## **NAME** : DHANSHREE B. RAJPUT & ADITI RANGANATHAN

## **CLASS** : M.Sc. DATA SCIENCE AND BIG DATA ANALYTICS

## **SUBJECT** : DATA VISUALIZATION USING TABLEAU

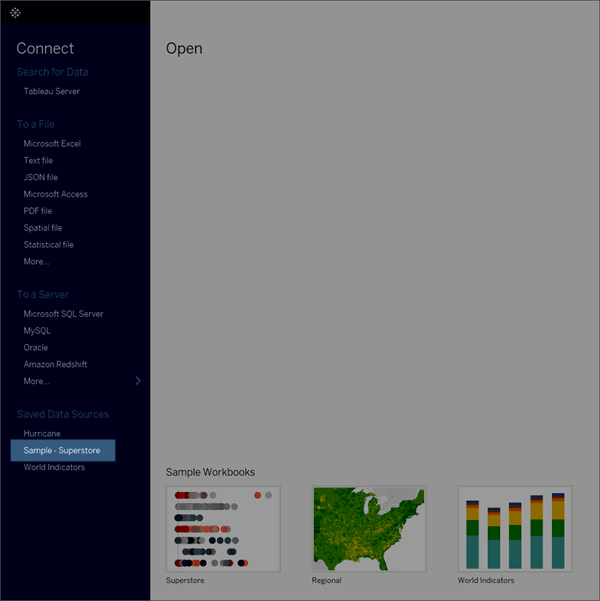
## **TOPIC** : INDIAN AGRICULTURAL CROP PRODUCTION

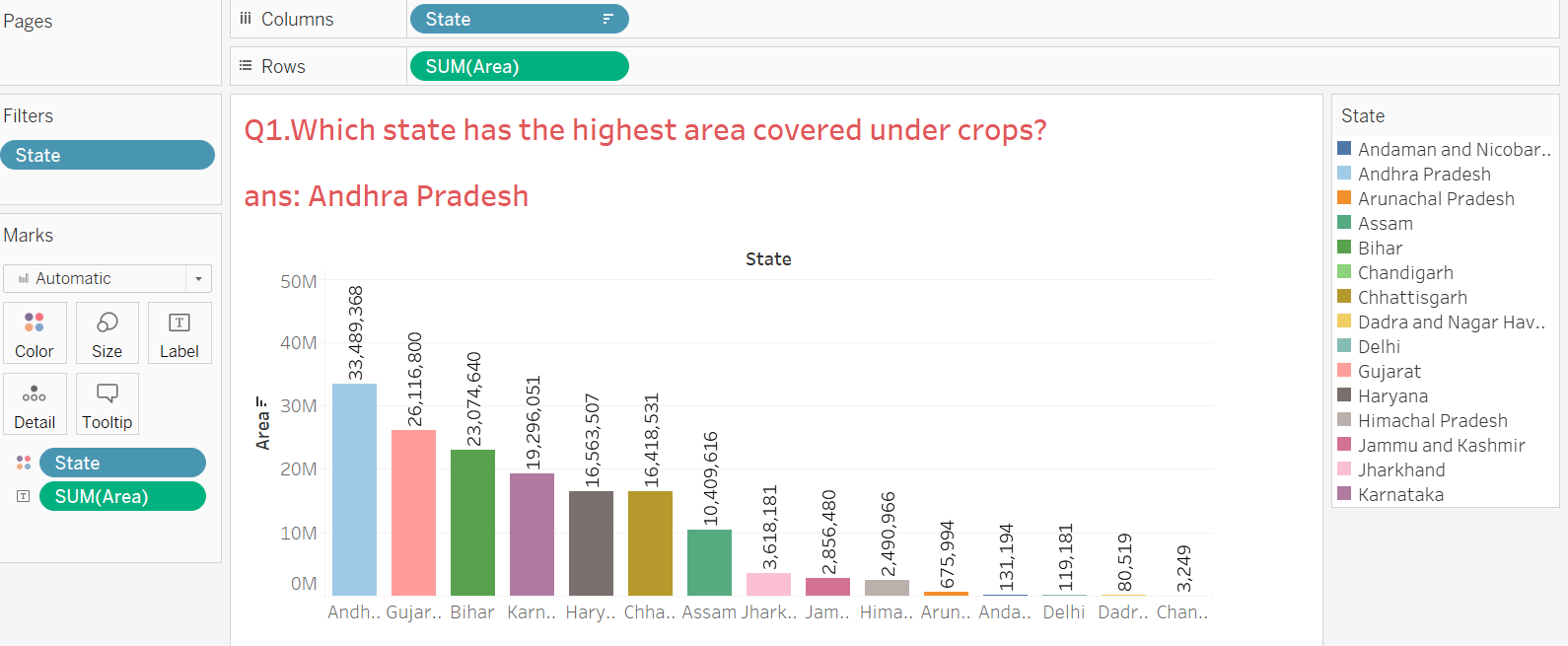
## **STUDENT ID** : 3848587,4044916

Q1.Which state has the highest area covered under crops?

Ans: Andra Pradesh

Steps:

* To create a bar chart that displays the area covered under crops, follow these steps:
  + Connect to the **Indian Agricultural Crop Production** data source.
  + 
  + Note: In Tableau 2020.2 and later, the Data pane no longer shows Dimensions and Measures as labels. Fields are listed by table or folder.
  + Drag the **State** dimension to **Columns** and drag the **Area** measure to **Rows**.
  + The **Area** measure is aggregated as a sum and an axis is created, while the column headers move to the bottom of the view. On the show me card, select **Bar chart** from the list.
  + The view changes to a bar chart.
  + The marks (which are bars in this case) are vertical because the axis is vertical. The length of each mark represents the sum of the area for that crop. The actual numbers you see here might not match the numbers you see—the sample data changes from time to time.
  + Drag the **State** dimension to **Color** on the Marks card.

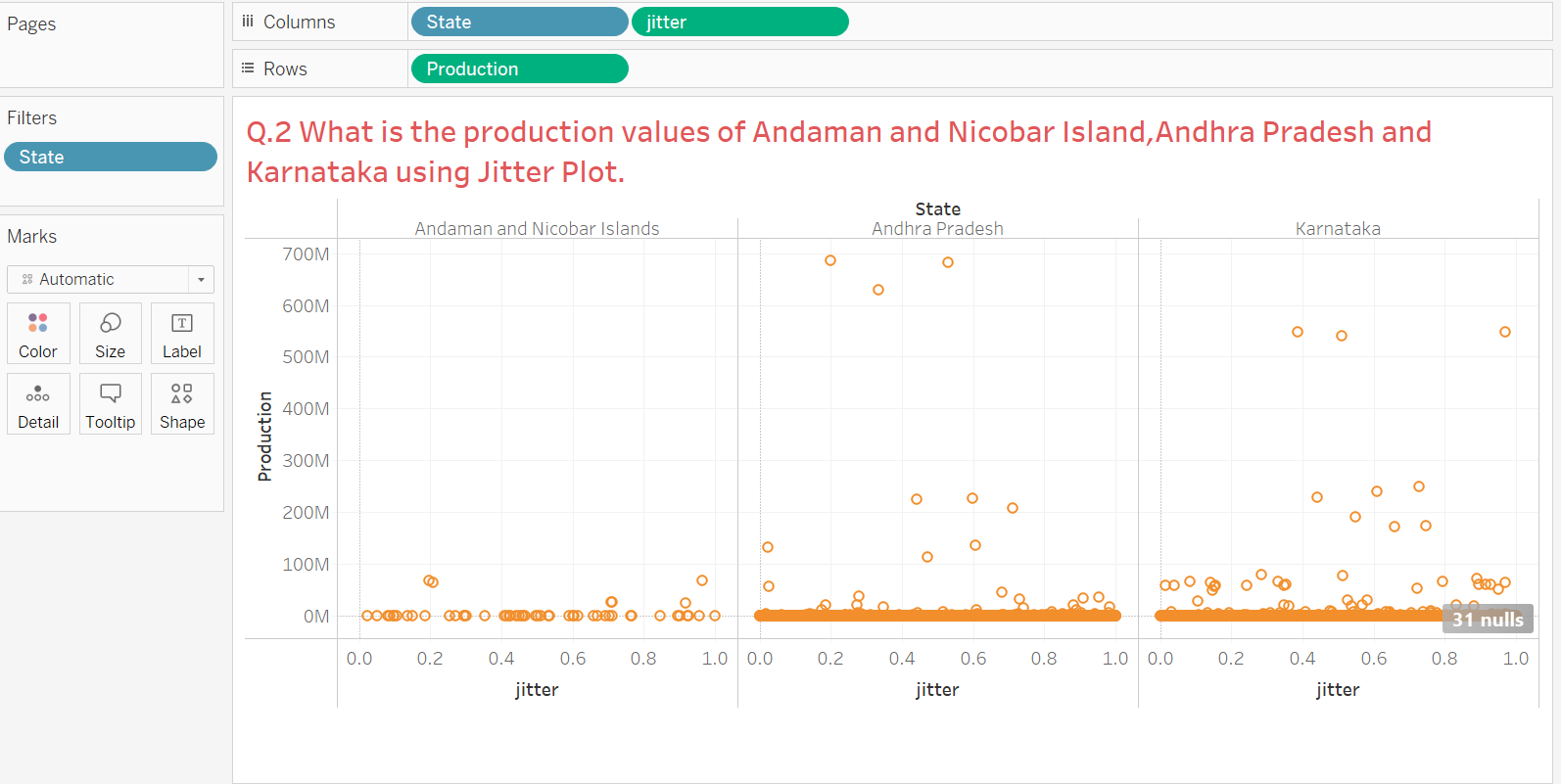


Q.2 What is the production values of Andaman and Nicobar Island,Andhra Pradesh and Karnataka using Jitter Plot.

Ans:

Steps:

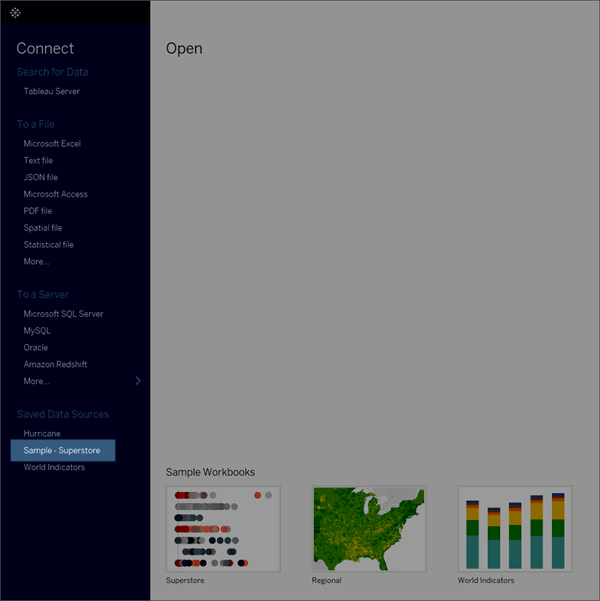
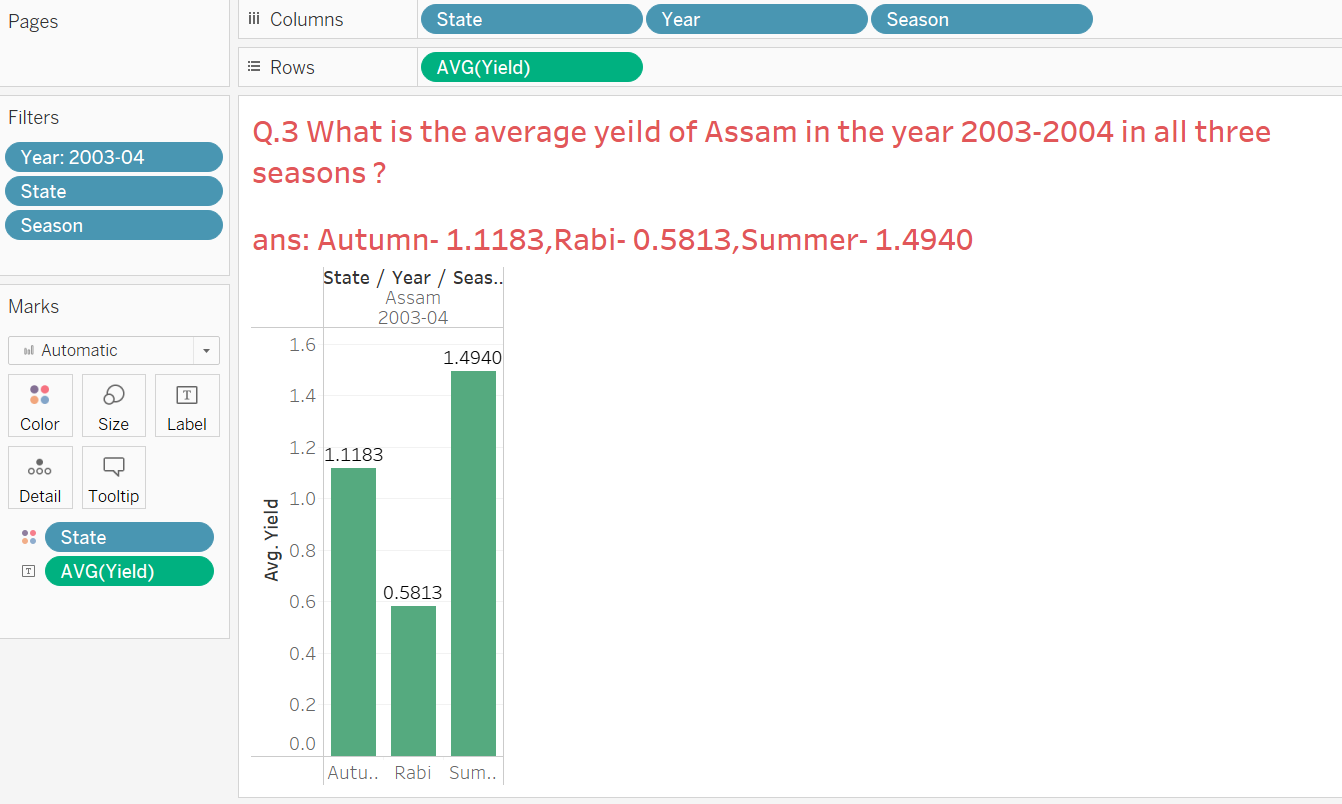
* Drag State to the column and Production to row.
* Drag State to filter and select only the states Anadaman and Nicobar Island, Andra Pradesh and Karnataka.
* Select Analysis and deselect aggregate measures.
* From the mark card select circle chart.
* You can see the difference in the default bar graph and the circle chart.
* To have a jitter plot we need to create a calculated field.
* Let’s name it as Jitter and have random() defined.
* A jitter field will be created. Drag it to the column.
* A jitter plot showing different points can be seen.



Q.3 What is the average yield of Assam in the year 2003-2004 in all three seasons?

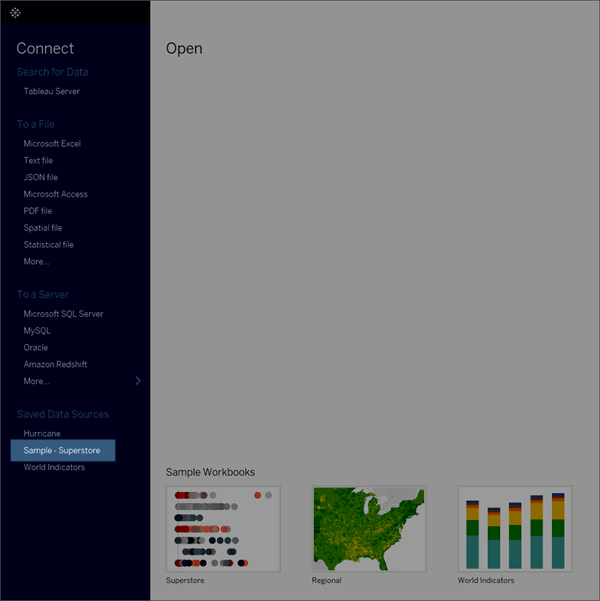
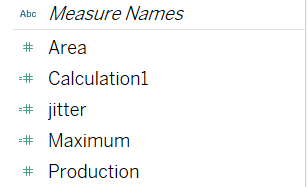
Ans: Autumn- 1.1183,Rabi- 0.5813,Summer- 1.4940

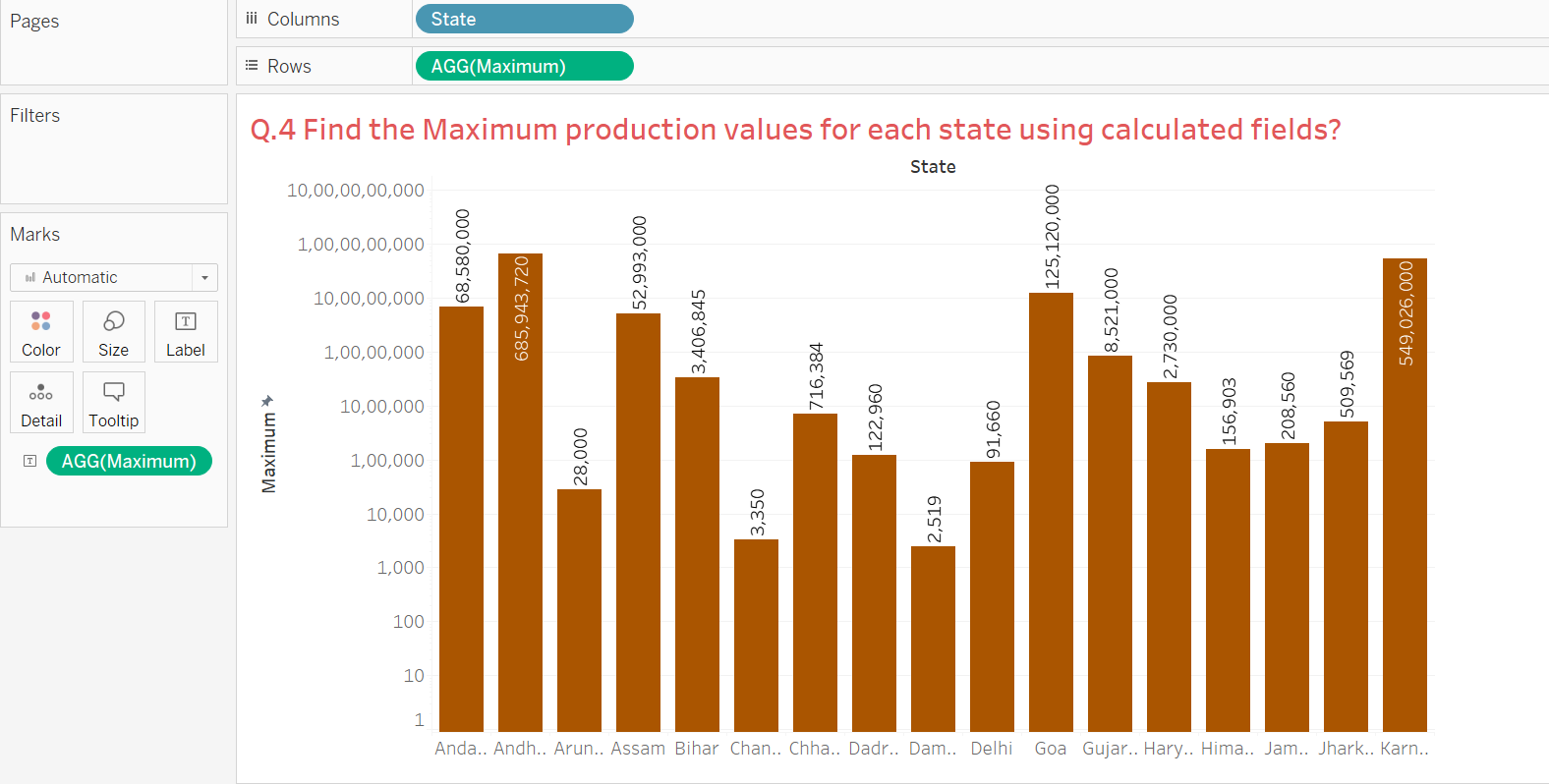
Steps:

* To create a bar chart that displays the area covered under crops, follow these steps:
  + Connect to the **Indian Agricultural Crop Production** data source.
  + 
  + Note: In Tableau 2020.2 and later, the Data pane no longer shows Dimensions and Measures as labels. Fields are listed by table or folder.
  + Drag the **Season** dimension to **Columns** and drag the **Yield** measure to **Rows**.
  + The **Yield** measure is aggregated as a sum and an axis is created, while the column headers move to the bottom of the view. Select measure to average.
  + Drag Year to filter and select only 2003-04. Again drag State to filter and using wildcard option of filter,type Assam in match value and select Exactly matches option. Further drag Season into filter and select Rabi, Summer and Autumn from the list.
  + On the show me card, select **Bar chart** from the list.
  + The view changes to a bar chart.
  + The marks (which are bars in this case) are vertical because the axis is vertical. The length of each mark represents the avg of the yield for Assam. The actual numbers you see here might not match the numbers you see—the sample data changes from time to time.
  + Drag the **State** dimension to **Color and Yield to the label** on the Marks card.
* 

Q.4 Find the Maximum production values for each state using calculated fields?

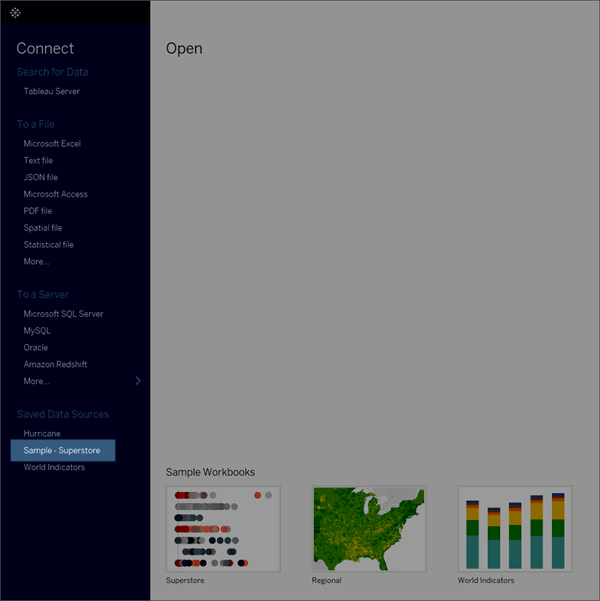
Ans:

* Steps: To create a bar chart that displays the maximum production values for each state, follow these steps:
  + Connect to the **Indian Agricultural Crop Production** data source.
  + 
  + Note: In Tableau 2020.2 and later, the Data pane no longer shows Dimensions and Measures as labels. Fields are listed by table or folder.
  + Drag the **State** dimension to **Columns**.
  + In Tableau, select Analysis > Create Calculated Field.
  + In the Calculation Editor that opens, do the following:
  + Enter a name for the calculated field. In this example, the field is called, Maximum.
  + Enter a formula.
  + To see a list of available functions, click the triangle icon on the right-side of the Calculation Editor.
  + Double-click a maximum function in the list to add it to the formula.
  + When finished, click **OK**.
  + The new calculated field is added to Measures in the Data pane because it returns a number. An equal sign (=) appears next to the data type icon. All calculated fields have equal signs (=) next to them in the Data pane.
  + 
  + Drag **Maximum** to the **Rows** shelf.   
    The view changes to a bar chart.
  + The marks (which are bars in this case) are vertical because the axis is vertical. The length of each mark represents the maximum production values for each state. The actual numbers you see here might not match the numbers you see—the sample data changes from time to time.
  + Drag the **Maximum** dimension to **Color and Label** on the Marks card.



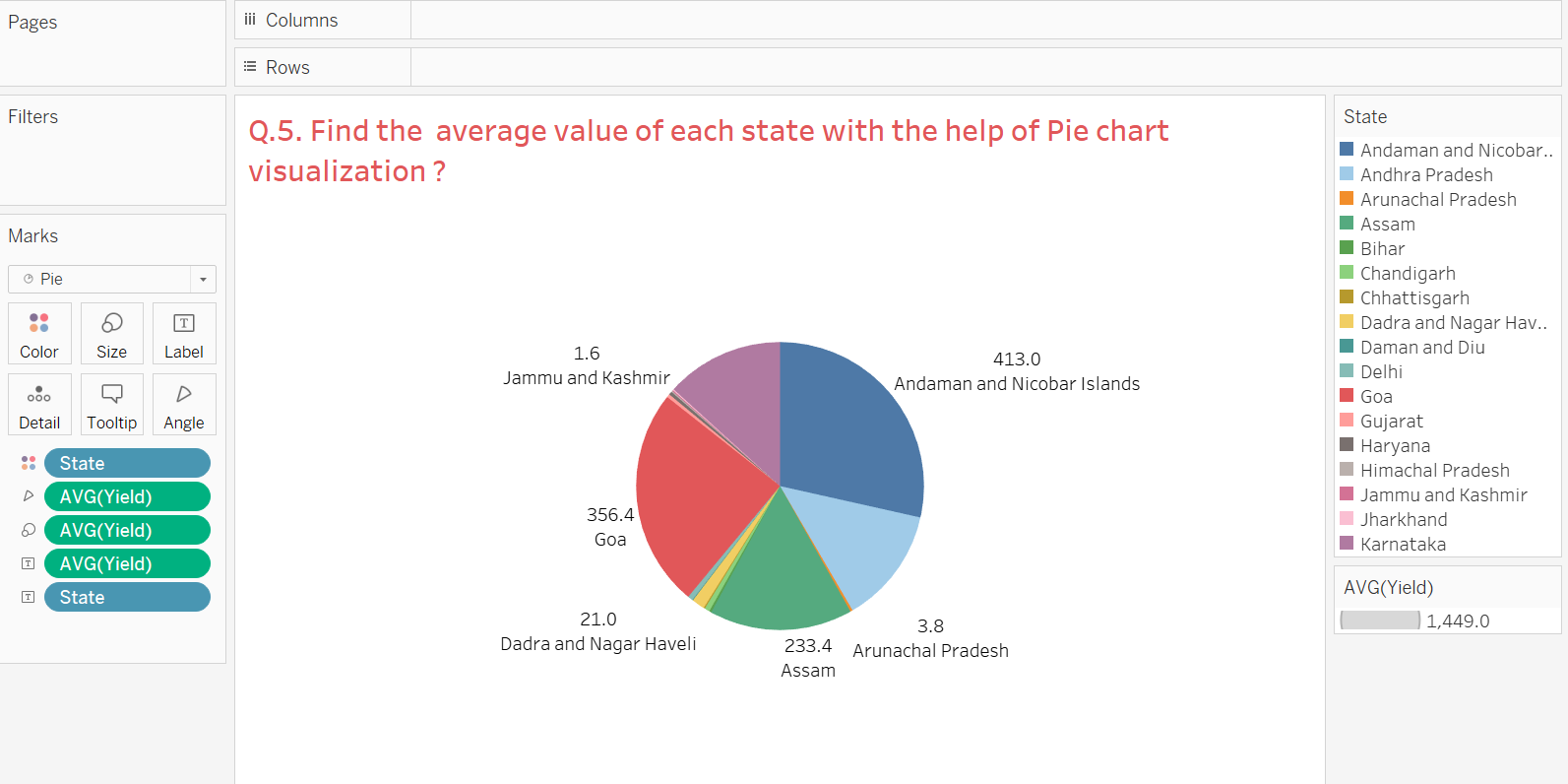
Q.5. Find the average value of each state with the help of Pie chart visualization ?

Ans:

* To create a pie chart view that shows how different product categories contribute to total sales, follow these steps:
* Connect to the **Indian Agricultural Crop Production** data source.
* Drag the **Yield** measure to **Columns** and drag the **State** dimension to **Rows**.
* Tableau aggregates the **Yield** measure as a sum. By default, Tableau displays a bar chart.  
  Click **Show Me** on the toolbar, then select the pie chart type. Pie charts require at least one or more dimensions and one or two measures. 

The result is a rather small pie. To make the chart bigger, hold down Ctrl + Shift nd press B several times.  
Add labels by dragging the **State and Yield** dimension from the **Data** pane to **Label** on the **Marks** card.

* The desired view of pie chart is ready.



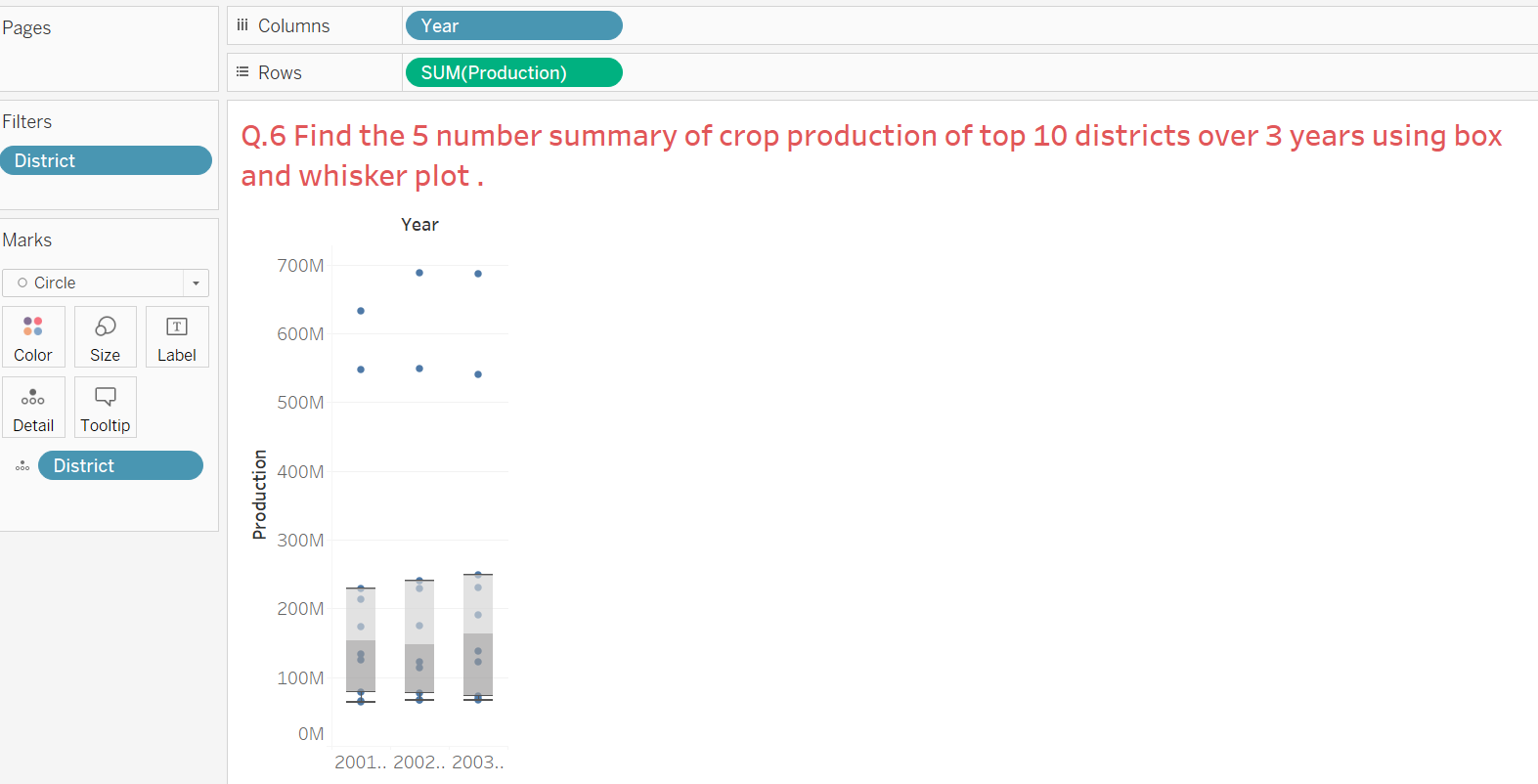
Q.6 Find the 5 number summary of crop production of top 10 districts over 3 years using box and whisker plot.

Ans:

* To create a box plot that shows 5 number summary of crop production of top 10 districts, follow these steps:
  + Connect to the **Indian Agricultural Crop Production** data source.
  + Drag the **Year** dimension to **Columns**.
  + Drag the **Production** measure to **Rows**.  
    Tableau creates a vertical axis and displays a bar chart—the default chart type when there is a dimension on the **Columns** shelf and a measure on the **Rows** shelf.
  + Click **Show Me** in the toolbar, then select the box-and-whisker plot chart type.  
    
  + Drag district to filter, and enter 10 in the by field option of the Top menu.

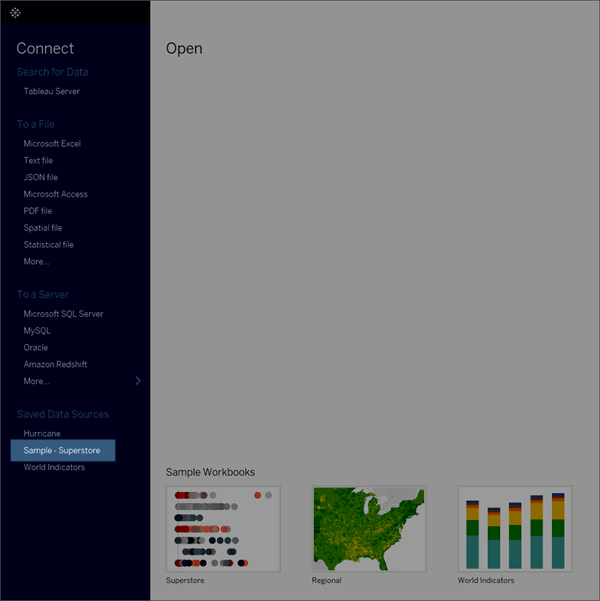
Drag District to detail.

* Tableau displays the a box plot with 5 number summary in detail:



Q.7 Which crop starting with 'G' has the highest yield over the years ?

Ans: Ginger

* Connect to the **Indian Agricultural Crop Production** data source.
* 
* Drag Crop dimension to **Columns**.
* Drag Yield to **Rows**.
* Drag Crop to filter, in filter select wildcard and then select starts with option from the list and enter ‘G’ in the match value box. The crops starting with ‘G’ will be filtered.
* Drag **Year** to **Color** on the Marks card.
* Drag **Crop** to **Label** on the Marks card.
* Click **Show Me** in the toolbar, then select the stacked bar type.

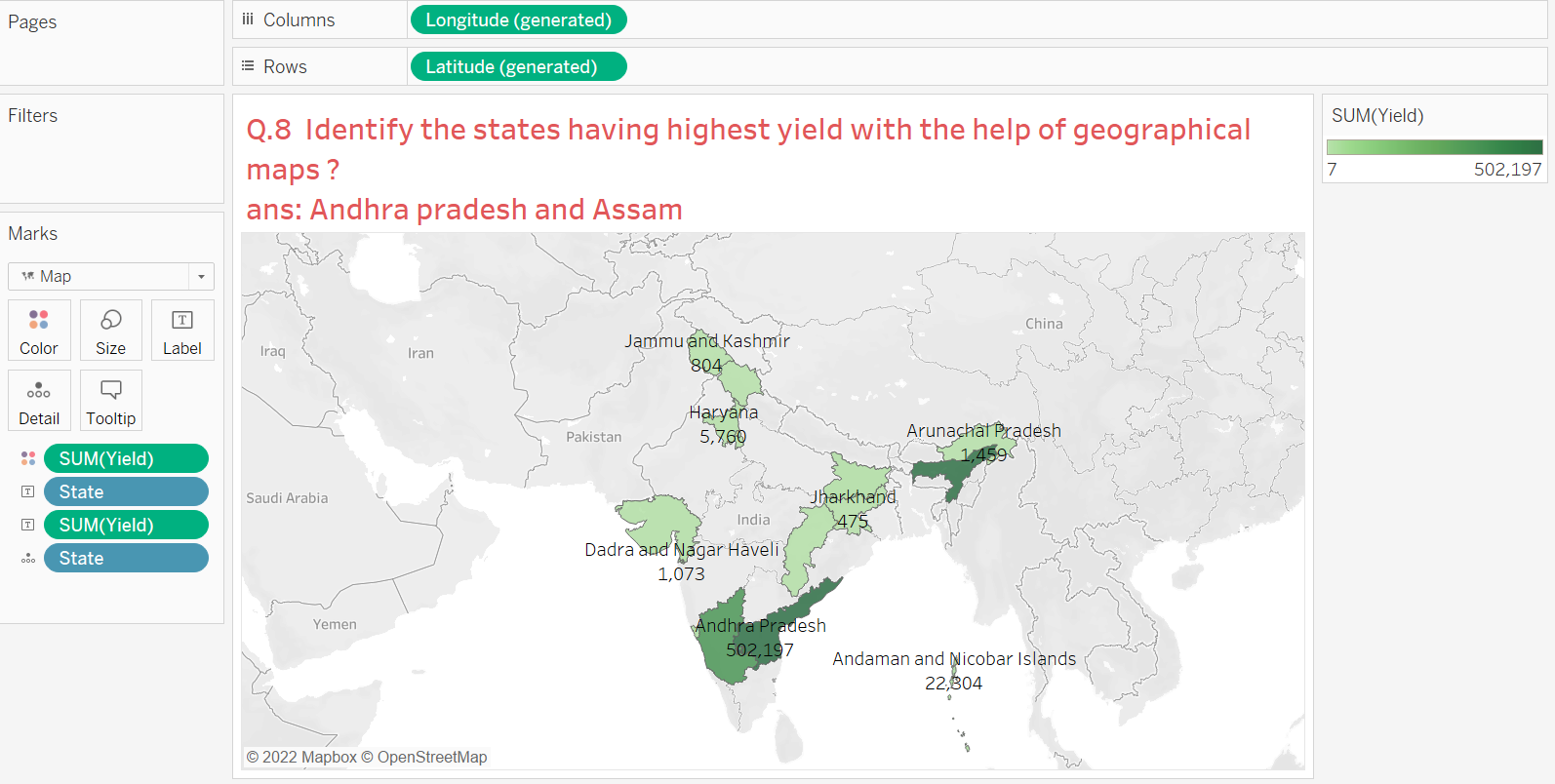


Q.8 Identify the states having the highest yield with the help of geographical maps ?

Ans: Andhra pradesh and Assam

Steps:

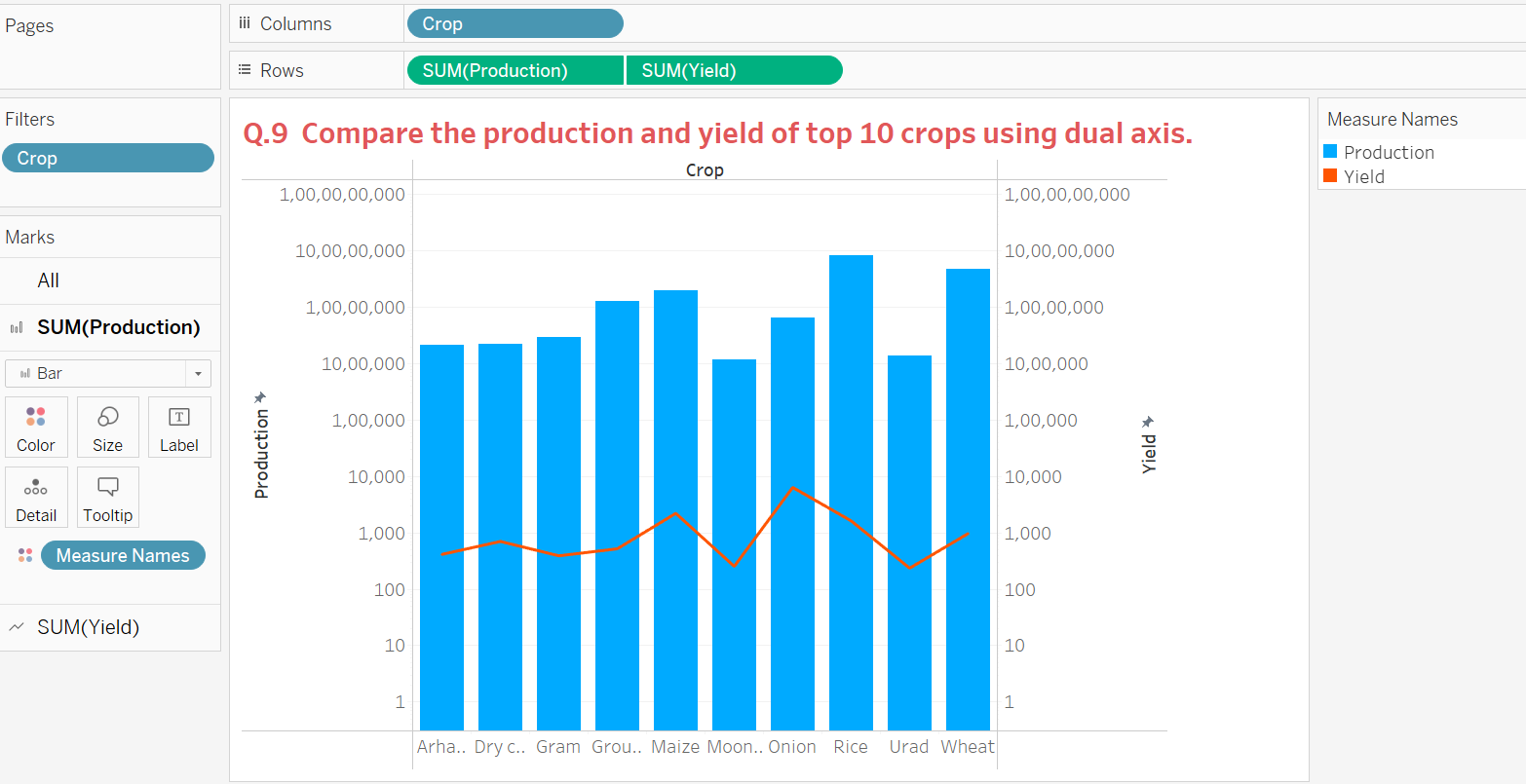
* **Build a simple filled (polygon) map:**
* Navigate to a worksheet.
* In the **Data** pane, under Tables, double-click **State** and **Country**
* On the Marks card, click the Mark Type drop-down and select **Map**.
* From Dimensions, drag **State** to **Color** on the Marks card.
* Drag State and Yield to label on the marks card.
* The map updates to show the sum of sales by state. The states with higher sales are darker blue, and the states with fewer sales are lighter blue.
* You can try with regions field dragged in color pane.



Q.9 Compare the production and yield of top 10 crops using dual axis.

Ans:

Steps:

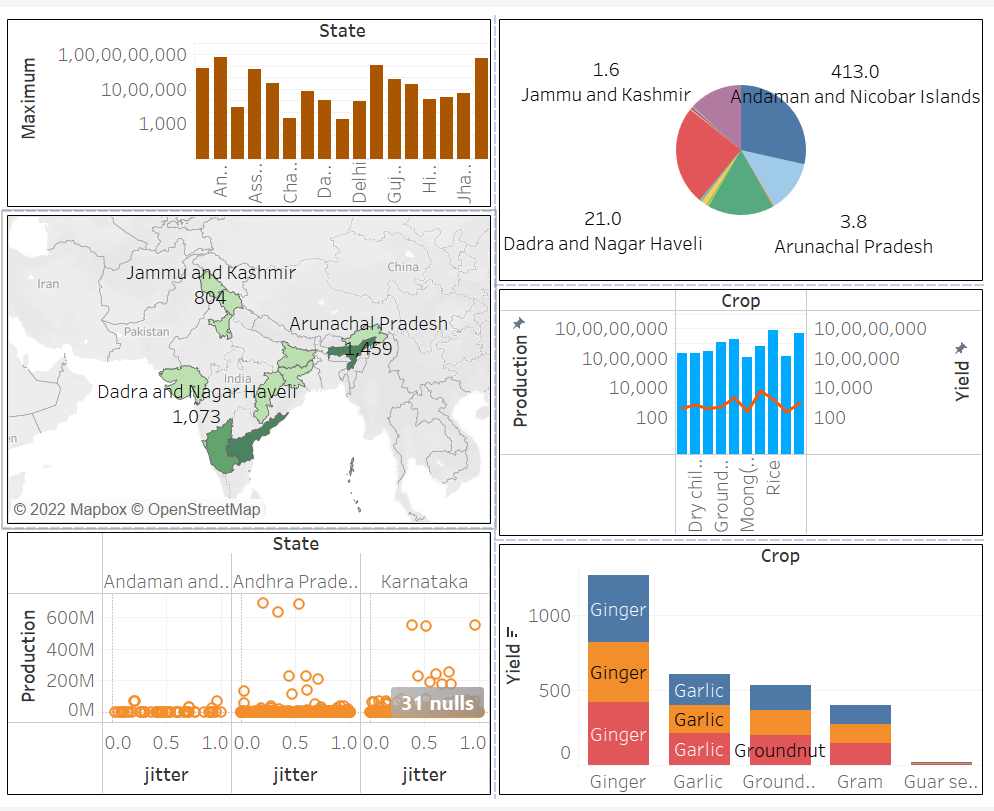
* Compare two measures using dual axes:
* You can compare multiple measures using dual axes, which are two independent axes that are layered on top of each other. Dual axes are useful for analyzing two measures with different scales.
* To add a measure as a dual axis, drag the field to the right side of the view and drop it when you see a black dashed line appear.
* You can also right-click (control-click on Mac) the measure on the Columns or Rows shelf and select **Dual Axis**.
* Change the mark type of Production axis to Bar and Yield axis to Line.
* Drag measure names to color on the marks card.
* The result is a dual axis view where the Production axis corresponds to the blue bar chart and the Yield axis corresponds to the red line.
* To align the two axes in a dual axes chart to use the same scale, right-click (control-click on Mac) the secondary axis, and select **Synchronize Axis**. This aligns the scale of the secondary axis to the scale of the primary axis.
* 

Q.10 Create a Dashboard of Indian Agricultural Crop Production ?

Ans:

Steps:

* Create a dashboard, and add or replace sheets
* You create a dashboard in much the same way you create a new worksheet.
* At the bottom of the workbook, click the **New Dashboard** icon:  
  
* From the **Sheets** list at left, drag views to your dashboard at right.
* To replace a sheet, select it in the dashboard at right. In the Sheets list at left, hover over the replacement sheet, and click the **Swap Sheets** button.
* Add interactivity
* You can add interactivity to dashboards to enhance users' data insights. Try these techniques:
* In the upper corner of sheet, enable the **Use as Filter** option to use selected marks in the sheet as filters for other sheets in the dashboard.
* You can also change the layouts like border, background for each container.
* Your dashboard is ready for visualization.



DATA SET :

<https://d.docs.live.net/08142f24a8cd5771/Desktop/indian%20agricultural%20production.csv>