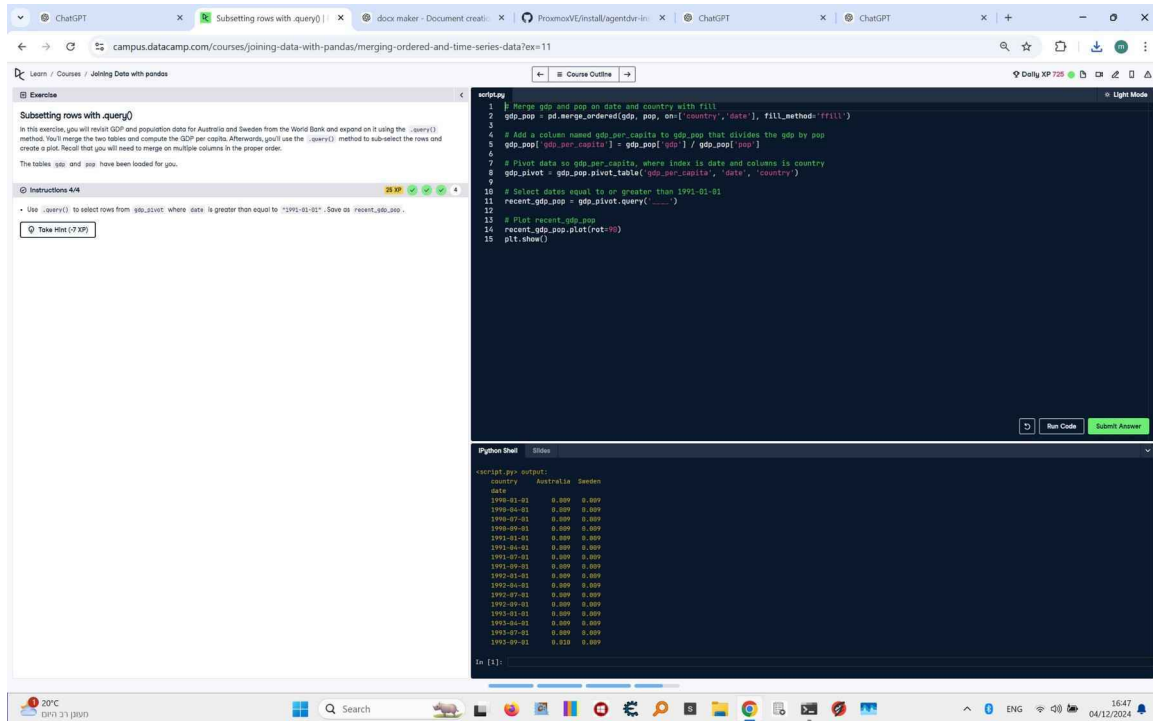


Subsetting rows with .query() - Step 4



The screenshot shows a web browser with a datacamp exercise page and a Jupyter Notebook. The exercise is titled "Subsetting rows with .query()". The Jupyter Notebook contains the following code:

```
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 data so gdp_per_capita, where index is date and columns is country
8 gdp_pivot = gdp_pop.pivot_table('gdp_per_capita', 'date', 'country')
9
10 # Select dates equal to or greater than 1991-01-01
11 recent_gdp_pop = gdp_pivot.query('date >= 1991-01-01')
12
13 # Plot recent_gdp_pop
14 recent_gdp_pop.plot(rot=90)
15 plt.show()
```

The output of the code is a table showing GDP per capita for Australia and Sweden from 1960 to 1993.

| date | Australia | Sweden |
|------------|-----------|--------|
| 1960-01-01 | 0.009 | 0.009 |
| 1961-01-01 | 0.009 | 0.009 |
| 1962-01-01 | 0.009 | 0.009 |
| 1963-01-01 | 0.009 | 0.009 |
| 1964-01-01 | 0.009 | 0.009 |
| 1965-01-01 | 0.009 | 0.009 |
| 1966-01-01 | 0.009 | 0.009 |
| 1967-01-01 | 0.009 | 0.009 |
| 1968-01-01 | 0.009 | 0.009 |
| 1969-01-01 | 0.009 | 0.009 |
| 1970-01-01 | 0.009 | 0.009 |
| 1971-01-01 | 0.009 | 0.009 |
| 1972-01-01 | 0.009 | 0.009 |
| 1973-01-01 | 0.009 | 0.009 |
| 1974-01-01 | 0.009 | 0.009 |
| 1975-01-01 | 0.009 | 0.009 |
| 1976-01-01 | 0.009 | 0.009 |
| 1977-01-01 | 0.009 | 0.009 |
| 1978-01-01 | 0.009 | 0.009 |
| 1979-01-01 | 0.009 | 0.009 |
| 1980-01-01 | 0.009 | 0.009 |
| 1981-01-01 | 0.009 | 0.009 |
| 1982-01-01 | 0.009 | 0.009 |
| 1983-01-01 | 0.009 | 0.009 |
| 1984-01-01 | 0.009 | 0.009 |
| 1985-01-01 | 0.009 | 0.009 |
| 1986-01-01 | 0.009 | 0.009 |
| 1987-01-01 | 0.009 | 0.009 |
| 1988-01-01 | 0.009 | 0.009 |
| 1989-01-01 | 0.009 | 0.009 |
| 1990-01-01 | 0.009 | 0.009 |
| 1991-01-01 | 0.009 | 0.009 |
| 1992-01-01 | 0.009 | 0.009 |
| 1993-01-01 | 0.009 | 0.009 |

Question:

Use .query() to select rows from gdp_pivot where date is greater than or equal to '1991-01-01'. Save as recent_gdp_pop and plot.

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'
)
```

```
# Use .query() to select rows where date >= '1991-01-01'  
recent_gdp_pop = gdp_pivot.query('date >= "1991-01-01"')
```

```
# Plot recent_gdp_pop  
recent_gdp_pop.plot(  
    rot=90  
)  
plt.show()
```