

# Welcome to Cert Prep Bootcamp

J/APAC





# Lucas Rafagnin

Google Cloud Partner Engineer - APAC

[linkedin.com/in/rafagnin](https://www.linkedin.com/in/rafagnin)

[rafagnin@google.com](mailto:rafagnin@google.com)

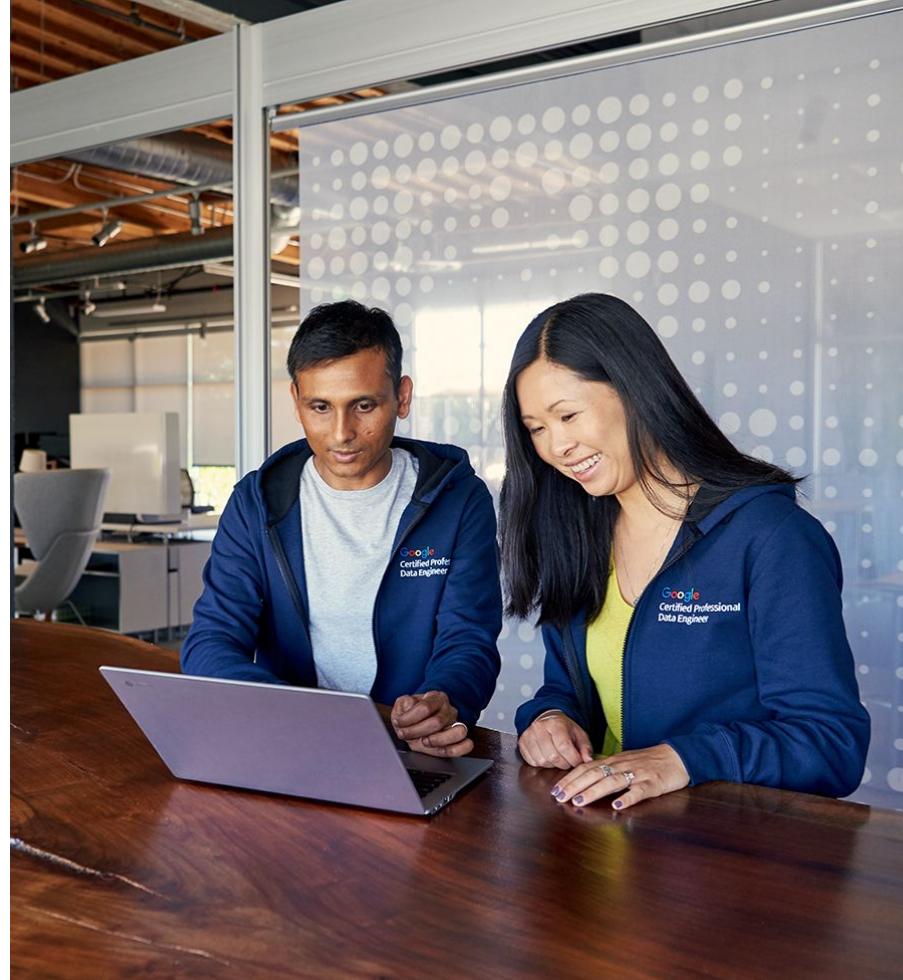


# Agenda:

- Why get Google Cloud Certified?
- Google Cloud Certification Overview
- Certification Preparation Resources for Partners
- Earning your Certification Exam Voucher
- Bootcamp: Practice Exam Question Review - Data Engineer
- How to Access your Qwiklabs Credits
- Q&A



# Why become Google Cloud Certified?





Gain industry recognition!



Validate your technical expertise!



Take your career to the next level!



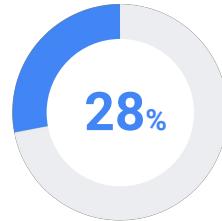
Wear cool Swag!

# Take your career to the next level

Certification value has skyrocketed.



Certifications help you validate, emphasize, and market the expertise that companies and customers are seeking to transform their businesses.



Most organizations are still in the early phases of their cloud journey. Only **28% of enterprises** consider themselves past the exploration phase of cloud adoption.



Lack of resources and expertise was still the #1 cloud challenge in 2017, with **25%** of enterprises citing it as a major concern.

# Boost your career

Certification value remains high, as experts with diverse skills are needed

**Google Professional  
Cloud Architect is the  
highest paying IT  
certification for 2019**

Source – [Global Knowledge](#)

**Solution architect and  
data engineer made the  
top 15 in LinkedIn's most  
promising jobs in 2018.**

# Pass your Exam. Get recognized with swag!

---

Show off your achievement with limited-edition swag and a digital badge that signals to the industry that you have validated skills. Pick out your favorite swag at the [Google Cloud Certification Perks Webstore](#).



# Google Cloud Certification Overview

# Google Cloud Certifications



Associate  
Cloud  
Engineer



Professional  
Collaboration  
Engineer



Professional  
Cloud  
Developer



Professional  
Cloud Architect



Cloud  
Network  
Engineer



Professional  
Data Engineer



Cloud  
Security  
Engineer

## Associate

Recommended  
6+ months hands-on  
experience with GCP

## Professional

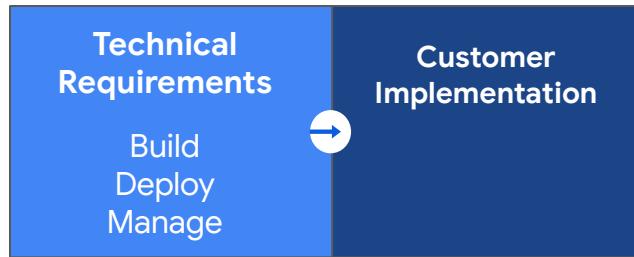
Recommended  
3+ years industry  
experience &  
1 year hands-on  
experience with GCP

# We'll help you reach your destination



# Scope of Google Cloud Certifications

Associate level certification



Professional level certifications





# Professional Data Engineer preparation journey

## 1 Self assess

[Review the exam guide](#)

## 2 Complete trainings

[Coursera training Free of charge for partners](#)

[Data Engineering on Google Cloud Platform Specialization](#)

[Preparing for the Google Cloud Professional Data Engineer Examination](#)

## 3 Hands-on practice

[Qwiklabs exam practice Free training credits](#)

[Data engineering quest](#)

[GCP Essentials](#)

## 4 Review sample solutions

[Detailed solutions](#)

[Sample architecture flow charts](#)

## 5 Assess your knowledge

[Take the practice exam](#)



# Certification Preparation Resources for Partners

# Partner Advantage Portal

## Training Options

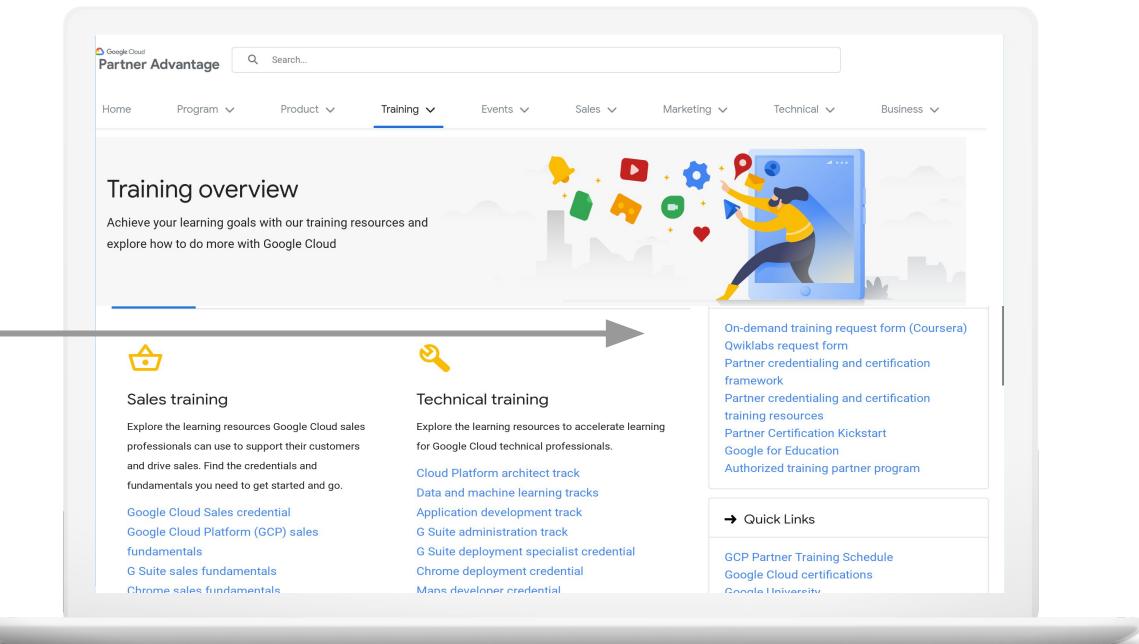
<http://partneradvantage.goog/>

### On-Demand: Learn at your own pace

- [On-demand training via Coursera](#)
- [Hands-on labs with Qwiklabs](#)
- [Practice exams](#)

### Instructor-Led: Interactive workshops

- [Partner-only training event schedule](#)



# Localized Google Cloud certification exams



German



French



Spanish



Portuguese



Japanese



Associate Cloud Engineer



Professional Cloud Architect



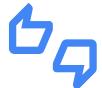
Professional Data Engineer



Professional Cloud Developer



# What to expect: certification exam details



The practitioners exam objectively measures your ability to demonstrate the critical job skills for the role



The exam must be taken in person at one of our testing-center locations



It's a multiple-choice format



Plan on two hours to complete the exam



The registration fee is \$200 USD

# Steps to Earn Certification Exam Voucher



## Attend the Certification Prep Bootcamp

Join us live here, or watch on demand to participate



## Complete the Required Qwiklab Quests

Complete the required Quests  
By 2019.12.20



## Earn your cert exam voucher

We will send you your exam voucher upon Quest completion



## Schedule your exam

Schedule your exam by 2020.01.17



# Revisiting GCP

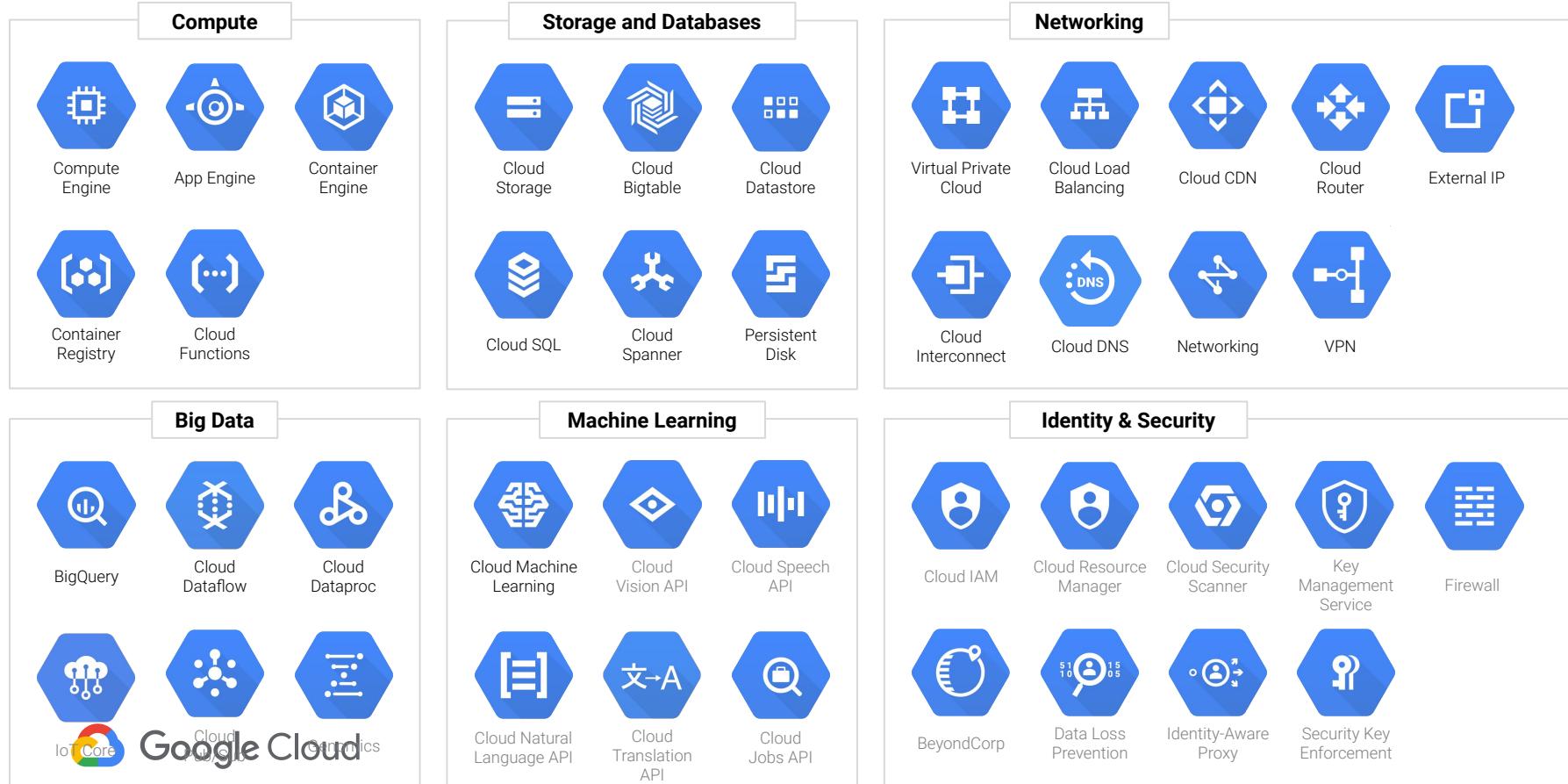




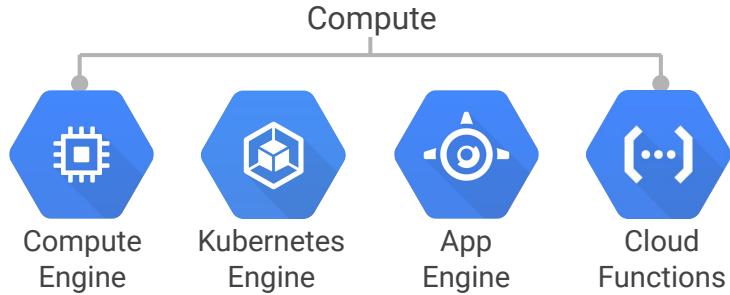
**[https://github.com/gregsramblings/  
google-cloud-4-words](https://github.com/gregsramblings/google-cloud-4-words)**



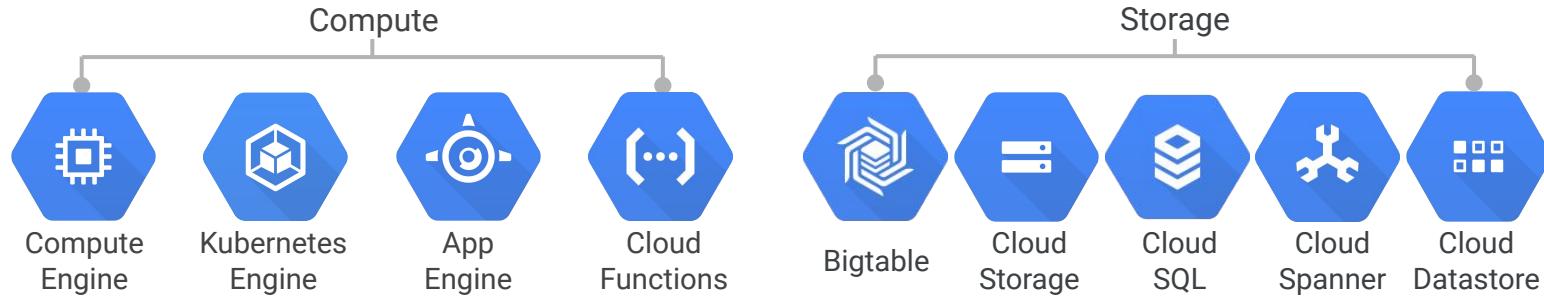
# Google Cloud



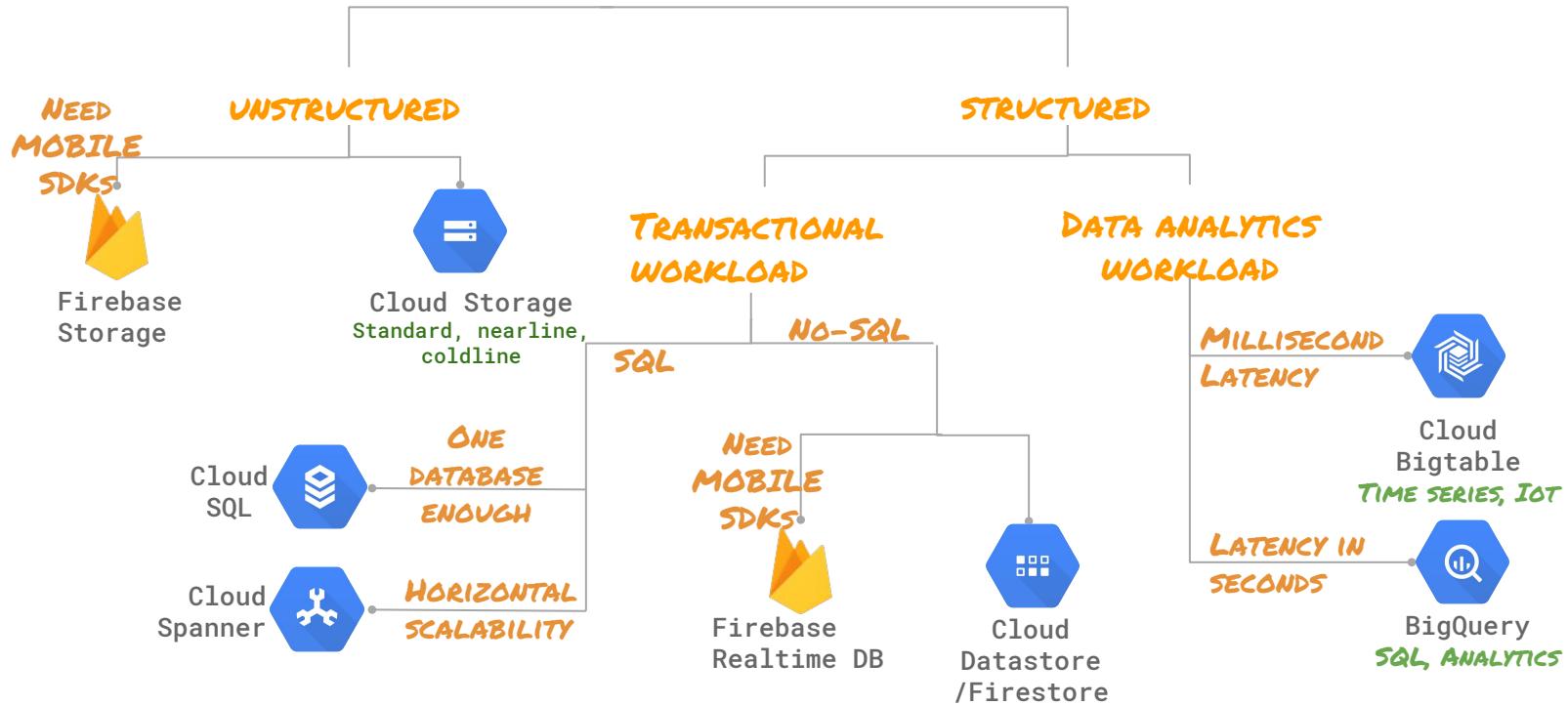
# Products on Google Cloud Platform



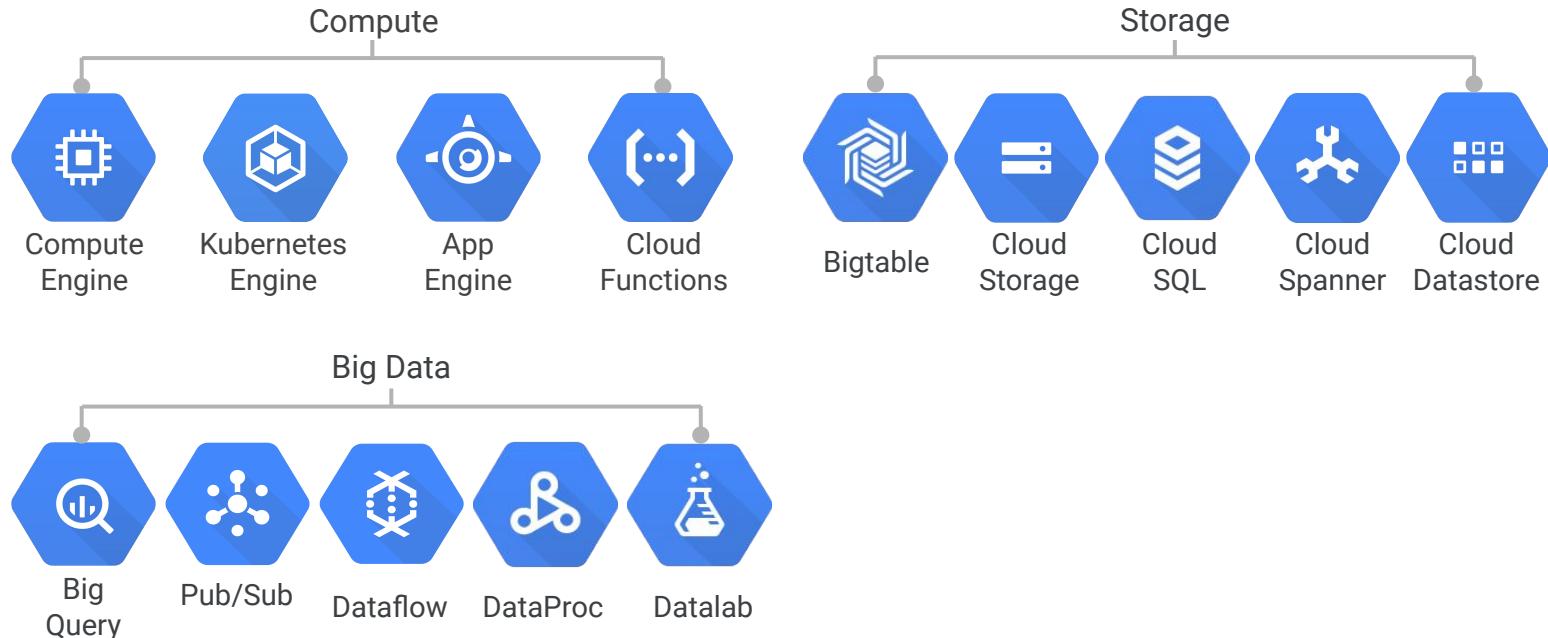
# Products on Google Cloud Platform



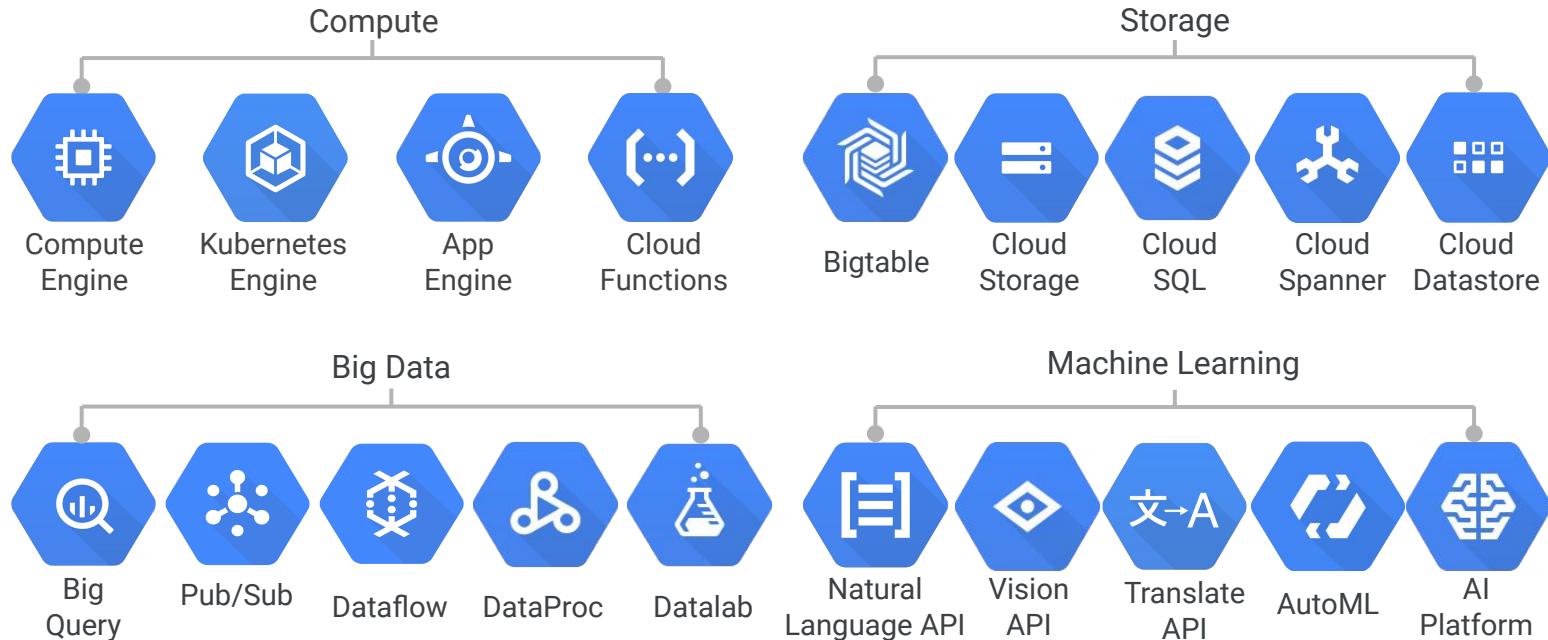
# Choosing where to store data on GCP



# Products on Google Cloud Platform

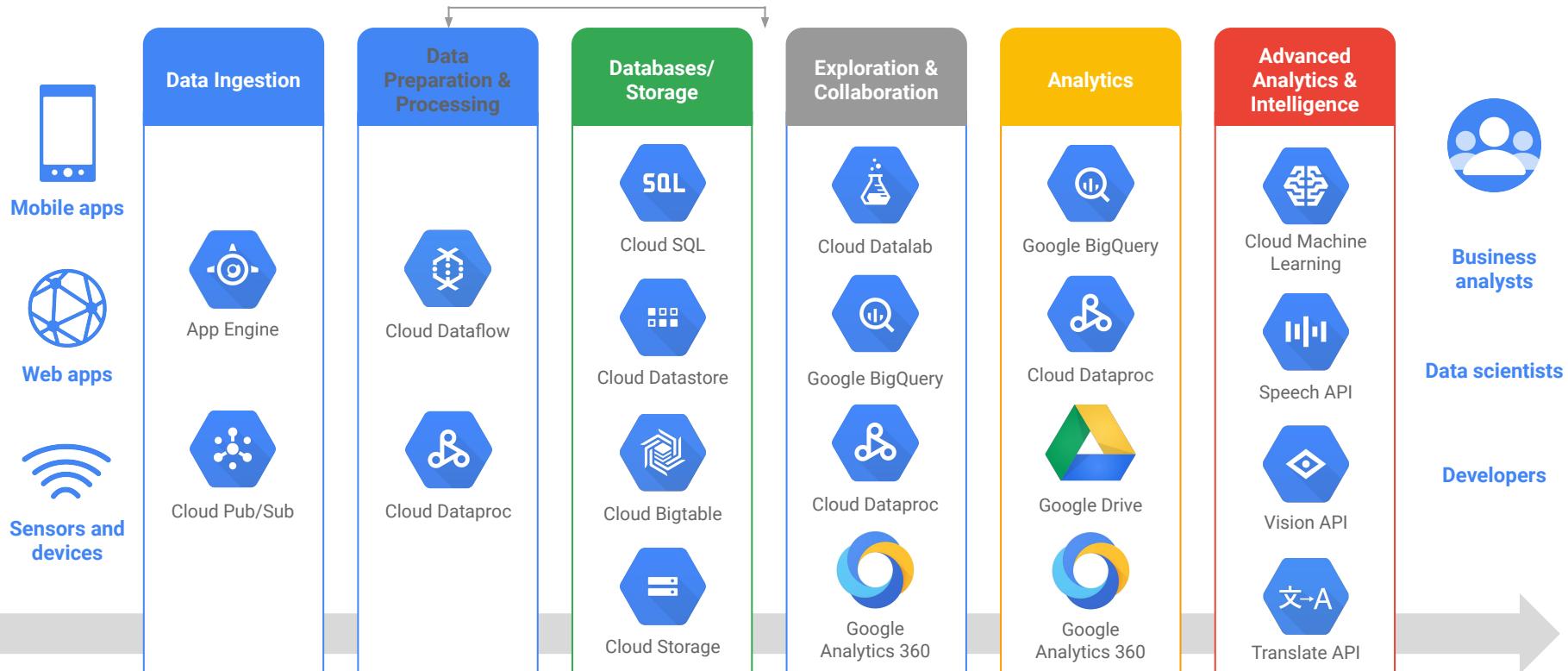


# Products on Google Cloud Platform

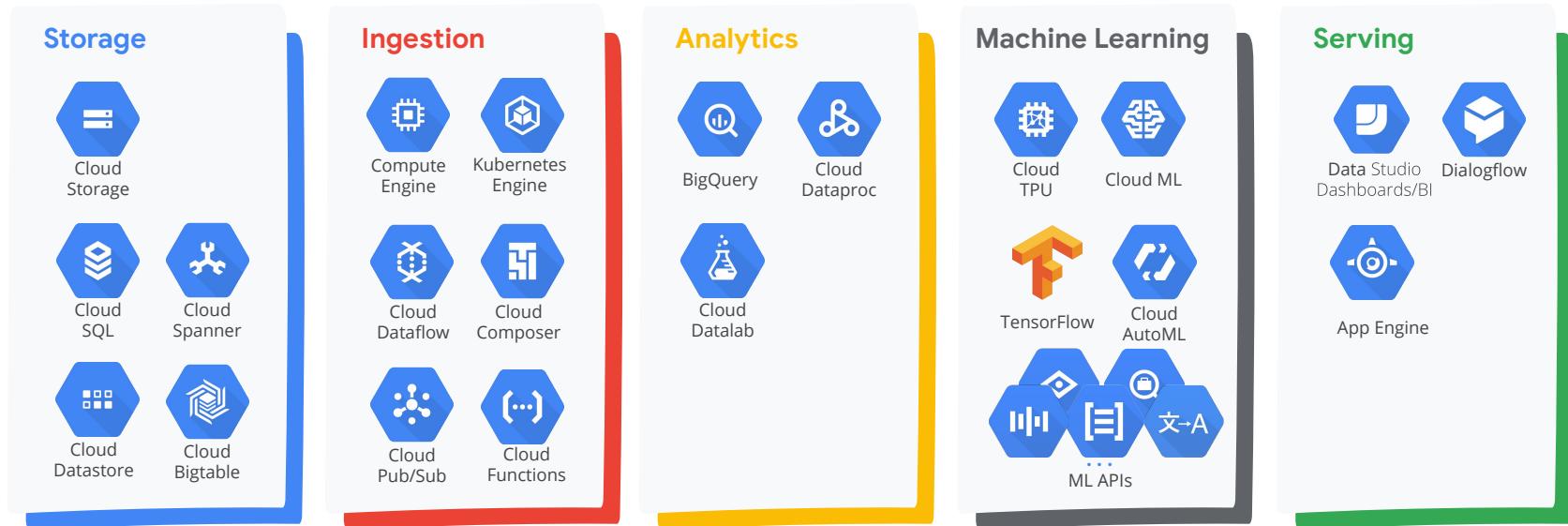


**Product  
Role**      **and**      **Possible Roles**  
                        **and**      **Possible Products**

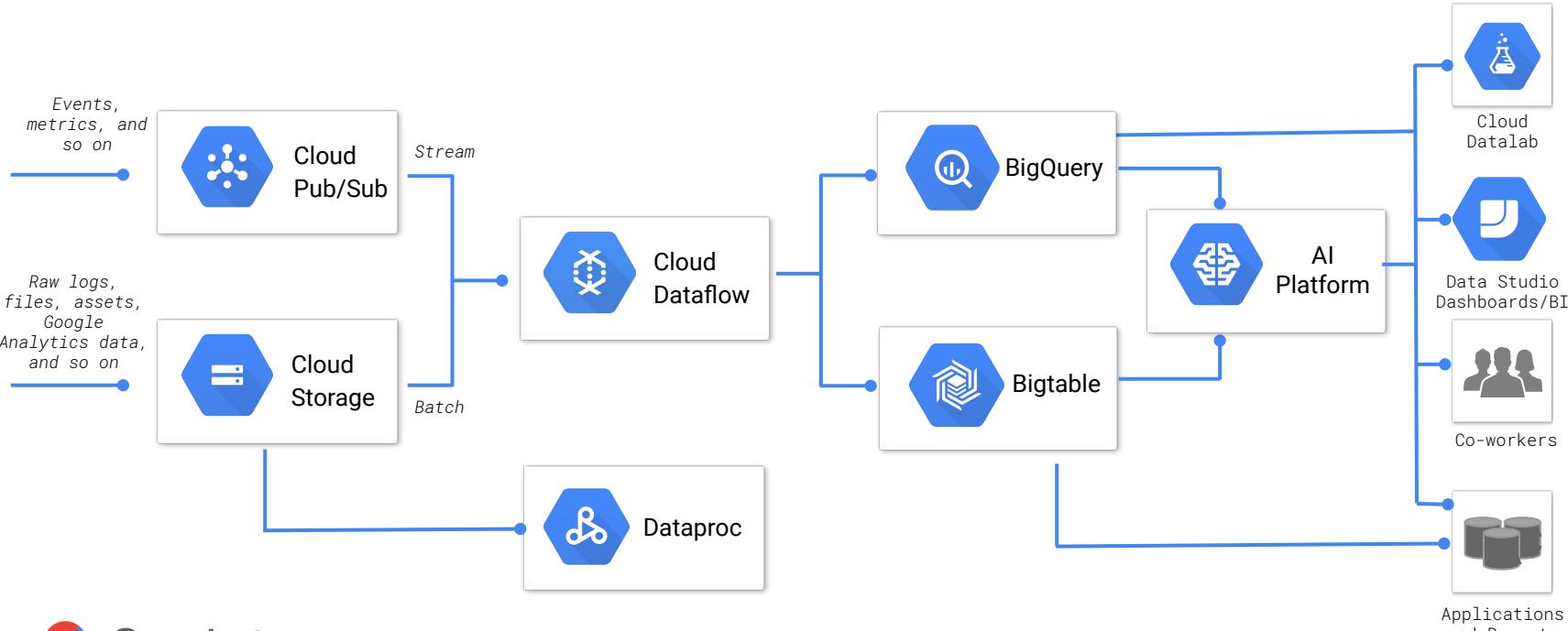
# Google Cloud Data Platform

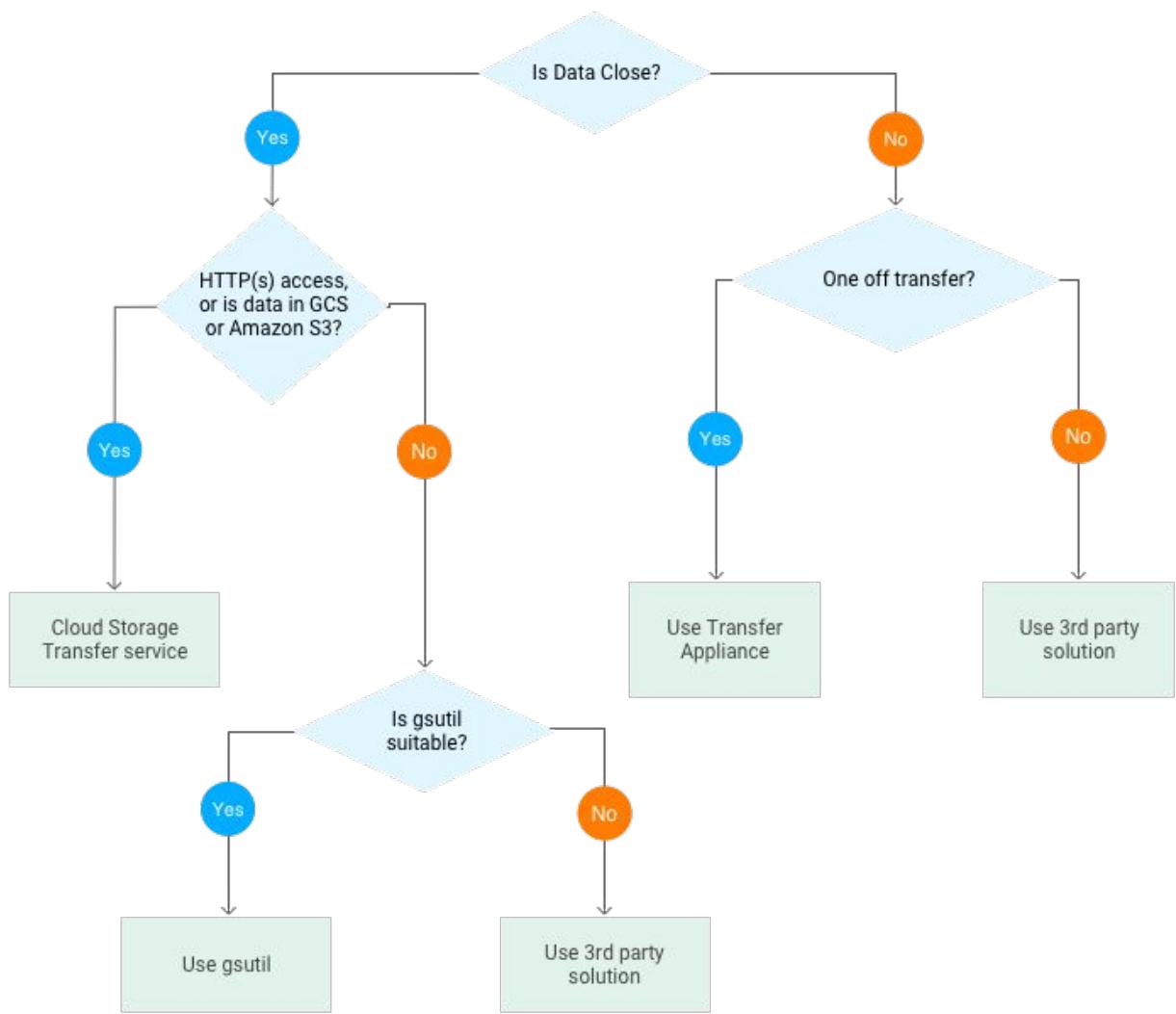


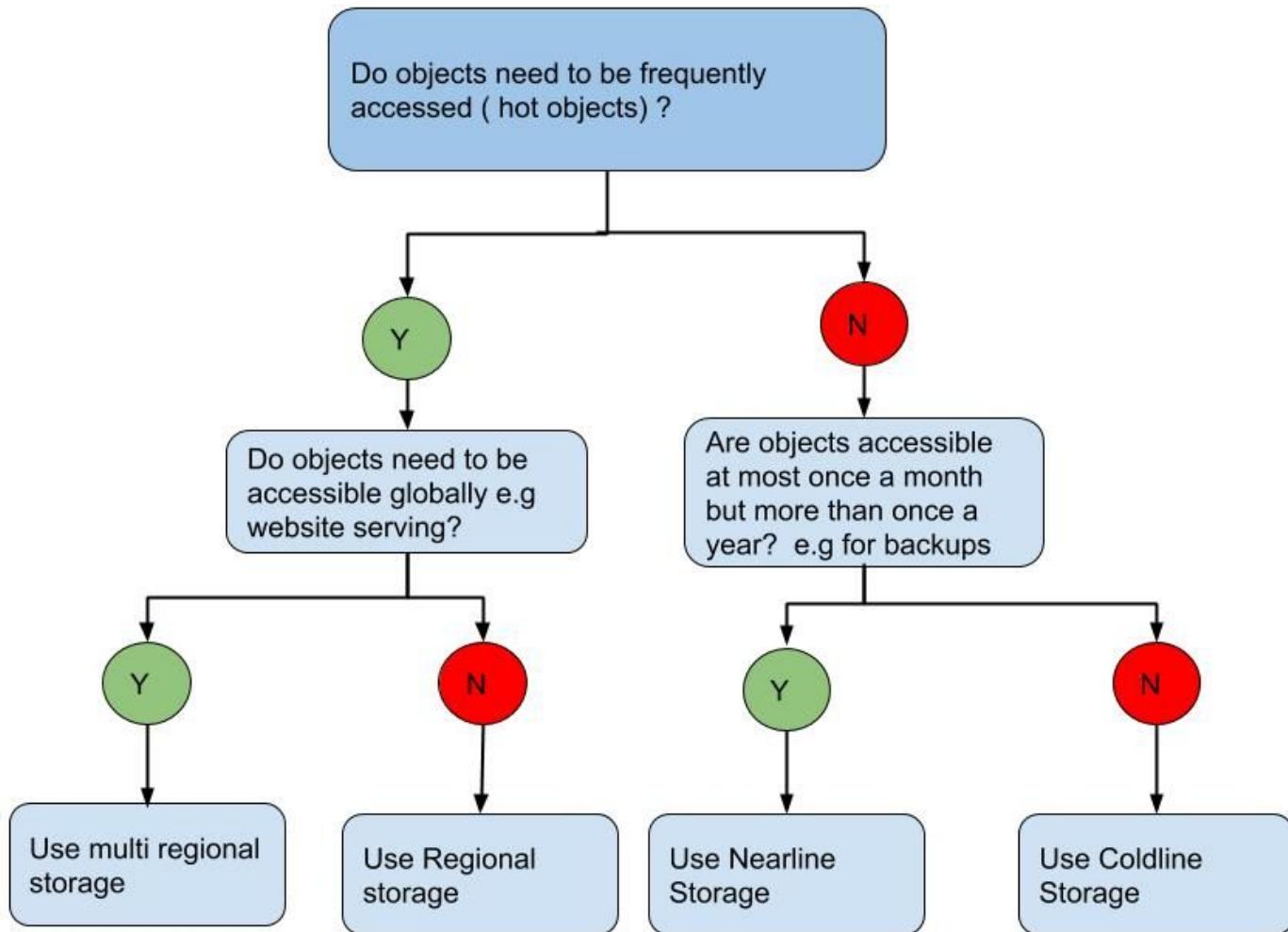
# The suite of big data products on Google Cloud Platform



# Reference Data Pipeline Solution



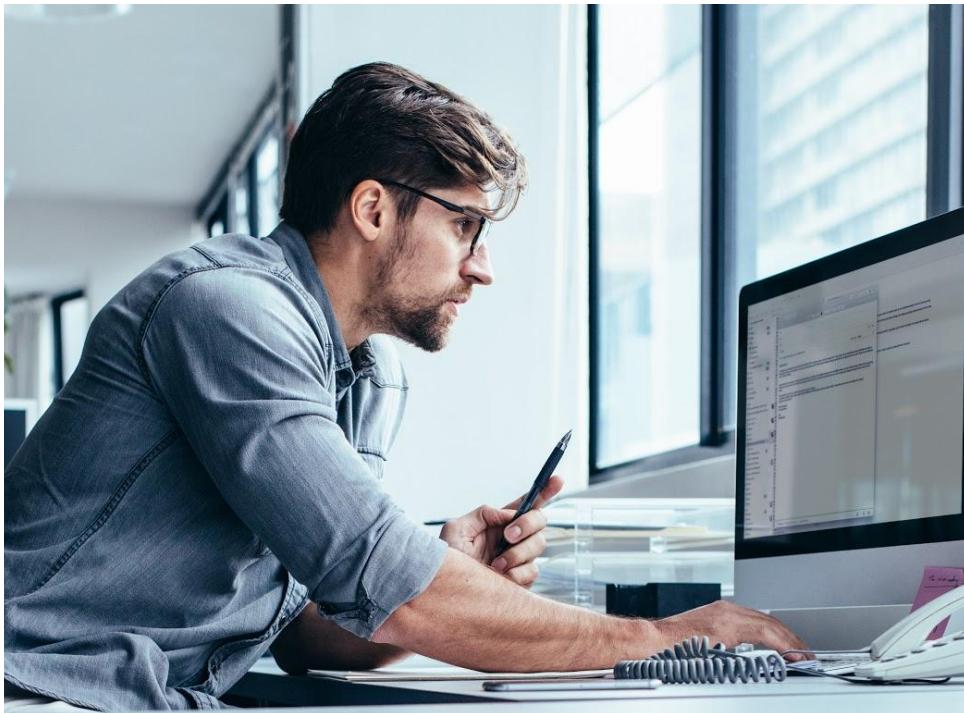




# Stages of Preparation

1. Study
  - a. Coursera + Labs
  - b. Documentation
    - i. Concepts & How to Guides
  - c. Gcp.solutions
  - d. Extra
    - i. Gcpweekly.com, google cloud medium publication
2. Practice Test Approach
3. Test the readiness (extra: LinuxAcademy quizzes)
4. Final Exam

# Bootcamp. Practice Exam Review



# Data Engineer

## Practice Exam Questions

# Reading a Question

1. Problem Statement
2. Constraints
  - a. Cost effective
  - b. Scalable/Future proof
  - c. No rewrite
  - d. NoOps / Managed Services
3. Current Solution

# Answering a Question: Possible Approach

1. Read the complete question
2. Read the complete answers
3. Use the answers
  - a. Eliminate
4. Revisit the question and answers
5. Don't assume
6. Mark for review
  - a. No pen/paper provided

**You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?**

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use [Cloud Storage](#) for storage. [Link data as permanent tables in BigQuery](#) and turn on the "Automatically detect" option in the Schema section of BigQuery.

You are building storage for files for a data pipeline on Google Cloud. You want to support JSON files. The schema of these files will occasionally change. Your analyst teams will use running aggregate ANSI SQL queries on this data. What should you do?

- A. Use BigQuery for storage. Provide format files for data load. Update the format files as needed.
- B. Use BigQuery for storage. Select "Automatically detect" in the Schema section.**
- C. Use Cloud Storage for storage. Link data as temporary tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.
- D. Use Cloud Storage for storage. Link data as permanent tables in BigQuery and turn on the "Automatically detect" option in the Schema section of BigQuery.

You use a Hadoop cluster both for serving analytics and for processing and transforming data. The data is currently stored on HDFS in Parquet format. The data processing jobs run for 6 hours each night. Analytics users can access the system 24 hours a day. Phase 1 is to quickly migrate the entire Hadoop environment without a major re-architecture. Phase 2 will include migrating to BigQuery for analytics and to Cloud Dataflow for data processing. You want to make the future migration to BigQuery and Cloud Dataflow easier by following Google-recommended practices and managed services. What should you do?

- A. Lift and shift Hadoop/HDFS to Cloud Dataproc.
- B. Lift and shift Hadoop/HDFS to Compute Engine.
- C. Create a single Cloud Dataproc cluster to support both analytics and data processing, and point it at a Cloud Storage bucket that contains the Parquet files that were previously stored on HDFS.
- D. Create separate Cloud Dataproc clusters to support analytics and data processing, and point both at the same Cloud Storage bucket that contains the Parquet files that were previously stored on HDFS.

You use a Hadoop cluster both for serving analytics and for processing and transforming data. The data is currently stored on HDFS in Parquet format. The data processing jobs run for 6 hours each night. Analytics users can access the system 24 hours a day. Phase 1 is to quickly migrate the entire Hadoop environment without a major re-architecture. Phase 2 will include migrating to BigQuery for analytics and to Cloud Dataflow for data processing. You want to make the future migration to BigQuery and Cloud Dataflow easier by following Google-recommended practices and managed services. What should you do?

- A. Lift and shift Hadoop/HDFS to Cloud Dataproc.
- B. Lift and shift Hadoop/HDFS to Compute Engine.
- C. Create a single Cloud Dataproc cluster to support both analytics and data processing, and point it at a Cloud Storage bucket that contains the Parquet files that were previously stored on HDFS.
- D. Create separate Cloud Dataproc clusters to support analytics and data processing, and point both at the same Cloud Storage bucket that contains the Parquet files that were previously stored on HDFS.**

You are building a new real-time data warehouse for your company and will use Google BigQuery streaming inserts. There is no guarantee that data will only be sent in once but you do have a unique ID for each row of data and an event timestamp. You want to ensure that duplicates are not included while interactively querying data. Which query type should you use?

- A. Include ORDER BY DESC on timestamp column and LIMIT to 1.
- B. Use GROUP BY on the unique ID column and timestamp column and SUM on the values.
- C. Use the LAG window function with PARTITION by unique ID along with WHERE LAG IS NOT NULL.
- D. Use the ROW\_NUMBER window function with PARTITION by unique ID along with WHERE row equals 1.

You are building a new real-time data warehouse for your company and will use Google BigQuery streaming inserts. There is no guarantee that data will only be sent in once but you do have a unique ID for each row of data and an event timestamp. You want to ensure that duplicates are not included while interactively querying data. Which query type should you use?

- A. Include ORDER BY DESC on timestamp column and LIMIT to 1.
- B. Use GROUP BY on the unique ID column and timestamp column and SUM on the values.
- C. Use the LAG window function with PARTITION by unique ID along with WHERE LAG IS NOT NULL.
- D. Use the ROW\_NUMBER window function with PARTITION by unique ID along with WHERE row equals 1.**

**You are designing a streaming pipeline for ingesting player interaction data for a mobile game. You want the pipeline to handle out-of-order data delayed up to 15 minutes on a per-player basis and exponential growth in global users. What should you do?**

- A. Design a Cloud Dataflow streaming pipeline with session windowing and a minimum gap duration of 15 minutes. Use "individual player" as the key. Use Cloud Pub/Sub as a message bus for ingestion.
- B. Design a Cloud Dataflow streaming pipeline with session windowing and a minimum gap duration of 15 minutes. Use "individual player" as the key. Use Apache Kafka as a message bus for ingestion.
- C. Design a Cloud Dataflow streaming pipeline with a single global window of 15 minutes. Use Cloud Pub/Sub as a message bus for ingestion.
- D. Design a Cloud Dataflow streaming pipeline with a single global window of 15 minutes. Use Apache Kafka as a message bus for ingestion.

**You are designing a streaming pipeline for ingesting player interaction data for a mobile game. You want the pipeline to handle out-of-order data delayed up to 15 minutes on a per-player basis and exponential growth in global users. What should you do?**

- A. Design a Cloud Dataflow streaming pipeline with session windowing and a minimum gap duration of 15 minutes. Use "individual player" as the key. Use Cloud Pub/Sub as a message bus for ingestion.**
- B. Design a Cloud Dataflow streaming pipeline with session windowing and a minimum gap duration of 15 minutes. Use "individual player" as the key. Use Apache Kafka as a message bus for ingestion.
- C. Design a Cloud Dataflow streaming pipeline with a single global window of 15 minutes. Use Cloud Pub/Sub as a message bus for ingestion.
- D. Design a Cloud Dataflow streaming pipeline with a single global window of 15 minutes. Use Apache Kafka as a message bus for ingestion.

**Your company is loading comma-separated values (CSV) files into Google BigQuery. The data is fully imported successfully; however, the imported data is not matching byte-to-byte to the source file. What is the most likely cause of this problem?**

- A. The CSV data loaded in BigQuery is not flagged as CSV.
- B. The CSV data had invalid rows that were skipped on import.
- C. The CSV data loaded in BigQuery is not using BigQuery's default encoding.
- D. The CSV data has not gone through an ETL phase before loading into BigQuery.

**Your company is loading comma-separated values (CSV) files into Google BigQuery. The data is fully imported successfully; however, the imported data is not matching byte-to-byte to the source file. What is the most likely cause of this problem?**

- A. The CSV data loaded in BigQuery is not flagged as CSV.
- B. The CSV data had invalid rows that were skipped on import.
- C. The CSV data loaded in BigQuery is not using BigQuery's default encoding.**
- D. The CSV data has not gone through an ETL phase before loading into BigQuery.

**Your company is migrating their 30-node Apache Hadoop cluster to the cloud. They want to re-use Hadoop jobs they have already created and minimize the management of the cluster as much as possible. They also want to be able to persist data beyond the life of the cluster. What should you do?**

- A. Create a Google Cloud Dataflow job to process the data.
- B. Create a Google Cloud Dataproc cluster that uses persistent disks for HDFS.
- C. Create a Hadoop cluster on Google Compute Engine that uses persistent disks.
- D. Create a Cloud Dataproc cluster that uses the Google Cloud Storage connector.

**Your company is migrating their 30-node Apache Hadoop cluster to the cloud. They want to re-use Hadoop jobs they have already created and minimize the management of the cluster as much as possible. They also want to be able to persist data beyond the life of the cluster. What should you do?**

- A. Create a Google Cloud Dataflow job to process the data.
- B. Create a Google Cloud Dataproc cluster that uses persistent disks for HDFS.
- C. Create a Hadoop cluster on Google Compute Engine that uses persistent disks.
- D. Create a Cloud Dataproc cluster that uses the Google Cloud Storage connector.**

You work for an economic consulting firm that helps companies identify economic trends as they happen. As part of your analysis, you use Google BigQuery to correlate customer data with the average prices of the 100 most common goods sold, including bread, gasoline, milk, and others. The average prices of these goods are updated every 30 minutes. You want to make sure this data stays up to date so you can combine it with other data in BigQuery as cheaply as possible. What should you do?

- A. Load the data every 30 minutes into a new partitioned table in BigQuery.
- B. Store and update the data in a regional Google Cloud Storage bucket and create a federated data source in BigQuery.
- C. Store the data in Google Cloud Datastore. Use Google Cloud Dataflow to query BigQuery and combine the data programmatically with the data stored in Cloud Datastore.
- D. Store the data in a file in a regional Google Cloud Storage bucket. Use Cloud Dataflow to query BigQuery and combine the data programmatically with the data stored in Google Cloud Storage.

You work for an economic consulting firm that helps companies identify economic trends as they happen. As part of your analysis, you use Google BigQuery to correlate customer data with the average prices of the 100 most common goods sold, including bread, gasoline, milk, and others. The average prices of these goods are updated every 30 minutes. You want to make sure this data stays up to date so you can combine it with other data in BigQuery as cheaply as possible. What should you do?

- A. Load the data every 30 minutes into a new partitioned table in BigQuery.
- B. Store and update the data in a regional Google Cloud Storage bucket and create a federated data source in BigQuery.**
- C. Store the data in Google Cloud Datastore. Use Google Cloud Dataflow to query BigQuery and combine the data programmatically with the data stored in Cloud Datastore.
- D. Store the data in a file in a regional Google Cloud Storage bucket. Use Cloud Dataflow to query BigQuery and combine the data programmatically with the data stored in Google Cloud Storage.

**You have 250,000 devices which produce a JSON device status event every 10 seconds. You want to capture this event data for outlier time series analysis. What should you do?**

- A. Ship the data into BigQuery. Develop a custom application that uses the BigQuery API to query the dataset and displays device outlier data based on your business requirements.
- B. Ship the data into BigQuery. Use the BigQuery console to query the dataset and display device outlier data based on your business requirements.
- C. Ship the data into Cloud Bigtable. Use the Cloud Bigtable cbt tool to display device outlier data based on your business requirements.
- D. Ship the data into Cloud Bigtable. Install and use the HBase shell for Cloud Bigtable to query the table for device outlier data based on your business requirements.

**You have 250,000 devices which produce a JSON device status event every 10 seconds. You want to capture this event data for outlier time series analysis. What should you do?**

- A. Ship the data into BigQuery. Develop a custom application that uses the BigQuery API to query the dataset and displays device outlier data based on your business requirements.
- B. Ship the data into BigQuery. Use the BigQuery console to query the dataset and display device outlier data based on your business requirements.
- C. Ship the data into Cloud Bigtable. Use the Cloud Bigtable cbt tool to display device outlier data based on your business requirements.**
- D. Ship the data into Cloud Bigtable. Install and use the HBase shell for Cloud Bigtable to query the table for device outlier data based on your business requirements.

**You are selecting a messaging service for log messages that must include final result message ordering as part of building a data pipeline on Google Cloud. You want to stream input for 5 days and be able to query the current status. You will be storing the data in a searchable repository. How should you set up the input messages?**

- A. Use Cloud Pub/Sub for input. Attach a timestamp to every message in the publisher.
- B. Use Cloud Pub/Sub for input. Attach a unique identifier to every message in the publisher.
- C. Use Apache Kafka on Compute Engine for input. Attach a timestamp to every message in the publisher.
- D. Use Apache Kafka on Compute Engine for input. Attach a unique identifier to every message in the publisher.

**You are selecting a messaging service for log messages that must include final result message ordering as part of building a data pipeline on Google Cloud. You want to stream input for 5 days and be able to query the current status. You will be storing the data in a searchable repository. How should you set up the input messages?**

- A. Use Cloud Pub/Sub for input. Attach a timestamp to every message in the publisher.
- B. Use Cloud Pub/Sub for input. Attach a unique identifier to every message in the publisher.
- C. Use Apache Kafka on Compute Engine for input. Attach a timestamp to every message in the publisher.
- D. Use Apache Kafka on Compute Engine for input. Attach a unique identifier to every message in the publisher.

You want to publish system metrics to Google Cloud from a large number of on-prem hypervisors and VMs for analysis and creation of dashboards. You have an existing custom monitoring agent deployed to all the hypervisors and your on-prem metrics system is unable to handle the load. You want to design a system that can collect and store metrics at scale. You don't want to manage your own time series database. Metrics from all agents should be written to the same table but agents must not have permission to modify or read data written by other agents. What should you do?

- A. Modify the monitoring agent to publish protobuf messages to Cloud PubSub. Use a Dataproc cluster or Dataflow job to consume messages from Pubsub and write to BigTable.
- B. Modify the monitoring agent to write protobuf messages directly to BigTable.
- C. Modify the monitoring agent to write protobuf messages to HBase deployed on GCE VM Instances
- D. Modify the monitoring agent to write protobuf messages to Cloud Pubsub. Use a Dataproc cluster or Dataflow job to consume messages from Pubsub and write to Cassandra deployed on GCE VM Instances.

You want to publish system metrics to Google Cloud from a large number of on-prem hypervisors and VMs for analysis and creation of dashboards. You have an existing custom monitoring agent deployed to all the hypervisors and your on-prem metrics system is unable to handle the load. You want to design a system that can collect and store metrics at scale. You don't want to manage your own time series database. Metrics from all agents should be written to the same table but agents must not have permission to modify or read data written by other agents. What should you do?

- A. Modify the monitoring agent to publish protobuf messages to Cloud PubSub. Use a Dataproc cluster or Dataflow job to consume messages from Pubsub and write to BigTable.
- B. Modify the monitoring agent to write protobuf messages directly to BigTable.
- C. Modify the monitoring agent to write protobuf messages to HBase deployed on GCE VM Instances
- D. Modify the monitoring agent to write protobuf messages to Cloud Pubsub. Use a Dataproc cluster or Dataflow job to consume messages from Pubsub and write to Cassandra deployed on GCE VM Instances.

**You are designing storage for CSV files and using an I/O-intensive custom Apache Spark transform as part of deploying a data pipeline on Google Cloud. You intend to use ANSI SQL to run queries for your analysts. How should you transform the input data?**

- A. Use BigQuery for storage. Use Cloud Dataflow to run the transformations.
- B. Use BigQuery for storage. Use Cloud Dataproc to run the transformations.
- C. Use Cloud Storage for storage. Use Cloud Dataflow to run the transformations.
- D. Use Cloud Storage for storage. Use Cloud Dataproc to run the transformations.

**You are designing storage for CSV files and using an I/O-intensive custom Apache Spark transform as part of deploying a data pipeline on Google Cloud. You intend to use ANSI SQL to run queries for your analysts. How should you transform the input data?**

- A. Use BigQuery for storage. Use Cloud Dataflow to run the transformations.
- B. Use BigQuery for storage. Use Cloud Dataproc to run the transformations.**
- C. Use Cloud Storage for storage. Use Cloud Dataflow to run the transformations.
- D. Use Cloud Storage for storage. Use Cloud Dataproc to run the transformations.

**You are designing a relational data repository on Google Cloud to grow as needed. The data will be transactionally consistent and added from any location in the world. You want to monitor and adjust node count for input traffic, which can spike unpredictably. What should you do?**

- A. Use Cloud Spanner for storage. Monitor storage usage and increase node count if more than 70% utilized.
- B. Use Cloud Spanner for storage. Monitor CPU utilization and increase node count if more than 70% utilized for your time span.
- C. Use Cloud Bigtable for storage. Monitor data stored and increase node count if more than 70% utilized.
- D. Use Cloud Bigtable for storage. Monitor CPU utilization and increase node count if more than 70% utilized for your time span.

**You are designing a relational data repository on Google Cloud to grow as needed. The data will be transactionally consistent and added from any location in the world. You want to monitor and adjust node count for input traffic, which can spike unpredictably. What should you do?**

- A. Use Cloud Spanner for storage. Monitor storage usage and increase node count if more than 70% utilized.
- B. Use Cloud Spanner for storage. Monitor CPU utilization and increase node count if more than 70% utilized for your time span.**
- C. Use Cloud Bigtable for storage. Monitor data stored and increase node count if more than 70% utilized.
- D. Use Cloud Bigtable for storage. Monitor CPU utilization and increase node count if more than 70% utilized for your time span.

You have a Spark application that writes data to Cloud Storage in Parquet format. You scheduled the application to run daily using DataProcSparkOperator and Apache Airflow DAG (Directed Acyclic Graph) by Cloud Composer. You want to add tasks to the DAG to make the data available to BigQuery users. You want to maximize query speed and configure partitioning and clustering on the table. What should you do?

- A. Use "BashOperator" to call "bq insert".
- B. Use "BashOperator" to call "bq cp" with the "--append" flag.
- C. Use "GoogleCloudStorageToBigQueryOperator" with "schema\_object" pointing to a schema JSON in Cloud Storage and "source\_format" set to "PARQUET".
- D. Use "BigQueryCreateExternalTableOperator" with "schema\_object" pointing to a schema JSON in Cloud Storage and "source\_format" set to "PARQUET".

You have a Spark application that writes data to Cloud Storage in Parquet format. You scheduled the application to run daily using DataProcSparkOperator and Apache Airflow DAG by Cloud Composer. You want to add tasks to the DAG to make the data available to BigQuery users. You want to maximize query speed and configure partitioning and clustering on the table. What should you do?

- A. Use "BashOperator" to call "bq insert".
- B. Use "BashOperator" to call "bq cp" with the "--append" flag.
- C. Use "GoogleCloudStorageToBigQueryOperator" with "schema\_object" pointing to a schema JSON in Cloud Storage and "source\_format" set to "PARQUET".
- D. Use "BigQueryCreateExternalTableOperator" with "schema\_object" pointing to a schema JSON in Cloud Storage and "source\_format" set to "PARQUET".

You have a website that tracks page visits for each user and then creates a Cloud Pub/Sub message with the session ID and URL of the page. You want to create a Cloud Dataflow pipeline that sums the total number of pages visited by each user and writes the result to BigQuery. User sessions timeout after 30 minutes. Which type of Cloud Dataflow window should you choose?

- A. A single global window
- B. Fixed-time windows with a duration of 30 minutes
- C. Session-based windows with a gap duration of 30 minutes
- D. Sliding-time windows with a duration of 30 minutes and a new window every 5 minute

You have a website that tracks page visits for each user and then creates a Cloud Pub/Sub message with the session ID and URL of the page. You want to create a Cloud Dataflow pipeline that sums the total number of pages visited by each user and writes the result to BigQuery. User sessions timeout after 30 minutes. Which type of Cloud Dataflow window should you choose?

- A. A single global window
- B. Fixed-time windows with a duration of 30 minutes
- C. Session-based windows with a gap duration of 30 minutes**
- D. Sliding-time windows with a duration of 30 minutes and a new window every 5 minute

**You are designing a basket abandonment system for an ecommerce company. The system will send a message to a user based on these rules: a). No interaction by the user on the site for 1 hour b). Has added more than \$30 worth of products to the basket c). Has not completed a transaction. You use Google Cloud Dataflow to process the data and decide if a message should be sent. How should you design the pipeline?**

- A. Use a fixed-time window with a duration of 60 minutes.
- B. Use a sliding time window with a duration of 60 minutes.
- C. Use a session window with a gap time duration of 60 minutes.
- D. Use a global window with a time based trigger with a delay of 60 minutes.

**You are designing a basket abandonment system for an ecommerce company. The system will send a message to a user based on these rules: a). No interaction by the user on the site for 1 hour b). Has added more than \$30 worth of products to the basket c). Has not completed a transaction. You use Google Cloud Dataflow to process the data and decide if a message should be sent. How should you design the pipeline?**

- A. Use a fixed-time window with a duration of 60 minutes.
- B. Use a sliding time window with a duration of 60 minutes.
- C. Use a session window with a gap time duration of 60 minutes.**
- D. Use a global window with a time based trigger with a delay of 60 minutes.

**You need to stream time-series data in Avro format, and then write this to both BigQuery and Cloud Bigtable simultaneously using Cloud Dataflow. You want to achieve minimal end-to-end latency. Your business requirements state this needs to be completed as quickly as possible. What should you do?**

- A. Create a pipeline and use ParDo transform.
- B. Create a pipeline that groups the data into a PCollection and uses the Combine transform.
- C. Create a pipeline that groups data using a PCollection and then uses Cloud Bigtable and BigQueryIO transforms.
- D. Create a pipeline that groups data using a PCollection, and then use Avro I/O transform to write to Cloud Storage. After the data is written, load the data from Cloud Storage into BigQuery and Cloud Bigtable.

**You need to stream time-series data in Avro format, and then write this to both BigQuery and Cloud Bigtable simultaneously using Cloud Dataflow. You want to achieve minimal end-to-end latency. Your business requirements state this needs to be completed as quickly as possible. What should you do?**

- A. Create a pipeline and use ParDo transform.
- B. Create a pipeline that groups the data into a PCollection and uses the Combine transform.
- C. **Create a pipeline that groups data using a PCollection and then uses Cloud Bigtable and BigQueryIO transforms.**
- D. Create a pipeline that groups data using a PCollection, and then use Avro I/O transform to write to Cloud Storage. After the data is written, load the data from Cloud Storage into BigQuery and Cloud Bigtable.

**Your company's on-premises Apache Hadoop servers are approaching end-of-life, and IT has decided to migrate the cluster to Google Cloud Dataproc. A like-for-like migration of the cluster would require 50 TB of Google Persistent Disk per node. The CIO is concerned about the cost of using that much block storage. You want to minimize the storage cost of the migration. What should you do?**

- A. Put the data into Google Cloud Storage.
- B. Use preemptible virtual machines (VMs) for the Cloud Dataproc cluster.
- C. Tune the Cloud Dataproc cluster so that there is just enough disk for all data.
- D. Migrate some of the cold data into Google Cloud Storage, and keep only the hot data in Persistent Disk.

Your company's on-premises Apache Hadoop servers are approaching end-of-life, and IT has decided to migrate the cluster to Google Cloud Dataproc. A like-for-like migration of the cluster would require 50 TB of Google Persistent Disk per node. The CIO is concerned about the cost of using that much block storage. You want to minimize the storage cost of the migration. What should you do?

- A. Put the data into Google Cloud Storage.
- B. Use preemptible virtual machines (VMs) for the Cloud Dataproc cluster.
- C. Tune the Cloud Dataproc cluster so that there is just enough disk for all data.
- D. Migrate some of the cold data into Google Cloud Storage, and keep only the hot data in Persistent Disk.

**You are designing storage for two relational tables that are part of a 10-TB database on Google Cloud. You want to support transactions that scale horizontally. You also want to optimize data for range queries on non-key columns. What should you do?**

- A. Use Cloud SQL for storage. Add secondary indexes to support query patterns.
- B. Use Cloud SQL for storage. Use Cloud Dataflow to transform data to support query patterns.
- C. Use Cloud Spanner for storage. Add secondary indexes to support query patterns.
- D. Use Cloud Spanner for storage. Use Cloud Dataflow to transform data to support query patterns.

**You are designing storage for two relational tables that are part of a 10-TB database on Google Cloud. You want to support transactions that scale horizontally. You also want to optimize data for range queries on non-key columns. What should you do?**

- A. Use Cloud SQL for storage. Add secondary indexes to support query patterns.
- B. Use Cloud SQL for storage. Use Cloud Dataflow to transform data to support query patterns.
- C. Use Cloud Spanner for storage. Add secondary indexes to support query patterns.**
- D. Use Cloud Spanner for storage. Use Cloud Dataflow to transform data to support query patterns.

**Your company is streaming real-time sensor data from their factory floor into Bigtable and they have noticed extremely poor performance. How should the row key be redesigned to improve Bigtable performance on queries that populate real-time dashboards?**

- A. Use a row key of the form <timestamp>.
- B. Use a row key of the form <sensorid>.
- C. Use a row key of the form <timestamp>#<sensorid>.
- D. Use a row key of the form <sensorid>#<timestamp>.

**Your company is streaming real-time sensor data from their factory floor into Bigtable and they have noticed extremely poor performance. How should the row key be redesigned to improve Bigtable performance on queries that populate real-time dashboards?**

- A. Use a row key of the form <timestamp>.
- B. Use a row key of the form <sensorid>.
- C. Use a row key of the form <timestamp>#<sensorid>.
- D. Use a row key of the form <sensorid>#<timestamp>.**

**You are developing an application on Google Cloud that will automatically generate subject labels for users' blog posts. You are under competitive pressure to add this feature quickly, and you have no additional developer resources. No one on your team has experience with machine learning. What should you do?**

- A. Call the Cloud Natural Language API from your application. Process the generated Entity Analysis as labels.
- B. Call the Cloud Natural Language API from your application. Process the generated Sentiment Analysis as labels.
- C. Build and train a text classification model using TensorFlow. Deploy the model using Cloud Machine Learning Engine. Call the model from your application and process the results as labels.
- D. Build and train a text classification model using TensorFlow. Deploy the model using a Kubernetes Engine cluster. Call the model from your application and process the results as labels.

**You are developing an application on Google Cloud that will automatically generate subject labels for users' blog posts. You are under competitive pressure to add this feature quickly, and you have no additional developer resources. No one on your team has experience with machine learning. What should you do?**

- A. Call the Cloud Natural Language API from your application. Process the generated Entity Analysis as labels.
- B. Call the Cloud Natural Language API from your application. Process the generated Sentiment Analysis as labels.
- C. Build and train a text classification model using TensorFlow. Deploy the model using Cloud Machine Learning Engine. Call the model from your application and process the results as labels.
- D. Build and train a text classification model using TensorFlow. Deploy the model using a Kubernetes Engine cluster. Call the model from your application and process the results as labels.

Your company is using WILDCARD tables to query data across multiple tables with similar names. The SQL statement is currently failing with the error shown below. Which table name will make the SQL statement work correctly?

```
# Syntax error: Expected end of statement but got "--" at [4:11]
SELECT age
FROM
    bigquery-public-data.noaa_gsod.gsod
WHERE
    age != 99
    AND _TABLE_SUFFIX = '1929'
ORDER BY
    age DESC
```

- A. `bigquery-public-data.noaa\_gsod.gsod`
- B. bigquery-public-data.noaa\_gsod.gsod\*
- C. 'bigquery-public-data.noaa\_gsod.gsod\*'
- D. `bigquery-public-data.noaa\_gsod.gsod\*`

Your company is using WILDCARD tables to query data across multiple tables with similar names. The SQL statement is currently failing with the error shown below. Which table name will make the SQL statement work correctly?

```
# Syntax error: Expected end of statement but got "--" at [4:11]
SELECT age
FROM
    bigquery-public-data.noaa_gsod.gsod
WHERE
    age != 99
    AND _TABLE_SUFFIX = '1929'
ORDER BY
    age DESC
```

- A. `bigquery-public-data.noaa\_gsod.gsod`
- B. bigquery-public-data.noaa\_gsod.gsod\*
- C. 'bigquery-public-data.noaa\_gsod.gsod\*'
- D. **bigquery-public-data.noaa\_gsod.gsod\***

You are working on an ML-based application that will transcribe conversations between manufacturing workers. These conversations are in English and between 30-40 sec long. Conversation recordings come from old enterprise radio sets that have a low sampling rate of 8000 Hz, but you have a large dataset of these recorded conversations with their transcriptions. You want to follow Google-recommended practices. How should you proceed with building your application?

- A. Use Cloud Speech-to-Text API, and send requests in a synchronous mode.
- B. Use Cloud Speech-to-Text API, and send requests in an asynchronous mode.
- C. Use Cloud Speech-to-Text API, but resample your captured recordings to a rate of 16000 Hz.
- D. Train your own speech recognition model because you have an uncommon use case and you have a labeled dataset.

You are working on an ML-based application that will transcribe conversations between manufacturing workers. These conversations are in English and between 30-40 sec long. Conversation recordings come from old enterprise radio sets that have a low sampling rate of 8000 Hz, but you have a large dataset of these recorded conversations with their transcriptions. You want to follow Google-recommended practices. How should you proceed with building your application?

- A. Use Cloud Speech-to-Text API, and send requests in a synchronous mode.
- B. Use Cloud Speech-to-Text API, and send requests in an asynchronous mode.
- C. Use Cloud Speech-to-Text API, but resample your captured recordings to a rate of 16000 Hz.
- D. Train your own speech recognition model because you have an uncommon use case and you have a labeled dataset.

**You are developing an application on Google Cloud that will label famous landmarks in users' photos. You are under competitive pressure to develop a predictive model quickly. You need to keep service costs low. What should you do?**

- A. Build an application that calls the Cloud Vision API. Inspect the generated MID values to supply the image labels.
- B. Build an application that calls the Cloud Vision API. Pass client image locations as base64-encoded strings.
- C. Build and train a classification model with TensorFlow. Deploy the model using Cloud Machine Learning Engine. Pass client image locations as base64-encoded strings.
- D. Build and train a classification model with TensorFlow. Deploy the model using Cloud Machine Learning Engine. Inspect the generated MID values to supply the image labels.

**You are developing an application on Google Cloud that will label famous landmarks in users' photos. You are under competitive pressure to develop a predictive model quickly. You need to keep service costs low. What should you do?**

- A. Build an application that calls the Cloud Vision API. Inspect the generated MID values to supply the image labels.
- B. Build an application that calls the Cloud Vision API. Pass client image locations as base64-encoded strings.**
- C. Build and train a classification model with TensorFlow. Deploy the model using Cloud Machine Learning Engine. Pass client image locations as base64-encoded strings.
- D. Build and train a classification model with TensorFlow. Deploy the model using Cloud Machine Learning Engine. Inspect the generated MID values to supply the image labels.

**You are building a data pipeline on Google Cloud. You need to select services that will host a deep neural network machine-learning model also hosted on Google Cloud. You also need to monitor and run jobs that could occasionally fail. What should you do?**

- A. Use Cloud Machine Learning to host your model. Monitor the status of the Operation object for 'error' results.
- B. Use Cloud Machine Learning to host your model. Monitor the status of the Jobs object for 'failed' job states.
- C. Use a Kubernetes Engine cluster to host your model. Monitor the status of the Jobs object for 'failed' job states.
- D. Use a Kubernetes Engine cluster to host your model. Monitor the status of Operation object for 'error' results.

**You are building a data pipeline on Google Cloud. You need to select services that will host a deep neural network machine-learning model also hosted on Google Cloud. You also need to monitor and run jobs that could occasionally fail. What should you do?**

- A. Use Cloud Machine Learning to host your model. Monitor the status of the Operation object for 'error' results.
- B. Use Cloud Machine Learning to host your model. Monitor the status of the Jobs object for 'failed' job states.**
- C. Use a Kubernetes Engine cluster to host your model. Monitor the status of the Jobs object for 'failed' job states.
- D. Use a Kubernetes Engine cluster to host your model. Monitor the status of Operation object for 'error' results.

**You work on a regression problem in a natural language processing domain, and you have 100M labeled examples in your dataset. You have randomly shuffled your data and split your dataset into training and test samples (in a 90/10 ratio). After you have trained the neural network and evaluated your model on a test set, you discover that the root-mean-squared error (RMSE) of your model is twice as high on the train set as on the test set. How should you improve the performance of your model?**

- A. Increase the share of the test sample in the train-test split.
- B. Try to collect more data and increase the size of your dataset.
- C. Try out regularization techniques (e.g., dropout or batch normalization) to avoid overfitting.
- D. Increase the complexity of your model by, e.g., introducing an additional layer or increasing the size of vocabularies or n-grams used to avoid underfitting.

You work on a regression problem in a natural language processing domain, and you have 100M labeled examples in your dataset. You have randomly shuffled your data and split your dataset into training and test samples (in a 90/10 ratio). After you have trained the neural network and evaluated your model on a test set, you discover that the root-mean-squared error (RMSE) of your model is twice as high on the train set as on the test set. How should you improve the performance of your model?

- A. Increase the share of the test sample in the train-test split.
- B. Try to collect more data and increase the size of your dataset.
- C. Try out regularization techniques (e.g., dropout or batch normalization) to avoid overfitting.
- D. Increase the complexity of your model by, e.g., introducing an additional layer or increasing the size of vocabularies or n-grams used to avoid underfitting.

You are using Cloud Pub/Sub to stream inventory updates from many point-of-sale (POS) terminals into BigQuery. Each update event has the following information: product identifier "prodSku", change increment "quantityDelta", POS identification "termId", and "messageId" which is created for each push attempt from the terminal. During a network outage, you discovered that duplicated messages were sent, causing the inventory system to over-count the changes. You determine that the terminal application has design problems and may send the same event more than once during push retries. You want to ensure that the inventory update is accurate. What should you do?

- A. Inspect the "publishTime" of each message. Make sure that messages whose "publishTime" values match rows in the BigQuery table are discarded.
- B. Inspect the "messageId" of each message. Make sure that any messages whose "messageId" values match corresponding rows in the BigQuery table are discarded.
- C. Instead of specifying a change increment for "quantityDelta", always use the derived inventory value after the increment has been applied. Name the new attribute "adjustedQuantity".
- D. Add another attribute orderId to the message payload to mark the unique check-out order across all terminals. Make sure that messages whose "orderId" and "prodSku" values match corresponding rows in the BigQuery table are discarded.

You are using Cloud Pub/Sub to stream inventory updates from many point-of-sale (POS) terminals into BigQuery. Each update event has the following information: product identifier "prodSku", change increment "quantityDelta", POS identification "termId", and "messageId" which is created for each push attempt from the terminal. During a network outage, you discovered that duplicated messages were sent, causing the inventory system to over-count the changes. You determine that the terminal application has design problems and may send the same event more than once during push retries. You want to ensure that the inventory update is accurate. What should you do?

- A. Inspect the "publishTime" of each message. Make sure that messages whose "publishTime" values match rows in the BigQuery table are discarded.
- B. Inspect the "messageId" of each message. Make sure that any messages whose "messageId" values match corresponding rows in the BigQuery table are discarded.
- C. Instead of specifying a change increment for "quantityDelta", always use the derived inventory value after the increment has been applied. Name the new attribute "adjustedQuantity".
- D. Add another attribute **orderId** to the message payload to mark the unique check-out order across all terminals. Make sure that messages whose "orderId" and "prodSku" values match corresponding rows in the BigQuery table are discarded.

You designed a database for patient records as a pilot project to cover a few hundred patients in three clinics. Your design used a single database table to represent all patients and their visits, and you used self-joins to generate reports. The server resource utilization was at 50%. Since then, the scope of the project has expanded. The database table must now store 100 times more patient records. You can no longer run the reports, because they either take too long or they encounter errors with insufficient compute resources. How should you adjust the database design?

- A. Add capacity (memory and disk space) to the database server by the order of 200.
- B. Shard the tables into smaller ones based on date ranges, and only generate reports with pre-specified date ranges.
- C. Normalize the master patient-record table into the patients table and the visits table, and create other necessary tables to avoid self-join.
- D. Partition the table into smaller tables, with one for each clinic. Run queries against the smaller table pairs, and use unions for consolidated reports.

You designed a database for patient records as a pilot project to cover a few hundred patients in three clinics. Your design used a single database table to represent all patients and their visits, and you used self-joins to generate reports. The server resource utilization was at 50%. Since then, the scope of the project has expanded. The database table must now store 100 times more patient records. You can no longer run the reports, because they either take too long or they encounter errors with insufficient compute resources. How should you adjust the database design?

- A. Add capacity (memory and disk space) to the database server by the order of 200.
- B. Shard the tables into smaller ones based on date ranges, and only generate reports with pre-specified date ranges.
- C. Normalize the master patient-record table into the patients table and the visits table, and create other necessary tables to avoid self-join.**
- D. Partition the table into smaller tables, with one for each clinic. Run queries against the smaller table pairs, and use unions for consolidated reports.

**Your startup has never implemented a formal security policy. Currently, everyone in the company has access to the datasets stored in Google BigQuery. Teams have the freedom to use the service as they see fit, and they have not documented their use cases. You have been asked to secure the data warehouse. You need to discover what everyone is doing. What should you do first?**

- A. Use Google Stackdriver Audit Logs to review data access.
- B. Get the identity and access management (IAM) policy of each table.
- C. Use Stackdriver Monitoring to see the usage of BigQuery query slots.
- D. Use the Google Cloud Billing API to see what account the warehouse is being billed to.

**Your startup has never implemented a formal security policy. Currently, everyone in the company has access to the datasets stored in Google BigQuery. Teams have the freedom to use the service as they see fit, and they have not documented their use cases. You have been asked to secure the data warehouse. You need to discover what everyone is doing. What should you do first?**

- A. Use Google Stackdriver Audit Logs to review data access.
- B. Get the identity and access management (IAM) policy of each table.
- C. Use Stackdriver Monitoring to see the usage of BigQuery query slots.
- D. Use the Google Cloud Billing API to see what account the warehouse is being billed to.

**You created a job which runs daily to import highly sensitive data from an on-premises location to Cloud Storage. You also set up a streaming data insert into Cloud Storage via a Kafka node that is running on a Compute Engine instance. You need to encrypt the data at rest and supply your own encryption key. Your key should not be stored in the Google Cloud. What should you do?**

- A. Create a dedicated service account, and use encryption at rest to reference your data stored in Cloud Storage and Compute Engine data as part of your API service calls.
- B. Upload your own encryption key to Cloud Key Management Service, and use it to encrypt your data in Cloud Storage. Use your uploaded encryption key and reference it as part of your API service calls to encrypt your data in the Kafka node hosted on Compute Engine.
- C. Upload your own encryption key to Cloud Key Management Service, and use it to encrypt your data in your Kafka node hosted on Compute Engine.
- D. Supply your own encryption key, and reference it as part of your API service calls to encrypt your data in Cloud Storage and your Kafka node hosted on Compute Engine.

**You created a job which runs daily to import highly sensitive data from an on-premises location to Cloud Storage. You also set up a streaming data insert into Cloud Storage via a Kafka node that is running on a Compute Engine instance. You need to encrypt the data at rest and supply your own encryption key. Your key should not be stored in the Google Cloud. What should you do?**

- A. Create a dedicated service account, and use encryption at rest to reference your data stored in Cloud Storage and Compute Engine data as part of your API service calls.
- B. Upload your own encryption key to Cloud Key Management Service, and use it to encrypt your data in Cloud Storage. Use your uploaded encryption key and reference it as part of your API service calls to encrypt your data in the Kafka node hosted on Compute Engine.
- C. Upload your own encryption key to Cloud Key Management Service, and use it to encrypt your data in your Kafka node hosted on Compute Engine.
- D. Supply your own encryption key, and reference it as part of your API service calls to encrypt your data in Cloud Storage and your Kafka node hosted on Compute Engine.**

You are working on a project with two compliance requirements. The first requirement states that your developers should be able to see the Google Cloud Platform billing charges for only their own projects. The second requirement states that your finance team members can set budgets and view the current charges for all projects in the organization. The finance team should not be able to view the project contents. You want to set permissions. What should you do?

- A. Add the finance team members to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only.
- B. Add the finance team members to the Billing Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project.
- C. Add the developers and finance managers to the Viewer role for the Project.
- D. Add the finance team to the Viewer role for the Project. Add the developers to the Security Reviewer role for each of the billing accounts.

You are working on a project with two compliance requirements. The first requirement states that your developers should be able to see the Google Cloud Platform billing charges for only their own projects. The second requirement states that your finance team members can set budgets and view the current charges for all projects in the organization. The finance team should not be able to view the project contents. You want to set permissions. What should you do?

- A. Add the finance team members to the default IAM Owner role. Add the developers to a custom role that allows them to see their own spend only.
- B. Add the finance team members to the Billing Administrator role for each of the billing accounts that they need to manage. Add the developers to the Viewer role for the Project.**
- C. Add the developers and finance managers to the Viewer role for the Project.
- D. Add the finance team to the Viewer role for the Project. Add the developers to the Security Reviewer role for each of the billing accounts.

# Earn your exam voucher

Partners must meet the following requirements in order to earn a FREE exam voucher:

- 1) attend this online bootcamp live on 22 November 2019
- 2) complete the hands-on labs Quests [Data Engineer](#) & [GCP Essentials](#)

**by 20 December 2019**

Instructions will be shared in the follow up emails.

Voucher will expire on **17 January 2020**



# Getting started with Qwiklabs

## Step 1

Complete the Qwiklabs credit request form,  
<https://qwiklabscredits.appspot.com/>

Qwiklabs credits registration

I have read and agree to the [Qwiklabs credits terms and conditions](#).

## Step 2

Upon submission, an email will be sent to you to confirm the request for Qwiklabs credits. The email will include an access link.

## Step 3

Upon selecting the link in the email, you will be asked to provide additional information to complete the request process.

## Step 4

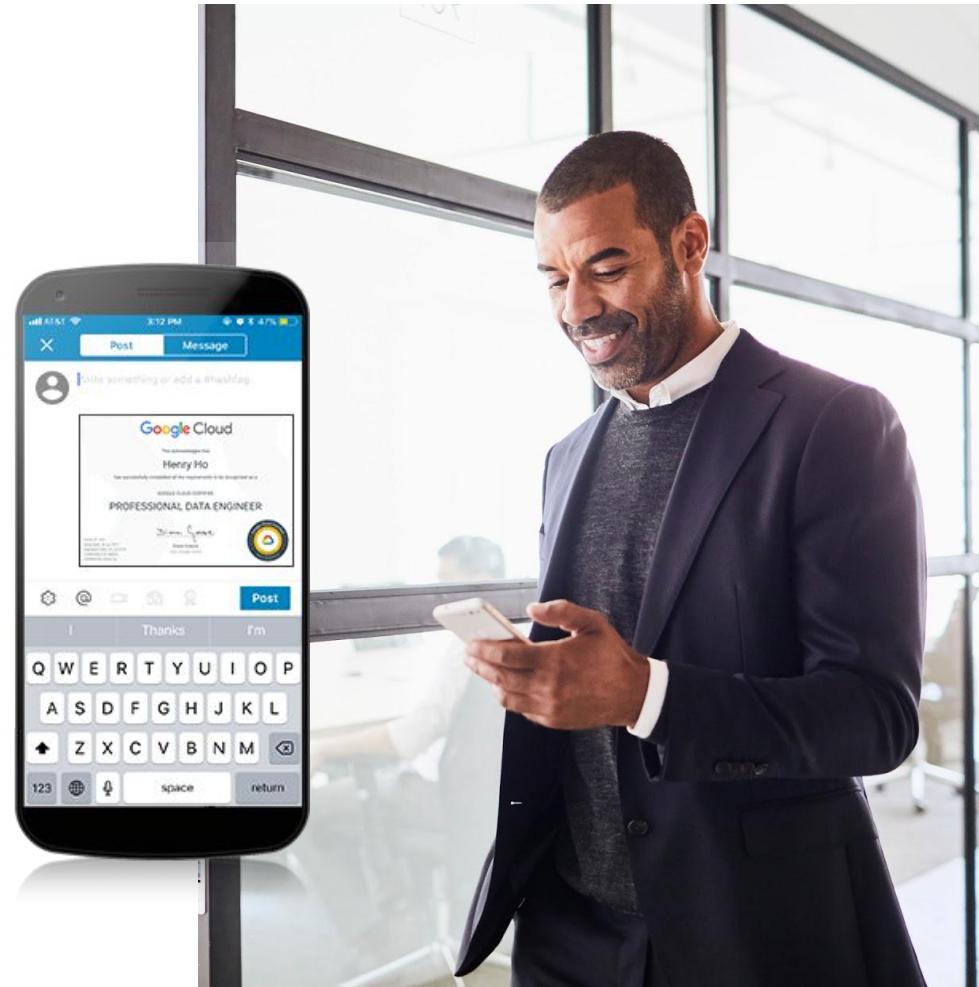
Upon submission, a final email will be sent which includes the steps to create your Qwiklabs account. You can also create your account by logging into Qwiklabs, via [google.qwiklabs.com](http://google.qwiklabs.com)



For questions and support, please reach out to [partner-training@google.com](mailto:partner-training@google.com)

# Track and share your progress socially

Follow us on  
[@Gcloudpartners](https://twitter.com/Gcloudpartners)  
Tag us with your  
[#googlecloudcertified](https://twitter.com/hashtag/googlecloudcertified)  
progress





# Questions?



# Thank You!

**Share your feedback with us!  
Keep an eye out for a feedback request form.**