Project Title:

Build an AI Agent to Answer E-commerce Data Questions

Datasets Provided

You will be given the following datasets:

- **Product-Level Ad Sales and Metrics**
- **Product-Level Total Sales and Metrics**
- **Product-Level Eligibility Table**

Objective

Your task is to **build an AI agent** that can:

- **Answer any question** related to the data provided.
- **Receive questions via API endpoints**, query the data, and **respond with accurate answers**.
- Bonus: If possible, **visualize the results** and provide **streamed responses** (like live typing effect).

Steps to Follow

1. **Convert the datasets into SQL tables.**

2. **Choose an LLM (Large Language Model)** that can run locally (downloadable and usable without internet).

3. **Write a codebase** that connects:

    - The LLM,
    - The SQL tables,
    - And the API endpoints to receive and respond to questions.

4. **Implement logic** so the AI agent can:

    - Understand the question,
    - Convert it into an SQL query,
    - Fetch the answer from the database,
    - And return it in a human-readable format.

5. **(Bonus)** Add:

    - Graphs/visuals for certain queries,
    - Event streaming responses to simulate real-time interaction.

Final Deliverables

- The complete **codebase** should be in the github and share the github link in the form.
- A **separate demo video** answering these example questions (recording must contain both the API call made and the output from terminal) - Upload in a drive and share it.
    1. *What is my total sales?*
    2. *Calculate the RoAS (Return on Ad Spend).*
    3. *Which product had the highest CPC (Cost Per Click)?*

Tips for Success

- Focus on structuring the data correctly in SQL.
- For LLM you can do any of the following:
    - Select an efficient, local LLM.
    - Use an already available free LLM API like Gemini 2.5 by google https://aistudio.google.com/apikey
- Make sure your system is modular: the LLM should translate the user's question to SQL, query the database, and send back a clean response.
- Add optional support for charts using libraries like **Matplotlib** or **Plotly** for bonus points.

Dataset:

1. Product-Level Eligibility Table (mapped)
2. Product-Level Ad Sales and Metrics (mapped)
3. Product-Level Total Sales and Metrics (mapped)

Submission Form:

https://forms.gle/QoDr7LUVvV47Pq2QA

Contact:

Thiruvikraman Anand - thiru.v@anarix.io

Ben Geo Abraham - ben.g@anarix.io