Assessment

Download the CSV file (from the previous lesson) and import the file into Excel. This is the data set that you will use to answer each multiple choice question in this assessment. Each row in the data set represents a building that was damaged by an earthquake in the Himalayas. The amount of damage is given in the variable "damage grade." A full list of variables and their meanings can be found in the previous lesson in the Code Book. We recommend that students download the data set and import into Excel prior to starting the multiple choice assessment. Choose the best answer from the answers provided to the questions.

Good Luck!

Multiple Choice Assessment

Question 1

1.0/1.0 point (graded)

Imagine that this dataset consisted of two data tables that you needed to join in order to do the analysis. One data table had information on the "damage_grade" for each building. The other data table had all the other information on each building. Which join should you perform to ensure you keep all rows from both files?

- Left join, using 'key' of `damage_grade`
- Right join, using 'key' of `building_id`
- Full join, using 'key' of `building_id` ✔
- Full join, using 'key' of `damage_grade`

Submit You have used 1 of 1 attempt

Question 2

1.0/1.0 point (graded)

A damage grade of `2` represents a medium amount of damage. A team seeking to improve the region's safety wants to know how many buildings had this level of damage. What percent of buildings have a grade of `2`?

- 56.32 ✔
- 9.38
- 34.26
- 85.12

Submit You have used 1 of 1 attempt

Question 3

1.0/1.0 point (graded)

The same team is wanting to describe the level of damage sustained in the region. They next need to know what the most common level of damage was. What was the most common damage grade?

- ○ 0
- ○ 1
- ◉ 2 ✔
- ○ 3

Submit

You have used 1 of 1 attempt

Question 4

1.0/1.0 point (graded)

The next objective is to assess the role of building age in safety. The variable “age” is the age of the building, in years. What is the average (mean) building age?

- ◉ 25.39 ✔
- ○ 14.12
- ○ 40.55
- ○ 18.74

Submit You have used 1 of 1 attempt

Question 5

1.0/1.0 point (graded)

The distribution of age is not symmetrical, meaning that the median may be a better indicator of 'average' age. What is the median age?

- (●) 15 ✔
- ( ) 20
- ( ) 25
- ( ) 18

Submit You have used 1 of 1 attempt

Question 6

1.0/1.0 point (graded)

Next, the team is interested in the size of buildings. The variable “area” represents the size of the building. The team is particularly interested in those buildings that are under 100 years of age. Of those buildings, what percent of buildings have an area greater than 50?

- (●) 18.3% ✔
- ( ) 81.7%
- ( ) 40.8%
- ( ) 59.2%

Submit You have used 1 of 1 attempt

Question 7

1.0/1.0 point (graded)

The team wants you to do further work with the 'area' variable, focusing on buildings with area less than 150. Make a histogram of “area” using only these buildings. How should this be described?

- ( ) No skew
- (●) Skew with long tail pointing in the positive direction ✔
- ( ) Skew with long tail pointing in the negative direction
- ( ) Unable to determine with information given

Submit You have used 1 of 1 attempt

Question 8

1.0/1.0 point (graded)

What variable type is `damage_grade`?

- ○ Date
- ◉ Numeric ✔
- ○ Currency
- ○ String

Submit

You have used 1 of 1 attempt

Question 9

0.0/1.0 point (graded)

The team is revisiting the relationship between area and age. For this analysis, the team wants *only* buildings that **have an age less than 150 and an area less than 200**. For these buildings, make a scatter plot of the relationship between age and area. Put age on the x-axis and area on the y-axis. What describes their relationship?

- There is a downward trend such that older buildings are smaller
- There is an upward trend such that older buildings are larger
- There is a curvilinear trend, such that older buildings and newer buildings are smallest ✖
- There is no discernible relationship

Submit You have used 1 of 1 attempt

Question 10

1.0/1.0 point (graded)

Refer to the scatter plot in the previous question. Assume that you are going to next quantify the relationship between age and area. That correlation should be ...

- A negative number between 0 and -1 ✔
- A positive number between 0 and +1
- A negative number between -1 and -2
- Approximately zero

Submit You have used 1 of 1 attempt

Question 11

1.0/1.0 point (graded)

The team next wishes to examine how well the buildings made of stone fared in the disaster. Stone buildings have a score of "1" on the `has_superstructure_stone_flag` variable. The team is worried that these buildings were particularly hard hit (`damage_grade` of 3). How many stone buildings had this level of damage?

- ○ 59%
- ○ 12%
- ◉ 44% ✔
- ○ 34%

Submit

You have used 1 of 1 attempt

Question 12

1.0/1.0 point (graded)

Continuing your work with stone buildings, what was the mean damage of buildings made of stone (possible range: 0-3)?

- (●) 2.42 ✔
- ( ) 1.98
- ( ) 2.24
- ( ) 2.99

Submit You have used 1 of 1 attempt

Question 13

1.0/1.0 point (graded)

The team is next interested in examining non-stone buildings. Non-stone buildings have a score of "0" on the `has_superstructure_stone_flag` variable. What percent of these buildings had a **damage great than 2**?

- ( ) 59%
- ( ) 12%
- ( ) 44%
- (●) 34% ✔

Submit You have used 1 of 1 attempt

Question 14

1.0/1.0 point (graded)

What was the **mean** damage of buildings not made of stone?

- ○ 2.42
- ○ 1.98
- ◉ 2.24 ✔
- ○ 2.99

Submit

You have used 1 of 1 attempt

Question 15

0.0/1.0 point (graded)

The team awaits your final conclusion regarding stone buildings. The best summary statement you could make about stone vs. non-stone buildings on the basis of this data?

- ○ Stone buildings survived the earthquake better
- ○ Stone buildings suffered worse damage
- ◉ There was little apparent difference ✖
- ○ Stone buildings were virtually undamaged

Submit You have used 1 of 1 attempt

Question 16

1.0/1.0 point (graded)

The researchers want to assess the impacts of other building materials. There are several `has_superstructure_` materials variables, each describing a different building material. Ignoring reinforced concrete buildings (the `rc_non_engineered`, `rc_engineered` ones), which of the other types had the lowest average (mean) damage grade?

- ○ Bamboo
- ◉ Brick ✔
- ○ Adobe mud
- ○ Timber

Submit You have used 1 of 1 attempt

Question 17

1.0/1.0 point (graded)

On average, buildings that have a secondary use (`has_secondary_use`) had...

- (●) Less damage ✔
- ( ) Zero damage
- ( ) More damage
- ( ) Essentially equal damage

Submit You have used 1 of 1 attempt

Question 18

0.0/1.0 point (graded)

Building damage may differ for multi-family dwellings; this is very important as the safety of multi-family dwellings can impact many families at once. Compare buildings that have one family in them vs more than one family (`count_families`). You should report to the team that buildings that have only one family had, on average, had...

- ( ) Slightly less damage (a difference within .25 points)
- ( ) Zero damage
- (●) More damage ✖
- ( ) Considerably less damage (a difference greater than .50 points)

Submit You have used 1 of 1 attempt

Question 19

0.0/1.0 point (graded)

Imagine you were tasked to evaluate each building material, as well as the role of building age and area, in earthquake damage. Imagine you were asked to produce a document analyzing those roles, this would be ...

- A dashboard
- A scorecard
- An analytic report
- An analysis ✖

Submit You have used 1 of 1 attempt

Question 20

0.0/1.0 point (graded)

Imagine you created a tool for damage surveyors in the field that would tally the damage and other building stats that updates as new data comes in. It might have several graphs and other statistics that update in real time. This would be:

- ○ Dashboard
- ○ Scorecard
- ◉ Analytic report ✖
- ○ Analysis

Submit You have used 1 of 1 attempt

Review Questions

The next 10 questions relate to the data analysis track as a whole.

Question 21

1.0/1.0 point (graded)

The *most important* goal of a data analyst working in any organization is to ...

- ◉ Find insights that the organization can use ✔
- ○ Find the most statistically significant results
- ○ Find the most nuanced descriptions of the data
- ○ None of the above are important

Submit You have used 1 of 1 attempt

Question 22

1.0/1.0 point (graded)

Which is a goal in healthcare data analytics?

- Optimize business goals
- Optimize health of patients
- Optimize efficiency of organization
- All of these are goals in healthcare analytics ✔

Submit You have used 1 of 1 attempt

Question 23

0.0/1.0 point (graded)

Why are bar graphs better than pie charts at showing distributions of categorical (category) data?

- You can more easily see the counts with bar graphs
- Pie charts are hard to read when you have many small groups
- You can still compare group size with bar graphs
- All of these are reason to prefer bar graphs over pie charts ✖

Submit You have used 1 of 1 attempt

Question 24

1.0/1.0 point (graded)

Which of the following quick calculations would tell you the revenue of a business for a given product?

- Quantity sold x value of goods
- (Value of goods / inventory of goods) - quantity sold
- Quantity of goods x price of goods ✔
- The sum of quantity of goods / price of goods

Submit You have used 1 of 1 attempt

Question 25

1.0/1.0 point (graded)

For analysts to draw insights from data, they first need to take data from a _____ and make a _____ with tools such as SQL.

- (●) Database; table ✔
- ( ) Table; database
- ( ) Table; graph
- ( ) Graph; table

Submit

You have used 1 of 1 attempt

Question 26

1.0/1.0 point (graded)

Rules such as "don't break the law" or "do no harm" are examples of _____ in data analytics ethics.

- ( ) Moral maximums
- (●) Moral minimums ✔
- ( ) Ethical hypotheses
- ( ) Ethical deductions

Submit You have used 1 of 1 attempt

Question 27

1.0/1.0 point (graded)

All of the following are types of graphs except:

- Pie chart
- Bar chart
- Butterfly chart
- Prism chart ✔

Submit You have used 1 of 1 attempt

Question 28

1.0/1.0 point (graded)

When crafting a data analytics story, one should always

- Prove a hypothesis
- Support one's opinion
- None of these
- Know one's audience ✔

Submit You have used 1 of 1 attempt

Question 29

1.0/1.0 point (graded)

Which should we do when improving our slides and other visual materials when crafting analytics stories?

- Have many visual borders
- Increase amount of text
- Eliminate white space
- Make text larger ✔

Submit You have used 1 of 1 attempt

Question 30

1.0/1.0 point (graded)

What does the following sigma notation instruct the data analyst to do?

$$\sum(x+7)$$

- (●) Add seven to every score, then sum the total ✔
- ( ) Add the sum of every score to seven
- ( ) Start with seven, sum every score, and add them together
- ( ) None of these

Submit

You have used 1 of 1 attempt