Python, Pandas & Polars Interview Questions

Python Programming (Core)

1. What are Python's key features?
2. Explain mutable vs immutable data types in Python with examples.
3. How do list comprehensions work? Give an example.
4. What are Python decorators? Can you write a simple one?
5. Explain Python's memory management (Garbage collection).
6. What is the difference between @staticmethod, @classmethod, and instance methods?
7. How do you handle exceptions in Python? Write a custom exception class.
8. What are iterators and generators? Show a generator function example.
9. Explain Python's GIL (Global Interpreter Lock).
10. How would you optimize Python code performance?

Pandas (Data Manipulation)

11. How do you handle missing data in a Pandas DataFrame?
12. Write code to group data by a column and calculate the mean of another column.
13. How do you merge two DataFrames in Pandas? Explain merge, join, and concat.
14. How to remove duplicate rows and duplicate columns in Pandas?
15. Write code to find the top 5 rows with the highest values in a column.
16. How do you handle large datasets efficiently with Pandas?
17. Explain apply() vs map() vs applymap() in Pandas.
18. How do you pivot or unpivot (melt) data in Pandas?
19. How do you write a Pandas DataFrame to CSV, Excel, and SQL?
20. How do you get column datatypes and memory usage in Pandas?

Polars (Efficient Data Processing)

21. What is Polars and how is it different from Pandas?
22. Explain lazy evaluation in Polars.
23. Write code to filter and group data using Polars.
24. How do you read large CSV/Parquet files efficiently using Polars?
25. Explain when to use Polars over Pandas.
26. How do you perform joins in Polars?
27. Explain the difference between eager and lazy APIs in Polars.
28. How to convert a Polars DataFrame to Pandas and vice versa?
29. How does Polars achieve better performance internally?
30. Write Polars code to compute summary statistics for all numeric columns.

Problem-Solving & Debugging

31. How do you debug a slow Python script?
32. Describe a time you solved a tricky data processing problem.
33. If a script is failing randomly in production, how do you troubleshoot it?
34. How do you write efficient algorithms for large datasets?
35. Explain Big O notation and analyze time complexity for a sorting algorithm.

Collaboration & Teamwork

36. How do you manage code reviews in your team?
37. How do you handle merge conflicts in Git?
38. What is your approach to writing clean and maintainable code?
39. How do you document your data pipelines and scripts?
40. Describe a situation where you collaborated cross-functionally to solve a problem.