

Dataprep: Qwik Start

GSP105



- Dataprep by Alteryx Designer Cloud offers intelligent, serverless data preparation for visual exploration, cleaning, and analysis at any scale.
- No infrastructure management required—use simple clicks for powerful data transformations without coding.
- This lab teaches importing datasets into Dataprep for immediate visual profiling and quality assessment.
- Practice correcting mismatched data types, applying transformations, and joining multiple datasets seamlessly.
- Master end-to-end data manipulation to prepare clean, analysis-ready datasets efficiently.

Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills Partner

What do you want to learn today?



5636

0



> Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik Start

Contents
Dataprep: Qwik Start

3. (Optional) You can list the active account name with this command:

gcloud auth list

4. Click Authorize.

Output:

```
ACTIVE: *
ACCOUNT: student-02-dc747dbe9199@qwiklabs.net
```

To set the active account, run:
\$ gcloud config set account `ACCOUNT`

5. (Optional) You can list the project ID with this command:

gcloud config list project

Output:

- Lab instructions and tasks 0/100
- GSP105
 - Overview
 - Setup and requirements
 - Task 1. Create a Cloud Storage bucket in your project
 - Task 2. Initialize Cloud Dataprep
 - Task 3. Create a flow
 - Task 4. Import datasets
 - Task 5. Prep the candidate file
 - Task 6. Wrangle the Contributions file and join it to the Candidates file
 - Task 7. Summary of data

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-87497481



Student Resources

Dataprep: Qwik Start - Qwiklabs Preview

Previous

Next

Dashboard – qwiklabs-gcp-00-8749748b7da9 x +

https://console.cloud.google.com/home/dashboard?project=qwiklabs-gcp-00-8749748b7da9&pli=1&cloudshell=true

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud qwiklabs-gcp-00-8749748b7da9 Search (/) for resources, docs, products, and more Search

CLOUD SHELL Terminal (qwiklabs-gcp-00-8749748b7da9) + Open Editor

Gemini CLI is available in Cloud Shell terminal! Type gemini to try it. [Learn more](#)

Don't show again Dismiss

Welcome to Cloud Shell! Type "help" to get started, or type "gemini" to try prompting with Gemini CLI.
Your Cloud Platform project in this session is set to **qwiklabs-gcp-00-8749748b7da9**.
Use `gcloud config set project [PROJECT_ID]` to change to a different project.
student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$ gcloud auth list
Credentialed Accounts

ACTIVE: *

ACCOUNT: student-02-dc747dbe9199@qwiklabs.net

To set the active account, run:
\$ gcloud config set account 'ACCOUNT'

student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$ gcloud config list project
[core]
project = qwiklabs-gcp-00-8749748b7da9

Your active configuration is: [cloudshell-18922]
student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$

9 21°C Sunny

Search

ENG IN 05:20 PM 08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills Partner

What do you want to learn today?



5636

[Home](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)Contents
Dataprep: Qwik Start

End Lab 01:27:14

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-87497481



Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)[Previous](#)

0/100

Lab instructions and tasks

GSP105

Overview

Setup and requirements

Task 1. Create a Cloud Storage bucket in your project

Task 2. Initialize Cloud Dataprep

Task 3. Create a flow

Task 4. Import datasets

Task 5. Prep the candidate file

Task 6. Wrangle the Contributions file and join it to the Candidates file

Task 7. Summary of data

Task 1. Create a Cloud Storage bucket in your project

1. In the Cloud Console, select **Navigation menu** (≡) > **Cloud Storage** > **Buckets**.
2. Click **Create bucket**.
3. In the **Create a bucket** dialog, **Name** the bucket a unique name. Leave other settings at their default value.
4. Uncheck **Enforce public access prevention on this bucket** for [Choose how to control access to objects](#).
5. Click **Create**.

You created your bucket. Remember the bucket name for later steps.

[Next](#)

Dashboard – qwiklabs-gcp-00-874... Create a bucket – Cloud Storage + https://console.cloud.google.com/storage/create-bucket?project=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud qwiklabs-gcp-00-8749748b7da9 Search (S) Search

Cloud Storage Create a bucket

Overview Buckets Monitoring Settings

Storage Intelligence Insights datasets Configuration

Marketplace Release Notes

9 21°C Sunny

Get Started

Pick a globally unique, permanent name. [Naming guidelines](#)

qwiklabs-gcp-00-8749748b7da9

Tip: Don't include any sensitive information

Labels (optional)

Continue

Choose where to store your data

Location: us (multiple regions in United States)
Location type: Multi-region

Choose how to store your data

Default storage class: Standard
Hierarchical namespace: Disabled
Anywhere Cache: Disabled

Choose how to control access to objects

Public access prevention: On
Access control: Uniform

Good to know

Location pricing

Storage rates vary depending on the storage class of your data and location of your bucket. [Pricing details](#)

Current configuration: Multi-region / Standard

Item	Cost
us (multiple regions in United States)	\$0.026 per GB-month
With default replication	\$0.020 per GB written

Estimate your monthly cost

ENG IN 05:21 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Create a bucket – Cloud Storage + https://console.cloud.google.com/storage/create-bucket?project=qwiklabs-gcp-00-8749748b7da9 | 1

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud Search (/) for resources, docs, products, and more Search

Cloud Storage Create a bucket

Name: qwiklabs-gcp-00-8749748b7da9

Choose where to store your data

Location: us (multiple regions in North America)

Location type: Multi-region

Choose how to store data

Default storage class: Standard

Hierarchical namespace: Disabled

Anywhere Cache: Disabled

Choose how to control access

Public access prevention: On

Access control: Uniform

Choose how to protect data

Soft delete policy: Default

Object versioning: Disabled

Bucket retention policy: Disabled

Object retention: Disabled

Encryption type: Google-managed

Processing... Cancel

Location pricing

Storage rates vary depending on the storage class of your data and location of your bucket. [Pricing details](#)

Cost

	Cost
Storage in United States	\$0.026 per GB-month
Storage outside United States	\$0.020 per GB written

Public access will be prevented

This bucket is set to prevent exposure of its data on the public internet.

Keep this setting enabled unless you have a use case that requires public access (such as static website hosting). You can change it now or later. [Learn more](#)

Enforce public access prevention on this bucket

Don't show this message again

Cancel Confirm



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills Partner

What do you want to learn today?



5636

[Home](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)

Dashboard



Catalog



Paths



Collections

Contents
Dataprep: Qwik Start

01:24:49

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-87497481



Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)[Previous](#)

50/100

Lab instructions and tasks

GSP105

Overview

Setup and requirements

Task 1. Create a Cloud Storage bucket in your project

Task 2. Initialize Cloud Dataprep

Task 3. Create a flow

Task 4. Import datasets

Task 5. Prep the candidate file

Task 6. Wrangle the Contributions file and join it to the Candidates file

Task 7. Summary of data

Task 2. Initialize Cloud Dataprep

1. Open **Cloud Shell** and run the following command:

```
gcloud beta services identity create --service=dataprep.googleapis.com
```

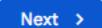


You should see a message saying the service identity was created.

2. In the Cloud console, go to the **Navigation menu**, click **View All Products** and under **Analytics** select **Alteryx Designer Cloud**.
3. Check to accept the Google Dataprep Terms of Service, then click **Accept**.
4. Check to authorize sharing your account information with Trifacta, then click **Agree and Continue**.
5. Click **Allow** to allow Trifacta to access project data.

Check complete. Points earned: 50. Message: Assessment Completed! Bucket count:

1. Bucket name(s): ['qwiklabs-gcp-00-8749748b7da9']



Dashboard – qwiklabs-gcp-00-8749748b7da9 – Cloud Storage

https://console.cloud.google.com/home/dashboard?project=qwiklabs-gcp-00-8749748b7da9&pli=1&cloudshell=true

Free AI Paraphrasing... Transcript of Case Study GCP-LAB Hydra datawarehouse

Google Cloud qwiklabs-gcp-00-8749748b7da9 Search (/) for resources, docs, products, and more Search

CLOUD SHELL Terminal (qwiklabs-gcp-00-8749748b7da9) Open Editor

Gemini CLI is available in Cloud Shell terminal! Type gemini to try it. [Learn more](#)

Welcome to Cloud Shell! Type "help" to get started, or type "gemini" to try prompting with Gemini CLI.
Your Cloud Platform project in this session is set to **qwiklabs-gcp-00-8749748b7da9**.
Use `gcloud config set project [PROJECT_ID]` to change to a different project.
student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$ gcloud auth list
Credentialated Accounts

ACTIVE: *
ACCOUNT: student-02-dc747dbe9199@qwiklabs.net

To set the active account, run:
\$ gcloud config set account 'ACCOUNT'

student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$ gcloud config list project
[core]
project = qwiklabs-gcp-00-8749748b7da9

Your active configuration is: [cloudshell-18922]
student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$ gcloud beta services identity create --service=dataprep.googleapis.com
Service identity created: service-78692667166@trifacta-gcloud-prod.iam.gserviceaccount.com
student_02_dc747dbe9199@cloudshell:~ (qwiklabs-gcp-00-8749748b7da9)\$

9 21°C Sunny

Search

ENG IN

05:23 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... qwiklabs-gcp-00-8749748b7da9 – Cloud +

https://console.cloud.google.com/storage/browser/qwiklabs-gcp-00-8749748b7da9;tab=objects?project=qwiklabs-gcp-00-8749748b7da9&prefi... | 1

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud

Cloud Hub > Bucket details

Cloud overview >

Solutions >

Recently visited >

Pinned products

Pin your top products here

Products

- Billing
- IAM & Admin
- Marketplace
- APIs & Services
- Vertex AI
- Compute Engine
- Kubernetes Engine

[View all products](#)

Search (/) for resources, docs, products, and more [Search](#)

Bucket details

Bucket: qwiklabs-gcp-00-8749748b7da9

Location: US (multiple regions in United States)

Storage class: Standard

Public access: Not public

Protection: Soft Delete

Objects Configuration Permissions Protection Lifecycle Observability New Inventory Reports Operations

Created bucket qwiklabs-gcp-00-8749748b7da9

Created bucket qwiklabs-gcp-00-8749748b7da9

https://console.cloud.google.com/products?project=qwiklabs-gcp-00-8749748b7da9

9 21°C Sunny ENG IN 05:23 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Solutions – qwiklabs-gcp-00-874... +

https://console.cloud.google.com/products?project=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud Solutions / Solutions

All products Jump Start Solutions Solution deployments

Categories Management

Compute • Storage Analytics • Networking Distributed Cloud Serverless Databases Observability Operations Security App Development

All products Explore products from Google partners at a glance

Management Control costs, establish identity and access management, and use APIs

Alteryx Designer Cloud Product · Visual data wrangling

BlueVector AI, LLC. Documentation · 100% dedicated to providing Google Cloud Platform solutions for Government...

Google Cloud Ready - BigQuery Partners Documentation · Alteryx Designer Cloud is a unified, workflow-based analytics platform with a...

Products & pages Alteryx Designer Cloud Product · Visual data wrangling

Documentation & tutorials

▼ Show more

Google Cloud Ready - AlloyDB Partners Documentation · These models are designed to be user-friendly, with always-on data reduction,...

Google Cloud Next 2023 wrap up Documentation · Looker Studio users now have native access to the Alteryx Designer Cloud for...

What's new for Looker and business intelligence at Next '23 Documentation · We are partnering with Alteryx to provide Looker Studio users with native...

Showing resource results for qwiklabs-gcp-00-8749748b7da9 only ?

Google Cloud Setup Set up and deploy a best-practice foundation

Created bucket qwiklabs-gcp-00-8749748b7da9 X

9 21°C Sunny

Search

ENG IN 05:23 PM 08-12-2025

Terms of Service / Service: dataprep

Welcome to Cloud Dataprep, a partner collaboration

This service is provided in collaboration with a third-party company, Trifacta. Check off the boxes below and tap “Accept” to agree to the Terms of Service and start using Dataprep.

Terms of Service

I have read and agree to the [Google Dataprep Terms of Service](#).

Accept

Created bucket qwiklabs-gcp-00-8749748b7da9 X



Dashboard – qwiklabs-gcp-00-874... Enable Data Sharing – Alteryx De +

https://console.cloud.google.com/dataprep/dataSharing?project=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud qwiklabs-gcp-00-8749748b7da9 Alteryx Designer Cloud X Search

Alteryx Designer Cloud / Home / Data sharing

Enable Data Sharing

⚠ Only project owners can consent to sharing account information with Trifacta.

Allow Trifacta to access project data?

By checking this box, you authorize Google LLC and its affiliates ("Google") to share your account information with Trifacta for purposes of sales attribution and technical support. You also represent that you have the authority to bind your company.

[?](#)

[Agree and Continue](#)

Created bucket qwiklabs-gcp-00-8749748b7da9



Dashboard – qwiklabs-gcp-00-874... Grant Access – Alteryx Designer +

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Google Cloud qwiklabs-gcp-00-8749748b7da9 Alteryx Designer Cloud Search

Alteryx Designer Cloud / Home / Grant access

Grant Access

Allow Trifacta to access project data?

So you can start transforming datasets, you hereby instruct Google to allow Trifacta, who provides the service Dataprep in collaboration with Google, to view and modify project data in Cloud Storage and BigQuery, run Dataflow jobs, and use all project service accounts.

Trifacta won't be able to access this data after you disable Dataprep. You can do this at any time in Project Settings.

[Learn more](#)

Allow

Dashboard – qwiklabs-gcp-00-874... Sign in - Google Accounts

https://accounts.google.com/signin/oauth/id?authuser=0&part=AJi8hAPEDtsB7e7-XxdhiBSF5yuk2kmu1qBSOtN0KGP7qamXf9NmCsRvuepl... 1

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Sign in with Google

 Sign in to Cloud Dataprep by Trifacta

student-02-dc747dbe9199@qwiklabs.net

Google will allow Cloud Dataprep by Trifacta to access this info about you

student 330fbea4
Name and profile picture

student-02-dc747dbe9199@qwiklabs.net
Email address

Review Cloud Dataprep by Trifacta's [Privacy Policy](#) and [Terms of Service](#) to understand how Cloud Dataprep by Trifacta will process and protect your data.

To make changes at any time, go to your [Google Account](#).

Learn more about [Sign in with Google](#).

Cancel Continue

Dashboard – qwiklabs-gcp-00-874... Sign in - Google Accounts

https://accounts.google.com/signin/oauth/consent?as=S-1859459910%3A1765194897589000&authuser=0&client_id=763568353131-3f82gu...

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

 Sign in with Google



Cloud Dataprep by Trifacta

wants to access your Google Account

 student-02-dc747dbe9199@qwiklabs.net

This will allow **Cloud Dataprep by Trifacta** to:

- View your Cloud Platform projects

Make sure you trust Cloud Dataprep by Trifacta

Review Cloud Dataprep by Trifacta's [Privacy Policy](#) and [Terms of Service](#) to understand how Cloud Dataprep by Trifacta will process and protect your data.

To make changes at any time, go to your [Google Account](#).

Learn how Google helps you [share data safely](#).

[Cancel](#) [Allow](#)

English (United States) ▾

Help Privacy Terms

Dashboard – qwiklabs-gcp-00-874... Accept License Agreement +

https://clouddataprep.com/accept-tos | 🏷

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

 Dataprep
by TRIFACTA

Cloud Dataprep by Trifacta is now **Generally Available** for all users and fully qualified for production use. This service is provided in collaboration with Google LLC. Google Cloud Platform support details can be found [here](#).

Please review and accept the Terms of Service below to start or continue use of Cloud Dataprep.

By checking this box, you agree to the [Trifacta Terms of Service](#).

[Deny & Log out](#) [Accept](#)



Dashboard – qwiklabs-gcp-00-874... Set Up: Choose storage location +

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

First time set up

 Dataprep
by TRIFACTA

Cloud Dataprep requires a location to save files you upload or create within Dataprep.

By default, the following storage location is used.

gs://dataprep-staging-6e84c374-2fc7-4e50-8444-20d51fd540d7/student-02... [Change](#)

Click **Continue** to complete setup for this project.

Project: qwiklabs-gcp-00-8749748b7da9 [Change project](#) [Continue](#)

9 21°C Sunny https://clouddataprep.com/setUp/storageLocation?url=%2F

Search

05:26 PM 08-12-2025 ENG IN



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Contents](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)

Dataprep: Qwik Start

01:20:43

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-87497481



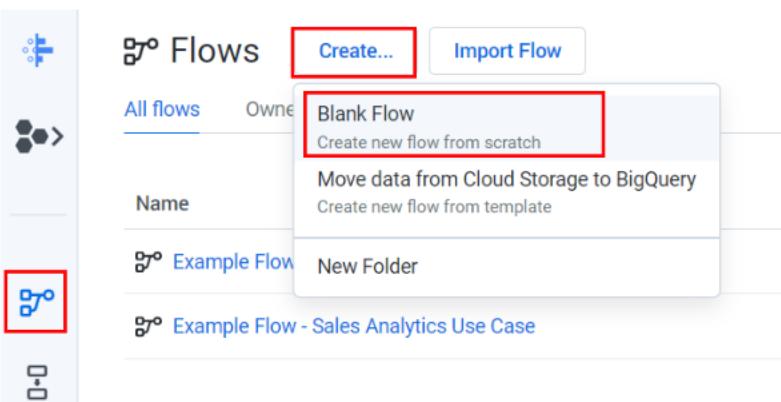
Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)

Task 3. Create a flow

Cloud Dataprep uses a `flow` workspace to access and manipulate datasets.

1. Click **Flows** icon, then the **Create** button, then select **Blank Flow**:



Lab instructions and tasks	50/100
GSP105	
Overview	
Setup and requirements	
Task 1. Create a Cloud Storage bucket in your project	
Task 2. Initialize Cloud Dataprep	
Task 3. Create a flow	
Task 4. Import datasets	
Task 5. Prep the candidate file	
Task 6. Wrangle the Contributions file and join it to the Candidates file	
Task 7. Summary of data	

[Previous](#)[Next](#)



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills Partner

What do you want to learn today?



5636

[Home](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)Contents
Dataprep: Qwik Start

End Lab 01:19:47

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-87497481



Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)

Example Flow

Example Flow - Sales Analytics Use Case



2. Click on **Untitled Flow**, then name and describe the flow. Since this lab uses 2016 data from the [United States Federal Elections Commission 2016](#), name the flow "FEC-2016", and then describe the flow as "United States Federal Elections Commission 2016".

3. Click **OK**.

The FEC-2016 flow page opens.

Task 4. Import datasets

In this section you import and add data to the FEC-2016 flow.

1. Click **Add Datasets**, then select the **Import Datasets** link.

[Previous](#)

100/100

Lab instructions and tasks	
GSP105	
Overview	
Setup and requirements	
Task 1. Create a Cloud Storage bucket in your project	
Task 2. Initialize Cloud Dataprep	
Task 3. Create a flow	
Task 4. Import datasets	
Task 5. Prep the candidate file	
Task 6. Wrangle the Contributions file and join it to the Candidates file	
Task 7. Summary of data	

[Next](#)

Dashboard – qwiklabs-gcp-00-874... Untitled Flow – 3 - Cloud Dataflow +

https://clouddataprep.com/flows/2579945?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Untitled Flow – 3

Dataset Recipe Output

Connect to your data Transform your data Export your data

100% Add datasets Share Schedule ...

9 21°C Sunny Search

ENG IN 05:28 PM 08-12-2025

```
graph LR; A[Dataset] --> B[Recipe]; B --> C[Output]
```





Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills Partner

What do you want to learn today?



5636

[Contents](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)

Dataflow

End Lab

01:18:13

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)

The FEC-2016 flow page opens.

Task 4. Import datasets

In this section you import and add data to the FEC-2016 flow.

1. Click **Add Datasets**, then select the **Import Datasets** link.
2. In the left menu pane, select **Cloud Storage** to import datasets from Cloud Storage, then click on the pencil to edit the file path.

Import Data and Add to Flow

Choose a file or folder

Upload Cloud Storage

Cloud Storage

Google Sheets

BigQuery

NAME	SIZE
dataprep-staging-51d7a1ba-e0c5-48e...	
qwiklabs-gcp-00-46cd4322783d	

[Previous](#)[Next](#)

Lab instructions and tasks	100/100
GSP105	
Overview	
Setup and requirements	
Task 1. Create a Cloud Storage bucket in your project	
Task 2. Initialize Cloud Dataprep	
Task 3. Create a flow	
Task 4. Import datasets	
Task 5. Prep the candidate file	
Task 6. Wrangle the Contributions file and join it to the Candidates file	
Task 7. Summary of data	



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > [Prepare Data for ML APIs on Google Cloud](#) > [Dataprep: Qwik Start](#)

Contents

Dataprep: Qwik Start



Dashboard

Catalog

Paths

Collections

01:17:53

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748I



Student Resources

[Dataprep: Qwik Start - Qwiklabs Preview](#)

In this section you import and add data to the FEC-2016 flow.

1. Click **Add Datasets**, then select the **Import Datasets** link.
2. In the left menu pane, select **Cloud Storage** to import datasets from Cloud Storage, then click on the pencil to edit the file path.

Import Data and Add to Flow

Choose a file or folder

Upload

Cloud Storage

NAME	SIZE
dataprep-staging-51d7a1ba-e0c5-48e...	
qwiklabs-gcp-00-46cd4322783d	

3. Type `gs://splz/gsp105` in the **Choose a file or folder** text box, then click **Go**.

You may have to widen the browser window to see the **Go** and **Cancel** buttons.

[Previous](#)[Next](#)ENG
IN05:29 PM
08-12-2025

Dashboard – qwiklabs-gcp-00-874... Home - Cloud Dataprep https://clouddataprep.com/home?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

Welcome, student 330fbea4!

Create a flow Import data

Resources Help center Tutorial video Community Documentation Training & certification Marketplace Contact support

Start from scratch Start from a template

+ Create a new flow Import data from Cloud Storage into BigQuery

Recently Updated

- FEC-2016 United States Federal Elections Commission 2016 Today at 5:29 PM
- Untitled Flow – 2 Today at 5:27 PM
- Untitled Flow Today at 5:27 PM
- Example Flow - Sales Analytics Use Case Today at 5:25 PM
- Example Flow - Event log data Today at 5:25 PM

View Flows • View Plans

Subscribe now! Your trial ends in 30 days. Subscribe now to continue with Dataprep!

View plans

9 21°C Sunny ENG IN 05:31 PM 08-12-2025

Search

Cloud Dataprep by TRIFACTA

Icons: Home, Search, Chat, Folder, Microsoft Edge, Firefox, Python, Power BI

Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep +

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Choose a file or folder

Cloud Storage Create Dataset with Parameters

Search... Show hidden

NAME	SIZE	LAST UPDATED
dataprep-staging-6e84c374-2fc7-4e5...		Today at 5:24 PM
qwiklabs-gcp-00-8749748b7da9		Today at 5:22 PM

0 New Datasets
Choose data to import.

Add to new flow

Cancel **Continue**

Import Data - Cloud Dataprep

Search

Cloud Dataprep

Cloud Storage

Google Sheets

BigQuery

New

Edit

21°C Sunny

ENG IN

05:32 PM 08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start[dataprep-staging-51d7a1ba-e0c5-48e...](#)[qwiklabs-gcp-00-46cd4322783d](#)

01:14:56

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

3. Type `gs://spl/gsp105` in the **Choose a file or folder** text box, then click **Go**.

You may have to widen the browser window to see the **Go** and **Cancel** buttons.

4. Click **us-fec/**.

5. Click the + icon next to `cn-2016.txt` to create a dataset shown in the right pane. Click on the title in the dataset in the right pane and rename it "Candidate Master 2016".

6. In the same way add the `itcont-2016-orig.txt` dataset, and rename it "Campaign Contributions 2016".

7. Both datasets are listed in the right pane; click **Import & Add to Flow**.

2 New Datasets

Clear All

< Previous

Next >

100/100

Lab instructions and tasks

GSP105

Overview

Setup and requirements

Task 1. Create a Cloud Storage bucket in your project

Task 2. Initialize Cloud Dataprep

Task 3. Create a flow

Task 4. Import datasets

Task 5. Prep the candidate file

Task 6. Wrangle the Contributions file and join



Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Choose a file or folder

Cloud Storage / spls / gsp105

Create Dataset with Parameters

Search... Show hidden

NAME	SIZE	LAST UPDATED
+	jobrun/	
+	temp/	
+	upload/	
+	us-fec/	

Parameterize Delete

0 New Datasets

Choose data to import.

Add to new flow

Cancel Continue

21°C Sunny

ENG IN 05:34 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep +

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Choose a file or folder

Upload / gsp105 / us-fec

Create Dataset with Parameters

Search... Show hidden

NAME	SIZE	LAST UPDATED	
+ NOPE-cn-2016.txt	717.52kB	11/04/2025	
+ c362ca52-5405-4761-9f10-666e...	694.48kB	11/04/2025	
+ c362ca52-5405-4761-9f10-666e...	82.81MB	11/04/2025	
+ ccl-2016.txt	257.29kB	11/04/2025	Parameterize Delete
+ ccl-2018.txt	186.54kB	11/04/2025	
+ cf39be63-0942-4889-8b3c-01d2...	694.48kB	11/04/2025	
+ cm-2016.txt	2.07MB	11/04/2025	

0 New Datasets

Choose data to import.

Add to new flow

Cancel Continue

Import Data - Cloud Dataprep

Search

Cloud Dataprep

GCP-LAB

Hydra

datawarehouse

Import Data

Search...

Upload

Cloud Storage

Google Sheets

BigQuery

New

Edit

21°C Sunny

ENG IN

05:34 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep +

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Choose a file or folder

Upload / gsp105 / us-fec

Create Dataset with Parameters

Search... / gsp105 / us-fec

Search... Show hidden

NAME SIZE LAST UPDATED

NAME	SIZE	LAST UPDATED
+ NOPE-cn-2016.txt	717.52kB	11/04/2025
+ c362ca52-5405-4761-9f10-666e...	694.48kB	11/04/2025
+ c362ca52-5405-4761-9f10-666e...	82.81MB	11/04/2025
+ ccl-2016.txt	257.29kB	11/04/2025
+ ccl-2018.txt	186.54kB	11/04/2025
+ cf39be63-0942-4889-8b3c-01d2...	694.48kB	11/04/2025
+ cm-2016.txt	2.07MB	11/04/2025

1 New Dataset Clear All

Candidate Master 2016

Add a Description

column2	column3	column4
C0008227	2016	2016
C0027466	2016	2016
C0042366	2016	2016
C0075820	2016	2016
C0099267	2016	2016

Edit settings

Add to new flow

Cancel Continue

9 21°C Sunny

Search

ENG IN 05:37 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep +

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Search... /

Choose a file or folder

Upload / gsp105 / us-fec

Create Dataset with Parameters

itcont-2016-orig.txt

Show hidden

NAME	SIZE	LAST UPDATED
7e3ffe8d-eb1c-4771-8ba1-e05e0...	82.81MB	11/04/2025
itcont-2016-orig.txt	3.55GB	11/04/2025

2 New Datasets

Clear All

itcont-2016-orig.txt

Add a Description

A ^B column2	X✓ column3	A ^B column4
C00004606	N	M4
C00004606	N	M4
C00004606	N	M4
C00452383	N	M4
C00452383	N	M4

Edit settings

Candidate Master 2016

Add a Description

A ^B column2	A ^B column3
HOAK00097	COX, JOHN R.
HOAL02087	ROBY, MARTHA
HOAL02095	JOHN, ROBERT E JR
HOAI 05049	CRAMFR RORFRT F

Add to new flow

Cancel Continue

Import Data - Cloud Dataprep

Search

Top Stories Operation Sagar... ENG IN 05:41 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Import Data - Cloud Dataprep +

https://clouddataprep.com/import/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

Import Data

Search... /

Choose a file or folder

Upload / gsp105 / us-fec

Create Dataset with Parameters

itcont-2016-orig.txt

Show hidden

NAME	SIZE	LAST UPDATED
7e3ffe8d-eb1c-4771-8ba1-e05e0...	82.81MB	11/04/2025
itcont-2016-orig.txt	3.55GB	11/04/2025

2 New Datasets

Campaign Contributions 20:

Add a Description

A ^B column2	X✓ column3	A ^B column4
C00004606	N	M4
C00004606	N	M4
C00004606	N	M4
C00452383	N	M4
C00452383	N	M4

Edit settings

Candidate Master 2016

Add a Description

A ^B column2	A ^B column3
HOAK00097	COX, JOHN R.
HOAL02087	ROBY, MARTHA
HOAL02095	JOHN, ROBERT E JR
HOAI 05049	CRAMFR RORFRT F

Add to new flow

Cancel Continue

Import Data - Cloud Dataprep

Search

Top Stories Operation Sagar... ENG IN 05:42 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Library - Cloud Dataprep https://clouddataprep.com/data?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

All Data Import Data Search Datasets /

Owned by me Shared with me All

Name	Owner	In flows	Source	Last updated
Campaign Contributions 2016	student 330fbea4	0	Cloud Storage	Today at 5:42 PM
Candidate Master 2016	student 330fbea4	0	Cloud Storage	Today at 5:42 PM
OPPORTUNITY OWNERS.csv	student 330fbea4	1	Cloud Storage	Today at 5:25 PM
SALES OPPORTUNITIES.csv	student 330fbea4	1	Cloud Storage	Today at 5:25 PM
USER EVENTS.csv	student 330fbea4	1	Cloud Storage	Today at 5:25 PM

Air: Poor Tomorrow

Search

05:42 PM 08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start

End Lab 01:03:55

You see both datasets listed as a flow.

Dashboard

Catalog

Paths

Collections

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



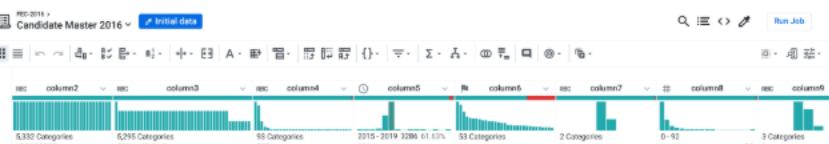
Student Resources

Task 5. Prep the candidate file

1. By default, the Candidate Master 2016 dataset is selected. In the right pane, click [Edit Recipe](#).



The Candidate Master 2016 Transformer page opens in the grid view.

[Previous](#)[Next](#)

GSP105
Overview
Setup and requirements
Task 1. Create a Cloud Storage bucket in your project
Task 2. Initialize Cloud Dataprep
Task 3. Create a flow
Task 4. Import datasets
Task 5. Prep the candidate file
Task 6. Wrangle the Contributions file and join



Dashboard – qwiklabs-gcp-00-874... FEC-2016 - Cloud Dataprep

https://clouddataprep.com/flows/2579945?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 United States Federal Elections Commission 2016

Dataset + Connect to your data

Add data from Library

Search... All (5) Imported (5) Reference (0)

	NAME	SOURCE	LAST UPDATED
<input type="checkbox"/>	Campaign Contributions 2016	Cloud Storage	Today at 5:42 ...
<input checked="" type="checkbox"/>	Candidate Master 2016	Cloud Storage	Today at 5:42 ...
<input type="checkbox"/>	OPPORTUNITY OWNERS.csv	Cloud Storage	Today at 5:25 ...
<input type="checkbox"/>	SALES OPPORTUNITIES.csv	Cloud Storage	Today at 5:25 ...

Import datasets Cancel Add

Import data from connections

BigQuery + Add connection

Upload data file

Air: Poor Tomorrow 9 ENG IN 05:43 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... FEC-2016 - Cloud Dataprep

https://clouddataprep.com/flows/2579945?recipe=14532067&tab=recipe&projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 United States Federal Elections Commission 2016

Dataset Recipe

Candidate Master 2016 Candidate Master 2016

Details

Candidate Master 2016

Edit recipe Add recipe

Recipe Data

Steps Preview

No Recipe Steps.
Edit the Recipe to wrangle.

Steps	0
Updated	Today at 5:46 PM
Created	Today at 5:46 PM

Trending videos The Hunger Ga...

Search

Cloud Dataprep icons

ENG IN 05:46 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?panel=recipe&projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Q ⌂ ⌂ <> Run

Recipe

...

column2 column3 column4 column5 column6 column7

	5,180 Categories	5,143 Categories	86 Categories	1990 - 2064	53 Categories	2 Categories
H0AK00097	COX, JOHN R.	REP	2014	AK	H	
H0AL02087	ROBY, MARTHA	REP	2016	AL	H	
H0AL02095	JOHN, ROBERT E JR	IND	2016	AL	H	
H0AL05049	CRAMER, ROBERT E "BUD" JR	DEM	2008	AL	H	
H0AL05163	BROOKS, MO	REP	2016	AL	H	
H0AL06088	COOKE, STANLEY KYLE	REP	2010	AL	H	
H0AL07086	SEWELL, TERRYCINA ANDREA	DEM	2016	AL	H	
H0AL07094	HILLIARD, EARL FREDERICK JR	DEM	2010	AL	H	
H0AL07177	CHAMBERLAIN, DON	REP	2012	AL	H	
H0AR01083	CRAWFORD, ERIC ALAN RICK	REP	2016	AR	H	
H0AR01091	GREGORY, JAMES CHRISTOPHER	DEM	2010	AR	H	
H0AR01109	CAUSEY, CHAD	DEM	2010	AR	H	
H0AR01125	SMITH, PRINCELLA D	REP	2010	AR	H	
H0AR02107	GRIFFIN, JOHN TIMOTHY	REP	2014	AR	H	
H0AR02131	ELLIOTT, JOYCE ANN	DEM	2010	AR	H	
H0AR03022	SKOCH, BERNARD KURT 'BERNIE'	REP	2010	AR	H	
H0AR03030	WHITAKER, DAVID JEFFREY	DEM	2010	AR	H	
H0AR03055	WOMACK, STEVE	REP	2016	AR	H	
H0AS00018	FALEOMAVAEGA, ENI	DEM	2014	AS	H	
H0AZ01184	FLAKE, JEFF MR.	REP	2012	AZ	H	
H0AZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H	
H0AZ01282	MUATA, STEVE	REP	2010	II		

Show / Hide data grid options 15 Columns 5,180 Rows 4 Data Types

9+ 19°C Clear

Search

ENG IN

05:47 PM 08-12-2025

The recipe is empty

Add New Step



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start

End Lab 00:58:52

The Transformer page is where you build your transformation recipe and see the results applied to the sample. When you are satisfied with what you see, execute the job against your dataset.

100/100

Lab instructions and tasks

Dashboard

Catalog

Paths

Collections

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

[Open Google Cloud console](#)

Username

student-02-dc747dbe9199e



Password

RcMw1vMgRQkJ

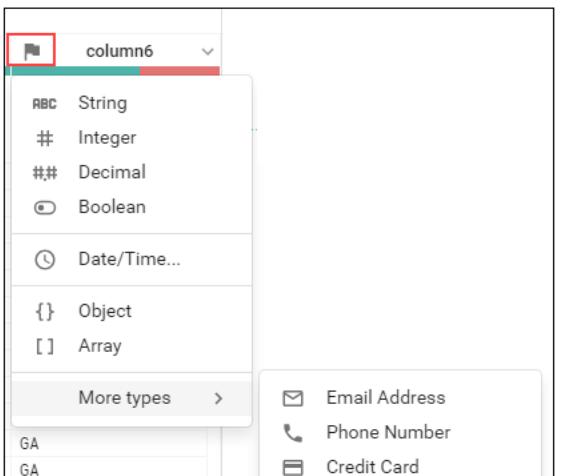


Project ID

qwiklabs-gcp-00-8749748l



Student Resources

[Previous](#)[Next](#)

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

<https://clouddataprep.com/data/2579945/14532067?panel=recipe&projectId=qwiklabs-gcp-00-8749748b7da9>

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

String Automatically update

Automatically update

	column2	column3	column4	column5	column6	column7
String	43 Categories	86 Categories	1990 - 2064	53 Categories	2 Categories	
Integer	X, JOHN R.	REP	2014	AK	H	
Decimal	BY, MARTHA	REP	2016	AL	H	
Boolean	HN, ROBERT E JR	IND	2016	AL	H	
Date/Time...	AMER, ROBERT E "BUD" JR	DEM	2008	AL	H	
	OOKS, MO	REP	2016	AL	H	
	OKE, STANLEY KYLE	REP	2010	AL	H	
	WELL, TERRYCINA ANDREA	DEM	2016	AL	H	
Object	LLIARD, EARL FREDERICK JR	DEM	2010	AL	H	
Array	AMBERLAIN, DON	REP	2012	AL	H	
	AWFORD, ERIC ALAN RICK	REP	2016	AR	H	
	EGORY, JAMES CHRISTOPHER	DEM	2010	AR	H	
More types	USEY, CHAD	DEM	2010	AR	H	
HOAR0125	SMITH, PRINCELLA D	REP	2010	AR	H	
HOAR02107	GRIFFIN, JOHN TIMOTHY	REP	2014	AR	H	
HOAR02131	ELLIOTT, JOYCE ANN	DEM	2010	AR	H	
HOAR03022	SKOCH, BERNARD KURT 'BERNIE'	REP	2010	AR	H	
HOAR03030	WHITAKER, DAVID JEFFREY	DEM	2010	AR	H	
HOAR03055	WOMACK, STEVE	REP	2016	AR	H	
HOAS00018	FALEOMAVAEGA, ENI	DEM	2014	AS	H	
HOAZ01184	FLAKE, JEFF MR.	REP	2012	AZ	H	
HOAZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H	
HOAZ01262	MUATA STEVE	REP	2010	

Show / Hide data grid options 15 Columns 5,180 Rows 4 Data Types

9+ 19°C Clear

Search

ENG IN

05:48 PM 08-12-2025

Recipe

The recipe is empty

Add New Step



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik Start

Contents

Dataprep: Qwik Start



00:59:23

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

GA	ZIP	Zip Code
ID		
IL	State	
IL		
IL	IP Address	
IL	URL	
IL		
IL	HTTP	HTTP Code

3. Notice also that when you click the name of the column, a **Details** panel opens on the right.

4. Click X in the top right of the Details panel to close the **Details** panel.

In the following steps you explore data in the grid view and apply transformation steps to your recipe.

1. Column5 provides data from 1990-2064. Widen column5 (like you would on a spreadsheet) to separate each year. Click to select the tallest bin, which represents the year 2016.

[Previous](#)[Next >](#)ENG
IN05:48 PM
08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik Start

Dashboard

Catalog

Paths

Collections

Contents

Dataprep: Qwik Start



00:57:37

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199e



Password

RcMw1vMgRQkJ



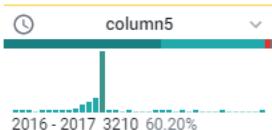
Project ID

qwiklabs-gcp-00-8749748l



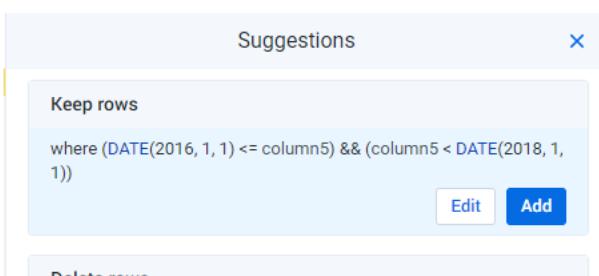
Student Resources

1. Column5 provides data from 1990-2064. Widen column5 (like you would on a spreadsheet) to separate each year. Click to select the tallest bin, which represents the year 2016.



This creates a step where these values are selected.

2. In the **Suggestions** panel on the right, in the **Keep rows** section, click **Add** to add this step to your recipe.

[Previous](#)[Next](#)

19°C

Clear



Search

ENG
IN05:49 PM
08-12-2025

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

<https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9>

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Details

column5
2008
Show more values...

Distribution

Patterns
{yyyy} 5,178
Show pattern details...

Show additional details >

Suggestions

Keep rows
where (DATE(2016, 1, 1) <= column5) && (column5 < DATE(2017, 1, 1))

Delete rows
where (DATE(2016, 1, 1) <= column5) && (column5 < DATE(2017, 1, 1))

9+ 19°C Clear Show / Hide data grid options 15 Columns 5,180 Rows 4 Data Types

Search

ENG IN 05:52 PM 08-12-2025

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	A ^B _C	column5	A ^B _C	column6	A ^B _C	column7
5,180 Categories	HOAK00097	5,143 Categories	COX, JOHN R.	86 Categories	REP	REP	1990 - 2064	AK	53 Categories	H	2 Categories
HOAL02087	ROBY, MARTHA	JOHN, ROBERT E JR	IND	2014	2016	AL	2016	AL	AL	H	
HOAL02095	CRAMER, ROBERT E "BUD" JR	BROOKS, MO	DEM	2016	2008	AL	2008	AL	AL	H	
HOAL05163	COOKE, STANLEY KYLE	SEWELL, TERRYCINA ANDREA	REP	2010	2010	AL	2010	AL	AL	H	
HOAL06088	HOAL07086	HILLIARD, EARL FREDERICK JR	DEM	2016	2010	AL	2010	AL	AL	H	
HOAL07094	CHAMBERLAIN, DON	CAUSEY, CHAD	DEM	2010	2012	AR	2010	AR	AR	H	
HOAL07177	CRAWFORD, ERIC ALAN RICK	SMITH, PRINCELLA D	REP	2016	2016	AR	2012	AR	AR	H	
HOAR01083	GREGORY, JAMES CHRISTOPHER	GRiffin, JOHN TIMOTHY	REP	2010	2010	AR	2016	AR	AR	H	
HOAR01091	HOAR01109	ELLIOTT, JOYCE ANN	DEM	2010	2014	AR	2010	AR	AR	H	
HOAR01125	SKOCH, BERNARD KURT 'BERNIE'	WHITAKER, DAVID JEFFREY	REP	2010	2010	AR	2010	AR	AR	H	
HOAR02107	HOAR02131	WOMACK, STEVE	DEM	2010	2010	AR	2010	AR	AR	H	
HOAR03022	FALEOMAVAEGA, ENI	FLAKE, JEFF MR.	REP	2016	2016	AS	2010	AS	AR	H	
HOAR03030	HOAZ01184	GOSAR, PAUL DR.	DEM	2014	2012	AZ	2014	AZ	AZ	H	
HOAR03055	HOAZ01259	HOAZ01262	HOAZ01262	2012	2016	HOAZ01262	2016	HOAZ01262	HOAZ01262	HOAZ01262	HOAZ01262



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start

2. In the **Suggestions** panel on the right, in the **Keep rows** section, click **Add** to add this step to your recipe.

Suggestions

Keep rows

```
where (DATE(2016, 1) <= column5) && (column5 < DATE(2018, 1, 1))
```

Edit **Add**

Delete rows

```
where (DATE(2016, 1) <= column5) && (column5 < DATE(2018, 1, 1))
```

Set

```
Set column5 to IF((DATE(2016, 1) <= column5) && (column5 < DATE(2018, 1, 1)), NULL(), $col)
```

Create a new column

```
(DATE(2016, 1) <= column5) && (column5 < DATE(2018, 1, 1))
```

Lab instructions and tasks

100/100

GSP105

Overview

Setup and requirements

Task 1. Create a Cloud Storage bucket in your project

Task 2. Initialize Cloud Dataprep

Task 3. Create a flow

Task 4. Import datasets

Task 5. Prep the candidate file

Task 6. Wrangle the Contributions file and join

< Previous

Next >

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Q ⌂ ⌂ <> Run

New Step Recipe

...

1 Keep rows where (DATE(2016, 1, 1) <= column5) && (column5 < DATE(2017, 1, 1))

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	⌚	column5	🚩	column6	A ^B _C	column7
3,064 Categories	HOAL02087	3,048 Categories	ROBY, MARTHA	79 Categories	REP	2016-2016	2016	53 Categories	AL	H	
HOAL02095	JOHN, ROBERT E JR	IND			2016		2016	AR			
HOAL05163	BROOKS, MO	REP			2016		2016	AR			
HOAL07086	SEWELL, TERRYCINA ANDREA	DEM			2016		2016	AR			
HOAR01083	CRAWFORD, ERIC ALAN RICK	REP			2016		2016	AZ			
HOAR03055	WOMACK, STEVE	REP			2016		2016	CA			
HOAZ01259	GOSAR, PAUL DR.	REP			2016		2016	CA			
HOCA03078	BERA, AMERISH	DEM			2016		2016	CA			
HOCA10073	DESAULNIER, MARK	DEM			2016		2016	CA			
HOCA10149	GARAMENDI, JOHN	DEM			2016		2016	CA			
HOCA11337	AMADOR, ANTONIO C	REP			2016		2016	CA			
HOCA15148	HONDA, MIKE	DEM			2016		2016	CA			
HOCA19173	DENHAM, JEFF	REP			2016		2016	CA			
HOCA22086	UEBERSAX, JOHN	NPA			2016		2016	CA			
HOCA27085	SCHIFF, ADAM	DEM			2016		2016	CA			
HOCA32101	CHU, JUDY	DEM			2016		2016	CA			
HOCA33117	BASS, KAREN	DEM			2016		2016	CA			
HOCA33158	GRIFFIN, MORRIS F	DEM			2016		2016	CA			
HOCA48024	ISSA, DARRELL	REP			2016		2016	CA			
HOCA49055	DAVIS, SUSAN	DEM			2016		2016	CT			
HOCT02140	NOVAK, DARIA IRENE	REP			2016		2016	CT			
HOCT02070	DELARIO ROSA L	DEM			2016		2016	CT			

Show / Hide data grid options 15 Columns 3,064 Rows 4 Data Types

9+ 19°C Clear Search ENG IN 05:52 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Q ⌂ ⌂ < > Run

New Step Recipe

...

1 Keep rows where (DATE(2016, 1, 1) <= column5) && (column5 < DATE(2017, 1, 1))

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	⌚	column5	🚩	column6	A ^B _C	column7
3,064 Categories	HOAL02087	3,048 Categories	ROBY, MARTHA	79 Categories	REP	2016-2016	2016	53 Categories	AL	2 Categories	H
HOAL02095	JOHN, ROBERT E JR	IND			2016		2016	AR			H
HOAL05163	BROOKS, MO	REP			2016		2016	AR			H
HOAL07086	SEWELL, TERRYCINA ANDREA	DEM			2016		2016	AZ			H
HOAR01083	CRAWFORD, ERIC ALAN RICK	REP			2016		2016	CA			H
HOAR03055	WOMACK, STEVE	REP			2016		2016	CA			H
HOAZ01259	GOSAR, PAUL DR.	REP			2016		2016	CA			H
HOCA03078	BERA, AMERISH	DEM			2016		2016	CA			H
HOCA10073	DESAULNIER, MARK	DEM			2016		2016	CA			H
HOCA10149	GARAMENDI, JOHN	DEM			2016		2016	CA			H
HOCA11337	AMADOR, ANTONIO C	REP			2016		2016	CA			H
HOCA15148	HONDA, MIKE	DEM			2016		2016	CA			H
HOCA19173	DENHAM, JEFF	REP			2016		2016	CA			H
HOCA22086	UEBERSAX, JOHN	NPA			2016		2016	CA			H
HOCA27085	SCHIFF, ADAM	DEM			2016		2016	CA			H
HOCA32101	CHU, JUDY	DEM			2016		2016	CA			H
HOCA33117	BASS, KAREN	DEM			2016		2016	CA			H
HOCA33158	GRIFFIN, MORRIS F	DEM			2016		2016	CA			H
HOCA48024	ISSA, DARRELL	REP			2016		2016	CA			H
HOCA49055	DAVIS, SUSAN	DEM			2016		2016	CA			H
HOCT02140	NOVAK, DARIA IRENE	REP			2016		2016	CT			H
HOCT02070	DELARIO, ROSA L	DEM			2016		2016	CT			H

Show / Hide data grid options 15 Columns 3,064 Rows 4 Data Types

9+ 19°C Clear Search ENG IN 05:53 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Preview

	column2	column3	column4	column5	column6	column7
3,064 Categories	3,048 Categories	79 Categories	2016 - 2016	53 Categories	2 Categories	
H0AL02087	ROBY, MARTHA	REP	2016	AL	H	
H0AL02095	JOHN, ROBERT E JR	IND	2016	AL	H	
H0AL05163	BROOKS, MO	REP	2016	AL	H	
H0AL07086	SEWELL, TERRYCINA ANDREA	DEM	2016	AL	H	
H0AR01083	CRAWFORD, ERIC ALAN RICK	REP	2016	AR	H	
H0AR03055	WOMACK, STEVE	REP	2016	AR	H	
H0AZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H	
H0CA03078	BERA, AMERISH	DEM	2016	CA	H	
H0CA10073	DESAULNIER, MARK	DEM	2016	CA	H	
H0CA10149	GARAMENDI, JOHN	DEM	2016	CA	H	
H0CA11337	AMADOR, ANTONIO C	REP	2016	CA	H	
H0CA15148	HONDA, MIKE	DEM	2016	CA	H	
H0CA19173	DENHAM, JEFF	REP	2016	CA	H	
H0CA22086	UEBERSAX, JOHN	NPA	2016	CA	H	
H0CA27085	SCHIFF, ADAM	DEM	2016	CA	H	
H0CA32101	CHU, JUDY	DEM	2016	CA	H	
H0CA33117	BASS, KAREN	DEM	2016	CA	H	
H0CA33158	GRIFFIN, MORRIS F	DEM	2016	CA	H	
H0CA48024	ISSA, DARRELL	REP	2016	CA	H	
H0CA49055	DAVIS, SUSAN	DEM	2016	CA	H	
H0CT02140	NOVAK, DARIA IRENE	REP	2016	CT	H	
H0CT02072	DELARIO, ROSA L	DEM	2016	CT		

Show / Hide data grid options 15 Columns 3,064 Rows 4 Data Types Show only affected Rows

Suggestions

Delete rows
where ISMISMATCHED(column6, ['State']) Edit Add

Keep rows
where ISMISMATCHED(column6, ['State'])

Create a new column
ISMISMATCHED(column6, ['State'])

Set
Set column6 to IFMISMATCHED(\$col, ['State'], NULL())

9+ 19°C Clear ENG IN 05:54 PM 08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start

00:51:48

End Lab

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199e



Password

RcMw1vMgRQkJ

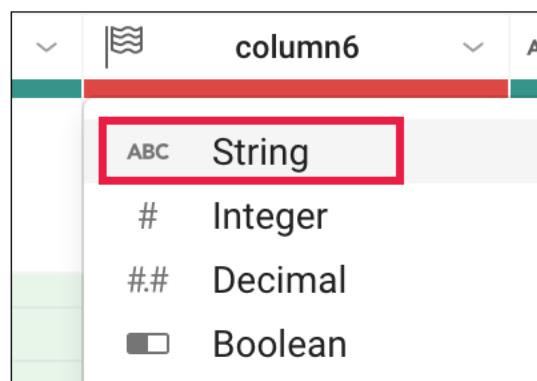


Project ID

qwiklabs-gcp-00-8749748l



Student Resources



Scroll down to the bottom (highlighted in red) find the mismatched values and notice how most of these records have the value "P" in column7, and "US" in column6. The mismatch occurs because column6 is marked as a "State" column (indicated by the flag icon), but there are non-state (such as "US") values.

4. To correct the mismatch, click X in the top of the Suggestions panel to cancel the transformation, then click on the flag icon in Column6 and change it to a "String" column.

Lab instructions and tasks	
GSP105	
Overview	
Setup and requirements	
Task 1. Create a Cloud Storage bucket in your project	
Task 2. Initialize Cloud Dataprep	
Task 3. Create a flow	
Task 4. Import datasets	
Task 5. Prep the candidate file	
Task 6. Wrangle the Contributions file and join	

[Previous](#)[Next >](#)

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Automatically update

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	⌚	column5	A ^B _C	column6	A ^B _C	column7	1 ² ₃	column8	A ^B _C	column9	A ^B _C	column10	
3,064 Categories	3,048 Categories	79 Categories	2016 - 2016	53 Categories	2 Categories	0 - 92	02	I	C	N	05	I	07	I	01	I	C	P
H0AL02087	ROBY, MARTHA	REP	2016	AL	H													
H0AL02095	JOHN, ROBERT E JR	IND	2016	AL	H													
H0AL05163	BROOKS, MO	REP	2016	AL	H													
H0AL07086	SEWELL, TERRYCINA ANDREA	DEM	2016	AL	H													
H0AR01083	CRAWFORD, ERIC ALAN RICK	REP	2016	AR	H													
H0AR03055	WOMACK, STEVE	REP	2016	AR	H													
H0AZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H													
H0CA03078	BERA, AMERISH	DEM	2016	CA	H													
H0CA10073	DESAULNIER, MARK	DEM	2016	CA	H													
H0CA10149	GARAMENDI, JOHN	DEM	2016	CA	H													
H0CA11337	AMADOR, ANTONIO C	REP	2016	CA	H													
H0CA15148	HONDA, MIKE	DEM	2016	CA	H													
H0CA19173	DENHAM, JEFF	REP	2016	CA	H													
H0CA22086	UEBERSAX, JOHN	NPA	2016	CA	H													
H0CA27085	SCHIFF, ADAM	DEM	2016	CA	H													
H0CA32101	CHU, JUDY	DEM	2016	CA	H													
H0CA33117	BASS, KAREN	DEM	2016	CA	H													
H0CA33158	GRIFFIN, MORRIS F	DEM	2016	CA	H													
H0CA48024	ISSA, DARRELL	REP	2016	CA	H													
H0CA49055	DAVIS, SUSAN	DEM	2016	CA	H													
H0CT02140	NOVAK, DARIA IRENE	REP	2016	CT	H													
H0CT02272	DELARIO ROSA L	DEM	2016	CT	H													

Show / Hide data grid options 15 Columns 3,064 Rows 4 Data Types

9+ 19°C Clear Search ENG IN 05:55 PM 08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

FEC-2016

Candidate Master 2016

Initial data

Q □ □ □ < > ↗

Run

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	A	column5	A ^B _C	column6	A ^B _C	column7	1 ² ₃	column8	A ^B _C	column9	A ^B _C	column
3,064 Categories	HOAL02087	3,048 Categories	ROBY, MARTHA	79 Categories	REP	JOHN, ROBERT E JR	2016 - 2016	Automatically update	String	Categories	1 - 92	0 - 92	3 Categories	3 Categories	3 Categories	3 Categories	
HOAL02095	HOAL05163	HOAL07086	BROOKS, MO	SEWELL, TERRYCINA ANDREA	IND	REP	2016	Integer	02	I	02	C	02	C	N	C	
HOAR01083	HOAR03055	HOAZ01259	CRAWFORD, ERIC ALAN RICK	WOMACK, STEVE	DEM	REP	2016	Decimal	05	I	01	I	07	I	C	C	
HOCA03078	HOCA10073	HOCA10149	BERA, AMERISH	GARAMENDI, JOHN	DEM	REP	2016	Boolean	07	I	03	I	01	I	C	P	
HOCA11337	HOCA15148	HOCA19173	DESAULNIER, MARK	GARAMENDI, JOHN	DEM	REP	2016	Date/Time...	03	I	04	I	07	I	C	C	
HOCA22086	HOCA27085	HOCA32101	AMADOR, ANTONIO C	HONDA, MIKE	DEM	DEM	2016	Object	11	I	09	C	17	I	C	C	
HOCA33117	HOCA33158	HOCA48024	DENHAM, JEFF	SCHIFF, ADAM	REP	DEM	2016	Array	03	I	09	C	10	I	C	C	
HOCA49055	HOCT02140	HOCT02072	UEBERSAX, JOHN	CHU, JUDY	NPA	DEM	2016	More types >	CA	H	CA	H	24	0	N	C	
NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	GRiffin, MORRIS F	BASS, KAREN	DEM	DEM	2016	CA	H	CA	H	28	I	27	I	C	
NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	ISSA, DARRELL	DAVIS, SUSAN	REP	DEM	2016	CA	H	CA	H	49	I	53	I	C	
NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	REP	DEM	2016	CA	H	CA	H	53	I	02	C	C	
NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	NOVAK, DARIA IRENE	REP	DEM	2016	CT	H	CT	H	02	C	33	T	C	

Show / Hide data grid options

15 Columns 3,064 Rows 4 Data Types





Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

FEC-2016

Candidate Master 2016

Initial data

Q □ □ □ < > ↗

Run

A ^B _C	column2	A ^B _C	column3	A ^B _C	column4	⌚	column5	A ^B _C	column6	A ^B _C	column7	1 ² ₃	column8	A ^B _C	column9	A ^B _C	column
3,064 Categories	HOAL02087	3,048 Categories	ROBY, MARTHA	79 Categories	REP	2016-2016	57 Categories	AL	H	2 Categories	0 - 92	02	I	3 Categories	C	3 Categories	N
HOAL02095	JOHN, ROBERT E JR	IND	2016	AL	H	02	C	05	I	07	I	01	I	03	I	04	P
HOAL05163	BROOKS, MO	REP	2016	AL	H	02	C	05	I	07	I	01	I	03	I	04	C
HOAL07086	SEWELL, TERRYCINA ANDREA	DEM	2016	AL	H	02	C	05	I	07	I	01	I	03	I	04	C
HOAR01083	CRAWFORD, ERIC ALAN RICK	REP	2016	AR	H	02	C	05	I	07	I	01	I	03	I	04	C
HOAR03055	WOMACK, STEVE	REP	2016	AR	H	02	C	05	I	07	I	01	I	03	I	04	C
HOAZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H	02	C	05	I	07	I	01	I	03	I	04	P
HOCA03078	BERA, AMERISH	DEM	2016	CA	H	02	C	05	I	07	I	11	I	03	I	04	C
HOCA10073	DESAULNIER, MARK	DEM	2016	CA	H	02	C	05	I	07	I	01	I	03	I	04	C
HOCA10149	GARAMENDI, JOHN	DEM	2016	CA	H	02	C	05	I	07	I	01	I	03	I	04	C
HOCA11337	AMADOR, ANTONIO C	REP	2016	CA	H	02	C	05	I	07	I	01	I	03	I	04	C
HOCA15148	HONDA, MIKE	DEM	2016	CA	H	02	C	05	I	07	I	01	I	03	I	04	C
HOCA19173	DENHAM, JEFF	REP	2016	CA	H	02	C	05	I	07	I	10	I	03	I	04	C
HOCA22086	UEBERSAX, JOHN	NPA	2016	CA	H	02	C	05	I	07	I	24	0	03	I	04	N
HOCA27085	SCHIFF, ADAM	DEM	2016	CA	H	02	C	05	I	07	I	28	I	03	I	04	C
HOCA32101	CHU, JUDY	DEM	2016	CA	H	02	C	05	I	07	I	27	I	03	I	04	C
HOCA33117	BASS, KAREN	DEM	2016	CA	H	02	C	05	I	07	I	37	I	03	I	04	C
HOCA33158	GRIFFIN, MORRIS F	DEM	2016	CA	H	02	C	05	I	07	I	44	0	03	I	04	N
HOCA48024	ISSA, DARRELL	REP	2016	CA	H	02	C	05	I	07	I	49	I	03	I	04	C
HOCA49055	DAVIS, SUSAN	DEM	2016	CA	H	02	C	05	I	07	I	53	I	03	I	04	C
HOCT02140	NOVAK, DARIA IRENE	REP	2016	CT	H	02	C	05	I	07	I	62	C	03	I	04	C
HOCT02272	DELARIO ROSA L	DEM	2016	CT	H	02	C	05	I	07	I	63	T	03	I	04	C

Show / Hide data grid options

15 Columns

3,064 Rows

4 Data Types



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik Start

Contents

Dataprep: Qwik Start

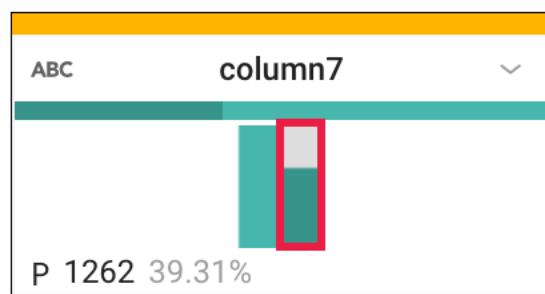


00:51:05

Boolean

There is no longer a mismatch and the column marker is now green.

5. Filter on just the presidential candidates, which are those records that have the value "P" in column7. In the histogram for column7, hover over the two bins to see which is "H" and which is "P". Click the "P" bin.



6. In the right Suggestions panel, click **Add** to accept the step to the recipe.

GSP105
Overview
Setup and requirements
Task 1. Create a Cloud Storage bucket in your project
Task 2. Initialize Cloud Dataprep
Task 3. Create a flow
Task 4. Import datasets
Task 5. Prep the candidate file
Task 6. Wrangle the Contributions file and join

< Previous

Next >

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

<https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9>

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Preview

	column2	column3	column4	column5	column6	column7
3,064 Categories	3,048 Categories	79 Categories	2016 - 2016	57 Categories	P 1128	36.81%
H0AL02087	ROBY, MARTHA	REP	2016	AL	H	
H0AL02095	JOHN, ROBERT E JR	IND	2016	AL	H	
H0AL05163	BROOKS, MO	REP	2016	AL	H	
H0AL07086	SEWELL, TERRYCINA ANDREA	DEM	2016	AL	H	
H0AR01083	CRAWFORD, ERIC ALAN RICK	REP	2016	AR	H	
H0AR03055	WOMACK, STEVE	REP	2016	AR	H	
H0AZ01259	GOSAR, PAUL DR.	REP	2016	AZ	H	
H0CA03078	BERA, AMERISH	DEM	2016	CA	H	
H0CA10073	DESAULNIER, MARK	DEM	2016	CA	H	
H0CA10149	GARAMENDI, JOHN	DEM	2016	CA	H	
H0CA11337	AMADOR, ANTONIO C	REP	2016	CA	H	
H0CA15148	HONDA, MIKE	DEM	2016	CA	H	
H0CA19173	DENHAM, JEFF	REP	2016	CA	H	
H0CA22086	UEBERSAX, JOHN	NPA	2016	CA	H	
H0CA27085	SCHIFF, ADAM	DEM	2016	CA	H	
H0CA32101	CHU, JUDY	DEM	2016	CA	H	
H0CA33117	BASS, KAREN	DEM	2016	CA	H	
H0CA33158	GRIFFIN, MORRIS F	DEM	2016	CA	H	
H0CA48024	ISSA, DARRELL	REP	2016	CA	H	
H0CA49055	DAVIS, SUSAN	DEM	2016	CA	H	
H0CT02140	NOVAK, DARIA IRENE	REP	2016	CT	H	
H0CT02070	DELARIO ROSA L	DEM	2016	CT	H	

Show / Hide data grid options 15 Columns 3,064 Rows 4 Data Types Show only affected Rows

Suggestions

Keep rows
where column7 == 'P'
[Edit](#) [Add](#)

Delete rows
where column7 == 'P'

Set
Set column7 to IF(column7 == 'P', NULL(), \$col)

Create a new column
column7 == 'P'

Deduplicate rows
where every value is the same

9+ 19°C Clear ENG IN 05:56 PM 08-12-2025

Dashboard – qwiklabs-gcp-00-874... Candidate Master 2016 - Cloud +

https://clouddataprep.com/data/2579945/14532067?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasing... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Candidate Master 2016 Initial data

Q ⌂ ⌂ <> ↗ Run

Initial data

column2 column3 column4 column5 column6 column7 column8 column9

1,128 Categories 1,126 Categories 66 Categories 2016 - 2016 1 Category 1 Category 0 - 0 2 Categories

P00000679 CARROLL, JERRY LEON NNE 2016 US P 00 0

P00001099 MARCINEK, ALOYSIUS R IND 2016 US P 00 0

P00003160 WELLS, THOMAS BAXTER OTH 2016 US P 00 0

P00003244 WINTERBOTTOM, THOMAS FRANCIS DEM 2016 US P 00 0

P00003392 CLINTON, HILLARY RODHAM / TIMOTHY MICHAEL KAINEN DEM 2016 US P 00 0

P00003640 LANCE-COUNCIL, TEMPERANCE ALESHA OTH 2016 US P 00 0

P00003715 ASHE, WARREN RODERICK DEM 2016 US P 00 0

P00003756 MARTISKO, MIKE BENJAMIN REP 2016 US P 00 0

P00003814 MALONE, JOAN REP 2016 US P 00 0

P00003897 WALLACE, WILLIAM LEONARD IND 2016 US P 00 0

P00003962 SCHRINGER, JOSEPH CHARLES UNK 2016 US P 00 0

P00004135 HANKINS, GERALD LAMAR MR DEM 2016 US P 00 0

P00004275 BROWN, HARLEY D NNE 2016 US P 00 0

P00004424 GREENSTEIN, MARK STEWART LIB 2016 US P 00 0

P00004507 EDWARDS, JENNY UNK 2016 US P 00 0

P00004523 HARRELL, SHARON WANDA UN 2016 US P 00 0

P00004580 BOYDSTON, J FATHER REV IND 2016 US P 00 0

P00004705 STRICKLAND, GLORIA DAWN IND 2016 US P 00 0

P00004846 IWACHIW, WALTER N REP 2016 US P 00 0

P00004861 BICKELMEYER, MICHAEL REP 2016 US P 00 0

P20000055 LEE, KIP LIB 2016 US P 00 0

PFULLER, LOWELL JACKSON REP 2016 US P 00 0

Show / Hide data grid options 15 Columns 1,128 Rows 4 Data Types

9+ 19°C Clear

Search

ENG IN 05:57 PM 08-12-2025



Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

Task 6. Wrangle the Contributions file and join it to the Candidates file

On the Join page, you can add your current dataset to another dataset or recipe based on information that is common to both datasets.

Before you join the Contributions file to the Candidates file, clean up the Contributions file.

1. Click on **FEC-2016** (the dataset selector) at the top of the grid view page.



2. Click to select the grayed out **Campaign Contributions 2016**.

3. In the right pane, click **Add > Recipe**, then click **Edit Recipe**.

4. Click the **recipe** icon at the top right of the page, then click **Add New Step**.

GSP105

Overview

Setup and requirements

Task 1. Create a Cloud Storage bucket in your project

Task 2. Initialize Cloud Dataprep

Task 3. Create a flow

Task 4. Import datasets

Task 5. Prep the candidate file

Task 6. Wrangle the Contributions file and join

< Previous

Next >

Dashboard – qwiklabs-gcp-00-874... FEC-2016 - Cloud Dataprep

https://clouddataprep.com/flows/2579945?recipe=14532142&tab=recipe&projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 United States Federal Elections Commission 2016

Dataset Recipe Candidate Master 2016 Candidate Master 2016

Dataset Recipe Campaign Contributions 2016 Untitled recipe

Untitled recipe

Edit recipe Add recipe ...

Recipe Data

Steps Preview

No Recipe Steps.
Edit the Recipe to wrangle.

Steps	0
Updated	Today at 5:58 PM
Created	Today at 5:58 PM

9+ 19°C Clear ENG IN 05:58 PM 08-12-2025

The screenshot shows the Cloud Dataprep interface with two main data flows. The first flow connects a 'Candidate Master 2016' dataset to a 'Candidate Master 2016' recipe. The second flow connects a 'Campaign Contributions 2016' dataset to an 'Untitled recipe'. On the right, a detailed view of the 'Untitled recipe' is shown, including its title, edit and add buttons, tabs for Recipe and Data, and a preview section indicating no steps have been added. Below the preview, there's a table with information about the recipe's creation and update times. The bottom of the screen features a taskbar with various icons and system status indicators.

Dashboard – qwiklabs-gcp-00-874... Untitled recipe - Cloud Dataprep +

https://clouddataprep.com/data/2579945/14532142?panel=recipe&projectId=qwiklabs-gcp-00-8749748b7da9 | 2

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Untitled recipe Initial data Recipe Run

Initial data

Loading data...

The recipe is empty

Add New Step

Show / Hide data grid options 15 Columns 1,128 Rows 4 Data Types

9+ 19°C Clear

Search

ENG IN 05:59 PM 08-12-2025

This screenshot shows the Cloud Dataprep interface. On the left is a vertical toolbar with various icons. The main workspace has a header bar with tabs like 'Untitled recipe - Cloud Dataprep' and a search bar. Below the header is a navigation bar with links to 'Free AI Paraphrasin...', 'Transcript of Case St...', 'GCP-LAB', 'Hydra', and 'datawarehouse'. The main area is titled 'FEC-2016 Untitled recipe' and contains a section labeled 'Initial data' with a progress bar. A large placeholder icon with the text 'The recipe is empty' and a 'Add New Step' button is visible. At the bottom, there's a data grid summary: 'Show / Hide data grid options', '15 Columns', '1,128 Rows', and '4 Data Types'. The bottom right corner shows system status: '9+ 19°C Clear', a search bar, and a date/time stamp '05:59 PM 08-12-2025'.



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Contents](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik Start

Dataprep: Qwik Start



The recipe is empty

[Add New Step](#)

00:48:08

Dashboard

Catalog

Paths

Collections

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.

[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

5. Insert the following Wrangle language command in the Search box:

```
replacepatterns col: * with: '' on: '{start}'|'{end}' global Copied!
```



The Transformation Builder parses the Wrangle command and populates the Find and Replace transformation fields.

Column

required

All

Find

{start}|{end}

[Previous](#)[Next](#)

Dashboard – qwiklabs-gcp-00-874... Untitled recipe - Cloud Dataprep +

https://clouddataprep.com/data/2579945/14532142?projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Untitled recipe Initial data

Replace text or patterns

Column required
All

Find
{start}|{end}

Replace with required
String

Advanced options ▾

Cancel Add

Loading data...

Show / Hide data grid options 21 Columns 61,046 Rows 5 Data Types

9+ 19°C Clear

Search

Cloud Dataprep

Run

ENG IN

06:01 PM

08-12-2025



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636



Contents

Dataprep: Qwik Start



6. Click Add to add the transform to the recipe.

7. Add another new step to the recipe. Click New Step, then type "Join" in the Search box.

End Lab
00:45:59Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)[Open Google Cloud console](#)

Username

student-02-dc747dbe9199e



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

< Recipe Search transformations X

Join

Join datasets

8. Click Join datasets to open the Joins page.

9. Click on "Candidate Master 2016" to join with Campaign Contributions 2016, then Accept in the bottom right.

< Previous

Next >

ENG
IN06:01 PM
08-12-2025

Dashboard – qwiklabs-gcp-00-874... Untitled recipe - Cloud Dataprep +

https://clouddataprep.com/data/2579945/14532208?panel=recipe&projectId=qwiklabs-gcp-00-8749748b7da9

Free AI Paraphrasin... Transcript of Case St... GCP-LAB Hydra datawarehouse

FEC-2016 Untitled recipe Initial data Recipe Run

Initial data

Not able to load the Data of “**Campaign Contributions 2016**”, poor performance

Loading data...

The recipe is empty

Add New Step

Show / Hide data grid options 0 Columns 0 Rows 0 Data Types

9+ 19°C Clear

Search

06:08 PM 08-12-2025 ENG IN



Free AI Paraphrasin...

Transcript of Case St...

GCP-LAB

Hydra

datawarehouse

Google Skills

Partner

What do you want to learn today?



★ 5636

0

[Home](#) > Prepare Data for ML APIs on Google Cloud > Dataprep: Qwik StartContents
Dataprep: Qwik Start

Advanced options ▾

Cancel

Add

Caution: When you are in the console, do not deviate from the lab instructions. Doing so may cause your account to be blocked.
[Learn more.](#)

[Open Google Cloud console](#)

Username

student-02-dc747dbe9199e



Password

RcMw1vMgRQkJ



Project ID

qwiklabs-gcp-00-8749748l



Student Resources

Recipe Search transformations



Join

Join datasets

[Previous](#)[Next >](#)ENG
IN06:08 PM
08-12-2025