

Fry Analysis

Ben Straub

9/29/2017

FIRST WEEK

```
day1 <- ndata_list[[1]]
day1 <- day1[,-2]
day1 <- day1[,-11]
day1 <- day1[,-1]
pcamod1 <- princomp(day1, scores=TRUE)
day1$Gender <- as.factor(day1$Gender)
day1$Samp. <- as.factor(day1$Samp.)
day1$Age <- as.factor(day1$Age)
mod1 <- lm(Overall.Liking ~ ., data=day1)
summary(mod1) # gender, appearance, taste, texture are significant
```

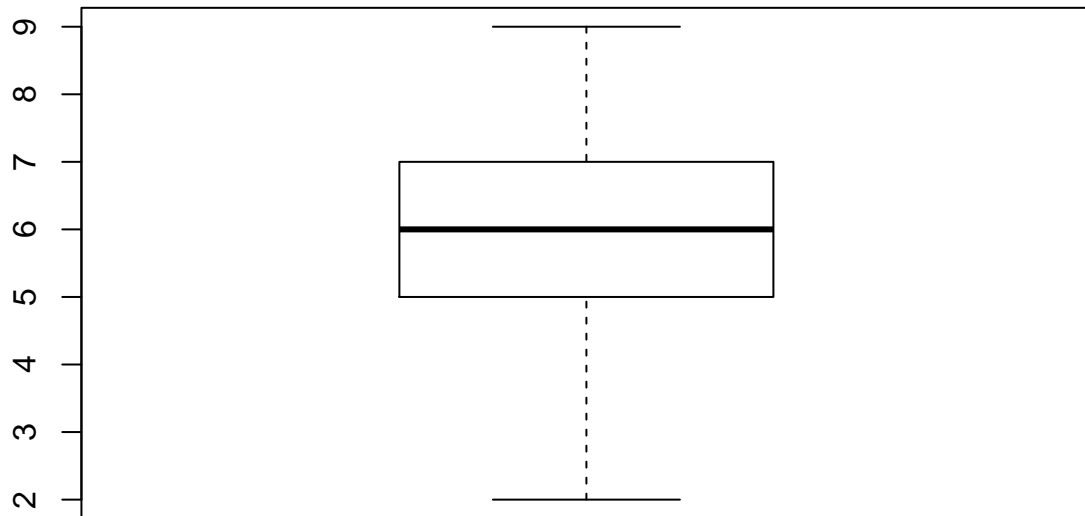
```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day1)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.10009 -0.27440 -0.00931  0.27294  1.83240
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.437111   0.488037  -0.896   0.3719
## Samp.2       0.069232   0.098734   0.701   0.4843
## Gender2      0.175815   0.104738   1.679   0.0954 .
## Age2         0.063818   0.142341   0.448   0.6546
## Age3        -0.003155   0.217315  -0.015   0.9884
## Age4         0.184396   0.149759   1.231   0.2202
## Age5         0.023961   0.154237   0.155   0.8768
## Temperature  0.101415   0.068880   1.472   0.1431
## Appearance   0.101873   0.047088   2.163   0.0322 *
## Color        0.001581   0.067726   0.023   0.9814
## Taste        0.635435   0.044169  14.387 < 2e-16 ***
## Texture      0.228276   0.046426   4.917 2.38e-06 ***
## Preference  -0.014497   0.064812  -0.224   0.8233
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6107 on 143 degrees of freedom
## Multiple R-squared:  0.846, Adjusted R-squared:  0.833
## F-statistic: 65.44 on 12 and 143 DF, p-value: < 2.2e-16
table(day1$Gender)# 58 male and 98 female

##
##  1  2
```

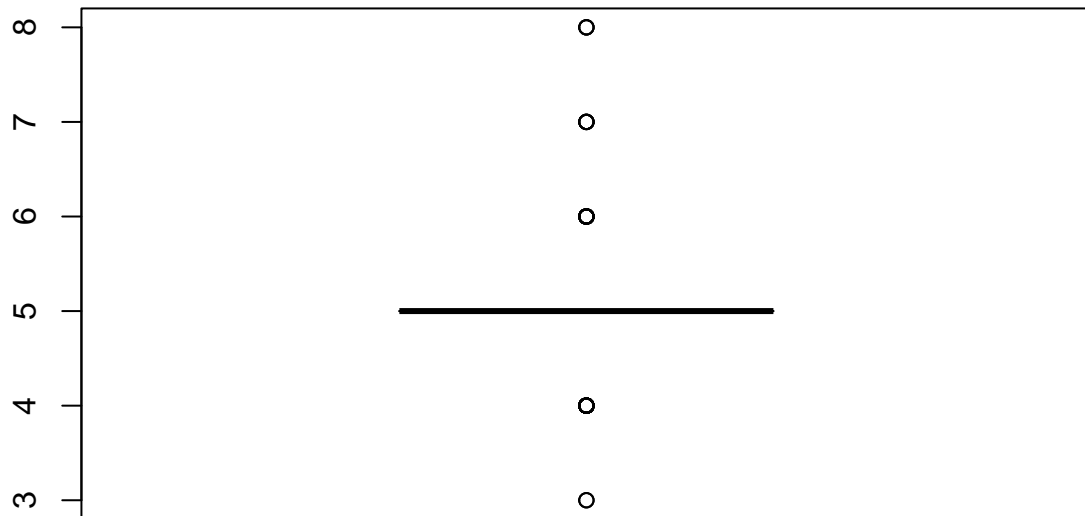
```
## 58 98
table(day1$Age) #1  2  3  4  5

##
##  1  2  3  4  5
## 42 38 10 34 32
#.....42 38 10 34 32
table(day1$Samp.)# Samples are balanced

##
##  1  2
## 78 78
boxplot(day1$Taste)
```



```
boxplot(day1$Temperature)
```



```
table(day1$Temperature)
```

```
##
##  3  4  5  6  7  8
##  2 22 110 14  5  3
```

```
day2 <- ndata_list[[2]]
day2 <- day2[,-2]
day2 <- day2[,-11]
day2 <- day2[,-1]
pcamod2 <- princomp(day2, scores=TRUE)
day2$Gender <- as.factor(day2$Gender)
day2$Samp. <- as.factor(day2$Samp.)
day2$Age <- as.factor(day2$Age)
mod2 <- lm(Overall.Liking~., data=day2)
summary(mod2) # appearance, taste, texture are significant, age4 significant
```

```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day2)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.8029 -0.2159  0.0093  0.2372  1.3582
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept) -0.05992    0.44906   -0.133    0.8940
## Samp.2      -0.01004    0.09866   -0.102    0.9191
## Gender2      0.03963    0.10088    0.393    0.6950
## Age2         0.17485    0.14705    1.189    0.2365
## Age3        -0.04807    0.17341   -0.277    0.7820
## Age4         0.26597    0.15581    1.707    0.0901 .
## Age5         0.10704    0.16770    0.638    0.5244
## Temperature -0.09827    0.06715   -1.463    0.1457
## Appearance   0.12327    0.05091    2.421    0.0168 *
## Color        0.05963    0.06653    0.896    0.3717
## Taste        0.61559    0.05285   11.648   < 2e-16 ***
## Texture      0.26202    0.05230    5.010   1.68e-06 ***
## Preference   0.06811    0.07652    0.890    0.3750
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5748 on 135 degrees of freedom
## Multiple R-squared:  0.862, Adjusted R-squared:  0.8498
## F-statistic: 70.29 on 12 and 135 DF,  p-value: < 2.2e-16
```

```
table(day2$Gender)# 74 males and 74 females
```

```
##
##  1  2
## 74 74
```

```
table(day2$Age)# 1  2  3  4  5
```

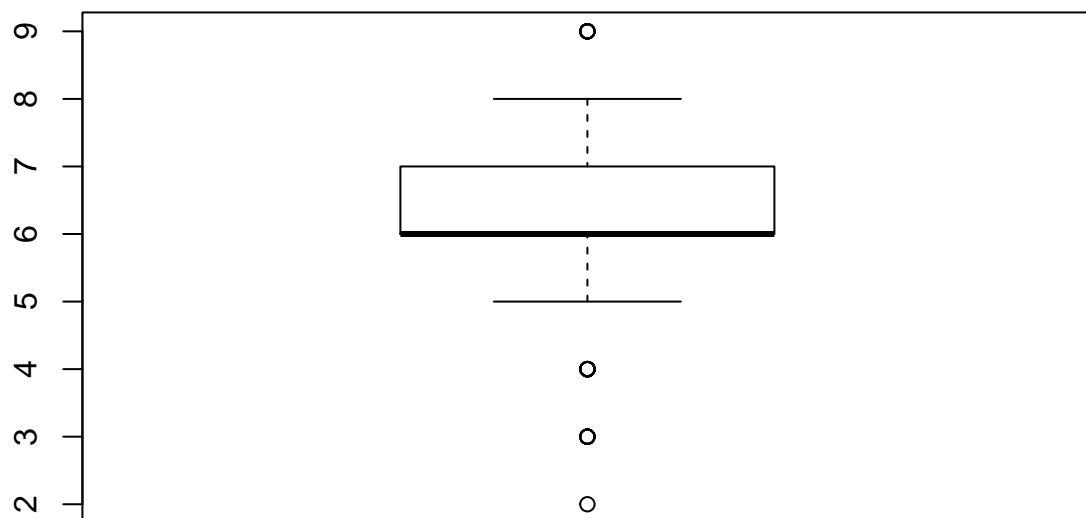
```
##
##  1  2  3  4  5
## 28 38 20 34 28
```

```
#.....28 38 20 34 28
```

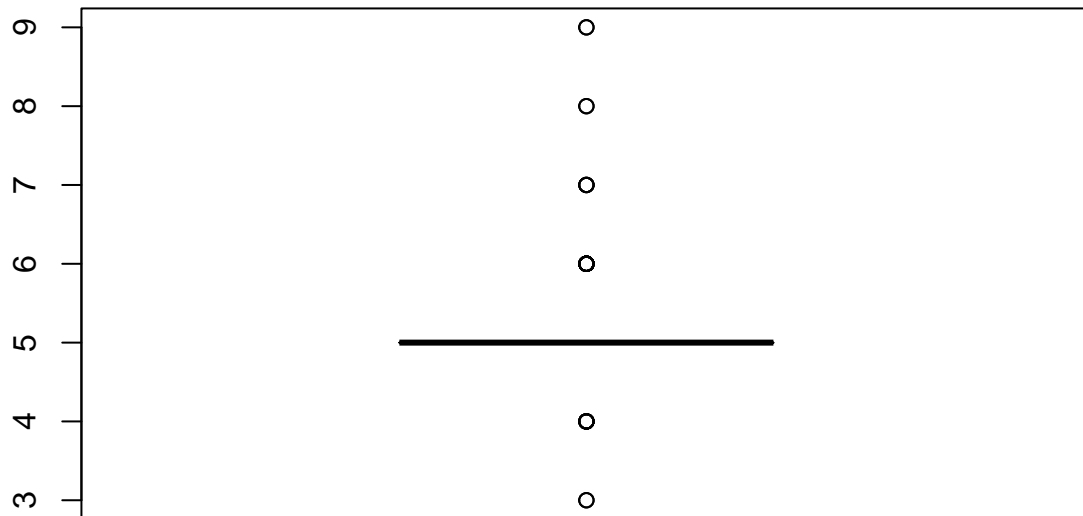
```
table(day2$Samp.)# Samples are balanced
```

```
##
##  1  2
## 74 74
```

```
boxplot(day2$Taste)
```



```
boxplot(day2$Temperature)
```



```

day3 <- ndata_list[[3]]
day3 <- day3[,-2]
day3 <- day3[,-11]
day3 <- day3[,-1]
pcamod3 <- princomp(day3, scores=TRUE)
day3$Gender <- as.factor(day3$Gender)
day3$Samp. <- as.factor(day3$Samp.)
day3$Age <- as.factor(day3$Age)
mod3 <- lm(Overall.Liking~.,data=day3) # appearance, taste, texture are significant
summary(mod3)

```

```

##
## Call:
## lm(formula = Overall.Liking ~ ., data = day3)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -2.15399 -0.31903 -0.01505  0.28073  1.87973
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.65753    0.53617   1.226  0.2222
## Samp.2      -0.07461    0.10827  -0.689  0.4920
## Gender2     -0.01099    0.11682  -0.094  0.9252
## Age2        0.11163    0.16615   0.672  0.5028
## Age3        0.00422    0.18605   0.023  0.9819

```

```
## Age4      0.02949    0.18184    0.162    0.8714
## Age5      0.15753    0.18263    0.863    0.3899
## Temperature 0.04573    0.08365    0.547    0.5855
## Appearance 0.10932    0.05561    1.966    0.0514 .
## Color      -0.05665    0.07366   -0.769    0.4432
## Taste      0.56275    0.05551   10.137   < 2e-16 ***
## Texture     0.25017    0.05675    4.408   2.13e-05 ***
## Preference -0.05625    0.07776   -0.723    0.4707
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6503 on 133 degrees of freedom
## Multiple R-squared:  0.8294, Adjusted R-squared:  0.814
## F-statistic: 53.89 on 12 and 133 DF,  p-value: < 2.2e-16
```

```
table(day3$Gender)# 62 males and 84 females
```

```
##
##  1  2
## 62 84
```

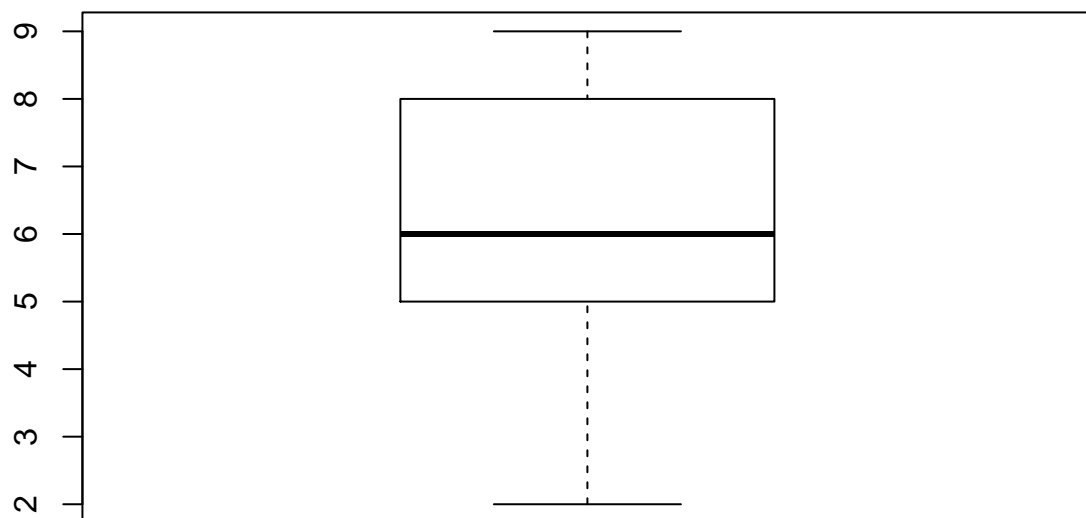
```
table(day3$Age)# 1  2  3  4  5
```

```
##
##  1  2  3  4  5
## 34 36 20 26 30
```

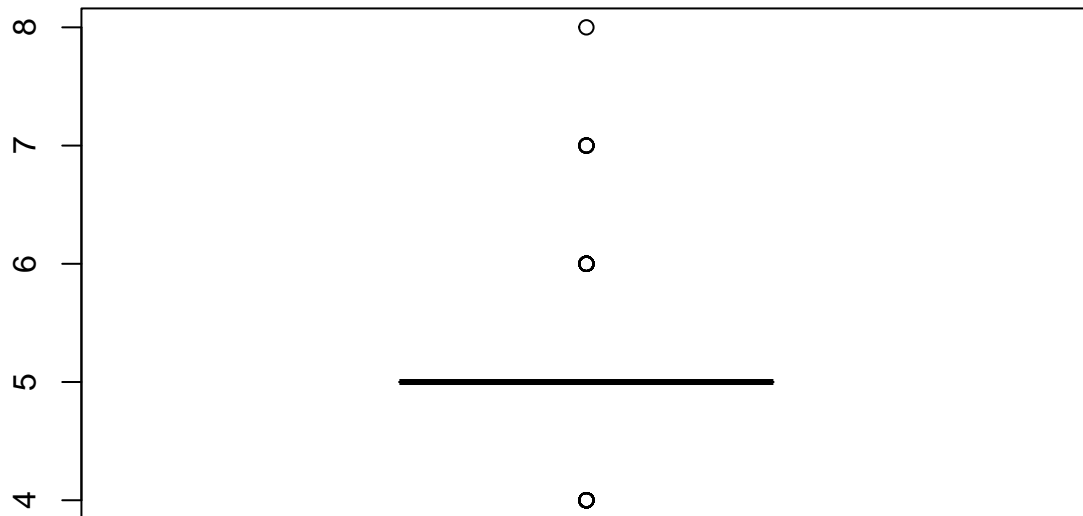
```
#.....34 36 20 26 30
table(day3$Samp.)# Samples are balanced
```

```
##
##  1  2
## 73 73
```

```
boxplot(day3$Taste)
```



```
boxplot(day3$Temperature)
```

```
day4 <- ndata_list[[4]]
day4 <- day4[,-2]
day4 <- day4[,-11]
day4 <- day4[,-1]
pcamod4 <- princomp(day4, scores=TRUE)
day4$Gender <- as.factor(day4$Gender)
day4$Samp. <- as.factor(day4$Samp.)
day4$Age <- as.factor(day4$Age)
mod4 <- lm(Overall.Liking~.,data=day4)
summary(mod4)#Appearance, Taste and Texture are significant
```

```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day4)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-2.17232	-0.29983	-0.02305	0.20418	1.60320

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	0.335160	0.446029	0.751	0.4534
Samp.2	-0.046808	0.086540	-0.541	0.5893
Gender2	-0.013600	0.093268	-0.146	0.8842
Age2	-0.132393	0.115751	-1.144	0.2543
Age3	0.168945	0.233869	0.722	0.4710

```
## Age4      0.141637  0.134064  1.056  0.2922
## Age5      -0.095416  0.146230 -0.653  0.5149
## Temperature 0.024162  0.057150  0.423  0.6730
## Appearance 0.099329  0.044512  2.231  0.0269 *
## Color      0.008583  0.051573  0.166  0.8680
## Taste      0.631681  0.045711 13.819 < 2e-16 ***
## Texture    0.227797  0.039889  5.711 4.67e-08 ***
## Preference -0.044654  0.059413 -0.752  0.4533
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5926 on 177 degrees of freedom
## Multiple R-squared:  0.8194, Adjusted R-squared:  0.8072
## F-statistic: 66.94 on 12 and 177 DF,  p-value: < 2.2e-16
```

```
table(day4$Gender)# 94 males and 96 females
```

```
##
##  1  2
## 94 96
```

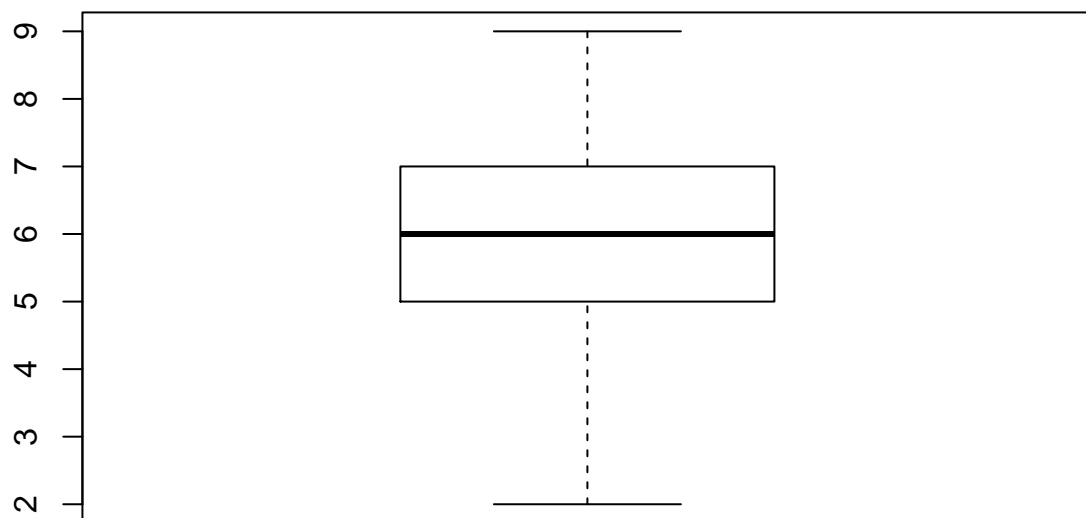
```
table(day4$Age)# 1  2  3  4  5
```

```
##
##  1  2  3  4  5
## 58 52  8 38 34
```

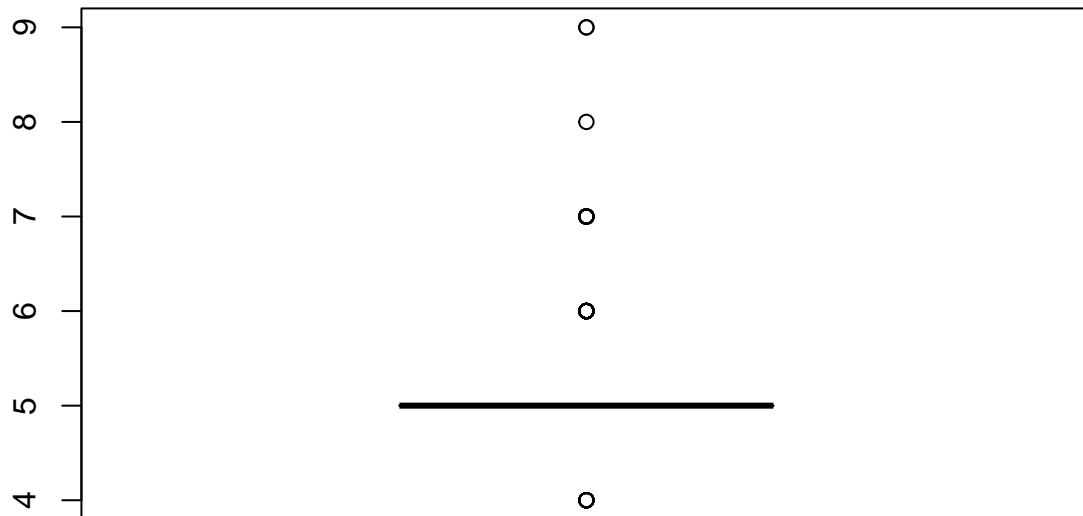
```
#.....58 52  8 38 34
table(day4$Samp.)# Samples are balanced
```

```
##
##  1  2
## 95 95
```

```
boxplot(day4$Taste)
```



```
boxplot(day4$Temperature)
```



```
day5 <- ndata_list[[5]]
day5 <- day5[,-2]
day5 <- day5[,-11]
day5 <- day5[,-1]
day5$Gender <- as.factor(day5$Gender)
day5$Samp. <- as.factor(day5$Samp.)
day5$Age <- as.factor(day5$Age)
mod5 <- lm(Overall.Liking~.,data=day5)
summary(mod5)#Female, Apperance, Taste and Texture are significant
```

```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day5)
##
## Residuals:
```

	Min	1Q	Median	3Q	Max
	-2.52025	-0.36208	-0.02896	0.34658	2.86273

```
##
## Coefficients:
```

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.26870	0.48471	-0.554	0.579994
Samp.2	0.10297	0.09272	1.111	0.268164
Gender2	-0.19624	0.10252	-1.914	0.057136 .
Age2	-0.04151	0.12117	-0.343	0.732295
Age3	-0.16618	0.18792	-0.884	0.377650
Age4	-0.11073	0.16566	-0.668	0.504712

```
## Age5          -0.13573    0.15663   -0.867  0.387277
## Temperature  0.03599    0.05863    0.614  0.540001
## Appearance   0.15595    0.04400    3.545  0.000497 ***
## Color        0.10709    0.07185    1.491  0.137771
## Taste        0.54466    0.04784   11.385  < 2e-16 ***
## Texture      0.26614    0.04219    6.307   2e-09 ***
## Preference   -0.03769    0.06230   -0.605  0.545994
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6407 on 187 degrees of freedom
## Multiple R-squared:  0.7805, Adjusted R-squared:  0.7664
## F-statistic: 55.41 on 12 and 187 DF,  p-value: < 2.2e-16
```

```
table(day5$Gender)# 95 males and 120 females
```

```
##
##    1    2
##  80 120
```

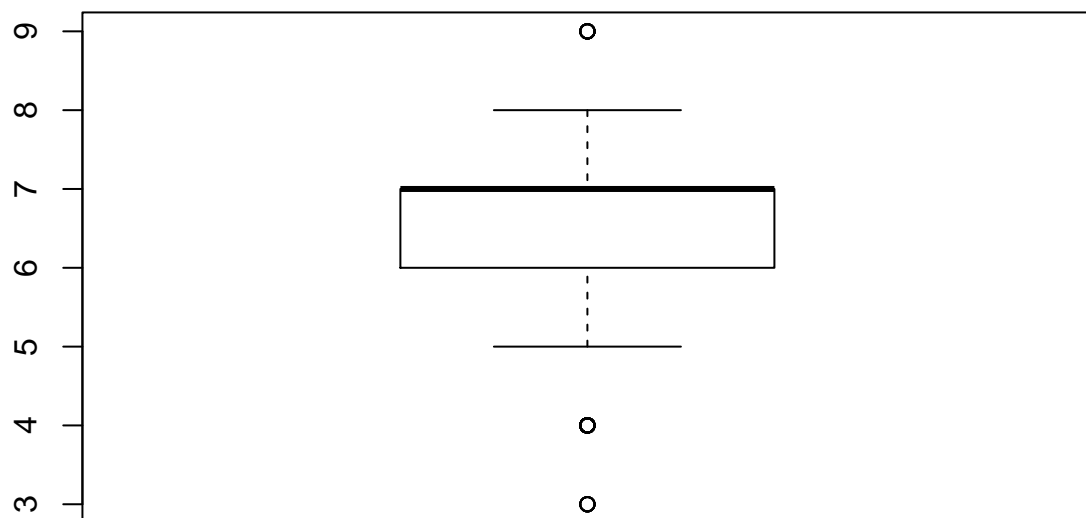
```
table(day5$Age)# 1  2  3  4  5
```

```
##
##    1  2  3  4  5
## 94 46 14 20 26
```

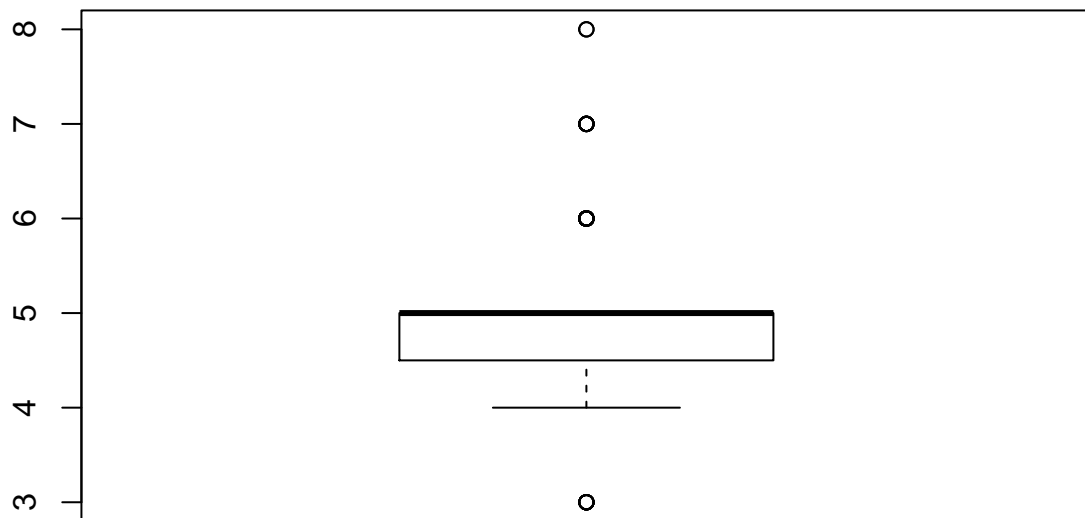
```
#.....94 46 14 20 26
table(day5$Samp.)# Samples are balanced
```

```
##
##    1    2
## 100 100
```

```
boxplot(day5$Taste)
```



```
boxplot(day5$Temperature)
```



```
histosforall <- function(x){
  dayx <- ndata_list[[x]]
  dayx <- dayx[,-2]
  dayx <- dayx[,-11]
  jpeg(paste0("Day",x,".jpg"))
  multi.hist(dayx, main = paste("Day", x, sep=""))
}
```

```
histosforall(1)
#\includegraphics[width=450pt]{Day1.jpg}
```

```
histosforall(2)
#\includegraphics[width=450pt]{Day2.jpg}
```

```
histosforall(3)
#\includegraphics[width=450pt]{Day3.jpg}
```

```
histosforall(4)
#\includegraphics[width=450pt]{Day4.jpg}
```

```
histosforall(5)
#\includegraphics[width=450pt]{Day5.jpg}
```

SECOND WEEK

```
data_list <- list(Day8, Day7, Day8, Day9, Day10)

#ndata_list <- lapply(data_list, `[`, -c(7:23))
ndata_list <- lapply(data_list, `[`, -c(4:5))
ndata_list <- lapply(ndata_list, `[`, -1)

day6 <- ndata_list[[1]]
  day6 <- day6[, -2]
  day6 <- day6[, -15]
  day6 <- day6[, -4]
  day6 <- day6[, -3]

day7 <- ndata_list[[2]]
  day7 <- day7[, -2]
  day7 <- day7[, -15]
  day7 <- day7[, -4]
  day7 <- day7[, -3]

day8 <- ndata_list[[3]]
  day8 <- day8[, -2]
  day8 <- day8[, -15]
  day8 <- day8[, -4]
  day8 <- day8[, -3]

day9 <- ndata_list[[4]]
  day9 <- day9[, -2]
  day9 <- day9[, -15]
  day9 <- day9[, -4]
  day9 <- day9[, -3]

day10 <- ndata_list[[5]]
  day10 <- day10[, -2]
  day10 <- day10[, -15]
  day10 <- day10[, -4]
  day10 <- day10[, -3]

mod6 <- lm(Overall.Liking~., data=day6)
summary(mod6)

##
## Call:
## lm(formula = Overall.Liking ~ ., data = day6)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73996 -0.27917 -0.02974  0.29164  2.27068
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.480854   0.473983   1.014   0.312
## Samp.Set    -0.001396   0.001564  -0.893   0.373
```



```
## Samp.      0.042873  0.083956  0.511  0.610
## Samp.BC    NA      NA      NA      NA
## Samp.Pos   0.033078  0.085540  0.387  0.699
## Gender     -0.133386  0.088102 -1.514  0.132
## Age        -0.057096  0.037532 -1.521  0.130
## Temperature -0.012379  0.061759 -0.200  0.841
## Appearance  0.230448  0.039980  5.764 3.40e-08 ***
## Color      -0.069681  0.058563 -1.190  0.236
## Taste       0.499300  0.048590 10.276 < 2e-16 ***
## Texture     0.320806  0.045572  7.039 3.71e-11 ***
## Preference -0.047551  0.063734 -0.746  0.457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5841 on 184 degrees of freedom
## Multiple R-squared:  0.8487, Adjusted R-squared:  0.8396
## F-statistic: 93.82 on 11 and 184 DF,  p-value: < 2.2e-16
```

```
mod7 <- lm(Overall.Liking~.,data=day7)
summary(mod7)
```

```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day7)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.83659 -0.20384 -0.01142  0.18156  2.47543
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.311e-01  4.024e-01   1.320  0.18817
## Samp.Set     4.583e-05  1.056e-03   0.043  0.96541
## Samp.        3.062e-02  7.346e-02   0.417  0.67715
## Samp.BC      NA      NA      NA      NA
## Samp.Pos     2.903e-03  7.321e-02   0.040  0.96840
## Gender       8.288e-02  7.490e-02   1.106  0.26965
## Age          6.073e-03  2.698e-02   0.225  0.82214
## Temperature -7.248e-02  4.555e-02  -1.591  0.11290
## Appearance   1.233e-01  3.908e-02   3.156  0.00181 **
## Color        -1.841e-02  5.281e-02  -0.349  0.72776
## Taste        5.191e-01  3.979e-02  13.045 < 2e-16 ***
## Texture      3.297e-01  3.717e-02   8.868 < 2e-16 ***
## Preference  -2.823e-02  5.544e-02  -0.509  0.61105
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.571 on 236 degrees of freedom
## Multiple R-squared:  0.8194, Adjusted R-squared:  0.811
## F-statistic: 97.36 on 11 and 236 DF,  p-value: < 2.2e-16
```

```
mod8 <- lm(Overall.Liking~.,data=day8)
summary(mod8)
```

```
##
```

```
## Call:
## lm(formula = Overall.Liking ~ ., data = day8)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.73996 -0.27917 -0.02974  0.29164  2.27068
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.480854   0.473983   1.014   0.312
## Samp.Set     -0.001396   0.001564  -0.893   0.373
## Samp.        0.042873   0.083956   0.511   0.610
## Samp.BC      NA         NA         NA      NA
## Samp.Pos     0.033078   0.085540   0.387   0.699
## Gender       -0.133386   0.088102  -1.514   0.132
## Age          -0.057096   0.037532  -1.521   0.130
## Temperature -0.012379   0.061759  -0.200   0.841
## Appearance   0.230448   0.039980   5.764 3.40e-08 ***
## Color        -0.069681   0.058563  -1.190   0.236
## Taste        0.499300   0.048590  10.276 < 2e-16 ***
## Texture      0.320806   0.045572   7.039 3.71e-11 ***
## Preference   -0.047551   0.063734  -0.746   0.457
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.5841 on 184 degrees of freedom
## Multiple R-squared:  0.8487, Adjusted R-squared:  0.8396
## F-statistic: 93.82 on 11 and 184 DF, p-value: < 2.2e-16
```

```
mod9 <- lm(Overall.Liking~.,data=day9)
summary(mod9)
```

```
##
## Call:
## lm(formula = Overall.Liking ~ ., data = day9)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.81989 -0.35002 -0.02726  0.25392  1.98503
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0091217  0.4432896   0.021   0.9836
## Samp.Set     -0.0002735  0.0014650  -0.187   0.8521
## Samp.        -0.0355358  0.0831155  -0.428   0.6695
## Samp.BC      NA         NA         NA      NA
## Samp.Pos     -0.1513898  0.0819780  -1.847   0.0663 .
## Gender       0.0678411  0.0829857   0.818   0.4147
## Age          -0.0162401  0.0322037  -0.504   0.6146
## Temperature  0.0582639  0.0540910   1.077   0.2828
## Appearance   0.2021070  0.0394986   5.117 7.51e-07 ***
## Color        -0.0366631  0.0550520  -0.666   0.5062
## Taste        0.5375433  0.0423714  12.686 < 2e-16 ***
## Texture      0.2888400  0.0395333   7.306 7.17e-12 ***
## Preference   -0.0199004  0.0567695  -0.351   0.7263
```

```

## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.582 on 192 degrees of freedom
## Multiple R-squared:  0.8639, Adjusted R-squared:  0.8561
## F-statistic: 110.8 on 11 and 192 DF,  p-value: < 2.2e-16

mod10 <- lm(Overall.Liking~.,data=day10)
summary(mod10)

##
## Call:
## lm(formula = Overall.Liking ~ ., data = day10)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -1.65768 -0.24985 -0.00662  0.28493  2.41453
##
## Coefficients: (1 not defined because of singularities)
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.759828   0.444116  -1.711  0.08891 .
## Samp.Set     0.004482   0.001765   2.540  0.01197 *
## Samp.       -0.048168   0.089433  -0.539  0.59086
## Samp.BC      NA         NA        NA      NA
## Samp.Pos     0.001216   0.090980   0.013  0.98936
## Gender       -0.035203   0.095868  -0.367  0.71392
## Age          -0.007504   0.035358  -0.212  0.83217
## Temperature  0.070506   0.050382   1.399  0.16349
## Appearance   0.113380   0.042824   2.648  0.00886 **
## Color        0.054906   0.064031   0.857  0.39237
## Taste        0.576026   0.045429  12.680 < 2e-16 ***
## Texture      0.332094   0.042498   7.814 5.25e-13 ***
## Preference  -0.070937   0.059205  -1.198  0.23250
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.6012 on 172 degrees of freedom
## Multiple R-squared:  0.8446, Adjusted R-squared:  0.8346
## F-statistic: 84.97 on 11 and 172 DF,  p-value: < 2.2e-16

#View(ndata_list[1])
histosforall <- function(x){
  dayx <- ndata_list[[x]]
  dayx <- dayx[,-2]
  dayx <- dayx[,-15]
  dayx <- dayx[,-4]
  dayx <- dayx[,-3]
  jpeg(paste0("Day",x,".jpg"))
  multi.hist(dayx, main = paste("Day", x, sep=""))
}

histosforall(1)
#\\includegraphics[width=450pt]{Day1.jpg}

histosforall(2)
#\\includegraphics[width=450pt]{Day2.jpg}

```

```
histosforall(3)
#\includegraphics[width=450pt]{Day3.jpg}
```

```
histosforall(4)
#\includegraphics[width=450pt]{Day4.jpg}
```

```
histosforall(5)
#\includegraphics[width=450pt]{Day5.jpg}
```

```
#### Only done for Day 1
# paste cells into one string, use ";" as separator
comments.string <- paste(Day1$Comments , collapse = " " )
# split string at ";"
comments.vector <- strsplit(comments.string , " " )[[1]]
# get rid of white space to prevent errors
comments.vector.clean <- gsub ( " " , "" , comments.vector )
# tabulate data
sort(table(comments.vector.clean),decreasing = TRUE)
```