**Week 4 QC Review**

1. **What is Tableau?**
2. **What is the difference between Dimensions and Measures?**

|  |  |
| --- | --- |
| Dimensions | Measures |
| Dimensions contain qualitative values (such as names, dates, or geographical data)  You can use dimensions to categorize, segment, and reveal the details in your data.  Example: Category, City, Country, Customer ID, Customer Name, Order Date, Order ID | Measures contain numeric, quantitative values that you can measure (such as Sales, Profit)  Measures can be aggregated  Example: Profit, Quantity, Rank, Sales, Sales per Customer, Total Orders |

1. **What is meant by ‘discrete’ and ‘continuous’ in Tableau?**
   1. Tableau represents data depending on whether the field is discrete (blue) or continuous (green).
      1. Discrete - "individually separate and distinct."
      2. Continuous - "forming an unbroken whole without interruption."
2. **What are the different types of filters in Tableau?**
   1. Extract filters
   2. Context filters
   3. Data source filters
   4. Filters on measures
   5. Filters on dimensions
   6. Table calculation filter
3. **What is the difference between a Joins and a Union?**
   1. Join: Combines columns from different tables based on a common field or key. Used to merge data horizontally.  
        
      Union: Stacks rows from multiple tables with similar structures into a single table. Used to combine data vertically.
4. **What is the difference between a Live Connection and an Extract?**
   1. Live Connection: Connects directly to the data source in real-time. Query is executed each time a view is accessed. Suitable for frequently updated data or large data sets.  
        
      Extract: Local copy of data stored in Tableau's format. Improves performance and allows for offline access. Useful for large data sets, complex calculations, and improved query performance.
5. **What is a calculated field? How do you create one?**
   1. A calculated field in Tableau is a new field that you create based on existing fields or formulas to perform calculations or manipulate data. You can create a calculated field by right-clicking on the data source or worksheet, selecting "Create Calculated Field," and then entering the desired formula using Tableau's calculation syntax.
6. **What is a group?**
   1. In Tableau, a group is a way to combine multiple dimension members into a single category or group. It allows you to aggregate and analyze data at a higher level of granularity. To create a group, you can either select the dimension members in the view, right-click, and choose "Group," or drag and drop the dimension onto the "Group" shelf in Tableau's interface.
7. **Why might you use a bar chart vs a line chart?**
8. **How would you handle null values?**
   1. In Tableau, you can handle null values by filtering them out, using functions like IFNULL(), or creating calculated fields with custom logic. Additionally, you can treat null values as a separate category if needed.
9. **What is a dashboard?**
10. **When and why would you as a data analyst need to perform data cleaning?**
11. **What could be some potential problems you could run into when analyzing a dataset? How would you approach these?**
    1. redundant, invalid or corrupt, incomplete, or inaccurate data
12. **Why is it important to identify outliers?**
    1. Identifying outliers in Tableau is important for maintaining data accuracy, gaining valuable insights, ensuring statistical integrity, improving visualization effectiveness, and supporting informed decision-making.

**1.Define Tableau?**

Tableau is a Business Intelligence software tool that visualizes and creates interactive, shareable dashboards and allows to connect to their respective data. One can create views of their data and share them with colleagues, customers, and partners. One can use Tableau to blend it with other data and can keep their data up to date automatically.

**2.Differentiate between Tableau and other traditional BI tools?**

Tableau provides easy-to-use, best-in-class, visual analytic capabilities, but it does not help with plumbing or data foundation. One can, for example, combine SQL Server with Tableau to get the complete package.Traditional BI tools can handle it all but with significant upfront costs, higher consulting, and hardware and software costs.

**3.What is Tableau Server?**

Tableau Server is a browser- and mobile-based insight that can be used by anyone. One can easily publish dashboards with Tableau Desktop and share them through Tableau Server. It is easy to set up and even easier to run.

**4.What is data visualization in Tableau?**

Data visualization refers to the techniques used to communicate data or information by encoding it as visual objects, for example, points, lines, or bars, contained in graphics.

**5.What are the products offered by Tableau?**

Tableau products support the entire cycle of self-service analytics, such as prep, analysis, sharing, etc., along with governance and data management every step of the way. Everything is integrated into the Tableau platform. There are five main products that are offered by Tableau:

Tableau Desktop

Tableau Server

Tableau Online

Tableau Reader

Tableau Public

**6.Differentiate Context Filter  from other filters?**

In Context Filter, Tableau will create a temporary table for that particular filter set, and the other filters will be applied on the Context Filter data such as cascade parameters.The Context Filter is not frequently changed by the user and if it is changed, then the database must be recomputed and the temporary table must be rewritten.

For example, one has created a Context Filter on Dubai and India, Tableau will create a temporary table for these two countries data. If one does not have the Context Filter and has other filters, then they will be applied on these two countries data and  each record will check for all filters.

**7.How to combine two Excel files with the same fields but different data like different years?**

If there are two or more sheets or workbooks that need to be summarized, the Consolidate command can help pull the data together onto a single sheet.

To decide which type of consolidation to use, make sure to check the sheets that are being combined. If the data in the different sheets is in inconsistent positions, then consolidate by position even if the row and column labels are not identical.

If the sheets use the same row and column labels for their categories, then consolidate by category even if the data does not have consistent positions.

**8. What are sets?**

Sets are custom fields that define a subset of data based on certain conditions. The condition can be a computed condition such as a list of customers with sales over a certain threshold. Computed sets update with change in data. Alternatively, a set can also be based on a specific data point in one’s view.

**9. What are groups?**

A group is a combination of dimension members that make higher-level categories. For instance, if one is working with a view that shows average test scores by major, then certain majors can be grouped together for major categories.

**10.How can you differentiate  joining and blending in Tableau?**

Joining is useful when combining data from a single source with several tables, sheets, or others.

Blending is the combination of data from two or more different sources,one can combine the data between two sources such as Oracle, SQL Server, Excel, and others.

**11.Explain data joining and data blending in detail with suitable examples?**

**12.Define data modeling?**

Data modeling is the analysis of data objects that are used in a business or other contexts and the identification of the relationships among these data objects.

**13.Differentiate between heat map and treemap?**

A heat map is a great way to compare categories using color and size. In a heat map, one can compare two different measures.

A treemap is a very powerful visualization, particularly used for illustrating hierarchical or tree-structured data. It can be used for visualizing a part of or a whole relationship.

**14.What is a blended axis?**

Blended axis is the axis where multiple measures are shown in a single axis and all the marks are shown in a single pane.

- Drag a dimension in a column.

- Drag the first measure in the column.

- Drag the second measure in the existing axis.

**15.How can you get values from two different sources as a single input into parameters?**

Tableau currently does not support multi-valued parameters.The dynamic-parameter-with-a-blend technique can be used to highlight a single value, but not multiple values because of the way it works. As Tableau parameters are not dynamic, one cannot filter the list of values at runtime.

**16.What is the use of the new custom SQL query in Tableau?**

The custom SQL query is written after connecting to data for pulling the data in a structured view. For example, suppose, one has 50 columns in a table, but they only need 10 columns. So, instead of taking 50 columns, one can write an SQL query. This will increase the performance.

**17.Suppose I want to design a view, without using a line or bar chart, to show the region-wise profit and sales. How should I go about doing it? Explain.**

- Generate a map using cities

- Then, drag the profit and sales to Details

- Add the state as a Quick filter

**18.Design a view in a map such that if a user selects any state, the profit and sales in the cities under that state show up?**

If a user wants to show the sales and profit of each and every city under the states in the same worksheet, then they should first have State, City, Sales, and Profit fields in their dataset.

- Double-click on the State field

- Drag City and drop into the Marks card, which is under the State field

- Drag Sales and drop into Size

- Drag Profit and drop into Color

- Click on Size legend and increase the size (75%)

- Right-click on the State field and select Show Quick filter

- Select any state and check whether you got the required view or not

In this, the view size indicates the number of sales and the color indicates the profit values.

**19.How can we combine a database and the flat file data in Tableau Desktop?**

Connect data twice, once for database tables and then for the flat file. The Data->Edit Relationships

Give a Join condition on the common column from DB tables to the flat file

**20.What is Tableau Public?**

Tableau Public is a free service that lets anyone publish interactive data to the web. Once on the web, anyone can interact with the data, download it, or create their own visualizations for it. No programming skills are required here. One can also check out the gallery to see some of the things people have been doing with Tableau Public.

**21.Explain the different connections that can be made with a dataset?**

One can either connect live to their data set or extract data to Tableau.

Live: Connecting live to a dataset leverages its computational processing and storage. New queries go to the database and are reflected as new or updated in the data.

Extract: It is a compressed snapshot of the data that is stored on disk and loaded into memory as required for use by Tableau’s data engine. The snapshot can be refreshed on a recurring schedule either as a whole or as incrementally appended data. These schedules are set up by using the Tableau server.

**22.What is Tableau Data Engine?**

Tableau Data Engine is an analytical database and a feature in Tableau that has been designed to achieve instant query response and predictive performance. It seamlessly integrates into existing data infrastructure and is not limited to loading complete datasets into memory.

Tableau Data Engine does take some time to import, create indexes, and sort large volumes of data. However, it eventually speeds up after these processes. Tableau Data Engine is not in-memory technology. The data is stored in the disk after it is imported, and the RAM is hardly used.

**23.How is a workbook published and scheduled in Tableau Server?**

First, create a schedule for a particular time and then create an Extract for the data source and publish the workbook on the server.

Before publishing it, there is an option called Scheduling and Authentication. Click on that and select the schedule from the drop-down and then publish. Also, publish data sources and assign the schedule. This schedule will automatically run for the assigned time and the workbook will get refreshed on a regular basis.

**24.How to view an SQL generated by Tableau Desktop?**

Tableau Desktop log files are located in C:\Users\MyDocuments\My Tableau Repository. If one has a live connection to the data source, then they need to check the log.txt and tabprotosrv.txt files. If one is using Extract, then they have to check the tdeserver.txt file. The tabprotosrv.txt file often shows detailed information about queries.

**25.How to create filled maps?**

-Build a Map View, double-click on a geographic field such as State, Area Code, Zip Code, etc.

- Select the Filled Map Mark Type; the automatic mark type will show this type of view as circles over a map. On the Marks card, select Filled Map to color the geographic areas.

- Drag a Field to the Color shelf, define how the locations are colored by dragging another field to the Color shelf.

**26.How is it possible to store a large amount of data in a memory engine using Tableau?**

It can store a large amount of data differently on the basis of the different configurations of Tableau Server implementation, such as 8 cores, 16 cores, etc. Not just if Hyper is used or not, but other factors such as server memory may also affect the amount of data that can be stored.When dumping a large amount of data on Tableau Server, one has to make sure that this data volume should not affect the dashboard’s performance, the response time, and the processing time for extracts. This is where Tableau’s efficiency is enhanced by Hyper.

**27.Differentiate INDEX and RANK in Tableau?**

RANK and INDEX come under table calculations in Tableau. INDEX mainly deals with a record’s physical position. Incremental numbers are assigned according to the record’s physical order.While, RANK deals with a record’s value. The highest value gets the highest rank and the lowest value gets the lowest rank.

**28.Differentiate quick filter and the normal filter in Tableau?**

The quick filter is used to view the filtering options and can be used to select options. While the normal filer is something with which one can limit the options from the list or use some conditions to limit the data by field or value.

**29.Explain Tableau Architecture?**

Data-Tableau can connect to any format of source data.

Data Connectors-Tableau provides over 40 optimized data connectors. To various data sources such as MS Excel, MS SQL Server, Google Big Query, Amazon RedShift, Oracle, and others.

It also provides a generic ODBC connector for systems without a native connector.

Data can be used in-memory or live.

Components - The following components handle the server operations

a)Application server. Application Server processes (wgserver.exe) handle content browsing, Server- Administration, and authentication to Tableau server web and mobile interfaces.

b)VizQL Server. When a user/client requests a visualization, it sends a request to the VizQL process (Vizqlserver.exe). The VizQL process in turn sends queries to the data source, returning a result set in the form of images.

c)Data Server. It facilitates the management of data sources on the server.

Gateway/Load balancer - Gateway directs requests to other components.

View - Users can view Tableau dashboards thru Tableau desktop or via zero-footprint HTML 5 in a web or mobile browser.

**30.Why Tableau is preferred by the business community?**

Tableau using powerful visualization helps in understanding the data.

Tableau dashboards are interactive and display the whole picture of the data. All the data is present, data analysis across different time periods and dimensions is faster.

Datasource used in the dashboard design can be shared with other developers and users, this helps in maintaining the single version of the truth.

Development is faster as compared to other applications. Tableau provides easy to use functionalities to create data hierarchy, calculated fields, filters, parameters, Sets, and Bins.

Development is not IT-centric. It gives power to business users. Users can design their own dashboards.

Creating worksheets, dashboards, and stories are easy.

Based on the data, Tableau’s "Show Me" feature suggests users the best visualization type to use. This helps novice users in creating meaningful charts and tables.

Tableau can handle huge amounts of data. Connection to the data can be live or to an extract.

For a quick review, workbooks can be emailed as packaged workbooks with data.

Publishing to the server and applying security is easy.

**31.How does Tableau development environment work?**

Tableau development environment works as follows:

-Tableau desktop is used for creating visualizations such as charts, tables, and Maps.

-Users can connect to any data source. If multiple tables are used, then these tables can be joined. Data Sources can be named and shared.

-When the data source is created, depending on the data type, Tableau automatically segregates data into measures and dimensions. Data elements can be transformed.

-These data elements are used to create charts, tables, and Maps by simple drag-and-drop of data elements into the development area.

-One visualization is created per “worksheet”. Multiple worksheets make one “dashboard”.

-If the user needs a “story”, worksheets and dashboards can be used to create a story.

- Worksheet, Dashboards, and Story is created by clicking on the icons at the bottom of the screen How does Tableau development environment works?

or by using the menu options

**32.Explain My Tableau Repository folder?**

When Tableau Desktop is installed, Tableau creates a folder called “My Tableau Repository” in the \Documents folder.

The folder contains all the files required by Tableau:

• Logs, the folder contains all the issue logs.

• Workbooks contain all the workbooks – twb and twbx files. Save all your workbooks in this folder.

• Datasources, this folder is used to keep all the data source files such as csv, excel, etc.

• Local Data, when custom geocoding is imported, it gets stored in this folder.

• Map sources, Tableau Map Source (.tms) file is stored in this folder.

• Bookmarks, with .tbm file extension, are stored in this folder.

• Shapes, this folder contains all the shapes provided by Tableau. To add your custom shapes, copy custom shapes in an image format and add to a new folder under this folder.

• Preference.tps, the file is used to add custom color palettes.

**33.Explain the functionality of Tableau Desktop as development environment?**

Tableau desktop provides functionality to develop interactive dashboards in a development environment.

Data window - Displays information about the data connection and fields in the data source.

Analytics -  Contains ready-to-use objects for faster analysis of data.

Workbook Name - The workbook consists of data connection, worksheet, dashboard, stories, and images. Worksheet name has an extension of .twb. If it is a packaged workbook then the extension is .twbx.

Cards/shelves - Views are created by placing fields on the cards or shelves. Mark cards have different shelves such as color, size, label, detail, tooltip. Fields can be placed on these shelves. Changing the Mark type like Automatic, Shape will change these shelves.

Toolbar - The toolbar provides quick access to different functionalities such as undo/redo, adding sheets, sorting, displaying labels, and so on.

Pills - Fields or calculations on the rows or columns are called Pills. Click on a pill to access pull-down menu options such as filter, Table calculations, etc. The dimension pill is blue in color and Measure pill is green.

View/Canvas - This space displays visualization created by the fields placed on the shelves.

Filters - A filter shelf is used to place filters that limit the data.

Pages - This shelf displays views on different pages. If a dimension is placed on Pages, it creates separate pages for each dimension. If a measure is used, then measure is converted to discreet measure.

Show Me - Depending on the field selection in the data window, Tableau suggests the best-suited visualization. Different visualizations can be selected in the "Show Me" box.

Status bar - Displays various attributes of the visualization in the current worksheet. It displays information such as the number of Marks, the the number of rows and columns, and aggregated measures.

Sheet Name - Displays the name of the current worksheet. Give meaningful names to the sheets, if multiple sheets are created. There are three types of sheets - worksheet, dashboards, and story.

**34.What are the responsibilities of a Tableau professional?**

-Understanding business requirements.

-Analyzing data sources and relationships.

-Extracting data from the required data sources. Making decisions in collaboration with Business users on connecting data live or as an extract.

-Applying required data transformations, creating calculations, Sets, Bins as driven by the requirements.

-Creating visualizations that help in answering business questions. These visualizations are created and displayed in Sheets, dashboards, and stories.

-Deploying the dashboards on the server

**35.What is Tableau Development life cycle?**

Typical Tableau implementation follows these principles:

- Requirement gathering and discussion with users to understand their data and visualization requirements.

- Analyze data sources and Gather information on different data sources and relationships among the data elements.

Create mockups of the dashboards. User review of mockups.

- Create a data source.

- Structure the data by renaming data fields to user-friendly columns. As required, create calculated fields, hierarchies, parameters, and other elements.

- Depending on the number of data elements in the data source, organize data into different folders.

- worksheets and dashboards as per the requirement.

- Unit Testing. Comparing the data in the visualizations with the source data. Review of a dashboard by the users.

- If the server option is available and desired deploy dashboards over the Server.

- Create data load schedules.

**36.Differentiate Dimensions and Measures fields in Tableau?**

- Dimension fields contain the textual attribute of the data. It provides the context to the Measure. Dimensions are generally used to create labels and filters. Dimensions are discreet and appear blue in color in the data pane and the view.

- Measure fields contain the measurable attribute of the data - such as Revenue, Profit, or Population. Measures are continuous. They are axis in the charts and appear green color. Measures are analyzed by dimensions.

**37.What are the best practices of data preparation in Tableau?**

Data should be prepared so that it improves user experience and helps in the development of visualizations.

-The Data source should be given a user-friendly name. Right-click on Data Source and select Rename.

-Dimension and Measure columns should be given user-friendly names. Right-click on a field and select Rename.

-The data types of the fields. If data types are incorrect then change the datatypes in Tableau.

-Related fields should be organized in Folders. Right-click on a field and select Group By/ Folder.

-Hierarchies help in drill-downs. Identify the columns to be included Hierarchy. Create Hierarchy by Shift or Ctrl select the columns to be included in Hierarchy, right-click, and select Hierarchy.

**38.What is a TDE file and how it is created?**

Tableau Data Extract (TDE) file is a data extract or snapshot of the data. TDE files can be created by right-clicking on the data source and selecting Extract Data.

**39.A data table has 100,000 rows of data. Every month around 10,000 new rows get added to this table. What option you will choose to load only the new rows?**

This can be done by creating an extract and choosing an incremental load option.

**40.What are the different types of Calculations available in Tableau?**

Calculations or calculated field/s help in enhancing your visualization and implementing business rules. Calculated fields are created by using different Tableau functions.

Types of calculations available are:

Regular Calculations - This calculation is sent to the data source for processing and the result is returned to Tableau.

Table Calculation - Calculation occurs on top of the returned result set or chart.

Quick Table Calculation - These calculations are predefined Table calculations provided by Tableau.

Level of Detail (LOD) calculation - This calculation computes aggregation that is out of the level of detail of the view.