1.
```
ANALYSIS OF VARIANCE
SOURCE     DF       SS       MS
TREATMENT   3     5530     1843
ERROR       8     2679      335
TOTAL      11     8209
```

2. (a)
```
> sheer=read.table("http://www.stat.wisc.edu/~st571-1/data/sheer.dat",
+                  col.names=c("type","strength"))
> tapply(sheer$strength, sheer$type, stem, scale=2)

  45 | 8
  46 | 38
  47 | 39
  48 |
  49 | 9
  50 | 4
  51 | 7

  45 | 5
  46 |
  46 | 6
  47 |
  47 | 6
  48 | 34
  48 |
  49 | 23
  49 | 8

  48 | 3
  49 |
  50 | 0029
  51 | 9
  52 | 8
  53 | 8

  49 | 5
  50 |
  51 | 13
  52 | 0
  53 | 3
  54 | 03
  55 |
  56 | 2
```

```
48 | 2
49 | 69
50 | 3
51 | 024
52 |
53 | 2
```

(b)
```
> sheer.lm=lm(strength~type,sheer)
> anova(sheer.lm)

Analysis of Variance Table
Response: strength
          Df  Sum Sq Mean Sq F value    Pr(>F)
type       4 12227.9  3057.0  9.1994 3.451e-05 ***
Residuals 35 11630.5   332.3
Total     39 23858.4
```

(c) $X_{ij} = \mu_i + e_{ij}$ where $e_{ij} \sim N(0, \sigma^2)$
$i = 1, 2, ..., 5, j = 1, 2, ...8.$
X_{ij} is the j^{th} measurement in the i^{th} glue type.
μ_i is the mean measurement of the i^{th} glue type.

(d) Hypotheses $H_0 : \mu_1 = \mu_2 = ... = \mu_k$ H_A :not all μ_i are equal

$$F = 9.1994$$

$$\text{P-value} = \Pr(F_{4,35} \geq 9.1994) = 0.000$$

(e) 99% C.I. for $\mu_1 - \frac{1}{4}(\mu_2 + \mu_3 + \mu_4 + \mu_5)$
$\lambda = (1, -1/4, -1/4, -1/4, -1/4)$

$$\begin{aligned}
\sum_{i=1}^{k} \lambda_i \bar{X}_{i.} \pm t_{(N-k,\frac{\alpha}{2})} \sqrt{\text{MSErr}} \sqrt{\sum_{i=1}^{k} \frac{\lambda_i^2}{n_i}} &= -23.3475 \pm t_{35,0.005} \times 18.23 \times 0.3952847 \\
&= -23.3475 \pm 19.62785 \\
&= [-42.97535, -3.71965]
\end{aligned}$$

$\bar{X}_{i.}$ can be obtained by

```
> tapply(sheer$strength,sheer$type, mean)
      A       B       C       D       E
482.625 480.875 509.875 527.125 506.000
```

New experimental glues appear to have stronger dry sheer strength than the “standard” glue.

(f) $H_0 : (1/2)(\mu_2 + \mu_4) - (1/2)(\mu_3 + \mu_5) = 0 \;\; vs \;\; H_A : (1/2)(\mu_2 + \mu_4) - (1/2)(\mu_3 + \mu_5) \neq 0$
$\lambda = (0, 1/2, -1/2, 1/2, -1/2)$

Test statistic

$$\begin{aligned} T &= \frac{\sum_{i=1}^{k} \lambda_i \bar{X}_i}{\sqrt{\text{MSErr}}\sqrt{\sum_{i=1}^{k} \frac{\lambda_i^2}{n_i}}} \\ &= \frac{-3.935}{18.23 \times 0.3535534} \\ &= -0.6105244 \end{aligned}$$

$$\text{P-value} = 2\Pr(T_{35} \geq 0.6105244) = 0.5454588$$

Very weak evidence againt H_0.

3. (a)

```
> y1=c(18.5, 21.0, 23.5,26.2,22.5,24.0,21.5,24.8,20.5,24.5,19.5,22.5,21.5)
> y2=c(19.0,18.0,21.3,19.5,18.3,18.8)
> y3=c(14.7,16.1,20.2,16.7,18.1,14.7,18.3,17.8,17.2,15.8,17.5)
> y=c(y1,y2,y3)
> group=c(rep(1,13),rep(2,6),rep(3,11))
> data=data.frame(y=y,group=factor(group))
> fit=lm(y~group,data)
> anova(fit)

Analysis of Variance Table
Response: y
          Df  Sum Sq Mean Sq F value    Pr(>F)
group      2 172.419  86.209  25.003 7.168e-07 ***
Residuals 27  93.096   3.448
Total     29 265.515
```

(b)

$$F = 25.00$$

$$\text{P-value} = \Pr(F_{2,27} \geq 25.003) = 0.000$$

Very strong evidence that RH affects the average weight of newly formed pipae.

(c) Let μ_1, μ_2, μ_3 denote the mean weights pupae at 70%, 20%, and 1% RH, respectively.
Then 95% C.I. for $\mu_2 - \mu_1$ is
(Notice $\lambda = (-1, 1, 0)$)

$$\begin{aligned} \sum_{i=1}^{k} \lambda_i \bar{X}_{i.} \pm t_{(N-k,\frac{\alpha}{2})} \sqrt{\text{MSErr}} \sqrt{\sum_{i=1}^{k} \frac{\lambda_i^2}{n_i}} &= 19.150 - 22.346 \pm t_{27,0.025} \times \sqrt{3.45} \times \sqrt{1/13 + 1/6} \\ &= -3.196 \pm 1.881 \\ &= [-5.077, -1.315] \end{aligned}$$

(d) Test for $H_0 : \sigma_1^2 = \cdots = \sigma_3^2$ Use Levene's Test:

```
medians: 22.5 18.9 17.2

abs deviances:
gp1(13): 4.0 1.5 1.0 3.7 0.0 1.5 1.0 2.3 2.0 2.0 3.0 0.0 1.0
gp2(8) : 0.1 0.9 2.4 0.6 0.6 0.1
gp3(11): 2.5 1.1 3.0 0.5 0.9 2.5 1.1 0.6 0.0 1.4 0.3

delete exactly one zero from gp 1(13obs) and gp3(11obs)
gp1(12): 4.0 1.5 1.0 3.7 1.5 1.0 2.3 2.0 2.0 3.0 0.0 1.0
gp2(8) : 0.1 0.9 2.4 0.6 0.6 0.1
gp3(10): 2.5 1.1 3.0 0.5 0.9 2.5 1.1 0.6 1.4 0.3

ANOVA table for the data above
 SOURCE     DF        SS        MS        F        p
 TREATMENT   2      5.28      2.64     2.44    0.108
 ERROR      25     27.09      1.08
 TOTAL      27     32.37
```

P-value is 0.108, which is not strongly against the homogeneity of variances.

(e) 99% C.I. for σ^2:

$$\begin{aligned}[\frac{(N-k)\text{MSErr}}{\chi^2_{(N-k,\frac{\alpha}{2})}}, \frac{(N-k)\text{MSErr}}{\chi^2_{(N-k,1-\frac{\alpha}{2})}}] &= [\frac{27\times 3.45}{\chi^2_{(27,0.005)}}, \frac{27\times 3.45}{\chi^2_{(27,0.995)}}] \\ &= [93.15/49.64, 93.15/11.81] \\ &= [1.877, 7.887]\end{aligned}$$

4. (a)

```
> y=c(12.4,15.7,10.1,11.7,12.8,
+     14.5,11.4,18.9,16.3,13.8,
+     14.2,18.0,20.5,15.1,18.5,
+     22.1,17.4,20.6,19.0,19.1,
+     26.1,19.7,23.1,21.0,24.2,
+     22.8,27.1,26.2,19.8,25.3)
> n=rep(5,6)
> group=rep(1:6,n)
> data=data.frame(weight=y,group=factor(group))
> fit=lm(weight~group,data)
> anova(fit)

Analysis of Variance Table
Response: weight
          Df Sum Sq Mean Sq F value    Pr(>F)
group      5 511.60  102.32  16.559 4.248e-07 ***
Residuals 24 148.30    6.18
Total     29 659.90
```

$$F = 16.56$$

$$\text{P-value} = \Pr(F_{5,24} \geq 16.56) = 0.000$$

Very strong evidence against the claim that the amount of nitrogen does not affect the weight of corn.

(b) From the plot it appears that the relationship between weight and units of N follow a linear trend: weight increases as units of N increase.

(c) Test for hypotheses:
$H_0 : -5\mu_1 - 3\mu_2 - 1\mu_3 + 1\mu_4 + 3\mu_5 + 5\mu_6 = 0 \quad vs \quad H_A : -5\mu_1 - 3\mu_2 - 1\mu_3 + 1\mu_4 + 3\mu_5 + 5\mu_6 \neq 0$
$\lambda = (-5, -3, -1, 1, 3, 5)$
Test statistic

$$\begin{aligned} T &= \frac{\sum_{i=1}^{k} \lambda_i \bar{X}_i}{\sqrt{\text{MSErr}}\sqrt{\sum_{i=1}^{k} \frac{\lambda_i^2}{n_i}}} \\ &= \frac{84.4}{\sqrt{6.18} \times 3.741657} \\ &= 9.073696 \end{aligned}$$

$$\text{P-value} = 2\Pr(T_{24} \geq 9.073696) = 0.0000$$

Very strong evidence againt H_0, indicating the linear trend is present.

5. a.
```
> cocoon=read.table("http://www.stat.wisc.edu/~st571-1/data/cocoon.dat",
+                   col.names=c("race","length"))
> cocoon.lm=lm(length~race,cocoon)
> anova(cocoon.lm)

Analysis of Variance Table
Response: length
          Df  Sum Sq Mean Sq F value    Pr(>F)
race       6 288.571  48.095  10.569 4.327e-08 ***
Residuals 63 286.700   4.551
Total     69 575.271
```

$k = 7$, $n_i = n = 10$, $N = 70$
$H_0 : \mu_A = \mu_B = \mu_C = \mu_D = \mu_E = \mu_F = \mu_G$
H_A: not all the means are equal
With $F = \frac{MSTrt}{MSErr} = \frac{48.095}{4.551} = 10.569$, we get $P - value = P(F_{6,63} > 10.569) = .000$. There's very strong evidence that there's some difference among the seven races.

b. Assumptions: independent samples from seven normal population with equal variances.
To verify the normality, we can make stem-and-leaf plots or normal score plots for **each** group(seperately).

c. LSD method: $LSD = t_{N-k,\alpha/2} s_p \sqrt{\frac{2}{n}} = 1.9983(2.133)\sqrt{\frac{2}{10}} = 1.906$

```
sort(tapply(cocoon$length,cocoon$race,mean))
```

```
Race F   Race C  Race B  Race A  Race G  Race D  Race E
29.000   29.200  31.200  31.700  33.600  33.900  34.300
---------------
                 --------------
                                 ----------------------
```

Q method: $QD = Q_{k,N-k,\alpha} s_p \sqrt{\frac{1}{n}} = 4.31(2.133)\sqrt{\frac{1}{10}} = 2.91$

```
Race F   Race C  Race B  Race A  Race G  Race D  Race E
29.000   29.200  31.200  31.600  33.600  33.900  34.300
-------------------------------
                 ------------------------------
                         ------------------------------
```

d. Test $H_{01}: \mu_A = \mu_B$, $H_{02}: \mu_A = \mu_C$, $H_{03}: \mu_A = \mu_D$, $H_{04}: \mu_A = \mu_E$, $H_{05}: \mu_A = \mu_F$, $H_{06}: \mu_A = \mu_G$.

Using Bonferroni method with $\alpha = .05$, $r = 6$, we can compute
$t_{N-k,\frac{\alpha}{2r}} s_p \sqrt{\frac{2}{n}} = 2.7241(2.133)\sqrt{\frac{2}{10}} = 2.599.$

```
Race F   Race C  Race B  Race A  Race G  Race D  Race E
29.000   29.200  31.200  31.600  33.600  33.900  34.300
         ------------------||------------------
```

Only race E and F are significantly different from race A.