

Subsetting rows with .query() - Step 3

Subsetting rows with .query()

In this exercise, you will merge GDP and population data for Australia and Sweden from the World Bank and expand on it using the `.query()` method. You'll merge the two tables and compute the GDP per capita. Afterwards, you'll use the `.query()` method to subselect the rows and create a plot. Recall that you will need to merge on multiple columns in the proper order.

The tables `gdp` and `pop` have been loaded for you.

Instructions 3/4

• Plot `gdp_pop` on `values="gdp_per_capita"`, `index="date"`, and `columns="country"`. Save as `gdp_pivot`.

Take hint (7 XP)

```
1 # Merge gdp and pop on date and country with fill
2 gdp_pop = pd.merge_ordered(gdp, pop, on=['country', 'date'], fill_method='ffill')
3
4 # Add a column named gdp_per_capita to gdp_pop that divides the gdp by pop
5 gdp_pop['gdp_per_capita'] = gdp_pop['gdp'] / gdp_pop['pop']
6
7 # Pivot table of gdp_per_capita, where index is date and columns is country
8 gdp_pivot = gdp_pop.pivot_table('gdp_per_capita', 'date', 'country')
```

	date	Australia	Sweden	
11	1992-01-01	167660.221	8108633	0.009
12	1993-01-01	168087.499	8108633	0.009
13	1994-01-01	168627.899	8108633	0.009
14	1995-01-01	169332.179	8108633	0.009
15	1996-01-01	169720.269	8108633	0.009
16	1996-01-01	79637.636	8108633	0.009
17	1996-01-01	80562.286	8108633	0.009
18	1996-01-01	79975.546	8108633	0.009
19	1996-01-01	80186.497	8108633	0.009
20	1997-01-01	79261.542	8117375	0.009
21	1997-01-01	79673.839	8117375	0.009
22	1997-01-01	79686.778	8117375	0.009
23	1997-01-01	79968.686	8117375	0.009
24	1997-01-01	79390.922	8108633	0.009
25	1997-01-01	79686.778	8108633	0.009
26	1997-01-01	79686.485	8108633	0.009
27	1997-01-01	79686.485	8108633	0.009
28	1997-01-01	79686.485	8108633	0.009
29	1997-01-01	79686.485	8108633	0.009
30	1997-01-01	79686.485	8108633	0.009
31	1997-01-01	79686.485	8108633	0.009

Question:

Pivot `gdp_pop` to values='gdp_per_capita', index='date', and columns='country'. Save as `gdp_pivot`.

Answer:

```
# Merge gdp and pop on date and country with fill
gdp_pop = pd.merge_ordered(
    gdp, pop,
    on=['country', 'date'],
    fill_method='ffill'
)
```

```
# Add a column named gdp_per_capita to gdp_pop that divides the gdp by
pop
gdp_pop['gdp_per_capita'] = gdp_pop['gdp'] / gdp_pop['pop']
```

```
# Pivot table of gdp_per_capita, where index is date and columns is country
gdp_pivot = gdp_pop.pivot_table(
    values='gdp_per_capita',
    index='date',
    columns='country'
)
```

```
# Print the resulting pivot table  
print(gdp_pivot)
```