Congratulations! You are the recipient of a collection of orthophotographs of Wake County, North Carolina! This collection of data is sure to bring your hours of enjoyment.

Silliness aside, this is a good collection to play with. There are a few bits of information you'll need to make any use of them, though. Things like "what are they", "where did they come from", and "how do I find anything?"

Take name, for instance. An "orthophoto" is a photograph, taken from the air, and then "distorted" in such a way that the effects of terrain are cancelled out. You know how everyone in a selfie looks like they have a gigantic head? That's predominantly because the camera is at different distances from their head and the rest of their body. If you warped the resulting selfie in the right ways, you could "un-warp" the distortion from being so close. There is software out there to do that, by the way.

So, if you want to make someone look skinnier, shoot them from a high angle above, but looking straight at the camera lens. They'll look like a bobblehead that hasn't had a good meal in a while. But I digress.

The problem of figuring out where the resulting image is in space is a much harder problem. It's kind like the geometry going on to correct for the terrain, except now you have to correct for the earth being a sphere.

Except the earth isn't really a sphere. It's more like a really, really fat football. Except it isn't symmetrical between the southern and northern hemispheres. So it's more like a chicken egg that's almost but not quite a sphere. And it turns out this matters.

The good thing about standards is that there are so many of them to choose from.

For historical reasons, the mapping function (there's that term again!) that converts spherical coordinates (latitude and longitude) to a point on the earth is called a **datum**. The first mostly useful datum for the US was formalized in 1927 and is called the North American Datum of 1927, or "NAD27" among *connoisseurs*.

NAD27 was basically, mostly, sort of good enough. It was absolutely dead-on at Meade Ranch, Kansas, a spot picked to be close to the center of the lower 48 states. Errors accumulated as you went toward the coasts, but they were managable. A few hundred feet along the west coast, but that's good enough to survey the width of the US to better than one part in 10,000. Not bad for some math that still treated the earth like it was a sphere!

Lots more work would be done over the years, and we fast forward to 1983. The NAD83 standard is created. This one is cool - it models the earth as an ellipsoid (instead of a spheroid) and it uses the earth's center of mass as the control point instead of some God foresaken spot in Kansas.

Upshot - the same lattitude and longitude surveyed by the NAD27 datum can be 80-something meters away from the same point under NAD83. A very readable, if fairly long, account of the state of practice in the US and how it came to be this way is in https://www.ngs.noaa.gov/PUBS_LIB/ManualNOSNGS5.pdf . Enjoy.

All this latitude and longitude stuff isn't very useful to engineers - they just want to put their bridge somewhere and go home. At the scale that bridges exist at, the earth may as well be flat. You can't pour concrete that smoothly anyway. So, a new coordinate system was created - The State Plane.

Actually, there isn't just one State Plane. Every state has one. Some states have a lot of them - Hawaii has 5, Alaska has 11, and even Georgia has 2. Fortunately, North Carolina makes do with just one.

A State Plane coordinate system is nothing more than a way to project a mostly-elipsoid earth onto a flat map. I'm going to skip the whole geometry of how this is done (it's interesting, but not that interesting) and cut to the chase:

State Plane coordinates are expressed as two numbers - the distance east of some legally defined point, and the distance north of the same point. These numbers are called "easting" and "northing".

Remember when I said North Carolina only had one State Plane? I lied. We have a bunch. We have the State Plane expressed in feet, expressed in meters, and both of those expressed in ways that correct for known errors in NAD83. For all practical purposes, NC State Plane Feet and NC State Plane Meters (both based on the regular, default NAD83 datum) are all you'll ever see. And 99% of the time it will be State Plane Feet, because Engineering is done in feet (except when it isn't - Google "space probe crash into mars" if you want a good chuckle some time).

Anyway, back to NC State Plane Feet. The coordinates (0,0) are in downtown Atlanta, Georgia. Yeah, I know. Places in North Carolina as thus expressed as some number of feet east of Atlanta (the easting) and some number of feet north of there. An easting of 2,105,000 feet and a northing of 74,500 feet will put you in Wake County.

Why in the heck does any of this matter? Basically because engineers and surveyors use State Plane coordinates, therefore the county Land Records offices do. It is in fact required by state law in most states, including NC. Normal People ™, on the other hand, prefer to use Lat/Lon coordinates. There is a nice set of tools called "GDAL" - Geospatial Data Abstraction Library. It's a very common way of dealing with these matters. You'll probably be very interested in a program called "gdaltransform". I'll just Santa Claus this one for you - the NC State Plane Feet coordinate system is "EPSG:2264", and you can specify NAD83 as… "NAD83". At least that was easy.

A very short aside on the rest of the earth outside the US - much of the rest of the world uses "World Geodetic System 1984", or just "WGS84". NAD83 and WGS84 usually agree to within an inch inside the Lower 48. WGS is updated at various times so it can seem to wander a fraction of an inch or so, but really it's just errors in NAD83 becoming more apparent. The physics of all this is *really cool*. The earth doesn't just rotate on its axis, but the axis moves, and it wobbles, and the wobble has a wobble of its own, and it doesn't quite rotate around its center of mass because the earth's density isn't all that consistent, and the earth has a delicious chewy center. There are a surprising number of people who measure and deal with all of this, many of whom have a hand in running navigation satellites, and at least a few who used to have business cards from the *coolest name of any organization ever*: The International Earth Rotation Service. I liked to imagine them out in space, working in shifts, keeping the old girl turning over smoothly day after day.

All this gets us to where we can use the orthos now. About time, I say.

In /shared/mapdata , you'll find the directories "disk1" through "disk4". In these directories, you'll find files with names such as this:

C2005_1704_016.sid

The first six characters of the name are pretty simple: the "c" means it's a color image, and the "2005" is the year the image was taken. The underscore is just a separator, of course.

The next four digits are coded in a format used for land records in North Carolina. Let's consider those four digits as W, X, Y, and Z. These are digits used to form easting and northing in NC State Plane Feet.

Step 1: Wake County is about 2 million feet east of Atlanta. So, the easting starts with a 2.
Step 2: The first digit, "W", is the "hundreds of thousands" digit in the easting.
Step 3: The X digit is the "hundreds of thousands" digit in the northing.
Step 4: The Y digit is the "tens of thousands" digit in the easting, and
Step 5: The Z digit is the tens of thousands digit in the northing.

So… "C2005_1704_016" has us at
Easting 2,100,000
Northing 74,000

And we're within 10,000 feet of where we want to be. That's a shade under 2 miles.

Now we have those next three digits - "017". You're going to love this.

The "1704", above, gave us the southwestern corner of a square that is 10,000 feet on a side. That square is divided into sixteen smaller squares, numbered 05 through 20.

05	06	07	08
09	10	11	12
13	14	15	16
17	18	19	20

So - “1704” got us to the bottom left (the southwest) corner of the box and now the “016” is getting us in the square three over and one up (square 16, right?). The squares are 5000 pixels on a side with a 6 inch resolution, so each square is 2500 feet on a side.

At this point, we know where that file represents - the bottom left corner of the image in file “C2005_1704_016.sid” is at:

Easting	2,100,000 + (3*2500) = 2,107,500
Northing	74,000 + 2500 = 76,500

If you’re curious, see:
https://web.archive.org/web/20070209234008/http://www.wakegov.com:80/gis/parcel.htm

And

https://web.archive.org/web/20060508160113/http://www.wakegov.com:80/NR/rdonlyres/A4142110-E56C-42F5-A0C5-74F81E0493BC/1982/lg50100grid.jpg

The .sid files are in an image file format called “MrSID”, and pronounced “Mister Sid.” This is a highly compressed image. It can be uncompressed with a command-line tool called “mrsiddecode”.

(This is actually all one line)
mrsiddecode -inputFile /shared/mapdata/disk1/c2005_0697_007.sid -outputFile chuck.png -outputFormat png

Notice that the directories also contain “.sdw” files with the same naming convention. These are “Mister SID World Files” and they describe the coordinates covered by the files (the “tiles”, in GIS parlance). For information on the format, see http://www.pcigeomatics.com/geomatica-help/references/gdb_r/gdb3N244.html . Be careful when you go through this - the .sdw file tells you where the center of the pixel in the top left corner is. First, this is top left when everything up to this point has been bottom left. Secondly, this is the center of a pixel that represents half a foot, so all the numbers come out “something and a quarter” or “something and three quarters”. Use scrap paper and draw pictures lest you end up as a data point for the *other* class project. :-)