Solution Architecture Overview

Azure Data Factory

Ingest — databricks

Transform — databricks

Analyze — databricks

Report

Ergast API

Azure Data Lake Raw Layer

Azure Data Lake Stage Layer

Azure Data Lake Presentation Layer

Power BI

I

Export data from the Ergast website

II

Import the data manually into Azure Data Lake "raw" Container

III

Take the data and process it using Databricks Notebooks to ingest into the stage Container within Azure Data Lake Storage

IV

Data in this layer is now columnar format parquet.

V

Ingested data will then be transformed via Databricks again & results will be stored in Azure Data Lake "presentation" container

VI

Similar to ingested data transform data is initially created in Parquet format

VII

Then we will use Databricks notebook to analyze the data & create Dashboards to satisfy the requirements for our analyses