

Guru Nanak Dev Engineering College

Training Diary – TR-102 Report

Name: Divanshi Goyal

URN: 2302513

CRN: 2315056

Day 14

Training Summary

On Day 14, we developed an **AI-powered Data Analyst Agent** capable of reading CSV files, analyzing data, and responding to user queries with relevant Python code and its output. This system was designed to streamline data exploration for non-technical users by generating code and visualizations automatically.

Project Overview: AI Data Analyst Agent

The goal was to build a Streamlit-based web application where users could upload datasets and ask questions in natural language. The agent would:

- **Understand the query**
- **Generate Python code** using the Gemini model
- **Execute the code** in a sandboxed environment
- **Return the result/output and the code**

✂ Tech Stack Used

- **Streamlit** – For building the web interface
- **Pandas** – For data manipulation
- **Matplotlib** – For data visualization
- **OpenAI_utils / Gemini API** – For generating Python code from natural language
- **Sandbox execution** – For running generated code safely

🔗 Key Functionalities

- Upload any CSV file for analysis
- Ask queries like:

"Show the average sales by region"

"Plot a pie chart of product categories"

- View:
 - The Python code generated for the query
 - The result/output (text, tables, or visualizations)
-

Example Scenario

Upload sales_data.csv

Ask: "What are the top 5 products by revenue?"

Output: A bar chart and a sorted table, with the Python code used to generate them.

Learning Outcomes

- Understood how to connect **LLMs with real-world data**
- Gained hands-on experience in **secure code execution**
- Strengthened skills in **data visualization and analytics automation**