

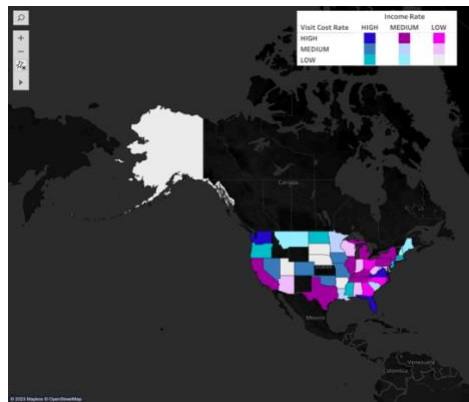
Tableau Bivariate Map Skill

Purpose: This demo shows how to create a Bivariate Map in Tableau. The Bivariate Map is useful when you want to create a map that is color coded based off multiple variables.

Question Addressed: Is there a correlation between average household income, and the total costs of visits in each state?

Dataset: HealthcareData.xlsx

Goal: Display the association between the average household income and the total amount of visit costs in each state. Below is a picture of what the completed chart looks like, giving you an idea of what our end goal is in this demo.



For a Bivariate Map, we need to identify two variables that we want to compare on a map simultaneously. If we only have one variable that we would like to see, then we can just make a regular choropleth map.

First, we must identify two variables that we would like to compare.

Picking our Variables

The Healthcare dataset has a measure called “Visit Cost,” which provides us with the total cost of each visit. This is an easy variable to measure and assess. For our second variable, I had to add to the dataset, as unsurprisingly, the other measure variables of copay and wait time had no variance from visit cost on a state-by-state basis. Thus, I created an income variable and randomly generated a household income value for every row in the set. This gives us a good variable to compare to visit cost in the dataset.

Prior to creating our map, we must do some manipulation to our variables.

Picking a Scale

Bivariate Maps can have different scales such as qualitative (variable that is not measured), binary (i.e. yes or no), diverging (i.e. above average, average, and below average), and sequential (i.e. high, medium, low). You must decide what scale represents your variable the best. For us, we are going to use a sequential scale for both of our variables.

Since our variables are currently in a numerical format, we are going to have to assign them a sequential value.

Calculation #1

We need to assign a sequential value, so we must figure out what our breaks will be for each value. We can do this by adding the “Visit Cost” variable into the worksheet, adding “Clinic State” as the detail, and sorting in descending order.

Clinic State	Visit Cost	
CA	2,387,998	Abc
PA	2,155,110	Abc
NY	2,146,913	Abc
TX	2,144,705	Abc
FL	1,647,573	Abc
IL	1,548,710	Abc
OH	1,412,583	Abc
MI	1,232,192	Abc
KY	1,215,437	Abc
NC	1,109,891	Abc
VA	1,011,679	Abc
WA	916,167	Abc
GA	903,707	Abc
TN	887,066	Abc
MO	816,623	Abc
IN	805,281	Abc
WI	784,010	Abc
AL	748,358	Abc
MN	699,676	Abc
MD	627,025	Abc
LA	615,585	Abc
WV	610,383	Abc
IA	582,331	Abc
OK	575,688	Abc
SC	543,660	Abc
CO	538,382	Abc
AZ	510,087	Abc
NV	487,360	Abc
NJ	479,969	Abc
AR	471,730	Abc
OR	443,735	Abc
MS	439,407	Abc
ME	364,873	Abc
NE	339,245	Abc
MA	295,982	Abc
UT	281,018	Abc
MT	213,298	Abc
SD	176,507	Abc
ND	176,395	Abc
NH	158,777	Abc
CT	142,843	Abc
AK	95,146	Abc

Now we can create our breaks based off this chart. Since we are going to evaluate our data on a state level, making this calculation a fixed level of detail will make our lives much easier later.

Visit Cost Rate

×

{FIXED [Clinic State]:

IF SUM([Visit Cost]) >= 850000 THEN 'HIGH'

ELSEIF SUM([Visit Cost]) >= 480000 THEN 'MEDIUM'

ELSE 'LOW'

END }

▶

The calculation is valid.

8 Dependencies ▼

Apply

OK

Now we can see that each state has been given a value based off their total visit costs.

We will then repeat the exact same process for our second variable, Household Income.

Calculation #2

Repeat the same process as calculation #1, but using average instead of sum:

Clinic State	Avg. Household Inc...
ND	211,343.14221219
CO	210,301.068081991
CT	209,891.569637883
NV	209,539.296492659
MO	209,265.070875665
NJ	208,764.470204082
MA	208,694.871485944
MS	208,570.62962963
IA	208,231.587618403
OK	208,092.778126088
VA	207,966.226525822
OR	207,686.323712256
WA	207,636.565564917
FL	207,609.695641734
MI	207,037.251762821
MN	206,858.613416714
KY	206,848.906794851
TN	206,706.947274352
PA	206,229.870044053
IL	206,229.111537969
NH	206,149.030864198
MT	206,137.887841659
TX	206,090.863657301
NY	206,016.915011547
SC	205,966.411850236
ME	205,957.687770563
LA	205,862.542570411
CA	205,392.606459131
GA	205,352.603105183
WV	205,300.665692008
IN	205,271.495093229
AZ	205,054.070243149
OH	204,952.165735828
AR	204,495.269086358
AL	204,162.599149389
WI	203,968.309127584
SD	203,627.595105673
NE	203,534.772063121
UT	203,372.256822953
MD	203,047.701384519
NC	202,921.268733392
AK	200,189.741414141

Income Rate

```

{FIXED [Clinic State]:
IF AVG([Household Income]) >= 207500 THEN "HIGH"
ELSEIF AVG([Household Income]) >= 205392 THEN "MEDIUM"
ELSE "LOW"
END }

```

The calculation is valid.
7 Dependencies
Apply
OK

Now that we have sequential values for our variables, we can combine the two variables to make one.

Combining Variables

Combining variables is very simple and can be done even if we had not used a fixed LOD in our calculation. We will simply create a string variable that consists of the two sequential values.

Cost -> Income

```

[Visit Cost Rate] + ' - ' + [Income Rate]

```

The calculation is valid.
4 Dependencies
Apply
OK

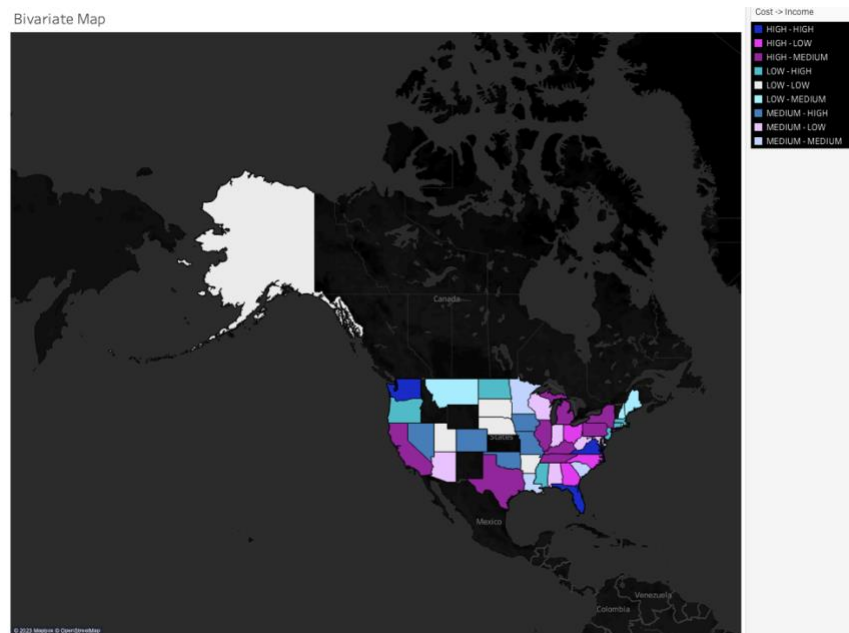
Resulting in a set of variables (below) that are assigned to each state.

HIGH - HIGH
HIGH - LOW
HIGH - MEDIUM
LOW - HIGH
LOW - LOW
LOW - MEDIUM
MEDIUM - HIGH
MEDIUM - LOW
MEDIUM - MEDIUM

Now that our variables are combined, we can create our map.

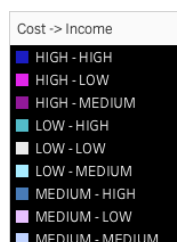
Creating the Bivariate Map

In a new worksheet, we will create the map, just like we would a regular choropleth map. Put the generated Latitude and Longitude on the Rows and Columns respectively. Then drag our combined variable, “Cost -> Income” on to colors, and “Clinic State” on to detail. You should get a bivariate map that looks similar to the one below:



You can make the map dark by selecting dark in themes under Map -> Background Layers. An important part to creating a good bivariate map is to choose a good color scheme. The colors need to match up with the variables represented in the map, and also contrast each other. Here, blue and pink contrast each other so I used the two to represent the contrasting variables, and put it on the dark theme so that the “Low-Low” states would not get confused for empty states.

Now, our Bivariate Map is complete. However, there is one more visual feature that will improve your map. The default legend for a Bivariate Map looks like this:












This is fine, but we can create our own legend that is far more visually appealing.

Bonus

Note: We can only create this other legend if we created our “Income Rate” and “Visit Cost Rate” variables as fixed LODs. If we did not, then we will be unable to create this legend.

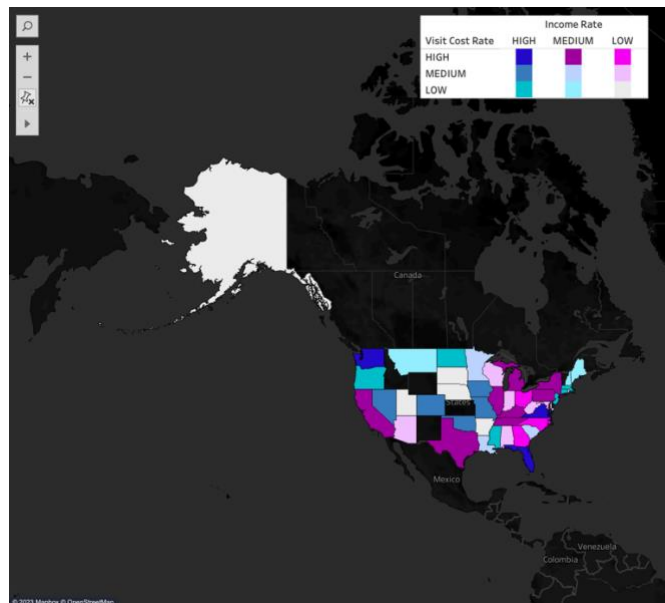
Creating a Legend

In a new worksheet, we will put our “Income Rate” and “Visit Cost Rate” on the Columns and Rows, and put “Cost -> Income” on the colors. This should put our sequential variables in a 3x3 grid:

Visit Cost Rate	Income Rate		
	HIGH	MEDIUM	LOW
HIGH			
MEDIUM			
LOW			

Select the colors for your color scheme for each square. Make sure “square” is selected in the marks section.

We now have a legend that is much more visually appealing and applicable to our map, making it worth it to create the fixed LOD. We can use this legend when we put out Bivariate Map on a dashboard. The result should look something like this:



Make sure to select “floating” on the legend so that you can place it on top of the map and select “Cost -> Income” under Highlight to make your legend interactive allowing users to click on the legend to highlight sections of the map.

