

Sorting Rows by Region and Family Members

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 DataCamp exercise titled "Sorting rows" for the course "Data Manipulation with pandas". The exercise instructions are as follows:

Sorting rows

Finding interesting bits of data in a DataFrame is often easier if you change the order of the rows. You can sort the rows by passing a column name to `.sort_values()`.

In cases where rows have the same value (this is common if you sort on a categorical variable), you may wish to break the ties by sorting on another column. You can sort on multiple columns in this way by passing a list of column names.

Sort on ... Syntax

one column	Syntax
	<code>df.sort_values("breed")</code>
multiple columns	<code>df.sort_values(["breed", "weight_kg"])</code>

By combining `.sort_values()` with `.head()`, you can answer questions in the form, "What are the top cases where...?".

homelessness is available and pandas is loaded as `pd`.

Instructions 3/3 30 XP

- Sort `homelessness` by the number of homeless individuals in the `individuals` column, from smallest to largest, and save this as `homelessness_ind`. Print the head of the sorted DataFrame.
- Sort `homelessness` by the number of homeless `family_members` in descending order, and save this as `homelessness_fam`.
- Sort `homelessness` first by region (ascending), and then by number of family members (descending). Save this as `homelessness_reg_fam`.

Take Hint (-9 XP)

The code editor shows the following Python code:

```
1 # Sort homelessness by region, then descending family members
2 homelessness_reg_fam = ____
3
4 # Print the top few rows
5 print(homelessness_reg_fam.head())
```

The output of the code is displayed in the IPython Shell:

```
<script.py> output:
  region      state  individuals  family_members  state_pop
32  Mid-Atlantic   New York    39827.0         52878.0   19530351
4    Pacific    California   109008.0         28964.0    39461588
21  New England  Massachusetts    6811.0         13257.0    6882635
9    South Atlantic   Florida   21443.0         9587.0   21244317
43 West South Central    Texas   19199.0         6111.0   28628666
```

Question

Sort the homelessness DataFrame first by region (in ascending order), then by the number of family members (in descending order). Save this as `homelessness_reg_fam`. Print the head of the sorted DataFrame.

Answer

```
# Sort homelessness by region, then descending family members
homelessness_reg_fam = homelessness.sort_values(
    by=["region", "family_members"], ascending=[True, False]
)
```

```
# Print the top few rows
print(homelessness_reg_fam.head())
```

Code Explanation

Explanation of the code:

1. ``homelessness.sort_values(by=["region", "family_members"], ascending=[True, False])``: Sorts the ``homelessness`` DataFrame by two criteria:
 - ``region``: Sorted in ascending order.
 - ``family_members``: Sorted in descending order (indicated by ``False``).
2. ``homelessness_reg_fam``: Stores the sorted DataFrame for further use.
3. ``print(homelessness_reg_fam.head())``: Displays the first five rows of the sorted DataFrame to verify the sorting.