Assignment 5 (Assignment 6: New final assignment)

1. Identify and describe three technical limits mentioned in this course that represent controls on map design.
   1. The three technical limits are printer, digital display and internet. When making a map we need to consider what quality the printer can handle (i.e. resolution), the size of the output (determined by paper size) and whether we have coloured ink/toner. With digital display we need to consider how the map is meant to be viewed, whether on a computer, phone or tablet. Finally, we need to consider the limitation of internet speeds, so that our map file is not too large which might take too long to load.
2. Describe three different methods that could be used in ArcMap to select fire hydrants (represented as points) from one vector file within a specified distance of a fire station (also represented as a point) stored in a second vector file.
   1. We can use vector buffers, raster distances, and spatial joins. Vector buffers let you specify a boundary around a point (fire station) to select another point within the area (fire hydrant), and the area can be a circle or various other shapes or polygons. Raster data works similar to vector buffers however it uses Euclidean distances to measure the exact distance for vector data. We can also use a spatial join of points to points which would allow us to see which fire hydrants are closest to certain fire stations as it automatically calculates the distances using Thiessen polygons.
3. You have been given two different data sets that cover the same area, but one uses NAD 27 and the other uses NAD 83.  Explain, in detail, what these terms mean and the significance of this in relation to working with these files in a GIS.
   1. NAD 27 and NAD 83 are datums, which specifies the ellipsoid being used and how it is connected to the Earth. North American Datum 23 (NAD 27) was adopted in 1927 and is commonly used for maps of North America, whereas NAD 83 was adopted in 1983 and is used for worldwide maps. Essentially, these datums take measurements in certain geographic regions to create a customized coordinate system based on the ellipsoid of that specific area. NAD 27 has measurements in the United States whereas NAD 83 has more measurements in more locations and is suitable in most cases. It is important to specify the datum in GIS which specifies the correct ellipsoid, and this affects where the points you create are located on the Earth. When mapping data points we need to make sure they have accurate locations.
4. Is it better to use total or derived values when creating a choropleth map?  Explain your answer.
   1. It is better to use derived values for choropleth maps. This is because we are usually working with area and if we are looking at quantities such as population, we need to make sure that the map is not biased by area. For instance, if the area is larger there may be a lot more people in that area. To fix this issue, we use derived values such as population density which is normalized by area so that we can look at which regions have a higher concentration of individuals. This tells a much more complete story than if we had used total population instead.
5. Explain why it is incorrect to state that 20°C is twice as warm as 10°C, and how this relates to interval and ratio levels of measurement.
   1. This is incorrect because temperature (on the degrees Celsius scale) does not have a natural starting point. Essentially, the difference between 20°C and 10°C is 10°C which is meaningful, but we can’t compare them with a ratio. Temperature is a type of interval measurement, where the differences are meaningful but there is no meaningful zero. On the other hand, ratio levels of measurement do start at a meaningful 0 so for ratio levels of measurement we could say that 20 is twice as large as 10.
6. Explain how you would create a buffer around roads in a map file where the length of the buffer varied according to road type, and what would be the most efficient way to organize your data so that the buffer lengths for each road could be easily changed.
   1. I would create a buffer by attribute (road type) first. Then I would create a new lookup table containing road types and their distances then join it by road type to my original road classification table. This is the most efficient method because if you need to change the buffer lengths for each road you just need to change the few values in the lookup table and join it back to the road classification table again.
7. Describe what is stored in an ArcGIS map document and why it is set up that way, and then briefly compare an ArcGIS layer to a map document.
   1. An ArcGIS map document stores what data to show, where to find the data and how to show the data (source: Documents and Layers video). It is set up this way in order to save space as you only need to the store the data once which can be used by different people with different map documents. This is an efficient storing method because you only need to update that one file rather than doing it for every single map document. An ArcGIS layer stores data for a certain feature that can be used for mapping. It can be points, lines, polygons that represent anything. The layer itself can be saved separately, so if you open a new map document and add the layer, it will not change.
8. Explain the purpose of visual hierarchical organization and provide three ways that visual hierarchy can be expressed graphically, without using colour or components of colour.
   1. The purpose of visual hierarchical organization is to help guide the map reader to the important subject of the map by visually indicating the relative importance of the map components. Generally, we want viewers to focus more on the geographic area of the map which should be central and large. We want viewers to focus less on the scale and legend which are less important. Three ways that visual hierarchy can be expressed graphically are: size, patterns, and shapes. Size is self explanatory, as the bigger the figure the more importance we place on it. Patterns are bit trickier, but you can have certain regions on a map stand out by assigning it a more prominent pattern. As for shapes, we can choose what shape to assign to points on the map, and some shapes pop out more than others (e.g. using a school icon versus a square for schools).
9. Define and compare the terms geoid, ellipsoid and datum.
   1. According to Esri, a geoid is the surface of the Earth’s gravity field; an ellipsoid is a three-dimensional shape that is derived from a two-dimensional ellipse which contains a major axis and minor axis; a datum is something that builds on top of the ellipsoid by incorporating variations in elevation. Essentially, an ellipsoid imitates the geoid for a specific location in the world, and the datum helps to specify which ellipsoid is being used and how it’s connected to the Earth. So we start with the geoid which is the entire Earth, we zoom into a location that is better defined by the ellipsoid, and zooming in further we use a datum that is even more accurate than the ellipsoid.
10. Define the following terms: target table, join table, primary key, foreign key. Explain why you should not use the OID field when joining tables.
    1. The target table is the table that we are joining the data to; the join table is the table that we are joining to the target table; the primary key is the unique identifier for each record in the target table; the foreign key is similar to the primary key in the join table, however it does not need to be unique like the primary key. You should not use the OID field when joining tables because it not a proper primary nor foreign key, but rather it represents the order in which the objects are created. It has no meaningful values that match up with keys in other tables, so if we were to use it to join two tables together the resulting table would have incorrect records.
11. You have been given two feature classes stored in UTM Zone 17, one called **Schools** that contains point locations of schools in Toronto, and one called **Neighbourhoods** that contains polygons representing Toronto neighbourhoods as polygons, explain how you would use ArcMap to calculate the density of schools per neighbourhood.  Neither feature class has any attributes.  
    **Note:** you only need to include the names of tools and/or main steps; you do not need to provide details such as clicking on the Ok button or right-clicking on something.
    1. The first thing to do is to use a spatial join, to join the feature classes based on location. So I would open the Join Data box from the Neighbourhoods attribute table, select “Join data from another layer based on spatial location” and choose the Schools layer to join to this layer, then select “Each point will be given all the attributes of the polygon given that: it falls inside”. According to the ArcGIS site, there will be a new field that is added to the output feature class, called Join\_Count which tells us how many join features match each target feature. This essentially gives us a count of how many schools are inside each neighbourhood. At this point we are done because we have the density of schools per neighbourhood.
12. Define metadata and explain its benefits to both data producers and data users.
    1. Metadata is the information about the dataset. It can include a summary of the dataset, along with a detailed description of its purpose, origin, scale, map projection, attributes, contact information and many more details. Metadata is important to data producers particularly in cases of staff turnover where if a new person comes in to use the dataset, they will need the metadata to understand what the data is about. Metadata also allows for data discovery so that many more people can find it easily for their own use. In the case of data users, they would save a lot of time and money in finding the dataset that they are looking for, especially since data acquisition is time consuming and expensive.
13. Define the terms standard line as well as the different types of case and aspect, as they all relate to map projections and distortion, using the cylindrical class of projections to explain your answer.
    1. A standard line is where a globe touches a developable surface (cone, cylinder, plane) and this line is important because there is no distortion on the projected two dimensional map at that line.  Using a cylinder wrapped around a globe, if we were to unfold the sheet of paper (i.e. project to a 2D map) then the standard line will run through roughly the middle of the page. In terms of project cases there are two main ones: tangent case and secant case. For cylindrical projections, the tangent case means that the cylinder intersects the globe once along a line, resulting in one standard line. In the secant case, the cylinder cuts through the globe at two places so there are two places where there are standard lines. A projection aspect describes where the developable surface touches the globe. There are three common aspects: normal, transverse and oblique. The normal aspect means that the cylinder is upright and so the standard line will occur at the equator. With the transverse aspect, the cylinder is placed horizontal, so the standard line is along the meridian. Finally, for the oblique aspect, the cylinder is placed at an angle so the standard line is also at an angle, perpendicular to the direction of the cylinder.