



# PR OFESSIONAL-DATA-ENGINEER<sup>Q&As</sup>

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 are designing an Apache Beam pipeline to enrich data from Cloud Pub/Sub with static reference data from BigQuery. The reference data is small enough to fit in memory on a single worker. The pipeline should write enriched results to BigQuery for analysis. Which job type and transforms should this pipeline use?

- A. Batch job, PubSubIO, side-inputs
- B. Streaming job, PubSubIO, JdbcIO, side-outputs
- C. Streaming job, PubSubIO, BigQueryIO, side-inputs
- D. Streaming job, PubSubIO, BigQueryIO, side-outputs

Correct Answer: C

---

**QUESTION 2**

You are using BigQuery and Data Studio to design a customer-facing dashboard that displays large quantities of aggregated data. You expect a high volume of concurrent users. You need to optimize the dashboard to provide quick visualizations with minimal latency. What should you do?

- A. Use BigQuery BI Engine with materialized views
- B. Use BigQuery BI Engine with streaming data.
- C. Use BigQuery BI Engine with authorized views
- D. Use BigQuery BI Engine with logical reviews

Correct Answer: A

<https://cloud.google.com/bigquery/docs/materialized-views-intro> In BigQuery, materialized views are precomputed views that periodically cache the results of a query for increased performance and efficiency. BigQuery leverages precomputed results from materialized views and whenever possible reads only delta changes from the base tables to compute up-to-date results. Materialized views can be queried directly or can be used by the BigQuery optimizer to process queries to the base tables.

Queries that use materialized views are generally faster and consume fewer resources than queries that retrieve the same data only from the base tables. Materialized views can significantly improve the performance of workloads that have the characteristic of common and repeated queries.

---

**QUESTION 3**

You create an important report for your large team in Google Data Studio 360. The report uses Google BigQuery as its data source. You notice that visualizations are not showing data that is less than 1 hour old. What should you do?

- A. Disable caching by editing the report settings.
- B. Disable caching in BigQuery by editing table details.



- C. Refresh your browser tab showing the visualizations.
- D. Clear your browser history for the past hour then reload the tab showing the virtualizations.

Correct Answer: A

Reference <https://support.google.com/datastudio/answer/7020039?hl=en>

---

#### QUESTION 4

You have a data pipeline with a Cloud Dataflow job that aggregates and writes time series metrics to Cloud Bigtable. This data feeds a dashboard used by thousands of users across the organization. You need to support additional concurrent users and reduce the amount of time required to write the data. Which two actions should you take? (Choose two.)

- A. Configure your Cloud Dataflow pipeline to use local execution
- B. Increase the maximum number of Cloud Dataflow workers by setting `maxNumWorkers` in `PipelineOptions`
- C. Increase the number of nodes in the Cloud Bigtable cluster
- D. Modify your Cloud Dataflow pipeline to use the Flatten transform before writing to Cloud Bigtable
- E. Modify your Cloud Dataflow pipeline to use the `CoGroupByKey` transform before writing to Cloud Bigtable

Correct Answer: BC

---

#### QUESTION 5

You are planning to migrate your current on-premises Apache Hadoop deployment to the cloud. You need to ensure that the deployment is as fault-tolerant and cost-effective as possible for long-running batch jobs. You want to use a managed service. What should you do?

- A. Deploy a Cloud Dataproc cluster. Use a standard persistent disk and 50% preemptible workers. Store data in Cloud Storage, and change references in scripts from `hdfs://` to `gs://`
- B. Deploy a Cloud Dataproc cluster. Use an SSD persistent disk and 50% preemptible workers. Store data in Cloud Storage, and change references in scripts from `hdfs://` to `gs://`
- C. Install Hadoop and Spark on a 10-node Compute Engine instance group with standard instances. Install the Cloud Storage connector, and store the data in Cloud Storage. Change references in scripts from `hdfs://` to `gs://`
- D. Install Hadoop and Spark on a 10-node Compute Engine instance group with preemptible instances. Store data in HDFS. Change references in scripts from `hdfs://` to `gs://`

Correct Answer: A

---



VCE & PDF

Pass4itSure.com

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

2024 Latest pass4itsure PROFESSIONAL-DATA-ENGINEER PDF and VCE  
dumps Download

---

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

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

[PROFESSIONAL-DATA-  
ENGINEER Exam  
Questions](#)