## GIS 5571 Lab 1

**Due:** 2 weeks from the date of assignment

**Goals**

1. Practice decomposing interfaces for spatial web API’s into informal conceptual models
2. Compare and contract different web API’s using informal conceptual models and custom-built extract, transform, and load (ETL) routines
3. Build an ETL pipeline with Open Source Tools in Esri’s Online and ArcPro Jupyter Notebook and integrate a two datasets via spatial join

**Deliverables**

Submit a lab report on Canvas as a PDF (see [report form](https://docs.google.com/document/u/0/d/1gOGBtTe3dQzrXCEMl644QIVdJgMp8ahN/?rtpof=true&usp=drive_fs)). Include all your code on Github.

**Specifics**

For this lab:

1. Write a lab report that does two things:
   1. Compare and contrast the conceptual models for the following API’s
      1. [Minnesota Geospatial Commons](https://gisdata.mn.gov/content/?q=help/api)
      2. [Google Places](https://developers.google.com/places/web-service/overview)
      3. [NDAWN](https://ndawn.ndsu.nodak.edu//)
   2. Create Jupyter notebooks that can programmatically get data from each of these APIs. Using Jupyter notebooks, build a pipeline that
      1. downloads two data sets,
      2. transform both datasets to the same [coordinate reference system](https://pro.arcgis.com/en/pro-app/latest/help/mapping/properties/coordinate-systems-and-projections.htm) (geographic and projected),
      3. spatially joins them,
      4. prints to screen the head of the table showing the merged attributes, and
      5. saves the integrated dataset to a geodatabase.
2. Make all code available on Github in your Lab 1 folder.

A few tips:

1. Before writing any code, start by using paper and pencil to unpack the dataset objects.
2. Look at other examples of how people designed ETL code.
   1. Towards Data Science [article](https://towardsdatascience.com/integrate-jupyter-into-your-data-pipeline-9a02fab3cee5) on ETL with CRON or Jupyter
      1. Google terms you don’t understand (there are a lot of resources)