ML Assignment

Gaurav Kudeshia

2023-09-09

```
# Load the 'caret' package
library(caret)
```

```
## Loading required package: ggplot2
```

```
## Loading required package: lattice
```

```
# Load the 'msleep' dataset from the 'ggplot2' package
data("msleep", package = "ggplot2")

# View the first few rows of the 'msleep' dataset
head(msleep)
```

```
## # A tibble: 6 × 11
##   name    genus vore  order conservation sleep_total sleep_rem sleep_cycle awak
e
##   <chr>   <chr> <chr> <chr> <chr>              <dbl>     <dbl>       <dbl> <dbl
>
## 1 Cheetah Acin… carni Carn… lc                   12.1      NA         NA      11.
9
## 2 Owl mo… Aotus omni  Prim… <NA>                 17         1.8       NA       7
## 3 Mounta… Aplo… herbi Rode… nt                   14.4       2.4       NA       9.
6
## 4 Greate… Blar… omni  Sori… lc                   14.9       2.3        0.133   9.
1
## 5 Cow     Bos   herbi Arti… domesticated         4          0.7        0.667  20
## 6 Three-… Brad… herbi Pilo… <NA>                 14.4       2.2        0.767   9.
6
## # ℹ 2 more variables: brainwt <dbl>, bodywt <dbl>
```

```
# Calculate summary statistics for quantitative variables
summary(msleep$sleep_total)
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.90    7.85   10.10   10.43   13.75   19.90
```

```
summary(msleep$bodywt)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##    0.005    0.174    1.670  166.136   41.750 6654.000
```

```
# Calculate Standard Deviation
sd(msleep$sleep_total)
```

```
## [1] 4.450357
```

```
sd(msleep$bodywt)
```

```
## [1] 786.8397
```

```
# Display tables for categorical variables
table(msleep$order)
```

```
##
##    Afrosoricida    Artiodactyla       Carnivora         Cetacea      Chiroptera
##               1               6              12               3               2
##       Cingulata Didelphimorphia   Diprotodontia  Erinaceomorpha      Hyracoidea
##               2               2               2               2               3
##      Lagomorpha     Monotremata  Perissodactyla          Pilosa        Primates
##               1               1               3               1              12
##     Proboscidea        Rodentia      Scandentia    Soricomorpha
##               2              22               1               5
```

```
table(msleep$vore)
```

```
##
##   carni   herbi insecti    omni
##      19      32       5      20
```

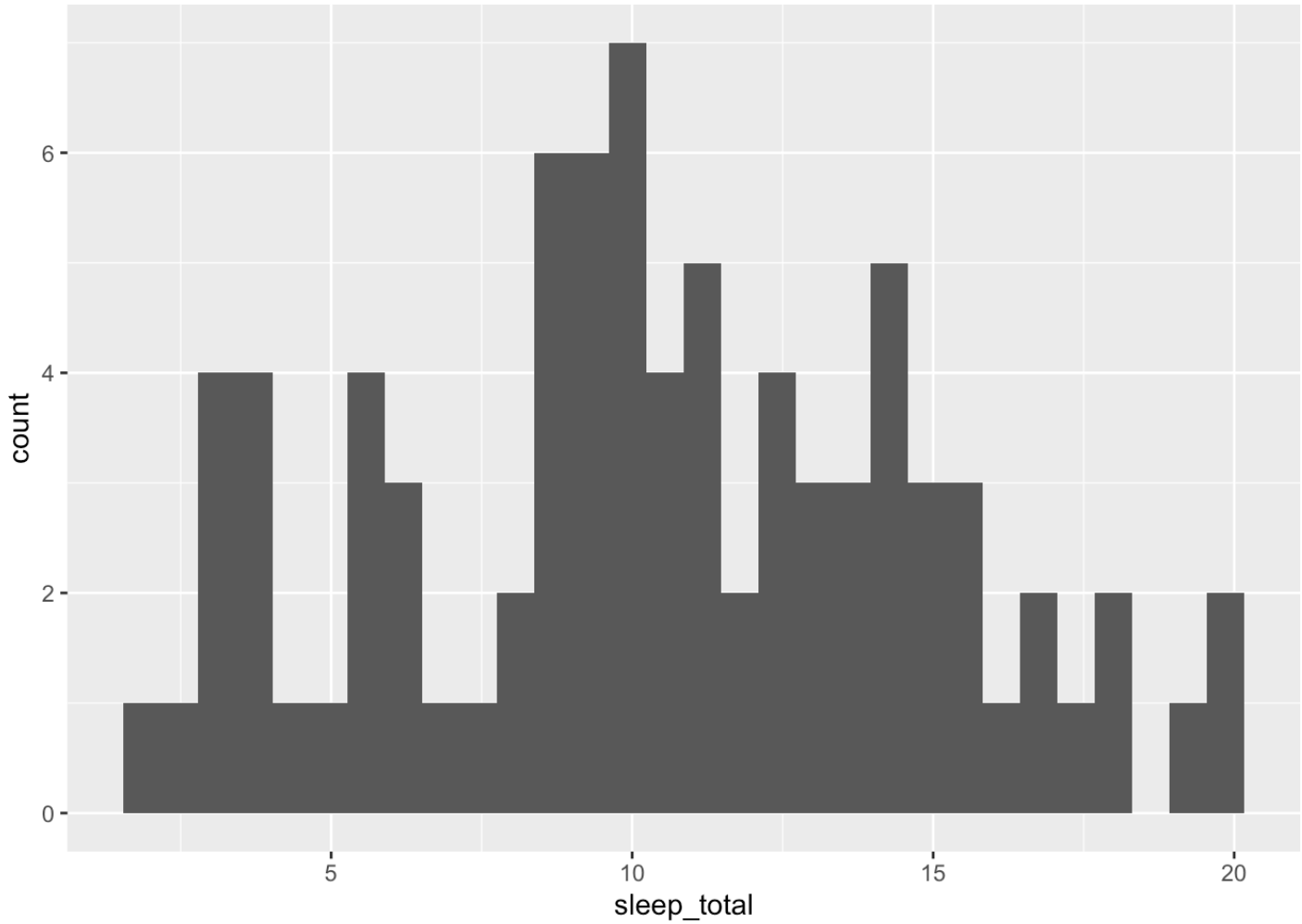
```
# Transform msleep$bodywt
msleep$bodywt_half <- msleep$bodywt / 2
msleep$bodywt_half
```

```
##  [1]   25.0000    0.2400    0.6750    0.0095  300.0000    1.9250   10.2450
##  [8]    0.0225    7.0000    7.4000   16.7500    0.3640    2.3750    0.2100
## [15]    0.0300    0.5000    0.0025    1.7500    1.4750    0.8500 1273.5000
## [22]    0.0115  260.5000   93.5000    0.3850    5.0000    0.0355    1.6500
## [29]    0.1000  449.9975  400.0000   42.5000    1.3125   31.0000    0.8350
## [36] 3327.0000    0.1850    3.4000    0.0265    0.0600    0.0175    0.0110
## [43]    0.0050    0.1330    0.7000    0.1050    0.0140    1.2500   27.7500
## [50]   26.1000   81.2820   50.0000   80.7495   12.6175    0.2750    0.5500
## [57]    0.0105    0.8100   43.0000   26.5900    0.5500   30.0000    1.8000
## [64]    0.1600    0.0220    0.3715    0.0375    0.0740    0.0610    0.4600
## [71]    0.0505    0.1025    0.0240   43.1250    2.2500    0.0560  103.7505
## [78]    0.4500    0.0520   86.6650    1.0000    1.6900    2.1150
```

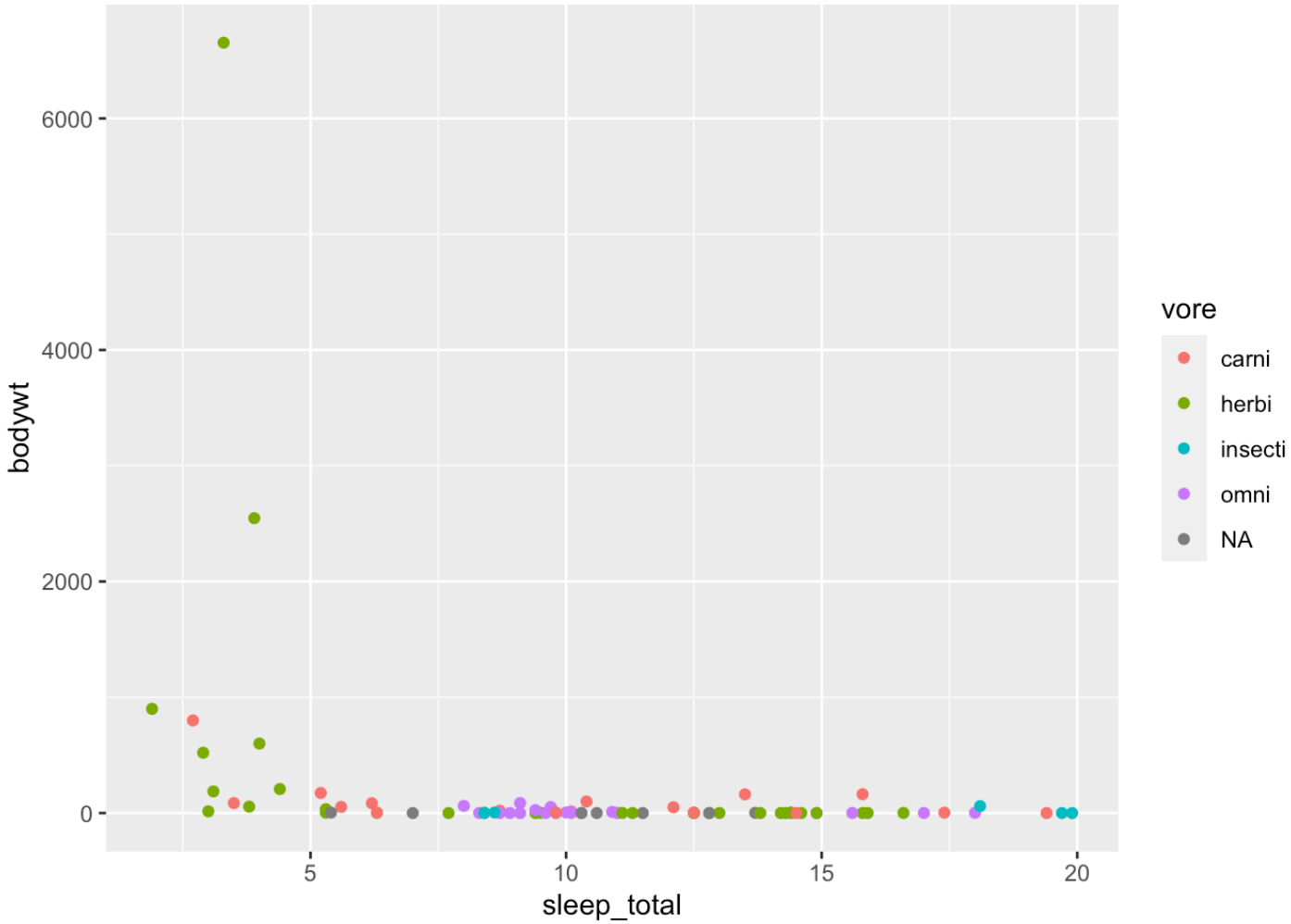
```
# Transform msleep$sleep_total to 'sleep_total' for better understanding
sleep_total <- msleep$sleep_total
bodywt <- msleep$bodywt

# Create a histogram with one variable
library(ggplot2)
ggplot(data = msleep, aes(x = sleep_total)) +
  geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
# Create a scatter plot
ggplot(data = msleep, aes(x = sleep_total, y = bodywt, color = vore)) +
  geom_point()
```


6000
4000
2000
0
bodywt
5
10
15
20
sleep_total
vore
carni
herbi
insecti
omni
NA