

Using Census Data to Calculate Social Inequality in Colombia

Written by Carolyn Talmadge, updated Oct 8, 2019 for 10.7.1 by Phil Gomez

What is Social Inequality?	1
Accessing the Data	2
Understanding the Metadata & Attribute Fields.....	2
Setting Symbology and Understanding Classification Methods	5
Ranking your Data using Field Calculator	6
Examining & Ranking NBI 2010	12
Examining and Ranking Gross Education Coverage	15
Using Field Calculator to Calculate Vulnerability.....	17
Using Symbology to Represent Vulnerability	18
Map Layout & Cartography Tips and Tricks.....	20
Further Research & Analysis	21

What is Social Inequality?

Social Inequality refers to the unequal distribution or access to goods, services, and opportunities across a population and can be quantified a number of different ways. There are numerous indicators for social inequality, ranging from employment, food access, education level, health status, income, adequate housing, sanitation, and many more. These indicators can be combined to calculate a measure of social inequality or vulnerability within a population.

Tutorial Introduction

The purpose of this tutorial is to examine several indicators of social inequality collected by the Colombian Census by *municipality* in order to calculate levels of vulnerability within the country's population.

The Social Indicators downloaded from Colombia's [National Administrative Department of Statistics \(DANE\)](#) used to determine vulnerability for this particular analysis are:

- Percentage of Households with **water sanitation services** per municipality, 2011
- Percentage of Households living in conditions of poverty or misery according to unsatisfied basic needs per municipality, 2010 (**Unmet Basic Needs Index – NBI**)
- Gross **Education** Index Per Municipality, 2013

In this tutorial, you will learn how to use the basic attribute table tools and calculators to create a vulnerability model within the Attribute Table at the municipality level. Functions to be covered:

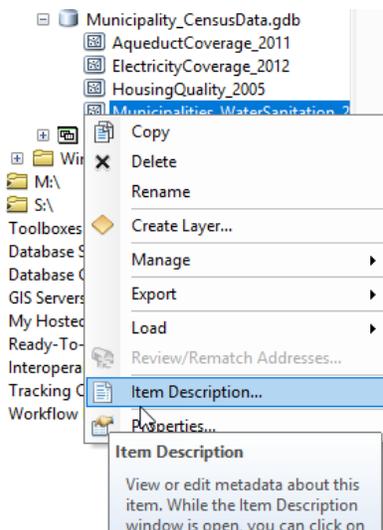
<i>Joins</i>	<i>Statistics</i>
<i>Add a Field</i>	<i>Summarize</i>
<i>Field Calculator</i>	<i>Classification Methods</i>

Accessing the Data

1. Open a blank ArcMap Session.
2. In **Catalog**, navigate to *S:\Tutorials & Tip Sheets\Tufts\Tutorial Data\SocialInequality_Vulnerability_Colombia*.
3. **Right click on the folder and copy** the entire folder to your **H: drive or USB**.
2. You need to make the copied folder “writable” – to do this, right-click on the folder in your H drive, choose Properties, and uncheck the Read-Only box. When prompted, say yes to *Apply changes to this folder, subfolders and files*.
3. Back in **ArcCatalog**, navigate to your H Drive and expand the **SocialInequality_Vulnerability_Colombia** folder. Inside you will see a **Geodatabase** called **Municipality_CensusData.gdb**. Double click again to expand the geodatabase to see the feature class data sets.
4. Drag the **Municipalities_WaterSanitation_2011** feature class data set into the ArcMap Session. The data for this tutorial is all at the **Municipality** level, which is their 2nd level administrative boundary in Colombia.

Understanding the Metadata & Attribute Fields

1. The first thing we want to do is understand what is in this dataset. Let’s explore the metadata for this feature class. In **Catalog**, right click on **Municipalities_WaterSanitation_2011** and select **Item Description**.

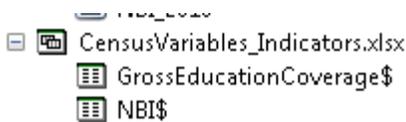


2. This pulls up some basic metadata, including a description of the dataset so you can understand what the data is measuring. In this case, **Municipalities_WaterSanitation_2011** has information on the **percent of households with Sanitation Services** (toilets, etc) in 2011. However, right now it is just showing the municipality boundaries since we have not set the symbology.
3. Right click on **Municipalities_WaterSanitation_2011** → **Open Attribute Table**
4. Examine the contents of the Attribute Table. Can you figure out what each of the fields are displaying? Which field holds the values for **Percent with Sanitation Services**? Make sure to remember the field-heading name.

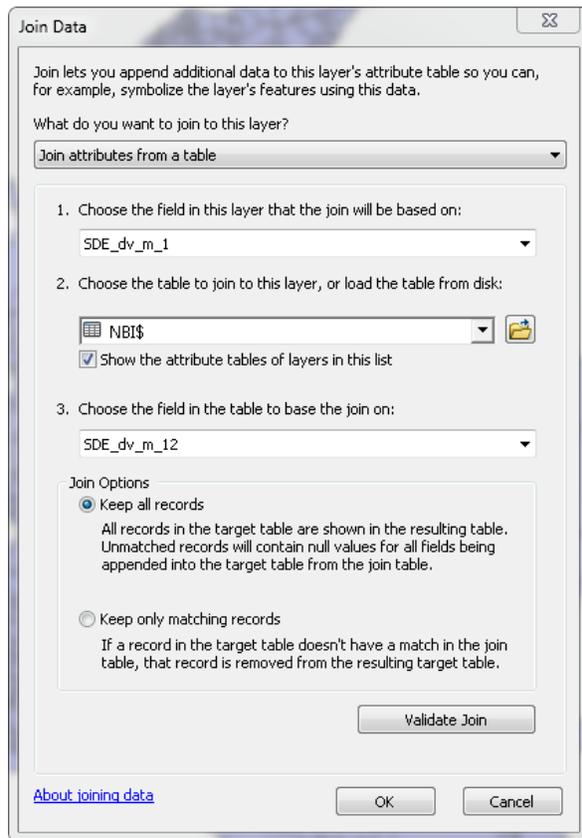
OBJECTID	Shape	SDE_dv_m_1	SDE_dv_m_3	SDE_dv_m_4	SDE_dv_m_5	dbo_V_Dato	dbo_V_Da_1	dbo_V_Da_2	Shape_Length	Shape_Area
1	Polygon	23189	Cabecera Municipal	CaRDOBA	CIRIACA DE ORO	2011	18.8	< 40	185694.470075	637488378.901902
2	Polygon	23570	Cabecera Municipal	CaRDOBA	PABLO NUEVO	2011	15	< 40	167091.371727	794996270.607521
3	Polygon	2306	Cabecera Municipal	CaRDOBA	PADEL	2011	17	< 40	228151.654587	1932619588.406299
4	Polygon	23570	Cabecera Municipal	CaRDOBA	BERNABERTEADOR	2011	17.3	< 40	280408.875116	1650731091.37717
5	Polygon	2306	Cabecera Municipal	CaRDOBA	BO	2011	2.2	< 40	168895.632958	472568643.049477
6	Polygon	231	Cabecera Municipal	CaRDOBA	SUCRE	2011	4.1	< 40	46900.864407	48973353.945247
7	Polygon	2323	Cabecera Municipal	SUCRE	SUCRE	2011	35.8	< 40	70751.798305	176253978.871982
8	Polygon	23820	Cabecera Municipal	SUCRE	SUCRE	2011	45.9	40.1 - 60	97812.896797	315355147.374673
9	Polygon	230713	Cabecera Municipal	SAN ONOFRE	SAN ONOFRE	2011	11.3	< 40	204608.251218	1071137577.183251
10	Polygon	2307473	Cabecera Municipal	MORROA	MORROA	2011	56.2	40.1 - 60	65409.436922	178554027.844101
11	Polygon	2307473	Cabecera Municipal	LOS PALMITOS	LOS PALMITOS	2011	31.4	< 40	81308.510437	20773777.408613
12	Polygon	2307473	Cabecera Municipal	SUCRE	SUCRE	2011	0.9	< 40	192837.58072	1066910635.891066
13	Polygon	2307473	Cabecera Municipal	SUCRE	MAJAGUAL	2011	0.6	< 40	187592.066085	971155353.345167
14	Polygon	2307473	Cabecera Municipal	DEPTO Distrito Turistico	CARTAGENA DE INDIAS (Distrito Turistico)	2011	76.7	60.1 - 80	302050.260239	614286474.336798
15	Polygon	2307473	Cabecera Municipal	BOLEVAR	SANTA CATALINA	2011	4.6	< 40	100771.78463	159708303.938372
16	Polygon	2307473	Cabecera Municipal	BOLEVAR	PINILLOS	2011	7	< 40	224462.8037	770781322.806667
17	Polygon	13074	Cabecera Municipal	BOLEVAR	BARRANCO DE LOBA	2011	1.3	< 40	166601.829995	430153348.187908
18	Polygon	13600	Cabecera Municipal	BOLEVAR	RIOVIEJO	2011	4.1	< 40	308853.340072	856989396.333285
19	Polygon	13473	Cabecera Municipal	BOLEVAR	MORALES	2011	40.8	40.1 - 60	209796.66299	1339226864.989113
20	Polygon	13688	Cabecera Municipal	BOLEVAR	SANTA ROSA DEL SUR	2011	46.6	40.1 - 60	349114.247189	2385046749.692915
21	Polygon	13160	Cabecera Municipal	BOLEVAR	CANTAGALLO	2011	35.2	< 40	174362.054989	867648209.988746
22	Polygon	13006	Cabecera Municipal	BOLEVAR	ACHE	2011	0.4	< 40	184837.649884	96927146.826773
23	Polygon	13030	Cabecera Municipal	BOLEVAR	ALTOS DEL ROSARIO	2011	0.2	< 40	103640.458329	305930304.8271
24	Polygon	13300	Cabecera Municipal	BOLEVAR	HATILLO DE LOBA	2011	0.4	< 40	79028.693046	156079001.219986
25	Polygon	13810	Cabecera Municipal	BOLEVAR	TIQUISIO (Puerto Rico)	2011	4.8	< 40	160604.650384	749088195.560191

- Notice that there is a unique identifier for each *municipality*, **SDE_dv_m_1**. This will be very important later in the tutorial.
- Since we are calculating a vulnerability index, it's important that we analyze several indicators of vulnerability. This particular indicator, *percent water sanitation coverage*, already came as a feature class. However, our other two indicators, **Unmet Basic Needs Index (NBI)** and **Gross Education Index**, were downloaded from the Colombian Census as **Excel tables**. This means we need to **JOIN** them to our existing municipality data with water sanitation information.

- In Catalog, expand the excel workbook by clicking the plus sign to show the sheets.



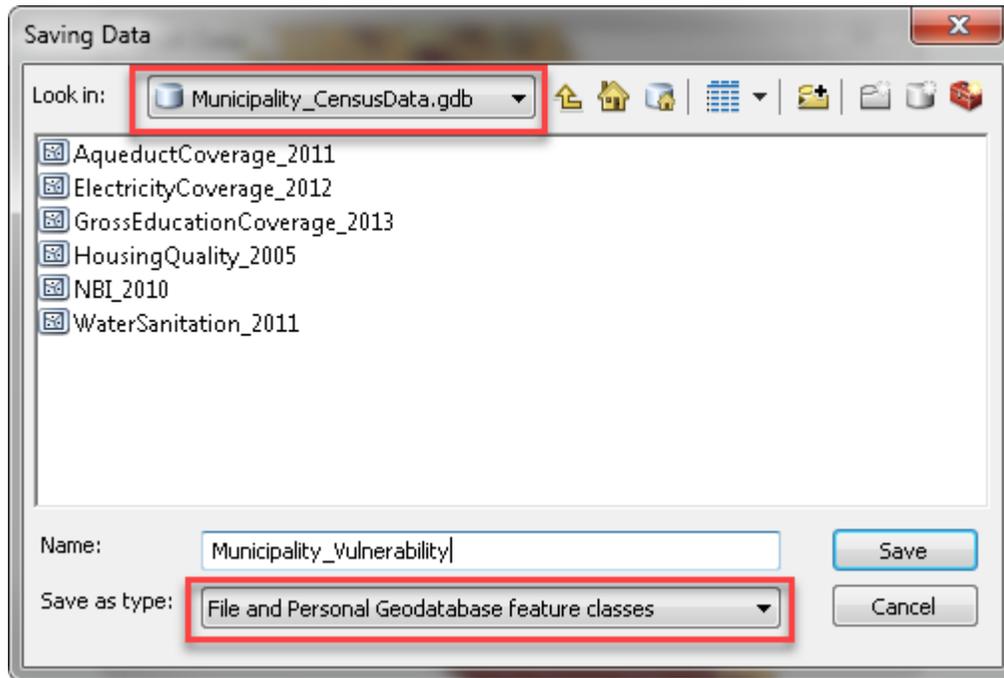
- Drag in both sheets into the data frame as you would a shapefile.
- In the Table of Contents, right click on **NBI\$** and **Open**. Here, we can examine the data that is in this excel table. We see three columns. The Municipality Code, the NBI Index and the year the data was compiled.
- It's really important that we know what the NBI is measuring. Let's refer to the metadata. In Windows Folders, navigate to your H Drive → SocialInequality_Vulnerability_Colombia → And open the **metadata** word document for NBI. What does the NBI measure? Is a high or low number more or less vulnerable?
- To simplify our work, we are going to **join** the **NBI Excel data** to the same **Municipalities_WaterSanitation_2011** feature class. Since they share the **same municipality unique identifier** (which is very important), we can join the NBI data to the same layer. **That way, all the data is in one attribute table and we do not need to combine it later.**
- Close the table and then **right click** on **Municipalities_WaterSanitation_2011** → **Joins and Relates** → **Join...**
- Make sure the NBI\$ is the selected table in step 2 of the Join window. For step 1, select **SDE_dv_m_1**, which was the Unique Identifier in the Municipalities layer. Now, it automatically recognizes the matching field in the excel table, **SDE_dv_m_12**. It is important to make sure that the matching identifiers are chosen in step 1 and 3. Otherwise, the join will not work. Ensure that the tool is filled out like the screen shot below:



14. Press Ok. Now, open the attribute table for **Municipalities_WaterSanitation_2011** and scroll to the right to ensure the data joined properly. You should see the 3 fields from the excel table now at the end.
15. Now, let's join the **GrossEducationCoverage\$** excel file to the same layer as well. It will be really helpful to have all this census data in the same attribute table.
16. But first, let's **open** the **GrossEducationCoverage\$** and make sure everything looks okay and we understand the metadata. You can find the **metadata word document** in the same folder. Read through the description of the data.
17. Now that we understand what this is measuring, right click on **Municipalities_WaterSanitation_2011** → **Joins and Relates** → **Join...**
18. Fill out the join table as we did for NBI, but **make sure the the GrossEducationCoverage\$ is selected for step 2**. Or else, we're just rejoining the NBI data. Press ok.
19. Open the attribute table for **Municipalities_WaterSanitation_2011** and scroll to the right to ensure the data joined properly. You should see the 6 fields from the 2 excel tables now at the end.
20. It's good practice to **Export** a Join into a new feature class or shapefile. This is because joins are **not permanent** and if you were to run some analysis on **Municipalities_WaterSanitation_2011**, it would drop the two tables you just joined. But if you export the data, it saves the join as the attribute table and will be permanent.
Right click on **Municipalities_WaterSanitation_2011** → **Data** → **Export Data**
21. Press the folder icon to choose where you are saving the new feature class.
22. Navigate to your **H drive** → **SocialInequality_Vulnerability_Colombia** folder → **Municipality_CensusData.gdb**
23. Name the file **Municipality_Vulnerability** and make sure it is being saved as **File and Personal Geodatabase Feature Class**. If you don't see the .gdb, it's because the Save as Type probably still says shapefile. **Press Save and then Ok**.

It's important to choose a name that tells what the feature class is mapping (**municipalities**) and by adding

Vulnerability, we know this is the feature class that has all of our indicator data in the attribute table.



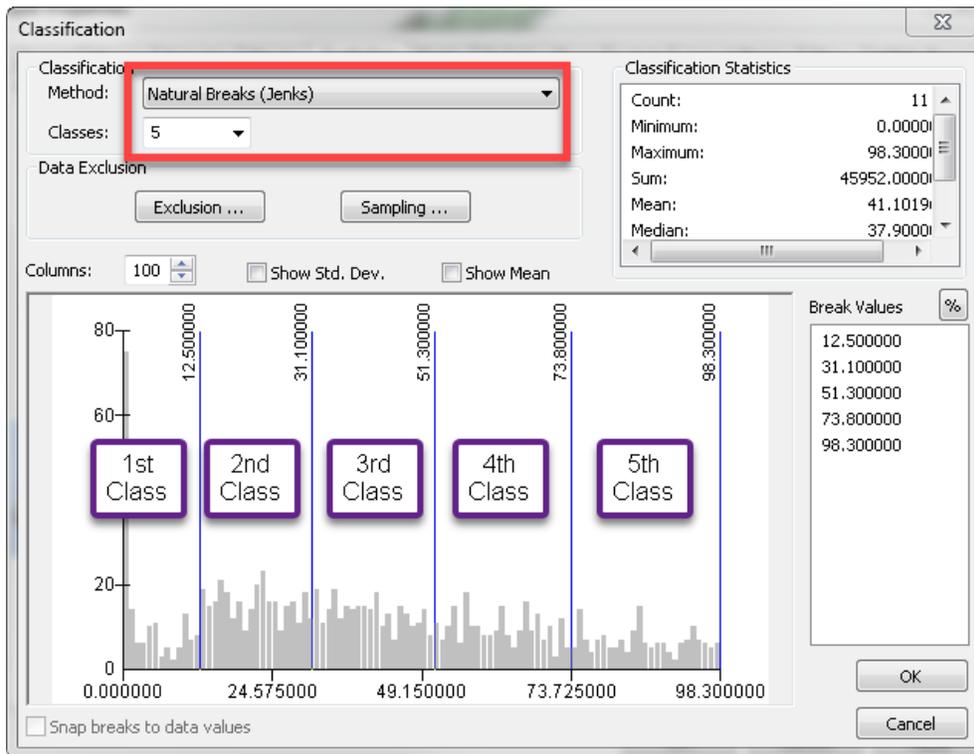
24. Click YES when asked if you want to add the exported data. Now you should see a new layer appears in the *Table of Contents* AND in *Catalog*.

Setting Symbology and Understanding Classification Methods

Eventually we will be combining the 3 indicators to create a **Vulnerability Model** for Social Inequality. However, we cannot just add the 3 fields together because the individual fields are on a different relative scale. We need to **rank** each indicator from low (0) to high (3) vulnerability. Then we can add each ranked field together to get a social inequality score from 0 – 9. **Examining the Symbology and the Classification Methods is a good way to understand how you are going to rank your data.**

Note: In your future analyses, it is up to you to decide how many classes you want to use to rank your data. In this case, we are using 4 classes (0, 1, 2, 3). However, it is not limited to 4. As the researcher, you must decide what is appropriate for your analysis.

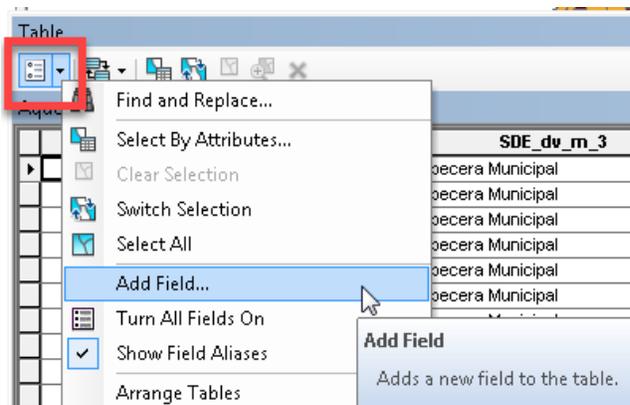
1. Let's start with the Water Sanitation data. Open the **symbology** properties for our new **Municipality_Vulnerability** layer.
2. We want to show the *percent of households with sanitation services per municipality*, which is a **Quantity**. Therefore, select **Quantities** on the left hand side. Then, set the field value to **dbo_V_Da_1**. Press Apply. The map has now updated to show the percent of households with sanitation services by graduated colors (low to high).
3. In the top right of **Symbology**, click on **Classify**. Now we can see the distribution of the values in this field by examining the **histogram**. This data is already broken up into **5 classes**. Each blue line represents a **class break**. For example, 0 – 12.5 is the 1st class, then 12.5 to 31.1 is a 2nd, and so forth and so on.



- Currently, the data is broken up using the **“Natural Breaks”** method. Natural breaks classes are based on natural groupings inherent in the data. Class breaks are identified that best group similar values and that maximize the differences between classes. In this exercise you will use Natural Breaks to determine your ranking classification from 0-3.
- Try a few of the different classification methods. Press OK and Apply to see how it changes the map. **How you group your data is VERY IMPORTANT** because where you break up the data can drastically alter the message your map is sending.
- Set the classification method back to **Natural Breaks** and **change the number of classes to 4**. You will use **these natural breaks to rank your data from low (0) to high (3) Vulnerability**. Then press OK and then OK to exit out of symbology.

Ranking your Data using Field Calculator

- Open the attribute table for **Municipality_Vulnerability**. First, we need to add a field to hold our vulnerability rankings for Sanitation Coverage. Click on the **Attribute Table options** → **Add Field...**



- Let's call this new field **Sanitation_Rank** and keep it as a **Short Integer**. Remember, no spaces or characters other than an underscore (_). Also, short integer means that the field will be a numeric field, but only allow for integers (aka no decimals). Press Ok.
- Scroll to the right in your attribute table. You will see that a new field has been added with our new name.

However, all the values are <Null> because we have not calculated them yet.

- Now, we need to decide how we are going to **Rank** Percent of households with sanitation services. We decided when exploring the **classification methods** in symbology that the natural breaks method was a good way of dividing the data. **Let's use those break values for the 4 classes we have already set.**

The important thing to remember is that we are measuring Inequality. **So LOW percent of sanitation coverage is actually HIGH inequality/vulnerability. This is very easy to get backwards, so be careful with EACH indicator.**

- Using the **natural breaks classification from Symbology**, our data on percent sanitation coverage will be ranked as follows:
 - 0.000000 - 20.800000 % = Maximum Vulnerability (3)
 - 20.800001 - 44.800000% = High Vulnerability (2)
 - 44.800001 - 70.500000% = Medium Vulnerability (1)
 - 70.500001 - 98.300000% = Low Vulnerability (0)

Municipalities_Vulnerability

dbo_V_Da_1

 0.000000 - 20.800000

 20.800001 - 44.800000

 44.800001 - 70.500000

 70.500001 - 98.300000

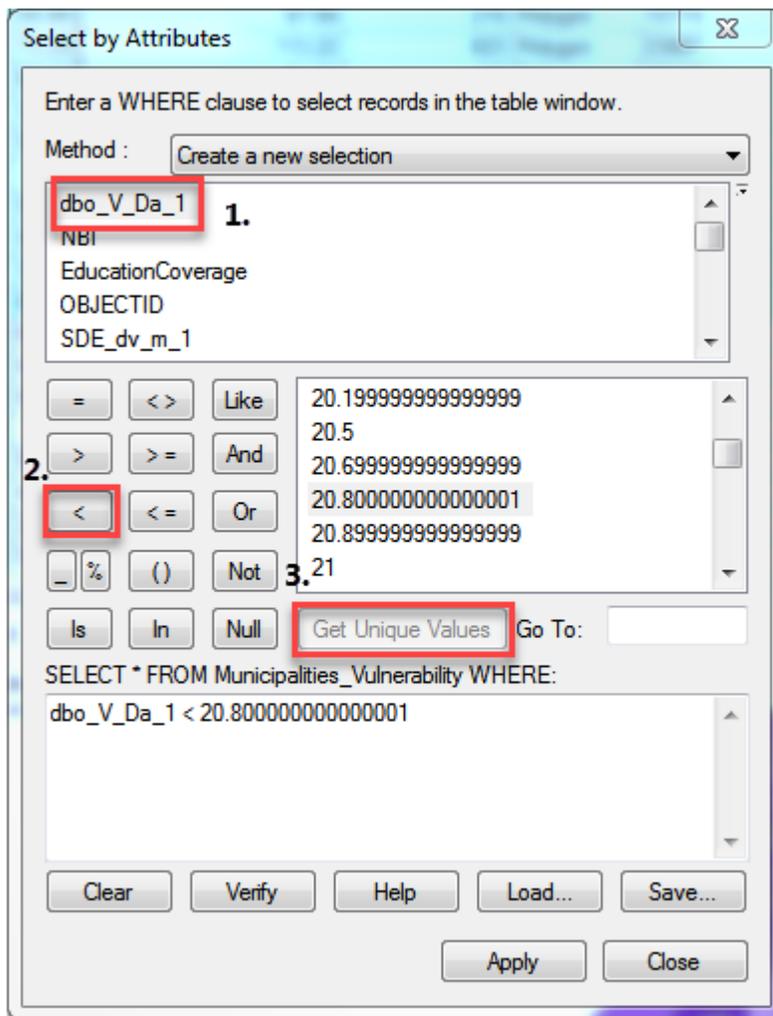
Note: If your break classification is slightly different, you can continue with your values in the next steps.

- In this new field, we need to assign a score of 3 to the MOST vulnerable municipalities. We can do this a few ways.
 - We can use **Select by Attributes** to Query the data and select all rows < 20.80
 - We can **Manually Select** the rows that have a value less than 20.80

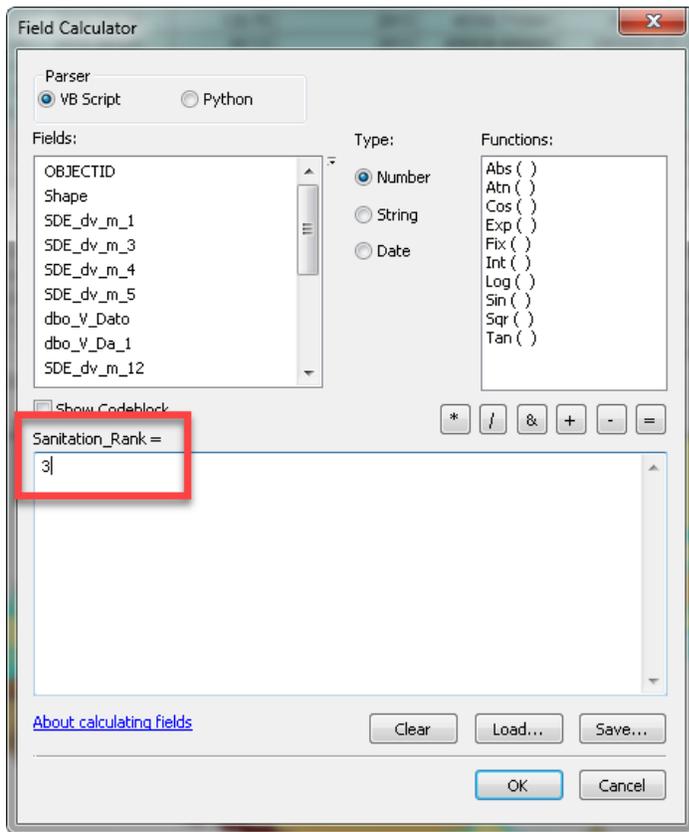
Option A is a more robust method, so let's use that. We will use a structured query to select the municipalities within each classification.

Note: Remember to **use the buttons** in the Select by Attribute section, as queries can be sensitive to formatting.

- Click on **Attribute Table options**  → **Select By Attributes**. Double click on **dbo_V_Da_1** to add it to the query window. Click on the less than button (<), then click "Get Unique Values." The values for **dbo_V_Da_1** will appear. Scroll down to 20.80 and double click on it. This should produce an expression that looks like: "dbo_V_Da_1 < 20.800000000000001".



8. Now we can see that 293 out of 1118 rows are selected and have values ranging from 0 – 20.8.
9. Now, with these values selected, right click on our new field, **Sanitation_Rank**, and select **Field Calculator...** **Check the box "Don't warn me again"**. Then press Yes to the warning.
10. This is where we will now assign a **Ranking/Score Value** to our selected municipalities. These are the municipalities with the **lowest** percentages with sanitation services, there for **they are the most vulnerable** and will be given a **score of 3**. In the Field Calculator, enter 3. The calculator already knows the expression is **Sanitation_Rank =** because we clicked on this field when opening the field calculator.



11. Press ok. Now all the selected municipalities have been given a score of 3 while the unselected rows remained <Null>.

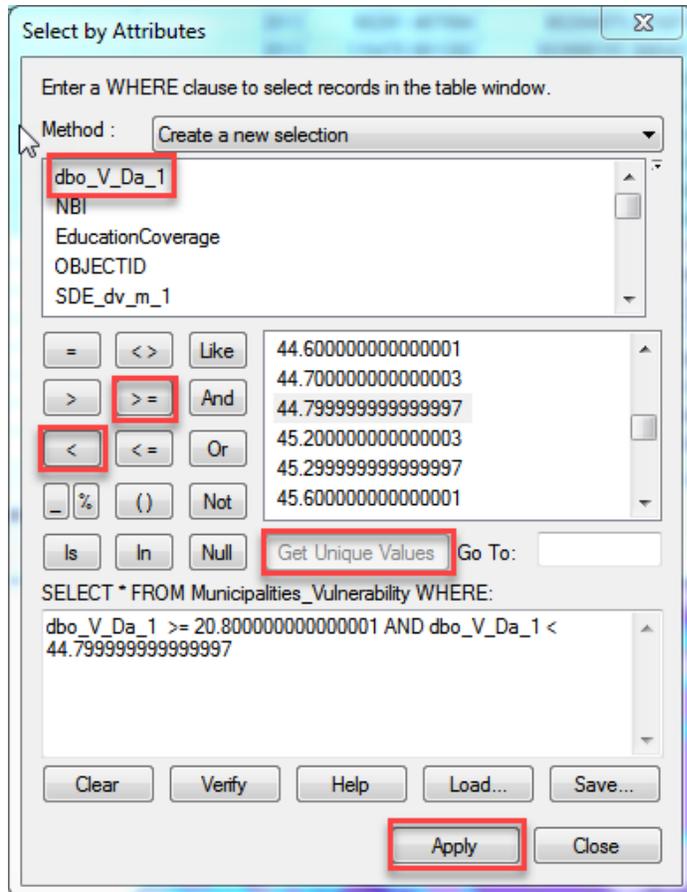
Note: When you run a tool or operation with selected values, it ONLY runs the tool on the selected rows. This is true for ALL tools, not just field calculator.

dbo_V_Da_1	NBI	SDE_dv_m_12	Date	SDE_dv_m_12_13	dbo_V_3_Date	Shape_Length	Shape_Area	Sanitation_Rank
19.6	80.28	19785	2010	19785	2013	78057.484075	136372925.922603	3
19.8	32.09	25095	2010	25095	2013	51173.307825	62274673.606578	3
20	55.95	15774	2010	15774	2013	82659.226343	183864066.063832	3
20	73.29	23807	2010	23807	2013	385389.340387	4929629727.298199	3
20.1	66.33	20310	2010	20310	2013	46028.781661	64196257.386503	3
20.1	37.61	25436	2010	25436	2013	49738.601871	108505762.870178	3
20.2	50.83	54398	2010	54398	2013	110498.396648	241902121.061783	3
20.2	56.8	68684	2010	68684	2013	41198.942393	76610915.286993	3
20.2	50.29	19743	2010	19743	2013	144658.161015	626924904.21382	3
20.5	42.98	68773	2010	68773	2013	167210.251133	483722357.279921	3
20.7	36.12	15090	2010	15090	2013	48815.979813	61966323.813712	3
20.7	60.77	18460	2010	18460	2013	175443.346107	1232002249.977216	3
20.8	68.93	13667	2010	13667	2013	139381.759746	462989665.285126	<Null>
20.8	40.62	86757	2010	86757	2013	142962.421503	386201397.289559	<Null>
20.9	44.97	41530	2010	41530	2013	66488.785027	171759736.068367	<Null>
20.9	36.9	15761	2010	15761	2013	37316.37007	57176667.644496	<Null>
21	56.53	54800	2010	54800	2013	264293.568717	933156192.223361	<Null>
21	44.09	25168	2010	25168	2013	66189.383547	171833396.311878	<Null>
21.2	27.27	25299	2010	25299	2013	46414.290749	108903735.501216	<Null>
21.3	57.91	91798	2010	91798	2013	570367.090028	9137708226.44182	<Null>

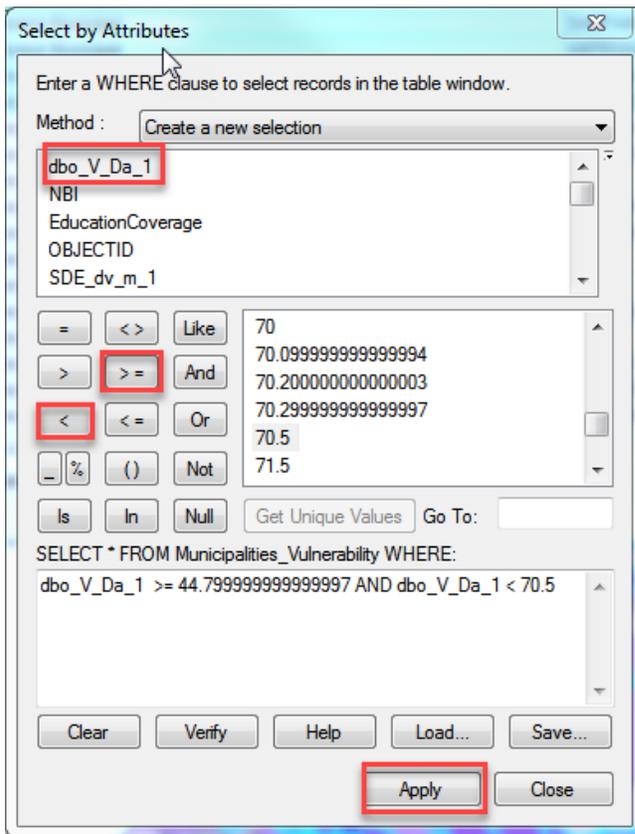
12. Now we need to assign the next values a score of 2 for medium-high vulnerability, which we have determined will be from 20.8 to 44.8. Clear your selection .

13. Repeat the same steps, this time selecting the rows from 20.8 down to 44.8. Click on **Attribute Table**

options → **Select By Attributes**. Double click on `dbo_V_Da_1` to add it to the query window. Click on the greater than or equal to button (`>=`), then click “Get Unique Values.” Scroll down to 20.80 and double click on it. Next, click on the And button (`And`). Next, double click on `dbo_V_Da_1` again, and less the less than button (`<`) and “Get Unique Values.” Scroll down to 44.79 and double click. This should produce an expression that looks like: “`dbo_V_Da_1 >= 20.800000000000001 AND dbo_V_Da_1 < 44.799999999999997`”.

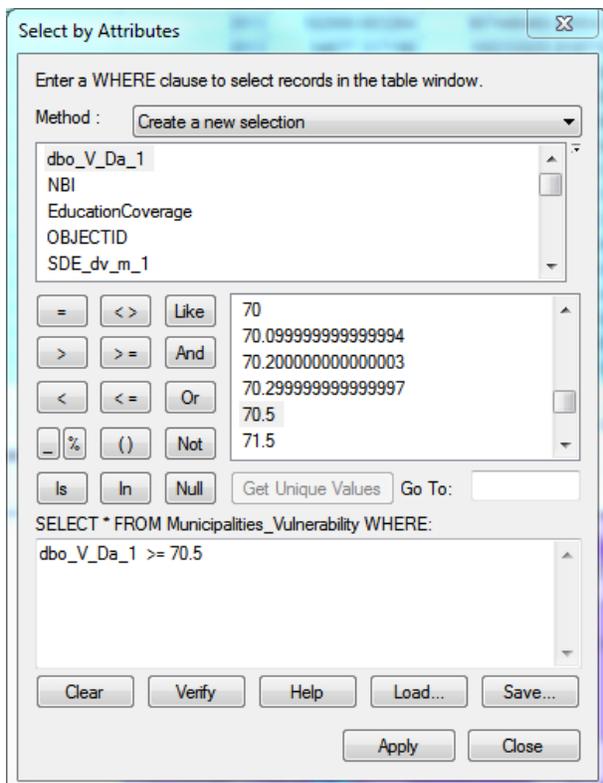


14. You should now have 361 out of 1118 rows selected.
15. Again, right click on **Sanitation_Rank** and select **Field Calculator...** Change the value from 3 to 2 and press ok. We are now assigning a score of 2 to the medium-high vulnerability municipalities. Press Ok and the selected cells will populate with 2.
16. Repeat the same steps, now using a query to select the rows between 44.79 and 70.5. You should have 276 rows selected and your expression should match the image below. Be sure to use the buttons.



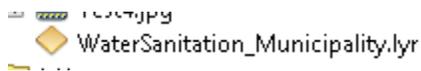
17. Again, Right click on **Sanitation_Rank** and select **Field Calculator...** Change the value from 2 to 1 and press ok. We are now assigning a score of 1 to the medium-low vulnerability municipalities. Press Ok and the selected cells will populate with 1.
18. Finally, we need to give the municipalities with high percentages of sanitation (70.5 – 100%) a score of 0 because they are least vulnerable. **Clear** the last selection.

Again, repeat the same steps, now using a query to select the rows between 70.5 and 98.3. You should have 188 rows selected and your expression should match the image below.



19. Return to the field calculator and give these values a score of 0. Press ok. After it populates, **clear** selection.
20. Now, the entire **Sanitation_Rank** Attribute Field should be populated with vulnerability scores from 0 -3. Make sure you did not miss any and there are no <Null> values.
21. We have completed the ranking/scoring for Water Sanitation Coverage. Let's save the symbology for the Water Sanitation Field. To do so, right click on Municipality_Vulnerability → **Save as Layer File...**
22. **Navigate to your H Drive → SocialInequality_Vulnerability_Colombia.** Name this layer WaterSanitation_Municipality and press **Save**.

This is a way of saving the **symbology properties** for a layer so you don't need to redo them if you want to change the symbology to a new field. You'll notice a new layer file has been added to your **Catalog** in your H Drive. This is not a shapefile. If you were to delete the **Municipality_Vulnerability** feature class from your catalog, this layer would not work. It is simply a way of saving the symbology properties, such as colors, groups, number formatting. If you were to drag this layer in, it will look exactly the same as we have it now.



Examining & Ranking NBI 2010

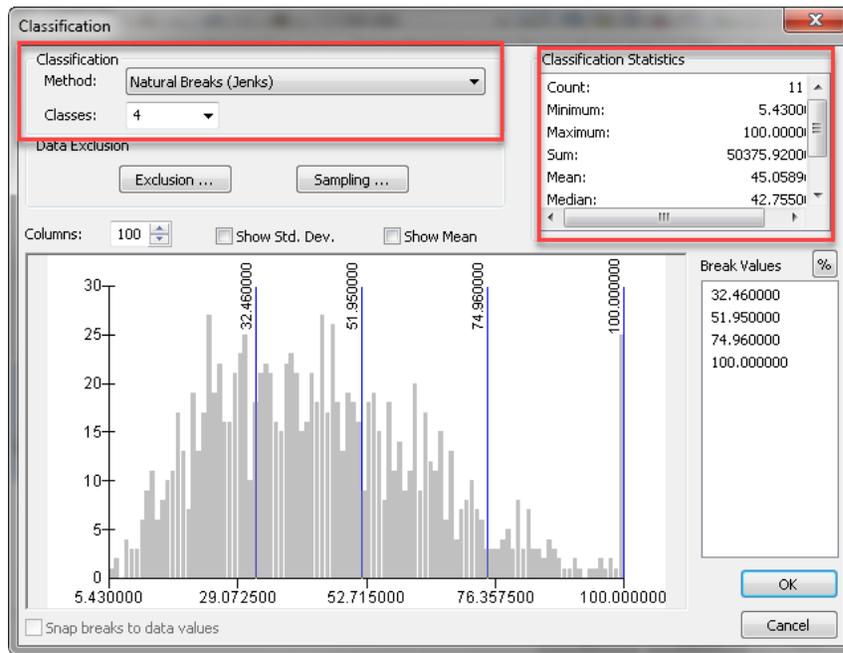
Now we need to repeat the process of assigning vulnerability scores to the other 2 indicators in this analysis.

1. Keep the Attribute Table open for **Municipality_Vulnerability** and scroll to the **NBI** field.

We learned from the **Metadata** that NBI_2010 is the **Unmet Basic Needs** Index. This index takes into account inadequate housing, housing with critical overcrowding, housing with inadequate services, and households with high levels of economic dependence.
2. Examine the contents of this field. What is the range of values?

Note: Right click on the NBI Field and **Freeze Column**. This will allow us to continuously view the NBI Rankings when we scroll to the right to see the NBI_Rank column.
3. Double click on **Municipality_Vulnerability** and then **Symbology**. Now we want to change the symbology properties so it is showing NBI. That way, we can set the **classification method and number of classes** to know how to best rank our data.
4. We want to show the NBI index, which is a Quantity. Therefore, select **Quantities** on the left hand side. Then, set the field value **NBI**. Press Apply. The map has now updated to show the NBI Index (low to high).
5. In the top right of **Symbology**, click on **Classify**. Again, we see the histogram and the class break values. Leave the **classification method** set to **Natural Breaks** and the number of **classes to 4**. We need the same number of rankings for all our indicators.

Additionally, in the top right, we see some **statistics** such as the minimum and maximum values and the mean NBI index.



6. Press OK and then OK again to exit **Symbology**.
7. Now that we have set our break values so we know how to rank/score our indicators, we can add a new field and assign the ranks.
8. Open the **attribute table**. Again, we need to go to **Attribute Table options**  to **Add Field...**, which will hold the rank/scores of the NBI Index.
9. Let's name this field **NBI_Rank** and make it a short integer. Press Ok.

Field Properties	
Alias	
Allow NULL Values	Yes
Default Value	

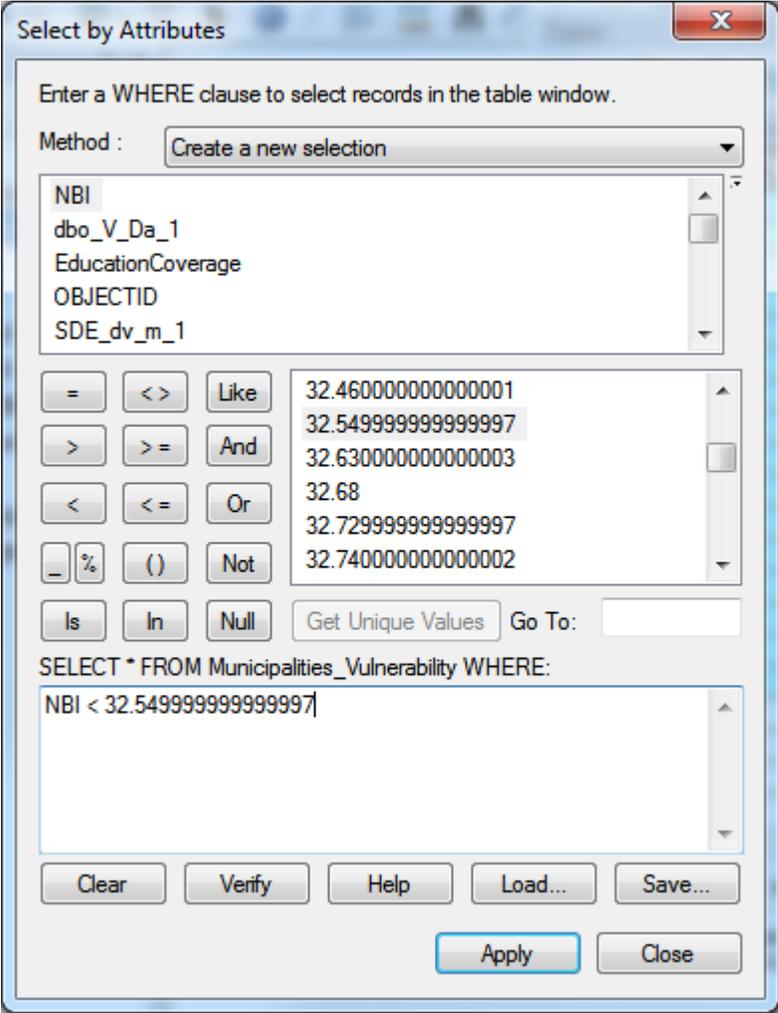
10. Once again, scroll to the right in your attribute table. You will see that a new field has been added with our new name and the values are <Null> because we have not calculated them yet.
11. **Before we start ranking our values, we need to think about what is considered high and low vulnerability.** Since we are ranking “unmet basic needs”, a score of **100** now represents the **HIGHLY Vulnerable population**. This is the exact opposite of our last indicator where a high percentage of sanitation services meant **low vulnerability**.

This means that if we use our natural breaks classification set in symbology, we need to rank it as follows:

- a. 0 – 32.46 = Low Vulnerability (0)
- b. 32.47 – 51.95 = Medium - Low Vulnerability (1)
- c. 51.96 – 74.96 = Medium – High Vulnerability (2)
- d. 74.97 - 100 = Highest Vulnerability (3)

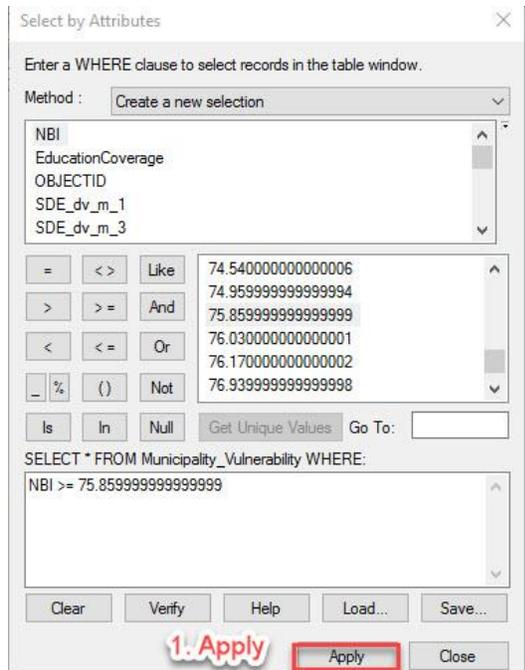
Again, it is VERY important to think about each indicator individually or else you could really mess up your vulnerability score!

- 12. Now that we have clarified how we need to rank the indicator, let's use **Select by Attribute** to select the values. *Double click* or right click → *sort ascending* on the field that holds the NBI Index to sort the values from lowest to highest.
- 13. Using the methods we learned in the last section (developing queries for each range), **select** the rows from 0 – 32.46. Click on **Attribute Table options** → **Select By Attributes**. Double click on NBI to add it to the query window. Click on the less than button (<), then click "Get Unique Values." The values for NBI will appear. Scroll down to 32.54999999999997 and double click on it. This should produce an expression that looks like: "NBI < 32.54999999999997".



- 14. You should have 350 selected records.
- 15. Now, with these values selected, right click on our new field header, **NBI_Rank**, and select **Field Calculator...**
- 16. In the field calculator, **set the value to 0** since we have selected the lowest vulnerable municipalities. Press ok and the selected cells will populate while the unselected cells remain <Null>.
- Reminder:** When you run a tool or operation with selected values, it **ONLY** runs the tool on the selected rows. This is true for ALL tools, not just field calculator.
- 17. **Clear** your selection .
- 18. Now we need to assign the Medium – Low Vulnerable Municipalities with values between 32.47 –51.95 a score of 1. Repeat the same steps, this time creating a query to select the rows from 32.47 –51.95. Produce an expression that looks like: "NBI >= 32.54999999999997 AND NBI < 52.00999999999998". You should now have 394 rows selected.

19. Again, right click on **NBI_Rank** and select **Field Calculator...** Change the value to 1 and press ok. We are now assigning a score of 1 to the medium vulnerability municipalities. Press Ok and the selected cells will populate with 1. **Clear** the last selection.
20. Next, we need to give the municipalities with Medium- high NBI Index (51.95– 74.96) a score of 2 because they are next most vulnerable. Again, create a selection to query these rows. This should produce an expression that looks like: "NBI >= 52.009999999999998 AND NBI < 75.859999999999999". There should be 278 rows selected.
21. Return to the field calculator and give these values a score of 2. Press ok.
22. Lastly, let's give the highest vulnerable municipalities with a NBI Index of 74.97 to 100 a score of 3 to represent the highest vulnerable populations. Be sure to unselect the Null Value in the Attribute Table – scroll down to the bottom, and right click on the row and click **Select/Unselect**.



23. The null value here means that this municipality has no data and therefore we can't include it in the ranking. There should be 96 selected rows.
24. Use Field Calculator to populate the last selected municipalities with a 3.
25. Now, the entire **NBI_Rank** Attribute Field should be populated with vulnerability scores from 0 -3. Make sure you did not miss any.
26. For now, we are done with the NBI_2010 field. **Clear** the selection.

Examining and Ranking Gross Education Coverage

1. Keep the Attribute Table open for **Municipality_Vulnerability** and scroll to the **EducationCoverage** field.
2. We learned from the **Metadata** that GrossEducationCoverage is the ratio of students per educational level over the population in the corresponding age range per municipality for 2013.
3. Examine the contents of this field. What is the range of values?
4. Right click on the EducationCoverage Field and **Freeze Column**. Unfreeze NBI if you no longer wish to see these values, or you can still keep it frozen.
5. Double click on **Municipality_Vulnerability** and then **Symbology**. Now we want to change the symbology properties so it is showing **EducationCoverage**. That way, we can set the **classification method and number of classes** to know how to best rank our data.
6. We want to show the **EducationCoverage**, which is a Quantity. Therefore, select **Quantities** on the left hand

side. Then, set the field value **EducationCoverage**. Press Apply. The map has now updated to show the **EducationCoverage** (low to high).

7. In the top right of **Symbology**, click on **Classify**. Again, we see the histogram and the class break values. Leave the **classification method** set to **Natural Breaks** and the number of **classes to 4**. We need the same number of rankings for all our indicators.
8. Check out the statistics in the top right too.
9. Press OK and then OK again to exit **Symbology**.
10. Now that we have set our break values so we know how to rank/score our indicators, we can add a new field and assign the ranks.
11. Open the **attribute table**. Again, we need to go to **table options** to **Add Field...**, which will hold the rank/scores of Education Coverage.
12. Let's name this field **Education_Rank** and make it a short integer. Press Ok.
13. Once again, scroll to the right in your attribute table. You will see that a new field has been added with our new name and the values are <Null> because we have not calculated them yet.
14. **Before you start ranking our values, you need to think about what is considered high and low vulnerability.** Since you are ranking "**Education Coverage**", a **high** value now represents the Lowest Vulnerability. And a low ratio of coverage represents Highest Vulnerability. Using **Natural Breaks**, you need to rank it as follows:
 - a. 131.08 – 250 = Low Vulnerability (0)
 - b. 103.42 - 131.07 = Medium - Low Vulnerability (1)
 - c. 78.54 - 103.41= Medium – High Vulnerability (2)
 - d. 23.82 - 78.54= Highest Vulnerability (3)

Again, it is VERY important to think about each indicator individually or else you could really mess up your vulnerability score!

15. Now that we have clarified how we need to rank the indicator, let's use field calculator to set the scores. *Double click* or right click → *sort ascending* on the field that holds the EducationCoverage to sort the values from lowest to highest. If you double click again, it will sort from **highest to lowest**, which might be easier for this field.
16. Using the methods we learned in the last section (developing queries for each range), **select** the rows from 131.08 – 250. This should produce an expression that looks like: "EducationCoverage > 131.06999999999999". You should have 97 selected records. Remember to use the buttons.
17. Now, with these values selected, right click on our new field header, **Education_Rank**, and select **Field Calculator...**
18. In the field calculator, **set the value to 0** since we have selected the lowest vulnerable municipalities (aka highest coverage of education). Press ok and the selected cells will populate while the unselected cells remain <Null>.
19. **Clear** your selection . Now we need to assign the Medium – Low Vulnerable Municipalities with values between 103.42 - 131.07 a score of 1. Repeat the same steps, this time selecting the rows from 103.42 - 131.07. This should produce an expression that looks like "EducationCoverage <= 131.06999999999999 AND EducationCoverage > 103.41".
You should now have 377 rows selected.
20. Again, right click on **Education_Rank** and select **Field Calculator...** Change the value to 1 and press ok. We are now assigning a score of 1 to the medium vulnerability municipalities. Press Ok and the selected cells will populate with 1. **Clear** the last selection.

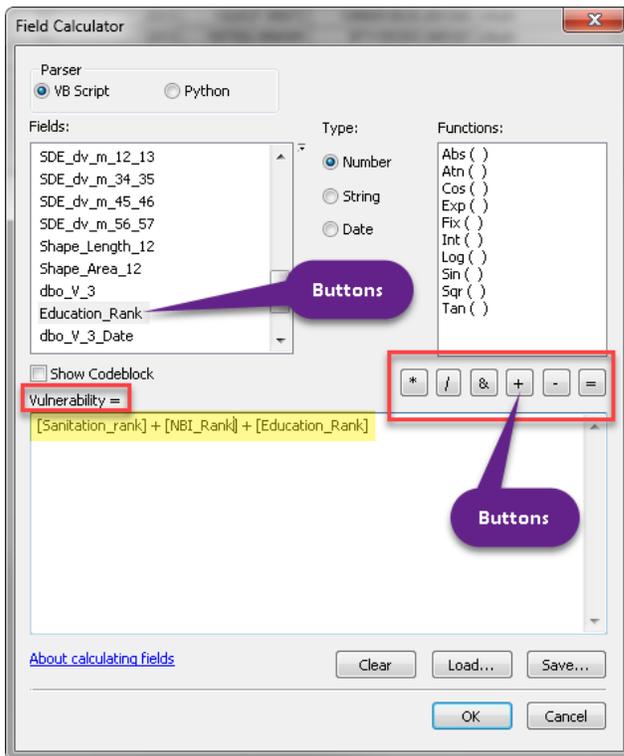
21. Next, we need to give the municipalities with Medium- high Education Coverage Index (78.55 - 103.41) a score of 2 because they are next most vulnerable. Again, select by attribute query to select all the rows with values 78.55 - 103.41. There should be 467 rows selected.
22. Return to the field calculator and give these values a score of 2. Press ok.
23. Lastly, let's give the highest vulnerable municipalities with an Education Coverage Index of 23.82 - 78.54 a score of 3 to represent the highest vulnerable populations. Do not select the Null Value – that means that this municipality has no data and therefore we can't include it in the ranking. There should be 176 selected rows.
24. Use Field Calculator to populate the last selected municipalities with a 3.
25. Now, the entire **Education_Rank** Attribute Field should be populated with vulnerability scores from 0 -3. Make sure you did not miss any (excluding the Null Value).
- 26. Clear your selection.**

Using Field Calculator to Calculate Vulnerability

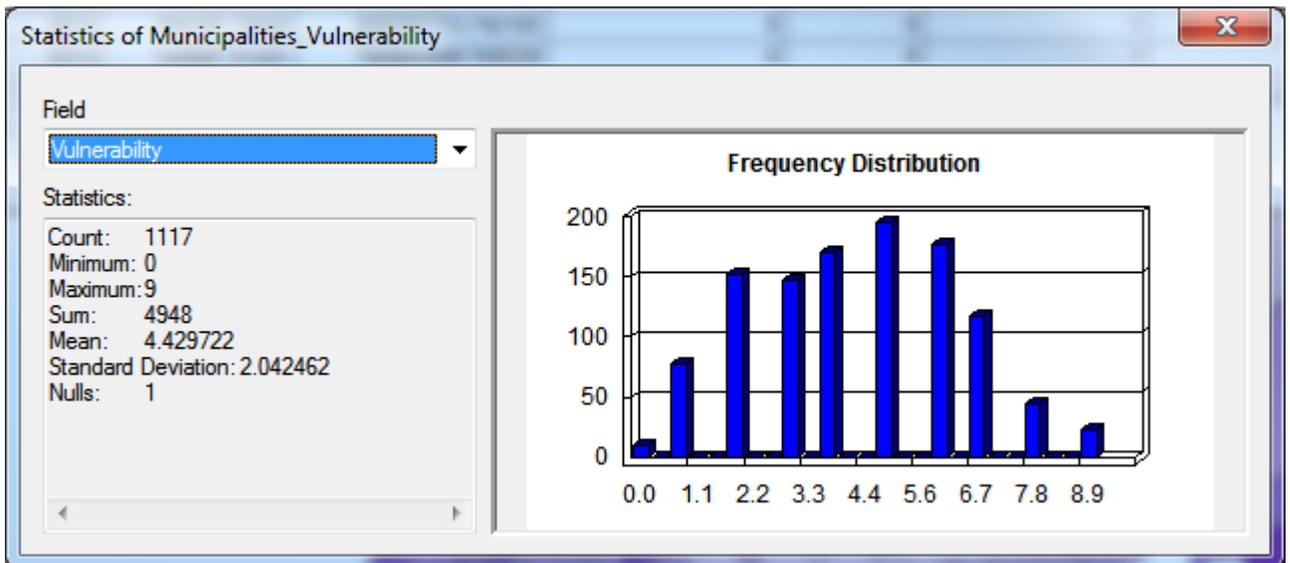
Now that we have completed ranking all 3 indicators from lowest vulnerable municipalities (0) to highest vulnerable populations (3), we can combine them using field calculator get an overall vulnerability score from 0 – 9.

1. Once again, the first thing we need to do is **Add a Field** to this attribute table to hold our vulnerability scores. Open the Attribute Table for **Municipality_Vulnerability**.
2. In the Table Options, select **Add a Field**...just like we've been doing for the other 3 ranking fields.
3. Let's name this new field **Vulnerability** and leave it as a **Short Integer**.
4. Once again, scroll all the way to the right of the **Municipality_Vulnerability attribute table** and you will see your new attribute field. Right click on the heading **Vulnerability → Field Calculator...**
 What should our calculation expression be to calculate vulnerability?
5. We simply need to add together the 3 fields that hold are ranking scores. This is why it was so important to give these fields identifiable names.
6. Double click on **Sanitation_Rank**, then click the **+ sign**, then double click on **NBI_Rank** then click the **+ sign**, then double click on **Education Rank**. Make sure your calculation looks like the screen shot below.

Note: It is important to use the “**buttons**” rather than typing in the expression. There are a lot of potential errors with the syntax if you try to type it the equation box directly, just like in the structured queries we used to select our ranks. The calculator is very particular about spaces and using the buttons helps minimize errors.



7. Press OK. Now our **Vulnerability** Field calculates with a score from 0 – 9 with 0 being lowest possible vulnerability (all 3 indicators equal 0) and 9 being highest possible vulnerability (all 3 indicators scored 3).
8. **Double Click** on Vulnerability field heading to sort it from low to high. Scroll down to get an idea of how the municipalities scored. You'll notice the field that had Nulls in the rankings remained Null in the final score.
9. *Right click* on **Vulnerability** → **Statistics...** Here, we can see the distribution of scores by looking at the histogram. We can see the min and max score and the mean vulnerability score.

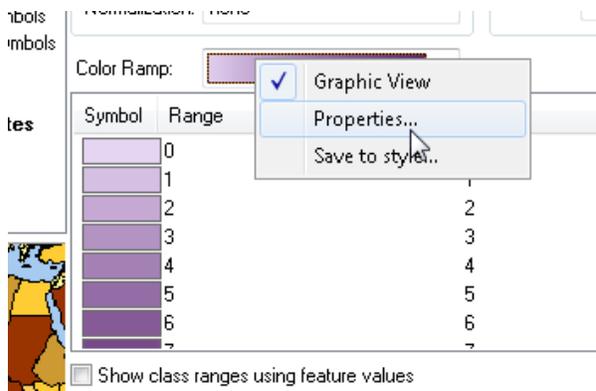


Using Symbolology to Represent Vulnerability

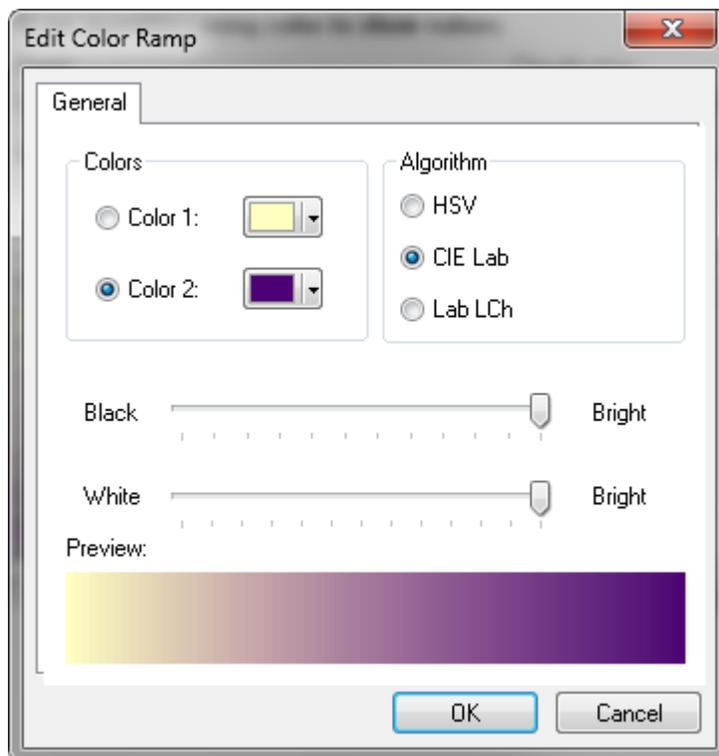
1. Close the statistics and the attribute table. Now, let's use **symbolology** to show our Social Inequality Vulnerability Scores. Open the **Symbolology** tab of this layer.
2. Are the scores considered **Categories** or **Quantities**? In this case, you could really choose either. However, because Quantities has better gradual color options from low to high, let's choose **quantities**. **Categories**

has more color options but not within a sequential order because they are meant to be distinct colors.

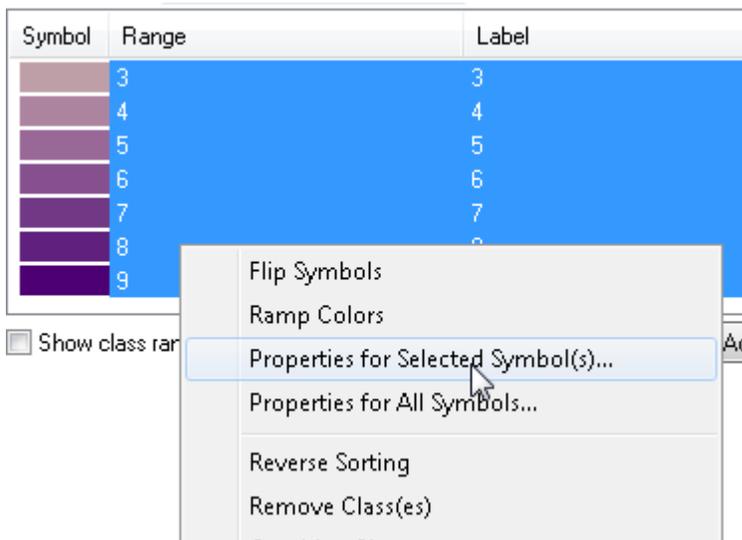
- Now, we'll need to select the field "Vulnerability" and change the **classes to 5**.
- Then, choose a color ramp that represents the colors well. Remember, 9 is high vulnerability...something like green wouldn't be a good color choice, where a light to dark red might make more sense.
- Explore the color options.** You can press Apply to see how it changes the map. Moreover, if you don't love any of the given color choices, you can **right click** directly on the color ramp bar and go to **properties**.



- This allows you to **adjust** the starting and ending colors for the ramp. It gives you a lot more options to make a color ramp that you feel accurately represents vulnerability. I'm going to make my dark purple just a bit darker and the light purple more of a yellow/tan, so there is more distinction between the color choices.



- When you decide on a ramp color you like, press OK and Apply. Don't exit symbology quite yet.
- Because the municipalities are so small, the municipality **outline** colors kind of get in the way of viewing the results. Let's remove the outline colors. Click on your lightest color, then hold down the shift button, then press on your darkest color. This should highlight all the colors. Then right click and select **Properties for Selected Symbol(s)**.



9. Here, you can change the outline color or take away the outline completely by making the width = 0. Press OK then Apply. Now the outlines disappear and it is much easier to see the data.

Map Layout & Cartography Tips and Tricks

1. Now it's time to style our map. Switch to **Layout View**.
2. First, pull the data frame out so it expands to the dashed lines (these are printing margins) and almost covers the whole page. Then, zoom into your data so it fills the page. Get used to taking advantage of all the space you can and really filling out your map.
3. Let's bring in some more data to give a bit more context to the map. In the SocialInequality_Vulnerability_Colombia folder, you should see another geodatabase called **Colombia_Data.gbd**. Expand that and bring in **Departmentos_2011**, **Capital_Bogota**, and **AllCountries_Deatailed**.
4. Pull **AllCountries** to the bottom and make it grey. In order to drag layers around to reorder, you may need to switch to "List by Drawing Order" if you were in "List by Source".
5. Bring **Departmentos_2011** to the top and make it see through, with black outlines at width 1.2pt. Now we can see which municipalities are within each Departments.
6. Pull **Capital_Bogota** all the way to the top. Make it a black star.

Legends

7. Now, let's bring in our map elements. First, let's add a **Legend**. Insert → Legend. Include only *Municipality_Vulnerability* and *Departmentos_2011* in the legend items. Then breeze through the set up wizard. We'll go back and edit it.
8. Once it's added, you'll see that it is using the names from the *table of contents*. They need clearer names. Click on *Departmentos_2011* and press F2. This is a quick way to change the name. Call it **Departments**.
9. Press on *Municipality_Vulnerability* and press F2. Change this to **Vulnerability Score**. Then press the heading, **Vulnerability**, and press F2. Change it to say **By Municipality**. It's important that we know what administrative boundary this is. Then press the 0 and hit F2. Change 0 to "**Low Vulnerability**". Do the same for 9, changing it to **High Vulnerability**. We need to know what 0 and 9 means, otherwise we don't know which is which.
10. The legend has already improved. Double click on the legend to get to its properties. Under **General**, uncheck **Show Legend**. We don't need it to say Legend. Still under general, press the **down arrow** to move *Departments* **under** *Vulnerability Score*. Click **Apply**. Much better.
11. Now, go to the **Items** Tab. You can pick different "**styles**" for each layer. Highlight *Vulnerability Score* and

press **Style** at the bottom. Select the top left style – horizontal bar. Or experiment with a style you like. Hit Apply. You'll see it removed the Layer Name – Vulnerability Score. To get it back, double click on Vulnerability Score → General Tab. Recheck Layer Name. Then ok and Apply. Now it's back. Pick a style that fits well with your layout.

12. To change the font and size (because why just accept defaults), highlight the layer you want to change. You can highlight both at the same time by using shift. Then, under the **font** drop down, select "Apply to all Labels" and pick a font and size you like. I'm going to pick **Bell MT and 12pt font**.
13. Now, let's make **Vulnerability Score** bigger than **By Municipality**. Make sure Vulnerability score is highlighted. Then change the font drop down to "**Apply to Layer Name**". Then change the font size to 15 and hit apply.
14. In the Legend **Layout** Tab, you can adjust the shape of the polygons. Under Area, switch it to an ellipse or rounded rectangle. In the **Frame** Tab, you can add a background. If you choose to add a background, make sure to make the X and Y 5pts to give it some space.
15. If you want to make the background blue to represent the oceans, here is a quick and easy way to do that. Double click on the *data frame*, aka **Layers**. Go to the **Frame** tab. Change the **Background** to a blue and press OK. Voila, water.

Draw Tool Bar

16. The Draw Toolbar is a really easy way to add text, labels and graphics. Go up to Customize → Toolbars → Draw. The toolbar will be added.
17. To Add text, such as a **title**, press the **A** and then click in the map where you would like to add text. **Add a title and press enter**. To quickly change the font, color, or size, click the text and use the drop downs in the tool bar to change the properties.
18. You can use this text box to add other text, such as Country Labels, Capital Labels Water Labels, etc. Sometimes it's nice to have annotation explaining which variables were used to determine vulnerability.
19. Add in your cartographer information and Data Sources. This is from Instituto Geografico "Agustin Codazzi" (IGAC). It tells you that in the metadata item description.

Additional Map Elements & Exporting your Map

20. Add in your **scalebar**. Double click on the scalebar to edit the properties, such as changing miles to kilometers or changing the number of divisions to 1 and subdivisions to 2.
21. You can use the draw tool bar to change the font of the scalebar. Just click the scalebar and use the font drop down.
22. Insert a north arrow.
23. Add in your cartographer information and Data Sources. This is from [National Administrative Department of Statistics \(DANE\)](#). It tells you that in the metadata item description.

Check out your final map of Social Inequality!

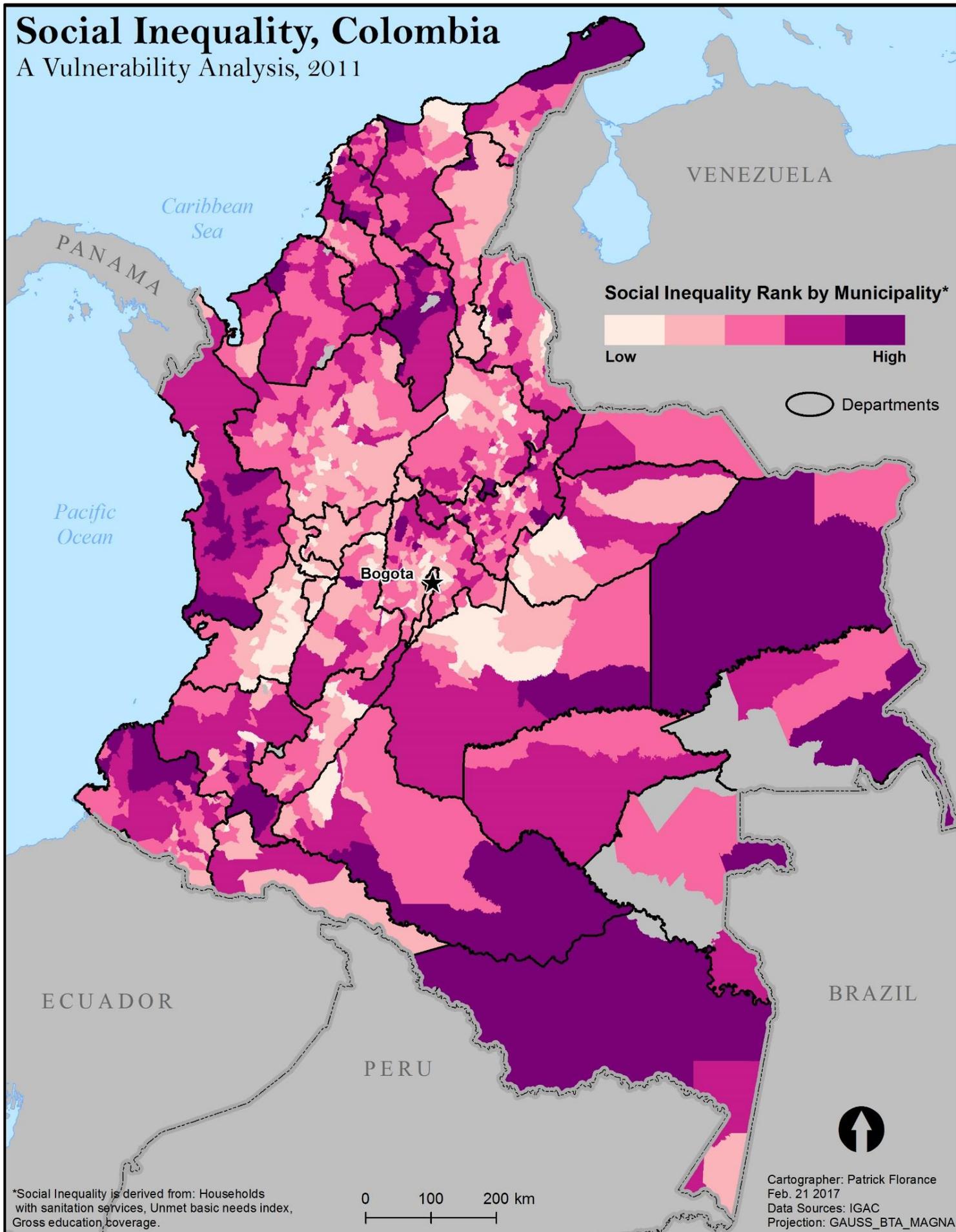
Further Research & Analysis

You'll notice that there are several other indicators of social inequality within the geodatabases, including **Housing Quality, Electricity Coverage, Aqueduct Coverage**. This analysis could be expanded to include many additional variables and is certainly not limited to only 3.

Think about what other variables you might want to include for a more in depth analysis of social inequality.

Social Inequality, Colombia

A Vulnerability Analysis, 2011



*Social Inequality is derived from: Households with sanitation services, Unmet basic needs index, Gross education coverage.