1. Import in data tables from Access
2. Currently this is manually done. Save tables from Access as .xlsx, and convert to .csv. Add to data folder in R project.
3. Open clean\_fish\_tables.Rmd and read tables into R.
4. Structural QA/QC (Clean tables)
5. Clean tables
   1. Remove columns that we don’t need
   2. Filter levels (e.g. stations) desired
   3. Clean up categorical naming
   4. Add columns
   5. Rename columns
   6. Remove NAs for important variables
6. Content QA/QC
7. Sampling information
   1. Join effort and sampling table.
   2. Remove samples that were not collected, duplicated samples from trap effort review.
   3. Join seine table.
   4. Export and evaluate sample comments
   5. Record changes to condition code based on comments, fill in missing Condition Codes
   6. Add flag for altered samples based on sample comments
   7. Make changes
   8. Separate effort table from environmental/event table
8. Water quality outliers
   1. Create regions for visualizations
   2. Separate out WQ, assign regions
   3. Look at mins and maxes of parameters
   4. Make data long for visualization
   5. Plot data and ID visual outliers
   6. Make outliers long, and identify visual outliers
      1. If visual outlier, Outlier\_vis = 3, Comment\_WQ = parameter that has an outlier
   7. Calculate Tukey, MAD outliers
      1. If extreme outlier, Tukey\_num or MAD\_num = 2. If mild outlier, Tukey\_num or MAD\_num = 1. Otherwise, assigned 0.
   8. Add Tukey, MAD, Outlier\_vis together to get an overall idea of how much each value is an outlier (range from 0-7)
   9. Classify each value into 1= acceptable (Outlier = 0-1), suspicious (Outlier = 1-4), highly suspicious (Outlier = 5-7).
   10. Plot outliers
   11. Convert outliers back to wide
   12. Combine outliers and outlier parameters into 1 column each.
   13. Combine WQ with rest of sample table

|  |  |  |  |
| --- | --- | --- | --- |
| Flag\_WQ | Definition | Outlier Value | Combinations of outlier types |
| 1 | Acceptable | 0-1 | “mild outlier” for either Tukey or MAD |
| 2 | Suspicious | 2-4 | All other combinations |
| 3 | Highly Suspicious | 5-7 | “extreme outlier” for either Tukey and/or MAD AND visual outlier |

1. Fish catch outliers
2. Write tables