Hands on Tableau Training for Data Science:

# Section -1- It’s super easy to get started.

## What is Tableau and Course Overview

Tableau:

Tableau is a very simple, yet powerful tool for everything to do with data. The company's mission is to help people see and understand data. Tableau is a completely drag and drop software and using Tableau, it is possible to create visuals sometimes ten times faster than what could be created in other programs.

And overall, Tableau is a new and innovative approach to business intelligence.



Table

Description automatically generated

Graphical user interface, application

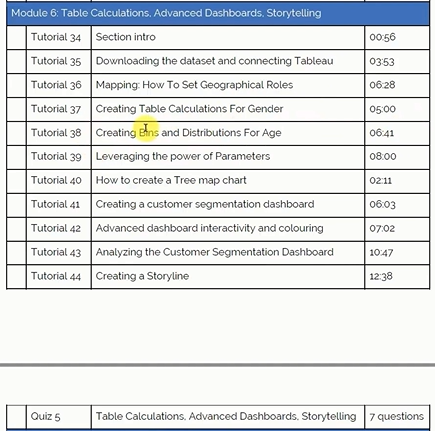
Description automatically generated with medium confidence

Table

Description automatically generated

Table

Description automatically generated

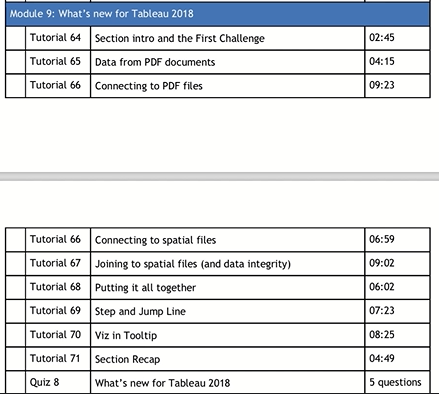


Table

Description automatically generated

Table

Description automatically generated



## Installation

Direct download of Tableau which is a 14-day trial

<https://www.tableau.com/>

Tableau Public – FREE

<https://public.tableau.com/app/discover>

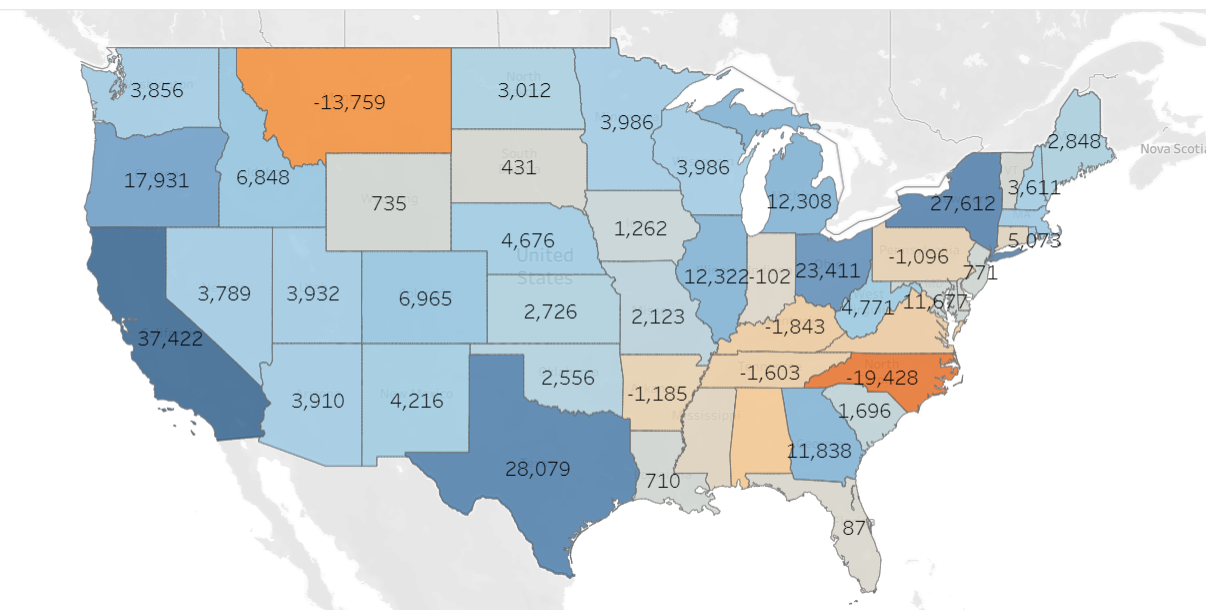
## Exercise

To get the Datasets for this course, click the below link

<https://www.superdatascience.com/pages/tableau>

* A Map to show which state has the greater and lesser profits

<https://public.tableau.com/app/profile/sayeesudha.senthil.velan.sayee./viz/SalesprofitinUS/Stateprofits>



INSIGHTS:

* It is clear that the least profitable state with most losses were incurred in North Carolina
* The most profitable state is California.
* According to the map, these southern eastern states are not doing well except for Georgia and the northern states.
* The western states are doing better when compared to Eastern State except for Montana.

## Get the Datasets here:

Please download the materials required for each section through the link below:

<https://www.artofvisualization.com/pages/tableau>

## Extra Resources:

<https://sdsclub.com/wp-content/uploads/2022/10/WHAT-CHART-TYPE-TO-CHOOSE-FOR-WHAT-DATA.pdf>

<https://sdsclub.com/wp-content/uploads/2022/10/7-Reasons-why-Tableau-Top-BI-tool.pdf>

# Section-2- Tableau Basics: First Bar Chart.

## The Business Challenge – Who gets the annual bonus?

BUSINESS PROBLEM:

It's end of financial year, and that means. It means it's time for annual bonuses.

The store operates in three regions, and only the top performing employee in each region qualifies for a bonus.

Find out which three employees are eligible to get a bonus to get bonuses for this year.

**Employees are measured on the total number or total value of sales.**

|  |
| --- |
| **CSV – Comma Separated Values** |

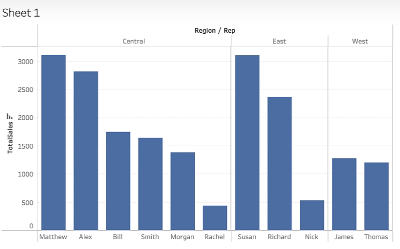
## Connecting Tableau to a Data File – CSV File

* CSV is opened as text file
* A single Tableau can many datasets and we form relationship between them in the connection manager screen
* Data -> New Data Source -> csv file

## Navigating Tableau

* Data table – it’s on the left and it has two divisions and they are
* Dimensions – quantitative data, independent variables
* Measures- qualitative data, dependent variable
* Worksheet – A single sheet where data analysis is performed
* Dashboards – A combination of worksheets
* Story – A combination of worksheets and dashboards. This is mostly the latest update of Tableau
* Analysis – How you want to perform analysis on the current sheet. Ex- Map
* Format – deals with formatting
* Worksheet
* Column
* Rows
* Show me

## Creating Calculated Fields

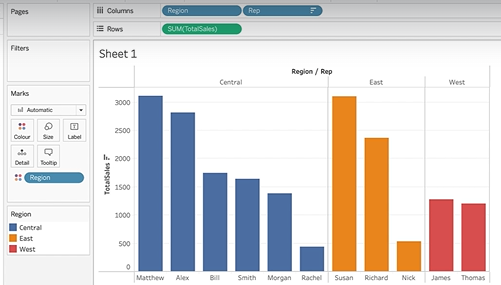
* 
* First the chart is broken as regions and following it the Rep data is broken
* Calculated fields – Right click on the unit section 🡪 Create Calculated fields
* By total sales
* 
* Bonus – Mathew, Susan and James
* By Unit Sales Result

Chart, bar chart

Description automatically generated

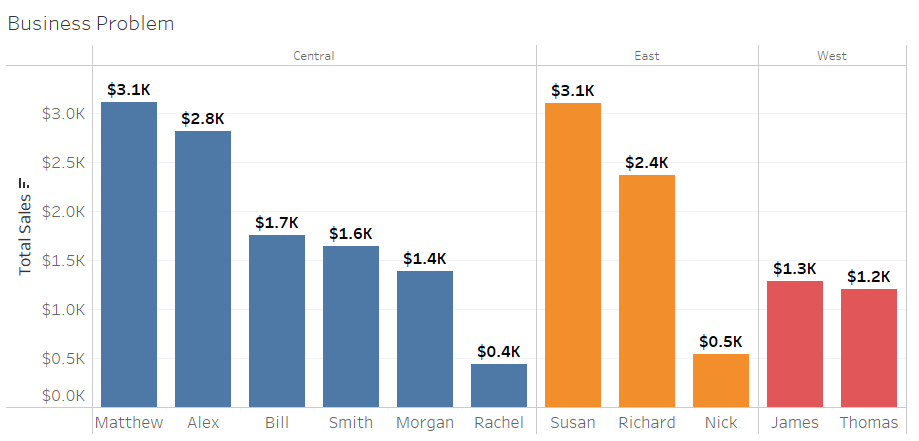
Bonus - Alex, Richard, James

## Adding Colours

* The Color button in Mark Section is where we could use to change color
* More color patterns can be provided by selecting: Colors 🡪 Edit Colors 🡪 Select Color Palette 🡪 Select anything 🡪 Assign Pallete 🡪 Apply 🡪 Ok
* TIPS – use Ctrl + select the field data in the Columns and Rows for which the colors are needed to be visualized. It is better than pulling the data from the Data sections of the Worksheet area
* 

## Adding Labels and Formatting

* How much Bill made?
* Here we can use the Label in Mark section
* Formatting – can be done using the format tab in the Menu bar and this format will be overridden by the Label format in Mark section



## Exporting Your Worksheet

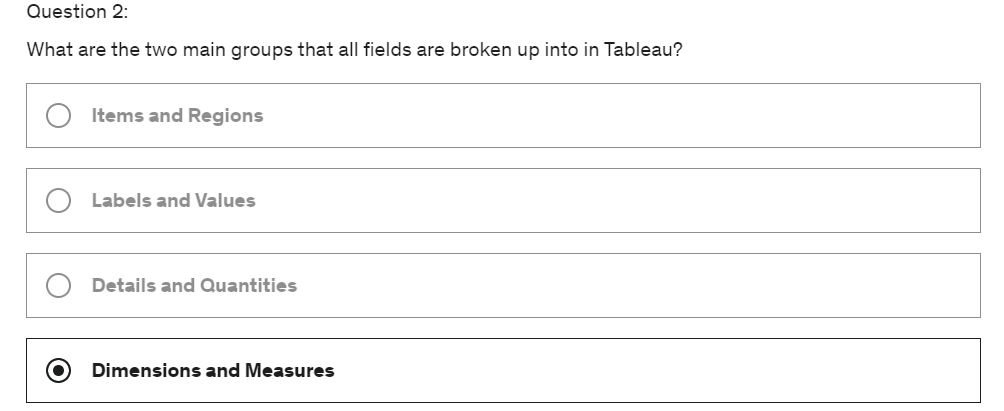
This cannot be done public version, it can be done in full Tableau Desktop version

* Worksheet 🡪 Copy 🡪 image
* Right click 🡪 copy 🡪 image

## Get the Viz

* Once the Visualized data is saved in public, then immediately your respective Viz would be created
* File 🡪 Save to Tableau Public

## Quiz 1:Tableau Basics



Graphical user interface, text, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Graphical user interface, text, application

Description automatically generated

Text, application

Description automatically generated

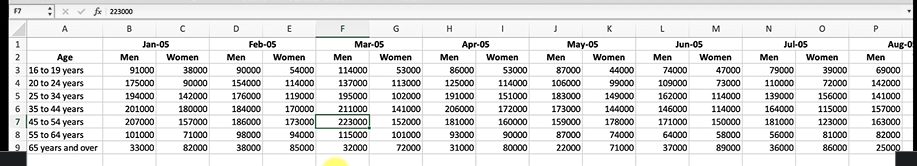
# Section-3- Time Series, Aggregation and Filters

## Section Intro

* Data Set – It is about long-time unemployment rate in the US for the past couple of years
* Time Series Data visualization
* Focus on Granularity and aggregation – concepts that underpins whole of Tableaus works
* How Granularity works, how level of details works in tableau in-order to do good visualization
* Filters and Quick Filters
* Area Chart and format it

## Working with Data Extracts in Tableau

* The Dataset has many repeated values or duplicates
* This more natural for an individual to understand the dataset



* The unstructured data is what Tableau or any BI tools preferred to modifying for the dataset to structured
* Extract Data (Paid version)– right click the data set in Data section 🡪 Click Extract data

This creates an extract for the tableau to work from

* Why do you use extract?

Using the live database will not be reliable as the data might be updated in the database and so extract the existing dataset and the evaluate the values and so if so any changes done in the extracted dataset. It will not affect the actual live dataset.

* To refresh the extract dataset: right click the data 🡪 click Extract 🡪 Refresh
* To return to the live connection: right click 🡪 uncheck Extract

## Working with Time Series

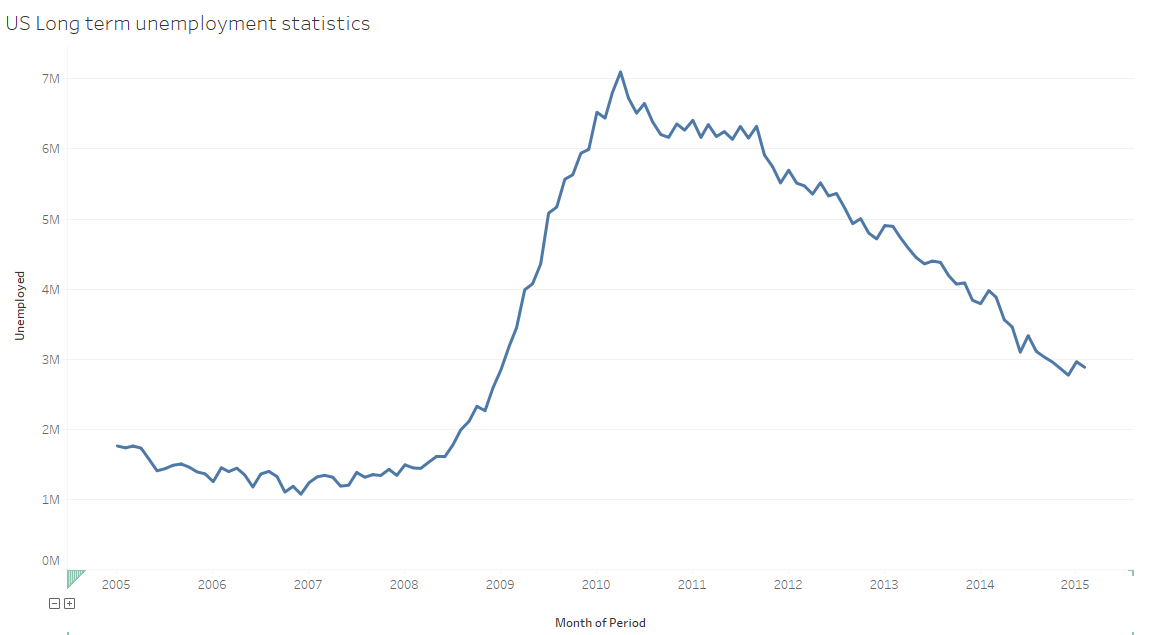
* Tableau automatically sums the number unemployment for each year (2007 – 2015)



* Now how to get the data of months (Granularity concept is utilized)
* TIPS: Dimensions are – ‘BLUE’ and measures – ‘GREEN’
* TIPS: double click the column or row data to see what calculation is performed by tableau



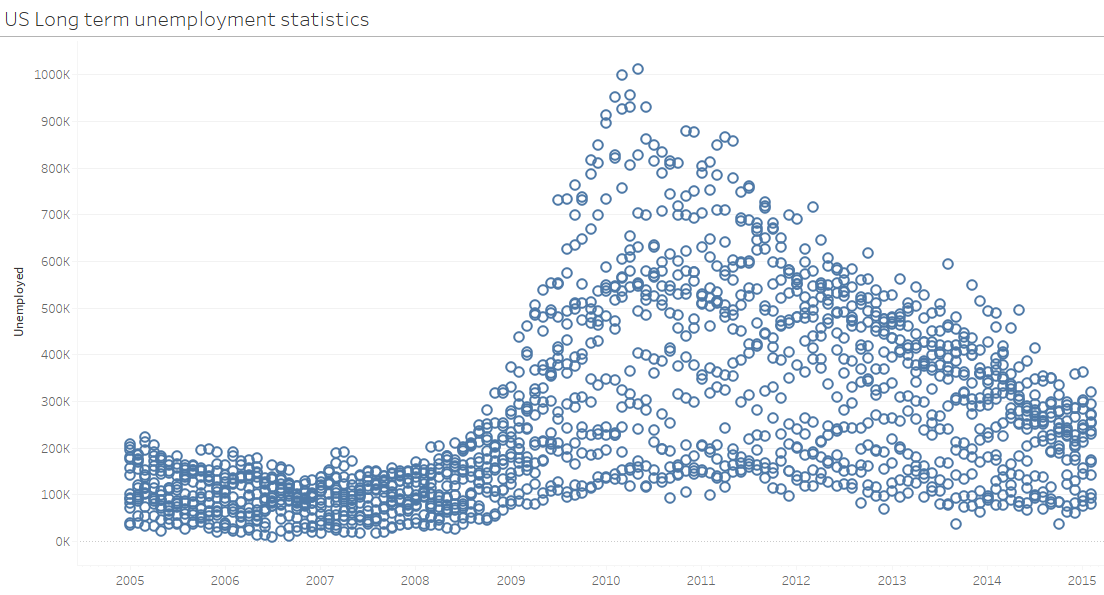
* In this case, the timeline must be evaluated as measures and not dimensions

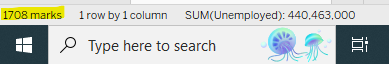


## Understanding Aggregation, Granularity, and Level of Details

* Month of period is the variable that governs granularity
* Measures get aggregated and dimensions mentions the level of granularity
* If you switch of the aggregation (Analysis 🡪 Untick the Aggregate measures)

Every single row of the data set is plotted in this graph. So it I possible to have lots of value on the same month and year



* In the bottom - 

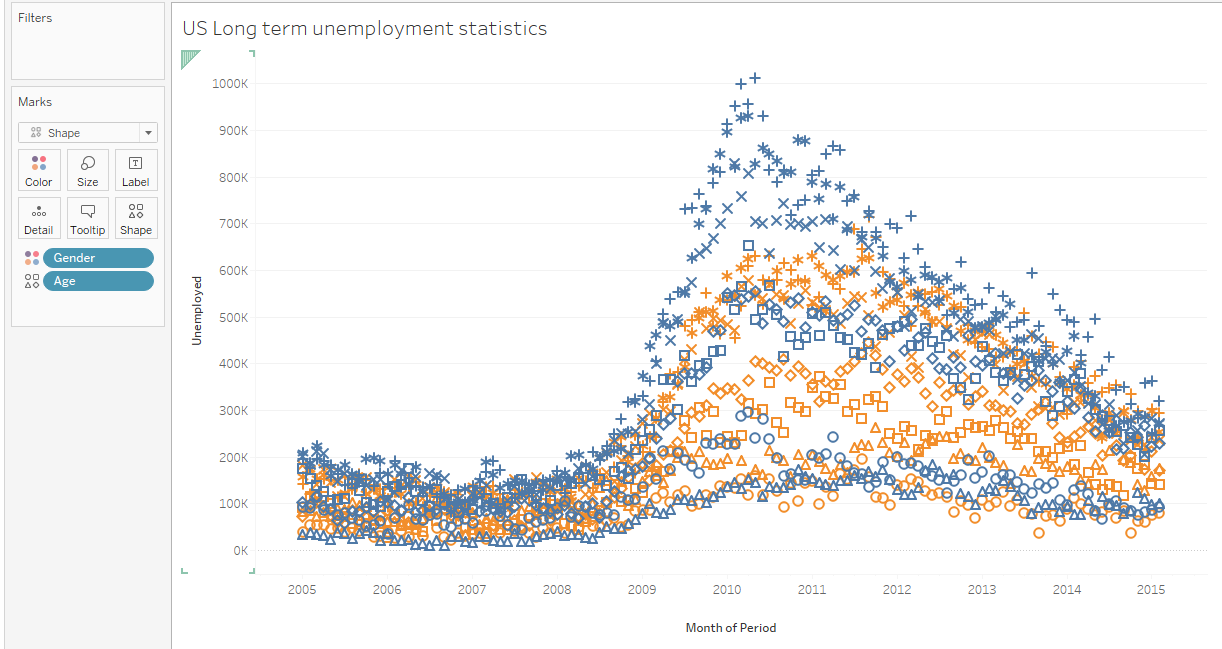
This is total number of rows that the dataset has, as we know our data set in monthly levels

* Gender – M and F who were unemployed in the specific months

Graphical user interface, chart, scatter chart

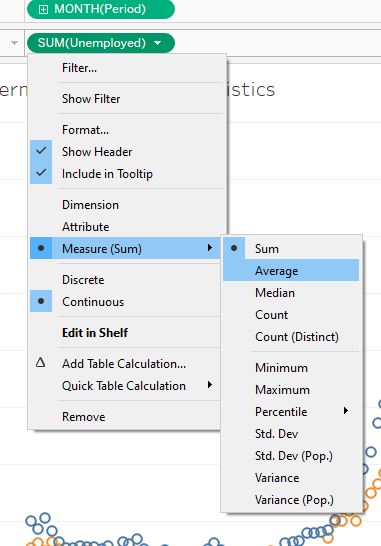
Description automatically generated

* Age is also granularized by using shape inf mark section



* Graphical user interface, application

  Description automatically generated
* Different types of aggregations:
* Avg



* Median

Chart, scatter chart

Description automatically generated

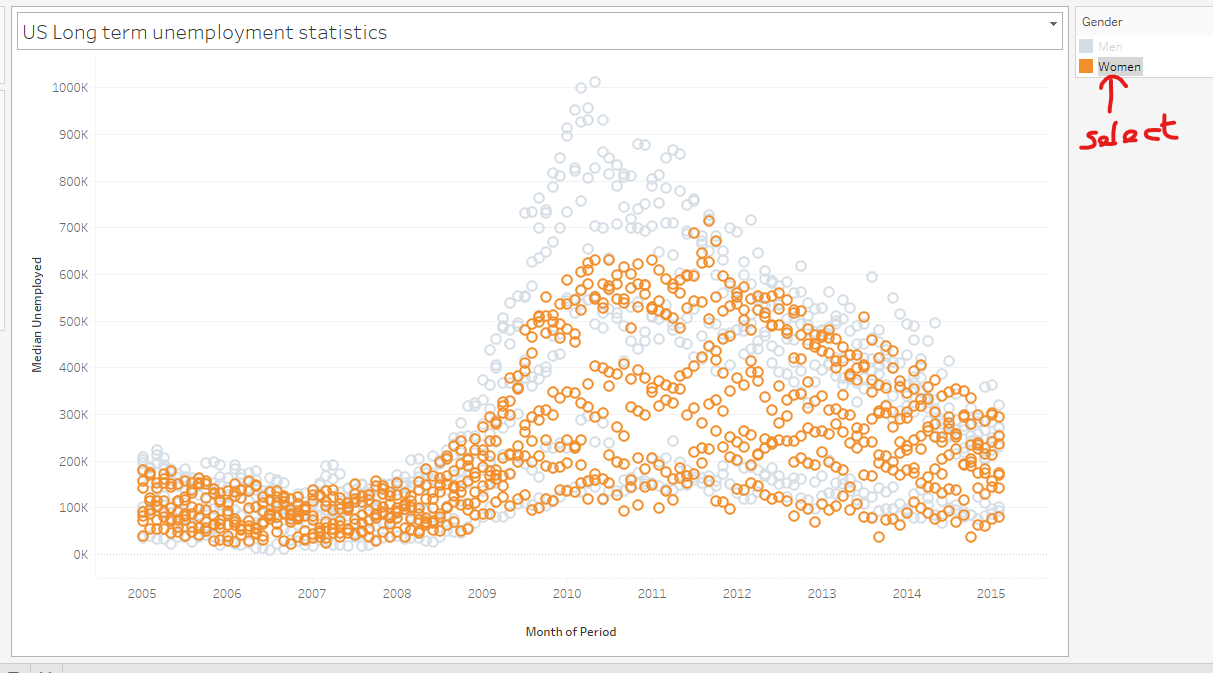
* More Details to be provided for granularity.
* Select option ‘Detail’ in mark section

Chart, scatter chart

Description automatically generated

## Creating an Area Chart & Learning about highlighting

* Need to know more about a particular Gender granularity.



* You can view the visualization of data of a particular group. For this Highlight select item option can be used

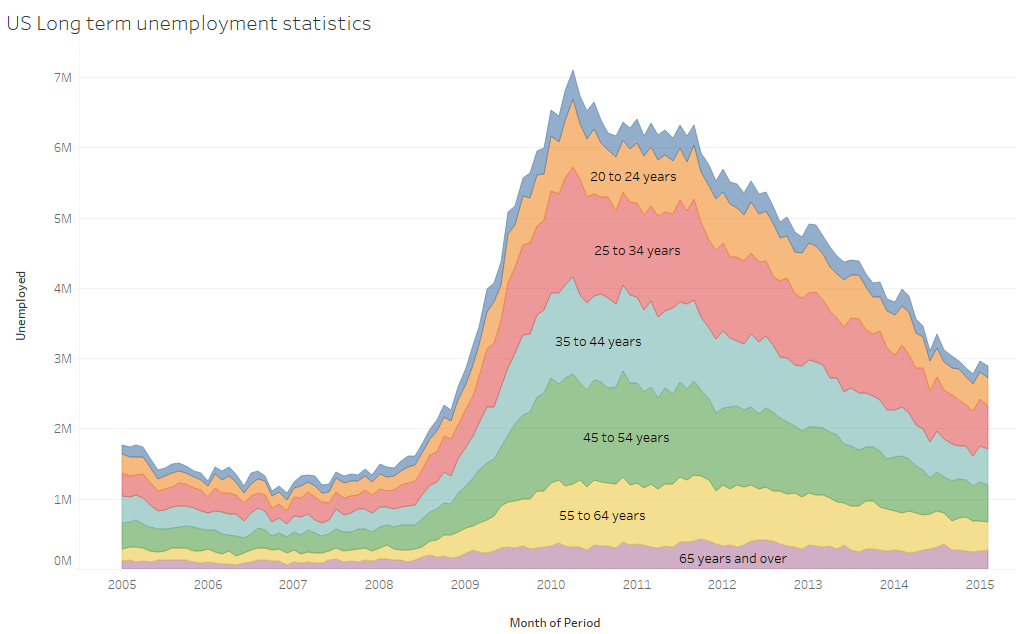
Graphical user interface, text, application

Description automatically generatedif we click on any of the data only that can be highlighted and visualized

Graphical user interface, chart, scatter chart

Description automatically generated

* Area Chart



## Adding a Filter and Quick Filter

* To do filter add the data field value to the filter section of the workseet

Chart, histogram

Description automatically generated

* Quick Filter – right click of the field value 🡪 show filter

Chart, histogram

Description automatically generated

* Multiple filters and quick filters

Chart, histogram

Description automatically generated

* Types of filters

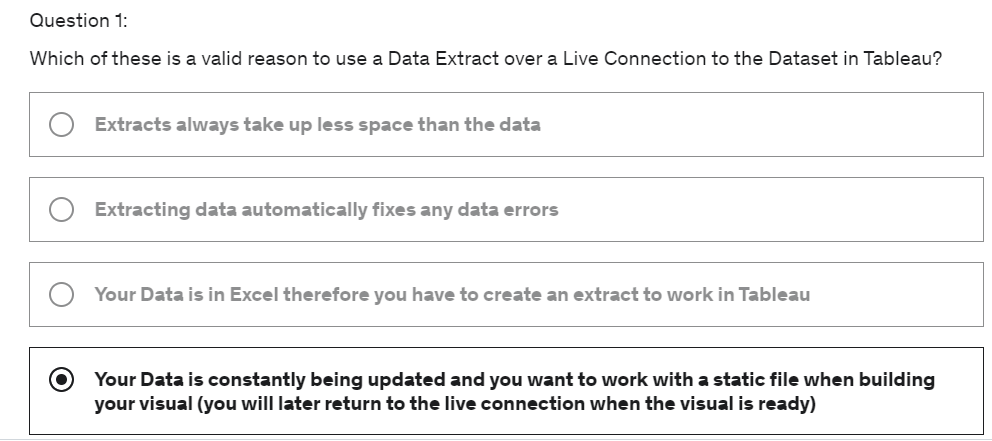
Graphical user interface, application

Description automatically generated

Chart

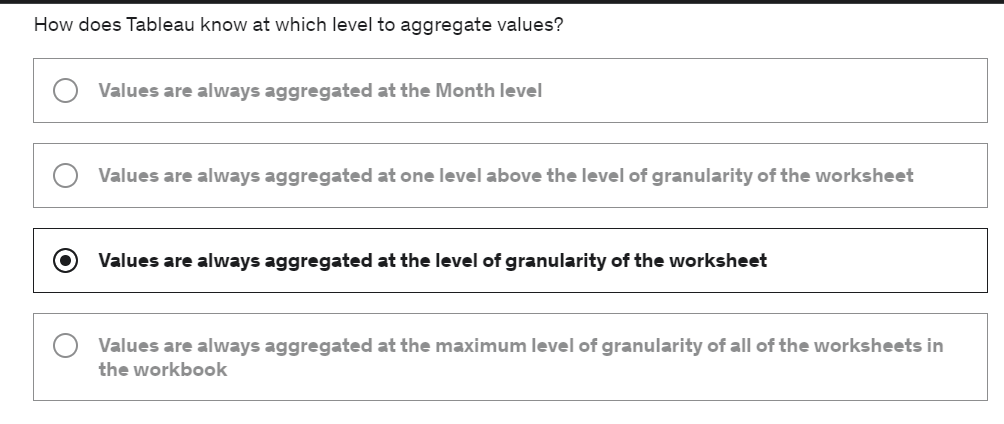
Description automatically generated

## QUIZ 2: Timeseries, Aggregation and Filters



Graphical user interface, text, application

Description automatically generated



Text

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

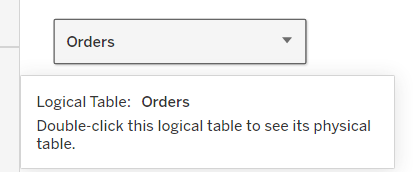
# Section-4- Maps, Scatterplots and your first Dashboard

## Section Intro

* We will doing Maps and Scatter plots
* Datasets – Sales data of a store in Europe and we will create a map to determine how it is performing across different regions and scatter plots for customers
* Create our first Dashboard which is interactive

## Relationships vs Joins

* V.2020.2
* Data Model - Where one dataset is connected with another dataset
* Relationships – in Logical Layer Tables are connected for relationships
* Joins – It is created by double clicking the table data set in the Logical layer, so it creates the physical layer



Physical layer

Graphical user interface, application

Description automatically generated

Drag the other data set in this layer in-order to create a join

Diagram

Description automatically generated

Once you close the physical layer, the logical layer will have the joined data set

Table

Description automatically generated with low confidence

## Joining Data in Tableau

* We need to join ‘List of Orders’ and ‘Order Breakdown’ tabs a

## Creating a Map, Working with Hierarchies

* Geographical data of the dataset are – Country, State and City
* We can create the natural hierarchy by dragging city from the dataset fields and place them next to country. Tableau will prompt us saying a hierarchy to be created

## Creating Scatter plot, Applying filters to Multiple Worksheets

## Lets create our first Dashboard

## Adding an Interactive Action – Filter

## Adding an interactive Action – Highlighting

## Quiz 3: Maps, Scatterplots and Your First Dashboard