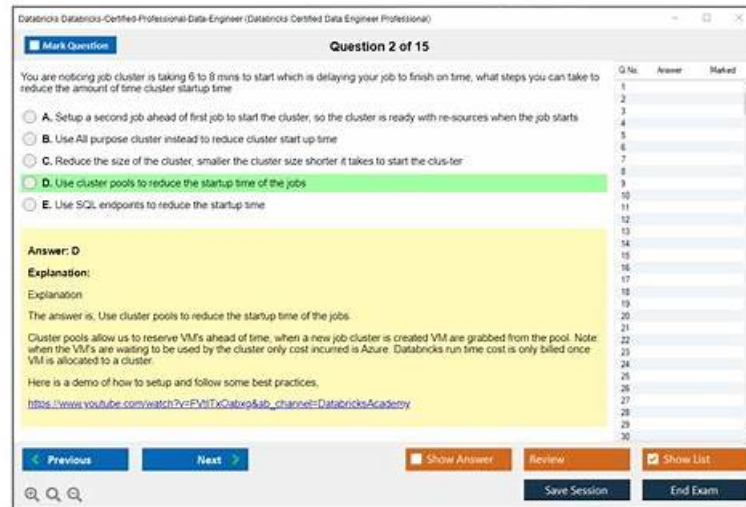


Accurate Professional-Data-Engineer Study Material, Latest Professional-Data-Engineer Exam Forum



P.S. Free 2025 Google Professional-Data-Engineer dumps are available on Google Drive shared by Actual4dump: <https://drive.google.com/open?id=1bwifQTHT4THmHFqGd76l1TWA8WmvBBCj>

The emerging Google field creates a space for Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) certification exam holders to accelerate their careers. Many unfortunate candidates don't get the Google Professional-Data-Engineer certification because they prepare for its Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) exam questions from a Google Professional-Data-Engineer exam that dumps outdated material. It results in a waste of time and money. You can develop your skills and join the list of experts by earning this Google Certified Professional Data Engineer Exam (Professional-Data-Engineer) certification exam.

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.

>> Accurate Professional-Data-Engineer Study Material <<

Google Certified Professional Data Engineer Exam Sure Exam Vce & Professional-Data-Engineer Training Torrent & Google Certified Professional Data Engineer Exam Latest Pdf

May be you will meet some difficult or problems when you prepare for your Professional-Data-Engineer exam, you even want to give it up. That is why I suggest that you must try our study materials. Because Professional-Data-Engineer guide torrent can help you to solve all the problems encountered in the learning process, Professional-Data-Engineer Study Tool will provide you with very flexible learning time so that you can easily pass the exam. I believe that after you try our products, you will love it soon.

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

NEW QUESTION # 233

Business owners at your company have given you a database of bank transactions. Each row contains the user ID, transaction type, transaction location, and transaction amount. They ask you to investigate what type of machine learning can be applied to the data. Which three machine learning applications can you use? (Choose three.)

- A. Reinforcement learning to predict the location of a transaction.
- B. Clustering to divide the transactions into N categories based on feature similarity.
- C. Supervised learning to determine which transactions are most likely to be fraudulent.
- D. Unsupervised learning to determine which transactions are most likely to be fraudulent.
- E. Unsupervised learning to predict the location of a transaction.
- F. Supervised learning to predict the location of a transaction.

Answer: A,B,D

Explanation:

Explanation/Reference:

NEW QUESTION # 234

Your company is performing data preprocessing for a learning algorithm in Google Cloud Dataflow. Numerous data logs are being generated during this step, and the team wants to analyze them. Due to the dynamic nature of the campaign, the data is growing exponentially every hour.

The data scientists have written the following code to read the data for a new key features in the logs.

```
BigQueryIO.Read
.named("ReadLogData")
.from("clouddataflow-readonly:samples.log_data")
```

You want to improve the performance of this data read. What should you do?

- A. Call a transform that returns TableRow objects, where each element in the PCollection represents a single row in the table.
- B. Use of both the Google BigQuery TableSchema and TableFieldSchema classes.
- C. Use .fromQuery operation to read specific fields from the table.
- D. Specify the TableReference object in the code.

Answer: A

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 is building a custom interface to share data. They have these requirements:

1. They need to do aggregations over their petabyte-scale datasets.
2. They need to scan specific time range rows with a very fast response time (milliseconds).

Which combination of Google Cloud Platform products should you recommend?

- A. Cloud Datastore and Cloud Bigtable
- B. BigQuery and Cloud Storage
- C. Cloud Bigtable and Cloud SQL
- D. BigQuery and Cloud Bigtable

Answer: D

NEW QUESTION # 236

You are deploying a new storage system for your mobile application, which is a media streaming service. You decide the best fit is Google Cloud Datastore. You have entities with multiple properties, some of which can take on multiple values. For example, in the entity 'Movie' the property 'actors' and the property 'tags' have multiple values but the property 'date released' does not. A typical query would ask for all movies with actor=<actormame> ordered by date_released or all movies with tag=Comedy ordered by date_released. How should you avoid a combinatorial explosion in the number of indexes?

A. Manually configure the index in your index config as follows:

Indexes:

-kind: Movie

Properties:

-name: actors

name: date_released

-kind: Movie

Properties:

-name: tags

name: date_released

B. Manually configure the index in your index config as follows:

Indexes:

-kind: Movie

Properties:

-name: actors

-name: tags

-name: date_published

C. Set the following in your entity options: exclude_from_indexes = 'actors, tags'

D. Set the following in your entity options: exclude_from_indexes = 'date_published'

- A. Option A
- B. Option C
- C. Option D
- D. Option B.

Answer: A

NEW QUESTION # 237

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

Answer: D

Explanation:

Explanation

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

• • • • •

Latest Professional-Data-Engineer Exam Forum: <https://www.actual4dump.com/Google/Professional-Data-Engineer-actualtests-dumps.html>

- DOWNLOAD the newest Actual4dump Professional-Data-Engineer PDF dumps from Cloud Storage for free:
<https://drive.google.com/open?id=1bwifQTHT4THmHFqGd76l1TWA8WmvBBCj>

