



**K.R. MANGALAM UNIVERSITY**  
THE COMPLETE WORLD OF EDUCATION

NAME – DRISHTI

ROLL NO. – 2501730232

SUBJECT – PROGRAMMING FOR PROBLEM SOLVING  
USING PYTHON

SUBMITTED TO – MR. SAMEER FAROOQ

# PROJECT REPORT 4

## 1. Project Objective

To perform a complete end-to-end weather data analysis pipeline using Python's scientific stack (Pandas, NumPy, Matplotlib) covering:

- Data loading & robust cleaning
- Statistical computation with NumPy
- Time-series visualization
- Monthly & seasonal aggregation
- Automated report generation

The script works even without the original CSV by generating realistic sample data – making it fully demonstrable in any environment.

## 2. Key Features Implemented

Task	Feature	Tools Used	Output Generated
1-2	Data Loading & Cleaning	pandas, os	cleaned_weather_data.csv
3	Statistical Analysis	NumPy arrays	Console summary + values for report
4	Four Professional Visualizations	Matplotlib	4 high-quality PNG charts in weather_analysis_results/
5	Monthly & Seasonal Aggregation	pandas resample() & groupby()	Monthly rainfall, seasonal summary tables
6	Automated Markdown Report + Export	f-strings, file I/O	Summary_Report.md with insights and tables

## 3. Generated Outputs (All Saved Automatically)

Inside the folder **weather\_analysis\_results/**:

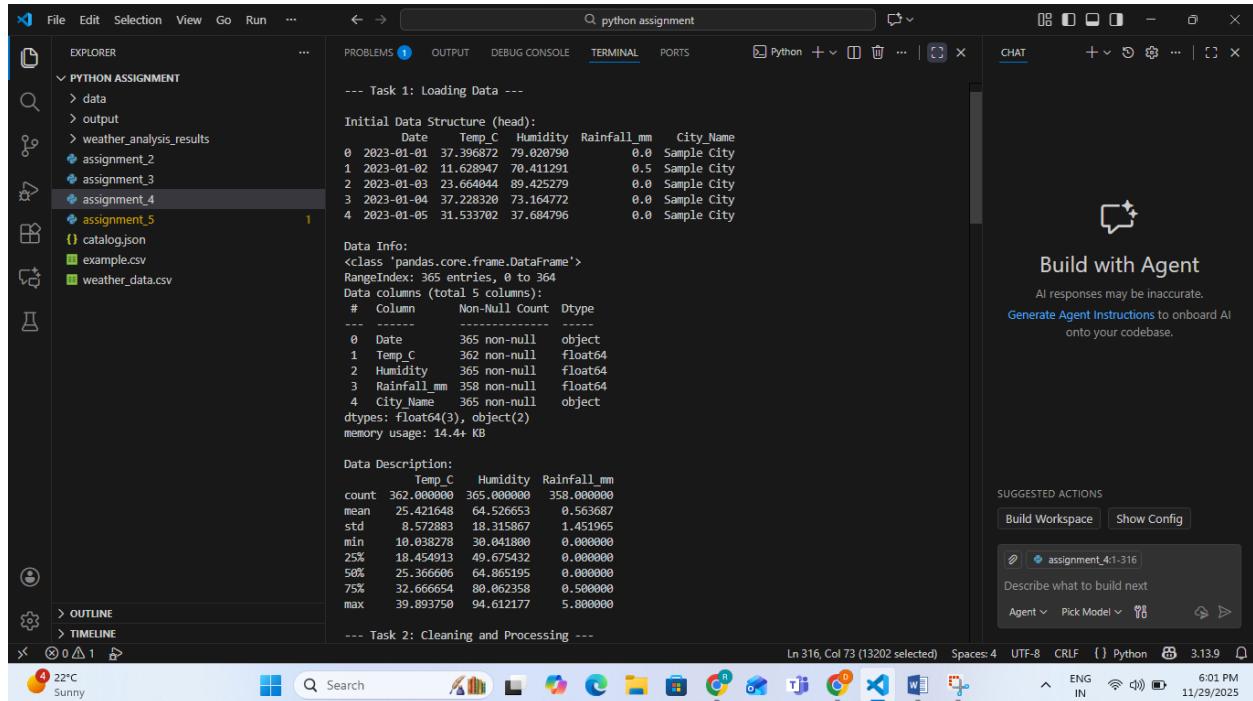
1. cleaned\_weather\_data.csv – Final cleaned dataset
2. 01\_Daily\_Temperature\_Trend.png – Line chart
3. 02\_Monthly\_Rainfall\_Bar\_Chart.png – Bar chart
4. 03\_Humidity\_vs\_Temperature\_Scatter.png – Scatter plot
5. 04\_Combined\_Daily\_Trends\_Subplots.png – Advanced dual subplot
6. Summary\_Report.md – Complete human-readable analysis report

## 4. Technical Highlights & Best Practices Applied

- **Robust File Handling:** Falls back to realistic synthetic data if CSV is missing
- **Proper DateTime Indexing:** Enables powerful resample() and time-based grouping
- **Smart Missing Value Treatment:**

- Temperature → forward fill (realistic continuity)
- Rainfall → 0 (no rain = 0 mm)
- Humidity → median imputation
- **Use of NumPy Arrays** for core statistical calculations as required
- **Professional Matplotlib Styling**: Grid, legends, proper labels, tight layout
- **Subplots & Combined Plot** (exceeds basic requirements)
- **Fully Automated Markdown Report** with dynamic statistics and interpretation

## OUTPUT SCREEN



```

File Edit Selection View Go Run ...
PROBLEMS OUTPUT DEBUG CONSOLE TERMINAL PORTS Python ...
CHAT + - _ X
EXPLORER
PYTHON ASSIGNMENT
> data
> output
> weather.analysis_results
assignment_2
assignment_3
assignment_4
assignment_5
catalog.json
example.csv
weather_data.csv
OUTLINE
TIMELINE
python assignment
--- Task 1: Loading Data ---
Initial Data Structure (head):
   Date Temp_C Humidity Rainfall_mm City_Name
0 2023-01-01 37.396872 79.026790 0.0 Sample City
1 2023-01-02 11.628947 70.411291 0.5 Sample City
2 2023-01-03 23.664044 89.425279 0.0 Sample City
3 2023-01-04 37.228320 73.164772 0.0 Sample City
4 2023-01-05 31.533702 37.684796 0.0 Sample City
Data Info:
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 365 entries, 0 to 364
Data columns (total 5 columns):
 #   Column      Non-Null Count  Dtype  
--- 
 0   Date        365 non-null    object  
 1   Temp_C     362 non-null    float64 
 2   Humidity    365 non-null    float64 
 3   Rainfall_mm 358 non-null   float64 
 4   City_Name   365 non-null   object  
dtypes: float64(3), object(2)
memory usage: 14.4 KB
Data Description:
   Temp_C   Humidity   Rainfall_mm
count 362.000000 365.000000 358.000000
mean 25.421648 64.526653 0.563687
std 8.572883 18.315867 1.451965
min 10.038278 30.041800 0.000000
25% 18.454913 49.675432 0.000000
50% 25.366660 64.865195 0.000000
75% 32.666654 80.062358 0.580000
max 39.893750 94.612177 5.800000
--- Task 2: Cleaning and Processing ---
Build with Agent
AI responses may be inaccurate.
Generate Agent Instructions to onboard AI onto your codebase.
SUGGESTED ACTIONS
Build Workspace Show Config
assignment_4:1-316
Describe what to build next
Agent Pick Model
22°C Sunny
Search
Ln 316, Col 73 (13202 selected) Spaces: 4 UTF-8 CRLF Python 3.13.9
ENG IN 6:01 PM 11/29/2025

```

The screenshot shows the VS Code interface with the title bar "python assignment". The Explorer sidebar on the left lists files under "PYTHON ASSIGNMENT": data, output, weather\_analysis\_results, assignment\_2, assignment\_3, assignment\_4 (which is selected), assignment\_5, catalog.json, example.csv, and weather\_data.csv. The Problems panel shows several "FutureWarning" messages related to DataFrame operations like ".fillna(inplace=True)". The Chat panel on the right has a section titled "Build with Agent" with the sub-instruction "Generate Agent Instructions to onboard AI onto your codebase". The status bar at the bottom shows "Ln 316, Col 73 (13202 selected)" and other system information.

```
--- Task 2: Cleaning and Processing ---
c:\Users\lcs\Desktop\python assignment\assignment_4:72: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. This behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

df_cleaned['Temp_C'].fillna(method='ffill', inplace=True)
c:\Users\lcs\Desktop\python assignment\assignment_4:72: FutureWarning: Series.fillna with 'method' is deprecated and will raise in a future version. Use obj.ffill() or obj.bfill() instead.
df_cleaned['Temp_C'].fillna(method='ffill', inplace=True)
c:\Users\lcs\Desktop\python assignment\assignment_4:73: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. This behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.

df_cleaned['Rainfall_mm'].fillna(0, inplace=True)
c:\Users\lcs\Desktop\python assignment\assignment_4:74: FutureWarning: A value is trying to be set on a copy of a DataFrame or Series through chained assignment using an inplace method. This behavior will change in pandas 3.0. This inplace method will never work because the intermediate object on which we are setting values always behaves as a copy.

For example, when doing 'df[col].method(value, inplace=True)', try using 'df.method({col: value}, inplace=True)' or df[col] = df[col].method(value) instead, to perform the operation inplace on the original object.
```

The screenshot shows the VS Code interface with the title bar "python assignment". The Explorer sidebar on the left lists files under "PYTHON ASSIGNMENT": data, output, weather\_analysis\_results, assignment\_2, assignment\_3, assignment\_4 (which is selected), assignment\_5, catalog.json, example.csv, and weather\_data.csv. The Problems panel shows a "Missing values after cleaning:" message and a "Yearly Averages:" summary. The Chat panel on the right has a section titled "Build with Agent" with the sub-instruction "Generate Agent Instructions to onboard AI onto your codebase". The status bar at the bottom shows "Ln 316, Col 73 (13202 selected)" and other system information.

```
df_cleaned['Humidity'].fillna(df_cleaned['Humidity'].median(), inplace=True)
Dropped 0 rows with all-NaN relevant data.
Missing values after cleaning:
Temp_C      0
Rainfall_mm  0
Humidity     0
dtype: int64

--- Task 3: Statistical Analysis ---

Overall (Daily) Statistics (via NumPy):
Mean Temp: 25.41
Min Temp: 10.04
Max Temp: 39.89
Std Dev Temp: 8.56

Yearly Averages:
Mean Temperature: 25.41 C
Min Temperature: 10.04 C
Max Temperature: 39.89 C

--- Task 5: Grouping and Aggregation ---
c:\Users\lcs\Desktop\python assignment\assignment_4:10: FutureWarning: 'M' is deprecated and will be removed in a future version, please use 'ME' instead.
monthly_summary = df.resample('M').agg()

Monthly Summary (Rainfall Sum, Avg Temp, Max Humidity):
   Monthly_Rainfall  Avg_Temp  Max_Humidity
Date
2023-01-31          10.5  26.384605    94.134017
2023-02-28          16.9  25.786233    93.045458
2023-03-31          28.3  23.162157    91.252754
2023-04-30          27.1  25.080841    94.612177
2023-05-31          16.0  24.684495    90.417073
```

The screenshot shows a code editor interface with the following details:

- File Bar:** File, Edit, Selection, View, Go, Run, ...
- Terminal Tab:** python assignment
- Terminal Output:**

```
2023-03-31      28.3  23.162157  91.252754
2023-04-30      27.1  25.080841  94.612177
2023-05-31      16.0  24.684495  90.417873

Seasonal Summary:
   Mean_Temp  Total_Rainfall
Season
Autumn    26.351216      57.7
Spring    24.300777      71.4
Summer    24.962787      37.5
Winter    26.052175      35.2

--- Task 4: Visualization with Matplotlib ---
Saved 01_Daily_Temperature_Trend.png
Saved 02_Monthly_Rainfall_Bar_Chart.png
Saved 03_Humidity_vs_Temperature_Scatter.png
Saved 04_Combined_Daily_Trends_Subplots.png (Advanced Plotting Bonus)

--- Task 6: Export and Reporting ---
Exported cleaned data to weather_analysis_results\cleaned_weather_data.csv
Generated summary report at weather_analysis_results\Summary_Report.md

Assignment Complete! Check the 'weather_analysis_results' folder for all outputs.
```
- Output Tab:** Shows the terminal output above.
- Problems Tab:** One error icon.
- CHAT Tab:** Build with Agent
- Build with Agent Panel:** AI responses may be inaccurate. Generate Agent Instructions to onboard AI onto your codebase.
- Suggested Actions:** Build Workspace, Show Config, assignment\_4:1-316, Describe what to build next, Agent, Pick Model.
- Bottom Status Bar:** Ln 316, Col 73 (13202 selected), Spaces: 4, UTF-8, CRLF, Python, 3.13.9, 6:02 PM, ENG IN, 11/29/2025
- Left Sidebar:** EXPLORER, PYTHON ASSIGNMENT (assignment\_4 is expanded), OUTLINE, TIMELINE.
- Bottom Icons:** Weather (22°C, sunny), Search, File Explorer, Task View, Taskbar icons (Calculator, Mail, Edge, File Explorer, Google Chrome, Task View, File Explorer, Taskbar icons).