Design Document

Key Technologies Used:

1. **Streamlit**: A framework to build interactive web applications easily, allowing users to upload data, input parameters, and view results without needing complex front-end code.
2. **Pandas**: A powerful library for data manipulation, used to load, clean, and process the data (both historical and real-time) in tabular format.
3. **Machine Learning (Isolation Forest)**: Anomaly detection model used to identify outliers in the data, helping flag unusual patterns or deviations from the norm.
4. **Python's OS Library**: Used to handle file operations like creating directories and saving output files.

Why They Are Used:

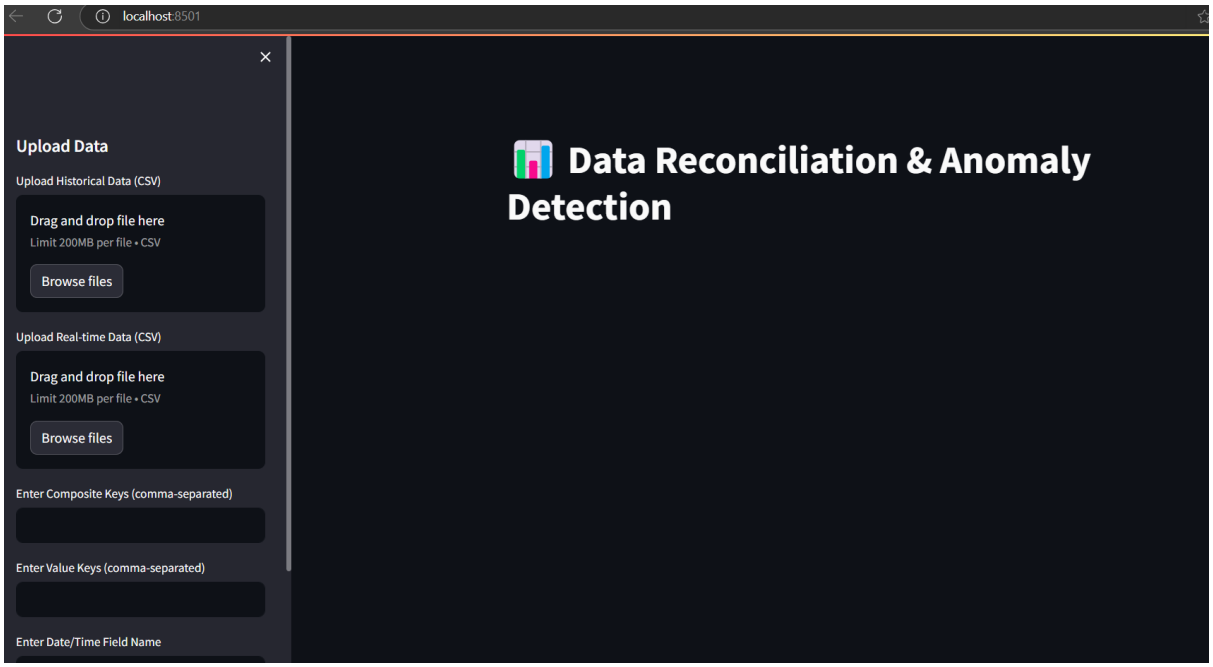
- **Streamlit** simplifies creating web apps and allows easy interaction with data.
- **Pandas** is essential for efficient data handling and analysis.
- **Isolation Forest** detects anomalies in data automatically.
- **OS** ensures the proper handling of files during processing.

These technologies together enable the creation of an easy-to-use web app for data reconciliation and anomaly detection.

Demo Screenshots

Step 1. Once we have executed streamlit run app.py command, below ui will pop-up. For full steps on how to run, please refer readme.md file inside code folder in repo.

UI URl – http://localhost:8501



Step2: We can upload historical_data.csv and real_time_data.csv file, and enter composite keys, value and date/time field name as highlighted below. Click on Process data to get anomaly result with anomaly reason.