Mean vs. Median Using .agg()
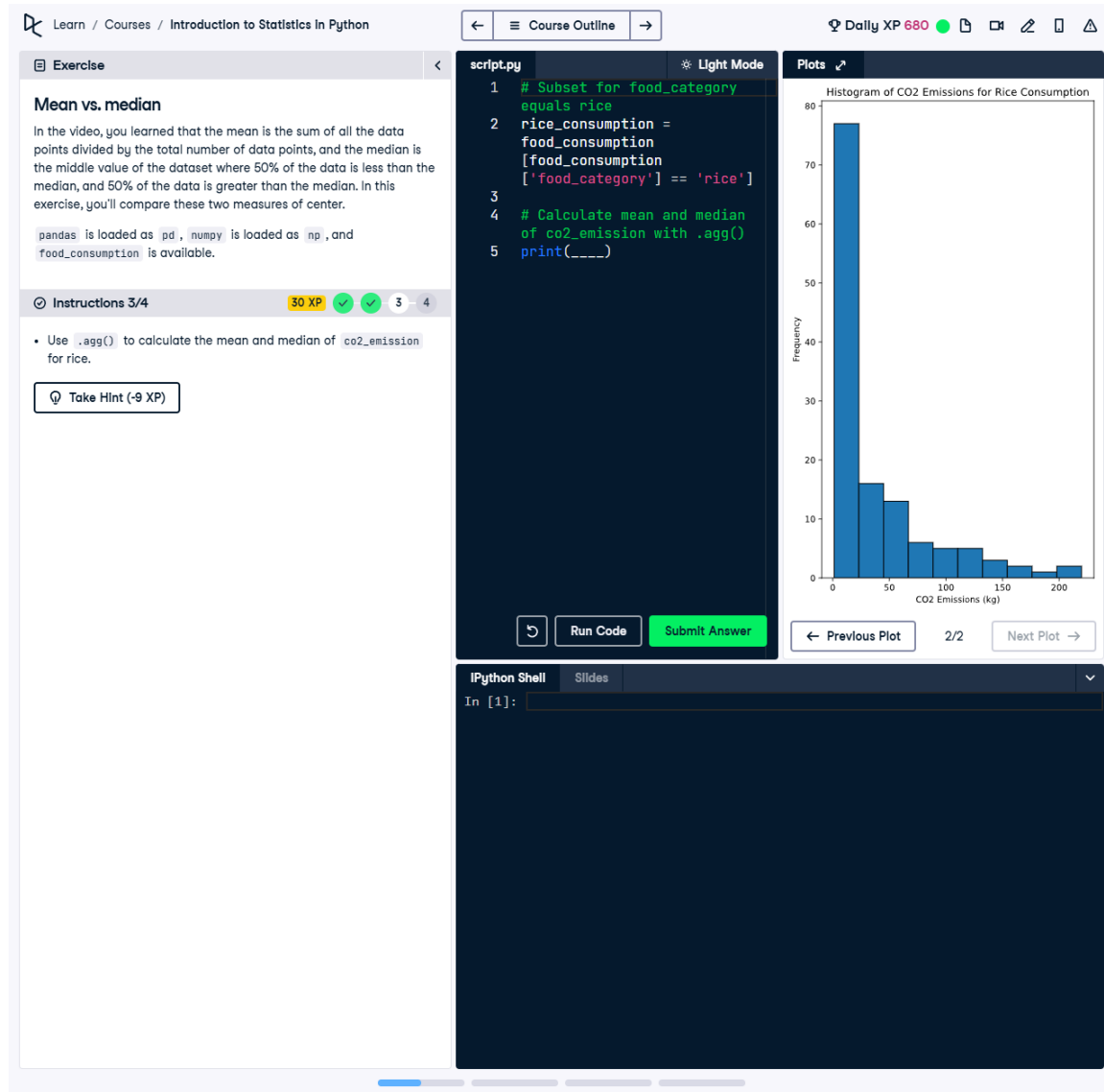
Exercise

Mean vs. median

In the video, you learned that the mean is the sum of all the data points divided by the total number of data points, and the median is the middle value of the dataset where 50% of the data is less than the median, and 50% of the data is greater than the median. In this exercise, you'll compare these two measures of center.

`pandas` is loaded as `pd`, `numpy` is loaded as `np`, and `food_consumption` is available.

Instructions 3/4   30 XP ✓ ✓ 3 4

- Use `.agg()` to calculate the mean and median of `co2_emission` for rice.

💡 Take Hint (-9 XP)

script.py   ☀ Light Mode

```python
1  # Subset for food_category
   equals rice
2  rice_consumption =
   food_consumption
   [food_consumption
   ['food_category'] == 'rice']
3
4  # Calculate mean and median
   of co2_emission with .agg()
5  print(____)
```

↺   Run Code   Submit Answer

Plots ↗

Histogram of CO2 Emissions for Rice Consumption

← Previous Plot   2/2   Next Plot →

IPython Shell   Slides

```
In [1]:
```

Figure 1: Screenshot showing the histogram and task to calculate mean and median using .agg().

Question

In this exercise, you will use the .agg() method to calculate the mean and median of CO2 emissions for rice consumption.
The task involves the following steps:

Instructions:
1. Subset the food_consumption DataFrame to get the rows where the

food_category is 'rice'.
2. Use the .agg() method to calculate the mean and median of the co2_emission column for the rice_consumption DataFrame.
3. Print the calculated values.

Corrected Code Solution

```
# Subset for food_category equals rice
rice_consumption = food_consumption[food_consumption['food_category'] == 'rice']

# Calculate mean and median of co2_emission with .agg()
print(rice_consumption['co2_emission'].agg(['mean', 'median']))
```

Answer Explanation

1. **Subset for rice consumption:** The rice_consumption DataFrame is created by filtering rows where the 'food_category' column equals 'rice'.
2. **Use .agg() method:** The .agg() method is applied to the 'co2_emission' column to calculate both the mean and the median in a single step.
   - 'mean': Computes the average CO2 emission for rice consumption.
   - 'median': Finds the middle value of the CO2 emission data, providing a measure of central tendency.
3. **Print the results:** The calculated mean and median are printed to provide insights into the dataset's central tendencies.