Data & AI Boot-Kon Event

Title: Data Canvas & Looker Studio

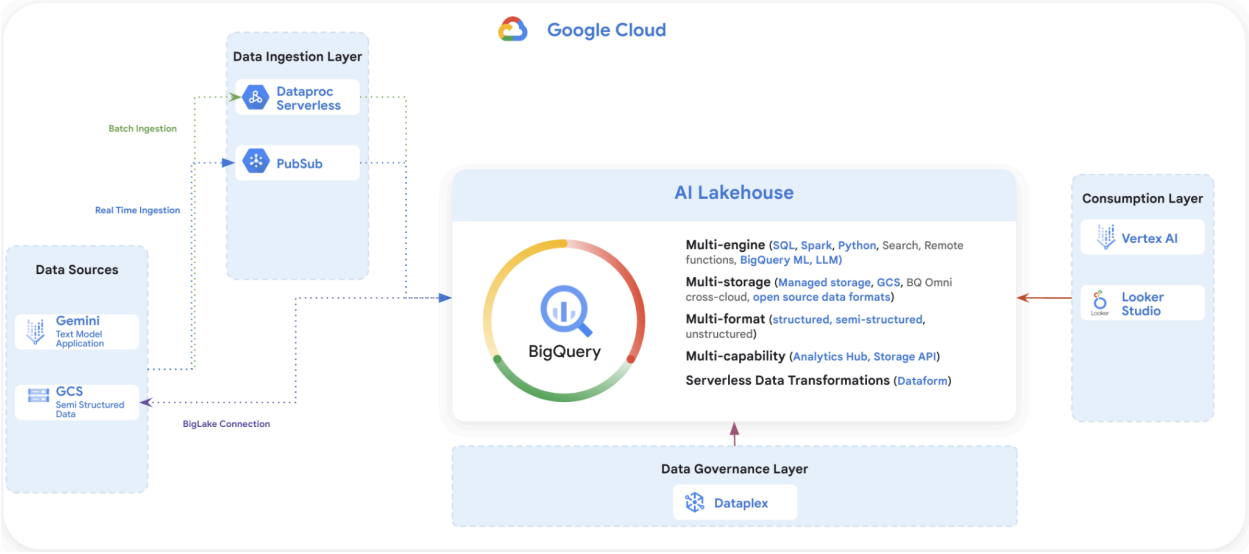
Goal of the lab

- To get data to insights with Natural Languages (NL)
- Without SQL Knowledge, only with NL to generate dashboards from the data

Author: Dinesh Sandra **Date**: 2024-04-01 **Estimated Completion Time**: 15 Minutes

CAUTION:
This lab is for educational purposes only and should be used with caution in production environments. Google Cloud Platform (GCP) products are changing frequently, and screenshots and instructions might become inaccurate over time. Always refer to the latest GCP documentation for the most up-to-date information.

Architecture Diagram:



What is Data Canvas?

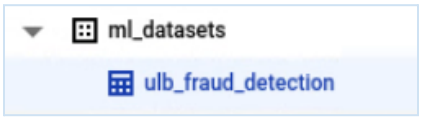
Data Canvas is a reimagined, NL-driven experience for data professionals to work with their data. Advances in transformer based models for LLMs, retrieval, NL2SQL present an opportunity to radically simplify the data to insights work that data professionals do.

The data to insights journey for data professionals is slow and inefficient. One of the main reasons is that it's a tools-first approach. People with specialized skills with specific tools need to collaborate to be able to get insights. This orchestration between people and tools creates a lot of friction, leading to many underutilized or unutilized opportunities to get insights from an enterprise's growing data corpus.

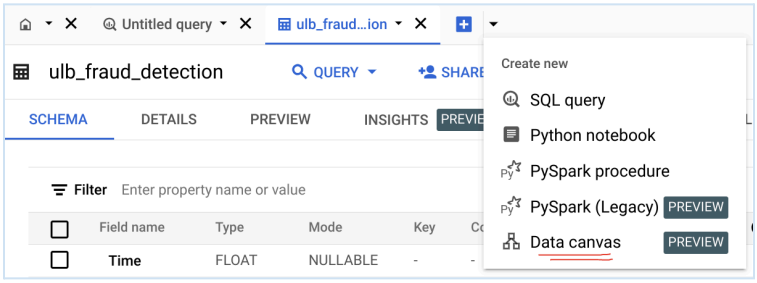
Data Canvas takes a *jobs-to-be-done* approach and offers a reimagined and context-rich user experience to help data professionals and enterprises get most out of their data. Users can use natural language to go through all the stages from data to insights. They can use it to discover data, build ingestion pipelines, prep or transform data, do advanced processing, compose analytics queries, do visualization and much more. And they can do all of this in a deeply collaborative manner.

[LAB] Data Canvas

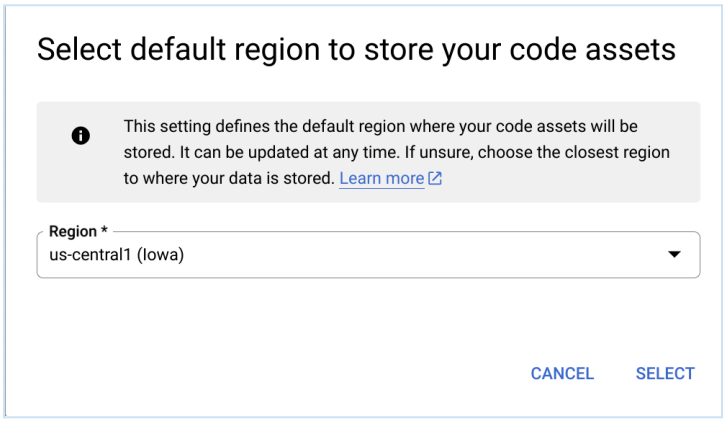
1. Goto Google Cloud Console, under BigQuery, navigate to Datasets, Tables(to the specific table ulb_fraud_detection)



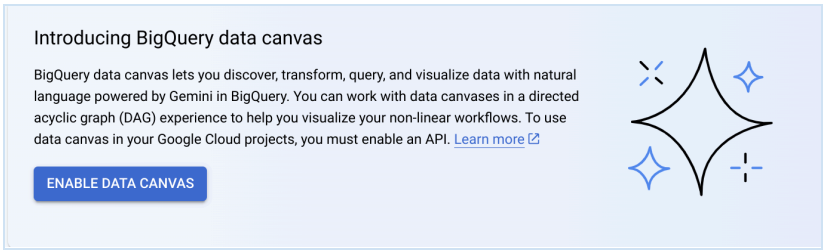
2. Click on dropdown button and select create data canvas as shown in the image below



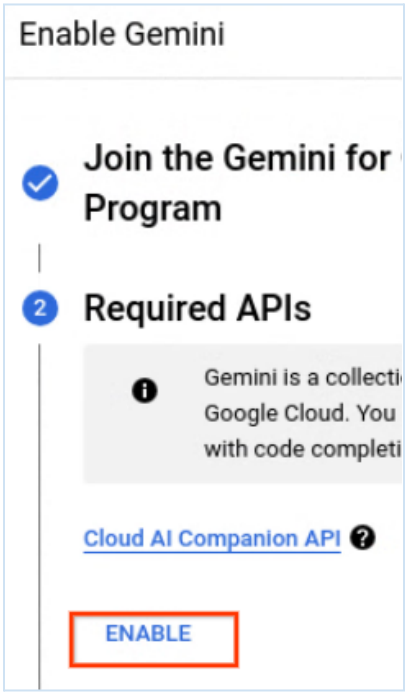
3. Select the Region for example: us-central1



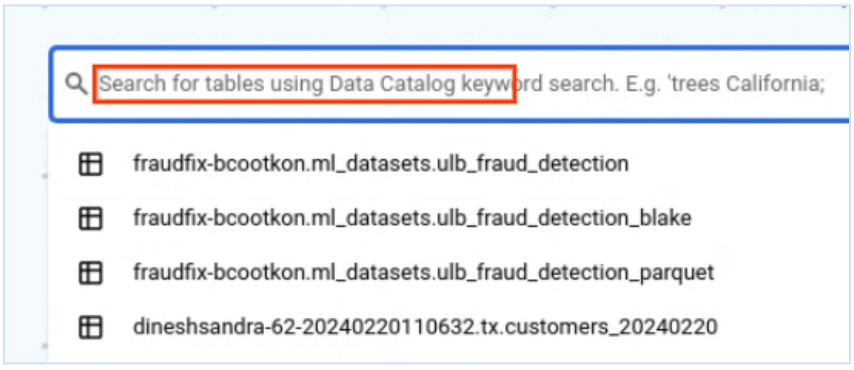
4. If you will be asked to enable the API, click on Enable API. Otherwise skip to step 5.



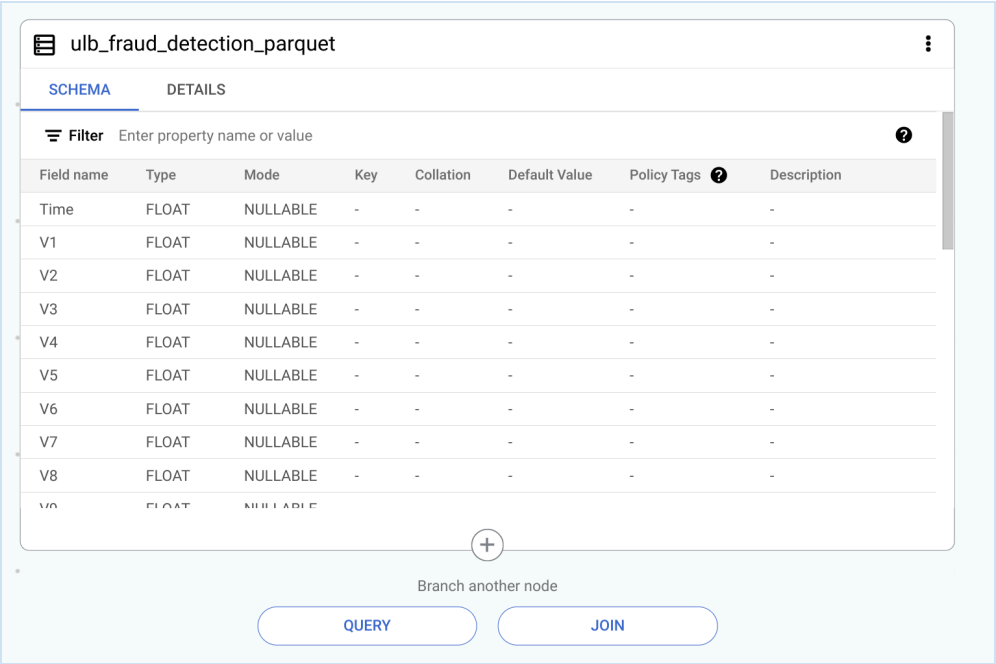
5. Enable Gemini APIs

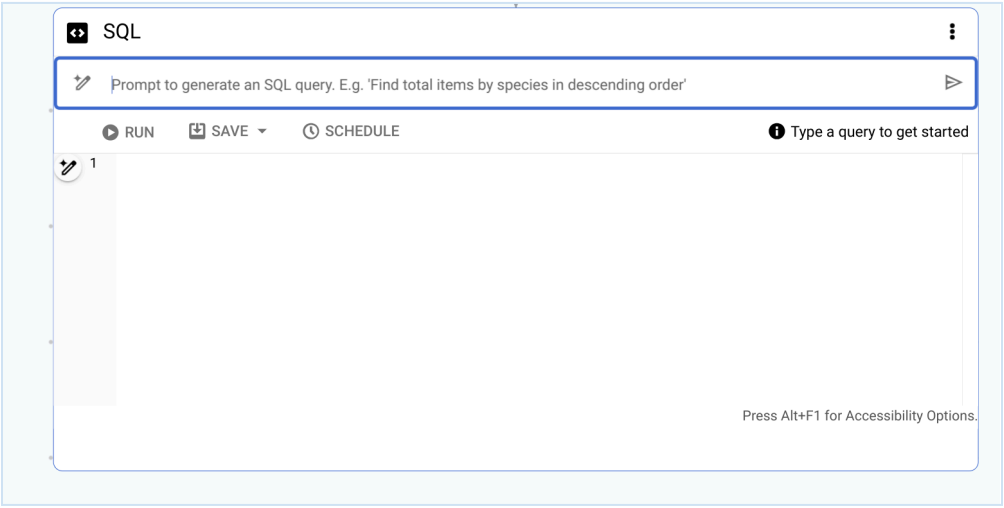


6. Type the **table name** that you want to explore data

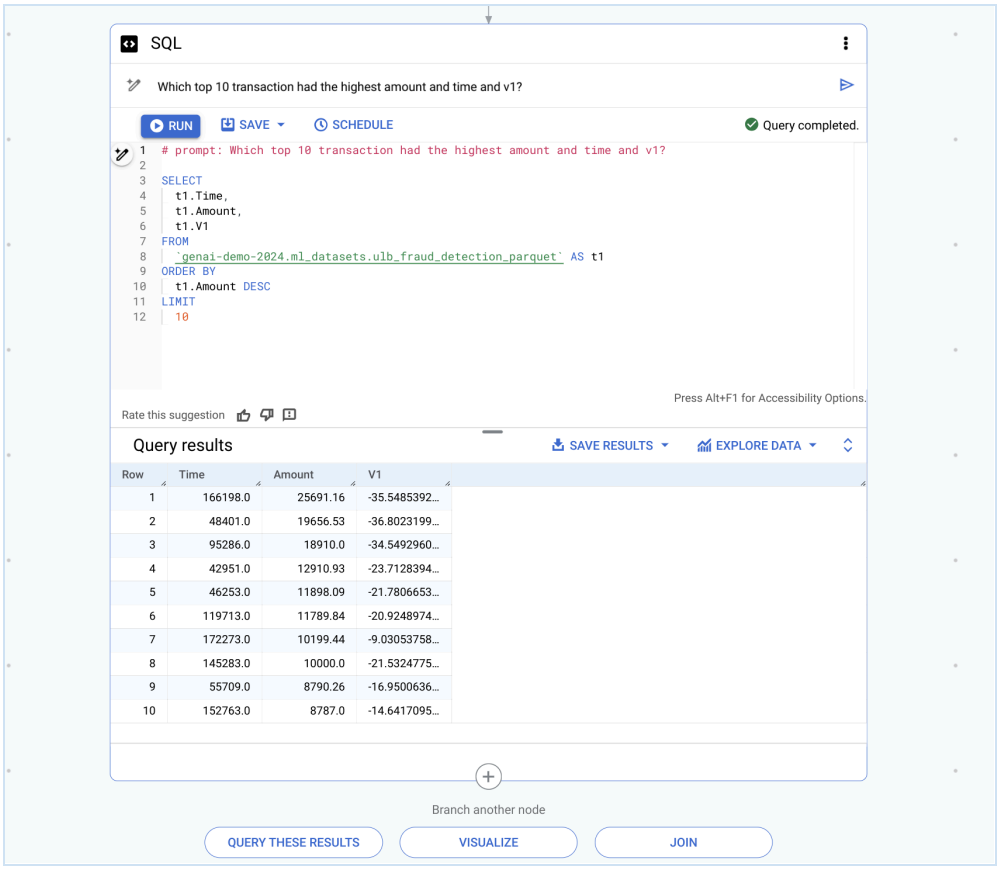


7.
8. Click on Query and type the question, let's ask the NL question, example...
   Which top 10 transaction had the highest amount and time and v1?


ulb_fraud_detection_parquet
SCHEMA
DETAILS
Filter
Enter property name or value
Field name
Type
Mode
Key
Collation
Default Value
Policy Tags
Description
Time FLOAT NULLABLE
V1 FLOAT NULLABLE
V2 FLOAT NULLABLE
V3 FLOAT NULLABLE
V4 FLOAT NULLABLE
V5 FLOAT NULLABLE
V6 FLOAT NULLABLE
V7 FLOAT NULLABLE
V8 FLOAT NULLABLE
Branch another node
QUERY
JOIN

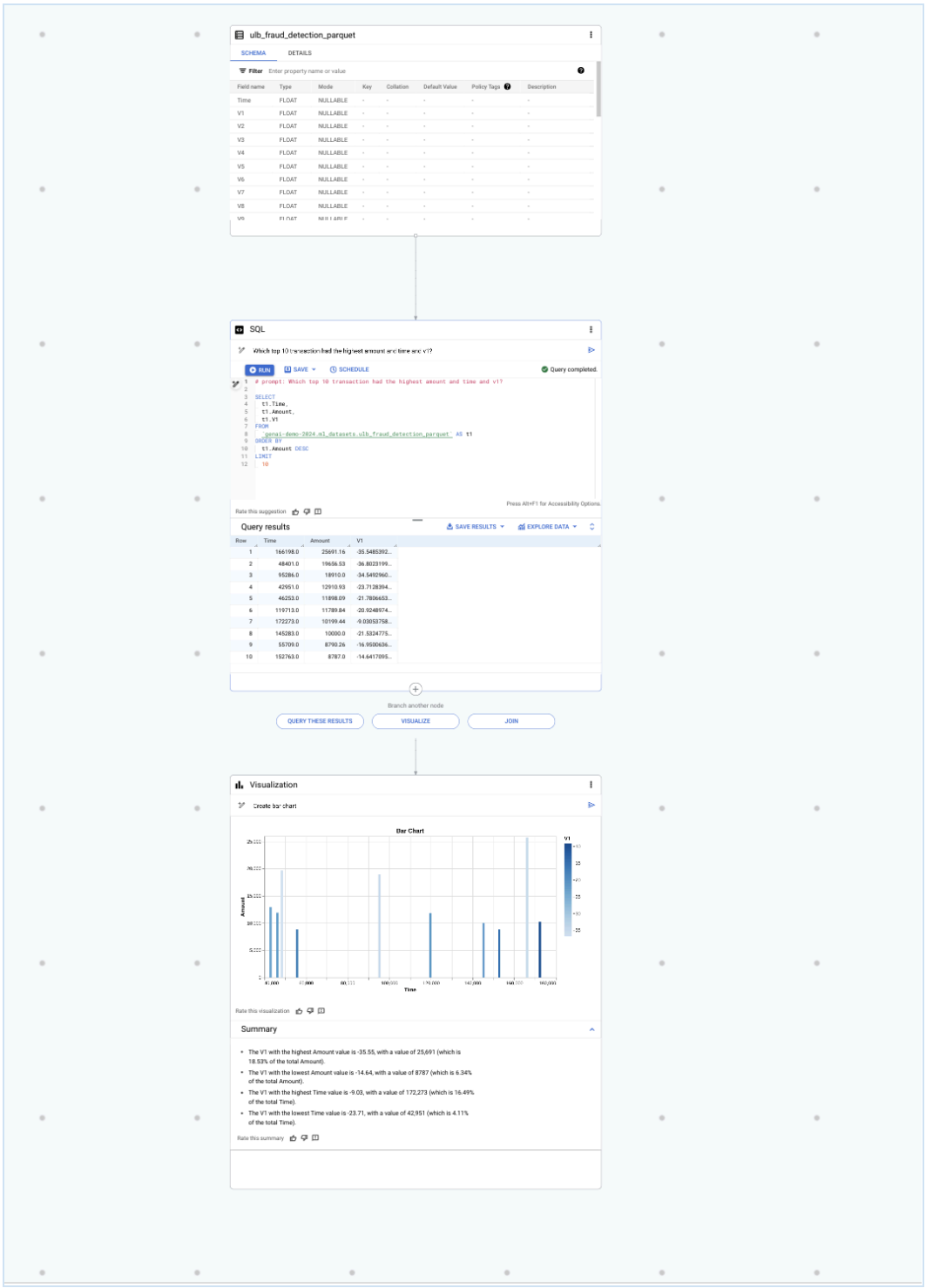

SQL
Prompt to generate an SQL query. E.g. 'Find total items by species in descending order'
RUN
SAVE
SCHEDULE
Type a query to get started
1
Press Alt+F1 for Accessibility Options.

9. Click **Run**



10. Then Click **Visualize,** select whatever chart you would like to visualize. Notice the Summary which gives you a quick summary of observations,

Conclusion: You see there are many options and possibilities to explore the data from BigQuery into reports via Natural Language as input.



🥳🥳Congratulations on completing Lab 7! 🥳🥳