

Valid Google Professional-Data-Engineer Test Questions & Professional-Data-Engineer Latest Exam Duration



P.S. Free & New Professional-Data-Engineer dumps are available on Google Drive shared by Dumpexams:
https://drive.google.com/open?id=1PR1cYB79vpni4ZWd4JCv866_HJFUh5ed

I know that the purpose of your test is definitely passing the Professional-Data-Engineer exam. So, buying our Professional-Data-Engineer guide quiz is definitely your best choice. Users who used Professional-Data-Engineer exam questions basically passed the exam. I believe that after you use our Professional-Data-Engineer Study Materials for a while, we will understand why we have a 99% pass rate. With the best quality and the latest version which we are always trying our best to develop, our Professional-Data-Engineer practice engine can help you pass the exam for sure.

Google Professional-Data-Engineer exam is intended for professionals who work with data engineering, data integration, or data analysis. Professional-Data-Engineer exam tests the candidate's knowledge and understanding of Google Cloud Platform tools and services, including BigQuery, Cloud Dataflow, Cloud Pub/Sub, Cloud Storage, and more. Professional-Data-Engineer exam consists of multiple-choice questions and practical scenarios that test the candidate's ability to apply their knowledge and skills to real-world problems. Passing the exam and obtaining the certification demonstrates the individual's proficiency in designing and implementing scalable and reliable data processing systems using Google Cloud Platform technologies.

To pass the Google Professional-Data-Engineer Exam, candidates must have a solid understanding of data engineering concepts and techniques, as well as practical experience working with the Google Cloud Platform. They must be able to design and implement data processing systems that are secure, scalable, and efficient, and have the ability to troubleshoot and optimize these systems as needed. Professional-Data-Engineer exam is challenging and comprehensive, but passing it can open up many career opportunities in data engineering, especially for those interested in working with Google Cloud Platform.

>> Valid Google Professional-Data-Engineer Test Questions <<

Professional-Data-Engineer Latest Exam Duration | Professional-Data-Engineer Dumps Free Download

Our company has dedicated ourselves to develop the Professional-Data-Engineer study materials for all candidates to pass the exam easier, also has made great achievement after more than ten years' development. As the certification has been of great value, a right Professional-Data-Engineer study material can be your strong forward momentum to help you pass the exam like a hot knife through butter. On the contrary, it might be time-consuming and tired to prepare for the Professional-Data-Engineer Exam without a specialist study material. So it's would be the best decision to choose our Professional-Data-Engineer study materials as your learning partner.

Google Certified Professional Data Engineer Exam Sample Questions (Q36-Q41):

NEW QUESTION # 36

You are designing BigQuery tables for large volumes of clickstream event data. Your data analyst team will most frequently query by specific event date ranges and filter by the user ID UUID. You want to optimize table structure for query cost and performance. What should you do?

- A. Cluster the table by both the event date and the user ID columns.
- **B. Partition the table by the event date column and cluster the table by the user ID column.**
- C. Partition the table by the user ID column and cluster the table by the event date column.
- D. Create an ingestion-time partitioned table and cluster it by the user ID column.

Answer: B

Explanation:

Comprehensive and Detailed Explanation:

This question is about applying the two primary optimization techniques in BigQuery: partitioning and clustering.

Partitioning divides a table into smaller segments based on a date, timestamp, or integer column. When you filter a query on the partition column, BigQuery performs "partition pruning," meaning it only scans the data in the relevant partitions. Since queries frequently filter by "specific event date ranges," partitioning by the event date column is the ideal strategy to reduce the amount of data scanned, which lowers cost and improves performance.

Clustering sorts the data within each partition based on the values in one or more columns. When you filter on a clustered column, BigQuery can use the sorted order to avoid scanning all the data within the relevant partitions. Since queries also filter by user ID, clustering by the user ID column will further improve performance and can reduce costs for those queries.

Conclusion: Option A is the correct answer because it correctly applies both techniques according to best practices: partition on the range-filtered column (event date) and cluster on the point-lookup/high-cardinality filtered column (user ID).

Option B is incorrect because you cannot partition a BigQuery table by a string column like a UUID.

Option C is incorrect because ingestion-time partitioning is less precise. Partitioning directly on the event date column is more effective for queries that filter on the event date itself.

Option D is incorrect because while clustering is helpful, partitioning is the more critical optimization for date range queries and provides the biggest cost savings. The combination of both is optimal.

Reference (Google Cloud Documentation Concepts):

The Google Cloud documentation for "Introduction to partitioned tables" and "Introduction to clustered tables" provides clear guidance. The best practice is to partition by a date or timestamp column that is commonly used as a filter to prune data. It then recommends to cluster by columns that are frequently used in WHERE clauses for filtering or in JOIN clauses. The combination of partitioning and clustering is a powerful optimization strategy.

NEW QUESTION # 37

You are building a model to make clothing recommendations. You know a user's fashion preferences are 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 a combination of existing data and the new data.**
- B. Train on the new data while using the existing data as your test set.
- C. Continuously retrain the model on just the new data.
- D. Train on the existing data while using the new data as your test set.

Answer: A

Explanation:

We have to use a combination of old and new test data as well as training data.

NEW QUESTION # 38

If a dataset contains rows with individual people and columns for year of birth, country, and income, how many of the columns are continuous and how many are categorical?

- A. 3 categorical
- **B. 2 continuous and 1 categorical**
- C. 1 continuous and 2 categorical
- D. 3 continuous

Answer: B

Explanation:

The columns can be grouped into two types-categorical and continuous columns:

A column is called categorical if its value can only be one of the categories in a finite set.

For example, the native country of a person (U.S., India, Japan, etc.) or the education level (high school, college, etc.) are categorical columns.

A column is called continuous if its value can be any numerical value in a continuous range. For example, the capital gain of a person (e.g. \$14,084) is a continuous column.

Year of birth and income are continuous columns. Country is a categorical column.

You could use bucketization to turn year of birth and/or income into categorical features, but the raw columns are continuous.

Reference: https://www.tensorflow.org/tutorials/wide#reading_the_census_data

NEW QUESTION # 39

Which of the following are feature engineering techniques? (Select 2 answers)

- A. Hidden feature layers
- B. Crossed feature columns
- C. Bucketization of a continuous feature
- D. Feature prioritization

Answer: B,C

Explanation:

Selecting and crafting the right set of feature columns is key to learning an effective model.

Bucketization is a process of dividing the entire range of a continuous feature into a set of consecutive bins/buckets, and then converting the original numerical feature into a bucket ID (as a categorical feature) depending on which bucket that value falls into.

Using each base feature column separately may not be enough to explain the data. To learn the differences between different feature combinations, we can add crossed feature columns to the model.

Reference:

https://www.tensorflow.org/tutorials/wide#selecting_and_engineering_features_for_the_model

NEW QUESTION # 40

You have enabled the free integration between Firebase Analytics and Google BigQuery. Firebase now automatically creates a new table daily in BigQuery in the format `app_events_YYYYMMDD`. You want to query all of the tables for the past 30 days in legacy SQL. What should you do?

- A. Use the `TABLE_DATE_RANGE` function
- B. Use `WHERE date BETWEEN YYYY-MM-DD AND YYYY-MM-DD`
- C. Use the `WHERE PARTITIONTIME` pseudo column
- D. Use `SELECT IF.(date >= YYYY-MM-DD AND date <= YYYY-MM-DD`

Answer: A

Explanation:

Explanation/Reference: <https://cloud.google.com/blog/products/gcp/using-bigquery-and-firebase-analytics-to-understand-your-mobile-app?hl=am>

NEW QUESTION # 41

.....

As long as you bought our Professional-Data-Engineer practice guide, then you will find that it cost little time and efforts to learn.

You can have a quick revision of the Professional-Data-Engineer learning quiz in your spare time. Also, you can memorize the

knowledge quickly. There almost have no troubles to your normal life. You can make use of your spare moment to study our

Professional-Data-Engineer Preparation questions. The results will become better with your constant exercises. Please have a brave attempt.

Professional-Data-Engineer Latest Exam Duration: <https://www.dumpexams.com/Professional-Data-Engineer-real-answers.html>

- What's more, part of that Dumpexams Professional-Data-Engineer dumps now are free: https://drive.google.com/open?id=1PR1cYB79vpni4ZWd4JCv866_HJFUh5ed

What's more, part of that Dumpexams Professional-Data-Engineer dumps now are free: https://drive.google.com/open?id=1PR1cYB79vpni4ZWd4JCv866_HJFUh5ed