

Master Your Next Data Analyst Interview



Python Basics Refresher

- What are Python's key features and advantages?
- Explain lists, tuples, and dictionaries.
- How do you handle exceptions in Python?
- Describe the difference between '==' and 'is'.
- What are list comprehensions?
- How do you comment your code?



Essential Data Libraries

- What is NumPy used for?
- Explain the Pandas library in data analysis.
- What is the primary data structure in Pandas?
- Differentiate between Series and DataFrame.
- Name some visualization libraries.
- What is the use of the Scipy library?



Pandas Data Manipulation 1

- How to create a DataFrame from a dictionary?
- Explain the `.head()` and `.tail()` methods.
- How to select a specific column in a DataFrame?
- What is boolean indexing?
- How to filter rows based on a condition?
- Describe the `.info()` and `.describe()` methods.



Pandas Data Manipulation 2

- How to handle missing values?
- Explain the `.dropna()` method.
- How to fill missing data with `.fillna()`?
- How to drop duplicate rows?
- What is the use of the `.groupby()` function?
- How to rename columns in a DataFrame?



Data Aggregation & Merging

- How to aggregate data using `.agg()`?
- Explain the difference between `.merge()` and `.join()`.
- What is concatenation in Pandas?
- How to create a pivot table?
- Describe the `.value_counts()` method.
- How to sort values in a DataFrame?



Data Cleaning Techniques

- How to change the data type of a column?
- Explain string manipulation with `.str`.
- How to apply a function to a column?
- What is feature encoding?
- How to use the `.map()` function?
- How to handle outliers in a dataset?



Statistical Analysis Questions

- How to calculate correlation in Pandas?
- Explain measures of central tendency.
- What is a normal distribution?
- How to perform a t-test in Python?
- Describe hypothesis testing steps.
- How to calculate percentiles?



Data Visualization Basics

- How to create a plot with Matplotlib?
- Explain the basics of a Seaborn plot.
- What is the difference between a bar plot and a histogram?
- How to create a scatter plot?
- How to customize plot labels and titles?
- How to show a plot in your environment?



SQL with Python

- How to connect to a SQL database?
- What library is commonly used for SQL connections?
- How to execute a SQL query from Python?
- How to load SQL results into a DataFrame?
- How to write back to a database?
- Explain parameterized queries for safety.



Scenario-Based Problems 1

- How would you analyze a messy CSV file?
- Describe your process for data validation.
- How to find and remove duplicate records?
- How to combine multiple data files?
- How to handle a column with mixed data types?
- How to sample a large dataset?



Scenario-Based Problems 2

- How to track changes in data over time?
- How to calculate growth rates or percentages?
- How to rank items within a group?
- How to handle date and time data?
- How to create a new feature from existing ones?
- How to export your cleaned data?



Efficiency & Best Practices

- What are some ways to make your code efficient?
- Explain the importance of code readability.
- How to use vectorization instead of loops?
- What are Python decorators?
- How to document your analysis code?
- Why is version control important?



Advanced Topics

- What is object-oriented programming in Python?
- Explain the use of lambda functions.
- How to use the `.apply()` method effectively?
- What are generators and iterators?
- How to work with APIs using the requests library?
- Briefly explain regular expressions.



Final Preparation Tips

- Practice writing code on a whiteboard.
- Be ready to explain your thought process.
- Review your past projects and code.
- Understand the business context of problems.
- Prepare your own questions for the interviewer.
- Stay calm and confident. You know this.



Follow for More



Like this post? Repost!

Share your thoughts in the comments!

I share little hacks and tips I actually use in
Data & AI — follow if you wanna see more!



Shoaib Akthar
@theshoaibakthar

Follow for more Data Analytics, Data Science and AI insights