Formulas Final

SSTO = $\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}^2 \quad -\frac{1}{IJ}\left(\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}\right)^2$

SSA = $\frac{1}{J}\sum_{i=1}^{I}\left(\sum_{j=1}^{J} x_{ij}\right)^2 - \frac{1}{IJ}\left(\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}\right)^2$

SSB = $\frac{1}{I}\sum_{i=1}^{I}\left(\sum_{j=1}^{J} x_{ij}\right)^2 - \frac{1}{IJ}\left(\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}\right)^2$

SSE = $\sum_{i=1}^{I}\sum_{j=1}^{J}(x_{ij} - \bar{x}_{i*} - \bar{x}_{*j} - \bar{x}_{**})^2$ or by differences

Source of Variation	df	SS	MS	F
Factor A	*I-1*	SSA	SSA/(*I-1*)	MSA/MSE
Factor B	*J-1*	SSB	SSA/(*J-1*)	MSB/MSE
Error	*(I-1)(J-1)*	SSE	SSE/ *(I-1)(J-1)*	
Total	*IJ - 1*	SSTO		

If both null hypotheses H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_I = \alpha$, and H_0: $\beta_1 = \beta_2 = \cdots = \beta_J = \beta$ are true:

E[MSTO] = E[MSA] = E[MSB] =E[MSE] = σ^2

Where σ^2 is the common variance of the normal populations from where each of the *IJ* data points.

If the null hypothesis H_0: $\alpha_1 = \alpha_2 = \cdots = \alpha_I = \alpha$ is **not** true:

E[MSE] = σ^2, but

E[MSA] = $\sigma^2 + \frac{J}{I-1}\sum_{i=1}^{I}(\alpha_i - \bar{\alpha})^2$

If the null hypothesis H_0: $\beta_1 = \beta_2 = \cdots = \beta_J = \beta$ is **not** true:

E[MSE] = σ^2, but

E[MSB] = $\sigma^2 + \frac{I}{J-1}\sum_{j=1}^{J}\left(\beta_i - \bar{\beta}\right)^2$

Note: E[MSTO] is not either equal to σ^2 if either of the two null hypotheses above is not true, but it is of no interest for purposes of constructing an ANOVA.
R code

Enter your data using the sub-indices of a matrix: For example, the matrix of data of Example 10.4.1 on p. 524 needs to be entered in a csv as follows:

Chem	Material	Score
1	1	3
1	2	7
1	3	6
2	1	9
2	2	11
2	3	8
3	1	2
3	2	5
3	3	7
4	1	7
4	2	9
4	3	8

Note that $x_{11} = 3$, $x_{12} = 7$, $x_{13} = 6$, $x_{21} = 9$, etc.

R code:

```
# Read csv file
datos <- read.csv("TheNameofYourFile.csv", header = TRUE)

# Visualize your data using boxplots

boxplot(Score ~ Chem, data=datos)
boxplot(Score ~ Material, data=datos)

# Eyeball the means for each level of each factor

with(datos, tapply(Score, Chem, mean))
with(datos, tapply(Score, Material, mean))

# create a factor variable FChem in object datos by using the variable Chem in object datos
datos$FLane <- factor(daos$Chem)

# create a factor variable FMaterial in object datos by using the variable Material in object datos
datos$Material <- factor(datos$Material)

# Fit the Model and find its ANOVA
model2 <- lm(Score ~ FChem + FMaterial, data = datos)
anova(model2)
```

```
# If the F statistics is significant for any of the factors, obtain the confidence interval for each
pair of means using the adjusted q statistic:

library(multcomp)
pairsChem <- glht(model2, linfct = mcp(FChem ="Tukey"), data = datos)
summary(pairsChem)
pairsMaterials <- glht(model21, linfct = mcp(FMaterial ="Tukey"), data = datos)
summary(pairsMaterial)
```