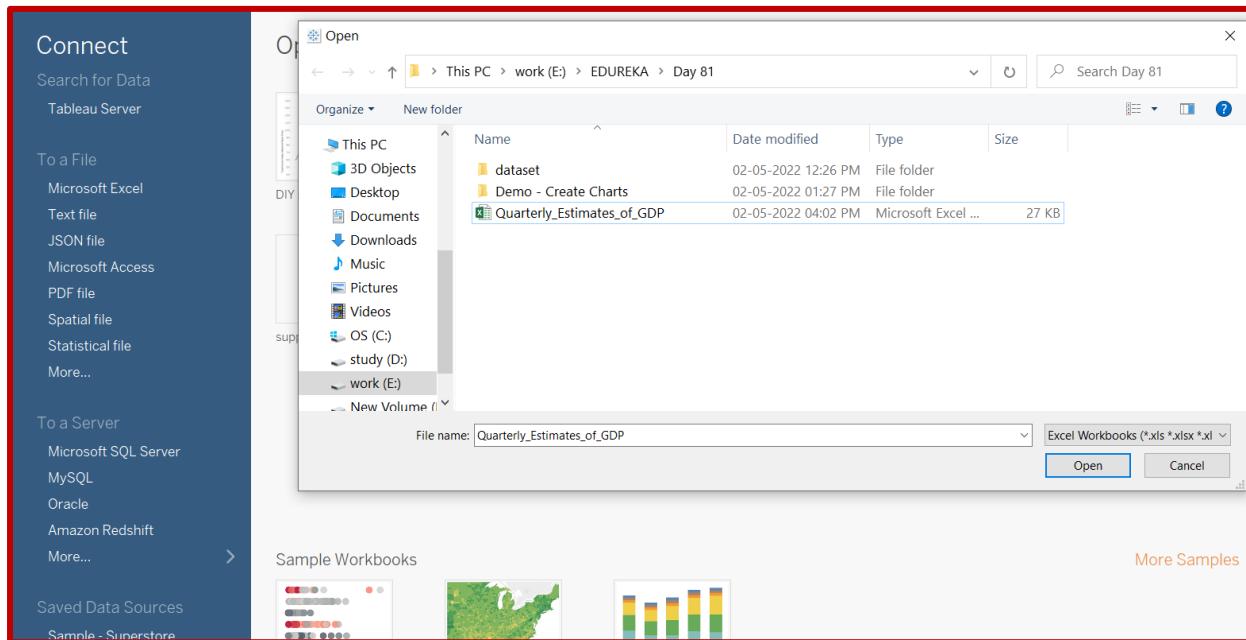


DATA SCIENCE AND MACHINE LEARNING INTERNSHIP PROGRAM

Assignment-8

Solution for Problem Statement 1

Task 1: Click on “Microsoft Excel” → Select “Quarterly_Estimates_of_GDP” → Open.



Drag and drop the “2011-12” sheet from the Quarterly_Estimates_of_GDP data onto the flow pane.

Data Science and Machine Learning Internship Program

The screenshot shows the Tableau Data Source interface. On the left, the 'Connections' pane is open, showing a connection to 'Quarterly_Estimates_of_GDP' (Microsoft Excel). The 'Sheets' pane lists several sheets: '1999-2000', '2004-05', '2011-12', and 'New Union'. The '2011-12' sheet is selected and previewed on the right. The preview area shows a table with 12 fields and 43 rows. A tooltip 'Go to Worksheet' points to the 'Sheet1' tab at the bottom. The top right shows connection status ('Live') and filter settings ('0 | Add').

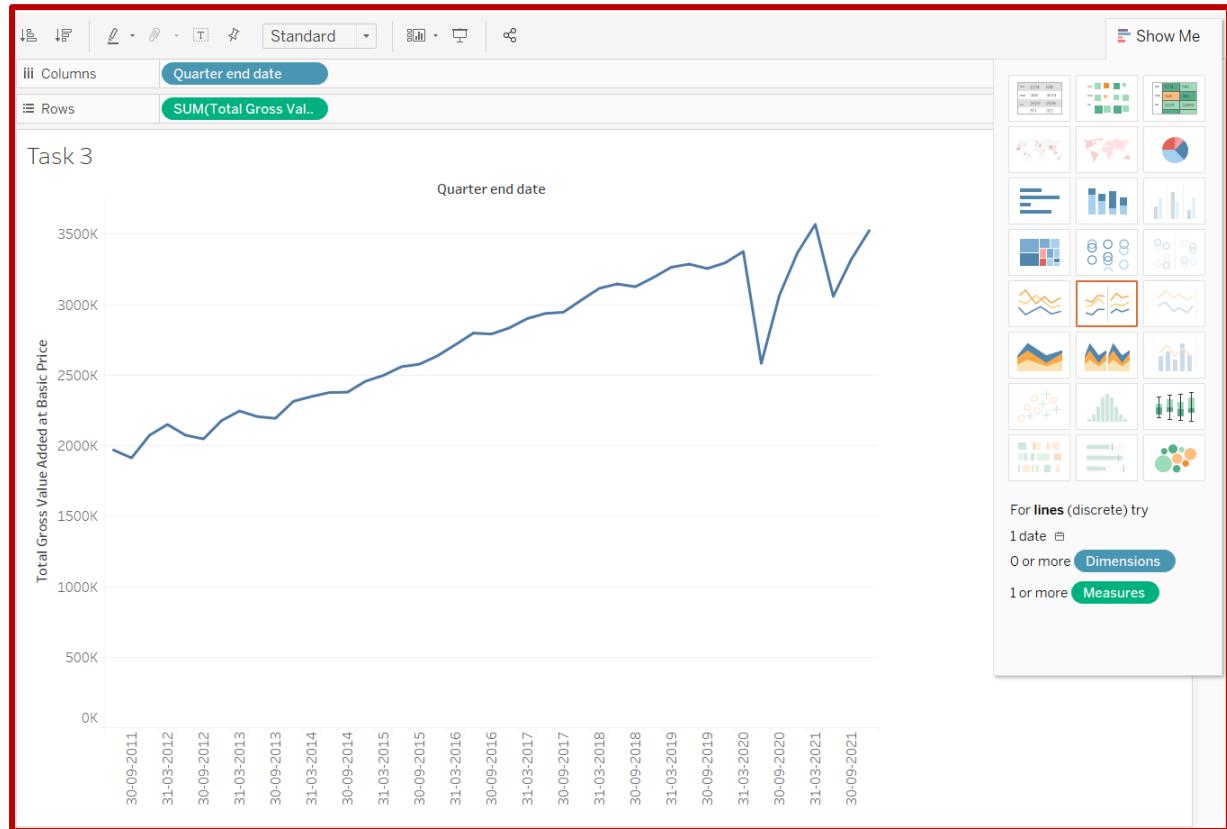
Task 2: Go to the table, locate “Quarter-end date,” select Abc, and then change its data type to date.

This screenshot is similar to the previous one, showing the '2011-12' sheet in Tableau. A context menu is open over the 'Quarter end date' field in the 'Fields' section. The menu options include: Number (decimal), Number (whole), Date & Time, Date, String, Spatial, Boolean, and Default. The 'Date' option is highlighted. The preview table on the right remains the same, showing data for quarters Q1 through Q4.

Task 3: Plot a line chart for an entire given timeline for “Total Gross Value Added”

Drag and drop the required fields from dimensions and measures to column and row shelf to visualize the Line chart.

1. “Total Gross Value Added ” to rows
2. “Quarter-end date” to columns
3. Select a line graph from the “Show me” menu

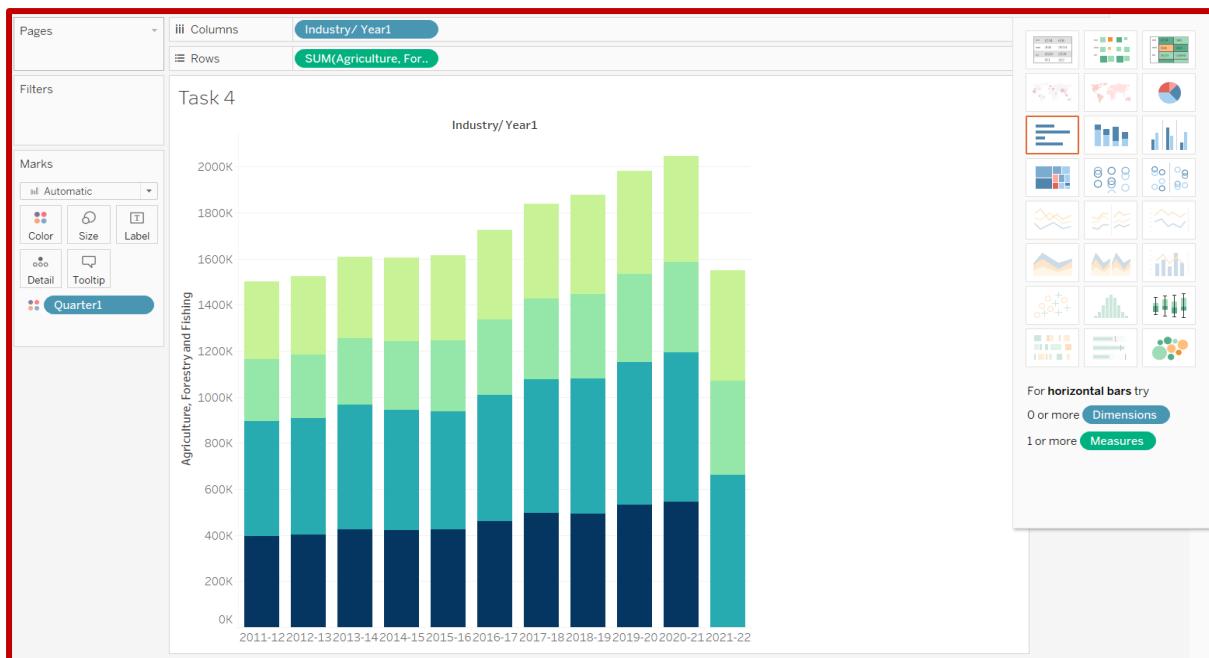


Click on a new sheet for new task.

Task 4: Finding out the total contribution of Agriculture, Forestry, and Fishing fields to GDP for each year and each quarter with help of a bar plot (use different color pellets to represent each quarter.)

Drag and drop the required fields from dimensions and measures to column and row shelf to visualize the bar chart. (Task 4)

1. “Agriculture, Forestry and Fishing” to rows
2. “Industry/Year1” to columns
3. Select a bar graph from the “Show me” menu and give color to the “Quater1 “ field to differentiate each Quarter.



Click on a new sheet for new task.

Task 5: Compare the contribution of manufacturing, mining, quarrying, and financial and real estate sectors and their contribution to GDP each year quarter-wise with the help of a bar plot

Drag and drop the required fields from dimensions and measures to column and row shelf to visualize the multiple bar chart. (task 5)

1. “Manufacturing”, “Mining and Quarrying”, and “Financial, Real Estate, and Professional Services” to rows
2. “Industry/Year1” to columns

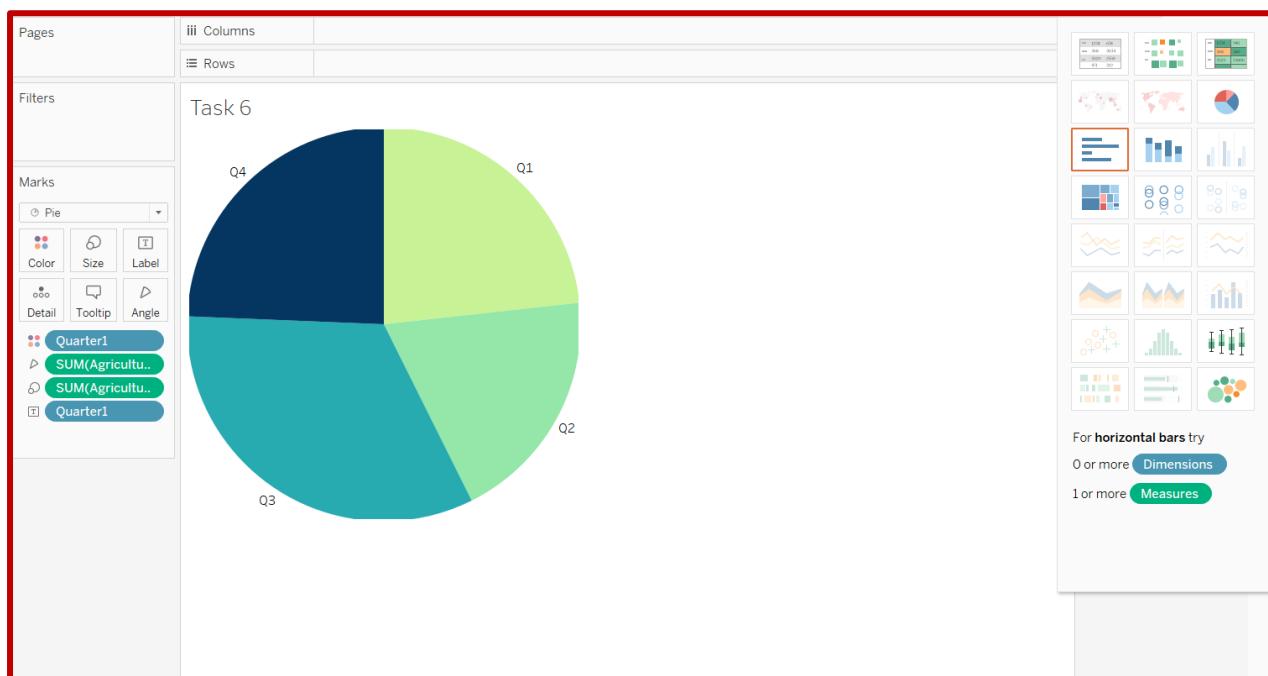


Click on a new sheet for new tasks.

Task 6: Find out in which quarter the agriculture and fishing sector contributed more to GDP in all years with help of a pie chart

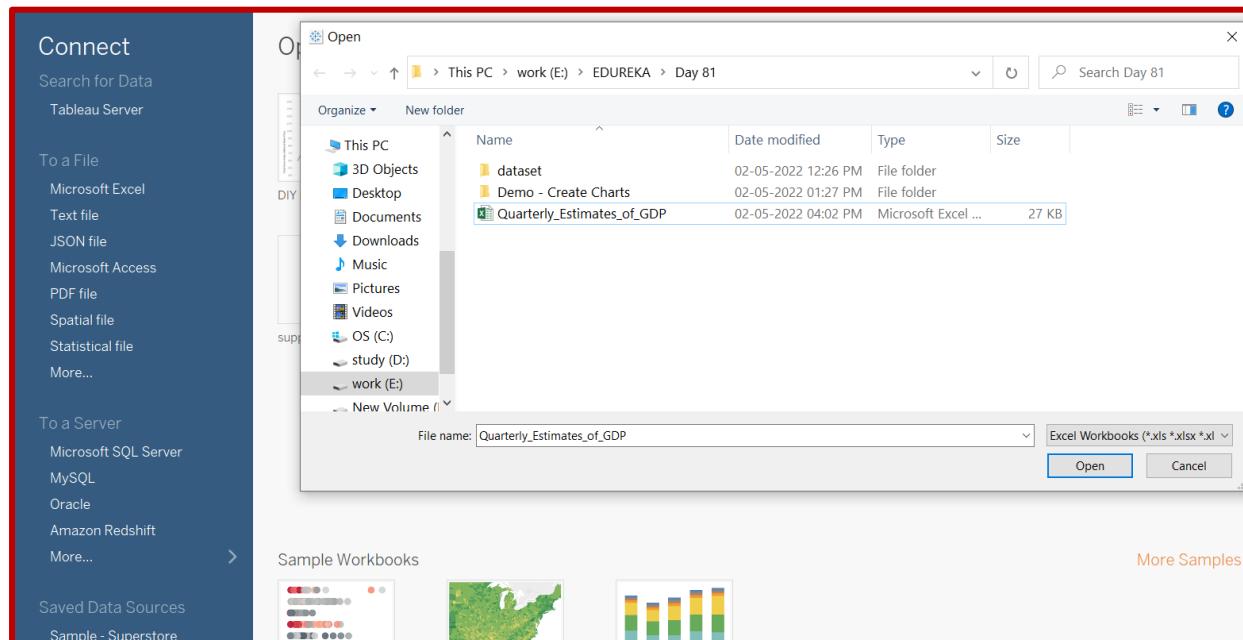
Drag and drop the required fields from dimensions and measures to column and row shelf to visualize the multiple bar chart. (task 6)

1. “Agriculture, Forestry and Fishing” to rows
2. “Quater1” to columns
3. Select a Pie chart from the “Show me” menu
4. Give labels and colors to the “Quater1” field



Solution for Problem Statement 2:

Task 1: Click on “Microsoft Excel” → Select “Quarterly_Estimates_of_GDP” → Open



Drag and drop the “2011-12” sheet from the *Quarterly_Estimates_of_GDP* data onto the flow pane.

The screenshot shows the Tableau interface with a connection to 'Quarterly_Estimates_of_GDP'. The 'Connections' section shows the connection is live. The 'Sheets' section lists '2011-12' as the active sheet. The main workspace displays the '2011-12 (Quarterly_Estimates_of_GDP)' sheet. The sheet has a title bar with '2011-12' and a small icon. Below the title, there's a note 'Need more data?' with a link 'Drag tables here to relate them. Learn more'. The data table has columns: 'Name', 'Quarter end date', 'Quarter', 'Industry/ Year', and '#'. The data rows are: Abc, 2011-12, 2011-12, 2011-12, Agriculture; 30-06-2011, Q1, 2011-12, , ; 30-09-2011, Q2, 2011-12, , ; 31-12-2011, Q3, 2011-12, , ; 31-03-2012, Q4, 2011-12, , . At the bottom, there are buttons for 'Data Source', 'Sheet 1', and other navigation icons.

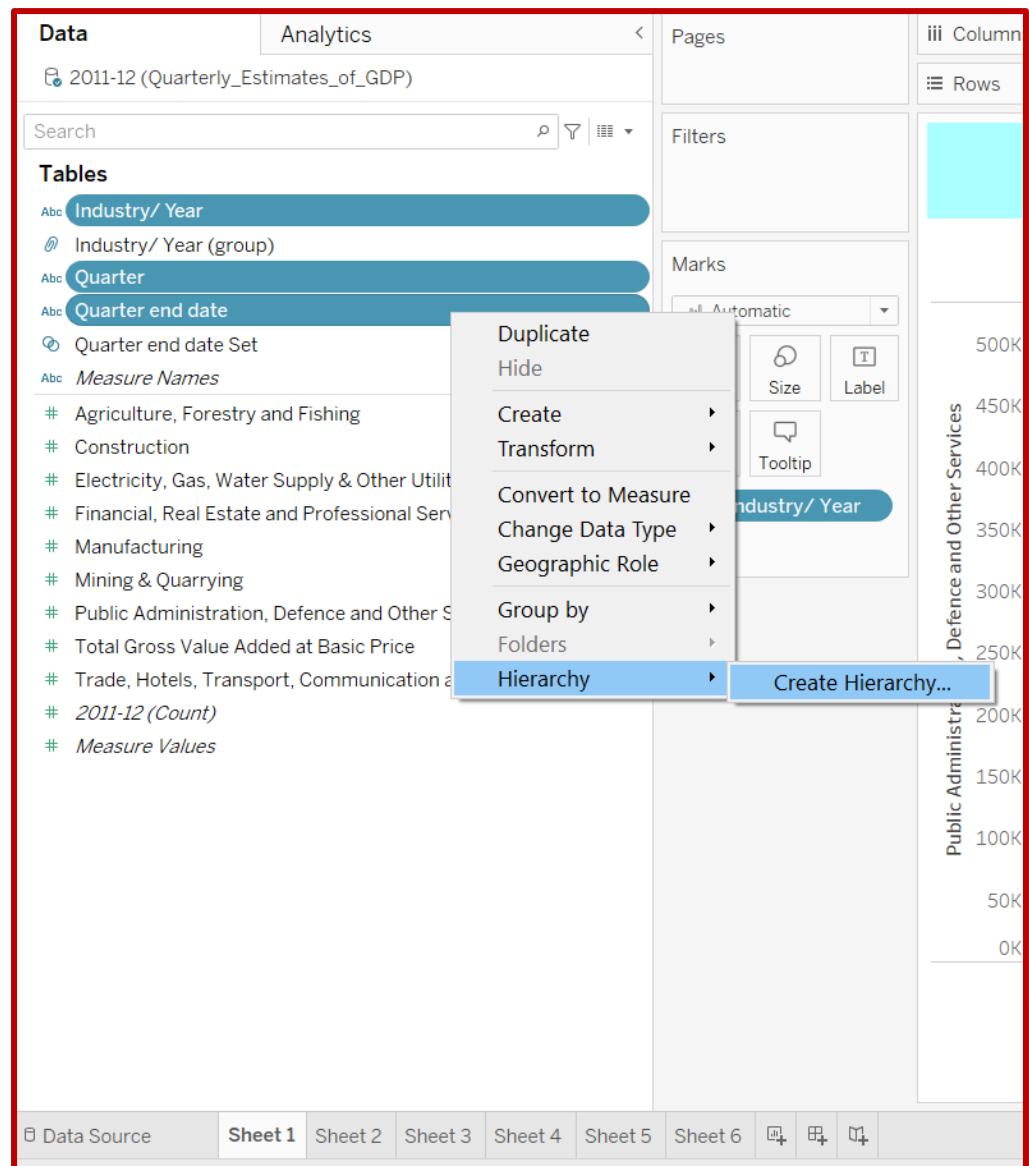
Task 2: Go to the table and locate “Quarter-end date”, select Abc, then change its data type to date.

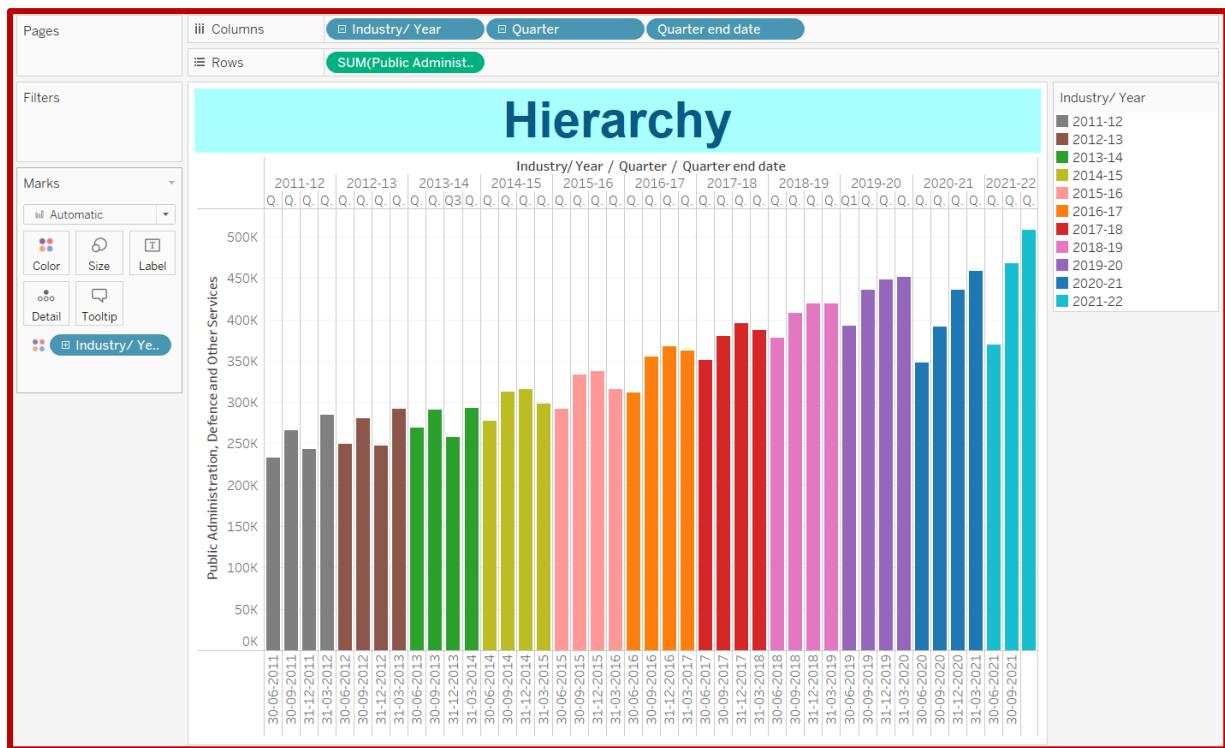
The screenshot shows the Tableau Data Source interface for the '2011-12 (Quarterly_Estimates_of_GDP)' connection. On the left, there's a sidebar with 'Connections' (selected), 'Sheets' (with options like '1999-2000', '2004-05', '2011-12', and 'New Union'), and a 'Data Interpreter' section. The main area shows a table with 12 fields and 43 rows. A context menu is open over a date field, with 'Date' selected from a list of types including Number (decimal), Number (whole), Date & Time, Date, String, Spatial, Boolean, Default, and Geographic Role. The table data includes columns for Quarter end date, Quarter, and Industry/ Year.

Task 3: Make a hierarchy of “year”, “quarter” and “quarter-end date” in the given order, and using a bar chart, observe “Public Administration, Defense, and Other Services”

To make a new group, press **Ctrl** and select all columns that you desired, then press right click and on Hierarchy click on Create Hierarchy

1. “Public Administration, Defense, and Other Services” to rows
2. Newly created group of columns

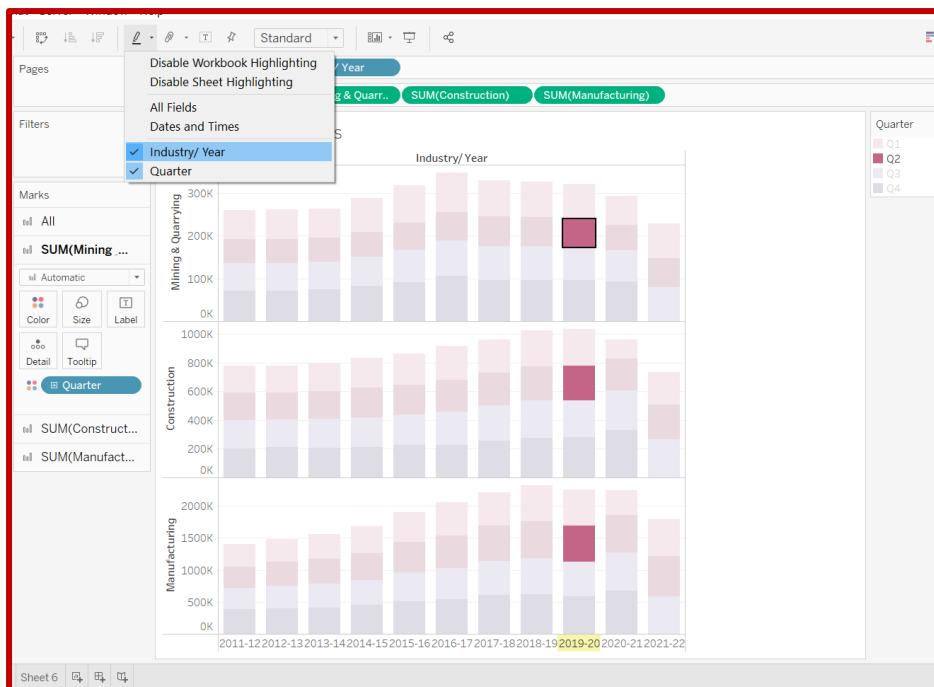




Click on a new sheet for new tasks.

Task 4: Plot a Multilevel bar chart for “manufacturing”, “construction”, “mining & quarrying”, and use Highlighted features on “quarter” and “industry/year”

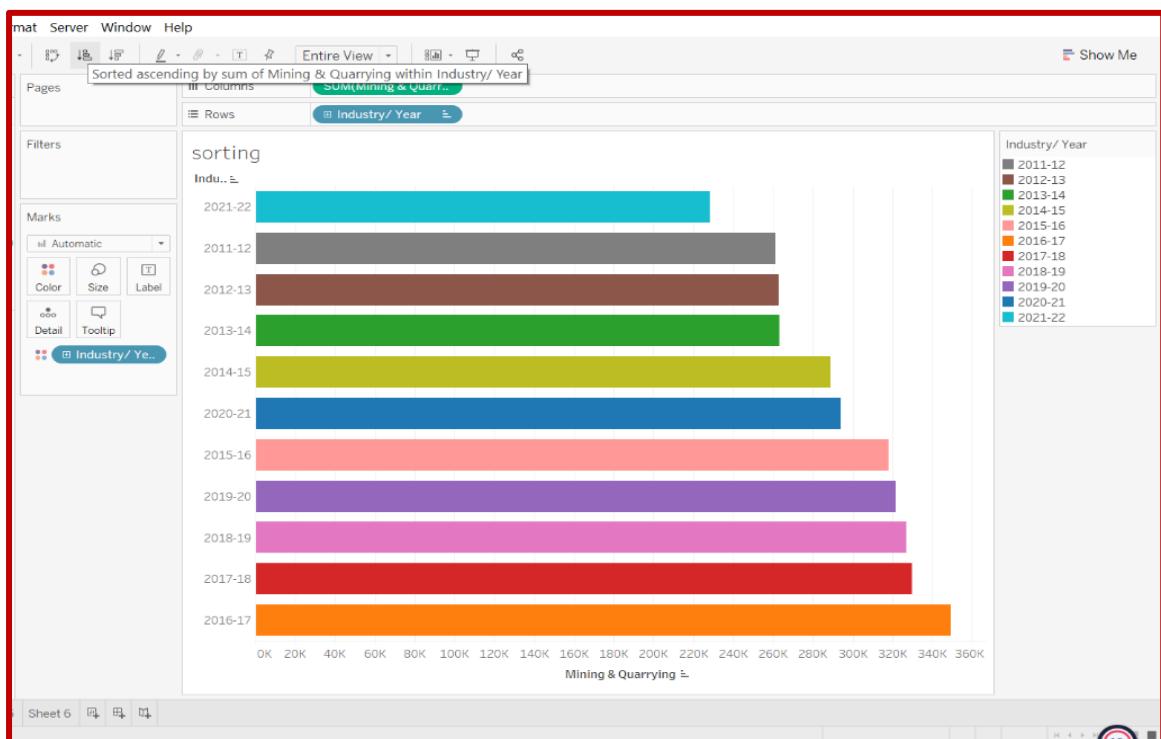
1. “Manufacturing”, “construction”, “mining & quarrying” to rows
2. “Industry/Year” to columns
3. Select a bar graph from the “Show me” menu and give color to the “Quater” field to differentiate each Quarter
4. Then From the formatting bar select the pen icon (highlight) and then select “industry/year”, “quarter”



Click on a new sheet for new tasks

Task 5: Find out in which year mining and quarrying contributed highest to GDP (Do sorting to find the answer visually)

1. "Mining and Quarrying" to columns
2. "Industry/Year" to rows
3. Then select the sorting option from the formatting bar



The year 2016-17 contributed highest to GDP

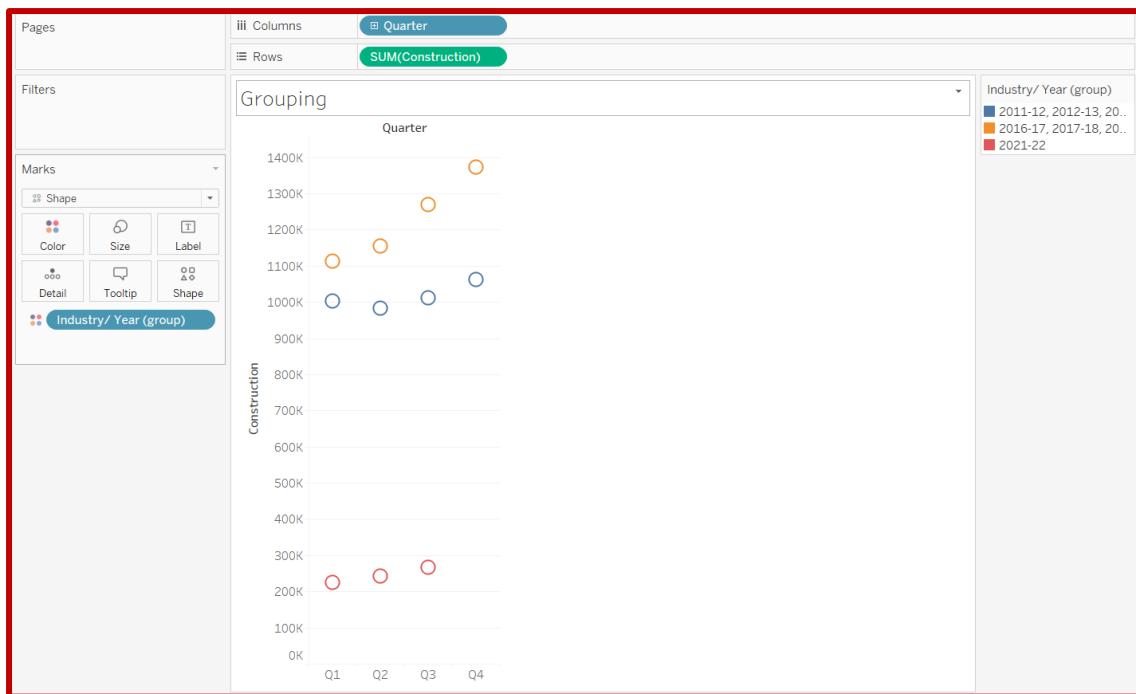
Click on a new sheet for new tasks

Task 6: Create two groups of the last decade in a 5-year gap on “industry/year” and compare construction activity (leave 2021-22 out of this because it is not part of the last decade)

1. “Construction” to rows
2. “Quater” to columns
3. select a Circle views from the “Show me” menu
4. Press on right-click, then go to create and click on the group
5. Then form a group as per your condition
6. And give color to a new group and observe

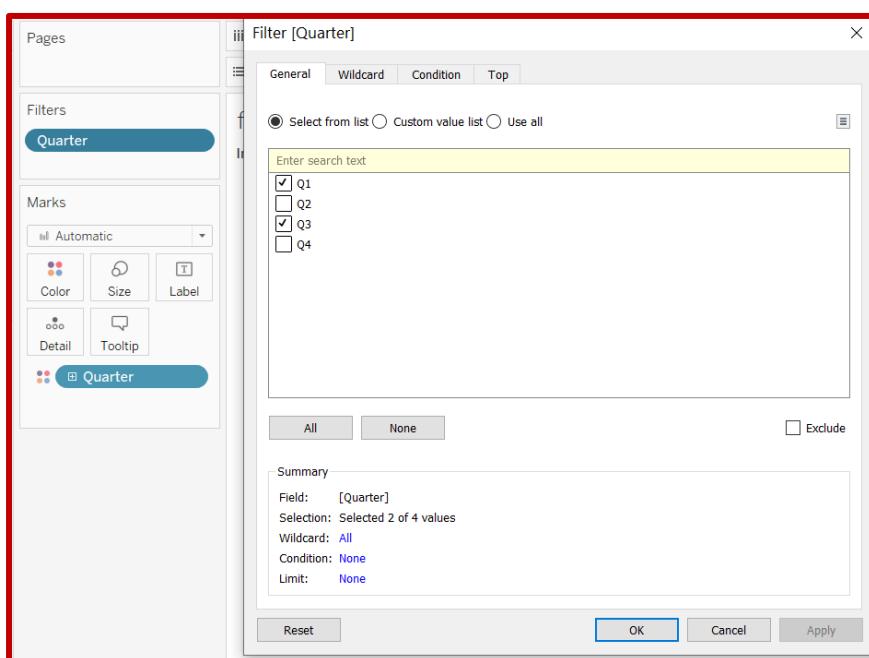
The screenshot shows the Tableau interface with the following details:

- Tableau - diy 82** is the active worksheet.
- Data** is selected in the top-left corner.
- Tables** pane: Contains 'Industry/ Year (group)', 'Industry/ Year, Quarter end date, Quarter', 'Quarter', 'Quarter end date', 'Quarter end date Set', 'Measure Names', and various industry categories like Agriculture, Construction, etc.
- Marks** pane: Shows various mark types: Size, Label, Tooltip, and Shape.
- Context Menu (Open at Industry/Year):**
 - Add to Sheet**: Sub-options include Show Filter, Duplicate, Rename, Hide, Aliases..., Create, Transform, Convert to Measure, Change Data Type, Geographic Role, Default Properties, Group by, Folders, Hierarchy, Replace References..., and Describe... (the 'Create' option is highlighted).
 - Create**: Sub-options include Calculated Field..., Group..., Set..., Parameter..., Group by, Folders, Hierarchy, Replace References..., and Describe...
- Create Group [Industry/ Year]** dialog box (highlighted with a red border):
 - Field Name:** Industry/ Year (group) 1
 - Groups:** A list of fiscal years from 2011-12 to 2020-21.
 - Add to:** A dropdown menu.
 - Buttons:** Group, Rename, Ungroup, OK, Cancel, Apply, Find >>, Reset, and a checkbox for 'Include 'Other''.
- Sheet** pane: Shows a chart area with a scatter plot.
- Bottom Navigation:** Data Source, Sheet 1, Sheet 2, Sheet 3, Sheet 4, Sheet 5, Sheet 6, and other sheet icons.



Task 7: Use the filter on quarter and select only q1 and q2 data only then plot horizontal bar plot for “Public Administration, Defence, and Other Services” by different color

1. “Industry/Year” to rows
2. “Public Administration, Defence, and Other Services” to columns
3. First, drag the quarter field to the filter section, then select q1 and q3 and apply the color quarter

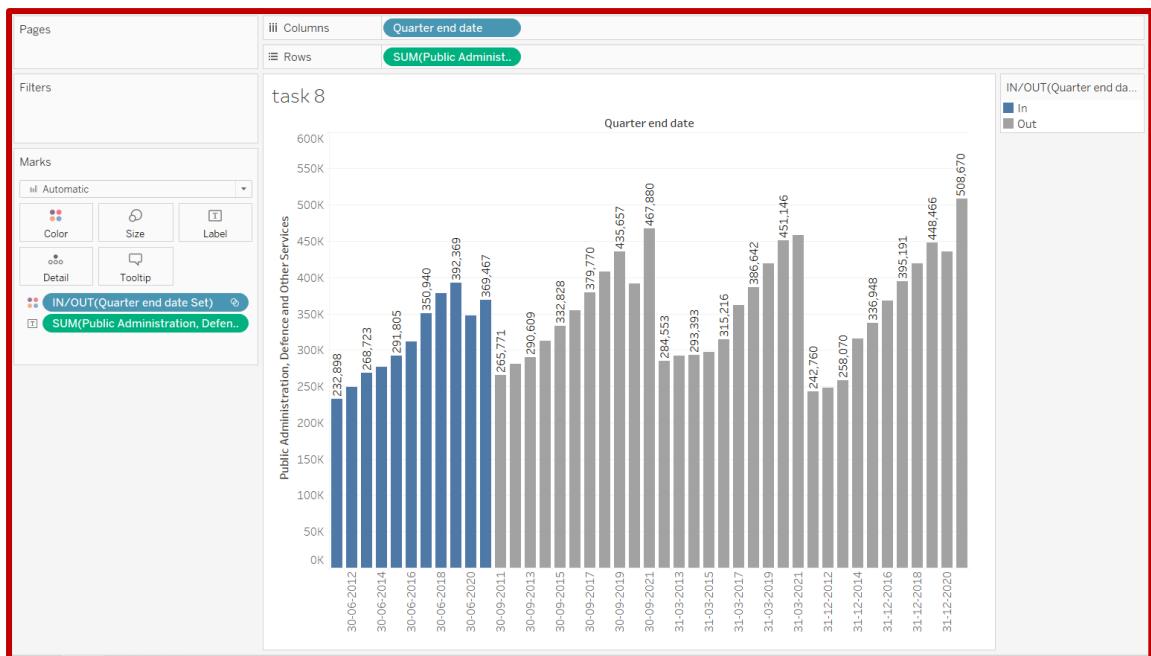




Task 8: Create a set of Q1 as per their end date on the "quarter-end date" (select all data from June month) and compare "electricity gas, water supply and other utility" and visualize using a bar chart.

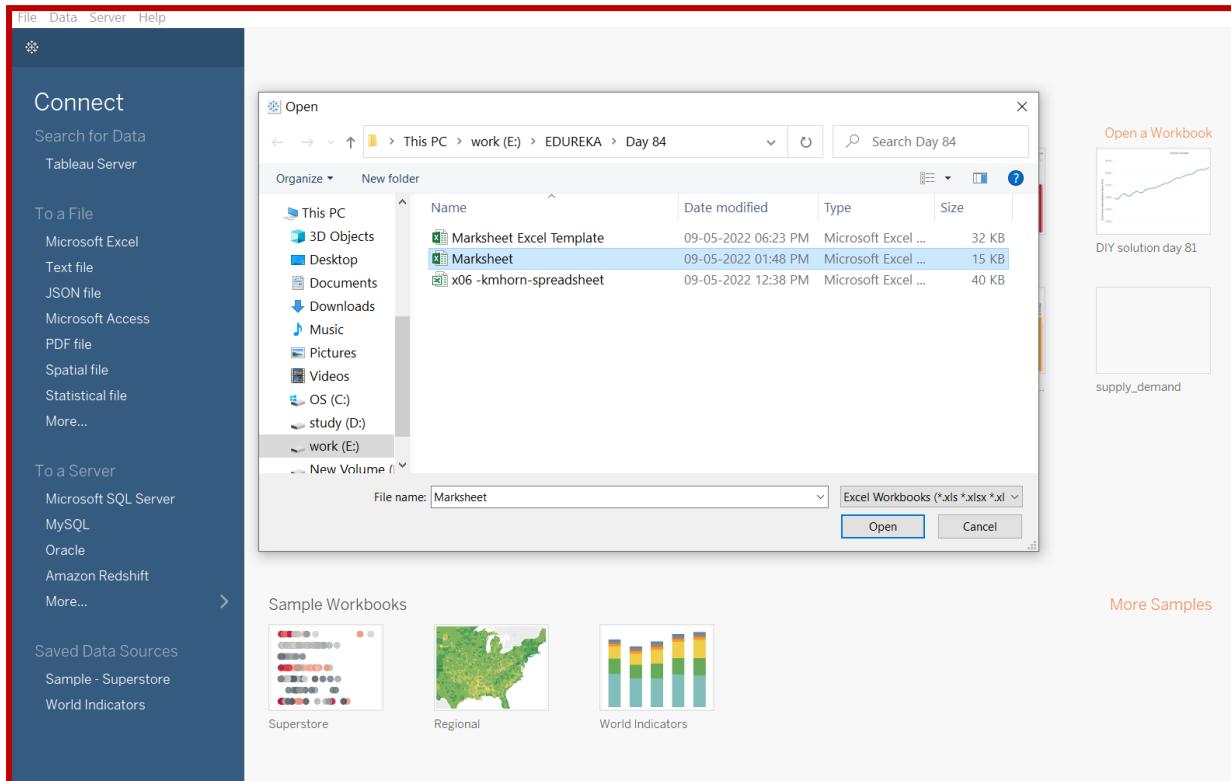
1. "Quarter-end date" to columns
2. "Public Administration, Defence, and Other Services" to rows
3. To create a new set, right-click on the field, then go to create and select set
4. Then select desired data as per your requirement and give color to the new set to observe the output

Data Science and Machine Learning Internship Program



Solution for Problem Statement 3:

Task 1: Click on “Microsoft Excel” → Select “Marksheet” → Open



Drag sheet1 to the work area and click on the sheet tab to start building your view.

Task 2: Create new calculated fields as “total” and add the total of all subjects

Right-click on the Data area → select “Create calculated fields” → name it “Total” use the formula as shown in the below figure → click on “OK”

The screenshot shows the Tableau Data Editor interface. On the left, the 'Tables' pane lists various dimensions and measures, including 'Total'. The main workspace shows a calculated field named 'Total' with the formula: `[English]+[Maths]+[Economics]+[Accountancy]+[Business Studies]`. A tooltip for the 'ABS' function is open, explaining it returns the absolute value of a number. The status bar at the bottom indicates 'The calculation is valid.' and shows 9 Dependencies.

Task 3: Create new calculated fields as “percentage” and find the student’s percentage in the exam with help of “total”

Right-click on the Data area → select “Create calculated fields” → and name it “Percentage” use the formula as shown in the below figure → , and click on “OK”

The screenshot shows the Tableau Data Editor interface. The calculated field 'Percentage' is defined with the formula: `[Total]*100/500`. A tooltip for the 'ABS' function is open, explaining it returns the absolute value of a number. The status bar at the bottom indicates 'The calculation is valid.' and shows 8 Dependencies.

Task 4: Create new calculated fields as “grade” and give them labels as if the percentage is < 40 then “F”, if < 50 then “D”, if < 60 then “C”, if < 70 then “B”, if < 80 then “A”, if <90 then “A+”, if >= 90 then “O”

Right-click on the Data area → select “Create calculated fields” → and name it “Grade” use the formula as shown in the below figure → , and click on “OK”

```
grade
IF [Percentage] < 40 THEN 'F'
ELSEIF 40<=[Percentage] AND [Percentage]< 50 THEN 'D'
ELSEIF 50<=[Percentage] AND [Percentage]< 60 THEN 'C'
ELSEIF 60<=[Percentage] AND [Percentage]< 70 THEN 'B'
ELSEIF 70<=[Percentage] AND [Percentage]< 80 THEN 'A'
ELSEIF 80<=[Percentage] AND [Percentage]< 90 THEN 'A+'
ELSEIF [Percentage] >= 90 THEN 'O'
END

The calculation is valid.
1 Dependency ▾ Apply OK
```

ABS (number)
Search
ABS Returns the absolute value of the given number.
ACOS
AND
AREA
ASCII
ASIN
ATAN
ATAN2
ATTR
AVG
BUFFER
CASE
CEILING
CHAR
COLLECT
CONTAINS
COOP

Task 5: Create new calculated fields as “result” and give a pass if the student passed in every subject,” ER” if the student just failed in a single subject, and “failed” if they got failed in more than one subject.

Right-click on the Data area → select “Create calculated fields”→ and name it “Result” use the formula as shown in the below figure → , and click on “OK”

```
Result
IF [Accountancy]<40 AND [Business Studies]< 40 THEN 'Fail'
ELSEIF [Accountancy]<40 AND [Economics]< 40 THEN 'Fail'
ELSEIF [Accountancy]<40 AND [English]< 40 THEN 'Fail'
ELSEIF [Accountancy]<40 AND [Maths]< 40 THEN 'Fail'

ELSEIF [Economics]<40 AND [Business Studies] < 40 THEN 'Fail'
ELSEIF [Economics]<40 AND [English]< 40 THEN 'Fail'
ELSEIF [Economics]<40 AND [Maths]< 40 THEN 'Fail'

ELSEIF [English]<40 AND [Business Studies] < 40 THEN 'Fail'
ELSEIF [English]<40 AND [Maths]< 40 THEN 'Fail'

ELSEIF [Accountancy]<40 OR [Business Studies]< 40 OR [Economics]< 40 OR [English]< 40 THEN 'ER'
ELSE 'Passed'
END

The calculation is valid.
1 Dependency ▾ Apply OK
```

ABS (number)
Search
ABS Returns the absolute value of the given number.
ACOS
AND
AREA
ASCII
ASIN
ATAN
ATAN2
ATTR
AVG
BUFFER
CASE
CEILING
CHAR
COLLECT
CONTAINS
COOP

Task 6: Use a packed bubble chart and figure out how many students are passed/fail or got ER

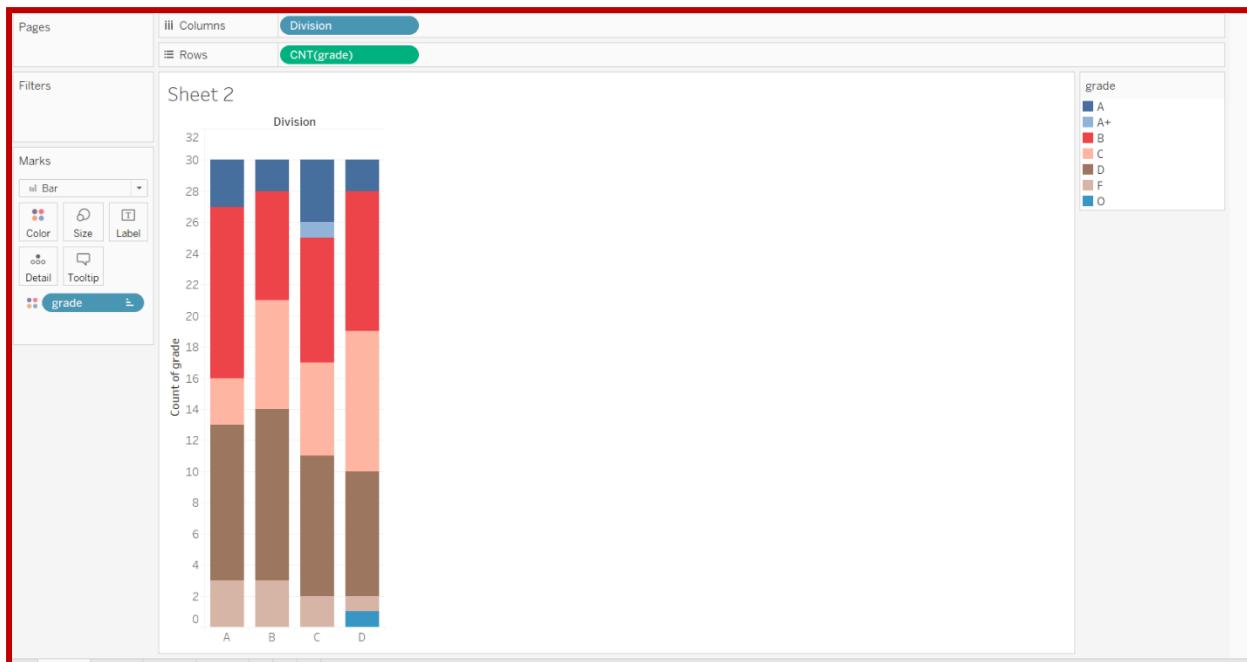
1. “Name” to rows
2. “Result” to columns

Right Click on “Name” in the rows field and set measure as “count” → Go on the show me the option and select “Packed bubble” graph



Task 7: Use the vertical bar plot to distinguish students by their grades in each division
(task 7)

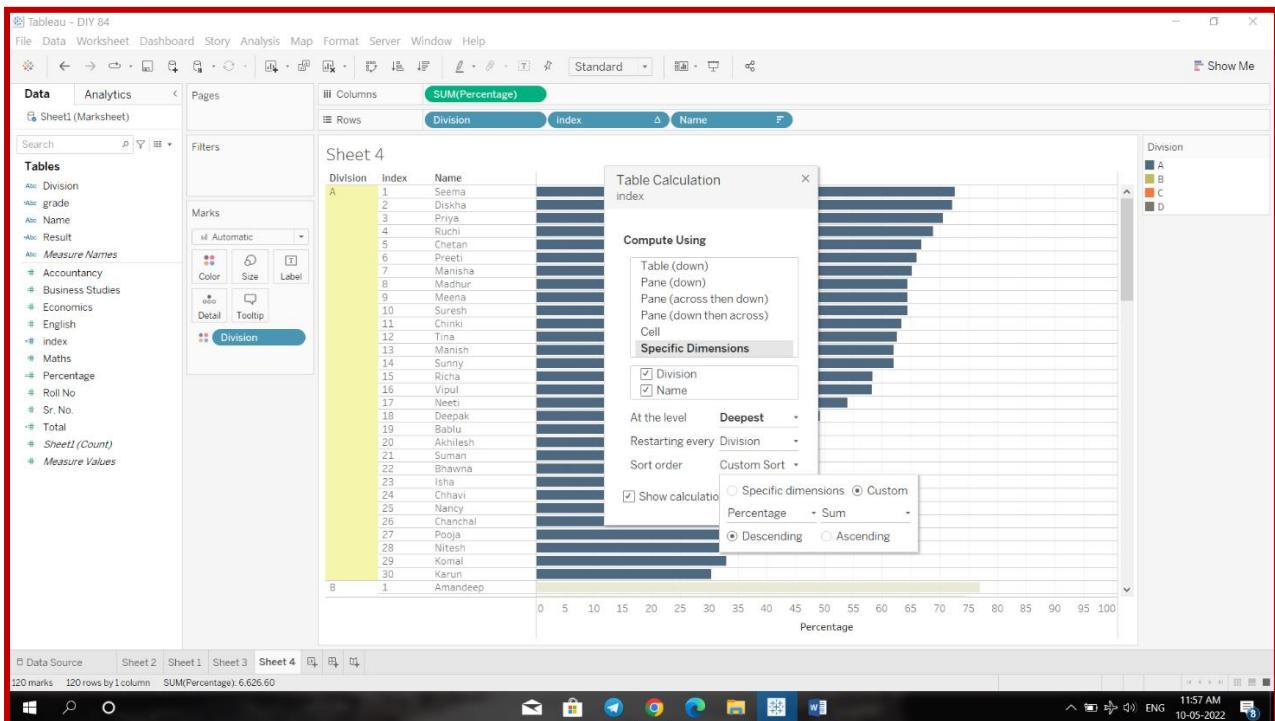
1. “grade” to rows, same as step 7 set mesure as count
2. “Division” to columns
3. Give color to “grades”



Task 8: Find out the top 5 students from each division

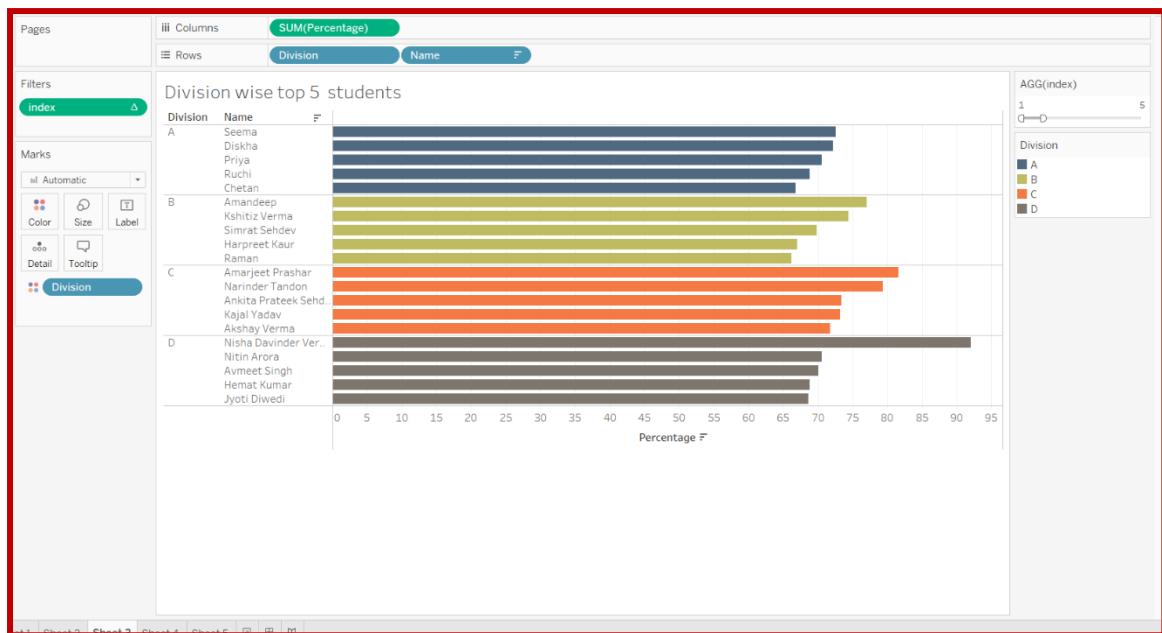
1. “Percentage” to column,
2. “Division”, “Name” to row
3. give color to “division”

Right-click on the Data area → select “Create calculated fields” → name it “Index” and use the index function → click on “OK” → drag index and drop in the Rows area → Right click on it → and select “Edit Table Calculation” → In compute using select “Specific Dimension” → In Restarting every select “Division” → In sort order select custom → percentage and descending → close →



Re drag index to filter area → Select all values and click ok → Right click on index and select continuous → Select range 1-5 and click ok →

Right-click on Name and select “sort” → select sort by as “field”, descending, → Field name as “percentage” → close → give color to “Division”



Task 9: Filter out data where students got more than 60% and less than 80% as result and sort them in descending order

1. “Percentage” to column,
2. “Name” to row

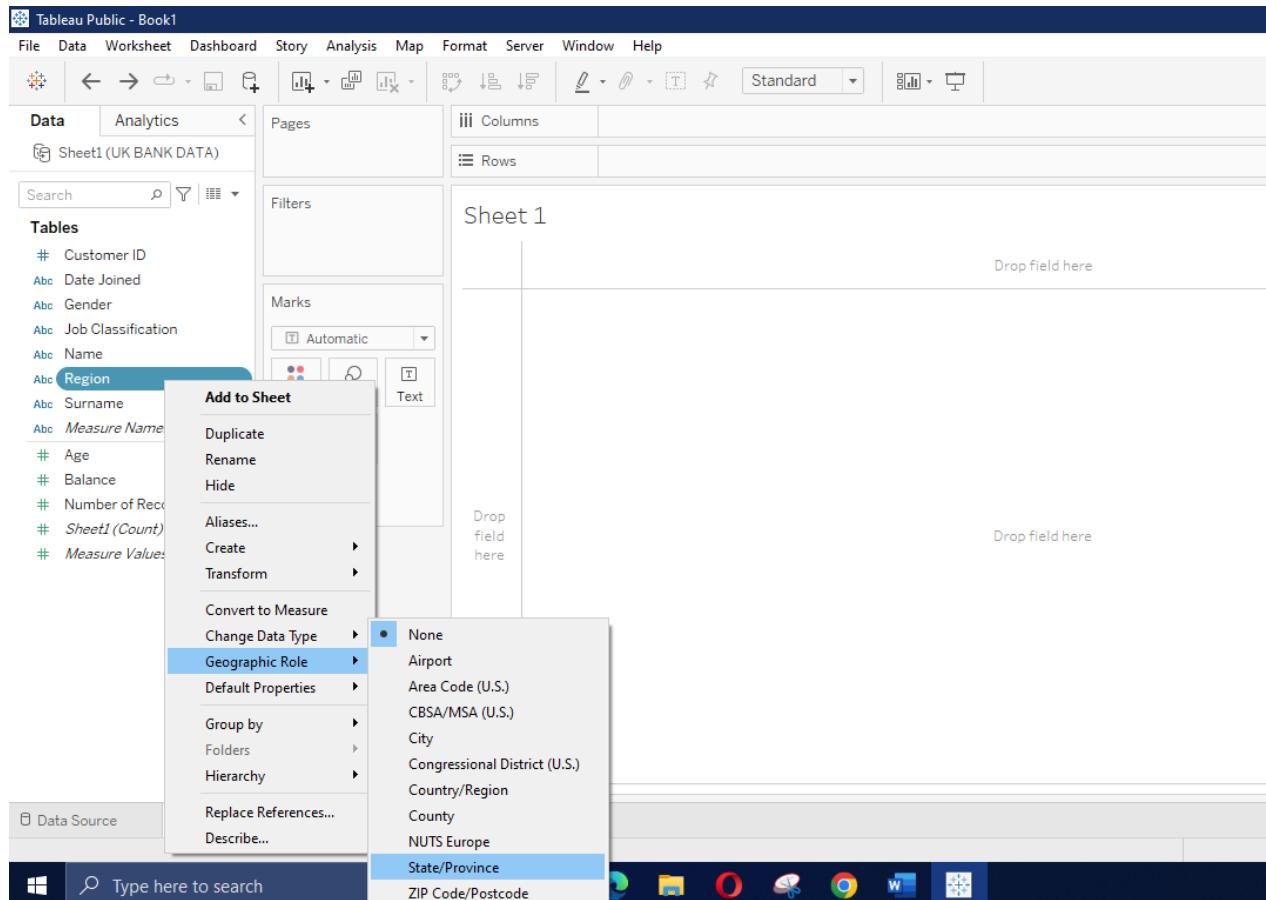
Drag “Percentage” to filter area → select “sum” for option → Select asked a range of percentage → Right-click on Name and select “sort” → select sort by as “field”, descending, → Field name as “percentage” → close → give color to “Division”

Solution for Problem Statement 4

Task 1: Create a Map chart that will display the number of transactions processed in each Country.

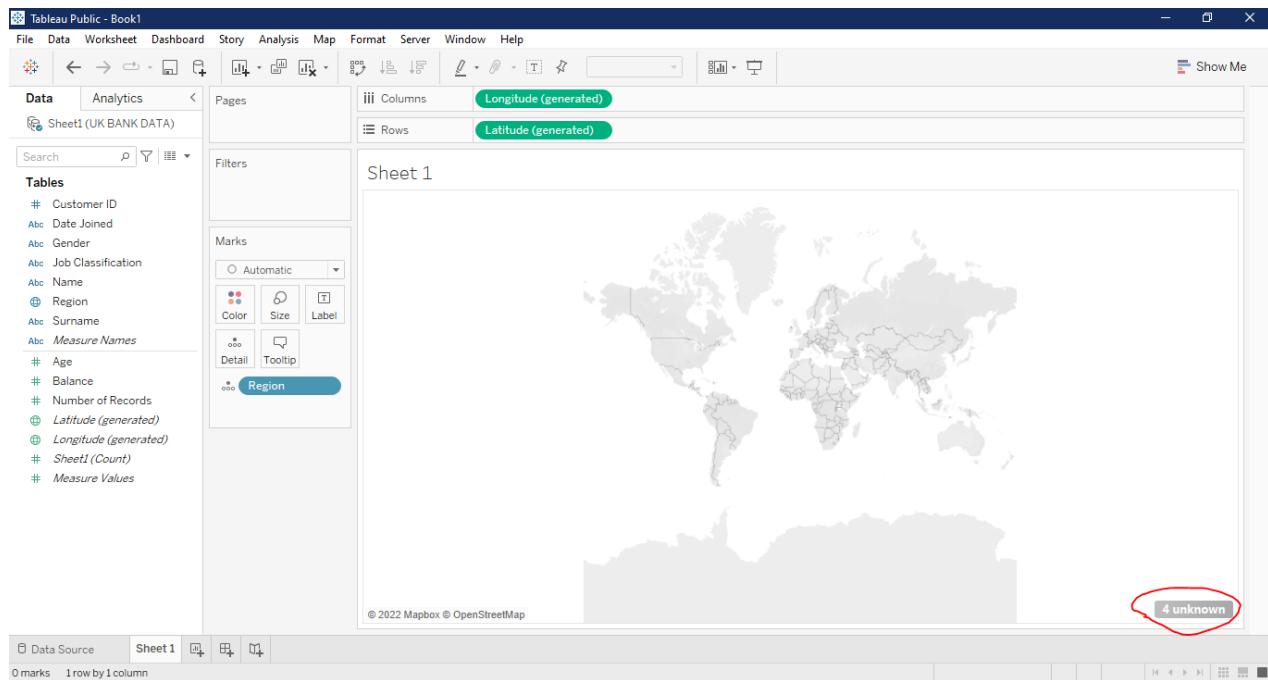
Start by assigning a Geographical role to Region. Since Tableau is considering as a string as of now.

Right-click on the Region field. Go to Geographic role within that select “State/Province”. Refer to the picture below: -

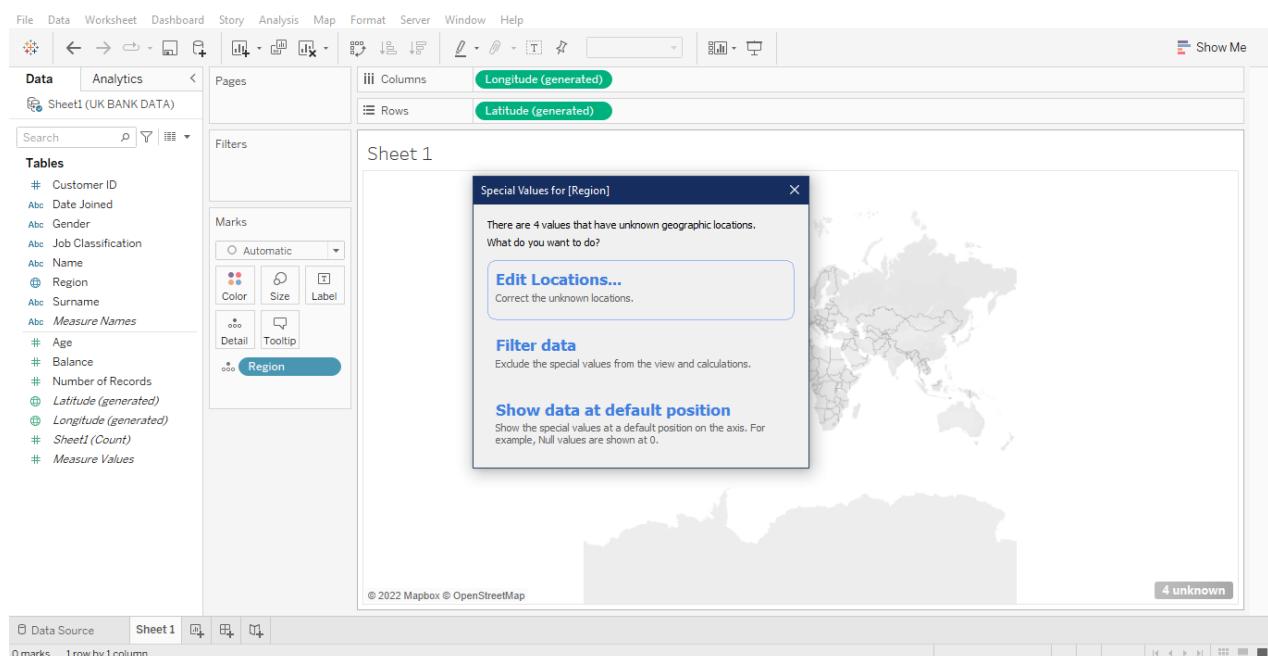


Drag region to the Pane. Select the unknown box located at the bottom right side of the visualization pane. Refer to the picture below: -

Data Science and Machine Learning Internship Program

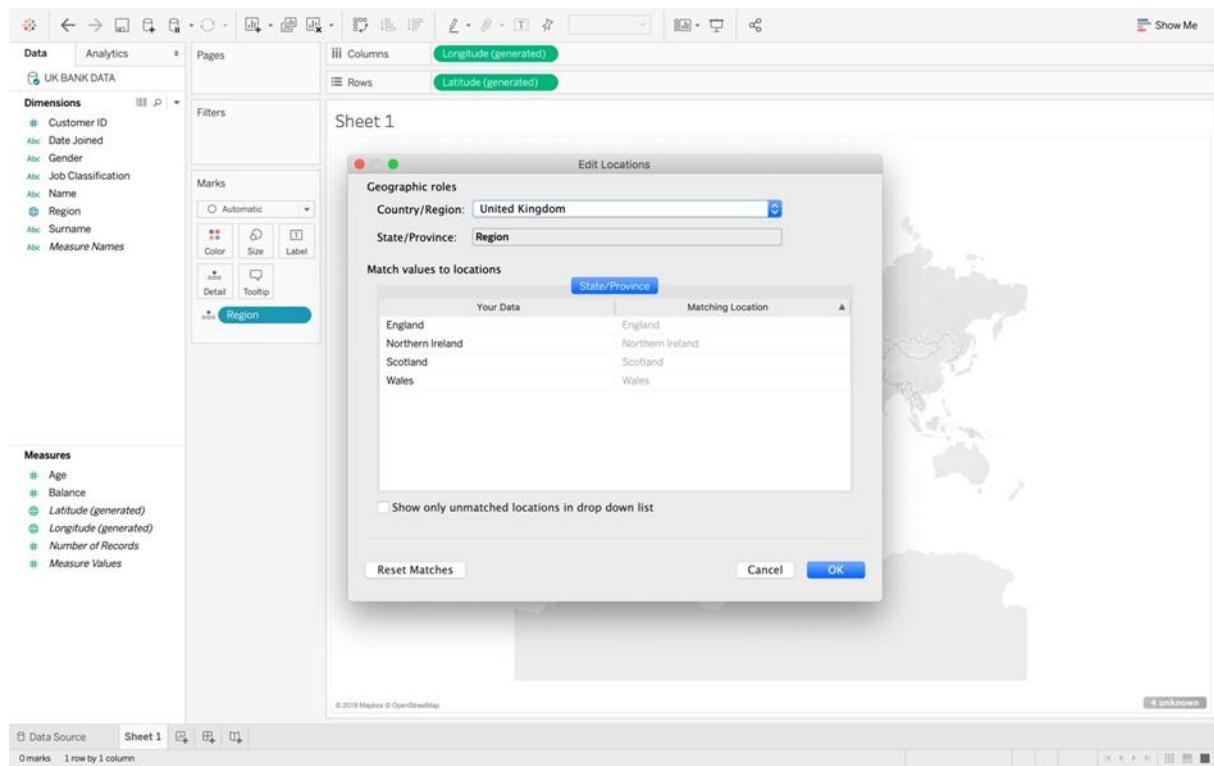


In the dialogue box, select Edit locations. Refer to the picture below:

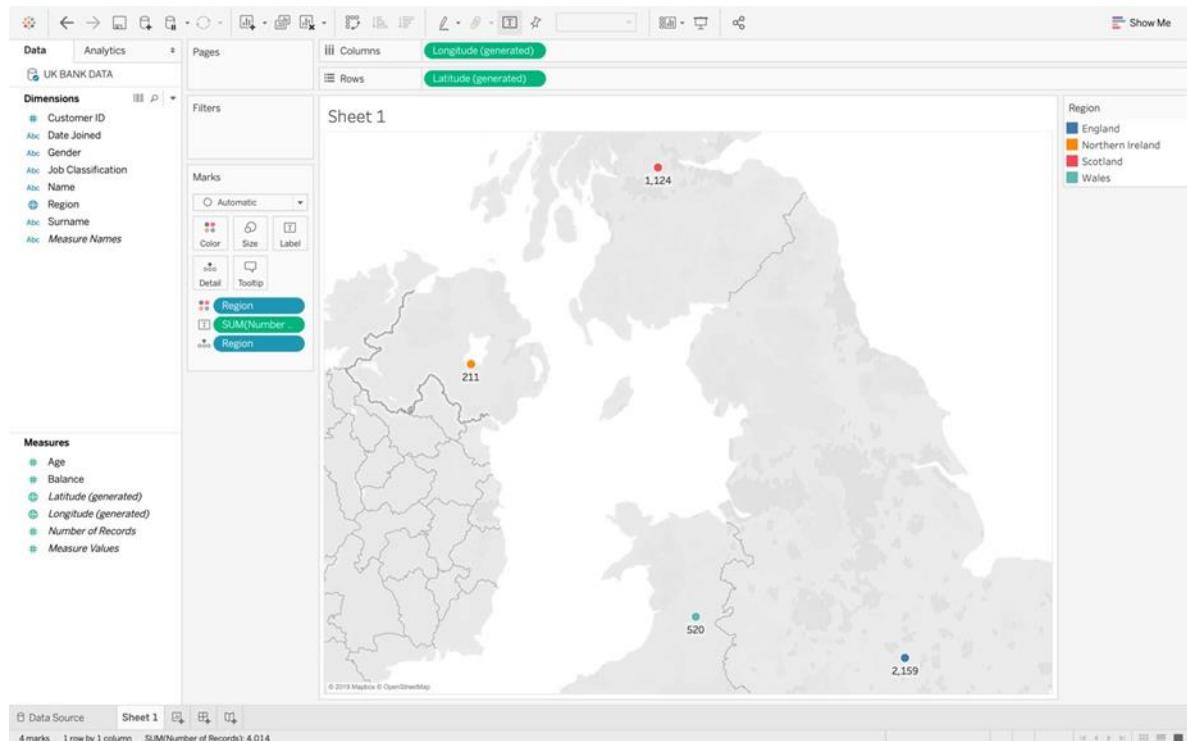


In the Country/Region menu change the Country to the United Kingdom. Refer to the picture below: -

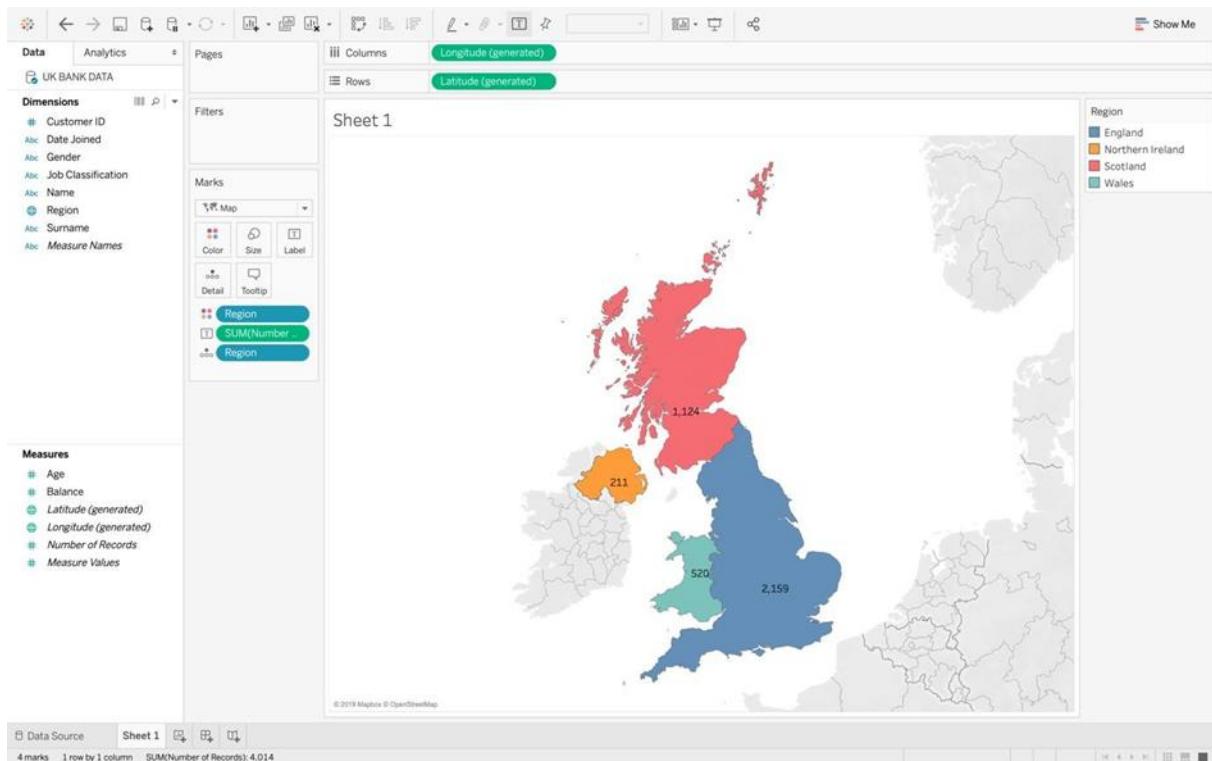
Data Science and Machine Learning Internship Program



Drag Region to Color and Number of records to Label marks card. This is how it should look like: -

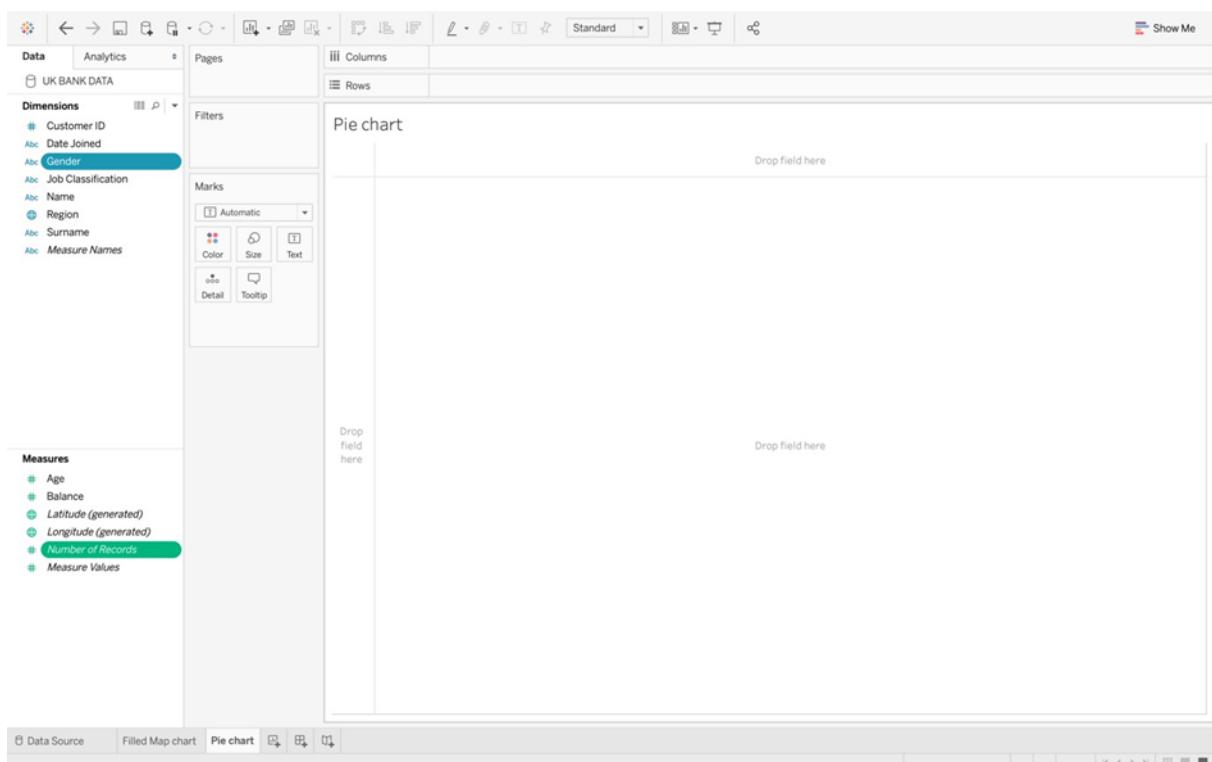


To make the visualization informative change the chart to a filled map. If for any reason the fields change revert them back. This is how it should look like: -



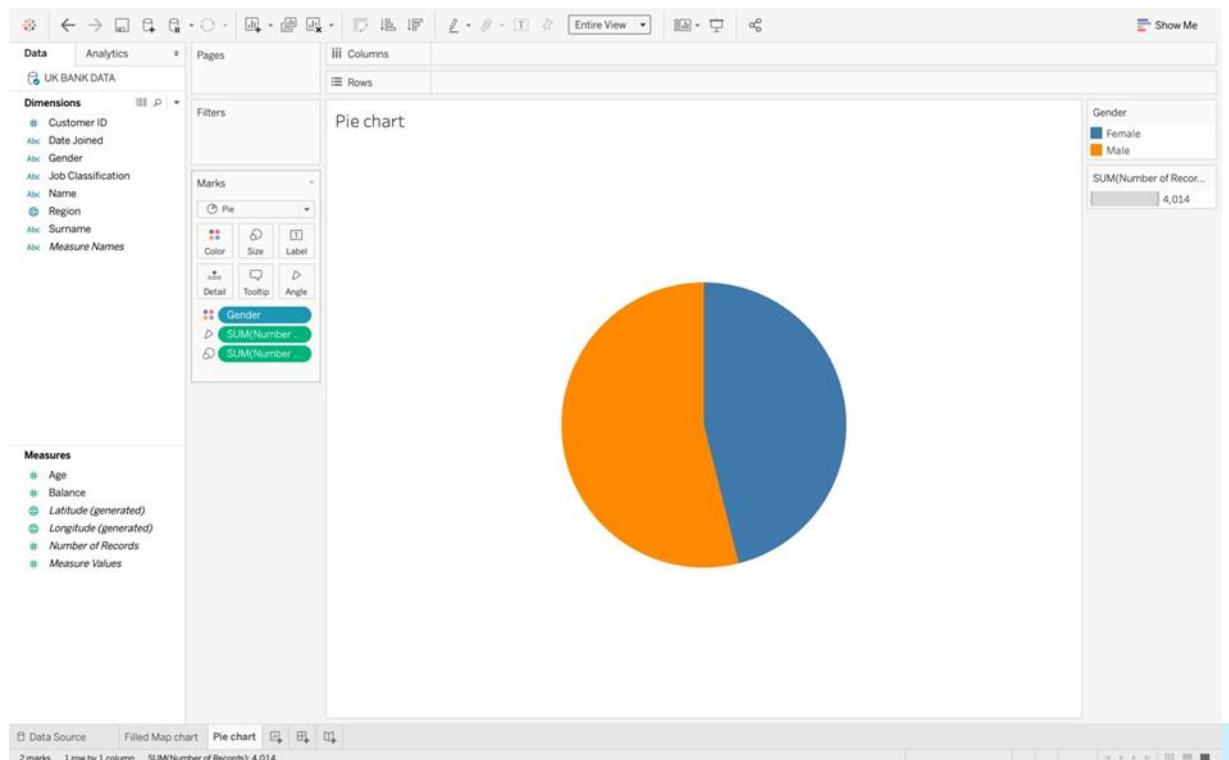
Task 2: Create a Pie chart displaying the percentage of each transaction processed between genders.

We will start by selecting the Gender and Number of transactions in the data pane section. Refer to the picture below: -

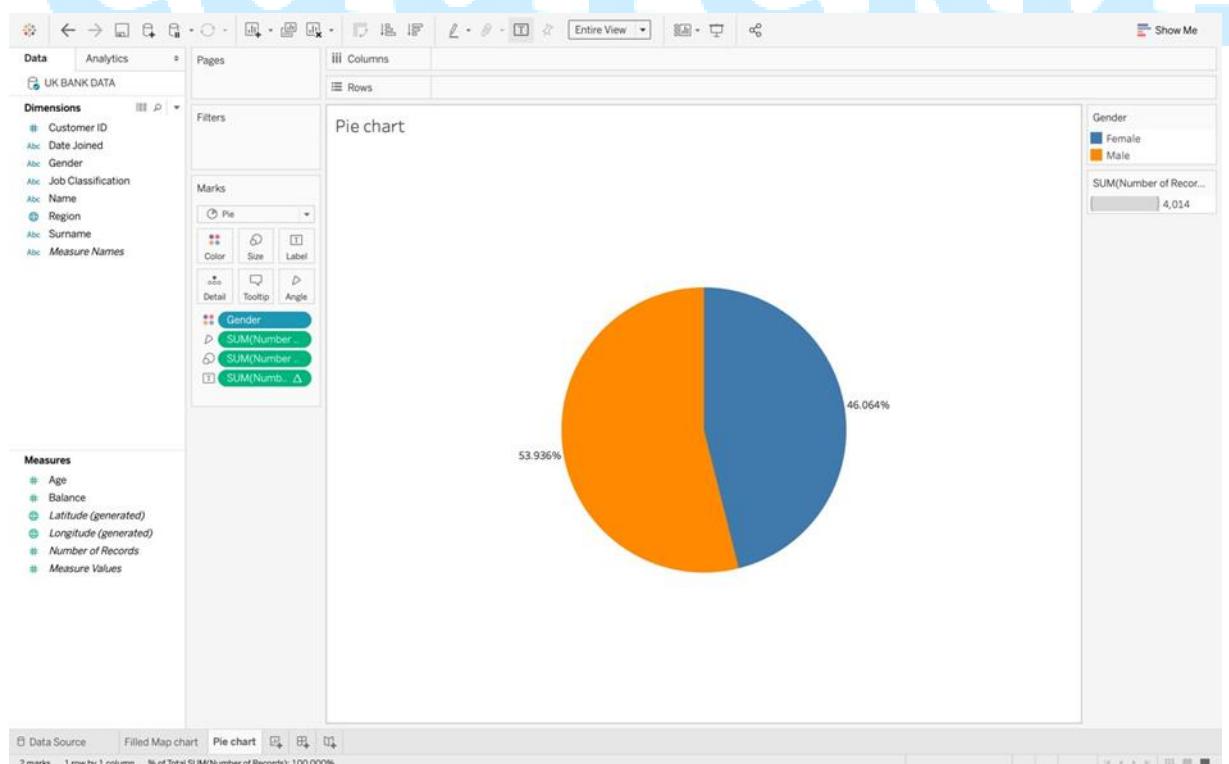


Data Science and Machine Learning Internship Program

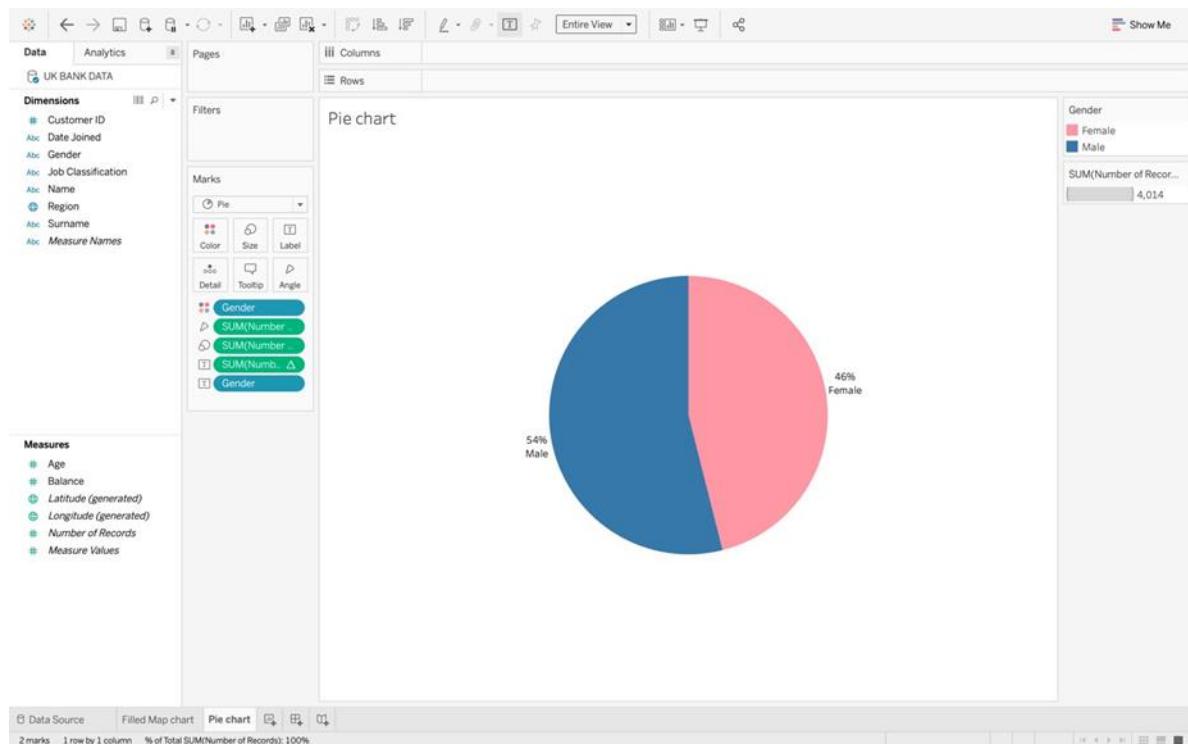
Go to “Show Me” section and select Pie chart also change the view to Entire View. This is it should look like: -



Drag Number of Records to Label marks card also apply “Percent of Total” table calculation. This is how it should look like:



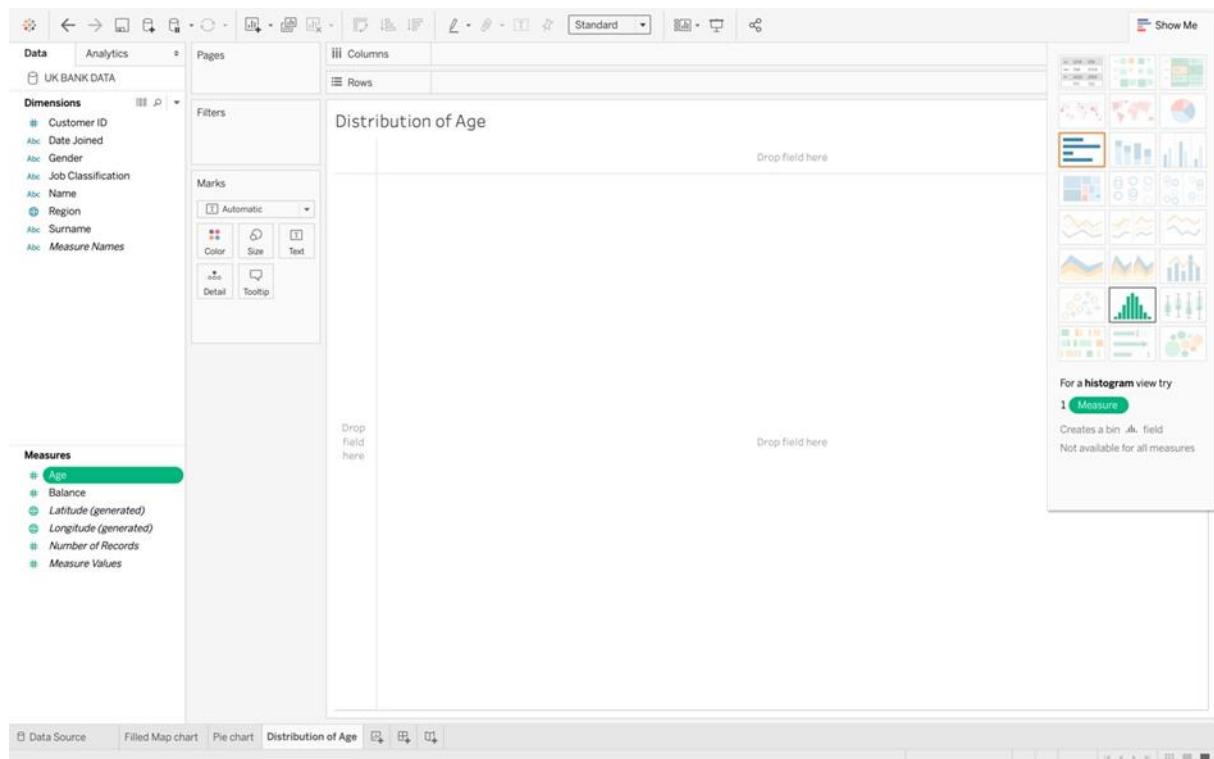
Add Gender to Label as well also Format percentage and Color if required. This is what it will look like:



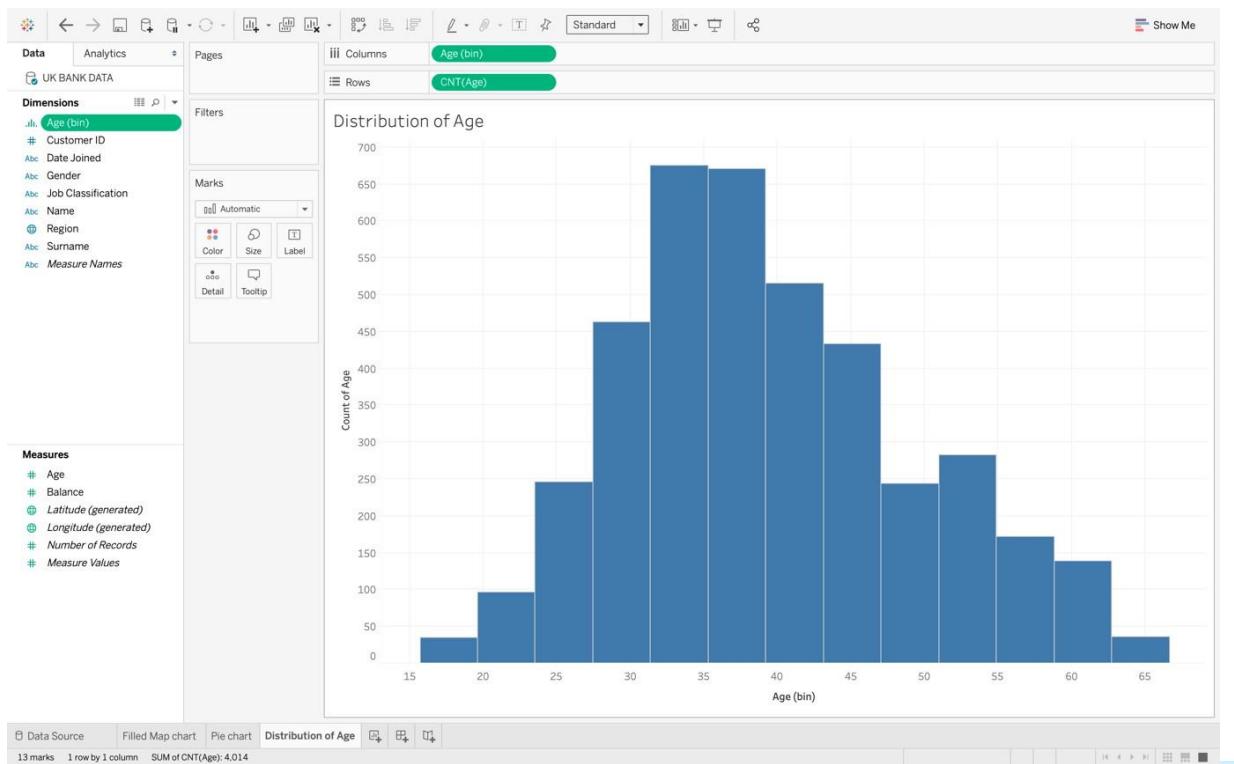
Task 3: Creating a Histogram displaying the distribution of age and transactions processed.

Start by creating a Histogram using Age

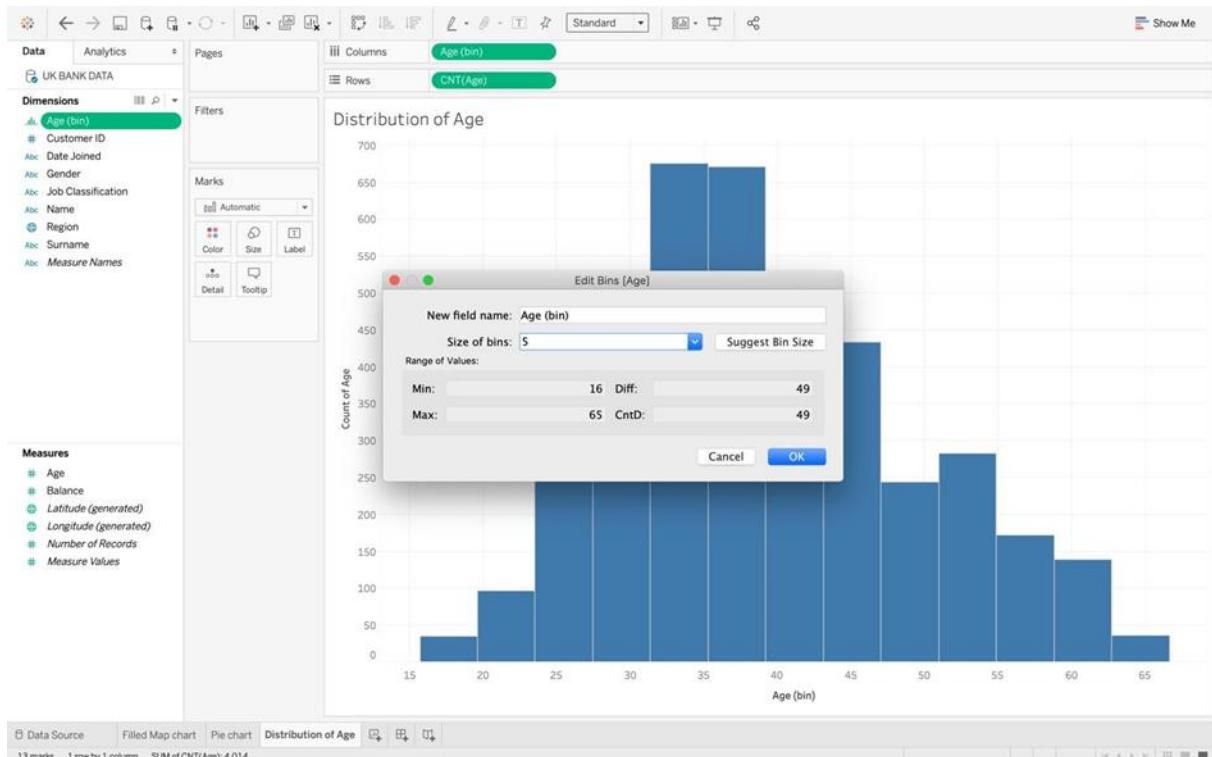
Select Age, go to Show Me, and select histogram. Refer to the picture below: -



Once selected this is how it should look like: -

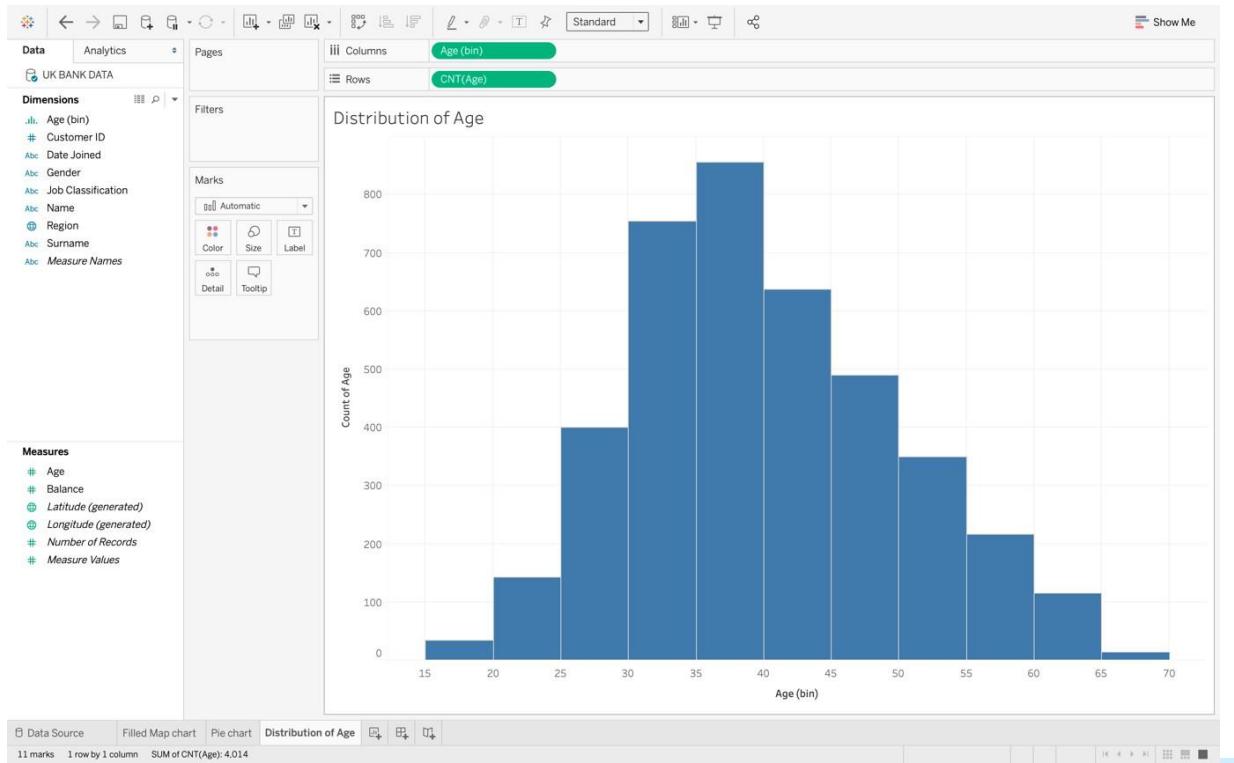


Edit the size of the Range by Right-clicking on the new Age(bin) field and selecting Edit. In the Dialogue box select 5 as Size of bins (Though it's not necessary as nothing is Mentioned). This is how it should look like: -

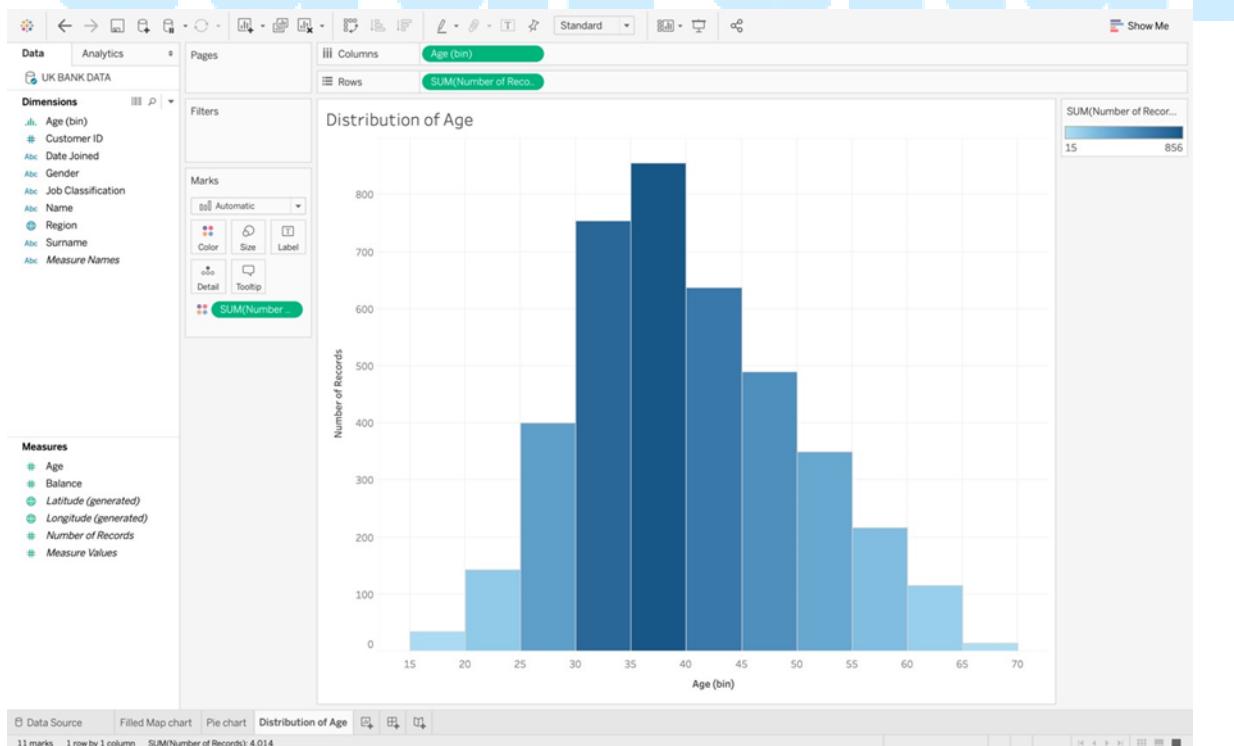


Data Science and Machine Learning Internship Program

Once changed this is how the chart should look like: -



Drag Number of Records to the Rows shelf and the Colors Marks card as well also remove CNT(Age) from the Rows shelf. This is how it should look like: -

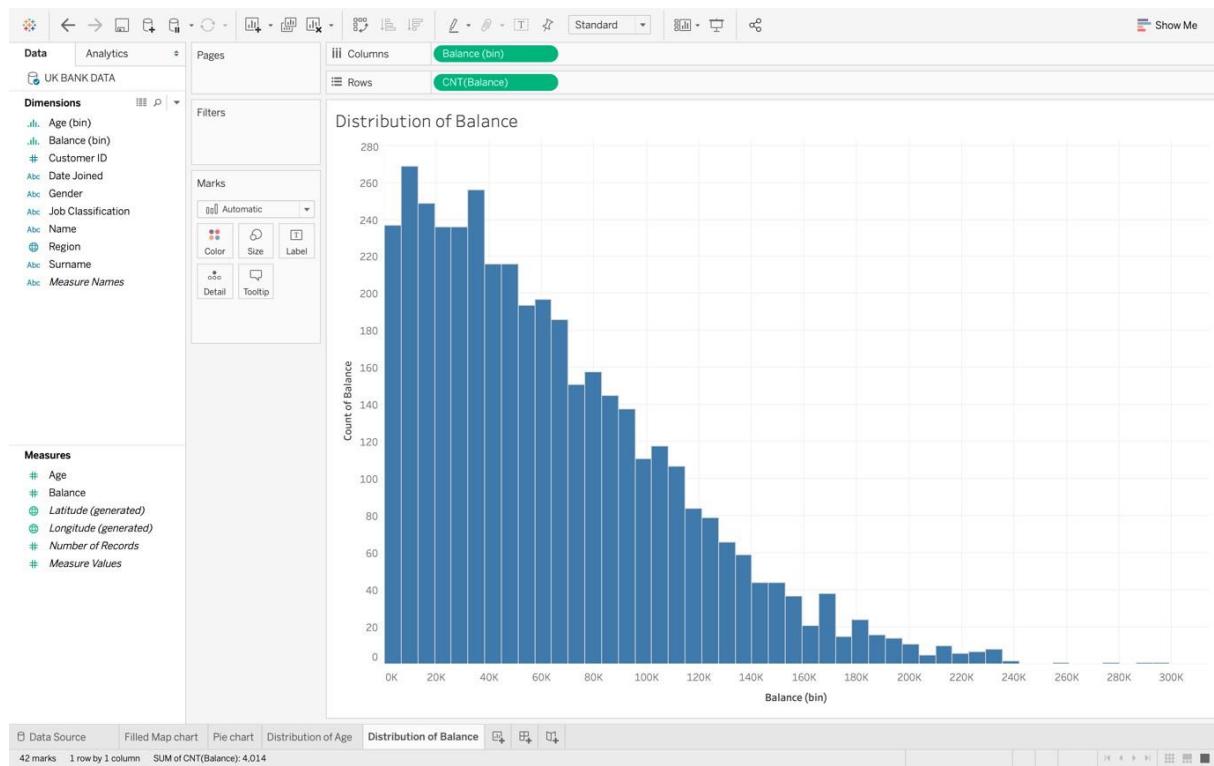


Data Science and Machine Learning Internship Program

Task 4: Create a Histogram displaying the distribution of balance and transactions processed.

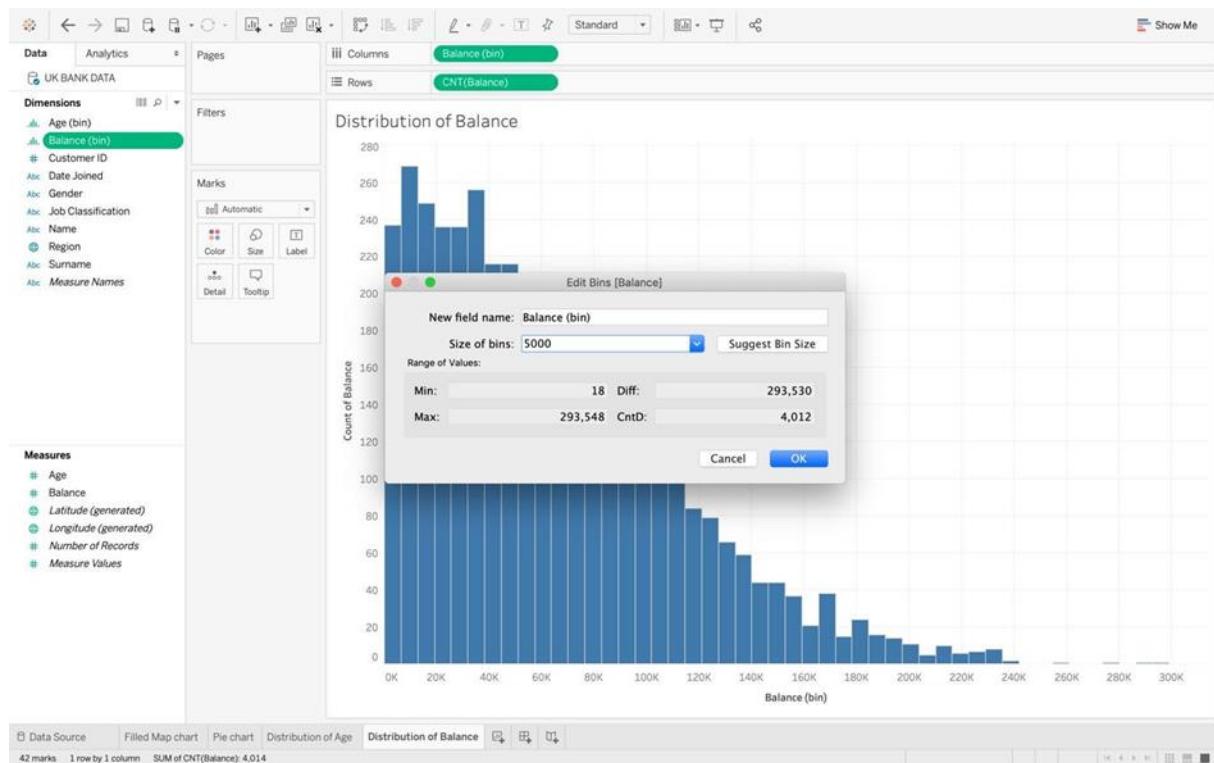
Start by creating a Histogram using Balance.

Select Balance, Go to Show me and select Histogram. Once created this is how the Histogram should look like: -

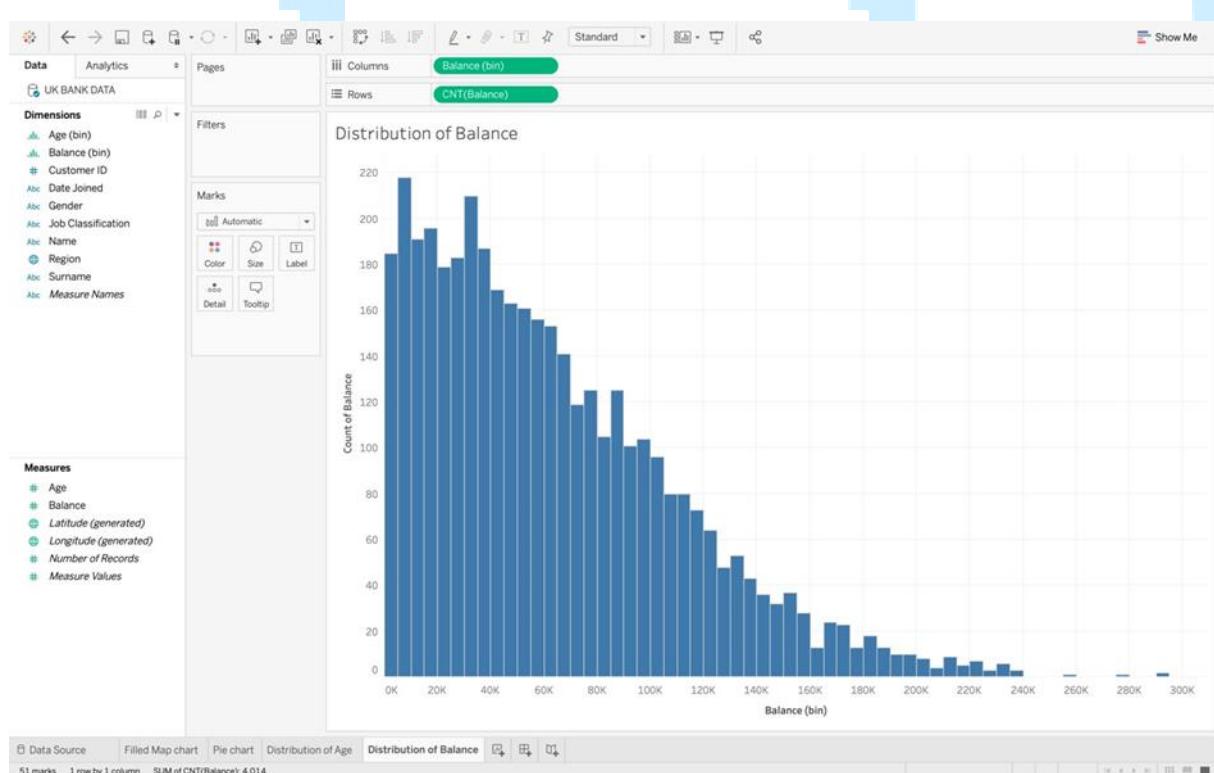


Edit the Balance (Bin) and set Bin size to 5000 (Though it's not necessary as nothing is Mentioned). This is how it should look like: -

Data Science and Machine Learning Internship Program

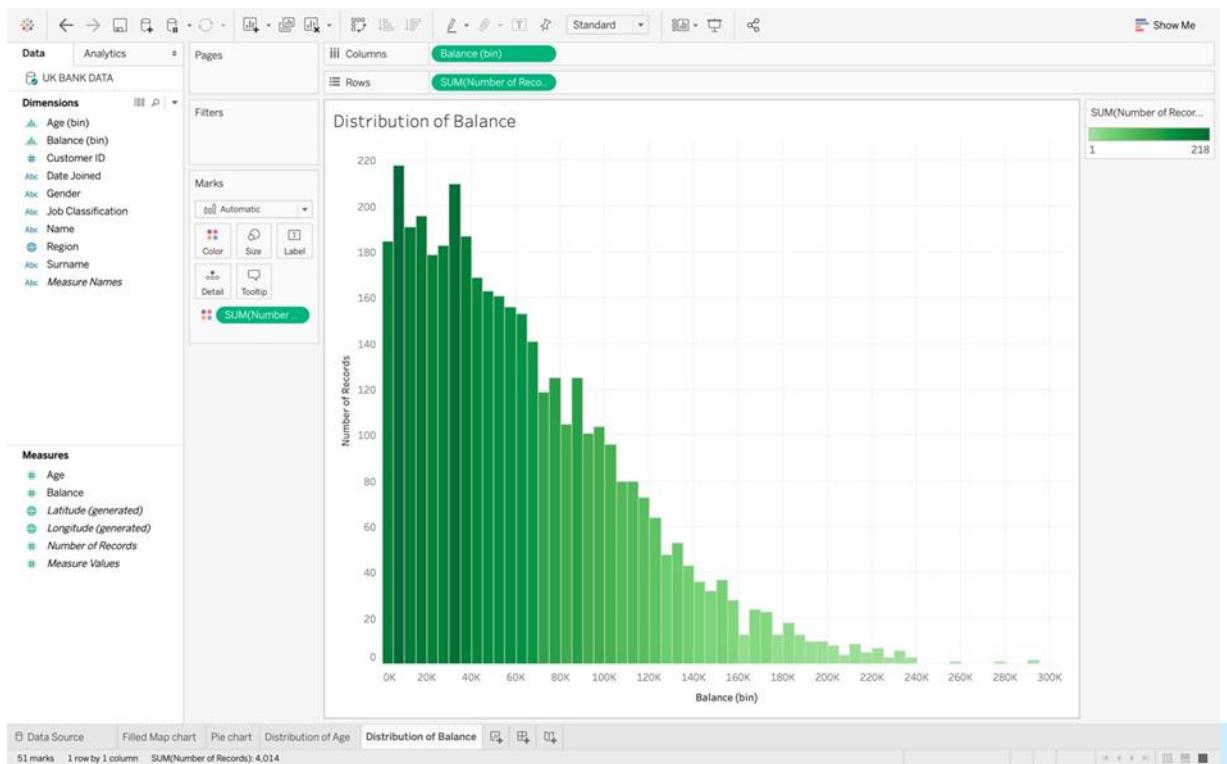


Once changed this is how it should look like:



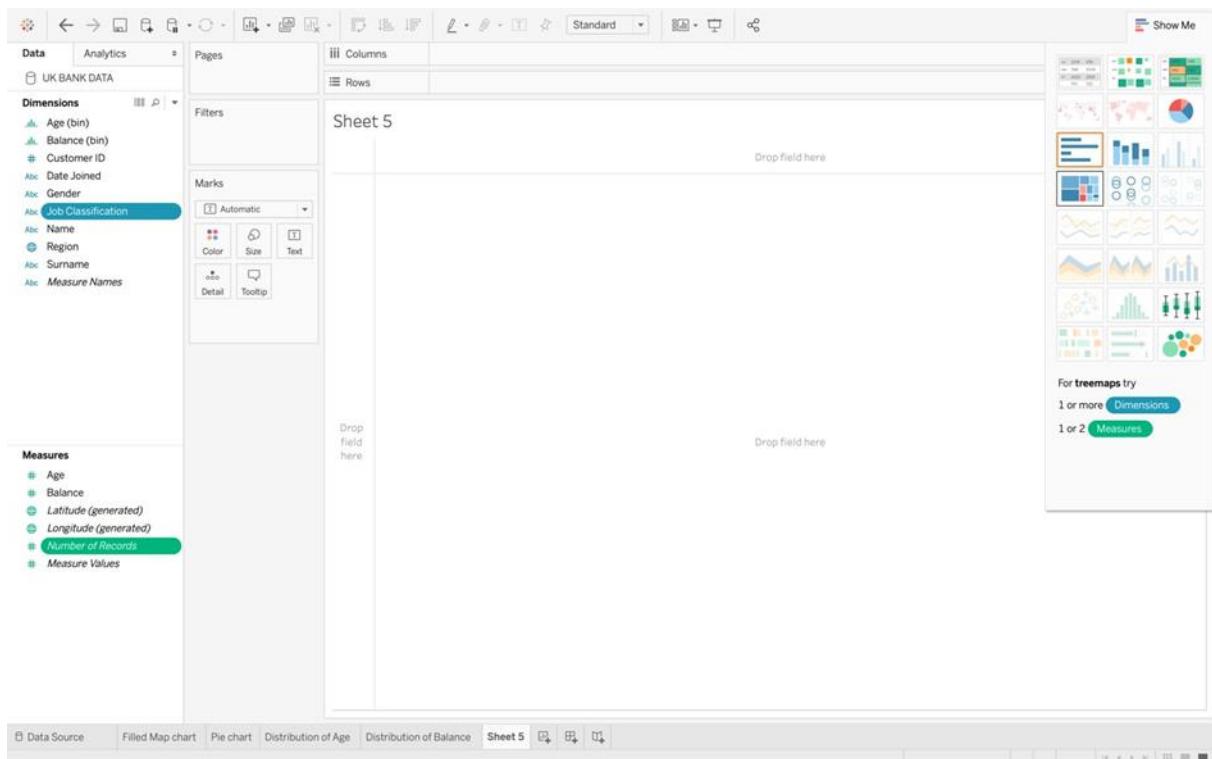
Data Science and Machine Learning Internship Program

Add the Number of Records to the Rows shelf and Colors marks card as well also change the Color of the chart. This is how it should look like: -



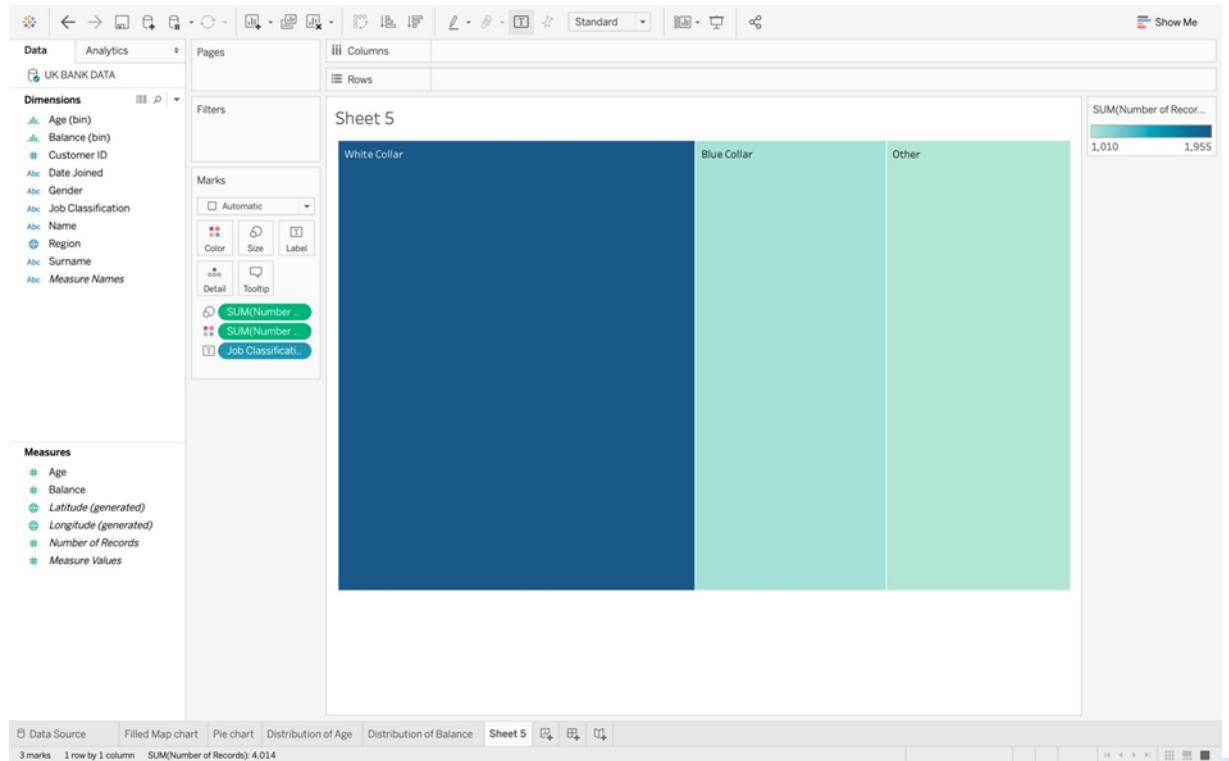
Task 5: Create a Tree-map displaying the number of transactions processed within the classification of each job.

Select Job classification and Number of Records in the Data Pane. Go to Show me section and select Tree-map. Refer to the picture below: -

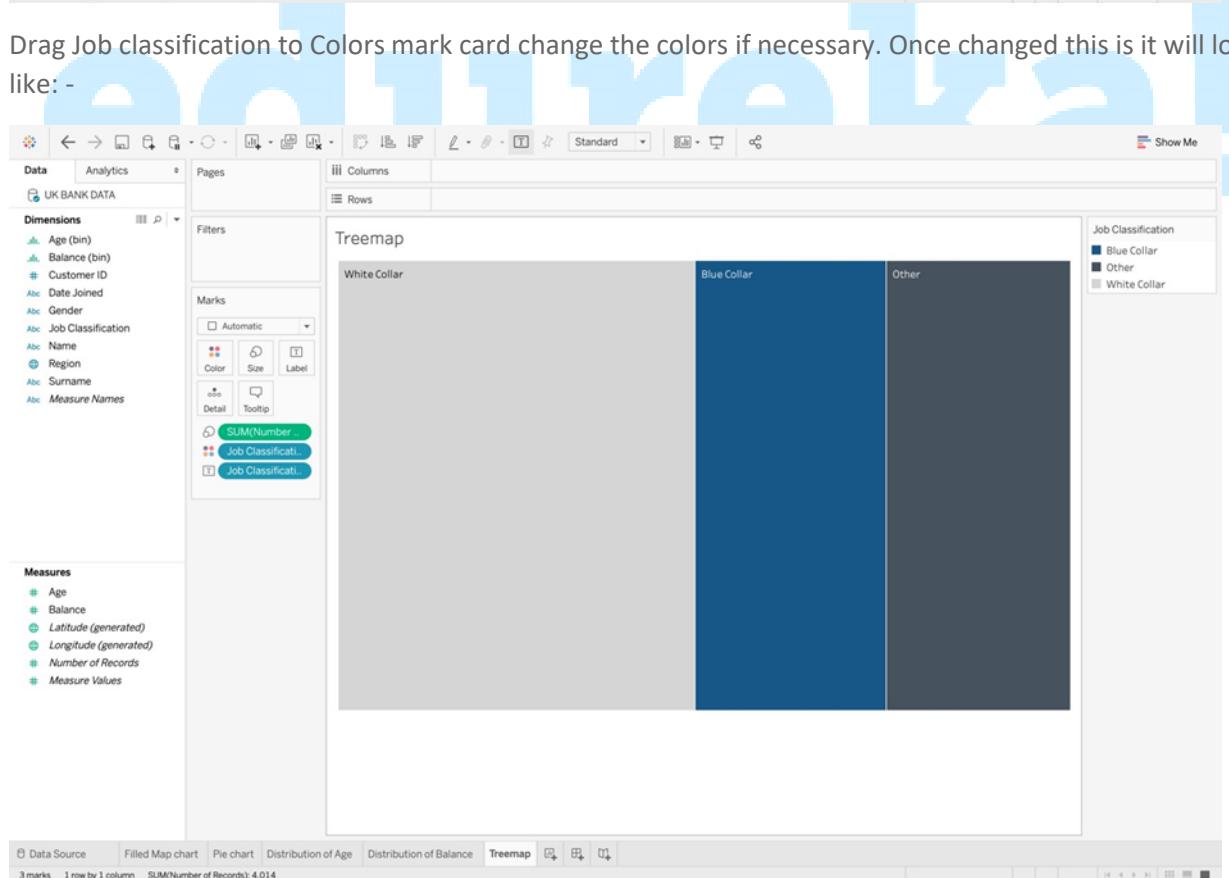


Data Science and Machine Learning Internship Program

Once selected this is how it should look like:

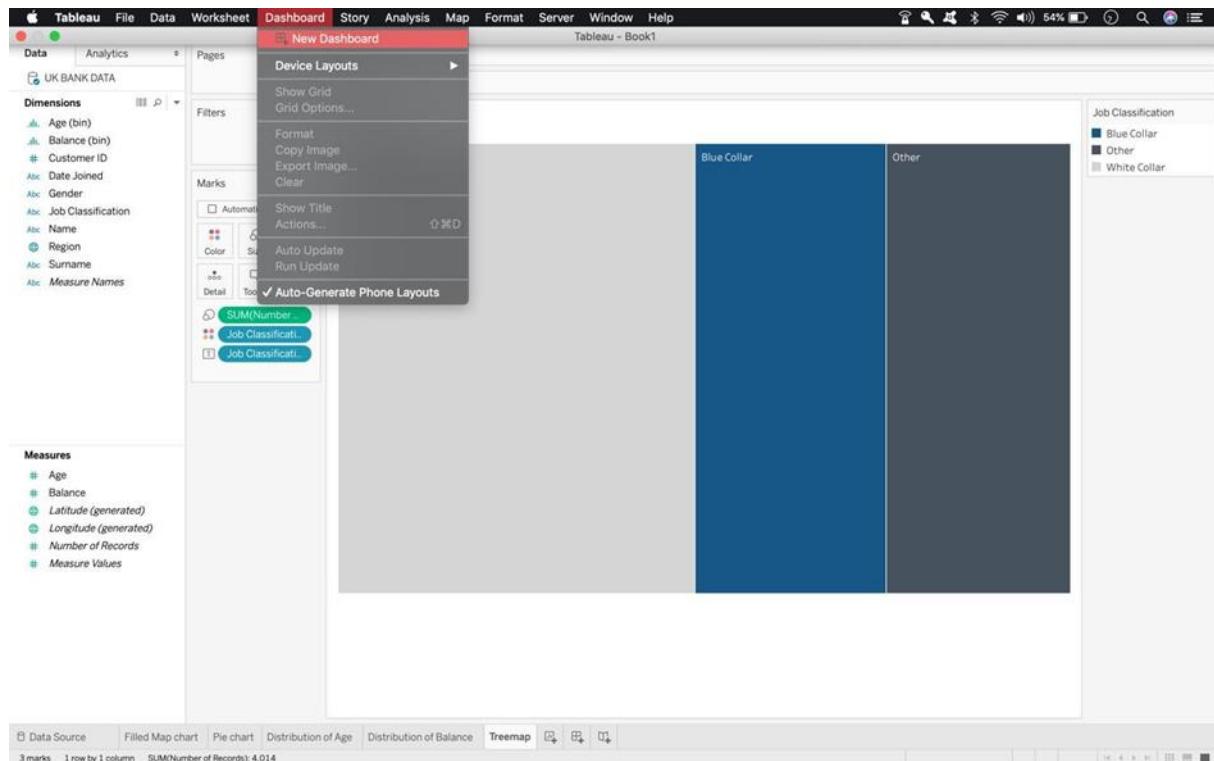


Drag Job classification to Colors mark card change the colors if necessary. Once changed this is it will look like: -

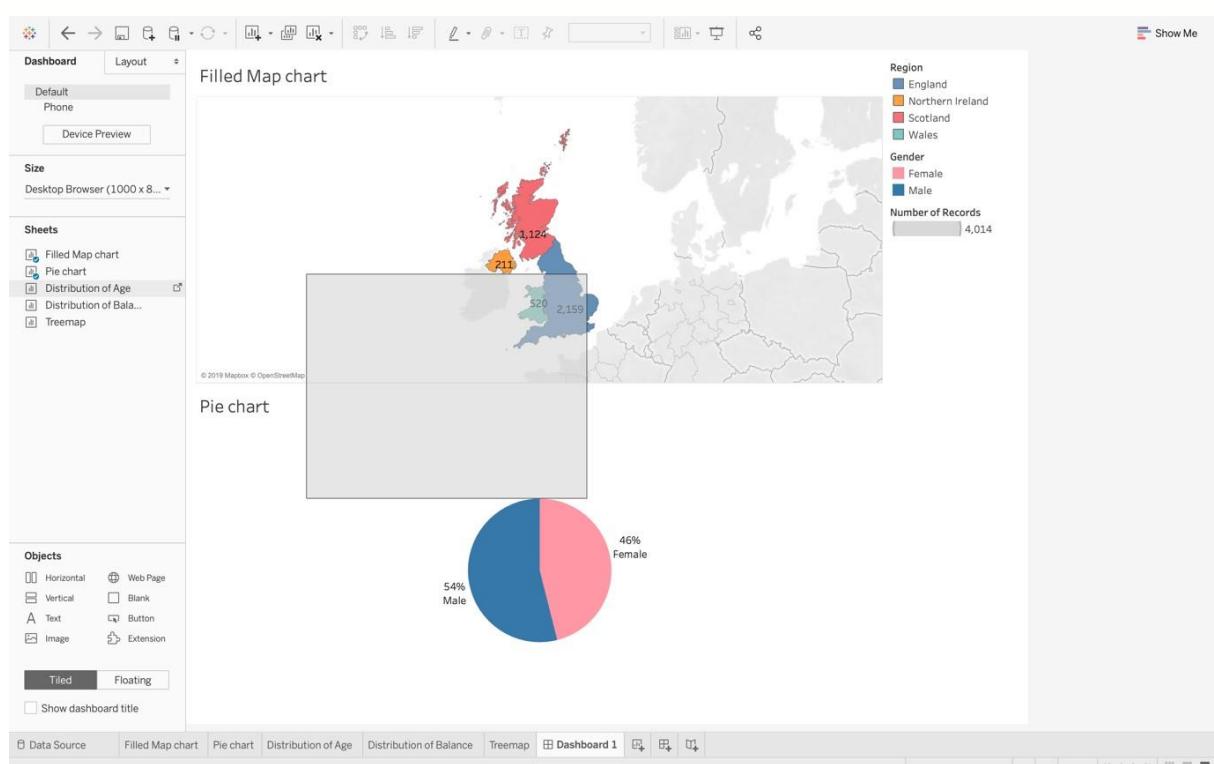


Task 6: Create a Dashboard using all the charts. Make use of Action Filters as it will further improve the analyzing experience for the management.

Go to Dashboard Menu and select “New Dashboard”. Refer to the picture below: -

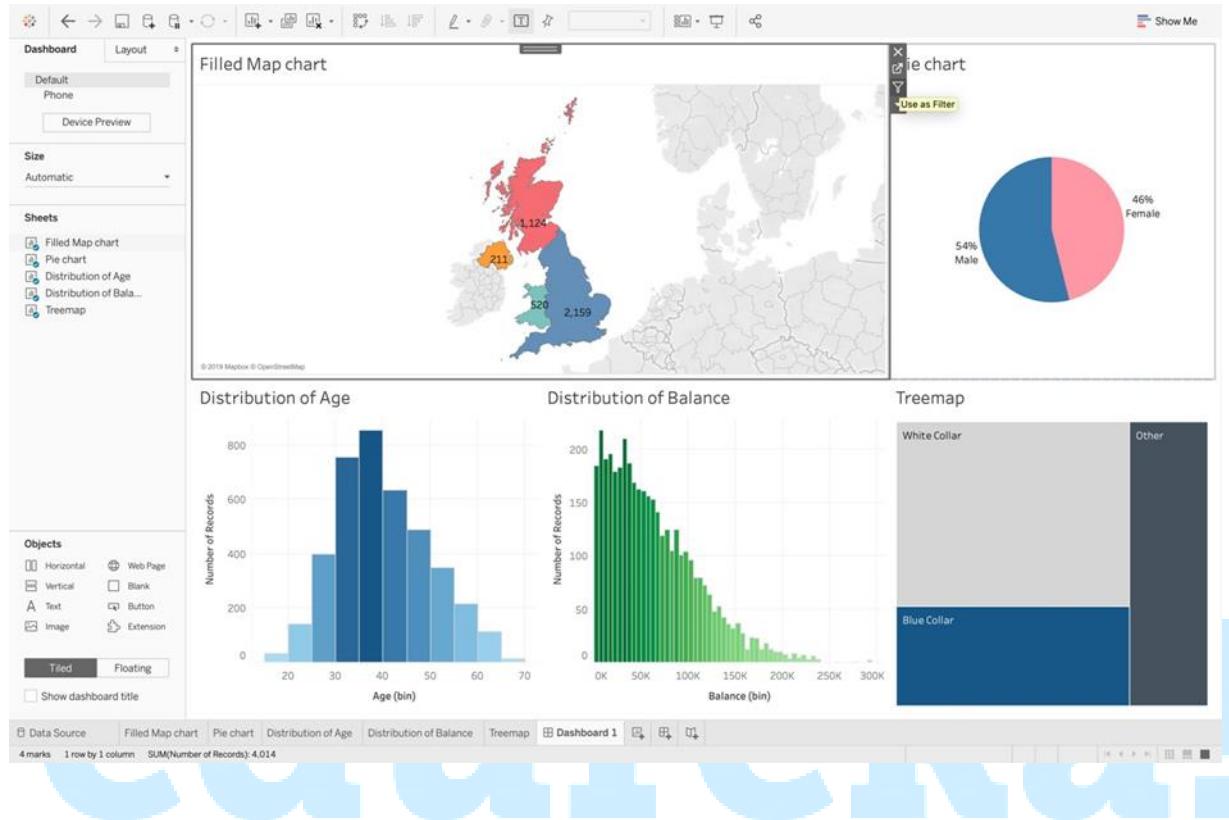


Manually Drag all the sheets to the Dashboard Container. Refer to the picture below: -



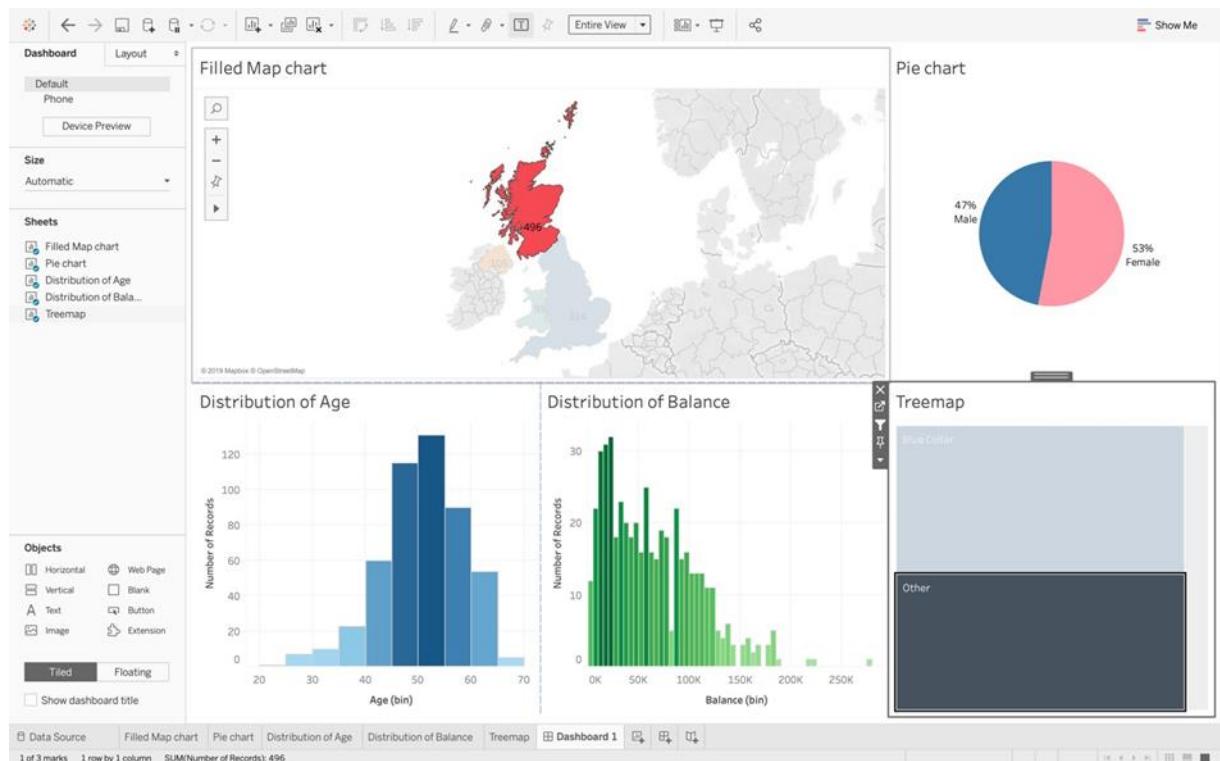
Once you add all the sheets change the size to Automatic. This is how it will look like: -

Manually select on each chart and Turn on “Use as Filter” button. Refer to the picture below: -



Doing this will make the whole Dashboard interactive. You can notice if I turn on Action Filter and Select Scotland in the Map chart and Other in Tree-map the remaining charts will also be adjusted accordingly. Use the picture for reference: -

Data Science and Machine Learning Internship Program



edureka!