

nla2007_&_2012_data_for_farnaz_may2017

B

May 9, 2017

To Do

- insult of the day: thou churlish bat-fouling flirt-gill

Introduction

- This document (nla2007_&_2012_data_for_farnaz_may2017) describes how a subset of the 2007 and 2012 National Lake Assessment data were assembled and processed.
- The raw data are available on the NARS website: <https://www.epa.gov/national-aquatic-resource-surveys>
- These data were processed and harmonized. See the following documents for more information:
 - nla2007_waterchem
 - nla2012_waterchem
 - nla_sites-and-nla_samples
- The output includes the following datasets:
 - farnaz201705.csv
 - farnaz201705_meta.csv

Data Notes

- Decide whether or not to keep duplicate samples (`...$duplicate=='D'`)
- Decide whether or not to keep samples from `...$visit_no=2`

Data Steps

- read in the raw data and data definitions (see: https://github.com/willbmisled/lakes_database/tree/master/output)
 - ‘nla2007_chem.csv’
 - ‘nla2012_chem.csv’
 - ‘nla_chem_data_definitions.csv’
 - ‘nla_chem_parameters.csv’
 - ‘nla_sites.csv’
 - ‘nla_sites_meta.csv’
 - ‘nla_samples.csv’
 - ‘nla_samples_meta.csv’
- spread chem2007 and chem2012 based on uid, duplicate, parameter, and result then `row_bind`
- merge sites, sample, and chem data
- select fields to include and reorder
- The final dataset and the data definitions saved as:
 - “farnaz201705.csv”
 - “farnaz201705_meta.csv”