Lesson: 1.5 Series and Central Tendency

In this lesson, students will learn how to create a use a Pandas Series. They will also learn and explore measures of central tendency including the mean, median, and mode.

Objective

Students will be able to:

Create a Series using the Pandas library

Compute the mean, median, and mode of a Series

Decide whether the mean, median, or mode is the best measure of central tendency for a specific dataset

Activities

These are all the activities included in the lesson.

Activity

1.5.1 Series and Central Tendency

1.5.2 Series and Central Tendency

1.5.3 Changing Indices

1.5.4 Series and Indices

1.5.5 Hot Dog Series

1.5.6 People Named Anna

1.5.7 Hot Dog Statistics

1.5.8 Critiquing Conclusions

Solution References

Refer to the solution reference for a more detailed look at exercise solutions.

Solution Reference

1.5.2 Series and Central Tendency

1.5.5 Hot Dog Series

1.5.7 Hot Dog Statistics

1.5.8 Critiquing Conclusions

Problem Guides

Refer to the problem guides for a more in-depth look at this lesson's problems.

Problem Guide

1.5.5 Hot Dog Series

1.5.7 Hot Dog Statistics

1.5.8 Critiquing Conclusions

Vocabulary

These are the key terms for this lesson.

Term Definition

series

A one-dimensional, labeled array (or list) that is formatted like a single column of a data table.

Handouts

Use handouts to supplement your class. Please note that there are handouts for teachers and for students.

Debugging Strategies

Planning Notes

The syntax used in this lesson can be difficult. Encourage students to use the DOCS tabs and be sure to always match an opening bracket/quote/parenthesis with a closing one. There is an optional debugging activity listed in the handouts for this lesson.

Teaching and Learning Strategies

Lesson Opener:

Have students brainstorm and write down answers to the discussion questions listed below. Students can work individually or in groups/pairs. Have them share their responses. [5 mins]

Activities:

Watch the lesson video and complete the corresponding quiz to check for understanding. [10 mins]

Explore the Changing Indices example. [10 mins]

Run the code and match the output to the code that printed it.

Add code to change the Pumpkin element to your favorite kind of pie.

Discuss the difference between pies[2] and new\_pies["C"]. Students should recognize that the first is an integer and the second is a string.

Explore the Series and Indices example. [10 mins]

Run the code and match the output to the code that printed it.

Notice how the value for the Colorado Rockies changes when it is printed.

Have students infer what NaN might stand for.

Complete the Hot Dog Series exercise. [10 mins]

Explore the People Named Anna example. [10 mins]

Complete the Hot Dog Statistics exercise. [15 min]

Students can be paired up or put into small groups to discuss the questions in this activity.

Complete the Critiquing Conclusions free-response activity. [15 mins]

Lesson Closer:

Have students reflect and discuss their responses to the end-of-class discussion questions. [5 mins]

Prior Knowledge

using Python variables, lists, and functions

importing Python libraries

finding the mean, median, and mode of a set of numbers

Video Slides

Discussion Questions

Beginning of Class:

Find the median and mode of the following data set: 1, 3, 6, 8, 9, 12, 15, 15, 67

The median is 9 and the mode is 15.

Find the mean (average) of the dataset above.

The mean is 15.11.

Which is most likely an outlier in the dataset above? Why?

The number 67 is most likely the outlier because it is further away from the other values.

End of Class:

When might it be a good decision to use the median as the best measure of central tendency?

The median is a good measure to use when there are outliers that negatively affect the mean as the best measure of central tendency.

Why does the mode print differently than the mean and median when using mode()?

There can be more than one mode, so it prints as a series or a list of all of the modes.

Write code that would result in the following series being displayed.

Answer:

sample = pd.Series(["Turtle", "Hare"], index=["1st", "2nd"])

print(sample)

Modification: Advanced

Have students explore the dtype and name parameters when creating a Series.

pd.Series(data=None, index=None, dtype=None, name=None)

Have students explore the different attributes that can be used with a Series.

Series Documentation

Modification: Special Education

Provide students with a template or skeleton for the syntax used in creating a series.

Have students keep a log or chart of all commands learned, including when and how to use them.

Modification: English Language Learners

Provide students with a template or skeleton for the syntax used in creating a series.

Have students keep a log or chart of all commands learned, including when and how to use them.