.

○ To consider the motets of William Byrd written between 1592 and 1597
in order to discover their distinctive contrapuntal characteristics.

Question 8

5 / 5 pts

It is necessary to understand why or how data/information is organized in order to obtain data required to conduct research.

Correct!

🔘 True

⚪ False

Question 9

5 / 5 pts

The search for data is very straightforward, just ask a research question and the required data will be readily available.

⚪ True

Correct!

🔘 False

Question 10

5 / 5 pts

Once a research question has been asked there will never by any reason to change it.

⚪ True

Correct!

   ⊙ False

Question 11

5 / 5 pts

We should always have a good idea what our conclusions will be before we ever enter the data processing pipeline.

○ True

Correct!

⊙ False

Question 12

5 / 5 pts

We really want our data to be "tidy," meaning:

○ Each separate record or observation in its own row

○ Each separate variable in its own column

Correct!

⊙ All of the answers given

○ Each type of observational unit in its own table, e.g. department in an organization

Question 13

5 / 5 pts

Raw data has not had any type of processing or pre-processing done on it.

Correct!

- ⦿ True

- ◯ False

Question 14

5 / 5 pts

It isn't necessary to record the steps taken to process data.

- ◯ True

Correct!

- ⦿ False

Question 15

5 / 5 pts

Once you have a data set there isn't any reason to know anything else about the data.

- ◯ True

Correct!

- ⦿ False

Question 16

0 / 5 pts

Which of the following commands allows you to look at the variable types of all of the variables in your dataframe?

○ D: vartypes()

You Answered

⦿ A:str()

○ E: All of the above

Correct Answer

○ C:A and B

○ B:glimbs()

Question 17

5 / 5 pts

Go to the website http://www-bcf.usc.edu/~gareth/ISL/data.html and download the data set College.csv and read it into R as "college," i.e. college <- read.csv("College.csv") to answer the following questions.

Using the R command str() how many observations and variables are there in this data set?

Correct!

⦿ 777 observations of 19 variables

○ 1545 observations of 8 variables

○ 2000 observations of 27 variables

○ 202 observations of 5 variables

Question 18   5 / 5 pts

Go to the website http://www-bcf.usc.edu/~gareth/ISL/data.html and download the data set College.csv and read it into R as "college," i.e. college <- read.csv("College.csv") to answer the following questions.

Using the summary() command, what is the maximum number of applications any college received?

○ 2003

Correct!   ◉ **48094**

○ 3624

○ 1558

Question 19   5 / 5 pts

Go to the website http://www-bcf.usc.edu/~gareth/ISL/data.html and download the data set College.csv and read it into R as "college," i.e.

college <- read.csv("College.csv") to answer the following questions.

Of all the colleges, what was the average number of applications accepted?

○ 1110

○ 604

○ 2424

Correct!   ⊙ 2019

Question 20   **5 / 5 pts**

In order to read data into R, you must download it onto your computer first before reading it in.

○ True

Correct!   ⊙ False

Quiz Score: **95** out of 100