## 1 mtcars example

Several data sets, intended as learning tools, are automatically installed when R is installed. Many more are installed within packages to complement learning to use those packages. One of these is the famous **mtcars** data set, which is used in many data mining exercises.

```
> data(mtcars)
> head(mtcars)
                  mpg cyl disp hp drat
                                          wt qsec vs am gear carb
Mazda RX4
                 21.0
                       6 160 110 3.90 2.620 16.46 0
Mazda RX4 Wag
                 21.0
                       6 160 110 3.90 2.875 17.02 0 1
Datsun 710
                 22.8
                       4 108 93 3.85 2.320 18.61 1 1
                                                                1
Hornet 4 Drive
                 21.4
                       6 258 110 3.08 3.215 19.44 1 0
                                                                1
Hornet Sportabout 18.7
                       8 360 175 3.15 3.440 17.02 0 0
                                                           3
                                                                2
                       6 225 105 2.76 3.460 20.22 1 0
Valiant
                 18.1
                                                                1
```

Suppose we fit a model with mpg (miles per gallon) as the response variable and cyl and wt (number of cylinders and weight of the car) as the predictor variables. We will call this fitted model fit.

```
fit <- lm(mpg ~ cyl + wt, data=mtcars)
```

# > summary(fit)

### Call:

lm(formula = mpg ~ cyl + wt, data = mtcars)

#### Residuals:

Min 1Q Median 3Q Max -4.2893 -1.5512 -0.4684 1.5743 6.1004

#### Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 39.6863 1.7150 23.141 < 2e-16 \*\*\*

cyl -1.5078 0.4147 -3.636 0.001064 \*\*

wt -3.1910 0.7569 -4.216 0.000222 \*\*\*

---

Signif. codes: 0 \*\*\* 0.001 \*\* 0.01 \* 0.05 . 0.1 1

Residual standard error: 2.568 on 29 degrees of freedom Multiple R-squared: 0.8302, Adjusted R-squared: 0.8185

F-statistic: 70.91 on 2 and 29 DF, p-value: 6.809e-12

## residuals(fit1)

```
R File Edit View Misc Packages Windows Help
> residuals(fit)
         Mazda RX4
                       Mazda RX4 Wag
                                              Datsun 710
                                                             Hornet 4 Drive
                                                                                Hornet Sportabout
                        -0.46544677
       -1.27914467
                                              -3.45202624
                                                                    1.01948376
                                                                                        2.05304242
       Merc 230
-0.80351937
                           Merc 280
-0.46254751
                                               Merc 280C -1.86254751
                                                                    Merc 450SE
                                                                                       Merc 450SL
1.57842434
                                                                    1.76335487
  Chrysler Imperial
                                              Honda Civic
                                                               Toyota Corolla
                                                                                   Toyota Corona
                          Fiat 125
5.76505710
Fiat X1-9
        4.13184435
                                                1.89833840
                                                                    6.10035227
                                                                                       -4.28933528
                                           Porsche 914-2
   Pontiac Firebird
                                                                  Lotus Europa
                                                                                 Ford Pantera L
        3.84538614
                          -0.18055052
                                               -0.82640123
                                                                    1.57285924
                                                                                       -1.70852005
>
              McAfee LiveSafe – Internet Security
```

Figure 1:

```
> sum(residuals(fit))
[1] 1.096345e-15

> #Shapiro-Wilk Test for Normality
> shapiro.test(resid(fit))

Shapiro-Wilk normality test

data: resid(fit)
W = 0.9375, p-value = 0.06341
```

