# HW2_yc4384

yc4384_Yangyang_Chen

2024-09-28

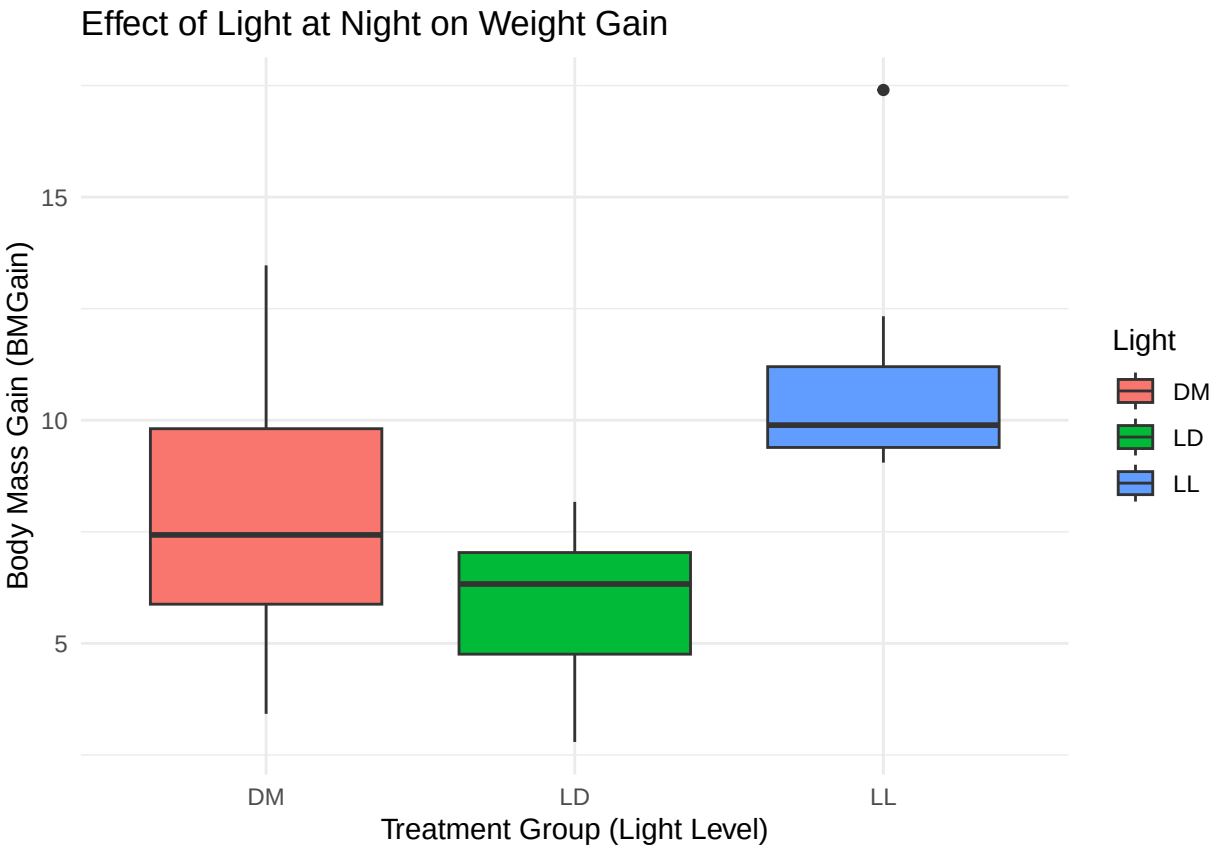
## Question 1

We are interested in the causal effect of light at night on weight gain. Plot the outcome by treatment group.

```
data = read_csv("light.csv")
```

```
## Rows: 27 Columns: 9
## -- Column specification --------------------------------------------------------
## Delimiter: ","
## chr (2): Light, GlucoseInt
## dbl (7): BMGain, Corticosterone, DayPct, Consumption, GTT15, GTT120, Activity
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

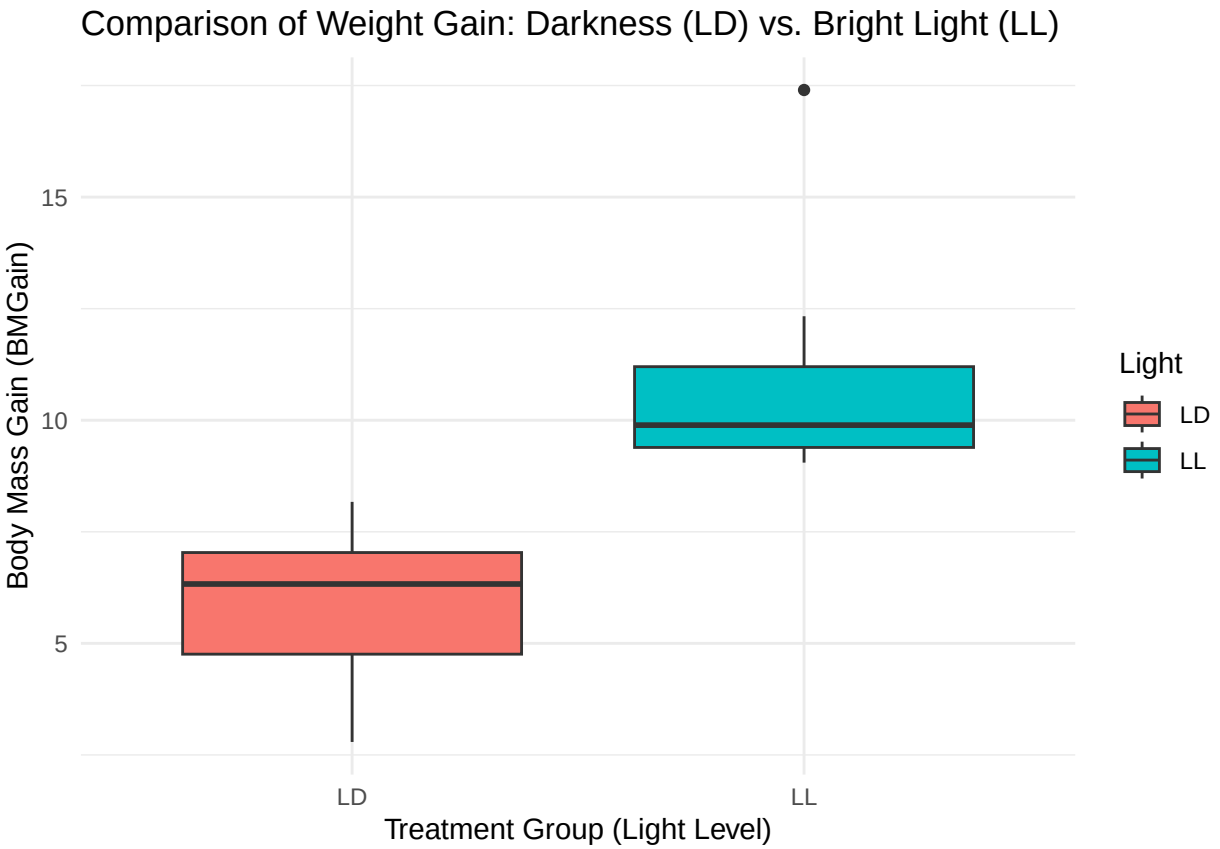
```
# Create the plot
ggplot(data, aes(x = Light, y = BMGain, fill = Light)) +
  geom_boxplot() +
  labs(title = "Effect of Light at Night on Weight Gain",
       x = "Treatment Group (Light Level)",
       y = "Body Mass Gain (BMGain)") +
  theme_minimal()
```




Conclusion: From this box plot, we can infer a potential causal relationship between light exposure at night and weight gain. Mice that were exposed to constant light (LL) or dim light (DM) during the night gained more weight than those exposed to low light (LD), implying that the presence of light at night could disrupt natural behaviors or metabolic processes, leading to increased weight gain.

## Question 2

```
# Subset the data to include only 'LL' (bright light) and 'LD' (darkness)
subset_data = data |> filter(Light %in% c("LL", "LD"))
ggplot(subset_data, aes(x = Light, y = BMGain, fill = Light)) +
  geom_boxplot() +
  labs(title = "Comparison of Weight Gain: Darkness (LD) vs. Bright Light (LL)",
       x = "Treatment Group (Light Level)",
       y = "Body Mass Gain (BMGain)") +
  theme_minimal()
```




The box plot indicates a causal relationship between night-time light exposure and weight gain, with the LL (bright light) group showing a greater increase in body mass than the LD (darkness) group. This suggests that bright light at night may play a role in promoting weight gain in nocturnal animals, possibly by altering their eating habits, metabolism, or sleep patterns.

## Question 3

Set up the data such that everything you will need has generic names (such as Yobs or whatever you want to call them). Everything specific to the context of your data (variable names, sample sizes) should only be in your R Script here. Everything else should be generic so you can copy/paste it for later use. What quantities will you need to evaluate the causal effect of light at night on weight gain?

```r
subset_data = subset_data  |>
  rename( Y_obs = BMGain,
          Treatment = Light)

n = nrow(data)

# Calculate mean outcome by treatment group
mean_outcome = subset_data |>
  group_by(Treatment) |>
  summarise(mean_Y_obs = mean(Y_obs, na.rm = TRUE))

# Calculate variance by treatment group
variance_outcome = subset_data  |>
  group_by(Treatment) |>
  summarise(var_Y_obs = var(Y_obs, na.rm = TRUE))
```

```r
# Calculate the effect size
effect_size =mean_outcome$mean_Y_obs[mean_outcome$Treatment == "LL"] -mean_outcome$mean_Y_obs[mean_outc

mean_outcome
```

```
## # A tibble: 2 x 2
##   Treatment mean_Y_obs
##   <chr>          <dbl>
## 1 LD              5.93
## 2 LL             11.0
```

```r
variance_outcome
```

```
## # A tibble: 2 x 2
##   Treatment var_Y_obs
##   <chr>         <dbl>
## 1 LD             3.61
## 2 LL             6.89
```

```r
effect_size
```

```
## [1] 5.08375
```

Conclusion: Quantities Needed for Causal Effect Evaluation are: * Observed Outcome (Y_obs) * Treatment Assignment (Treatment) * Sample Size (n) * Average Outcome by Treatment Group * Variance of Outcome by Group * Difference in Means (Effect Size)

## Question 4

Suppose we want the statistic to be the difference in means between the two treatment groups. Calculate Tobs.

```r
# Calculate mean outcome by treatment group
effect_size =mean_outcome$mean_Y_obs[mean_outcome$Treatment == "LL"] -mean_outcome$mean_Y_obs[mean_outc
effect_size
```

```
## [1] 5.08375
```

The difference in means between the 2 treatment groups is c(“LD”, “LL”), c(5.92625, 11.01).

## Question 5

How many different possibilities are there for A? Enumerate all of these possibilities in a matrix. (Hint: it’s probably easiest to first install the ri or perm package, have a look at the function chooseMatrix in R, it may come in handy.)

```r
amat = perm::chooseMatrix(19, 10)

nrow(amat)
```

```
## [1] 92378
```