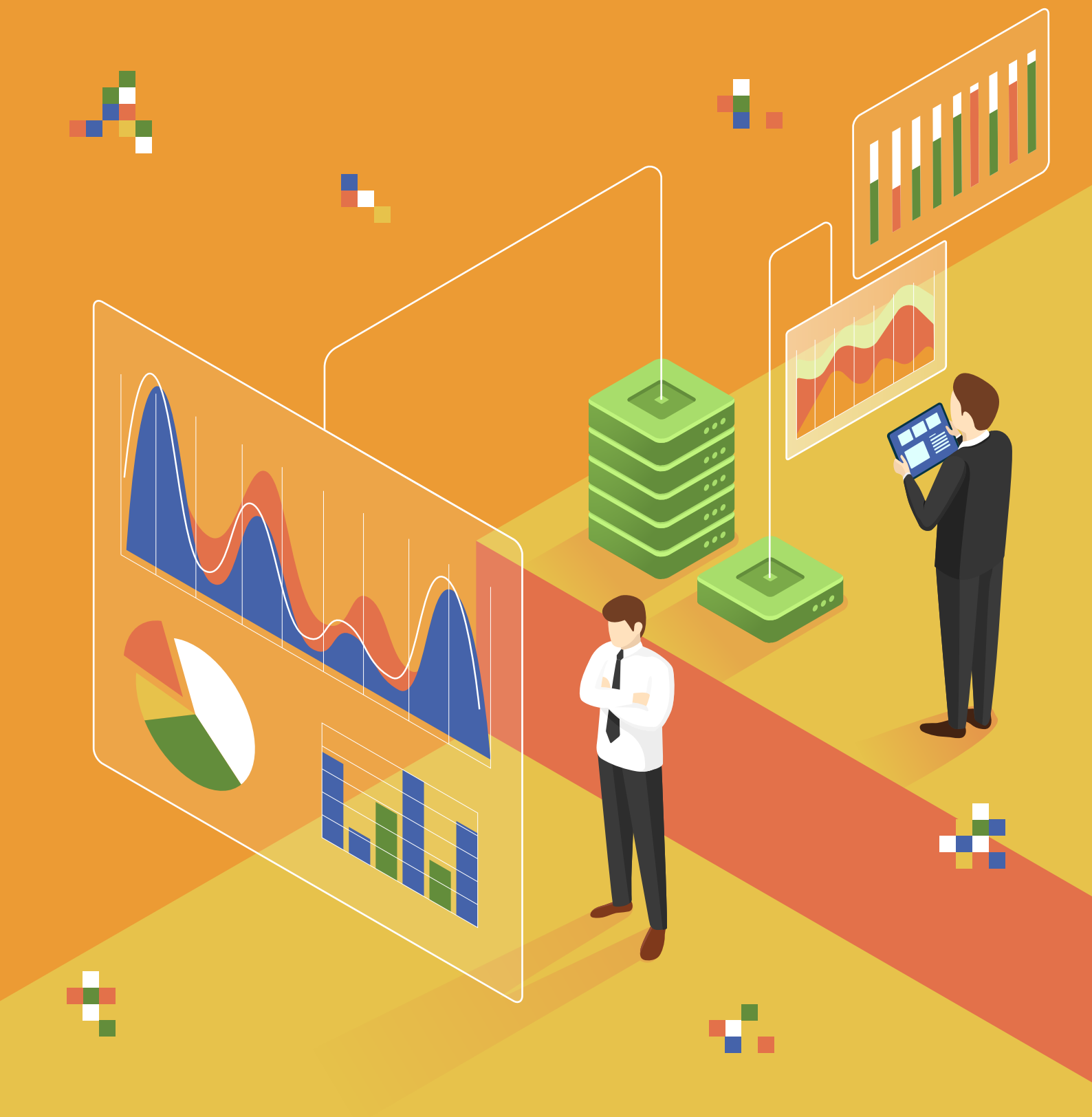
SuperDataScience

AWS Machine Learning
Practical Exams Reward



1. **John has stored most of his work-related documents in his S3 bucket in CSV format. He wants to run analytics on the data and also visualize them. Which are the AWS service that would help you achieve this?**

   A AWS Glue, AWS Athena and QuickSight
   B AWS Sagemaker, AWS RedShift and QuickSight
   C AWS Glue, AWS Polly and AWS EMR
   D AWS Sagemaker, AWS Athena and AWS EMR

2. **You are working on a large S3 data lake containing both structured and unstructured data. You want to transform the data and wish to analyse it using SQL queries. Which AWS service would help with this?**

   A Use AWS DataStream to organize the unstructured data and use AWS EMR to do the query
   B Use AWS DataStream to organize the unstructured data and use AWS Redshift to do the query
   C Use AWS Glue to organize the unstructured data and use AWS Athena to do the query
   D Use AWS Glue to organize the unstructured data and use AWS ColudSearch to do the query

3. **You want to analyze unstructured data stream logs using Hive on EMR, then add this data to the structured data from RDS and finally upload it to Redshift . Which AWS tool helps you to run this particular process on a pre-defined schedule?**

   A AWS Step Function
   B AWS Athena
   C AWS Data Pipeline
   D AWS Quicksight

4. **You want to migrate data from an on-premises database to the AWS cloud and you want to reduce the time taken to transfer the data. Which AWS service would help with this?**

   A AWS DMS
   B AWS Glue
   C AWS Athena
   D AWS RDS

5. **Which of the function from AWS Glue can be used to populate the AWS Glue Data Catalog with tables?**

   A ETL
   B Extractor
   C Migrator
   D Crawler

6. **Jack is working on a project which has access to data stream from DynamoDB, but he wants to convert the file format from JSON to CSV file, so that he can store it in S3 and use it in his Sagemaker Model. Which AWS Service helps him achieve this in a effective way?**

   A AWS DMS to convert from JSON to csv and store it in S3
   B AWS Athena to convert from JSON to csv and store it in S3
   C AWS Glue to convert from JSON to csv and store it in S3
   D AWS RDS to convert from JSON to csv and store it in S3

6. **You are given access to data and you are tasked to build a AWS Glue Catalog. But on running the pre-defined crawler, it was not able to produce any tables. Why is it happening and how can it be solved?**

   A AWS Glue build-in classifiers were not able to find the schema of the data and you can use custom created classifier

   B Data from the source is missing and add more data

   C Data that is supplied to the AWS is corrupted. So, you have to remove the corrupted data

   D Use AWS Athena to create queries and table

8. **You are tasked to move data from DynamoDb to S3 bucket. So that it can be used later for any analysis or processing. How can you achieve this?**

   A Use AWS Quickshift

   B Use Datapipeline

   C Use Redshiftspectrum

   D Use AWS Athena

9. **You are well-versed in python and been using python script to perform data transformation on the local machine and would like to use this in AWS to perform ETL and also would like to schedule the transformation jobs. Which is the most efficient way to go about it?**

   A Use AWS Glue job which uses python shell as the job and also automate the transformation using AWS Glue

   B Use AWS Glue job which uses spark shell as the job and also automate the transformation using AWS Glue

   C Use EMR cluster which uses python shell as the job and also automate the transformation using EMR

   D Use EMR cluster which uses spark shell as the job and also automate the transformation using EMR

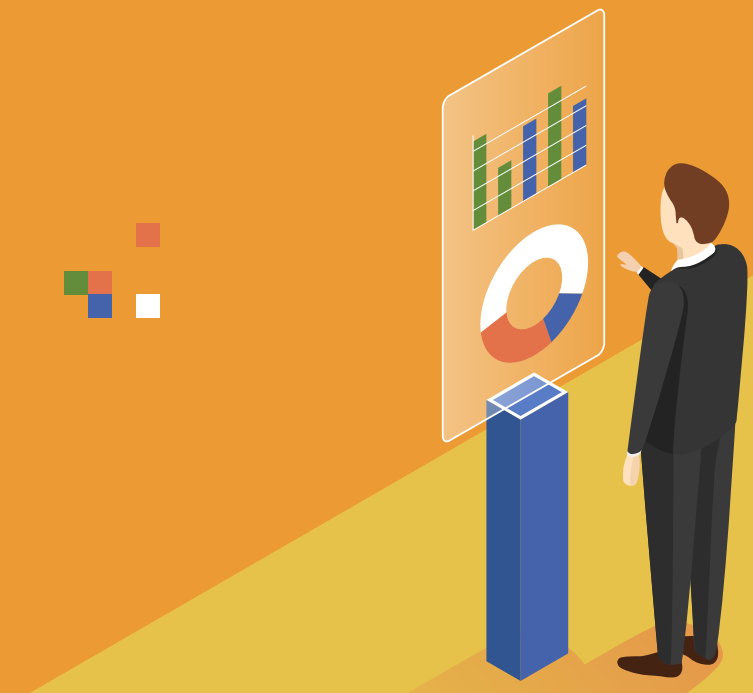
10. What is the simplest AWS tool that can make sure that long-running ETL jobs are executed successfully instead of orchestrating manually and can also help in standardising ML workflow?

A AWS Batch

B AWS RDS

C AWS Step Function

D AWS Athena



ANSWER KEY:

1. A	2. C	3. C	4. A	5. D
6. C	7. A	8. B	9. A	10. C