

Inspecting a DataFrame - Descriptive Statistics

This document includes the question, the solution, and a breakdown of the code provided in the screenshot.

Uploaded Screenshot

Below is the screenshot of the task:

The screenshot shows a web browser window with the URL campus.datacamp.com/courses/data-manipulation-with-pandas/transforming-dataframes?ex=2. The page is titled "Inspecting a DataFrame" and is part of a course on "Data Manipulation with pandas". The instructions for the exercise are as follows:

- When you get a new DataFrame to work with, the first thing you need to do is explore it and see what it contains. There are several useful methods and attributes for this.
- `.head()` returns the first few rows (the "head" of the DataFrame).
- `.info()` shows information on each of the columns, such as the data type and number of missing values.
- `.shape` returns the number of rows and columns of the DataFrame.
- `.describe()` calculates a few summary statistics for each column.

The `homelessness` is a DataFrame containing estimates of homelessness in each U.S. state in 2018. The `individual` column is the number of homeless individuals not part of a family with children. The `family_members` column is the number of homeless individuals part of a family with children. The `state_pop` column is the state's total population.

The instructions for the exercise are:

- Print some summary statistics that describe the `homelessness` DataFrame.

The code editor shows the following code:

```
1 # Print the head of the homelessness data
2 print(homelessness.head())
3
4 # Print information about homelessness
5 print(homelessness.info())
6
7 # Print the shape of homelessness
8 print(homelessness.shape)
9
10 # Print a description of homelessness
11 print(_____)
```

The IPython Shell output shows the result of the code execution:

```
Data columns (total 5 columns):
# Column      Non-Null Count  Dtype
---  ---
0 region      51 non-null    object
1 state       51 non-null    object
2 individuals  51 non-null    float64
3 family_members  51 non-null    float64
4 state_pop   51 non-null    int64
dtypes: float64(2), int64(1), object(2)
memory usage: 2.4+ KB
None
(51, 5)
```

Question

Print some summary statistics that describe the homelessness DataFrame.

Answer

```
# Print the head of the homelessness data
print(homelessness.head())
```

```
# Print information about homelessness
print(homelessness.info())
```

```
# Print the shape of homelessness
print(homelessness.shape)
```

```
# Print a description of homelessness
print(homelessness.describe())
```

Code Explanation

Explanation of the code:

1. ``print(homelessness.head())``: Displays the first five rows of the DataFrame for a quick preview.
2. ``print(homelessness.info())``: Provides detailed information about the columns, including data types and non-null counts.
3. ``print(homelessness.shape)``: Outputs a tuple ``(number_of_rows, number_of_columns)`` representing the DataFrame dimensions.
4. ``print(homelessness.describe())``: Computes and prints summary statistics for the numeric columns in the DataFrame, such as mean, standard deviation, min, max, and quartiles.