MATH 324 Homework 9

Mehmet Arslan

4/19/2021

Problem 1

```
Aluminum = c(75, 77, 76, 79, 74, 77, 75)
Stainless_Steel = c(74, 76, 75, 78, 74, 77, 75, 77)
Alloy_1 = c(73, 74, 72, 74, 70, 73, 74, 71)
Alloy_2 = c(71, 74, 74, 73, 74, 73, 71)

na = length(Aluminum)
nb = length(Stainless_Steel)
nc = length(Alloy_1)
nd = length(Alloy_2)

corrosion_data = c(Aluminum, Stainless_Steel, Alloy_1, Alloy_2)

corrosion_rates = factor(rep((c("A","B","C","D")), (c(na,nb,nc,nd))))

corrosion_dat = data.frame(corrosion_data, corrosion_rates)

corrosion_dat
```

```
##    corrosion_data corrosion_rates
## 1              75               A
## 2              77               A
## 3              76               A
## 4              79               A
## 5              74               A
## 6              77               A
## 7              75               A
## 8              74               B
## 9              76               B
## 10             75               B
## 11             78               B
## 12             74               B
## 13             77               B
## 14             75               B
## 15             77               B
## 16             73               C
## 17             74               C
## 18             72               C
## 19             74               C
## 20             70               C
## 21             73               C
## 22             74               C
```

```
## 23             71              C
## 24             71              D
## 25             74              D
## 26             74              D
## 27             73              D
## 28             74              D
## 29             73              D
## 30             71              D
```

Problem 1 Part A

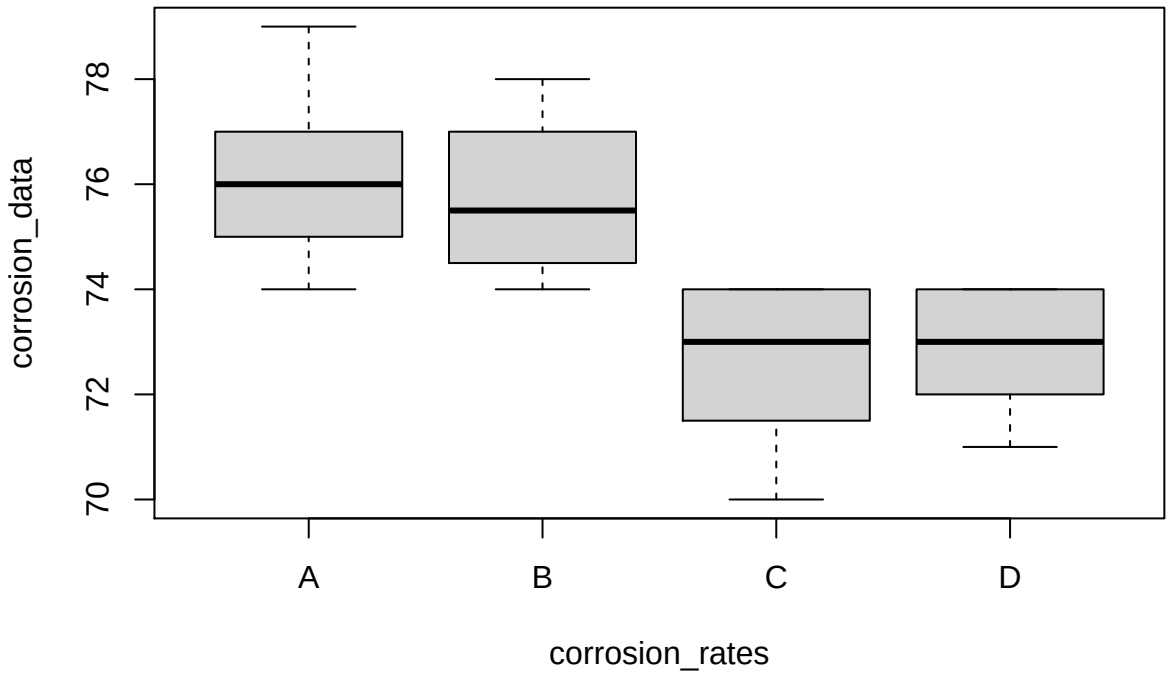
```
means = tapply(corrosion_data, corrosion_rates, mean)

variances = tapply(corrosion_data, corrosion_rates, var)

sampsize = tapply(corrosion_data, corrosion_rates, length)

round(rbind(means, variances, sampsize), 2)
```

```
##                A     B     C     D
## means      76.14 75.75 72.62 72.86
## variances   2.81  2.21  2.27  1.81
## sampsize    7.00  8.00  8.00  7.00
```

```
grandmean = mean(corrosion_data)
```

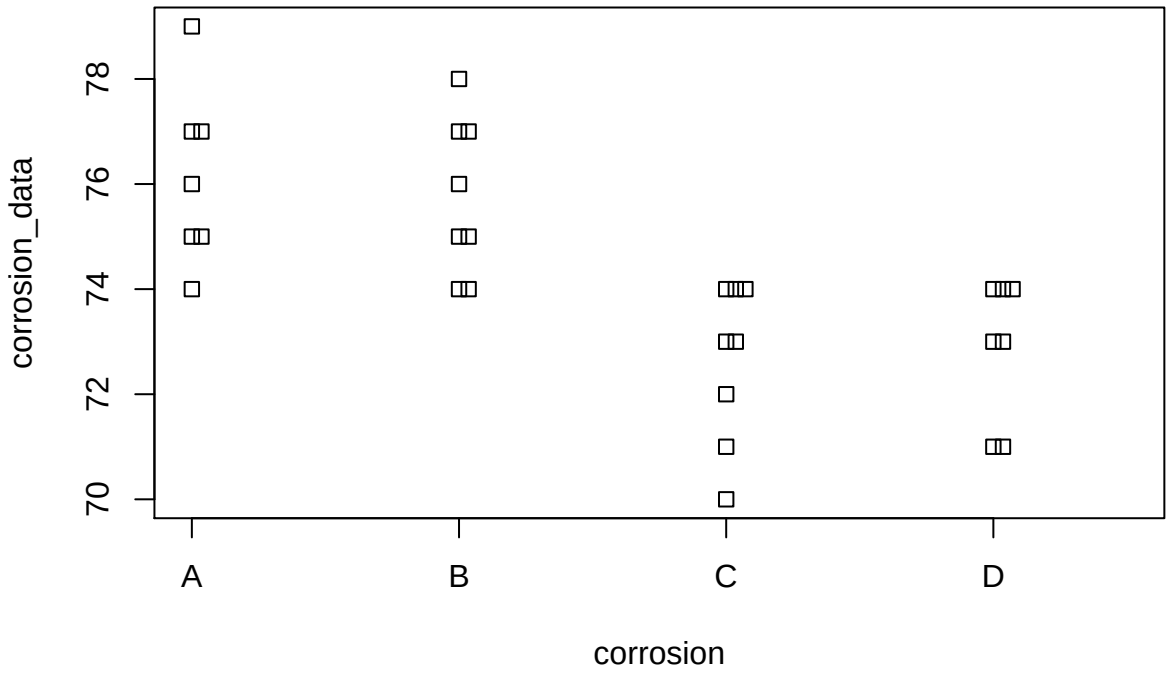
Problem 1 Part B

```
boxplot(corrosion_data ~ corrosion_rates, data = corrosion_dat)
```



Problem 1 Part C

```
stripchart(corrosion_data ~ corrosion_rates, data = corrosion_dat, vertical = T,
           method = "stack", xlab = "corrosion")
```



Problem 1 Part D

```
anov1 = lm(corrosion_data ~ corrosion_rates)
anova(anov1)
```

```
## Analysis of Variance Table
##
## Response: corrosion_data
##                 Df Sum Sq Mean Sq F value    Pr(>F)
## corrosion_rates  3 77.577 25.8591  11.378 5.971e-05 ***
## Residuals       26 59.089  2.2727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Problem 1 Part E

```
(N = length(corrosion_data))
```

```
## [1] 30
```

```
(grandmean = mean(corrosion_data))
```

```
## [1] 74.33333
```

```r
(TSS = (N-1)*var(corrosion_data))
```

```
## [1] 136.6667
```

```r
anov2 = lm(corrosion_data ~ corrosion_rates)
anova(anov2)
```

```
## Analysis of Variance Table
##
## Response: corrosion_data
##                 Df Sum Sq Mean Sq F value    Pr(>F)
## corrosion_rates  3 77.577 25.8591  11.378 5.971e-05 ***
## Residuals       26 59.089  2.2727
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
t_val = qt(1.99/2, df = 26)
round(t_val, 3)
```

```
## [1] 2.779
```

```r
qtukey(0.99, 4, 26)
```

```
## [1] 4.865002
```

```r
qtukey(0.99, 4, 26)/sqrt(2)
```

```
## [1] 3.440076
```

```r
TukeyHSD(aov(anov2), conf.level = 0.99)
```

```
##   Tukey multiple comparisons of means
##     99% family-wise confidence level
##
## Fit: aov(formula = anov2)
##
## $corrosion_rates
##           diff       lwr        upr     p adj
## B-A -0.3928571 -3.076887  2.2911726 0.9575153
## C-A -3.5178571 -6.201887 -0.8338274 0.0006706
## D-A -3.2857143 -6.057768 -0.5136602 0.0020358
## C-B -3.1250000 -5.718019 -0.5319809 0.0017095
## D-B -2.8928571 -5.576887 -0.2088274 0.0051765
## D-C  0.2321429 -2.451887  2.9161726 0.9906214
```

Problem 2

```r
y_means = c(10.5, 8.4, 11.6)
y_vars = c(4.6, 3.5, 5.5)
ybar_grand = 10.2
n_i = c(8, 6, 6)
N_1 = sum(n_i)
I = length(y_means)
N_1
```

```
## [1] 20
```

```r
I
```

```
## [1] 3
```

```r
mean(y_means) # this was to confirm the means for the y-values.
```

```
## [1] 10.16667
```

```r
SSG = sum(n_i*(y_means - ybar_grand)^2)
SSG
```

```
## [1] 31.92
```

```r
RSS = sum((n_i-1)*y_vars)
RSS
```

```
## [1] 77.2
```

```r
(TSS = SSG + RSS)
```

```
## [1] 109.12
```

```r
(MS = 31.92/2)
```

```
## [1] 15.96
```

```r
(S2 = 77.2/17)
```

```
## [1] 4.541176
```

```r
(Fstat = 15.96/4.54)
```

```
## [1] 3.515419
```

```r
(p_value = 1-pf(3.52, df1 = 2, df2 = 19))
```

```
## [1] 0.05006911
```