

🔗 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 26 DISCUSSION

Actual exam question from Google's Professional Data Engineer

Question #: 26

Topic #: 1

[\[All Professional Data Engineer Questions\]](#)

You are working on a sensitive project involving private user data. You have set up a project on Google Cloud Platform to house your work internally. An external consultant is going to assist with coding a complex transformation in a Google Cloud Dataflow pipeline for your project. How should you maintain users' privacy?

- A. Grant the consultant the Viewer role on the project.
- B. Grant the consultant the Cloud Dataflow Developer role on the project.
- C. Create a service account and allow the consultant to log on with it.
- D. Create an anonymized sample of the data for the consultant to work with in a different project.

Show Suggested Answer

by [jvg637](#) at March 11, 2020, 7:18 p.m.

### Comments

Type your comment...

Submit

🗨️ [jvg637](#) Highly Voted 5 years, 1 month ago

The Answer should be B. The Dataflow developer role will not provide access to the underlying data.

👍 🗳️ 🗳️ upvoted 76 times

upvoted 70 times

  **willymac2** 2 years, 10 months ago

The answer should be D.  
You do not need any DataFlow permission to implement a pipeline.  
If needed, you can test using the DirectRunner which runs locally:

[https://cloud.google.com/dataflow/docs/concepts/access-control#example\\_role\\_assignment](https://cloud.google.com/dataflow/docs/concepts/access-control#example_role_assignment)

   upvoted 3 times

  **willymac2** 2 years, 10 months ago

Sorry I did a wrong copy/paste on the link, I wanted to send:

[https://cloud.google.com/dataflow/docs/concepts/security-and-permissions#security\\_and\\_permissions\\_for\\_local\\_pipelines](https://cloud.google.com/dataflow/docs/concepts/security-and-permissions#security_and_permissions_for_local_pipelines)



<https://cloud.google.com/dataflow/docs/guides/setting-pipeline-options#LocalExecution>

   upvoted 2 times

  **cleroy** 5 years, 1 month ago



Remember he's an external consultant. You need to create a service account for him, you can't grant before that... I think C is correct in this case.

   upvoted 5 times

  **Rajuuu** 4 years, 10 months ago



Service account is between applications and non human entry.

   upvoted 36 times

  **Tanzu** 3 years, 3 months ago

u can enable a service account as user so that externals can use to login

   upvoted 1 times

  **Tanzu** 3 years, 3 months ago

u can enable a service account as user so that externals can use to login.

but the problem is service account is about login. not the minimum resources to do the dataflow related stuffs. so C is not enough !.

so the answer should be B.

if the question was about "doing the 1st thing", then yeah may be creating a service account could be the 1st thing.

   upvoted 3 times

  **VincentMenzel** 1 year, 9 months ago

Im not sure how you expect the consultant to implement a pipeline without having access to any data that is being processed. Having test data is a prerequisite.

   upvoted 4 times

  **ThorstenStaerk** 2 years ago

and now? For seeing test data, (D) would be right. And the system tells me (C) is the right answer. What shall I click in the exam?

   upvoted 5 times

  **[Removed]**  5 years, 1 month ago

Answer: B

Description: Provides the permissions necessary to execute and manipulate Dataflow jobs.

   upvoted 18 times

  **Abizi**  2 months ago

**Selected Answer: D**

logical answer for me


   upvoted 1 times

  **baimus** 7 months, 2 weeks ago

**Selected Answer: D**

The answer cannot be B, because B is too restrictive, it can only create and manage dataflow jobs, but cannot view data. I acknowledge that is secure, but no consultant can do the job without seeing representative test data. D is the only one that provides enough to do the job, while still remaining totally private.

   upvoted 5 times

  **mouthwash** 7 months, 2 weeks ago

D cannot be the answer because the question clearly states the developer has to work in your project. Creating another project is not in scope and is a waste of time. Correct answer is B. Developer role has developer rights only, no view rights.

   upvoted 2 times

  **iooj** 9 months, 1 week ago

**Selected Answer: D**

A. Grant the consultant the Viewer role on the project.

This role provides read-only access to all resources in the project, which could expose sensitive data to the consultant, violating privacy principles.

B. Grant the consultant the Cloud Dataflow Developer role on the project.

This role allows the consultant to create and manage Dataflow jobs but does not give them access the underlying data, it is not sufficient, the developer still needs data.

- C. Create a service account and allow the consultant to log on with it.

Allowing the consultant to log on with a service account could grant them access to sensitive data if the service account has broad permissions. This approach does not address the need to limit data exposure.

- D. Create an anonymized sample of the data for the consultant to work with in a different project.

- this fits the requirements

   upvoted 3 times

  **Shash\_88** 1 year, 1 month ago

D

B is a good option to maintain privacy of sensitive data, but he also need some test data to validate the transformation logic right, so creating sample data and allow him to test in another project seems good.

   upvoted 1 times

  **hamzad\_basha** 1 year, 2 months ago

**Selected Answer: B**

Data flow data privacy rules cant allow the developer to see what the data, He/she just designs the pipelines and the flow as the interdependent tasks for the composer

   upvoted 1 times

  **MaxNRG** 1 year, 4 months ago

**Selected Answer: B**

B as the Dataflow developer role would help provide the third-party consultant access to create and work on the Dataflow pipeline. However, it does not provide access to view the data, thus maintaining user's privacy.

Refer GCP documentation - Dataflow roles:

<https://cloud.google.com/dataflow/docs/concepts/access-control#roles>

Option A is wrong as it would not allow the consultant to work on the pipeline.

Option C is wrong as the consultant cannot use the service account to login.


Option D is wrong as it does not enable collaboration.

   upvoted 3 times

  **Jconnor** 1 year, 5 months ago

C and A will not maintain user's privacy so out. B without data will be enough. D will give a good sample data, maintain privacy and the consultant will help creating the dataflow pipe for the project as requested. so D.

   upvoted 1 times

  **axantroff** 1 year, 5 months ago

**Selected Answer: D**

I follow the corresponding logic choosing between B and D:

Yes, with the Dataflow Developer role it is possible to execute and manipulate Dataflow jobs, but do we need to execute it?

Based on my understanding we only need to ask for help to write it down. Is it possible without having access to test the data? I don't think so. At the same time, we need to perform an anonymization on it. So the answer D is more appropriate for me

   upvoted 1 times

  **rocky48** 1 year, 6 months ago

**Selected Answer: D**

By creating an anonymized sample of the data, you can provide the consultant with a realistic dataset that doesn't contain sensitive or private information. This way, the consultant can work on the project without direct access to sensitive data, reducing privacy risks.

Options A and B involve granting the consultant access to the project, which may expose sensitive data, even if they have limited permissions.

Option C involves creating a service account, but it doesn't address the need to anonymize the data or provide a separate,

safe environment for the consultant to work with.

Option D provides a controlled environment that allows the consultant to work effectively while maintaining data privacy.

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **rtcpost** 1 year, 6 months ago

**Selected Answer: D**

D. Create an anonymized sample of the data for the consultant to work within a different project.

By creating an anonymized sample of the data, you can provide the consultant with a realistic dataset that doesn't contain sensitive or private information. This way, the consultant can work on the project without direct access to sensitive data, reducing privacy risks.

Options A and B involve granting the consultant access to the project, which may expose sensitive data, even if they have limited permissions.

Option C involves creating a service account, but it doesn't address the need to anonymize the data or provide a separate, safe environment for the consultant to work with.

Option D provides a controlled environment that allows the consultant to work effectively while maintaining data privacy.

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **imran79** 1 year, 7 months ago

D. Creating an anonymized sample of the data for the consultant to work with in a different project is the safest option. This way, the consultant can develop and test the transformation logic without accessing the real, sensitive data.

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **ruben82** 1 year, 6 months ago

The question says "with coding a complex transformation", so I don't think that a sample of data is enough. I think that the most suitable way is C, 'cos with a service account you can handle access fine-grained

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **navioshi** 1 year, 7 months ago

I think C would be correct, as the question says external consultants want to do some work and how we can maintain the 'external consultant' user privacy. Question didn't mention about the company user data or customer information.

👍 ↩ 🚩 upvoted 2 times

🗨️ 👤 **hxy8** 1 year, 8 months ago

Answer: C

👍 ↩ 🚩 upvoted 1 times

🗨️ 👤 **madhu15** 1 year, 8 months ago

Dataflow Developer  
(roles/dataflow.developer)

Provides the permissions necessary to execute and manipulate Dataflow jobs.

👍 ↩ 🚩 upvoted 1 times

[Load full discussion...](#)



## Platform

> [Home](#)

> [Examtopics PRO](#)

> [All Exams](#)

> [Training Courses](#)

