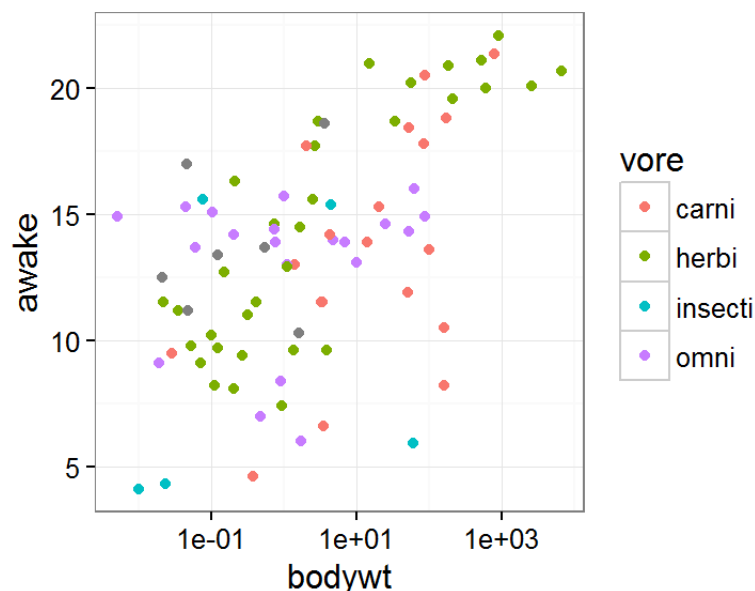


Lab Worksheet 2

We will be working with the `msleep` data set that is provided with `ggplot2`. The data set contains information about the sleep habits of 83 mammals. Enter `?msleep` on the R command line to learn more about the dataset.

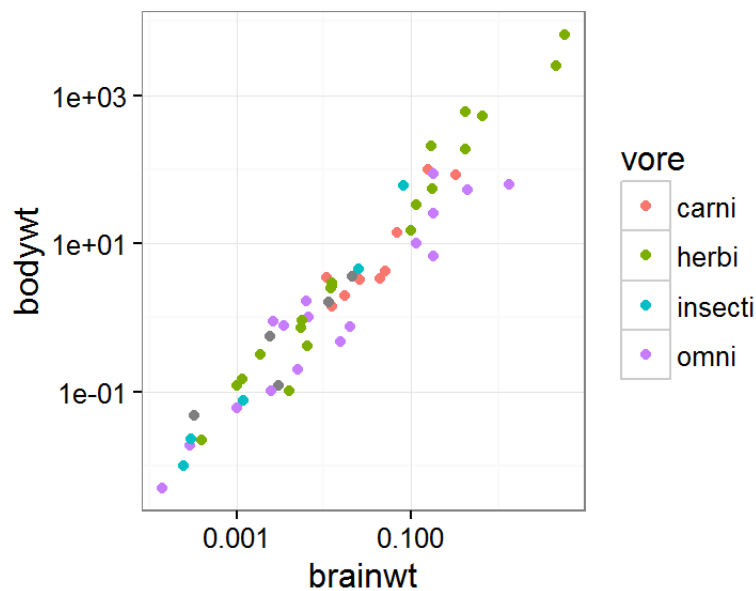
Question 1: Make the following plots: (i) a plot of total time awake vs. body weight, colored by “vore” (carnivore, herbivore, etc.); (ii) a plot of body weight vs. brain weight, colored by “vore”. When you plot body weight and/or brain weight, consider whether a linear scale or a logarithmic scale seems more appropriate, and explain your reasoning in 1-2 sentences. **HINT:** Use the functions `scale_x_log10()` and `scale_y_log10()` to change the scales.

```
ggplot(msleep, aes(x=bodywt, y=awake, color=vore)) + geom_point()+scale_x_log10()
```



```
ggplot(msleep, aes(x=brainwt, y=bodywt, color=vore)) + geom_point() + scale_y_log10() + scale_x_log10()
```

```
## Warning: Removed 27 rows containing missing values (geom_point).
```



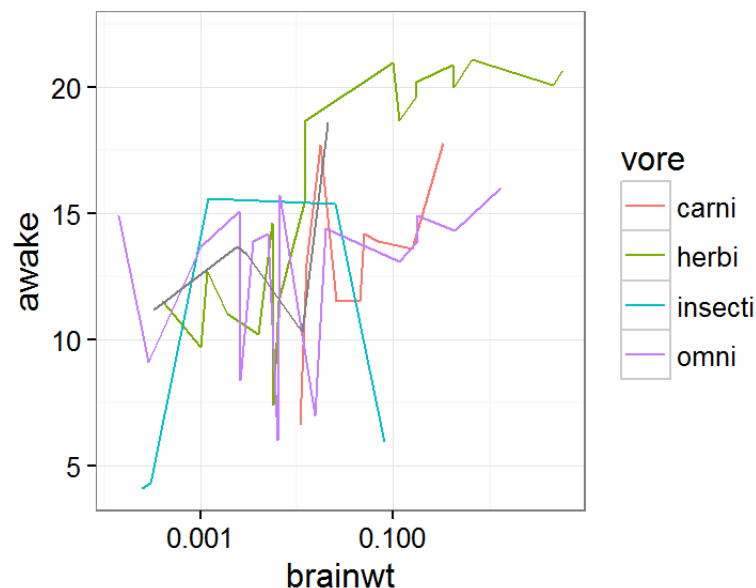
It is more appropriate to use the logarithmic scale in order to incorporate the data points in a reasonable scale that displays the data appropriately.

Question 2: Explain the difference between `geom_line()` and `geom_path()`. Make up a simple data set (5-10 data points), plot it twice, once using `geom_line()` and once using `geom_path()`, and explain why each plot looks the way it does.

`geom_path` connects observations by the order in which they appear, while `geom_line` connects them in order of variable on the x-axis

```
# your R code goes here
sleep <- msleep[, 1:10]
ggplot(sleep, aes(x=brainwt, y=awake, color=vore)) + geom_line() + scale_x_log10()
```

```
## Warning: Removed 27 rows containing missing values (geom_path).
```



```
ggplot(sleep, aes(x=brainwt, y=awake, color=vore)) + geom_path() + scale_x_log10()
```

Warning: Removed 4 rows containing missing values (geom_path).

