Intro to Mapping and GIS: Module 4, part 2 Geocoding

Hello everybody, welcome to the final week of our class Intro to Mapping and GIS for journalists. In this final week we're going to do a little bit more hands-on stuff with QGIS. We're also going to do some things outside of QGIS. You just got done listening to Darla talking about design principles and population maps and how to avoid doing those. We're going to top cover three topics and four videos here basically.

First one in this video is going to be geocoding and then we're going to talk about simplifying polygons using two different methods, and then the last is going to be sort of things to keep in mind when exporting data outside of QGIS. So the first thing we're going to do is geocoding and in order to geocode we're going to have to go outside of QGIS, but you may be wondering what is geocoding. So let me explain, is a fairly simple process.

So let's say you have a data sheet, let's say you have locations or hospitals or whatever and inside of the spreadsheet is address information and city information, but there's not that handy latitude and longitude information. Remember when we were working with data earlier, we had the latitude and longitude coordinates already in the spreadsheet, which we gave to QGIS to put dots on the map.

Let's say you don't have that latitude and longitude information in your spreadsheet, geocoding is basically taking those addresses for the data points in your spreadsheet and creating latitude and longitude points, adding it to the spreadsheet, that way we can put it into QGIS and actually make dots on a map using that latitude and longitude information.

So if you're missing that, this video is for you. So you need to do geocoding outside of QGIS. So my favorite place to go is Geocodio, oh, I guess that's how you pronounce it, I'm not really sure, but the website is www.geocod.io, so go ahead and type that in, I will make a note, there's many many different websites you can use to geocode data, this is what I prefer.

I did some looking around, you almost always need to have some sort of account. You don't have to pay if you aren't geocoding a lot of data, which is nice, but you do need to set up an account for every website I found.

So for this one, I'm going to go ahead and use this website and in order to use it you're going to have to sign up for an account, but you don't have to pay anything, you don't have to put in a credit card information, you don't have to put any personal information at all, which is nice, which is what I like about this.

So go ahead and sign up for an account and get signed in. I'm already signed in and once you get signed up, and go to your dashboard, it should look something like this, but before we do that, let's actually get some data downloaded.

So the other tab I had open is this Austin, Texas government website. What we're going to do here is we're going to download some hospital data, specifically Public Health locations in the city of Austin, Texas. Excuse me, go ahead and hit YES. So go and type that in and go ahead and search and the first thing you should pop up here is Austin Public Health locations.

So if you want to use a different data set feel free, I use this obviously because I'm in Austin, but you can use a different data set if you want just make sure it's not too long, geocodio does have limits for how much you can how many rows you can geocode for free and make sure obviously the latitude and longitude information is not already in there and that there are addresses in there because you need addresses to give the geocodio to geocode to get the latitude and longitude information.

So this is a perfect data set for that, it's very small, it has the addresses, and it doesn't have the latitude and longitude information already. So that's what we're going to use for this course. So once you get that typed in you'll see this little screen, go ahead and click on the first thing that pops up. in this is the data that we want. You can view it, you can read more about it, there's a lot of information here.

A lot of cities have data websites like this. If you live in the United States data.gov is a hub, but there's other hubs as well and other countries as well. So if you want to find something more local go for it, this is what I'm going to use for this example. So we got the public health locations, go ahead and hit the export button, you're going to want to hit a CSV. Again that's important also if you're finding data, make sure you get a CSV.

And it will download into my folder, so I'm going to go ahead and open it up here and you'll see it is a very simple spreadsheet not long at all, it's got some name on the facility hours website, but the most important part is this street address column which we can use to give to geocodio to figure out the latitude and longitude point.

So let's go ahead and do that now actually, so we've got this downloaded. I'm going to go back and pull up geocodio. Go to my dashboard, go ahead and hit upload list, take a second to load here. Once this pulls up you will not have any spreadsheets that you've geocoded I've used website a few times before so that's why I have things listed here, but you can go ahead and just ignore that.

Go ahead and hit upload list, and then this is a really nice feature of geocodio, you can actually drag your spreadsheet here or you can just copy and paste it. So that's actually what I'm going to do. I'm going to copy and paste it right in here. Make sure you give it a name, so this is "Austin Public Health locations", hit continue, and so this is really important, you're going to tell geocodio what is the street address column? what is the city column? what is the state column? what is the zip code column?

You know, fortunately street address is called street address in the column. If it's something weird or not super obvious, I can be a little annoying, but for this spreadsheet that I downloaded was very obvious, so it's already been selected, and I have already guessed that that's what street address ment and they guessed right.

Same with the zip code, so if they guess wrong or you're using a different sheet, make sure you include the zip code and then if you have the city and state in your spreadsheet make sure that you include those columns as well when you're going through this process.

My spreadsheet actually does not have city or state, but that's fine because it has zip code and has street address, so this will be this will probably be enough to properly geocode. Although if you want to double-check when you're done that's always a good idea. So go ahead and hit preview, this is actually one of my favorite parts of the website.

You can actually preview before you go through the whole process of geocoding and all that. You can actually see where these little dots show up on the map. So if you're like working with Austin data, and you see like, you know a dot over in India, then obviously something screwy happened.

But it looks like everything worked out pretty well here. Everything is within the Austin area and then you can also click on the dots and check out exactly what row that is in the spreadsheet as well as the latitude and longitude that it's going to split out for that. So this looks all good to me, yes, it looks good, thank you. I don't need anything else, go ahead and hit continue, and then finalize going to start.

And this will take a few seconds maybe a few minutes depending on how much data you have, so just be patient. Okay, when that's done you might want to refresh the page like I did just to double-check every few minutes or few seconds depending on how big your spreadsheet is, but mine's a pretty small spreadsheet, so it didn't take very long. So it's completed, so I can hit this little button to the right which is download and this is going to be my Austin Public Health location spreadsheet with latitude and longitude information.

And some of this other stuff is related to more information about how it geocoded, what guesses it made, and you can go ahead and ignore all that for now and just focus on that latitude and longitude columns for now. I'm going to close out of my original spreadsheet as well, I'm not going to save this, and I'm going to go ahead and change the name of this a little bit too, and instead of that really long string at the end, I'm just going to do "Austin Public Health".

Okay, now we've got our locations geocoded in that spreadsheet. We're going to go through QGIS, and put them on the map just like we've done before, very similar process. So go ahead and have QGIS opened up, and let's go ahead and open up that same country shapefile that we have used now for the third module on a row, go ahead and open that up just to kind of get oriented on the map, and then once you have that, go ahead and go to "layer", "add layer" and

again, we're going to go to "add delimited text layer", we're gonna go searching around for that file that we just created, and that is actually in my downloads folder, open it and we're going to make sure that we tell it has points coordinates, we have that newly-created longitude and longitude field, which is great.

And once we input that we can just hit "add", go ahead and hit "close" and we're going to zoom in on Austin, Texas and you will see that we have dots on the map, so that is great. So if you have a data set and you have addresses, and you don't have latitude and longitude, don't worry, you can still get them on a map in QGIS if you geocode them first.

So we just walk through that, that's it for this video. Next steps are simplifying shapes, which will explain why that's important and what that is and the next two videos and then after that we will kind of go through some exporting options as well. So that's it, thanks for tuning in.