SKY STATES

DATA SCIENCE FULL MOCK INTERVIEW QUESTIONS

(Fresher to Experienced Level – Skills Wise Complete Preparation)

Created by: Ravi Kumar Singh

SECTION – A : DATA SCIENCE FUNDAMENTALS

1- What is Data Science?

2- Difference between Data Science, AI, ML, and Deep Learning.

3- Explain Data Analytics vs Data Science.

4- What is the Data Science lifecycle?

5- What is supervised learning?

6- What is unsupervised learning?

7- What is reinforcement learning?

8- What is overfitting and underfitting?

9- Explain train-test split.

10- What is model evaluation?

11- What is feature selection?

12- What is feature engineering?

13- What is dimensionality reduction?

14- What is the difference between classification and

regression?

15- What is sampling?

16- What is bias in Machine Learning?

17- What is variance in Machine Learning?

18- Explain bias-variance tradeoff.

19- What is cross-validation?

20- What is data leakage?

SECTION – B : PYTHON INTERVIEW QUESTIONS

Basic Python

1. What are variables in Python?
2. Difference between list, tuple, set, and dictionary.
3. What are mutable and immutable objects?
4. Explain data types in Python.
5. What is type casting?
6. What is slicing?
7. Difference between append() and extend().
8. Difference between remove(), pop(), and del.
9. What is list comprehension?
10. What is dictionary comprehension?

Functions

11. What is a function?
12. What are lambda functions?

13. Difference between *args and **kwargs.
14. What is recursion?
15. What are decorators?

OOPs Concepts

16. What is Object-Oriented Programming?
17. Explain class and object.
18. What is inheritance?
19. What is polymorphism?
20. What is encapsulation?
21. What is abstraction?
22. Difference between method overloading and overriding.

Advanced Python

23. What are iterators?
24. What are generators?
25. Difference between module and package.
26. What is exception handling?
27. Explain try, except, else, finally.
28. What is multithreading?
29. Difference between deep copy and shallow copy.
30. What is memory management in Python?

SECTION – C : NUMPY INTERVIEW QUESTIONS

1. What is NumPy?
2. Why is NumPy faster than lists?
3. What is ndarray?
4. Explain broadcasting.
5. Difference between reshape() and flatten().
6. What is vectorization?
7. Difference between zeros(), ones(), and empty().
8. Explain indexing and slicing in NumPy.
9. What is axis in NumPy?
10. Difference between vstack() and hstack().
11. Explain random module in NumPy.
12. What is masking?
13. How to handle missing values in NumPy?
14. Explain aggregation functions.
15. Difference between copy() and view().

SECTION – D : PANDAS INTERVIEW QUESTIONS

1. What is Pandas?
2. Difference between Series and DataFrame.
3. What is iloc and loc?
4. How to read CSV and Excel files?
5. How to remove duplicates?

6. How to handle missing values?
7. Difference between merge(), join(), and concat().
8. What is groupby()?
9. Explain pivot table.
10. What is apply() function?
11. Difference between map() and apply().
12. How to rename columns?
13. What is value_counts()?
14. Explain sorting in Pandas.
15. Difference between drop() and drop_duplicates().
16. How to filter rows?
17. What is boolean indexing?
18. Explain datetime handling in Pandas.
19. How to export data?
20. What is multi-indexing?

SECTION – E : SQL INTERVIEW QUESTIONS

SQL Basics

1. What is SQL?
2. Difference between SQL and NoSQL.
3. What are primary keys and foreign keys?
4. What are constraints?

5. Explain normalization.
6. What is denormalization?
7. Difference between CHAR and VARCHAR.
8. What is indexing?
9. What is composite key?
10. What is auto increment?

SQL Queries

11. Difference between WHERE and HAVING.
12. Difference between DELETE, DROP, and TRUNCATE.
13. What is GROUP BY?
14. What is ORDER BY?
15. What is DISTINCT?
16. What are aggregate functions?
17. Difference between UNION and UNION ALL.
18. What is CASE statement?
19. Explain COALESCE().
20. What is NULL handling?

SQL Joins

21. Explain INNER JOIN.
22. Explain LEFT JOIN.
23. Explain RIGHT JOIN.
24. Explain FULL JOIN.

25. Difference between JOIN and UNION.

Advanced SQL

26. What are subqueries?
27. What are CTEs?
28. What are window functions?
29. Explain ROW_NUMBER(), RANK(), DENSE_RANK().
30. Write query for 2nd highest salary.
31. Write query to remove duplicate rows.
32. What are stored procedures?
33. What are triggers?
34. Explain ACID properties.
35. What is transaction management?

SECTION – F : STATISTICS INTERVIEW QUESTIONS

1. What is statistics?
2. Difference between descriptive and inferential statistics.
3. What is mean, median, and mode?
4. What is variance?
5. What is standard deviation?
6. What is covariance?
7. What is correlation?
8. Difference between covariance and correlation.

9. What is probability?
10. Explain normal distribution.
11. What is skewness?
12. What is kurtosis?
13. What is hypothesis testing?
14. What is p-value?
15. What is confidence interval?
16. What is z-test?
17. What is t-test?
18. Difference between Type 1 and Type 2 error.
19. What is central limit theorem?
20. Explain ANOVA.

SECTION – G : MACHINE LEARNING INTERVIEW QUESTIONS

ML Fundamentals

1. What is Machine Learning?
2. Types of Machine Learning.
3. Explain supervised learning.
4. Explain unsupervised learning.
5. What is reinforcement learning?
6. What is training data and testing data?
7. What is model accuracy?

8. What is precision and recall?
9. What is F1-score?
10. What is ROC-AUC?

Algorithms

11. Explain Linear Regression.
12. Explain Logistic Regression.
13. Explain Decision Tree.
14. Explain Random Forest.
15. Explain Naive Bayes.
16. Explain KNN.
17. Explain SVM.
18. Explain K-Means Clustering.
19. Explain PCA.
20. Explain XGBoost.

Model Improvement

21. What is hyperparameter tuning?
22. What is GridSearchCV?
23. What is RandomizedSearchCV?
24. What is feature scaling?
25. Difference between normalization and standardization.
26. What is regularization?

27. Explain L1 and L2 regularization.
28. What is multicollinearity?
29. What is ensemble learning?
30. Explain bagging and boosting.

SECTION – H : DEEP LEARNING INTERVIEW QUESTIONS

1. What is Deep Learning?
2. Difference between ML and Deep Learning.
3. What is ANN?
4. What is activation function?
5. What is ReLU?
6. What is sigmoid function?
7. What is softmax?
8. What is backpropagation?
9. What is gradient descent?
10. What is epoch and batch size?
11. What is CNN?
12. What is RNN?
13. Difference between CNN and RNN.
14. What is LSTM?
15. What is transfer learning?
16. What is dropout?

17. What is vanishing gradient problem?
18. Explain TensorFlow and Keras.
19. What is GPU training?
20. Explain computer vision and NLP.

SECTION – I : POWER BI INTERVIEW QUESTIONS

1. What is Power BI?
2. Components of Power BI.
3. What is Power Query?
4. What is DAX?
5. Difference between measures and calculated columns.
6. What are KPIs?
7. What are slicers?
8. What is drill-through?
9. What is drill-down?
10. What is star schema?
11. What is snowflake schema?
12. Explain relationships in Power BI.
13. What is row-level security?
14. Difference between Import and DirectQuery.
15. What is incremental refresh?
16. What is dashboard vs report?

17. Explain bookmarks.
18. What are cards and visuals?
19. What is gateway?
20. How to optimize Power BI reports?

SECTION – J : EXCEL INTERVIEW QUESTIONS

1. What is VLOOKUP?
2. Difference between VLOOKUP and XLOOKUP.
3. Explain INDEX and MATCH.
4. What is Pivot Table?
5. What is Conditional Formatting?
6. What is Data Validation?
7. Explain IF formula.
8. Explain nested IF.
9. Explain IF with AND/OR.
10. Difference between COUNT, COUNTA, COUNTIF.
11. What is Goal Seek?
12. What is Solver?
13. What is Power Query in Excel?
14. What are macros?
15. What is dashboard creation in Excel?
16. Explain charts in Excel.

17. What is freeze panes?
18. What is text-to-columns?
19. What is duplicate removal?
20. What is flash fill?

SECTION – K : DATA VISUALIZATION INTERVIEW QUESTIONS

1. What is data visualization?
2. Best chart for trend analysis.
3. Best chart for comparison.
4. Difference between histogram and bar chart.
5. What is dashboard design?
6. What are KPIs?
7. Explain storytelling with data.
8. What is heatmap?
9. What is scatter plot?
10. How do you choose the right chart?

SECTION – L : BIG DATA / CLOUD QUESTIONS

1. What is Big Data?
2. Explain 5Vs of Big Data.
3. What is Hadoop?
4. What is Spark?

5. Difference between Hadoop and Spark.
6. What is cloud computing?
7. Explain AWS in Data Science.
8. What is Azure Data Factory?
9. What is Databricks?
10. What is ETL pipeline?

SECTION – M : PROJECT-BASED QUESTIONS

1. Explain your project architecture.
2. What problem were you solving?
3. Which dataset did you use?
4. How did you clean the data?
5. Which ML algorithm did you choose and why?
6. What challenges did you face?
7. How did you improve model accuracy?
8. Which evaluation metric did you use?
9. How did you deploy the project?
10. What business impact did your project create?

SECTION – N : REAL INDUSTRY SCENARIO QUESTIONS

1. How do you handle missing values?
2. How do you handle imbalanced datasets?

3. What will you do if accuracy is low?
4. How do you reduce overfitting?
5. How do you choose the best algorithm?
6. What if your client changes requirements?
7. How do you explain technical insights to non-technical people?
8. What will you do if the dataset is huge?
9. How do you optimize SQL queries?
10. How do you automate reporting?

SECTION – O : HR + BEHAVIORAL QUESTIONS

1. Tell me about yourself.
2. Why do you want to become a Data Scientist?
3. Explain your strengths and weaknesses.
4. Why should we hire you?
5. Describe a challenge you faced.
6. Explain teamwork experience.
7. How do you manage deadlines?
8. Where do you see yourself in 5 years?
9. Why are you switching careers?
10. What motivates you?

ADVANCED EXPERIENCE-LEVEL QUESTIONS (2–5 Years)

1. How do you design scalable ML pipelines?
2. Explain MLOps.
3. What is model drift?
4. How do you monitor production models?
5. Explain CI/CD in ML.
6. What is A/B testing?
7. How do you optimize large Power BI dashboards?
8. Explain distributed computing.
9. What is feature store?
10. Explain real-time analytics systems.

RAPID FIRE QUESTIONS

1. Difference between Al and ML.
2. Difference between list and tuple.
3. Difference between loc and iloc.
4. Difference between JOIN and UNION.
5. Difference between mean and median.
6. Difference between variance and standard deviation.
7. Difference between supervised and unsupervised learning.
8. Difference between bagging and boosting.

9. Difference between CNN and RNN.
10. Difference between Power BI dashboard and report.

CODING ROUND QUESTIONS

Python

1. Reverse a string.
2. Find duplicate elements in list.
3. Check palindrome.
4. Fibonacci series.
5. Prime number check.
6. Count frequency of characters.
7. Sort dictionary by value.
8. Find largest number in list.
9. Remove duplicates from list.
10. Find factorial using recursion.

SQL Coding

1. Find nth highest salary.
2. Remove duplicate records.
3. Find employees without manager.
4. Fetch last 3 records.
5. Find department-wise highest salary.

6. Calculate running total.
7. Find duplicate emails.
8. Rank employees by salary.
9. Find monthly sales growth.
10. Write query using CASE statement.

FINAL INTERVIEW MASTER QUESTIONS

1. Explain end-to-end Data Science workflow.
2. How do you deploy ML models?
3. How do you improve model performance?
4. Explain one real-time project deeply.
5. How do businesses use Data Science?
6. Explain complete ETL pipeline.
7. What tools have you worked on?
8. Which project are you most proud of?
9. Explain business impact using numbers.
10. What makes you different from other candidates?