Pandas Data Input Methods

Date: 2025-10-09

Pandas provides multiple methods to input (read) data from various sources. Below are the commonly used input methods:

- **1. read_csv()** - Reads data from a CSV (Comma-Separated Values) file.
- **2. read_excel()** - Reads data from Excel files (.xls, .xlsx). Requires openpyxl or xlrd library.
- **3. read_json()** - Reads data from a JSON file or JSON string.
- **4. read_html()** - Parses tables from HTML pages into DataFrames.
- **5. read_sql()** - Reads data directly from an SQL database query.
- **6. read_sql_table()** - Reads an entire SQL table into a DataFrame.
- **7. read_sql_query()** - Executes an SQL query and returns the result as a DataFrame.
- **8. read_clipboard()** - Reads data from the system clipboard (useful for quick copy-paste).
- **9. read_pickle()** - Reads a pickled (serialized) pandas object from a file.
- **10. read_parquet()** - Reads Parquet files (columnar storage format, very efficient for large data).
- **11. read_feather()** - Reads Feather format files for fast I/O between Python and R.
- **12. read_stata()** - Reads Stata (.dta) files used in statistical analysis.
- **13. read_sas()** - Reads SAS files (.sas7bdat, .xpt).
- **14. read_hdf()** - Reads data from HDF5 files.
- **15. read_orc()** - Reads Apache ORC (Optimized Row Columnar) files.
- **16. read_xml()** - Reads data from XML files.
- **17. read_table()** - Reads general delimited text files (similar to read_csv, but more flexible).

Example Usage:

```
import pandas as pd # Reading CSV file df = pd.read_csv("data.csv") # Reading Excel
file df = pd.read_excel("data.xlsx") # Reading JSON file df =
pd.read_json("data.json") # Reading SQL query result import sqlite3 conn =
sqlite3.connect("database.db") df = pd.read_sql_query("SELECT * FROM users", conn)
```