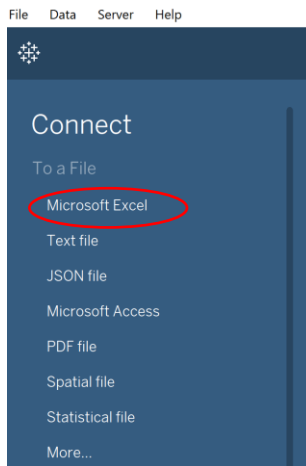# Notes for Tableau Data Visualization
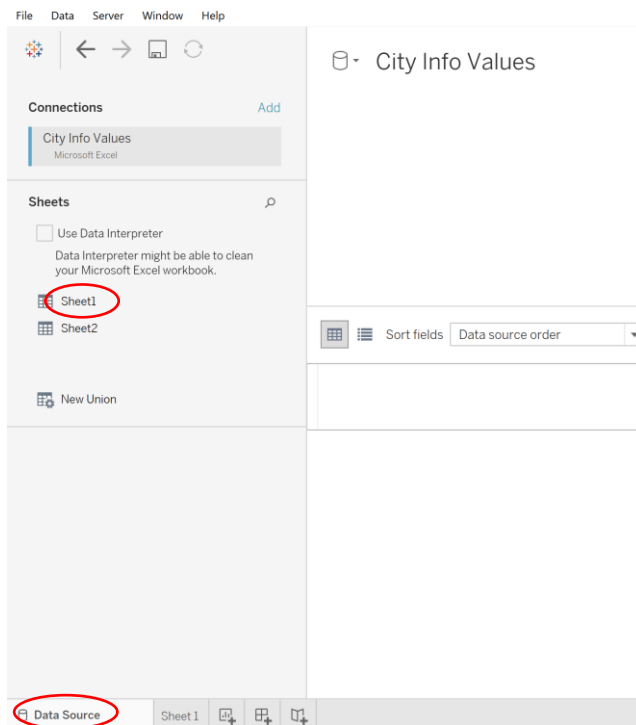
By Prof. Sujin Kim

In this class, we will practice Tableau using the city information datasheet <u>city info values (Excel)</u> to visualize trends about cities, income, and population.

0. Load data

- Load Tableau in your PC. Click Microsoft Excel and load the city information data into Tableau.

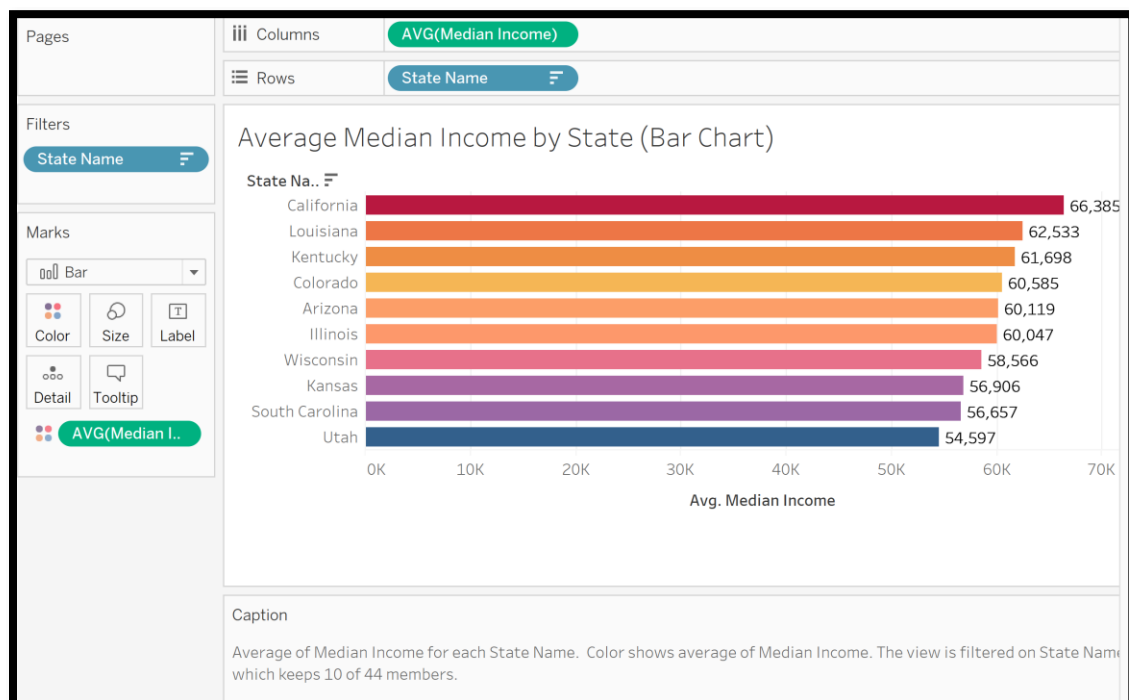- Data source tab: Drag 'Sheet1' to the right or just click it.

Now we are ready to create plots! Go to 'Sheet 1'.

1. Plots for summary statistics (You may add more variables to columns)

We will create plots for 'Median Income' by state.

(i) Bar chart

Figure 1-(i): Create a bar chart of **average** median income by state (top 10 states only)

- Move variables to the Columns and The Rows shelves: Median income to column and State Name to row. Then, a bar chart appears in the sheet.
- Open the drop down menu of 'SUM(Median income)' and change the data format
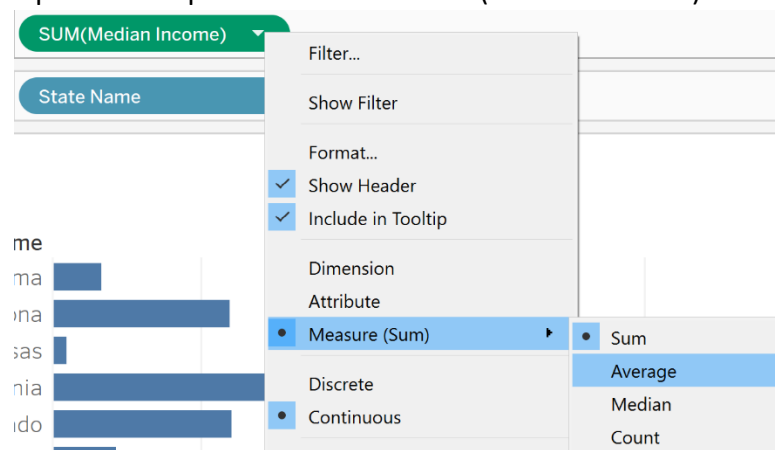
- Add 'State Name' to the Filters shelf. Move to the Top tab and keep the top 10 states by average median income.
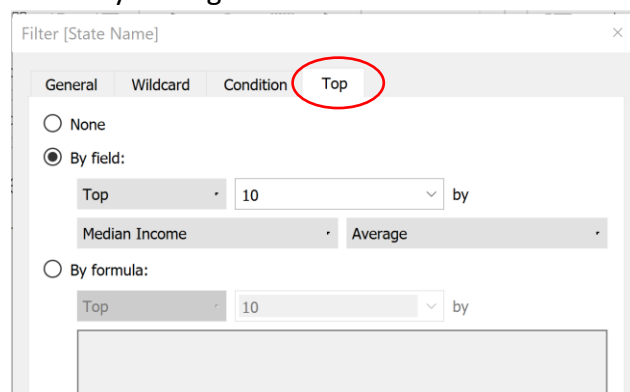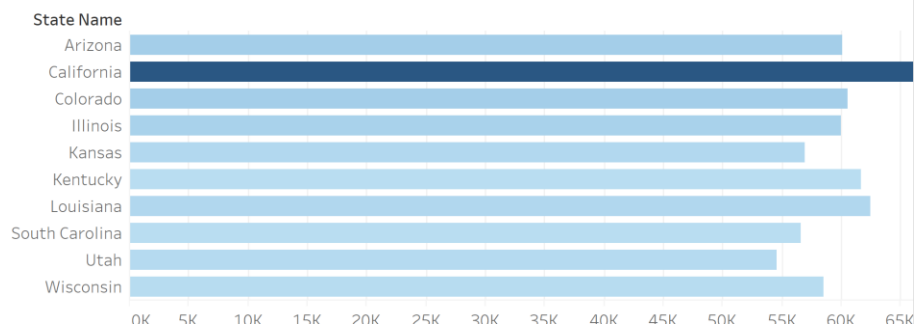
- Let's change the colors: Move 'Median Income' to the Color and change the measure using the drop-down menu.
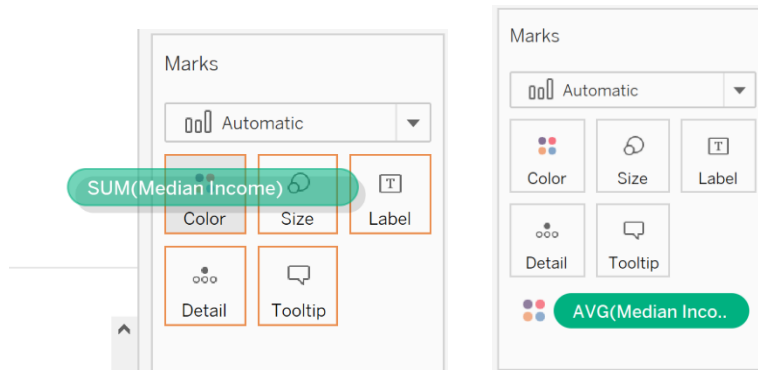
- Edit the color using the editing tool on the right-hand side panel. If you do not see the editing panel, unselect 'Show Me' on the top to deactivate it.
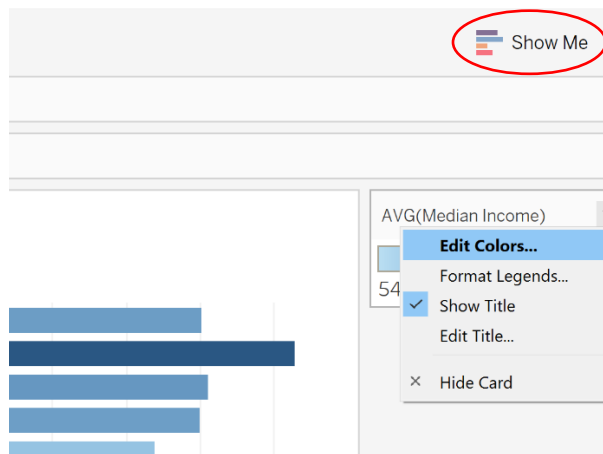
- Sort the chart by average median income:
  Method 1: Go to 'State Name' in the chart and use the drop-down menu
  Method 2: Click 'Nested Sort' on the top ribbon.
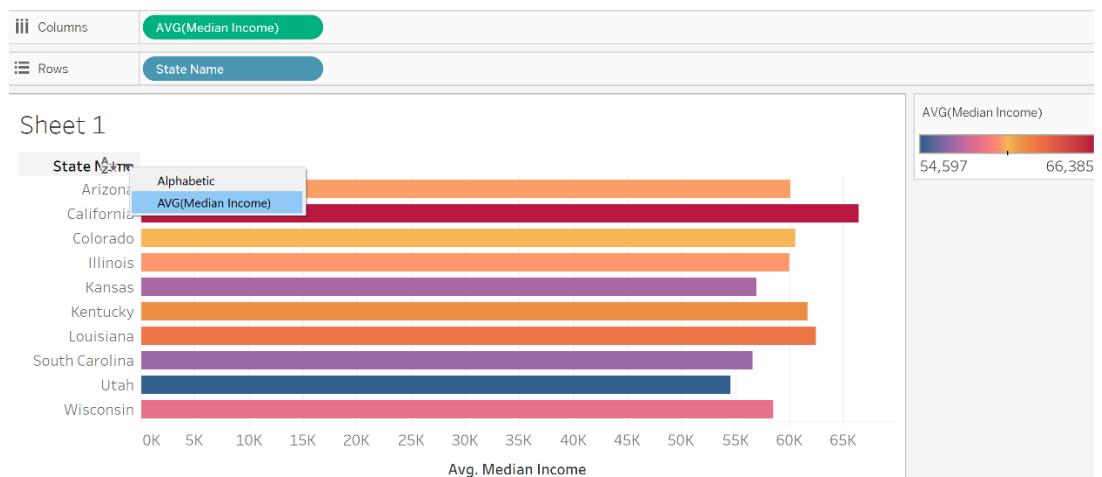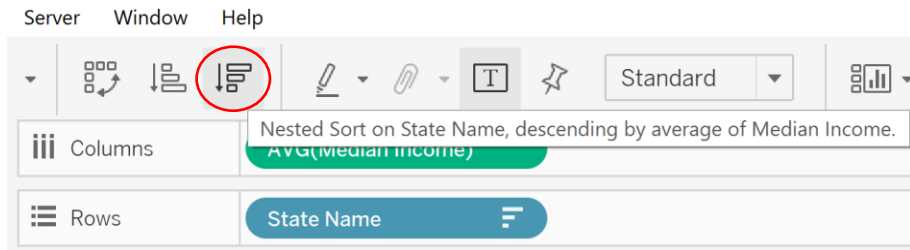
– Label the chart using the Label card or the Analysis tab
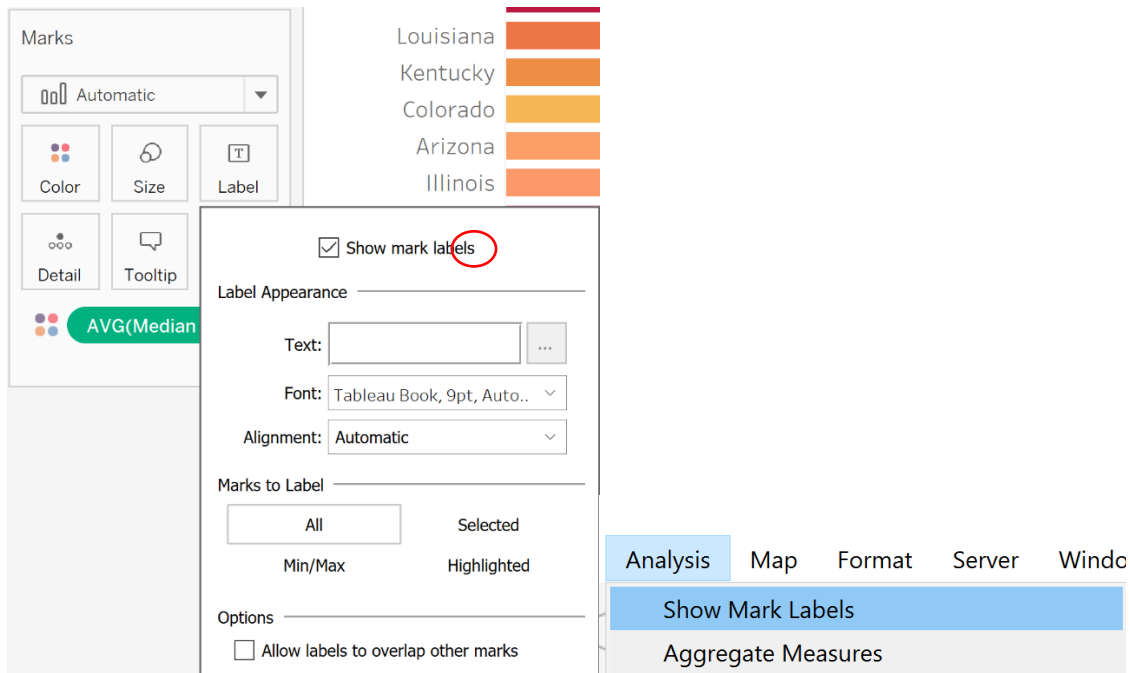
– Edit the title: Double click the title.
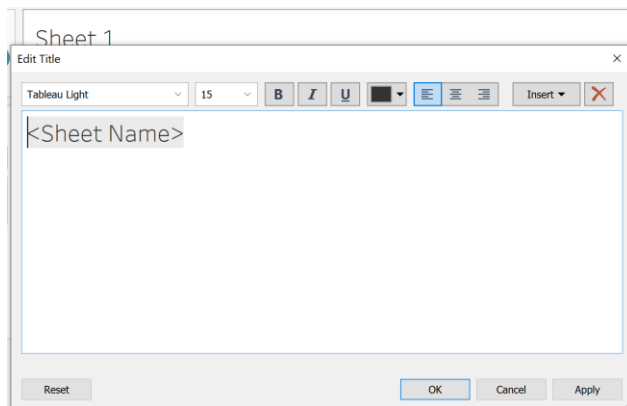
- Turn on caption, title and summary: Right click on the sheet or go to the 'Worksheet' tab
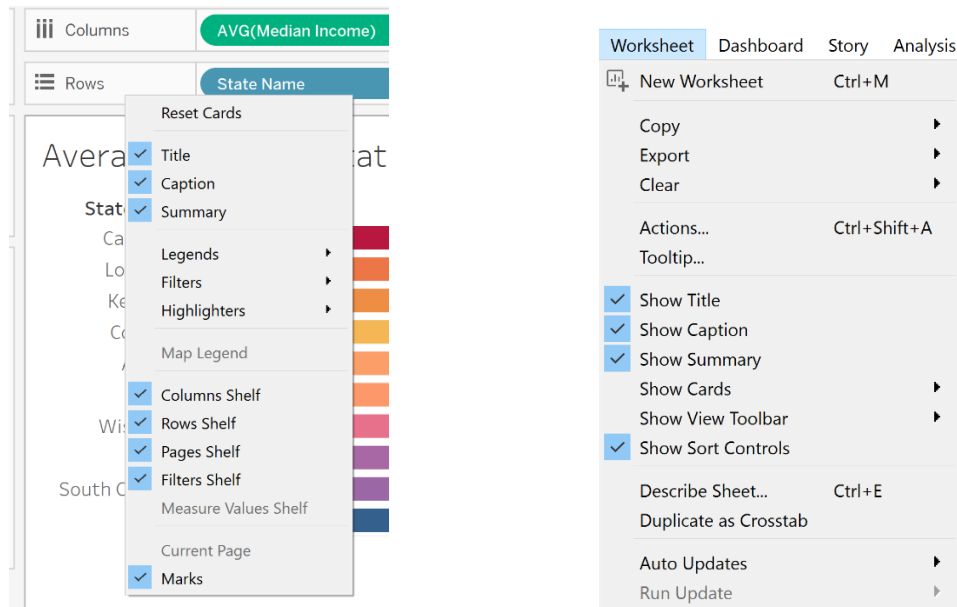
- You can edit the Summary using the drop down menu.
- Also, place cusor on the chart to view the summary statistics.
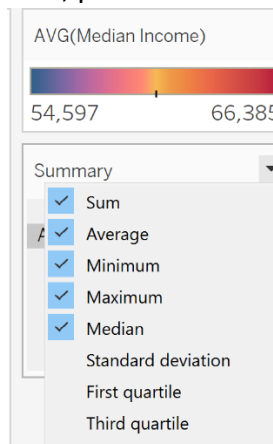
- You can copy or export the worksheet
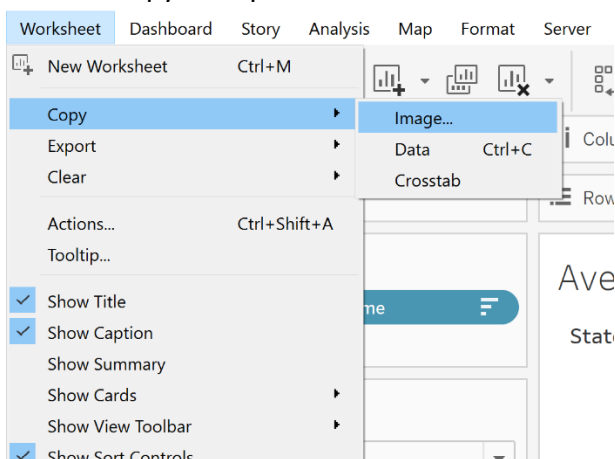
(ii) Box plot

Figure 1-(ii): Create a box plot of Median Income by state (top 10 states only)

Average Median Income by State (Box- Plot)

Median Income for each State Name. Color shows sum of Median Income. The view is filtered on State Name, which keeps 10 of 44 members.
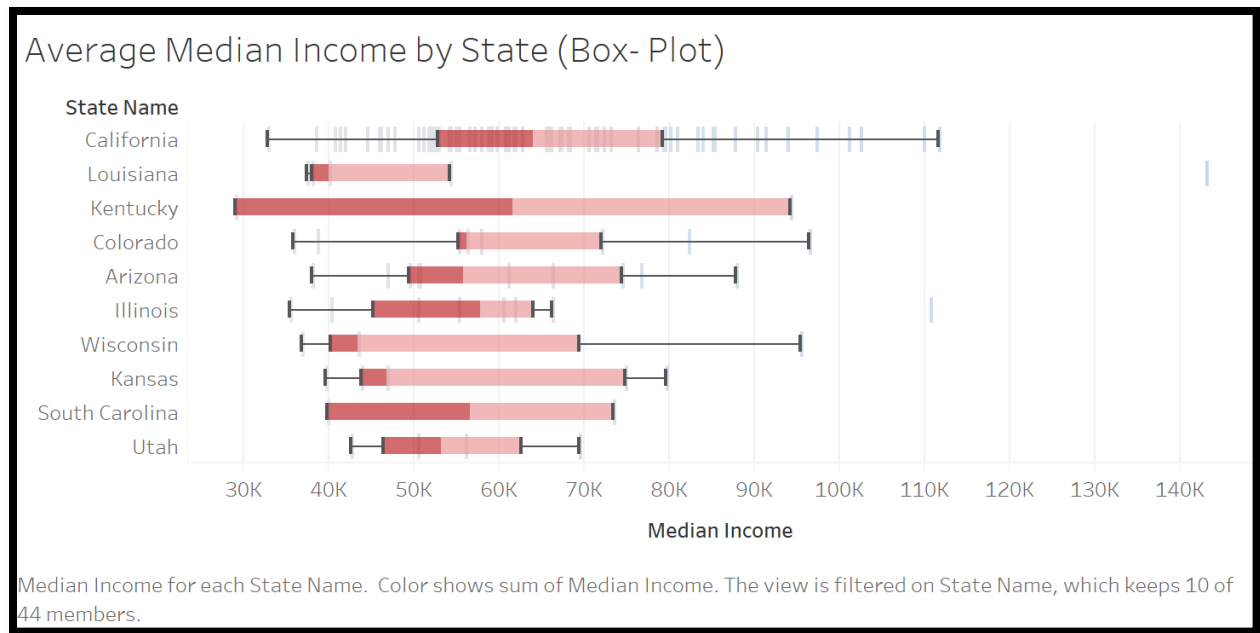
   —   Duplicate the bar chart: Right click 'Sheet 1' tab or click 'Duplicate' on the top ribbon.
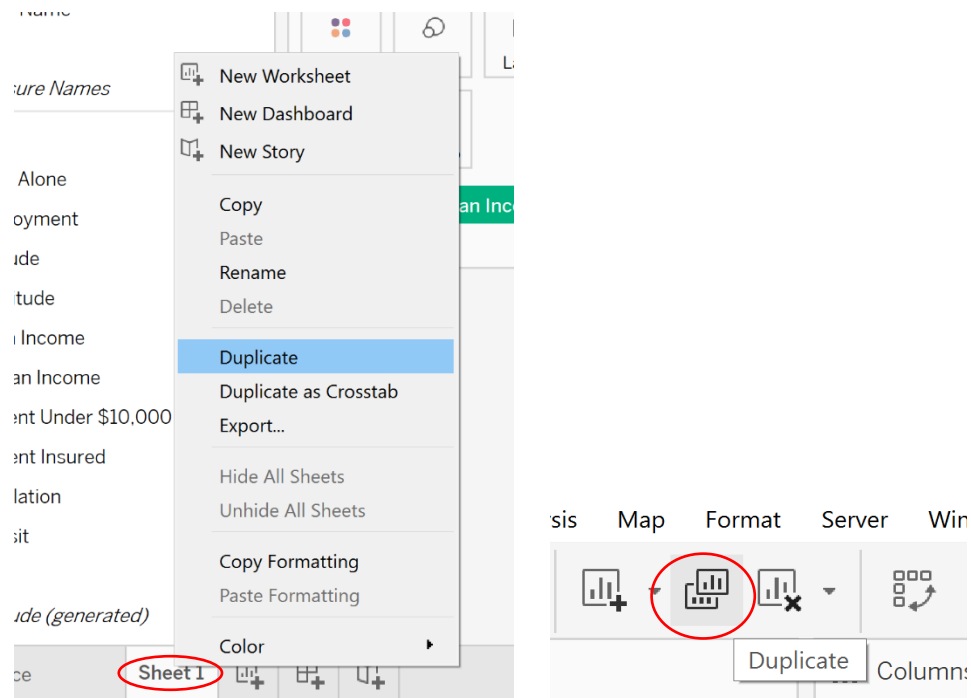
– On the Mark shelf, select 'Gantt Bar'
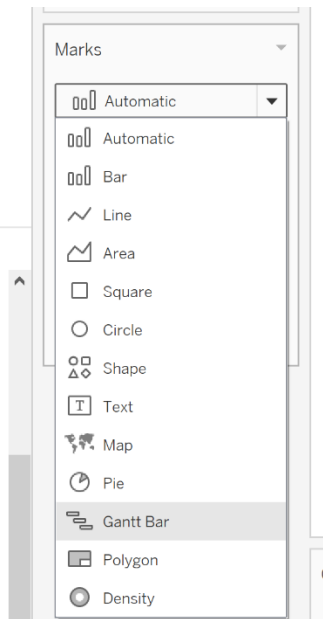
– Change the format of 'Median Income':
Method 1: Go to the Columns shelf and double click 'Dimension'
Method 2: Go to the Analysis tab and unselect 'Aggregate Measures'.
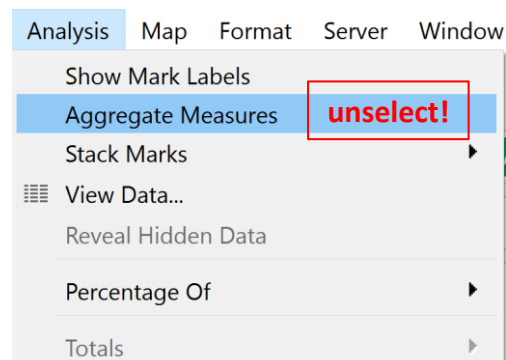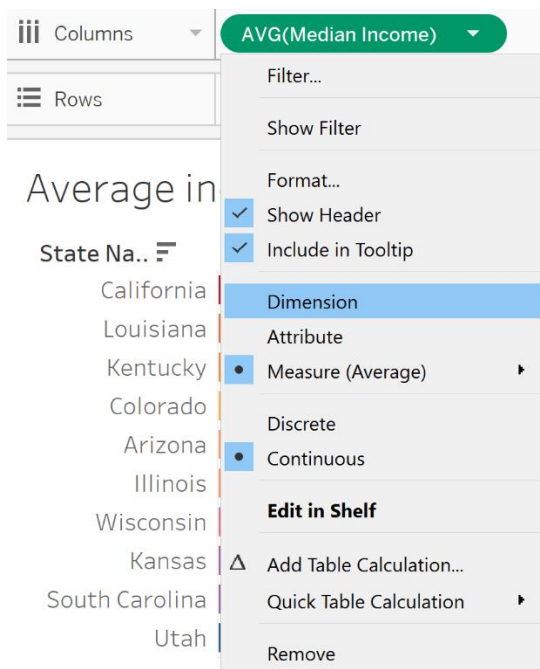
- Create a boxplot:
  Method 1: Go to the Analytics tab on the left and click Box plot
  Method 2: Click the Show Me tab on the top right and select box-and-whisker plot.
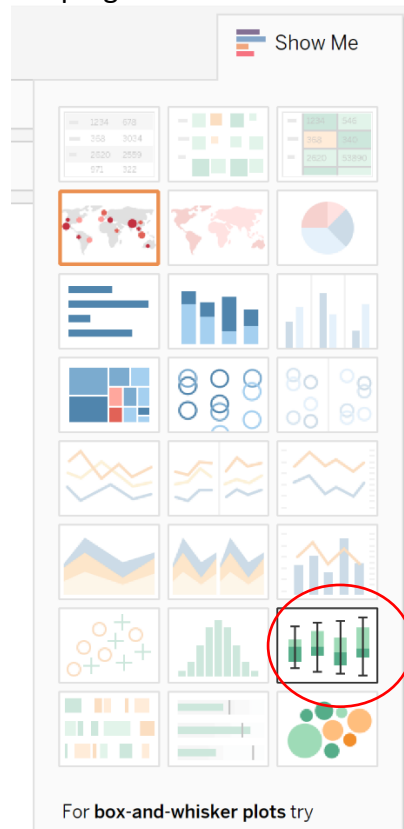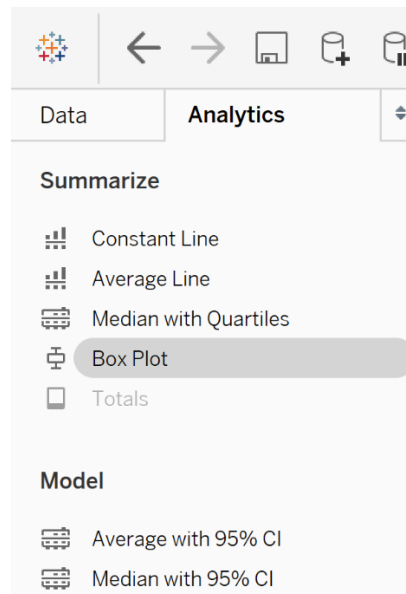
- You can swap the column and the row of the box plot
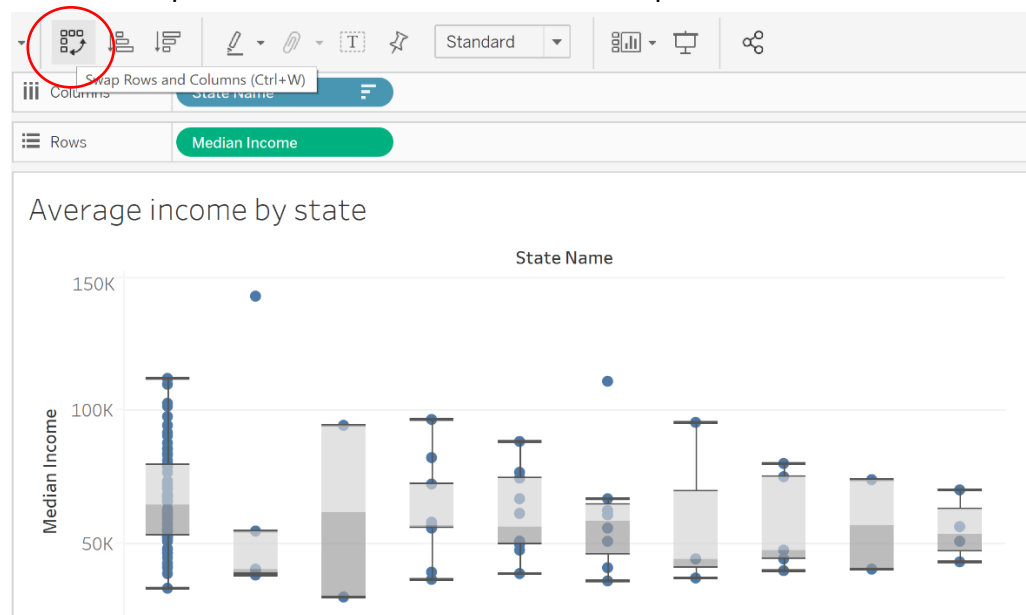
- Right click a bar and change the color using the editing tool.
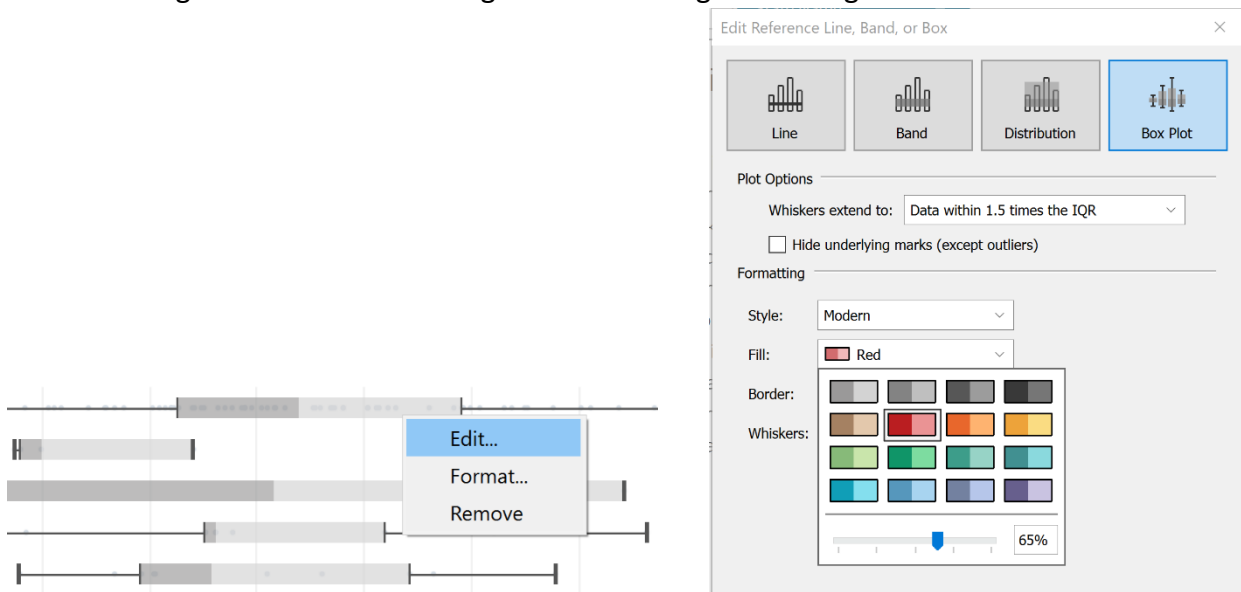
(iii) Histogram

Figure 1-(iii): Create a histogram of Median Income
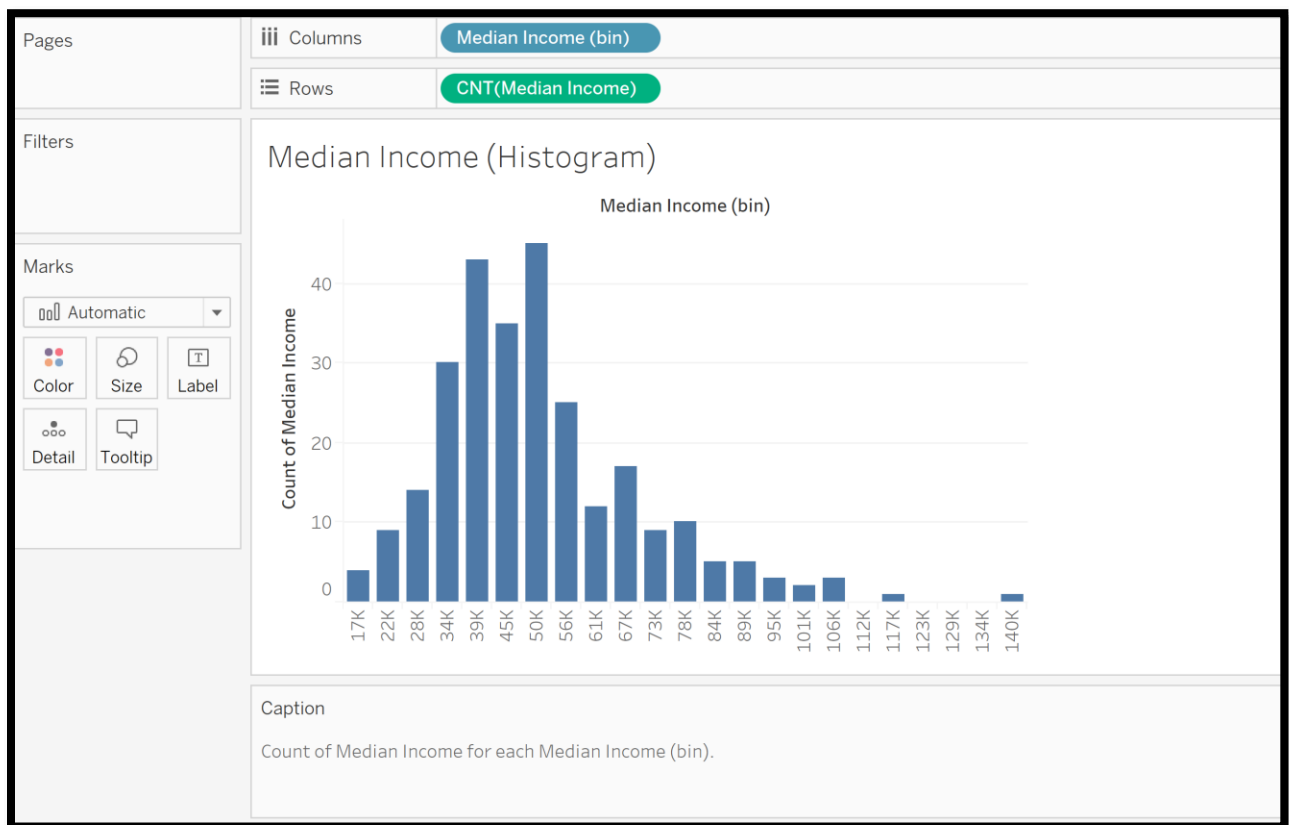
– Add Median Income to the Columns shelf

– Select histogram on the Show Me panel.
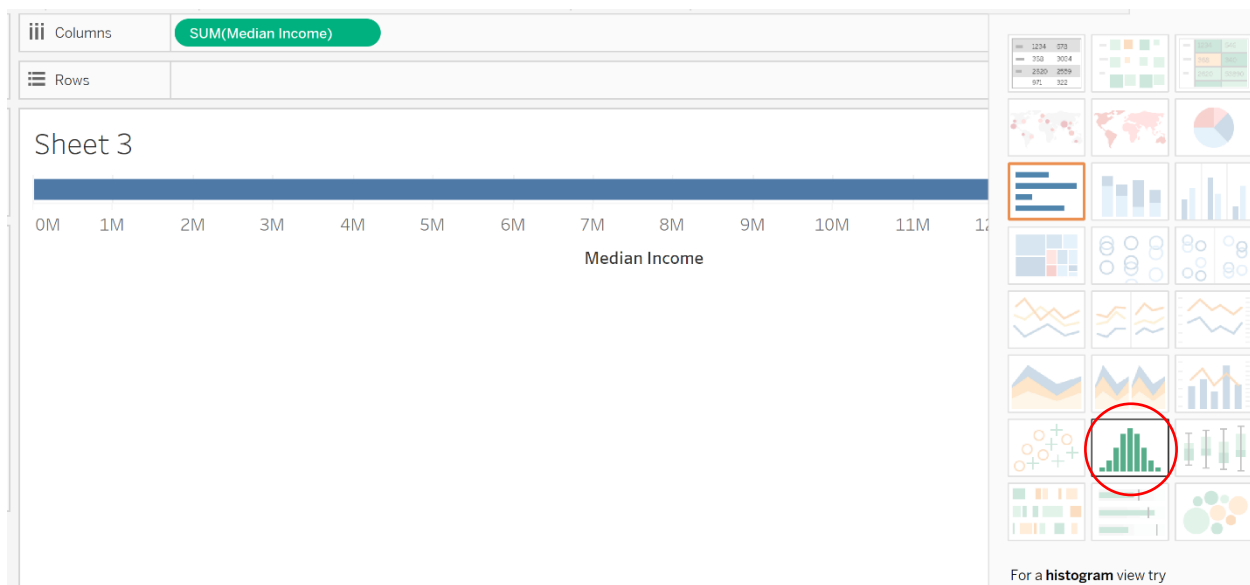
2. Scatter plot

Figure 2: Create a scatter plot of Median Income and LN(population)
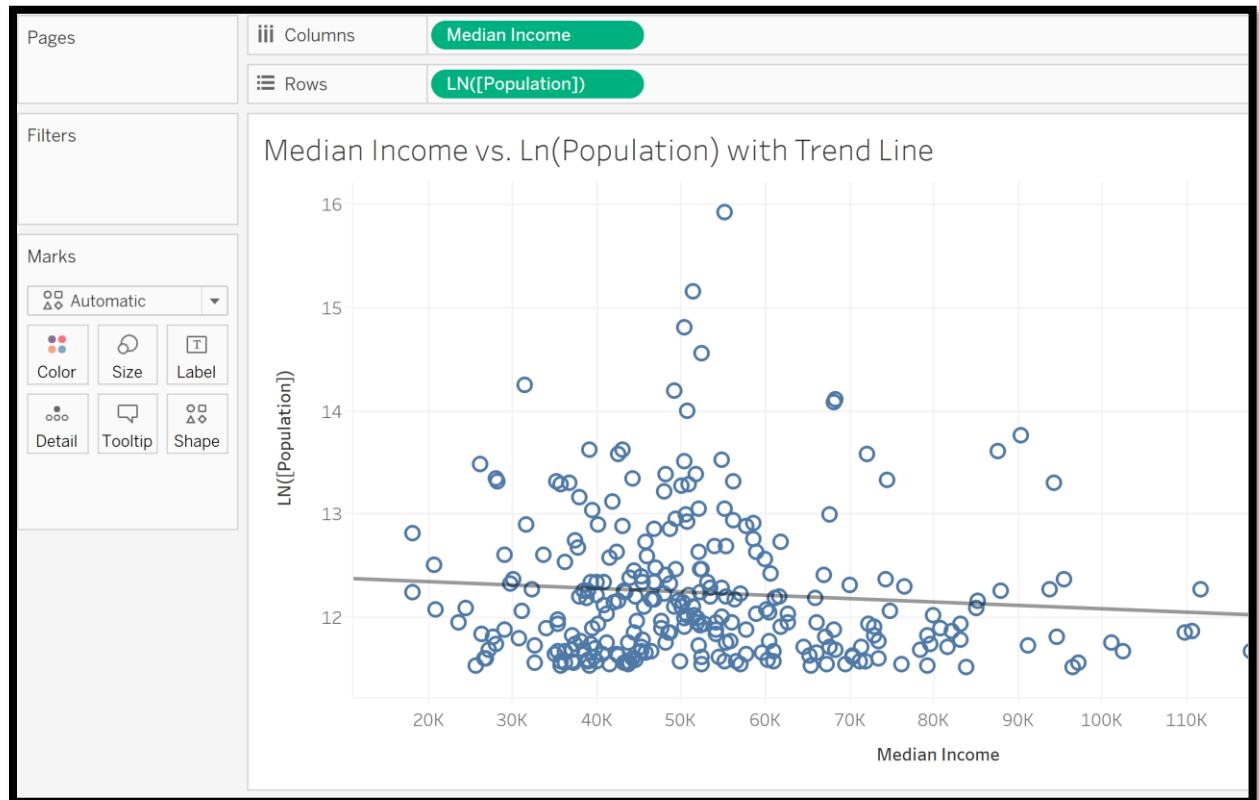
- Add Median Income and Population to the Column and Row shelves
- Change the variable format: Dimension
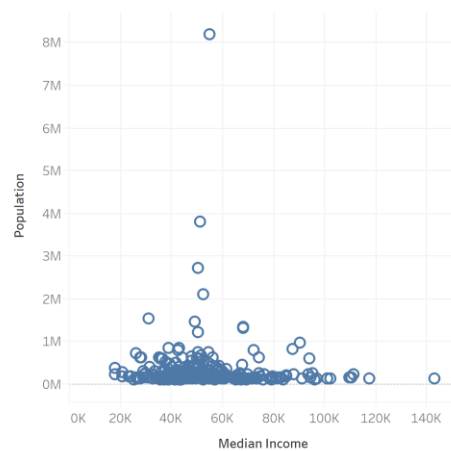
This plot does not present a clear relationship between Median Income and Population. Therefore, we consider a variable transformation.

- Double click Population – Change the formula
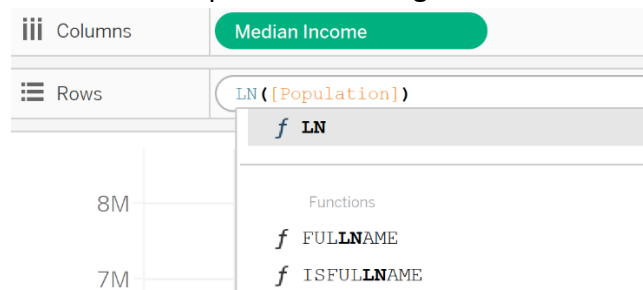
- Double click the y-axis and change the range.

− Right click the plot and add a trend line:
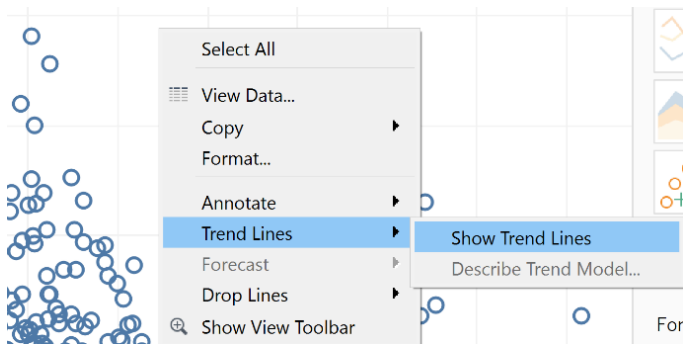
We observe a linear relationship between Median Income and LN(Population).

3. Map

(i) Present Median Income and Population on a map

Figure 3-(i)
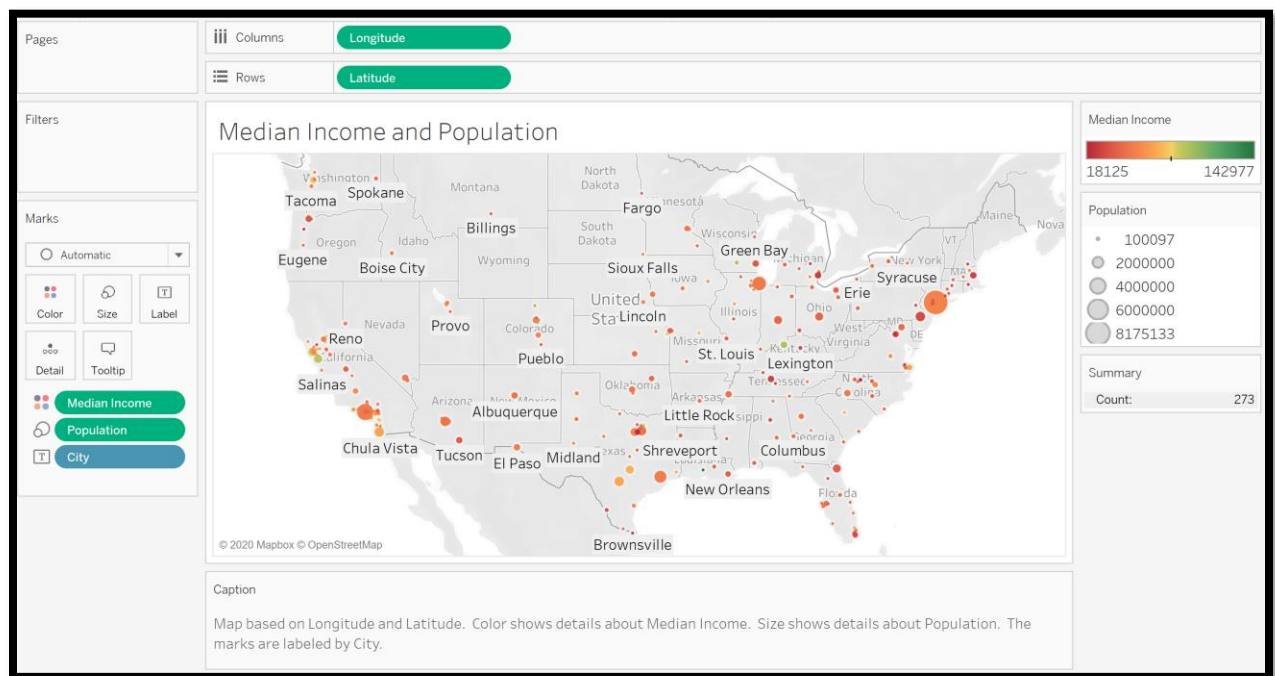
Create a map to present all cities:

Method 1: Drag city onto a sheet. Longitudes and latitudes are automatically generated by Tableau.
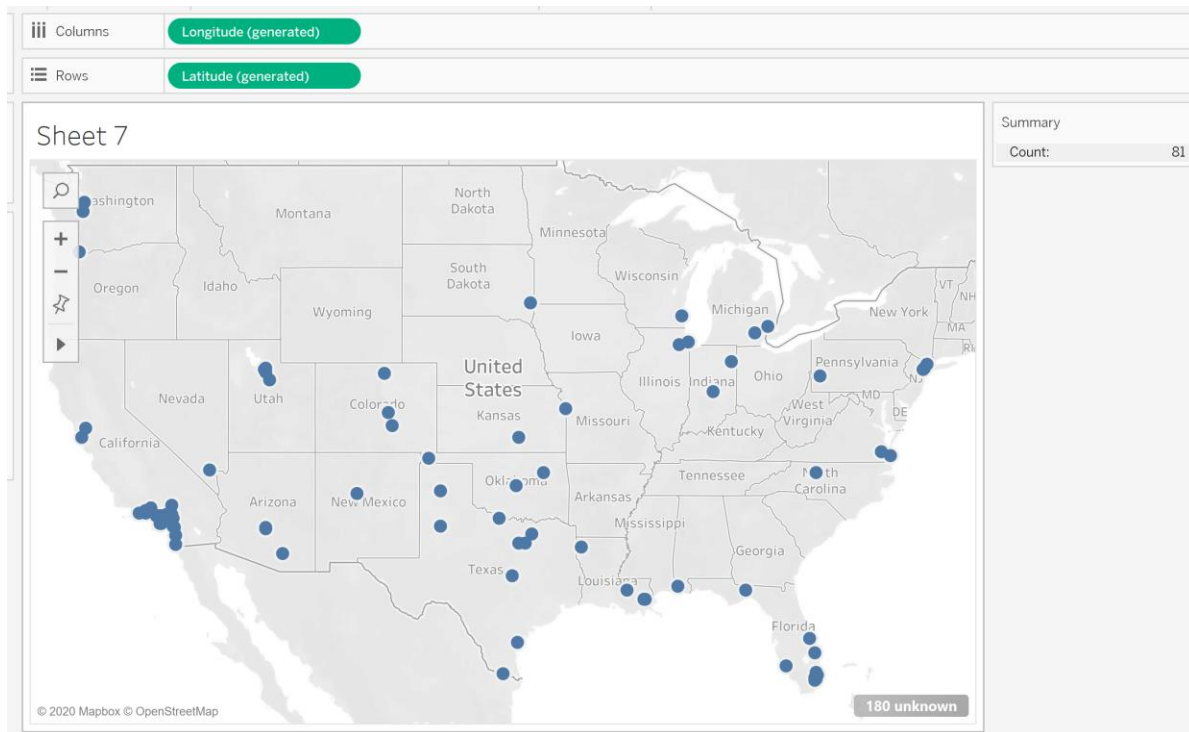
Method 2:

o Move Longitude and Latitude to the Columns and the Rows shelves
o **Variable format: Dimension**
o If a map does not appear, click the map on the right 'Show Me' panel or go to the Marks shelf and click Map.
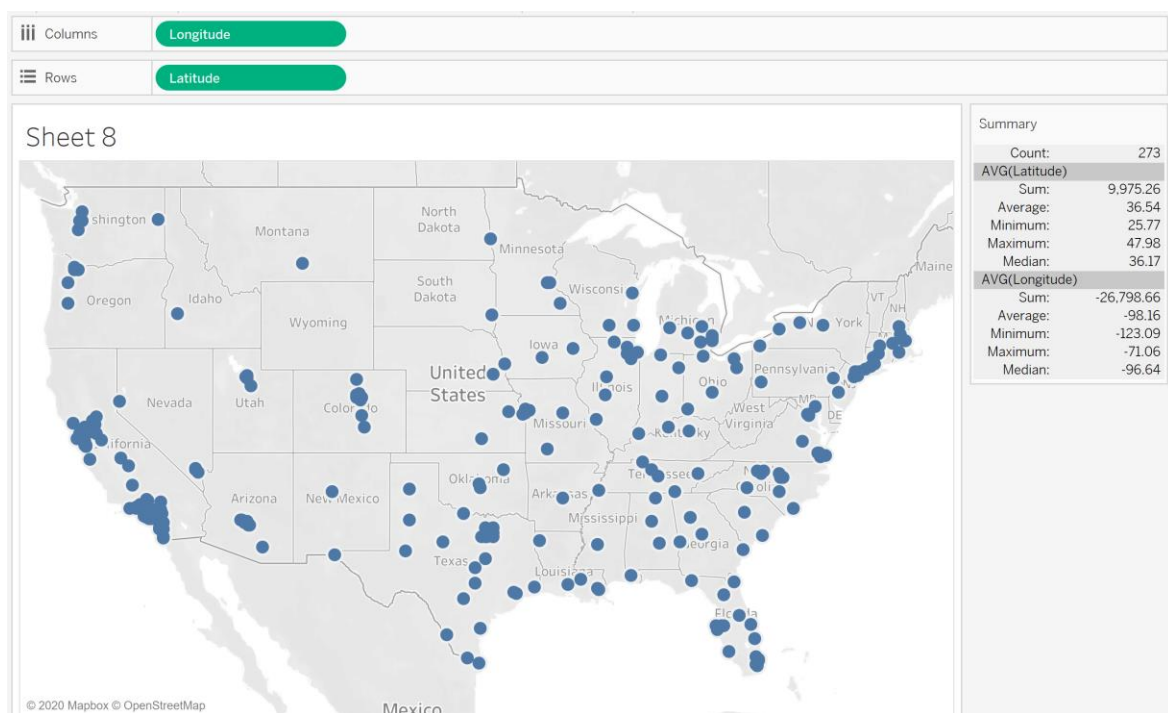
Note that the two maps are different.

o   Method 1: 81 cities
o   Method 2: 273 cities
We will go with the **Method 2.**

Add variables to the Marks shelf:

o   Median income (color)
o   Population (size)
o   Add city (label)

(ii) Sets and Groups for filtering

Figure 3-(ii): Present big cities in California
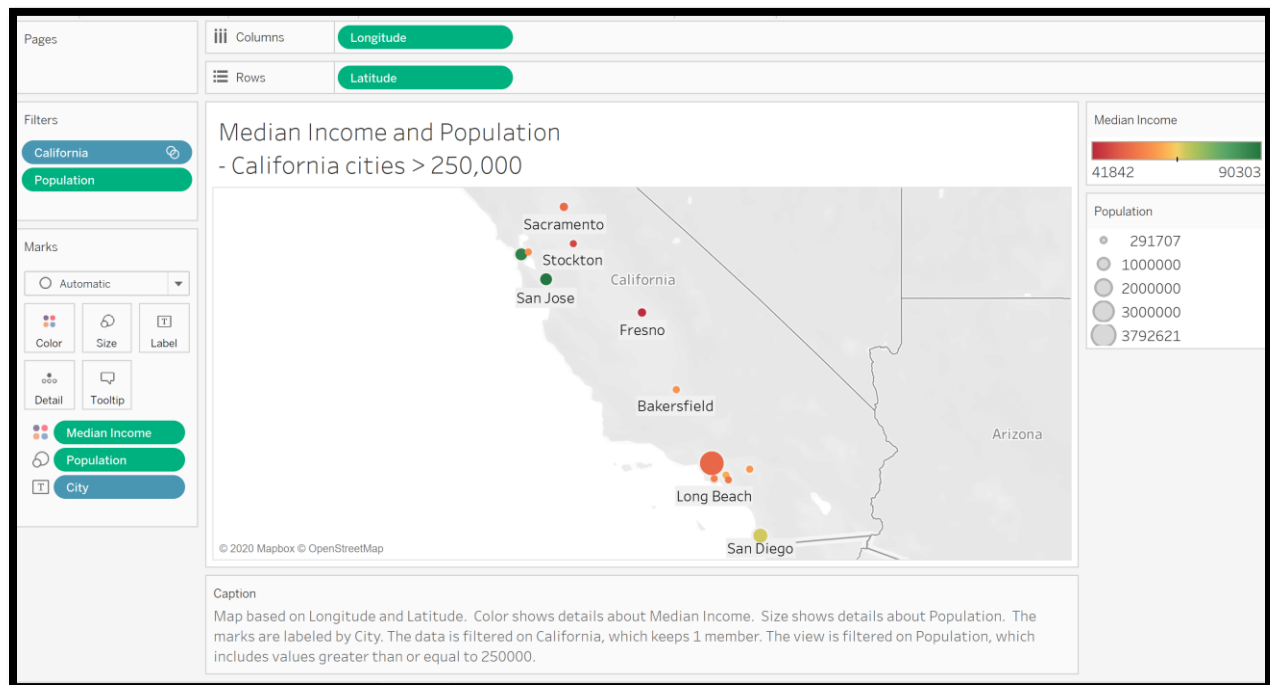
Copy the previous worksheet

In Tableau, we can create a filter using Sets. We will create a set to filter out non California states

   –   State Name-> open the dropdown menu > create 'set' -> Use the filtering tool – only
        show California-> Save it as a set: name = California
   –   The set California appears on the bottom

Median Income (bin)
Abc Source
State Id
State Name
Zip
Abc Measure Names

Measures

County Fips
Drive Alone
Employment
Latitude
Longitude
Mean Income
Median Income

Automatic

Add to Sheet
Show Filter

Duplicate
Rename
Hide

Aliases...

Create   ▶   Calculated Field...
Transform   ▶   Group...
                         Set...
Convert to Measure   Parameter...
Change Data Type ▶
Geographic Role ▶

Create Set   ×

Name: California

General | Condition | Top

◉ Select from list ○ Custom value list ○ Use all

Enter search text

☐ Alabama
☐ Arizona
☐ Arkansas
☑ California
☐ Colorado
☐ Connecticut
☐ District of Columbia
☐ Florida
☐ Georgia
☐ Idaho
☐ Illinois

All   None   ☐ Exclude

Latitude (generated)
Longitude (generated)

Sets

California

Data Source   Sheet 1   Sheet 1 (3)

− Move the set to the Filters shelf

Pages

Columns   Longitude
Rows   Latitude

Filters
California

Sheet 8

Marks

Automatic

Color Size Label
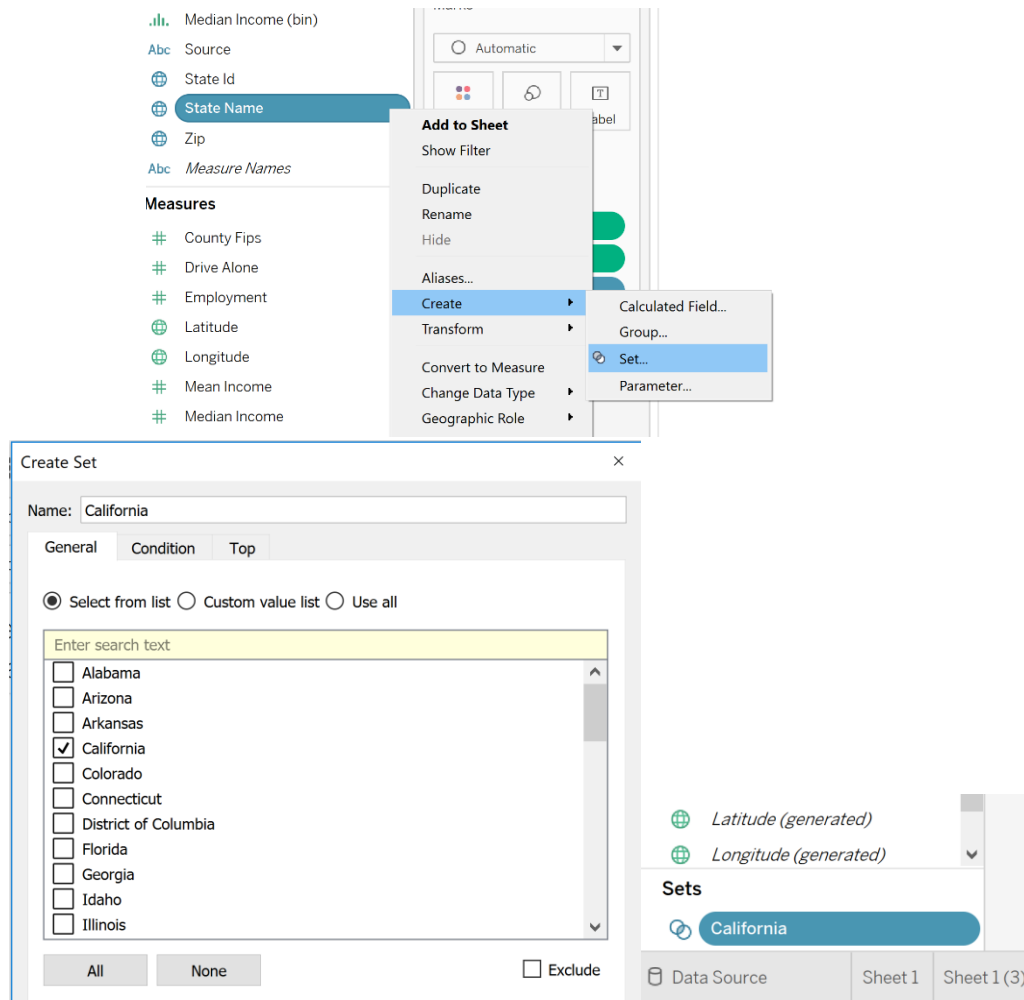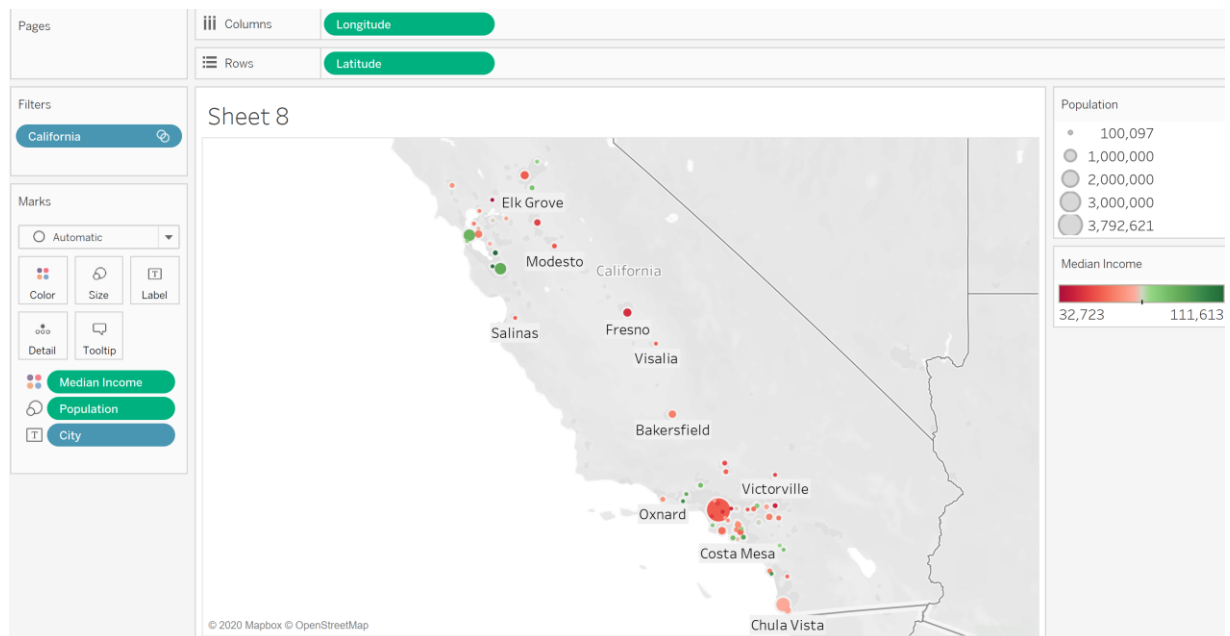Detail Tooltip

Median Income
Population
City

Population
• 100,097
• 1,000,000
• 2,000,000
• 3,000,000
• 3,792,621

Median Income
32,723   111,613

Elk Grove
Modesto
California
Salinas
Fresno
Visalia
Bakersfield
Victorville
Oxnard
Costa Mesa
Chula Vista

© 2020 Mapbox © OpenStreetMap

– Add Population to the Filter shelf: Present cities with population > 250,000

- Right click the sheet and view data. You can also export this data.
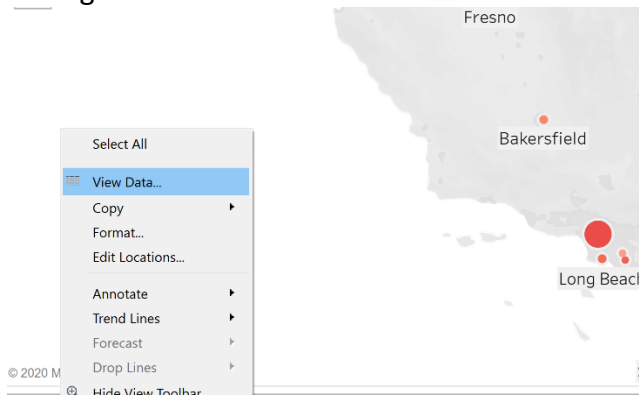
| City	Latitude	Longitude	Median Income	Population
Stockton	37.95770	-121.29078	46,033	291,707
Santa Ana	33.74557	-117.86783	54,062	324,528
San Jose	37.33939	-121.89496	90,303	945,942
San Francisco	37.77493	-122.41942	87,701	805,235
San Diego	32.71533	-117.15726	68,117	1,307,402
Sacramento	38.58157	-121.49440	52,071	466,488
Riverside	33.95335	-117.39616	58,979	303,871
Oakland	37.80437	-122.27080	57,778	390,724
Los Angeles	34.05223	-118.24368	51,538	3,792,621
Long Beach	33.76696	-118.18923	55,151	462,257
Fresno	36.74773	-119.77237	41,842	494,665
Bakersfield	35.37329	-119.01871	58,669	347,483
Anaheim	33.83529	-117.91450	61,826	336,265

– Sets and Groups are very similar. Unlike a Set, a Group is defined as a dimension.
– A set is a binary variable (In/Out), but a Group is a categorical variable

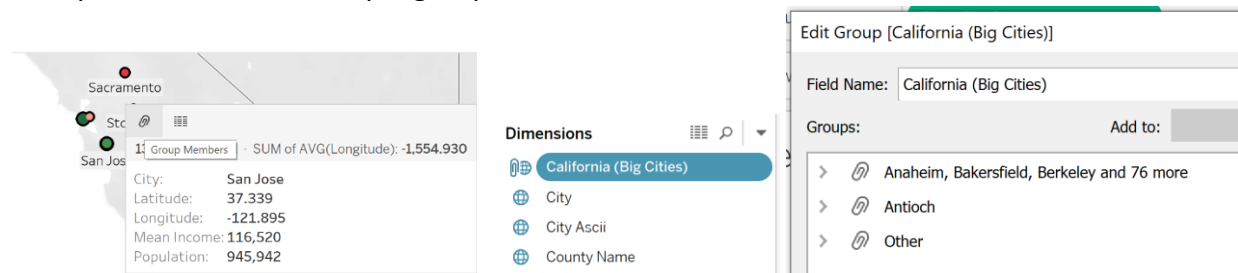− Select all and save it as a group or a set: name = Caalifornia (Big Cities)
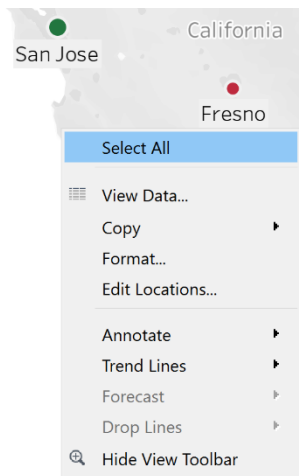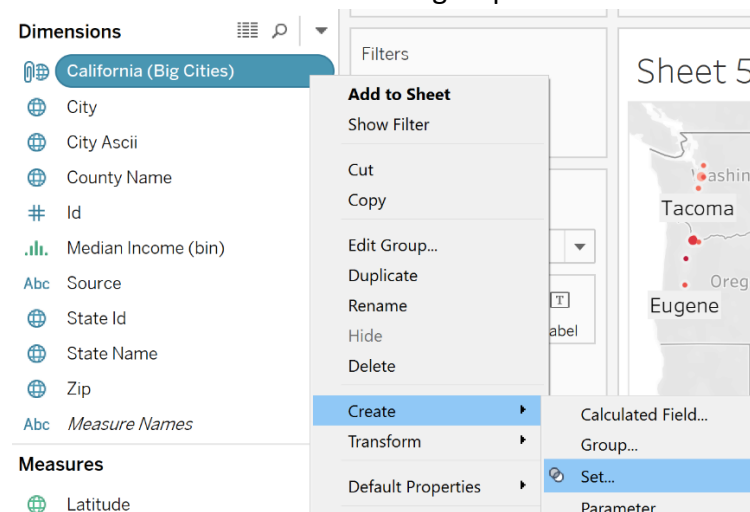

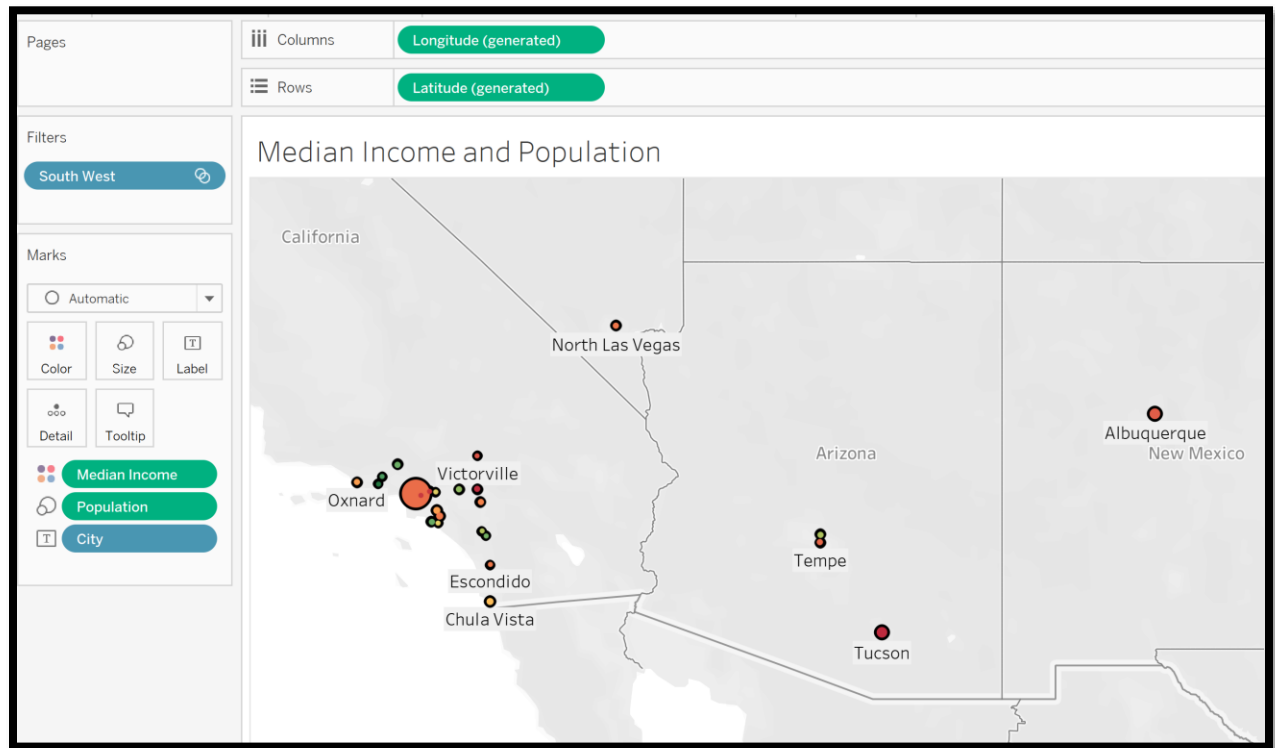**Set:**
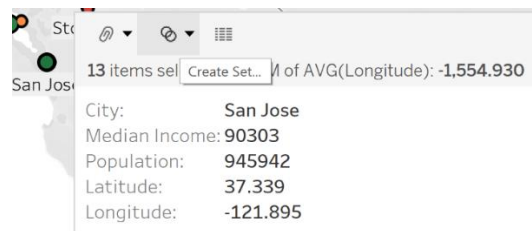
Sets

⊘ California

⊘ California (Big Cities)

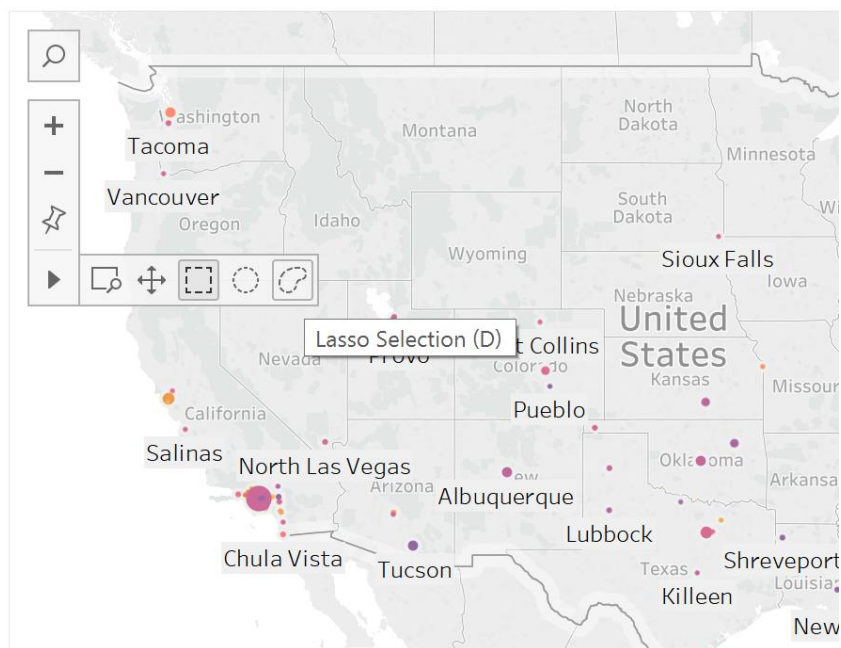**Group: You can make multiple groups**

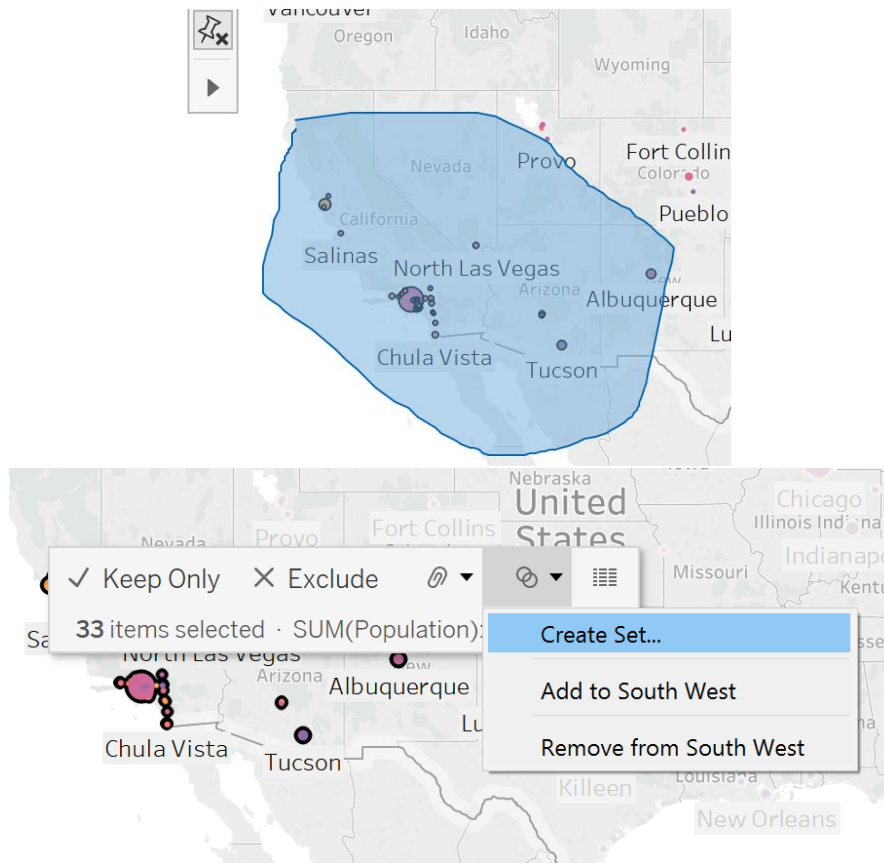**You can also create a set from a group:**

(iii) Lasso
Figure 3-(iii)

- Copy the first map (3-(i))
- Use the lasso tool – just select South West region
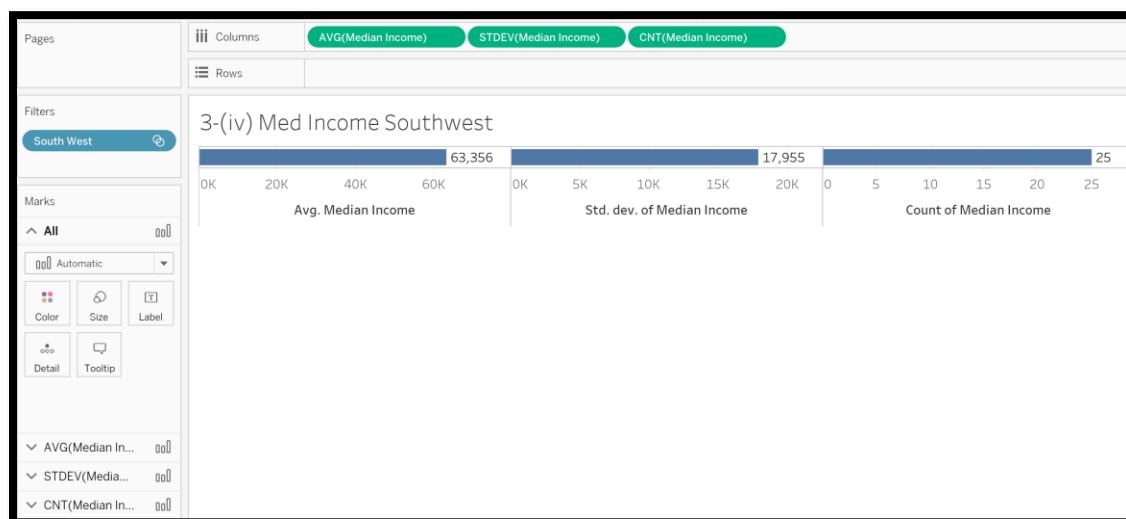   - Click the Lasso Selection

- Select a region and save it as a **set** or a **group**

- Move the set to the Filters shelf

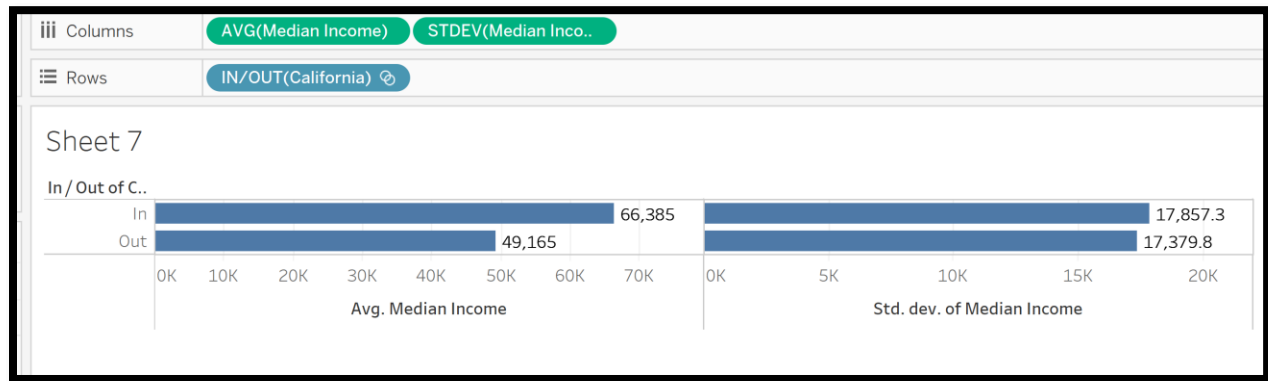**(iv) Comparisons: Summary statistics of South West Median Income**

Figure 3-(iv): Average, Stdev, and Count (Sample Size)

- Add Median income thre times to the Columns shelf. Change the format: average, stdev, and count
- Move the SouthWest set, to the Filters shelf

(v) Comparisons: California vs all other states

Figure 3-(v)

Compare the median income of California to all other states

- Add Median income two times to the Columns shelf. Change the format: average and stdev
- Move the California set, to the Rows shelf