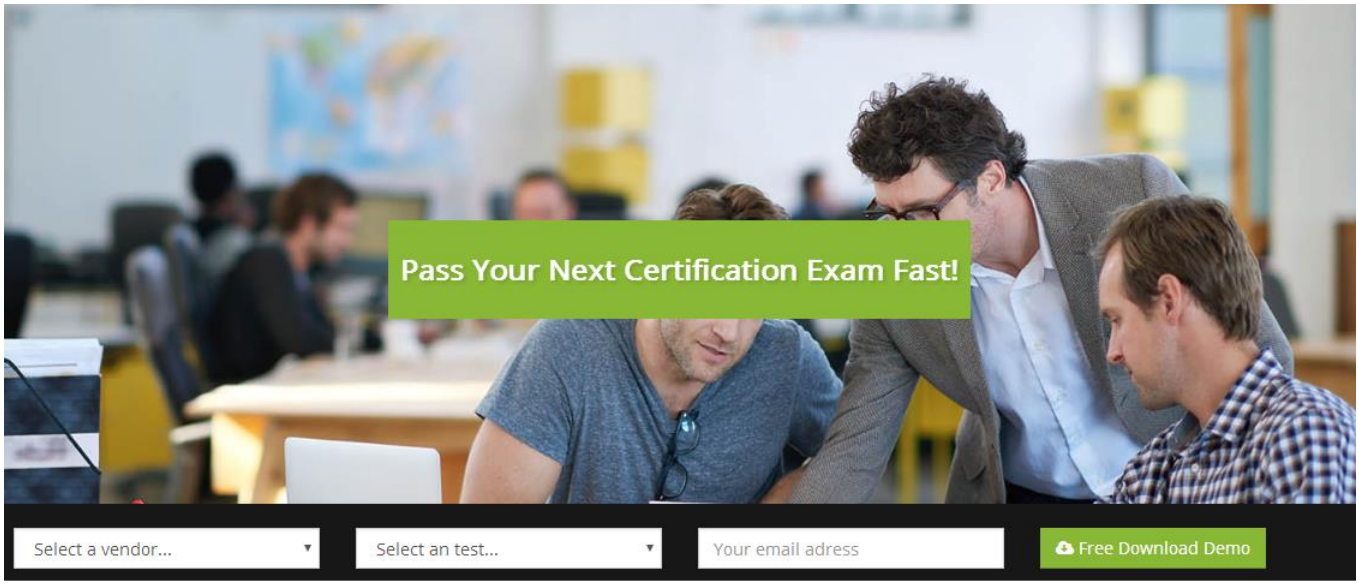
ITPassLeader



Instant Download | 365 Days Free Updates | Money Back Guarantee | Security & Privacy

Choose the version that fits your needs	PDF Version	Desktop Test Engine	Online Test Engine
Latest and Up-to-Date exam dumps with real exam questions answers.	✓	✓	✓
Get 12-Months free updates without any extra charges.	✓	✓	✓
Experience same exam environment before appearing in the certification exam.	✗	✓	✓
100% exam passing guarante in the first attempt.	✓	✓	✓
20% discount on more than one license and 30% discount on 5+ license purchases.	✗	✓	✓
100% secure purchase on SSL.	✓	✓	✓
Completely private purchase without sharing your personal info with anyone.	✓	✓	✓

http://www.itpassleader.com

High-praise Exam Dumps Questions grant you success by high pass rate - ITPassLeader

Exam : **Associate-Data-Practitioner**

Title : Google Cloud Associate Data Practitioner

Vendor : Google

Version : DEMO

NO.1 You created a curated dataset of market trends in BigQuery that you want to share with multiple external partners. You want to control the rows and columns that each partner has access to. You want to follow Google-recommended practices. What should you do?

A. Publish the dataset in Analytics Hub. Grant dataset-level access to each partner by using subscriptions.

B. Create a separate Cloud Storage bucket for each partner. Export the dataset to each bucket and assign each partner to their respective bucket. Grant bucket-level access by using 1AM roles.

C. Grant each partner read access to the BigQuery dataset by using 1AM roles.

D. Create a separate project for each partner and copy the dataset into each project. Publish each dataset in Analytics Hub. Grant dataset-level access to each partner by using subscriptions.

Answer: A

Explanation:

Comprehensive and Detailed in Depth Explanation:

Why A is correct:Analytics Hub allows you to share datasets with external partners while maintaining control over access.

Subscriptions allow granular control.

Why other options are incorrect:B: Cloud storage is for files, not bigquery datasets.

C: IAM roles do not allow for granular row and column level control.

D: Creating a separate project for each partner is complex and not scalable.

NO.2 You are using your own data to demonstrate the capabilities of BigQuery to your organization's leadership team. You need to perform a one-time load of the files stored on your local machine into BigQuery using as little effort as possible. What should you do?

A. Write and execute a Python script using the BigQuery Storage Write API library.

B. Create a Dataproc cluster, copy the files to Cloud Storage, and write an Apache Spark job using the spark-bigquery-connector.

C. Execute the bq load command on your local machine.

D. Create a Dataflow job using the Apache Beam FileIO and BigQueryIO connectors with a local runner.

Answer: C

Explanation:

Comprehensive and Detailed In-Depth Explanation:

A one-time load with minimal effort points to a simple, out-of-the-box tool. The files are local, so the solution must bridge on-premises to BigQuery easily.

* Option A: A Python script with the Storage Write API requires coding, setup (authentication, libraries), and debugging-more effort than necessary for a one-time task.

* Option B: Dataproc with Spark involves cluster creation, file transfer to Cloud Storage, and job scripting-far too complex for a simple load.

* Option C: The bq load command (part of the Google Cloud SDK) is a CLI tool that uploads local files (e.g., CSV, JSON) directly to BigQuery with one command (e.g., bq load --source_format=CSV dataset. table file.csv). It's pre-built, requires no coding, and leverages existing SDK installation, minimizing effort.

NO.3 Your team needs to analyze large datasets stored in BigQuery to identify trends in user behavior. The analysis will involve complex statistical calculations, Python packages, and

visualizations. You need to recommend a managed collaborative environment to develop and share the analysis. What should you recommend?

A. Create a Colab Enterprise notebook and connect the notebook to BigQuery. Share the notebook with your team. Analyze the data and generate visualizations in Colab Enterprise.

B. Create a statistical model by using BigQuery ML. Share the query with your team. Analyze the data and generate visualizations in Looker Studio.

C. Create a Looker Studio dashboard and connect the dashboard to BigQuery. Share the dashboard with your team. Analyze the data and generate visualizations in Looker Studio.

D. Connect Google Sheets to BigQuery by using Connected Sheets. Share the Google Sheet with your team. Analyze the data and generate visualizations in Gooqle Sheets.

Answer: A

Explanation:

Using a Colab Enterprise notebook connected to BigQuery provides a managed, collaborative environment ideal for complex statistical calculations, Python packages, and visualizations. Colab Enterprise supports Python libraries for advanced analytics and offers seamless integration with BigQuery for querying large datasets. It allows teams to collaboratively develop and share analyses while taking advantage of its visualization capabilities. This approach is particularly suitable for tasks involving sophisticated computations and custom visualizations.

NO.4 You are a database administrator managing sales transaction data by region stored in a BigQuery table. You need to ensure that each sales representative can only see the transactions in their region. What should you do?

A. Grant the appropriate 1AM permissions on the dataset.

B. Create a data masking rule.

C. Create a row-level access policy.

D. Add a policy tag in BigQuery.

Answer: C

NO.5 Your organization needs to implement near real-time analytics for thousands of events arriving each second in Pub/Sub. The incoming messages require transformations. You need to configure a pipelinethat processes, transforms, and loads the data into BigQuery while minimizing development time. What should you do?

A. Use a Google-provided Dataflow template to process the Pub/Sub messages, perform transformations, and write the results to BigQuery.

B. Create a Cloud Data Fusion instance and configure Pub/Sub as a source. Use Data Fusion to process the Pub/Sub messages, perform transformations, and write the results to BigQuery.

C. Load the data from Pub/Sub into Cloud Storage using a Cloud Storage subscription. Create a Dataproc cluster, use PySpark to perform transformations in Cloud Storage, and write the results to BigQuery.

D. Use Cloud Run functions to process the Pub/Sub messages, perform transformations, and write the results to BigQuery.

Answer: A

Explanation:

Using aGoogle-provided Dataflow templateis the most efficient and development-friendly approach

to implement near real-time analytics for Pub/Sub messages. Dataflow templates are pre-built and optimized for processing streaming data, allowing you to quickly configure and deploy a pipeline with minimal development effort. These templates can handle message ingestion from Pub/Sub, perform necessary transformations, and load the processed data into BigQuery, ensuring scalability and low latency for near real- time analytics.

NO.6 You manage a large amount of data in Cloud Storage, including raw data, processed data, and backups. Your organization is subject to strict compliance regulations that mandate data immutability for specific data types.
You want to use an efficient process to reduce storage costs while ensuring that your storage strategy meets retention requirements. What should you do?
A. Configure lifecycle management rules to transition objects to appropriate storage classes based on access patterns. Set up Object Versioning for all objects to meet immutability requirements.
B. Move objects to different storage classes based on their age and access patterns. Use Cloud Key Management Service (Cloud KMS) to encrypt specific objects with customer-managed encryption keys (CMEK) to meet immutability requirements.
C. Create a Cloud Run function to periodically check object metadata, and move objects to the appropriate storage class based on age and access patterns. Use object holds to enforce immutability for specific objects.
D. Use object holds to enforce immutability for specific objects, and configure lifecycle management rules to transition objects to appropriate storage classes based on age and access patterns.
Answer: D
Explanation:
Usingobject holdsandlifecycle management rulesis the most efficient and compliant strategy for this scenario because:
* Immutability: Object holds (temporary or event-based) ensure that objects cannot be deleted or overwritten, meeting strict compliance regulations for data immutability.
* Cost efficiency: Lifecycle management rules automatically transition objects to more cost-effective storage classes based on their age and access patterns.
* Compliance and automation: This approach ensures compliance with retention requirements while reducing manual effort, leveraging built-in Cloud Storage features.

NO.7 You are constructing a data pipeline to process sensitive customer data stored in a Cloud Storage bucket. You need to ensure that this data remains accessible, even in the event of a single-zone outage. What should you do?
A. Set up a Cloud CDN in front of the bucket.
B. Enable Object Versioning on the bucket.
C. Store the data in a multi-region bucket.
D. Store the data in Nearline storaqe.
Answer: C
Explanation:
Storing the data in amulti-region bucketensures high availability and durability, even in the event of a single- zone outage. Multi-region buckets replicate data across multiple locations within the selected region, providing resilience against zone-level failures and ensuring that the data remains accessible. This approach is particularly suitable for sensitive customer data that must remain available without

interruptions.
A single-zone outage requires high availability across zones or regions. Cloud Storage offers location-based redundancy options:
* Option A: Cloud CDN caches content for web delivery but doesn't protect against underlying storage outages-it's for performance, not availability of the source data.
* Option B: Object Versioning retains old versions of objects, protecting against overwrites or deletions, but doesn't ensure availability during a zone failure (still tied to one location).
* Option C: Multi-region buckets (e.g., us or eu) replicate data across multiple regions, ensuring accessibility even if a single zone or region fails. This provides the highest availability for sensitive data in a pipeline.

NO.8 Your organization sends IoT event data to a Pub/Sub topic. Subscriber applications read and perform transformations on the messages before storing them in the data warehouse. During particularly busy times when more data is being written to the topic, you notice that the subscriber applications are not acknowledging messages within the deadline. You need to modify your pipeline to handle these activity spikes and continue to process the messages. What should you do?

A. Retry messages until they are acknowledged.
B Implement flow control on the subscribers
B. Forward unacknowledged messages to a dead-letter topic.
C. Seek back to the last acknowledged message.

Answer: B
Explanation:
Implementingflow control on the subscribersallows the subscriber applications to manage message processing during activity spikes by controlling the rate at which messages are pulled and processed. This prevents overwhelming the subscribers and ensures that messages are acknowledged within the deadline. Flow control helps maintain the stability of your pipeline during high-traffic periods without dropping or delaying messages unnecessarily.

NO.9 You manage a web application that stores data in a Cloud SQL database. You need to improve the read performance of the application by offloading read traffic from the primary database instance. You want to implement a solution that minimizes effort and cost. What should you do?

A. Use Cloud CDN to cache frequently accessed data.
B. Store frequently accessed data in a Memorystore instance.
C. Migrate the database to a larger Cloud SQL instance.
D. Enable automatic backups, and create a read replica of the Cloud SQL instance.

Answer: D
Explanation:
Enabling automatic backups and creating a read replica of the Cloud SQL instance is the best solution to improve read performance. Read replicas allow you to offload read traffic from the primary database instance, reducing its load and improving overall performance. This approach is cost-effective and easy to implement within Cloud SQL. It ensures that the primary instance focuses on write operations while replicas handle read queries, providing a seamless performance boost with minimal effort.

NO.10 Your organization has decided to migrate their existing enterprise data warehouse to

BigQuery. The existing data pipeline tools already support connectors to BigQuery. You need to identify a data migration approach that optimizes migration speed. What should you do?

A. Create a temporary file system to facilitate data transfer from the existing environment to Cloud Storage. Use Storage Transfer Service to migrate the data into BigQuery.

B. Use the Cloud Data Fusion web interface to build data pipelines. Create a directed acyclic graph (DAG) that facilitates pipeline orchestration.

C. Use the existing data pipeline tool's BigQuery connector to reconfigure the data mapping.

D. Use the BigQuery Data Transfer Service to recreate the data pipeline and migrate the data into BigQuery.

Answer: C

Explanation:

Since your existing data pipeline tools already support connectors to BigQuery, the most efficient approach is touse the existing data pipeline tool's BigQuery connectorto reconfigure the data mapping. This leverages your current tools, reducing migration complexity and setup time, while optimizing migration speed. By reconfiguring the data mapping within the existing pipeline, you can seamlessly direct the data into BigQuery without needing additional services or intermediary steps.

NO.11 You manage an ecommerce website that has a diverse range of products. You need to forecast future product demand accurately to ensure that your company has sufficient inventory to meet customer needs and avoid stockouts. Your company's historical sales data is stored in a BigQuery table. You need to create a scalable solution that takes into account the seasonality and historical data to predict product demand. What should you do?

A. Use the historical sales data to train and create a BigQuery ML time series model. Use the ML.FORECAST function call to output the predictions into a new BigQuery table.

B. Use Colab Enterprise to create a Jupyter notebook. Use the historical sales data to train a custom prediction model in Python.

C. Use the historical sales data to train and create a BigQuery ML linear regression model. Use the ML.

PREDICT function call to output the predictions into a new BigQuery table.

D. Use the historical sales data to train and create a BigQuery ML logistic regression model. Use the ML.PREDICT function call to output the predictions into a new BigQuery table.

Answer: A

Explanation:

Comprehensive and Detailed In-Depth Explanation:

Forecasting product demand with seasonality requires a time series model, and BigQuery ML offers a scalable, serverless solution. Let's analyze:

* Option A: BigQuery ML's time series models (e.g., ARIMA_PLUS) are designed for forecasting with seasonality and trends. The ML.FORECAST function generates predictions based on historical data, storing them in a table. This is scalable (no infrastructure) and integrates natively with BigQuery, ideal for ecommerce demand prediction.

* Option B: Colab Enterprise with a custom Python model (e.g., Prophet) is flexible but requires coding, maintenance, and potentially exporting data, reducing scalability compared to BigQuery ML's in-place processing.

* Option C: Linear regression predicts continuous values but doesn't handle seasonality or time series patterns effectively, making it unsuitable for demand forecasting.