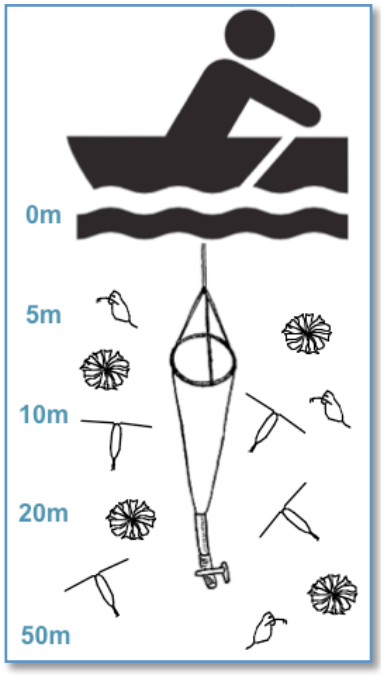
Instructions:



Background: Plankton are microscopic organisms that form the base of many aquatic food webs – fueling the growth of fish and other larger organisms. It's common to sample them using a net or another container that can be controlled to collect water just from certain depths; so you can see how plankton collected at the surface (0 meters) might be different from plankton at another depth (e.g. 10 meters below the surface).
(For more information:

http://en.wikipedia.org/wiki/Phytoplankton and

http://en.wikipedia.org/wiki/Zooplankton.)

They are identified and counted under a microscope, and usually their numbers are reported as individuals per liter or milliliter.

Frequently, aquatic scientists collect plankton samples during both day (e.g. noon) and night (e.g. 2am) because plankton change their distributions from day to night, and not all species alter their distributions in the same way. (For more information, search "diel vertical migration" on the web.)

You should have a fictional data file: pond2010.csv.

The data is part of a long-term study where the investigators wanted to examine the day-night distribution of 2 species of zooplankton across multiple years. Ther are accompanying data sets that you do not have access to. The type of zooplankton they studied is called rotifers generally, and specifically the genus *Conochilus,* in which groups of individual rotifers stick together in colonies (see http://eol.org/pages/43393/overview).

This exercise is to write metadata that describes the data set.

Activity 1

Review the data and the material in this handout to complete the grey boxes in the metadata table overleaf.

- Write an appropriate title for this data set.
- Enter the time period the data was collected.
- Identify some appropriate theme keywords for this dataset.

Activity 2

Upload this example dataset to the development instance of the Knowledge Network for Biocomplexity (KNB) and write the accompanying metadata. Instructions will be presented in class.

Pond2010 Metadata

This is some (fictional) information about the (fictional) data set called pond2010.xlsx. The data set can be used to fill in metadata fields in a formal record, such as the one below, but note that there may also be additional important metadata within the pond2010 file.

Title of the Data set	
Originator/Dataset Author	Anna Sassin Dan D. Lyons
Abstract	This dataset is one of a collection of four population survey datasets documenting colony growth, reproduction, and survival of two rotifer species (*Conochilus unicornis* and *Conochilus hippocrepis*) at four time periods of the year. This dataset describes population data for the summer season. Samples of both species were taken at Littlevick pond, Surrey, UK. Measurements taken include depth, temperature, colony density and colony diameter.
Purpose	Data were collected to evaluate how temperature and depth affect the survival of rotifer colonies in ponds within the UK.
Publication	*Publisher:* International Rotifer Recovery Science Center *Place:* Surrey, UK *Publication_Date:* 12/08/2012 *Series Name:* Four Season Rotifer Survey *Name of Issue*: Summer Survey
Larger_Work_Citation	*Originator:* Sassin, Anna and Lyons, Dan .D. *Publication_Date:* 12/08/2012 *Title:* Relationships between population and temperature: Tracking rotifers over the course of four seasons in the United Kingdom. *Publisher:* Rotifer Conservation *Place:* UK *Volume;Issue;Pages:* 4(2): 325-340
Time Period of Content	Begin Date: End Date:
CurrentnessReference	Ground Condition
Progress/status:	Complete
Maintenance_and_Update_Frequency	None planned
Geographic coverage	Littlevick Pond Natural Reserve, Surrey, UK.

Bounding_Coordinates:	*West_Bounding_Coordinate:* -0.92456818028327 *East_Bounding_Coordinate:* 0.371818538415 *North_Bounding_Coordinate:* 51.511581803063 *South_Bounding_Coordinate:* 50.808817656094
Keywords (theme)	
Keywords (place)	Surrey UK International Littlevick Pond Natural Reserve
Keywords (temporal)	summer, June
Data Access_Constraints	No legal or policy restriction for accessing this dataset.
Data Use_Constraints:	Must properly cite originator if used in publications, reports, presentations, etc. Please cite data set according to DataCite.org standards
Contact_Person_Primary:	*Contact_Person:* Tad Pohl (Data steward) *Contact_Organization:* International Rotifer Recovery Science Center *Address:* 5638 Independence Way *City:* Guildford *State_or_Province:* Surrey, UK *Contact_Telephone:* +44 (0) 888-8888
Data_Set_Credit	Funding was provided by International Rotifer Foundation
Analytical_Tools	SAS, R, MatLab
Data_Quality_Information	
Attribute_Accuracy_Report	Temperature instrument was tested and calibrated for accuracy before each sampling. Density and colony counts were conducted according to the Standard Plate Count procedure. Counts were conducted by two data counters. Each technicians count was verified by the second technician. Counting accuracy was found to be 95% accurate.
Completeness_Report	The data set is generally complete although the temperature for one sample depth could not be recorded due to instrument malfunction. Colony and density counts are also mostly complete except for two instances where the data is missing and is therefore unknown. Statistical summary (boxplot) of the data was performed and no outstanding outliers or potentially erroneous values were found.
Positional_Accuracy:	Positional Accuracy was not assessed
Process_Step: *Process_Description:*	Data was collected by 2 people the first week and by the same 2 people the following week. Water samples and temperature were

	taken at five different depths. In order to account for variability in sample measurements, 6 water samples were taken at each depth. These 6 samples were later randomly divided into two even groups of three. The two groups were randomly assigned a rotifer species name whereby data counters would perform the density and colony counts for the particular species.
Entity and Attribute Information	
Detailed_Description *Entity_Type*	*Entity_Type_Label:* pond2010.xlsx *Entity_Type_Definition:* Rotifer population survey at various depths and temperature
Attribute	*Attribute_Label:* z *Attribute_Definition:* Depth in centimeters from the surface *Attribute_Domain_Values:* *Enumerated_Domain:* *Enumerated_Domain_Value:* 0.5 *Enumerated_Domain_Value_Definition:* 0.5 cm below surface *Enumerated_Domain_Value:* 5 *Enumerated_Domain_Value_Definition:* 5 cm below surface *Enumerated_Domain_Value:* 10 *Enumerated_Domain_Value_Definition:* 10 cm below surface *Enumerated_Domain_Value:* 25 *Enumerated_Domain_Value_Definition:* 25 cm below surface *Enumerated_Domain_Value:* 50 *Enumerated_Domain_Value_Definition:* 50 cm below surface
Attribute	*Attribute_Label:* Temperature *Attribute_Definition:* Temperature of water in Celsius *Attribute_Domain_Values:* *Unrepresentable_Domain*
Attribute	*Attribute_Label:* Density *Attribute_Definition:* Number of individuals per colony *Attribute_Domain_Values:* *Unrepresentable_Domain*
Attribute	*Attribute_Label:* Colony Diameter *Attribute_Definition:* Length of longest colony diameter in millimeters *Attribute_Domain_Values:* *Unrepresentable_Domain*
Attribute	*Attribute_Label:* Species *Attribute_Definition:* Rotifer species *Attribute_Domain_Values:* *Enumerated_Domain_Value:* cuni *Enumerated_Domain_Value_Definition:* Conochilus unicornis

	Enumerated_Domain_Value: chippo *Enumerated_Domain_Value_Definition:* Conochilus hippocrepis
Distribution Information	
Distributor *Contact_Information* *Contact_Organization_Primary*	*Contact_Organization:* Rotifer Network for Biocomplexity (RNB) *Contact_Person:* Metadata Coordinator *Address:* 6534 Biodata Way *City:* Novel Jersey *State_or_Province:* New Jersey *Postal_Code:* 97564 *Contact_Voice_Telephone:* 555-555-1034 *Contact_Email:* info@rnb.net
Distribution_Liability	The Rotifer Network for Biocomplexity (RNB) shall not be held liable for improper or incorrect use of the data described and/or contained herein. It is the responsibility of the data user to use the data appropriately and consistent within the limitations of the data.