Homework 4

This homework is due on the deadline posted on edX. Please submit a .pdf file of your output and upload a .zip file with your .Rmd file.

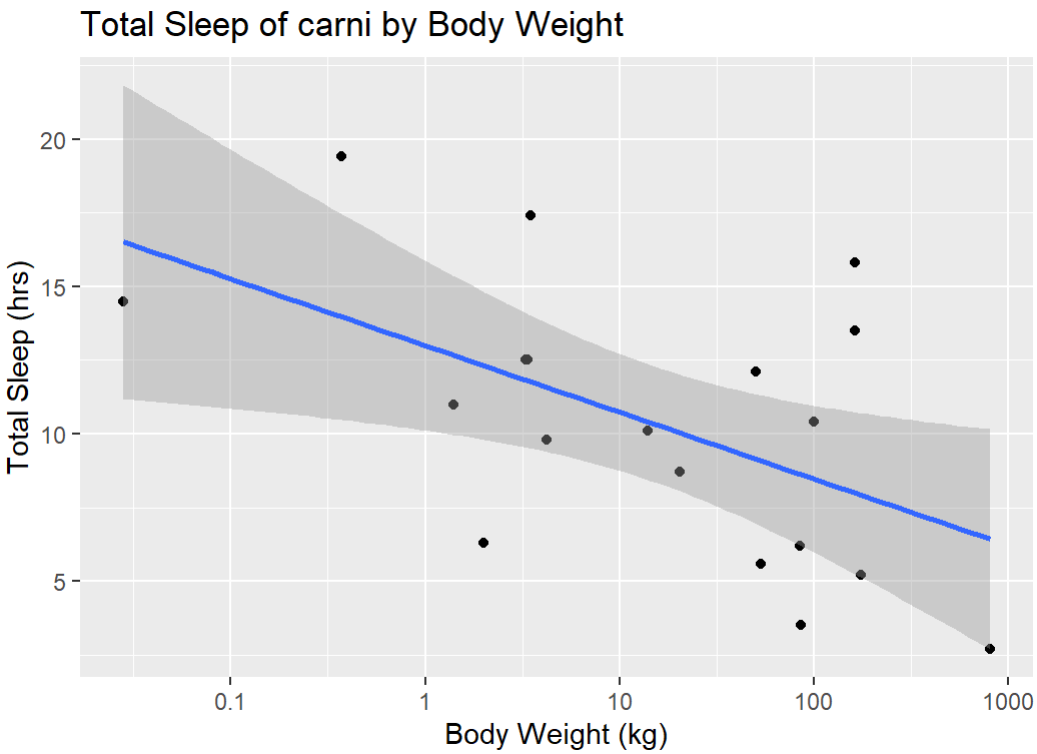
For this homework, we will use the `msleep` dataset provided by **ggplot2**. See here for details: https://ggplot2.tidyverse.org/reference/msleep.html (https://ggplot2.tidyverse.org/reference/msleep.html)

```
glimpse(msleep)
```

```
## Rows: 83
## Columns: 11
## $ name         <chr> "Cheetah", "Owl monkey", "Mountain beaver", "Greater shor~
## $ genus        <chr> "Acinonyx", "Aotus", "Aplodontia", "Blarina", "Bos", "Bra~
## $ vore         <chr> "carni", "omni", "herbi", "omni", "herbi", "herbi", "carn~
## $ order        <chr> "Carnivora", "Primates", "Rodentia", "Soricomorpha", "Art~
## $ conservation <chr> "lc", NA, "nt", "lc", "domesticated", NA, "vu", NA, "dome~
## $ sleep_total  <dbl> 12.1, 17.0, 14.4, 14.9, 4.0, 14.4, 8.7, 7.0, 10.1, 3.0, 5~
## $ sleep_rem    <dbl> NA, 1.8, 2.4, 2.3, 0.7, 2.2, 1.4, NA, 2.9, NA, 0.6, 0.8, ~
## $ sleep_cycle  <dbl> NA, NA, NA, 0.1333333, 0.6666667, 0.7666667, 0.3833333, N~
## $ awake        <dbl> 11.9, 7.0, 9.6, 9.1, 20.0, 9.6, 15.3, 17.0, 13.9, 21.0, 1~
## $ brainwt      <dbl> NA, 0.01550, NA, 0.00029, 0.42300, NA, NA, NA, 0.07000, 0~
## $ bodywt       <dbl> 50.000, 0.480, 1.350, 0.019, 600.000, 3.850, 20.490, 0.04~
```

Problem 1: Visualize the relationship between total amount of sleep and body weight in each mammal classified as a carnivore (`vore == "carni"`). Your plot should include raw data points as well as a linear trendline with confidence interval. What do you observe?

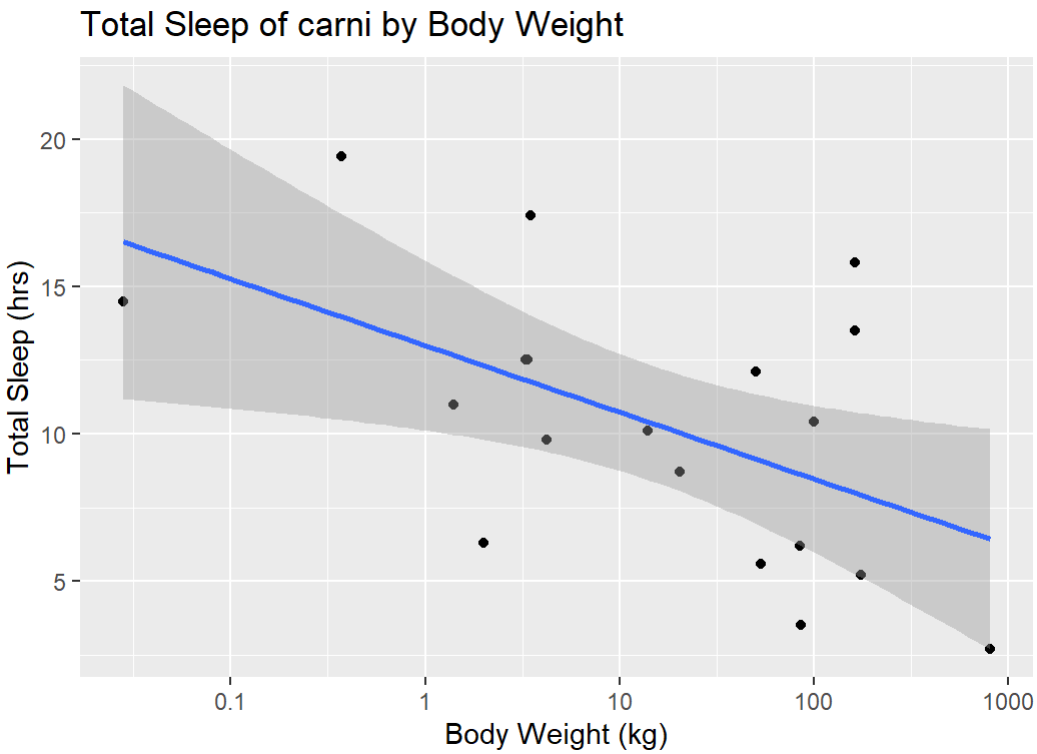
```
msleep %>%
  filter(vore == "carni") %>%
  ggplot(aes(bodywt, sleep_total)) +
  geom_point() +
  geom_smooth(method="lm", formula = y ~ x) +
  scale_x_continuous(
    trans = "log10",
    breaks = c(0.1, 1, 10, 100, 1000),
    labels = c("0.1", "1", "10", "100", "1000")
  ) +
  labs(
    x = "Body Weight (kg)",
    y = "Total Sleep (hrs)",
    title = "Total Sleep of carni by Body Weight"
  )
```



By the plot it can be seen that the heavier the carnivore, the more likely they are to sleep less in a given 24 hour period. There is a tighter confidence interval near the 1 to 10 kg weight range meaning we have a greater likelihood of correctly predicting the number of hours these carnivores will sleep than we are able to predict on lighter or heavier carnivores.

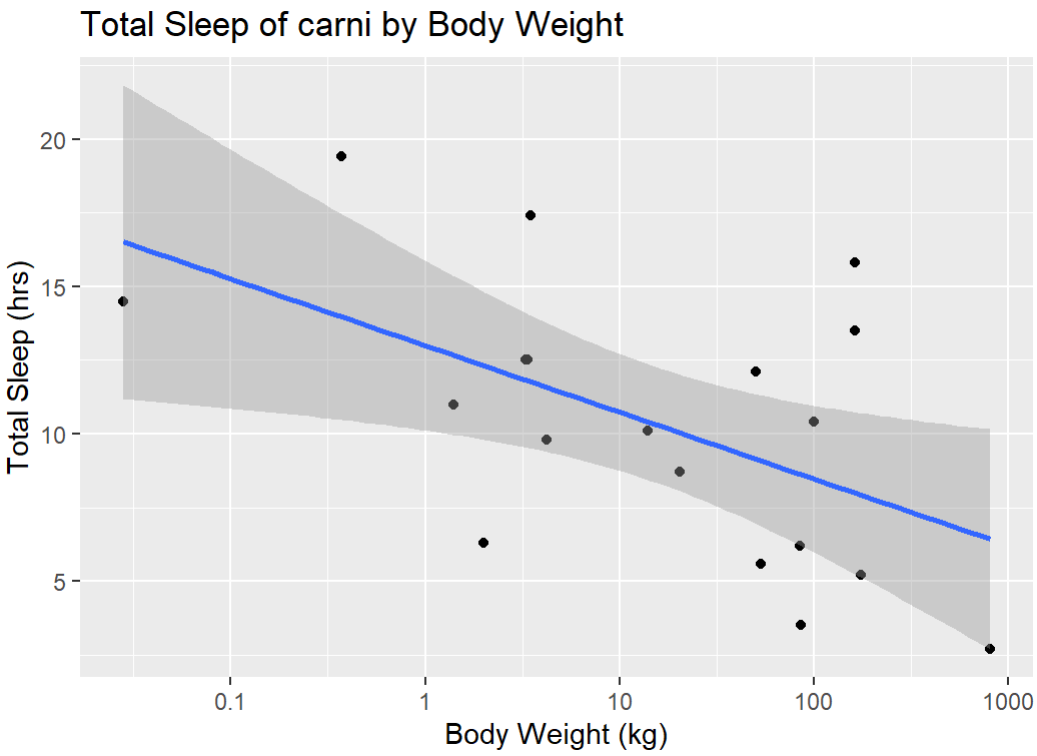
Problem 2: Write a function to create the plot above. Your function should have two inputs: `data`, which is the dataset to plot, and `vore`, which is a string indicating the vore type, such as `"carni"`. Reproduce the plot using your new function.

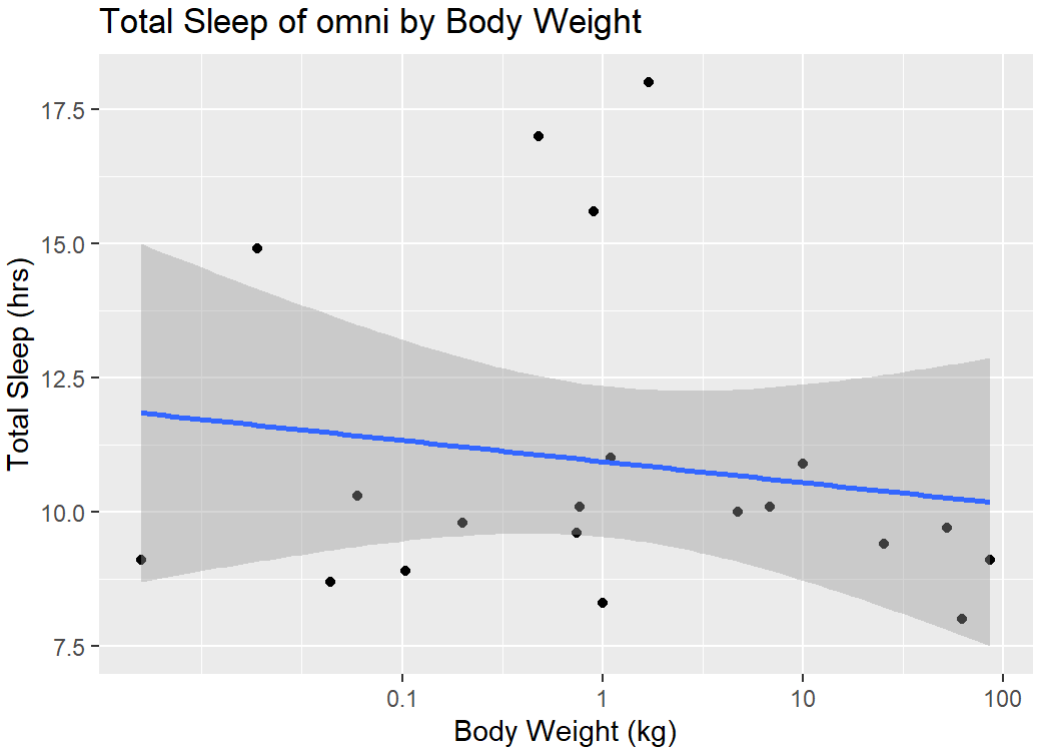
```
make_vore_plot <- function(data, vore) {
  data %>%
    filter(.data$vore == .env$vore) %>%
    ggplot(aes(bodywt, sleep_total)) +
    geom_point() +
    geom_smooth(method="lm", formula = y ~ x) +
    scale_x_continuous(
      trans = "log10",
      breaks = c(0.1, 1, 10, 100, 1000),
      labels = c("0.1", "1", "10", "100", "1000")
    ) +
    labs(
      x = "Body Weight (kg)",
      y = "Total Sleep (hrs)",
      title = glue("Total Sleep of {vore} by Body Weight")
    )
}

make_vore_plot(msleep, "carni")
```

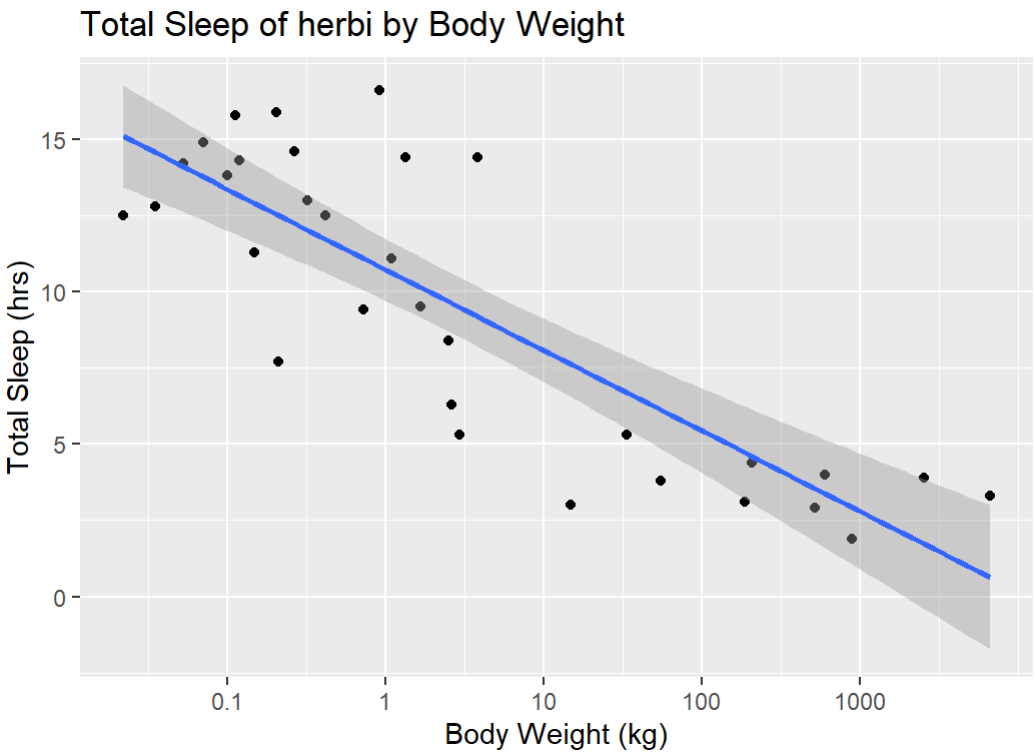


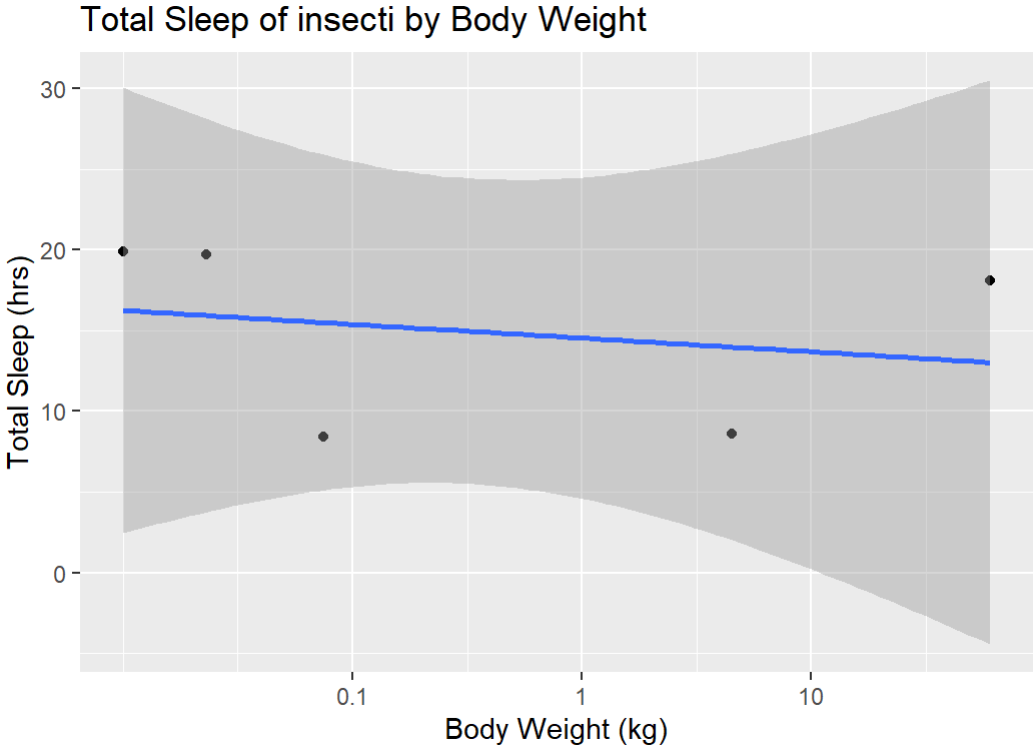
Problem 3: Write code that automatically applies the function you created in Problem 2 to all vore types (you can exclude `NA` values). **Do not write a `for` loop.** How does the relationship between body weight and total amount of sleep vary across vores?

```
make_vore_plot_v2 <- function(vore) {
  make_vore_plot(msleep, vore)
}

msleep %>%
  filter(!is.na(vore)) %>%
  nest(data = -vore) %>%
  mutate(plots = map(vore, make_vore_plot_v2)) %>%
  pull(plots) %>%
  walk(print)
```


Total Sleep of carni by Body Weight
Total Sleep (hrs)
20
15
10
5
0.1
1
10
100
1000
Body Weight (kg)


Total Sleep of omni by Body Weight
Total Sleep (hrs)
17.5
15.0
12.5
10.0
7.5
0.1
1
10
100
Body Weight (kg)





While the number of hours of sleep consistently decreases with an increase in weight, the rate at which the decrease happens varies by the animal's diet. With herbivores, the decrease is the steepest and the confidence interval is the tightest. We can very closely predict the number of hours of sleep they will exhibit based on their weight. For insectivores, they are the least reliable. While the regression does suggest a slight decrease, it is nearly level and the confidence interval is gigantic. It is very possible that there is actually no relation between insectivore's number of hours of sleep in a 24 hour period and their body weight.