Kafka with Spring Boot - Cheatsheet & Interview Guide

# Kafka Core Concepts

* Producer, Consumer, Topic, Partition, Offset
* Messages are immutable; partitions are append-only logs
* Kafka is horizontally scalable and durable

# Spring Boot Kafka Integration

* Use KafkaTemplate to send messages
* Use @KafkaListener to consume messages
* Configure bootstrap-servers, serializers/deserializers in application.yml

# Structured Messaging

## JSON

* Use JsonSerializer and JsonDeserializer

## Avro + Schema Registry

* Use kafka-avro-serializer
* Define .avsc schema, generate POJO using avro-maven-plugin
* Configure Schema Registry URL

# Error Handling & Retry

* Set AckMode (MANUAL, RECORD, etc.)
* Use DefaultErrorHandler with FixedBackOff
* Use DeadLetterPublishingRecoverer for failed messages

# Kafka Transactions

* Enable transactions with transaction-id-prefix
* Use kafkaTemplate.executeInTransaction() for atomic operations
* Set enable.idempotence=true for EOS

# Kafka Streams

* Use StreamsBuilder and KStream to process topics
* Supports map, filter, join, window, aggregate
* Run inside Spring @Configuration class

# Kafka Security

* Enable SSL/TLS for encryption
* Use SASL/PLAIN or SCRAM for authentication
* Enable ACLs for topic-level authorization

# Monitoring Tools

* Use JMX Exporter + Prometheus + Grafana
* Use Kafdrop or AKHQ for topic UI
* Monitor lag, ISR, under-replicated partitions, GC

# Kafka Interview Questions & Answers

* Q: What is the difference between Kafka and a message queue?
* A: Kafka is log-based, supports pub-sub, and stores data durably; MQs are queue-based and delete messages after consumption.
* Q: What is the role of partitions in Kafka?
* A: Partitions enable parallelism and scaling of reads/writes across brokers.
* Q: How does Kafka ensure message durability?
* A: Messages are written to disk, replicated across brokers, and acknowledged based on acks settings.
* Q: What is Kafka’s exactly-once delivery mechanism?
* A: Enable idempotent producer and use transactional writes with isolation.level=read\_committed.
* Q: How do you monitor Kafka in production?
* A: Use JMX metrics, Prometheus/Grafana dashboards, and alert on lag, ISR, GC, disk.
* Q: What is schema evolution in Avro?
* A: The ability to update schema while maintaining compatibility using backward, forward, or full compatibility modes.