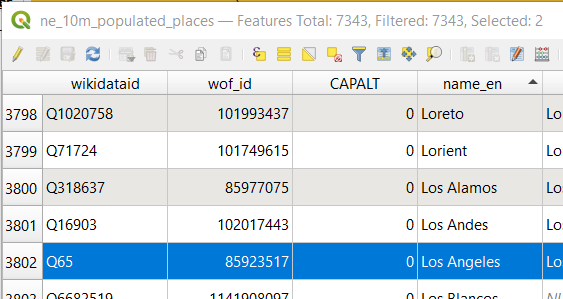
**Part 1: comparing distance and projections**

1. From <https://www.naturalearthdata.com/>, download the Populated Places and States, Provinces datasets. **Get the Data >** Under Large scale data, click **Cultural** > Scroll to **Populated Places** and **States, Provinces.** Note: We downloaded states and provinces last week.
2. Load **States and Provinces** and **Populated Places** shapefiles. Remember; you can load a shapefile by either dragging the file.shp into QGIS, or using the “Add Vector Layer” tool.

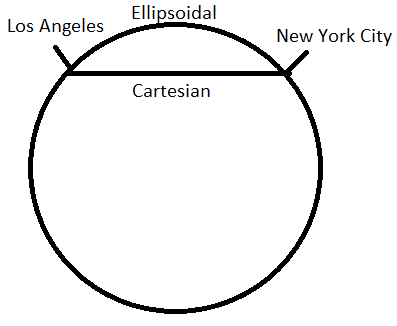
It is important to know that the Populated Places dataset is a layer representing point locality data of cities. Point data are often also loaded as a Delimited Text Layer ().

1. Identify the points that represent New York City and Los Angeles. Save these as a new layer.
   1. In the Table of Contents, right-click on the Populated Places layer and open the Attribute Table. This table has a lot of information about each city, including in many languages. Scroll over to the “name\_en” column. This column shows each city name in English. Click the column head to sort the column alphabetically. Scroll down to “New York City” and click the row label that corresponds with NYC. Then scroll up to “Los Angeles” and do the same. REMEMBER: hold crtl (Windows) or cmd (Mac) before clicking on Los Angeles. Two features should be selected:



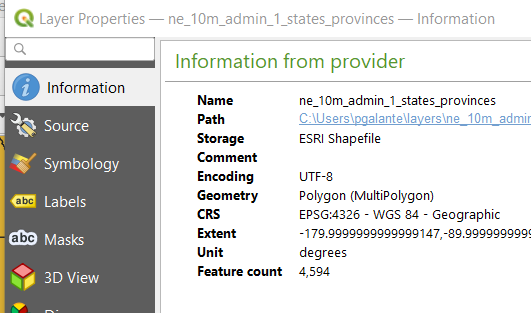
Then, in the Table of Contents, right-click on the layer. Click on **Export > Save Selected Features As…** Save the file as an ESRI shapefile as “TwoCities.shp”. It should automatically load into the Table of Contents. Turn off the Populated Places layer.

1. **Cartesian vs Ellipsoidal distances.**
   1. Calculate the distance between NYC and LA.
      1. Click on the “Measure Line” tool (). This tool allows you to measure straight-line or path distances between mouse-clicks. Make sure the units are set to kilometers.
      2. You’ll notice that there are two radio buttons: Cartesian and Ellipsoidal. Cartesian distance is measured on an assumed flat earth. Ellipsoidal distance is measured on a mostly spherical shape – a mathematical model that better approximates the actual curvature of Earth. Set the radio button to **Cartesian.**

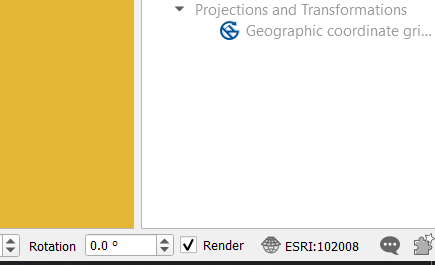
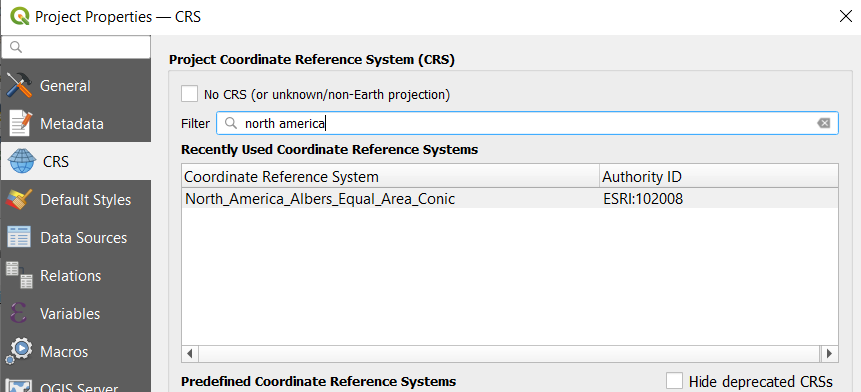


* + 1. First, click on NYC then LA. You should see that a line extends from your first click to your second. Each time you click, the distance between your previous click and the distance of the cursor is calculated. What is the distance between the two cities?

1. This distance is calculated in an *unprojected* map. Double-click on the States and Provinces layer in the Table of Contents and click on the **Information**tab. You’ll see the CRS is WGS84 – Geographic.



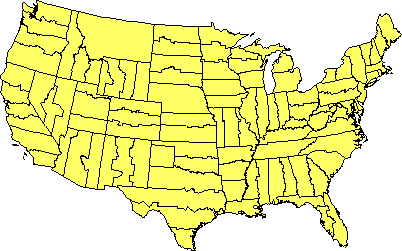
1. Set the Coordinate Reference System for the map to one that is made for North America.
   1. Click on the map CRS settings button and in the new window, navigate to the “filter” box and type “North America”. Then click on “North\_America\_Albers\_Equal\_Area\_Conic”. Click Apply then OK.

1. Take a moment to zoom out to a global view. You’ll notice that this projection represents North America pretty well, but distortion increases as you move away – even to the point that Asia is split. Zoom back in and measure the distance between the two cities. Notice that they are now nearer. NYC is actually about 3,930 km from LA. This exemplifies the great importance of proper projections in your data.

**State Plane Coordinate System**

1. Each state has several, very accurate projections called the State Plane Coordinate System. Each zone has a Cartesian coordinate system with an established origin placed west and south of the zone. All points within the zone are measured in feet from that origin (easting and northing).



1. Using the “Select Features by Area” tool (), create a new layer of just New York State.
   1. Once the state is highlighted, right click on the layer in the Table of Contents and select **Export > Save selected features as…** Name the file NewYorkState.shp and save it in your directory. Make sure that the file format remains “ESRI Shapefile”.
2. Reproject the NewYorkState layer to the state plane coordinate system for Long Island.
   1. In the **Processing Toolbox**, search for “reproject”. Under **Vector general > Reproject layer.**
   2. Make sure that the input layer is NewYorkState. Set the Target CRS to the correct coordinate system. An easy way to find this is to Filter for: “ESRI:103495”. Click run, then close.
3. You’ll notice that the map has not changed. That is because QGIS is transforming the layer to the geodetic coordinate system (unprojected). The *project* projection still needs to be set. Click on the map coordinate system properties button in the lower right of the screen. It should say “EPSG:4326”. Again, filter for “ESRI:103495” to set the map CRS. The preview map pane in this window will show you the area for which this projection was created.
4. All of New York State is now projected in this state plane projection. However, it is only accurate within the Long Island region. Try to find the coordinate system’s origin (where the latitude and longitude measures are both 0).
5. State Plane Systems are measured in feet. How wide is Manhattan? How tall? This coordinate system is useful because map units can be interpreted as more familiar units, making measuring distances on the map more intuitive.

**Part 2: (de)construct a proj4String**

A proj.4 string is a common projection documentation format that we will encounter when we begin GIS in R. Others to know are: EPSG and WKT.

1. European Petroleum Survey Group. We often encounter this shorthand method of documenting the projection in QGIS. EPSG 4326 is the unprojected geographic coordinate system. EPSG 102008 is the North American Albers Equal Area Conic projection.
2. Well-Known-Text: A geometric method for representing how shapes should be drawn. This is perhaps a less-common projection method, especially in Biology. The projection definitions are long and cumbersome.

In this course, we will focus on EPSG and proj4.

The proj4 string is made up of many components. Here, lets dissect a common proj4: UTM zone 11 projection. This type of projection would be idea for a study region near California, Nevada, Idaho, and Alberta, Canada.

+proj=utm +zone=11 +datum=WGS84 +units=m +no\_defs +ellps=WGS84 +towgs84=0,0,0

+proj=utm: the projection is Universal Transverse Marcator

+zone=11: UTM has many zones for both northern and southern hemispheres. Zone 11 represents roughly the California longitude from the equator to the north pole.

+datum=WGS84: the datum is WGS 84

+units=m: the units of measurements are meters

+ellps=WGS84: Specifies that WGS84 is the geodetic datum

+towgs84=0,0,0: describes the datum transformation

Another common string type is:

+proj=longlat +datum=WGS84 +ellps=WGS84

+proj=longlat: this means the data are in a geographic (unprojected) coordinate reference system

+datum=wgs84: As above, using the most common datum

+ellps=WGS84: The ellipsoid on which the projection is based is WGS84

You’ll notice that there are no units associated here. That is because geographic coordinate systems are measured in decimal degrees.

**Your turn:**

1. Using the same data from the previous lab, zoom into your region of interest. Now, find a suitable projection that minimizes distortion. Depending on your region and project, you may be interested in maintaining direction, or you may be interested in preserving area. The type of projection you use will depend on the nature of your project.