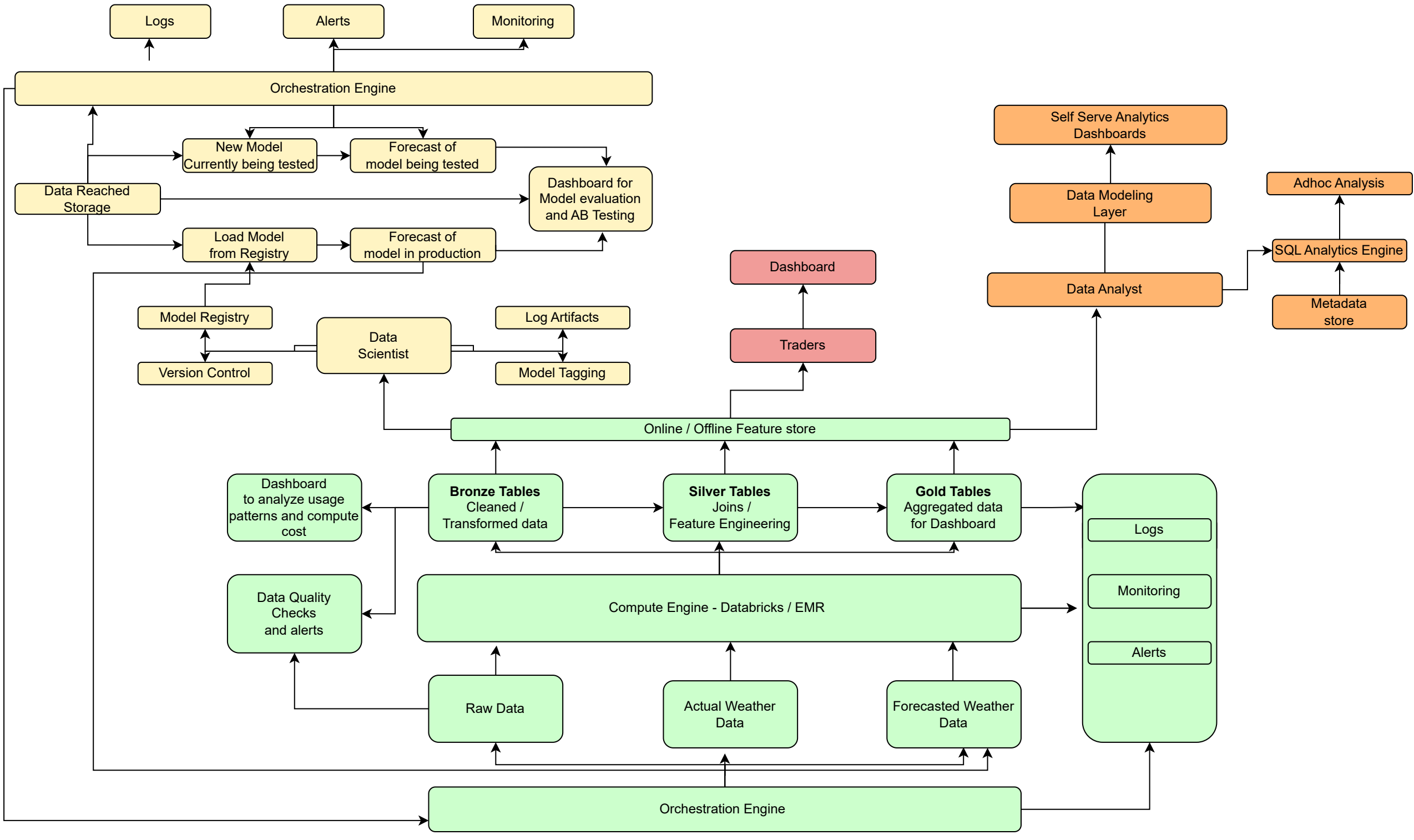
Weather Forecast Management Platform



Data processing module - The raw data along with acutal and forecasted weather data is cleaned and transformed in the bronze table. Different bronze tables are joined, features required for modeling and analytics are created and stored in silver table. The data in silver table is aggregated and stored in gold tables. Orchestration engine will take care of dependencies between the pipelines and their execution. Data quality checks will be carried out using packages like great expectations and users are alerted in case of any discrepancy. The logs for processing are stored and monitoring will be setup for the data processing. The computation required will be selected such that the **Gold tables will be created within 15 mins of forecasted data availablity.**

Data science Module - It can access data from Broze, Silver and Gold tables or via the feature store. A computation platform like Databricks or EMR will be used for modelling purpose. As a data scientist will be creating a model for each grid, the hyperparameters during trainng will be logged to Mlflow, each model will be versioned and tagged. The artifiacts for each model will be stored. The data and code will be verisioned and saved. Best performing model will be versioned and saved in registry. **Orchestration engine will be configured to run the pipeline every 6 hours.** The weather data will be pulled from storage, model will be loaded from registry and predictions will be generated. Once the predictions are stored, it will trigger the data processing pipeline which will update the Bronze, Silver and Gold tables. **This will ensure that data is available for dashboards within 15 mins of forecast generation**

Gold tables / feature store will be connected to dashboards for consumption by Traders

All the tables will be made available for adhoc analysis using SQL engine. The tables will be made available for data modeling in the BI tool of choice. Analyst can use either the sql engine or use the BI tool for analysis