Most Asked Interview Questions for Data Analysts

»»

SQL

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

Power BI

- What are the key components of Power BI, and how do they interact?
- How do you handle relationships between tables in Power BI?
- What is the difference between calculated columns and measures in DAX?
- How do you optimize Power BI reports for performance?
- 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?
- 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?
- What is the difference between DirectQuery and Import Mode? When should you use each?
- How would you handle large datasets in Power BI without affecting performance?
-

Tableau

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

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)

Python Basics

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

Data Cleaning

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

Problem Solving

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

Statistics

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

Soft Skills

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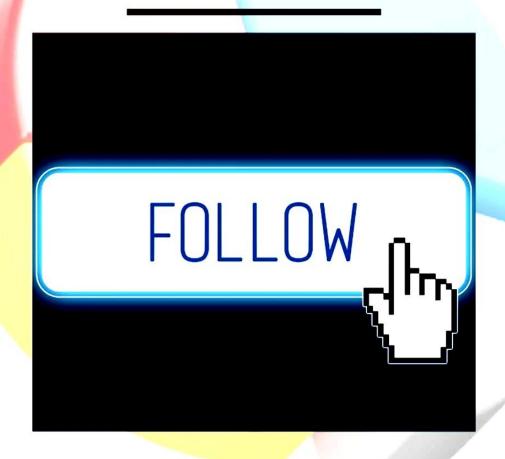
Scenario Based

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

Found Helpful ?

FOLLOW

Comment Yes to get pdf in your dm