

100% Pass Google - Useful Professional-Data-Engineer Valid Test Syllabus



What's more, part of that ITexamReview Professional-Data-Engineer dumps now are free: <https://drive.google.com/open?id=1Kb6oX6E8-rtHQJxXmC8MAR3tT8b6nB7K>

Professional-Data-Engineer materials trends are not always easy to forecast, but they have predictable pattern for them by ten-year experience who often accurately predict points of knowledge occurring in next Professional-Data-Engineer preparation materials. Our professional experts can give you the latest and the most accurate Professional-Data-Engineer Training Material for that they have been in this field for so many years and know every aspect of the change of Professional-Data-Engineer practice questions. You can trust in our Professional-Data-Engineer learning braindump for sure.

Google Professional-Data-Engineer Certification is a highly respected credential that validates the knowledge and skills of professionals working in the field of data engineering. Google Certified Professional Data Engineer Exam certification is designed to test the ability of candidates to design, build, operate, and manage data processing systems that are scalable, secure, and reliable. Google Certified Professional Data Engineer Exam certification exam is conducted by Google and is intended for individuals who have a good understanding of the Google Cloud Platform and data engineering best practices.

Professional Data Engineer Exam Details

Like other Google exams, this exam also consists of multiple choice and multiple select questions. Consider the fact that you need to pay \$200 for the registration. After that, you will access the test for 2 hours which is presented either in English or Japanese. Moreover, you can either take the exam online or have to find a test center near your place to take this test.

There is no formal prerequisite for the exam but it is recommended to have 3-4 years of experience within the data engineering field and to be responsible for the tasks related to data engineering and machine learning. So, on the final test day, you need to have exhaustive knowledge about these domains to perform your best.

- Operationalizing machine learning models
- Providing solution quality
- Designing data processing systems
- Building data processing systems

Reliable Google Professional-Data-Engineer Practice Materials | Latest Professional-Data-Engineer Study Materials

The Professional-Data-Engineer pdf format of the ITexamReview product is easy-to-use. It contains actual Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) exam questions. You can easily download and use Google Professional-Data-Engineer pdf on laptops, tablets, and smartphones. ITexamReview regularly updates Professional-Data-Engineer Exam Questions' pdf version so that you always have the latest material. Furthermore, the Google Professional-Data-Engineer pdf can be printed enabling paper study.

Google Certified Professional Data Engineer Exam Sample Questions (Q329-Q334):

NEW QUESTION # 329

Case Study: 2 - MJTelco

Company Overview

MJTelco is a startup that plans to build networks in rapidly growing, underserved markets around the world. The company has patents for innovative optical communications hardware. Based on these patents, they can create many reliable, high-speed backbone links with inexpensive hardware.

Company Background

Founded by experienced telecom executives, MJTelco uses technologies originally developed to overcome communications challenges in space. Fundamental to their operation, they need to create a distributed data infrastructure that drives real-time analysis and incorporates machine learning to continuously optimize their topologies. Because their hardware is inexpensive, they plan to overdeploy the network allowing them to account for the impact of dynamic regional politics on location availability and cost. Their management and operations teams are situated all around the globe creating many-to-many relationship between data consumers and provides in their system. After careful consideration, they decided public cloud is the perfect environment to support their needs.

Solution Concept

MJTelco is running a successful proof-of-concept (PoC) project in its labs. They have two primary needs:

Scale and harden their PoC to support significantly more data flows generated when they ramp to more than 50,000 installations.

Refine their machine-learning cycles to verify and improve the dynamic models they use to control topology definition.

MJTelco will also use three separate operating environments ?development/test, staging, and production ?

to meet the needs of running experiments, deploying new features, and serving production customers.

Business Requirements

Scale up their production environment with minimal cost, instantiating resources when and where needed in an unpredictable, distributed telecom user community. Ensure security of their proprietary data to protect their leading-edge machine learning and analysis.

Provide reliable and timely access to data for analysis from distributed research workers Maintain isolated environments that support rapid iteration of their machine-learning models without affecting their customers.

Technical Requirements

Ensure secure and efficient transport and storage of telemetry data Rapidly scale instances to support between 10,000 and 100,000 data providers with multiple flows each.

Allow analysis and presentation against data tables tracking up to 2 years of data storing approximately 100m records/day

Support rapid iteration of monitoring infrastructure focused on awareness of data pipeline problems both in telemetry flows and in production learning cycles.

CEO Statement

Our business model relies on our patents, analytics and dynamic machine learning. Our inexpensive hardware is organized to be highly reliable, which gives us cost advantages. We need to quickly stabilize our large distributed data pipelines to meet our reliability and capacity commitments.

CTO Statement

Our public cloud services must operate as advertised. We need resources that scale and keep our data secure. We also need environments in which our data scientists can carefully study and quickly adapt our models. Because we rely on automation to process our data, we also need our development and test environments to work as we iterate.

CFO Statement

The project is too large for us to maintain the hardware and software required for the data and analysis.

Also, we cannot afford to staff an operations team to monitor so many data feeds, so we will rely on automation and infrastructure.

Google Cloud's machine learning will allow our quantitative researchers to work on our high-value problems instead of problems

with our data pipelines.

You need to compose visualization for operations teams with the following requirements:

Telemetry must include data from all 50,000 installations for the most recent 6 weeks (sampling once every minute) The report must not be more than 3 hours delayed from live data. The actionable report should only show suboptimal links.

Most suboptimal links should be sorted to the top.

Suboptimal links can be grouped and filtered by regional geography. User response time to load the report must be <5 seconds.

You create a data source to store the last 6 weeks of data, and create visualizations that allow viewers to see multiple date ranges, distinct geographic regions, and unique installation types.

You always show the latest data without any changes to your visualizations. You want to avoid creating and updating new visualizations each month. What should you do?

- A. Export the data to a spreadsheet, compose a series of charts and tables, one for each possible combination of criteria, and spread them across multiple tabs.
- B. Load the data into relational database tables, write a Google App Engine application that queries all rows, summarizes the data across each criteria, and then renders results using the Google Charts and visualization API.
- **C. Look through the current data and compose a small set of generalized charts and tables bound to criteria filters that allow value selection.**
- D. Look through the current data and compose a series of charts and tables, one for each possible combination of criteria.

Answer: C

NEW QUESTION # 330

You migrated your on-premises Apache Hadoop Distributed File System (HDFS) data lake to Cloud Storage. The data scientist team needs to process the data by using Apache Spark and SQL. Security policies need to be enforced at the column level. You need a cost-effective solution that can scale into a data mesh. What should you do?

- A. 1. Load the data to BigQuery tables.
2. Create a taxonomy of policy tags in Data Catalog.
3. Add policy tags to columns.
4. Process with the Spark-BigQuery connector or BigQuery SQL.
- **B. 1. Apply an Identity and Access Management (IAM) policy at the file level in Cloud Storage
2. Define a BigQuery external table for SQL processing
3. Use Dataproc Spark to process the Cloud Storage files.**
- C. 1. Define a BigLake table.
2. Create a taxonomy of policy tags in Data Catalog.
3. Add policy tags to columns.
4. Process with the Spark-BigQuery connector or BigQuery SQL.
- D. 1. Deploy a long-living Dataproc cluster with Apache Hive and Ranger enabled.
2. Configure Ranger for column level security.
3. Process with Dataproc Spark or Hive SQL.

Answer: B

Explanation:

For automating the CI/CD pipeline of DAGs running in Cloud Composer, the following approach ensures that DAGs are tested and deployed in a streamlined and efficient manner.

Use Cloud Build for Development Instance Testing:

Use Cloud Build to automate the process of copying the DAG code to the Cloud Storage bucket of the development instance.

This triggers Cloud Composer to automatically pick up and test the new DAGs in the development environment.

Testing and Validation:

Ensure that the DAGs run successfully in the development environment.

Validate the functionality and correctness of the DAGs before promoting them to production.

Deploy to Production:

If the DAGs pass all tests in the development environment, use Cloud Build to copy the tested DAG code to the Cloud Storage bucket of the production instance.

This ensures that only validated and tested DAGs are deployed to production, maintaining the stability and reliability of the production environment.

Simplicity and Reliability:

This approach leverages Cloud Build's capabilities for automation and integrates seamlessly with Cloud Composer's reliance on Cloud Storage for DAG storage.

By using Cloud Storage for both development and production deployments, the process remains simple and robust.

Google Data Engineer Reference:

Cloud Composer Documentation

Using Cloud Build

Deploying DAGs to Cloud Composer

Automating DAG Deployment with Cloud Build

By implementing this CI/CD pipeline, you ensure that DAGs are thoroughly tested in the development environment before being automatically deployed to the production environment, maintaining high quality and reliability.

NEW QUESTION # 331

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

Explanation:

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

NEW QUESTION # 332

When a Cloud Bigtable node fails, _____ is lost.

- A. the last transaction
- **B. no data**
- C. all data
- D. the time dimension

Answer: B

Explanation:

A Cloud Bigtable table is sharded into blocks of contiguous rows, called tablets, to help balance the workload of queries. Tablets are stored on Colossus, Google's file system, in SSTable format. Each tablet is associated with a specific Cloud Bigtable node. Data is never stored in Cloud Bigtable nodes themselves; each node has pointers to a set of tablets that are stored on Colossus. As a result:

Rebalancing tablets from one node to another is very fast, because the actual data is not copied. Cloud Bigtable simply updates the pointers for each node. Recovery from the failure of a Cloud Bigtable node is very fast, because only metadata needs to be migrated to the replacement node.

When a Cloud Bigtable node fails, no data is lost

Reference: <https://cloud.google.com/bigtable/docs/overview>

NEW QUESTION # 333

You are a head of BI at a large enterprise company with multiple business units that each have different priorities and budgets. You use on-demand pricing for BigQuery with a quota of 2K concurrent on-demand slots per project. Users at your organization sometimes don't get slots to execute their query and you need to correct this. You'd like to avoid introducing new projects to your account.

What should you do?

- **A. Switch to flat-rate pricing and establish a hierarchical priority model for your projects.**
- B. Convert your batch BQ queries into interactive BQ queries.
- C. Increase the amount of concurrent slots per project at the Quotas page at the Cloud Console.
- D. Create an additional project to overcome the 2K on-demand per-project quota.

Answer: A

Explanation

Reference <https://cloud.google.com/blog/products/gcp/busting-12-myths-about-bigquery>

NEW QUESTION # 334

• • • • •

The web-based Google Professional-Data-Engineer practice exam is compatible with all browsers like Chrome, Mozilla Firefox, MS Edge, Internet Explorer, Safari, Opera, and more. Unlike the desktop version, it requires an internet connection. The Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) practice exam will ask real Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) exam questions. Consistent practice with it relieves exam stress and boosts self-confidence. The web-based Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) practice exam does not require additional software installation. All operating systems also support this Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) practice test.

Reliable Professional-Data-Engineer Practice Materials: <https://www.itexamreview.com/Professional-Data-Engineer-exam-dumps.html>

- [illegible]

DOWNLOAD the newest ITexamReview Professional-Data-Engineer PDF dumps from Cloud Storage for free:
<https://drive.google.com/open?id=1Kb6oX6E8-rtHOJxXmC8MAR3t8B6nB7K>

