

## Infosys Data Analyst Interview Questions – SQL & Power BI (30 Q&A)

---

### SQL – 15 Questions

1. **Q:** What is the difference between INNER JOIN and LEFT JOIN?

**A:**

- **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; unmatched right rows return NULL.
- 

2. **Q:** How do you find the **second highest salary** in a table?

**A:**

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

---

3. **Q:** What is the difference between WHERE and HAVING?

**A:**

- WHERE filters rows before aggregation.
  - HAVING filters aggregated results after GROUP BY.
- 

4. **Q:** How to find **duplicate records** in SQL?

**A:**

```
SELECT column_name, COUNT(*)  
FROM table_name
```

GROUP BY column\_name

HAVING COUNT(\*) > 1;

---

**5. Q:** Write a query to get **top 5 customers by sales**.

**A:**

SELECT customer\_id, SUM(sales) AS total\_sales

FROM orders

GROUP BY customer\_id

ORDER BY total\_sales DESC

LIMIT 5;

---

**6. Q:** Explain RANK() vs DENSE\_RANK().

**A:**

- RANK() leaves gaps in ranking when there are ties.
  - DENSE\_RANK() gives consecutive ranks without gaps.
- 

**7. Q:** How do you handle **NULL values** in SQL?

**A:**

Use functions like COALESCE() or ISNULL() to replace NULL with default values.

---

**8. Q:** Write a query to calculate **running total of sales**.

**A:**

SELECT order\_date, sales,

SUM(sales) OVER (ORDER BY order\_date) AS running\_total

FROM orders;

---

**9. Q:** What is the difference between UNION and UNION ALL?

**A:**

- UNION removes duplicates.
  - UNION ALL keeps duplicates.
- 

**10. Q:** How to find **customers who didn't place any order?**

**A:**

```
SELECT c.customer_id, c.customer_name
FROM customers c
LEFT JOIN orders o ON c.customer_id = o.customer_id
WHERE o.customer_id IS NULL;
```

---

**11. Q:** Explain the purpose of **indexes** in SQL.

**A:**

Indexes speed up data retrieval by reducing the number of rows scanned.

---

**12. Q:** Write a query to get **monthly sales for 2024**.

**A:**

```
SELECT MONTH(order_date) AS month, SUM(sales) AS total_sales
FROM orders
WHERE YEAR(order_date) = 2024
GROUP BY MONTH(order_date)
ORDER BY month;
```

---

**13. Q:** What is the difference between TRUNCATE and DELETE?

**A:**

- DELETE removes rows but logs each transaction and can have a WHERE clause.
  - TRUNCATE removes all rows without logging each row and resets identity.
- 

**14. Q:** How to find the **percentage contribution of each category** to total sales?

**A:**

SELECT category,

SUM(sales) AS category\_sales,

(SUM(sales) \* 100.0 / SUM(SUM(sales)) OVER()) AS percentage\_contribution

FROM orders

GROUP BY category;

---

**15. Q:** Explain CTE and when to use it.

**A:**

- **CTE (Common Table Expression)** is a temporary result set defined using WITH for better readability and reusability in complex queries.
- 

## Power BI – 15 Questions

**16. Q:** What is the difference between **calculated column** and **measure** in Power BI?

**A:**

- **Calculated column** is stored in the model and computed row-by-row.
  - **Measure** is calculated on the fly based on filter context.
- 

**17. Q:** How do you handle **data refresh** in Power BI?

**A:**

- Schedule refresh in Power BI Service.

- Ensure gateway is configured for on-premises sources.
- 

**18. Q:** Explain **Row-Level Security (RLS)** in Power BI.

**A:**

RLS restricts data visibility based on filters defined in roles, so users only see authorized data.

---

**19. Q:** What is the difference between **Import** and **DirectQuery** mode?

**A:**

- **Import** stores data in Power BI for faster performance.
  - **DirectQuery** queries the source directly, useful for large datasets.
- 

**20. Q:** How to create **year-to-date (YTD) sales** in DAX?

**A:**

YTD Sales =

TOTALYTD(SUM(Orders[Sales]), 'Orders'[Order Date])

---

**21. Q:** How do you improve **Power BI dashboard performance**?

**A:**

- Reduce columns and rows in data model.
  - Use measures instead of calculated columns where possible.
  - Apply aggregations and query folding.
- 

**22. Q:** Explain **relationship cardinality** in Power BI.

**A:**

- **One-to-One (1:1)**, **One-to-Many (1:\*)**, and **Many-to-Many (:)** define how tables relate.
-

**23. Q:** How to create a **previous year sales comparison** in Power BI?

**A:**

Previous Year Sales =

```
CALCULATE(SUM(Orders[Sales]), SAMEPERIODLASTYEAR('Orders'[Order Date]))
```

---

**24. Q:** What is a **Star Schema** in Power BI?

**A:**

A data model design with one **fact table** linked to multiple **dimension tables**, improving query performance.

---

**25. Q:** How do you use **Bookmarks** in Power BI?

**A:**

Bookmarks capture a report's current state, allowing toggling between views for storytelling or navigation.

---

**26. Q:** How to create a **Top 5 products by sales** measure in DAX?

**A:**

Top 5 Products Sales =

```
TOPN(5, SUMMARIZE(Products, Products[Product Name], "Sales",  
SUM(Orders[Sales])), [Sales], DESC)
```

---

**27. Q:** What is **Query Folding** in Power BI?

**A:**

Query Folding pushes transformations back to the data source, reducing load time and improving performance.

---

**28. Q:** How to create a **Profit Margin %** in Power BI?

**A:**

Profit Margin % =

DIVIDE(SUM(Orders[Profit]), SUM(Orders[Sales]), 0)

---

**29. Q:** How do you publish and share a Power BI report securely?

**A:**

- Publish to Power BI Service.
  - Share via workspaces, apps, or RLS-based sharing.
- 

**30. Q:** How do you handle **different time zones** in Power BI reports?

**A:**

- Convert UTC to local time in Power Query using DateTimeZone.SwitchZone.
- 

Dipankar Pal

[Dippal351@gmail.com](mailto:Dippal351@gmail.com)

[www.linkedin.com/in/dipankar-analyst](https://www.linkedin.com/in/dipankar-analyst)