Car_viz_h

Carolyn Wright

```
# call built-in data mtcars.
data(mtcars)
```

Documentation on mtcars dataset

Motor Trend Car Road Tests

- Description:
  - The data was extracted from the 1974 Motor Trend US magazine, and comprises fuel consumption and 10 aspects of automobile design and performance for 32 automobiles (1973–74 models).

Usage

- mtcars–>

Format

- A data frame with 32 observations on 11 (numeric) variables.
  - [, 1] mpg Miles/(US) gallon
  - [, 2] cyl Number of cylinders
  - [, 3] disp Displacement (cu.in.)
  - [, 4] hp Gross horsepower
  - [, 5] drat Rear axle ratio
  - [, 6] wt Weight (1000 lbs)
  - [, 7] qsec 1/4 mile time
  - [, 8] vs Engine (0 = V-shaped, 1 = straight)
  - [, 9] am Transmission (0 = automatic, 1 = manual)
  - [,10] gear Number of forward gears
  - [,11] carb Number of carburetors

Note

- Henderson and Velleman (1981) comment in a footnote to Table 1: Hocking [original transcriber]s noncrucial coding of the Mazdas rotary engine as a straight six-cylinder engine and the Porsches flat engine as a V engine, as well as the inclusion of the diesel Mercedes 240D, have been retained to enable direct comparisons to be made with previous analyses.

Source

- Henderson and Velleman (1981), Building multiple regression models interactively. Biometrics, 37, 391–411.

```r
# Select only car models where mpg<20
mtcars_mpg2 <- mtcars[mtcars$mpg < 20,]

# Reduce the variables to mpg, cyl, disp, hp, gears
mtcars_mpg2 <- mtcars_mpg2[, c(1,2,3,4,10)]

# read the R file hand_functions.R so that it can be used
# note I have changed echo to false so that the source code will not be displayed
source(file = "hand_functions.R", echo = FALSE)

# Now use the function from hand_functions.R
sp_out <- sum_special(mtcars_mpg2)
```

Look at some results from 'sum_special' function

- Means:

```r
#call the sp_means that is a list of means included in sp_out
sp_out$sp_means
```

```
##        mpg        cyl       disp         hp       gear
##  15.900000   7.555556 313.811111 191.944444   3.444444
```

- Variance:

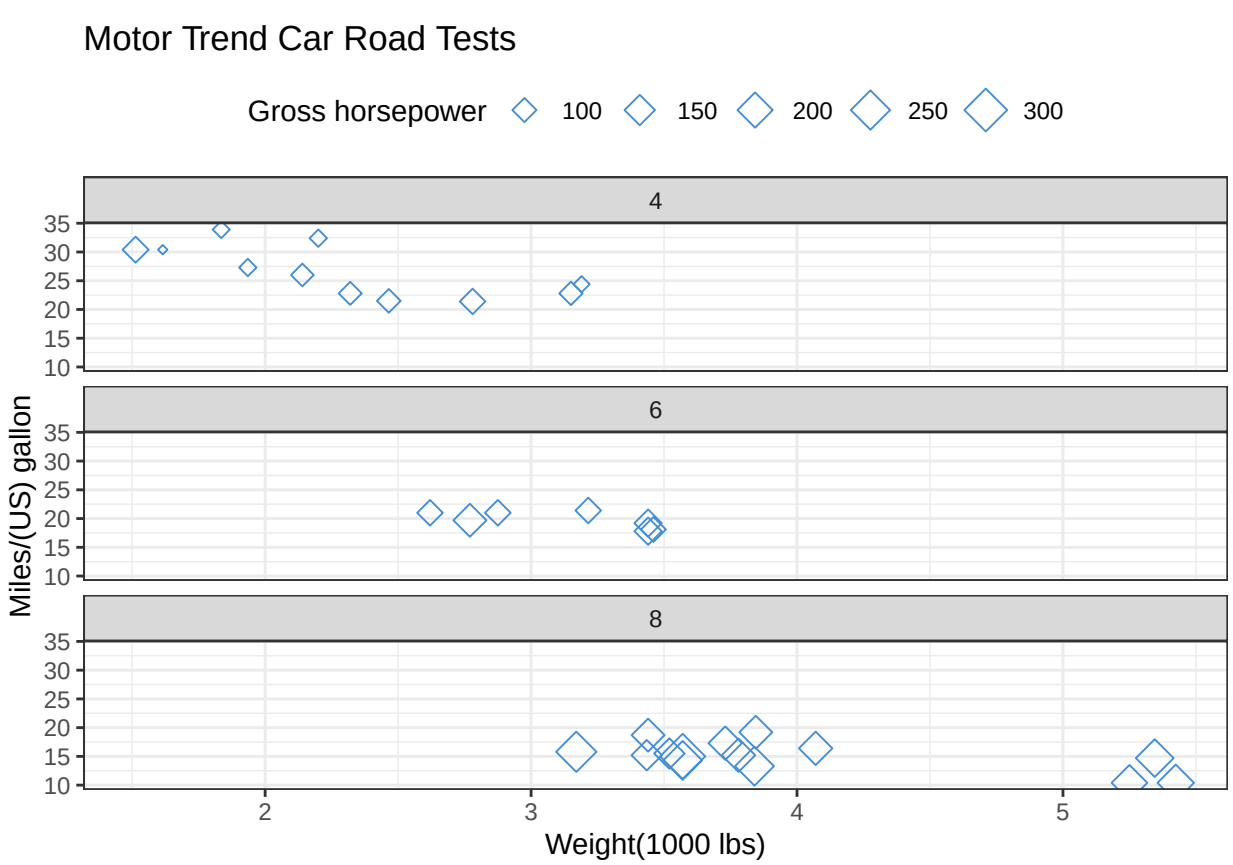
```r
#call the sp_var that is a list of variances included in sp_out
sp_out$sp_var
```

```
##          mpg          cyl         disp           hp         gear
##    7.5258824    0.7320261 9438.7645752 3253.5849673    0.6143791
```

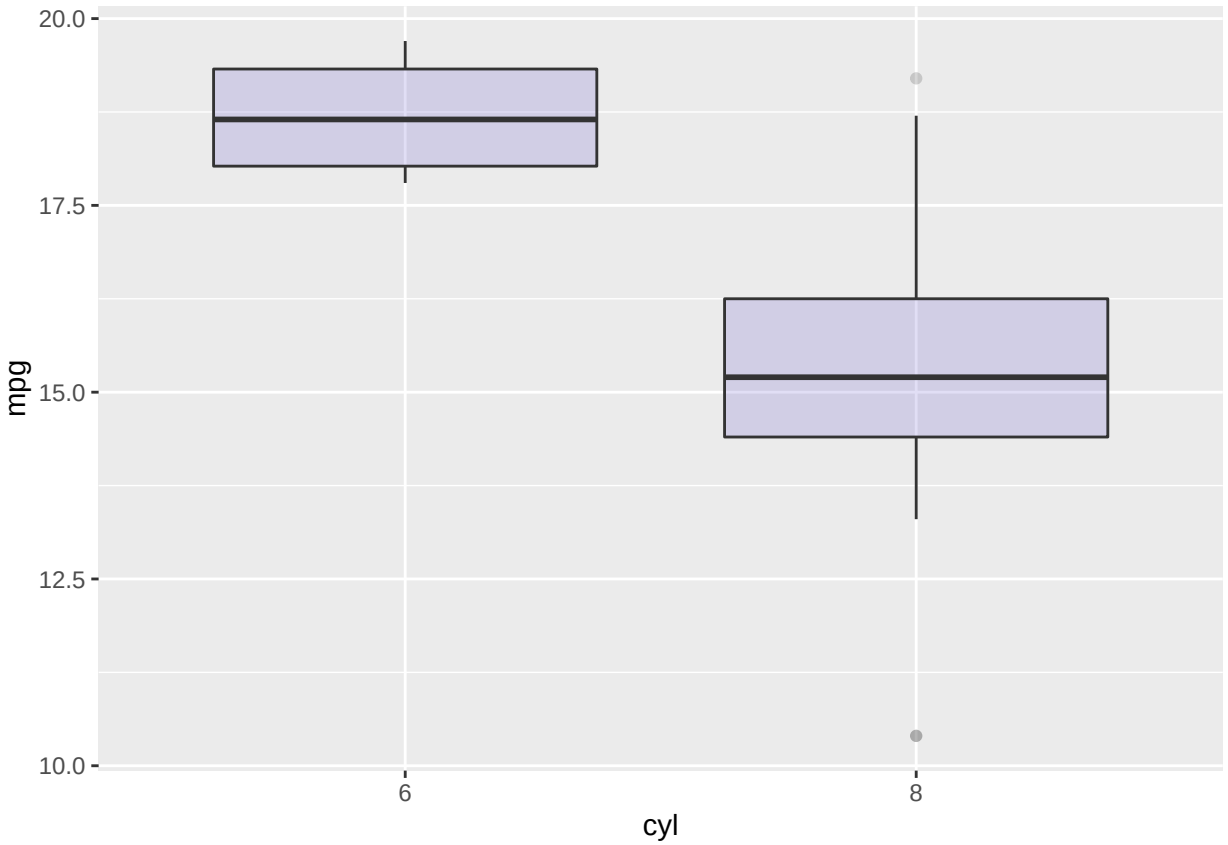
Investigate the esquisse package

```r
#esquisser(data = mtcars, viewer = "browser")

#include code that is created through the esquisse package
ggplot(mtcars) +
 aes(x = wt, y = mpg, size = hp) +
 geom_point(shape = "diamond filled", colour = "#458ED2") +
 labs(x = "Weight(1000 lbs)", y = "Miles/(US) gallon",
      title = "Motor Trend Car Road Tests", size = "Gross horsepower") +
 theme_bw() +
 theme(legend.position = "top") +
 facet_wrap(vars(cyl), ncol = 1L)
```




```
# note that this boxplot cannot be made with esquisse() unless
# the data is adjusted.  What adjustment is needed?
ggplot(mtcars_mpg2, aes(x=as.factor(cyl), y=mpg)) +
  geom_boxplot(fill="slateblue", alpha=0.2) +
  xlab("cyl")
```



Discussion of above boxplot

- The above boxplot cannot be created using esquisse() due to the fact that the variable 'cyl' is not defined as a factor in the raw data. Below is the code for changing 'cyl' into a factor variable in the raw data.

```
#convert 'cyl' into a factor in the raw data
mtcars$cyl <-as.factor(mtcars$cyl)

#check variable type--> output should be "TRUE" if the conversion worked properly
is.factor(mtcars$cyl)
```

```
## [1] TRUE
```