In general, I didn’t remove any individual fish records from the database. The ones I did remove had errors that applied to the whole sample (e.g., wrong lake code, no year, etc.).

1. Cleaned
   1. Lake Code
      1. All lake codes were cleaned (case sensitivity, null values, etc.)
      2. All lake codes in database are present in the lake code table.
   2. Year
      1. Years have been cleaned (records with null values or nonsense years removed)
   3. Month
      1. All months have been cleaned (no null values and all between 1 and 12)
   4. Gear Code
      1. All gear codes cleaned (no null values). All gear codes have been updated to new decimal codes based on years. All gear codes in database are present in the gear code table, and records with nonsense gear codes were removed.
   5. Number.of.individuals
      1. Just made zeroes into nulls…did not remove any outliers (didn’t want to split up sample).
   6. TL\_mm
      1. Just made zeroes into nulls…did not remove any outliers (didn’t want to split up sample).
   7. Wt\_g
      1. Just made zeroes into nulls…did not remove any outliers (didn’t want to split up sample).
   8. Species Code
      1. All species codes cleaned (no null values and records with nonsense species codes were removed from the database).
2. Not cleaned - important
   1. Effort
   2. Gear.Length
   3. Station
   4. See User’s Guide (Data Input Considerations) for reasoning of importance. No way to clean really.
3. Not cleaned – unimportant
   1. Time
   2. Pool Elevation
   3. Surface Temp
      1. Units are mixed between C and F in database
   4. Secchi
   5. Conductivity
   6. Gear Length
   7. Habitat