GIS2

PROJECTING GEOGRAPHIC COORDINATES ONTO FLAT MAPPING SYSTEMS

BACKGROUND

The distance between two points is a fundamental property of their spatial relationship. We can, for example, use the Euclidean distance between two cities to estimate the cost and time needed to fly between them. Or, we can also use the network distance between two cities to estimate the cost and time needed to drive between them. Because distance matters, how we measure distance matters.

The distance between two points can be calculated with their coordinates (Eq. 1)(Bolstad, 2017: p119, 174, 399, 470, 530, 626, 629, and 687)(Kimmerling et al., 2012: Chapter 11).

$$\boldsymbol{Distance}_{1\to 2} = \sqrt{(\boldsymbol{x}_2 - \boldsymbol{x}_1)^2 + (\boldsymbol{y}_2 - \boldsymbol{y}_1)^2} \qquad \text{Eq. 1}$$

where:

1,2	a counter {from point 1, to point 2}
x	the x or easting value
y	the y or northing value
Distance	the Euclidean distance between two points

The total length of a line composed of segments can be calculated if we know: a) how many vertices define the line; and b) the coordinates of those vertices (Equation 2).

$$\boldsymbol{ShapeLength}_{1\to n} = \sum \sqrt{(\boldsymbol{x}_{i+1} - \boldsymbol{x}_i)^2 + (\boldsymbol{y}_{i+1} - \boldsymbol{y}_i)^2} \qquad \text{Eq. 2}$$

where:

i	a counter {from vertex 1, … , to vertex n}
x	the x or easting value
y	the y or northing value
shapeLength	the calculated distance, or length, of a line

The area of a polygon can be calculated if we know: a) how many vertices define its perimeter; and b) the coordinates of each vertex (Equation 3).

$$\boldsymbol{ShapeArea}_{1\to n\to 1} = \frac{1}{2}\sum\left[\,(\boldsymbol{x}_1 \times \boldsymbol{y}_{i+1}) - (\boldsymbol{x}_{1+1} \times \boldsymbol{y}_i)\,\right] \qquad \text{Eq. 3}$$

where:

i	a counter {from vertex 1, … , to vertex *n*, and return to vertex 1}
x	the x or easting value
y	the y or northing value
ShapeArea	the calculated area of the polygon

One more method

The US Census Bureau produces an official count of the US population every ten years (in the years ending with 0s) and it samples the population monthly to build inter-census estimates. Census estimates and official census counts are aggregated up to the national, state, and county levels for people to use.

Calculating the population density for a county is easy (Equation 4). You need an accurate population value (either a count or estimate) and an accurate area value (usually in sq.mi or sq.km). The US Census Bureau publishes reliable population data, but we have to use GIS correctly to make sure our county areas are just as reliable.

$$\boldsymbol{PopDensity_i = \frac{pop_i}{area_i}} \qquad \text{Eq. 4}$$

where:

i	a counter {county 1, county 2, county 3, … , county n}
pop	the population count or estimate for county i
area	the area of county i (see Eq. 3 for details)
PopDensity	the calculated population density

##############

I'm showing you the methods used to calculate shape lengths (Eq. 2) and shape areas (Eq. 3) because I need you to understand that each method relies on each shape's coordinates and only on each shape's coordinates. No other variables are used to calculate ShapeLength or ShapeArea.

Ideally, when the geographic coordinates of two points are projected onto a flat mapping system, when want their flat map properties to mimic their true geodetic properties (i.e., planar distance = geodetic distance). Unfortunately, it is impossible to preserve the geodetic distance between every pair of points during projection – especially when projecting hundreds or thousands of points scattered over large extents. Some projected points will land farther apart then they should; the distorted lengths and areas we calculate from their projected coordinates will be inflated. Some points will be projected closer together than they should; the lengths and areas we calculate from those projected coordinates will be deflated. Those distortions can propagate as errors in subsequent calculations (e.g., in Eq. 4).

##############

PURPOSE

The purpose of this lab is threefold: a) to help you develop a working knowledge of our NAD83 geographic coordinate system; b) to help you develop a working knowledge of commonly used planar mapping systems; and c) to help you understand how projecting geographic coordinates onto planar mapping systems can create nasty distortions that lurk in your projected data. Take your time and take good notes. Work carefully and ask questions.

OBJECTIVES AND METHODS

Obj. 1. Get familiar with the geospatial data you downloaded
The data you need for this lab is available on the course website.

`https://webspace.ship.edu/sadrzy/geo363/`

For this lab, we're going to use a copy of the Census Bureau's 1:500,000 (500K) scale ***Cartographic Boundary Shapefiles*** for US counties (less AK and HI). I added two text fields to the county polygon attribute table so that you can see which county fits within each zone of the US State Plane mapping system (see Kimmerling et al., 2017; Fig. 4.5 on p67)(see Bolstad, 2016; p684).

We're also going to use two shapefiles that I created, which represent the direct flights for and the airports that are reachable via Harrisburg International Airport.

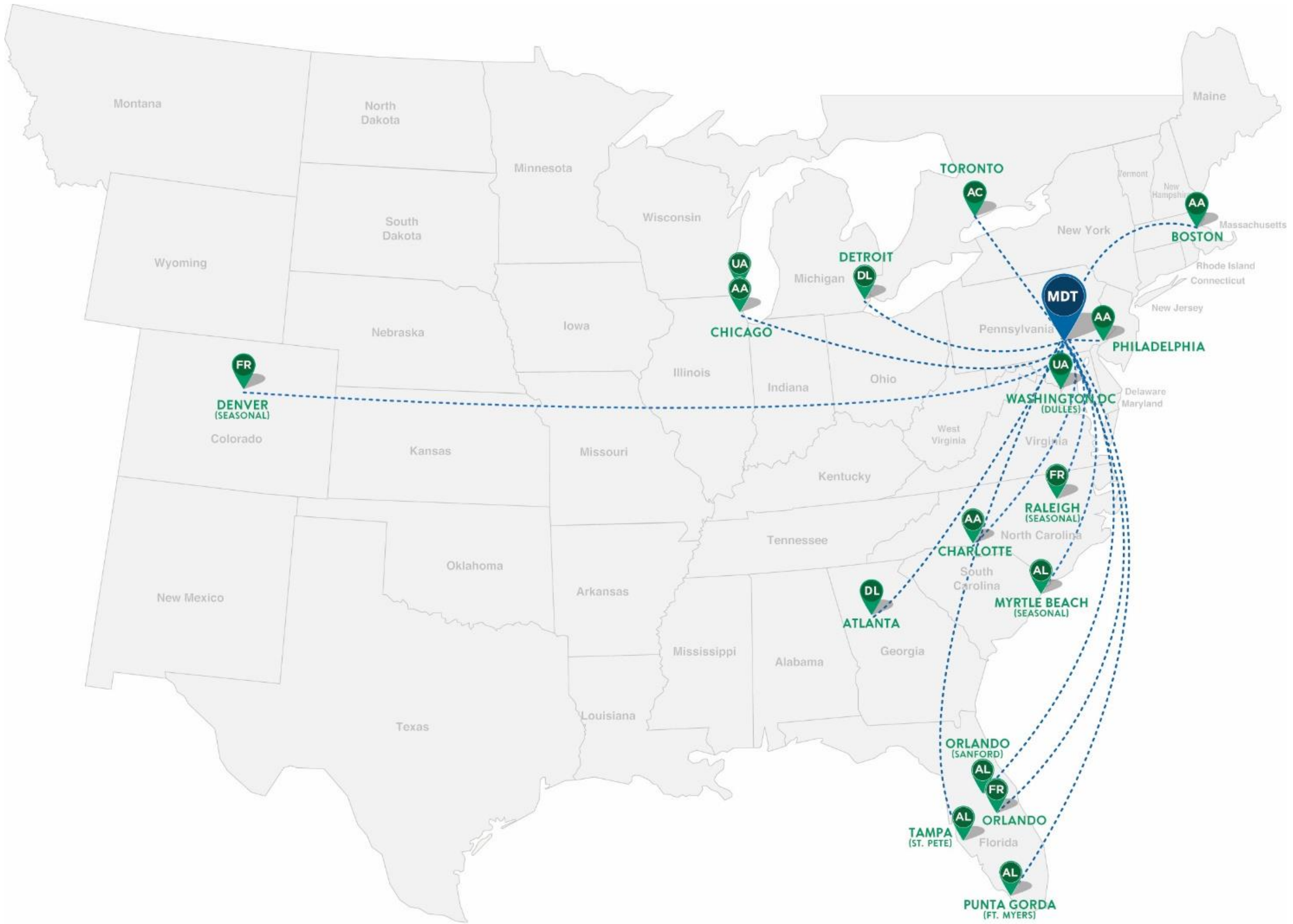


Figure 1. A stylized map of routes from Harrisburg International Airport to fifteen destinations (last updated in July 2018).

Open **ArcCatalog** and use the Catalog Tree to navigate to your workspace.

Select your airports shapefile - `airports.shp` - so that you can:

1. **Preview** your point geography;
2. **Preview** your point attribute table; and
3. read the metadata **Description**.

To reveal all the metadata, use **Customize > ArcCatalog Options...> Metadata** and change the *Metadata Style* from the sparse "Item Description" to "FGDC CSDGM Metadata". Apply your changes, refresh your view, then find out more about the geographic **Extent ▶**, the **Resource Points of Contact ▶**, and the **Spatial Reference ▶** for these data.

Question 1: According to the attribute table, which three airports are the northernmost airports?

Select your non-stop routes shapefile - `directRoutes.shp` - so that you can:

1. **Preview** your line geography;
2. **Preview** your line attribute table; and
3. read the metadata **Description**.

Question 2: According to the metadata, what is the geographic extent covered by these data? Report your answer using both the given DD *and* the DMS format for geographic coordinates.

Select your US county shapefile - `cb_2016_us_county_500k.shp` - so that you can:

1. **Preview** your polygon geography;
2. **Preview** your polygon attribute table; and
3. read the metadata **Description**.

Question 3: According to the metadata, what is the geographic extent covered by your **county** data? Report your answer using both the given DD *and* the DMS format for geographic coordinates.

Question 4: If you had a question about these **county** data, then how would you contact the responsible party by phone or email?

Obj. 2. Convert your data from shapefile format to geodatabase feature class format.

In **ArcCatalog**, **Create a new File Geodatabase** called **`Lab4data_projections.gdb`**.

Export the polygons in your **`cb_2016_us_county_500k.shp`** shapefile to make new polygon features in your geodatabase. Name your output feature class **county**.

In a geodatabase, polygon perimeter (`Shape_Length, Eq.2`) and area (`Shape_Area, Eq.3`) values are calculated automatically for each polygon.

Question 5: Preview your new **county** attribute table. Compare the auto-calculated shape area of Miami-Dade County, FL and the given GEODETIC area of (sq.km).

Next, **Export** the flight lines in your shapefile **`directRoutes.shp`** to make new line features in your geodatabase. Name your output line feature class **connections**.

New shape attributes values will be calculated automatically from the geographic coordinates that define your lines and added to your line attribute table.

Question 6: Preview your new **connections** attribute table. What are the perimeter (Shape_Length) and area (Shape_Area) values for the Harrisburg to Denver connection?

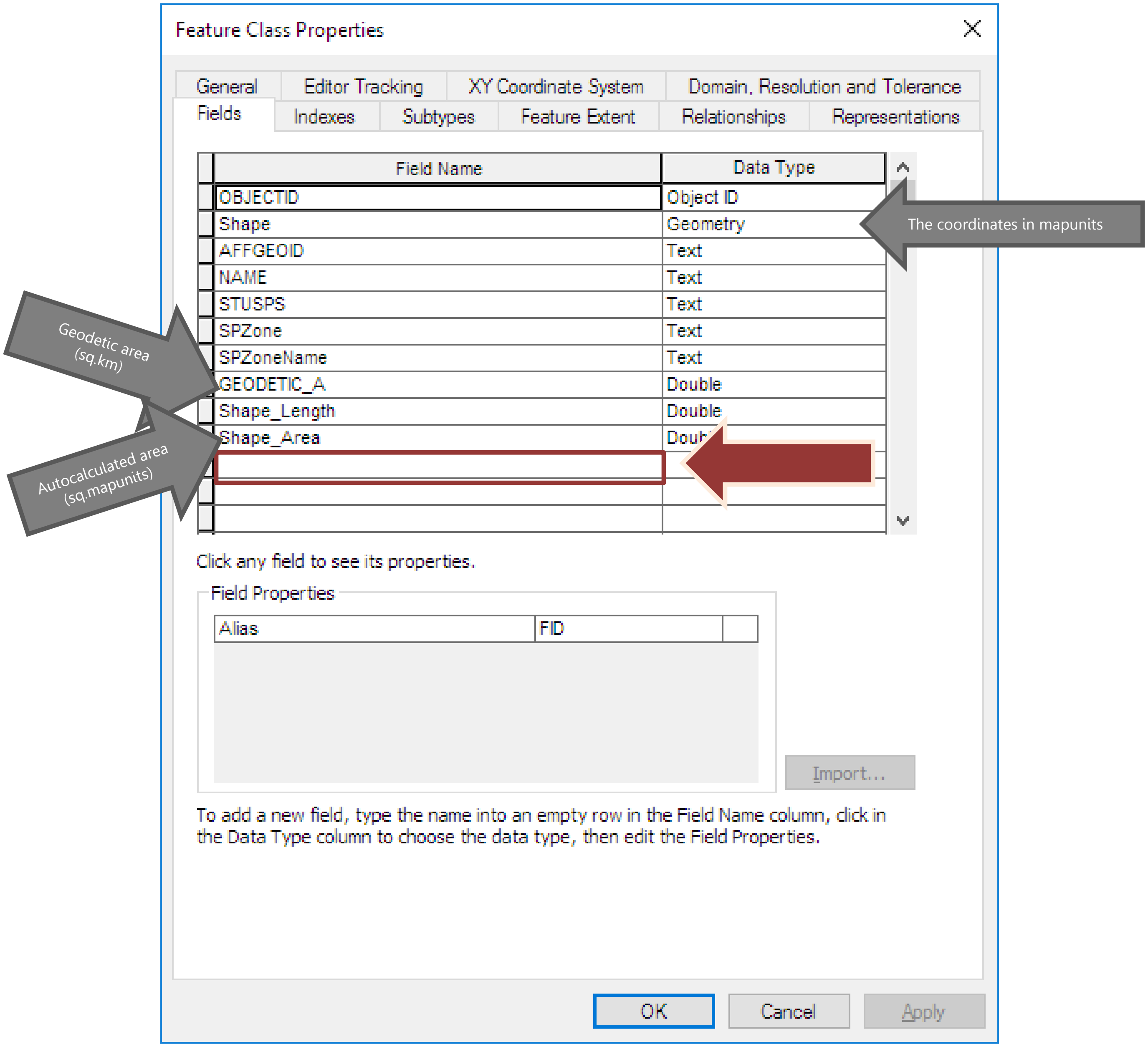
Next, **Export** the airports in your shapefile **`airports.shp`** to make new point features in your geodatabase. Name your output point feature class **airports**.

Obj 3. Prepare your attribute tables to store calculated values.

Continue working in **ArcCatalog**. Do not use ArcMap.

Next, we need to be able to keep track of how the `Shape_Length` and `Shape_Area` values become distorted during projection. So, we need to add some new fields to your attribute table.

In the Catalog Tree, right-click your **county** feature class to access its **Properties…**, then select the **Fields** tab. If you followed all the instructions properly so far, your **Feature Class Properties** should look exactly like Figure 2.



Field Name	Data Type
OBJECTID	Object ID
Shape	Geometry
AFFGEOID	Text
NAME	Text
STUSPS	Text
SPZone	Text
SPZoneName	Text
GEODETIC_A	Double
Shape_Length	Double
Shape_Area	Double

The coordinates in mapunits

Geodetic area (sq.km)

Autocalculated area (sq.mapunits)

Click any field to see its properties.

Field Properties

Alias	FID

To add a new field, type the name into an empty row in the Field Name column, click in the Data Type column to choose the data type, then edit the Field Properties.

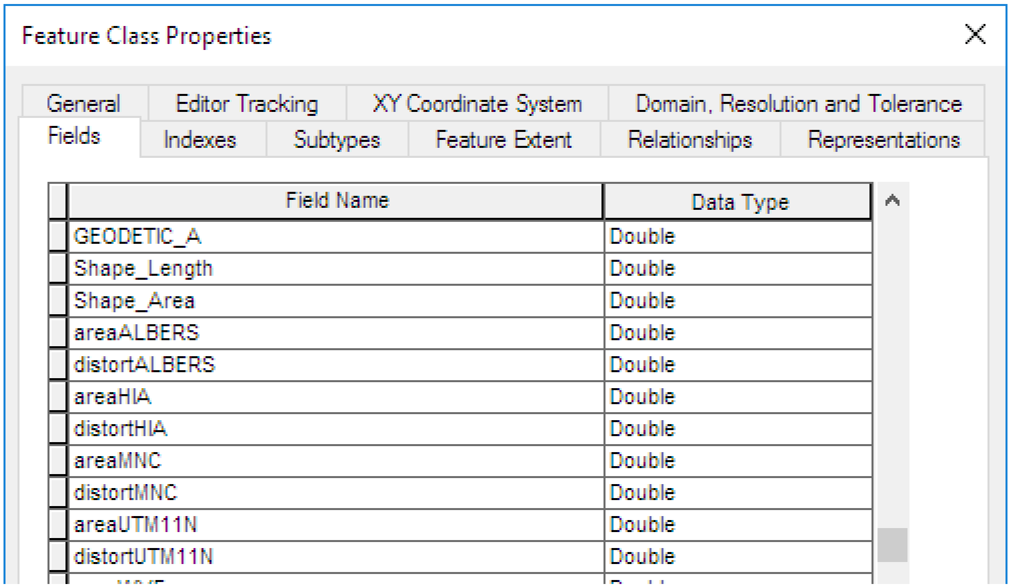
Figure 2. Feature Class Properties dialog box.

Add a new field by selecting the first empty box in the **Field Name** column (see the red arrow in Figure 2) and start typing the word “areaALBERS”. Under the **Data Type** column choose the **Double** data type, which will let the field hold large numbers with decimals. **Apply** your change. Next, add nine more fields (see Table 1).

Table 1. The ten new fields we’re adding to the county polygon attribute table.

Field name	Data type
~~areaALBERS~~	~~double~~
distortALBERS	double
areaHIA	double
distortHIA	double
areaMNC	double
distortMNC	double
areaUTM17N	double
distortUTM17N	double
areaNME	double
distortNME	double

When done, your new Feature Class Properties should look like Figure 3. **Preview** your attribute table again to verify that your new fields were added.



Field Name	Data Type
GEODETIC_A	Double
Shape_Length	Double
Shape_Area	Double
areaALBERS	Double
distortALBERS	Double
areaHIA	Double
distortHIA	Double
areaMNC	Double
distortMNC	Double
areaUTM11N	Double
distortUTM11N	Double

Figure 3. New feature class properties.

In the Catalog Tree, right-click your **connections** feature class to access its **Properties…**; then select the **Fields** tab. Add the ten new fields listed in Table 2.

Table 2. The new fields we're adding to the connections attribute table.

Field name	Data type
`lengthALBERS`	double
`distortALBERS`	double
`lengthHIA`	double
`distortHIA`	double
`lengthMNC`	double
`distortMNC`	double
`lengthUTM17N`	double
`distortUTM17N`	double
`lengthNME`	double
`distortNME`	double

Last, **Preview** your line attribute table again to double-check your work. You should see a lot of new NULL values.

In the Catalog Tree, right-click your **airports** feature class to access its **Properties…**; then select the **Fields** tab. Add the five fields listed in Table 3.

Table 3. The new fields we're adding to the airports attribute table.

Field name	Data type
`northAlbers`	long integer
`northHIA`	long integer
`northMNC`	long integer
`northUTM17N`	long integer
`northNME`	long integer

Last, **Preview** your point attribute table again to double-check your work. You should see a lot of new NULL values.

Congratulations. You just finished preparing your data for analysis. Close **ArcCatalog**.

Obj. 4. Calculate the distortions that are created during map projection.

Use ArcMap to **open** the map document I gave you, then make your geodatabase the **default geodatabase**.

Your map document should open in "data view" mode and looking at the layers in the "US system – Albers Equal Area method" data frame (Figure 4). All the data you see are shown using the software's default rectangular map projection. Notice how the lines of latitude do not converge at the poles like they should. Notice how distorted Antarctica looks because the lines don't converge.

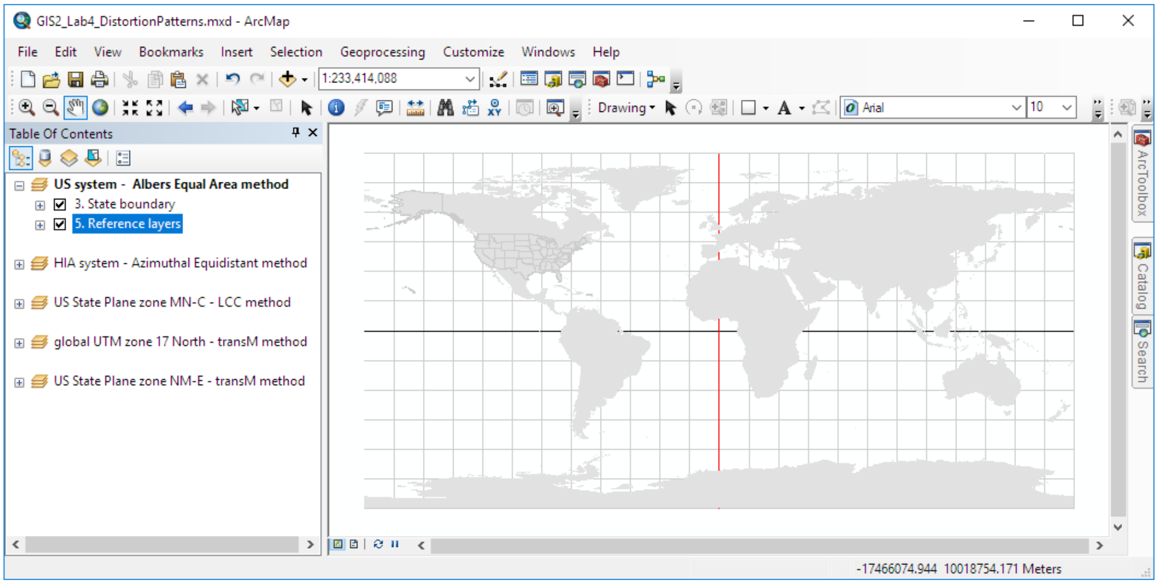


Figure 4: Open map document in data view mode with the ugly default rectangular projection.

Add your airports, connections, and counties as the 1st, 2nd, and 4th layers (Figure5), respectively.

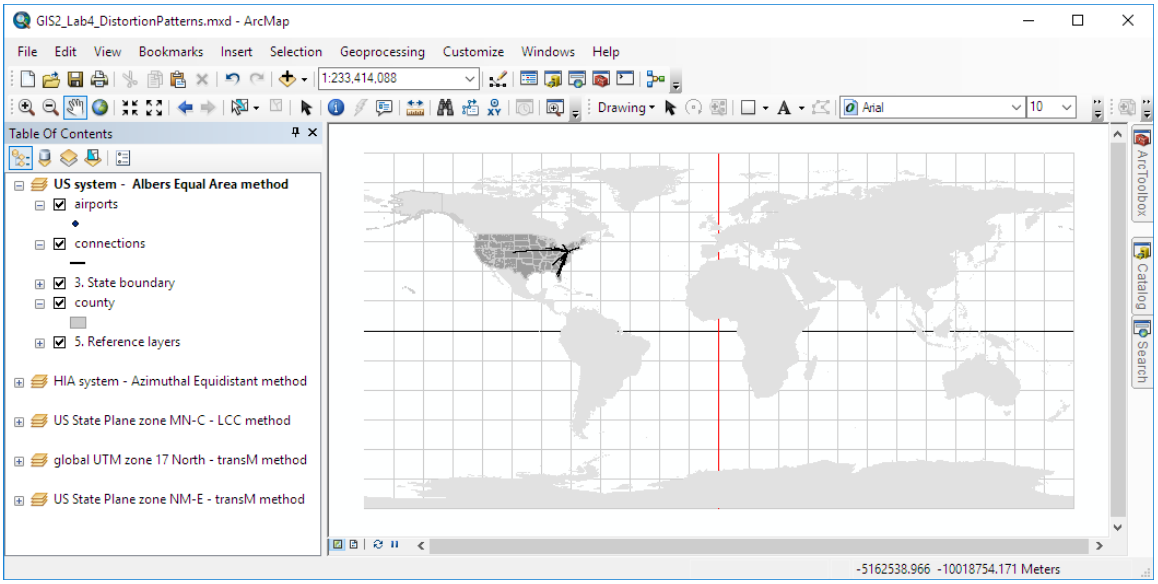


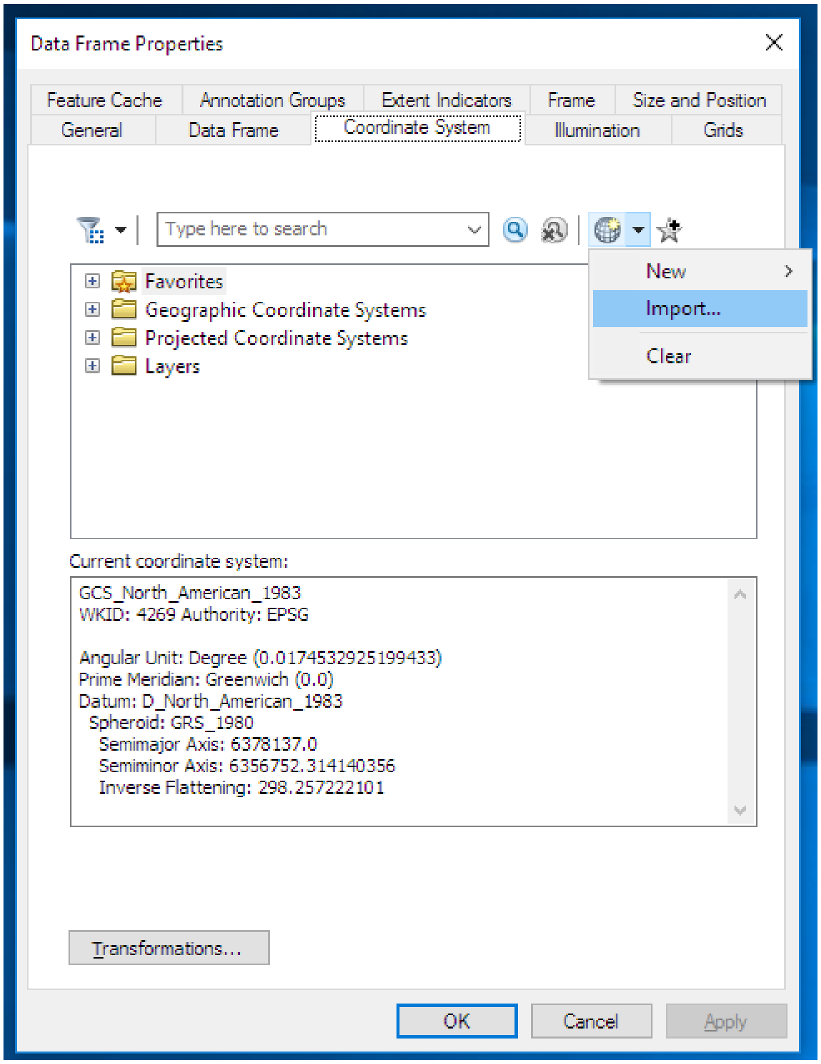
Figure 5. Airports, connections, and counties added as new layers.

Okay, now we're setup and ready to do some analysis.

US mapping system – Albers Equal Area projection method

Right-click your "US system" data frame to access its **Properties…**, then select the **Coordinate System tab**. You'll notice that the *Current coordinate system* is "GCS_North_American_1983", which is the geographic coordinate system used to reference locations on the "GRS_1980" ellipsoid.

You can project all the data in your data frame by changing the *Current coordinate system* of your data frame (Figure 6). **Import…** the spatial reference system that's stored in the projection file named "`NAD83 Contiguous USA Albers Equal Area.prj`". **Apply**ing the Albers projection method will re-project the data onto your screen. This particular instance of the Albers projection method was customized for the contiguous 48 states region.



Current coordinate system:

```
GCS_North_American_1983
WKID: 4269 Authority: EPSG

Angular Unit: Degree (0.0174532925199433)
Prime Meridian: Greenwich (0.0)
Datum: D_North_American_1983
  Spheroid: GRS_1980
    Semimajor Axis: 6378137.0
    Semiminor Axis: 6356752.314140356
    Inverse Flattening: 298.257222101
```

Figure 6. Projecting your data via the Data Frame Properties dialog box.

Next, **open** the **airports** attribute table. **Right-click** the `northALBERS` field to access the **Calculate Geometry** tool.

Calculate the “Y Coordinate of Point” in “meters” using the “coordinate system of the data frame.” You should now see the new northings for these points after they’ve been projected onto this metric mapping system.

Right-click the `northALBERS` field to sort your new northing values in descending order.

Question 7: On the USA mapping system, which three airports were projected as the three northernmost airports?

Next, **close** the **airports** attribute table and **open** the **connections** attribute table.

Right-click the `lengthALBERS` field to access the **Calculate Geometry** tool.

Calculate the “Length” of each feature in “Nautical miles US [nm]” using the “coordinate system of the data frame.” You should now see the new metric lengths of these lines after they been projected onto the US mapping system.

Right-click the `distortALBERS` field to access the **Calculate Field** tool.

Calculate the amount of distance distortion in your flight lines using Equation 5:

`[lengthALBERS] - [GEODETIC_L]` Eq. 5

Right-click the `distortALBERS` field to sort your values in descending order.

Question 8: On this US mapping system, which connection suffered the most inflation? Least distortion? The most deflation? Use your data to support your answer.

Next, **close** the **connections** attribute table and **open** the **county** attribute table. **Right-click** the `areaALBERS` field to access the **Calculate Geometry** tool.

Calculate the “Area” of each feature in “Square kilometers [sq km]” using the “coordinate system of the data frame.” You should now see the new metric areas of these polygons after they’ve been projected onto this US mapping system.

Right-click the `distortALBERS` field to access the **Calculate Field** tool.

Calculate the amount of areal distortion in your projected counties using Equation 6:

`[areaALBERS] - [GEODETIC_A]` Eq. 6

Right-click the `distortALBERS` field to sort your values in descending order.

Question 9: On this USA mapping system, which county area suffered the most inflation? The most deflation? The least distortion? Use your data to support your answer.

Okay – we're done making calculations in this data frame, so you can close your attribute tables. Now we want to visualize the pattern of projected polygon area distortions.

Right-click the **county** layer to access its **Properties…** and select the **Symbology** tab.

[Import…] the symbols stored in the layer file named `countyAreaDistortions.lyr`. **[OK]**

Set the **Value Field** = `areaALBERS` and set the **Normalization Field** = `GEODETIC_A`. **[OK][OK]**

If you did everything right so far, then every county on your map should be symbolized with a light yellow hue.

If everything is yellow, then congratulations!
If not, then go back to the red paper airplane on page 10 and try again.

If you need a break, then this point is a natural break in the lab.

HIA mapping and coordinate system – Azimuthal Equidistant projection method

Activate your “HIA zone” data frame.

Right-click your “HIA zone” data frame to access its **Properties…**, then select the **Coordinate System tab**. You’ll notice that the *Current coordinate system* is “GCS_North_American_1983”, which is used to reference locations on the “GRS_1980” ellipsoid.

You can project all the data in your data frame by changing the *Current coordinate system* of your data frame (Figure 6). **Import…** the spatial reference system that’s stored in the projection file named “`NAD83_Harrisburg_Intl_Airport_metric.prj`”. **Apply**ing the Azimuthal Equidistant projection method will change how your data are projected on screen. This particular instance of the Azimuthal Equidistant projection method was customized for the airport control tower at Harrisburg International Airport.

Next, **open** the **airports** attribute table. **Right-click** the `northingHIA` field to access the **Calculate Geometry** tool.

Calculate the “Y Coordinate of Point” in “meters” using the “coordinate system of the data frame.” You should now see the new metric northings of these points after they’ve been projected onto the HIA mapping system.

Right-click the `northingHIA` field to sort your new northing values in descending order.

Question 10: On the HIA-centered mapping and coordinate system, which three airports were projected as the three northernmost airports?

Next, **close** the **airports** attribute table and **open** the **connections** attribute table.

Right-click the `lengthHIA` field to access the **Calculate Geometry** tool.

Calculate the “Length” of each feature in “Nautical miles US [nm]” using the “coordinate system of the data frame.” You should now see the new metric lengths of these lines after they been projected onto the USA mapping system.

Right-click the `distortHIA` field to access the **Calculate Field** tool.

Calculate the amount of distance distortion in these flight lines using Equation 7:

`[lengthHIA] - [GEODETIC_L]` Eq. 7

Right-click the `distortHIA` field to sort your values in descending order.

Question 11: On the HIA mapping and coordinate system, which connection length suffered the most inflation? The most deflation? Use your data to support your answer.

Next, **close** the **connections** attribute table and **open** the **county** attribute table. **Right-click** the `areaHIA` field to access the **Calculate Geometry** tool.

Calculate the "Area" of each feature in "Square kilometers [sq km]" using the "coordinate system of the data frame." You should now see the new metric areas of these polygons after they've been projected onto the HIA mapping and coordinate system.

Right-click the `distortHIA` field to access the **Calculate Field** tool.

Calculate the amount of areal distortion in these flight lines using Equation 8:

`[areaHIA] - [GEODETIC_A]` Eq. 8

Right-click the `distortHIA` field to sort your values in descending order.

Question 12: On this HIA mapping and coordinate system, which county suffered the most inflation? The most deflation? The least distortion? Use your data to support your answer.

Okay – we're done making calculations in this data frame, so you can close your attribute tables. Now we want to visualize the pattern of county polygon area distortions across the county.

Right-click the **county** layer to access its **Properties…** and select the **Symbology** tab.

[Import…] the symbols stored in the layer file named `countyAreaDistortions.lyr`. **[OK]**

Set the **Value Field** = `areaHIA` and set the **Normalization Field** = `GEODETIC_A`. **[OK][OK]**

If you did everything right so far, then every county on your map should NOT be symbolized with a light yellow color. You should see a projection that preserved geodetic distances from Harrisburg, yet increasingly distorted shape areas with distance from Harrisburg.

If yes, then congratulations! If no, then go back to the paper airplane on page 13 and try again.

State Plane mapping and coordinate system for Minnesota's Central Zone – Lambert's Conformal Conic projection method[1]

Activate your "State Plane zone MN-C" data frame.

If you successfully completed the work in the first two data frames, then you've probably recognized a standard workflow.

a. Project all the data in your data frame.
b. Calculate new geometric attributes based on the current projection method.
c. Calculate distortion attributes to see what got messed up.
d. Symbolize the county polygons to reveal the national pattern.

Use the instructions I provided for the HIA mapping and coordinate system as a guide for processing the data for Minnesota's Central mapping zone. Just be sure to substitute "MNC" every time you see "HIA" in the instructions.

Question 13: On the mapping and coordinate system for Minnesota's Central Zone, which three airports were projected as the three northernmost airports?

Question 14: On the mapping and coordinate system for Minnesota's Central Zone, which connection suffered the most inflation? Least distortion? The most deflation? Use your data to support your answer.

Question 15: On the mapping and coordinate system for Minnesota's Central Zone, which county area suffered the most inflation? The most deflation? The least distortion? Use your data to support your answer.

The mapping and coordinate system for Zone 17N in the global UTM system – the transverse cylindrical Mercator projection method

Activate your "UTM Zone 17N" data frame.

Question 16: On the mapping and coordinate system for UTM zone 17N, which three airports were projected as the three northernmost airports?

Question 17: On the mapping and coordinate system for UTM zone 17N, which connection length suffered the most inflation? The most deflation? Use your data to support your answer.

Question 18: On the mapping and coordinate system for UTM zone 17N, which county area suffered the most inflation? The most deflation? The least distortion? Use your data to support your answer.

---

[1] Why Minnesota's Central mapping zone? Because that's the zone that Bolstad (2016) discusses on pages 126-127.

State Plane mapping and coordinate system for New Mexico's East Zone – a different instance of the transverse cylindrical Mercator projection method but for a smaller zone

Activate your "State Plane zone NM-E" data frame.

Question 19: On the mapping and coordinate system for New Mexico's East Zone, which three airports were projected as the three northernmost airports?

Question 20: On the mapping and coordinate system for New Mexico's East Zone, which connection length suffered the most inflation? The most deflation? Use your data to support your answer.

Question 21: On the mapping and coordinate system for New Mexico's East Zone, which county area suffered the most inflation? The most deflation? The least distortion? Use your data to support your answer.

Final layout and final questions

Change your document **View** from Data View to Layout View. You should see all five data frames, one symbol key (for all five data frames), and a little abstract.

Make it easier to compare apples-to-apples by changing the extent (pan) and map scale (zoom level) of each data frame so that they're all showing the same stuff. You could, for example, right-click each **connections** layer and **Zoom to layer**.

Question 22a: The latest population estimate for Miami-Dade County, FL, is 2,751,796 people.[2]
So, if you need to calculate the population density [Eq. 4] for Miami-Dade County, then which of your five projected Miami-Dade County polygons would give you an accurate result? Which planar mapping systems would lead you to overestimate population density? Underestimate it? Report and explain your results, and use calculated population densities to support your claims.

DELIVERABLES

Given the long length of this lab and the relatively large number of questions, I am not requiring a complete written report for this lab. Instead, I've given you a map layout to print (on legal size paper – 8.5" x 14") and a same-sized worksheet to complete. You should see lines or boxes for each question listed above. Staple the printed layout behind your answer sheet before submitting them.

---

[2] Source: U.S. Census Bureau. 2018. *Population estimates, July 1, 2017*. Last accessed on March 23, 2018 at `https://www.census.gov/quickfacts/fact/table/miamidadecountyflorida/POP060210`