**Assignment Report**

**LLM-Powered Booking Analytics & QA System**

Objective: Develop a system that processes hotel booking data, extracts insights, and enables retrieval-augmented question answering (RAG).

Work-flow of the approach is same as mentioned in assignment, which is:-

1. Data Preprocessing
2. Analytics
3. Retrieval-Augmented Question Answering (RAG)
4. API Development
5. Performance Evaluation
6. Deploymnet

Data Processing and Analytics: The sample dataset of hotel booking which was given that only is used to solve the problem. It requires preprocessing before doing any kind of operations. Some preprocessing techniques like handling missing values, formatting of features were involved which doing preprocessing. For this libraries like- NumPy, Pandas, Matplotlib, Seaborn, Plotly etc. were used.

Retrieval-Augmented Question Answering (RAG): Now comes next part to build RAG model using csv dataset, it is a challenging task to work csv dataset as usually RAG model uses text data to develop model. The process starts with installing necessaries frameworks and libraries, langchain and huggingface are the frameworks which are used to do necessaary processing like- pipeline creation, converting csv data into text format, embedding, vector storing , creating chain, prompt ensuring and choosing open- source LLM model for embedding and reponse generation. The RAG-model answers the query asked by the user and also has feature of chat history by which user can see what he has serached. Below is sample queries which were run and their output.

API Development: For API development Flask is used which serve as the backend of the RAG model and provide support for deployment.

Deployment: For user-interface streamlit framework is used which is an open source deployment framework and provide nice interface.

[Github\_Repository](https://github.com/Anurag17singh/LLM_QA_System.git)