

Apache Spark - Interview Sheet

Top 10 Apache Spark Interview Questions - Data Engineering

worksheetdownload.com

Top 10 Interview Questions

Q1. What is Apache Spark and what primary problem does it solve in the Data Engineering space?

Answer:

Q2. Explain the core architecture or components of Apache Spark and how they work together.

Answer:

Q3. What are the main alternatives to Apache Spark and when would you choose it over them?

Answer:

Q4. How do you install, configure, and get started with Apache Spark for a new project?

Answer:

Q5. What are the key configuration options or settings that affect Apache Spark's behavior in production?

Answer:

Q6. How does Apache Spark handle scalability, reliability, and high availability?

Answer:

Q7. What observability does Apache Spark provide: metrics, logs, traces, or health endpoints?

Answer:

Q8. What are the common performance bottlenecks in Apache Spark and how do you address them?

Answer:

Q9. What security best practices should you follow when deploying Apache Spark?

Answer:

Q10. What are the most common pitfalls or mistakes teams make when adopting Apache Spark?

Answer:
