



PR PROFESSIONAL-DATA-ENGINEER^{Q&As}

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



VCE & PDF

GeekCert.com

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

2024 Latest geekcert PROFESSIONAL-DATA-ENGINEER PDF and VCE
dumps Download

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





QUESTION 1

Which of the following statements is NOT true regarding Bigtable access roles?

- A. Using IAM roles, you cannot give a user access to only one table in a project, rather than all tables in a project.
- B. To give a user access to only one table in a project, grant the user the Bigtable Editor role for that table.
- C. You can configure access control only at the project level.
- D. To give a user access to only one table in a project, you must configure access through your application.

Correct Answer: B

For Cloud Bigtable, you can configure access control at the project level. For example, you can grant the ability to:

1.

Read from, but not write to, any table within the project.

2.

Read from and write to any table within the project, but not manage instances.

3.

Read from and write to any table within the project, and manage instances. Reference:
<https://cloud.google.com/bigtable/docs/access-control>

QUESTION 2

Your company uses a proprietary system to send inventory data every 6 hours to a data ingestion service in the cloud. Transmitted data includes a payload of several fields and the timestamp of the transmission. If there are any concerns about a transmission, the system re-transmits the data. How should you deduplicate the data most efficiently?

- A. Assign global unique identifiers (GUID) to each data entry.
- B. Compute the hash value of each data entry, and compare it with all historical data.
- C. Store each data entry as the primary key in a separate database and apply an index.
- D. Maintain a database table to store the hash value and other metadata for each data entry.

Correct Answer: A

QUESTION 3

When using Cloud Dataproc clusters, you can access the YARN web interface by configuring a browser to connect through a ____ proxy.



- A. HTTPS
- B. VPN
- C. SOCKS
- D. HTTP

Correct Answer: C

When using Cloud Dataproc clusters, configure your browser to use the SOCKS proxy. The SOCKS proxy routes data intended for the Cloud Dataproc cluster through an SSH tunnel. Reference:
<https://cloud.google.com/dataproc/docs/concepts/cluster-web-interfaces#interfaces>

QUESTION 4

An organization maintains a Google BigQuery dataset that contains tables with user-level data. They want to expose aggregates of this data to other Google Cloud projects, while still controlling access to the user-level data. Additionally, they need to minimize their overall storage cost and ensure the analysis cost for other projects is assigned to those projects. What should they do?

- A. Create and share an authorized view that provides the aggregate results.
- B. Create and share a new dataset and view that provides the aggregate results.
- C. Create and share a new dataset and table that contains the aggregate results.
- D. Create dataViewer Identity and Access Management (IAM) roles on the dataset to enable sharing.

Correct Answer: A

Reference: <https://cloud.google.com/bigquery/docs/access-control>

QUESTION 5

You work for a large fast food restaurant chain with over 400,000 employees. You store employee information in Google BigQuery in a Users table consisting of a FirstName field and a LastName field. A member of IT is building an application and asks you to modify the schema and data in BigQuery so the application can query a FullName field consisting of the value of the FirstName field concatenated with a space, followed by the value of the LastName field for each employee. How can you make that data available while minimizing cost?

- A. Create a view in BigQuery that concatenates the FirstName and LastName field values to produce the FullName.
- B. Add a new column called FullName to the Users table. Run an UPDATE statement that updates the FullName column for each user with the concatenation of the FirstName and LastName values.
- C. Create a Google Cloud Dataflow job that queries BigQuery for the entire Users table, concatenates the FirstName value and LastName value for each user, and loads the proper values for FirstName, LastName, and FullName into a new table in BigQuery.
- D. Use BigQuery to export the data for the table to a CSV file. Create a Google Cloud Dataproc job to process the CSV file and output a new CSV file containing the proper values for FirstName, LastName and FullName. Run a BigQuery load job to load the new CSV file into BigQuery.



Correct Answer: B

BigQuery has no quota on the DML statements. (Search Google - does bigquery have quota for update).

Why not C: This is a one time activity and SQL is the easiest way to program it. DataFlow is way overkill for this. You will need to find an engineer who can develop DataFlow pipelines. Whereas, SQL is so much more widely known and easier. One of the great features about BigQuery is its SQL interface. Even for BigQueryML services.

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

[PROFESSIONAL-DATA-ENGINEER Practice Test](#)

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