How-To Library

Informatica

Getting Started with Informatica Intelligent Cloud Services on Google Cloud Platform

Abstract

This article explains how to get started with Informatica Intelligent Cloud Services℠ on Google Cloud Platform (GCP). The instructions in this article apply to the IICS instance on the GCP POD only.

Supported Versions

- Informatica Intelligent Cloud Services GCP Beta (Dec 2019)

Table of Contents

Introduction

Welcome to Getting Started with Informatica Intelligent Cloud Services on Google Cloud Platform.

Informatica Intelligent Cloud Services (IICS) is the only complete platform for cloud integration and data management. IICS Data Integration lets you easily connect to a variety of cloud, on-premises, mobile, and social data sources to ensure you can efficiently share your relevant and trustworthy business information. To help you better engage with your customers and realize the potential of your SaaS investments, IICS goes beyond traditional point-to-point integration vendors. Informatica offers the most complete suite of enterprise data management in the Integration Platform as a Service (iPaaS) market.

As the proven leader in cloud integration with hundreds of connectors, more than 9,500 customers, and having processed trillions of transactions per month in the cloud (and counting), the IICS visual design interface is designed for non-technical business users and incorporates best practices and templates that accelerate your deployment of cloud apps.

This article will provide you an opportunity to experience the power of IICS for Google Cloud Platform (GCP). You will learn how easy it is to synchronize data from a CSV file to Google Cloud Storage and BigQuery without writing a single line of code.

Getting started

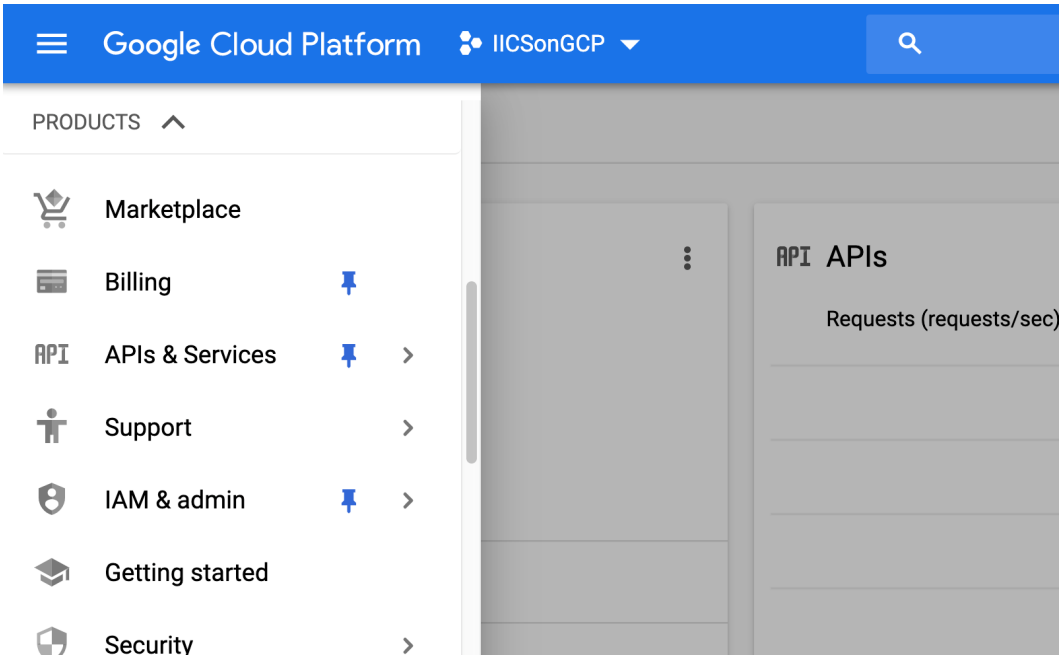
Getting started with IICS on GCP consists of just a few steps.

Step 1. Activate IICS from GCP marketplace

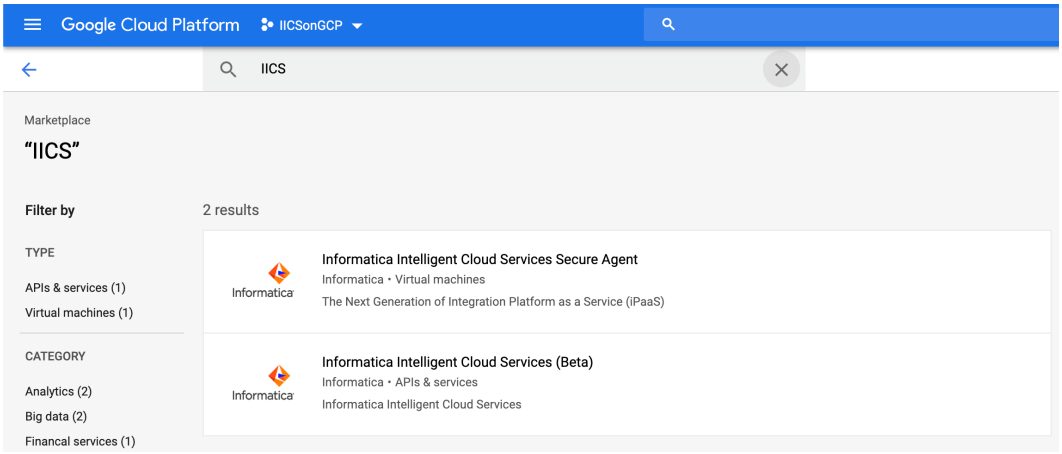
IICS is listed in GCP Marketplace. Go to the Marketplace to activate IICS.

1. Using the Google Chrome browser, log in to GCP Console.

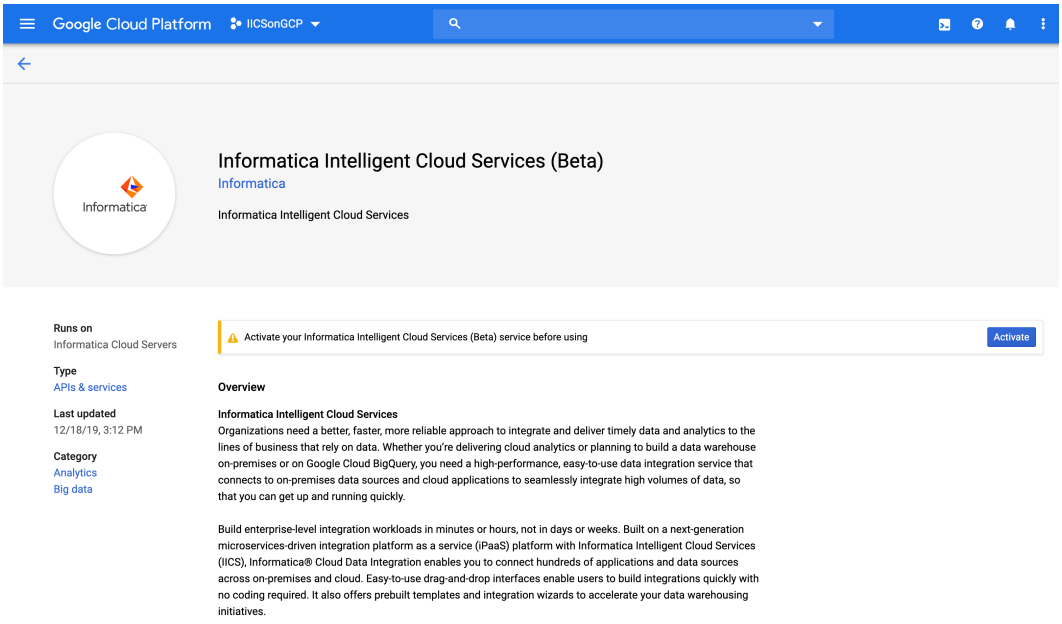
2. Select **Marketplace** from the menu:



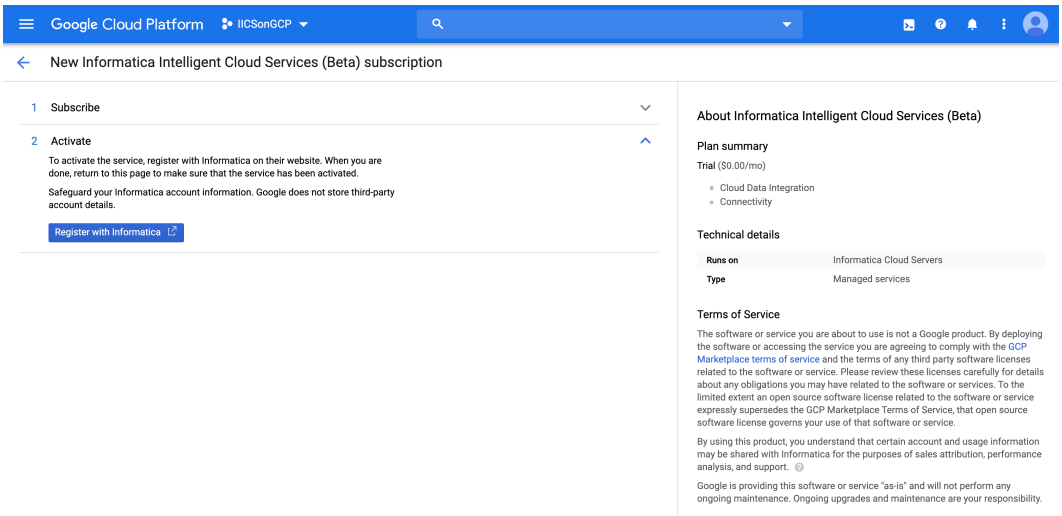
3. Search for `IICS` and click **Informatica Intelligent Cloud Services (Beta)**:



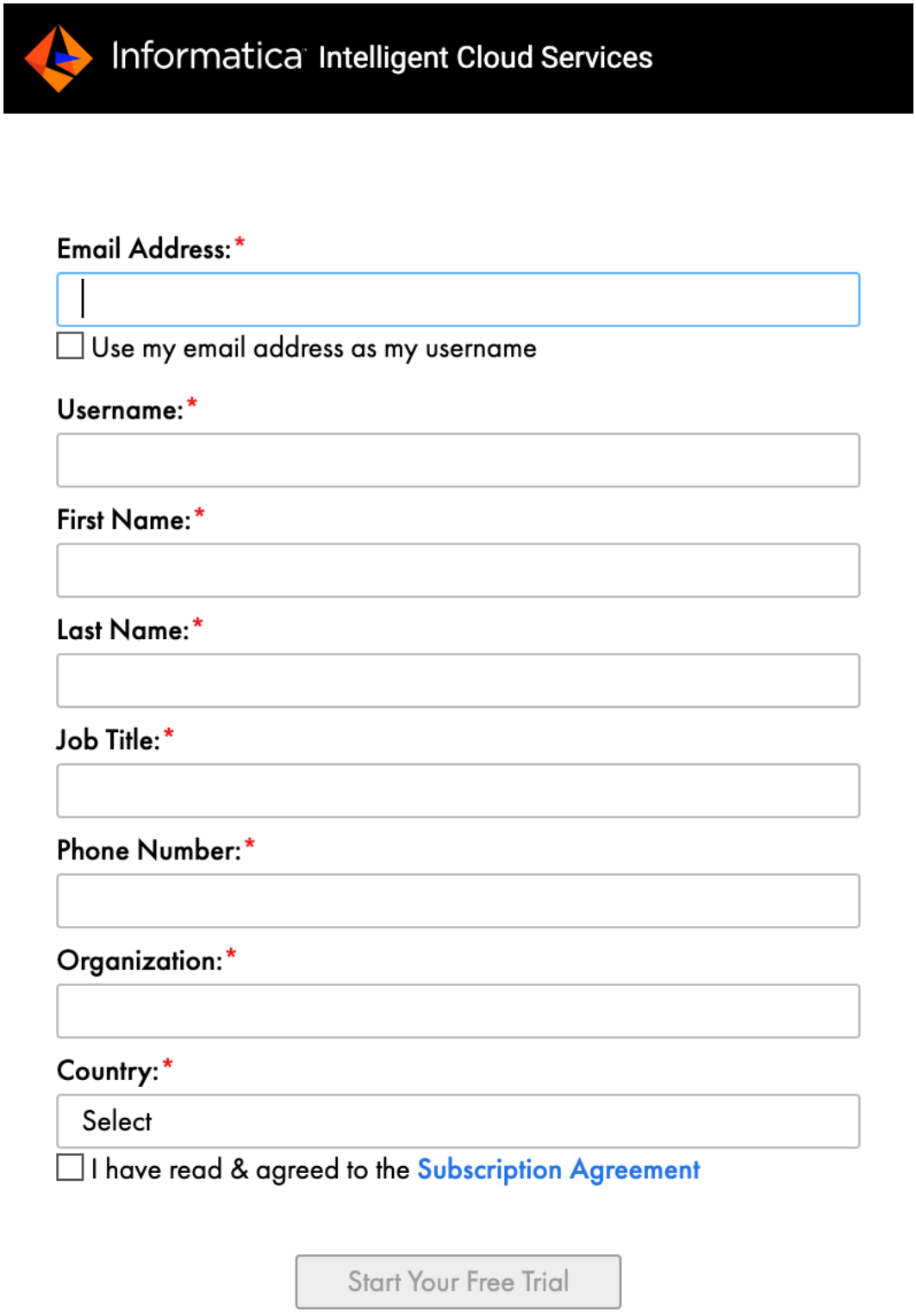
4. Click **Activate**:



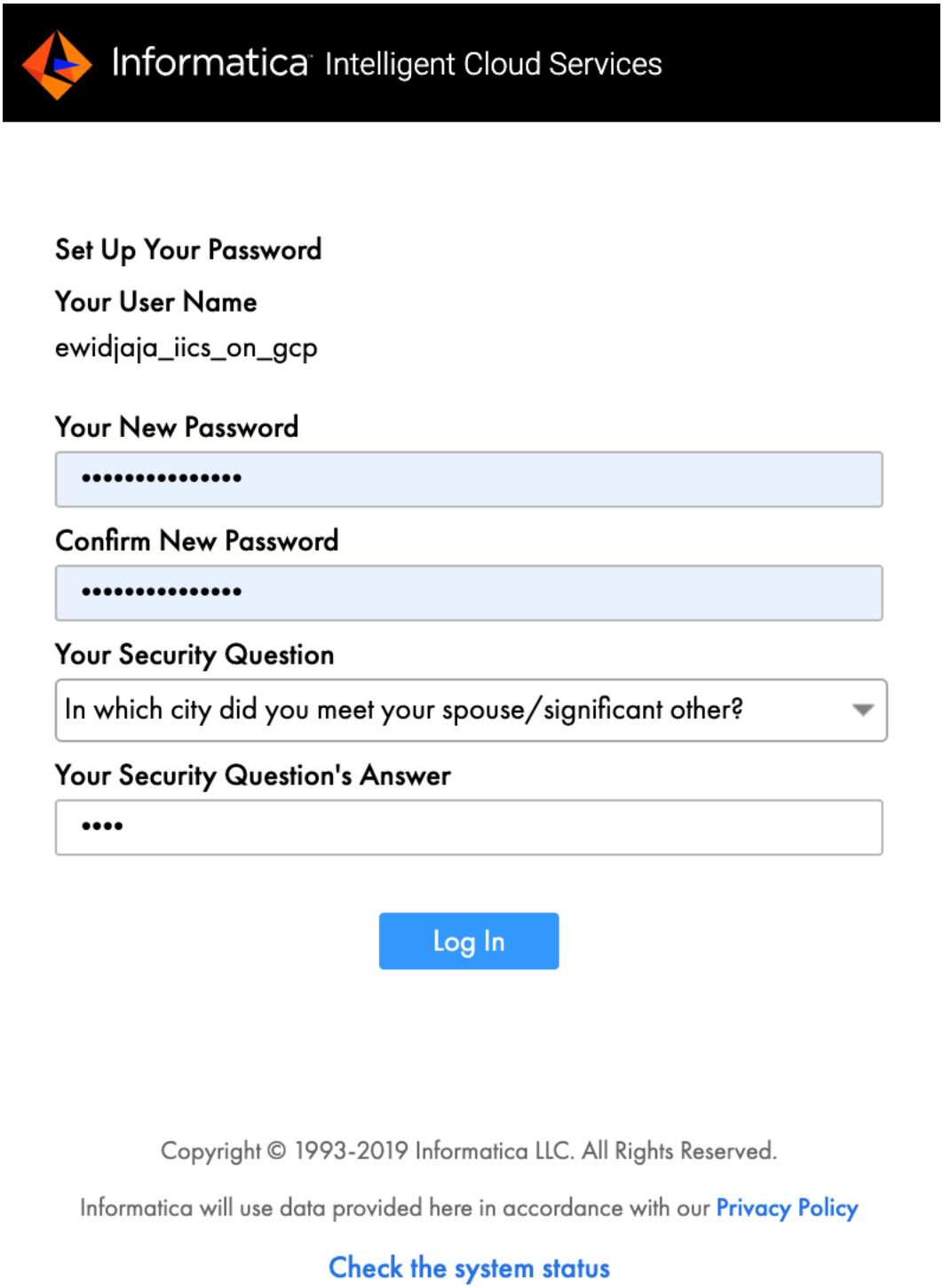
5. Click **Register with Informatica**:



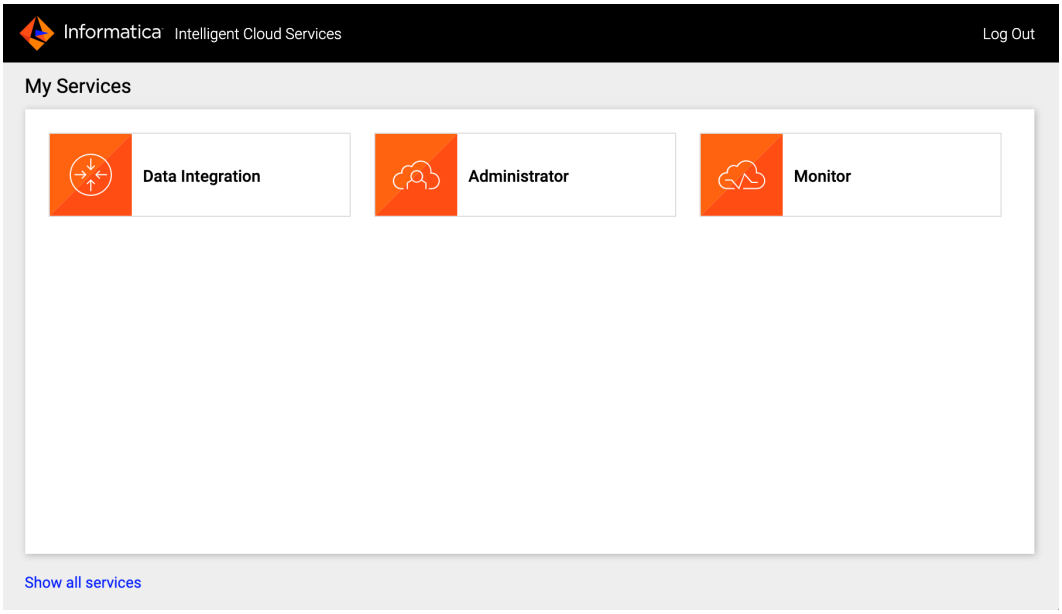
6. Fill in the form, and click **Start Your Free Trial**:



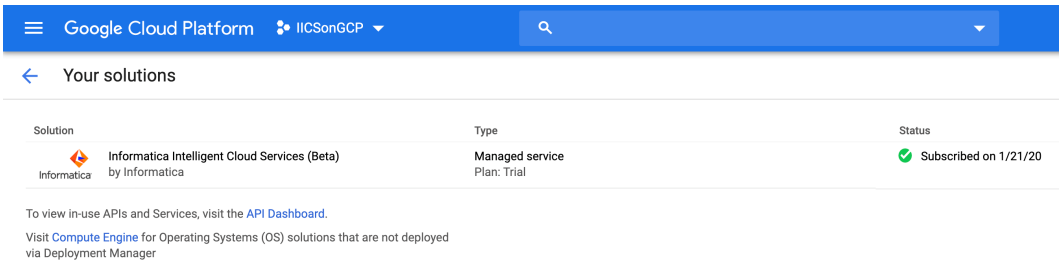
7. Check your email, and click **Confirm Account** in the email.
8. Set your password, security question, and answer, and then click **Log In**:



9. After successful login, you see the **My Services** page:



10. In GCP Marketplace, click **Your solutions** to view your IICS subscription:



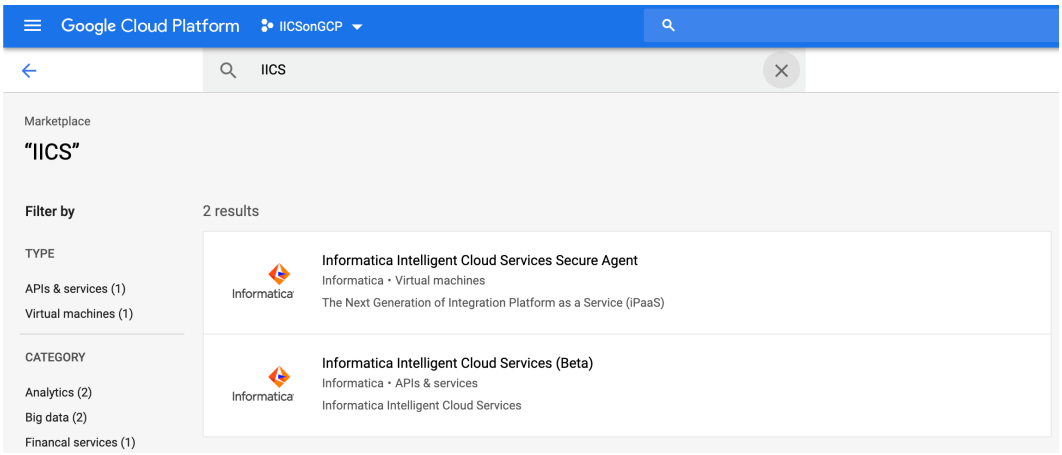
Step 2. Install a Secure Agent (runtime environment)

A Secure Agent is a small footprint application that enables secure communication across the firewall between your organization and IICS. It enables IICS to get through the firewall to access application, relational database, and file sources and targets in your on-premises network.

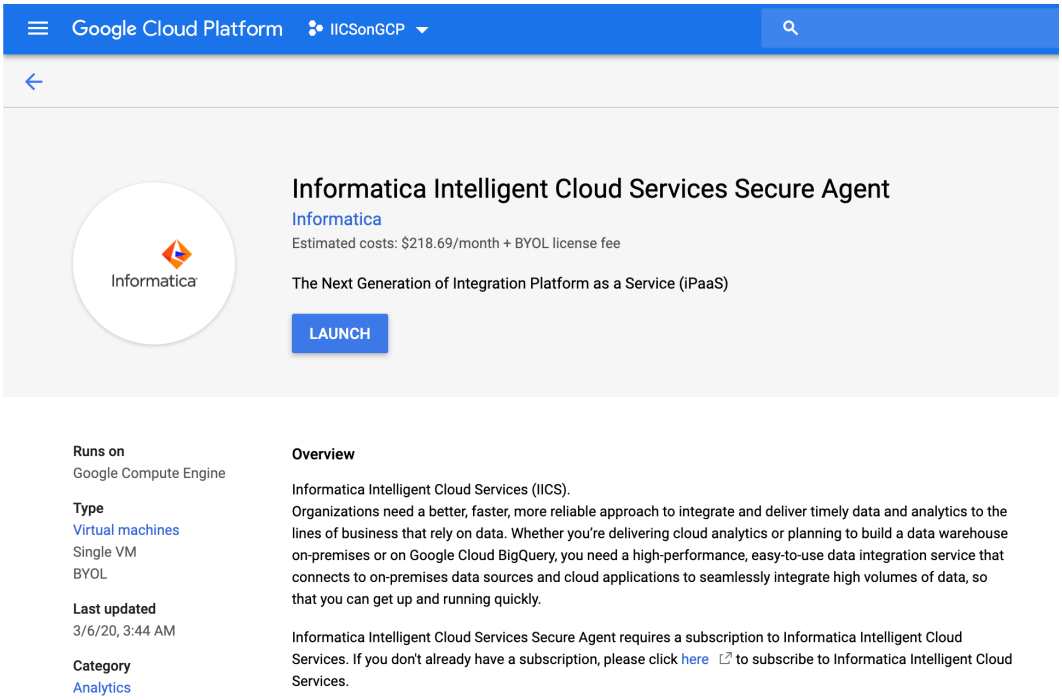
In this step, you launch a Secure Agent on a Google Compute Engine from the Marketplace.

The Secure Agent is listed in GCP Marketplace. Go to the Marketplace to launch it.

1. Search for `IICS` and click **Informatica Intelligent Cloud Services Secure Agent**:



2. Click **LAUNCH**:



3. Enter the following information on the **New Informatica Intelligent Cloud Services Secure Agent deployment** page and accept the default values for all other properties:

Property	Description
Region	Select a region or accept the default value.
Operating System	Select Linux.
Username	Enter your IICS user name.
Install Token	Paste the Secure Agent install token that you get from the IICS Administrator service.*

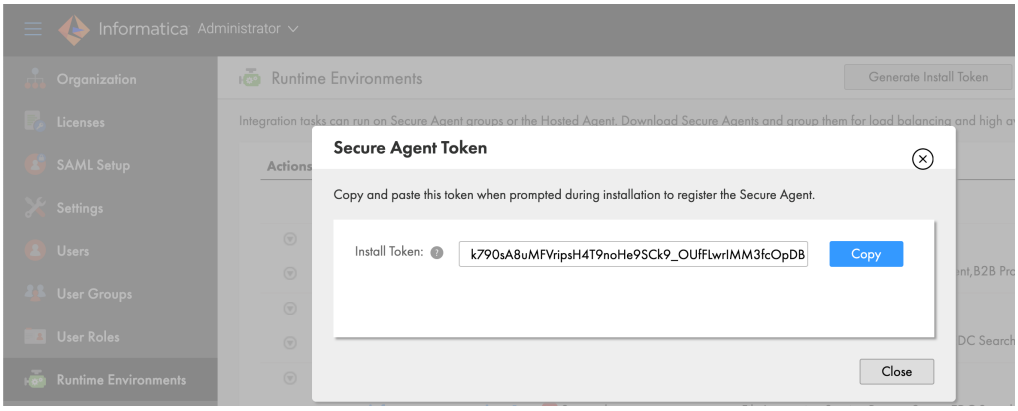
* Perform the following steps to get the install token:

a. Log in to IICS.

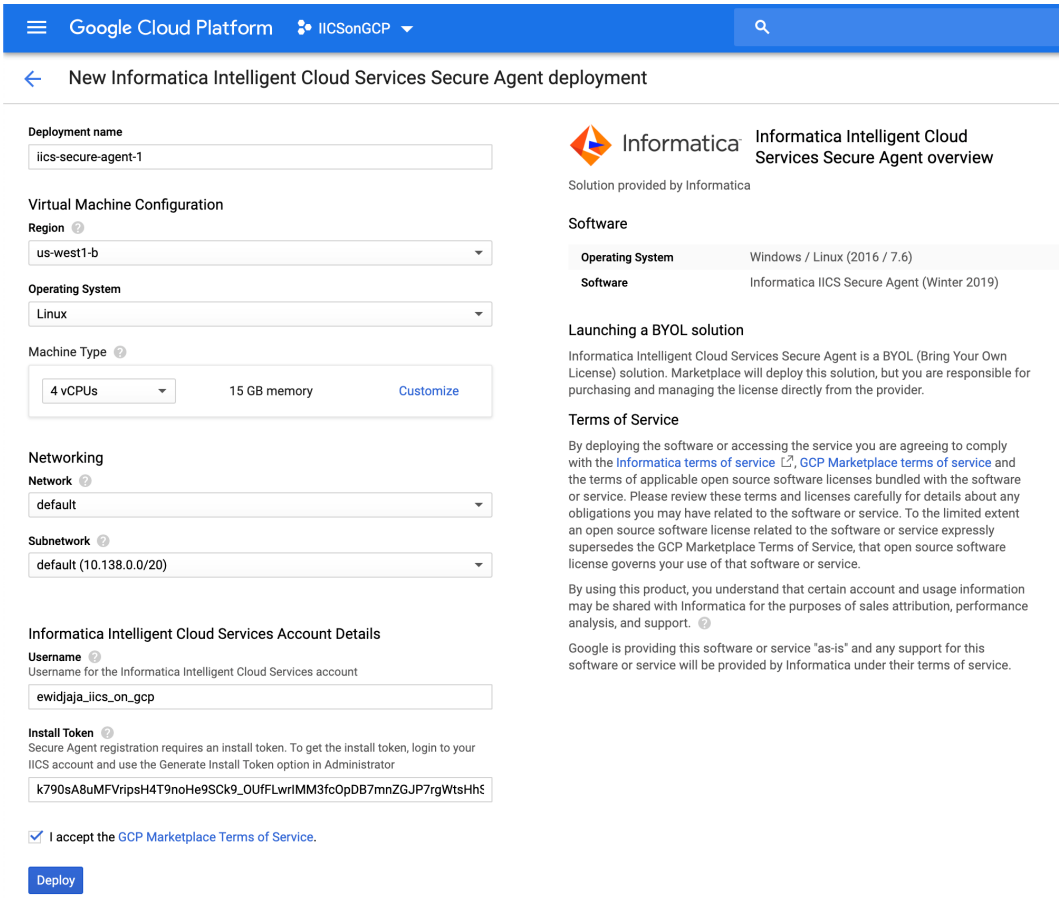
b. Select **Administrator**, and then click **Runtime Environments**.

c. Click **Generate Install Token**.

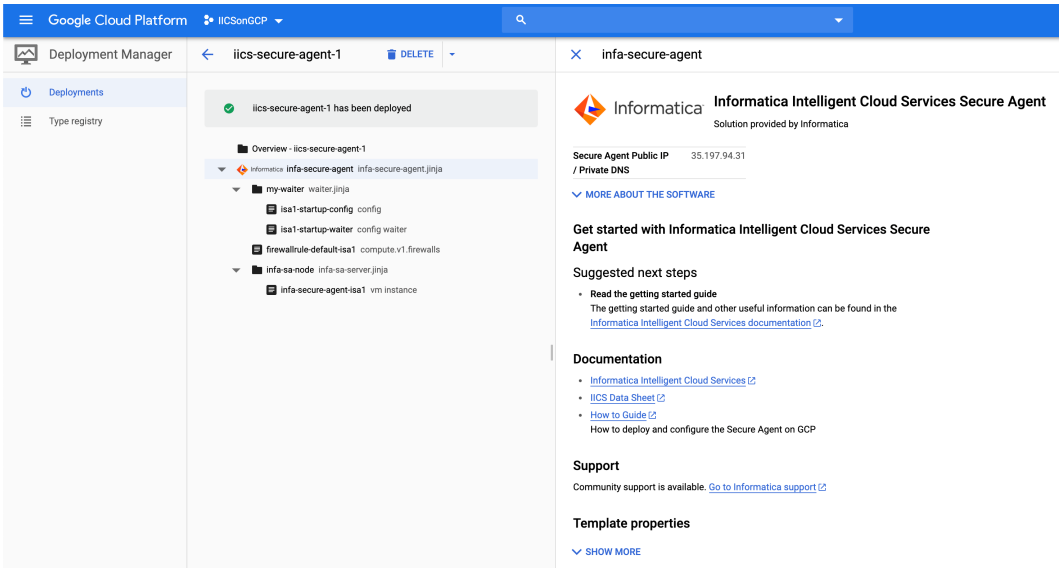
d. Click **Copy** to copy the install token string:



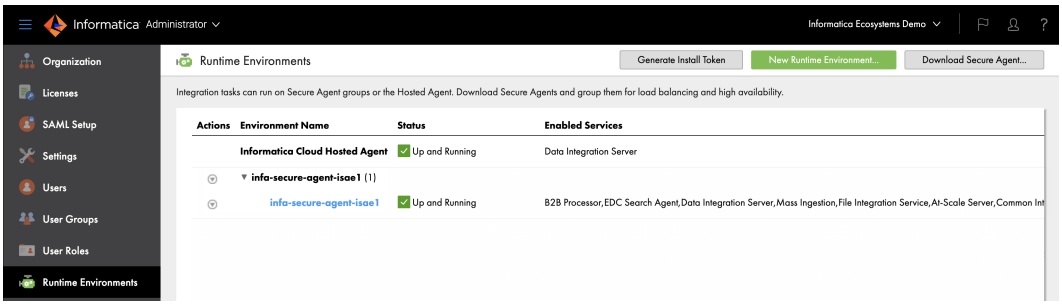
4. Click **Deploy**:



When the deployment completes, you will see the following page:



You should also see the Secure Agent listed on the **Runtime Environments** page in Administrator:



Note: Your runtime environment should be listed with the status "Up and Running" before you continue to the next step. If you see the status "Not all the services are running," wait 5-10 minutes for the agent to finish starting up and installing updates.

Step 3. Configure connections

Informatica supports more than 300 connectors. A connector is a pre-built integration that allows you to connect to data sources and targets.

In this step, you configure Flat File, Google Cloud Storage, and Google BigQuery connections.

Configure a flat file connection

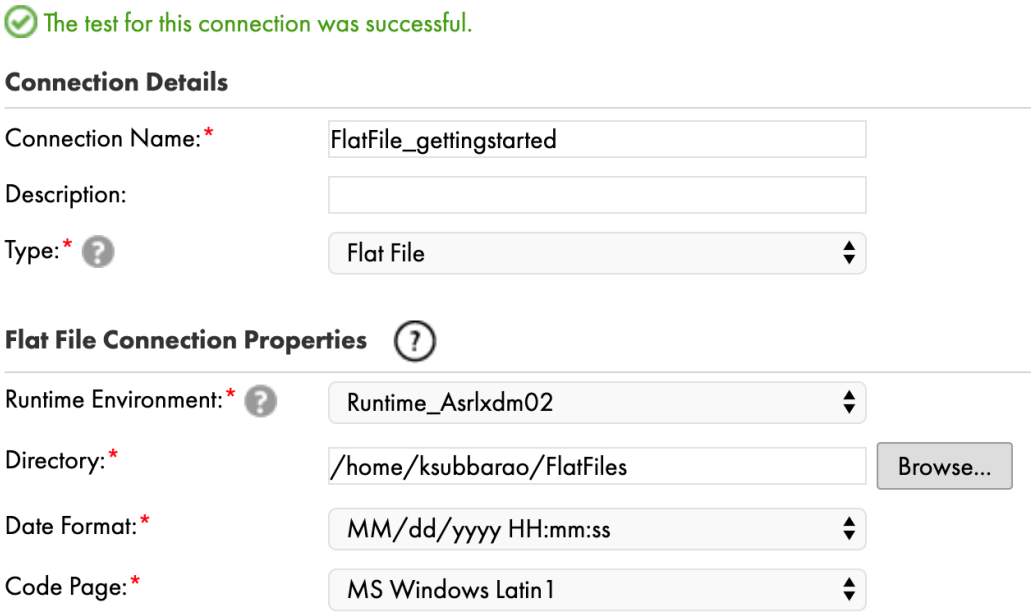
Configure a flat file connection on the **Connections** page in the Administrator service.

1. In Administrator, click **Connections**.
2. Click **New Connection** in the top right corner of the screen.
3. Enter a name and optional description for the connection.
4. Select **Flat File** as the connection type.
5. Configure the following properties:

Property	Description
Runtime Environment	Select the runtime environment that you created.
Directory	Enter the directory where the flat files are stored or click **Browse** to select the directory. The directory must be accessible by all Secure Agents in the selected runtime environment. **Note:** If you click **Browse** and get an error that says, "The runtime environment cannot get the requested file system information...," see KB article 498540 on Informatica Network.
Date Format	Select the date format for the files that will use this connection.
Code Page	Select the code page for the files that will use this connection. If you use a flat file connection with the Shift-JIS code page and a UTF data object, be sure to install fonts that fully support Unicode.

6. To test the connection, click **Test Connection**.

You should see the message, "The test for this connection was successful":



7. Click **Save** to save the connection.

For more information about configuring a flat file connection, see this topic in the IICS *Connections* guide.

Configure a Google Cloud Storage V2 connection

Configure a Google Cloud Storage V2 connection on the **Connections** page in the Administrator service.

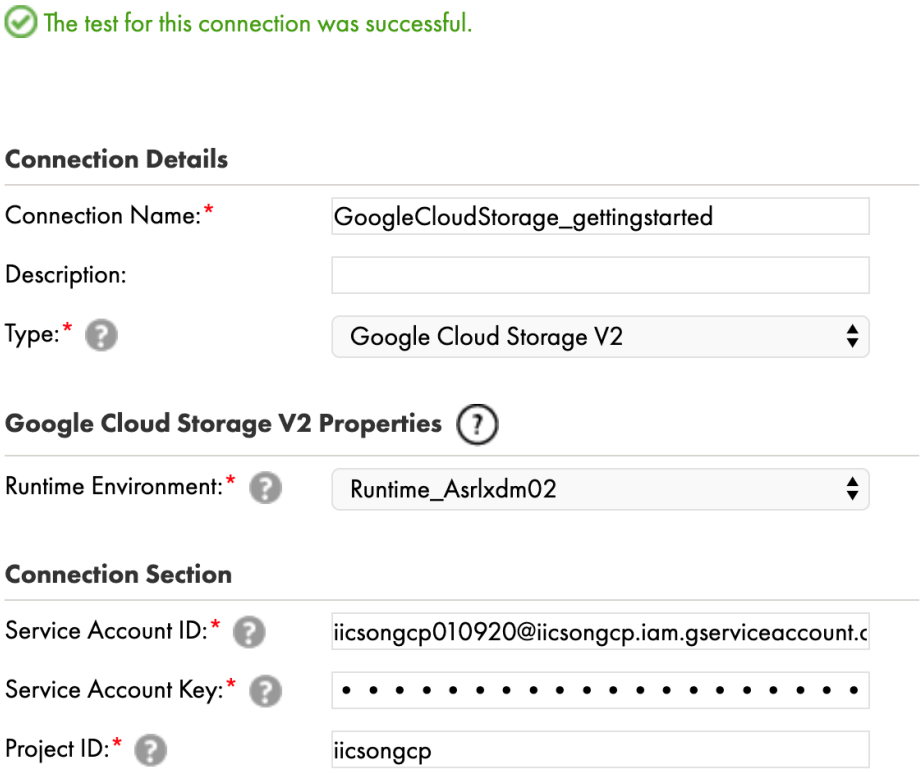
1. In Administrator, click **Connections**.
2. Click **New Connection** in the top right corner of the screen.
3. Enter a name and optional description for the connection.
4. Select **Google Cloud Storage V2** as the connection type.
5. Configure the following properties:

Property	Description
Runtime Environment	Select the runtime environment that you created.
Service Account ID	Enter the client_email value present in the JSON file that you download after you create a service account.
Service Account Key	Enter the private_key value present in the JSON file that you download after you create a service account.
Project ID	Enter the project_id value present in the JSON file that you download after you create a service account. If you create multiple projects with the same service account, enter the ID of the project that contains the bucket that you want to connect to.

For information about creating service accounts, see this GCP guide.

6. To test the connection, click **Test Connection**.

You should see the message, "The test for this connection was successful":



7. Click **Save** to save the connection.

For more information about configuring a Google Cloud Storage V2 connection, see this topic in the IICS *Connections* guide.

Configure a Google BigQuery V2 connection

Configure a Google BigQuery V2 connection on the **Connections** page in the Administrator service.

1. In Administrator, click **Connections**.
2. Click **New Connection** in the top right corner of the screen.
3. Enter a name and optional description for the connection.
4. Select **Google BigQuery V2** as the connection type.
5. Configure the following properties:

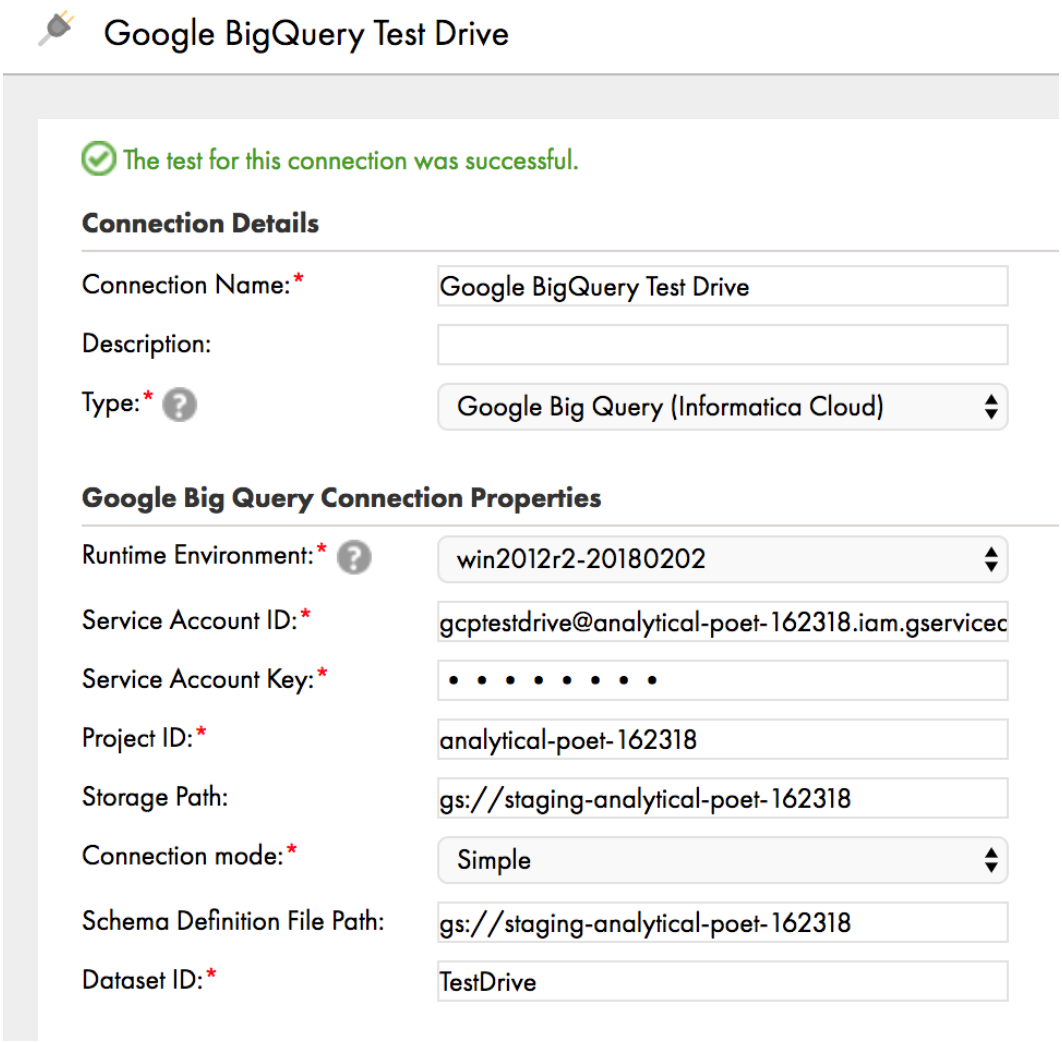
Property	Description
Runtime Environment	Select the runtime environment that you created.
Service Account ID	The client_email value present in the JSON file that you download after you create a service account.

Property	Description
Service Account Key	The private_key value present in the JSON file that you downloaded after you create a service account. **Note:** Ensure that the service account key contains no leading or trailing spaces.
Project ID	Specifies the project_id value present in the JSON file that you download after you create a service account.
Storage Path	Path in Google Cloud Storage where the Secure Agent creates a local stage file to store the data temporarily. This is required when the write mode is set to bulk instead of streaming.
Connection Mode	Mode that you want to use to read data from or write data to Google BigQuery.
Schema Definition File Path	Directory on the Secure Agent machine where the Secure Agent must create a JSON file with the sample schema of the Google BigQuery table. The JSON file name is the same as the Google BigQuery table name.
Dataset ID	Name of the dataset that contains the source table and target table that you want to connect to.

For information about creating service accounts, see this GCP guide.

6. To test the connection, click **Test Connection**.

You should see the message, "The test for this connection was successful":



7. Click **Save** to save the connection.

For more information about configuring a Google BigQuery V2 connection, see this topic in the IICS *Connections* guide.

Step 4. Create a mapping to load a CSV file into Google Cloud Storage

The Data Integration Mapping Designer allows you to define data integration flow in a visual designer interface. It starts from defining the data source to read data from, applying transformations based on your requirements on that data set, and then writing the resulting data set to the data target.

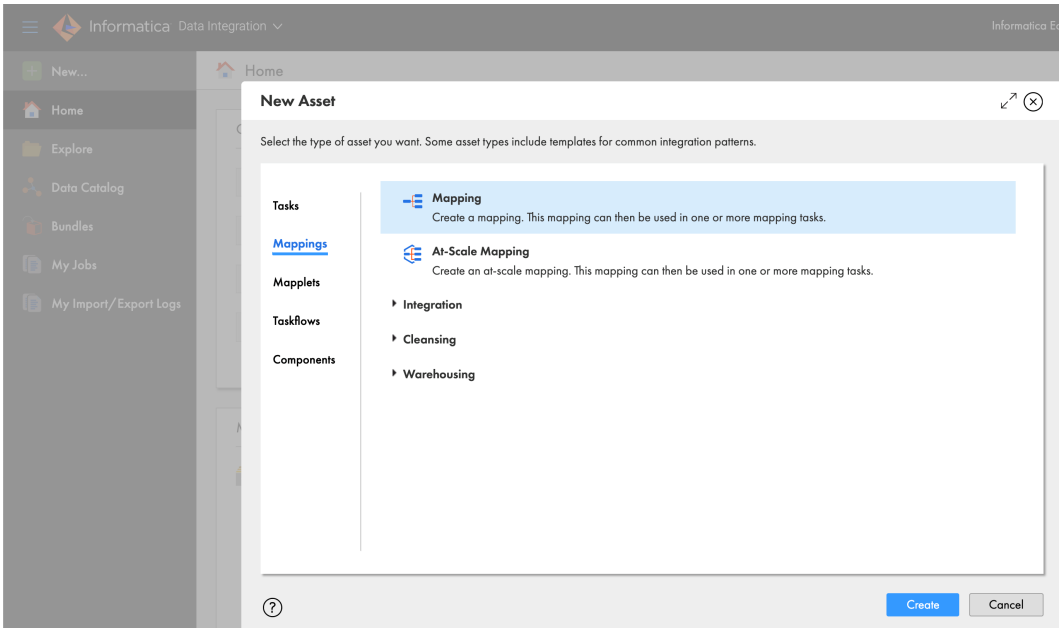
In this step, you load a CSV file from the flat file directory configured in flat file connection and write it to Google Cloud Storage. Please download the CSV file and place it in the flat file directory.

For information about transferring files to Google Compute Engine, see this GCP guide.

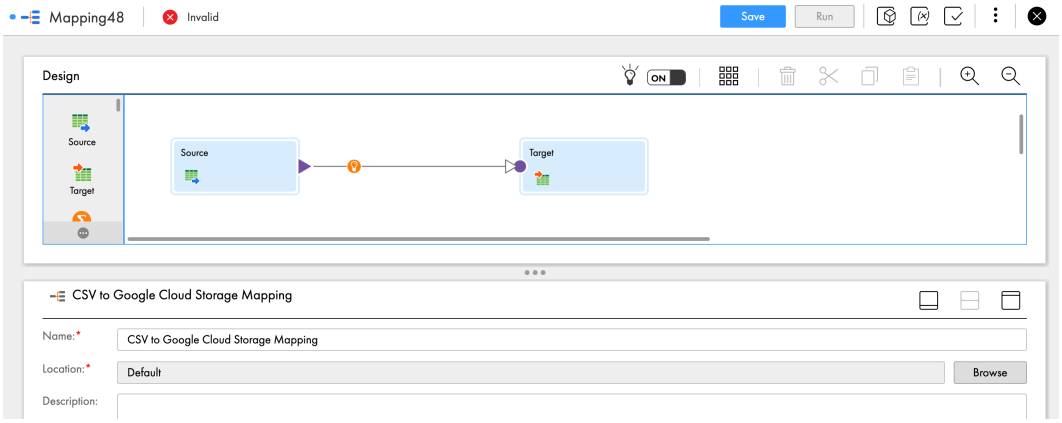
If you still have the Administrator service open, click **Administrator** next to the Informatica logo in the top left corner of the screen. Then select **Data Integration** from the menu.

1. In Data Integration, click **New** to open the **New Asset** dialog box.

2. In the **New Asset** dialog box, click **Mappings** in the menu on the left, select **Mapping**, and click **Create**:



3. When the mapping opens, enter the mapping name in the **Name** field:

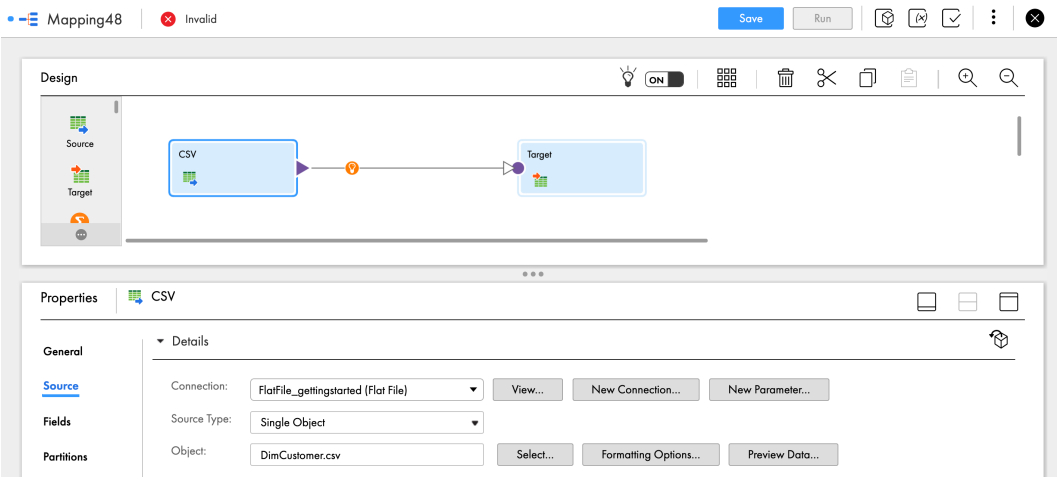


4. Click the Source transformation in the canvas.
5. On the **General** tab, update the name of the Source transformation to reflect the actual object.

   In this case, the source is a CSV file.
6. Click the **Source** tab, and perform the following steps:
   a. In the **Connection** field, select the flat file connection that you created.
   b. Click **Select** next to the **Object** field to select the CSV file.

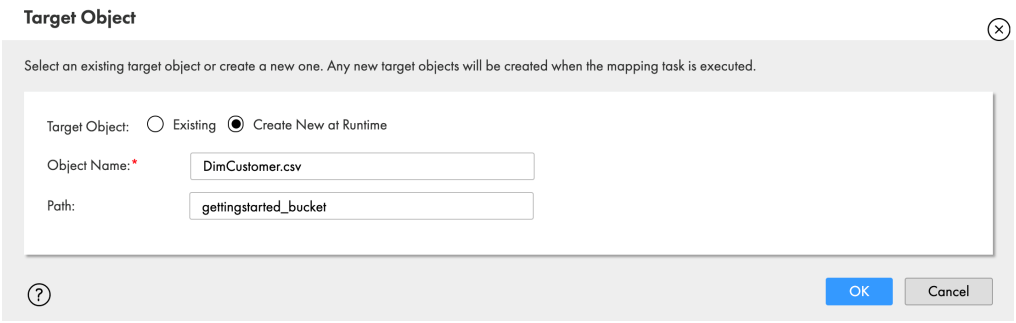
      IICS treats flat files as CSV by default. You can change the formatting by clicking **Formatting Options**.

c. Optionally, click **Preview Data** if you want to view the first few rows of the data file.



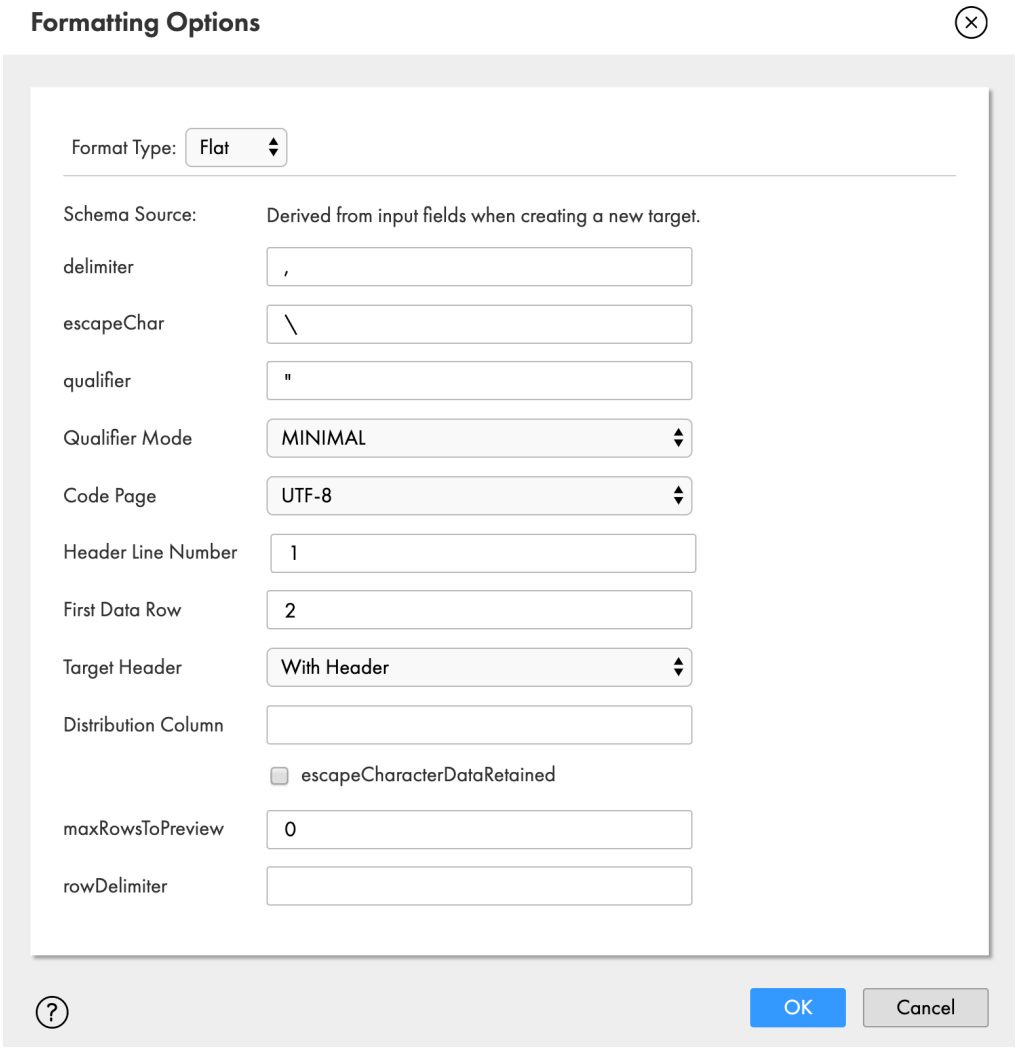
7. Click the Target transformation in the canvas.
8. On the **General** tab, update the name of the Target transformation to reflect the actual object.

   In this case, the target is Google Cloud Storage.
9. Click the **Target** tab, and perform the following steps:
   a. In the **Connection** field, select the Google Cloud Storage V2 connection that you created.
   b. Click **Select** next to the **Object** field to create the target file.
   c. Select **Create New at Runtime**.
   d. In the **Object Name** field, enter the name of the target file that will be created.
   e. In the **Path** field, enter the name of the Google Cloud Storage bucket where the file will be stored:



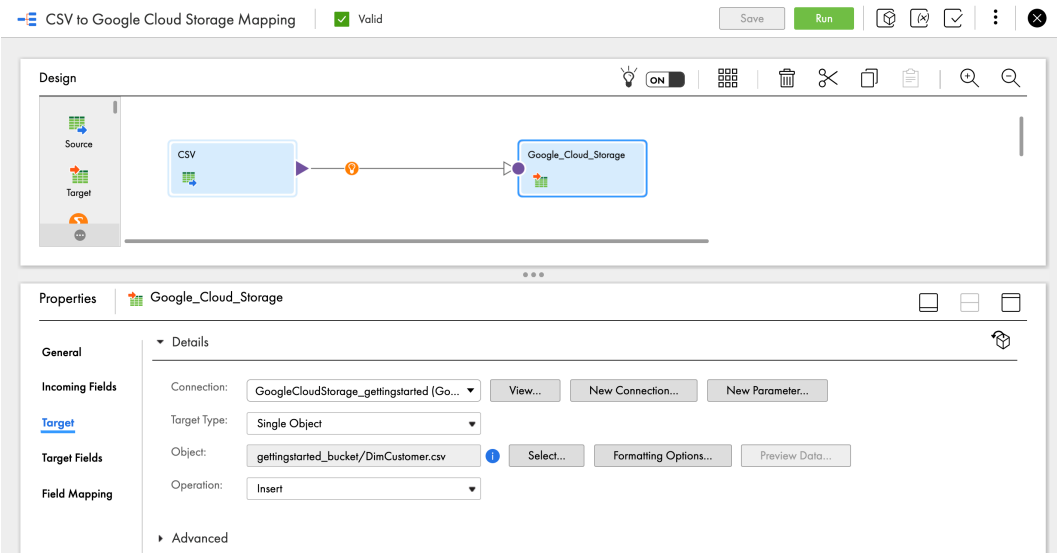
For information about creating Google Cloud Storage buckets, see this GCP guide.

f. Click **Formatting Options**, select **Flat** for the format type, and then click **OK**:

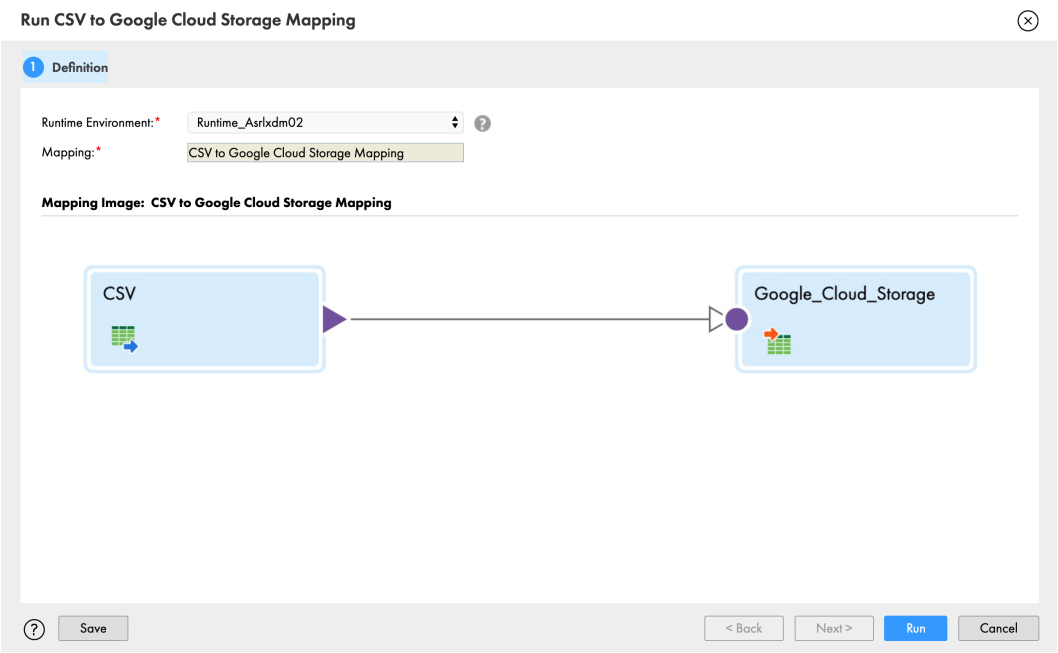


10. Click **Save**.

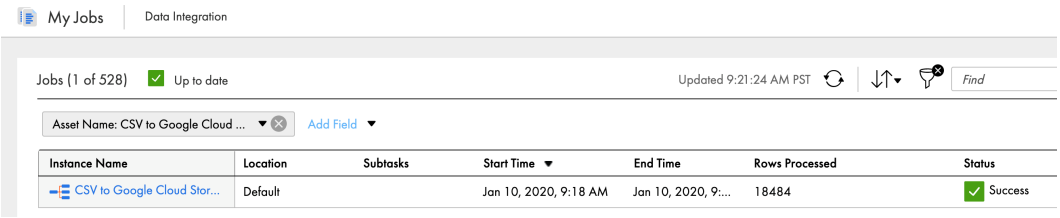
11. Click **Run** in the top right corner of the screen:



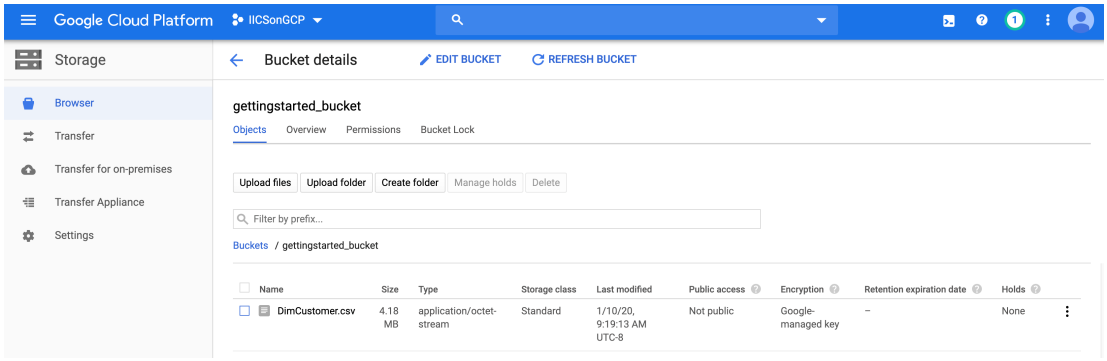
12. Select the runtime environment, and then click **Run**:



13. Click **My Jobs** to open the job activity page:



The target file is created in the Google Cloud Storage bucket:

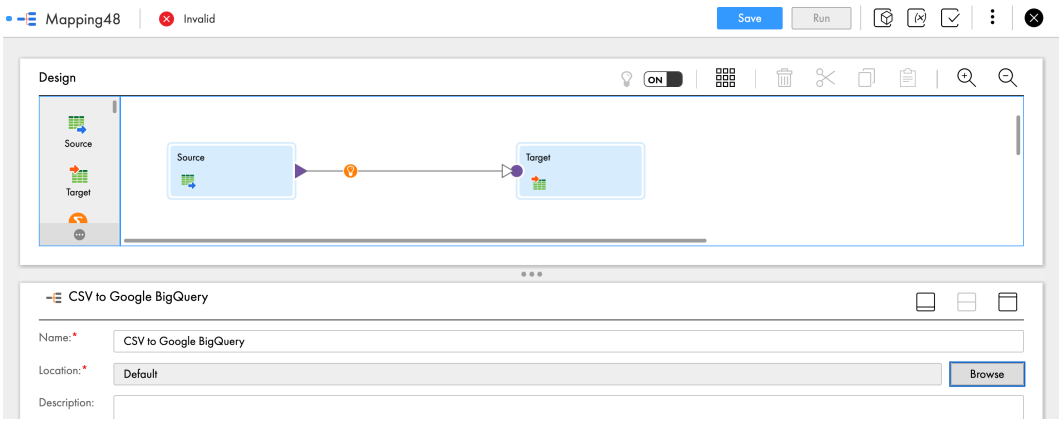


For more information about creating mappings, see the Data Integration Mappings guide.

Step 5. Create a mapping to load a CSV file into Google BigQuery

In this step, you load a CSV file from the flat file directory configured in flat file connection and write it to a table created during runtime in Google BigQuery. Use the same CSV file that you used in the previous step.

1. In Data Integration, click **New** to open the **New Asset** dialog box.
2. In the **New Asset** dialog box, click **Mappings** in the menu on the left, select **Mapping**, and click **Create**.
3. When the mapping opens, enter the mapping name in the **Name** field:

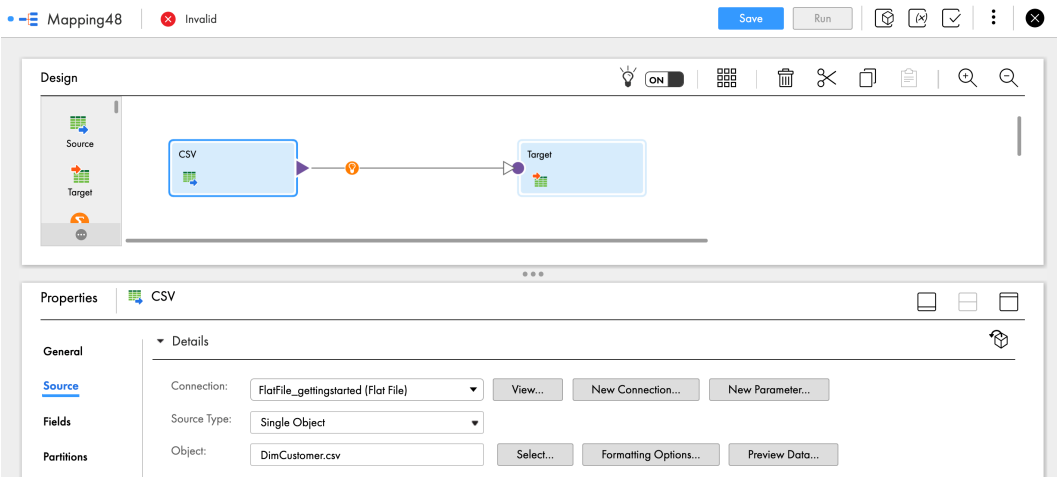


4. Click the Source transformation in the canvas.
5. On the **General** tab, update the name of the Source transformation to reflect the actual object.

   In this case, the source is a CSV file.
6. Click the **Source** tab, and perform the following steps:
   a. In the **Connection** field, select the flat file connection that you created.
   b. Click **Select** next to the **Object** field to select the CSV file.

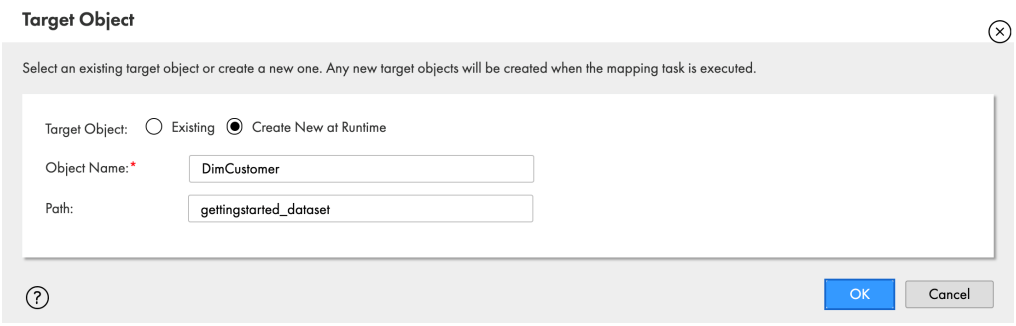
      IICS treats flat files as CSV by default. You can change the formatting by clicking **Formatting Options**.

c. Optionally, click **Preview Data** if you want to view the first few rows of the data file.



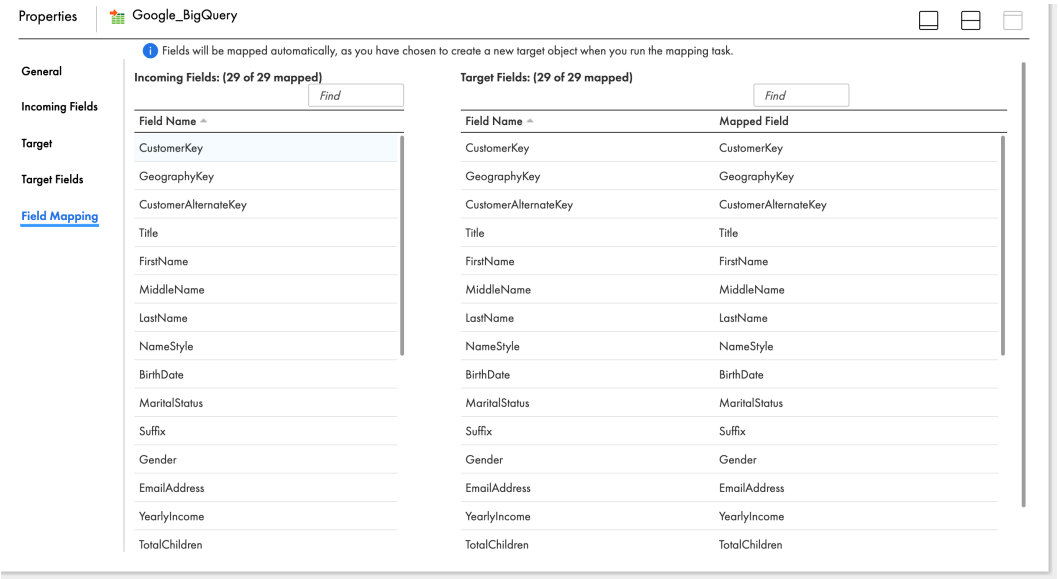
7. Click the Target transformation in the canvas.
8. On the **General** tab, update the name of the Target transformation to reflect the actual object.

   In this case, the target is Google BigQuery.
9. Click the **Target** tab, and perform the following steps:
   a. In the **Connection** field, select the Google BigQuery V2 connection that you created.
   b. Click **Select** next to the **Object** field to create the target file.
   c. Select **Create New at Runtime**.
   d. In the **Object Name** field, enter the name of the target file that will be created.
   e. In the **Path** field, enter the name of the Google BigQuery dataset where the table will be created:

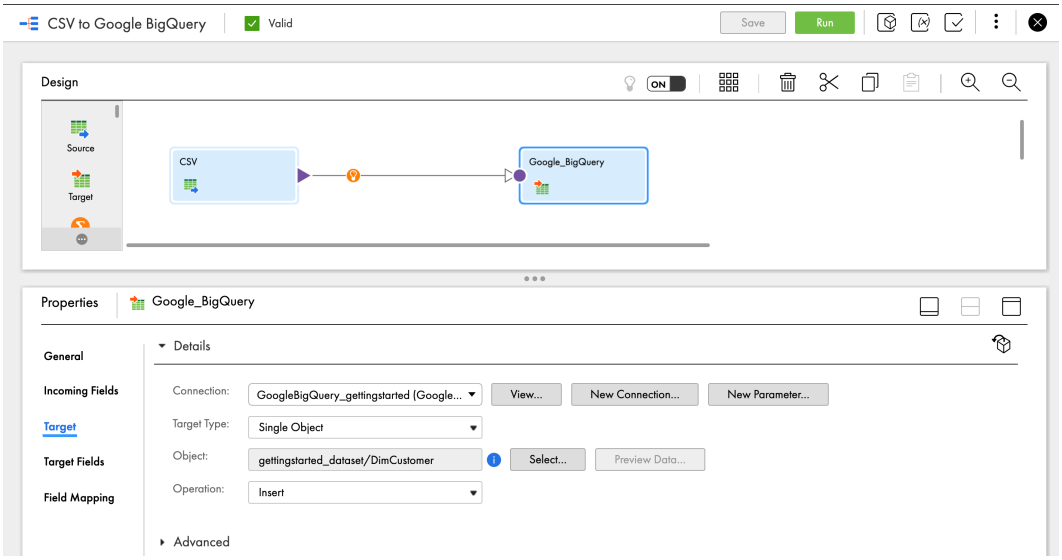


   For information about creating Google BigQuery datasets, see this GCP guide.
10. Click **Save**.
11. Click the **Field Mapping** tab.

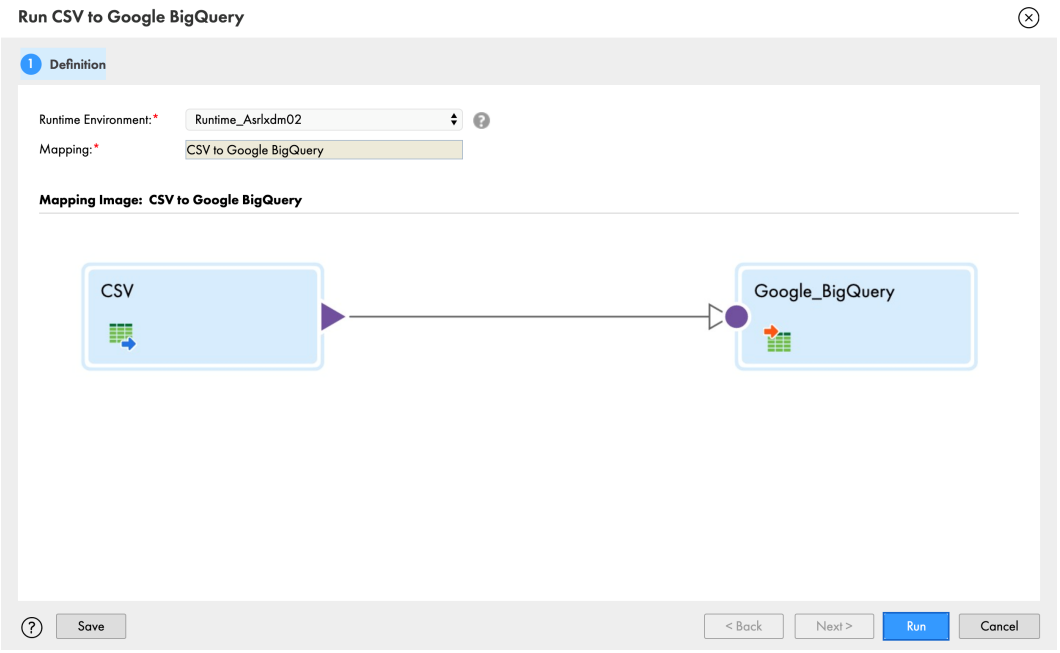
When you create a new target object, IICS automatically creates and maps fields from incoming fields, which come from the source dataset in this case, to the target:



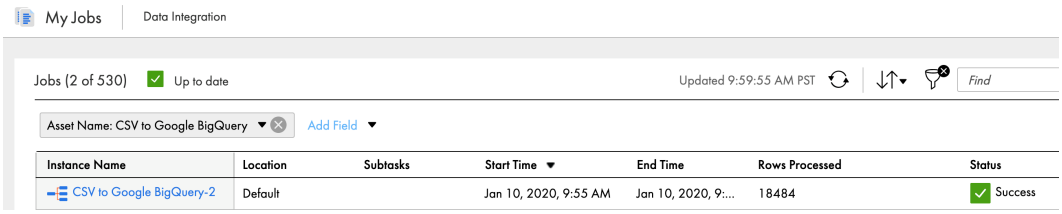
12. Click **Run** in the top right corner of the screen:



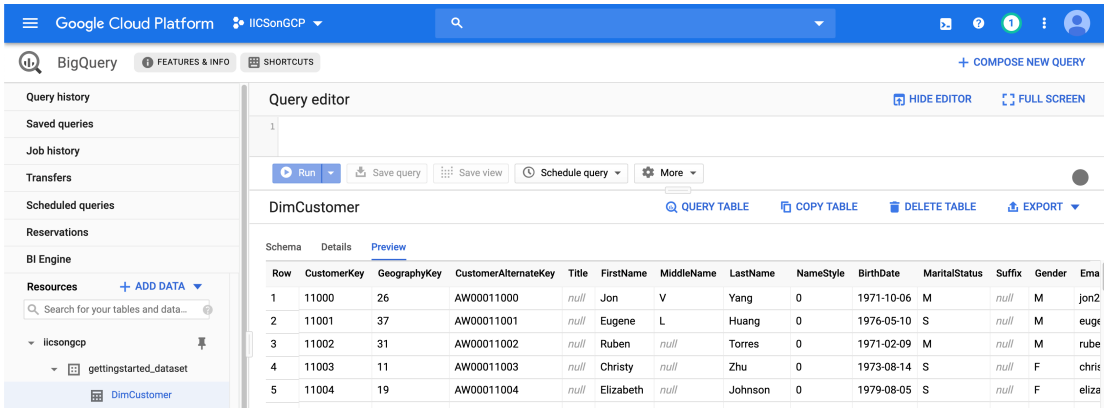
13. Select the runtime environment, and then click **Run**:



14. Click **My Jobs** to open the job activity page:



The target table is created in the Google BigQuery dataset:



Next steps

Congratulations on completing these Getting Started tasks!

There is much more that you can do with IICS. Please check out the Cloud Data Integration documentation on Informatica Network to find out more.

If you want to learn more about Informatica Intelligent Cloud Services on Google Cloud Platform, please contact us.

Author

Informatica Cloud R&D Team