

Professional-Data-Engineer Reliable Exam Tutorial, Reliable Professional-Data-Engineer Test Prep



What's more, part of that Pass4SureQuiz Professional-Data-Engineer dumps now are free: <https://drive.google.com/open?id=1TvBK0RGt3KoubWfen3pQX8j-pIZKIVJZ>

Pass4SureQuiz Professional-Data-Engineer practice test simulates the real Google Professional-Data-Engineer exam environment. This situation boosts the candidate's performance and enhances their confidence. After attempting the Professional-Data-Engineer practice exams, candidates become more familiar with a real Google Certified Professional Data Engineer Exam Professional-Data-Engineer Exam environment and develop the stamina to sit for several hours consecutively to complete the Professional-Data-Engineer exam. This way, the actual Google Certified Professional Data Engineer Exam Professional-Data-Engineer exam becomes much easier for them to handle.

Google Professional-Data-Engineer exam is a certification offered by Google that validates the skills and knowledge required for designing, building, and managing data processing systems on the Google Cloud Platform. Google Certified Professional Data Engineer Exam certification is intended for data professionals who want to demonstrate their expertise in designing and managing data solutions on Google Cloud. Professional-Data-Engineer Exam covers various topics such as designing data processing systems, implementing data storage solutions, managing data processing infrastructure, and ensuring data security and compliance.

>> Professional-Data-Engineer Reliable Exam Tutorial <<

Reliable Professional-Data-Engineer Test Prep, Test Professional-Data-Engineer Objectives Pdf

If you do not know how to pass the exam more effectively, I'll give you a suggestion is to choose a good training site. This can play a multiplier effect. Pass4SureQuiz site has always been committed to provide candidates with a real Google Professional-Data-Engineer Certification Exam training materials. The Pass4SureQuiz Google Professional-Data-Engineer Certification Exam software are authorized products by vendors, it is wide coverage, and can save you a lot of time and effort.

Google Certified Professional Data Engineer Exam Sample Questions (Q233-Q238):

NEW QUESTION # 233

Which of the following job types are supported by Cloud Dataproc (select 3 answers)?

- A. Hive
- B. Spark
- C. Pig

- D. YARN

Answer: A,B,C

Explanation:

Cloud Dataproc provides out-of-the box and end-to-end support for many of the most popular job types, including Spark, Spark SQL, PySpark, MapReduce, Hive, and Pig jobs.

NEW QUESTION # 234

An online retailer has built their current application on Google App Engine. A new initiative at the company mandates that they extend their application to allow their customers to transact directly via the application.

They need to manage their shopping transactions and analyze combined data from multiple datasets using a business intelligence (BI) tool. They want to use only a single database for this purpose. Which Google Cloud database should they choose?

- A. Cloud Datastore
- **B. Cloud BigTable**
- C. BigQuery
- D. Cloud SQL

Answer: B

Explanation:

reference: <https://cloud.google.com/solutions/business-intelligence/>

NEW QUESTION # 235

MJTelco Case Study

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.

MJTelco needs you to create a schema in Google Bigtable that will allow for the historical analysis of the last 2 years of records.

Each record that comes in is sent every 15 minutes, and contains a unique identifier of the device and a data record. The most common query is for all the data for a given device for a given day. Which schema should you use?

- A. Rowkey: device_id
Column data: date, data_point
- B. Rowkey: date#device_id
Column data: data_point
- C. Rowkey: date
Column data: device_id, data_point
- D. Rowkey: date#data_point
Column data: device_id
- E. Rowkey: data_point
Column data: device_id, date

Answer: E

NEW QUESTION # 236

A shipping company has live package-tracking data that is sent to an Apache Kafka stream in real time. This is then loaded into BigQuery. Analysts in your company want to query the tracking data in BigQuery to analyze geospatial trends in the lifecycle of a package. The table was originally created with ingest-date partitioning.

Over time, the query processing time has increased. You need to implement a change that would improve query performance in BigQuery. What should you do?

- A. Re-create the table using data partitioning on the package delivery date.
- B. Tier older data onto Google Cloud Storage files and create a BigQuery table using GCS as an external data source.
- C. Implement clustering in BigQuery on the package-tracking ID column.
- D. Implement clustering in BigQuery on the ingest date column.

Answer: C

NEW QUESTION # 237

A shipping company has live package-tracking data that is sent to an Apache Kafka stream in real time. This is then loaded into BigQuery. Analysts in your company want to query the tracking data in BigQuery to analyze geospatial trends in the lifecycle of a package. The table was originally created with ingest-date partitioning.

Over time, the query processing time has increased. You need to implement a change that would improve query performance in BigQuery. What should you do?

- A. Tier older data onto Cloud Storage files, and leverage extended tables.
- B. Re-create the table using data partitioning on the package delivery date.
- C. Implement clustering in BigQuery on the ingest date column.

- Answer: C**

• • • • •

Reliable Professional-Data-Engineer Test Prep: <https://www.pass4surequiz.com/Professional-Data-Engineer-exam-quiz.html>

- What's more, part of that Pass4SureQuiz Professional-Data-Engineer dumps now are free: <https://drive.google.com/open?id=1TvBK0RGt3KoubWfen3pOX8j-pIZKIVJZ>