



## Reading: Google Fiber datasets

---

By now, you are getting ready to take the next steps with your Course 2 end-of-course project. To work with the Google Fiber data, you will need to upload your data to the appropriate workspace. If you plan on using BigQuery or Dataflow, upload the files to your project space to JOIN them. Additionally, because this data is already clean, you can connect these datasets in Tableau directly and merge them there.

Your interviewers have provided three CSV files:

- [Market\\_1](#)
- [Market\\_2](#)
- [Market\\_3](#)

### Upload to BigQuery

First, navigate to your BigQuery console. Go to the BigQuery homepage or go directly to [the console](#).

# BigQuery

Cost-effective, serverless, multicloud enterprise data warehouse to power your data-driven innovation.

New customers get \$300 in free credits to spend on BigQuery. All customers get 10 GB storage and up to 1 TB queries free per month, not charged against their credits.

[Go to console](#)

[Contact sales](#)

- ✓ Run analytics at scale with 27% lower three-year TCO than [cloud data warehouse alternatives](#)
- ✓ Democratize insights with built-in business intelligence and machine learning
- ✓ Power business decisions from data across clouds with a flexible, multicloud analytics solution
- ✓ BigQuery is at the core of Google's unified [data cloud](#) to help you drive data transformation



VIDEO

Learn BigQuery in a minute, including how it works and common use cases

01:26

## BENEFITS

### Gain insights with real-time and predictive analytics

Query streaming data in real time and get up-to-date information on all your business processes. Predict business outcomes easily with built-in machine learning—without the need to move data.

### Protect your data and operate with trust

Rely on BigQuery's robust security, governance, and reliability controls that offer high availability and a 99.99% uptime SLA. Protect your data with encryption by default and customer-managed encryption keys.

### Break data silos and prevent lock-in

Embrace the ecosystem of partners, open data standards, and common industry data models to have the choice to work with data across any platform or environment.

Click on the + ADD DATA button in the Explorer menu pane to open the Add Data menu.

## Add data



### Source

Search for data sources

#### Popular sources



##### Local file

Upload a local file



##### Google Cloud Storage

Google object storage service



##### Connections to external data sources

Connection from BigQuery to an external data source

#### Additional sources

Viewing all 24 results.

Search for and star a project

Search for a BigQuery project and add it to the Explorer

Star a project by name

Add a BigQuery project to the Explorer by project name

Analytics Hub

Discover and subscribe to public, commercial or privately shared datasets

Google Drive

Google storage service

Amazon S3 - Data Transfer

Amazon object storage service, via the Data Transfer Service

Public Datasets

BigQuery public datasets from the Google Cloud Public Dataset Program

Cloud Bigtable

Highly-scalable NoSQL database

Informatica Data Loader

Load without code. 100% free covering over 35 sources and no capacity limits

Fivetran Data Pipelines

Automate data flows into BigQuery with near real-time source connectors

Pub/Sub Subscription

Create a Pub/Sub subscription to write data to an existing BigQuery table

Datastream

Real time CDC replication from relational databases such as MySQL, PostgreSQL, and Oracle

Amazon S3 - Omni

Amazon object storage service, via BigQuery Omni

CLOSE

From here, select Local file to upload the CSV or Google Cloud Storage to choose the sheet from your personal Drive. However you add the file, you will need to fill out the necessary fields in the Create Table menu. If you haven't already, the Create table menu will also prompt you to create a dataset to house this table.

### Create table

---

**Source**

Create table from \_\_\_\_\_

Select file \* \_\_\_\_\_

File format \_\_\_\_\_

**Destination**

Project \* \_\_\_\_\_

Dataset \* \_\_\_\_\_  
! Dataset is required

Table \* \_\_\_\_\_  
Unicode letters, marks, numbers, connectors, dashes or spaces allowed.

Table type \_\_\_\_\_

**Schema**

Auto detect

Edit as text

+

**Partition and cluster settings**

Partitioning \_\_\_\_\_

Clustering order \_\_\_\_\_  
Clustering order determines the sort order of the data. Clustering can be used on both partitioned and non-partitioned tables.

**Advanced options**

### Create dataset

---

Project ID \_\_\_\_\_

[CHANGE](#)

Dataset ID \* \_\_\_\_\_  
Letters, numbers, and underscores allowed

**Location type** ?  
 Multi-region  
Letting BigQuery select a region within a group of regions provides higher quota limits

Multi-region \* \_\_\_\_\_

**Default table expiration**  
 Enable table expiration ?

Default maximum table age \_\_\_\_\_

Days

**Advanced options**

Select CREATE NEW DATASET and name the dataset appropriately for this project. Leave the data location set to default. Once you have completed filling out this information, click Create Dataset.

Now, finish filling out the information for your table. Name your table appropriately for your project and select CSV under file type. Finally, select Auto detect for the schema. Once done, select Create Table. The new table should appear under your dataset in the Explorer pane momentarily.

The screenshot shows the BigQuery Explorer interface. The top navigation bar includes a magnifying glass icon, the word "Explorer", a "+ ADD DATA" button, and a back arrow icon. A search bar with the placeholder "Type to search" and a help icon is located below the navigation. The main content area displays a list of datasets and tables. A sidebar on the left contains various icons for different features like external connections, datasets, and users. The datasets listed are: "root-station-350516" (with "External connections" and "cyclistic" tables), "fiber" (with "market\_1", "market\_2", "market\_3", "mydataset", and three starred tables), "airy-shuttle-315515", "bigquery-public-data", and "steptratr". The "fiber" dataset and its sub-tables are highlighted with a light blue background, indicating they are the current focus.

From here, explore the schema, preview the data, and familiarize yourself with this table. Once you have uploaded this dataset, you will be ready to continue with your project!

---