

# **DATA ANALYTICS AND VISUALIZATION**

## **INTERVIEW GUIDE**

### **QUESTIONS & ANSWERS**



Minal Pandey 



## Excel:

Excel Questions to practice: [Excel Practice](#)

## Steps to create powerful visualization in Power BI:

### Discovering Custom Visualizations:

- Ditching the Ordinary: Say goodbye to standard charts; Power BI's custom visualizations offer unique options.
- Marketplace Exploration: Dive into the extensive Power BI marketplace, where numerous custom visuals await experimentation.
- Precision in Selection: Choose visuals that fit seamlessly with your data and storytelling goals.

### Unveiling the Power of Advanced Analytics:

- Beyond Basic Charts: Explore advanced features like forecasting, clustering, and key influencers for a more sophisticated analysis.
- Insightful Revelations: Delve deeper into your data, revealing hidden patterns and relationships that basic charts might miss.
- Holistic Data Understanding: Use advanced analytics tools to gain a nuanced and comprehensive understanding of your dataset.

### Elevating Reports with Advanced Formatting and Interaction:

- Crafting Visual Masterpieces: Hone your formatting skills to transform your reports into visually stunning presentations.
- Spotlight on Significance: Emphasize critical data points using conditional formatting techniques.
- User-Centric Interaction: Improve user experience through interactive features such as drill-through options and report tooltips.
- Artistry in Design: Infuse creativity into your reports with thoughtful choices in color schemes, fonts, and layout designs.

### The Continuous Journey of Learning:

- Staying Curious: Approach Power BI visualization as an ongoing learning journey, always seeking new techniques and approaches.
- Community Engagement: Connect with the Power BI community for insights, tips, and collaborative learning.
- Real-World Application: Apply your newfound skills to real-world scenarios, refining your expertise through practical experience.

## **Q1. What is Ribbon in Excel and where does it appear?**

Answer:

The Ribbon is basically your key interface with Excel, and it appears at the top of the Excel window. It allows users to access many of the most important commands directly. It consists of many tabs such as File, Home, View, Insert, etc. You can also customize the ribbon to suit your preferences. To customize the Ribbon, right-click on it and select the “Customize the Ribbon” option.

## **Q2. How can we deal with problems that arise when the data flows in from a variety of sources?**

Answer:

There are many ways to go about dealing with multi-source problems. However, these are done primarily to solve the problems of:

- Identifying the presence of similar/same records and merging them into a single record
- Re-structuring the schema to ensure there is good schema integration

## **Q3. How to convert pdf to excel?**

Answer:

Open the PDF document you want to convert in XLSX format in Acrobat DC.

Go to the right pane and click on the “Export PDF” option.

Choose spreadsheet as the Export format.

Select “Microsoft Excel Workbook.”

Now click “Export.”

Download the converted file or share it.

## **Q4. How to enable macros in excel?**

Answer:

Click the file tab and then click “Options.”

A dialog box will appear. In the “Excel Options” dialog box, click on the “Trust Center” and then “Trust Center Settings.”

Go to the “Macro Settings” and select “enable all macros.”

Click OK to apply the macro settings.

## **Q5. What is the shortcut to add a filter to a table in EXCEL?**

Answer:

The filter mechanism is used when you want to display only specific data from the entire dataset.

By doing so, there is no change being made to the data. The shortcut to add a filter to a table is Ctrl+Shift+L.

## Q6. How to round off in excel?

Answer:

You can use the ROUNDUP function to round off in Excel. Here is how –

Select a blank cell.

Use the ROUND function by entering the syntax – “=ROUNDUP(number,num\_digits).

Where number denotes the number to round up and num\_digits refers to a few digits to which it should be rounded up.

Press Return.

Using the button, you can round off in Excel by following these steps –

Select the cells you want to round off.

Go to the Home tab, click Increase Decimal or Decrease Decimal.

Press Enter.

## Q7. How to compare two columns in excel?

Answer:

You can compare two columns in Excel using the simple IF formula. The syntax to be used is – “=IF-(A2=B2,”Match,” Mismatch”)

If the values to be compared are case-sensitive, use these following formula –

=IF(EXACT(A2,B2),”Match”,”Mismatch”)

## Q8. How are nested IF statements used in Excel?

Answer:

The function IF () can be nested when we have multiple conditions to meet. The FALSE value in the first IF function is replaced by another IF function to make a further test.

## Q9. How to use excel formulas?

Answer:

Formulas in Excel are expressions that can make the calculation based on the information in your spreadsheet. The formula can be used in Excel by first selecting an empty cell where you want to enter the formula. The way to add formula syntax is by typing the “=” equal sign and then followed by operators to be used in the calculation.

The relevant operators used in the formula should be kept between opening and closing parenthesis.

Press Enter to get the results.

## **Q10. What is macro in excel?**

Answer:

Macro refers to an algorithm or a set of actions that help automate a task in Excel by recording and playing back the steps taken to complete that task. Once the steps are stored, you create a Macro, and it can be edited and played back as many times as the user wants.

Macro is great for repetitive tasks and eliminates errors. For example, suppose an account manager must share reports regarding the company employees for non-payment of dues. In that case, it can be automated using a Macro and doing minor changes every month, as needed.

## **Q11. How to maintain accounts in excel sheet format?**

Answer:

You can maintain accounts in an Excel sheet by using its rows and columns for record-keeping and keep track of account-related information such as account numbers, invoicing, and receipt of payments.

The cells can be used to input client or account holder information. The expenses can be calculated using the formula =SUM (F3:F6). You can also use tools such as expense tracking tools and contract tools for auto-invoicing, performing tracking, and billing of expenses.

## **Q12. How to create a serial number generator in excel?**

Answer:

You can create serial numbers in Excel using the Autofill method. Here are the steps –

Insert value (let's say 1) in a row and select it as the active cell.

You will notice that a small box or filler icon at the bottom right of the active cell will appear.

Double click on that box icon to drag to the desired cell.

Now go back to the second row and insert another value (let's say 2).

Now select both the cells at once, and the box or filler icon will appear again.

Drag down the icon to the desired row, and you will notice that this time it shows serial numbers starting with the first value we inserted (1).

## **Q13. What is the order of operations in Excel?**

Answer:

Excel follows PEMDAS: parentheticals, exponents, multiplication, division, addition, and then subtraction. If you type in “=1+2/4” the answer will be 3/2 rather than  $\frac{3}{4}$ .

**Q14. What is Relative cell referencing in excel?**

Answer:

In Relative referencing, there is a change when copying a formula from one cell to another cell concerning the destination. cells' address Meanwhile, there is no change in Absolute cell referencing when a formula is copied, irrespective of the cell's destination. This type of referencing is there by default. Relative cell referencing doesn't require a dollar sign in the formula.

**Q15. Gantt chart in Tableau ?**

Answer:

A Tableau Gantt chart illustrates the duration of events as well as the progression of value across the period. Along with the time axis, it has bars. The Gantt chart is primarily used as a project management tool, with each bar representing a project job.

**Q16. What in Excel is a macro?**

Answer:

An Excel macro is an algorithm or a group of steps that helps automate an operation by capturing and replaying the steps needed to finish it. Once the steps have been saved, you may construct a Macro that the user can alter and replay as often as they like.

**Q17. How are nested IF statements used in Excel?**

Answer:

The function IF() can be nested when we have multiple conditions to meet. The FALSE value in the first IF function is replaced by another IF function to make a further test.

**Q18. How would you track down the last line and segment in VBA?**

Answer:

To track down the last column, utilize the underneath lines code in the VBA module:

```
Sub FindingLastRow()
    Faint lastRow As Long
    lastRow = ActiveSheet.Cells.SpecialCells(xlLastCell).Row
    MsgBox (lastRow)
End Sub
```

To track down the last section, utilize the beneath lines code in the VBA module:

```
Sub FindingLastColumn()
    Faint lastColumn As Long
    lastColumn = ActiveSheet.Cells.SpecialCells(xlLastCell).Column
    MsgBox (lastColumn)
End Sub
```

## **Q19. How to prevent someone from copying the cell from your worksheet in excel?**

Answer:

If you want to protect your worksheet from being copied, go into Menu bar > Review > Protect sheet > Password.

By entering password, you can prevent your worksheet from getting copied.

There are three report formats available in Excel; they are:

1. Compact Form
2. Outline Form
3. Tabular Form

## **Q20. How are Pivot tables used to filter data in Excel?**

Answer:

You can filter data according to your requirements with Excel Pivot tables. Place the field on which you want the data to be filtered. Then open the drop-down list of the field you put in the Filter area from the pivot table and choose your line.

## **Q21: What is the purpose of using a pivot table in Excel?**

Answer:

Pivot tables are used to summarize, analyze, and visualize large datasets in a structured format.

They allow users to perform calculations, rearrange data, and generate customizable summaries to gain insights from the data.

## **Q22: How does conditional formatting enhance data visualization in Excel?**

Answer:

Conditional formatting applies formatting styles such as colors, data bars, and icons to cells based on specified criteria. This helps highlight important trends, patterns, or outliers in the data, making it easier for users to identify key insights briefly.

## **Q23: Can you explain how VLOOKUP function works in Excel?**

Answer:

VLOOKUP function searches for a value in the first column of a table array and returns a corresponding value from a specified column in the same row. It is commonly used for data retrieval, such as looking up employee names based on their ID numbers in a database.

## **Q24: How do you use macros to automate tasks in Excel?**

Answer:

Macros record a sequence of user actions and allow users to play back those actions to automate repetitive tasks. They can be created using the “Record Macro” feature, assigned to buttons or keyboard shortcuts, and executed to perform tasks like data cleaning, formatting, or analysis.

**Q25: What are some common data cleaning techniques in Excel?**

Answer:

Common data cleaning techniques include removing duplicates, sorting and filtering data, using text functions to manipulate strings, and converting data formats. These techniques help ensure data accuracy and consistency for further analysis.

**Q26: How can you perform statistical analysis in Excel?**

Answer:

Excel provides built-in functions for basic statistical analysis, such as calculating mean, median, standard deviation, and correlation coefficients. Additionally, users can use data analysis add-ins for more advanced statistical procedures like regression analysis or analysis of variance (ANOVA).

**Q27: What is the purpose of data validation in Excel?**

Answer:

Data validation ensures data integrity by restricting the type of data that can be entered into a cell. It can be used to create drop-down lists, apply custom validation rules, or set data limits, preventing entry of incorrect or invalid data into spreadsheets.

---

## POWER BI

### Q1. How to create filters in Power BI?

Answer:

Filters are an integral part of Power BI reports. They are used to slice and dice the data as per the dimensions we want.

Filters are created in a couple of ways.

**Using Slicers:** A slicer is a visual under Visualization Pane. This can be added to the design view to filter our reports. When a slicer is added to the design view, it requires a field to be added to it. For example- Slicer can be added for Country fields. Then the data can be filtered based on countries.

**Using Filter Pane:** The Power BI team has added a filter pane to the reports, which is a single space where we can add different fields as filters. These fields can be added depending on whether you want to filter only one visual (Visual level filter), all the visuals in the report page (Page level filters), or applicable to all the pages of the report (report level filters).

### Q2. How to sort data in Power BI?

Answer:

Sorting is available in multiple formats. In the data view, a common sorting option of alphabetical order is there. Apart from that, we have the option of Sort by column, where one can sort a column based on another column. The sorting option is available in visuals as well. Sort by ascending and descending option by the fields and measure present in the visual is also available.

### Q3. What is DAX in Power BI?

Answer:

DAX stands for Data Analysis Expressions. It's a collection of functions, operators, and constants used in formulas to calculate and return values. In other words, it helps you create new info from data you already have.

### Q4. Which two cross filter directions are available in Power BI table relationships?

Answer:

When a relationship is created between two different tables in Power BI, then the relationship asks us the cross-filter direction. There are two options available for cross-filtering.

Single – When the cross-filter direction is single, then the filtering between tables happens from left table to right table. It is the default setting. The first table can be used to filter the data in the second table.

Both – When the cross-filter direction is both, then the filtering between the tables will work in both ways. Either table can be used to filter the other table.

## Q5. What is Power BI and how do you use it?

Answer:

Power BI is a business analytics tool by Microsoft used to visualize and analyse data. It allows users to connect to various data sources, create interactive reports and dashboards, and share insights with others. It is used for data modelling, data transformation, and data visualization.

## Q6. How do you connect to data sources in Power BI?

Answer:

In Power BI, you can connect to data sources such as Excel files, SQL databases, SharePoint lists, online services like Salesforce, and many others. You can connect by selecting the data source type, providing connection details, and importing or connecting to the data.

## Q7. What is a measure in Power BI, and how do you create one?

Answer:

A measure in Power BI is a calculation applied to data in a report, such as sums, averages, counts, or custom calculations. Measures are created using the DAX (Data Analysis Expressions) language. To create a measure, you use the “New Measure” option in Power BI, enter the DAX expression, and give the measure a name.

## Q8. What are the differences between calculated columns and measures in Power BI?

Answer:

Calculated columns are calculated at the row level and stored in the data model, while measures are calculated based on the context of the visualization and are not stored in the data model. Calculated columns are static, while measures are dynamic and respond to filters and slicers applied to the report.

## Q9. Can you explain the concept of drill-through in Power BI and give an example?

Answer:

Drill-through in Power BI allows users to navigate from one report page to another to get more detailed information. For example, if you have a summary dashboard with sales data, users can drill through to a detailed sales report for a specific product category or time.

## Q10. How do you create a dashboard in Power BI, and what are the best practices to follow?

Answer:

To create a dashboard in Power BI, you can pin visualizations from different report pages to a dashboard canvas. Best practices include using consistent formatting, organizing content logically, using meaningful titles and descriptions, and keeping the dashboard focused on key metrics.

## **Q11. How do you use Power Query in Power BI to clean and transform data?**

Answer:

Power Query is used in Power BI to clean and transform data before loading it into the data model.

You can use Power Query to remove duplicates, filter rows, rename columns, create custom columns, pivot and unpivot data, and perform other data cleaning and transformation tasks.

## **Q12. What is DAX, and how do you use it in Power BI?**

Answer:

DAX (Data Analysis Expressions) is a formula language used in Power BI to create custom calculations, measures, and calculated columns. DAX functions are used to perform calculations such as aggregations, comparisons, filtering, and more.

## **Q13. Can you explain the difference between a table and a matrix visual in Power BI?**

Answer:

A table visual in Power BI displays data in a tabular format with rows and columns, while a matrix visual is like a pivot table and allows you to summarize data by rows and columns, with subtotals and totals.

## **Q14. How do you use Power BI to create a report that updates automatically based on new data?**

Answer:

In Power BI, you can set up data refresh schedules to automatically update your reports and dashboards with new data from the data source. You can configure data refresh settings in the Power BI service or use DirectQuery mode to query data in real-time.

## **Q15. How do you share your reports and dashboards with others in Power BI?**

Answer:

- Use data modelling techniques like data summarization, aggregation, and filtering to reduce the dataset size.
- Employ Power BI's query folding feature to push data transformation operations back to the data source, reducing the amount of data loaded into memory.
- Utilize incremental data loading to only bring in new or updated data instead of loading the entire dataset each time.
- Implement data compression techniques to reduce the storage footprint of the dataset within Power BI.
- Consider using DirectQuery mode or connecting to a live dataset for real-time analysis without importing large datasets into Power BI.

## Q16. How do you handle large datasets in Power BI?

Answer:

- An example could involve multiple fact tables and dimension tables with intricate relationships, such as a retail analytics model with sales, inventory, customer demographics, and product hierarchies.
- Complex DAX calculations and measures may be necessary to derive meaningful insights, such as year-over-year comparisons, cohort analysis, or predictive analytics.

## Q17. Can you give an example of a complex data model you have built in Power BI?

Answer:

- Set up scheduled refreshes to automatically update datasets at defined intervals, ensuring data remains current.
- Configure refresh options such as incremental refresh, which updates only new or modified data, reducing refresh time for large datasets.
- Monitor refresh performance and optimize queries to improve efficiency.
- Utilize Power BI Gateway to refresh on-premises data sources securely.
- Explore options for refreshing datasets manually or on-demand when immediate updates are required.

## Q18. How do you handle data refreshes in Power BI, and what options are available?

Answer:

- Implement role-based security to control access to specific datasets, reports, or dashboards based on user roles or groups.
- Utilize row-level security (RLS) to restrict data access at the row level based on user identity or attributes.
- Secure sensitive data by encrypting data at rest and in transit, leveraging Power BI's integration with Azure Active Directory (AAD) for authentication and authorization.
- Ensure compliance with data privacy regulations by applying data masking or anonymization techniques where necessary.

## Q19. How do you handle security and permissions in Power BI?

Answer:

- **Data volume limitations:** Employ data modeling techniques to reduce dataset size and utilize incremental refresh for large datasets.
- **Data source connectivity:** Use custom connectors or consider using Power BI Gateway to connect to on-premises data sources.
- **Performance issues:** Optimize DAX calculations, minimize visuals and queries, and partition large datasets for better performance.
- **Visualization limitations:** Explore custom visuals from AppSource or build custom visuals using the Power BI SDK to extend visualization capabilities.

## **Q20. What are some limitations of Power BI, and how would you work around them?**

Answer:

- **Data volume limitations:** Employ data modeling techniques to reduce dataset size and utilize incremental refresh for large datasets.
- **Data source connectivity:** Use custom connectors or consider using Power BI Gateway to connect to on-premises data sources.
- **Performance issues:** Optimize DAX calculations, minimize visuals and queries, and partition large datasets for better performance.
- **Visualization limitations:** Explore custom visuals from AppSource or build custom visuals using the Power BI SDK to extend visualization capabilities.

## **Q21. How do you create a drill-through report in Power BI?**

Answer:

- Define drill-through actions on report elements (e.g., tables, matrices) to navigate to another report page or URL when users interact with those elements.
- Designate fields as drill-through filters to pass selected values from one report page to another for deeper analysis.
- Create a separate drill-through page with detailed information related to the selected data point, allowing users to explore specific details without cluttering the main report.

## **Q22. How do you create a custom visualization in Power BI?**

Answer:

- Use Power BI custom visuals from AppSource or develop custom visuals using the Power BI Visuals SDK and TypeScript or R language.
- Customize visual appearance, behaviour, and interactivity to meet specific requirements or address unique data visualization needs.
- Leverage community-contributed custom visuals or build proprietary custom visuals tailored to organizational requirements.

## **Q23. How do you optimize the performance of a Power BI report?**

Answer:

- Minimize data model complexity by removing unnecessary columns, tables, or relationships.
- Optimize DAX calculations and measures for efficiency, avoiding complex or nested expressions where possible.
- Reduce the number of visuals on a report page to improve rendering speed.
- Limit the use of custom visuals, particularly those with heavy computational overhead.
- Opt for DirectQuery or live connection mode for real-time data analysis, especially for large datasets requiring frequent updates.

## Q24. Can you explain the difference between DirectQuery and Import modes in Power BI?

Answer:

- **Import Mode:** Data from the data source is imported and stored within the Power BI dataset. Queries and calculations operate on this imported data.
- **DirectQuery Mode:** Power BI connects directly to the data source at query time, retrieving data dynamically without importing it. This mode enables real-time or near-real-time analysis but may impact performance depending on the data source and query complexity.

## Q25. Steps to create powerful visualization in Power BI:

Answer:

### Discovering Custom Visualizations:

- **Ditching the Ordinary:** Say goodbye to standard charts; Power BI's custom visualizations offer unique options.
- **Marketplace Exploration:** Dive into the extensive Power BI marketplace, where numerous custom visuals await experimentation.
- **Precision in Selection:** Choose visuals that fit seamlessly with your data and storytelling goals.

### Unveiling the Power of Advanced Analytics:

- **Beyond Basic Charts:** Explore advanced features like forecasting, clustering, and key influencers for a more sophisticated analysis.
- **Insightful Revelations:** Delve deeper into your data, revealing hidden patterns and relationships that basic charts might miss.
- **Holistic Data Understanding:** Use advanced analytics tools to gain a nuanced and comprehensive understanding of your dataset.

### Elevating Reports with Advanced Formatting and Interaction:

- **Crafting Visual Masterpieces:** Hone your formatting skills to transform your reports into visually stunning presentations.
- **Spotlight on Significance:** Emphasize critical data points using conditional formatting techniques.
- **User-Centric Interaction:** Improve user experience through interactive features such as drill-through options and report tooltips.
- **Artistry in Design:** Infuse creativity into your reports with thoughtful choices in color schemes, fonts, and layout designs.

### The Continuous Journey of Learning:

- **Staying Curious:** Approach Power BI visualization as an ongoing learning journey, always seeking new techniques and approaches.
- **Community Engagement:** Connect with the Power BI community for insights, tips, and collaborative learning.

- **Real-World Application:** Apply your newfound skills to real-world scenarios, refining your expertise through practical experience.

Questions on formulas like: SUMIFS, COUNTIFS, LOOKUPS, INDEX & MATCH, AVERAGEIFS. Plus, some basic questions on pivot tables, conditional formatting, data validation, and charts.

## **Q26. How are nested IF statements used in Excel?**

Answer:

The function IF() can be nested when we have multiple conditions to meet. The FALSE value in the first IF function is replaced by another IF function to make a further test.

## **Q27. What is x-velocity in Power Pivot?**

Answer:

X-Velocity is the in-memory analytics engine behind Power Pivot that loads and handles huge data in Power BI. It stores data in columnar storage that results in faster processing.

## **Q28. How to calculate the average in Power BI?**

Answer:

Average can be calculated in two ways-one is when we add a measure to a visual; by default, it summarizes any measure. When we click on the drop-down for the measure- we can change from Sum to Average. This gives us an average.

The second one is creating a calculated measure for average using the AVERAGE() DAX function.

## **Q29. Where is the data stored in Power BI?**

Answer:

The nation or area for the business identity is chosen by the first user in your organization who signs up for Power BI or Microsoft 365. The cloud's shared identity and access management service, Azure Active Directory (AAD), establishes a tenant in the data center region nearest to the chosen nation or area. AAD is a multi-tenant service, with each enterprise represented in the data center as a separate tenant. The data is saved in the area you choose at sign-up. This region will be the same for all users in your organization, regardless of their location. The chosen region should, ideally, be in the same geographical area as most of your consumers.

---

## Tableau

### Q1. What are groups in Tableau?

Answer:

A group is a combination of dimension members that make higher level categories. For example, if you are working with a view that shows average test scores by major, you may want to group certain majors together to create major categories.

### Q2. How to add the logo in the tableau dashboard?

Answer:

In the objects pane in dashboard there is an option to import image. Make sure that you have selected the floating type instead of by default tiled option. Drag and drop image object to dashboard and then select the logo from your system saved in .jpg, .png or .jpeg format. To use it as a background user can use the small drop-down option on the right side of the image and select send to back and then increase the size of image to make it as a watermark logo.

### Q3. What do Tableau's sets and groups mean?

Answer:

Data is grouped using sets and groups according to predefined criteria. The primary distinction between the two is that although a set can have only two options—either in or out—a group can divide the dataset into several groups. A user should decide which group or sets to apply based on the conditions.

### Q4. What is Tableau?

Answer:

Tableau is a powerful data visualization tool that allows users to create interactive and shareable dashboards, reports, and visualizations from various data sources. It enables users to explore and analyze data visually to gain actionable insights.

### Q5. What are the different products offered by Tableau?

Answer:

Tableau offers various products including Tableau Desktop, Tableau Server, Tableau Online, Tableau Public, and Tableau Prep Builder.

### Q6. Explain Tableau Desktop.

Answer:

Tableau Desktop is a data visualization tool used to create interactive dashboards, reports, and visualizations. It allows users to connect to different data sources, prepare and clean data, and create visualizations using drag-and-drop functionality.

## **Q7. How do you connect to data sources in Tableau?**

Answer:

In Tableau, you can connect to data sources such as Excel, CSV files, databases (like SQL Server, MySQL), cloud services (like Google Analytics, Salesforce), and more. You can connect to data sources using the “Connect to Data” option and selecting the appropriate connector.

## **Q8. What is Tableau Server?**

Answer:

Tableau Server is an enterprise-level platform that allows users to share, collaborate, and manage Tableau workbooks and dashboards securely within an organization. It provides centralized data governance, security, and scalability.

## **Q9. Explain the difference between Tableau Desktop and Tableau Server.**

Answer:

Tableau Desktop is a desktop application used to create visualizations and dashboards, while Tableau Server is a web-based platform used to share, collaborate, and manage Tableau content within an organization. Tableau Desktop is used by analysts and data scientists to create visualizations, whereas Tableau Server is used by business users to access and interact with these visualizations through a web browser.

## **Q10. What is Tableau Public?**

Answer:

Tableau Public is a free version of Tableau that allows users to create and share visualizations and dashboards publicly on the web. It is often used by individuals, students, and journalists to create interactive data stories and visualizations.

## **Q11. What is a dashboard in Tableau?**

Answer:

A dashboard in Tableau is a collection of visualizations (such as charts, graphs, and maps) and other objects (such as images and web pages) arranged on a single canvas. Dashboards allow users to view and interact with multiple visualizations simultaneously to gain insights from their data.

## **Q12. How do you create calculated fields in Tableau?**

Answer:

In Tableau, you can create calculated fields using formulas to perform calculations or manipulate data. You can create calculated fields by right-clicking in the data pane and selecting “Create Calculated Field” or by clicking on the drop-down arrow next to a field in the data pane and selecting “Create Calculated Field.”

## **Q13. What are the different types of joins supported in Tableau?**

Answer:

Tableau supports different types of joins including inner join, left join, right join, and full outer join. These joins allow users to combine data from multiple tables based on a common field or key.

These questions cover a range of topics related to Tableau and can help you prepare for interviews effectively. It's essential to understand the concepts thoroughly and be able to demonstrate your knowledge and skills with practical examples.

## **Q14. What is data blending in Tableau?**

Answer:

Data blending is a feature in Tableau that allows users to combine data from multiple data sources within a single visualization. It enables users to create relationships between different datasets based on common dimensions and blend the data together seamlessly.

## **Q15. Explain the concept of LOD (Level of Detail) expressions in Tableau.**

Answer:

LOD expressions in Tableau allow users to compute aggregations at different levels of granularity independently of the visualization's level of detail. There are three types of LOD expressions: FIXED, INCLUDE, and EXCLUDE. These expressions enable users to control the level of aggregation in their visualizations and perform complex calculations.

## **Q16. What are Tableau actions?**

Answer:

Tableau actions are interactive features that allow users to create dynamic and linked dashboards. Actions enable users to filter, highlight, or navigate between different visualizations based on user interactions such as clicking, hovering, or selecting data points.

## **Q17. How do you optimize Tableau performance for large datasets?**

Answer:

To optimize Tableau performance for large datasets, you can:

Aggregate data at the data source level before importing it into Tableau.

Limit the number of marks displayed in visualizations by applying filters or using aggregated extracts.

Use data extracts to improve query performance.

Optimize workbook design by reducing the number of worksheets and using efficient data blending and joins.

## **Q18. Explain the concept of data densification in Tableau.**

Answer:

Data densification is a process in Tableau where missing data points are filled in to create a continuous axis or grid in visualizations. Tableau automatically densifies data to ensure that visualizations are displayed accurately, even when data is sparse or missing.

## **Q19. What are the different types of Tableau maps?**

Answer:

Tableau supports different types of maps including symbol maps, filled maps, and density maps.

Symbol maps represent data points using symbols or markers, filled maps display data using color shading, and density maps show the concentration or density of data points using color gradients.

## **Q20. How do you implement advanced analytics in Tableau?**

Answer:

Tableau integrates with R, Python, and other statistical tools to perform advanced analytics and predictive modeling. Users can leverage Tableau's integration with these tools to create calculated fields, predictive models, and statistical analyses directly within Tableau visualizations.

## **Q21. Explain the use of Tableau's parameter feature.**

Answer:

Tableau parameters allow users to create dynamic and interactive dashboards by enabling users to change certain values interactively. Parameters can be used to control various aspects of visualizations such as filter values, aggregation methods, and reference lines dynamically.

## **Q22. What are Tableau extensions?**

Answer:

Tableau extensions are custom-built add-ons that enhance Tableau's functionality by providing additional features and integrations. Extensions can be used to integrate with third-party applications, add new visualizations, or extend Tableau's capabilities.

## **Q23. How do you deploy and manage Tableau dashboards in an enterprise environment?**

Answer:

Tableau Server or Tableau Online can be used to deploy and manage Tableau dashboards in an enterprise environment. These platforms provide centralized administration, security, and governance features to manage user access, permissions, and content distribution effectively.

## **Q24. What is Tableau Prep Builder, and how does it differ from Tableau Desktop?**

Answer:

Tableau Prep Builder is a data preparation tool that allows users to visually and interactively clean, shape, and combine data for analysis. It is used for data preparation tasks such as cleaning, reshaping, and blending data from multiple sources. Tableau Desktop, on the other hand, is primarily used for data visualization and analysis.

## **Q25. Explain Tableau Extracts and their benefits.**

Answer:

Tableau Extracts are a snapshot of data that is optimized for analysis and storage in Tableau. Extracts can improve performance by reducing the amount of data transferred from the data source to Tableau and by enabling offline access to data. Extracts also allow for faster analysis and visualization of large datasets.

## **Q26. What is Tableau Data Server, and how does it differ from Tableau Server?**

Answer:

Tableau Data Server is a component of Tableau Server that stores and manages Tableau data sources centrally within an organization. It allows users to publish and share data sources securely across teams and departments. Tableau Server, on the other hand, is a web-based platform used to share, collaborate, and manage Tableau workbooks and dashboards.

## **Q27. Explain Tableau's Level of Detail (LOD) expressions.**

Answer:

Level of Detail (LOD) expressions in Tableau allow users to compute aggregations at different levels of granularity within a visualization. LOD expressions enable users to control the level of detail at which calculations are performed, providing more flexibility and granularity in analysis.

## **Q28. How do you optimize Tableau dashboards for performance?**

Answer:

To optimize Tableau dashboards for performance, you can take several steps such as minimizing the number of visualizations, reducing the complexity of calculations, using extracts instead of live connections, optimizing data source queries, and leveraging Tableau Server features such as caching and incremental refreshes.

## **Q29. What are the different ways to share Tableau visualizations with others?**

Answer:

Tableau provides several ways to share visualizations including publishing to Tableau Server or Tableau Online, embedding visualizations in web pages or applications, exporting visualizations as static images or PDFs, and sharing Tableau Public visualizations publicly on the web.

## **Q30. Explain Tableau's geospatial capabilities and how they can be used in data visualization.**

Answer:

Tableau's geospatial capabilities allow users to create maps and visualize spatial data such as geographic locations, boundaries, and routes. Users can plot latitude and longitude coordinates on maps, create custom geocoding, perform spatial analysis, and visualize spatial relationships in data.

## **Q31. How do you handle large datasets in Tableau?**

Answer:

To handle large datasets in Tableau, you can use techniques such as data blending, data extracts, incremental refreshes, and server-level optimizations. Additionally, you can use filters, aggregations, and parameters to reduce the amount of data displayed in visualizations and improve performance.

## **Q32. Explain Tableau's integration with R and Python.**

Answer:

Tableau integrates with R and Python to allow users to perform advanced analytics and statistical modelling directly within Tableau. Users can use R and Python scripts to create calculated fields, generate predictions, perform clustering, and create custom visualizations using Tableau's analytics extension API.

## **Q33. How do you approach storytelling with data in Tableau?**

Answer:

Storytelling with data in Tableau involves using visualizations and narratives to communicate insights effectively. To create compelling data stories, you can structure your dashboards with a logical flow, use annotations and captions to provide context, and incorporate storytelling techniques such as framing, conflict, and resolution to engage your audience.

## **Q34. Explain the concept of Tableau Actions.**

Answer:

Tableau Actions allow users to create interactive dashboards by enabling actions such as filtering, highlighting, URL actions, and parameter actions. These actions allow users to interact with visualizations dynamically, drill down into data, and navigate between dashboards.

## **Q35. What are Tableau Parameters, and how can they be used in visualizations?**

Answer:

Tableau Parameters are dynamic values that users can control to modify aspects of a visualization, such as filters, calculations, and reference lines. Parameters enable users to create interactive and flexible visualizations that allow for exploration and analysis.

## **Q36. Explain the concept of Tableau Sets and their applications.**

Answer:

Tableau Sets are custom fields that group data based on specified conditions or criteria. Sets allow users to create subsets of data dynamically and perform analysis on these subsets. Sets can be used for segmentation, cohort analysis, outlier detection, and creating custom groupings.

**Q37. What is Tableau's Table Calculation feature, and how does it differ from regular calculations?**

Answer:

Tableau's Table Calculation feature allows users to perform calculations on the results of a visualization rather than on the underlying data. Table calculations are computed based on the visualization's layout, including filters, sorting, and grouping. They differ from regular calculations, which are computed on the raw data before visualization.

**Q38. Explain the concept of Data Blending in Tableau.**

Answer:

Data Blending in Tableau is a technique used to combine data from multiple sources in a single visualization. It allows users to blend data from different data sources on a common field or key. Data blending is useful when data resides in separate data sources and needs to be combined for analysis.

**Q39. How do you handle missing or null values in Tableau?**

Answer:

In Tableau, missing or null values can be handled using techniques such as data cleaning, data imputation, filtering out null values, replacing null values with default values, or using calculated fields to handle missing values dynamically.

**Q40. Explain the concept of Data Governance in Tableau and its importance.**

Answer:

Data Governance in Tableau refers to the processes and policies for managing and ensuring the quality, security, and integrity of data used in Tableau. It involves establishing standards, roles, and responsibilities for data management, ensuring compliance with regulations, and maintaining data privacy and security.

**Q41. What are the different deployment options for Tableau Server?**

Answer:

Tableau Server can be deployed on-premises, in the cloud (Tableau Online), or in a hybrid environment. On-premises deployment involves installing Tableau Server on local servers within an organization's infrastructure, while Tableau Online is a cloud-based deployment managed by Tableau. Hybrid deployment combines on-premises and cloud-based components.

**Q42. How do you monitor and optimize performance in Tableau Server?**

Answer:

To monitor and optimize performance in Tableau Server, you can use tools such as Server Monitoring, Performance Recording, and Tableau Server Administrator Views. Additionally, you can optimize server settings, manage extract schedules, and monitor resource usage to improve performance.

**Q43. What are the best practices for designing effective Tableau dashboards?**

Answer:

Best practices for designing effective Tableau dashboards include designing for the end user, keeping dashboards simple and focused, using consistent formatting and design principles, optimizing performance, providing interactivity and context, and iterating based on user feedback.

**Q44. Explain the concept of LOD Expressions in Tableau and provide examples of when they might be used.**

Answer:

Level of Detail (LOD) Expressions allow users to compute aggregations at different levels of granularity in Tableau. They can be used to perform calculations independent of the visualization's level of detail or to aggregate data at a specific level of granularity. For example, you can use FIXED LOD expressions to calculate the total sales for a specific category, regardless of other dimensions in the visualization.

**Q45. What is the difference between blending and joining data in Tableau? When would you use each technique?**

Answer:

Data blending in Tableau involves combining data from different data sources on a common field or key, without modifying the original data sources. Joining data, on the other hand, combines data from different tables within the same data source based on a common field. Blending is typically used when data resides in separate data sources, while joining is used when data resides in the same data source.

**Q46. Explain the concept of spatial joins in Tableau and provide an example of when you might use them.**

Answer:

Spatial joins in Tableau involve combining spatial data from one dataset with attribute data from another dataset based on a spatial relationship, such as proximity or containment. For example, you might use a spatial join to combine a dataset of store locations with a dataset of customer addresses to determine which customers are closest to each store location.

**Q47. What are the different types of filters in Tableau, and when would you use each type?**

Answer:

Tableau offers several types of filters, including Extract Filters, Data Source Filters, Context Filters, Dimension Filters, Measure Filters, Top N Filters, and Relative Date Filters. The choice of filter depends on the specific filtering requirements and the level at which the filter should be applied.

**Q48. Explain the concept of dashboard actions in Tableau and provide examples of how they can be used to create interactive dashboards.**

Answer:

Dashboard actions in Tableau allow users to create interactive dashboards by enabling actions such as filtering, highlighting, URL actions, and parameter actions. For example, you can use a dashboard action to filter multiple visualizations based on user selections in one visualization, creating a dynamic and interactive user experience.

**Q49. How does Tableau handle security and permissions for data access?**

Answer:

Tableau provides several security features, including user authentication, data encryption, row-level security, and permission-based access controls. Administrators can define user roles and permissions to control access to data sources, workbooks, and dashboards, ensuring that sensitive data is only accessible to authorized users.

**Q50. Explain the concept of performance optimization in Tableau and provide examples of techniques for improving dashboard performance.**

Answer:

Performance optimization in Tableau involves identifying and addressing bottlenecks that can impact dashboard performance, such as slow data connections, inefficient queries, and complex calculations. Techniques for improving performance include using data extracts, optimizing data source connections, minimizing the number of visualizations, and leveraging server-level optimizations.

**Q51. What are the limitations of Tableau Public, and when might you choose to use Tableau Desktop instead?**

Answer:

Tableau Public is a free version of Tableau that allows users to create and share visualizations publicly on the web. However, it has limitations such as data size restrictions, limited connectivity to data sources, and lack of support for certain advanced features. Tableau Desktop, on the other hand, is a full-featured version of Tableau that offers more advanced functionality and flexibility, making it suitable for professional use cases.

**Q52. How does Tableau integrate with other data visualization and analytics tools?**

Answer:

Tableau integrates with a variety of data visualization and analytics tools, including R, Python, MATLAB, and SAS. These integrations allow users to leverage advanced analytics and statistical modeling capabilities within Tableau, enhancing the depth and sophistication of their analyses.

## **Q53. Can you explain the process of publishing Tableau workbooks to Tableau Server, and how permissions are managed?**

Answer:

Publishing Tableau workbooks to Tableau Server involves saving the workbook to the Tableau Server repository, where it can be accessed by other users with appropriate permissions. Administrators can manage permissions at the workbook, project, and server levels, controlling who can view, edit, or publish content on the server.

## **Q54. What is difference between Power BI and Tableau?**

Answer:

- Tableau is known for its intuitive user interface and powerful data visualization capabilities. It offers a wide range of charts and graphs, as well as drag-and-drop functionality that makes it easy to create interactive dashboards. Tableau is also known for its ability to handle large amounts of data, making it a great choice for big data analytics.
- Power BI, on the other hand, is known for its seamless integration with other Microsoft products, such as Excel and SharePoint. It offers a similar range of data visualization options as Tableau, as well as advanced data modeling capabilities. Power BI also offers robust collaboration features, making it a great choice for teams working on data analysis projects.
- In terms of cost, Power BI is often seen as more affordable than Tableau, particularly for small to medium-sized businesses.
- Ultimately, the choice between Tableau and Power BI will depend on your specific needs and preferences, as well as your budget. If you value ease-of-use and powerful visualization capabilities, Tableau might be the better choice. If you value collaboration and integration with other Microsoft products, Power BI might be the better choice.

## **Q55. How to increase size of pie in tableau?**

Answer:

Creating a pie chart requires atleast one measure and one dimension in row and shelf column. Then you can select a pie chart from either the show me option on the right side of the screen or from the marks card change automatic to pie. Then give some detailing to the pie chart by using a dimension in colour and measure in angle. Option to increase the size also comes under marks card. Click on the size option and then move the slider towards right to increase its size.

## **Q56. What are groups in Tableau?**

Answer:

A group is a combination of dimension members that make higher level categories. For example, if you are working with a view that shows average test scores by major, you may want to group certain majors together to create major categories.

## Q57. What Are the Data Types Supported in Tableau?

Answer:

Following data types are supported in Tableau:

Text (string) values

Date values

Date and time values

Numerical values

Boolean values (relational only)

Geographical values (used with maps)

## Q58. How Do You Calculate the Daily Profit Measures Using LOD?

Answer:

LOD expressions allow us to easily create bins on aggregated data such as profit per day.

Scenario: We want to measure our success by the total profit per business day.

Create a calculated field named LOD - Profit per day and enter the formula:

`FIXED [Order Date] : SUM ([Profit])`

Create another calculated field named LOD - Daily Profit KPI and enter the formula:

`IF [LOD - Profit per day] > 2000 then "Highly Profitable."`

`ELSEIF [LOD - Profit per day] <= 0 then "Unprofitable"`

`ELSE "Profitable"`

`END`

To calculate daily profit measure using LOD, follow these steps to draw the visualization:

Bring YEAR(Order Date) and MONTH(Order Date) to the Columns shelf

Drag Order Id field to Rows shelf. Right-click on it, select Measure and click on Count(Distinct)

Drag LOD - Daily Profit KPI to the Rows shelf

Bring LOD - Daily Profit KPI to marks card and change mark type from automatic to area.

## Q59. What is the Difference Between Joining and Blending in Tableau?

Answer:

Combining the data from two or more different sources is data blending, such as Oracle, Excel, and SQL Server. In data blending, each data source contains its own set of dimensions and measures. Combining the data between two or more tables or sheets within the same data source is data joining. All the combined tables or sheets contain a common set of dimensions and measures.

## **Q60. How is the grid search parameter different from the random search?**

Answer:

Model Hyperparameter tuning is very useful to enhance the performance of a machine learning model. The only difference between both the approaches is in grid search we define the combinations and do training of the model whereas in RandomizedSearchCV the model selects the combinations randomly. Both are very effective ways of tuning the parameters that increase the model generalizability. Random search is a technique where random combinations of the hyperparameters are used to find the best solution for the built model. The drawback of random search is that it yields high variance during computing. Since the selection of parameters is completely random; and since no intelligence is used to sample these combinations, luck plays its role.

## **Q61. How should you maintain a deployed model?**

Answer:

A deployed model needs to be retrained after a while to improve the performance of the model. Since deployment, a track should be kept of the predictions made by the model and the truth values. Later this can be used to retrain the model with the new data. Also, root cause analysis for wrong predictions should be done.

## **Q62. Explain the Difference Between Tableau Worksheet, Dashboard, Story, and Workbook?**

Answer:

Tableau uses a workbook and sheet file structure, much like Microsoft Excel.  
A workbook contains sheets, which can be a worksheet, dashboard, or a story.  
A worksheet contains a single view along with shelves, legends, and the Data pane.  
A dashboard is a collection of views from multiple worksheets.  
A story contains a sequence of worksheets or dashboards that work together to convey information.

## **Q63. How to create a calculated field in Tableau?**

Answer:

Click the drop down to the right of Dimensions on the Data pane and select “Create > Calculated Field” to open the calculation editor.  
Name the new field and create a formula.

## **Q64. What is the difference between joining and blending in Tableau?**

Answer:

Joining term is used when you are combining data from the same source, for example, worksheet in an Excel file or tables in Oracle database While blending requires two completely defined data sources in your report.

## **Q65. What are sets in Tableau?**

Answer:

Sets are custom fields that define a subset of data based on some conditions. A set can be based on a computed condition, for example, a set may contain customers with sales over a certain threshold. Computed sets update as your data changes. Alternatively, a set can be based on specific data point in your view.

## **Q66. What do Tableau's sets and groups mean?**

Answer:

Data is grouped using sets and groups according to predefined criteria. The primary distinction between the two is that although a set can have only two options—either in or out—a group can divide the dataset into several groups. A user should decide which group or sets to apply based on the conditions.

## **Q67. What Would You Do If Some Countries/Provinces (Any Geographical Entity) are Missing and Displaying a Null When You Use Map View in Tableau?**

Answer:

When working with maps and geographical fields, unknown or ambiguous locations are identified by the indicator in the lower right corner of the view.

Click the indicator and choose from the following options:

Edit Locations - correct the locations by mapping your data to known locations

Filter Data - exclude the unknown locations from the view using a filter. The locations will not be included in calculations

Show Data at Default Position - show the values at the default position of (0, 0) on the map.

## **Q68. What is the difference between joining and blending in Tableau?**

Answer:

Joining term is used when you are combining data from the same source, for example, worksheet in an Excel file or tables in Oracle database. While blending requires two completely defined data sources in your report.

---

## Database

**Q1. Write a sql query to get emp id and department for each department who recently joined the organization and still in working.**

**table name - employee**

**columns - emp id, first name, last name, date of join, date of exit, department.**

Answer:

```
SELECT emp_id, department
FROM employee
WHERE date_of_join = (SELECT MAX(date_of_join) FROM employee WHERE department =
e.department AND date_of_exit IS NULL)
```

**Q2. Write a sql query to get the days which is having temperature greater than previous day**

**table name - A**

**columns - temp , days**

Answer:

```
SELECT days
FROM (
  SELECT days, temp,
    LAG(temp) OVER (ORDER BY days) AS prev_temp
  FROM A
) AS subquery
WHERE temp > prev_temp;
```

**Q3. Write a sql query to get the sum of earnings of each employee excluding the last month salary.**

**table name - A**

**columns - Employee, month, earning**

Answer:

```
SELECT Employee, SUM(earning) AS Total_Earnings
FROM A
WHERE month <> (SELECT MAX(month) FROM A)
GROUP BY Employee;
```

**Q4. Outputs of Left, Right, Inner and outer join from two tables having duplicate rows.**

Answer:

For Example:-

Left table A-

Column

1  
1  
1  
2  
2  
3  
4  
5

Right table B-

Column

1  
1  
2  
2  
2  
3  
3  
3  
4

Tell the no. of rows as output of all joins for above table A &amp; B

**Q5. Write a query to get mean, median and mode for earning?**

table - A

columns - emp, earning

Answer:

SELECT

AVG(earning) AS mean\_earning,

PERCENTILE\_CONT(0.5) WITHIN GROUP (ORDER BY earning) AS median\_earning,

MODE() WITHIN GROUP (ORDER BY earning) AS mode\_earning

FROM

A;

**Q6: Consider a table named “Orders” with the following columns: OrderID, OrderDate, CustomerID. Write a SQL query to find the customers who have placed orders on consecutive days.**

Answer:

```
SELECT DISTINCT o1.CustomerID
FROM Orders o1
JOIN Orders o2 ON o1.CustomerID = o2.CustomerID
AND DATEDIFF(o2.OrderDate, o1.OrderDate) = 1;
```

**Q7: You have three tables: “Employees,” “Departments,” and “Salaries.” The “Employees” table has the following columns: EmployeeID, EmployeeName, DepartmentID. The “Departments” table has the following columns: DepartmentID, DepartmentName. The “Salaries” table has the following columns: EmployeeID, Salary, EffectiveDate. Write a SQL query to retrieve the employee who has had the highest salary increase within the last year, along with their name, department, and the percentage increase.**

Answer:

```
SELECT e.EmployeeName, d.DepartmentName,
((MAX(s.Salary) - MIN(s.Salary)) / MIN(s.Salary)) * 100 AS SalaryIncreasePercentage
FROM Employees e
JOIN Departments d ON e.DepartmentID = d.DepartmentID
JOIN Salaries s ON e.EmployeeID = s.EmployeeID
WHERE s.EffectiveDate >= DATEADD(year, -1, GETDATE())
GROUP BY e.EmployeeName, d.DepartmentName
ORDER BY SalaryIncreasePercentage DESC
LIMIT 1;
```

**Q8: You have a table named “Logs” with the following columns: LogID, LogTime, UserID. Write a SQL query to find the top 5 users who have logged in the most consecutive days, along with the number of consecutive days.**

Answer:

```
SELECT UserID, COUNT(*) AS ConsecutiveDays
FROM (
  SELECT UserID,
  DATE(LogTime) - ROW_NUMBER() OVER (PARTITION BY UserID ORDER BY LogTime) AS Grp
  FROM Logs
) AS Subquery
GROUP BY UserID, Grp
ORDER BY COUNT(*) DESC
LIMIT 5;
```

**Q9:** Consider a table named “Transactions” with the following columns: TransactionID, TransactionDate, Amount, UserID. Write a SQL query to calculate the average transaction amount for each user, including users who have no transactions, and display the result as zero for those users.

Answer:

```
SELECT UserID,
       COALESCE(AVG(Amount), 0) AS AverageTransactionAmount
  FROM Transactions
 GROUP BY UserID;
```

**Q10:** You have two tables: “Customers” and “Purchases.” The “Customers” table has the following columns: CustomerID, CustomerName. The “Purchases” table has the following columns: PurchaseID, PurchaseDate, CustomerID, ProductID. Write a SQL query to find the customers who have purchased all products.

Answer:

```
SELECT c.CustomerID, c.CustomerName
  FROM Customers c
 WHERE NOT EXISTS (
   SELECT p.ProductID
     FROM Products p
 WHERE NOT EXISTS (
   SELECT *
     FROM Purchases pu
    WHERE pu.CustomerID = c.CustomerID
      AND pu.ProductID = p.ProductID
  )
 );
 );
```

**Q11. What is concurrency control in DBMS?**

Answer:

This is a process of managing simultaneous operations in a database so that database integrity is not compromised. The following are the two approaches involved in concurrency control:

- Optimistic approach – Involves versioning
- Pessimistic approach – Involves locking

## **Q12. What is a checkpoint in DBMS and when does it occur?**

Answer:

A checkpoint is a mechanism where all the previous logs are removed from the system and are permanently stored on the storage disk. So, basically, checkpoints are those points from where the transaction log record can be used to recover all the committed data up to the point of crash.

## **Q13. List the different types of relationships in SQL.**

Answer:

One-to-One - This can be defined as the relationship between two tables where each record in one table is associated with the maximum of one record in the other table.

One-to-Many & Many-to-One - This is the most used relationship where a record in a table is associated with multiple records in the other table.

Many-to-Many - This is used in cases when multiple instances on both sides are needed for defining a relationship.

Self-Referencing Relationships - This is used when a table needs to define a relationship with itself.

## **Q14. What is the case when statement in SQL?**

Answer:

The CASE expression goes through conditions and returns a value when the first condition is met (like an if-then-else statement). So, once a condition is true, it will stop reading and return the result. If no conditions are true, it returns the value in the ELSE clause. If there is no ELSE part and no conditions are true, it returns NULL.

## **Q15. Explain the update statement in SQL.**

Answer:

The SQL UPDATE Query is used to modify the existing records in a table. You can use the WHERE clause with the UPDATE query to update the selected rows, otherwise all the rows would be affected.

Syntax

The basic syntax of the UPDATE query with a WHERE clause is as follows –

UPDATE table\_name

SET column1 = value1, column2 = value2..., columnN = valueN

WHERE [condition];

## **Q16. What are the two types of cursors?**

Answer:

There are 2 types of Cursors: Implicit Cursors, and Explicit Cursors. These are explained as following below.

### **Implicit Cursors:**

Implicit Cursors are also known as Default Cursors of SQL SERVER. These Cursors are allocated by SQL SERVER when the user performs DML operations.

### **Explicit Cursors :**

Explicit Cursors are Created by Users whenever the user requires them. Explicit Cursors are used for Fetching data from Table in Row-By-Row Manner.

## **Q17. What are the aggregate functions in SQL?**

Answer:

In database management an aggregate function is a function where the values of multiple rows are grouped together as input on certain criteria to form a single value of more significant meaning.

### Various Aggregate Functions

1. Count()
2. Sum()
3. Avg()
4. Min()
5. Max()

## **Q18. What are the advantages and disadvantages of views in the database?**

Answer:

### Advantages of Views:

As there is no physical location where the data in the view is stored, it generates output without wasting resources.

Data access is restricted as it does not allow commands like insertion, updation, and deletion.

### Disadvantages of Views:

The view becomes irrelevant if we drop a table related to that view.

Much memory space is occupied when the view is created for large tables.

## **Q19. What is a Recursive Stored Procedure?**

Answer:

A stored procedure that calls itself until a boundary condition is reached, is called a recursive stored procedure. This recursive function helps the programmers to deploy the same set of code several times as and when required.

## **Q20. What are the advantages and disadvantages of views in the database?**

Answer:

Advantages of Views:

As there is no physical location where the data in the view is stored, it generates output without wasting resources.

Data access is restricted as it does not allow commands like insertion, updation, and deletion.

Disadvantages of Views:

The view becomes irrelevant if we drop a table related to that view.

Much memory space is occupied when the view is created for large tables.

## **Q21. Describe the Difference Between Window Functions and Aggregate Functions in SQL.**

Answer:

The main difference between window functions and aggregate functions is that aggregate functions group multiple rows into a single result row; all the individual rows in the group are collapsed and their individual data is not shown. On the other hand, window functions produce a result for each individual row. This result is usually shown as a new column value in every row within the window.

## **Q22. Write a SQL query to find the top three products with the highest revenue in the last quarter from a sales database.**

Answer:

A SQL query to find the top three products with the highest revenue in the last quarter:

```
SELECT TOP 3 ProductID, SUM(Revenue) AS TotalRevenue
FROM Sales
WHERE OrderDate >= DATEADD(QUARTER, DATEDIFF(QUARTER, 0, GETDATE()) - 1, 0)
GROUP BY ProductID
ORDER BY TotalRevenue DESC;
```

## **Q23. What are the different subsets of SQL?**

Answer:

Data Definition Language (DDL) – It allows you to perform various operations on the database such as CREATE, ALTER, and DELETE objects.

Data Manipulation Language(DML) – It allows you to access and manipulate data. It helps you to insert, update, delete and retrieve data from the database.

Data Control Language(DCL) – It allows you to control access to the database. Example – Grant, Revoke access permissions.

## **Q24. List the different types of relationships in SQL.**

Answer:

There are different types of relations in the database:

One-to-One – This is a connection between two tables in which each record in one table corresponds to the maximum of one record in the other.

One-to-Many and Many-to-One – This is the most frequent connection, in which a record in one table is linked to several records in another.

Many-to-Many – This is used when defining a relationship that requires several instances on each side.

Self-Referencing Relationships – When a table has to declare a connection with itself, this is the method to employ.

## **Q25. What is a Stored Procedure?**

Answer:

A stored procedure is a subroutine available to applications that access a relational database management system (RDBMS). Such procedures are stored in the database data dictionary. The sole disadvantage of stored procedure is that it can be executed nowhere except in the database and occupies more memory in the database server.

## **Q26. What is Pattern Matching in SQL?**

Answer:

SQL pattern matching provides for pattern search in data if you have no clue as to what that word should be. This kind of SQL query uses wildcards to match a string pattern, rather than writing the exact word. The LIKE operator is used in conjunction with SQL Wildcards to fetch.

## **Q27. What is a Subquery in SQL? What are its types?**

Answer:

A subquery is a query within another query, also known as a nested query or inner query. It is used to restrict or enhance the data to be queried by the main query, thus restricting or enhancing the output of the main query respectively. There are two types of subqueries.

## **Q28. Correlated and non-correlated.**

Answer:

A correlated subquery cannot be considered as an independent query, but it can refer to the column in a table listed in the FROM of the main query.

A non-correlated subquery can be considered as an independent query and the output of the subquery is substituted in the main query.

## **Q29. What are Query and Query language?**

Answer:

A query is nothing, but a request sent to a database to retrieve data or information. The required data can be retrieved from a table or many tables in the database.

Query languages use various types of queries to retrieve data from databases. SQL, Datalog, and AQL are a few examples of query languages; however, SQL is known to be the widely used query language.

## **Q30. What is concurrency control in DBMS?**

Answer:

This is a process of managing simultaneous operations in a database so that database integrity is not compromised. The following are the two approaches involved in concurrency control:

Optimistic approach – Involves versioning

Pessimistic approach – Involves locking

## **Q31. What is a checkpoint in DBMS and when does it occur?**

Answer:

A checkpoint is a mechanism where all the previous logs are removed from the system and are permanently stored on the storage disk. So, basically, checkpoints are those points from where the transaction log record can be used to recover all the committed data up to the point of crash.

## **Q32. What is CTE in SQL?**

Answer:

A CTE (Common Table Expression) is a one-time result set that only exists for the duration of the query. It allows us to refer to data within a single SELECT, INSERT, UPDATE, DELETE, CREATE VIEW, or MERGE statement's execution scope. It is temporary because its result cannot be stored anywhere and will be lost as soon as a query's execution is completed.

## **Q33. Explain character-manipulation functions? Explains its different types in SQL.**

Answer:

Change, extract, and edit the character string using character manipulation routines. The function will do its action on the input strings and return the result when one or more characters and words are supplied into it.

The character manipulation functions in SQL are as follows:

- A. CONCAT (joining two or more values): This function is used to join two or more values together. The second string is always appended to the end of the first string.
- B. SUBSTR: This function returns a segment of a string from a given start point to a given endpoint.
- C. LENGTH: This function returns the length of the string in numerical form, including blank spaces.
- D. INSTR: This function calculates the precise numeric location of a character or word in a string.
- E. LPAD: For right-justified values, it returns the padding of the left-side character value.

- F. RPAD: For a left-justified value, it returns the padding of the right-side character value.
- G. TRIM: This function removes all defined characters from the beginning, end, or both ends of a string. It also reduced the amount of wasted space.
- H. REPLACE: This function replaces all instances of a word or a section of a string (substring) with the other string value specified.

## **Q34. What are Superkey and candidate key?**

Answer:

A super key may be a single or a combination of keys that help to identify a record in a table. Know that Super keys can have one or more attributes, even though all the attributes are not necessary to identify the records.

A candidate key is the subset of Superkey, which can have one or more than one attributes to identify records in a table. Unlike Superkey, all the attributes of the candidate key must be helpful to identify the records.

Note that all the candidate keys can be Super keys, but all the super keys cannot be candidate keys.

## **Q35. What is a Self-Join?**

Answer:

A self-join is a type of join that can be used to connect two tables. As a result, it is a unary relationship. Each row of the table is attached to itself and all other rows of the same table in a self-join. As a result, a self-join is mostly used to combine and compare rows from the same database table.

## **Q36. What is OLTP?**

Answer:

OLTP, or online transactional processing, allows huge groups of people to execute massive amounts of database transactions in real time, usually via the internet. A database transaction occurs when data in a database is changed, inserted, deleted, or queried.

## **Q37. What are the different integrity rules present in the DBMS?**

Answer:

The different integrity rules present in DBMS are as follows:

**Entity Integrity:** This rule states that the value of the primary key can never be NULL. So, all the tuples in the column identified as the primary key should have a value.

**Referential Integrity:** This rule states that either the value of the foreign key is NULL or it should be the primary key of any other relation.

### **1. Explain data cleansing.**

Data cleaning, also known as data cleansing or data scrubbing or wrangling, is basically a process of identifying and then modifying, replacing, or deleting the incorrect, incomplete, inaccurate, irrelevant, or missing portions of the data as the need arises. This fundamental element of data science ensures data is correct, consistent, and usable.

## 2. What is an Affinity Diagram?

An Affinity Diagram is an analytical tool used to cluster or organize data into subgroups based on their relationships. These data or ideas are mostly generated from discussions or brainstorming sessions and are used in analyzing complex issues.

## **Q38. What is an Alias in SQL?**

Answer:

An alias is a feature of SQL that is supported by most, if not all, RDBMSs. It is a temporary name assigned to the table or table column for the purpose of a particular SQL query. In addition, aliasing can be employed as a confusion technique to secure the real names of database fields. A table alias is also called a correlation name.

An alias is represented explicitly by the AS keyword but in some cases, the same can be performed without it as well.

## **Q39. How to change a table name in SQL?**

Answer:

This is the command to change a table name in SQL:

```
ALTER TABLE table_name  
RENAME TO new_table_name;
```

We will start off by giving the keywords ALTER TABLE, then we will follow it up by giving the original name of the table, after that, we will give in the keywords RENAME TO and finally, we will give the new table name.

## **Q40. How to use LIKE in SQL?**

Answer:

The LIKE operator checks if an attribute value matches a given string pattern. Here is an example of LIKE operator

```
SELECT * FROM employees WHERE first_name like 'Steven';
```

With this command, we will be able to extract all the records where the first name is like "Steven".

## **Q41. If we drop a table, does it also drop related objects like constraints, indexes, columns, default, views and sorted procedures?**

Answer:

Yes, SQL server drops all related objects, which exists inside a table like constraints, indexes, columns, defaults etc. But dropping a table will not drop views and sorted procedures as they exist outside the table.

## **Q42. Explain SQL Constraints.**

Answer:

SQL Constraints are used to specify the rules of data type in a table. They can be specified while creating and altering the table. The following are the constraints in SQL: NOT NULL CHECK DEFAULT UNIQUE PRIMARY KEY FOREIGN KEY

## **Q43. How to fetch unique records from a table in SQL?**

Answer:

SQL DISTINCT clause is used to remove the duplicated columns from the result set.

The distinct keyword is used with the select keyword in conjunction. It is helpful when we avoid duplicate values present in the specific columns/tables. The unique values are fetched when we use the distinct keyword.

To handle missing values in SQL queries, you can use the COALESCE or ISNULL function to replace NULL values with a specified default value, or you can use the WHERE clause to filter out rows with missing values using the IS NULL or IS NOT NULL condition.

## **Q44. What is Data Integrity?**

Answer:

Data Integrity is the assurance of accuracy and consistency of data over its entire life-cycle and is a critical aspect of the design, implementation, and usage of any system which stores, processes, or retrieves data. It also defines integrity constraints to enforce business rules on the data when it is entered into an application or a database.

## **Q45. What is the difference between NOW() and CURRENT\_DATE() in SQL?**

Answer:

NOW() returns a constant time that indicates the time at which the statement began to execute. (Within a stored function or trigger, NOW() returns the time at which the function or triggering statement began to execute.

The simple difference between NOW() and CURRENT\_DATE() is that NOW() will fetch the current date and time both in format 'YYYY-MM\_DD HH:MM:SS' while CURRENT\_DATE() will fetch the date of the current day 'YYYY-MM\_DD'.

## **Q46. What is DBSCAN Clustering?**

Answer:

DBSCAN groups 'densely grouped' data points into a single cluster. It can identify clusters in large spatial datasets by looking at the local density of the data points. The most exciting feature of DBSCAN clustering is that it is robust to outliers. It also does not require the number of clusters to be told beforehand, unlike K-Means, where we have to specify the number of centroids.

## **Q47. What are the different forms of joins in a table?**

Answer:

SQL has many kinds of different joins including INNER JOIN, SELF JOIN, CROSS JOIN, and OUTER JOIN. In fact, each join type defines the way two tables are related in a query. OUTER JOINS can further be divided into LEFT OUTER JOINS, RIGHT OUTER JOINS, and FULL OUTER JOINS. Constraints are used to specify the rules concerning data in the table. It can be applied for single or multiple fields in an SQL table during the creation of the table or after creating using the ALTER TABLE command. The constraints are: NOT NULL - Restricts NULL value from being inserted into a column. CHECK - Verifies that all values in a field satisfy a condition. DEFAULT - Automatically assigns a default value if no value has been specified for the field. UNIQUE - Ensures unique values to be inserted into the field. INDEX - Indexes a field providing faster retrieval of records. PRIMARY KEY - Uniquely identifies each record in a table. FOREIGN KEY - Ensures referential integrity for a record in another table.

## **Q48. Describe the Difference Between Window Functions and Aggregate Functions in SQL.**

Answer:

The main difference between window functions and aggregate functions is that aggregate functions group multiple rows into a single result row; all the individual rows in the group are collapsed and their individual data is not shown. On the other hand, window functions produce a result for each individual row. This result is usually shown as a new column value in every row within the window.

## **Q49. What is the difference between the RANK() and DENSE\_RANK() functions?**

Answer:

The RANK() function in the result set defines the rank of each row within your ordered partition. If both rows have the same rank, the next number in the ranking will be the previous rank plus a number of duplicates. If we have three records at rank 4, for example, the next level indicated is 7. The DENSE\_RANK() function assigns a distinct rank to each row within a partition based on the provided column value, with no gaps. If we have three records at rank 4, for example, the next level indicated is 5.

## **Q50. What are the different forms of joins in a table?**

Answer:

SQL has many kinds of different joins including INNER JOIN, SELF JOIN, CROSS JOIN, and OUTER JOIN. In fact, each join type defines the way two tables are related in a query. OUTER JOINS can further be divided into LEFT OUTER JOINS, RIGHT OUTER JOINS, and FULL OUTER JOINS.

## **Q51. List the different types of relationships in SQL.**

Answer:

One-to-One - This can be defined as the relationship between two tables where each record in one table is associated with the maximum of one record in the other table.

One-to-Many & Many-to-One - This is the most used relationship where a record in a table is associated with multiple records in the other table.

Many-to-Many - This is used in cases when multiple instances on both sides are needed for defining a relationship.

Self-Referencing Relationships - This is used when a table needs to define a relationship with itself.

## **Q52. What are Nested Triggers?**

Answer:

Triggers may implement DML by using INSERT, UPDATE, and DELETE statements. These triggers that contain DML and find other triggers for data modification are called Nested Triggers.

---

## **Alteryx for Data Analytics Preparation:**

### **Introduction to Alteryx**

- Overview of Alteryx's capabilities
- Key differences between Alteryx Designer, Alteryx Server, and Alteryx Analytics Gallery
- Navigating the Alteryx interface and understanding tool categories

### **Getting Started with Data in Alteryx**

- Importing data from various sources (Excel, databases, cloud platforms, etc.)
- Understanding data types and roles in Alteryx
- Preparing and cleaning data using Input and Data Cleansing tools

### **Data Blending and Preparation**

- Tools for data blending: Join, Union, Append, etc.
- Data transformation tools: Select, Formula, Filter, Sample, etc.
- Handling missing data and data type conversion

### **Advanced Data Preparation**

- Utilizing Parse tools for complex data manipulation (Text to Columns, RegEx, XML and JSON parsing)
- Data parsing and transformation with the Multi-Field Formula tool
- Advanced data preparation techniques using Fuzzy Matching and Data Cleansing tools

### **Workflow Creation and Management**

- Designing efficient workflows using various tools and macros
- Managing and organizing workflows with Containers and Tool Palette customization
- Utilizing Macros for reusable workflow components

### **Analytical Applications in Alteryx**

- Predictive analytics with the Predictive tools suite
- Spatial analytics for geospatial data processing and visualization
- Utilizing the R tool for custom R scripts within workflows

### **Building and Sharing Analytic Apps**

- Creating interactive analytic applications (Apps) in Alteryx
- Configuring interface elements (dropdowns, checkboxes, etc.) for user inputs
- Publishing and sharing workflows and apps on Alteryx Server and Alteryx Analytics Gallery

## Optimization and Performance

- Best practices for workflow performance optimization
- Techniques for large data handling and in-database processing
- Tips for efficient workflow design and tool usage

## Integration and Automation

- Automating workflows with Alteryx Scheduler
- Integrating Alteryx workflows with external systems and APIs
- Leveraging Alteryx Server for collaboration and centralized workflow management

## Advanced Analytical Techniques

- Time-series forecasting using the TS Forecast tool
- Clustering and classification using the Predictive Grouping tools
- Advanced analytics with the Python tool for custom Python scripts

## Real-world Projects and Case Studies

- Practice datasets and project scenarios for hands-on experience
- Developing a portfolio of workflows and analytic applications
- Analyzing real-world data challenges and proposing solutions using Alteryx

Excel:

---

## **Basic Excel Skills:**

1. What is Excel and what are its primary uses?
2. How do you open a new workbook in Excel?
3. Explain the difference between a workbook and a worksheet.
4. How do you navigate between different worksheets in Excel?
5. How do you insert/delete rows and columns in Excel?
6. What are Excel formulas? Give an example.
7. How do you copy formulas in Excel without changing their references?
8. Explain the purpose of Excel functions.
9. Differentiate between relative and absolute cell references.
10. How do you apply cell formatting in Excel?
11. What is the shortcut key to save an Excel workbook?
12. How do you rename a worksheet in Excel?
13. Explain how to freeze panes in Excel.
14. What is conditional formatting and how is it used in Excel?
15. How do you create a chart in Excel?

## ***Intermediate Excel Skills:***

16. Describe the VLOOKUP function and when you would use it.
17. How do you use the IF function in Excel?
18. Explain the difference between COUNT, COUNTA, COUNTBLANK, and COUNTIF functions.
19. What is the purpose of the CONCATENATE function in Excel?
20. How do you use the SUMIF and SUMIFS functions?
21. Describe the INDEX and MATCH functions and when they are used.
22. What is the purpose of the PivotTable in Excel?
23. Explain the concept of data validation in Excel.
24. How do you create a drop-down list in Excel?
25. What is the purpose of the RAND and RANDBETWEEN functions?
26. How do you remove duplicates in Excel?

27. Explain the purpose of the TRANSPOSE function. 28.

How do you protect cells in Excel from being edited? 29.

What is the purpose of the CONCAT function? 30. How do you create a named range in Excel? 31. Explain the purpose of the HLOOKUP function. 32. How do you perform text-to-columns in Excel?

## ***Advanced Excel Skills:***

33. Describe the purpose of the SUMPRODUCT function and provide an example.

34. How do you use the TEXT function in Excel?\

35. Explain the purpose of the INDIRECT function.

36. How do you use array formulas in Excel?

37. Describe the purpose of the OFFSET function.

38. What is the purpose of the XNPV function in Excel?

39. Explain the concept of goal-seeking in Excel.

40. How do you use the GETPIVOTDATA function?

41. What is the purpose of the Excel Data Model?

42. How do you create a calculated field in a PivotTable?

43. Explain the purpose of the LOOKUP function.

44. How do you use the scenario manager in Excel?

45. Describe the purpose of the EDATE function.

46. How do you use the TREND function in Excel?

47. Explain the purpose of the CORREL function.

48. How do you create a macro in Excel?

49. Describe the purpose of the DATEDIF function.

50. What is the purpose of the Goal Seek tool in Excel?

## ***Data Analysis and Visualization:***

51. How do you import data from external sources into Excel?
52. Describe the purpose of the Excel Data Model.
53. How do you use the Power Query Editor in Excel?
54. Explain the purpose of the Data Analysis ToolPak.
55. How do you perform regression analysis in Excel?
56. Describe the purpose of the Solver tool in Excel.
57. How do you create a histogram in Excel?
58. What is a sparkline and how do you create one in Excel?
59. Explain the purpose of the slicer tool in Excel.
60. How do you create a pivot chart in Excel?
61. Describe the purpose of the Scenario Manager tool.
62. How do you use the Goal Seek tool in Excel?
63. Explain the purpose of the Forecast Sheet feature in Excel.
64. How do you use the Trendline feature in Excel charts?
65. Describe the purpose of the Analysis ToolPak in Excel.
66. How do you create a waterfall chart in Excel?

## ***Excel Shortcuts:***

67. What is the shortcut to paste special in Excel?
68. How do you insert a new row/column in Excel using a shortcut?
69. Explain the shortcut to open the Excel formula editor.
70. What is the shortcut to autosum in Excel?
71. How do you apply the currency format to a cell using a shortcut?
72. What is the shortcut to select an entire column in Excel?
73. How do you select non-adjacent cells in Excel using shortcuts?
74. What is the shortcut to switch between relative, absolute, and mixed cell references in a formula?
75. How do you create a new worksheet in Excel using a shortcut?
76. What is the shortcut to insert a new chart in Excel?

## ***Excel VBA (Visual Basic for Applications) Questions:***

1. What is VBA and how is it used in Excel?
2. Explain the difference between a subroutine and a function in VBA.
3. How do you record a macro in Excel?
4. What is the VBA Editor and how do you access it in Excel?
5. How do you create a new module in the VBA Editor?
6. Explain the purpose of variables in VBA.
7. What is the difference between a public and a private variable in VBA?
8. How do you declare a variable in VBA?
9. Describe the purpose of loops in VBA and provide examples.
10. What is the difference between a For...Next loop and a Do...Loop in VBA?
11. How do you use conditional statements (If...Then...Else) in VBA?
12. Explain the purpose of the Range object in VBA.
13. How do you refer to a specific cell in VBA?
14. Describe the purpose of the ActiveSheet object in VBA.
15. How do you use the Cells property in VBA?
16. Explain the purpose of the Offset property in VBA.
17. How do you write a message to the user in VBA?
18. What is the purpose of the InputBox function in VBA?
19. Describe the purpose of the MsgBox function in VBA.
20. How do you use the Select Case statement in VBA?
21. Explain the purpose of error handling in VBA.
22. How do you use the On Error statement in VBA?
23. What is the purpose of the Resume statement in VBA error handling?
24. Describe the purpose of functions in VBA.
25. How do you call a function in VBA?
26. What is the difference between ByVal and ByRef in VBA function arguments?
27. How do you create a custom function (UDF) in VBA?

28. Explain the purpose of the Worksheet object in VBA.
29. How do you loop through all worksheets in a workbook using VBA?
30. Describe the purpose of the Workbook object in VBA.
31. How do you open another workbook using VBA?
32. Explain the purpose of events in VBA.
33. What is the Worksheet\_Change event in VBA?
34. How do you use the Workbook\_Open event in VBA?
35. Describe the purpose of user forms in VBA.
36. How do you create a new user form in VBA?
37. Explain the purpose of controls on a user form in VBA.
38. How do you add controls to a user form in VBA?
39. What is the purpose of the CommandButton control in VBA?
40. How do you write VBA code to handle events on a user form?

### ***Excel Scenario-Based Questions:***

1. Imagine you have a dataset of sales transactions. How would you use Excel to calculate the total sales revenue?
2. Suppose you have a list of employee names and their respective salaries. How would you find the average salary using Excel?
3. You have a dataset with missing values. How would you handle missing data in Excel?
4. Imagine you have a list of product prices and quantities sold. How would you calculate the total revenue using Excel?
5. Suppose you have a dataset with duplicate entries. How would you remove duplicates using Excel?
6. You need to create a report summarizing quarterly sales performance. How would you use Excel to accomplish this task?
7. Imagine you have a dataset with multiple sheets. How would you consolidate the data using Excel?
8. You have a list of customer names and their corresponding purchase dates. How would you calculate the average time between purchases using Excel?
9. Suppose you have a dataset with text entries in uppercase. How would you convert them to lowercase using

## **Tableau Questions:**

- Q1. What are the different types of joins available in Tableau, and when would you use each type?
- Q2. How do you handle data security and permissions in Tableau Server?
- Q3. Explain the concept of dashboard actions in Tableau and provide examples of how they can be used to enhance interactivity.
- Q4. What are the different types of data connections available in Tableau?
- Q5. Can you explain the Tableau data engine?
- Q6. What are the different file types that can be connected in Tableau?
- Q7. What is a Tableau workbook and what does it contain?
- Q8. What is the difference between a measure and a dimension in Tableau?
- Q9. What is a calculated field in Tableau?
- Q10. Can you explain how to create a hierarchy in Tableau?
- Q11. How do you blend data in Tableau?
- Q12. What is the difference between a join and a blend in Tableau?
- Q13. How do you handle missing values in Tableau?
- Q14. Can you explain Tableau parameters?
- Q15. How do you create a map in Tableau?
- Q16. Can you explain the difference between a heat map and a tree map in Tableau?
- Q17. How do you create a calculated field using a LOD (Level of Detail) expression in Tableau?
- Q18. Can you explain the difference between a table calculation and a calculated field in Tableau?
- Q19. How do you create a dashboard in Tableau?
- Q20. Can you explain Tableau actions?
- Q21. What is a context filter in Tableau?
- Q22. How do you optimize performance in Tableau?
- Q23. Can you explain the Tableau Server architecture?
- Q24. How do you publish a workbook in Tableau Server?
- Q25. What is a Tableau extract?
- Q26. Can you explain the difference between a live connection and an extract in Tableau?
- Q27. What is a Tableau data source?
- Q28. How do you schedule refreshes in Tableau Server?
- Q29. What is the difference between a quick filter and a normal filter in Tableau?
- Q30. How do you create a parameter control in Tableau?
- Q31. Can you explain the Tableau file types and extensions?
- Q32. How do you create a combined axis in Tableau?
- Q33. Can you explain Tableau groups?
- Q34. What is a reference line in Tableau?
- Q35. How do you create a calculated field using a table calculation in Tableau?
- Q36. How do you create a dual axis chart in Tableau?
- Q37. How do you create a histogram in Tableau?
- Q38. What is a dimension hierarchy in Tableau?
- Q39. How do you create a calculated field using a date function in Tableau?
- Q40. How do you create a drill-down report in Tableau?
- Q41. How do you create a scatter plot in Tableau?

- Q42. How do you create a calculated field using an aggregate function in Tableau?
- Q43. Can you explain the Tableau Desktop interface?
- Q44. How do you use filters in Tableau?
- Q45. How do you create a bar chart in Tableau?
- Q46. Can you explain the difference between a discrete and continuous field in Tableau?
- Q47. How do you create a calculated field using a logical function in Tableau?
- Q48. What is a parameter control in Tableau and how is it used?
- Q49. How do you create a box-and-whisker plot in Tableau?
- Q50. Can you explain the difference between a trend line and a reference line in Tableau?
- Q51. How do you create a calculated field using a string function in Tableau?
- Q52. How do you create a pie chart in Tableau?

## **Practical interview questions for an entry-level data analyst role in Tableau:**

1. Data Import Scenario: Describe how you would import data from various sources (Excel, SQL Server, CSV) into Tableau.
2. Data Cleaning Exercise: In Tableau, how would you handle a dataset with missing values and inconsistent formats to prepare it for analysis?
3. Handling Large Datasets: If you're working with a very large dataset in Tableau that is causing performance issues, what strategies would you use to optimize the data processing?
4. Calculated Fields Challenge: Explain how you would use calculated fields in Tableau to analyze year-over-year growth.
5. Data Blending Case: You have sales data in one table and customer data in another. How would you blend these data sources in Tableau to analyze customer purchase behavior?
6. Visualizations Task: Describe your approach to visualizing sales data in Tableau to highlight trends over time across different product categories.
7. Dashboard Optimization: A Tableau dashboard is loading slowly. What steps would you take to diagnose and improve its performance?
8. Data Refreshes Scheduling: How would you set up and manage automatic data refreshes for a weekly sales report in Tableau?
9. User Filters for Security: How would you implement user-level security in Tableau for a report that needs different access levels for various users?
10. Troubleshooting a Calculation: If a calculated field in Tableau is not returning the expected results, how would you go about troubleshooting it?
11. Integration with Other Tools: Describe a scenario where you integrated Tableau with another tool or service (like Excel or a web API).
12. Interactive Reports Creation: How would you design a Tableau report that allows user interaction, such as using filters or drill-down features?
13. Adapting to Data Source Changes: If there are structural changes in a primary data source (like addition or removal of columns), how would you update your Tableau reports and dashboards?
14. Managing Tableau Server/Online: Explain how you would share a report with your team and set up access controls using Tableau Server or Tableau Online.
15. SQL Queries in Tableau: How do you use SQL queries in Tableau for advanced data transformation or analysis?
16. Error Handling in Data Sources: How do you manage and resolve errors in data sources or calculations in Tableau?

17. Custom Visuals Usage: Have you used custom visuals in Tableau? Describe the scenario and the benefits.
18. Tableau Template Application: Provide an example of a situation where you created or used a Tableau template. What advantages did this offer?
19. Collaboration in Tableau Projects: Discuss how you have worked with others on a Tableau project. What collaboration tools or features within Tableau did you utilize?

## PowerBI Questions

1. What is Power BI, and how does it differ from other business intelligence tools?
2. Can you explain the various components of Power BI Desktop and their respective functions?
3. How do you import data into Power BI, and what types of data sources does it support?
4. Describe the process of transforming data in Power BI, including common data cleaning and shaping techniques.
5. What are relationships in Power BI, and how do you manage them to create effective data models?
6. How do you create calculated columns and measures in Power BI, and what are the differences between them?
7. Can you explain the concept of DAX (Data Analysis Expressions) in Power BI, and how is it used?
8. What are the different visualization types available in Power BI, and when would you use each one?
9. How do you customize and format visualizations in Power BI to enhance their appearance and usability?
10. Explain the process of creating calculated tables and calculated fields in Power BI.
11. How do you create and manage hierarchies in Power BI, and why are they useful?
12. What are slicers and how do you use them to filter data in Power BI reports?
13. Describe the process of creating and using bookmarks in Power BI to save and navigate between report views.
14. What is the role of Power Query Editor in Power BI, and how do you use it to manipulate data?
15. How do you publish Power BI reports to the Power BI Service, and what are the benefits of doing so?
16. Can you explain the concept of row-level security in Power BI, and how is it implemented?
17. What are the different deployment options available for Power BI, and how do they differ?
18. How do you schedule data refreshes in Power BI to ensure that reports are up to date?
19. What are Power BI Apps, and how do they facilitate collaboration and sharing of reports within organizations?
20. Describe some advanced features or functionalities in Power BI that you have used in your projects, and how did they add value?

## Database Questions:

1. What is a database, and why are databases important in software development?
  2. Can you explain the difference between a relational database and a non-relational (NoSQL) database?
  3. What is SQL, and how is it used in database management?
  4. Describe the main components of a database management system (DBMS).
  5. What is a primary key, and why is it important in database design?
  6. Explain the concept of normalization in database design and why it is necessary.
  7. What are the different types of database joins, and how do they work?
  8. How do you optimize database performance, and what techniques can be used?
  9. What is indexing in databases, and how does it improve query performance?
  10. Describe the ACID properties of database transactions and their importance.
  11. What is a foreign key, and how is it used to establish relationships between tables?
  12. Can you explain the differences between a clustered and non-clustered index?
  13. What is the purpose of stored procedures, and how are they different from SQL queries?
  14. Describe the concept of database transactions and explain their role in ensuring data consistency.
  15. What are the different types of database backups, and how do you implement them?
  16. Explain the concept of database replication and its use cases.
  17. What is database normalization, and why is it important in database design?
  18. Can you explain the differences between OLTP (Online Transaction Processing) and OLAP (Online Analytical Processing) databases?
  19. What is data warehousing, and how does it differ from traditional database systems?
  20. Describe the process of database migration and the challenges associated with it.
-