🔍

⊘ **Google Discussions**

Exam Professional Data Engineer All Questions

View all questions & answers for the Professional Data Engineer exam

Go to Exam

📄 **EXAM PROFESSIONAL DATA ENGINEER TOPIC 1 QUESTION 10 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 10

Topic #: 1

[All Professional Data Engineer Questions]

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.

Show Suggested Answer

by [deleted] at *March 15, 2020, 8:44 a.m.*

Comments

Type your comment...

Submit

samdhimal `Highly Voted` 3 years, 3 months ago

correct option -> B. Restrict access to tables by role.
Reference: https://cloud.google.com/bigquery/docs/table-access-controls-intro

correct option -> D. Restrict BigQuery API access to approved users.
Only approved users will have access which means other users will have minimum amount of information required to do their job.
Reference: https://cloud.google.com/bigquery/docs/access-control

correct option -> F. Use Google Stackdriver Audit Logging to determine policy violations.
Reference: https://cloud.google.com/bigquery/docs/table-access-controls-intro#logging

A. Disable writes to certain tables. ---> Read is still available(not minimal access)
C. Ensure that the data is encrypted at all times. ---> Data is encrypted by default.
E. Segregate data across multiple tables or databases. ---> Normalization is of no help here.

👍 ↩ ⚑ upvoted 39 times

---

samdhimal 7 months, 1 week ago

I was WRONG. I am not sure why s o many upvotes lol.

I think this is the correct answer:
B. Restrict access to tables by role.
D. Restrict BigQuery API access to approved users.
E. Segregate data across multiple tables or databases.

Restrict access to tables by role: You can use BigQuery's access controls to restrict access to specific tables based on user roles. This allows you to ensure that users can only access the data they need to do their job.
Restrict BigQuery API access to approved users: By using Cloud Identity and Access Management (IAM) you can control who has access to the BigQuery API, and what actions they are allowed to perform. This will help to ensure that only authorized users can access the data.
Segregate data across multiple tables or databases: You can use multiple tables or databases to separate different types of data, so that users can only access the data they need. This will prevent users from seeing data they shouldn't have access to.

👍 ↩ ⚑ upvoted 27 times

---

directtoking 4 months ago

If there is "Restrict access to tables by role", then what is the requirement for "Segregate data across multiple tables or databases"?

👍 ↩ ⚑ upvoted 1 times

---

samdhimal 2 years, 3 months ago

Option A is incorrect because disabling writes to certain tables would prevent users from updating the data which is not in line with the goal of providing access to the minimum amount of information required to do their jobs.
Option C is incorrect because while data encryption is important for security it doesn't specifically help with providing users access to the minimum amount of information required to do their jobs.
Option F is incorrect because while Google Stackdriver Audit Logging can help to determine policy violations it does not help to enforce the access controls and segregation of data.

👍 ↩ ⚑ upvoted 6 times

[Removed] 2 years, 2 months ago

There is no database in Bigquery, only datasets. I would pick it if it says "tables and datasets".

👍 ↩ ⚑ upvoted 4 times

---

billalltf 11 months, 3 weeks ago

What about cloud SQL and Spanner

👍 ↩ ⚑ upvoted 1 times

IsaB `Highly Voted` 4 years, 8 months ago

Yes. Access control on table level is now possible in BigQuery : https://cloud.google.com/bigquery/docs/table-access-controls-intro

👍 ↩ ⚑ upvoted 15 times

---

axantroff 1 year, 6 months ago

Thanks. For me, this type of answer is more valuable because even as time passes, I can revisit existing solutions and ideas and refresh the concepts of the initial question. It helped me on the ACE exam

👍 ↩ ⚑ upvoted 1 times

---

cqrm3n `Most Recent` 3 months, 3 weeks ago

Selected Answer: BDE

B - Use IAM to define granular permissions.
D - Only authorised users or systems can query or manipulate BigQuery data.
E - By segregating data into different tables or datasets, specific permissions can be assigned to each data subset.

👍 ↩ 🏳 upvoted 1 times

☐ 👤 **NicolasN** 7 months, 1 week ago

Selected Answer: BDE

I disagree with [F]. It's too late for a "highly regulated industry" to detect access violations by audit logs.
[E] is a more reasonable answer, since it is a kind of row-level security, especially the times when BigQuery row-level security wasn't available.
It is a practice still recommended (even with row-level sec. available) for the extreme scenario that:
(Through repeated observation of query duration when querying tables with row-level access policies,) "a user could infer the values of rows that otherwise might be protected by row-level access policies"
"If you are sensitive to this level of protection, we recommend using separate tables to isolate rows with different access control requirements, instead."
Source:
https://cloud.google.com/bigquery/docs/best-practices-row-level-security#limit-side-channel-attacks

👍 ↩ 🏳 upvoted 3 times

☐ 👤 **Asheesh1909** 7 months, 1 week ago

Selected Answer: BDE

E seems more sensible to be as the question concentrates more on table access restriction than access violation, policy violations can only be determined through stackdriver how ever we cant restrict the access to tables.Probably option E should be considered as, by segregating the data into diferrent tables , we can restrict access to tables.

👍 ↩ 🏳 upvoted 3 times

  ☐ 👤 **odacir** 2 years, 4 months ago

  E says segregate across multiple tables or databases, this is not the pattern of BigQuery, in BQ there is only one database, and you can organize your data in datasets...

  👍 ↩ 🏳 upvoted 1 times

☐ 👤 **rtcpost** 7 months, 1 week ago

Selected Answer: BDE

B. Restrict access to tables by role: You can define roles in BigQuery and grant specific permissions to these roles to control who can access particular tables.

D. Restrict BigQuery API access to approved users: You can control access to the BigQuery API and, consequently, to the underlying data by ensuring that only approved users or services can make API requests.

E. Segregate data across multiple tables or databases: You can separate data into different tables or databases based on user access requirements, which allows you to limit users' access to specific data sets.

These approaches, when used together, can help you enforce data access controls in a regulated environment. Options A, C, and F are also important considerations but are not direct methods for enforcing fine-grained access control to specific data.

👍 ↩ 🏳 upvoted 4 times

☐ 👤 **SatyamKishore** 7 months, 1 week ago

B. Restrict access to tables by role.

Use IAM roles and permissions to control access to specific datasets or tables based on the user's role.
D. Restrict BigQuery API access to approved users.

Limit API access to only those users or services that need it, ensuring that unauthorized users cannot interact with the data.
E. Segregate data across multiple tables or databases.

Organize data in a way that separates sensitive information, allowing more granular control over who has access to specific datasets.
These options directly contribute to enforcing the principle of least privilege, ensuring users can only access the data necessary for their roles.

👍 ↩ 🏳 upvoted 1 times

☐ 👤 **MaxNRG** 1 year, 4 months ago

Selected Answer: BDE

You want to enforce this requirement with Google BigQuery -> BDE

👍 ↩ 🏳 upvoted 1 times

☐ 👤 **RT_G** 1 year, 5 months ago

Selected Answer: BDF

BDF. We are fairly unanimous with options B and D. I'm going with F because it does help identifying policy violations which

is also one aspect to be considered when designing access controls. Option D only indicates segregating into multiple tables and databases which may or may not help with controlling access leaving it open-ended for the architect to decide.

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **rocky48** 1 year, 6 months ago

Selected Answer: BDF

In Google BigQuery, you can organize and segregate data across multiple tables within the same dataset, but you cannot directly segregate data into separate databases. BigQuery uses a flat namespace structure where data is organized into datasets and tables within those datasets. Datasets are the highest level of organization within BigQuery.

So i'm sticking with BDF

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **RheaZzang** 1 year, 6 months ago

Selected Answer: BDE

B. Restrict access to tables by role.
D. Restrict BigQuery API access to approved users.
E. Segregate data across multiple tables or databases.

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **AnonymousPanda** 1 year, 8 months ago

Selected Answer: BDF

BDF as per other answers

👍 ↩ 🚩 upvoted 2 times

☐ 👤 **nescafe7** 1 year, 9 months ago

Selected Answer: BDF

Regarding E or F, opinions seem to be divided into two parts.

I think E is insufficient because it seems that appropriate conditions must be additionally described for table or dataset separation.

F is also emphasized in Google's official textbook. You need to ensure that it is operating well as set up through monitoring.

So, BDF!

👍 ↩ 🚩 upvoted 4 times

☐ 👤 **Liting** 1 year, 9 months ago

Selected Answer: DEF

Why B is correct? Access control can only be applied on dataset and views, not on partitions and tables. => So it is not possible to restrict access to table, but only to dataset. Can someone help me understand why in this scenario B is correct?

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **FP77** 1 year, 9 months ago

I was thinking the same thing. I thought dataset access gave you access to all tables within it, and that you couldn't restrict access on the table level.

👍 ↩ 🚩 upvoted 1 times

☐ 👤 **KK0202** 1 year, 10 months ago

Selected Answer: BDF

Option E says "...or databases". The data housing service in question in BigQuery and the context is to design that support BigQuery access delegation. Seems random to include moving to another database as an option. If it did not mention databases and stopped at just tables, then E would also be the right option

👍 ↩ 🚩 upvoted 2 times

☐ 👤 **Oleksandr0501** 2 years ago

B. Restrict access to tables by role: This approach can be used to control access to tables based on user roles. Access controls can be set at the project, dataset, and table level, and roles can be customized to provide granular access controls to different groups of users.

D. Restrict BigQuery API access to approved users: This approach involves using IAM (Identity and Access Management) to control access to the BigQuery API. Access can be granted or revoked at the project or dataset level, and policies can be customized to control access based on user roles, IP addresses, and other factors.

E. Segregate data across multiple tables or databases: This approach involves breaking down large datasets into smaller, more manageable tables or databases. This helps to ensure that individual users have access only to the minimum amount of information required to do their jobs, and reduces the risk of data breaches or policy violations.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **juljobs** 2 years, 1 month ago

juliobs 2 years, 1 month ago

F won't avoid undesired access, only detect after it already happened.
E makes it easier to control access.

👍 ↩ 🚩 upvoted 5 times

Load full discussion...