

# Marketing Analytics - Week 1

---

## Introduction to Google Cloud Platform

In week 1, we will:

- [Analytics Challenges Faced by Data Analysts](#)
- [Compare Data On-Premises vs on the Cloud](#)
- [Learn from Real-World Use Cases of Companies Transformed through Analytics on the Cloud](#)
- [Navigate Google Cloud Platform Project Basics](#)
- [Activate Your Google Cloud Free Sandbox Account](#)
  - [Task 1: Activate your account and create a project](#)
  - [Task 2: Star the lab project in BigQuery](#)
  - [Task 3: Explore e-commerce data](#)

## Analytics Challenges Faced by Data Analysts

Some comments from the analysts:

- “My queries are taking way too long to run and is stalling my analysis.”
- “I have no easy way to combine and query all the data I’ve collected”
- “We’re a data department, not an infrastructure department. Maintaining and upgrading our own servers is unsustainable.”
- “My on-premise clusters aren’t scaling with my analysis”
- “We can only afford to store a subset of the data our business generates”
- “We don’t have a central data analytics warehouse or set of tools”

### Q1: Based on the above comments and your experience, what are the challenges faced by data analysts?

There are three challenges faced by data analysts: Querying, Infrastructure, and Storage. Two of the most common barriers a data analyst will run into are either too much data or data that is not connected together.

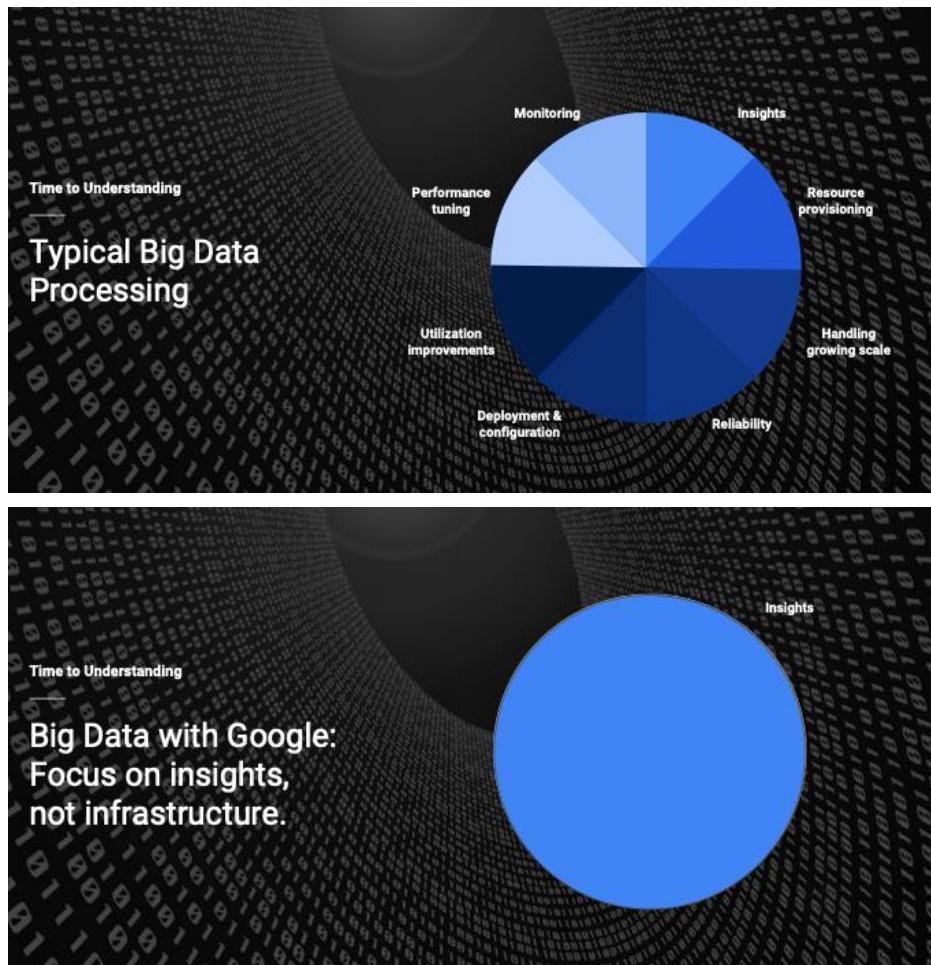
## Compare Data On-Premises vs on the Cloud

Reasons why Google Cloud Platform is used for Data Analysis

- Storage is Cheap
- Focus on Queries, not infrastructure
  - Traditional big data platforms require an investment in infrastructure



- Although hard drives are cheap, they’re not the only thing you need to have to query big data.
- Essentials:
  - 1 – Storage
  - 2 - Computing Power
  - 3 – Networking
  - 4 - Admin and hardware teams to maintain and upgrade your infrastructure (Not to mention software and software license costs)
- Massive Scalability



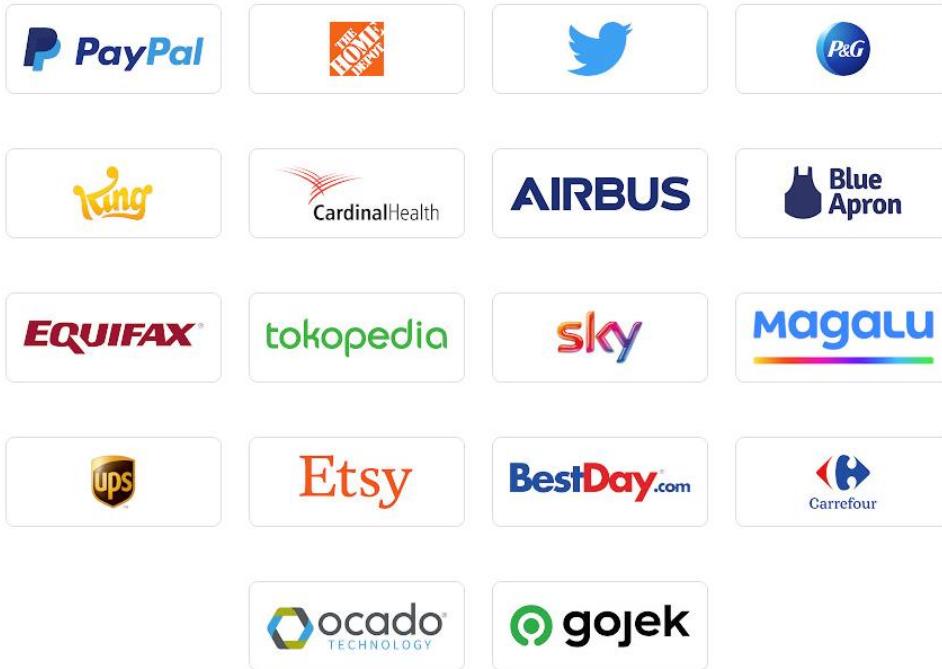
Key insights from the Comparisons

- 1) **scale with data growth** even as it explodes
- 2) are **managed** so that you aren't wasting time on dealing with all the underlying complexities
- 3) are just generally **magically awesome**, so you can get back to **data insights**

## Learn from Real-World Use Cases of Companies Transformed through Analytics on the Cloud

- **Twitter**
- **JB Hi-Fi**
  - Increasing revenue and improving the relevance of recommended products with Recommendations AI.
- **Target**
  - “We want to be a positive force behind millions of people, every day. If we’re going to keep doing that, we need cloud technology that focuses on experiences. That’s elegant, simple, and adaptive. That doesn’t lock us in.” (Krishnan Srinivasan, Vice President of Cloud and Compute Infrastructure Platforms, Target).

- Leading companies around the world are choosing Google Cloud



## Navigate Google Cloud Platform Project Basics

### 1. Projects

### 2. Resources

### 3. Billing

The screenshot shows the Google Cloud Platform dashboard for a project named 'data-to-insights'. The 'Project info' section displays the project name, ID, and number. The 'Resources' section lists 'Cloud Storage' with 1 bucket ('BigQuery') and 2 datasets. The 'Billing' section shows estimated charges of \$0.00 for the billing period Aug 1 – 3, 2017.

## Background on: Project resource

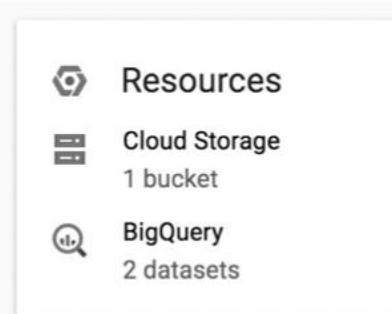
1. Projects organize and govern your activities in the cloud

First of all, a project is required to use Google Cloud Platform, and forms the basis for creating, enabling, and using all Cloud Platform services, managing APIs, enabling billing, adding and removing collaborators, and managing permissions.

All projects consist of the following:

- Two identifiers:
  - Project ID, which is a unique identifier for the project.
  - Project number, which is automatically assigned when you create the project. It is read-only.
- One mutable display name.
- The lifecycle state of the project; for example, ACTIVE or DELETE\_REQUESTED.

- A collection of labels that can be used for filtering projects.
  - The time when the project was created.
2. Resources are what you are using in the cloud  
Commonly used by data analysts:
- **Storage** in Google Cloud Storage
    - Example: You use a Bucket for uploading large CSV files to ingest later for analysis
  - **Datasets** in Google BigQuery
    - Example: You perform analysis on raw data and create a brand-new dataset



- The Cloud Storage Bucket is your go to for scalable storage
- Buckets are scalable containers that hold your data.
  - You can create and upload files to your buckets within your Cloud Console
3. You are billed for the resources you use (currently we are using a free Google sandbox)
- Commonly used by data analysts:
- **Storage in Google Cloud Storage**
    - Billed for Bucket Storage
  - **Datasets in Google BigQuery**
    - Billed for Query processing
    - Billed for Table Storage

A walkthrough guide on enabling billing to look at BigQuery usage:

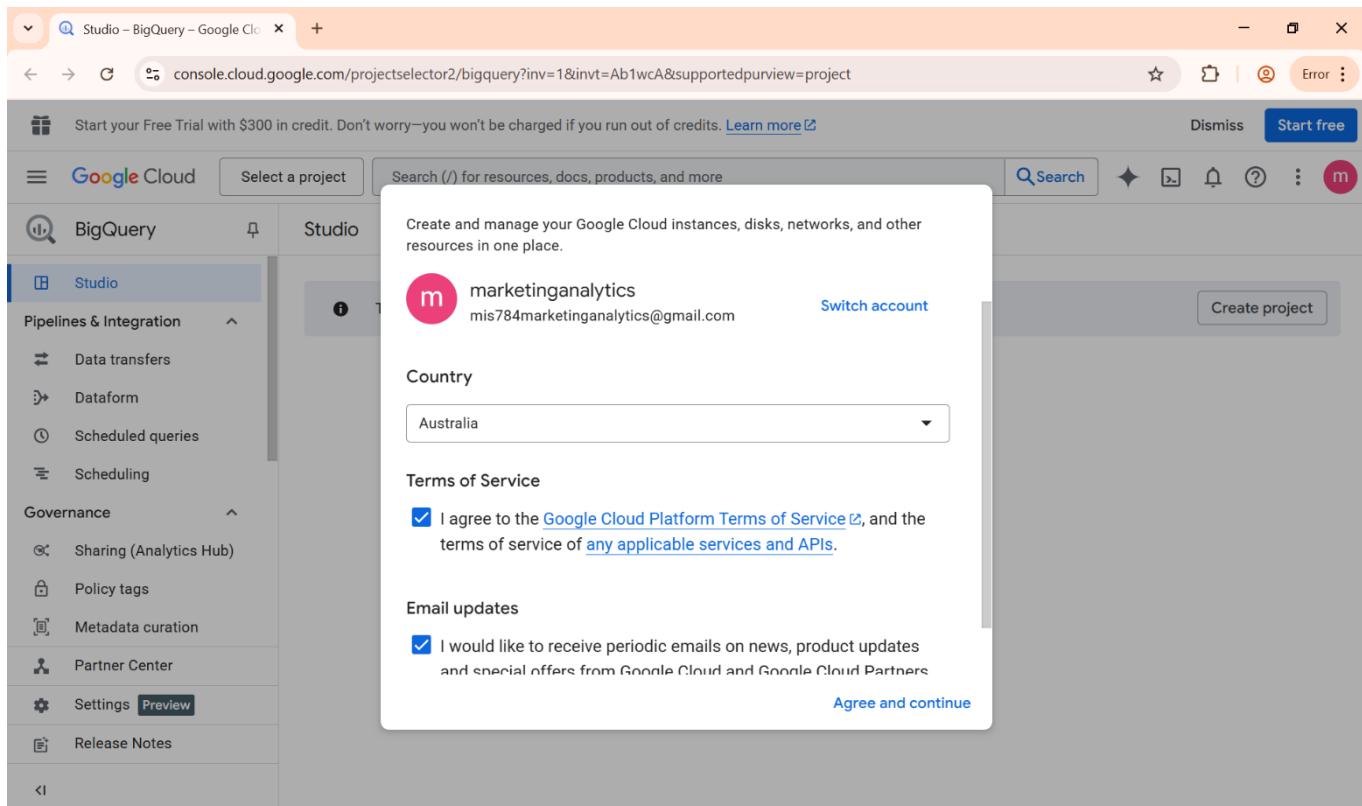
<https://medium.com/google-cloud/visualize-gcp-billing-using-bigquery-and-data-studio-d3e695f90c08>

# Activate Your Google Cloud Free Sandbox Account (BigQuery sandbox) - A Step-by-Step Guide

The BigQuery sandbox lets you experience BigQuery without providing a credit card or creating a billing account for your project.

## Task 1: Activate your account and create a project

- **Step 1:** Sign in to your Google account. If you don't have a Google account, you can [create one](#).
- **Step 2:** Once you sign in to your Google account, you can open BigQuery in the Google Cloud console by entering the following URL in your browser.  
<https://console.cloud.google.com/bigquery>
- **Step 3:** On the welcome page, do the following:
  - For Country, select your country as "Australia".
  - For Terms of Service, select the checkbox if you agree to the terms of service.
  - Optional: If you are asked about email updates, select the checkbox if you want to receive email updates.
  - Click Agree and continue (while you're using the sandbox, you do not need to create a billing account, and you do not need to attach a billing account to the project).



- **Step 4:** Click Create project (to use the BigQuery sandbox, you must create a Cloud project).

- For Project name, enter a name for your project.
- For Location, select “No organization”.
- Click Create. You are redirected back to the BigQuery page in the Google Cloud console (after you create a Cloud project, the Google Cloud console displays the sandbox banner).

You have 12 projects remaining in your quota. Request an increase or delete projects. [Learn more](#)

[Manage Quotas](#)

Project name \* MyProject1

Project ID: myproject1-464811. It cannot be changed later. [Edit](#)

Location \* No organization [Browse](#)

Parent organization or folder

[Create](#) [Cancel](#)

Welcome to BigQuery Studio!

Create new

[SQL query](#) [Notebook](#) [Notebook with Spark](#) [Data canvas](#) [Pipeline](#)

[Data preparation](#) [Table](#)

Try with sample data

Try the Google Trends Demo Query

This simple query generates the top search terms in the US from the Google Trends public dataset.

[Open query](#) [View dataset](#)

Repository [Preview](#)

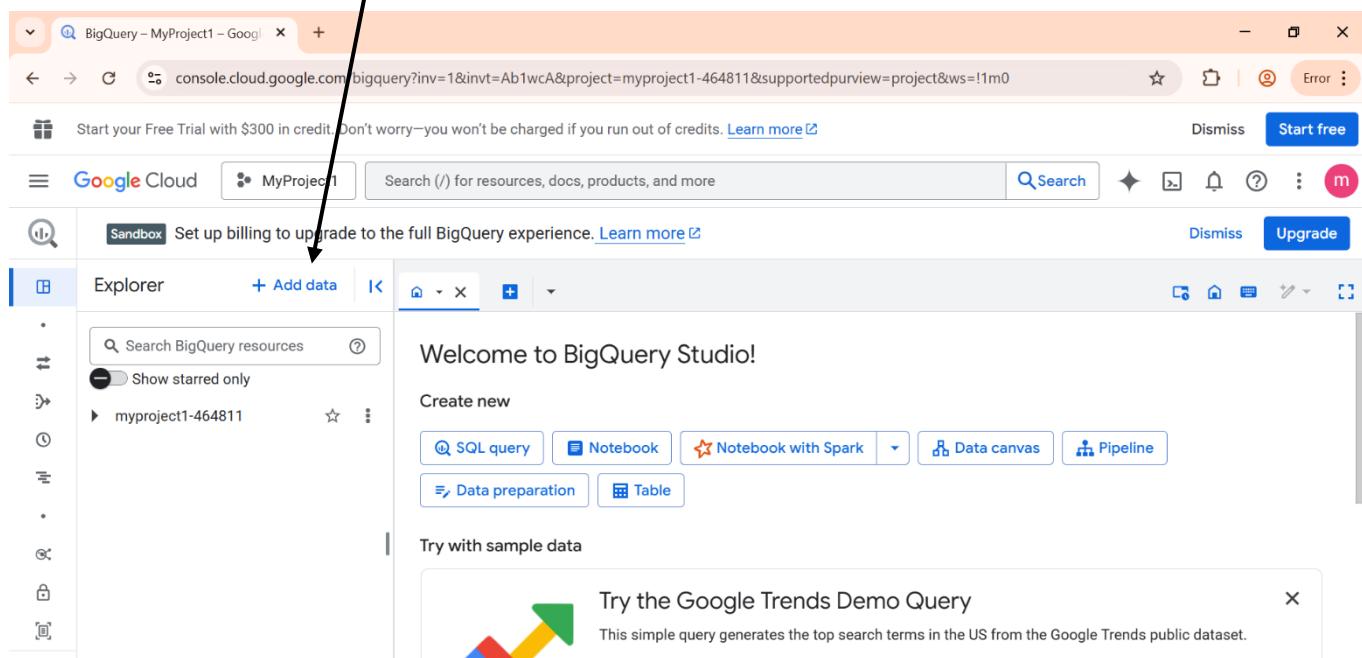
Select a repository and a workspace to view its content.

Job history [Refresh](#)

**Note:** [Learn more about additional Google Cloud benefits for students](#)

## Task 2: Star the lab project in BigQuery

- Step 1: Click + Add data.



Welcome to BigQuery Studio!

Create new

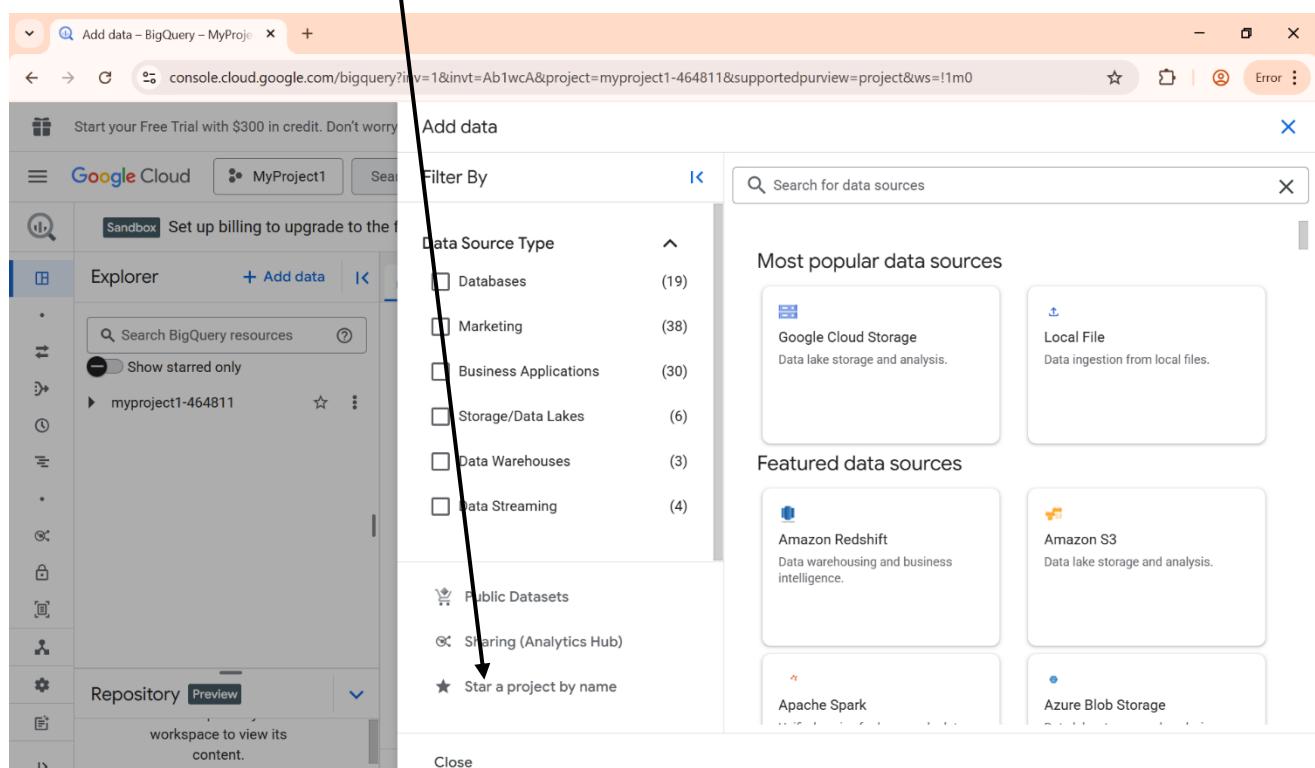
- SQL query
- Notebook
- Notebook with Spark
- Data canvas
- Pipeline

Try with sample data

Try the Google Trends Demo Query

This simple query generates the top search terms in the US from the Google Trends public dataset.

- Step 2: Select Star a project by name.



Add data

Filter By

- Data Source Type
  - Databases (19)
  - Marketing (38)
  - Business Applications (30)
  - Storage/Data Lakes (6)
  - Data Warehouses (3)
  - Data Streaming (4)
- Public Datasets
- Sharing (Analytics Hub)
- ★ Star a project by name

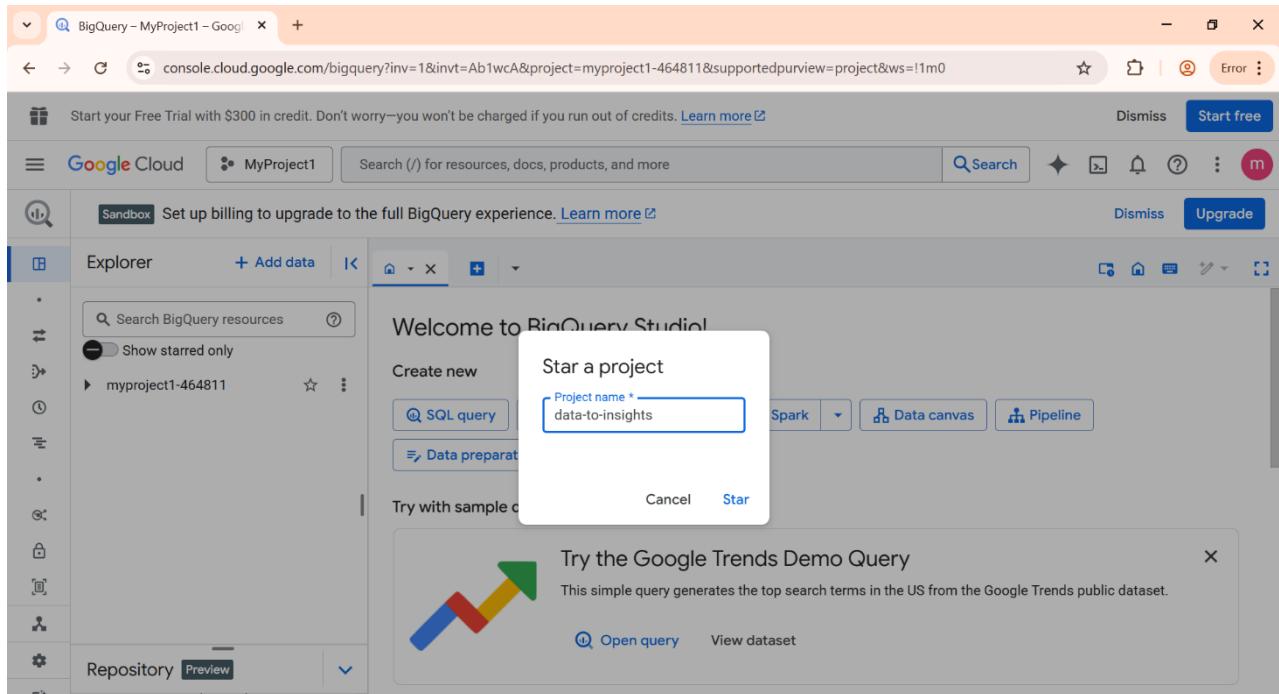
Most popular data sources

- Google Cloud Storage
- Local File

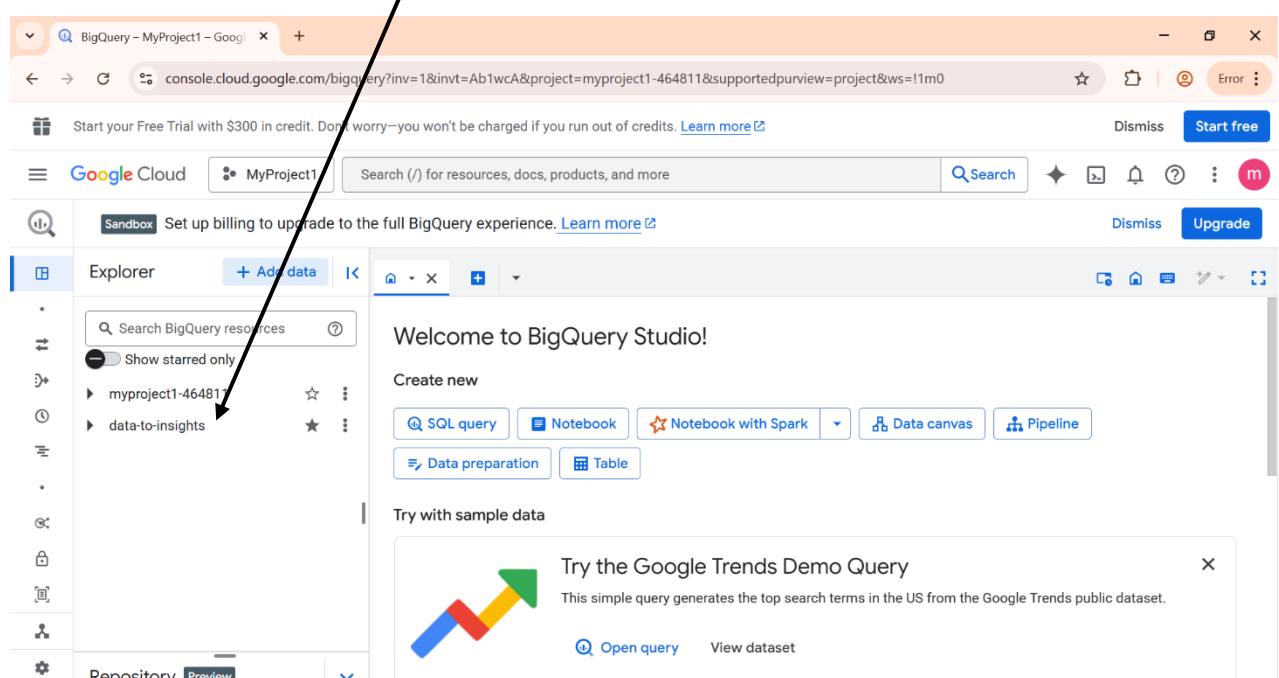
Featured data sources

- Amazon Redshift
- Amazon S3
- Apache Spark
- Azure Blob Storage

- **Step 3:** For Project name, enter "data-to-insights" and click Star.



- **Step 4:** In the left Panel, under Viewing pinned projects, you will see the data-to-insights project stared.

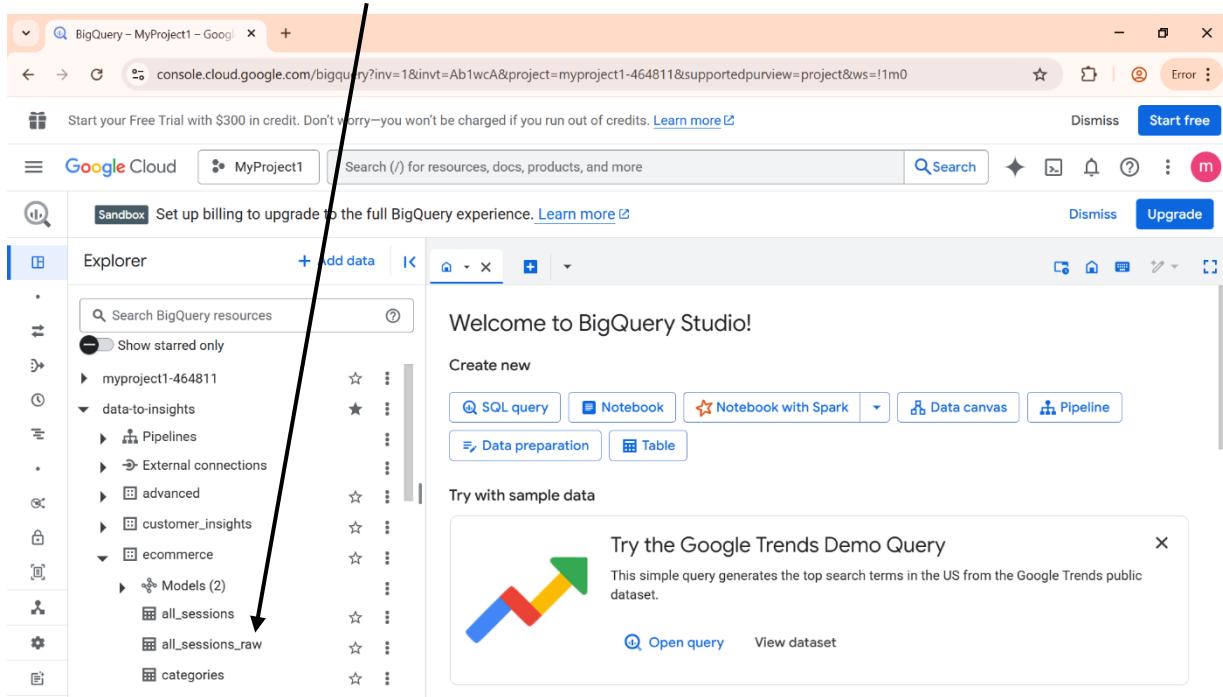


Note: The “data-to-insights” project is a publicly available dataset in Google BigQuery, commonly used in educational settings and sandbox environments to help users practice data analysis and SQL querying.

## Task 3: Explore e-commerce data

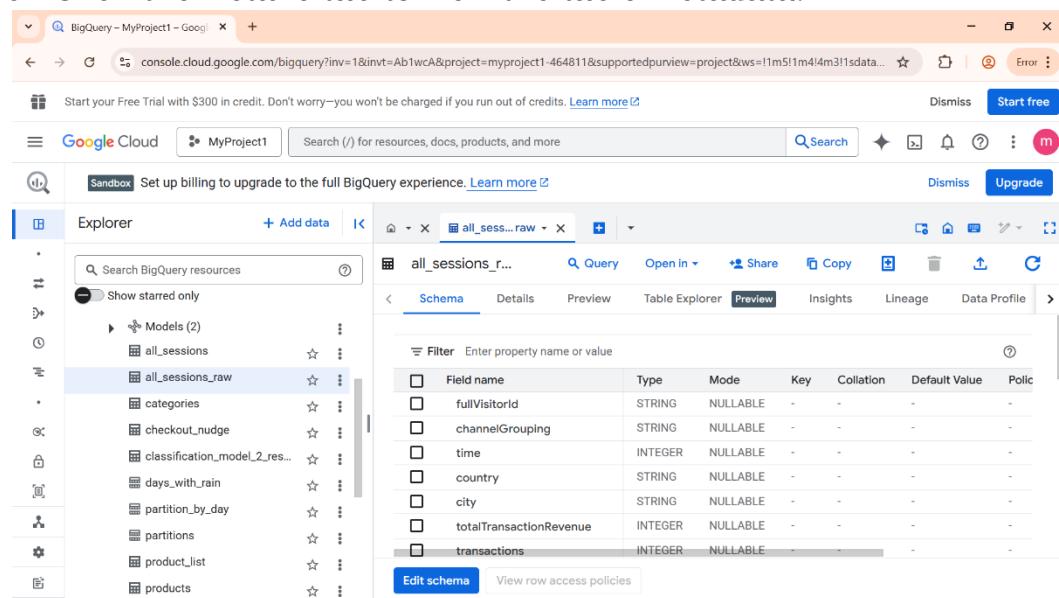
**Scenario:** Your data analyst team exported Google Analytics logs for an e-commerce website into BigQuery and created a new table of all the raw e-commerce visitor session data.

- **Step1:** Explore all\_sessions\_raw table data:
  - Expand **data-to-insights** project.
  - Expand **ecommerce**.
  - Click **all\_sessions\_raw**.



The screenshot shows the BigQuery Studio interface. On the left, the Explorer sidebar displays the 'data-to-insights' project, specifically the 'ecommerce' dataset, which contains the 'all\_sessions\_raw' table. A large black arrow points from the text 'Click all\_sessions\_raw.' towards this table entry in the sidebar.

- **Step 2:** In the right panel, a section opens that provides 3 views of the table data:
  - Schema tab: Field name, Type, Mode, and Description; the logical constraints used to organize the data.
  - Details tab: Table metadata.
  - Preview tab: Table preview.
  - Click the Details tab to view the table metadata.



The screenshot shows the BigQuery Studio interface with the 'all\_sessions\_raw' table selected. The right panel displays three tabs: Schema, Details, Preview, and Table Explorer. The Schema tab is currently active, showing a table of fields with their respective types and modes. The table includes fields like fullVisitorId (STRING, NULLABLE), channelGrouping (STRING, NULLABLE), time (INTEGER, NULLABLE), country (STRING, NULLABLE), city (STRING, NULLABLE), totalTransactionRevenue (INTEGER, NULLABLE), and transactions (INTEGER, NULLABLE).

Field name	Type	Mode	Key	Collation	Default Value	Policy
fullVisitorId	STRING	NULLABLE	-	-	-	-
channelGrouping	STRING	NULLABLE	-	-	-	-
time	INTEGER	NULLABLE	-	-	-	-
country	STRING	NULLABLE	-	-	-	-
city	STRING	NULLABLE	-	-	-	-
totalTransactionRevenue	INTEGER	NULLABLE	-	-	-	-
transactions	INTEGER	NULLABLE	-	-	-	-