

TCS Data Analyst interview

TCS (Tata Consultancy Services) Data Analyst interview questions and answers based on recent patterns in 2025.

◆ **1. What is the difference between INNER JOIN and LEFT JOIN in SQL?**

Answer:

- INNER JOIN: Returns only matching rows from both tables.
- LEFT JOIN: Returns all rows from the left table and matching rows from the right table. Non-matching right-side rows will be NULL.

-- Example:

```
SELECT A.customer_id, B.order_id  
FROM Customers A  
LEFT JOIN Orders B  
ON A.customer_id = B.customer_id;
```

◆ **2. How do you handle missing data in a dataset?**

Answer:

- Methods:
 - Drop missing rows/columns (`dropna()` in Python).
 - Impute using mean/median/mode.
 - Predict missing values using regression or KNN imputer.
- Depends on context: if <5% missing → drop; if important → impute.

◆ **3. What are the key steps in a data analysis process?**

Answer:

1. Define the problem.
2. Collect data.
3. Clean and preprocess data.

4. Analyze using statistics, SQL, or Python.
 5. Visualize using Power BI or Excel.
 6. Interpret results and present insights.
-

◆ **4. What is the difference between OLAP and OLTP?**

Answer:

Feature	OLTP	OLAP
Purpose	Transactional	Analytical
Speed	Fast writes	Fast reads
Normalization	Highly normalized	Denormalized
Examples	ATM, e-commerce	Sales dashboard

◆ **5. Explain the use of CALCULATE() in Power BI.**

Answer:

- CALCULATE() modifies the context of a calculation using filters.
- Example:

Total2023Sales = CALCULATE(SUM(Sales[Amount]), Sales[Year] = 2023)

◆ **6. How do you identify outliers in data?**

Answer:

- Statistical Methods:
 - IQR Method: Values below $Q1 - 1.5 \times IQR$ or above $Q3 + 1.5 \times IQR$
 - Z-score: Data points with $|z| > 3$ are outliers
- Visualization:
 - Boxplots, scatter plots

◆ **7. What is normalization in SQL databases?**

Answer:

- Process of organizing data to reduce redundancy.
 - Common normal forms:
 - 1NF: Atomic values
 - 2NF: No partial dependency
 - 3NF: No transitive dependency
-

◆ **8. Explain the difference between WHERE and HAVING in SQL.**

Answer:

- WHERE: Filters rows **before** grouping.
- HAVING: Filters groups **after** aggregation.

```
SELECT dept, COUNT(*)
```

```
FROM Employees
```

```
GROUP BY dept
```

```
HAVING COUNT(*) > 5;
```

◆ **9. What are measures and dimensions in Power BI?**

Answer:

- **Dimension:** Descriptive data (e.g., Product, Region).
 - **Measure:** Numeric calculations (e.g., Sales, Profit).
-

◆ **10. Scenario: Sales dropped in a region, how will you analyze it?**

Answer:

1. Compare Y-o-Y or M-o-M trends using visuals.
 2. Filter by region → check product mix, quantity.
 3. Analyze customer churn, pricing, or competition.
 4. Correlate with external factors (holidays, supply chain).
-

◆ **11. Write a SQL query to get the second highest salary.**

```
SELECT MAX(salary) AS Second_Highest  
FROM employees  
WHERE salary < (SELECT MAX(salary) FROM employees);
```

◆ **12. How do you optimize a large dataset in Power BI?**

Answer:

- Remove unused columns.
 - Use star schema.
 - Use DAX over M where possible.
 - Aggregate at query level.
 - Turn off auto-date/time.
-

◆ **13. What is data wrangling?**

Answer:

- The process of cleaning, transforming, and mapping raw data into a usable format.
 - Tools: Power Query, Pandas, SQL.
-

◆ **14. What is the use of GROUP BY in SQL?**

Answer:

- Groups rows with the same values into summary rows.

```
SELECT department, COUNT(*) FROM employees GROUP BY department;
```

◆ **15. Explain a project where you used data to drive business decisions.**

Answer:

(Customize based on your project)

"I worked on an e-commerce sales dashboard. I analyzed top-selling products, profit margins, and customer segments. Found that returns from a particular region were high. After root cause analysis, we adjusted logistics and reduced return rate by 18%."

 **Tips for TCS Data Analyst Interviews:**

- Practice case-based thinking: "How will you approach this business problem?"
 - Be confident in **SQL + Excel + Power BI/Python**.
 - Show communication skills & storytelling with data.
-

Dipankar Pal

(Power BI Developer)

Email – Dippal351@gmail.com