Google Professional-Data-Engineer Prep Guide, Professional-Data-Engineer Valid Dumps Demo

BTW, DOWNLOAD part of PassSureExam Professional-Data-Engineer dumps from Cloud Storage: https://drive.google.com/open?id=1kr0byNZZkoywpAIo1E9nDw3h8jloYUHa

Those who are ambitious to obtain Professional-Data-Engineer certification mainly include office workers; they expect to reach a higher position and get handsome salary, moreover, a prosperous future. All of these requirements our Professional-Data-Engineer exam materials can meet. Our Professional-Data-Engineer study materials can help you pass the exam successful. Before you decide to buy our Professional-Data-Engineer Exam Torrent, you can free download the demo of our Professional-Data-Engineer exam questions, which contains a few of questions and answers of our Professional-Data-Engineer training guide.

Great concentrative progress has been made by our company, who aims at further cooperation with our candidates in the way of using our Professional-Data-Engineer exam engine as their study tool. Owing to the devotion of our professional research team and responsible working staff, our training materials have received wide recognition and now, with more people joining in the Professional-Data-Engineer Exam army, we has become the top-raking Professional-Data-Engineer training materials provider in the international market. Believe in our Professional-Data-Engineer study guide, you will succeed in your exam!

>> Google Professional-Data-Engineer Prep Guide <<

Professional-Data-Engineer Valid Dumps Demo - New Professional-Data-Engineer Braindumps Questions

As we all know, if the content of your exam materials is complex and confusing, then if you want to pass the exam, you will be quite worried. Our Professional-Data-Engineer study guide helps the candidates to easily follow the needed contents with simplified languages and skillfully explanations according the perfect designs of the professional experts. Preparing with the help of our Professional-Data-Engineer Exam Questions frees you from getting help from other study sources, and you can pass the exam with 100% success guarantee.

Google Professional-Data-Engineer Certification Exam is an online exam that tests the candidate's knowledge and skills in various areas such as data engineering, data analysis, machine learning, and big data processing. Professional-Data-Engineer exam consists of multiple-choice questions and requires a thorough understanding of the Google Cloud Platform, its services, and its features.

Candidates need to demonstrate their proficiency in designing and implementing data processing systems that meet business requirements.

Google Certified Professional Data Engineer Exam Sample Questions (Q26-Q31):

NEW QUESTION # 26
Flowlogistic Case Study
Company Overview
Flowlogistic is a leading logistics and supply chain provider. They help businesses throughout the world manage their resources and transport them to their final destination. The company has grown rapidly, expanding their offerings to include rail, truck, aircraft, and oceanic shipping.
Company Background
The company started as a regional trucking company, and then expanded into other logistics market.
Because they have not updated their infrastructure, managing and tracking orders and shipments has become a bottleneck. To improve operations, Flowlogistic developed proprietary technology for tracking shipments in real time at the parcel level. However, they are unable to deploy it because their technology stack, based on Apache Kafka, cannot support the processing volume. In addition, Flowlogistic wants to further analyze their orders and shipments to determine how best to deploy their resources.
Solution Concept
Flowlogistic wants to implement two concepts using the cloud:
Use their proprietary technology in a real-time inventory-tracking system that indicates the location of
their loads
Perform analytics on all their orders and shipment logs, which contain both structured and unstructured
data, to determine how best to deploy resources, which markets to expand info. They also want to use predictive analytics to learn earlier when a shipment will be delayed.
Existing Technical Environment
Flowlogistic architecture resides in a single data center:
Databases
- 8 physical servers in 2 clusters
- SQL Server - user data, inventory, static data
- 3 physical servers
- Cassandra - metadata, tracking messages
10 Kafka servers - tracking message aggregation and batch insert
Application servers - customer front end, middleware for order/customs
- 60 virtual machines across 20 physical servers
- Tomcat - Java services
- Nginx - static content
- Batch servers
Storage appliances
- iSCSI for virtual machine (VM) hosts
- Fibre Channel storage area network (FC SAN) - SQL server storage
Network-attached storage (NAS) image storage, logs, backups
10 Apache Hadoop /Spark servers
- Core Data Lake
- Data analysis workloads
20 miscellaneous servers
- Jenkins, monitoring, bastion hosts,
Business Requirements
Build a reliable and reproducible environment with scaled panty of production.
Aggregate data in a centralized Data Lake for analysis
Use historical data to perform predictive analytics on future shipments
Accurately track every shipment worldwide using proprietary technology
Improve business agility and speed of innovation through rapid provisioning of new resources
Analyze and optimize architecture for performance in the cloud

Migrate fully to the cloud if all other requirements are met
Technical Requirements
Handle both streaming and batch data
Migrate existing Hadoop workloads
Ensure architecture is scalable and elastic to meet the changing demands of the company.
Use managed services whenever possible
Encrypt data flight and at rest
Connect a VPN between the production data center and cloud environment
SEO Statement
We have grown so quickly that our inability to upgrade our infrastructure is really hampering further growth and efficiency. We are efficient at moving shipments around the world, but we are inefficient at moving data around.
We need to organize our information so we can more easily understand where our customers are and what they are shipping.
CTO Statement
IT has never been a priority for us, so as our data has grown, we have not invested enough in our technology. I have a good staff to manage IT, but they are so busy managing our infrastructure that I cannot get them to do the things that really matter, such as organizing our data, building the analytics, and figuring out how to implement the CFO' s tracking technology.
CFO Statement
Part of our competitive advantage is that we penalize ourselves for late shipments and deliveries. Knowing where out shipments are at all times has a direct correlation to our bottom line and profitability.
Additionally, I don't want to commit capital to building out a server environment.
Flowlogistic's management has determined that the current Apache Kafka servers cannot handle the data volume for their real-time inventory tracking system. You need to build a new system on Google Cloud Platform (GCP) that will feed the proprietary tracking software. The system must be able to ingest data from a variety of global sources, process and query in real-time, and store the data reliably. Which combination of GCP products should you choose?

- A. Cloud Pub/Sub, Cloud Dataflow, and Cloud Storage
- B. Cloud Pub/Sub, Cloud SQL, and Cloud Storage
- C. Cloud Dataflow, Cloud SQL, and Cloud Storage
- D. Cloud Load Balancing, Cloud Dataflow, and Cloud Storage
- E. Cloud Pub/Sub, Cloud Dataflow, and Local SSD

Answer: B

Explanation:
Explanation/Reference:

NEW QUESTION # 27
Which of these statements about BigQuery caching is true?

- A. Query results are cached even if you specify a destination table.
- B. BigQuery caches query results for 48 hours.
- C. There is no charge for a query that retrieves its results from cache.
- D. By default, a query's results are not cached.

Answer: C

Explanation:
When query results are retrieved from a cached results table, you are not charged for the query.
BigQuery caches query results for 24 hours, not 48 hours.
Query results are not cached if you specify a destination table.
A query's results are always cached except under certain conditions, such as if you specify a destination table.

NEW QUESTION # 28
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 visualizations for operations teams with the following requirements:
Which approach meets the requirements?

- A. Load the data into Google Sheets, use formulas to calculate a metric, and use filters/sorting to show only suboptimal links in a table.
- B. Load the data into Google BigQuery tables, write Google Apps Script that queries the data, calculates the metric, and shows only suboptimal rows in a table in Google Sheets.
- C. Load the data into Google Cloud Datastore tables, write a Google App Engine Application that queries all rows, applies a function to derive the metric, and then renders results in a table using the Google charts and visualization API.
- D. Load the data into Google BigQuery tables, write a Google Data Studio 360 report that connects to your data, calculates a metric, and then uses a filter expression to show only suboptimal rows in a table.

Answer: C

NEW QUESTION # 29
Your company is in a highly regulated industry. One of your requirements is to ensure individual users have access only to the minimum amount of information required to do their jobs. You want to enforce this requirement with Google BigQuery. Which three approaches can you take? (Choose three.)

- A. Restrict access to tables by role.
- B. Ensure that the data is encrypted at all times.
- C. Disable writes to certain tables.
- D. Segregate data across multiple tables or databases.
- E. Use Google Stackdriver Audit Logging to determine policy violations.
- F. Restrict BigQuery API access to approved users.

Answer: A,E,F

NEW QUESTION # 30
You used Cloud Dataprep to create a recipe on a sample of data in a BigQuery table. You want to reuse this recipe on a daily upload of data with the same schema, after the load job with variable execution time completes. What should you do?

- A. Export the recipe as a Cloud Dataprep template, and create a job in Cloud Scheduler.
- B. Export the Cloud Dataprep job as a Cloud Dataflow template, and incorporate it into a Cloud Composer job.
- C. Create an App Engine cron job to schedule the execution of the Cloud Dataprep job.
- D. Create a cron schedule in Cloud Dataprep.

Answer: B

Explanation:
Topic 1, Flowlogistic Case Study
Company Overview
Flowlogistic is a leading logistics and supply chain provider. They help businesses throughout the world manage their resources and transport them to their final destination. The company has grown rapidly, expanding their offerings to include rail, truck, aircraft, and oceanic shipping.
Company Background
The company started as a regional trucking company, and then expanded into other logistics market. Because they have not updated their infrastructure, managing and tracking orders and shipments has become a bottleneck. To improve operations, Flowlogistic developed proprietary technology for tracking shipments in real time at the parcel level. However, they are unable to deploy it because their technology stack, based on Apache Kafka, cannot support the processing volume. In addition, Flowlogistic wants to further analyze their orders and shipments to determine how best to deploy their resources.
Solution Concept
Flowlogistic wants to implement two concepts using the cloud:
Use their proprietary technology in a real-time inventory-tracking system that indicates the location of their loads Perform analytics on all their orders and shipment logs, which contain both structured and unstructured data, to determine how best to deploy resources, which markets to expand info. They also want to use predictive analytics to learn earlier when a shipment will be delayed.
Existing Technical Environment
Flowlogistic architecture resides in a single data center:
Databases
8 physical servers in 2 clusters
SQL Server - user data, inventory, static data
3 physical servers
Cassandra - metadata, tracking messages
10 Kafka servers - tracking message aggregation and batch insert
Application servers - customer front end, middleware for order/customs
60 virtual machines across 20 physical servers
Tomcat - Java services
Nginx - static content
Batch servers
Storage appliances
iSCSI for virtual machine (VM) hosts
Fibre Channel storage area network (FC SAN) - SQL server storage
Network-attached storage (NAS) image storage, logs, backups
Apache Hadoop /Spark servers
Core Data Lake
Data analysis workloads
20 miscellaneous servers
Jenkins, monitoring, bastion hosts,
Business Requirements

Build a reliable and reproducible environment with scaled panty of production.
Aggregate data in a centralized Data Lake for analysis
Use historical data to perform predictive analytics on future shipments Accurately track every shipment worldwide using proprietary technology Improve business agility and speed of innovation through rapid provisioning of new resources Analyze and optimize architecture for performance in the cloud Migrate fully to the cloud if all other requirements are met Technical Requirements Handle both streaming and batch data Migrate existing Hadoop workloads Ensure architecture is scalable and elastic to meet the changing demands of the company.
Use managed services whenever possible
Encrypt data flight and at rest
Connect a VPN between the production data center and cloud environment
SEO Statement
We have grown so quickly that our inability to upgrade our infrastructure is really hampering further growth and efficiency. We are efficient at moving shipments around the world, but we are inefficient at moving data around.
We need to organize our information so we can more easily understand where our customers are and what they are shipping.
CTO Statement
IT has never been a priority for us, so as our data has grown, we have not invested enough in our technology. I have a good staff to manage IT, but they are so busy managing our infrastructure that I cannot get them to do the things that really matter, such as organizing our data, building the analytics, and figuring out how to implement the CFO' s tracking technology.
CFO Statement
Part of our competitive advantage is that we penalize ourselves for late shipments and deliveries. Knowing where out shipments are at all times has a direct correlation to our bottom line and profitability. Additionally, I don't want to commit capital to building out a server environment.

NEW QUESTION # 31
.....

In the matter of quality, our Professional-Data-Engineer practice engine is unsustainable with reasonable prices. Despite costs are constantly on the rise these years from all lines of industry, our Professional-Data-Engineer learning materials remain low level. That is because our company beholds customer-oriented tenets that guide our everyday work. The achievements of wealth or prestige is no important than your exciting feedback about efficiency and profession of our Professional-Data-Engineer Practice Engine. So our Professional-Data-Engineer practice materials are great materials you should be proud of and we are!

Engineer Study Tool

- Reliable Professional-Data-Engineer Test Question □ Professional-Data-Engineer Pass Guide □ Professional-Data-Engineer Valid Test Answers □ Search for □ Professional-Data-Engineer □ and obtain a free download on □ www.pdfvce.com □ □Exam Professional-Data-Engineer PDF
- Professional-Data-Engineer Reliable Test Questions □ Professional-Data-Engineer Free Study Material □ Sample Professional-Data-Engineer Exam □ Search for ⇛ Professional-Data-Engineer ⇚ and obtain a free download on ➠ www.prepawayexam.com □ □Sample Professional-Data-Engineer Exam
- haimakrpt930731.answerblogs.com, bookmarkity.com, ummalife.com, hannabfgr058475.spintheblog.com, roryrxst756303.blogsidea.com, emilyasgt457813.scrappingwiki.com, joshakvj883993.aboutyoublog.com, www.stes.tyc.edu.tw, louiseeduf870726.onzeblog.com, carlyevny728112.blogrenanda.com, Disposable vapes

DOWNLOAD the newest PassSureExam Professional-Data-Engineer PDF dumps from Cloud Storage for free: https://drive.google.com/open?id=1kr0byNZZkoywpAIo1E9nDw3h8jloYUHa