FISH FARMING RESEARCH AND ANALYSIS REPORT
SPECIES BIO COMPARISON
CHRISTINA LU JIN | OCT 2020

PART A – NUMERIC DESCRIPTIVE CHARACTERISTICS ANALYSIS

By first checking on the top and bottom 3 rows of the dataset, it gives us a general idea on what are we dealing with. Looks like we are handling a dataset about length and weight regarding different species of fishes. I've also noticed that fishes with longer total length and heavier weight are marked as 'true' for the "scale" attribute, thus, I'm assuming that only the matured fishes have been scaled, not the juvenile ones.

```
> headtail(bio,n=3)
    netID fishID        species   tl      w  tag scale
1      12     16       Bluegill   61    2.9       FALSE
2      12     23       Bluegill   66    4.5       FALSE
3      12     30       Bluegill   70    5.2       FALSE
674   110    863 Black Crappie  307  415.0 1783   TRUE
675   129    870 Black Crappie  279  344.0 1789   TRUE
676   129    879 Black Crappie  302  397.0 1792   TRUE
```

It had also been brought to my attention that there are two species, Bluegill and Largemouth Bass, have way more records than the rest. With just these two together, it makes a 66% of the total data for entire 8 species.

```
> d
          Species RelFreq Counts CumFreq CumCounts
1  Largemouth Bass    34 %    228    34 %       228
2         Bluegill    33 %    220    66 %       448
3 Bluntnose Minnow    15 %    103    82 %       551
4     Yellow Perch     6 %     38    87 %       589
5    Black Crappie     5 %     36    92 %       625
6      Iowa Darter     5 %     32    97 %       657
7      Pumpkinseed     2 %     13    99 %       670
8   Tadpole Madtom     1 %      6   100 %       676
```

In addition, a little data clean was applied for this dataset to first keep only the records that are with acutal values (remove all 'N/A's). By looking at some additional descriptive data below, the length attribute appears to be pretty spread out through the span, whereas the weight attribute is extremely right skewed. Most likely that this dataset contains a great amount of juvenile or light-weighted species of fishes, and at the same time there are some outliers with extremely heavy weight.

```
> summary(bio_clean)
   species                tl                  w                 scale
 Length:511        Min.   : 28.0    Min.   :   0.2    Mode :logical
 Class :character  1st Qu.: 61.0    1st Qu.:   2.0    FALSE:200
 Mode  :character  Median :150.0    Median :  54.5    TRUE :311
                   Mean   :156.6    Mean   : 126.8
                   3rd Qu.:226.0    3rd Qu.: 190.5
                   Max.   :429.0    Max.   :1070.0
```

Thus, furthermore, I filtered out the juvenile fishes (weight under 10), in order to deliver a better picture of what the portions are of both mature and juvenile fishes for each species. It is quite surprising to see how many juvenile fishes are there for the two dominated species - Largemouth Bass and Bluegill. And also some other species like Blutnose Minnow, Iowa Darter, and Tadpole Madtom don't even have any mature fishes, or they are possibly just species with genrally small size and don't weight much. Further investigation is need for this.

```
> arrange(tmp,desc(count))          > arrange(bio_mature_count
           species count                       species count
1  Largemouth Bass   228            1         Bluegill   157
2         Bluegill   220            2 Largemouth Bass    88
3 Bluntnose Minnow   103            3   Black Crappie    25
4     Yellow Perch    38            4    Yellow Perch    25
5    Black Crappie    36            5     Pumpkinseed     8
6      Iowa Darter    32
7      Pumpkinseed    13
8    Tadpole Madtom     6
```

PART B – GRAPHIC DESCRIPTIVE CHARACTERISTICS ANALYSIS

In this section, we will be interpreting some graphic descriptive character using the following plots, where more information can be revealed in regards of how each species are distributed, as well as how are the length and weight related. These will help in making strategical decision of what species are better choices to be cultivated and harvested with in the given water farm.

The first bar plot below shows the relative frequency of each species of fishes, and the second pareto graph shows a combination of counts, cumulative counts and cumulative percentage for all types. Both graphs showed a domination of two species in the research data. I believe further researches are needed to determine the reason for such imbalance, whether the water body is not the perfect inhabitance for the rest of the species or whether these two types of fishes prey the rest.
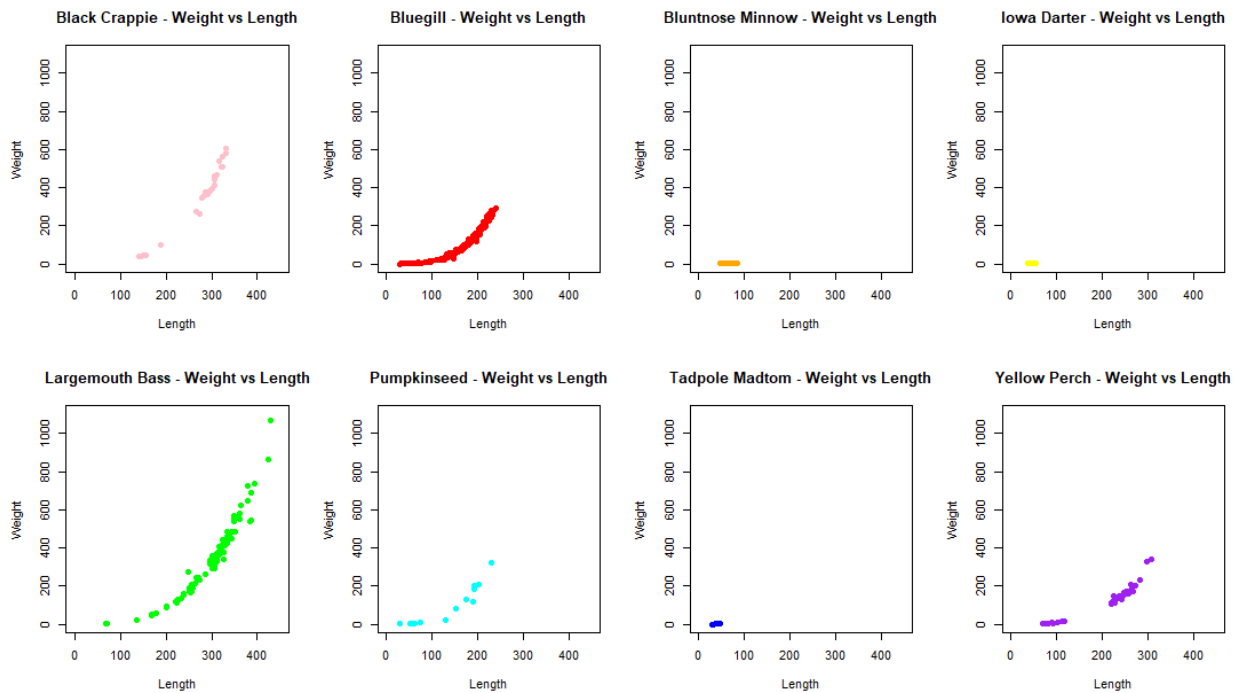
Fish Relative Frequency
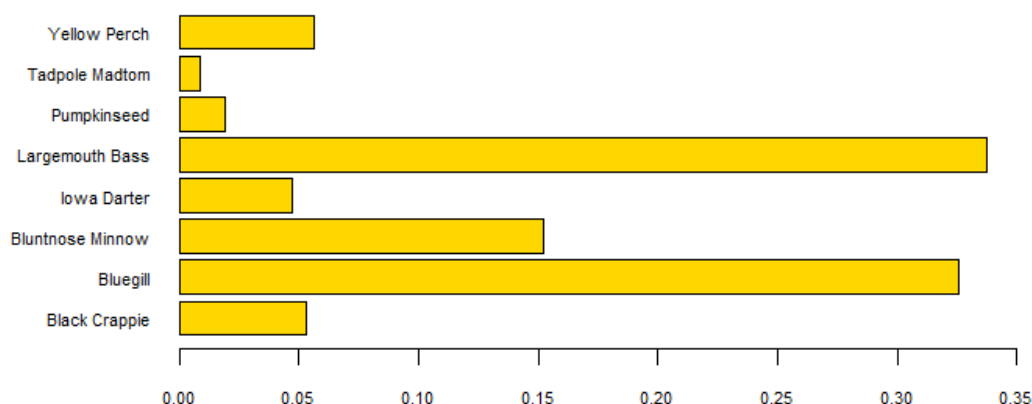
Species Pareto

by Christina Lu Jin
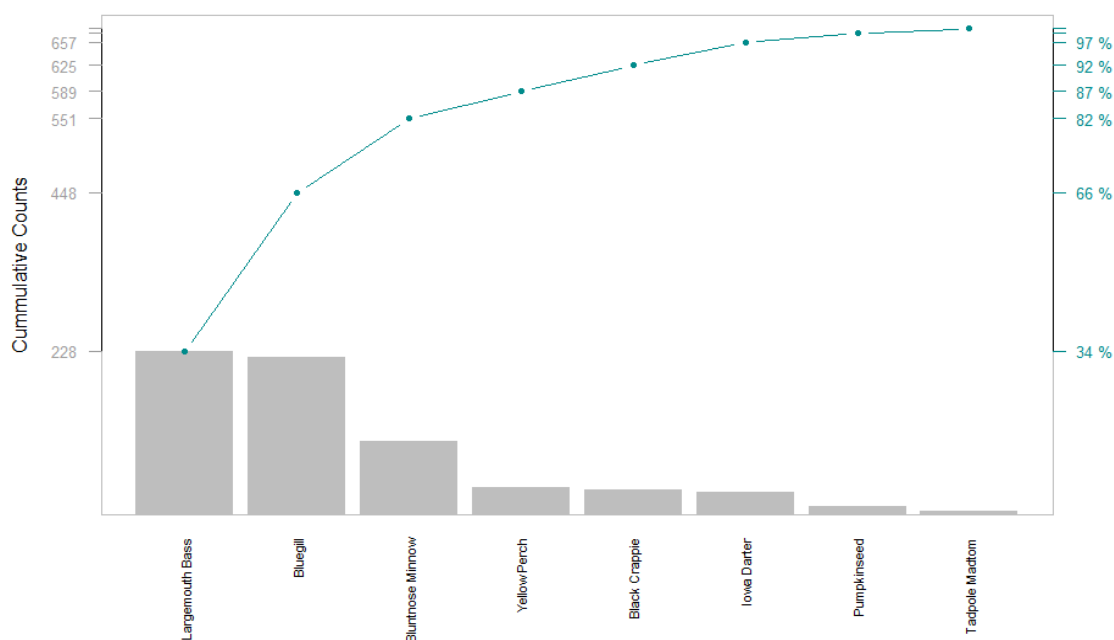
Following series of scatter plots were generated to analyze the relationship between length and weight among each species, a drastic difference in size is shown. Weight doesn't seem to change much over length at early stage (shorter than 100). Largemouth Bass surprisingly grows quite big, there are definitely a major reason causing the extreme positive skewness in the average weight. Black Crappie seems to grow quite big as well, while Yellow Perch grows to similar length but does not weight as much. Both Bluegill and Pumpkinseed have a denser population at a relatively shorter length and lighter weight. Whereas the rest three species are still at very early stage and are super tiny.

Black Crappie - Weight vs Length | Bluegill - Weight vs Length | Bluntnose Minnow - Weight vs Length | Iowa Darter - Weight vs Length

Largemouth Bass - Weight vs Length | Pumpkinseed - Weight vs Length | Tadpole Madtom - Weight vs Length | Yellow Perch - Weight vs Length

SUMMARY

Therefore, it is suggested to cultivate more Largemouth Bass, Black Crappie and Yellow Perch in this water body. It worth a little further research on Bluntnose Minnow, Iowa Darter, and Tadpole Madtom to figure out the reason there are barely any mature ones among them. It will also be beneficial to cultivate these three types separately from the rest species to prevent any predation. If the result shows that they are just generally tiny in size or hard to live in this water body, then maybe consider cultivating other types of fishes. Also, overfishing should be prohibited to allow longer time for all fishes to grow.