# Finding Outliers Using IQR - Step 3/4

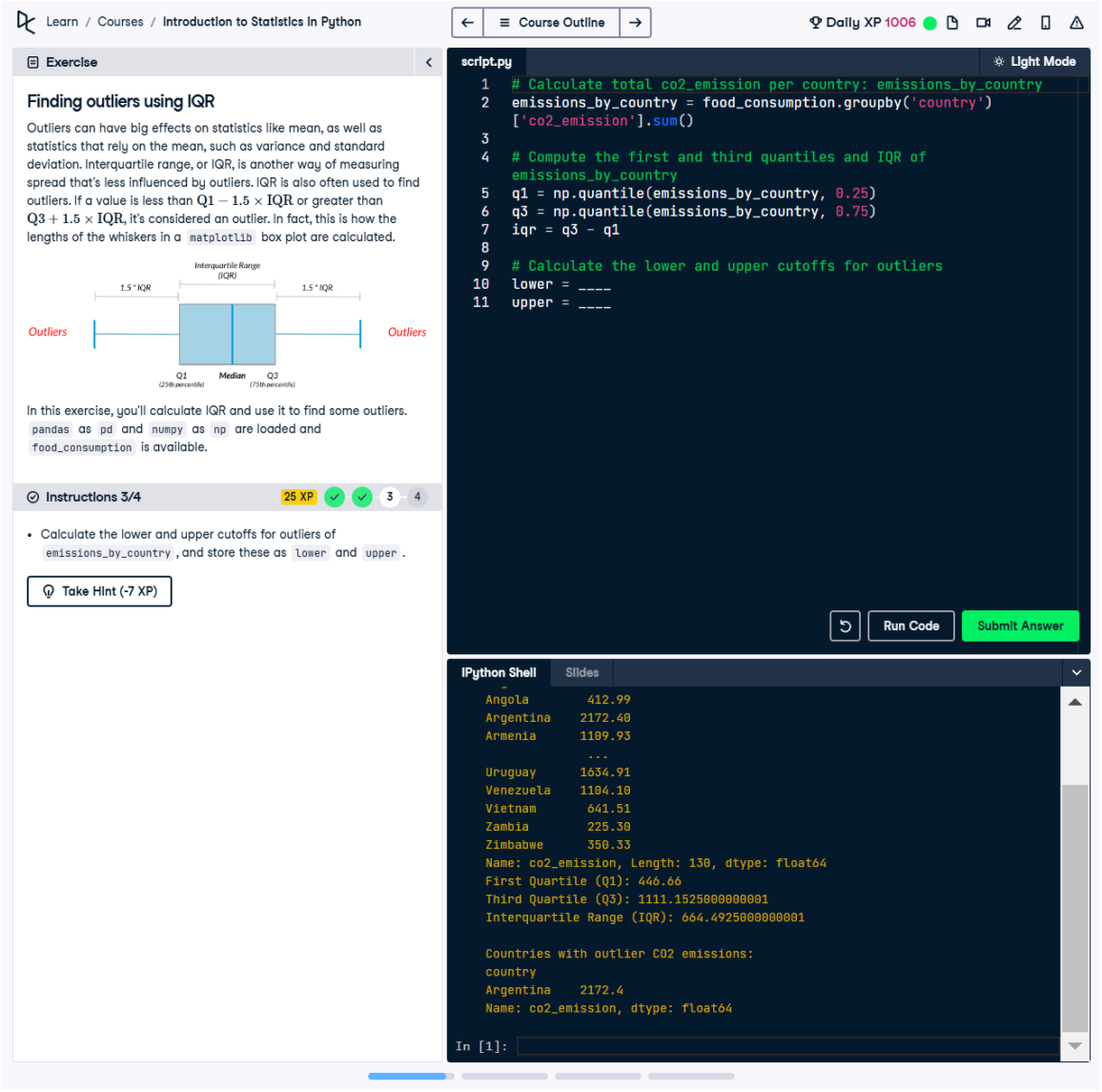


Figure 3: Screenshot showing the calculation of lower and upper cutoffs for outliers using IQR.

## Question

Outliers can have big effects on statistical measures such as mean, variance, and standard deviation. Interquartile range (IQR) is another way of measuring spread that is less influenced by outliers. In this step, you'll calculate the lower and upper cutoffs for outliers of `emissions_by_country` and store these as `lower` and `upper`.

**Instructions:**
1. Calculate the lower cutoff for outliers using Q1 - 1.5 * IQR.
2. Calculate the upper cutoff for outliers using Q3 + 1.5 * IQR.
3. Print the lower and upper cutoffs for identifying outliers.

## Corrected Code Solution

```
import pandas as pd
import numpy as np

# Calculate total CO2 emissions per country
emissions_by_country = food_consumption.groupby('country')
['co2_emission'].sum()

# Compute Q1 and Q3
q1 = np.quantile(emissions_by_country, 0.25)
q3 = np.quantile(emissions_by_country, 0.75)
iqr = q3 - q1

# Calculate lower and upper cutoffs for outliers
lower = q1 - 1.5 * iqr
upper = q3 + 1.5 * iqr

# Print the results
print(f"First Quartile (Q1): {q1}")
print(f"Third Quartile (Q3): {q3}")
print(f"Interquartile Range (IQR): {iqr}")
print(f"Lower cutoff for outliers: {lower}")
print(f"Upper cutoff for outliers: {upper}")
```

### Answer Explanation

1. **Lower and Upper Cutoffs:** The cutoffs are calculated as:
   - Lower = Q1 - 1.5 * IQR.
   - Upper = Q3 + 1.5 * IQR.
   These thresholds help identify values that deviate significantly from the bulk of the data.

2. **Results:** The lower and upper cutoffs are printed, providing a numerical range for identifying outliers in the `emissions_by_country` dataset.

3. **Output Analysis:** Any country with total CO2 emissions below the lower cutoff or above the upper cutoff is considered an outlier. These cutoffs assist in isolating extreme values for further investigation.