

Subsetting Rows - Individuals Greater Than 10,000

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 titled 'Data Manipulation with pandas'. The exercise is 'Subsetting rows'. The instructions are as follows:

- 1. Filter `homelessness` for cases where the number of `individuals` is greater than ten thousand, assigning to `ind_gt_10k`. View the printed result.
- 2. Filter `homelessness` for cases where the USA Census `region` is "Mountain", assigning to `mountain_reg`. View the printed result.
- 3. Filter `homelessness` for cases where the number of `family_members` is less than one thousand and the `region` is "Pacific", assigning to `fam_lt_1k_pac`. View the printed result.

The code editor on the right shows the following code:

```
1 # Filter for rows where individuals is greater than 10000
2 ind_gt_10k = 
3 
4 # See the result
5 print(ind_gt_10k)
```

The interface also includes a 'Run Code' button and a 'Submit Answer' button.

Question

Filter `homelessness` for cases where the number of `individuals` is greater than ten thousand, assigning the result to `ind_gt_10k`. View the printed result.

Answer

```
# Filter for rows where individuals is greater than 10000
ind_gt_10k = homelessness[homelessness["individuals"] > 10000]
```

```
# See the result
print(ind_gt_10k)
```

Code Explanation

Explanation of the code:

1. ``homelessness["individuals"] > 10000``: This creates a boolean mask, where each row evaluates to ``True`` if the value in the ``individuals`` column is greater than 10,000, and ``False`` otherwise.
2. ``homelessness[homelessness["individuals"] > 10000]``: Filters the ``homelessness`` DataFrame by keeping only rows where the mask is ``True``.
3. ``ind_gt_10k``: Stores the filtered DataFrame for further use.
4. ``print(ind_gt_10k)``: Prints the resulting DataFrame to verify the filter.