Exercise: **Data Frames**

Day 1, Part C

For this exercise, we will use the `msleep` data frame from the `ggplot2` library. This can be loaded using the following code:

```r
library(ggplot2)
data("msleep")
msleep <- data.frame(msleep)
```

Some information about the data set is available in the help file for `msleep`.

1. Check that `msleep` is a data frame.

   ```r
   > is.data.frame(msleep)
   [1] TRUE
   ```

   ```r
   > class(msleep)
   [1] "data.frame"
   ```

2. Use each of the following functions to explore the data frame: `head()`, `tail()`, `summary()`, and `View()`.

   ```r
   > head(msleep)
   name       genus    vore        order conservation sleep_total
   1                     Cheetah   Acinonyx carni    Carnivora           lc        12.1
   2                 Owl monkey      Aotus  omni     Primates         <NA>        17.0
   3            Mountain beaver  Aplodontia herbi     Rodentia           nt        14.4
   4 Greater short-tailed shrew    Blarina  omni Soricomorpha           lc        14.9
   5                        Cow         Bos herbi Artiodactyla domesticated         4.0
   6           Three-toed sloth    Bradypus herbi       Pilosa         <NA>        14.4
     sleep_rem sleep_cycle awake brainwt  bodywt
   1       NA          NA  11.9      NA  50.000
   2      1.8          NA   7.0 0.01550   0.480
   3      2.4          NA   9.6      NA   1.350
   4      2.3   0.1333333   9.1 0.00029   0.019
   5      0.7   0.6666667  20.0 0.42300 600.000
   6      2.2   0.7666667   9.6      NA   3.850
   ```

   ```r
   > tail(msleep)
   name       genus    vore         order conservation sleep_total
   78                  Tenrec     Tenrec  omni Afrosoricida         <NA>        15.6
   79              Tree shrew     Tupaia  omni    Scandentia         <NA>         8.9
   80 Bottle-nosed dolphin   Tursiops carni       Cetacea         <NA>         5.2
   81                   Genet    Genetta carni     Carnivora         <NA>         6.3
   82               Arctic fox     Vulpes carni     Carnivora         <NA>        12.5
   83                 Red fox     Vulpes carni     Carnivora         <NA>         9.8
     sleep_rem sleep_cycle awake brainwt  bodywt
   78      2.3          NA   8.4  0.0026   0.900
   79      2.6   0.2333333  15.1  0.0025   0.104
   80       NA          NA  18.8      NA 173.330
   81      1.3          NA  17.7  0.0175   2.000
   82       NA          NA  11.5  0.0445   3.380
   83      2.4   0.3500000  14.2  0.0504   4.230
   ```

```
> summary(msleep)
 name               genus               vore                order
 Length:83          Length:83           Length:83           Length:83
 Class :character   Class :character    Class :character    Class :character
 Mode  :character   Mode  :character    Mode  :character    Mode  :character




 conservation        sleep_total        sleep_rem         sleep_cycle          awake
 Length:83          Min.   : 1.90     Min.   :0.100     Min.   :0.1167     Min.   : 4.10
 Class :character   1st Qu.: 7.85     1st Qu.:0.900     1st Qu.:0.1833     1st Qu.:10.25
 Mode  :character   Median :10.10     Median :1.500     Median :0.3333     Median :13.90
                    Mean   :10.43     Mean   :1.875     Mean   :0.4396     Mean   :13.57
                    3rd Qu.:13.75     3rd Qu.:2.400     3rd Qu.:0.5792     3rd Qu.:16.15
                    Max.   :19.90     Max.   :6.600     Max.   :1.5000     Max.   :22.10
                                      NA's   :22        NA's   :51
    brainwt            bodywt
 Min.   :0.00014   Min.   :   0.005
 1st Qu.:0.00290   1st Qu.:   0.174
 Median :0.01240   Median :   1.670
 Mean   :0.28158   Mean   : 166.136
 3rd Qu.:0.12550   3rd Qu.:  41.750
 Max.   :5.71200   Max.   :6654.000
 NA's   :27

> View(msleep)
```

3. How many rows and columns does this data frame have?

```
> nrow(msleep)
[1] 83
```

```
> ncol(msleep)
[1] 11
```

```
> dim(msleep)
[1] 83 11
```

4. What are the names of the columns? What data type is each column?

```
> names(msleep)
 [1] "name"         "genus"        "vore"         "order"         "conservation"
 [6] "sleep_total"  "sleep_rem"    "sleep_cycle"  "awake"         "brainwt"
[11] "bodywt"
```

```
> str(msleep)
'data.frame':   83 obs. of  11 variables:
 $ name        : chr  "Cheetah" "Owl monkey" "Mountain beaver" "Greater short-tailed shrew" ...
 $ genus       : chr  "Acinonyx" "Aotus" "Aplodontia" "Blarina" ...
 $ vore        : chr  "carni" "omni" "herbi" "omni" ...
 $ order       : chr  "Carnivora" "Primates" "Rodentia" "Soricomorpha" ...
 $ conservation: chr  "lc" NA "nt" "lc" ...
 $ sleep_total : num  12.1 17 14.4 14.9 4 14.4 8.7 7 10.1 3 ...
 $ sleep_rem   : num  NA 1.8 2.4 2.3 0.7 2.2 1.4 NA 2.9 NA ...
```

```
 $ sleep_cycle : num  NA NA NA 0.133 0.667 ...
 $ awake       : num  11.9 7 9.6 9.1 20 9.6 15.3 17 13.9 21 ...
 $ brainwt     : num  NA 0.0155 NA 0.00029 0.423 NA NA NA 0.07 0.0982 ...
 $ bodywt      : num  50 0.48 1.35 0.019 600 ...
```

5. Select the `name` column three different ways.

```
> msleep[, 1]
 [1] "Cheetah"                    "Owl monkey"
 [3] "Mountain beaver"            "Greater short-tailed shrew"
 [5] "Cow"                        "Three-toed sloth"
 [7] "Northern fur seal"          "Vesper mouse"
 [9] "Dog"                        "Roe deer"
[11] "Goat"                       "Guinea pig"
[13] "Grivet"                     "Chinchilla"
[15] "Star-nosed mole"            "African giant pouched rat"
[17] "Lesser short-tailed shrew"  "Long-nosed armadillo"
[19] "Tree hyrax"                 "North American Opossum"
[21] "Asian elephant"             "Big brown bat"
[23] "Horse"                      "Donkey"
[25] "European hedgehog"          "Patas monkey"
[27] "Western american chipmunk"  "Domestic cat"
[29] "Galago"                     "Giraffe"
[31] "Pilot whale"                "Gray seal"
[33] "Gray hyrax"                 "Human"
[35] "Mongoose lemur"             "African elephant"
[37] "Thick-tailed opposum"       "Macaque"
[39] "Mongolian gerbil"           "Golden hamster"
[41] "Vole "                      "House mouse"
[43] "Little brown bat"           "Round-tailed muskrat"
[45] "Slow loris"                 "Degu"
[47] "Northern grasshopper mouse" "Rabbit"
[49] "Sheep"                      "Chimpanzee"
[51] "Tiger"                      "Jaguar"
[53] "Lion"                       "Baboon"
[55] "Desert hedgehog"            "Potto"
[57] "Deer mouse"                 "Phalanger"
[59] "Caspian seal"               "Common porpoise"
[61] "Potoroo"                    "Giant armadillo"
[63] "Rock hyrax"                 "Laboratory rat"
[65] "African striped mouse"      "Squirrel monkey"
[67] "Eastern american mole"      "Cotton rat"
[69] "Mole rat"                   "Arctic ground squirrel"
[71] "Thirteen-lined ground squirrel" "Golden-mantled ground squirrel"
[73] "Musk shrew"                 "Pig"
[75] "Short-nosed echidna"        "Eastern american chipmunk"
[77] "Brazilian tapir"            "Tenrec"
[79] "Tree shrew"                 "Bottle-nosed dolphin"
[81] "Genet"                      "Arctic fox"
[83] "Red fox"

> msleep[, "name"]
 [1] "Cheetah"                    "Owl monkey"
```

```
 [3] "Mountain beaver"               "Greater short-tailed shrew"
 [5] "Cow"                           "Three-toed sloth"
 [7] "Northern fur seal"             "Vesper mouse"
 [9] "Dog"                           "Roe deer"
[11] "Goat"                          "Guinea pig"
[13] "Grivet"                        "Chinchilla"
[15] "Star-nosed mole"               "African giant pouched rat"
[17] "Lesser short-tailed shrew"     "Long-nosed armadillo"
[19] "Tree hyrax"                    "North American Opossum"
[21] "Asian elephant"                "Big brown bat"
[23] "Horse"                         "Donkey"
[25] "European hedgehog"             "Patas monkey"
[27] "Western american chipmunk"     "Domestic cat"
[29] "Galago"                        "Giraffe"
[31] "Pilot whale"                   "Gray seal"
[33] "Gray hyrax"                    "Human"
[35] "Mongoose lemur"                "African elephant"
[37] "Thick-tailed opposum"          "Macaque"
[39] "Mongolian gerbil"              "Golden hamster"
[41] "Vole "                         "House mouse"
[43] "Little brown bat"              "Round-tailed muskrat"
[45] "Slow loris"                    "Degu"
[47] "Northern grasshopper mouse"    "Rabbit"
[49] "Sheep"                         "Chimpanzee"
[51] "Tiger"                         "Jaguar"
[53] "Lion"                          "Baboon"
[55] "Desert hedgehog"               "Potto"
[57] "Deer mouse"                    "Phalanger"
[59] "Caspian seal"                  "Common porpoise"
[61] "Potoroo"                       "Giant armadillo"
[63] "Rock hyrax"                    "Laboratory rat"
[65] "African striped mouse"         "Squirrel monkey"
[67] "Eastern american mole"         "Cotton rat"
[69] "Mole rat"                      "Arctic ground squirrel"
[71] "Thirteen-lined ground squirrel" "Golden-mantled ground squirrel"
[73] "Musk shrew"                    "Pig"
[75] "Short-nosed echidna"           "Eastern american chipmunk"
[77] "Brazilian tapir"              "Tenrec"
[79] "Tree shrew"                    "Bottle-nosed dolphin"
[81] "Genet"                         "Arctic fox"
[83] "Red fox"
```

```
> msleep$name
[1] "Cheetah"                        "Owl monkey"
 [3] "Mountain beaver"               "Greater short-tailed shrew"
 [5] "Cow"                           "Three-toed sloth"
 [7] "Northern fur seal"             "Vesper mouse"
 [9] "Dog"                           "Roe deer"
[11] "Goat"                          "Guinea pig"
[13] "Grivet"                        "Chinchilla"
[15] "Star-nosed mole"               "African giant pouched rat"
[17] "Lesser short-tailed shrew"     "Long-nosed armadillo"
[19] "Tree hyrax"                    "North American Opossum"
```

```
[21] "Asian elephant"                "Big brown bat"
[23] "Horse"                         "Donkey"
[25] "European hedgehog"             "Patas monkey"
[27] "Western american chipmunk"     "Domestic cat"
[29] "Galago"                        "Giraffe"
[31] "Pilot whale"                   "Gray seal"
[33] "Gray hyrax"                    "Human"
[35] "Mongoose lemur"                "African elephant"
[37] "Thick-tailed opposum"          "Macaque"
[39] "Mongolian gerbil"             "Golden hamster"
[41] "Vole "                         "House mouse"
[43] "Little brown bat"             "Round-tailed muskrat"
[45] "Slow loris"                    "Degu"
[47] "Northern grasshopper mouse"    "Rabbit"
[49] "Sheep"                         "Chimpanzee"
[51] "Tiger"                         "Jaguar"
[53] "Lion"                          "Baboon"
[55] "Desert hedgehog"              "Potto"
[57] "Deer mouse"                    "Phalanger"
[59] "Caspian seal"                  "Common porpoise"
[61] "Potoroo"                       "Giant armadillo"
[63] "Rock hyrax"                    "Laboratory rat"
[65] "African striped mouse"        "Squirrel monkey"
[67] "Eastern american mole"        "Cotton rat"
[69] "Mole rat"                      "Arctic ground squirrel"
[71] "Thirteen-lined ground squirrel" "Golden-mantled ground squirrel"
[73] "Musk shrew"                    "Pig"
[75] "Short-nosed echidna"          "Eastern american chipmunk"
[77] "Brazilian tapir"              "Tenrec"
[79] "Tree shrew"                    "Bottle-nosed dolphin"
[81] "Genet"                         "Arctic fox"
[83] "Red fox"
```

6. Select just rows where the `order` column is "Carnivora."

```
> msleep[msleep$order == "Carnivora", ]
name           genus   vore     order conservation sleep_total sleep_rem
1          Cheetah    Acinonyx carni Carnivora           lc        12.1        NA
7   Northern fur seal  Callorhinus carni Carnivora           vu         8.7       1.4
9              Dog        Canis carni Carnivora domesticated        10.1       2.9
28      Domestic cat       Felis carni Carnivora domesticated        12.5       3.2
32        Gray seal Haliochoerus carni Carnivora           lc         6.2       1.5
51            Tiger     Panthera carni Carnivora           en        15.8        NA
52           Jaguar     Panthera carni Carnivora           nt        10.4        NA
53             Lion     Panthera carni Carnivora           vu        13.5        NA
59      Caspian seal        Phoca carni Carnivora           vu         3.5       0.4
81            Genet      Genetta carni Carnivora         <NA>         6.3       1.3
82        Arctic fox       Vulpes carni Carnivora         <NA>        12.5        NA
83          Red fox       Vulpes carni Carnivora         <NA>         9.8       2.4
   sleep_cycle awake brainwt  bodywt
1          NA  11.9      NA  50.000
7   0.3833333  15.3      NA  20.490
9   0.3333333  13.9  0.0700  14.000
```

5

```
28   0.4166667  11.5  0.0256    3.300
32          NA  17.8  0.3250   85.000
51          NA   8.2      NA  162.564
52          NA  13.6  0.1570  100.000
53          NA  10.5      NA  161.499
59          NA  20.5      NA   86.000
81          NA  17.7  0.0175    2.000
82          NA  11.5  0.0445    3.380
83   0.3500000  14.2  0.0504    4.230
```

7. Select just rows for animals with body weight > 200kg.

```
> msleep[msleep$bodywt > 200, ]
name            genus  vore            order  conservation sleep_total
5                 Cow            Bos herbi   Artiodactyla domesticated         4.0
21    Asian elephant        Elephas herbi     Proboscidea           en         3.9
23              Horse          Equus herbi Perissodactyla domesticated         2.9
30            Giraffe        Giraffa herbi   Artiodactyla           cd         1.9
31       Pilot whale Globicephalus carni        Cetacea           cd         2.7
36  African elephant      Loxodonta herbi     Proboscidea           vu         3.3
77   Brazilian tapir        Tapirus herbi Perissodactyla           vu         4.4
   sleep_rem sleep_cycle awake brainwt    bodywt
5        0.7   0.6666667 20.00   0.423   600.000
21        NA          NA 20.10   4.603  2547.000
23       0.6   1.0000000 21.10   0.655   521.000
30       0.4          NA 22.10      NA   899.995
31       0.1          NA 21.35      NA   800.000
36        NA          NA 20.70   5.712  6654.000
77       1.0   0.9000000 19.60   0.169   207.501
```

8. Create a new variable for the total amount of sleep in minutes (`sleep_total_min`).

```
> msleep$sleep_total_min <- msleep$sleep_total * 60
```

a. What is the longest time any animal sleeps (in minutes)?

```
> max(msleep$sleep_total_min)
[1] 1194
```

b. Which animal sleeps this long?

```
> msleep[msleep$sleep_total_min == max(msleep$sleep_total_min), "name"]
[1] "Little brown bat"
```

```
> msleep$name[msleep$sleep_total_min == max(msleep$sleep_total_min)]
[1] "Little brown bat"
```

9. Create a new variable for the proportion of total sleep spent in REM (`prop_rem`).

```
> msleep$prop_rem <- msleep$sleep_rem/msleep$sleep_total
```

a. What is the minimum proportion? (hint, look at the argument options in the help file for `min()`)

```
> min(msleep$prop_rem, na.rm = T)
[1] 0.03703704
```

```
> min(msleep$prop_rem[!is.na(msleep$prop_rem)])
[1] 0.03703704
```

b. How long is REM sleep for the animal with the smallest proportion REM? (hint, this requires two conditions be met: `prop_rem` is NOT missing, and `prop_rem` is the minimum value)

```
> msleep$sleep_rem[msleep$prop_rem == min(msleep$prop_rem, na.rm = T) & !is.na(msleep$prop_r
[1] 0.1
```

10. Create a new variable for the ratio of body weight to brain weight (`bdy_brn_ratio`).

```
> msleep$bdy_brn_ratio <- msleep$bodywt/msleep$brainwt
```

a. What is the median body-to-brain weight ratio?

```
> median(msleep$bdy_brn_ratio, na.rm = T)
[1] 143.7799
```

b. What is the variance of the body-to-brain weight ratio?

```
> var(msleep$bdy_brn_ratio, na.rm = T)
[1] 90841.86
```

11. Drop the `conservation`, `sleep_cycle`, and `bodywt` variables from the `msleep` object.

```
> msleep[, c("conservation", "sleep_cycle", "bodywt")] <- NULL
```

12. Create a new object (`msleep_carni`) that contains just the data for carnivores and only the `name`, `genus`, `order`, `sleep_total`, and `brainwt` variables.

```
> msleep_carni <- msleep[msleep$vore == "carni", c("name", "genus", "order", "sleep_total",
+     "brainwt")]
```