The dplyr Library for Data Manipulation Part-1

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
# Load the Motor Trend Car Road Tests (mtcars) dataset
carData = read.csv('mtcars.csv')
#carData = data('mtcars')

# Print the first five rows (or samples) in the data frame
head(carData, 5)
```

```
##                   X  mpg cyl disp  hp drat    wt  qsec vs am gear carb
## 1         Mazda RX4 21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## 2     Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## 3        Datsun 710 22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## 4    Hornet 4 Drive 21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## 5 Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
```

```
# Print the structure of the data frame
str(carData)
```

```
## 'data.frame':    32 obs. of  12 variables:
##  $ X   : chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
##  $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
##  $ cyl : int  6 6 4 6 8 6 8 4 4 6 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : int  110 110 93 110 175 105 245 62 95 123 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
##  $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
##  $ qsec: num  16.5 17 18.6 19.4 17 ...
##  $ vs  : int  0 0 1 1 0 1 0 1 1 1 ...
##  $ am  : int  1 1 1 0 0 0 0 0 0 0 ...
##  $ gear: int  4 4 4 3 3 3 3 4 4 4 ...
##  $ carb: int  4 4 1 1 2 1 4 2 2 4 ...
```

```
# Print the names of the columns (features or variables)
colnames(carData)
```

```
##  [1] "X"    "mpg"  "cyl"  "disp" "hp"   "drat" "wt"   "qsec" "vs"   "am"
## [11] "gear" "carb"
```

```r
# Print the number of samples (rows) and features (columns) in the data frame
nrow(carData)
```

```
## [1] 32
```

```r
ncol(carData)
```

```
## [1] 12
```

```r
# Create a vector of categorical columns
categorical_cols = c('vs', 'am')

# Convert the columns to factor type
carData[categorical_cols] = lapply(carData[categorical_cols], as.factor)

# Print the structure of the resulting dataframe
str(carData)
```

```
## 'data.frame':    32 obs. of  12 variables:
##  $ X   : chr  "Mazda RX4" "Mazda RX4 Wag" "Datsun 710" "Hornet 4 Drive" ...
##  $ mpg : num  21 21 22.8 21.4 18.7 18.1 14.3 24.4 22.8 19.2 ...
##  $ cyl : int  6 6 4 6 8 6 8 4 4 6 ...
##  $ disp: num  160 160 108 258 360 ...
##  $ hp  : int  110 110 93 110 175 105 245 62 95 123 ...
##  $ drat: num  3.9 3.9 3.85 3.08 3.15 2.76 3.21 3.69 3.92 3.92 ...
##  $ wt  : num  2.62 2.88 2.32 3.21 3.44 ...
##  $ qsec: num  16.5 17 18.6 19.4 17 ...
##  $ vs  : Factor w/ 2 levels "0","1": 1 1 2 2 1 2 1 2 2 2 ...
##  $ am  : Factor w/ 2 levels "0","1": 2 2 2 1 1 1 1 1 1 1 ...
##  $ gear: int  4 4 4 3 3 3 3 4 4 4 ...
##  $ carb: int  4 4 1 1 2 1 4 2 2 4 ...
```

```r
# Select only the feature wt
select(carData, wt)
```

```
##       wt
## 1  2.620
## 2  2.875
## 3  2.320
## 4  3.215
## 5  3.440
## 6  3.460
## 7  3.570
## 8  3.190
## 9  3.150
## 10 3.440
## 11 3.440
## 12 4.070
## 13 3.730
## 14 3.780
## 15 5.250
## 16 5.424
## 17 5.345
## 18 2.200
## 19 1.615
## 20 1.835
## 21 2.465
## 22 3.520
## 23 3.435
## 24 3.840
## 25 3.845
## 26 1.935
## 27 2.140
## 28 1.513
## 29 3.170
## 30 2.770
## 31 3.570
## 32 2.780
```

```
# Select everything except the feature wt
select(carData, -wt)
```

```
##                        X  mpg cyl  disp  hp drat  qsec vs am gear carb
## 1            Mazda RX4 21.0   6 160.0 110 3.90 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0   6 160.0 110 3.90 17.02  0  1    4    4
## 3           Datsun 710 22.8   4 108.0  93 3.85 18.61  1  1    4    1
## 4       Hornet 4 Drive 21.4   6 258.0 110 3.08 19.44  1  0    3    1
## 5    Hornet Sportabout 18.7   8 360.0 175 3.15 17.02  0  0    3    2
## 6              Valiant 18.1   6 225.0 105 2.76 20.22  1  0    3    1
## 7           Duster 360 14.3   8 360.0 245 3.21 15.84  0  0    3    4
## 8            Merc 240D 24.4   4 146.7  62 3.69 20.00  1  0    4    2
## 9             Merc 230 22.8   4 140.8  95 3.92 22.90  1  0    4    2
## 10            Merc 280 19.2   6 167.6 123 3.92 18.30  1  0    4    4
## 11           Merc 280C 17.8   6 167.6 123 3.92 18.90  1  0    4    4
## 12          Merc 450SE 16.4   8 275.8 180 3.07 17.40  0  0    3    3
## 13          Merc 450SL 17.3   8 275.8 180 3.07 17.60  0  0    3    3
## 14         Merc 450SLC 15.2   8 275.8 180 3.07 18.00  0  0    3    3
## 15  Cadillac Fleetwood 10.4   8 472.0 205 2.93 17.98  0  0    3    4
## 16 Lincoln Continental 10.4   8 460.0 215 3.00 17.82  0  0    3    4
## 17   Chrysler Imperial 14.7   8 440.0 230 3.23 17.42  0  0    3    4
## 18            Fiat 128 32.4   4  78.7  66 4.08 19.47  1  1    4    1
## 19         Honda Civic 30.4   4  75.7  52 4.93 18.52  1  1    4    2
## 20      Toyota Corolla 33.9   4  71.1  65 4.22 19.90  1  1    4    1
## 21       Toyota Corona 21.5   4 120.1  97 3.70 20.01  1  0    3    1
## 22    Dodge Challenger 15.5   8 318.0 150 2.76 16.87  0  0    3    2
## 23         AMC Javelin 15.2   8 304.0 150 3.15 17.30  0  0    3    2
## 24          Camaro Z28 13.3   8 350.0 245 3.73 15.41  0  0    3    4
## 25    Pontiac Firebird 19.2   8 400.0 175 3.08 17.05  0  0    3    2
## 26           Fiat X1-9 27.3   4  79.0  66 4.08 18.90  1  1    4    1
## 27       Porsche 914-2 26.0   4 120.3  91 4.43 16.70  0  1    5    2
## 28        Lotus Europa 30.4   4  95.1 113 3.77 16.90  1  1    5    2
## 29      Ford Pantera L 15.8   8 351.0 264 4.22 14.50  0  1    5    4
## 30        Ferrari Dino 19.7   6 145.0 175 3.62 15.50  0  1    5    6
## 31       Maserati Bora 15.0   8 301.0 335 3.54 14.60  0  1    5    8
## 32          Volvo 142E 21.4   4 121.0 109 4.11 18.60  1  1    4    2
```

```
# Select only the features cyl and wt
select(carData, c(cyl, wt)) # Quotes not required for column names because dplyr knows this from the da
```

```
##    cyl    wt
## 1    6 2.620
## 2    6 2.875
## 3    4 2.320
## 4    6 3.215
## 5    8 3.440
## 6    6 3.460
## 7    8 3.570
## 8    4 3.190
## 9    4 3.150
## 10   6 3.440
## 11   6 3.440
## 12   8 4.070
## 13   8 3.730
## 14   8 3.780
## 15   8 5.250
## 16   8 5.424
## 17   8 5.345
## 18   4 2.200
## 19   4 1.615
## 20   4 1.835
## 21   4 2.465
## 22   8 3.520
## 23   8 3.435
## 24   8 3.840
## 25   8 3.845
## 26   4 1.935
## 27   4 2.140
## 28   4 1.513
## 29   8 3.170
## 30   6 2.770
## 31   8 3.570
## 32   4 2.780
```

```
# Select everything except the features cyl and wt
select(carData, -c(cyl, wt))
```

```
##                       X  mpg  disp  hp drat  qsec vs am gear carb
## 1            Mazda RX4 21.0 160.0 110 3.90 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0 160.0 110 3.90 17.02  0  1    4    4
## 3           Datsun 710 22.8 108.0  93 3.85 18.61  1  1    4    1
## 4       Hornet 4 Drive 21.4 258.0 110 3.08 19.44  1  0    3    1
## 5    Hornet Sportabout 18.7 360.0 175 3.15 17.02  0  0    3    2
## 6              Valiant 18.1 225.0 105 2.76 20.22  1  0    3    1
## 7           Duster 360 14.3 360.0 245 3.21 15.84  0  0    3    4
## 8            Merc 240D 24.4 146.7  62 3.69 20.00  1  0    4    2
## 9             Merc 230 22.8 140.8  95 3.92 22.90  1  0    4    2
## 10            Merc 280 19.2 167.6 123 3.92 18.30  1  0    4    4
## 11           Merc 280C 17.8 167.6 123 3.92 18.90  1  0    4    4
## 12          Merc 450SE 16.4 275.8 180 3.07 17.40  0  0    3    3
## 13          Merc 450SL 17.3 275.8 180 3.07 17.60  0  0    3    3
## 14         Merc 450SLC 15.2 275.8 180 3.07 18.00  0  0    3    3
## 15  Cadillac Fleetwood 10.4 472.0 205 2.93 17.98  0  0    3    4
## 16 Lincoln Continental 10.4 460.0 215 3.00 17.82  0  0    3    4
## 17   Chrysler Imperial 14.7 440.0 230 3.23 17.42  0  0    3    4
## 18            Fiat 128 32.4  78.7  66 4.08 19.47  1  1    4    1
## 19         Honda Civic 30.4  75.7  52 4.93 18.52  1  1    4    2
## 20      Toyota Corolla 33.9  71.1  65 4.22 19.90  1  1    4    1
## 21       Toyota Corona 21.5 120.1  97 3.70 20.01  1  0    3    1
## 22    Dodge Challenger 15.5 318.0 150 2.76 16.87  0  0    3    2
## 23         AMC Javelin 15.2 304.0 150 3.15 17.30  0  0    3    2
## 24          Camaro Z28 13.3 350.0 245 3.73 15.41  0  0    3    4
## 25    Pontiac Firebird 19.2 400.0 175 3.08 17.05  0  0    3    2
## 26           Fiat X1-9 27.3  79.0  66 4.08 18.90  1  1    4    1
## 27       Porsche 914-2 26.0 120.3  91 4.43 16.70  0  1    5    2
## 28        Lotus Europa 30.4  95.1 113 3.77 16.90  1  1    5    2
## 29      Ford Pantera L 15.8 351.0 264 4.22 14.50  0  1    5    4
## 30        Ferrari Dino 19.7 145.0 175 3.62 15.50  0  1    5    6
## 31       Maserati Bora 15.0 301.0 335 3.54 14.60  0  1    5    8
## 32          Volvo 142E 21.4 121.0 109 4.11 18.60  1  1    4    2
```

```
# Filter cars with V-shaped engine
filter(carData, vs == 0)
```

```
##                      X  mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1            Mazda RX4 21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## 3    Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## 4           Duster 360 14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## 5           Merc 450SE 16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## 6           Merc 450SL 17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## 7          Merc 450SLC 15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## 8   Cadillac Fleetwood 10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## 9  Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## 10   Chrysler Imperial 14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## 11    Dodge Challenger 15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## 12         AMC Javelin 15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## 13          Camaro Z28 13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## 14    Pontiac Firebird 19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## 15       Porsche 914-2 26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## 16      Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## 17        Ferrari Dino 19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## 18       Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
```

```
# Filter cars with V-shaped engine and manual transmission
filter(carData, vs == 0 & am == 1)
```

```
##                X  mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1      Mazda RX4 21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## 2  Mazda RX4 Wag 21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## 3  Porsche 914-2 26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## 4 Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## 5   Ferrari Dino 19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## 6  Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
```

```
# Filter cars with V-shaped engine and manual transmission
# and hp greater than 150 or less than 100
filter(carData, vs == 0 & am == 1 & (hp < 100 | hp > 150))
```

```
##                X  mpg cyl  disp  hp drat   wt qsec vs am gear carb
## 1  Porsche 914-2 26.0   4 120.3  91 4.43 2.14 16.7  0  1    5    2
## 2 Ford Pantera L 15.8   8 351.0 264 4.22 3.17 14.5  0  1    5    4
## 3   Ferrari Dino 19.7   6 145.0 175 3.62 2.77 15.5  0  1    5    6
## 4  Maserati Bora 15.0   8 301.0 335 3.54 3.57 14.6  0  1    5    8
```

```r
# Select only the feature cyl and wt for cars with v-shaped engine
carData %>% filter(vs == 0) %>% select(cyl, wt)
```

```
##    cyl    wt
## 1    6 2.620
## 2    6 2.875
## 3    8 3.440
## 4    8 3.570
## 5    8 4.070
## 6    8 3.730
## 7    8 3.780
## 8    8 5.250
## 9    8 5.424
## 10   8 5.345
## 11   8 3.520
## 12   8 3.435
## 13   8 3.840
## 14   8 3.845
## 15   4 2.140
## 16   8 3.170
## 17   6 2.770
## 18   8 3.570
```

```r
# Select only the feature cyl and wt for cars with v-shaped engine
# and hp greater than 150
carData %>% filter(vs == 0 & hp > 150) %>% select(cyl, wt)
```

```
##    cyl    wt
## 1    8 3.440
## 2    8 3.570
## 3    8 4.070
## 4    8 3.730
## 5    8 3.780
## 6    8 5.250
## 7    8 5.424
## 8    8 5.345
## 9    8 3.840
## 10   8 3.845
## 11   8 3.170
## 12   6 2.770
## 13   8 3.570
```

```r
# Create a new column called wtton
mutate(carData, wtton = 0.45*wt)
```

```
##                      X  mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1            Mazda RX4 21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## 3           Datsun 710 22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## 4       Hornet 4 Drive 21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## 5    Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## 6              Valiant 18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## 7           Duster 360 14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## 8            Merc 240D 24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## 9             Merc 230 22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## 10            Merc 280 19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## 11           Merc 280C 17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## 12          Merc 450SE 16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## 13          Merc 450SL 17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## 14         Merc 450SLC 15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## 15  Cadillac Fleetwood 10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## 16 Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## 17   Chrysler Imperial 14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## 18            Fiat 128 32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## 19         Honda Civic 30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## 20      Toyota Corolla 33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## 21       Toyota Corona 21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## 22    Dodge Challenger 15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## 23         AMC Javelin 15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## 24          Camaro Z28 13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## 25    Pontiac Firebird 19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## 26           Fiat X1-9 27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## 27       Porsche 914-2 26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## 28        Lotus Europa 30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## 29      Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## 30        Ferrari Dino 19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## 31       Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
## 32          Volvo 142E 21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
##      wtton
## 1  1.17900
## 2  1.29375
## 3  1.04400
## 4  1.44675
## 5  1.54800
## 6  1.55700
## 7  1.60650
## 8  1.43550
## 9  1.41750
## 10 1.54800
## 11 1.54800
## 12 1.83150
## 13 1.67850
## 14 1.70100
## 15 2.36250
## 16 2.44080
## 17 2.40525
## 18 0.99000
## 19 0.72675
## 20 0.82575
## 21 1.10925
## 22 1.58400
## 23 1.54575
## 24 1.72800
## 25 1.73025
## 26 0.87075
## 27 0.96300
## 28 0.68085
## 29 1.42650
## 30 1.24650
## 31 1.60650
## 32 1.25100
```

```
# Create a new column called wtton by ensuring
# cars with NA weight values are negelected
carData = carData %>% filter(!is.na(wt)) %>% mutate(wtton = 0.45*wt)
```

```
# Check if wtton is a new column in the dataframe
colnames(carData)
```

```
##  [1] "X"     "mpg"   "cyl"   "disp"  "hp"    "drat"  "wt"    "qsec"  "vs"   
## [10] "am"    "gear"  "carb"  "wtton"
```

```
# Add a new column called cyltype with value High
# is cyl is greater than 4 and Low otherwise
carData %>% mutate(cyltype = ifelse(cyl > 4, 'High', 'Low'))
```

```
##                      X  mpg cyl  disp  hp drat    wt  qsec vs am gear carb
## 1            Mazda RX4 21.0   6 160.0 110 3.90 2.620 16.46  0  1    4    4
## 2        Mazda RX4 Wag 21.0   6 160.0 110 3.90 2.875 17.02  0  1    4    4
## 3           Datsun 710 22.8   4 108.0  93 3.85 2.320 18.61  1  1    4    1
## 4       Hornet 4 Drive 21.4   6 258.0 110 3.08 3.215 19.44  1  0    3    1
## 5    Hornet Sportabout 18.7   8 360.0 175 3.15 3.440 17.02  0  0    3    2
## 6              Valiant 18.1   6 225.0 105 2.76 3.460 20.22  1  0    3    1
## 7           Duster 360 14.3   8 360.0 245 3.21 3.570 15.84  0  0    3    4
## 8            Merc 240D 24.4   4 146.7  62 3.69 3.190 20.00  1  0    4    2
## 9             Merc 230 22.8   4 140.8  95 3.92 3.150 22.90  1  0    4    2
## 10            Merc 280 19.2   6 167.6 123 3.92 3.440 18.30  1  0    4    4
## 11           Merc 280C 17.8   6 167.6 123 3.92 3.440 18.90  1  0    4    4
## 12          Merc 450SE 16.4   8 275.8 180 3.07 4.070 17.40  0  0    3    3
## 13          Merc 450SL 17.3   8 275.8 180 3.07 3.730 17.60  0  0    3    3
## 14         Merc 450SLC 15.2   8 275.8 180 3.07 3.780 18.00  0  0    3    3
## 15  Cadillac Fleetwood 10.4   8 472.0 205 2.93 5.250 17.98  0  0    3    4
## 16 Lincoln Continental 10.4   8 460.0 215 3.00 5.424 17.82  0  0    3    4
## 17   Chrysler Imperial 14.7   8 440.0 230 3.23 5.345 17.42  0  0    3    4
## 18            Fiat 128 32.4   4  78.7  66 4.08 2.200 19.47  1  1    4    1
## 19         Honda Civic 30.4   4  75.7  52 4.93 1.615 18.52  1  1    4    2
## 20      Toyota Corolla 33.9   4  71.1  65 4.22 1.835 19.90  1  1    4    1
## 21       Toyota Corona 21.5   4 120.1  97 3.70 2.465 20.01  1  0    3    1
## 22    Dodge Challenger 15.5   8 318.0 150 2.76 3.520 16.87  0  0    3    2
## 23         AMC Javelin 15.2   8 304.0 150 3.15 3.435 17.30  0  0    3    2
## 24          Camaro Z28 13.3   8 350.0 245 3.73 3.840 15.41  0  0    3    4
## 25    Pontiac Firebird 19.2   8 400.0 175 3.08 3.845 17.05  0  0    3    2
## 26           Fiat X1-9 27.3   4  79.0  66 4.08 1.935 18.90  1  1    4    1
## 27       Porsche 914-2 26.0   4 120.3  91 4.43 2.140 16.70  0  1    5    2
## 28        Lotus Europa 30.4   4  95.1 113 3.77 1.513 16.90  1  1    5    2
## 29      Ford Pantera L 15.8   8 351.0 264 4.22 3.170 14.50  0  1    5    4
## 30        Ferrari Dino 19.7   6 145.0 175 3.62 2.770 15.50  0  1    5    6
## 31       Maserati Bora 15.0   8 301.0 335 3.54 3.570 14.60  0  1    5    8
## 32          Volvo 142E 21.4   4 121.0 109 4.11 2.780 18.60  1  1    4    2
##      wtton cyltype
## 1  1.17900    High
## 2  1.29375    High
## 3  1.04400     Low
## 4  1.44675    High
## 5  1.54800    High
## 6  1.55700    High
## 7  1.60650    High
## 8  1.43550     Low
## 9  1.41750     Low
## 10 1.54800    High
## 11 1.54800    High
## 12 1.83150    High
## 13 1.67850    High
## 14 1.70100    High
## 15 2.36250    High
## 16 2.44080    High
## 17 2.40525    High
## 18 0.99000     Low
## 19 0.72675     Low
## 20 0.82575     Low
## 21 1.10925     Low
## 22 1.58400    High
## 23 1.54575    High
## 24 1.72800    High
## 25 1.73025    High
## 26 0.87075     Low
## 27 0.96300     Low
## 28 0.68085     Low
## 29 1.42650    High
## 30 1.24650    High
## 31 1.60650    High
## 32 1.25100     Low
```

```
# Create a new data frame with the cyltype and wtton added
carData.new = carData %>% mutate(cyltype = ifelse(cyl > 4, 'High', 'Low'), wtton = 0.45*wt)
head(carData.new, 5)
```

```
##                   X  mpg cyl disp  hp drat    wt  qsec vs am gear carb   wtton
## 1         Mazda RX4 21.0   6  160 110 3.90 2.620 16.46  0  1    4    4 1.17900
## 2     Mazda RX4 Wag 21.0   6  160 110 3.90 2.875 17.02  0  1    4    4 1.29375
## 3        Datsun 710 22.8   4  108  93 3.85 2.320 18.61  1  1    4    1 1.04400
## 4    Hornet 4 Drive 21.4   6  258 110 3.08 3.215 19.44  1  0    3    1 1.44675
## 5 Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2 1.54800
##   cyltype
## 1    High
## 2    High
## 3     Low
## 4    High
## 5    High
```

```
# Return the mean weight (in tons) of the cars
carData.new %>% summarise(mean(wtton))
```

```
##   mean(wtton)
## 1    1.447763
```

```
# Return the mean weight (in tons) and mean or minimum displacement
# of the cars
carData.new %>% summarise(mean(wtton), mean(disp))
```

```
##   mean(wtton) mean(disp)
## 1    1.447763   230.7219
```

```
carData.new %>% summarise(mean(wtton), min(disp))
```

```
##   mean(wtton) min(disp)
## 1    1.447763      71.1
```

```
# Group cars according to cyltype and calculate mean weight and
# mean displacement
carData.new %>% group_by(cyltype) %>% summarise(mean(wtton), mean(disp))
```

```
## # A tibble: 2 x 3
##   cyltype `mean(wtton)` `mean(disp)`
##   <chr>           <dbl>        <dbl>
## 1 High             1.67         297.
## 2 Low              1.03         105.
```

```
# Group cars according to cyltype and calculate mean weight and
# minimum displacement
carData.new %>% group_by(cyltype) %>% summarise(mean(wtton), min(disp))
```

```
## # A tibble: 2 x 3
##   cyltype `mean(wtton)` `min(disp)`
##   <chr>           <dbl>       <dbl>
## 1 High             1.67       145
## 2 Low              1.03        71.1
```

```
# Group cars according to cyltype and calculate mean weight and
# minimum displacement followed by arranging in increasing order
# of mean weight

carData.new %>% group_by(cyltype) %>% summarise(mwt = mean(wtton), mdisp = min(disp)) %>% arrange(mwt)
```

```
## # A tibble: 2 x 3
##   cyltype   mwt mdisp
##   <chr>   <dbl> <dbl>
## 1 Low      1.03  71.1
## 2 High     1.67 145
```

```
# Group cars according to cyltype and calculate mean weight and
# minimum displacement followed by arranging in decreasing order
# of mean weight

carData.new %>% group_by(cyltype) %>% summarise(mwt = mean(wtton), mdisp = min(disp)) %>% arrange(desc(
```

```
## # A tibble: 2 x 3
##   cyltype   mwt mdisp
##   <chr>   <dbl> <dbl>
## 1 High     1.67 145
## 2 Low      1.03  71.1
```

```
# Return number of samples for each cyltype
carData.new %>% group_by(cyltype) %>% count()
```

```
## # A tibble: 2 x 2
## # Groups:   cyltype [2]
##   cyltype     n
##   <chr>   <int>
## 1 High       21
## 2 Low        11
```

```
# Return number of samples for each cyltype and vs
carData.new %>% count(cyltype, vs)
```

```
##   cyltype vs  n
## 1    High  0 17
## 2    High  1  4
## 3     Low  0  1
## 4     Low  1 10
```

```
# Return number of samples for each cyltype and sort by
# increasing order of count
carData.new %>% count(cyltype) %>% arrange(n)
```

```
##   cyltype  n
## 1     Low 11
## 2    High 21
```