**Steps 2024-5-16**

These steps provide general guidance for cleaning the North American Breeding Bird Survey Dataset (BBS 1966 - 2022)

* Download raw BBS data: <https://www.sciencebase.gov/catalog/item/64ad9c3dd34e70357a292cee>

**Background**

* BBS is a comprehensive dataset that monitors the status and trends of North American bird populations. Started 1966, the BBS collects annual data through roadside surveys conducted during the peak breeding season. Expert observers follow a standardized protocol, recording bird species and their abundances along designated routes.
* We will use BBS to build measures of bird biodiversity and study how they potentially respond to Million Dollar Plant openings.

**Data processing**

* The URL above contains a list of files:

A screenshot of a computer

Description automatically generated

* The ultimate output dataset is a **county x year panel data** with the following information

- County FIPS: 5-digit FIPS code of the county

- Year: calendar year (1966-2022)

- **Abundance**: total number of individual birds observed (regardless of species)

- **Richness**: total number of different bird species observed

- Survey-related variables, including

- Number of surveying routes in the FIPS county

- Survey duration (minutes)

- .. and potentially other variables that may determine observed bird counts and quality of the counts

* The main task is to make sense of the relationships between the raw files, and how to combine them to produce the panel dataset.
* For reference, in the DB folder, I uploaded a cleaned BBS dataset but for period 1997-2021 (**“./data/bbs/example.dta”**). See if the dataset you produce can replicate information in this file.

A screenshot of a computer

Description automatically generated