⬅ **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 75 DISCUSSION**

Actual exam question from Google's Professional Data Engineer

Question #: 75

Topic #: 1

[All Professional Data Engineer Questions]

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.

Show Suggested Answer

by 👤 **AWSandeep** at *Sept. 3, 2022, 1:51 p.m.*

Comments

Type your comment...

Submit

⊟ 👤 **midgoo** [Highly Voted 👍] 1 year, 2 months ago

A is the answer. Don't be confused by the documentation saying "Authorized views should be created in a different dataset". It is a best practice but not a technical requirement. And we don't create a new dataset for each authorized view. If you are not clear on this, try in the system, don't just read the documentation without understanding.

B is wrong when saying we must SHARE Dataset. Although creating a dataset and view in it will not incur extra cost, but sharing dataset is something we always try not to do.

At for the project that run the query it the project to be billed, that is standard behaviour. View only give access to data, whoever run the view will need pay for the query cost

👍 ↩ 🚩 upvoted 17 times

⊟ 👤 **DAYAGOWDA** 1 year, 2 months ago

https://cloud.google.com/bigquery/docs/authorized-views#:~:text=An%20authorized%20view%20and%20authorized,users%20are%20able%20to%20query.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **Yiouk** 9 months, 3 weeks ago

Have to consider where the billing goes to:
https://stackoverflow.com/questions/52201034/bigquery-authorized-view-cost-billing-account
hence anwer is B

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **Mathew106** 9 months, 2 weeks ago

Did you even read the answer in the SO link you shared?
Part of the answer is below:
"""After a deeper investigation and some test scenarios, I have confirmed that the billing charges related to the query jobs are applied to the Billing account associated to the project that executes the query; however, the view owner keeps getting the charges related to the storage of the source data."""

Soo, if you create an authorized view, the users from the other project that has access to the view will get billed for the querying.

The only reason to pick up B over A is that it's the recommended approach to store views in a different dataset than the base data.

👍 ↩ 🚩 upvoted 4 times

⊟ 👤 **[Removed]** `Highly Voted 👍` 1 year, 6 months ago

The link on authorized views (https://cloud.google.com/bigquery/docs/share-access-views) explicitly states "Authorized views should be created in a different dataset from the source data. That way, data owners can give users access to the authorized view without simultaneously granting access to the underlying data." therefore B is the correct answer because we are to create a new dataset and view within that dataset.

👍 ↩ 🚩 upvoted 13 times

⊟ 👤 **desertlotus1211** `Most Recent ⊘` 1 month, 3 weeks ago

When other projects query the authorized view, the query costs are billed to the project that runs the query

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **loki82** 3 months, 2 weeks ago

I think both A and B are totally valid answers, making this a fairly stupid question. But as a DBA, B would be easier to implement, easier to manage, easier to audit. So even if it's the wrong answer, it's still the right solution, so I can't help but choose B.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **desertlotus1211** 1 month, 3 weeks ago

this approach typically involves duplicating data or managing a separate dataset that isn't as tightly controlled, Answer A is better...

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **inamm** 4 months, 1 week ago

B is the correct Ans

You are correct in noting that creating a view within an existing dataset could potentially expose other tables within that dataset if the dataset-level permissions are not carefully managed. To ensure that only the aggregate results are shared and to avoid inadvertently exposing other tables, it is indeed a good practice to create a new dataset specifically for the view.

Revised Answer:

B. Create and share a new dataset and view that provides the aggregate results.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **rocky48** 5 months, 1 week ago

Selected Answer: A

A. Create and share an authorized view that provides the aggregate results.

An authorized view is a BigQuery feature that allows you to share only a specific subset of data from a table, while still keeping the original data private. This way, the organization can expose only the aggregate data to other projects, while still controlling access to the user-level data. By using an authorized view, the organization can minimize their overall storage cost as the aggregate data takes up less storage space than the original data. Additionally, by using authorized view, the analysis cost for other projects is assigned to those projects.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **odiez3** 9 months, 2 weeks ago

I thing that is B because for security yo need to create a new data set when share a view, apart when yo grant access the top level is a data set if you share a view in same dataset that you have your tables, that access can see all tables inside dataset.

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **baht** 10 months, 3 weeks ago

Selected Answer: B

"Authorized views should be created in a different dataset from the source data. That way, data owners can give users access to the authorized view without simultaneously granting access to the underlying data."
https://cloud.google.com/bigquery/docs/share-access-views?hl=en#console_5

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **[Removed]** 1 year ago

Selected Answer: A
B is wrong by itself. why do you need to create a view, if you have already created an aggregated dataset?

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **MrMone** 1 year, 1 month ago

Selected Answer: A

"they need to minimize their overall storage cost". Also, you are sharing the aggregate's results, not the underlying table

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **bha11111** 1 year, 1 month ago

Selected Answer: A

minimize cost so view

👍 ↩ 🚩 upvoted 1 times

⊟ 👤 **Paritosh07** 1 year, 2 months ago

Selected Answer: A

A should be the answer, as we need to separate costs according to projects. As in the following SO question (and the attached google resources), the 'project that runs the queries is the project that gets billed.'
So we can generate a view and give it's access to the other project to run the analysis
https://stackoverflow.com/questions/52201034/bigquery-authorized-view-cost-billing-account

👍 ↩ 🚩 upvoted 3 times

⊟ 👤 **musumusu** 1 year, 2 months ago

I will go with A, as i wanna save cost, dont need to create separate dataset for permanent storage.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **samdhimal** 1 year, 3 months ago

A. Create and share an authorized view that provides the aggregate results.

An authorized view is a BigQuery feature that allows you to share only a specific subset of data from a table, while still keeping the original data private. This way, the organization can expose only the aggregate data to other projects, while still controlling access to the user-level data. By using an authorized view, the organization can minimize their overall storage cost as the aggregate data takes up less storage space than the original data. Additionally, by using authorized view, the analysis cost for other projects is assigned to those projects.

👍 ↩ 🚩 upvoted 2 times

⊟ 👤 **samdhimal** 1 year, 3 months ago

B. Creating and sharing a new dataset and view that provides the aggregate results is also a correct option but not as optimal as authorized view as it creates a copy of the data and increases the storage costs.
C. Creating and sharing a new dataset and table that contains the aggregate results is also a correct option but not as optimal as authorized view as it creates a copy of the data and increases the storage costs.
D. Creating dataViewer Identity and Access Management (IAM) roles on the dataset to enable sharing is not the best

D. Creating dataViewer Identity and Access Management (IAM) roles on the dataset to enable sharing is not the best option as it would give access to the user-level data, not just the aggregate data.

👍 ↩ 🚩 upvoted 2 times

☐ 👤 **Rupendra06** 1 year, 3 months ago

Ensure the analysis cost for other projects is assigned to those projects indicates B is the correct answer.

👍 ↩ 🚩 upvoted 2 times

☐ 👤 **GCPpro** 1 year, 3 months ago

B is the correct answer

👍 ↩ 🚩 upvoted 2 times

☐ 👤 **Kyr0** 1 year, 4 months ago

I would say 1 too

👍 ↩ 🚩 upvoted 1 times

Load full discussion...