Mean vs. Median: Best Measure of Central Tendency Learn / Courses / Introduction to Statistics in Python ← ≡ Course Outline → Ф Daily XP 710 ● 🕒 🖂 🗘 🗓 🛕 ■ Exercise Light Mode script.py Histogram of CO2 Emissions for Rice Consumption 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. Question Given the skew of this data, what measure of central tendency best summarizes the kilograms of ${\rm CO}_2$ emissions per person per year for rice? Mean O Median Both mean and median Submit Answer @ Take Hint (-6 XP) り Run Code ← Previous Plot Next Plot \rightarrow IPython Shell mean 37.592 median 15.200 Name: co2_emission, dtype: float64

Figure 1: Screenshot showing the histogram and task to identify the best measure of central tendency.

Question

Given the skew of this data, what measure of central tendency best summarizes the kilograms of CO2 emissions per person per year for rice?

- 1. Mean
- 2. Median

^{**}Possible Answers:**

3. Both mean and median

Correct Answer

Median

Explanation

In right-skewed data, such as the CO2 emissions for rice consumption, the mean is heavily influenced by extreme values (outliers) on the higher end of the distribution. As a result, the mean does not accurately represent the center of the data.

The median, being the middle value when the data is ordered, is less affected by these outliers. It provides a better representation of the central tendency for skewed distributions, making it the preferred measure in this case.