Data Science Interview Questions

1. What do you know about these positions?
2. How would your past employers describe you?
3. What are your favourite tools to use when analysing data?
4. What are the qualifications as a data scientist?
5. What do you do to avoid selection bias?
6. How do I prepare for a data science interview?
7. What aspects of data analysis do you find most challenging? How do you overcome these challenges in your work?
8. Describe your experience in leading technical teams?
9. How do you ensure you are aware of recent developments in data management?
10. How do you determine the types of data with which you work?
11. How do you organize an extensive set of data?
12. How do you identify whether there is a correlation or a cause in a dataset?
13. What programs are you proficient with for data management and analysis?
14. If our finance team wanted to create a five-year forecast, would you apply linear or logistic regression? Why?
15. Describe your experience integrating data for deep learning?
16. What machine learning framework are you familiar with using?
17. How have you implemented A/B testing in past data projects?
18. Which programming language do you feel is most suitable for text analytics?
19. What is one project you worked on that required cluster sampling? What was the outcome?
20. Share an example of an iteration from one of your successful data projects. What made it successful?
21. What parameters do you establish when building an artificial neural network?
22. How do you ensure the confidentiality and integrity of critical data?
23. Have you created your own algorithms before? How did you develop them?
24. How do you approach data cleaning in Python?
25. When working with a large dataset, how do you account for outliers? Missing values or transformation?
26. What are some sorting algorithms you have used in the R language?
27. If your data team uses Hadoop, how would you integrate it with R for enhanced data analysis?
28. Describe your proficiency in SQL?
29. What are some successful projects you have completed in SQL? What made them successful?
30. What applications would you use the recommender system with?
31. How do you apply univariate analysis?
32. What process do you use to define the number of cluster values within a clustering algorithm?
33. Explain your understanding of auto-encoders?
34. How would you apply the batch normalisation process to organise and analyse data systems?
35. What machine learning library do you feel is most beneficial for supervised learning projects?
36. What steps do you take before you apply machine learning algorithms?
37. How would you resolve unbalanced binary classifications?
38. What are the benefits of box plots when visualization big data?
39. Describe the regularisation method you can apply when implementing training data?
40. How do you select metrics for cross-validation?
41. How would you evaluate a predictive model from multiple regression analysis?
42. When would you use random forests over support vector machines?
43. How do you apply dimension reduction to data?
44. What types of problems would require the use of principal component analysis?
45. What is an alternative you would use instead of mean square error?
46. How do you ensure that the regression models you develop fit with the data?
47. How would you support finance teams in creating a decision tree?
48. What probability fundamentals are crucial to incorporate in data analysis?
49. What is the relevance of the Markov chain in fixed data locations?
50. How can an experimental design support investigation into behaviour?
51. When would you apply mean imputation in your analysis?