Big Data Analysis Question Bank

1. Discuss the four characteristics of Big Data?
2. Compare Traditional RDBMS and Big Data Approach. Justify with example Which one to use?
3. Give an overview about the big data being generated by Instagram.
4. Explain Hadoop Ecosystem with diagram
5. Explain its Physical architecture of Hadoop.
6. Discuss the classification of Big Data
7. Compare structured, semi-structured and un-structured data in terms of Technology use, Flexibility, Scalability, Robustness, Query performance.
8. Discuss the core components of Hadoop framework
9. What is Shuffling in MapReduce? Explain with word count problem
10. How does NameNode tackle DataNode failure?
11. Explain How Map and Reduce Work with the help of MapReduce Pipeline?
12. Why is HDFS more suitable for applications having large datasets and not when there are small files?
13. What is a JobTracker and TaskTracker?
14. Explain Matrix Vector Multiplication by MapReduce.
15. Imagine you're designing a data-intensive e-commerce website. Assess the potential challenges and benefits of using a NoSQL database over a traditional relational database in this context.
16. Different architecture style of NOSQL.
17. Suppose you have a high-traffic social media platform. Which type of NoSQL database would you choose, and why? Provide specific reasons for your choice.
18. What are the main types of NoSQL databases?
19. What are the advantages of using NoSQL databases?
20. What is eventual consistency in NoSQL databases?
21. What are some challenges of using NoSQL databases?
22. Compare and contrast the advantages and disadvantages of using a column-family NoSQL database and a graph NoSQL database in a scenario where data relationships are crucial.