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

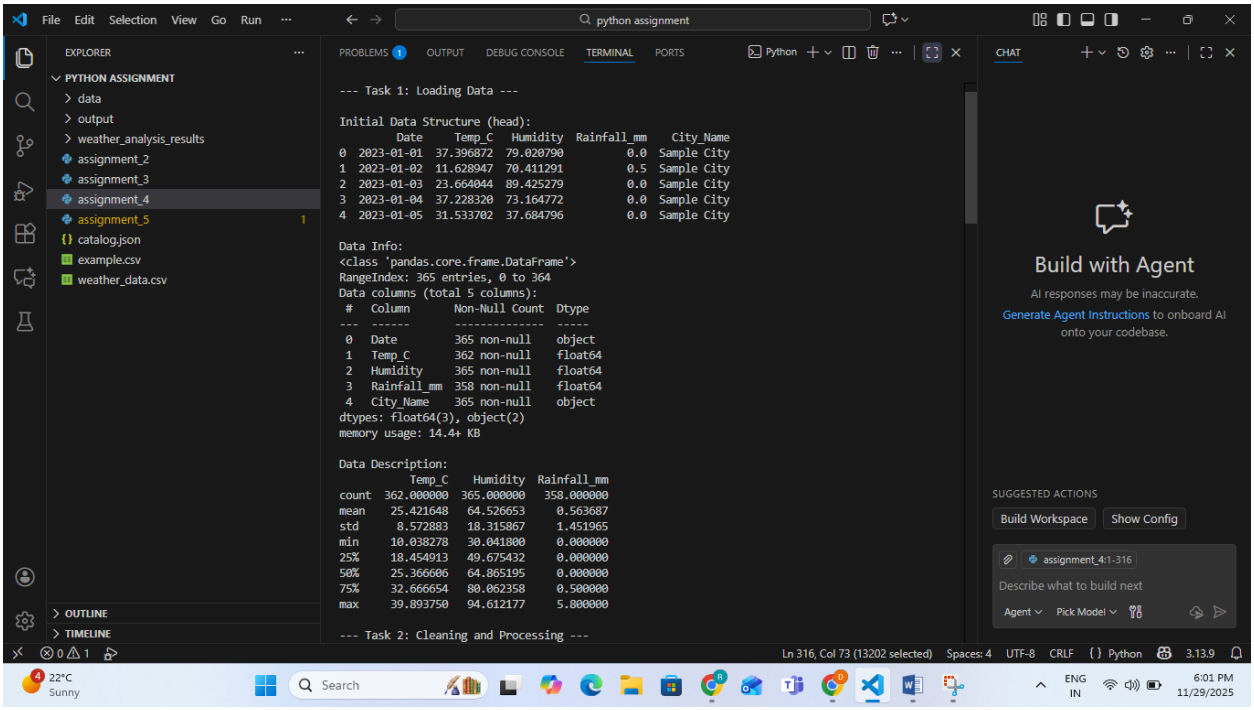
4. Technical Highlights & Best Practices Applied

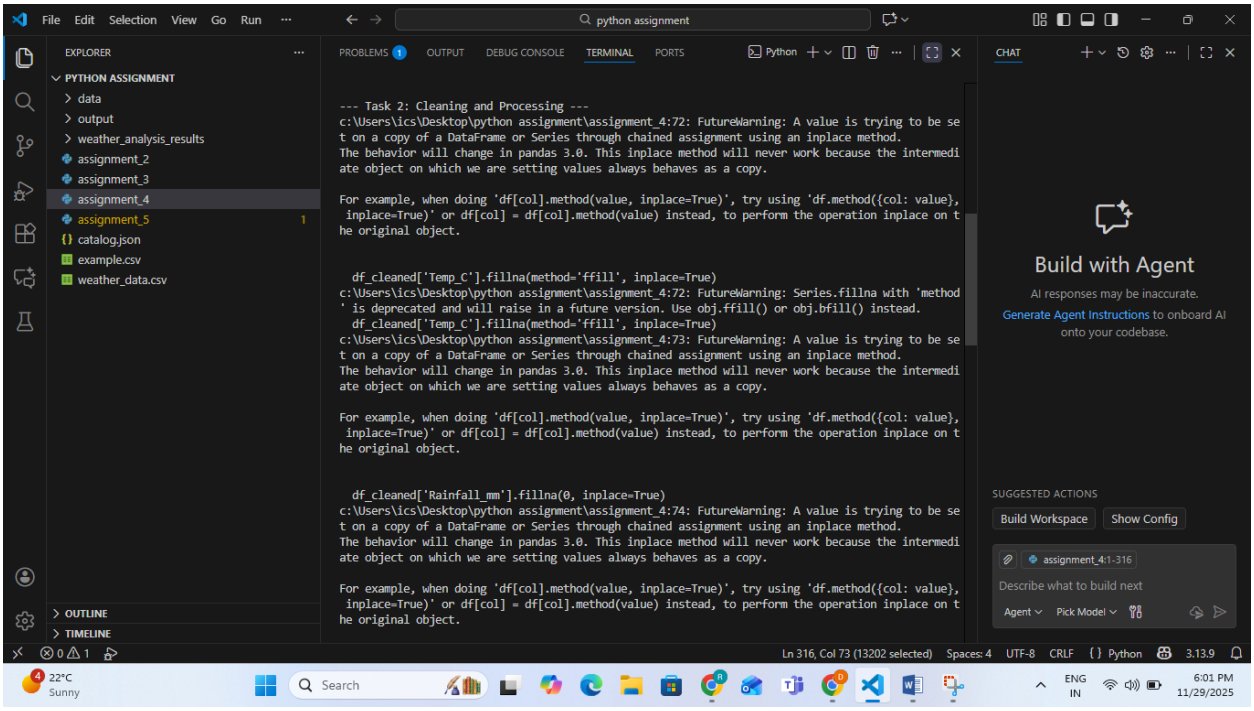
- **Robust File Handling**: Falls back to realistic synthetic data if CSV is missing
- **Proper DateTime Indexing**: Enables powerful resample() and time-based grouping
- **Smart Missing Value Treatment**:

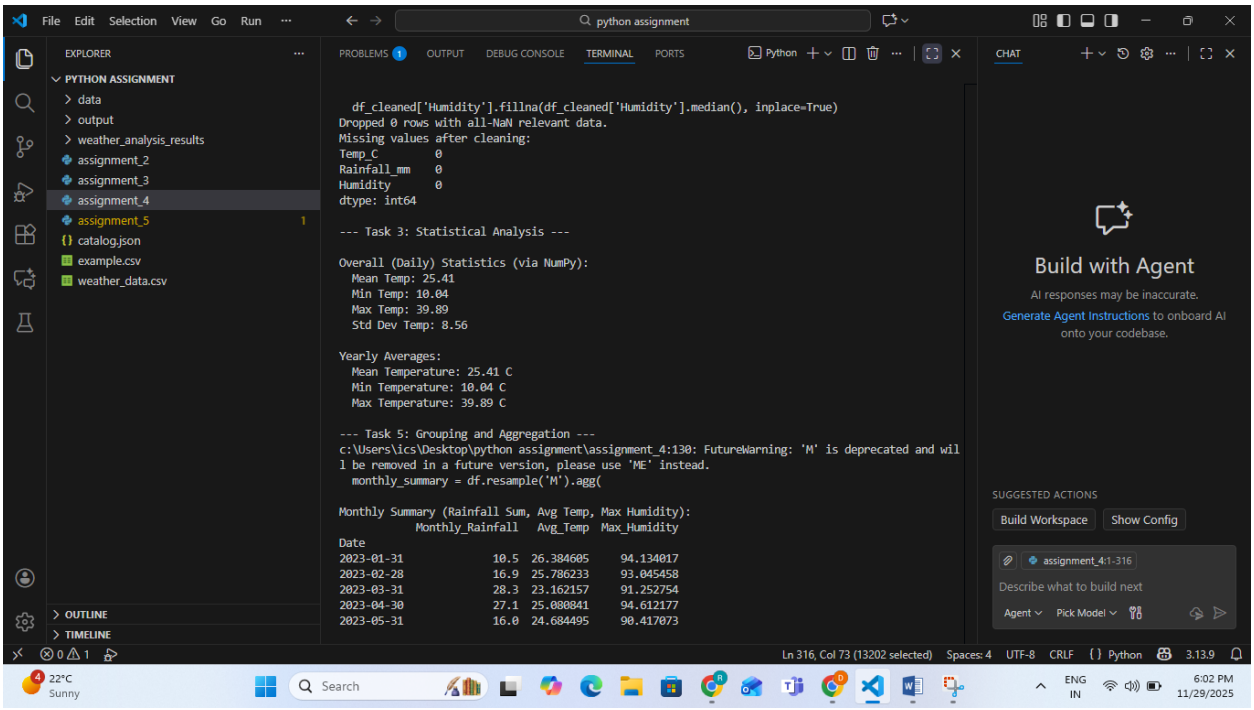
- Temperature → forward fill (realistic continuity)
- Rainfall → 0 (no rain = 0 mm)
- Humidity → median imputation

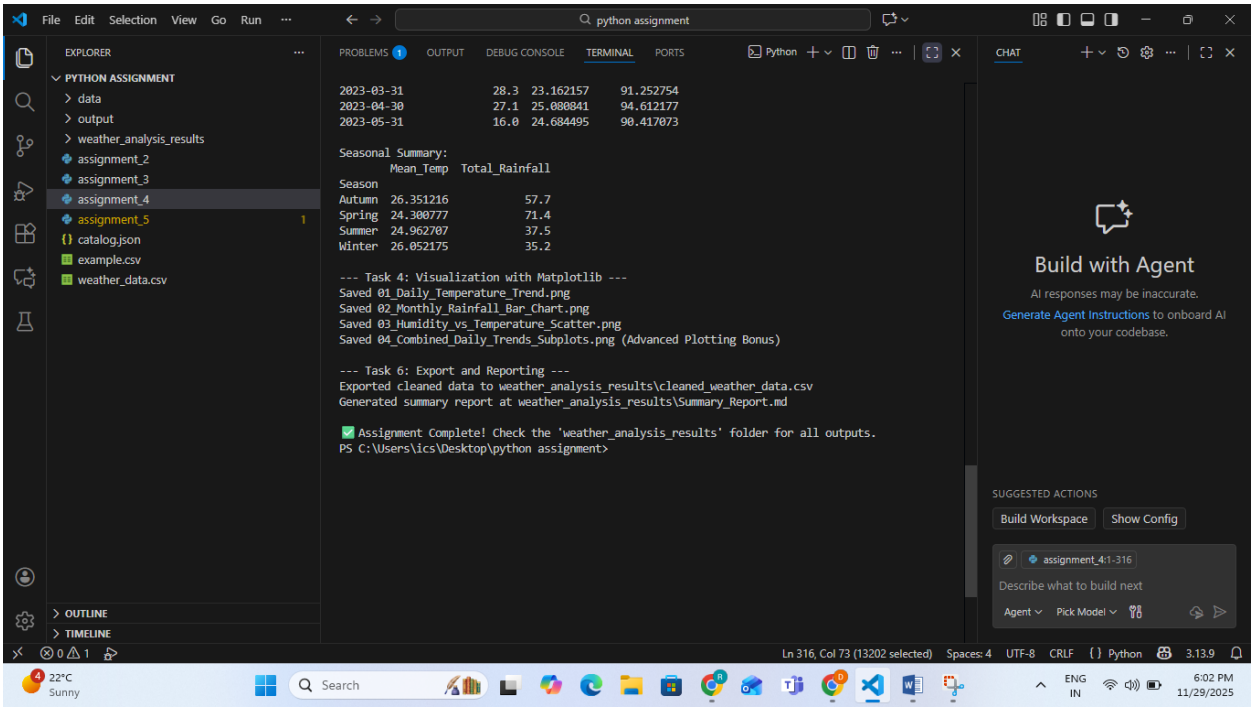
- **Use of NumPy Arrays** for core statistical calculations as required
- **Professional Matplotlib Styling**: Grid, legends, proper labels, tight layout
- **Subplots & Combined Plot** (exceeds basic requirements)
- **Fully Automated Markdown Report** with dynamic statistics and interpretation

OUTPUT SCREEN




File Edit Selection View Go Run
python assignment
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
PROBLEMS 1
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Python
--- Task 2: Cleaning and Processing ---
c:\Users\ics\Desktop\python assignment\assignment_4:72: FutureWarning: A value is trying to be se
t on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermedi
ate 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 t
he original object.
df_cleaned['Temp_C'].fillna(method='ffill', inplace=True)
c:\Users\ics\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\ics\Desktop\python assignment\assignment_4:73: FutureWarning: A value is trying to be se
t on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermedi
ate 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 t
he original object.
df_cleaned['Rainfall_mm'].fillna(0, inplace=True)
c:\Users\ics\Desktop\python assignment\assignment_4:74: FutureWarning: A value is trying to be se
t on a copy of a DataFrame or Series through chained assignment using an inplace method.
The behavior will change in pandas 3.0. This inplace method will never work because the intermedi
ate 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 t
he original object.
CHAT
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
Ln 316, Col 73 (13202 selected)
Spaces: 4
UTF-8
CRLF
Python
3.13.9
22°C
Sunny
Search
ENG
IN
6:01 PM
11/29/2025


File Edit Selection View Go Run
python assignment
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
PROBLEMS 1
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Python
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\ics\Desktop\python assignment\assignment_4:130: FutureWarning: 'M' is deprecated and wil
l 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
CHAT
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
Ln 316, Col 73 (13202 selected)
Spaces: 4
UTF-8
CRLF
Python
3.13.9
22°C
Sunny
Search
ENG
IN
6:02 PM
11/29/2025


File Edit Selection View Go Run
python assignment
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
PROBLEMS 1
OUTPUT
DEBUG CONSOLE
TERMINAL
PORTS
Python
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
Seasonal Summary:
Mean_Temp Total_Rainfall
Season
Autumn 26.351216 57.7
Spring 24.300777 71.4
Summer 24.962707 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.
PS C:\Users\ics\Desktop\python assignment>
CHAT
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
Ln 316, Col 73 (13202 selected)
Spaces: 4
UTF-8
CRLF
Python
3.13.9
22°C
Sunny
Search
ENG
IN
6:02 PM
11/29/2025