# Challenge 00 - Overview

## Scenario

You are working with the New York City Taxi and Limousine Commission. They have several years’ of data about taxi trips in New York City, but have no way to analyze it.

You will build a modern data estate to enable analytic and real-time data analysis and visualization. Your task is to use Azure data technologies to build this data estate, including:

* Ingesting and storing raw data
* Cleaning and enriching data
* Loading data into a serving layer
* Consuming data in a BI layer
* Orchestrating the overall data flow for repeatability

## Notes for OpenHack

The OpenHack is time-boxed to a few hours. Due to the short timeframe, the architecture to build is provided here for you, and the challenges will be specific and prescriptive regarding technologies and configurations.

## Architecture



## Challenges

* Prepare Resources
* Ingest and Store Source Data in Blob Storage
* Clean and Merge Data in Databricks
* Ingest cleaned/merged data from blob storage in Azure Data Warehouse using Polybase and Azure Data Factory
* Create Data model in Azure Analysis Services using SQL DW as source
* Consume Data in Power BI from Databricks and Azure Analysis Services
* Access a Weather API, store weather data in Cosmos DB, join to transaction data in Power BI

Some of the challenges can be worked on in parallel; others require preceding challenges to be completed. We suggest you read all the challenges before starting work and discuss/assign them as a team.

## Outcomes

An end-to-end solution spanning the data lifecycle, with functional Power BI dashboard(s).

## Resources and Artifacts

Challenge documentation is available during the event on the OpenHack Team. Documents, proctor guides, and code will be made available after the OpenHack.