DE Zoomcamp 2023

Week 6 Homework

1. Please select the statements that are correct

\*\*Answer:\*\* Kafka Node is responsible to store topics;

Zookeeper is removed form Kafka cluster starting from version 4.0;

Retention configuration ensures the messages not get lost over specific period of time;

Group-Id ensures the messages are distributed to associated consumers

\*\*Code:\*\*

```

N/A

```

2. Please select the Kafka concepts that support reliability and availability

\*\*Answer:\*\* Topic Replication;

Topic Paritioning;

Consumer Group Id;

Ack All

\*\*Code:\*\* N/A

3. Please select the Kafka concepts that support scaling

\*\*Answer:\*\* Topic Replication;

Topic Paritioning;

Consumer Group Id;

Ack All

\*\*Code:\*\*

4. Please select the attributes that are good candidates for partitioning key. Consider cardinality of the field you have selected and scaling aspects of your application

\*\*Answer:\*\* vendor\_id, tpep\_pickup\_datetime, tpep\_dropoff\_datetime

\*\*Code:\*\*

5. Which configurations below should be provided for Kafka Consumer but not needed for Kafka Produce

\*\*Answer:\*\* Deserializer Configuration;

Topics Subscription;

Bootstrap Server;

Group-Id;

Offset

\*\*Code:\*\* N/A