

Subsetting Columns - Individuals and State

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 a DataCamp course page. The page title is 'Subsetting columns'. The instructions on the left side of the page are as follows:

Subsetting columns

When working with data, you may not need all of the variables in your dataset. Square brackets (`[]`) can be used to select only the columns that matter to you in an order that makes sense to you. To select only `"col_a"` of the DataFrame `df`, use

```
df["col_a"]
```

To select `"col_a"` and `"col_b"` of `df`, use

```
df[["col_a", "col_b"]]
```

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

Instructions 3/3 30 XP

- ✓ Create a Series called `individuals` that contains only the `individuals` column of `homelessness`.
- ✓ Create a DataFrame called `state_fam` that contains only the `state` and `family_members` columns of `homelessness`, in that order.
- ❓ Create a DataFrame called `ind_state` that contains the `individuals` and `state` columns of `homelessness`, in that order.

[Take Hint \(-9 XP\)](#)

The right side of the screenshot shows a code editor with the following code:

```
1 # Select only the individuals and state columns, in that order
2 ind_state = ____
3
4 print(ind_state.head())
```

Below the code editor is a 'Python Shell' window showing the output of the code:

```
<script.py> output:
state family_members
0 Alabama 864.0
1 Alaska 582.0
2 Arizona 2686.0
3 Arkansas 432.0
4 California 28964.0
```

The bottom of the screenshot shows a Windows taskbar with the date and time 17:05 24/11/2024.

Question

Create a DataFrame called `ind_state` that contains the `individuals` and `state` columns of `homelessness`, in that order. Print the first few rows of the DataFrame.

Answer

```
# Select only the individuals and state columns, in that order
ind_state = homelessness[["individuals", "state"]]
```

```
# Print the first few rows
print(ind_state.head())
```

Code Explanation

Explanation of the code:

1. ``homelessness[["individuals", "state"]]`: Selects the ``individuals`` and ``state`` columns from the ``homelessness`` DataFrame.
 - The double square brackets ``[[]]` are used to select multiple columns.
2. ``ind_state``: Stores the resulting DataFrame with the selected columns.
3. ``print(ind_state.head())``: Displays the first five rows of the ``ind_state`` DataFrame to check its contents.