Exercise 2. Building a Base Dataset of the San Marcos Basin

GIS in Water Resources
Fall 2013

Prepared by David R. Maidment

Goals of the Exercise

This exercise is intended for you to build a base data set of geographic information for a watershed using the San Marcos Basin in South Texas as an example. The base dataset comprises watershed boundaries and streams from the National Hydrography Dataset Plus (NHDPlus) and soils from the SSURGO soils database. A geodatabase is created to hold all these primary data layers. In addition, you will create a point Feature Class of stream gage sites by inputting latitude and longitude values for the gages in an Excel table that is added to ArcMap and the geodatabase. You also compare the locations of the San Marcos basin surface boundaries, and the Edwards aquifer subsurface boundaries.

Computer and Data Requirements

To complete this exercise, you'll need to run ArcGIS 10.2 from a PC. You will download map packages of hydrologic and soils information to do this exercise from ArcGIS Online.   If you have trouble accessing these packages, there is a backup at
http://www.caee.utexas.edu/prof/maidment/giswr2013/Ex2/Ex2Data.zip

Procedure for the Assignment

Getting Started
We'll begin by getting the input data for Water Resource Region 12, and creating a new, empty geodatabase into which you'll put data for the San Marcos basin, which is a small drainage area within this region.

Login to ArcGIS Online and Search for **GISWR2013Ex2**

And …. Aaaah… Disaster, you can't find anything! Ouch.
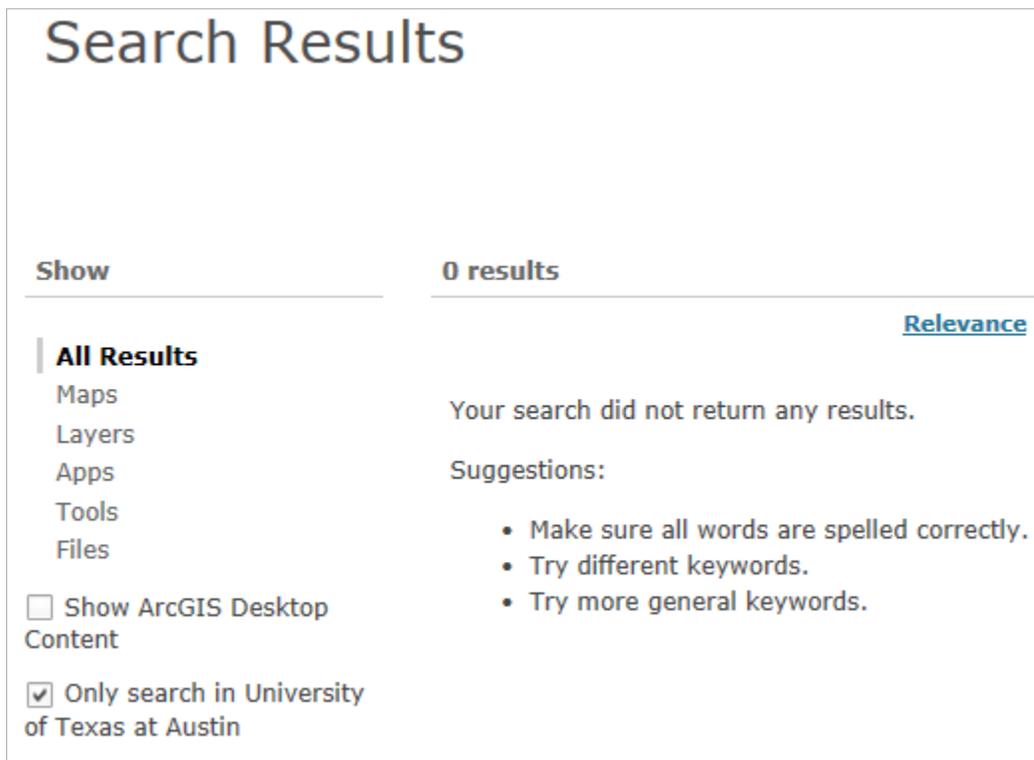
Notice in the lower left of the response screen that **Show ArcGIS Desktop Content** is clicked off and **Only Search in University of Texas at Austin** is clicked on. Let's reverse that, and Success!! We have found our Map Package. Isn't technology wonderful (and wonderfully obscure at times!)
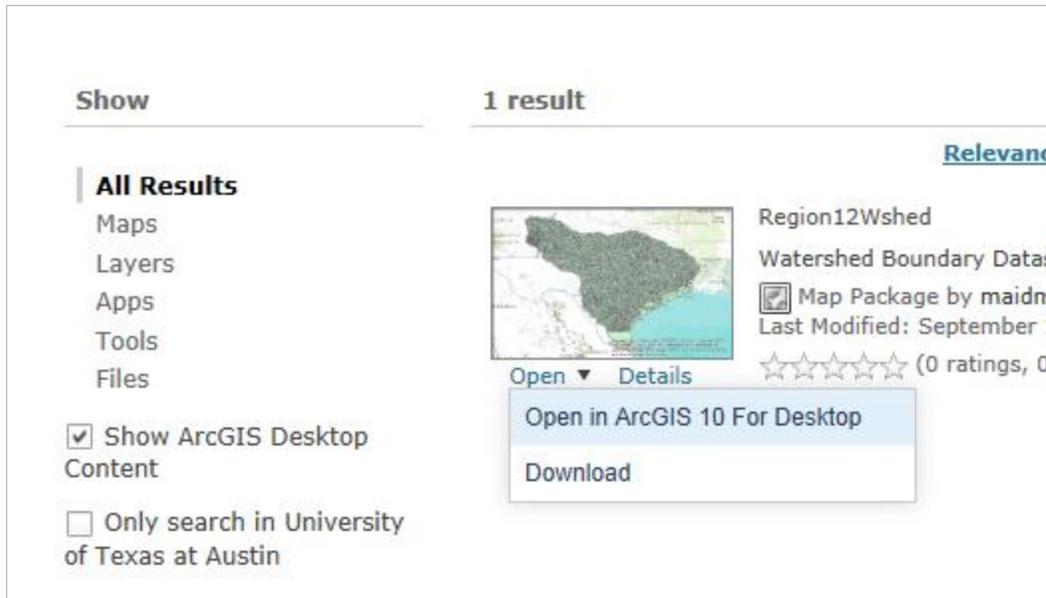
If you don't get this package successfully from ArcGIS Online, pick it up at:
http://www.caee.utexas.edu/prof/maidment/giswr2013/Ex2/Ex2Data.zip

Once you have located the map package **Region12Wshed** in ArcGISOnline using the tag **GISWR2013Ex2** open it in ArcGIS and you should see a display like this
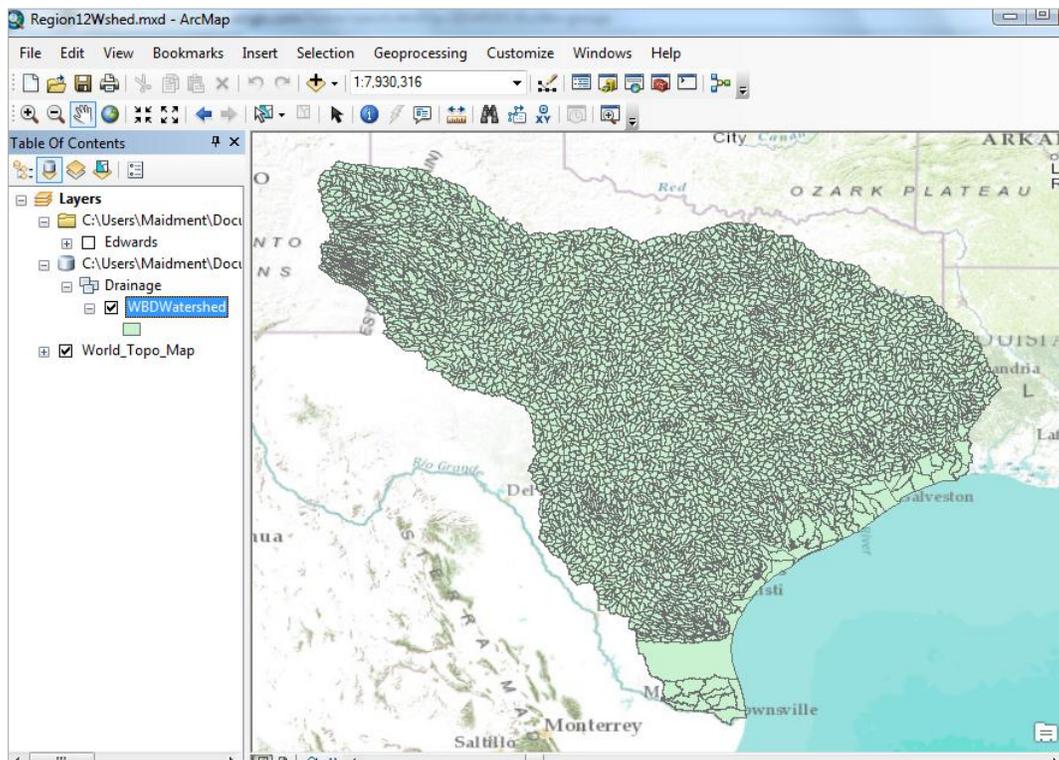
This map package contains a feature class **WBDWatersheds** of the HUC-12 Subwatersheds in Water Resource Region 12 that we'll focus on first in this exercise. There is also a feature class for the **Edwards** Aquifer that we'll use near the end of the exercise.

From ArcMap, open **ArcCatalog**, navigate through **Folder Connections** to a place where you want to have a workspace,
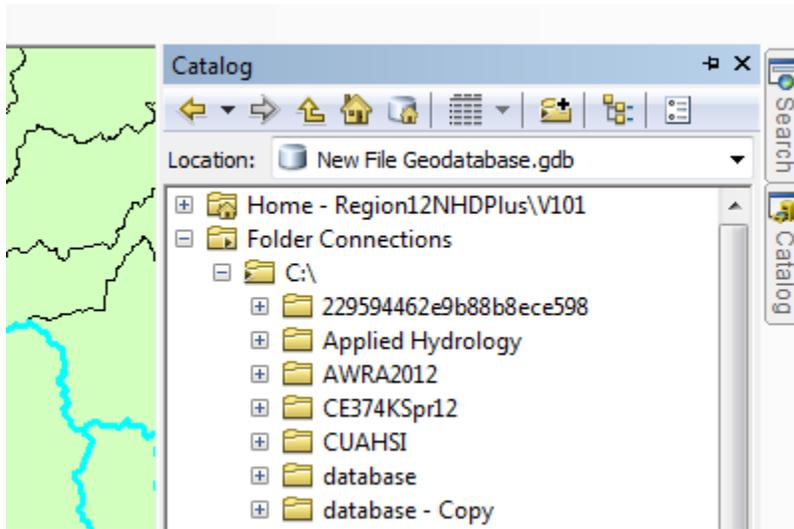
and create a new **File Geodatabase**:
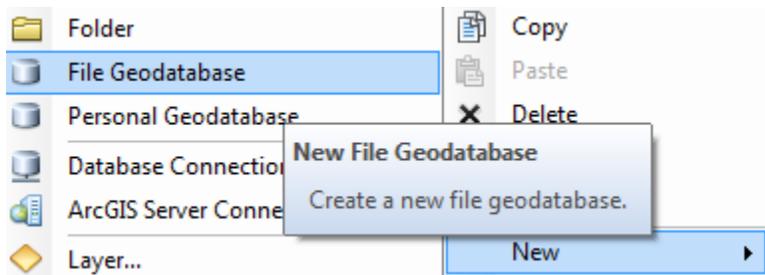
And call this **SanMarcos.**gdb. Within this, create a new **Feature Dataset**
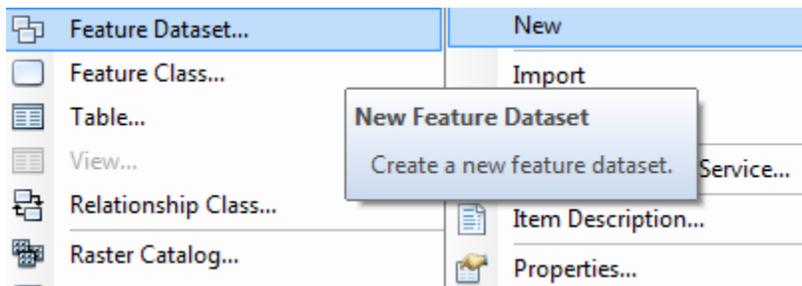
and call it **BaseData**
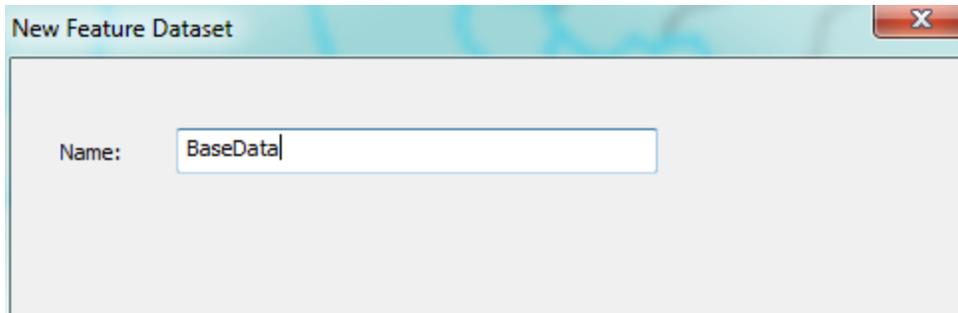
choose a Geographic Coordinate System
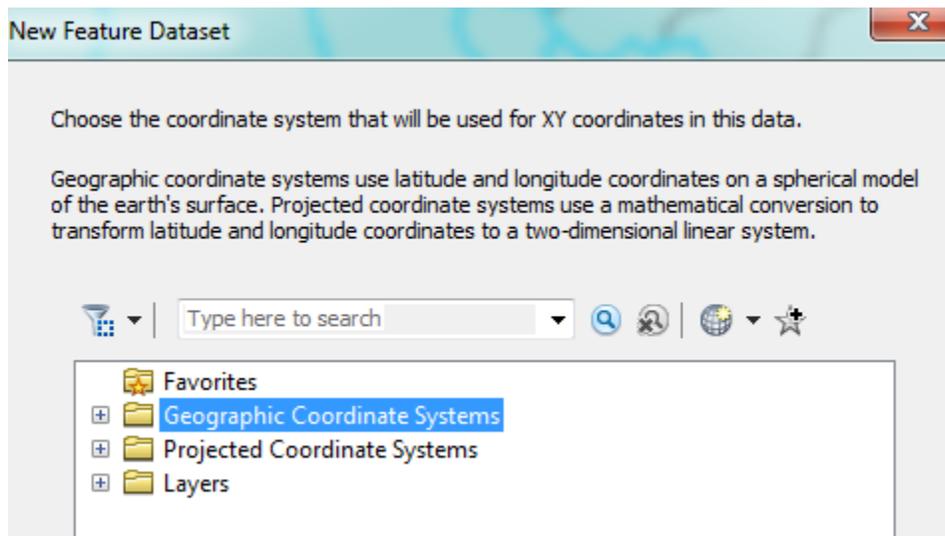
within North America,
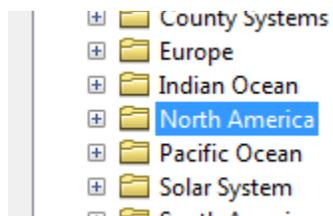
select the NAD83 coordinate system
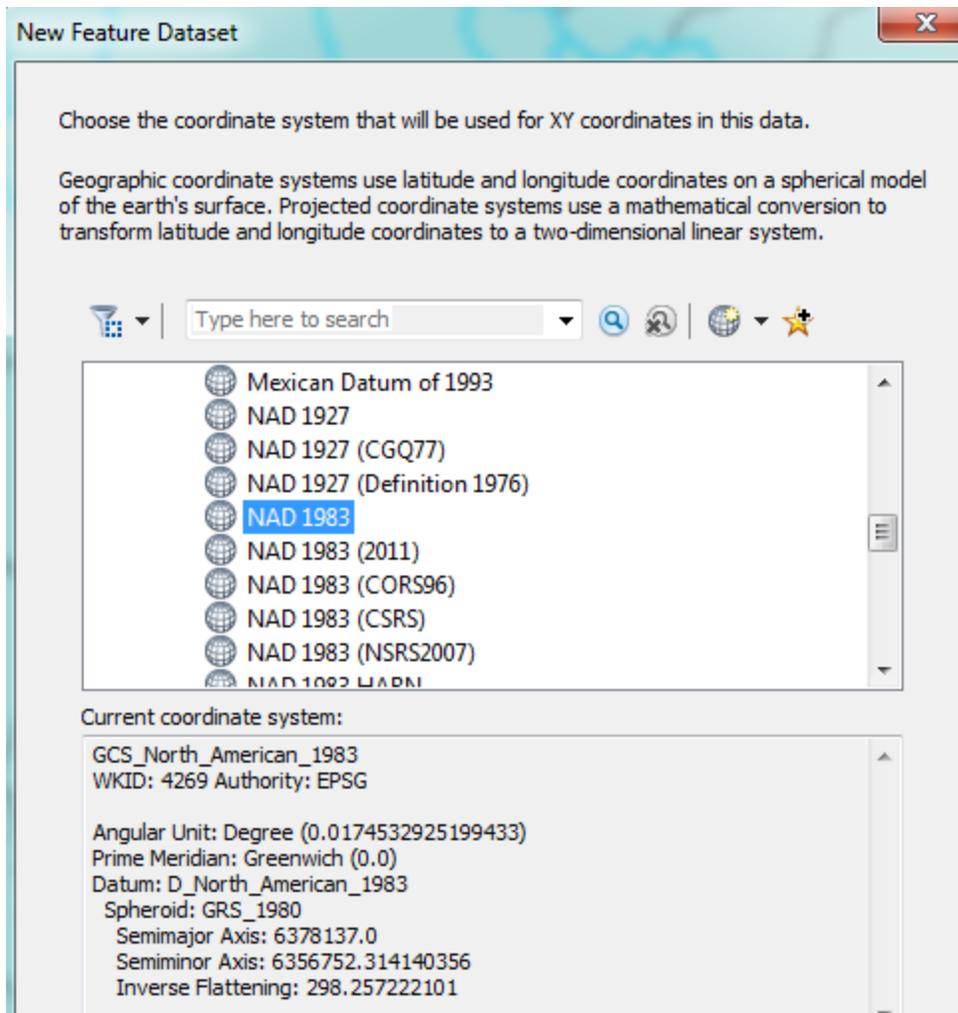
Hit **Next**, and **Next** again to bypass having a Vertical Coordinate system, and then **Finish** to complete creating the Feature Dataset, leaving the tolerance information at the default values.
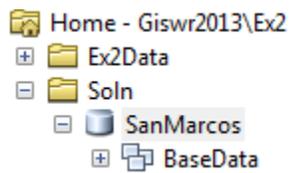
This **BaseData** feature dataset within the **SanMarcos** geodatabase will hold the data that you create for the San Marcos Basin.

Selecting the Watersheds in the San Marcos Basin

Let's zoom into the San Marcos basin.

We want all the HUC12 subwatersheds that lie within the San Marcos subbasin, which has a HUC8 value of [HUC_8 = 12100203].

Open the Attribute Table of the Watershed Boundary Dataset (**WDBWatershed**)

At the top left corner of the **Table**, click on the **Select by Attributes** tool
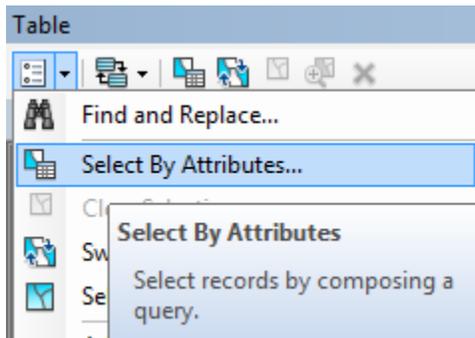
Click on "HUC8", "=", **Get Unique Values** and then type 12100203 in the **Go To** box, double click on the resulting '12100203' to form the expression

"HUC_8" = '12100203'

In the selection window. Be careful about how you do this since the form of the expression is important. Click **Apply** and **Close** the Select by Attributes window**.**
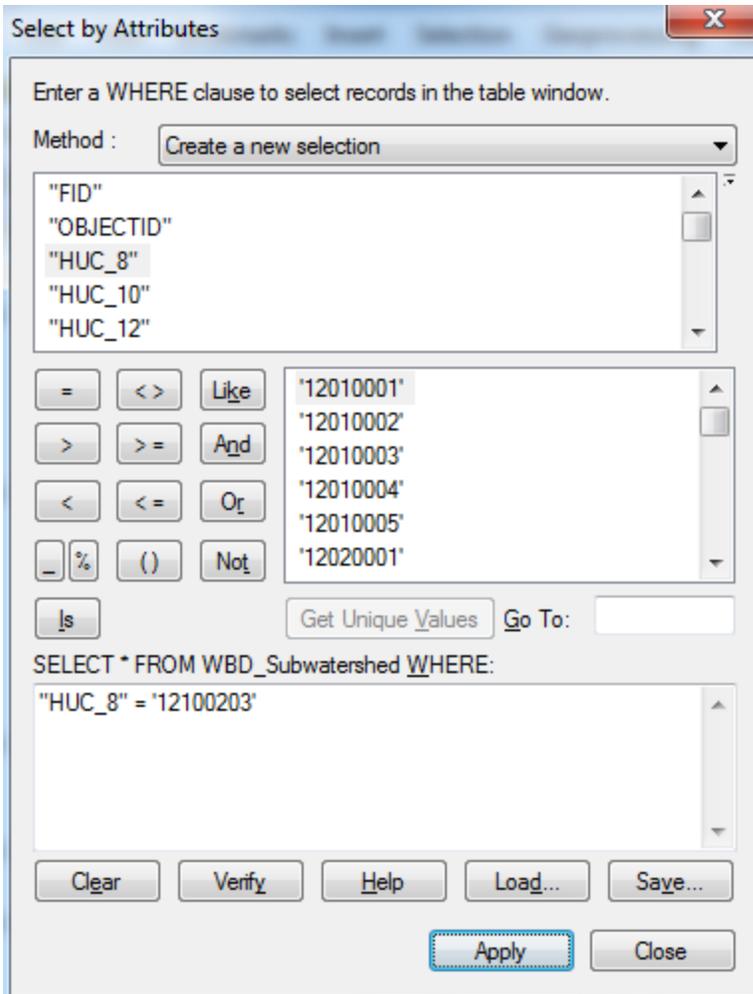
You'll see that this selects 32 of the HUC-12 Subwatersheds that lie within the San Marcos basin (one HUC-8 Subbasin). If you hit the **Selected** button at the bottom of the Table, you'll see the selected records, and also their highlighted images in the map.
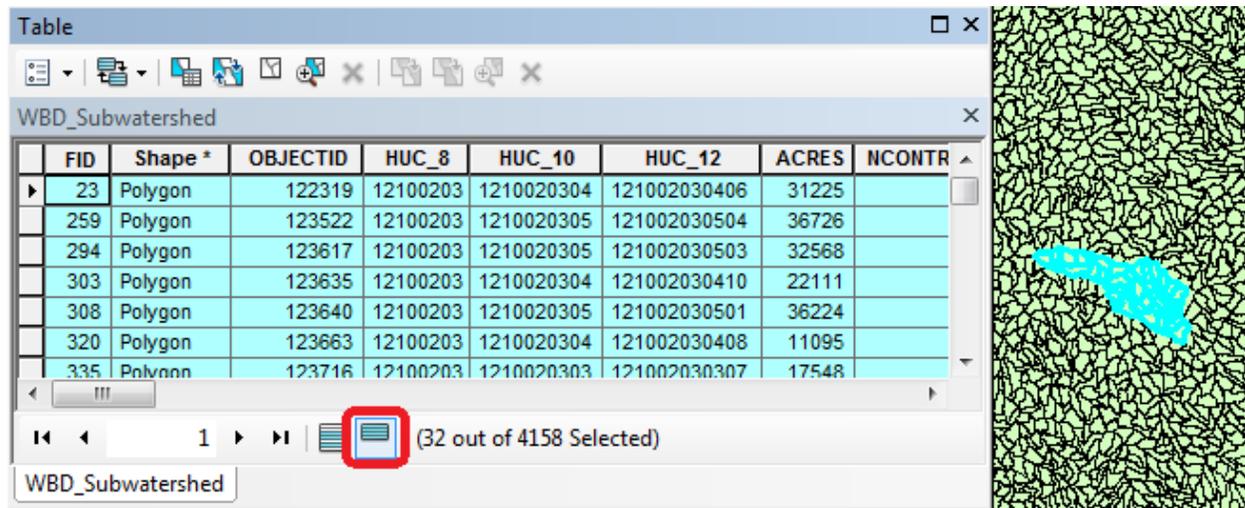
Use **Selection/Zoom to Selected Features**:
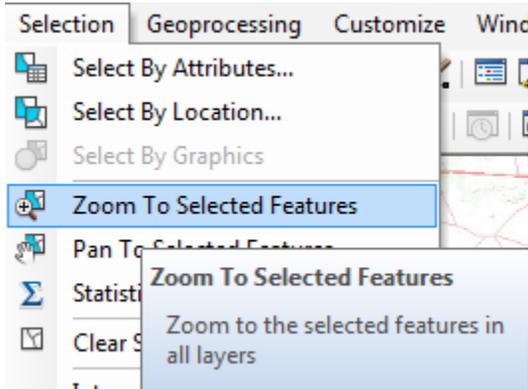
Close the **WBD_Subwatershed** table to get it out of the way. Right Click on the watersheds layer (**WBD_Subwatershed**) and select **Data/Export Data** to produce a new Feature Class.
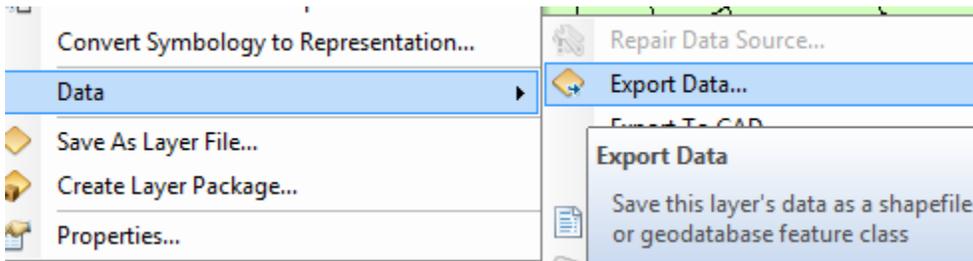
Be sure to navigate to where you established the SanMarcos geodatabase earlier and don't just accept the default geodatabase presented to you, which is somewhere deep in the file system that you may never find again! Browse inside the SanMarcos geodatabase you created to the **BaseData** Feature dataset and name this new feature class as **Watershed** and click **Save**. (Note that you may have to change the Save as Type to File and Personal Geodatabase feature classes).
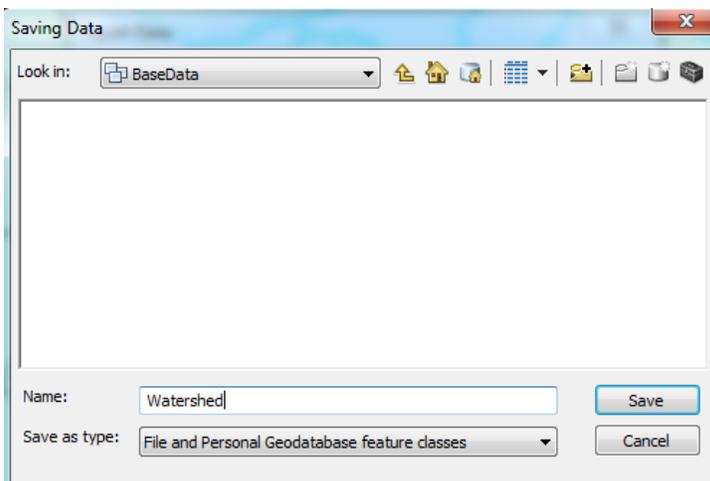
At the next screen click OK
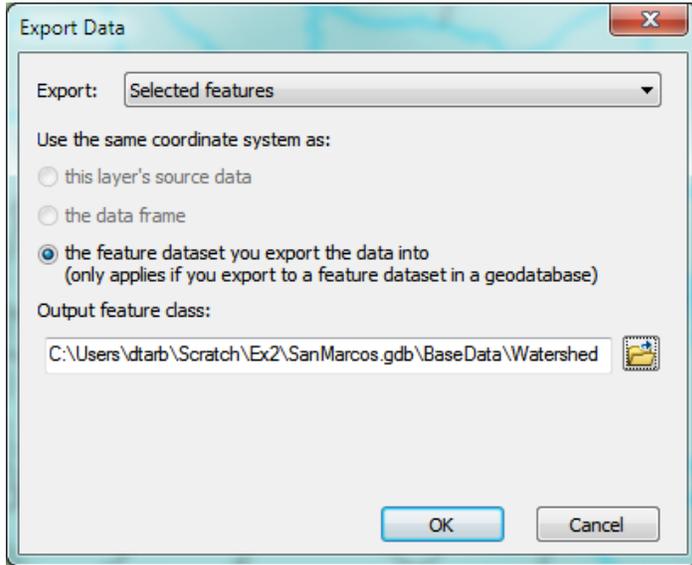
You will be prompted to whether add this theme to the Map, click **Yes**. In ArcMap, Use **Selection/Clear Selected Features** to clear the selection you just made.
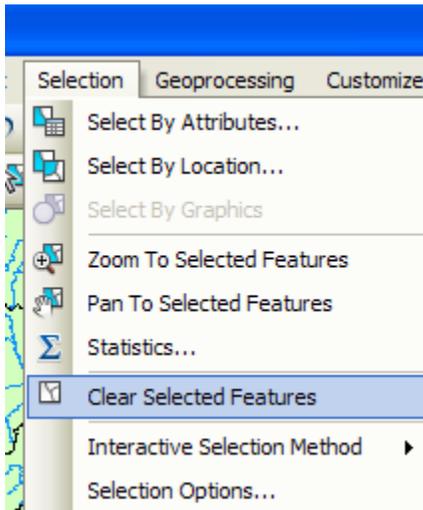
And then Zoom to Layer to focus in on your selected Watersheds. You can click off the little check mark by the **WBDWatershed** layer and **Basemap** so that you just see the selected watersheds displayed. Change their symbology so that they are colored green if necessary. Watersheds are always green!
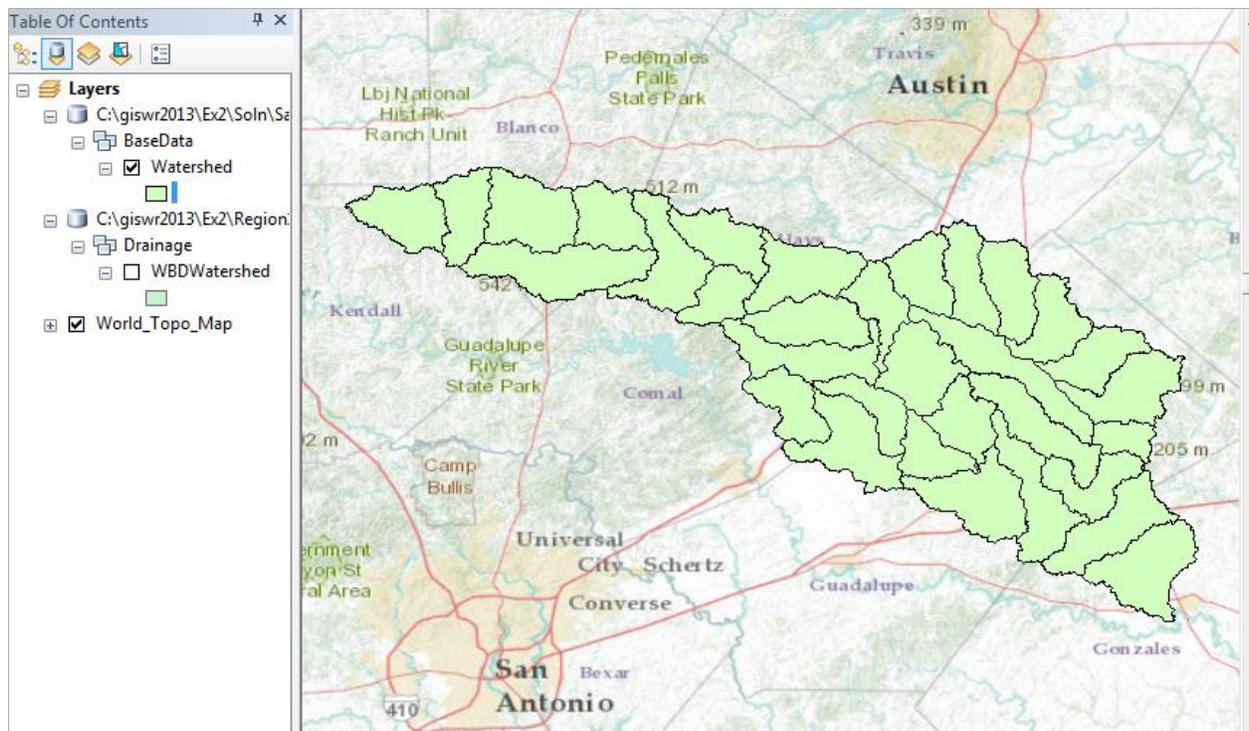
Lets make our basin a bit more interesting.   Right click on the Watershed feature class, and
select Properties/Symbology. Select Categories Unique values and use HUC-10 as the Value
Field, hit Add All Values to give each HUC-10 watershed a different color. Hit **Apply** and **OK**
to get this color scheme applied to the map.
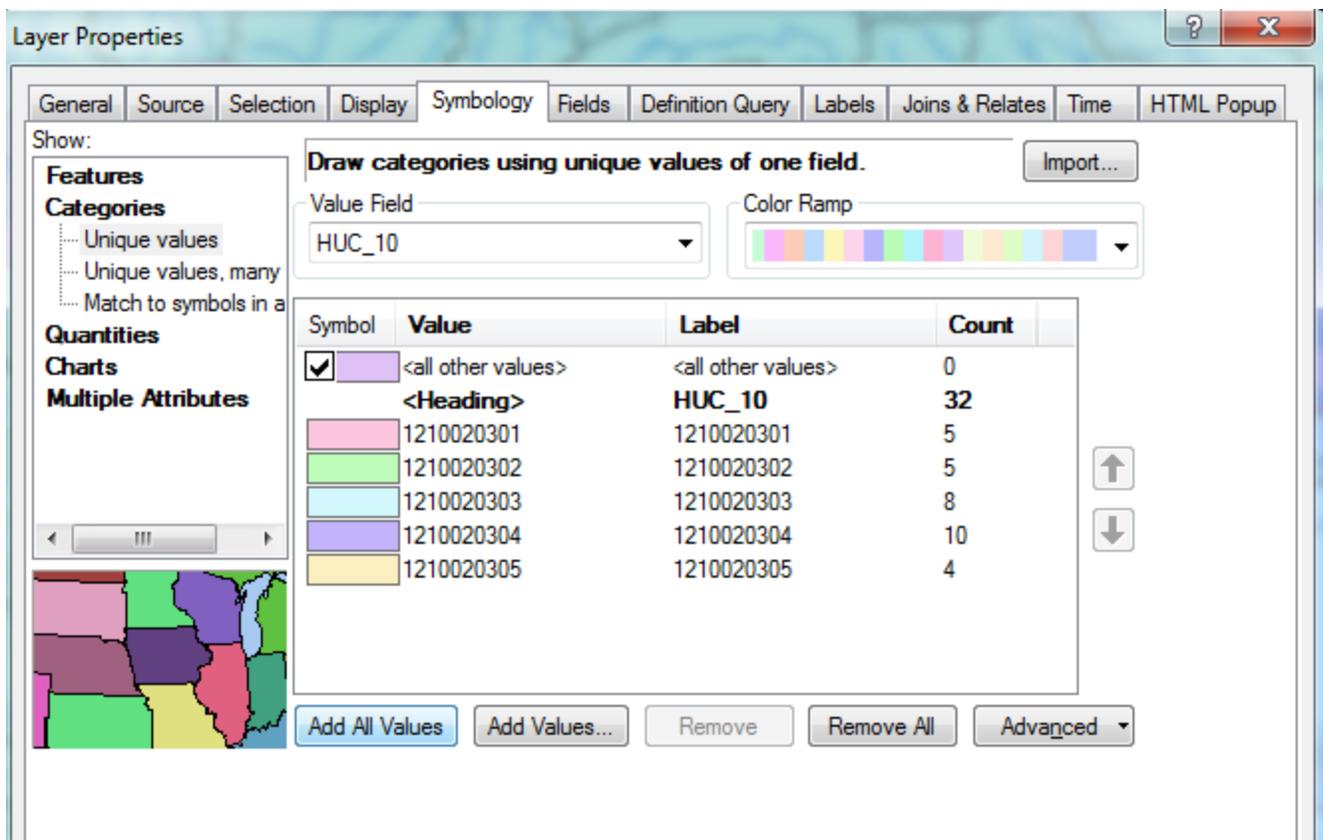
Lets focus on the Watersheds feature class by turning off the display of the other feature classes using the check box in the Table of Contents.
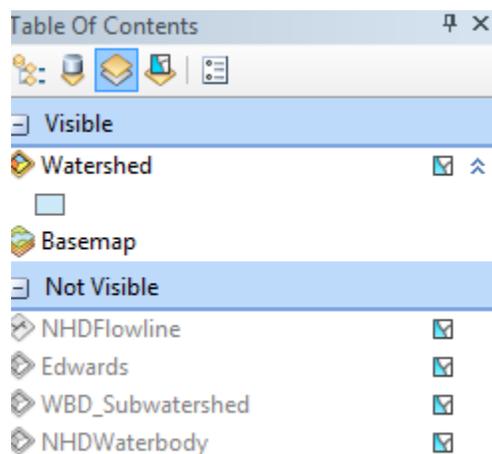
And you'll get this nicely colored map of the watersheds and subwatersheds of the San Marcos basin.
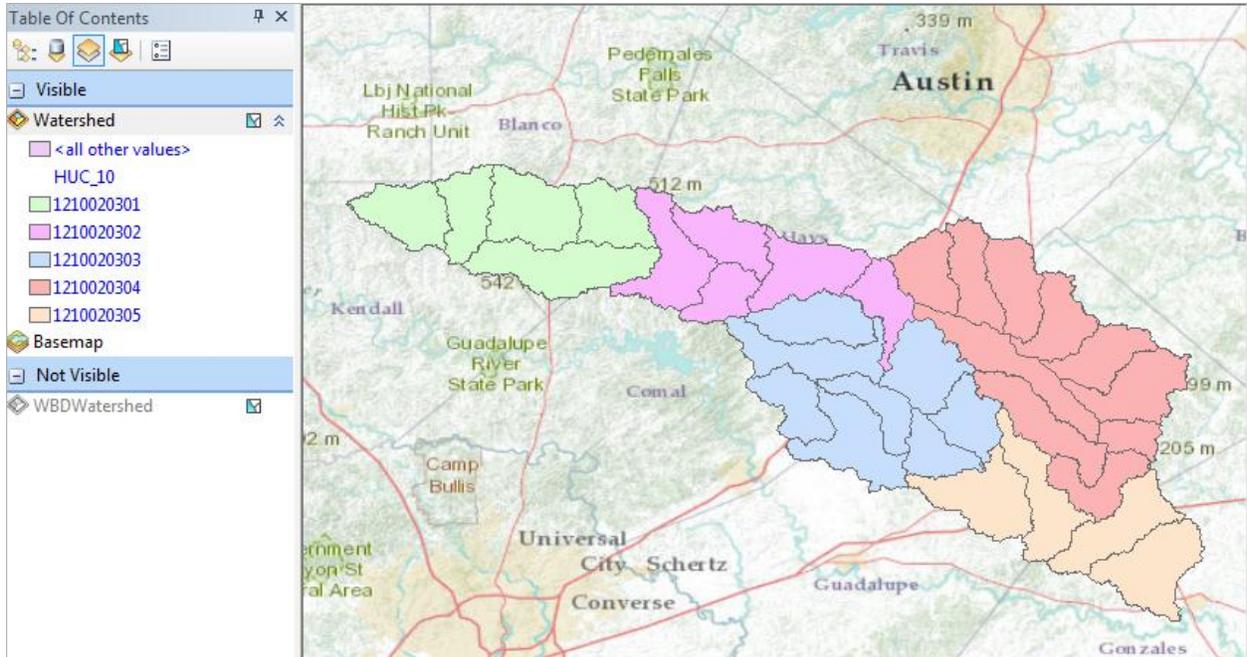
Notice that the 32 HUC-12 *subwatersheds* have been grouped into five *watersheds* within the San Marcos *subbasin* (I am here using the Watershed Boundary Set nomenclature to refer to the drainage area hierarchy in its formal sense).

Select the **Identify** tool, go up near the top of the San Marcos Basin, and click on one of the HUC-12 subwatersheds. You'll see its attributes pop up. Notice the hierarchy of numbers for the HUC_8, HUC_10, and HUC_12 attributes.
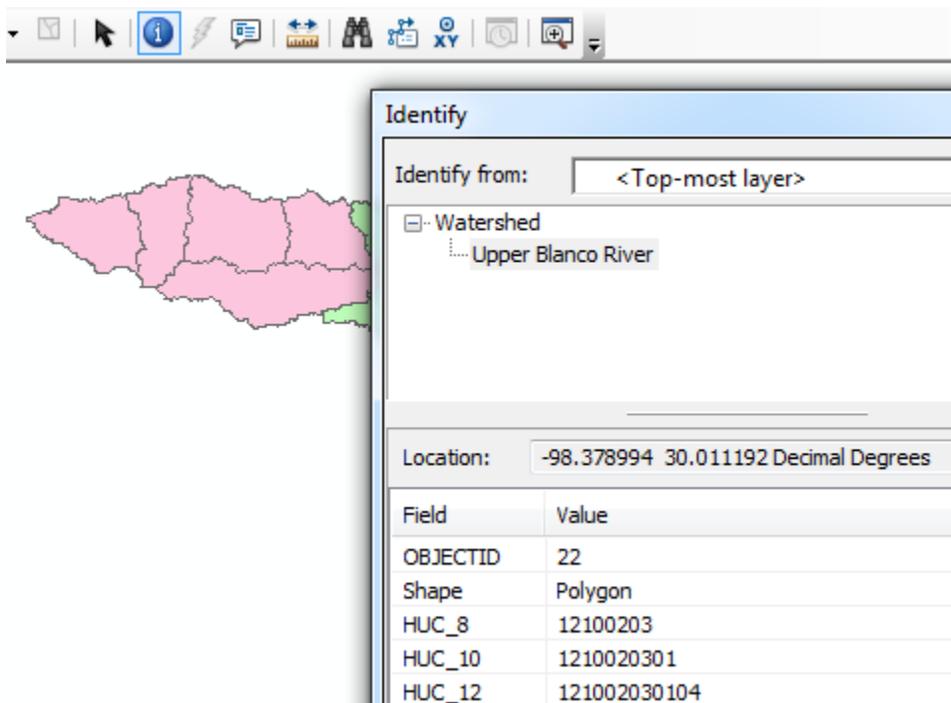
Use **File/Save As** to save your map file as **Ex2.mxd** with the new information that you've created (and to keep it distinct from the Map Document Region12NHDPlus.mxd opened from ArcGIS Online).

Where is My Stuff?

Right click on **Watershed** and select **Properties** and select the **Source** tab. Notice that this Feature Class you created is in the **BaseData** Feature Dataset in the **SanMarcos.gdb** Geodatabase in the location where you created it.   It comprises Simple Features (no topology), that are Polygons (have X, Y values) but have no Z values or M values which deal with elevation and measure, respectively, that we'll encounter in a later exercises. It has a **Geographic Coordinate System** using the **North American 1983** datum. You'll learn more about these shortly also.
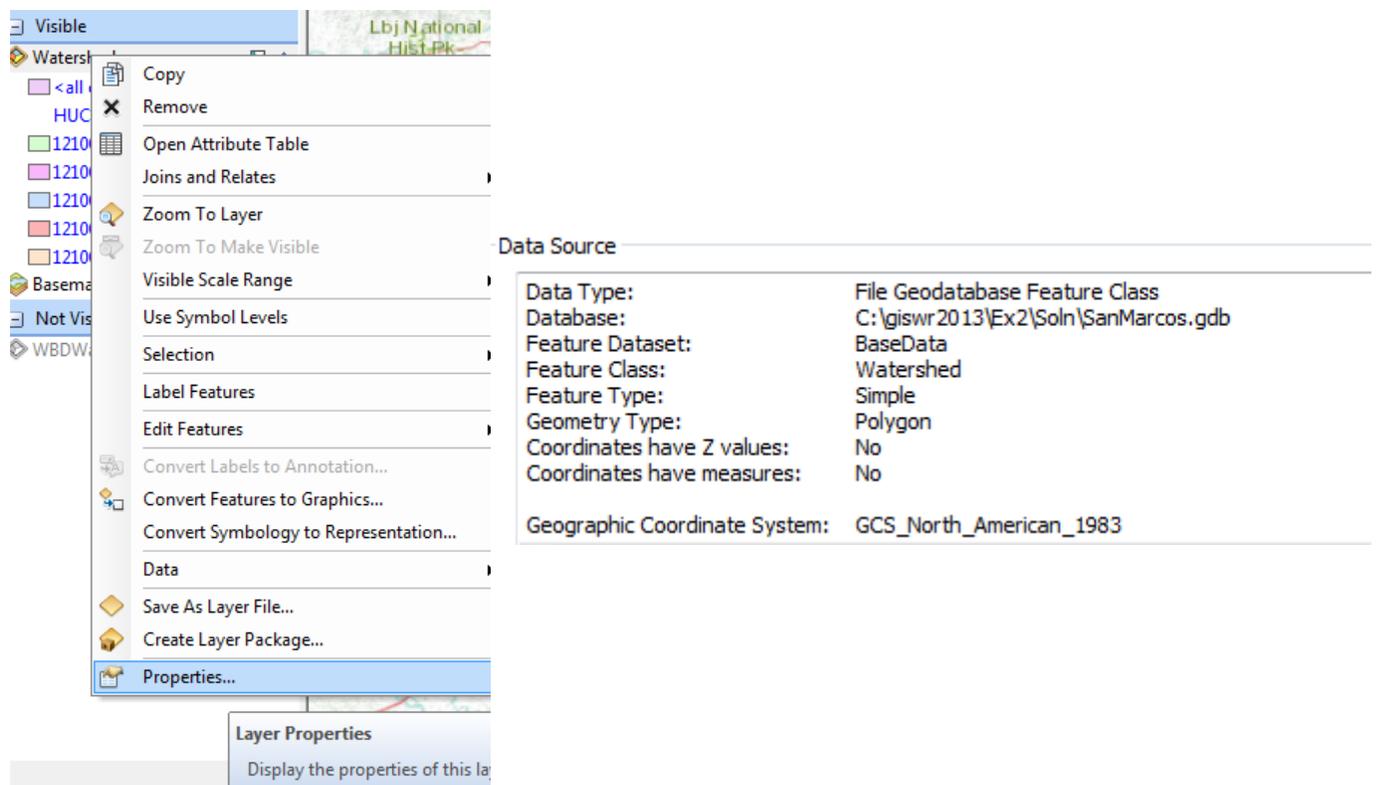
Creating a San Marcos Basin Boundary

It is useful to have a single polygon that is the outline of the San Marcos Basin. Click on the **Search** button in ArcMap and within the Search box that opens up on the right hand side of the ArcMap display, click on **Tools** and then type **Dissolve**. You will see the autocomplete tool gives you several options and select **Dissolve (Data Management)**
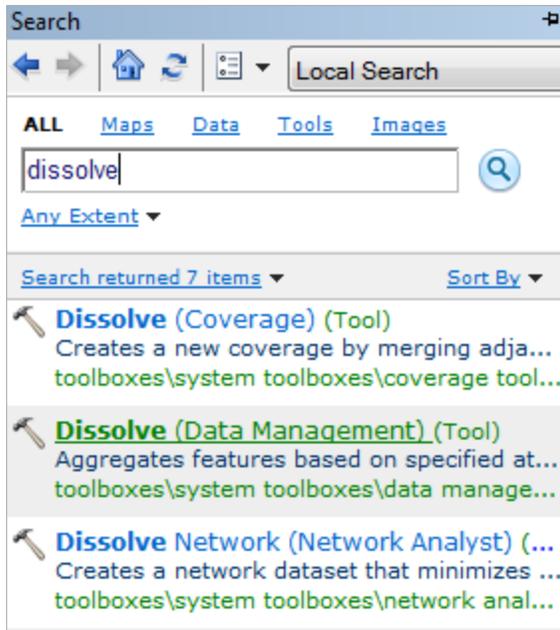
You'll see a **Dissolve** tool window appear. You can drag and drop the **Watershed** feature class from the Table of Contents into the **Input Features** area of this window. Click on **HUC_8** as your **Dissolve_Field**. This means that all Watersheds with the same HUC8 number (12100203) will be merged together. Hit Ok to execute the function. Notice that the Output Feature Class is being set automatically to go to a default geodatabase location and called **Watershed_Dissolve1**
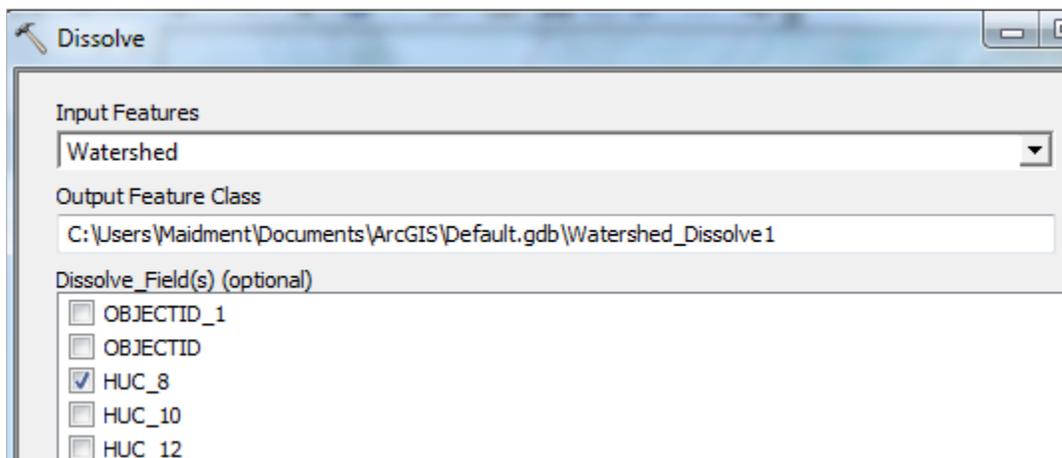
There'll be no apparent activity for a while and then you'll see some blue scrolling text at the bottom right and a pop up indicating completion and the **Watershed_Dissolve1** will appear.

Let's put this new feature class into the SanMarcos geodatabase and call it Basin**.** Right click on the **Basin** feature class and select **Export Data**
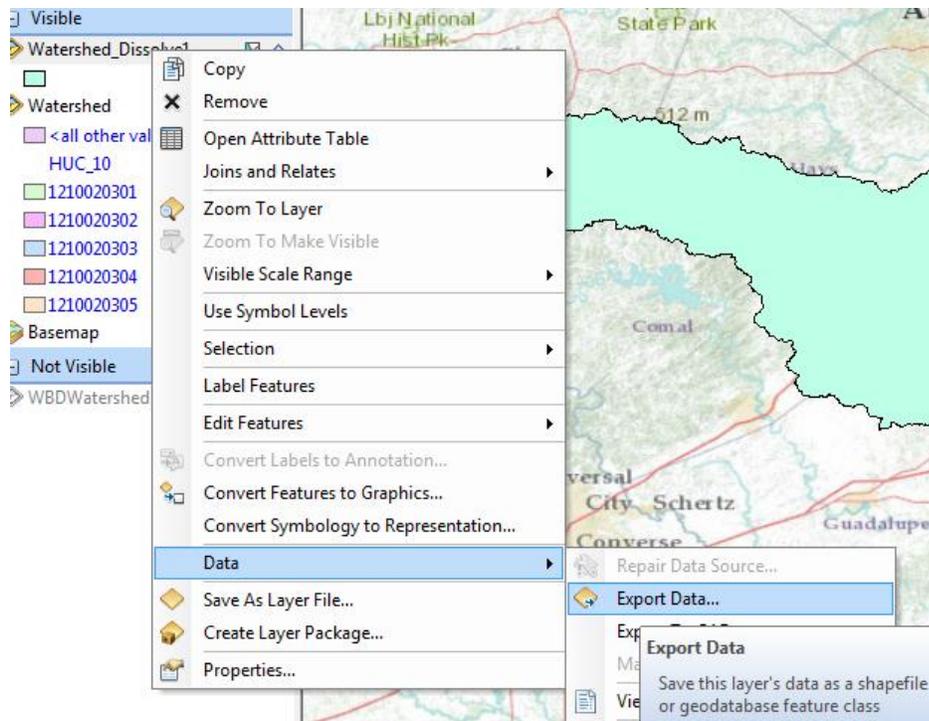
Name the new feature class Basin
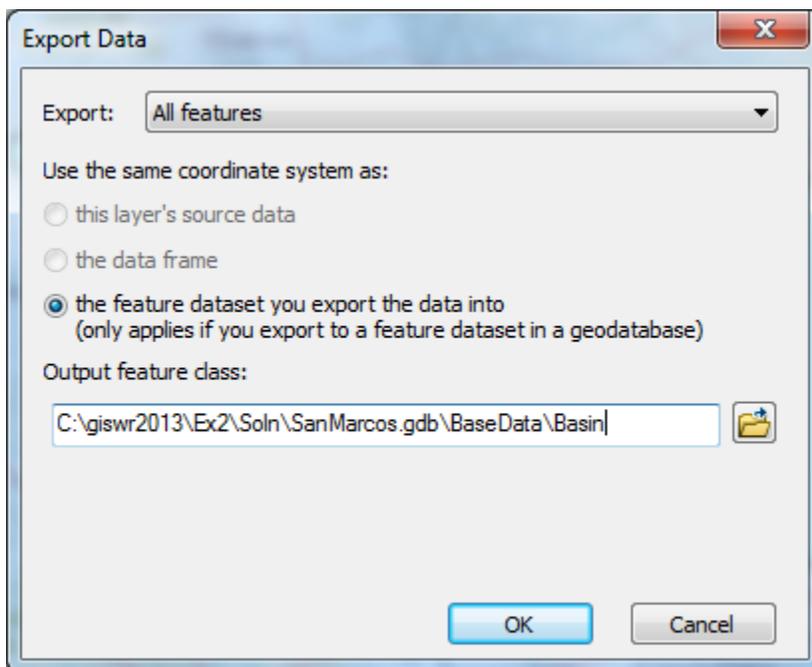
Lets alter the map display to make the **Basin** layer just an outline. Click on the Symbol for the

Basin layer ☑ Basin ▢▮ and select **Hollow** for the shape, Green for the **Outline Color** and 2 for the **Outline Width**.
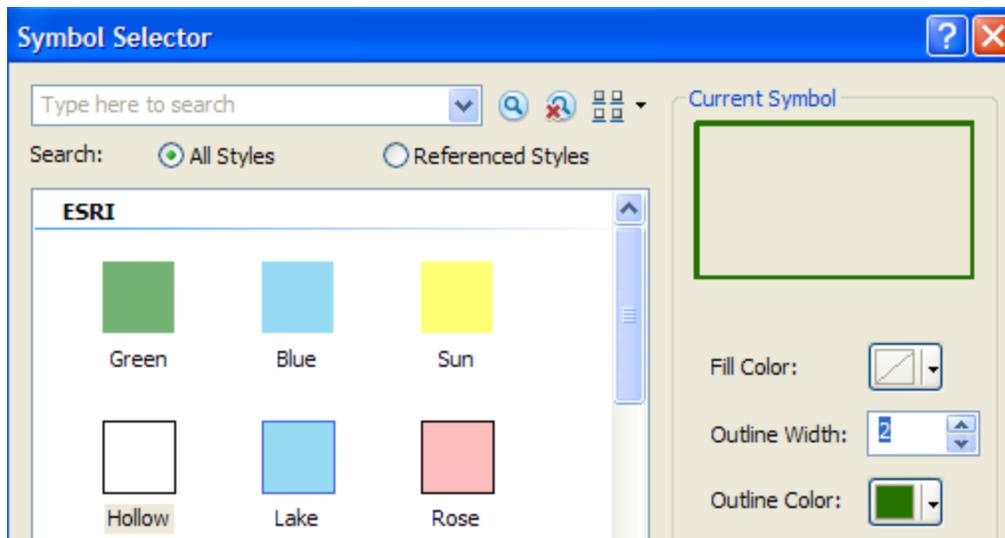
And you'll get a very nice looking map of the San Marcos Basin with its constituent subdrainage areas. Let's also remove the **Watershed_Dissolve1** feature class since we don't need that any more in our map display. Right click on that feature class and select **Remove**
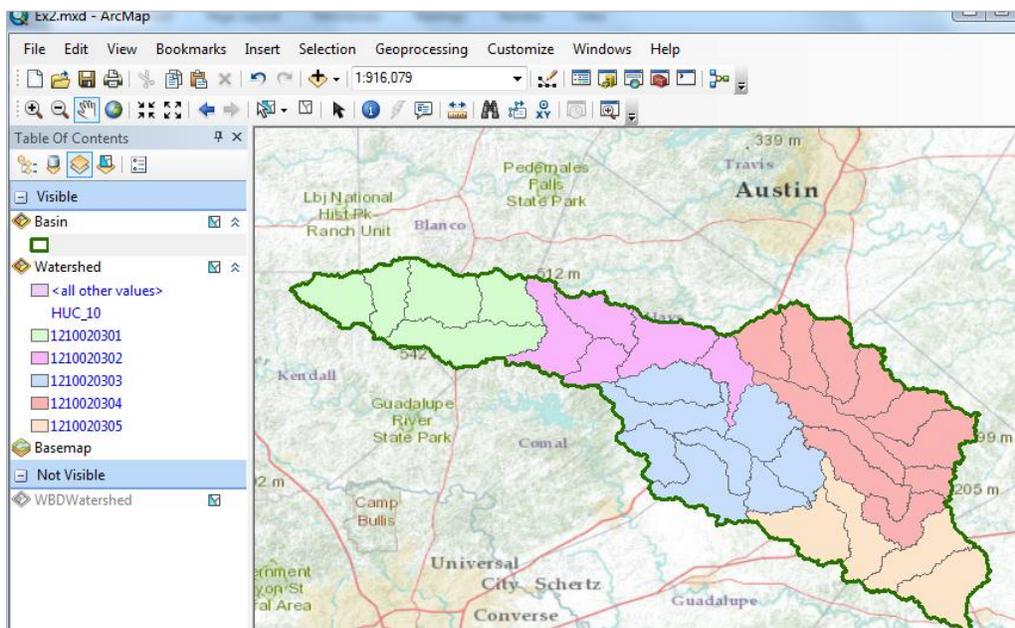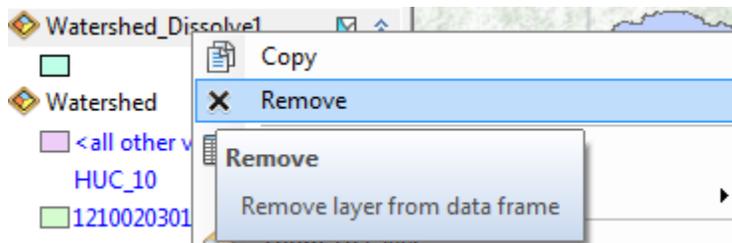
Right click on the Basin feature class and open its Attribute Table. Notice that the Basin feature class has only one Polygon and it is identified with the HUC_8 = 12100203, which is the 8-digit number all the HUC_12 subwatersheds had in common.

	OBJECTID_1 *	Shape *	HUC_8	Shape_Length	Shape_Area
▶	1	Polygon	12100203	4.405801	0.32871

Save your ArcMap document to the file **Ex2Basin.mxd.** Note that this is a different name than used earlier, so you can retrieve the former configuration or this one separately.

Click on the **Catalog** window in ArcMap and navigate to your **BaseData** feature dataset. Notice how you've now got the **Watershed** and **Basin** feature classes that you've just created stored inside it.
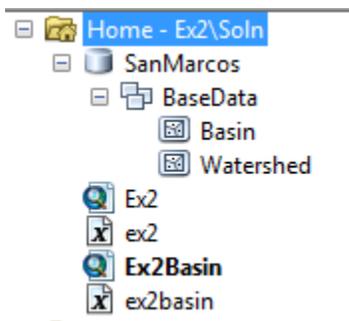
To be turned in: Make a map of the San Marcos basin with its HUC-10 and HUC-12 watersheds and subwatersheds. How many HUC-10 and HUC-12 units exist in the San Marcos Basin?

Soil Information for the San Marcos Basin

Now, we are going to use some of the new data services to find some soil properties of the San Marcos basin.

In ArcCatalog, select **Add GIS Server** and accept **Use GIS Services**
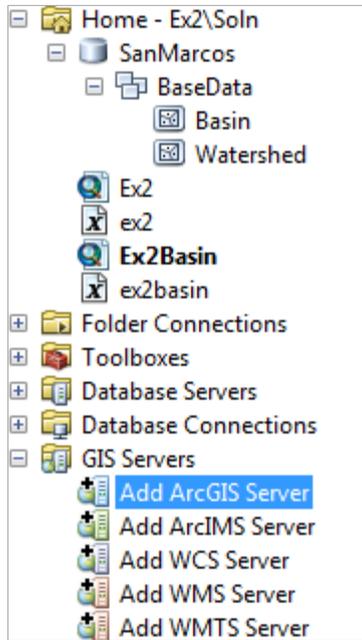
For the server URL use http://landscape2.arcgis.com/arcgis/services Enter your ArcGIS Online **User Name** and **Password** and hit Finish.
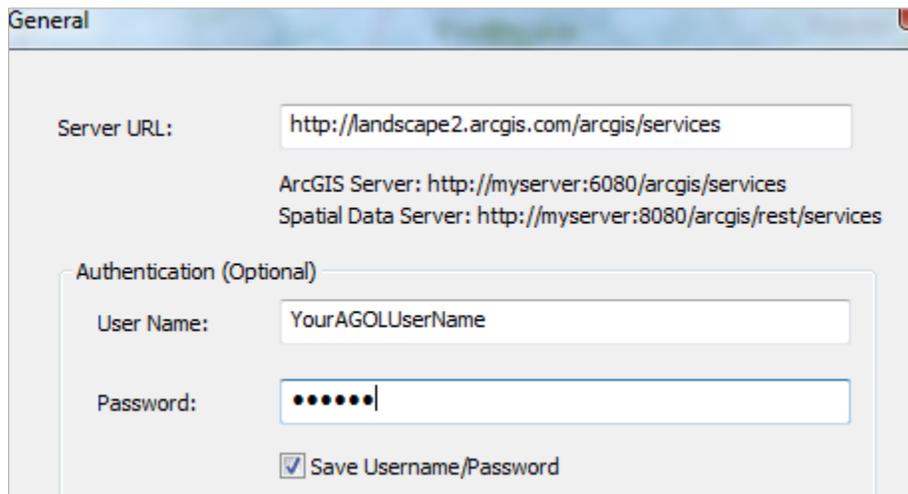
If you click on the + sign on the service that appears, you'll see an entry for **USA_Soils_Available_Water_Storage**. This is a data layer derived from USDA SSURGO that shows the amount of the soil volume that is available for water storage. What we will be using are data for the top 1m of soil and the results are reported in cm, so 10cm means that in the top 1 meter of soil, 10% of the soil volume is available for storing water between the wilting point and field capacity of the soil, which are the lower and upper limits of the soil water storage that is available to vegetation through their roots.
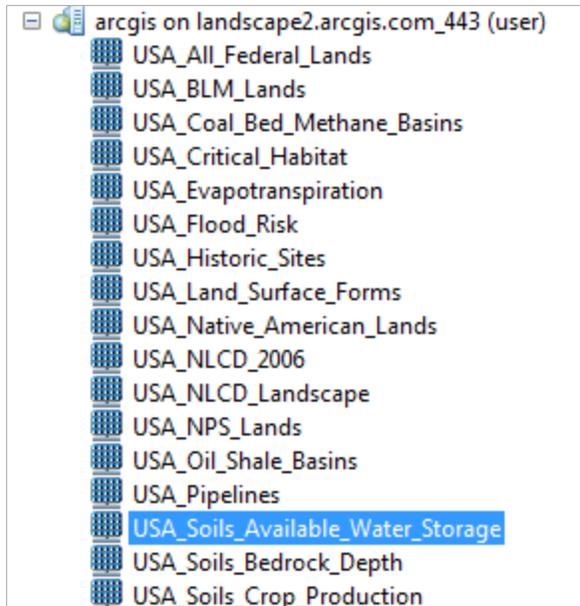
Drag the **USA_Soils_Available_Water_Storage** layer into your map and you'll see it shows up with a nice blue color where more blue means more avaialable water storage, and vice versa. Click on List by Drawing Order to get back your original Table of Contents display and click off the **Watershed** layer, so that you only have the **Basin** displayed over the soil water data.
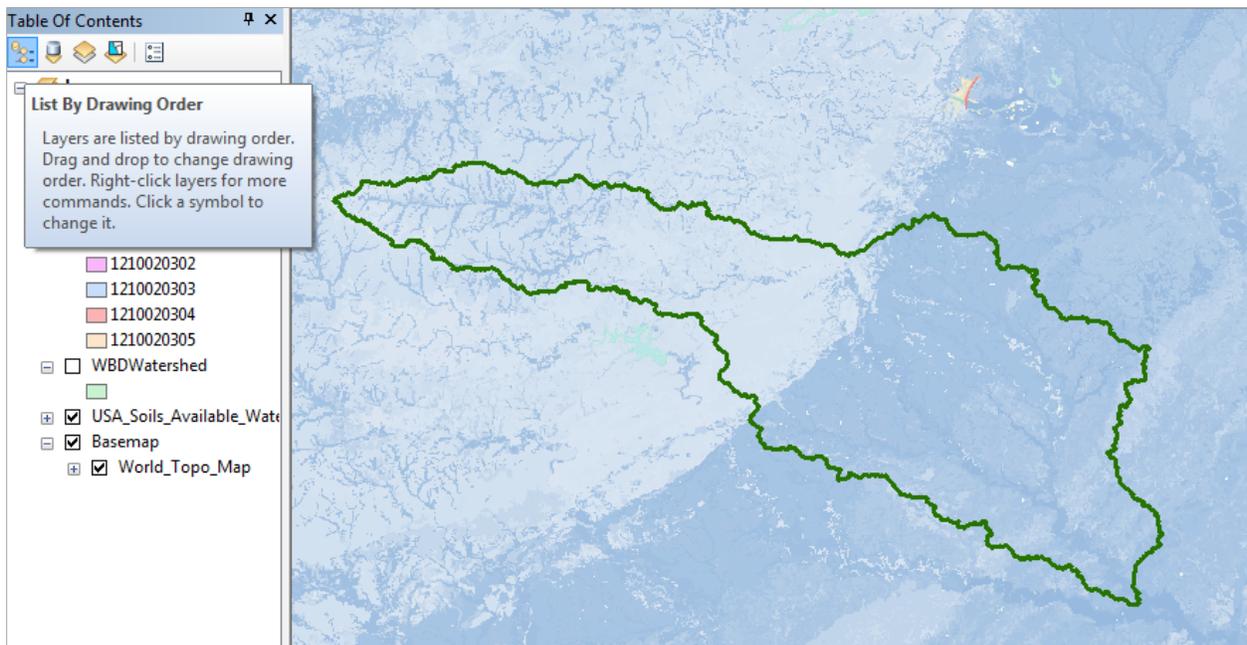
Click on the Identify tool , and in the resulting display select USA_Soils_Available_Water_Storage as the Layer to be Identified, and click around the map. The **Pixel** value is the Available Water Storage value, in this example, 24 cm, which means that

about 24% of the soil volume could be filled with water that is accessible to vegetation. Notice the striking spatial pattern of this soil characteristic.
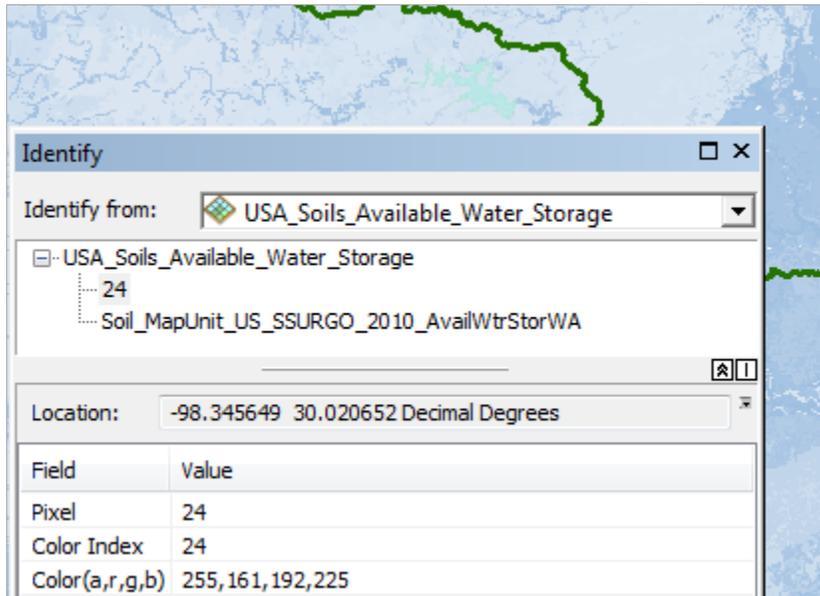
In order to do the next computation, you must have the Spatial Analyst extension of ArcGIS active. Click on **Customize/Extensions** and make sure that **Spatial Analyst** is checked on
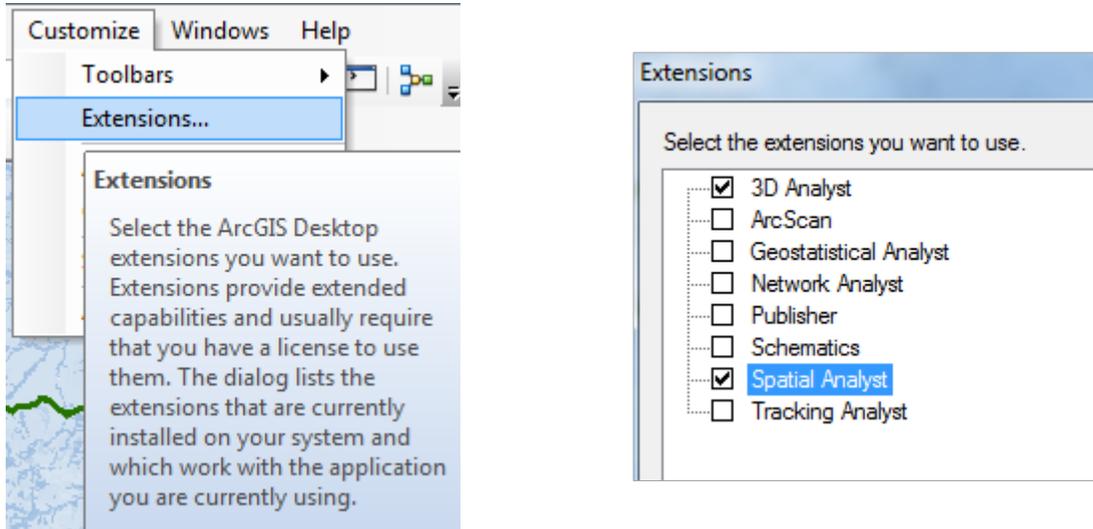
Use the **Search** to identify tools for **Zonal Statistics** and select **Zonal Statistics as Table** tool (click on the Green entry at the top by the little hammer).
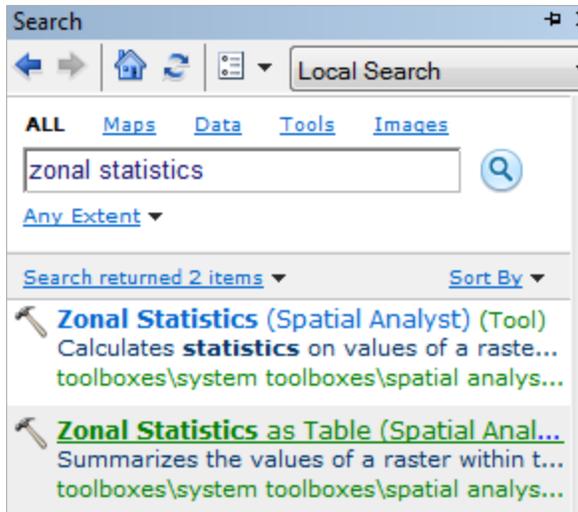
Choose **Basin** for the Input Feature Zone, leave **HUC_8** as the Zone field, and put **USA_Soils_Available_Water_Storage** as the Input value raster. What this tool does is to summarize the statistics of the soil water storage for the whole San Marcos Basin (HUC_8 = 12100203). Let the output table go to the Default Geodatabase since we really don't need to retain it, just get data from it.
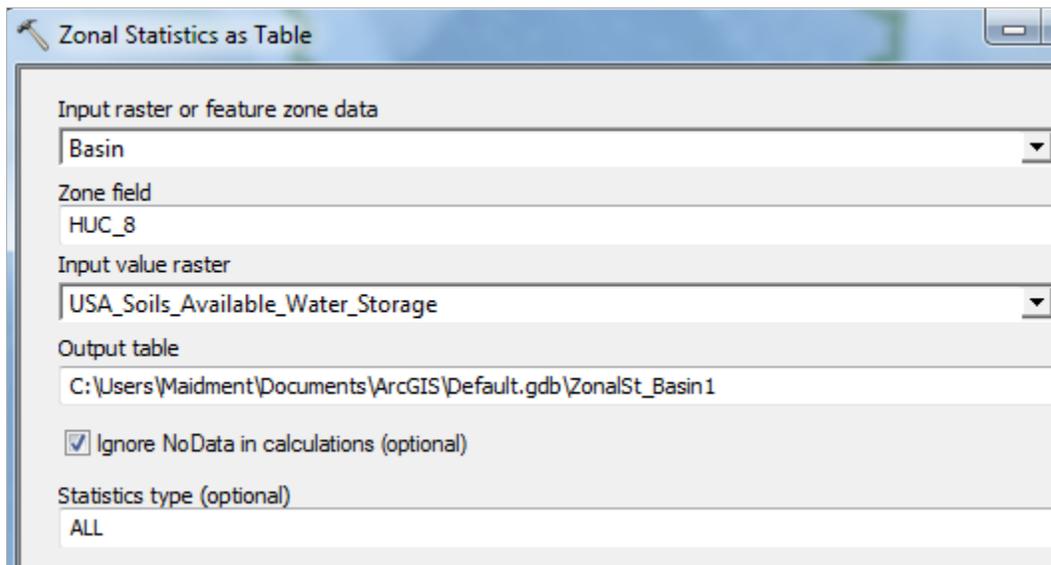
The resulting table **ZonalSt_Basin1** shows that there are 388,578 (**Count**) soil cells in the basin, whose area is 3497205600 (**Area**) square meters or 3497.2 square kilometers, available water storage varies between 0 (**Min**) and 28 (**Max**) cm with a mean (**Mean**) of 14.32 cm and a standard deviation (**Std**) of 8.68 cm. Very cool!

Save your map as Ex2Soils.mxd.

To be turned in: Make a map of the soil water storage variation over the San Marcos Basin. Discuss the spatial pattern of soil moisture storage that is shown in your map. Why is it like this? If the area of the basin is 3497.2 square kilometers, what volume of water (km³) could be stored in the top 1m of soil in the San Marcos basin and be available to vegetation?

Obtaining the San Marcos Flowlines

Now, let's get the flowlines for the streams in the San Marcos basin from a dynamic map service.   Repeat the same steps as before to Add GIS Server, except that this time enter http://landscape1.arcgis.com/arcgis/services as the Server URL.
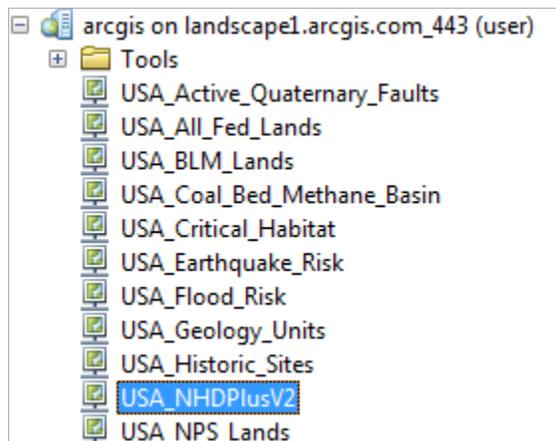
Drag the **USA_NHDPlusV2** service into your map. Initially you don't see anything because the NHDPlus data only show up when you zoom in. Switch the Table of Contents to **List by Drawing Order** and drag the USA_NHDPlusV2 entry in the Table of contents above the Watershed layer and turn the Watershed layer on. Make the ArcMap display as large as possible on your screen.
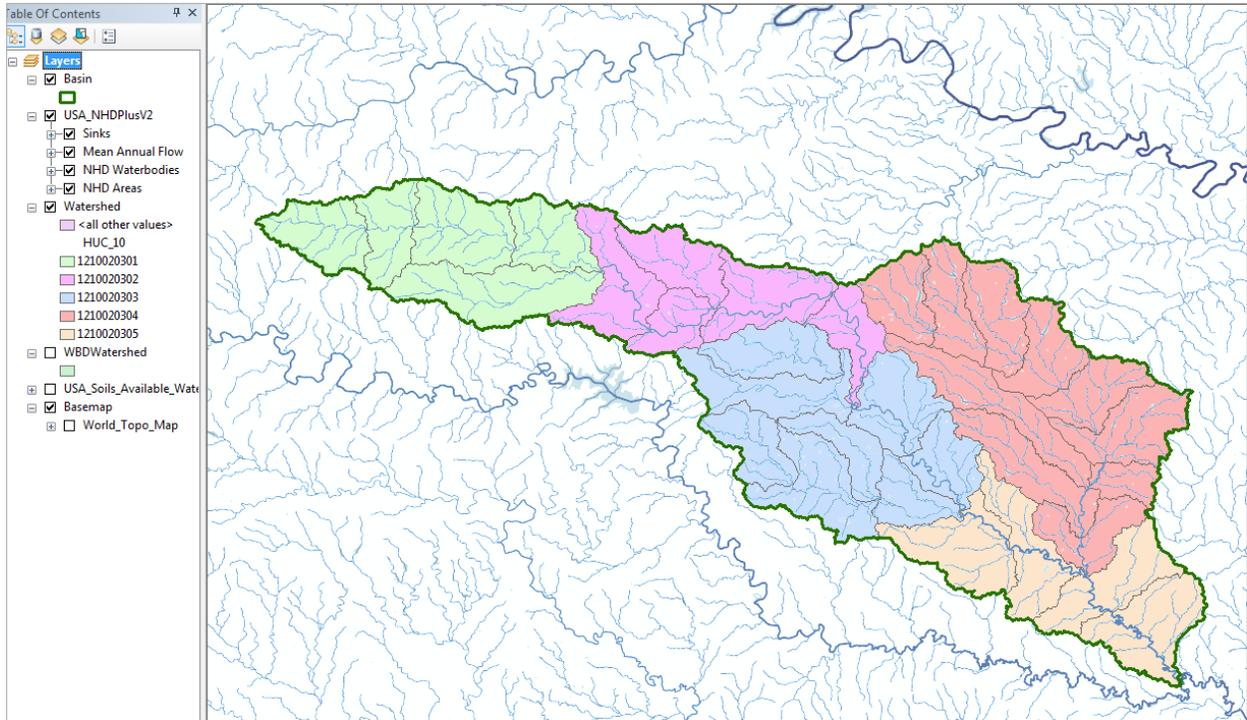
Now you can see the streams over the basin. Zoom into the upper end of the basin. Click on the Identify button again, and choose Watershed as the layer to be displayed.   Click around in this green area, and you'll see that this is a HUC_10 watershed called the **Upper Blanco River**.
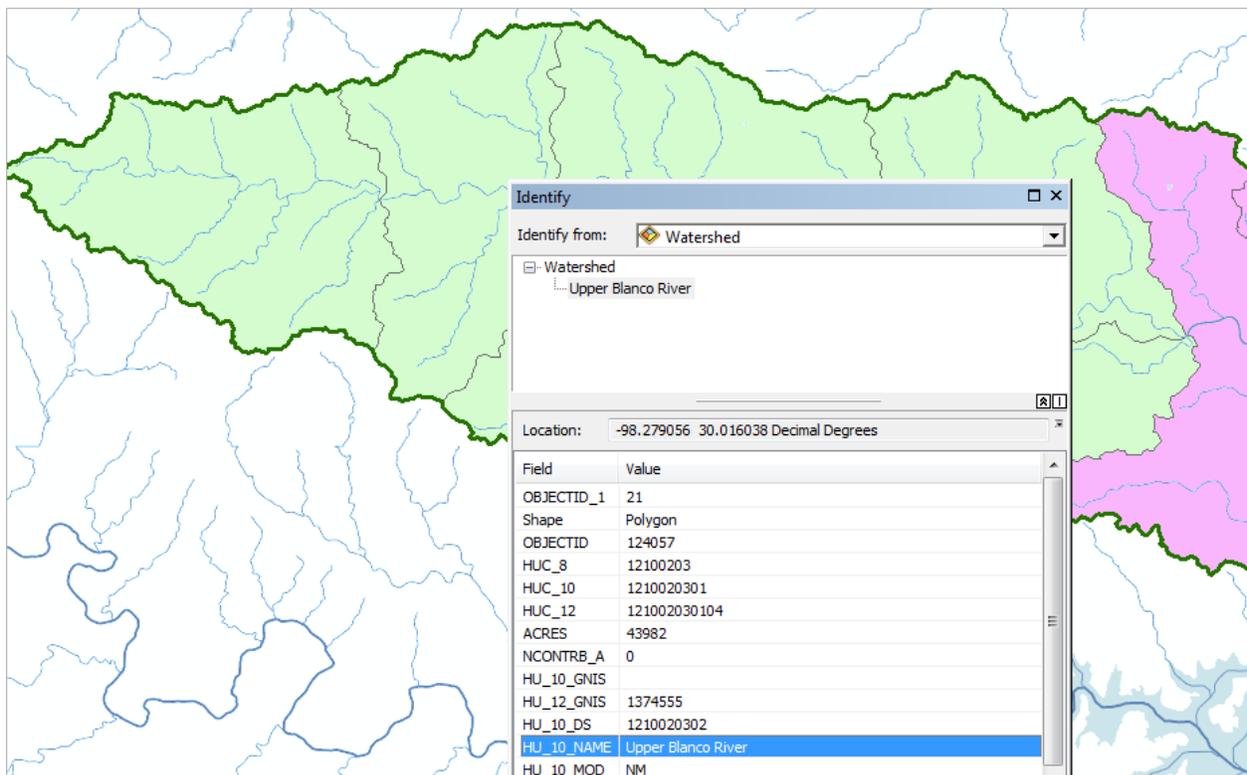
Now let's zoom into the outlet of this watershed and see what we can learn from the NHDPlus Flowlines. Use Identify again, with Mean Annual Flow as the layer to be queried, and click on the Blanco River as it just enters the purple area (ie after the junction of the two branches upstream.   You'll see that you identify some characteristics of the Blanco River itself.
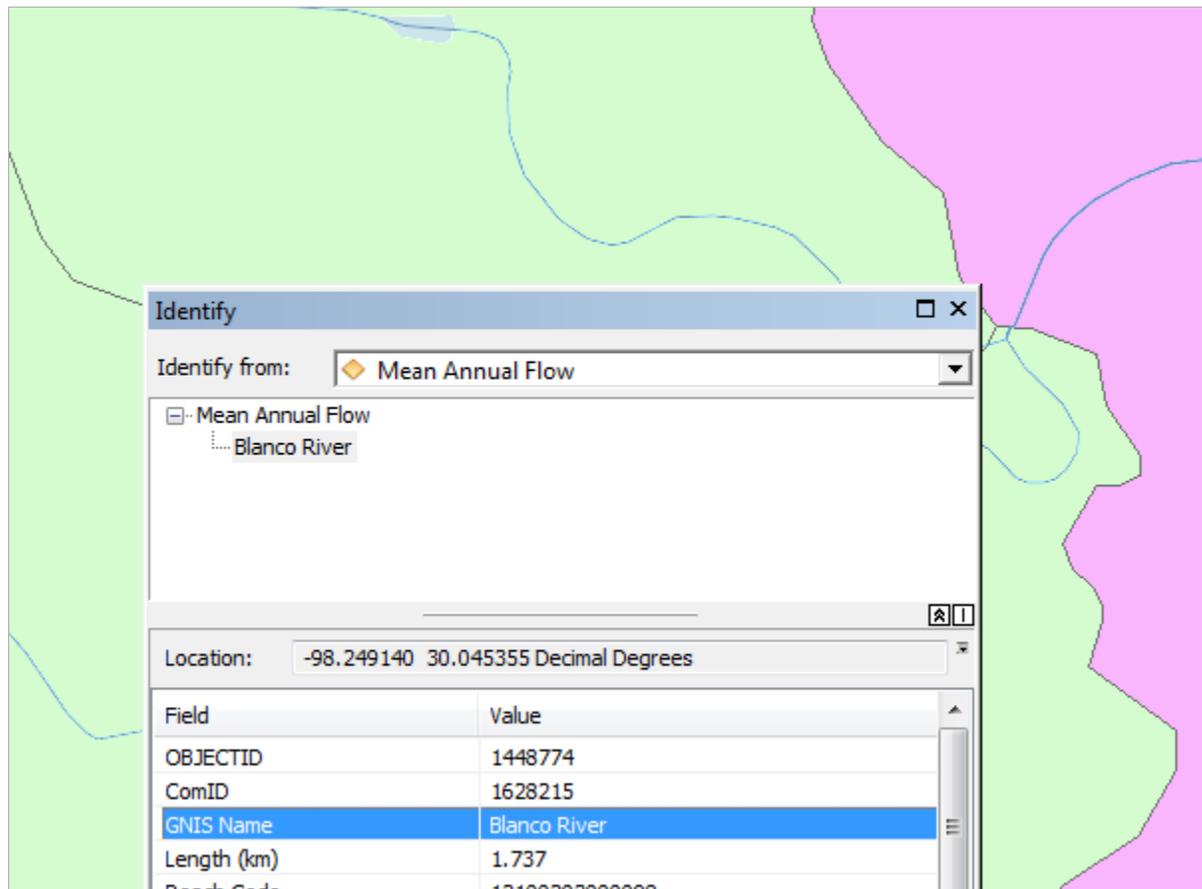
Scroll down a bit further and you'll see some characteristics of the Blanco River at this location. Catchment refers to the local drainage area around this particular stream reach. In this instance, the **Catchment Area** is 3.05 km^2, the mean annual **Temperature** is 19.2°C, and the mean annual **Precipitation** is 912mm. Now, for the upper Upper Blanco River watershed upstream (including this local catchment area), the **Total Drainage Area** is 618.96 km^2, and the mean annual flow (**Gage Adjusted Flow E**) is estimated to be 68.13 cubic feet per second. There is a long story behind these various flow estimates (A, B, C, D, E) that we'll discuss later in the class.   This is actually pretty useful information because we can start to get a sense of how much water we can expect to have in this river, what uses that can support, and so on.

Now let's create a map and do some summarization of watershed attributes. Right click in the grey area. Open the **Draw** Toolbar
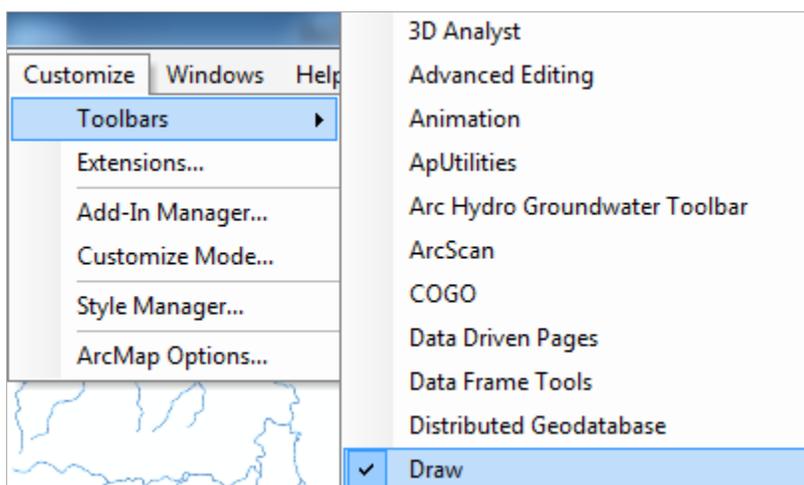
And select a **Callout** tool
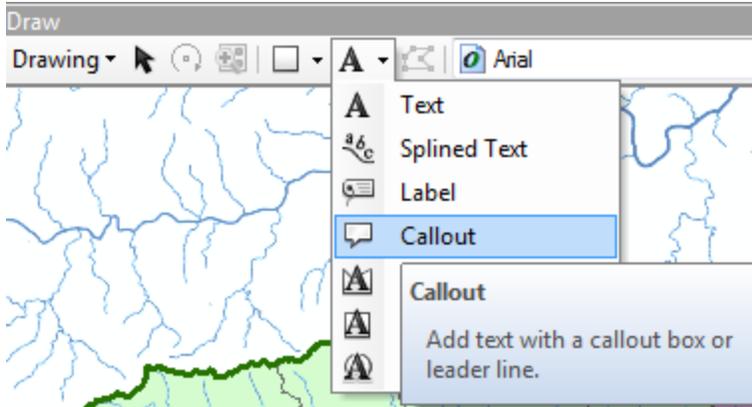
Click somewhere in the Upper Blanco River watershed and drag the callout away to create a connection with that point. Type in **Upper Blanco River** as the text. Do the same for the other four HUC-10 watersheds to create a map like this.

Save your map document as **Ex2Flow.mxd**
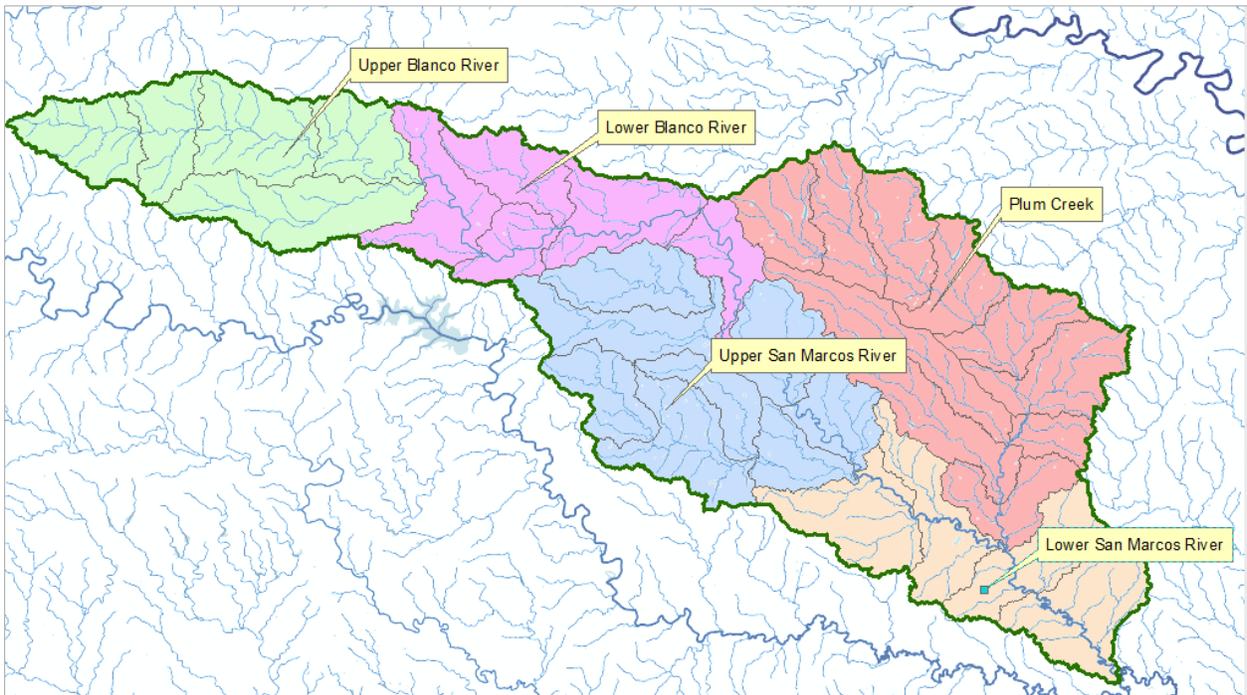
Zoom into the outlet of each of these five watersheds and query the attributes of the NHDPlus flowlines.

To be turned in: Make a map of the five HUC-10 watersheds. Prepare a table that shows for the five watersheds, the watershed Name, the Total Drainage Area at its outlet, and the Incremental Area that results from this Watershed alone. Determine the Mean Annual Flow, Temperature and Precipitation at that location. Compute the Mean Annual Flow per unit of Total Drainage Area in units of mm, and compare that to the Precipitation. Discuss your results.

Creating a Point Feature Class of Stream Gages

Now you are going to build a new Feature Class yourself of stream gage locations in the San Marcos basin. I have extracted information from the USGS site information at http://waterdata.usgs.gov/tx/nwis/si

SiteID	SiteName	Latitude	Longitude	DASqMile	MAFlow
08171000	Blanco Rv at Wimberley, Tx	29° 59' 39"	98° 05' 19"	355	142
08171300	Blanco Rv nr Kyle, Tx	29° 58' 45"	97° 54' 35"	412	165
08172400	Plum Ck at Lockhart, Tx	29° 55' 22"	97° 40' 44"	112	49
08173000	Plum Ck nr Luling, Tx	29° 41' 58"	97° 36' 12"	309	114
08172000	San Marcos Rv at Luling, Tx	29° 39' 58"	97° 39' 02"	838	408
08170500	San Marcos Rv at San Marcos, Tx	29° 53' 20"	97° 56' 02"	49	176

(a) Define a table containing an ID and the long, lat coordinates of the gages

The coordinate data is in geographic degrees, minutes, & seconds. These values need to be converted to digital degrees, so go ahead and perform that computation for the 8 pairs of longitude and latitude values. This is something that has to be done carefully because any errors in conversions will result in the stations lying well away from the San Marcos basin. I suggest that you prepare an Excel table showing the gage longitude and latitude in degrees, minutes and seconds, convert it to long, lat in decimal degrees using the formula

Decimal Degrees (DD) = Degrees + Min/60 + Seconds/3600

Remember that West Longitude is negative in decimal degrees. Shown below is a table that I created. **Be sure to format the columns containing the Longitude and Latitude data in decimal degrees (LongDD and LatDD) so that they explicitly have Number format with 4 decimal places using Excel format procedures. Format the column SITEID as Text or it will not retain the leading zero in the SiteID data**. Add the additional information about the USGS SiteID, SiteName and Mean Annual Flow (MAF). Note the name of the worksheet that you have stored the data in. I have called mine **latlong.xlsx**. Close Excel before you proceed to ArcMap.

SiteID	SiteName	Latitude	Longitude	LatDeg	LatMin	LatSec	LongDeg	LongMin	LongSec	LatDD	LongDD	DASqMile	MAFlow
08171000	Blanco Rv at Wimberley, Tx	29° 59' 39"	98° 05' 19"	29	59	39	98	5	19	29.9942	-98.0886	355	142
08171300	Blanco Rv nr Kyle, Tx	29° 58' 45"	97° 54' 35"	29	58	45	97	54	35	29.9792	-97.9097	412	165
08172400	Plum Ck at Lockhart, Tx	29° 55' 22"	97° 40' 44"	29	55	22	97	40	44	29.9228	-97.6789	112	49
08173000	Plum Ck nr Luling, Tx	29° 41' 58"	97° 36' 12"	29	41	58	97	36	12	29.6994	-97.6033	309	114
08172000	San Marcos Rv at Luling, Tx	29° 39' 58"	97° 39' 02"	29	39	58	97	39	2	29.6661	-97.6506	838	408
08170500	San Marcos Rv at San Marcos, Tx	29° 53' 20"	97° 56' 02"	29	53	20	97	56	2	29.8889	-97.9339	49	176

(b) Creating and Projecting a Feature Class of the Gages

(1) Open **ArcMap** and the **Ex2Flow.mxd** file you created earlier in this exercise. Select the add data button and navigate to your Excel spreadsheet
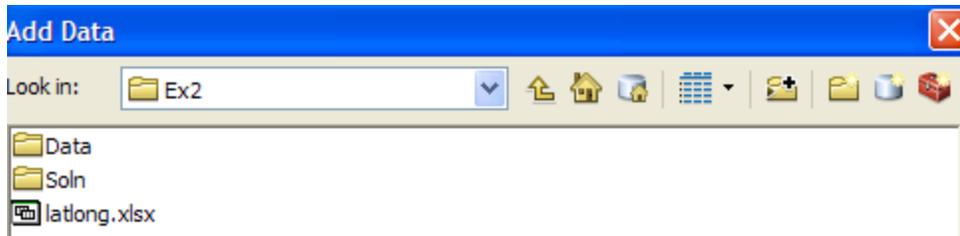
Double click on the spreadsheet to identify the individual worksheet within the spreadsheet that you want to add to ArcMap (it's a coincidence that they have the same name in this example and that is not necessary in general).
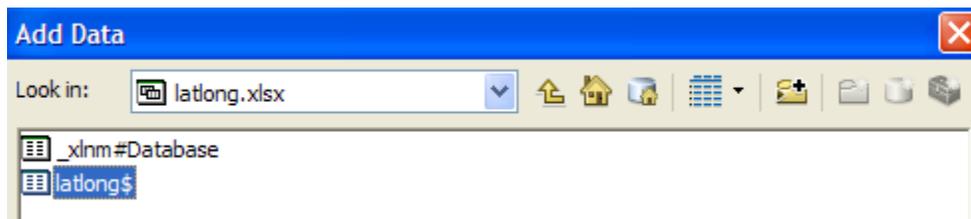
Hit **Add** and your spreadsheet will be added to ArcMap. It's possible that you'll get an error here that says "Cannot make a Database Connection". If that is the case, save your Excel file in the formal **Comma Separated Variable** (.csv) and use that format instead to add your data to ArcMap. This is a simpler format that is more easily read than a straight Excel Spreadsheet.

SiteID	SiteName	Latitude	Longitude	LatDeg	LatMin	LatSec	LongDeg	LongMin	LongSec	LatDD	LongDD	DASqMile	MAFlow
08170500	San Marcos Rv at San Marcos, Tx	29□ 53' 20"	97□ 56' 02"	29	53	20	97	56	2	29.888889	-97.933889	49	176
08173000	Plum Ck nr Luling, Tx	29□ 41' 58"	97□ 36' 12"	29	41	58	97	36	12	29.699444	-97.603333	309	114
08172000	San Marcos Rv at Luling, Tx	29□ 39' 58"	97□ 39' 02"	29	39	58	97	39	2	29.666111	-97.650556	838	408
08172400	Plum Ck at Lockhart, Tx	29□ 55' 22"	97□ 40' 44"	29	55	22	97	40	44	29.922778	-97.678889	112	49
08171300	Blanco Rv nr Kyle, Tx	29□ 58' 45"	97□ 54' 35"	29	58	45	97	54	35	29.979167	-97.909722	412	165
08171000	Blanco Rv at Wimberley, Tx	29□ 59' 39"	98□ 05' 19"	29	59	39	98	5	19	29.994167	-98.088611	355	142

Now we are going to convert the tabular data in the spreadsheet to points in the ArcMap display.

(2) **Right click** on the new table, **latlong$**, and select **Display XY Data**
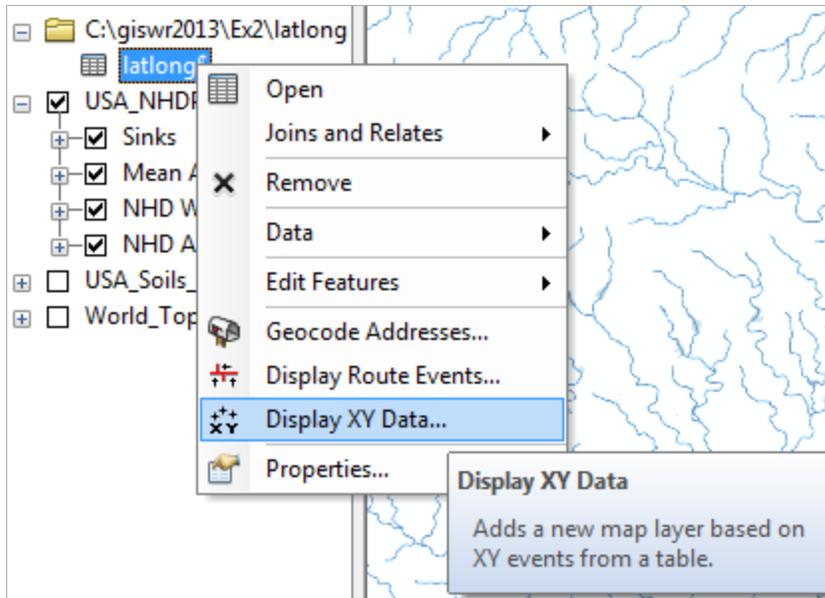
(3) Set the X Field to **LongDD (or Longitude)**, the Y Field to **LatDD (or Latitude)**, *[Be careful here, don't accept the default choices presented to you]* Hit **Edit** to change the spatial coordinate system. Scroll to the folder **Layers** at the bottom of the list to see the Spatial References of the Layers in the Map. Expand the folder to see that NHDFlowline, Watershed and other layers have the Spatial Reference **GCS_North_American_1983**.
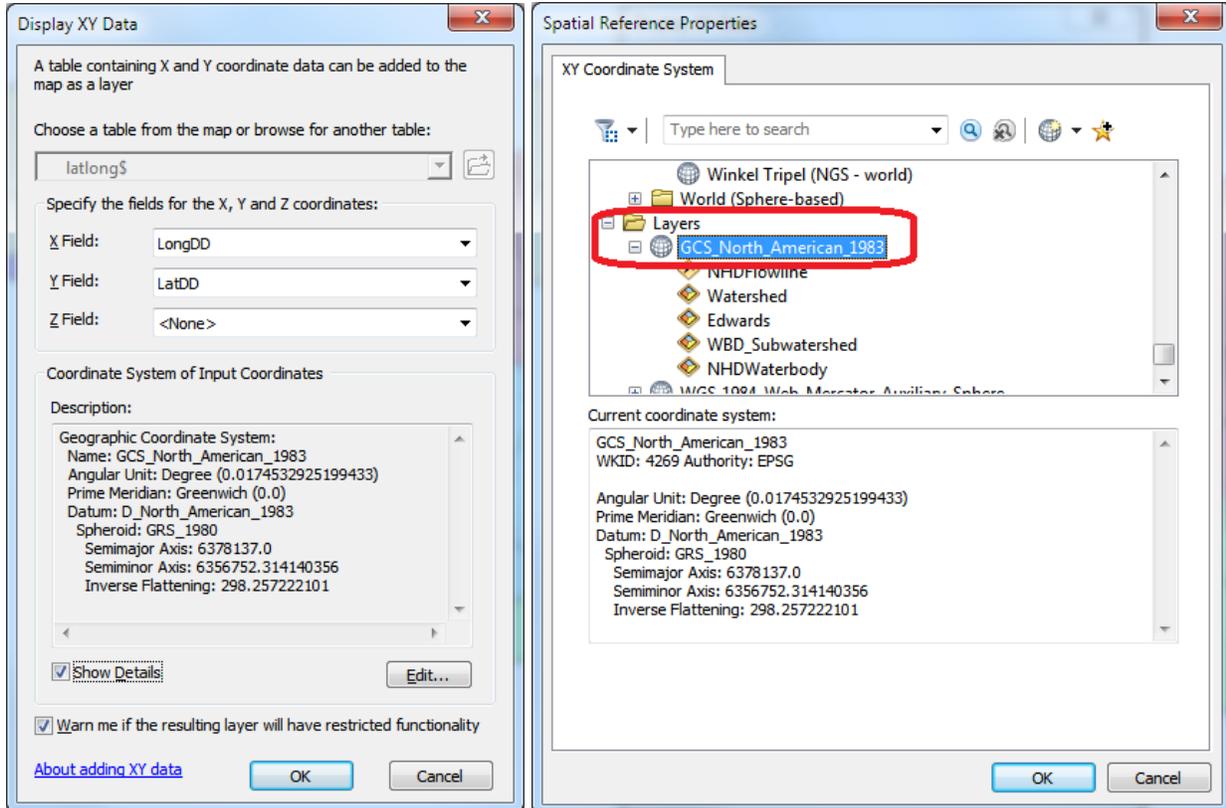
Hit **OK**, to complete it and you'll get a warning message about your table not having an ObjectID. Just hit Ok and and voila! Your gage points show up on the map right along the San Marcos River just like they should. Magic.   I remember the first time I did this I was really thrilled. This stuff really works. I can create data points myself! If you don't see any points, don't be dismayed. Check back at your spreadsheet to make sure that the correct X field and Y field have been selected as the ones that have your data in decimal degrees.

Now let's store these points in our geodatabase as a real feature class, called **Gages**. Right click on **Latlong$ Events** and Export Data the points to a Gages feature class in the San Marcos Geodatabase, as you did earlier with Basin and Watershed.
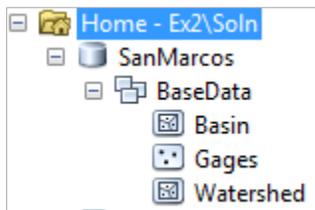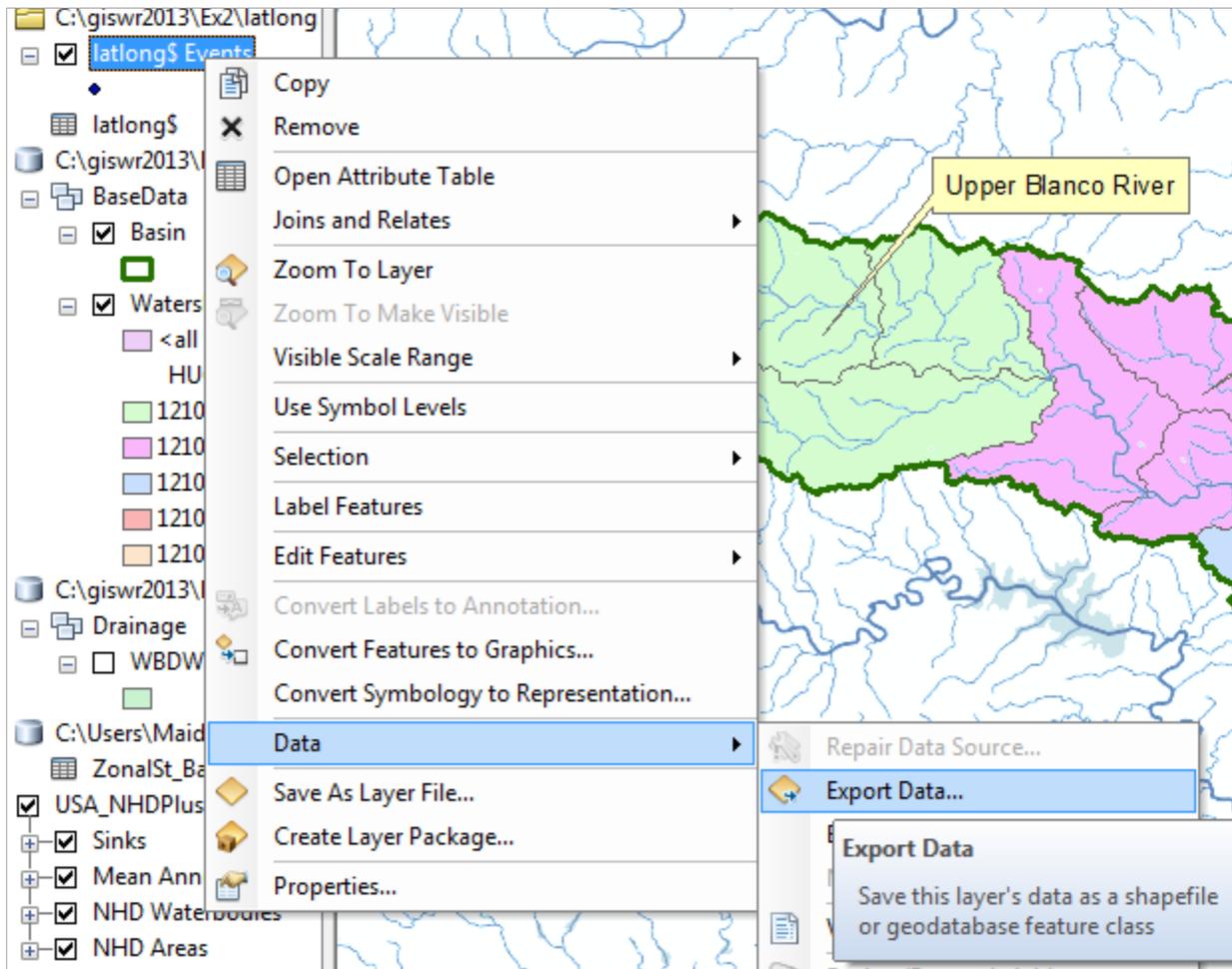
Add the resulting Gages back into your map and recolor and resize them so you can see them clearly.
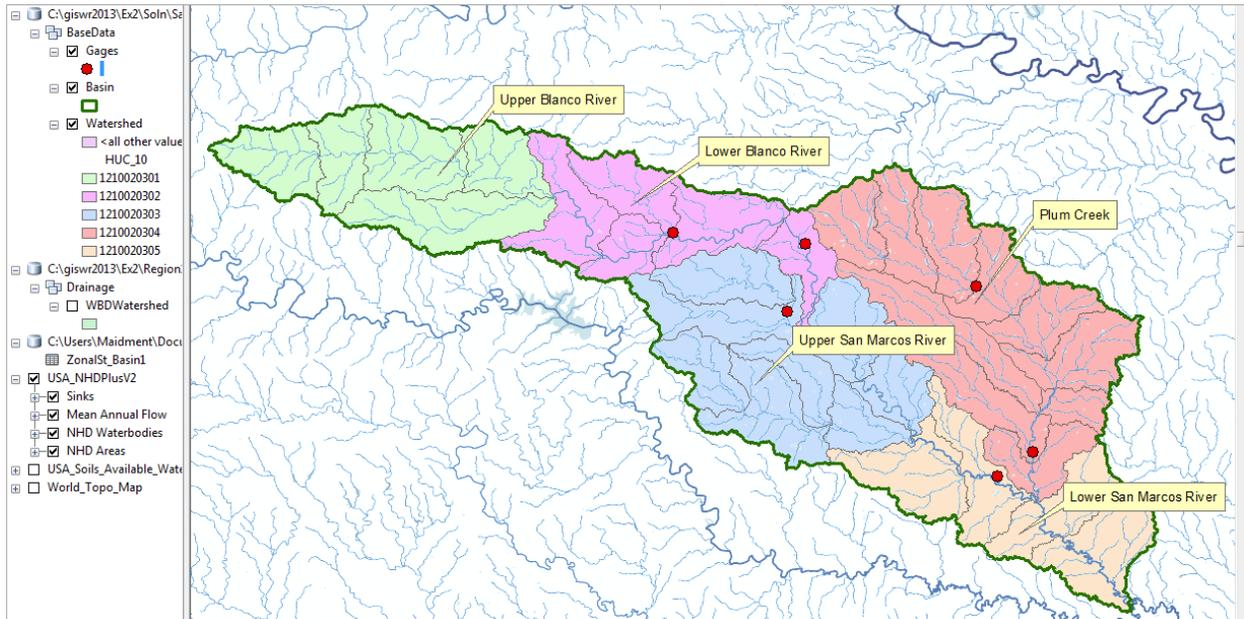
Now lets label the Gages with their Names.   Right click on the Gage feature class and click on Label Features. You'll see some labels show up in small lettering.   It can occur that some labels don't show up because they display where you've got your Watershed Callouts created earlier. Drag those Callout boxes to another location and the gage labels will appear.
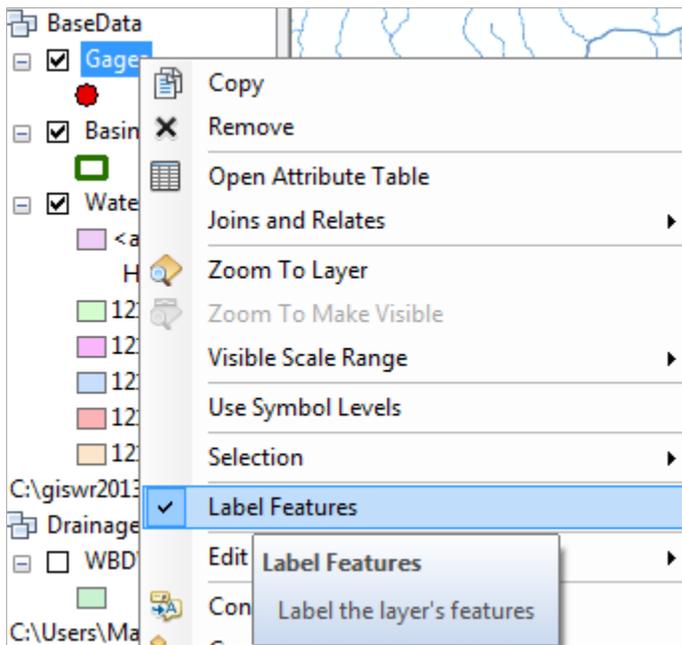
To resize the labels, right click on Gage and select Properties/ Labels, and then select 12 point type.
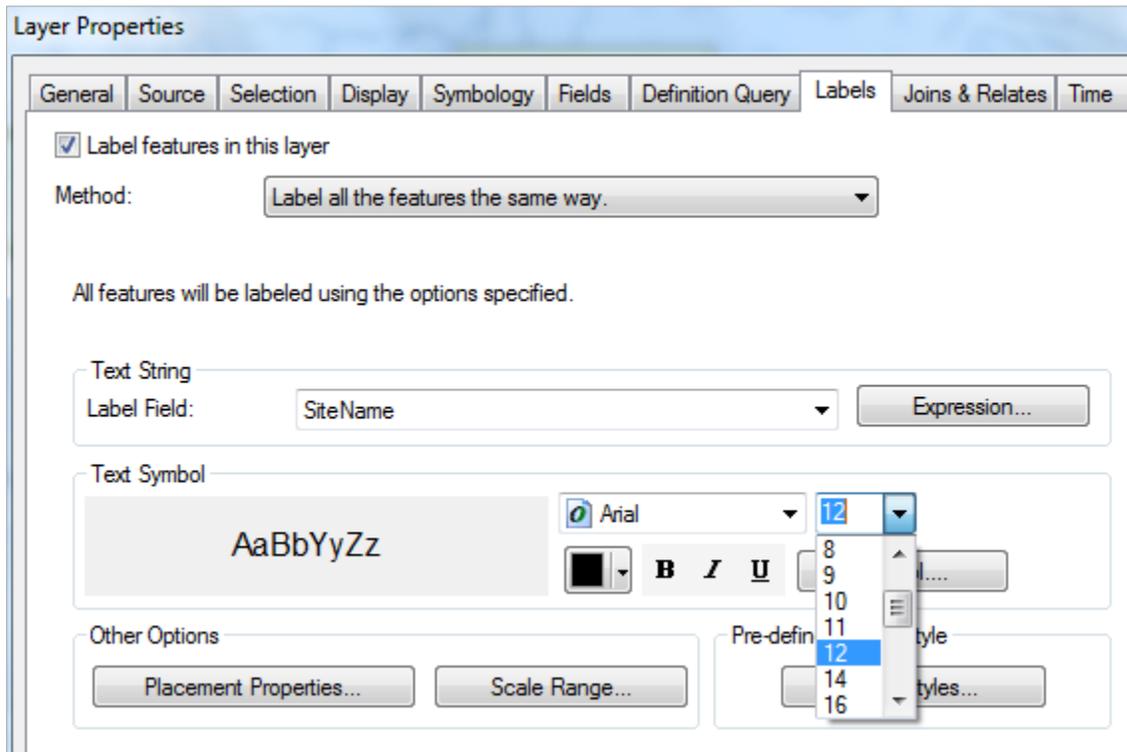
Open the Attribute Table of the Gages, and right click on the fields that you really don't need to see and hide them and reorder the records so that the gages are displayed from 1-6 in the table.
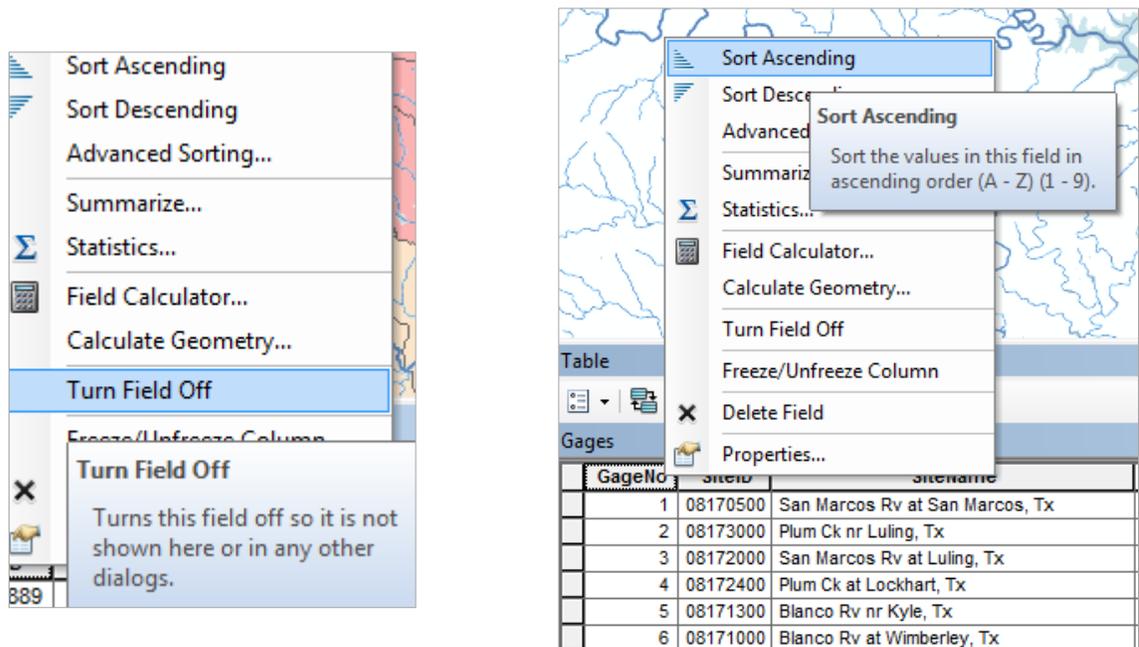
To make a display like that shown below.
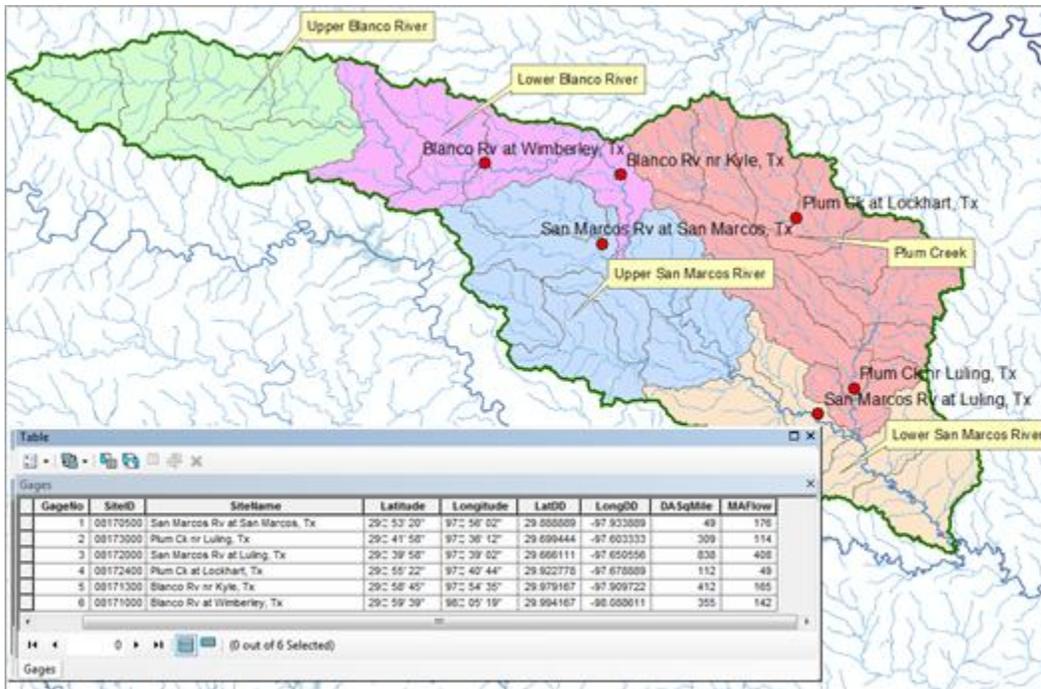
Delete the Watershed name Callouts, and add to the map the Edwards aquifer feature class **Edwards** which you downloaded with the original map package. The green areas in the aquifer are its Outcrop areas where it touches the land surface. The brown areas are its downdip area where it goes below the ground surface.
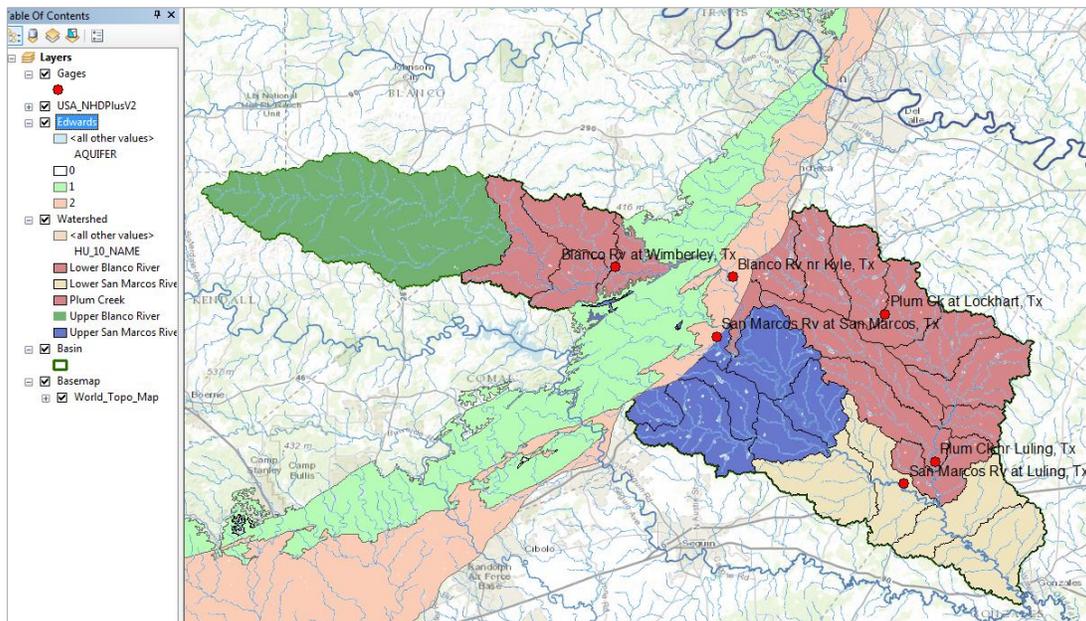
The interaction of the groundwater flow from the Edwards Aquifer and the surface water flow from the surface watersheds creates a spatially varied picture of surface and groundwater balance in this basin.

To be turned in. Make a map showing the labeled gages and their attribute table. Zoom into each of your gages, and compare the Drainage Area and the Mean Annual Flow from between the gage values and those given on the NHDPlus.   Prepare a table for your six gages which shows these comparisons. Discuss your results. What causes the large discrepancies in flow that you observe at some sites between the gaged values and the NHDPlus values?

Ok, you're done!

Resave your **Ex2Gages.mxd** file.

Summary of Items to be Turned in:

1. *Make a map of the San Marcos basin with its HUC-10 and HUC-12 watersheds and subwatersheds. How many HUC-10 and HUC-12 units exist in the San Marcos Basin?*
2. *Make a map of the soil water storage variation over the San Marcos Basin. Discuss the spatial pattern of soil moisture storage that is shown in your map. Why is it like this? If the area of the basin is 3497.2 square kilometers, what volume of water (km^3) could be stored in the top 1m of soil in the San Marcos basin and be available to vegetation?*
3. *Make a map of the five HUC-10 watersheds. Prepare a table that shows for the five watersheds, the watershed Name, the Total Drainage Area at its outlet, and the Incremental Area that results from this Watershed alone. Determine the Mean Annual Flow, Temperature and Precipitation at that location. Compute the Mean Annual Flow per unit of Total Drainage Area in units of mm, and compare that to the Precipitation. Discuss your results.*
4. *Make a map showing the labeled gages and their attribute table. Zoom into each of your gages, and compare the Drainage Area and the Mean Annual Flow from between the gage values and those given on the NHDPlus.   Prepare a table for your six gages which shows these comparisons. Discuss your results. What causes the large discrepancies in flow that you observe at some sites between the gaged values and the NHDPlus values?*