**Q1: Problem Statement**

**Let’s say you have two CSVs:**

1. **employees.csv**: employee\_id, name, department\_id, work\_mode, hours\_per\_week, productivity\_score, experience
2. **departments.csv:** department\_id, department\_name, head\_of\_department

**You want to allow questions like:**

**“What is the average productivity by department for remote workers with more than 5 years experience?”**

You are tasked with building an **AI-enabled interactive dashboard** where users can:

1. Upload HR data (employees.csv) and organizational structure data (departments.csv).
2. Ask **natural language questions** (e.g., *“Which department has the highest average productivity for remote employees?”*).
3. Automatically convert those questions into **SQL queries using an AI model**.
4. Execute the query and summarize the result in **human-friendly answers**.

Your application should support:

* **Multi-table SQL queries** (JOINs, GROUP BY, WHERE conditions)
* **Logical operations** (e.g., AND, OR, >, <)
* **Flexible natural language understanding** via LangChain + local LLM (llama3)
* A fully interactive **Streamlit** UI

**Sample User Questions**

* “List all employees who are in Engineering and have productivity > 8.”
* “Which department has the most hybrid workers?”
* “Show average hours per week by role for remote employees.”
* “Which head of department leads the team with the highest average experience?”

**Q2: Problem Statement**

Try same for below files

Area.csv

Invoices.csv

Offering.csv

**Sample User Questions**

* Calculate total actual profit per menu\_item.
* Calculate total stock used (total-wastage) per state per menuid
* Count number of invoices per year with target revenue more than 1000
* "Show total actual revenue by market."
* "Which menu category had the highest average profit?"
* "Compare target vs actual profit by area."
* "List top 5 areas with highest stock wastage."