**INFERA DATA SCIENCE INTERNSHIP**

**Daily Progress Report**

**Date:** 11th June 2025  
**Intern Name:** Aniruddh Vijayvargia  
**Project:** Multi-Dataset Analysis – Week 1 Assignment  
**Report Day:** Day 1

**1. Executive Summary**

Today marked the initiation of my internship project. I understood the assignment requirements, created a structured plan for the 3-day sprint, reviewed the datasets, and prepared the development environment and project structure.

**2. Tasks Completed Today**

* Reviewed assignment brief and deliverables.
* Planned and structured a 3-day execution plan.
* Explored both sensor and weather datasets (sample screenshots attached).
* Created a professional folder and GitHub structure for the project.
* Drafted this daily report format for upcoming days.

**3. Technical Work Details**

* Explored sensor data: ~100 locations with hourly foot traffic.
* Explored weather data: daily environmental variables (temperature, rainfall, wind, pressure)
* Identified relevant features for preprocessing and modelling
* Created directory structure with folders for data, notebooks, reports, etc.
* Set up Python environment (pandas, numpy, matplotlib, seaborn, sklearn, streamlit)

**4. Key Learnings & Insights**

* Real-world datasets are often messy and must be reshaped
* Consistent folder structure and daily planning are critical for multi-day assignments
* Time-series + environmental correlation is a powerful analysis opportunity

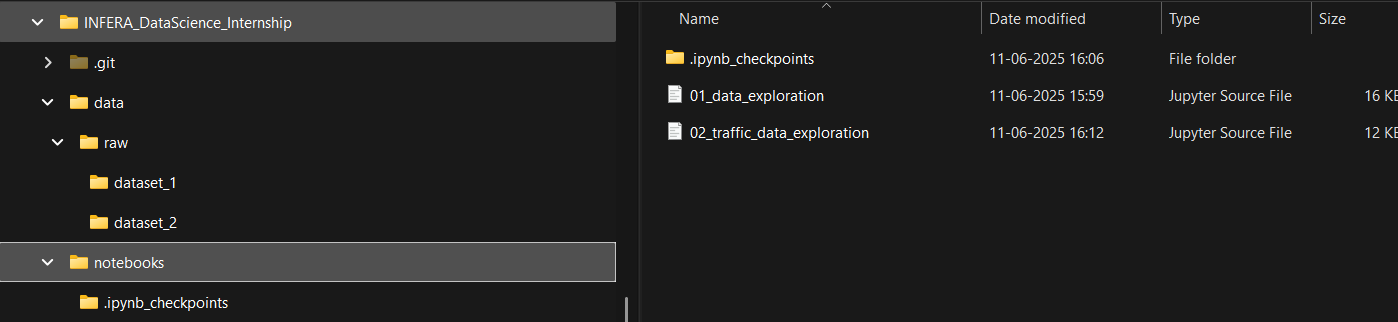
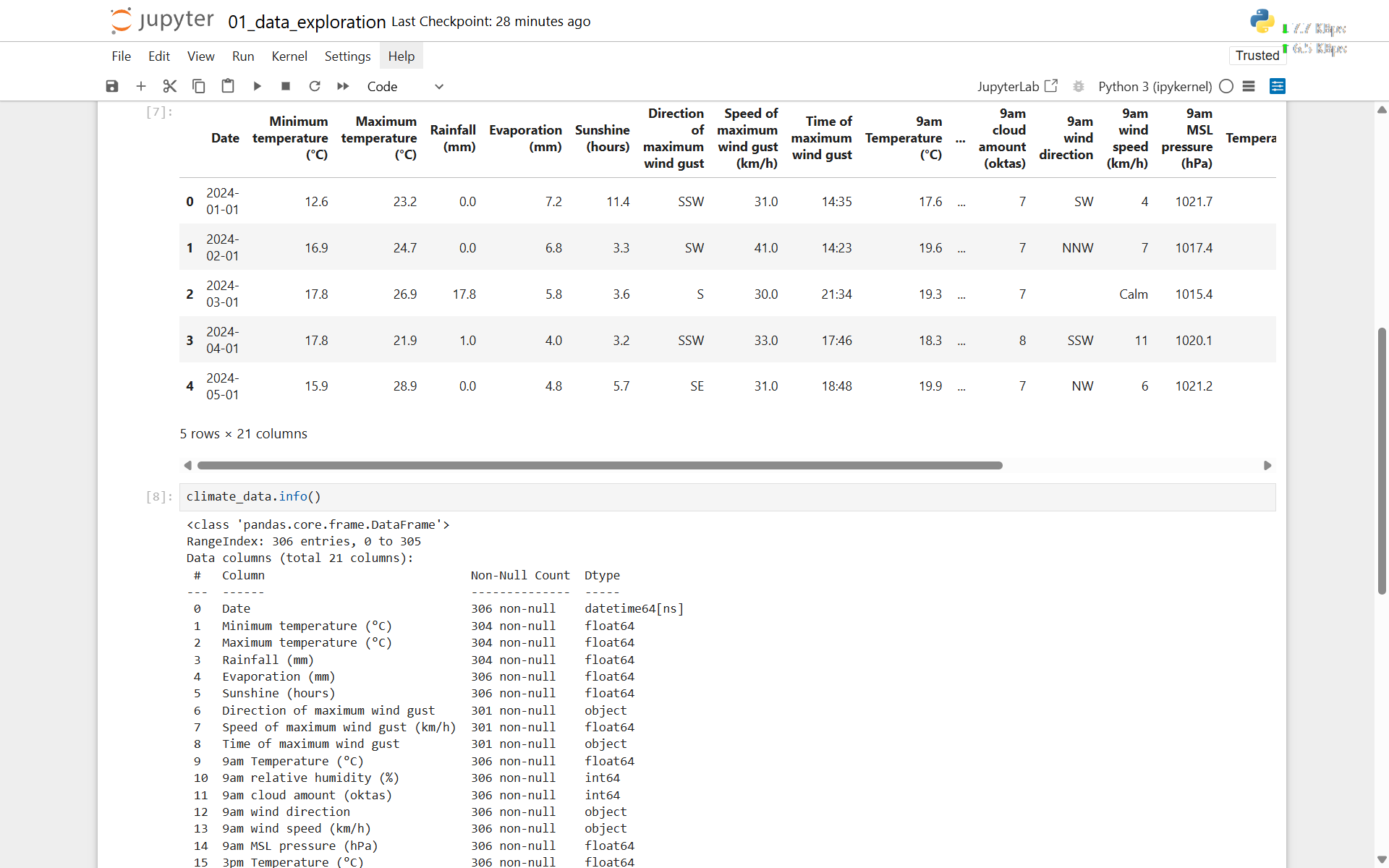
**5. Challenges Encountered**

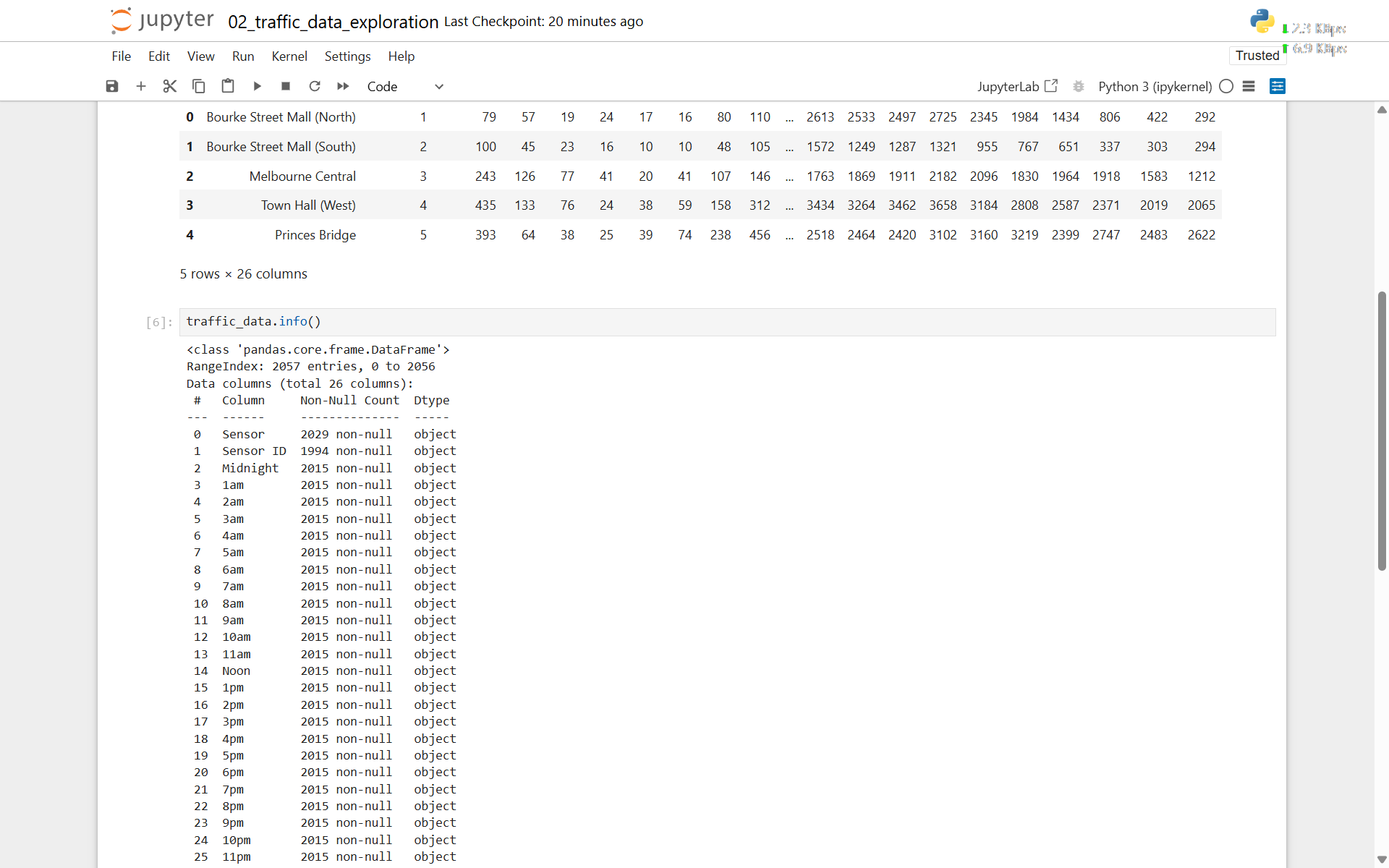
* PowerShell syntax error with && delayed GitHub structure creation
* Visual analysis of large dataset was initially overwhelming
* No critical blockers currently; full focus tomorrow on code

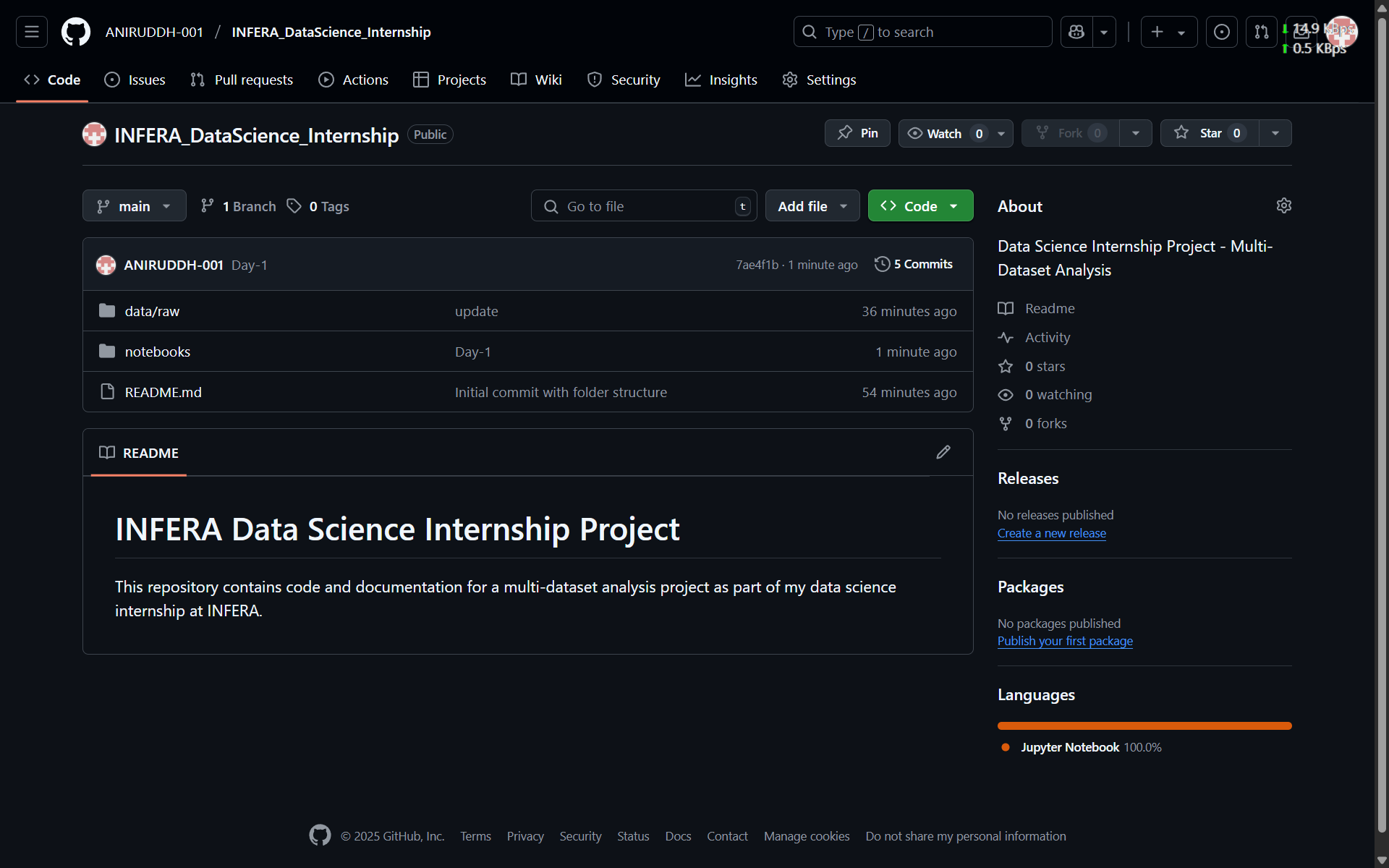
**6. Progress Metrics**

| **Task** | **Status** | **Completion %** |
| --- | --- | --- |
| Dataset Understanding | ✅ Done | 100% |
| Environment Setup | ✅ Done | 100% |
| Folder Structure & Planning | ✅ Done | 100% |
| Actual Coding (EDA/Cleaning) | ⏳ Planned | 0% |

**7. Screenshots & Evidence**





**8. Tomorrow’s Action Plan**

* Load all 14 Excel files into Pandas
* Melt and reshape sensor data for hourly analysis
* Clean weather data, handle missing values
* Begin plotting EDA visuals
* Start preparing for data merging

**9. Questions & Support Needed**

* None for today; awaiting further feedback after data exploration and reshaping tomorrow.