MOST ASKED INTERVIEW QUESITONS FOR DATA ANALYSTS

SQL

1. How do you find duplicate records in a table?
2. Explain the difference between INNER JOIN, LEFT JOIN, RIGHT JOIN, and FULL OUTER JOIN.
3. How do you optimize a slow-running query?
4. What are window functions, and how do you use them for ranking, running totals, or moving averages?
5. What are Common Table Expressions (CTEs), and how do they simplify complex queries?
6. How would you filter results using a subquery in another table?
7. How do you calculate a rolling 7-day average in SQL?
8. What's the difference between WHERE and HAVING? When should you use each?
9. How do you pivot (transpose) a table in SQL?
10. How do indexes improve query performance, and when should you use them?

Swipe →

Power BI

1. What are the key components of Power BI, and how do they interact?
2. How do you handle relationships between tables in Power BI?
3. What is the difference between calculated columns and measures in DAX?
4. How do you optimize Power BI reports for performance?
5. Explain row-level security (RLS) in Power BI and how to implement it. What are the different types of filters in Power BI, and when should you use each?
6. How do you use the CALCULATE function in DAX, and why is it important? How do you create a dynamic date range (e.g., last 7 days) in Power BI?
7. What is the difference between DirectQuery and Import Mode? When should you use each?
8. How would you handle large datasets in Power BI without affecting performance?

Tableau

1. What are the different types of joins in Tableau, and how do they work?
2. What is the difference between a live connection and an extract in Tableau? When should you use each?
3. How do you create calculated fields in Tableau? Can you provide an example?
4. What are Level of Detail (LOD) expressions, and how do FIXED, INCLUDE, and EXCLUDE differ?
5. How do you create dynamic parameters and filters in Tableau?
6. What are the different types of charts available in Tableau, and how do you decide which one to use?
7. How do you optimize Tableau dashboards for better performance?
8. How can you implement row-level security (RLS) in Tableau?
9. What is the difference between table calculations and calculated fields? When should you use each?
10. How do you create a dual-axis chart in Tableau, and when is it useful?

Python Basics

1. How do you read and manipulate data in Pandas? Can you provide examples of common operations (e.g., filtering, grouping, merging)?
2. What is the difference between a list, tuple, set, and dictionary in Python? When should you use each?
3. How do you handle missing values in a dataset using Pandas?
4. What is the difference between apply(), map(), and lambda functions in Pandas?
5. How do you optimize a slow-running Pandas operation for large datasets?
6. Explain the difference between NumPy arrays and Pandas DataFrames. What are list comprehensions, and how do they improve performance in Python?
7. How do you write an efficient SQL query using Python (e.g., with sqlite3 or SQLAlchemy)?
8. How do you visualize data using Matplotlib and Seaborn?
9. Can you provide an example? What are Python's built-in functions for working with date and time?
10. How do you manipulate date-time objects in Pandas?

Data Cleaning

1. How do you handle missing values in Pandas? (dropna(), fillna(), interpolation)
2. How do you detect and remove duplicate rows in a DataFrame? (duplicated(), drop_duplicates())
3. How do you standardize column names and data formats in Pandas? (str.lower(), str.replace(), astype()) ?
4. How do you handle outliers in a dataset using Pandas? (quantile(), IQR, z-score)
5. How do you split a column into multiple columns? (str.split(), expand=True)
6. How do you merge and join datasets efficiently in Pandas? (merge(), concat(), join())
7. How do you convert data types in Pandas? (astype(), to_datetime(), pd.to_numeric())
8. How do you filter and remove unwanted characters from text columns? (str.replace(), regex patterns)
9. How do you group and aggregate data effectively in Pandas? (groupby(), agg())
10. How do you reshape a dataset (pivot, melt, transpose) in Pandas? (pivot(), melt(), T)

Statistics

1. What is the difference between descriptive and inferential statistics?
2. How do you calculate and interpret mean, median, and mode?
3. What is standard deviation, and how does it relate to variance?
4. Explain the Central Limit Theorem and its significance in statistics?
5. What are the different types of probability distributions (normal, binomial, Poisson, etc.)?
6. How do you detect and handle outliers in a dataset?
7. What is correlation vs. causation? How do you interpret correlation coefficients?
8. Explain hypothesis testing, p-values, and confidence intervals?
9. What is the difference between Type I and Type II errors?
10. What are z-score and t-score, and when should you use them?

Excel

- What are the different types of LOOKUP functions in Excel, and how do they work? (VLOOKUP(), HLOOKUP(), XLOOKUP(), INDEX-MATCH())
- How do you use Pivot Tables for data analysis? (Grouping, Slicers, Calculated Fields)
- What are Array Formulas, and how do they work? (SEQUENCE(), FILTER(), SORT(), UNIQUE())
- How do you use Conditional Formatting for data visualization? (Highlighting duplicates, top/bottom values, and custom rules)
- How do you clean and preprocess data in Excel? (TRIM(), CLEAN(), TEXT() functions)
- What are Excel Macros and VBA, and how can they be used to automate tasks?
- How do you perform What-If Analysis in Excel? (Goal Seek, Data Tables, Solver)
- How do you use Power Query for data transformation? (Merging, Cleaning, and Automating Data Imports)

Soft Skills

1. How do you explain complex data insights to a non-technical audience?
2. Can you describe a time when you had to collaborate with a cross-functional team (e.g., marketing, finance, or engineering)?
3. How do you handle conflicting priorities and tight deadlines?
4. Tell me about a time when you made a mistake in your analysis. How did you handle it?
5. How do you approach problem-solving when faced with incomplete or messy data?
6. Can you give an example of how you used data to drive a business decision?
7. How do you handle feedback or criticism of your analysis?
8. Describe a time when you had to convince stakeholders to take action based on your data insights?
9. How do you stay up-to-date with industry trends and continuously improve your skills?
10. Tell me about a challenging project and how you managed it from start to finish.

Problem Solving

1. You receive a dataset with missing values in critical columns. How would you handle it?
2. Your SQL query is taking too long to run. How would you optimize it?
3. A report you generated is showing unexpected trends. How would you investigate and validate the results?
4. Your dashboard performance is slow due to a large dataset. How do you improve efficiency?
5. A stakeholder requests a KPI that is not directly available in the data. How do you derive or approximate it?
6. You need to analyze customer churn, but you only have limited historical data. How would you approach this?
7. Your analysis contradicts the assumptions made by leadership. How do you verify and present your findings?
8. You need to segment customers for targeted marketing, but there are no predefined categories. How do you create meaningful segments?

Scenario Based

1. You found an error in a report after it was sent to leadership. How would you handle the situation?
2. A stakeholder is questioning the accuracy of your analysis. How would you respond?
3. You're given a dataset with missing and inconsistent values. How do you clean and present reliable insights?
4. A manager asks for an urgent report, but you realize the required data isn't available. What do you do?
5. Two stakeholders have conflicting requirements for a dashboard. How do you manage expectations and find a solution?
6. You've identified a trend in the data that contradicts a leadership decision. How would you communicate your findings?
7. Your team is working on a high-priority project, but a last-minute request comes in. How do you prioritize your tasks?
8. A non-technical stakeholder is struggling to understand your visualization. How do you make the insights clearer?