Temp/Size Analysis

Kathleen Durkin

2023-10-12

I'm doing this analysis under the assumption that all three of the SL columns (SL_11212022, SL_12272022, and SL_mm) are measuring some consistent size/length value over three measurement dates. Similarly, I'm assuming the the columns WWT_11212022, WWT_12272022, and WholeBodyWW_g are measurements of body weight (maybe whole body wet weight?) on the same three measurement dates.

Data Munging

```
# read in the data
codTempData <- read.csv("../../data/temp-experiment.csv")
head(codTempData)
```

```
##   Microchip.ID SL_11212022 WWT_11212022 Tank Temperature SL_12272022
## 1          620          93         8.53    1          16         101
## 2         1164          88         7.06    1          16          96
## 3         1476         102        10.70    1          16         108
## 4         9387          87         7.83    1          16          95
## 5         9407         100        11.51    1          16         117
## 6         9415          92         8.68    1          16         100
##   WWT_12272022 MortDate DissectionDate SL_mm WholeBodyWW_g TOTAL_Liver_WW_mg
## 1        11.12                 2/8/23    119         16.15            0.4945
## 2         8.64                 2/8/23    105         10.89            0.1997
## 3        12.25                 2/8/23    110         12.97            0.1715
## 4        10.16                 2/8/23    116         15.40            0.3625
## 5        14.98                 2/8/23    127         17.98            0.3482
## 6        10.96                 2/8/23    114         14.02            0.2343
##   LiverforLipids_WW_mg MuscleWWforLipids_mg GeneticSamplingCount
## 1               0.1546               0.3495                    8
## 2               0.1091               0.3328                    5
## 3               0.1107               0.3834                    4
## 4               0.1681               0.3262                    6
## 5               0.1210               0.3434                    2
## 6               0.1342               0.2776                    9
##   DissectionComments
## 1
## 2
## 3
## 4
## 5
## 6
```

```
# Create two new columns indicating change from Nov.2022 measurement to Feb.2022 measurement, for both
codTempData_plus <- transform(codTempData,
```

```
                              # create column for change in size
                              sizeChange_mm = SL_mm - SL_11212022,
                              # create column for change in weight
                              weightChange_g = WholeBodyWW_g - WWT_11212022) %>%
  # change type of Temperature variable to an ordered factor
  mutate(codTempData, Temperature = relevel(as.factor(Temperature), "0", "5", "9", "16"))
head(codTempData_plus)
```

```
##   Microchip.ID SL_11212022 WWT_11212022 Tank Temperature SL_12272022
## 1          620          93         8.53    1          16         101
## 2         1164          88         7.06    1          16          96
## 3         1476         102        10.70    1          16         108
## 4         9387          87         7.83    1          16          95
## 5         9407         100        11.51    1          16         117
## 6         9415          92         8.68    1          16         100
##   WWT_12272022 MortDate DissectionDate SL_mm WholeBodyWW_g TOTAL_Liver_WW_mg
## 1        11.12                   2/8/23   119         16.15            0.4945
## 2         8.64                   2/8/23   105         10.89            0.1997
## 3        12.25                   2/8/23   110         12.97            0.1715
## 4        10.16                   2/8/23   116         15.40            0.3625
## 5        14.98                   2/8/23   127         17.98            0.3482
## 6        10.96                   2/8/23   114         14.02            0.2343
##   LiverforLipids_WW_mg MuscleWWforLipids_mg GeneticSamplingCount
## 1               0.1546               0.3495                    8
## 2               0.1091               0.3328                    5
## 3               0.1107               0.3834                    4
## 4               0.1681               0.3262                    6
## 5               0.1210               0.3434                    2
## 6               0.1342               0.2776                    9
##   DissectionComments sizeChange_mm weightChange_g
## 1                               26           7.62
## 2                               17           3.83
## 3                                8           2.27
## 4                               29           7.57
## 5                               27           6.47
## 6                               22           5.34
```

```
# Reformatted data with single column for size values and single column for measurement values (and add
#
# Sample of how data is being reformatted:
# Original data
# fishID | size_date1 | size_date2 | weight_date1 | weight_date2
#---------------------------------------------------------------
#  001   |     s11    |     s12    |     w11      |     w12
#  002   |     s21    |     s22    |     w21      |     w22
#
# Reformatted data
# fishID |  date  | size | weight
#--------------------------------
#  001   | date1 |  s11 |  w11
#  001   | date2 |  s12 |  w12
#  002   | date1 |  s21 |  w21
#  002   | date2 |  s22 |  w22
```

```r
# Note I renamed the final size and weight measurements to include the date 02/08/2023 -- this is just

codTempData_reformat <- codTempData_plus %>%
  # Rename final size/weight variables to include date
  rename(WWT_02082023=WholeBodyWW_g) %>%
  rename(SL_02082023=SL_mm) %>%
  # Reformat data
  pivot_longer(
    cols = c("SL_11212022", "SL_12272022", "SL_02082023", "WWT_11212022", "WWT_12272022", "WWT_02082023
    names_to = "var",
    values_to = "value"
  ) %>%
  separate(var, into = c("var", "date"), sep = "_") %>%
  pivot_wider(
    names_from = "var",
    values_from = "value"
  )
# Set the date variable to have desired (chronological) order
codTempData_reformat$date <- factor(codTempData_reformat$date, levels = c("11212022", "12272022", "0208
```

Plots

```r
# Plot size measurements for all temperature treatments across the time of the study
codTempData_reformat %>%
  ggplot(aes(x=Temperature,
             y=SL,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.2,
              height = 0.2,
              size = 1.5) +
  xlab("Temperature Treatment (*C)") +
  ylab("Size (mm)") +
  facet_wrap(~date)
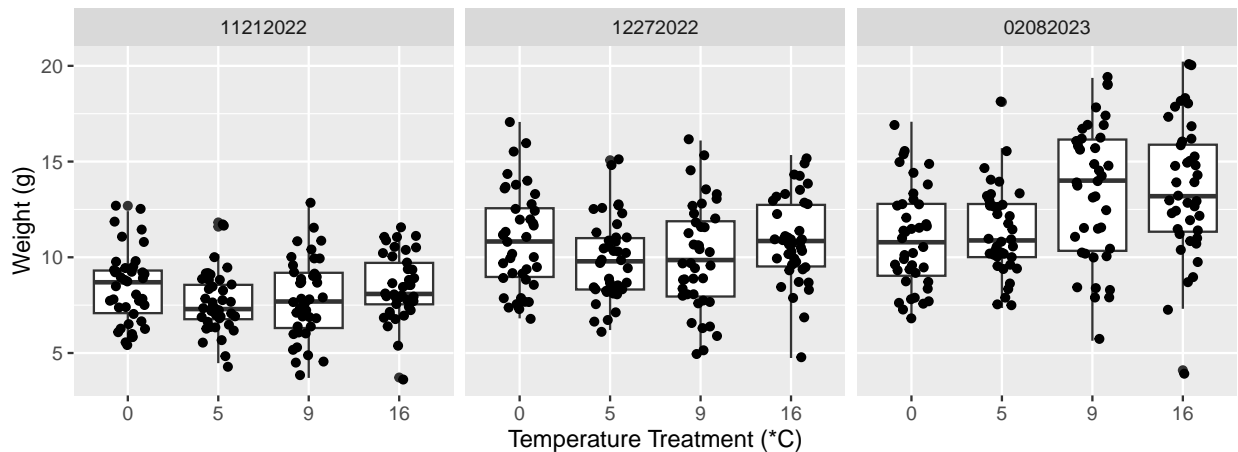```



```r
ggsave(
    "01_sizeVtreatment-all-dates.png",
```

```
    plot = last_plot(),
    path = "../../output/01_temp-size-analysis"
  )

# Plot weight measurements for all temperature treatments across the time of the study
codTempData_reformat %>%
  ggplot(aes(x=Temperature,
             y=WWT,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.2,
              height = 0.2,
              size = 1.5) +
  xlab("Temperature Treatment (*C)") +
  ylab("Weight (g)") +
  facet_wrap(~date)
```
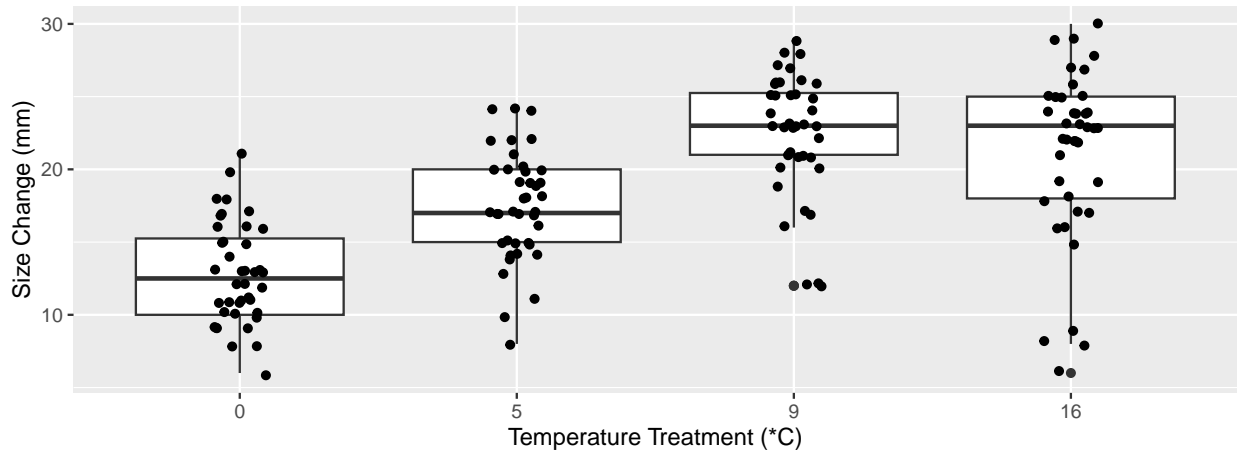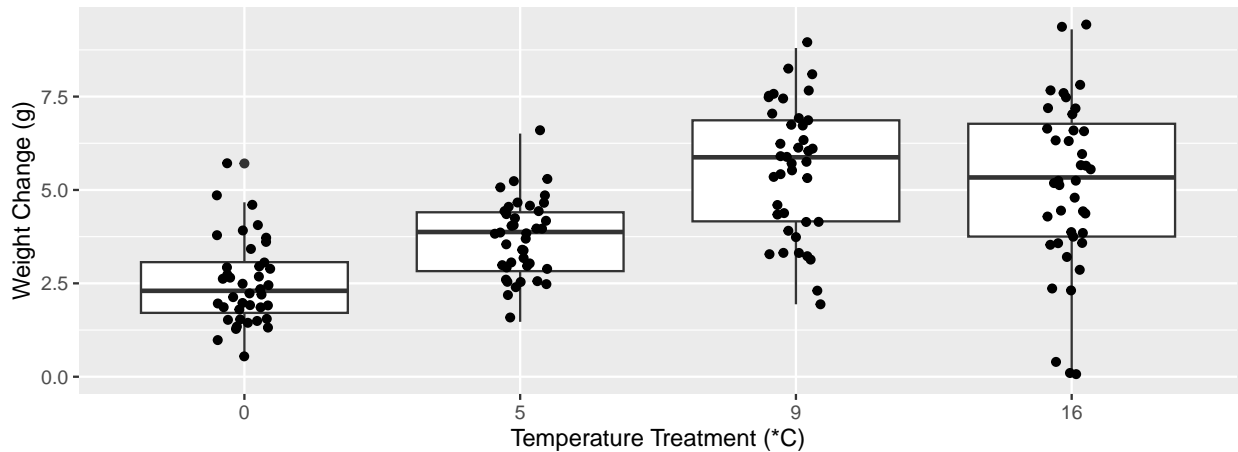


```
ggsave(
    "02_weightVtreatment-all-dates.png",
    plot = last_plot(),
    path = "../../output/01_temp-size-analysis"
  )
```

Size/Weight Change

```
# Plot *change* in size from beginning to end of study for all temperature treatments
codTempData_plus %>%
  ggplot(aes(x=Temperature,
             y=sizeChange_mm,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.1,
              height = 0.2,
              size = 1.5) +
  xlab("Temperature Treatment (*C)") +
  ylab("Size Change (mm)")
```
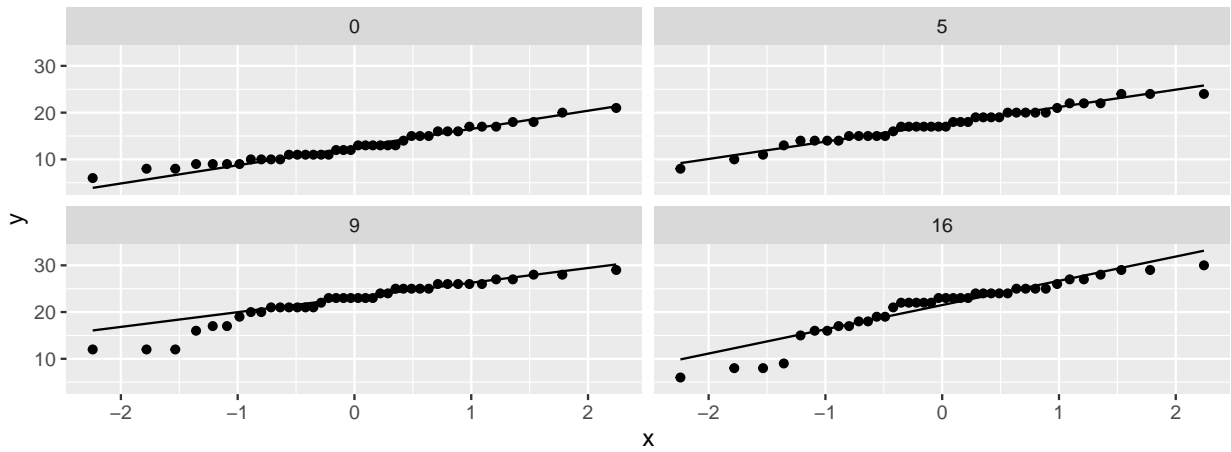
```r
# Plot *change* in weight from beginning to end of study for all temperature treatments
codTempData_plus %>%
  ggplot(aes(x=Temperature,
             y=weightChange_g,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.1,
              height = 0.2,
              size = 1.5) +
  xlab("Temperature Treatment (*C)") +
  ylab("Weight Change (g)")
```


Looking at these plots visually, there seems to be a difference in change in both size and weight over time among the treatment temperatures. Let's test this statistically.
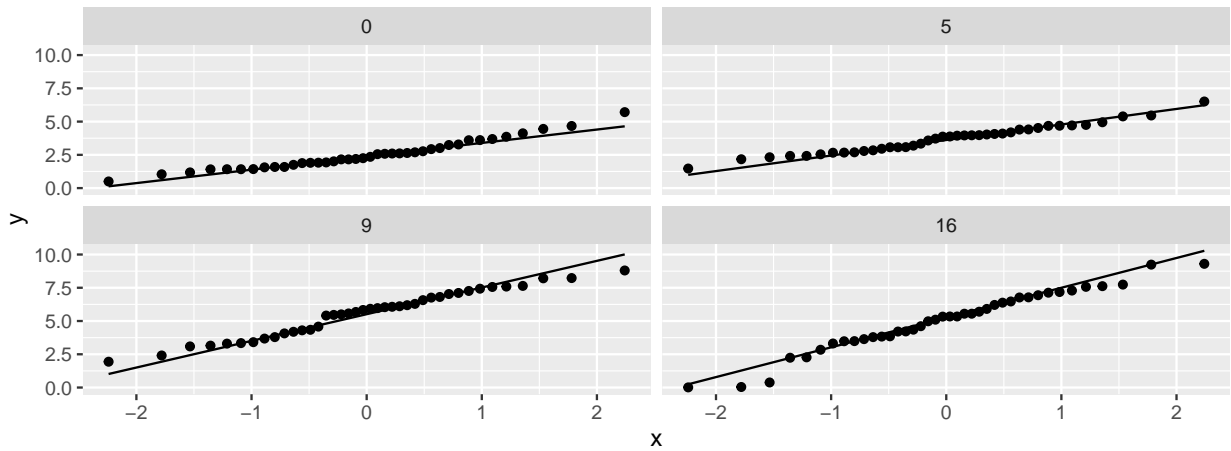
Check Assumptions

```r
# Check conditions for ANOVA

# Normality
  # Not perfect, but normalish enough that I feel comfortable using ANOVA
codTempData_plus %>%
  ggplot(aes(sample = sizeChange_mm)) +
```

```
stat_qq() +
stat_qq_line() +
facet_wrap(~Temperature)
```



```
codTempData_plus %>%
  ggplot(aes(sample = weightChange_g)) +
  stat_qq() +
  stat_qq_line() +
  facet_wrap(~Temperature)
```



```
# Variance
  # For both sizeChange and weightChange, the largest SD is, at most, ~2x the smallest SD. This is suff
codTempData_plus %>%
  group_by(Temperature) %>%
  summarize(meanSizeChange = mean(sizeChange_mm),
            sdSizeChange = sd(sizeChange_mm),
            meanWeightChange = mean(weightChange_g),
            sdWeightChange = sd(weightChange_g))
```

```
## # A tibble: 4 x 5
##   Temperature meanSizeChange sdSizeChange meanWeightChange sdWeightChange
##   <fct>                <dbl>        <dbl>            <dbl>          <dbl>
## 1 0                     12.8         3.49             2.51           1.09
## 2 5                     17.4         3.70             3.69           1.05
```

```
## 3 9                        22.4            4.28            5.57                1.74
## 4 16                       21.2            5.86            5.05                2.24
```
```
# Assuming data are independent (part of experimental design)
```

ANOVA

Size Change

```
# ANOVA
sizeANOVA <- aov(sizeChange_mm~Temperature, data=codTempData_plus)

tidySizeANOVA <- tidy(sizeANOVA)
tidySizeANOVA
```
```
## # A tibble: 2 x 6
##   term           df sumsq meansq statistic   p.value
##   <chr>       <dbl> <dbl>  <dbl>     <dbl>     <dbl>
## 1 Temperature     3 2234.   745.      37.9  1.79e-18
## 2 Residuals     156 3062.   19.6       NA    NA
```
```
# Calculate R^2 (how much of the variation in the data is explained by the treatment)
r_squared <- tidySizeANOVA$sumsq[1]/(tidySizeANOVA$sumsq[1]+tidySizeANOVA$sumsq[2])
r_squared
```
```
## [1] 0.4218882
```

$p = 1.79e{-}18 \ll 0.05$, so there is a significant relationship between treatment (temperature) and size growth (change in size). $R^2=0.422$, indicating ~42% of variance in size change is explained by the temperature treatment.

```
# Tukey HSD
sizeANOVA %>%
  TukeyHSD() %>%
  tidy() %>%
  select(contrast, estimate, adj.p.value) %>%
  arrange(adj.p.value)
```
```
## # A tibble: 6 x 3
##   contrast estimate adj.p.value
##   <chr>       <dbl>       <dbl>
## 1 9-0          9.57     3.89e-14
## 2 16-0         8.38     1.54e-13
## 3 9-5          4.97     8.19e- 6
## 4 5-0          4.6      4.25e- 5
## 5 16-5         3.78     1.13e- 3
## 6 16-9        -1.20     6.21e- 1
```
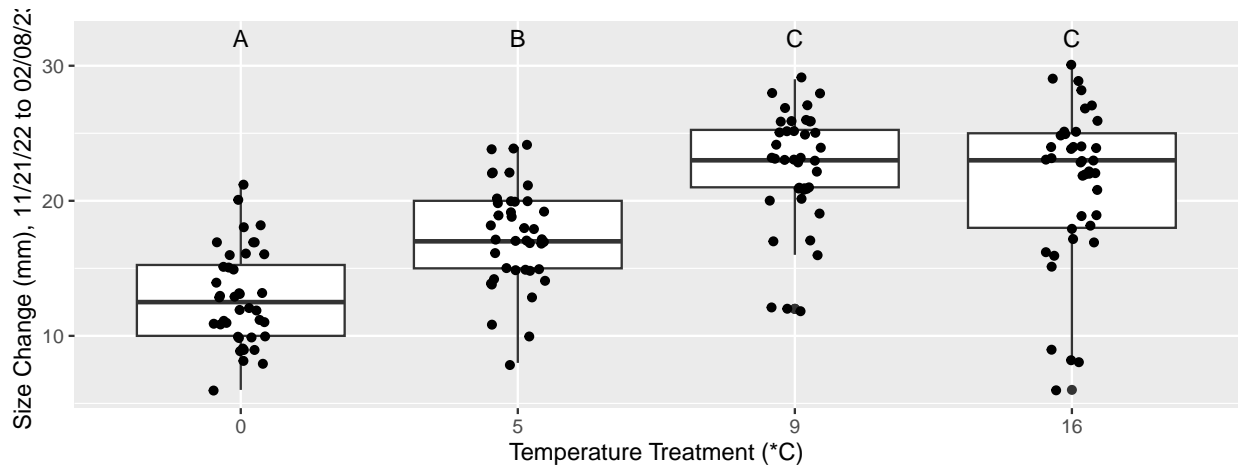```
codTempData_plus %>%
  ggplot(aes(x=Temperature,
             y=sizeChange_mm,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.1,
              height = 0.2,
              size = 1.5) +
```

```
annotate(geom = "text", x = 1:4, y = 32, label = c("A","B","C","C")) +
xlab("Temperature Treatment (*C)") +
ylab("Size Change (mm), 11/21/22 to 02/08/23")
```



```
ggsave(
  "03_size-change-TukeyHSD-plot.png",
  plot = last_plot(),
  path = "../../output/01_temp-size-analysis"
)
```

Weight Change

```
# ANOVA
weightANOVA <- aov(weightChange_g~Temperature, data=codTempData_plus)

tidyWeightANOVA <- tidy(weightANOVA)
tidyWeightANOVA
```

```
## # A tibble: 2 x 6
##   term          df sumsq meansq statistic  p.value
##   <chr>      <dbl> <dbl>  <dbl>     <dbl>    <dbl>
## 1 Temperature    3  229.  76.3       29.5 3.73e-15
## 2 Residuals    156  404.   2.59       NA   NA
```

```
# Calculate R^2 (how much of the variation in the data is explained by the treatment)
r_squared <- tidyWeightANOVA$sumsq[1]/(tidyWeightANOVA$sumsq[1]+tidyWeightANOVA$sumsq[2])
r_squared
```

```
## [1] 0.361786
```

p = 3.73e-15 « 0.05, so there is a significant relationship between treatment (temperature) and weight change. R^2=0.362, indicating ~36% of variance in weight change is explained by the temperature treatment.
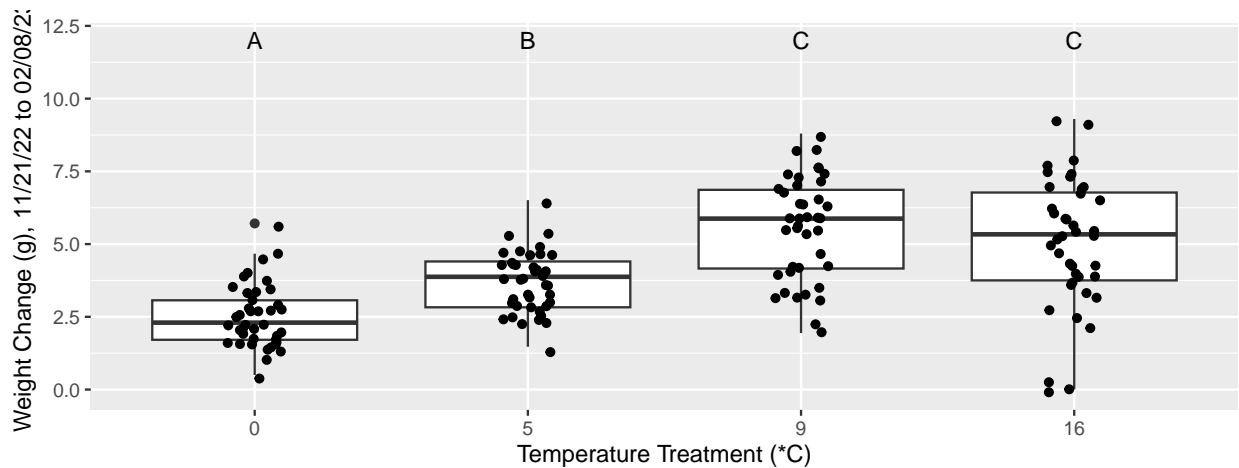
```
# Tukey HSD
weightANOVA %>%
  TukeyHSD() %>%
  tidy() %>%
  select(contrast, estimate, adj.p.value) %>%
  arrange(adj.p.value)
```

```
## # A tibble: 6 x 3
##    contrast estimate adj.p.value
##    <chr>       <dbl>       <dbl>
## 1 9-0          3.06     1.24e-13
## 2 16-0         2.54     2.90e-10
## 3 9-5          1.88     3.28e- 6
## 4 16-5         1.36     1.22e- 3
## 5 5-0          1.18     6.90e- 3
## 6 16-9        -0.516    4.79e- 1
```

```
codTempData_plus %>%
  ggplot(aes(x=Temperature,
             y=weightChange_g,
             group=Temperature)) +
  geom_boxplot() +
  geom_jitter(width = 0.1,
              height = 0.2,
              size = 1.5) +
  annotate(geom = "text", x = 1:4, y = 12, label = c("A","B","C","C")) +
  xlab("Temperature Treatment (*C)") +
  ylab("Weight Change (g), 11/21/22 to 02/08/23")
```



```
ggsave(
    "04_weight-change-TukeyHSD-plot.png",
    plot = last_plot(),
    path = "../../output/01_temp-size-analysis"
  )
```

For both size and weight, growth from 11/21/22 to 02/08/23 significantly differed among all temperature treatments, with the exception of the 9 degree and 16 degree treatments. For the 9 and 16 degree treatments, changes in size and weight were statistically similar. In other words, growth increased with the treatment temperature until the 16 degree treatment, for which growth was not significantly different from the 9 degree treatment in either size or weight.