data_exploration

Erica Griffin

2025-09-09

SICB Project Data Exploration

This is a collection of simple plots that I ran of the raw data to begin to observe patterns. I wanted to look at the stress series to see how it's correlating with temperature, as well as look as some co-variations between variables to start to see what may be important. Overall, here are some notes / patterns that I see:

- similar patterns between plots when using average daytime temperature versus average 3 day temperature. but a lot less observations below the 18.5 deg threshold when using the 3 day average
- for all sex differences, the lines are so close it seems like there likely is no significant difference, but overall see a similar pattern that baseline cort spikes once temps drop. there are also seems to be a similar pattern when looking at stress-induced cort.
- I used encounter day of year (not sure if this is an equal measure to clutch initiation date), but it does look like this does have some sort of relationship with baseline cort
- there seems to be a correlation between body mass and baseline cort
- I looked at how average daytime temp differs (just base stats) between years - 2013, 2014, 2019, and 2021 look to have colder days and more variation

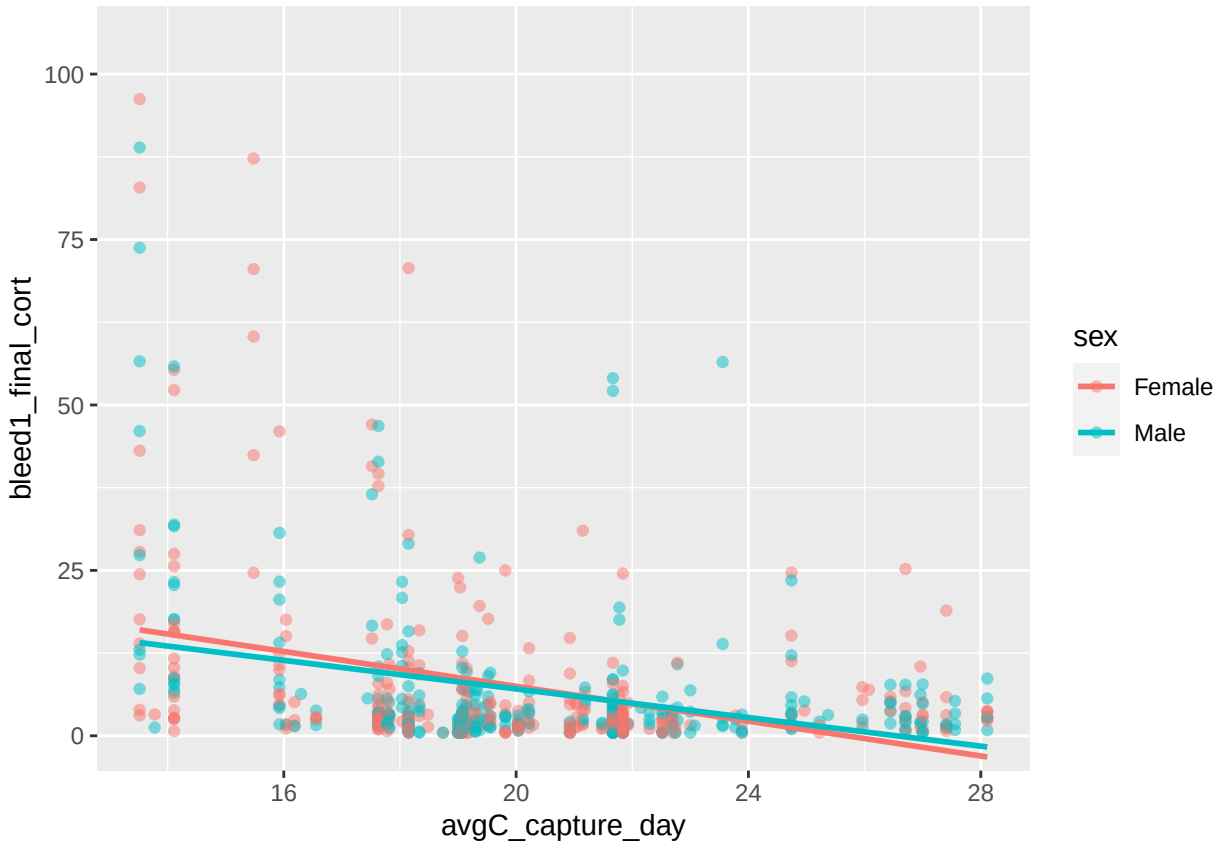
Plot 1: Baseline cort ~ average daytime temp

```
vis1 <- ggplot(df, aes(x = avgC_capture_day, y = bleed1_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis1
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```

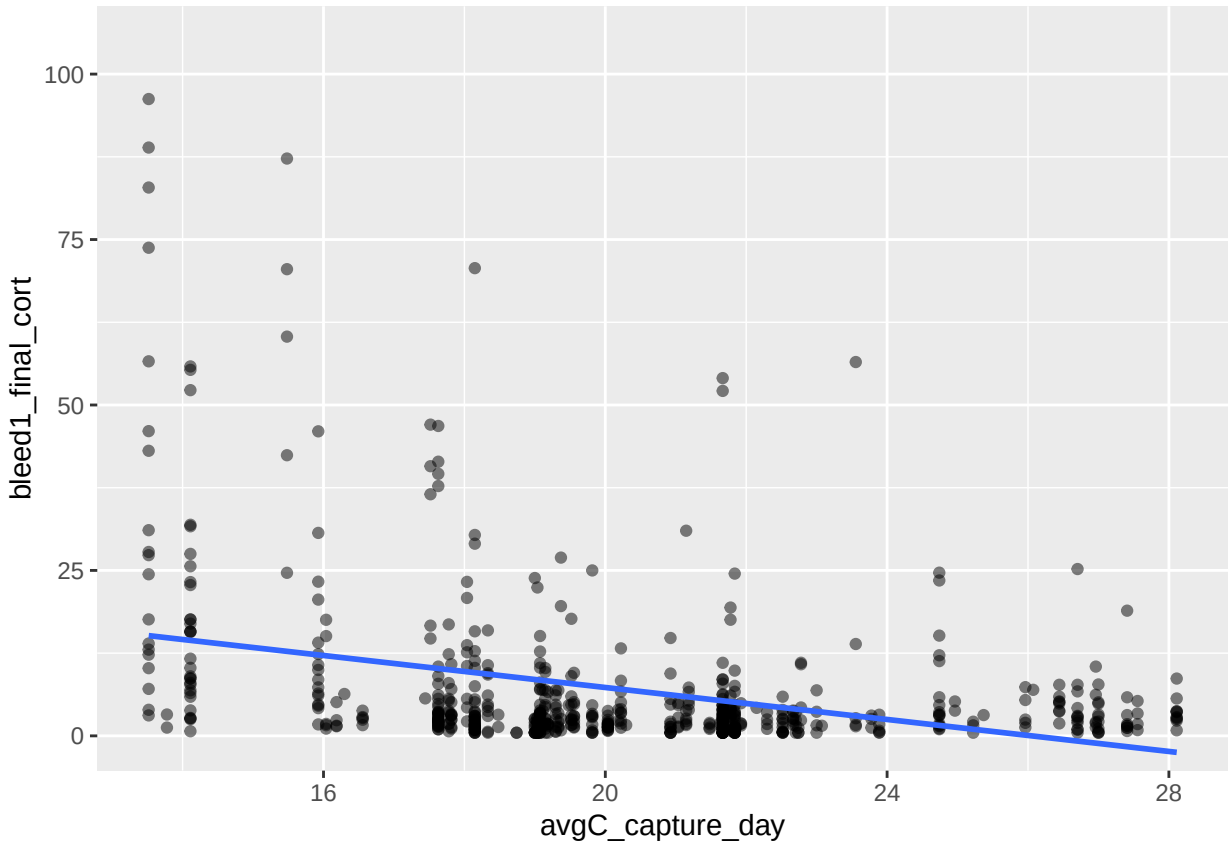



```
vis1.1 <- ggplot(df, aes(x = avgC_capture_day, y = bleed1_final_cort)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis1.1
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```



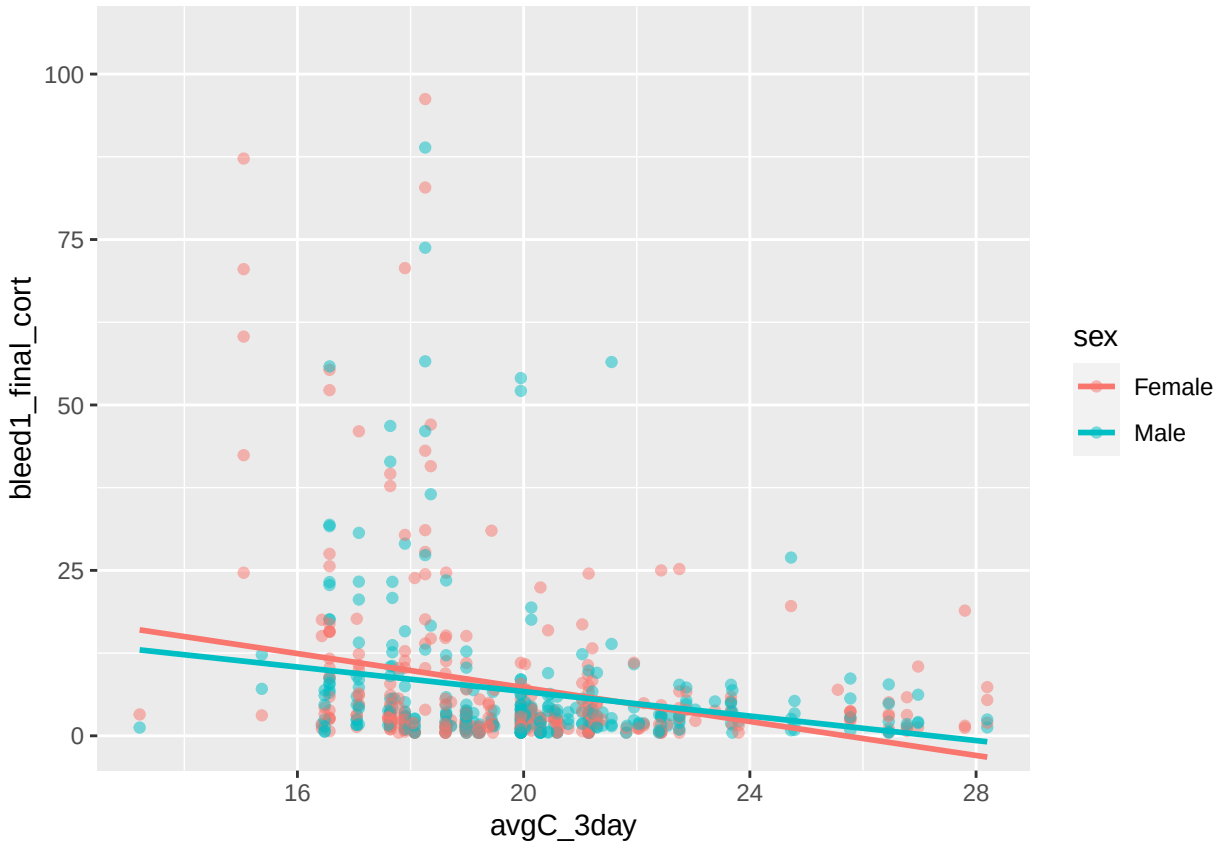
Plot 2: Baseline cort ~ average 3 day temp

```
vis2 <- ggplot(df, aes(x = avgC_3day, y = bleed1_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis2
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```



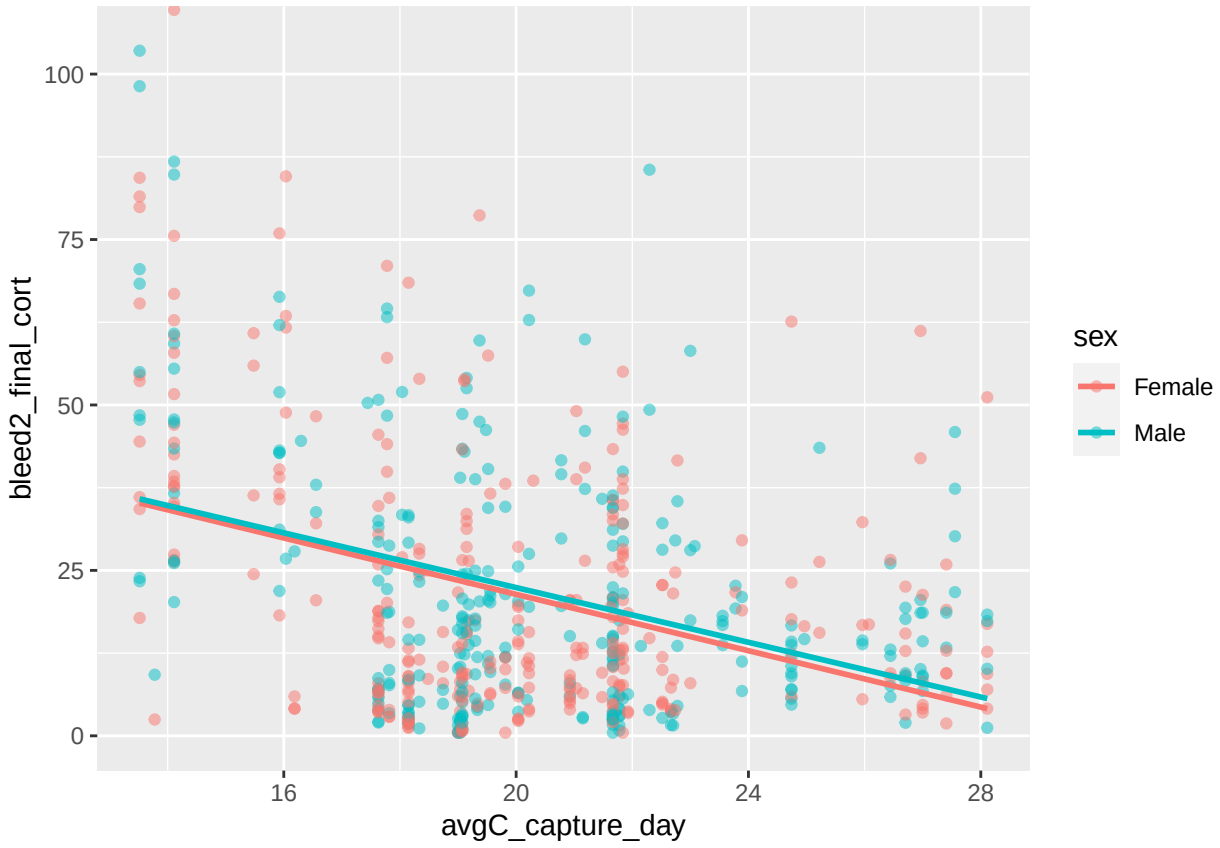
Plot 3: Stress-induced cort ~ average daytime temp

```
vis3 <- ggplot(df, aes(x = avgC_capture_day, y = bleed2_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis3
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 36 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 36 rows containing missing values (`geom_point()`).
```



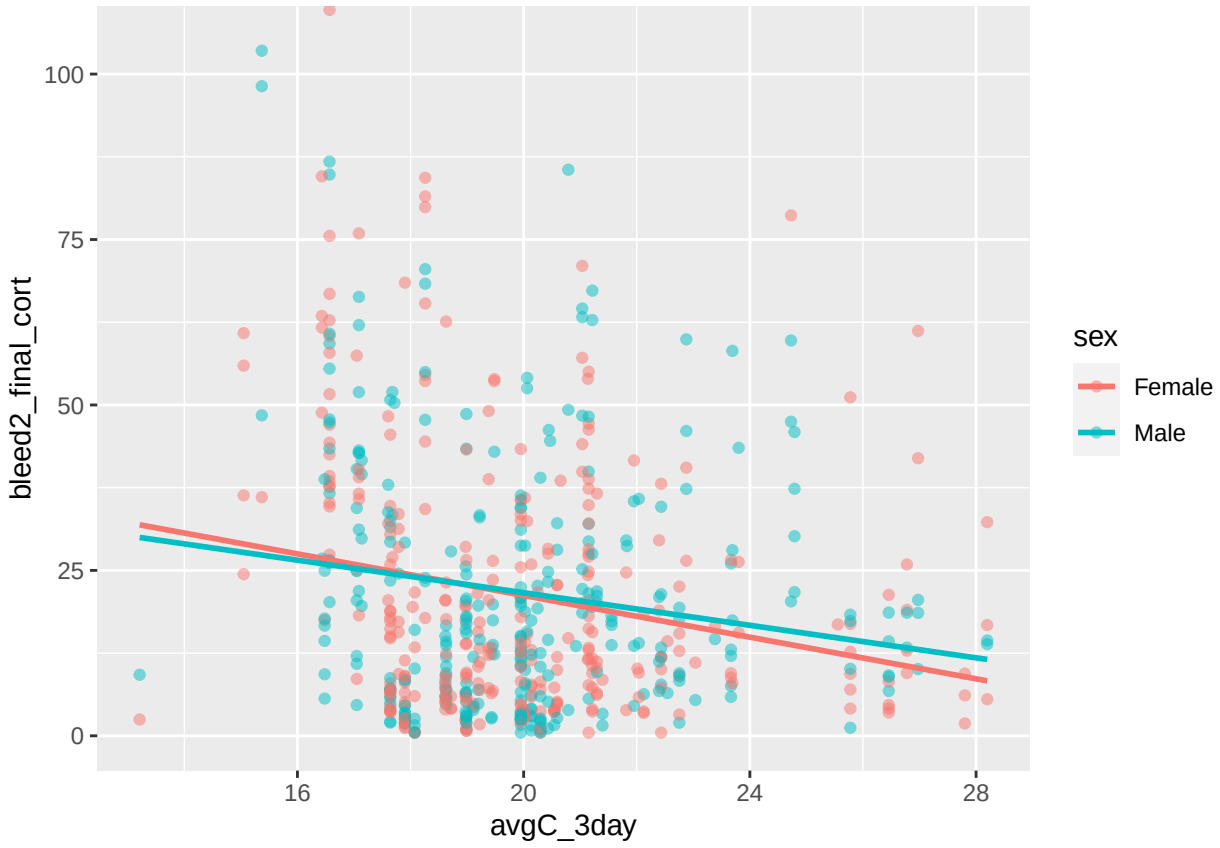
Plot 4: Stress-induced cort ~ average 3 day temp

```
vis4 <- ggplot(df, aes(x = avgC_3day, y = bleed2_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis4
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 36 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 36 rows containing missing values (`geom_point()`).
```



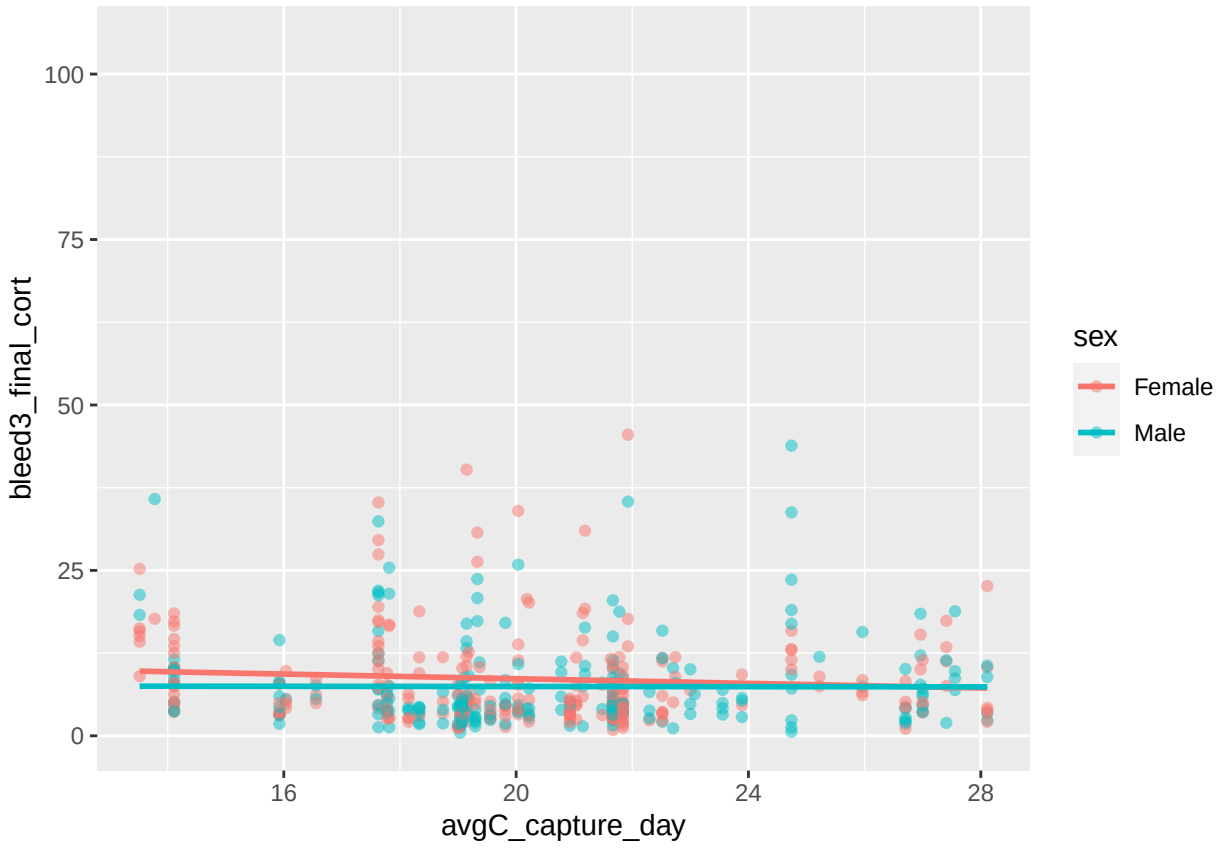
Plot 5: Dex ~ average daytime temp

```
vis5 <- ggplot(df, aes(x = avgC_capture_day, y = bleed3_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis5
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 170 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 170 rows containing missing values (`geom_point()`).
```



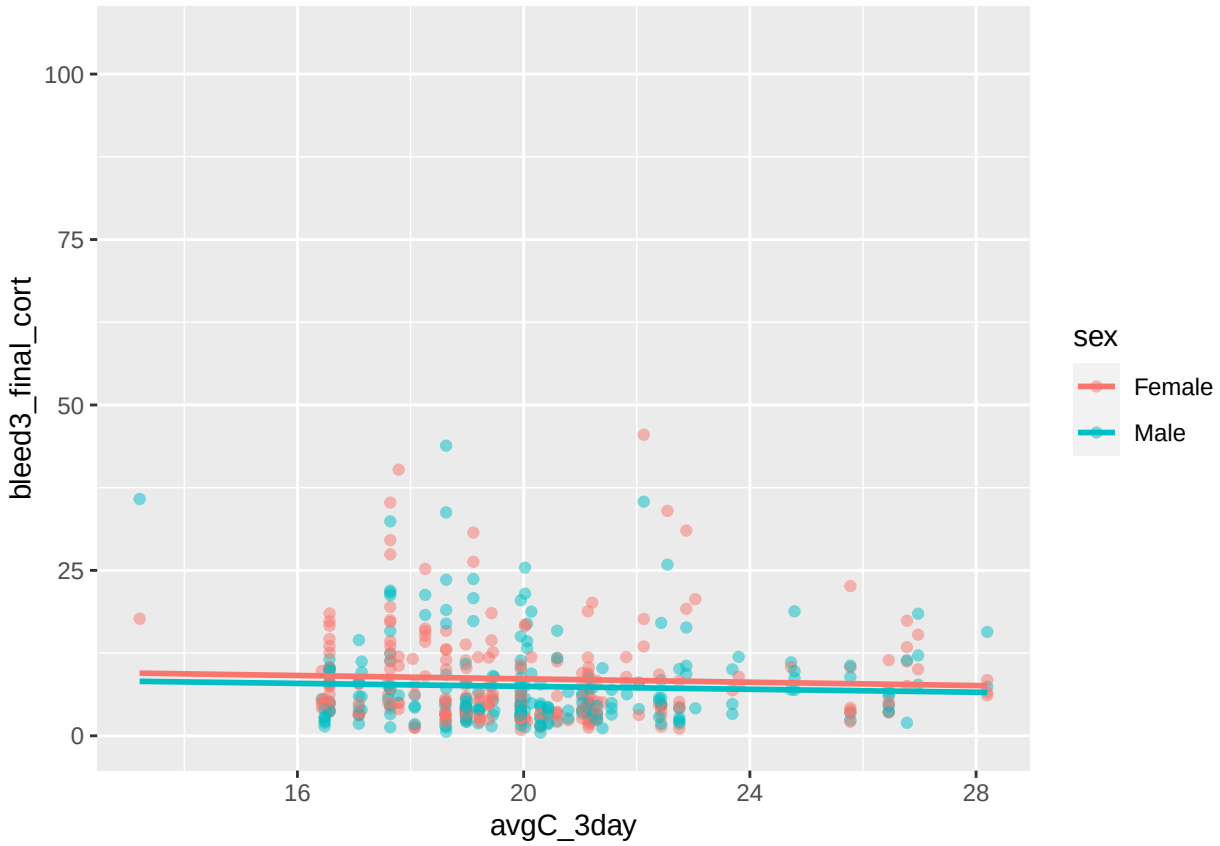
Plot 6: Dex ~ average 3 day temp

```
vis6 <- ggplot(df, aes(x = avgC_3day, y = bleed3_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis6
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 170 rows containing non-finite values (`stat_smooth()`).
```

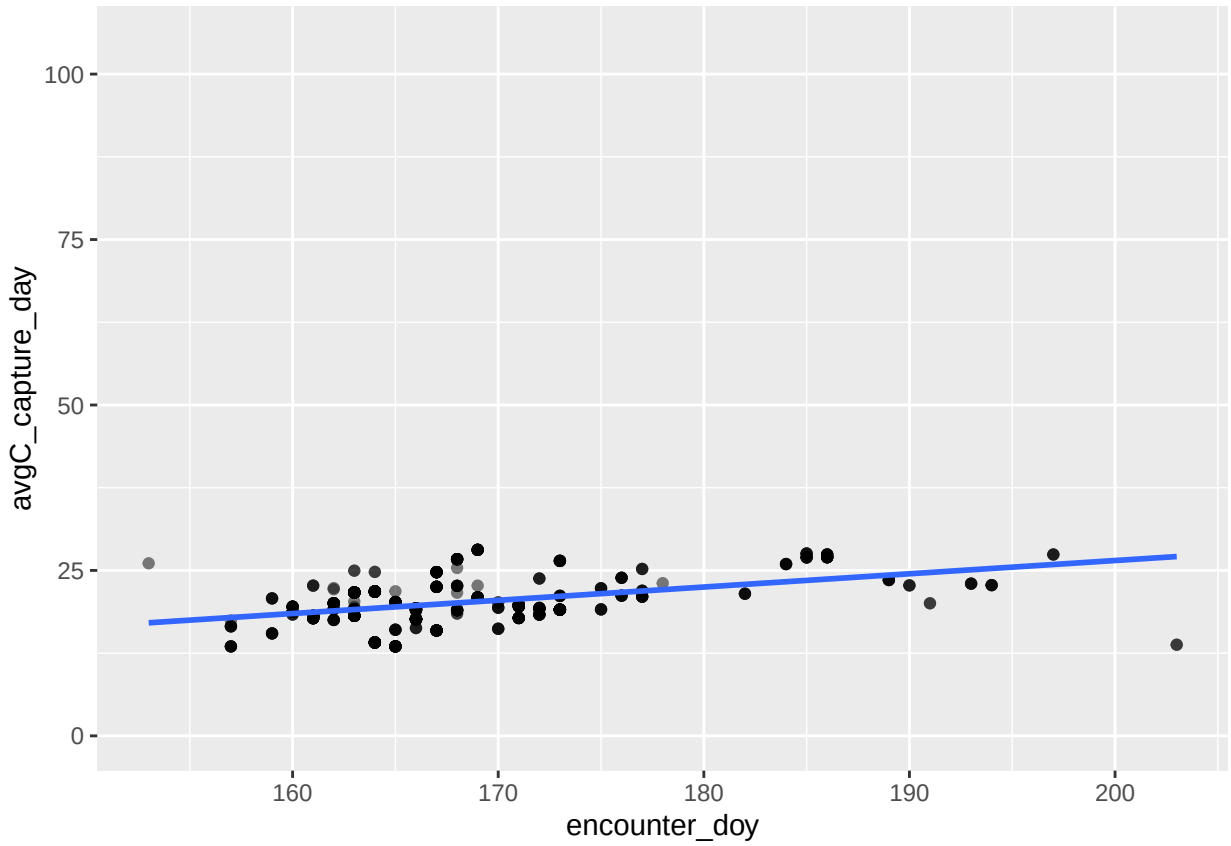
```
## Warning: Removed 170 rows containing missing values (`geom_point()`).
```



Plot 7: Encounter day of year ~ average daytime temp

```
vis7 <- ggplot(df, aes(x = encounter_doy, y = avgC_capture_day)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis7
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



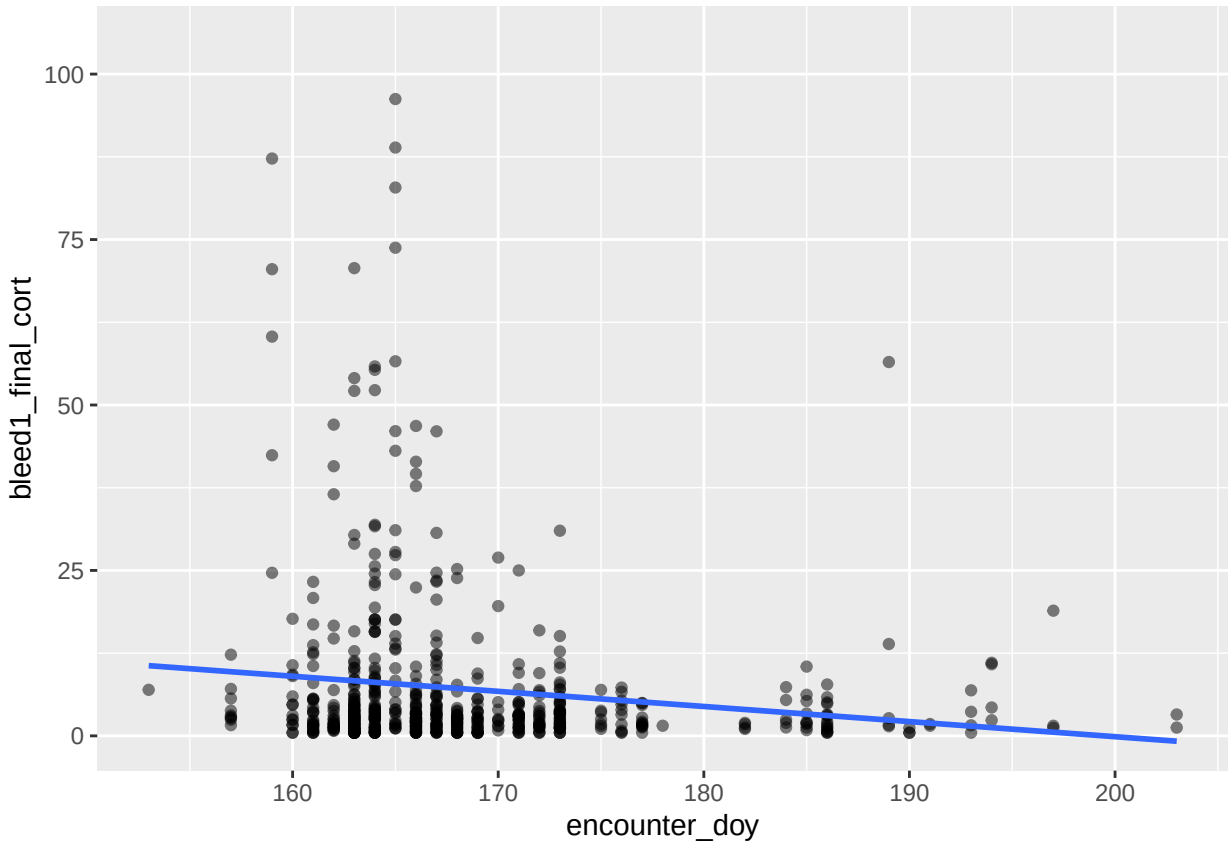
Plot 8: Encounter day of year ~ baseline cort

```
vis8 <- ggplot(df, aes(x = encounter_doy, y = bleed1_final_cort)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis8
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```



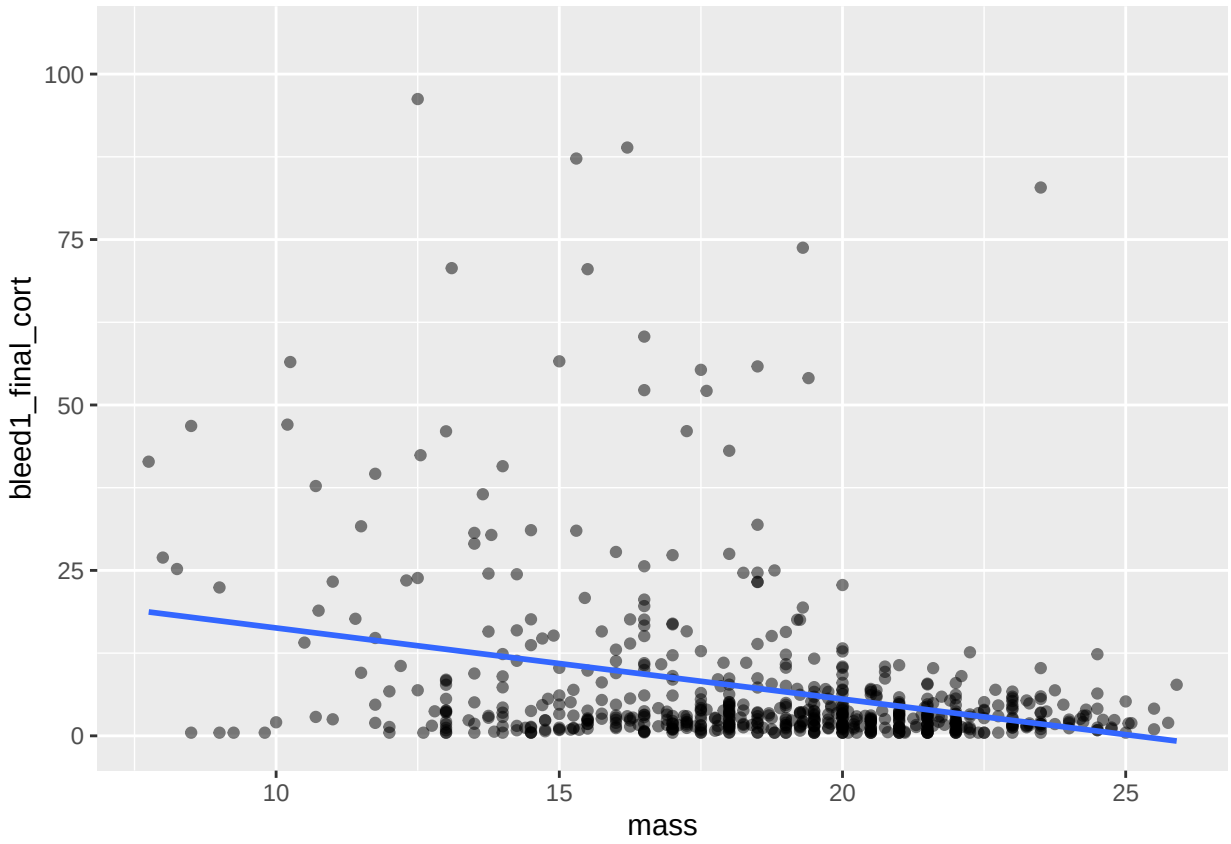
Plot 9: Body mass ~ baseline cort

```
vis9 <- ggplot(df, aes(x = mass, y = bleed1_final_cort)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis9
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

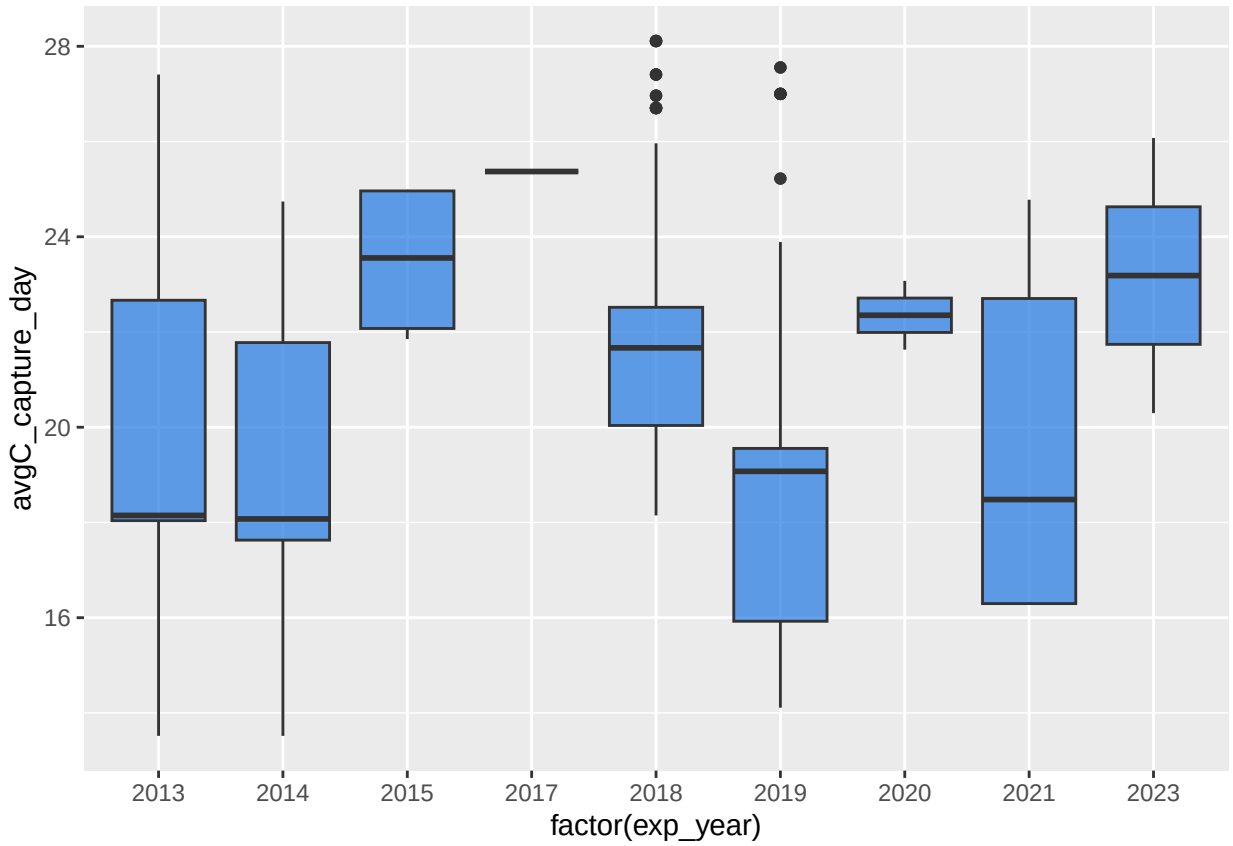
```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```



Plot 10: Observation year ~ average daytime temp

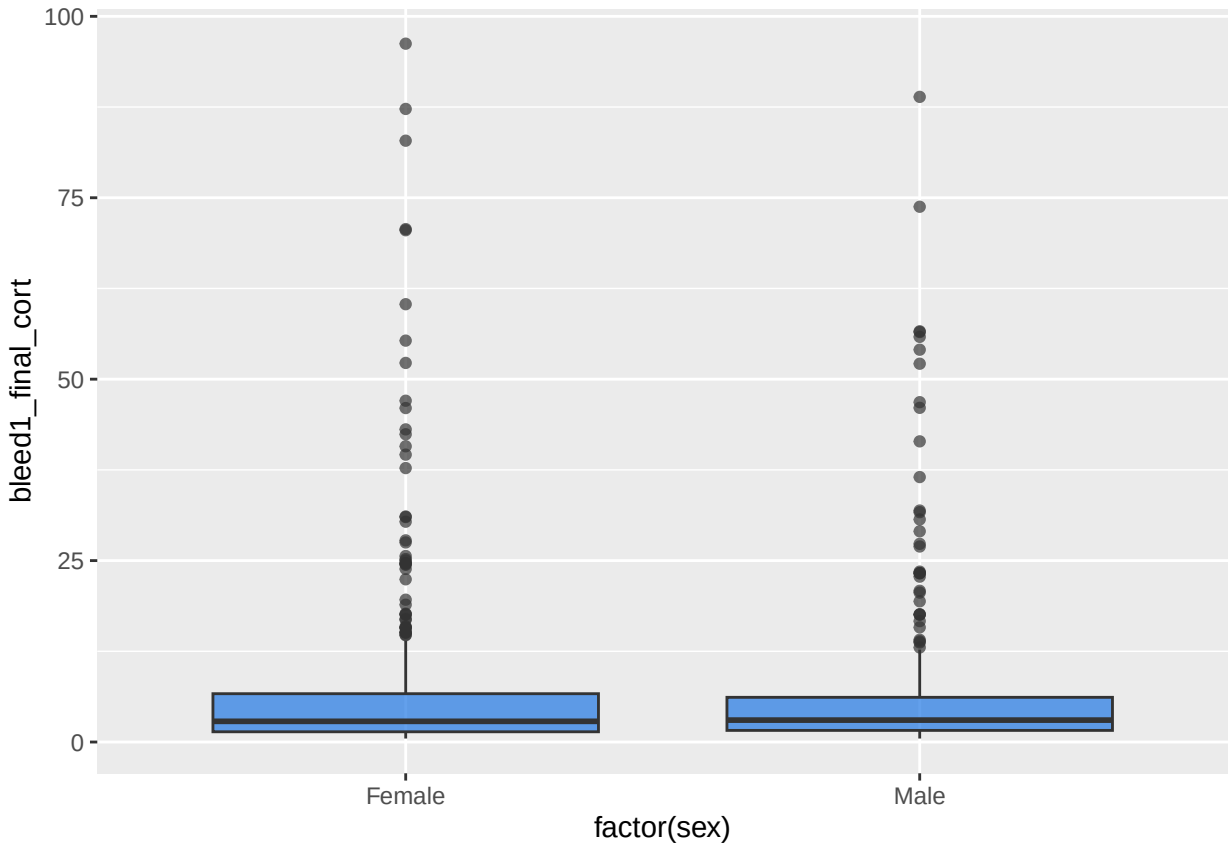
```
vis10 <- ggplot(df, aes(x = factor(exp_year), y = avgC_capture_day)) +
  geom_boxplot(fill = "#1f77e4", alpha = 0.7)
vis10
```



Plot 11: Baseline cort ~ sex

```
vis11 <- ggplot(df, aes(x = factor(sex), y = bleed1_final_cort)) +
  geom_boxplot(fill = "#1f77e4", alpha = 0.7)
vis11
```

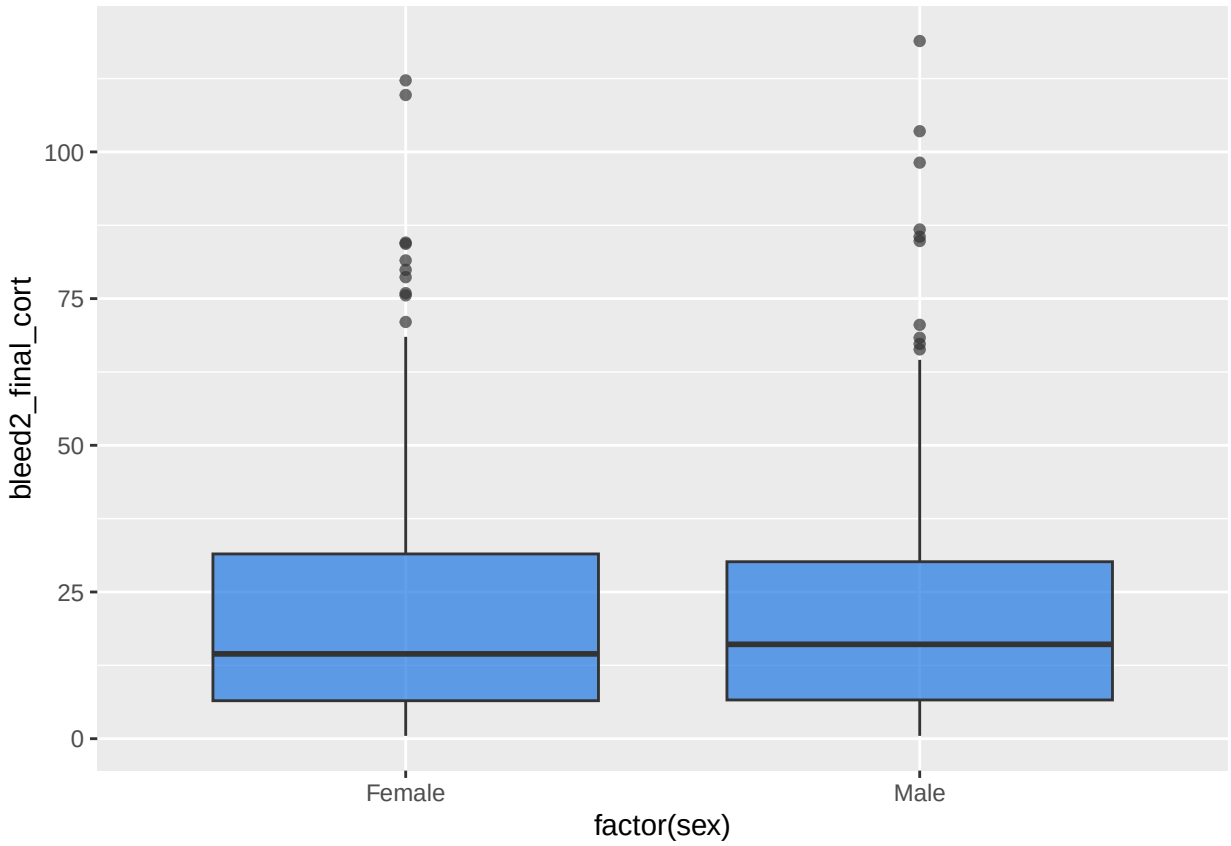
```
## Warning: Removed 28 rows containing non-finite values (`stat_boxplot()`).
```



Plot 12: Stress-induced cort ~ sex

```
vis12 <- ggplot(df, aes(x = factor(sex), y = bleed2_final_cort)) +
  geom_boxplot(fill = "#1f77e4", alpha = 0.7)
vis12
```

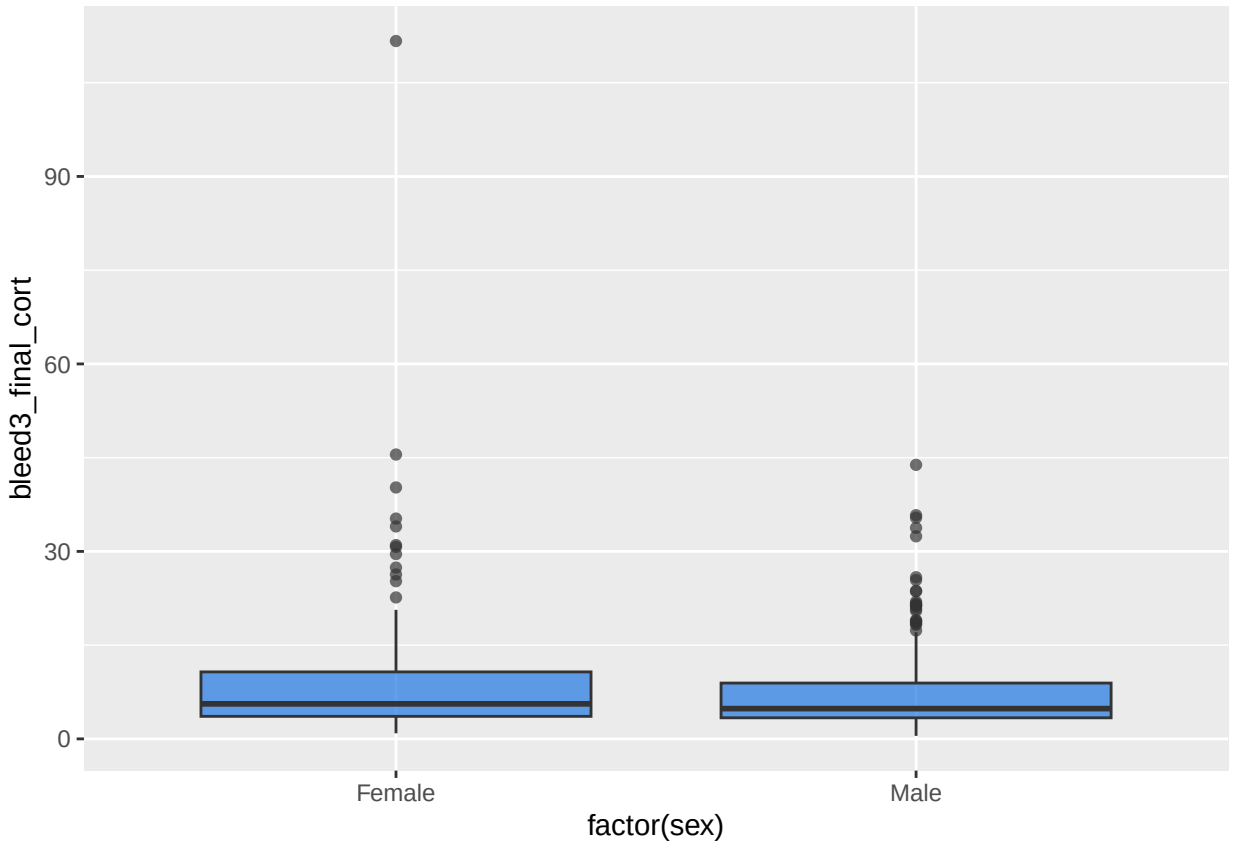
```
## Warning: Removed 36 rows containing non-finite values (`stat_boxplot()`).
```



Plot 13: Dex ~ sex

```
vis13 <- ggplot(df, aes(x = factor(sex), y = bleed3_final_cort)) +
  geom_boxplot(fill = "#1f77e4", alpha = 0.7)
vis13
```

```
## Warning: Removed 170 rows containing non-finite values (`stat_boxplot()`).
```



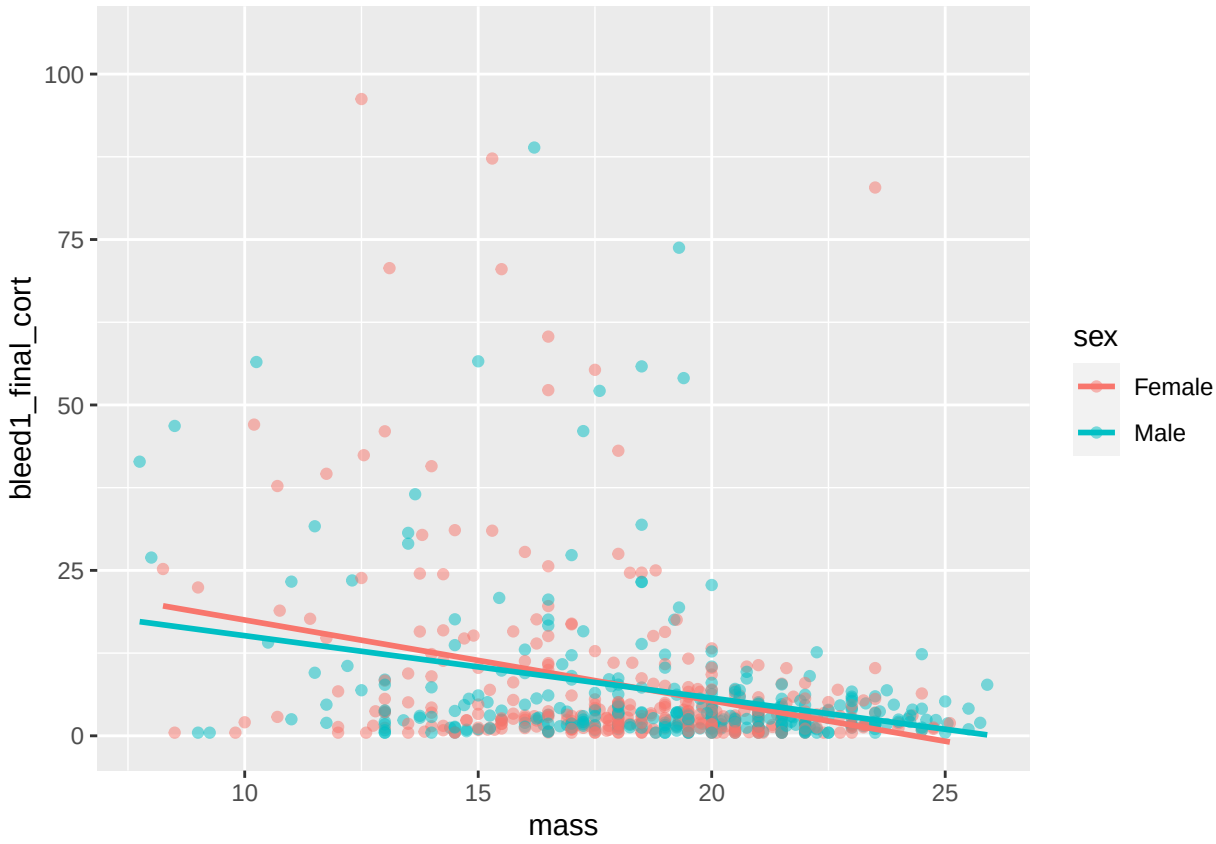
Plot 14: Baseline cort ~ mass by sex

```
vis14 <- ggplot(df, aes(x = mass, y = bleed1_final_cort, color = sex)) +
  geom_point(alpha = 0.5) +
  coord_cartesian(ylim = c(0, 105)) +
  geom_smooth(method = "lm", se = FALSE)
vis14
```

```
## `geom_smooth()` using formula = 'y ~ x'
```

```
## Warning: Removed 28 rows containing non-finite values (`stat_smooth()`).
```

```
## Warning: Removed 28 rows containing missing values (`geom_point()`).
```


bleed1_final_cort
100
75
50
25
0
10
15
20
25
mass
sex
Female
Male