



**Get to Know the Google Cloud  
Certification Paths:  
Professional  
Data Engineer**

# Data Engineer

Level: Professional



The Professional Data Engineer knows how to build scalable, reliable data pipelines and applications. Anyone **intending on taking** this exam should also be **familiar with** selecting, monitoring, and troubleshooting machine learning models.

3+ years industry and 1+ years GCP experience recommended.  
Exam length: Two hours | Available in: English, Japanese



## Who Is It For?

While the professional-level certifications generally correlate to **job title**, be aware of other roles that may benefit from this certification, such as:



Data Engineer



Cloud Architect



Data Scientist

...and many more!  
Data Warehouse Engineer  
Big Data Engineer  
Data Science Engineer  
Software Engineer, Data  
Data Integration Engineer  
Analytics Engineer  
Database Architect  
ML Engineer  
BI Developer

## Market Value

The Google Cloud Professional Data Engineer placed as the



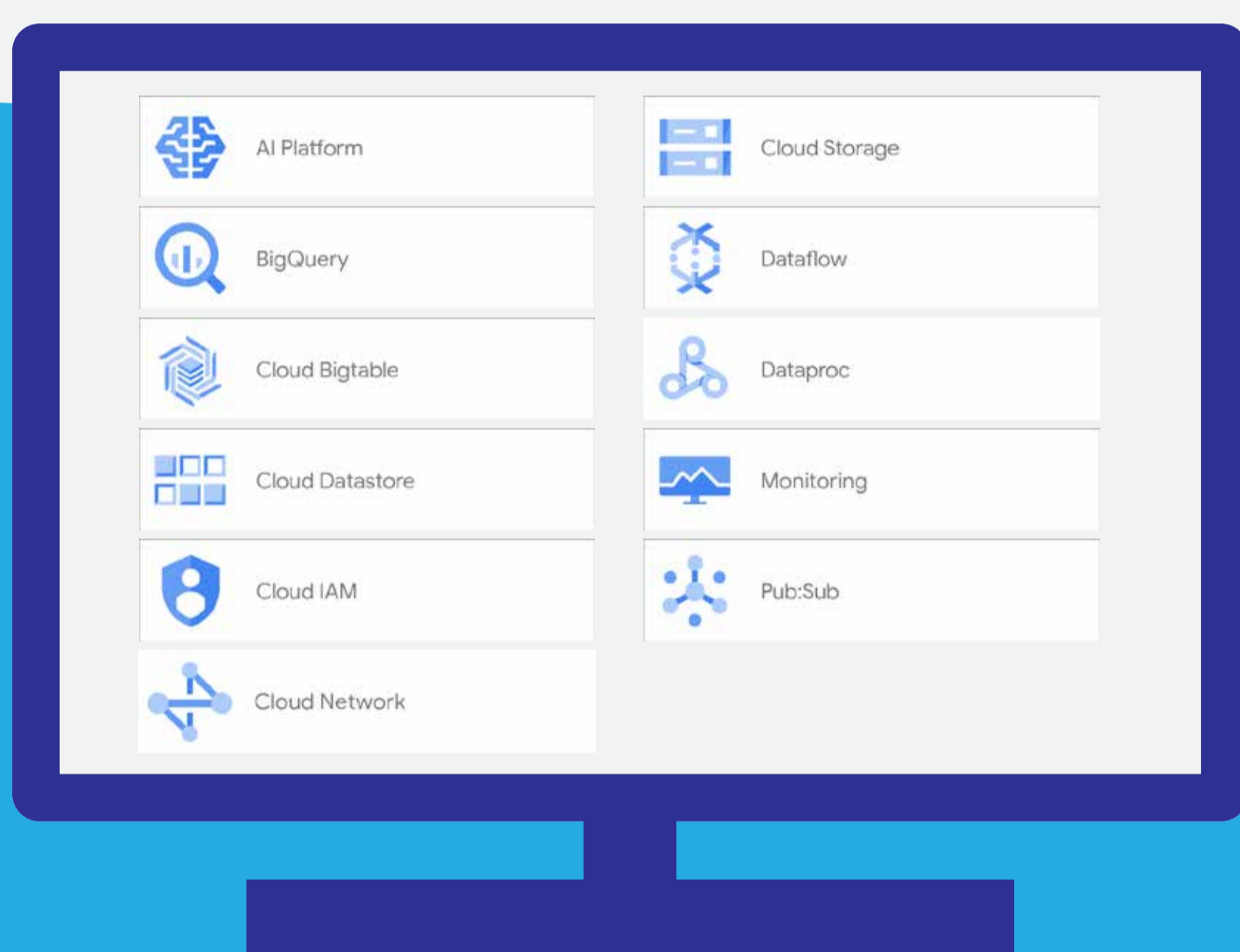
In North America. Source: Global Knowledge

Data Engineer is the fastest-growing tech occupation.

Data Engineer	50%
Back End Developer	38%
Senior Data Scientist	32%
CRM Developer	29%
UI Developer	24%

Year-over-year growth, United States. Source: DICE

## Exam Topics



### Google Cloud Products and Services

Exam topics cover products under AI and ML, data analytics, databases, storage, operations, networking, and related APIs.

### Data Processing Systems



#### Designing

Meeting business requirements  
Data modeling  
Schema design

Publishing and visualization  
Batch and streaming data  
Automation and orchestration

Architecture options  
System availability  
Hybrid cloud and edge computing

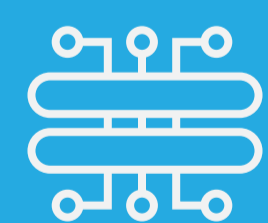
#### Storage Systems



#### Pipelines



#### Infrastructure



#### Building



Using managed services  
Cost and performance  
Data life cycle management

Data cleansing  
Acquisition and import  
Integrating with new data sources

Provisioning resources  
Monitoring and adjusting pipelines  
Testing and quality control

### Operationalizing Machine Learning Models



#### Pre-built ML Models

ML APIs, customization, and conversational experiences



#### Deploying ML Pipelines

Data ingestion, retraining, and continuous evaluation



#### Training and Serving

Distributed vs. single machine, edge compute, and accelerators



#### Monitoring and Troubleshooting

Terminology, errors, and dependencies

### Ensuring Solution Quality



#### Security and Compliance

Identity and access management, encryption, privacy, data loss prevention, and legal compliance

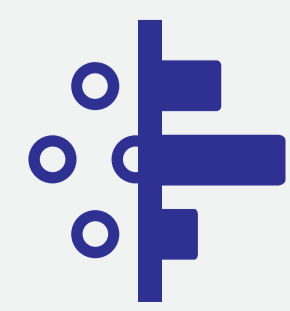


#### Scalability and Efficiency

Building test suites, monitoring, improving, resizing, and autoscaling data resources

#### Reliability and Fidelity

Data preparation, quality control, verification, and data recovery



#### Flexibility and Portability

Multicloud, data residency requirements, staging, cataloging and discovery

