**Lab 1: Dry (and not so dry) counties**

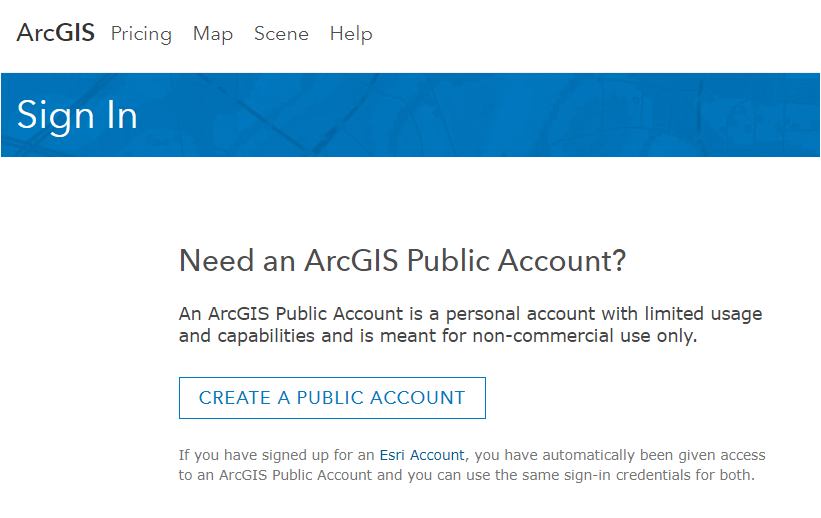
*Geog2011: Introduction to GIScience*

**Introduction**

In this lab, you’ll be doing analysis of climate data from NASA’s Daymet dataset ([more information here](https://daymet.ornl.gov/)). Based on climate observations, Daymet provides daily data on temperature, humidity, and precipitation for every point within the United States going back several decades. In this lab, you’ll be looking just at precipitation for one month--January 2015--for counties in the continental United States. The unit for this measurement is millimeters per month. Using descriptive statistics and online mapping, you’ll analyze regional trends in temperature during that month.

**Setting up a free ArcGIS account**

We will be using ESRI’s ArcGIS Online service for most of our web mapping this semester. ESRI offers a free account with storage and other limitations for this service. You can sign up [on their website](https://www.arcgis.com/home). Click on “Sign In” in the upper right, click on “Create a Public Account,” and follow the instructions.



**Lab data**

The files for the lab are the same you used in Lab 0, so continue to use what you worked with for that assignment.

**Summarizing the data and identifying outliers**

Open the “Daymet\_2015.xls” file in Excel (or a similar program). For this lab, you’ll specifically be looking at ***prcp***, the total precipitation in each county in January 2015. Specifically, we want to identify the number of outliers for the prcp variable by census region. We will do so using filter and pivot tables in Excel.

First, you need to determine the threshold for outliers. You can start by using Excel’s functions to calculate summary statistics for these counties. We went over this in Lab 0, and the functions you want are listed below.

* Mean: =AVERAGE(variable range)
* Median: =MEDIAN(variable range)
* 1st quartile: =QUARTILE(variable range,1)
* 3rd quartile: =QUARTILE(variable range,3)
* IQR: Subtract the 1st quartile from the 3rd quartile.
* Lower outlier: 1st quartile - (1.5 \* IQR)
* Upper outlier threshold: 3rd quartile + (1.5 \* IQR)

Calculate each of these values for the whole dataset. You may want to jot down the values for the upper and lower outliers. (Spoiler alert! Since the value for lower outliers is lower than 0, you won’t worry about it for this lab. There’s no such thing as negative precipitation.)

Now we want to identify the upper outliers counties. To do so, you’ll need to create *dummy variables*--a column that has a 0 (not an outlier) or 1 (an outlier) for each county. [See this video tutorial](https://youtu.be/pnPECPo4fUU) for an example of how this works. The basic steps are:

1. Create a new dummy variable (call it “outlier”).
2. Set all values to 0
3. Filter the data based on the outlier value
4. Set value of filtered observations to 1 for the outlier variable

Once you’ve created those dummy variables, you’ll use pivot tables to sum up the “1”s by the region variable. [Here](https://support.office.com/en-us/article/create-a-pivottable-to-analyze-worksheet-data-a9a84538-bfe9-40a9-a8e9-f99134456576)’s Microsoft’s help page for pivot tables. You can also create them in Google Sheets--[here’s a link](https://support.google.com/a/users/answer/9308944?hl=en). They can do some pretty powerful analysis but also can be confusing to use. We’re doing something simple here--summing up the count of outlier counties by region. To do that, you’ll need to:

1. Insert a pivot table
2. Select the region and outlier columns
3. Make sure that the regions define the rows and the values are the sum of outliers

[This video](https://youtu.be/C-LAMdbPPc4) walks you through the basic process.

That’s it for the Excel portion of this lab! Now we’ll move onto mapping these data.

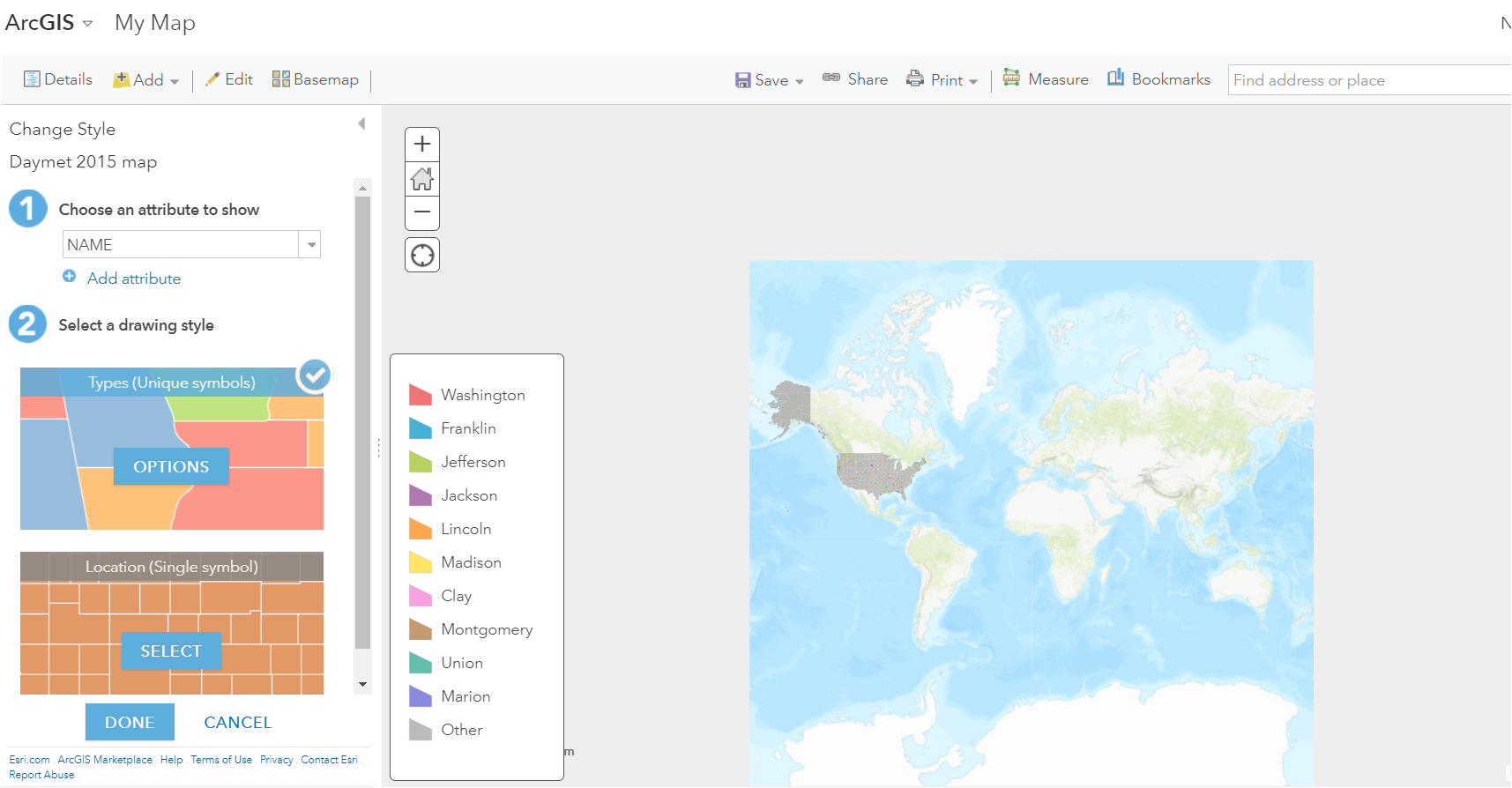
**Mapping the data**

You’re now ready to upload the shapefile to ArcGIS Online. Once you’ve logged in, choose “Map” from the top menu bar on the home page.



A new blank canvas will open. You can now add layers to the map. ESRI provides an overview of how to upload data from a file [on this support guide](http://doc.arcgis.com/en/arcgis-online/create-maps/add-layers.htm). Just click on the link for “add layers from files” to scroll down to the relevant guide.

You should now have something that looks like this:



Use the zoom control or your mouse to zoom to just the 48 continental states--you won’t be looking at Alaska or Hawaii for this analysis.

As shown above, ArcGIS will try to style the data based on the first attribute, county name. You want the prcp variable. Changing the variable and style for the map is relatively easy. ESRI’s guide is found [at this link](https://doc.arcgis.com/en/arcgis-online/create-maps/change-style.htm). You can also watch a short video tutorial that styles the temperature variable.

* [Link to video tutorial](https://youtu.be/fX3AWyG3GxA) Note--this tutorial uses the tmax variable, **NOT** the prcp variable. You will need to use the prcp variable for your final map.

Create a map that breaks the data into five classes using a manual classification scheme. You want the breaks for these classes to be:

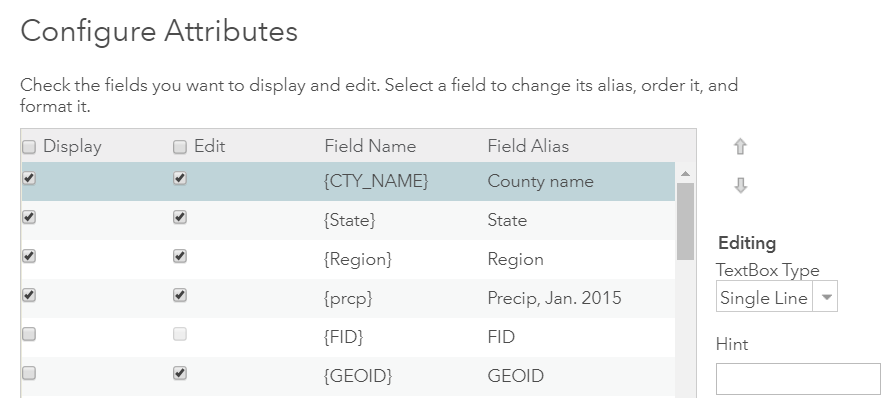
* Break 1: 1st quartile value
* Break 2: Median
* Break 3: 3rd quartile value
* Break 4: Value that defines the upper outliers

The video walkthrough above shows how to adjust these breaks. You should be able to get the values from the analysis you did in Excel earlier in this lab.

**Adding pop-ups**

How will users get more information on each of these counties? One easy solution in creating pop-up windows. ESRI’s guide to this feature is [available here](https://doc.arcgis.com/en/arcgis-online/create-maps/configure-pop-ups.htm). To configure pop-ups, click on the three dot menu next to the Daymet layer on the left sidebar, and choose “Configure pop-up.”

There’s lots of pop-up options, but you want to focus on the field attributes to display--basically which data you want to be visible. You can adjust these by clicking on the “Configure attributes” link. You want just the four variables to be selected under “Display”. Use the arrows on the side to move them to the right order. Click on “Field Alias” to change the label for each variable in the pop-up and use the labels shown below.



**Finishing the map**

Click on the Save button to save the map. Give the map this title: “Monthly precipitation (mm)-Jan. 2015”. This gives the content of the map, the unit of measurement (millimeters) and the date. Use the tag “geog2011\_fa20” and any others you think would be helpful. Then write a one or two sentence description of the map content that gives the variable shown, data source, date, and area for this map.

Once you’re done, share the map using the “Share” button at the top of the page. Make sure you have allowed public access using the “Everyone (public)” checkbox. Then copy the link provided for your map and paste into your Word document.

**Lab deliverables**

*Note: Please save this file as: LastName\_Lab1.doc*

In a Word compatible document, provide the following:

1. (5 points) Copy and paste your pivot table from Excel or include it as a screenshot. Make sure it fits neatly on your page and includes all required information.
2. (10 points) The URL of your map on ArcOnline. This map should include all U.S. counties showing monthly precipitation as described above. Double check your map for the following:
   1. The correct variable
   2. The correct classification for the prcp variable
   3. Pop ups should also be set up as described above
   4. You should check the box under “Share” for public access
3. (3 points) Is the precipitation variable *skewed* at all? How can you tell? Be sure to make specific reference to topics covered in the class lectures.
4. (4 points) Your map adapts a *quantile* classification scheme for these data. Pick one of the alternative classification schemes listed in the Manson, et al. text. In your own words explain how that scheme is different from the one you used from this map.
5. (3 points) What one finding or pattern you identified through this analysis that you found interesting or surprising? Explain your answer using specific evidence from the map and/or data.