



# PR PROFESSIONAL-DATA-ENGINEER<sup>Q&As</sup>

Professional Data Engineer on Google Cloud Platform

**Pass Google PROFESSIONAL-DATA-ENGINEER  
Exam with 100% Guarantee**

Free Download Real Questions & Answers **PDF** and **VCE** file from:

<https://www.geekcert.com/professional-data-engineer.html>

100% Passing Guarantee  
100% Money Back Assurance

Following Questions and Answers are all new published by Google  
Official Exam Center



- ⚙️ **Instant Download** After Purchase
- ⚙️ **100% Money Back** Guarantee
- ⚙️ **365 Days** Free Update
- ⚙️ **800,000+** Satisfied Customers





### QUESTION 1

MJTelco's Google Cloud Dataflow pipeline is now ready to start receiving data from the 50,000 installations.

You want to allow Cloud Dataflow to scale its compute power up as required.

Which Cloud Dataflow pipeline configuration setting should you update?

- A. The zone
- B. The number of workers
- C. The disk size per worker
- D. The maximum number of workers

Correct Answer: D

---

### QUESTION 2

Your company is in a highly regulated industry. One of your requirements is to ensure individual users have access only to the minimum amount of information required to do their jobs. You want to enforce this requirement with Google BigQuery. Which three approaches can you take? (Choose three.)

- A. Disable writes to certain tables.
- B. Restrict access to tables by role.
- C. Ensure that the data is encrypted at all times.
- D. Restrict BigQuery API access to approved users.
- E. Segregate data across multiple tables or databases.
- F. Use Google Stackdriver Audit Logging to determine policy violations.

Correct Answer: BDF

---

### QUESTION 3

You receive data files in CSV format monthly from a third party. You need to cleanse this data, but every third month the schema of the files changes. Your requirements for implementing these transformations include:

1.  
Executing the transformations on a schedule
- 2.



Enabling non-developer analysts to modify transformations

3.

Providing a graphical tool for designing transformations What should you do?

- A. Use Cloud Dataprep to build and maintain the transformation recipes, and execute them on a scheduled basis
- B. Load each month's CSV data into BigQuery, and write a SQL query to transform the data to a standard schema. Merge the transformed tables together with a SQL query
- C. Help the analysts write a Cloud Dataflow pipeline in Python to perform the transformation. The Python code should be stored in a revision control system and modified as the incoming data's schema changes
- D. Use Apache Spark on Cloud Dataproc to infer the schema of the CSV file before creating a Dataframe. Then implement the transformations in Spark SQL before writing the data out to Cloud Storage and loading into BigQuery

Correct Answer: A

you can use dataprep for continuously changing target schema In general, a target consists of the set of information required to define the expected data in a dataset. Often referred to as a "schema," this target schema information can include:

Names of columns Order of columns Column data types Data type format Example rows of data

A dataset associated with a target is expected to conform to the requirements of the schema. Where there are differences between target schema and dataset schema, a validation indicator (or schema tag) is displayed.

[https://cloud.google.com/dataprep/docs/html/Overview-of-RapidTarget\\_136155049](https://cloud.google.com/dataprep/docs/html/Overview-of-RapidTarget_136155049)

#### QUESTION 4

Which of these sources can you not load data into BigQuery from?

- A. File upload
- B. Google Drive
- C. Google Cloud Storage
- D. Google Cloud SQL

Correct Answer: D

You can load data into BigQuery from a file upload, Google Cloud Storage, Google Drive, or Google Cloud Bigtable. It is not possible to load data into BigQuery directly from Google Cloud SQL. One way to get data from Cloud SQL to BigQuery would be to export data from Cloud SQL to Cloud Storage and then load it from there.

Reference: <https://cloud.google.com/bigquery/loading-data>

#### QUESTION 5

Which of the following statements about Legacy SQL and Standard SQL is not true?



- A. Standard SQL is the preferred query language for BigQuery.
- B. If you write a query in Legacy SQL, it might generate an error if you try to run it with Standard SQL.
- C. One difference between the two query languages is how you specify fully-qualified table names (i.e. table names that include their associated project name).
- D. You need to set a query language for each dataset and the default is Standard SQL.

Correct Answer: D

You do not set a query language for each dataset. It is set each time you run a query and the default query language is Legacy SQL.

Standard SQL has been the preferred query language since BigQuery 2.0 was released.

In legacy SQL, to query a table with a project-qualified name, you use a colon, :, as a separator. In standard SQL, you use a period, ., instead.

Due to the differences in syntax between the two query languages (such as with project-qualified table names), if you write a query in Legacy SQL, it might generate an error if you try to run it with Standard SQL.

Reference:

<https://cloud.google.com/bigquery/docs/reference/standard-sql/migrating-from-legacy-sql>

[PROFESSIONAL-DATA-ENGINEER PDF Dumps](#)

[PROFESSIONAL-DATA-ENGINEER Study Guide](#)

[PROFESSIONAL-DATA-ENGINEER Exam Questions](#)