# Data Workflows Cheatsheet (Python 3.13 · VS Code)

Chris N John · Cleveland, OH

January 12, 2026

Refresher/Intermediate

## 1) File I/O for Data

Read/write CSV, Excel, JSON using pandas and pathlib.

```
import pandas as pd
from pathlib import Path

# Read CSV
df = pd.read_csv('data.csv')  # read_csv: load tabular data

# Write Excel
df.to_excel('output.xlsx', index=False)  # to_excel: export DataFrame to Excel
```

## 2) pandas Essentials

Create DataFrames, select/filter rows/columns.

```
data = {'name': ['Alice', 'Bob'], 'score': [10, 20]}
df = pd.DataFrame(data)  # DataFrame: tabular structure
print(df[['name']])  # select column
print(df[df['score'] > 15])  # filter rows
```

## 3) Data Cleaning Patterns

Normalize strings, parse dates, handle missing values.

```
df['name'] = df['name'].str.strip().str.casefold()  # normalize
df['date'] = pd.to_datetime(df['date'], errors='coerce')  # parse dates
df.fillna(0, inplace=True)  # fill missing values
```

## 4) Joining & Merging

Combine datasets using merge and concat.

```
merged = pd.merge(df1, df2, on='id', how='inner')  # merge: join on key
concat_df = pd.concat([df1, df2])  # concat: stack vertically
```

## 5) Quick Visualization

Basic charts using pandas.plot().

```
df['score'].plot(kind='bar', title='Scores')  # plot: quick visualization
```

## Mini-Project Script

```
# Mini-Project: Load CSV -> Clean -> Aggregate -> Export Excel -> Plot
import pandas as pd
df = pd.read_csv('scores.csv')
df['user'] = df['user'].str.strip().str.casefold()
df['score'] = pd.to_numeric(df['score'], errors='coerce').fillna(0).astype(int)
agg = df.groupby('user', as_index=False)['score'].sum()
agg.to_excel('report.xlsx', index=False)
agg.plot(kind='bar', x='user', y='score', title='User Scores')
```

### *VS Code Launch Configuration (launch.json)*

```
{
  "version": "0.2.0",
  "configurations": [
    {
      "name": "Data Workflow Mini-Project",
      "type": "python",
      "request": "launch",
      "program": "${workspaceFolder}/data_workflow.py",
      "console": "integratedTerminal"
    }
  ]
}
```