1. The model adopted is

\[ Y_{ij} = \mu + \tau_i + \varepsilon_{ij} \quad i = 1, \ldots, 3, j = 1, \ldots, 8, \]

where the parameter \( \mu \) is the overall mean, the parameters \( \tau_i \) are the location effects and the \( \varepsilon_{ij} \) are random errors, assumed \( \text{NID}(0, \sigma^2) \) with \( \sigma^2 \) unknown. The constraint \( \sum_{i=1}^{3} \tau_i = 0 \) is imposed. Using a one-way ANOVA, we test the null hypothesis, \( H_0 : \tau_1 = \tau_2 = \tau_3 = 0 \), against the alternative \( H_1 : \tau_i \neq 0 \) for at least one \( i \). The following S+ output gives a \( p \)-value of 0.008, so that we reject \( H_0 \) at the 1\% significance level. There is very strong evidence of differences among the effects of the locations.

\[ \begin{array}{l}
> \text{bdA} \leftarrow \text{c}(81.6, 81.3, 82., 79.6, 78.4, 81.8, 80.2, 80.7) \\
> \text{bdB} \leftarrow \text{c}(81.8, 84.7, 82., 85.6, 79.9, 83.2, 84.1, 85.) \\
> \text{bdC} \leftarrow \text{c}(82.1, 79.6, 83.1, 80.7, 81.8, 79.9, 82.6, 81.9) \\
> \text{bd} \leftarrow \text{c(bdA, bdB, bdC)} \\
> \text{loc} \leftarrow \text{rep(1:3, rep(8, 3))} \\
> \text{brix} \leftarrow \text{data.frame(bd, location)} \\
> \text{rm(bdA, bdB, bdC, bd, loc, location)} \\
> \text{brix.aov} \leftarrow \text{aov(bd ~ location, data = brix)} \\
> \text{summary(brix.aov)} \\
\end{array} \]

\[ \begin{array}{l}
\text{Df} \quad \text{Sum of Sq} \quad \text{Mean Sq} \quad \text{F Value} \quad \text{Pr(F)} \\
\text{location} \quad 2 \quad 28.28583 \quad 14.14292 \quad 6.150686 \quad 0.007896784 \\
\text{Residuals} \quad 21 \quad 48.28750 \quad 2.29940 \\
\text{Residuals} \quad 21 \quad 48.28750 \quad 2.29940 \\
\end{array} \]

2. \( A \leftarrow \text{c}(12.8, 13.4, 11.2, 11.6, 9.4, 10.3, 14.1, 11.9, 10.5, 10.4) \)
\( B \leftarrow \text{c}(8.1, 10.3, 4.2, 7.8, 5.6, 8.1, 12.7, 6.8, 6.9, 6.4) \)
\( C \leftarrow \text{c}(9.8, 10.6, 9.1, 4.3, 11.2, 11.6, 8.3, 8.9, 9.2, 6.4) \)
\( D \leftarrow \text{c}(16.4, 8.2, 15.1, 10.4, 7.8, 9.2, 12.6, 11., 8., 9.8) \)
\( \text{content} \leftarrow \text{c(A, B, C, D)} \)
\( \text{sl} \leftarrow \text{rep(1:4, rep(10, 4))} \)
\( \text{content} \leftarrow \text{c(A, B, C, D)} \)
\( \# \text{not wise to use s, since s is already a function name} \)
\( \text{soil} \leftarrow \text{factor(sl)} \)
\( \text{moisture.data} \leftarrow \text{data.frame(content, soil)} \)
\( \text{rm(A, B, C, D, content, sl, soil)} \)
\( \text{moisture.data.aov} \leftarrow \text{aov(content ~ soil, data = moisture.data)} \)
\( \text{summary(moisture.data.aov)} \)

\[ \begin{array}{l}
\text{Df} \quad \text{Sum of Sq} \quad \text{Mean Sq} \quad \text{F Value} \quad \text{Pr(F)} \\
\text{soil} \quad 3 \quad 93.854 \quad 31.28467 \quad 5.722165 \quad 0.002616618 \\
\text{Residuals} \quad 36 \quad 196.822 \quad 5.46728 \\
\end{array} \]

The \( p \)-value of 0.003 for the F statistic in the ANOVA shows that there is very strong evidence for differences among the the mean moisture contents of the soils.
3. (i) > pos <- rep(1:4,rep(9,4))
> position <- factor(pos)
> wear <- c(20.935,17.123, ...
... > tyres <- data.frame(position,wear)
> by(wear,position,summary)

INDICES: 1
Min. 1st Qu. Median Mean 3rd Qu. Max.
15.29 17.12 20.33 21.62 28.09 29.59

INDICES: 2
Min. 1st Qu. Median Mean 3rd Qu. Max.
11.28 14.82 19.39 17.52 20.10 21.20

INDICES: 3
Min. 1st Qu. Median Mean 3rd Qu. Max.
28.00 29.18 30.53 32.58 37.23 40.02

INDICES: 4
Min. 1st Qu. Median Mean 3rd Qu. Max.
18.79 20.18 28.70 27.71 34.34 36.31

> plot(position,wear)

It appears that rear tyres experience more wear than front tyres and that offside tyres experience more wear than nearside tyres.

(ii) > tyres <- data.frame(position,wear)
> tyres.aov <- aov(wear ~ position, data = tyres)
> summary(tyres.aov)

Df Sum of Sq Mean Sq F Value Pr(>F)
position 3 1189.014 396.3379 13.73909 6.279273e-006
Residuals 32 923.119 28.8475

There is very strong evidence of overall differences in wear among the four positions for the tyres ($p \ll 0.01$).