

# FMDB Size Structure Hands-On

Derek H. Ogle

---

## Hands-On Exercise

Create a script that performs the following tasks (*Note that you will need to load several packages to perform these tasks*).

1. Load and prepare (add a month, add properly capitalized species names, and repeat records that represent multiple fish) a data.frame in R that contains the Sawyer County Raw Fish Data from the FM Database.
  2. Further prepare the data for analysis by reducing the variables to only the species name (properly capitalized), waterbody name, gear, year, month, and length (in inches and after expansion) and adding Gabelhouse length categories (“stock”, “quality”, etc.) for all species for which these categories have been defined.
  3. Create a subset of data that consists of a single year for which there is a “large” overall sample size in the “Fall.” [*HINT: First, make a two-way frequency table with sampling year as rows and months as columns.*]
  4. Identify a gamefish species and waterbody for which a “large” sample size exists. [*HINT: Make two one-way frequency tables.*]
  5. Create three new subsets of data – (1) your gamefish species (for all waterbodies, gears, etc.), (2) your waterbody (for all species, gears, etc.), and (3) your species in your waterbody.
  6. Summarize the total length (graphically, mean/SD, and PSD-X values) for your fish in your waterbody. Select a non-Gabelhouse length category to include in your analysis.
  7. Summarize the total length (mean/SD and PSD-X values) for your fish in all waterbodies (in which that fish was sampled). Select a non-Gabelhouse length category to include in your analysis.
  8. Summarize the total length (mean/SD and PSD-X values) for all fish in your waterbody.
- 

## Time Permitting

### Extensions to the Above

1. Recreate summaries of the total lengths for your fish in all waterbodies, but include quantiles. (*HINT: `quantile(X, 0.50, na.rm=TRUE)` may be used to compute the median (i.e., 50% quantile) of the data in  $X$ .*)
2. Create histograms of the total length for your fish in all waterbodies. (*HINT: Histograms of  $X$  separated by groups in  $G$  may be constructed with `hist(X~G, data=df)`. See `?hist.formula` for more help on the use of this function.*)

### Repetition of (Similar to) the Above

1. Repeat the analyses above but by choosing a different time frame (e.g., summer instead of fall), year, or gamefish species.
-