[External] Databricks Data Intelligence & Power BI Integration Demo

This demo walks through the latest capabilities of the Databricks Data Intelligence Platform as well as our deep integration with Power BI.

Capabilities shown include: Unity Catalog, DB SQL Serverless, DatabricksIQ, Genie Data Rooms, and Publish to Power BI.

———————————————————————————————————

Recorded demo

———————————————————————————————————

Reference

DI Platform & PBI Integration Demo.dbc - download and import this DBC to your workspace/repo folder.

- Setup Notebook: 00_Setup
- Cleanup Notebook: 01_Cleanup
- Catalog: main
    - Schema: airlinedata_your_username
        - Tables:
            - airports
            - flights
            - lookupcodes
        - Volumes
            - raw_airline_files

———————————————————————————————————

Setup

1. DI Platform & PBI Integration Demo.dbc - download and import this DBC to your workspace.
2. Run the 00_Setup notebook
3. Run through the steps of the demo prior to walking the customer through to actually create the Power BI semantic model, Power BI dashboard, and Genie Space that you will pull up during the demo.

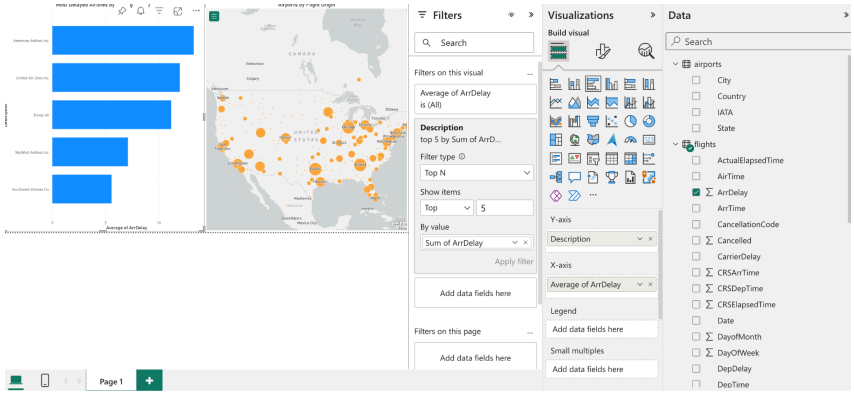
———————————————————————————————————

Demo

1. **UC/DBSQL/PBI Integration**

1. Navigate to Catalog Explorer - centralized data & AI catalog to easily browse & govern D&AI assets
    1. Catalog: main
    2. Schema: airlinedata_your_username
    3. Tables: ***briefly walk through what is in each table***
        1. airports: airports in each city and their IATA code
        2. flights: flight data including origin and destination time and location, unique carrier codes
        3. lookupcodes: mapping of unique carrier code to airline name
    4. Volumes
        1. raw_airline_files
            1. With Volumes, you can even interact with unstructured files, and easily create Delta Tables
            2. ***Click create table UI***
2. Flights table
    1. AI to auto-generate the table description and column descriptions
        1. ***Click AI Generate***
    2. ***Click on Sample Data*** - browse sample data without writing any queries
    3. ***Click on Permissions*** - easily grant permissions through the UI or SQL grants to users or groups. We even support RLS/CLS
    4. ***View Relationships*** - *can see the primary and foreign key constraints I set on this table*
    5. ***Do you use Power BI?***
        1. Let's publish this full UC schema to PowerBI > airlinedata > ***Publish to Power BI Workspace*** > ***workspace_name*** (don't publish, already did in setup) - Direct Query and Import mode supported
            1. Note: make sure the Fabric capacity is running
            2. Alternatively can open in Power BI Desktop
        2. Databricks PBI Flight Demo
            1. ***Open Data Model***
                - Made relationships between airport IATA and flight origin. Lookupcodes UniqueCode and flights UniqueCarrier
            2. ***Navigate to the Semantic Model***
            3. ***Create a Power BI Report***
                - *Map visual. Location: city, Size: origin*

- *Clustered bar chart visual. Y-axis: Description, X-axis: Average of ArrDelay. Add a filter on Description as Top N, 5, on Average of ArrDelay*



6. Flights table > ***Create Query*** - can query in SQL query editor
    1. Execute your queries with our Serverless SQL Warehouses which have best-in-class performance with instant spinup
        1. ***Choose Shared Unity Catalog Warehouse > Format Query > Run***
7. Maybe we want to know how many flights per year
    1. ***Assistant > Give me a query for the total number of flights per year***

2. **<u>Genie Spaces</u>** - new capability, democratize insights to non-technical teams, ask natural language questions to and get results in natural language or in a chart
    1. Flights table
        1. ***Create Genie Space***
            1. Title: ***Airline Demo UserName***
            2. Tables:
                1. Flights is already there
                2. ***Add main.airlinedata.airports, main.airlinedata.lookupcodes > Save***
    2. Data Room
        1. ***Click on explain the dataset button***
        2. ***Which airlines are delayed the most? Give me the top 5 by average delay***
        3. ***Click on show generated code*** - can easily use this to turn into a production pipeline

4. ***Now join this up with the lookup code table to give me the actual airline name***
5. ***Now graph this as a bar chart sorted most to least delayed***

3. **<u>Clean Up</u>** - if you are no longer using the demo
    1. Run the Cleanup Notebook
    2. Delete the Genie Space
    3. Delete the Power BI semantic model and Dashboard from the Fabric workspace