TITLE OF PROJECT

DATA PREPARATION Vignette

Name: Joe Blow

Date: March 24, 2013

Study Site(s): e.g. Reach Name(s)

Class: WATS 4930 or 6921 (delete one)

INSTRUCTIONS – Delete these… This is a template for a data preparation vignette. The vignette is intended to summarize everything you did to get your data organized and prepared for analyses. Some of the figures you produce here may end up in your final manuscript, but much of the detail here will not be necessary or need to be condensed for your manuscript (depends on specific journal).

# PROBLEM & Data NEEDS

In as little as 1-2 sentences, but no more then 1-2 paragraphs, reiterate the problem from your proposal and describe the data needs associated with your project. For example, if you need raster inputs of mean annual precipitation to run a model, identify that you will need ‘Mean Annual Precipitation’ but do not digress here into details (save for below).

# DATA SCOPE & Geographic Application

Provide a brief description of the spatial extents of where you plan on conducting your study. A common problem that students have for this course is being too ambitious in the spatial scope of their projects. A simple way to fix this is to treat this project as a pilot study or proof of concept for your original more ambitious idea. For example, do you need to run your model for the entire state of Utah, or could you run it for the Logan River Watershed and that would be sufficient?

Please describe the basic location and characteristics of where you are conducting your model. Include at least one Project Location Map and site vicinity map (if appropriate). Denote on the map the spatial extent that will be necessary for all your inputs.

## Coordinate System

Define the coordinate system you will use for your project inclusive of horizontal projection, vertical datum and units. This should be the coordinate system you use for all your data frames in your map documents. You may also wish to specify whether it is critical that all input data is transformed into this coordinate system.

## Data Extents

Describe (with reference to your location map if necessary) the extents of your data in the coordinate system you defined above. If necessary, specify whether a data-extent mask is also necessary within those data extents (e.g. a watershed boundary):

|  |  |  |  |
| --- | --- | --- | --- |
| Left | Top | Right | Bottom |
|  |  |  |  |

# GEOPROCESSING INPUTS

Your project is required to include some form of geoprocessing or spatial modeling. All geoprocessing requires inputs. Using whatever format you feel is most effective (e.g. table, subsections or bulleted list), please go through each of the inputs you will need to conduct your analyses (include spatial and non-spatial data) and describe:

* What the input is
* What the data type you need is (if spatial is it raster or vector, is it continuous or categorical, integers or floating points, etc.)
* What attributes need to be included in it (if any)
* What units (if applicable) the data values need to be in or if categorical data what classification system it needs).

# DATA SOURCES

Please pull together a list or table of where you will get every input. For each input, explain if the data is available:

* In the exact format you need it
* For the exact spatial extents you need

For every data source, provide a full citation. For unpublished or non-publicly available data, describe concisely who the data is from, how it was collected or produced, and any other relevant details.

# Preliminary Testing of Geoprocessing Steps

I highly recommend you attempt some sloppy practice runs of the geoprocessing steps you intend to run with a very small subset of your data to make sure you know exactly what the input requirements our *really* need are prior to implementing them all. Many of the data conversion and pre-processing data steps can be very time consuming. Make sure you really need to do them and make sure they will work prior to wasting a ton of time cleaning or prepping something you don’t need. Use that testing to help you define and refine exactly what your data needs are. Please report your methods and findings briefly here. For example, if I was going to interpolate a cost surface, I might try the method first with a subset (if large) of the raw input data, to get a better understanding for how it will work.

# DATA CONVERSION & PRE-PROCESSING

Describe here any and all data conversion or pre-processing you completed or needed to complete to get your data sources translated into the exact format, extent and types you need to make them inputs to your geoprocessing. For example, you may consider the following questions (partial list… you don’t necessarily need to answer all these):

* Are data are available as raw points and you need to interpolate them to surfaces before using them?
* If data are in the wrong projection, do they need to be transformed to a new coordinate system?
* Do data need to be combined from multiple tiles or layers into one layer with a common extent (do they need to be orthogonal or concurrent if rasters)?
* Do data need to be clipped down to an area of interest?
* Do raster data need to be resampled or reinterpolated from raw?

Please describe concisely the methods and/or geoprocessing tasks you will perform to do these conversions and/or pre-processing steps.

# File Management & Naming conventions

A major stumbling block for many students is simply getting organized and making sure you know exactly where your data is. Please outline your naming standards and where you will keep your project data. You may consider:

* Having a ‘Raw’ data folder where all the original data is left untouched.
* A ‘Context’ folder for commonly used context layers (e.g. hillshades) that are really just for cartographic context.
* Different folders for different analyses or steps in your analyses
* A ‘Working’ folder where draft versions of analyses and attempts of analyses get saved temporarily.
* A ‘Final’ folder for your final inputs and outputs.

Please refer back to the [File Management](http://gis.joewheaton.org/assignments/labs/lab01/getting-organized-and-oriented/file-management) guidelines in [Lab 1](http://gis.joewheaton.org/assignments/labs/lab01) for more tips.

# Cartographic Standards

For any of your input data that it will be necessary to produce maps and figures of, define some basic cartographic standards that you stick to consistently throughout your project. These are basically the symbology settings you will employ and save as style files and apply to all data of the similar types. You can do this concisely with words or example map/figures if you wish. If appropriate, please also give some thought to how your outputs should be displayed.

# Unanticipated Problems

Without feeling an obligation to robustly back everything up with lots of citations, just highlight your main interpretations from doing this work. What does it tell you? What does it not tell you? Did it work? Feel free to include conjecture that is helpful for exploring new ideas. This is an opportunity for you to articulate your preliminary findings and get feedback from your supervisor and/or committee on some work.

# Questions & additional data PREPARATION needs

If everything went smoothly this week for you, congratulations. Typically however, there are many small problems that arise and you may list these as items you need to address or need help on.

# References

Use EndNote if possible and include all references.