



GOOGLE

Cloud Certified Professional Data Engineer

Exam Questions & Answers

Total Questions: **05**

[Click Here to Buy All Questions →](#)

<https://itexamsquiz.com/product/professional-data-engineer-practice-tests/>



About ITExamsQuiz

At ITExamsQuiz, we believe every IT professional deserves access to high-quality exam preparation without breaking the bank. Our expert team curates accurate, up-to-date study material for a wide range of certification exams.

With ITExamsQuiz you get expert-vetted study material, real exam insights, and a fair shot at success.

- ✓ **Lifetime Free Updates**
- ✓ **Money-Back Guarantee**
- ✓ **100% Verified Questions**
- ✓ **24/7 Support via Email**

Doubt Support

We resolve 400+ doubts every day with an average rating of 4.8/5.

 <https://itexamsquiz.com>
 support@itexamsquiz.com

★ ★ ★ ★ ★ **Logan**

I passed with 920/1000! More than 45 questions were from your PDFs. Highly recommended!

★ ★ ★ ★ ★ **Hailey**

Updated material and correct answers. Passed AZ-900 on first attempt. Thank you ITExamsQuiz!

★ ★ ★ ★ ★ **Brandon**

Amazing accuracy. Out of 52 questions, 50 were from your dumps. Passed with 870 marks!

★ ★ ★ ★ ★ **Madison**

Customer support is very responsive. The PDFs are super accurate. Will recommend to friends!

Question: 1

ITExamsQuiz

Your company built a TensorFlow neural-network model with a large number of neurons and layers. The model fits well for the training data. However, when tested against new data, it performs poorly. What method can you employ to address this?

- A. Threading
- B. Serialization
- C. Dropout Methods
- D. Dimensionality Reduction

✓ **Answer: C**

Explanation:

<https://itexamsquiz.com> | ACE YOUR EXAM

Question: 2

ITExamsQuiz

You are building a model to make clothing recommendations. You know a user's fashion preference is likely to change over time, so you build a data pipeline to stream new data back to the model as it becomes available. How should you use this data to train the model?

- A. Continuously retrain the model on just the new data.
- B. Continuously retrain the model on a combination of existing data and the new data.
- C. Train on the existing data while using the new data as your test set.
- D. Train on the new data while using the existing data as your test set.

✓ **Answer: D**

Explanation:

<https://itexamsquiz.com> | ACE YOUR EXAM

Question: 3

ITExamsQuiz

You designed a database for patient records as a pilot project to cover a few hundred patients in three clinics. Your design used a single database table to represent all patients and their visits, and you used self-joins to generate reports. The server resource utilization was at 50%. Since then, the scope of the project has expanded. The database must now store 100 times more patient records. You can no longer run the reports, because they either take too long or they encounter errors with insufficient compute resources. How should you adjust the database design?

- A. Add capacity (memory and disk space) to the database server by the order of 200.
- B. Shard the tables into smaller ones based on date ranges, and only generate reports with prespecified date ranges.
- C. Normalize the master patient-record table into the patient table and the visits table, and create other necessary tables to avoid self-join.
- D. Partition the table into smaller tables, with one for each clinic. Run queries against the smaller table pairs, and use unions for consolidated reports.

✓ **Answer: B**

Explanation:

<https://itexamsquiz.com> | ACE YOUR EXAM

Question: 4

ITExamsQuiz

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.

✓ **Answer: A**

Explanation:

<https://itexamsquiz.com> | ACE YOUR EXAM

Question: 5

ITExamsQuiz

An external customer provides you with a daily dump of data from their database. The data flows into Google Cloud Storage GCS as comma-separated values (CSV) files. You want to analyze this data in Google BigQuery, but the data could have rows that are formatted incorrectly or corrupted. How should you build this pipeline?

- A. Use federated data sources, and check data in the SQL query.
- B. Enable BigQuery monitoring in Google Stackdriver and create an alert.
- C. Import the data into BigQuery using the gcloud CLI and set max_bad_records to 0.
- D. Run a Google Cloud Dataflow batch pipeline to import the data into BigQuery, and push errors to another dead-letter table for analysis.

✓ **Answer: D**

Explanation:

Thank you for trying the Cloud Certified Professional Data Engineer Exam Dumps PDF Demo

Practice smarter with our updated Cloud Certified Professional Data Engineer PDF Questions

If you want to buy our premium Practice Questions, please follow the link below and get started now!

<https://itexamsquiz.com/product/professional-data-engineer-practice-tests/>