Codebook

Data Set Name	data.clean
Column count	12
Row count	2,371
Updated date	2/26/2024

Overview of Data:

From 1989-2019, monitors with the U.S. Fish and Wildlife Service took samples at Back Bay National Wildlife Refuge in Virginia Beach, VA, to ensure water quality standards were being met. They were concerned that allowing water quality parameters to "remain at undesirable levels for a prolonged time frame will result in mortalities to aquatic vegetation and organisms".

Sources and Methodology:

The government department responsible, recruited volunteers to assist in this monitoring. For twenty years, volunteers diligently collected water samples once every two weeks at several points in the bay and in holding ponds connected to it. That collection of data points can be found at https://ecos.fws.gov/ServCat/Reference/Profile/117348 .

Note on Missing Values:

This information, being gathered by volunteers, contained missing values due to poor inputs. Function and character variables (site.id and volunteer.name, read.date and time, respectively) have been marked with "Not Recorded" in instances of a blank entry. Missing numerical variable information (all others) are marked by N/A.

Variable Descriptions and Characteristics are presented on the next page.

Variables Presented in Order of File Contents:

Variable Name: record.id
Variable Type: Character
Description: Record number

Variable Name: site.id
Variable Type: Factor
Description: Site of bodies of water in the Back Bay National Wildlife Refuge in Virginia Beach, Virgina from which volunteers are taking samples from.

Code	Full Name	Count
A	A-Pool	434
B	B-Pool	437
Bay	Back Bay	794
C	C-Pool	264
D	D-Pool (fishing pond)	441
Not Recorded		1

Variable Name: read.date
Variable Type: Character
Description: Date of water testing.

Variable Name: time
Variable Type: Character
Description: Time of day when water was being tested

Variable Name: volunteer.name
Variable Type: Factor
Description: Name of volunteer conducting the testing. Volunteers occasionally went out together, resulting in multiple recordings in these instances, which were separately assigned in the following counts.

Volunteer Names	Count
Phillips, John	54
Feldman, Mary	207
Pease	66
Poe, Susan	800
Strader	92
Not Recorded	1,264

Variable Name: salinity
Variable Type: Numerical
Description: Salinity is the dissolved salt content of a body of water measured by parts per thousand.

Summary of Variable:

Summary	Value
Minimum	0.00
25th Percentile	0.00
Mean	0.71
Median	0.00
75th Percentile	1.00
Maximum	9.00

Variable Name: water.depth
Variable Type: Numerical
Description: The depth of water measured by fathoms.

Summary of Variable:

Summary	Value
Minimum	0.01
25th Percentile	0.40
Mean	0.76
Median	0.65
75th Percentile	0.95
Maximum	12.00

Variable Name: water.temp
Variable Type: Numerical
Description: Temperature of the water being tested in Celsius.

Summary of Variable:

Summary	Value
Minimum	0.00
25th Percentile	11.00
Mean	18.06
Median	19.00
75th Percentile	25.00
Maximum	74.00

Variable Name: air.temp
Variable Type: Numerical
Description: Temperature of the air in Celsius at testing site.

Summary of Variable:

Summary	Value
Minimum	0.00
25th Percentile	9.00
Mean	16.44
Median	15.00
75th Percentile	21.70
Maximum	74.00

Variable Name: ph
Variable Type: Numerical
Description: pH scale ranges from 0(most acidic) to 14(most basic), with 7 being neutral (www.usgs.gov).

Summary of Variable:

Summary	Value
Minimum	0.30
25th Percentile	6.50
Mean	7.17
Median	7.00
75th Percentile	7.50
Maximum	9.90

Variable Name: oxygen
Variable Type: Numerical
Description: The amount of oxygen measured in mg that is present in a liter of water.

Summary of Variable:

Summary	Value
Minimum	0.00
25th Percentile	4.80
Mean	6.65
Median	6.50
75th Percentile	8.50
Maximum	15.10