Angel Carrasquillo

BIO 539

**Final Project**

**Introduction**

This project looks at oysters from various wild populations and selectively bred commercial lines in and how they perform in New England. The four wild population stocks used in this project are named Martha’s Vineyard (MV), Narrow River (NR), Green Hill Pond (GH) and Connecticut (CT), based off of where they were collected. The two commercially bred lines were created from genetic material in the mid-Atlantic and are named Commercial Lines 1 & 2 (CL1 & CL2) to protect the anonymity of our providers.

**Methods**

Oyster Growout and Measuring

These lineages were grown over the course of two years at Roger Williams University. Each lineage started with 3 bags with 55 oysters per bag. The oysters were born April 2021 and measured in July 2021, August 2021, October 2021, May 2022 and October 2022 when the experiment ended and the oysters were sacrificed. At the end of the experiment, the data collected will be used to determine yield, percent mortality and size of oysters in each lineage. At each measurement day, the oysters were recorded into a master datasheet. This means that each row of the master datasheet contained information about the oyster on the date it was measured. On each measurement date the oyster ID, its lineage, shell length and if it was alive, dead, or missing was recorded into the sheet.

Determining Yield

For this project, yield is defined as the sum of the shell lengths of living oysters at the end of the experiment. The yield of each bag was calculated in R using oysters that had been measured and recorded alive in October 2022. Some oysters were alive at the end of the experiment, but for various reasons were unable to be measured. Unfortunately, some of our bags had been lost during the experiment, possibly due to storms leading to some lineages only having one bag. In addition, one of the bags from the NR line was open at the end of the experiment, and it is likely that a significant number of oysters had gone missing from the bag opening. For this reason, this bag was not included in yield measurements. The yield of each line was graphed using a boxplot in R, but because each lineage had at most three replicates, no statistical analysis could be done to show statistical significance. Thus, the results of this analysis were based on visual interpretation of the graph.

Determining Percent Mortality

Mortality was measured using the Kaplan-Meier estimate. This statistic evaluates survival in a group from the beginning of a treatment group to the end of a treatment group. This analysis was done using the SPSS software program from IBM. SPSS requires a dataframe with the individual, group, event and time when the individual died or exited the study to perform a Kaplan-Meier estimate. R was used to edit the original master datasheet to include the necessary information for this method in a form that SPSS could analyze. Missing oysters were given the same status as alive oysters at the time they went missing indicating that this was the time they exited the analysis as opposed to being reported as dead. The age of the oyster when it exited the analysis or died was calculated based on the months after the spawning of the oysters. Oysters that survived the experiment were reported as exiting the analysis after 19 months. Oysters from the NR lineage had to be excluded from the analysis because of one of the bags was reported as missing before the end of the experiment which would have led to unreliable results. The dataframe generated from R was imported into SPSS in which age was the time variable, status was the status variable with 1 being the event and lineage was the factor. The log rank test statistic was used, and a survival table was generated. SPSS also performed a Chi-square Test to determine if the difference in mortality between each lineage was significant.

Analyzing Size

The final metric used to evaluate performance was the size of oysters at the end of the experiment. As with yield, this analysis looked at the size of oysters that were alive and measured at the end of the experiment. However, for this analysis the oysters were not separated by bag, instead a histogram of oyster lengths was generated for each lineage. Using a Shapiro-Wilk normality test, it was found that the sizes of oysters were not normally distributed and to determine if the sizes of the oysters in each lineage was statistically significant, a Kruskal-Wallis test was used in place of an ANOVA.

**Results**

Chart, box and whisker chart

Description automatically generated

Figure 1: Boxplot of average oyster yield in each lineage based on the cumulative length of all oyster shells per bag at the end of the experiment. Visually, the CT and GH stocks performed similarly to the commercial lines. Both commercial lines performed similarly, but the wild lineages performed differently, due to the variety of conditions from where they were collected.

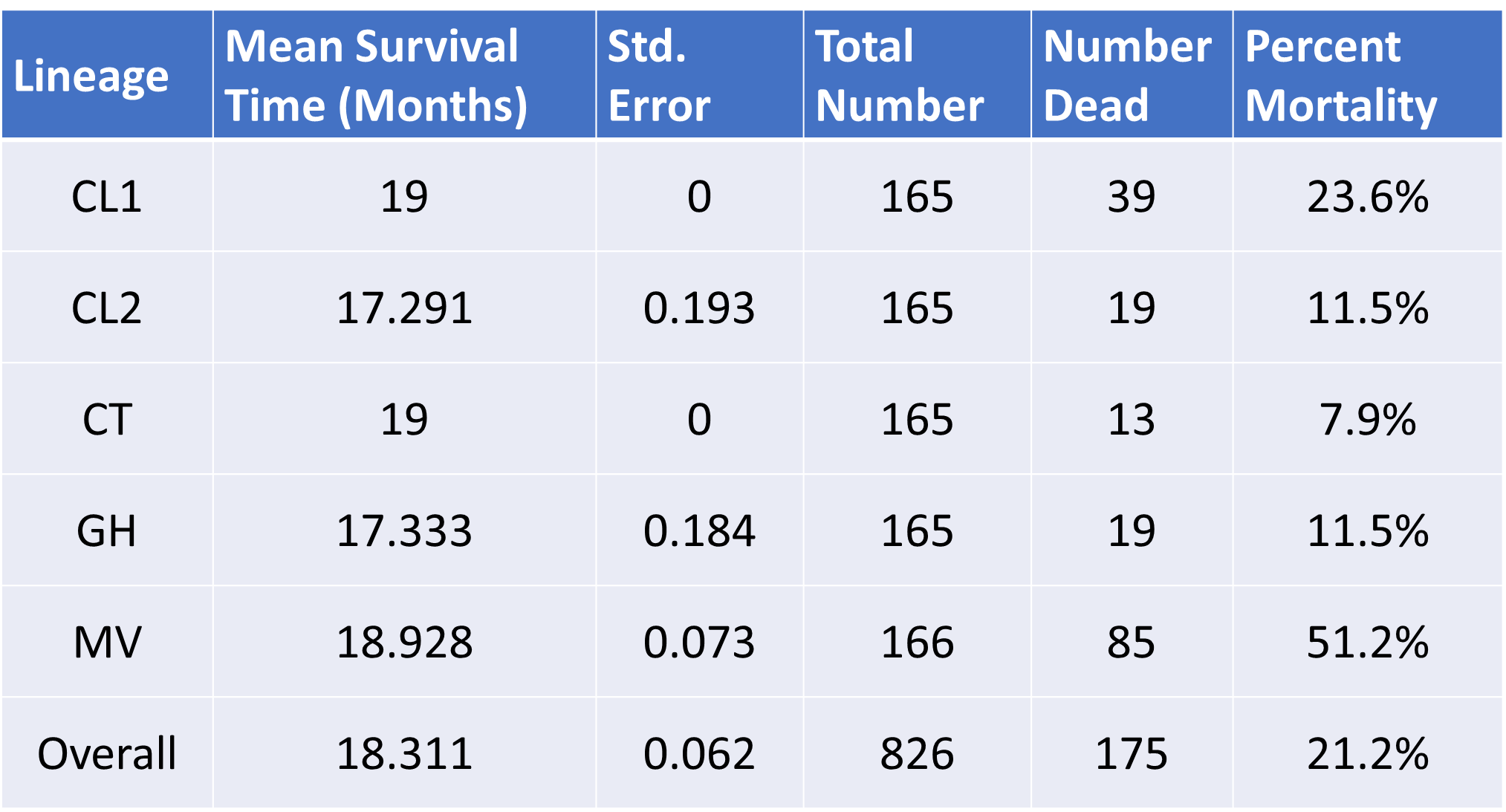


Table 1: Percent survival in each lineage measured at each collection time. The best performing line was CT, which had a mortality of 11.5% while MV was the worst with a mortality of 51.1%. The values in this table were generated in SPSS. By the end of the growout period there was a statistically significant difference in mortality (Chi-square Test, p < 0.01) between groups.

Chart, histogram

Description automatically generated

Figure 2: Histogram of oyster sizes from each lineage. These histograms were generated in R. The charts in red indicate commercial lines while the charts in blue are wild stocks. The two commercial lines had the largest oysters and widest distributions of shell sizes. Each lineage performed significantly different (Kruskal-Wallis Test, p<2.2e-16).