



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.pass4itsure.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

Pass4itSure.com

<https://www.pass4itsure.com/professional-data-engineer.html>

2024 Latest pass4itsure 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**

You want to use a BigQuery table as a data sink. In which writing mode(s) can you use BigQuery as a sink?

- A. Both batch and streaming
- B. BigQuery cannot be used as a sink
- C. Only batch
- D. Only streaming

Correct Answer: A

When you apply a BigQueryIO.Write transform in batch mode to write to a single table, Dataflow invokes a BigQuery load job. When you apply a BigQueryIO.Write transform in streaming mode or in batch mode using a function to specify the destination table, Dataflow uses BigQuery's streaming inserts Reference:

<https://cloud.google.com/dataflow/model/bigquery-io>

QUESTION 2

You are planning to use Google's Dataflow SDK to analyze customer data such as displayed below. Your project requirement is to extract only the customer name from the data source and then write to an output PCollection.

Tom,555 X street

Tim,553 Y street

Sam, 111 Z street

Which operation is best suited for the above data processing requirement?

- A. ParDo
- B. Sink API
- C. Source API
- D. Data extraction

Correct Answer: A

In Google Cloud dataflow SDK, you can use the ParDo to extract only a customer name of each element in your PCollection. Reference: <https://cloud.google.com/dataflow/model/par-do>

QUESTION 3

You are migrating a large number of files from a public HTTPS endpoint to Cloud Storage. The files are protected from unauthorized access using signed URLs. You created a TSV file that contains the list of object URLs and started a transfer job by using Storage Transfer Service. You notice that the job has run for a long time and eventually failed. Checking the logs of the transfer job reveals that the job was running fine until one point, and then it failed due to HTTP



403 errors on the remaining files You verified that there were no changes to the source system You need to fix the problem to resume the migration process. What should you do?

- A. Set up Cloud Storage FUSE, and mount the Cloud Storage bucket on a Compute Engine Instance Remove the completed files from the TSV file Use a shell script to iterate through the TSV file and download the remaining URLs to the FUSE mount point.
- B. Update the file checksums in the TSV file from using MD5 to SHA256. Remove the completed files from the TSV file and rerun the Storage Transfer Service job.
- C. Renew the TLS certificate of the HTTPS endpoint Remove the completed files from the TSV file and rerun the Storage Transfer Service job.
- D. Create a new TSV file for the remaining files by generating signed URLs with a longer validity period. Split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel.

Correct Answer: D

A signed URL is a URL that provides limited permission and time to access a resource on a web server. It is often used to grant temporary access to protected files without requiring authentication. Storage Transfer Service is a service that allows you to transfer data from external sources, such as HTTPS endpoints, to Cloud Storage buckets. You can use a TSV file to specify the list of URLs to transfer. In this scenario, the most likely cause of the HTTP 403 errors is that the signed URLs have expired before the transfer job could complete. This could happen if the signed URLs have a short validity period or the transfer job takes a long time due to the large number of files or network latency. To fix the problem, you need to create a new TSV file for the remaining files by generating new signed URLs with a longer validity period. This will ensure that the URLs do not expire before the transfer job finishes. You can use the Cloud Storage tools or your own program to generate signed URLs. Additionally, you can split the TSV file into multiple smaller files and submit them as separate Storage Transfer Service jobs in parallel. This will speed up the transfer process and reduce the risk of errors. References: Signed URLs | Cloud Storage Documentation V4 signing process with Cloud Storage tools V4 signing process with your own program Using a URL list file What Is a 403 Forbidden Error (and How Can I Fix It)?

QUESTION 4

You want to archive data in Cloud Storage. Because some data is very sensitive, you want to use the "Trust No One" (TNO) approach to encrypt your data to prevent the cloud provider staff from decrypting your data. What should you do?

- A. Use `gcloud kms keys create` to create a symmetric key. Then use `gcloud kms encrypt` to encrypt each archival file with the key and unique additional authenticated data (AAD). Use `gsutil cp` to upload each encrypted file to the Cloud Storage bucket, and keep the AAD outside of Google Cloud.
- B. Use `gcloud kms keys create` to create a symmetric key. Then use `gcloud kms encrypt` to encrypt each archival file with the key. Use `gsutil cp` to upload each encrypted file to the Cloud Storage bucket. Manually destroy the key previously used for encryption, and rotate the key once and rotate the key once.
- C. Specify customer-supplied encryption key (CSEK) in the `.boto` configuration file. Use `gsutil cp` to upload each archival file to the Cloud Storage bucket. Save the CSEK in Cloud Memorystore as permanent storage of the secret.
- D. Specify customer-supplied encryption key (CSEK) in the `.boto` configuration file. Use `gsutil cp` to upload each archival file to the Cloud Storage bucket. Save the CSEK in a different project that only the security team can access.

Correct Answer: B



QUESTION 5

You are building an application to share financial market data with consumers, who will receive data feeds. Data is collected from the markets in real time. Consumers will receive the data in the following ways:

1.

Real-time event stream

2.

ANSI SQL access to real-time stream and historical data Batch historical exports

Which solution should you use?

A. Cloud Dataflow, Cloud SQL, Cloud Spanner

B. Cloud Pub/Sub, Cloud Storage, BigQuery

C. Cloud Dataproc, Cloud Dataflow, BigQuery

D. Cloud Pub/Sub, Cloud Dataproc, Cloud SQL

Correct Answer: A

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

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

[PROFESSIONAL-DATA-ENGINEER Study Guide](#)