

Try again once you are ready  
Grade received 75%  
To pass 80% or higher  
Try again

1. Fill in the blank: Missing data has a value that is not stored for a \_\_\_\_\_ in a dataset.

1 / 1 point

- ☐ row
- ☒ variable
- ☐ visualization
- ☐ column



Correct

Missing data has a value that is not stored for a variable in a dataset. It is typically encoded as N/A, NaN, or a blank.

2. A data professional requests additional information from a dataset's original owner. Unfortunately, they are not able to provide the information. Therefore, the data professional creates a NaN category in the dataset. What concept does this scenario describe?

1 / 1 point

- ☐ Mapping variables in a dataset
- ☐ Ensuring two datasets are compatible
- ☐ Managing big data
- ☒ Solving the problem of missing data



Correct

This scenario describes solving the problem of missing data. There are four common ways to do this: Request the missing values from the owner of the data; delete the missing columns, rows, or values; create a NaN category; or derive new representative values.

3. When merging data, a data professional uses the following code:

0 / 1 point

```
df_joined = df.merge(df_zip, how='left',  
on=['date', 'center_point_geom'])
```

What is the function of the parameters `how` and `on` in this code?

- ☐ To tell Python how to find missing values in the rows and columns
- ☐ To tell Python how to place the appropriate values on the top row of the dataset
- ☒ To tell Python which datasets should be merged
- ☐ To tell Python which way to join the data and which column to join from



Incorrect

The parameters `how` and `on` tell Python which way to join the data and which column to join from. `how` tells Python which way to join the data, and `on` tells Python which column to start from.

4. Non-null count is the total number of blank data entries within a data column.

1 / 1 point

- ☐ True
- ☒ False

✓ Correct  
Non-null count is the total number of data entries for a data column that are not blank.