Draft Proposal Guidelines

You should use the following outline for your proposal, which shouldn't need more than about one page.  The proposal draft (which will be reviewed by your peers; see below) is worth up to 10 points; the revised proposal (which will be reviewed by the course instructors; also see below) is also worth up to 10 points.

1. Summarize your proposed project in a few sentences.

* What is your proposed project and why are you proposing it?
* What are the question(s) you want to answer, or goal to achieve?

# Project Summary

## Summary

### Our proposed project is to use Argo data (a data set of ocean floats that collect variables such as salinity and temperature in oceans around the world) and map the change of salinity and temperature in the Estuary and Gulf of St. Lawrence, and see if there is a correlation to fish populations in the same area.

## Questions

### How has ocean temperatures and salinity changed over a period of ten years in the Estuary and Gulf of St. Lawrence?

### How has the fish populations (pelagic primarily, but also possibly demersal fish populations as well) changed in the Estuary and Gulf of St. Lawrence in the same time period?

### Is there a meaningful correlation between temperatures/salinity/other factors and population to further study?

2. Describe two different data sources you plan to access, manipulate and bring together.  The data sources must require different access mechanisms and/or use different data formats. (For example, you might pick one data source that uses a Web API that returns JSON, and the second might use SQL to query a database, or fetch and parse an HTML page.)

For each data set, you should summarize these properties:

* Name
* Short description (i.e., 1-3 sentences)
* Size (in records and/or bytes)
* Location (give the URL or other access method)
* Format
* Access method

# Data Sets

## Data Set 1

### Name:

### Short description

### Size

### Location: <ftp://usgodae.org/pub/outgoing/argo>

### Format:

### Access method: FTP

## Data Set 2

### Name: Pelagic fish species abondance in the Estuary and Gulf of St. Lawrence between 2009 and 2018

### Short Description:

### Size: Roughly 1,800 lines of data

### Location: <https://open.canada.ca/data/en/dataset/f1fc359c-0ed1-4045-a421-adef2497b68d>

### Format: csv/REST API

### Access method: Likely just download it via csv and import to relational database via pymysql

3. Describe with 1-3 sentences for each point below what data manipulation is likely to be needed:

* What initial processing will have to be done on each?
* How will you combine the datasets, and what will be produced as output?
* What new information will result from combining them?

# Data Manipulation Needed

## Initial processing

4. Describe in 1-3 sentences one interesting visualization you could include in a final presentation and report that would show the value/answer a particular question in your final output dataset (that would not be possible with either of the original datasets alone).

5. Indicate the contribution that each team member will make to the project.

