

Multiple Grouped Summaries

The screenshot shows a web browser window with multiple tabs. The active tab is 'campus.datacamp.com/courses/data-manipulation-with-pandas/aggregating-dataframes?ex=12'. The page content includes a header for 'Learn / Courses / Data Manipulation with pandas', a 'Course Outline' button, and a 'Daily XP 100' indicator. The main section is titled 'Exercise' and 'Multiple grouped summaries'. It contains a text block explaining the .agg() method and NumPy functions. Below this is an 'Instructions' section with a list of tasks and a 'Take Hint (-30 XP)' button. To the right is a code editor with a Python script. At the bottom is an IPython shell showing a message about session inactivity.

You are using an unsupported command-line flag: --unsafely-treat-insecure-origin-as-secure=http://54.173.176.93:4444. Stability and security will suffer.

Learn / Courses / Data Manipulation with pandas

Exercise

Multiple grouped summaries

Earlier in this chapter, you saw that the `.agg()` method is useful to compute multiple statistics on multiple variables. It also works with grouped data. NumPy, which is imported as `np`, has many different summary statistics functions, including: `np.min`, `np.max`, `np.mean`, and `np.median`.

`sales` is available and `pandas` is imported as `pd`.

Instructions 100 XP

- Import `numpy` with the alias `np`.
- Get the min, max, mean, and median of `weekly_sales` for each store type using `.groupby()` and `.agg()`. Store this as `sales_stats`. Make sure to use `numpy` functions!
- Get the min, max, mean, and median of `unemployment` and `fuel_price_usd_per_l` for each store type. Store this as `unemp_fuel_stats`.

Take Hint (-30 XP)

```
1 # Import numpy with the alias np
2 ----
3
4 # For each store type, aggregate weekly_sales: get min, max, mean, and median
5 sales_stats = ____
6
7 # Print sales_stats
8 print(sales_stats)
9
10 # For each store type, aggregate unemployment and fuel_price_usd_per_l: get min, max, mean,
11 and median
12 unemp_fuel_stats = ____
13
14 # Print unemp_fuel_stats
15 print(unemp_fuel_stats)
```

Run Code Submit Answer

IPython Shell Slides

We closed your session due to inactivity.
If the problem persists, please report an issue.

Restart Session

Earlier in this chapter, you saw that the `.agg()` method is useful to compute multiple statistics on multiple variables. It also works with grouped data. NumPy, which is imported as `np`, has many different summary statistics functions, including: `np.min`, `np.max`, `np.mean`, and `np.median`.

`sales` is available and `pandas` is imported as `pd`.

Final Answer

```
# Import numpy with the alias np
import numpy as np
```

```
# For each store type, aggregate weekly_sales: get min, max, mean, and
median
sales_stats = sales.groupby("type")["weekly_sales"].agg([np.min, np.max,
np.mean, np.median])
```

```
# Print sales_stats
print(sales_stats)
```

```
# For each store type, aggregate unemployment and fuel_price_usd_per_l:
get min, max, mean, and median
```

```
unemp_fuel_stats = sales.groupby("type")[["unemployment",  
"fuel_price_usd_per_l"]].agg([np.min, np.max, np.mean, np.median])  
  
# Print unemp_fuel_stats  
print(unemp_fuel_stats)
```