EXAMTOPICS

- Expert Verified, Online, **Free**.

 MENU

Google Discussions

Exam Associate Cloud Engineer All Questions
View all questions & answers for the Associate Cloud Engineer exam

Go to Exam

 EXAM ASSOCIATE CLOUD ENGINEER TOPIC 1 QUESTION 93 DISCUSSION

Actual exam question from Google's Associate Cloud Engineer

Question #: 93

Topic #: 1

[All Associate Cloud Engineer Questions]

Your company uses BigQuery for data warehousing. Over time, many different business units in your company have created 1000+ datasets across hundreds of projects. Your CIO wants you to examine all datasets to find tables that contain an employee_ssn column. You want to minimize effort in performing this task.
What should you do?

A. Go to Data Catalog and search for employee_ssn in the search box.

B. Write a shell script that uses the bq command line tool to loop through all the projects in your organization.

C. Write a script that loops through all the projects in your organization and runs a query on INFORMATION_SCHEMA.COLUMNS view to find the employee_ssn column.

D. Write a Cloud Dataflow job that loops through all the projects in your organization and runs a query on INFORMATION_SCHEMA.COLUMNS view to find employee_ssn column.

Show Suggested Answer

by poogcp at *June 10, 2020, 12:59 p.m.*

Comments

Type your comment...

poogcp `Highly Voted` 4 years, 4 months ago

Its A.

upvoted 40 times

filco72 `Highly Voted` 4 years, 3 months ago

Correct is A.
I tested on my account following this procedure: https://cloud.google.com/bigquery/docs/quickstarts/quickstart-web-ui?authuser=4
I created a data set and through Data Catalog I easily and effortlessly searched for the column name "gender"

upvoted 32 times

JB28 `Most Recent` 9 months, 2 weeks ago

The best option to minimize effort in performing this task is A. Go to Data Catalog and search for employee_ssn in the search box.

Google Cloud's Data Catalog is a fully managed and scalable metadata management service that empowers organizations to quickly discover, manage, and understand all their data in Google Cloud. It offers a simple and easy-to-use search interface for data discovery. By searching for "employee_ssn" in the Data Catalog, you can quickly find all tables across all datasets and projects that contain this column. This approach is more efficient and requires less effort compared to writing and maintaining scripts or jobs.

Please note that access to Data Catalog and the visibility of datasets, tables, and columns are subject to permissions and roles in IAM policy. Make sure you have the necessary permissions to view the metadata.

upvoted 3 times

ogerber 10 months, 3 weeks ago

Selected Answer: A

Data Catalog lets you search and tag entries such as BigQuery tables with metadata. Some examples of metadata that you can use for tagging include public and private tags, data stewards, and rich text overview.

https://cloud.google.com/data-catalog/docs/tag-bigquery-dataset

upvoted 4 times

Vik96 11 months, 2 weeks ago

I am preparing for the GCP-ACE exam, I was able to access 96 questions only, if anyone has the entire questions please share them with my vittoriaprovenza@tiscali.it address. I have exam on next week,pls share Thanks in advance. I would be forever grateful.

upvoted 2 times

BAofBK 11 months, 2 weeks ago

I will go for A

upvoted 1 times

BAofBK 11 months, 2 weeks ago

The correct answer is A

upvoted 1 times

Captain1212 1 year, 1 month ago

Selected Answer: A

Answer is A, as it requires the less effort and other options are more time consuming and error prone

upvoted 1 times

deadsong 1 year, 5 months ago

The most efficient approach to identify tables that contain an employee_ssn column in BigQuery would be to query the INFORMATION_SCHEMA.COLUMNS view, which provides metadata about all columns in all tables in a given dataset. Therefore, options C and D are both possible solutions.

Option A, searching for the column name in Data Catalog, may not be efficient if there are too many datasets to search through manually.

Option B, writing a shell script to loop through all the projects in your organization, may work, but it would require more effort and time than options C and D. Also, it would be more error-prone since the script would need to handle authentication and authorization, handle exceptions and errors, and collect the results.

Therefore, options C and D are better choices, but option D, using Cloud Dataflow, might be overkill for this specific task. Option C, looping through all projects and querying INFORMATION_SCHEMA.COLUMNS view, is the simplest and most effective solution to minimize effort.

upvoted 1 times

Neeyo 1 year, 6 months ago

Hi All, I have my GCP ACE exam scheduled for tomorrow. However, I am only being able to access 96 questions. Can anyone kindly share the entire list of questions as I have hardly anytime left before my exam. oniyi6@yahoo.com. Thank you all so much

upvoted 1 times

---

arnika98 1 year, 5 months ago

Did you pass the exam? If so any questions from here came? Please let us know so that it will be helpful

upvoted 1 times

R4F 1 year, 7 months ago

C is IMO the correct answer (https://stackoverflow.com/questions/68746567/big-query-find-all-column-name-containing-surname-across-all-tables)

upvoted 1 times

Buruguduystunstugudunstuy 1 year, 8 months ago

Selected Answer: A

Answer A is the correct answer. Go to Data Catalog and search for employee_ssn in the search box.

Data Catalog is a fully managed and scalable metadata management service that allows you to discover, understand, and manage your data. It provides search functionality that allows you to search for datasets, tables, columns, and other metadata across your organization. Therefore, you can simply go to Data Catalog and search for "employee_ssn" in the search box to find all datasets that contain this column. This is the most efficient and straightforward solution to the problem.

Answers B, C, and D are not ideal solutions.

Answer B requires writing a shell script and using the bq command line tool to loop through all the projects, which is time-consuming and error-prone.

Answer C requires writing a script that loops through all the projects and runs a query on INFORMATION_SCHEMA.COLUMNS view, which is also time-consuming and error-prone.

Answer D involves writing a Cloud Dataflow job, which is unnecessary and OVERKILL for this simple task.

upvoted 8 times

Kirangm 1 year, 8 months ago

The answer is A

upvoted 1 times

Ushnishm 1 year, 9 months ago

Hello, I have my GCP ACE exam scheduled early next week. However, I am only being able to access 96 questions. Can anyone kindly share the entire list of questions as I have hardly anytime left before my exam.

upvoted 2 times

---

leogor 1 year, 11 months ago

Selected Answer: A

Data Catalog

upvoted 1 times

Charumathi 2 years ago

A is the correct answer, Data Catalog can be used to search the column with keyword:value pair,

Filter your search by adding a keyword:value to your search terms in the search box:

Keyword Description
name: Match data asset name
***column: Match column name or nested column name
description: Match table description

upvoted 4 times

---

mdvp 2 years, 1 month ago

ITS A COLUMN WE ARE SEARCHING FOR AND I DONT SEE DATA CATALOGUE CAN SEARCH FOR COLUMN

upvoted 1 times

Load full discussion...

Start Learning for free

EXAMTOPICS

We are the biggest and most updated IT certification exam material website.

Using our own resources, we strive to strengthen the IT professionals community for free.