Linux Academy (https://linuxacademy.com/cp)   35 🔴   1 🔔   Support ⚙   hieupm.it90@gmail.com 👤

🏠
Home (https://linuxacademy.com/cp)

🎓
Training (https://app.linuxacademy.com/search)

☁
Playground

🕐
Quick Training (https://app.linuxacademy.com/search?duration=%3A7200)

⚗
Hands-On Labs (https://app.linuxacademy.com/search?type=Hands-On%20Lab~Live%20Environment%20Learning%20Activity)

📍
Learning Paths (https://linuxacademy.com/cp/learningpaths)

💬
Community ()

🔍

Congratulations, you passed!

100.0%

Great job on the challenge, you just earned rewards to use on more challenges.

in Add to profile (https://www.linkedin.com/profile/add?startTask=QUIZ: Data Ingest and Processing)

Linux Academy   **Go Back**

Report Card

Expectations	Score
1. Google Cloud Data Engineer - Data Ingest and Processing	100 %

Exam Breakdown

Google Cloud Data Engineer - Data Ingest and Processing   ⌃

1. What is a requirement for running preemptible workers on a Dataproc cluster? ⌃

A. You can only have a maximum of a 2:1 ratio of preemptible to standard worker nodes.	
B. Preemptible nodes are only available on high availability master node configurations.	
C. The cluster must also have at least two standard worker nodes	✓ **Correct**
D. You must configure the startup and shutdown scripts to gracefully handle preemptible nodes being shut down.	

👍 👎

2. Which of the below options is the preferred method to connect your Dataflow pipeline to Bigtable for output? ⌃

A. You cannot connect Bigtable to Dataflow.	
B. Publish Dataflow data to a Pub/Sub topic. Configure Bigtable to subscribe to the topic to pull data.	
C. Use Cloud Storage as a staging ground for outputting into Bigtable.	
D. Use the Cloud Dataflow connector for Bigtable.	✓ **Correct**

👍 👎

3. You need to correct streaming messages that arrive out of order due to latency. Which Google Cloud service would you use to resolve this? ⌃

A. BigQuery	
B. Cloud Dataflow	✓ **Correct**
C. Cloud Pub/Sub	
D. Cloud SQL	

👍 👎

4. Which of these open source technologies is the direct equivalent to Google BigQuery? ⌃

A. Hive	✓ **Correct**
B. Spark	
C. Hadoop	
D. Pig	

👍 👎

5. You need to extract an address field from a multi-column element using Dataflow. Which mechanism is able to help with this task? ⌃

A. Window	
B. ParDo	✓ **Correct**
C. PCollection	
D. Transform	

👍 👎

6. You have a streaming Dataflow pipeline that you need to shut down. You want data already in the pipeline to finish and be sent to output before shutting down. Which shutdown option should you use to complete the shutdown process? ⌃

A. Stop	
B. Drain	✓ **Correct**
C. Shutdown	
D. Cancel	

👍 👎

7. What is the open source equivalent to Cloud Pub/Sub? ⌃

A. Apache Beam	
B. Apache Hadoop	
C. Kubernetes	
D. Apache Kafka	✓ **Correct**

👍 👎

8. Which dataflow concept determines when the contents of a window should be submitted, based on certain conditions being met?

A. Watermark	
B. Trigger	✓ **Correct**
C. Condition	
D. Windows	

👍 👎