



**Free Questions for Professional-Data-Engineer by
actualtestdumps**

Shared by Spears on 07-06-2022

For More Free Questions and Preparation Resources

Check the Links on Last Page

Question 1

Question Type: MultipleChoice

Your company built a TensorFlow neural-network model with a large number of neurons and layers. The model fits well for the training data. However, when tested against new data, it performs poorly. What method can you employ to address this?

Options:

- A) Threading
- B) Serialization
- C) Dropout Methods
- D) Dimensionality Reduction

Answer:

C

Explanation:

Reference <https://medium.com/mlreview/a-simple-deep-learning-model-for-stock-price-prediction-using-tensorflow-30505541d877>

Question 2

Question Type: MultipleChoice

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= 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'`

Options:

A) Option A

B) Option B.

C) Option C

D) Option D

Answer:

A

Question 3

Question Type: MultipleChoice

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?

Options:

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

Answer:

D

Question 4

Question Type: MultipleChoice

You are working on a sensitive project involving private user data. You have set up a project on Google Cloud Platform to house your work internally. An external consultant is going to assist with coding a complex transformation in a Google Cloud Dataflow pipeline for your project. How should you maintain users' privacy?

Options:

- A) Grant the consultant the Viewer role on the project.
- B) Grant the consultant the Cloud Dataflow Developer role on the project.
- C) Create a service account and allow the consultant to log on with it.
- D) Create an anonymized sample of the data for the consultant to work with in a different project.

Answer:

C

Question 5

Question Type: MultipleChoice

Your company's customer and order databases are often under heavy load. This makes performing analytics against them difficult without harming operations. The databases are in a MySQL cluster, with nightly backups taken using mysqldump. You want to perform analytics with minimal impact on operations. What should you do?

Options:

- A) Add a node to the MySQL cluster and build an OLAP cube there.
- B) Use an ETL tool to load the data from MySQL into Google BigQuery.
- C) Connect an on-premises Apache Hadoop cluster to MySQL and perform ETL.
- D) Mount the backups to Google Cloud SQL, and then process the data using Google Cloud Dataproc.

Answer:

C

Question 6

Question Type: MultipleChoice

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?

Options:

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

Answer:

C

Question 7

Question Type: MultipleChoice

Flowlogistic's CEO wants to gain rapid insight into their customer base so his sales team can be better informed in the field. This team is not very technical, so they've purchased a visualization tool to simplify the creation of BigQuery reports. However, they've been overwhelmed by all the data in the table, and are spending a lot of money on queries trying to find the data they need. You want to solve

their problem in the most cost-effective way. What should you do?

Options:

- A) Export the data into a Google Sheet for virtualization.
- B) Create an additional table with only the necessary columns.
- C) Create a view on the table to present to the virtualization tool.
- D) Create identity and access management (IAM) roles on the appropriate columns, so only they appear in a query.

Answer:

C

To Get Premium Files for Professional-Data-Engineer Visit

<https://www.p2pexams.com/products/professional-data-engineer>

For More Free Questions Visit

<https://www.p2pexams.com/google/pdf/professional-data-engineer>

