Analysis of Fishhead Company: Fish Population Study
Basic Trends

The original fish competition data has several interesting trends as shown in the graph “Results of Competition 1990-1999.” The most distinct trend is the number of fish caught. This category peaked in 1992 and has been on the decline since. Also, the general pattern of the number tagged and returned has declined after peaking in 1992. Interestingly, the number over twelve inches without tags has remained constant, which suggests that the majority of the fish have been tagged in some way. In addition, the number over twelve inches with tags has declined, but has increased slightly in recent years. Lastly, the number already tagged and returned has dramatically declined since 1996.

![Results of Competition 1990-1999](image_url)

After examining the basic trend of the competition results, an analysis of the composition of the catch will give a more in-depth perspective. As seen in the graph below titled “Composition of Catch,” the overall flow of the chart is the same as the number of fish caught in the above graph. As a notation, the one on the x-axis represents 1990, while the ten is 1999. In the first few years, 1990-1993, the majority of the composition was made up of the number of fish tagged and returned. Then as the years went on, this category declined and no longer contributed the most to the composition because the number with tags increased. Thus, as the number tagged and returned decreased in the later years, the number caught with tags increased. A similar pattern develops in the relationship between the number of fish already tagged and returned and the number of fish caught without tags. The number already tagged and returned increased until 1994 through 1996 and has been decreasing since then. Conversely, the number caught without tags had decreased until 1994 through 1996 and has been increasing since then. Lastly, the number already tagged and returned has decreased incredibly from 1997-1999 to a point where it is almost nonexistent in the composition.
Through an initial observation of the data presented, we can conclude that the number of fish caught is on the decline. This conclusion is seen in the previous graph titled “Results of Competition 1990-1999.” Although the data is not linear overall, it appears to be linear since 1992. This change in the pattern of the graph insinuates that something dramatic occurred in 1992 to cause the fish population to decrease. In order to continue with our analysis and prediction we would calculate the least-squares regression equation for this time span. After determining the regression coefficients, slope and intercept, we concluded that the equation is: $y = 737.53 - 52.369x$, where $y$ is the number of fish and $x$ is time in years. From this, we can later compute the predicted figures for
the number of fish. As seen in the above graph, “Regression of Fish Caught,” the regression line follows the overall pattern of the fish caught from 1992 to 1999. Please note on the graph that number eight is 1999, while one is 1992.

Confidence Intervals

Following the calculation of the regression line, confidence intervals were computed to determine the reliability of the prediction. In making the computations, the following equations were used to produce the 95% confidence intervals.

A confidence interval for the intercept is \( b_0 \pm t^* s_{b0} \) where \( t^* \) is the t-score associated with a 95% confidence interval and the standard error

\[
S_{b_0} = s \sqrt{\frac{1}{n} + \frac{-2}{x \sum (x_i - \bar{x})^2}}.
\]

A confidence interval for the slope is \( b_1 \pm t^* s_{b1} \) where \( t^* \) is the t-score associated with a 95% confidence interval and the standard error

\[
S_{b_1} = \frac{s}{\sqrt{\sum (x_i - \bar{x})^2}}.
\]

The upper and lower intervals for the regression intercept were 785.2738 and 689.7976 respectively. Likewise, the upper and lower confidence intervals for the regression slope were –42.9155 and –61.82259, respectively. By looking at the numbers of the confidence intervals, we see that there is a wide range of predictions that could fit within these intervals. However, the size of these intervals is skewed due to the lack of data. If there were more data, then the confidence intervals would provide a better security of the prediction.

Prediction

Despite the immense confidence intervals, we still used the regression line for making our predictions because the number of fish caught followed the predicted line. Thus, as long as our prediction is made in a timely manner, meaning within a few years of the original data, then we can assume accuracy. The number of fish caught in the next three years is predicted to be (2000,453), (2001,442), (2002,431). Yes, the population seems to be declining as the years progress; however, we believe that the number of fish will begin to level off in the near future. By looking at the original fish data, one can see that the amount of regression has slowed down in recent years, which leads us to believe that the number will begin to even out.

In addition, the correlation between the number of fish caught and the number of fish tagged and returned show that there is a possible leveling off of the population. Once again looking at the “Results of Competition 1992-1999,” we see that the number of fish tagged and returned has leveled in recent years. By computing the correlation between the number tagged and returned and the number of fish caught, we find a correlation of .917943. Since this number is near positive one, we see that the two variables have a strong positive correlation. Thus, the number of tagged and returned have began to level off which has allowed us to predict that the number of fish caught will follow in a similar manner.
Additional Information

After examining the given data, as a statistical company, we have come to the conclusion that there are some possible reasons for the decline and there is some more information that could enhance our prediction. For instance, additional data on the number of participants in the competitions would be beneficial to us. If the number of participants are declining then a conclusion can be drawn that the participants are not catching the same amount as previous years because there are not enough people. Likewise, if there is an increase in the number of participants, then a conclusion can be drawn that due to the abundance of people on the lake, then the fish may be frightened away from the surface. In addition, the creation of new technology, such as sonars and other fishing devices, may have caused over-fishing to occur in the lake. Likewise, other technology, such as motorboats and jet skis, may have had an impact on the population.

In addition, more information concerning the environment surrounding the lake would be beneficial. For instance, the pH of the water in each year may be a factor. If the pH has fluctuated significantly then the fish may have decreased as a result. Also, any development of industries in the area since the start of the competitions may have an impact on the fish population. For instance, if a new industry arose in the area and is dumping chemicals in the water, then the fish population may be depleting as a result.

Also, the fish may not be reaching their full maturity level in order to reproduce. Since the fish may not be reproducing as much as expected, then the population will decrease. With additional information, we would be more inclined to develop a better prediction can be made. As seen in the many alternatives to the decline in the fish population, time series data does not provide sufficient information to make a prediction of the population of the fish. However, based on the information we received, we have produced the best prediction within our knowledge and experience.