

Week 2 — pandas Foundations (Series, DataFrames, Indexing)

Goal: Build comfort with pandas fundamentals: loading data, inspecting dtypes, selecting rows/columns, filtering, assigning new columns, and basic aggregations.

Dataset Choices

Option A: NYC Airbnb (CSV) —

<https://raw.githubusercontent.com/justmarkham/pandas-videos/master/data/airbnb.csv>

Option B: Netflix Titles (CSV) —

https://raw.githubusercontent.com/vega/vega-datasets/master/data/netflix_titles.csv

Task 1 — Load & Inspect

- 1 Load the CSV into a DataFrame.
- 2 Show shape, columns, dtypes, and the first 5 rows.
- 3 Identify obvious missing values.

Task 2 — Column Selection & Renaming

- 1 Select 3–5 relevant columns.
- 2 Rename to snake_case if needed.
- 3 Briefly explain why these columns matter.

Task 3 — Row Filtering

- 1 Create at least 3 different filters (equality, contains, numeric, date).
- 2 Report the counts that match each filter.
- 3 Keep output tidy (use head() as needed).

Task 4 — Derived Columns (Assign)

- 1 Create 1–2 new columns derived from existing ones.
- 2 Explain the logic in 1–2 sentences.

Task 5 — Basic Aggregations

- 1 Compute simple aggregates (count, mean/median, min/max).
- 2 If few numeric columns, show category frequencies instead.
- 3 Display results in a small, readable table.

Task 6 — Groupby WarmUp

- 1 Group by 1 categorical column and compute 1–2 aggregates.
- 2 Sort results; show top/bottom groups.
- 3 Add one sentence interpreting the result.

(Optional) Task 7 — Tiny Visualization

- 1 Create a single bar/line plot from your groupby output.
- 2 Include labels and a title.

Deliverables (end of Week 2)

- 1 Notebook: Week_2_Exercises.ipynb (with narrative Markdown + tidy outputs).
- 2 README.md in week_02/ with dataset choice, how to run, and 3–5 takeaways.
- 3 Push to your GitHub repo under week_02/.