Individual project pre analysis

Claire Campion

5/1/2022

1. Overview

Bumble bee populations are increasingly under threat from habitat fragmentation, pesticides, pathogens, and climate change. Climate is likely a prime driver of declines in abundance and distribution, as bumble bees are limited by their thermal tolerance (Kerr et al. 2015). However, the tolerance of whole organisms often exceeds that of their gametes; for example, insects can be sterilized by temperatures below their upper thermal tolerances (Heerwaarden and Sgrò 2021; David et al. 2005). In Bombus, males are independent from the colony and can withstand extreme temperatures, but whether these temperatures compromise spermatozoa is unclear. Using commercially-reared Bombus impatiens males, we measured how exposure to sublethal temperatures near male critical thermal minimum and maximum impact spermatozoa viability. We measured temperature effects on spermatozoa in intact males to determine if males are potentially protecting spermatozoa. A LIVE/DEAD™ Cell Imaging Kit (Invitrogen™) and cell counter (Nexcelom Cellometer Spectrum Image Cytometry System) were used to estimate sperm count and viability in males exposed to 45°C for 85 minutes and 4°C for 85 min and 48 hours. Cells were stained with LIVE/DEAD kit immediatly after temperature exposures and were compared to control males which were held at room temperature (22°C) throughout experiment.The purpose of this study is to determine how exposure time may influence temperature effects on viability. This gap is important to address as bumble bee populations continue to decline and potential sterility to males could be a contributing factor.

My goal for this project is to be able to create box plots and violin graphs to display these results and begin to visualize what is happening in my datasets with effects of mass and temperature.

2. Dataset

The dataset I will be using for this experiment include cell counts of male spermatozoa. The data consisits of live cells, dead cells, total cells, and percent viability(live/total). Other information included in the datsets are treatment, date, colony, species, date, mass, beeid, image number, and sex. I get my data from a automated microscope and I take 3 images of each sample and average them to get a accurate calculation of the cell count. I have 2 datasets, one for cold and one for heat treatment as explained above.

I anticipate there not being many issues with this data but I want to explore some addition asepcts I could add to my boxplot such as changing the y axis to log scale or creating a histogram to see if differences are due to mass or other variables.

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.1 --
```

```
## v ggplot2 3.3.5     v purrr   0.3.4
```

```
## v tibble  3.1.6     v dplyr   1.0.7
## v tidyr   1.1.4     v stringr 1.4.0
## v readr   2.1.2     v forcats 0.5.1

## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(readxl)
library(ggplot2)
```

```
chillvia <- read_xlsx("chillcoma exp.xlsx")
heatvia <- read_xlsx("heat exp.xlsx")
```

3. Tidying Data

In this section, my goal is to tidy data, make sure all columns are correct data type, removing any potential NAs, calculating the mean live, dead, and total counts for each unique beeID, and grouping by bee ID.

```
#For the chill dataset, I using the mutate function to assign each column is the correct data type
chillvia <- chillvia %>%
  mutate(treat=factor(treat),
         time=as.numeric(time),
         live=as.numeric(live),
         dead=as.numeric(dead),
         total=as.numeric(total),
         pc_viability=as.numeric(pc_viability))

#Next I am grouping the dataset by BeeID and using summarize to calculate the means for viability, live
cvia1 <- chillvia %>%
  group_by(beeid) %>%
  summarize(via = mean(pc_viability, na.rm=TRUE),
            live = mean(live, na.rm=TRUE),
            dead = mean(dead, na.rm=TRUE),
            total= live+dead,
            treat = unique(treat),
            time = mean(time, na.rm=T),
            mass = mean(mass, na.rm=TRUE),
            colony = unique(colony))
cvia1
```

```
## # A tibble: 31 x 9
##    beeid   via  live  dead total treat      time   mass colony
##    <dbl> <dbl> <dbl> <dbl> <dbl> <fct>     <dbl>  <dbl> <chr>
##  1     1  86.2  914  155   1069  Control       0 0.118  simon
##  2     2  78.0  448. 124.   572  Chillcoma    48 0.0989 simon
##  3     3  79.1  480. 123    603. Chillcoma    48 0.123  fleetwood
##  4     4  77.4  821  432   1253  Chillcoma    48 0.154  fleetwood
##  5     5  82.0  882. 193.  1075  Chillcoma    48 0.181  mac
##  6     6  74.9  321  101    422  Control       0 0.119  mac
##  7     8  94.6 1559.  87   1646. Chillcoma    48 0.0805 fleetwood
```

```
##  8     9  86.2 1066  175   1241  Control       0 0.113  fleetwood
##  9    10  86   1624. 267.  1891  Chillcoma    48 0.104  fleetwood
## 10    11  96.2 1773.  69.7 1842. Control       0 0.140  mac
## # ... with 21 more rows
```

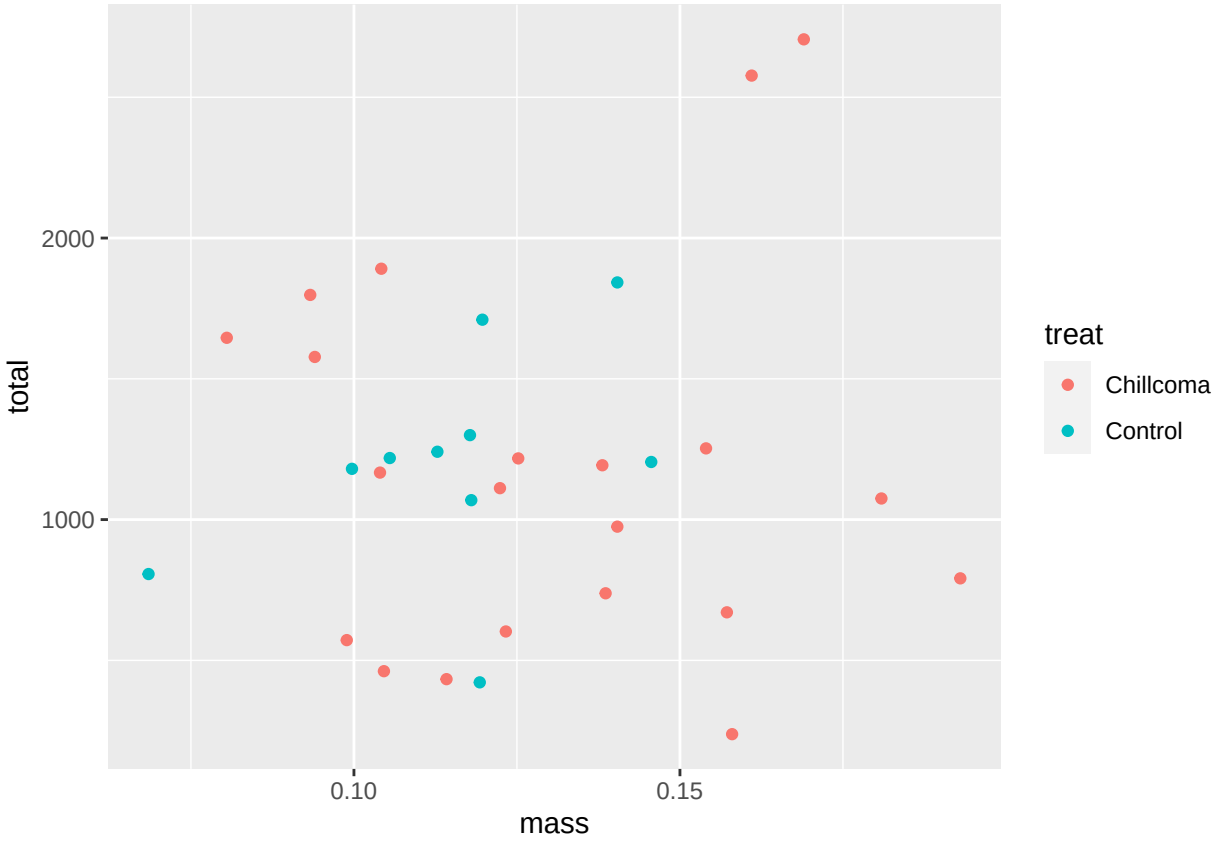
```
#For the heat dataset, I using the mutate function to assign each column is the correct data type
heatvia <- heatvia %>%
  mutate(treat=factor(treat),
         time=as.numeric(time),
         live=as.numeric(live),
         dead=as.numeric(dead),
         total=as.numeric(total),
         pc_viability=as.numeric(pc_viability))

#Next I am grouping the dataset by BeeID and using summarize to calculate the means for viability, live
hvia1 <-  heatvia %>%
  group_by(beeid) %>%
  summarize(via = mean(pc_viability, na.rm=TRUE),
            live = mean(live, na.rm=TRUE),
            dead = mean(dead, na.rm=TRUE),
            total= live+dead,
            treat = unique(treat),
            time = mean(time, na.rm=T),
            mass = mean(mass, na.rm=TRUE),
            colony = unique(colony))
hvia1
```

```
## # A tibble: 39 x 9
##    beeid   via   live  dead total treat    time   mass colony
##    <dbl> <dbl>  <dbl> <dbl> <dbl> <fct>   <dbl>  <dbl> <chr>
##  1     1  80   137.    34.3 171   control     0 0.161  aspen
##  2     2  75.1 497.   163.  661.  hs         29 0.166  willow
##  3     3  81.3 389     90.7 480.  heat       29 0.102  aspen
##  4     4  83.2 275.    56   331.  control     0 0.138  aspen
##  5     5  47.6 279    307   586   heat       29 0.197  aspen
##  6     6  69.6 244.   106.  351.  hs         29 0.105  aspen
##  7     7  80.2 732.   184   916.  control     0 0.128  willow
##  8     8  54.4 145.   121   266.  heat       36 0.0988 aspen
##  9     9  24.4   7.33  22.3  29.7 hs         36 0.0655 aspen
## 10    10  86.9 401.    62   463.  control     0 0.209  aspen
## # ... with 29 more rows
```

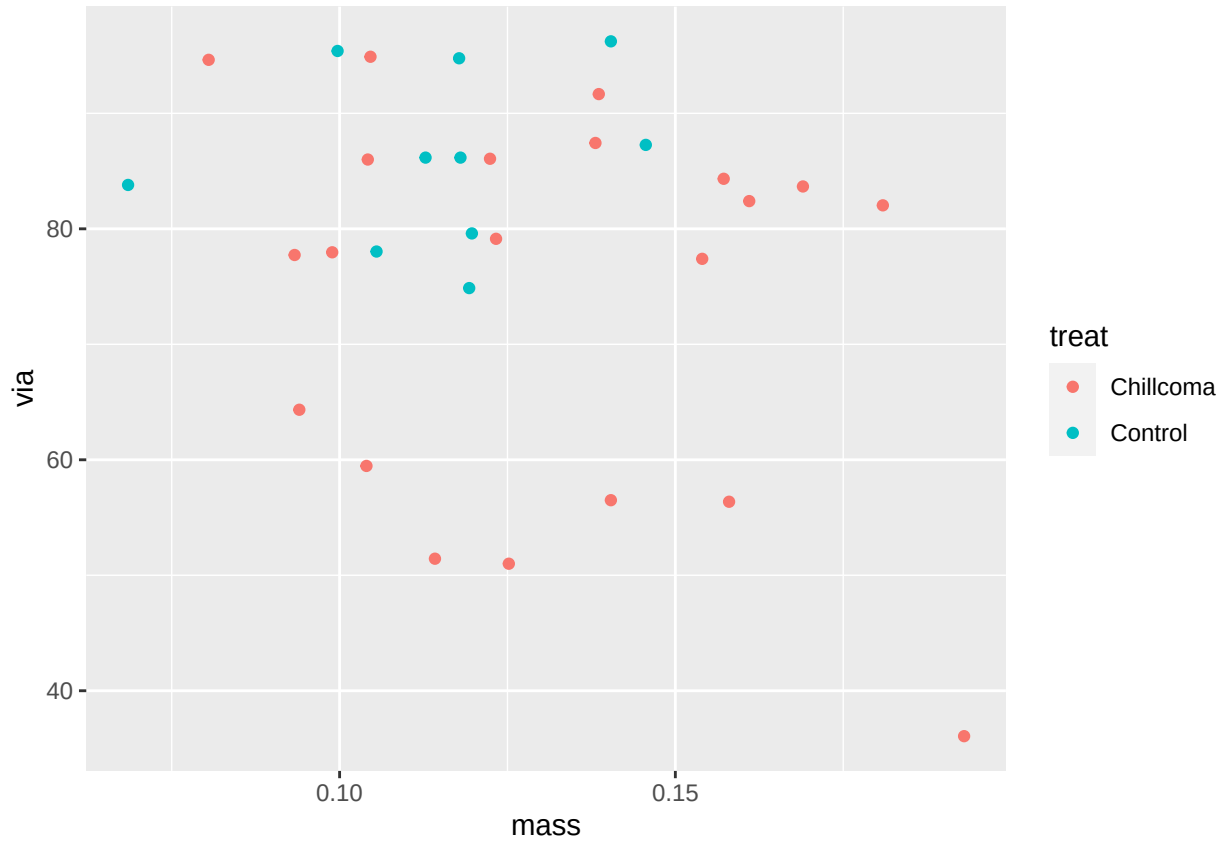
4. Scatterplots of mass

Using scatterplot, I will visualize how mass varies among treatments in both datasets. This is to ensure any results are not due to mass differences.

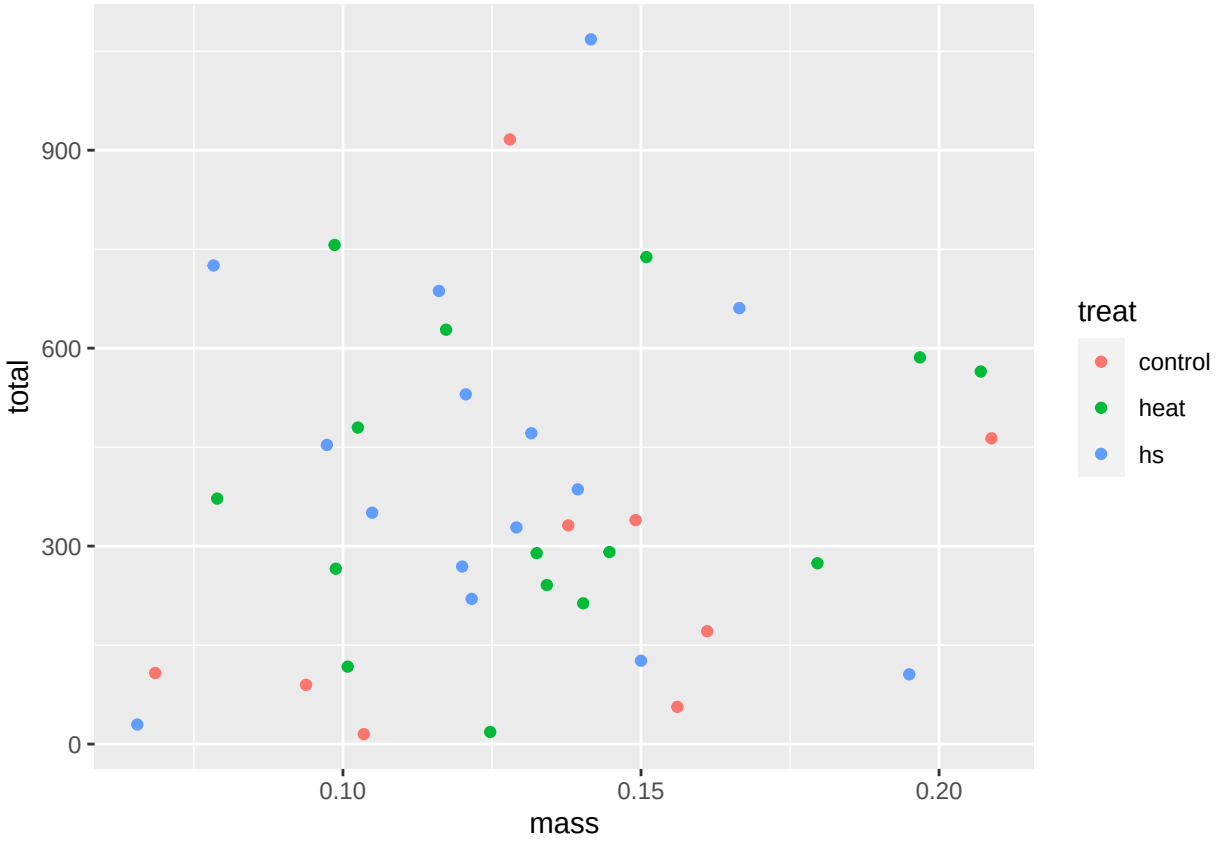
```
#comparing mass in cold treatment experiments by total cell count and % viability

cvia1 %>%
  ggplot(aes(x=mass, y=total, col=treat)) +
  geom_point()
```



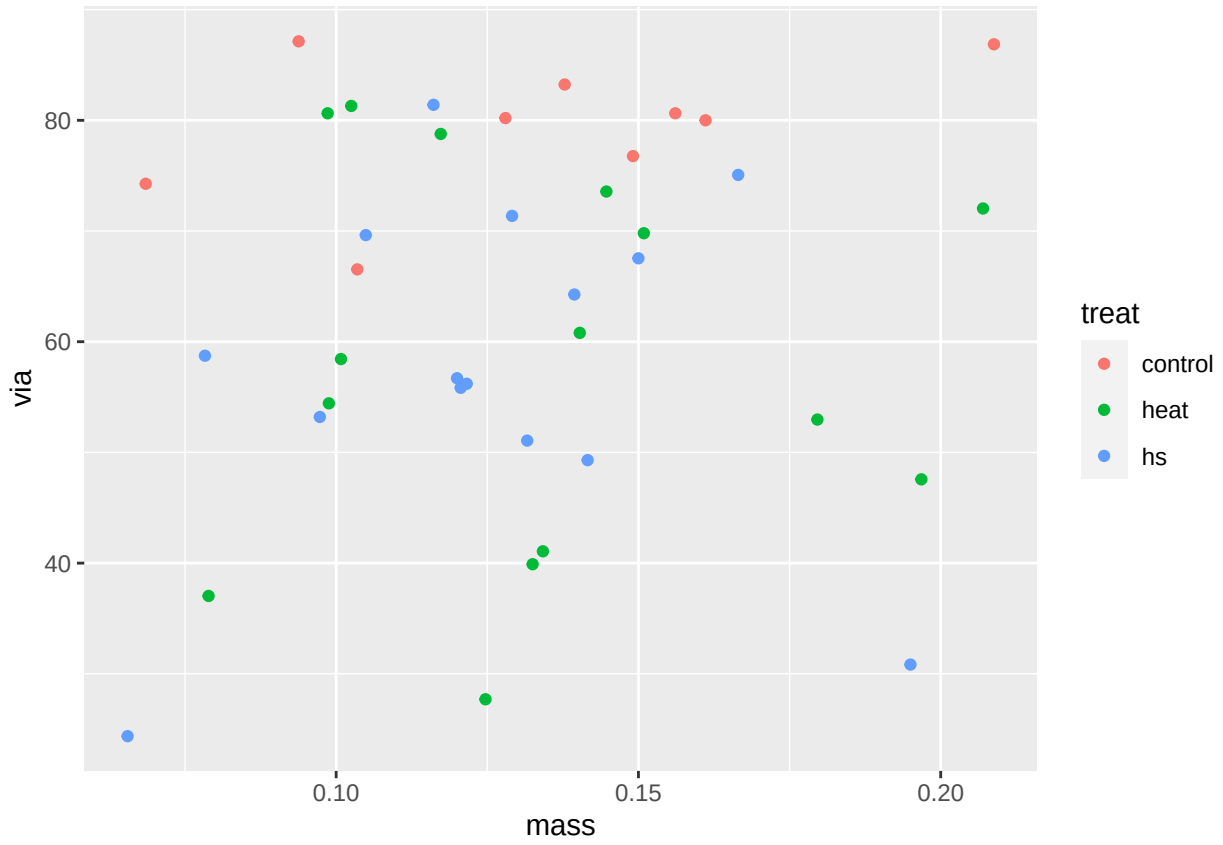

```
cvia1 %>%
  ggplot(aes(x=mass, y=via, col=treat)) +
  geom_point()
```




```
#comparing mass in heat treatment experiments by total cell count and % viability

hvia1 %>%
  ggplot(aes(x=mass, y=total, col=treat)) +
  geom_point()
```



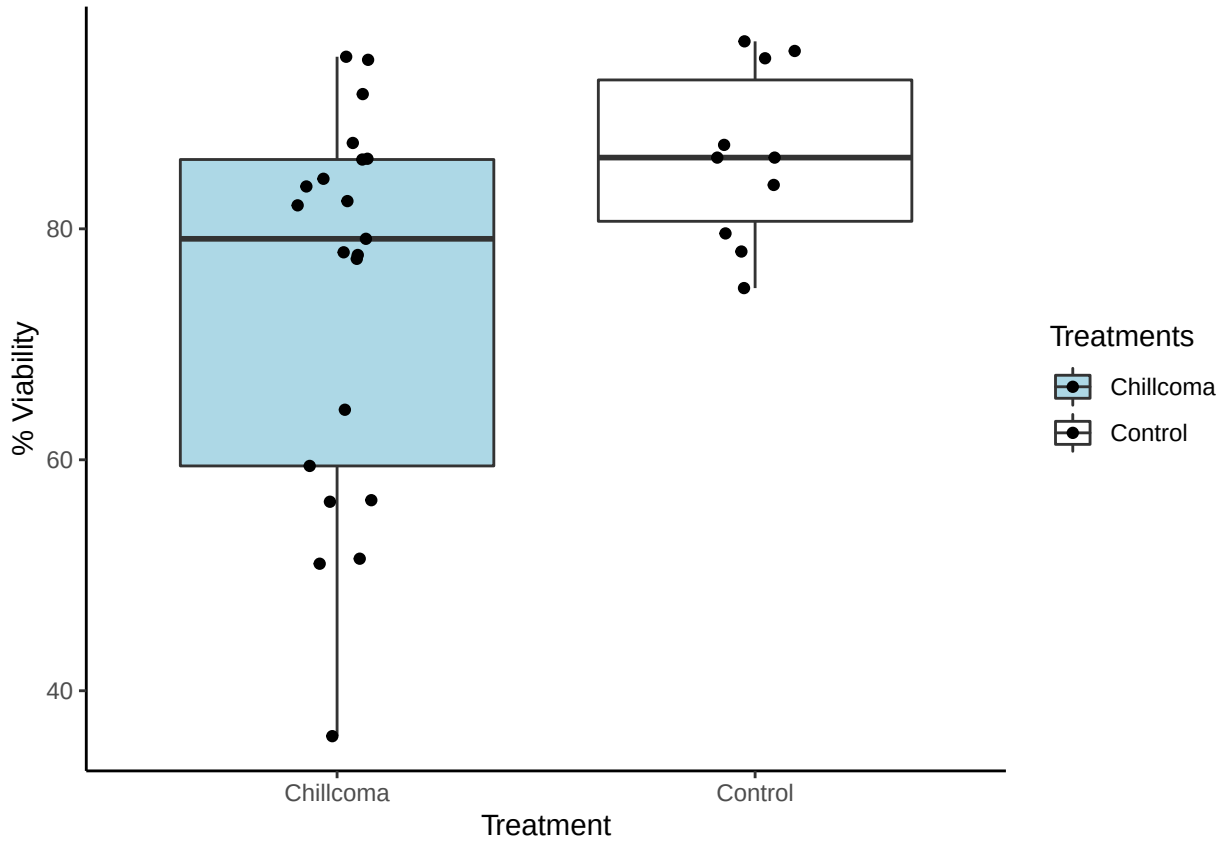

```
hvia1 %>%
  ggplot(aes(x=mass, y=via, col=treat)) +
  geom_point()
```



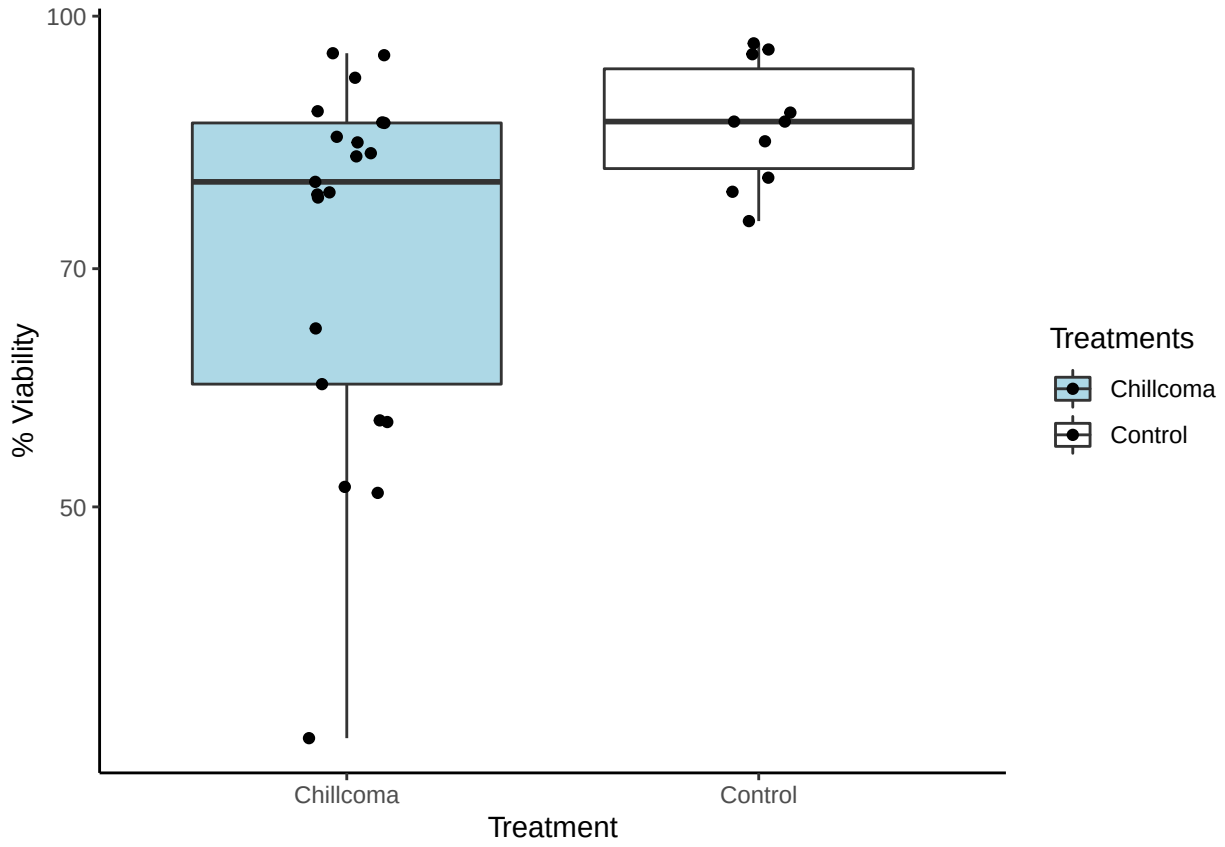
5. Plotting Viability in both datasets with boxplots

Using boxplots, I will visualize how temperature impacts spermatozoa viability in male bumble bees. The first box plot will compare males in control group and males exposed to 4C for 48 hours. This group is called 'coma.' The second plot will compare males who were exposed to 45C and either hit heat stupor (hs) or were just exposed to heat and did not hit heat stupor (heat).

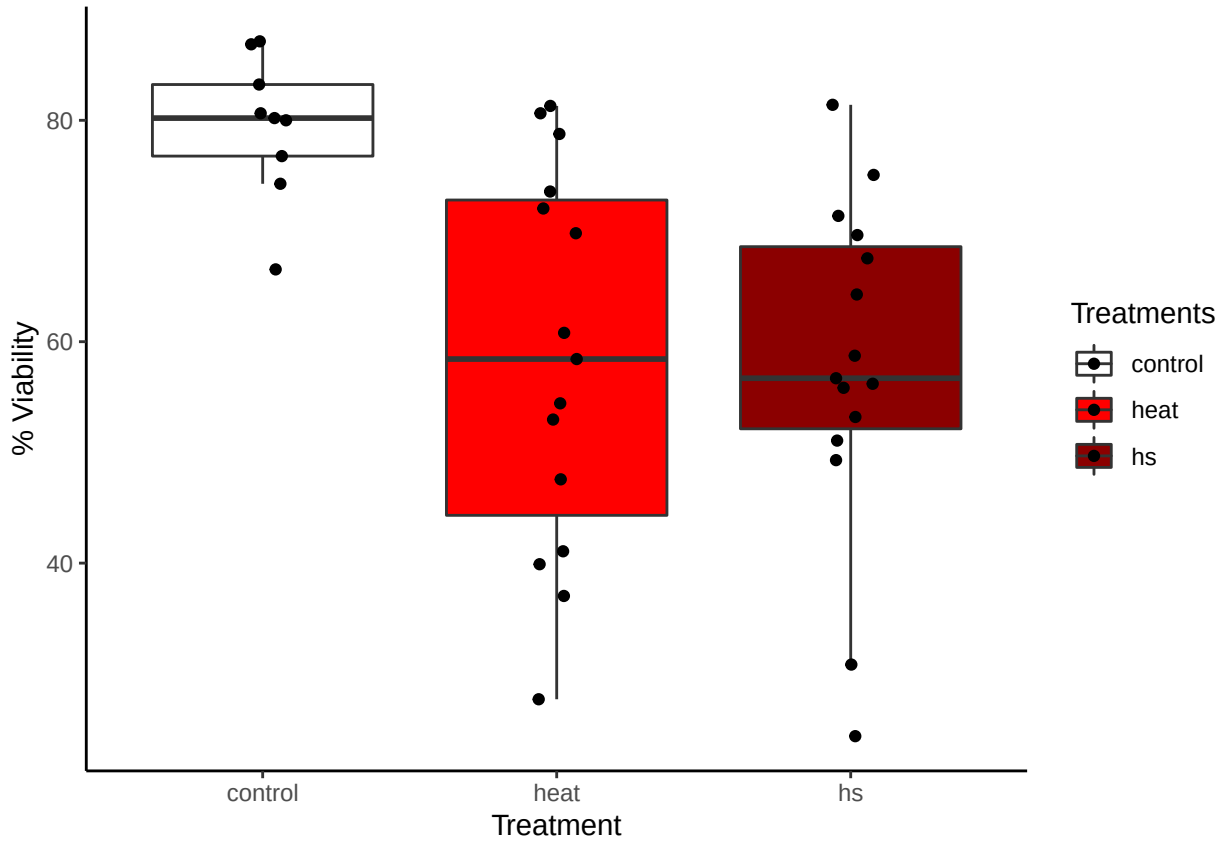
```
#cold treatment boxplot
p1 <- cvia1 %>%
  ggplot(aes(x=treat, y=via, fill=treat)) +
  geom_boxplot(outlier.shape = NA) +
  theme_classic() +
  scale_fill_manual(values=c("light blue","white")) +
  labs(x="Treatment", y="% Viability") +
  labs(fill='Treatments') +
  geom_jitter(position = position_jitterdodge())

p1
```



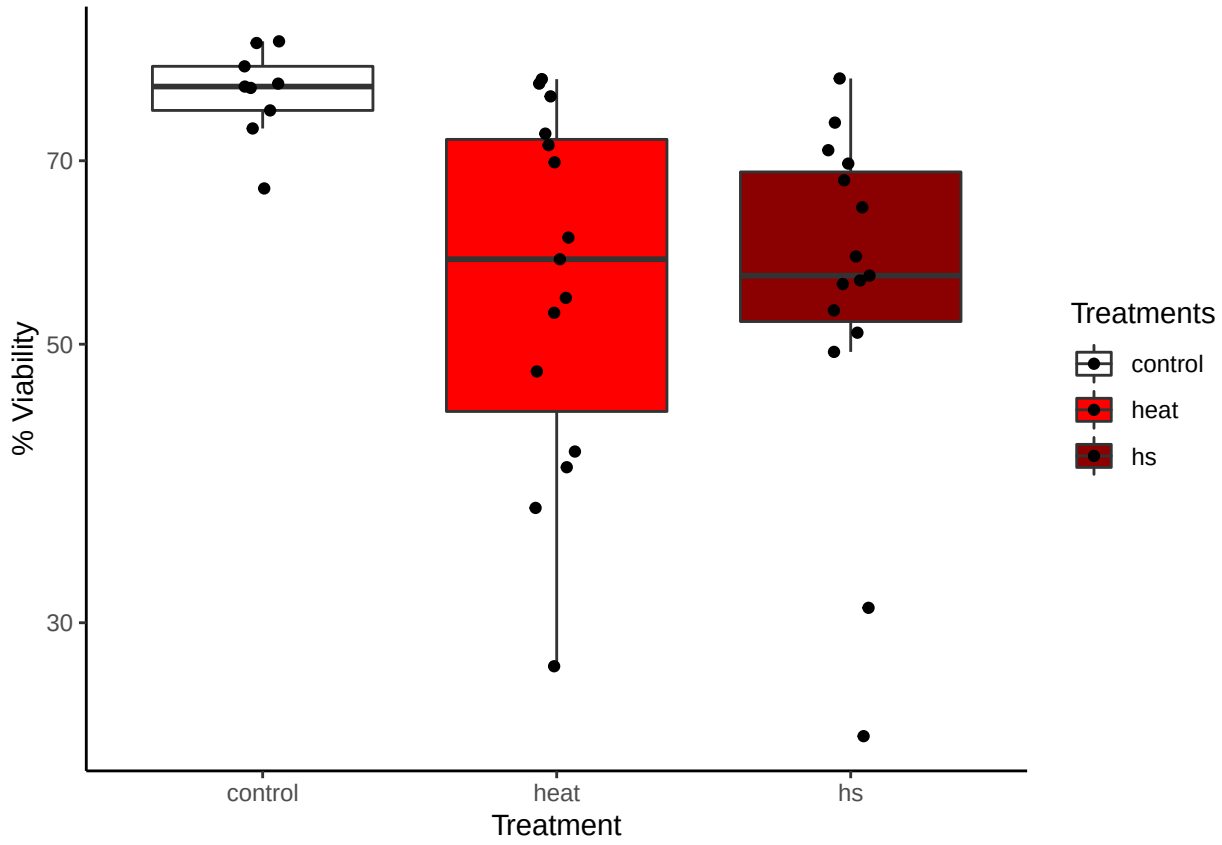

```
#cold treatment boxplot with y axis log scale to see if data is better visualized
p2 <- cvia1 %>%
  ggplot(aes(x=treat, y=via, fill=treat)) +
  geom_boxplot(outlier.shape = NA) +
  theme_classic() +
  scale_fill_manual(values=c("light blue","white")) +
  labs(x="Treatment", y="% Viability") +
  labs(fill='Treatments') +
  scale_y_continuous(trans='log10') +
  geom_jitter(position = position_jitterdodge())

p2
```




```
#heat treatment boxplot
p3 <- hvia1 %>%
 ggplot(aes(x=treat, y=via, fill=treat)) +
  geom_boxplot(outlier.shape = NA) +
  scale_fill_manual(values=c("white","red", "dark red")) +
  theme_classic() +
  labs(x="Treatment", y="% Viability") +
  labs(fill='Treatments') +
  geom_jitter(position = position_jitterdodge())

p3
```



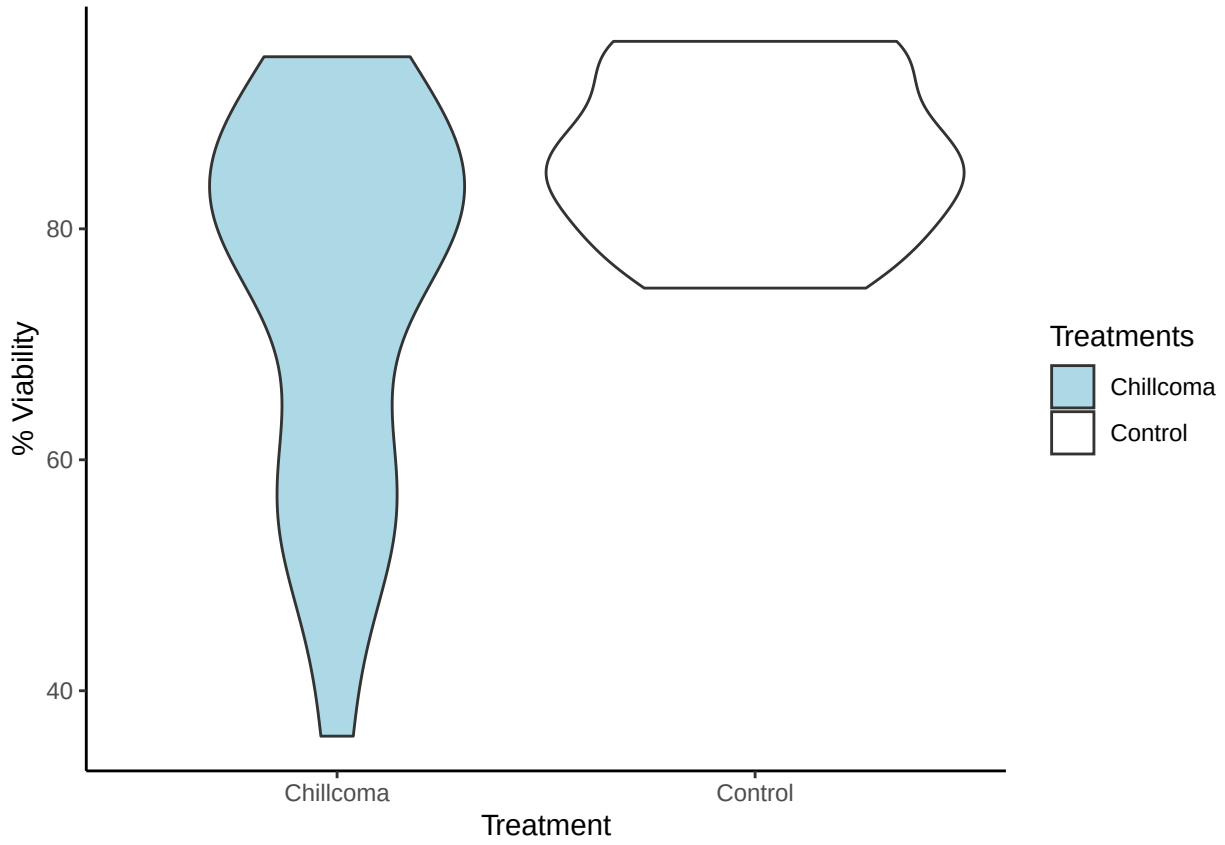

```
#heat treatment boxplot with y axis log scale to see if data is better visualized
p4 <- hvia1 %>%
 ggplot(aes(x=treat, y=via, fill=treat)) +
  geom_boxplot(outlier.shape = NA) +
  scale_fill_manual(values=c("white","red", "dark red")) +
  theme_classic() +
  labs(x="Treatment", y="% Viability") +
  labs(fill='Treatments') +
  scale_y_continuous(trans='log10') +
  geom_jitter(position = position_jitterdodge())

p4
```



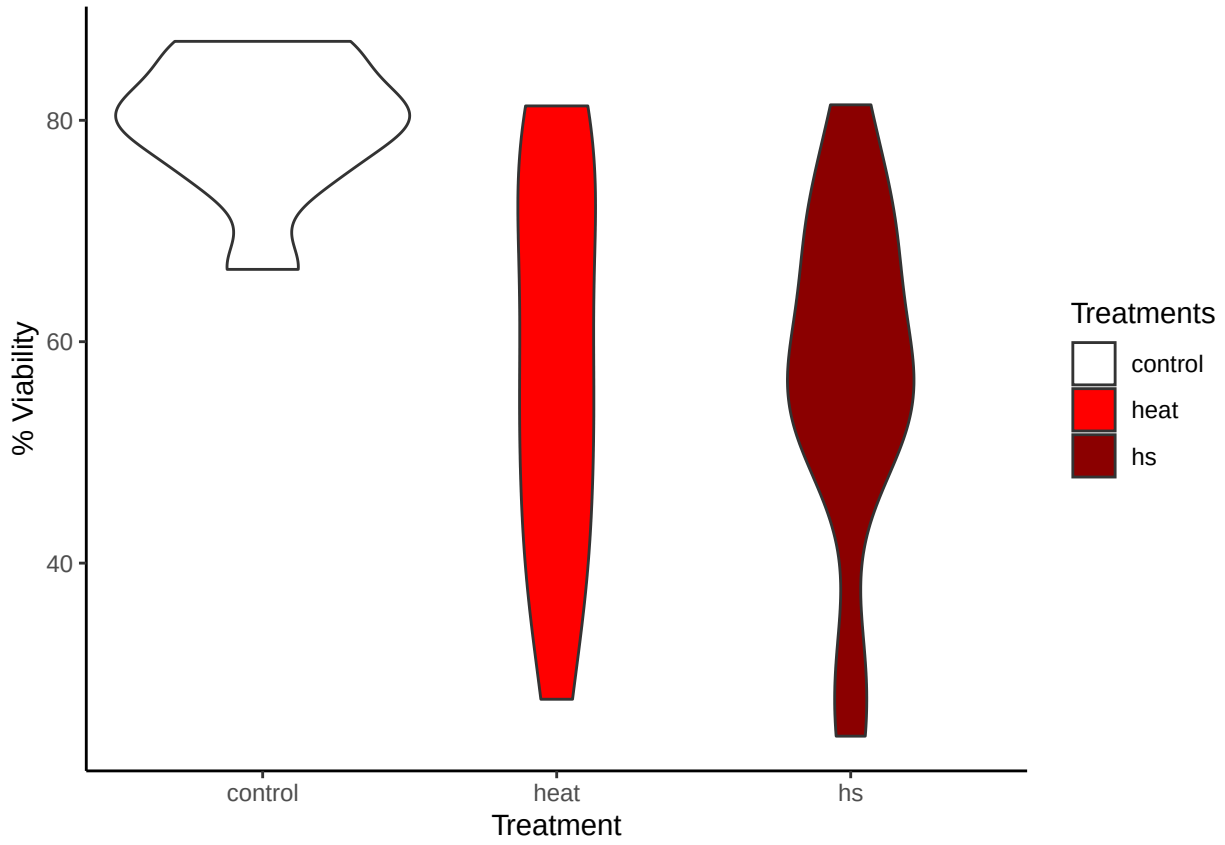
6. Plotting Viability in both datasets with violin plots

Another way to visualize this dataset could be using violin plots.

```
#cold treatment violin plot
cvia1 %>%
 ggplot(aes(x=treat, y=via, fill=treat)) +
    geom_violin(width=1) +
    theme_classic() +
    scale_fill_manual(values=c("light blue","white")) +
    labs(x="Treatment", y="% Viability") +
    labs(fill='Treatments')
```




```
#heat treatment violin plot
hvia1 %>%
 ggplot(aes(x=treat, y=via, fill=treat)) +
    geom_violin(width=1) +
    theme_classic() +
    scale_fill_manual(values=c("white","red", "dark red")) +
    labs(x="Treatment", y="% Viability") +
    labs(fill='Treatments')
```



7. References

David, J. R., L. O. Araripe, M. Chakir, H. Legout, B. Lemos, G. Pétavy, C. Rohmer, D. Joly, and B. Moreteau. 2005. "Male Sterility at Extreme Temperatures: A Significant but Neglected Phenomenon for Understanding Drosophila Climatic Adaptations." *Journal of Evolutionary Biology* 18 (4): 838–46. https://doi.org/10.1111/j.1420-9101.2005.00914.x.

Heerwaarden, Belinda van, and Carla M. Sgrò. 2021. "Male Fertility Thermal Limits Predict Vulnerability to Climate Warming." *Nature Communications* 12 (1): 2214. https://doi.org/10.1038/s41467-021-22546-w.

Kerr, J. T., A. Pindar, P. Galpern, L. Packer, S. G. Potts, S. M. Roberts, P. Rasmont, et al. 2015. "Climate Change Impacts on Bumblebees Converge Across Continents." *Science* 349 (6244): 177–80. https://doi.org/10.1126/science.aaa7031.