

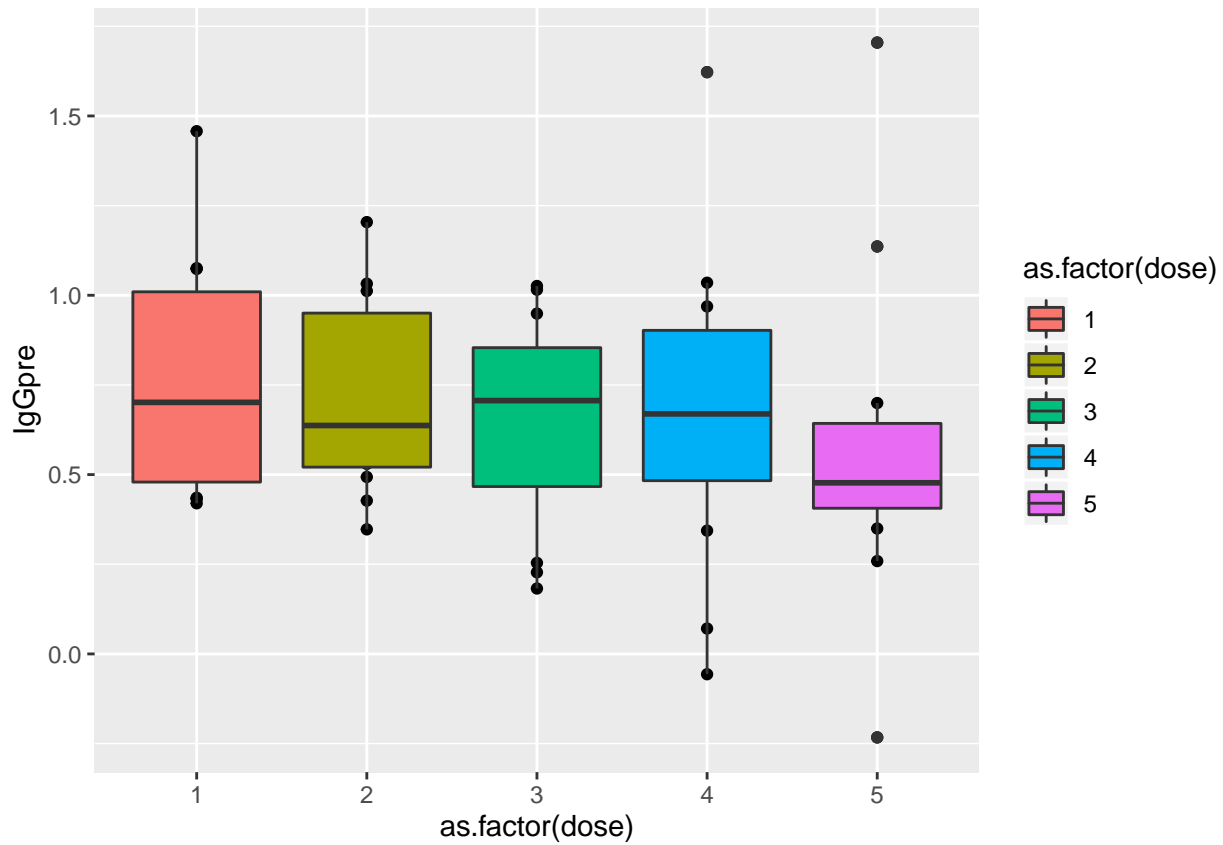
# Vaccine\_Project

Patrick Poleshuk

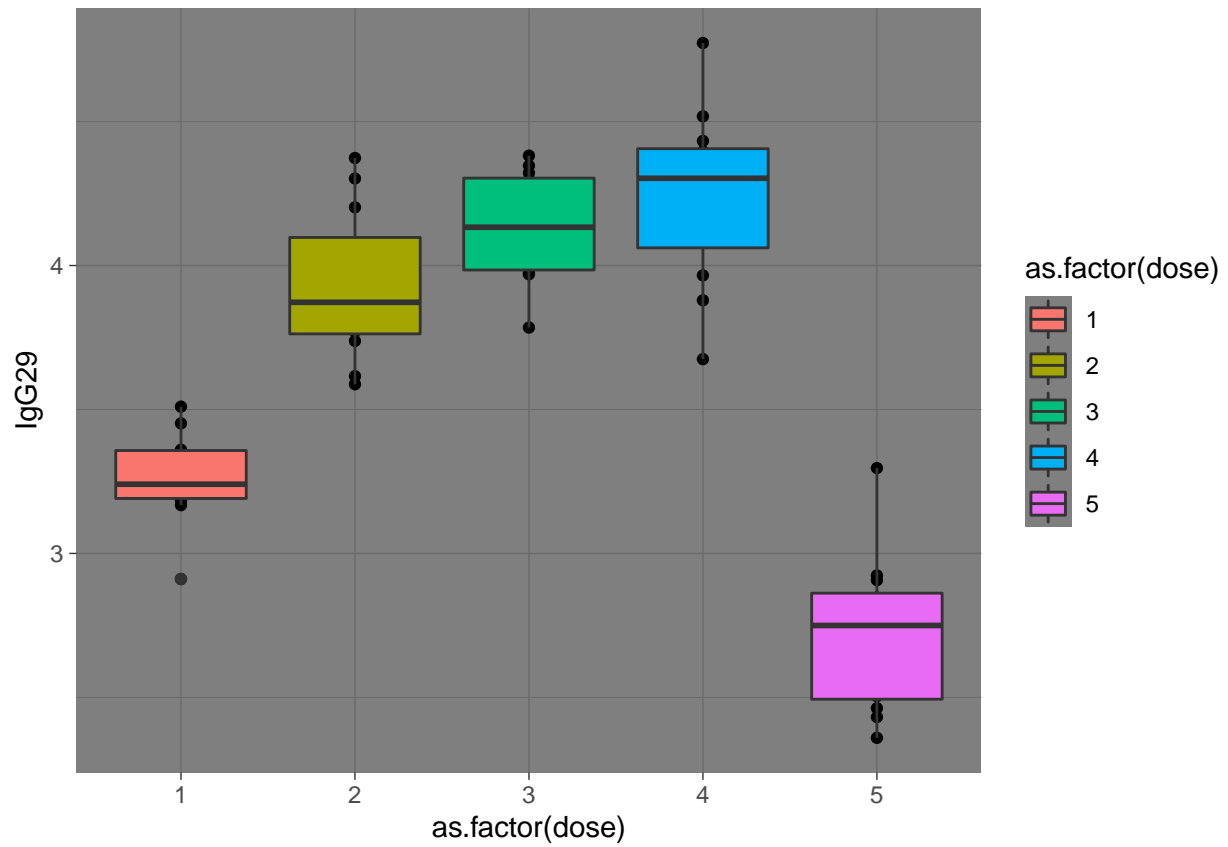
3/2/2021

```
library(ggplot2)
immune_test = read.csv("/Users/patrickpoleshuk/Downloads/immuneResponse.csv", sep = ",", header = T)

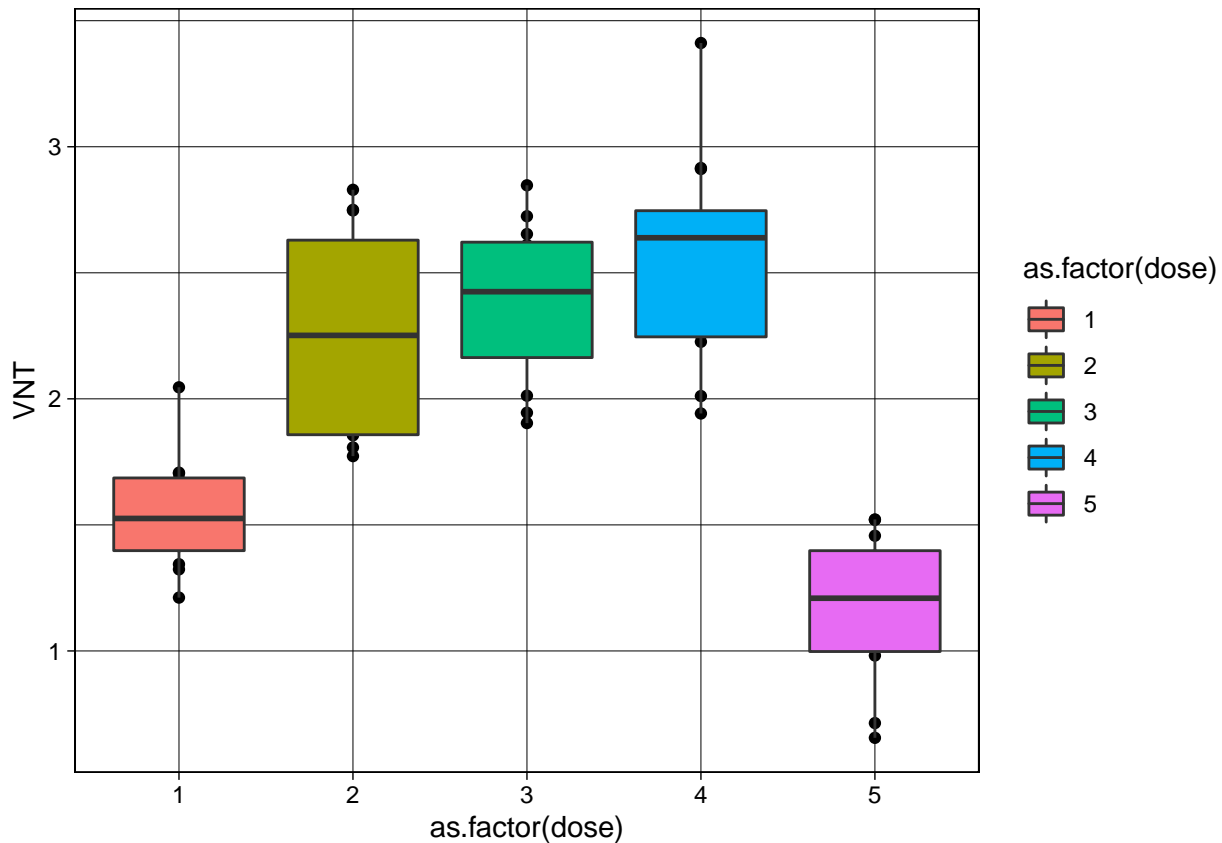
ggplot(data = immune_test, aes(x = as.factor(dose), y = IgGpre, fill = as.factor(dose))) +
  geom_point() + geom_boxplot()
```



```
ggplot(data = immune_test, aes(x = as.factor(dose), y = IgG29, fill = as.factor(dose))) +
  geom_point() + geom_boxplot() + theme_dark()
```



```
ggplot(data = immune_test, aes(x = as.factor(dose), y = VNT, fill = as.factor(dose))) +  
  geom_point() + geom_boxplot() + theme_linedraw()
```



```
var = c("IgGpre", "IgG29", "VNT")
outlist = c()
fn_list <- function(x){
  outlist = c(outlist, mean(x), sd(x), IQR(x), median(x))
}
l = tapply(immune_test[, var[1]], immune_test[, 4], fn_list)
IgGPRE_df = as.data.frame(l)

for (i in 1:NROW(IgGPRE_df)){
  rownames(IgGPRE_df)[i] <- paste0("Dose Group: ", i)
}
colnames(IgGPRE_df) <- c("MEAN", "SD", "IQR", "MEDIAN")
IgGPRE_df
```

```
##          MEAN          SD          IQR          MEDIAN
## Dose Group: 1 0.7650815, 0.3308874, 0.5304835, 0.7014505
## Dose Group: 2 0.7079494, 0.2729822, 0.4294729, 0.6368983
## Dose Group: 3 0.6543556, 0.2984409, 0.3871994, 0.7063217
## Dose Group: 4 0.6692356, 0.4479995, 0.4191508, 0.6692866
## Dose Group: 5 0.5712575, 0.4746484, 0.2365727, 0.4771336
```

```
s = tapply(immune_test[, var[2]], immune_test[, 4], fn_list)
IgG29_df = as.data.frame(s)

for (i in 1:NROW(IgG29_df)){
  rownames(IgG29_df)[i] <- paste0("Dose Group: ", i)
}
colnames(IgG29_df) <- c("MEAN", "SD", "IQR", "MEDIAN")
```

```
IgG29_df
```

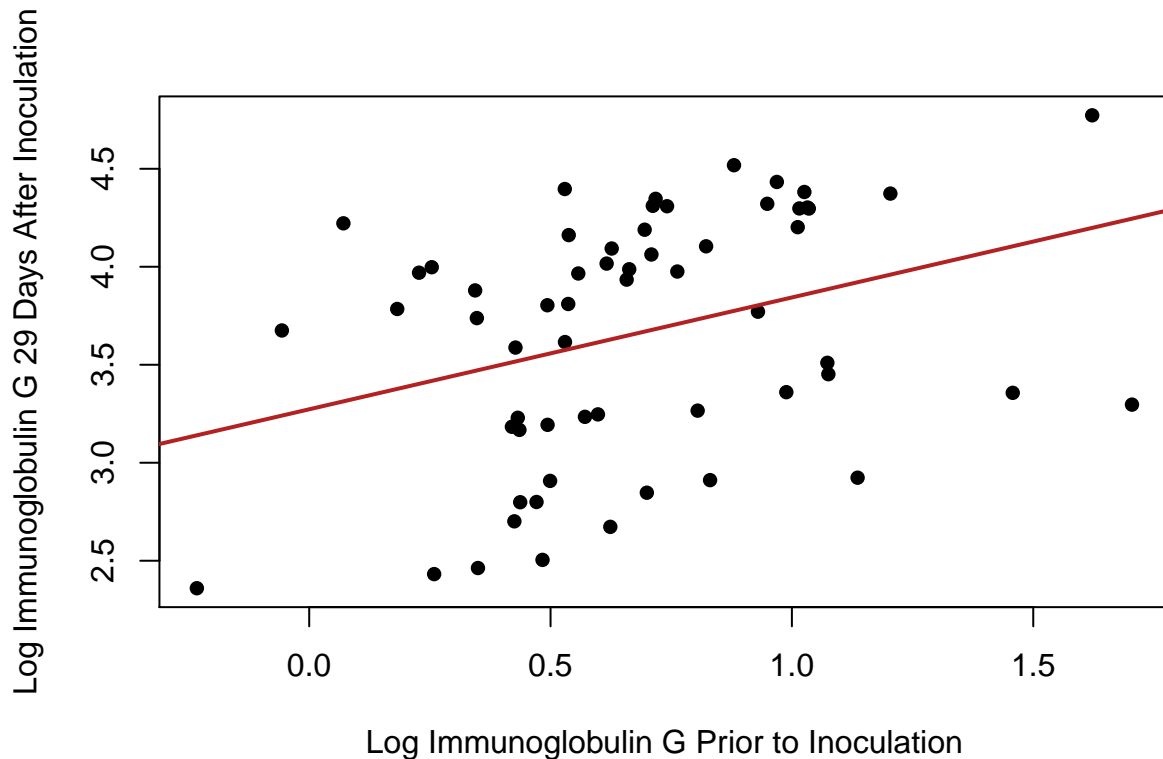
```
##           MEAN          SD          IQR          MEDIAN
## Dose Group: 1 3.2591991, 0.1539724, 0.1665585, 3.2404413
## Dose Group: 2 3.9347705, 0.2598049, 0.3348957, 3.8722545
## Dose Group: 3 4.1264243, 0.1874612, 0.3191539, 4.1329675
## Dose Group: 4 4.2393304, 0.2988536, 0.3444530, 4.3030041
## Dose Group: 5 2.7254719, 0.2638021, 0.3682269, 2.7497923
```

```
t = tapply(immune_test[, var[3]], immune_test[, 4], fn_list)
VNT_df = as.data.frame(t)
```

```
for (i in 1:NROW(VNT_df)){
  rownames(VNT_df)[i] <- paste0("Dose Group: ", i)
}
colnames(VNT_df) <- c("MEAN          SD          IQR          MEDIAN")
VNT_df
```

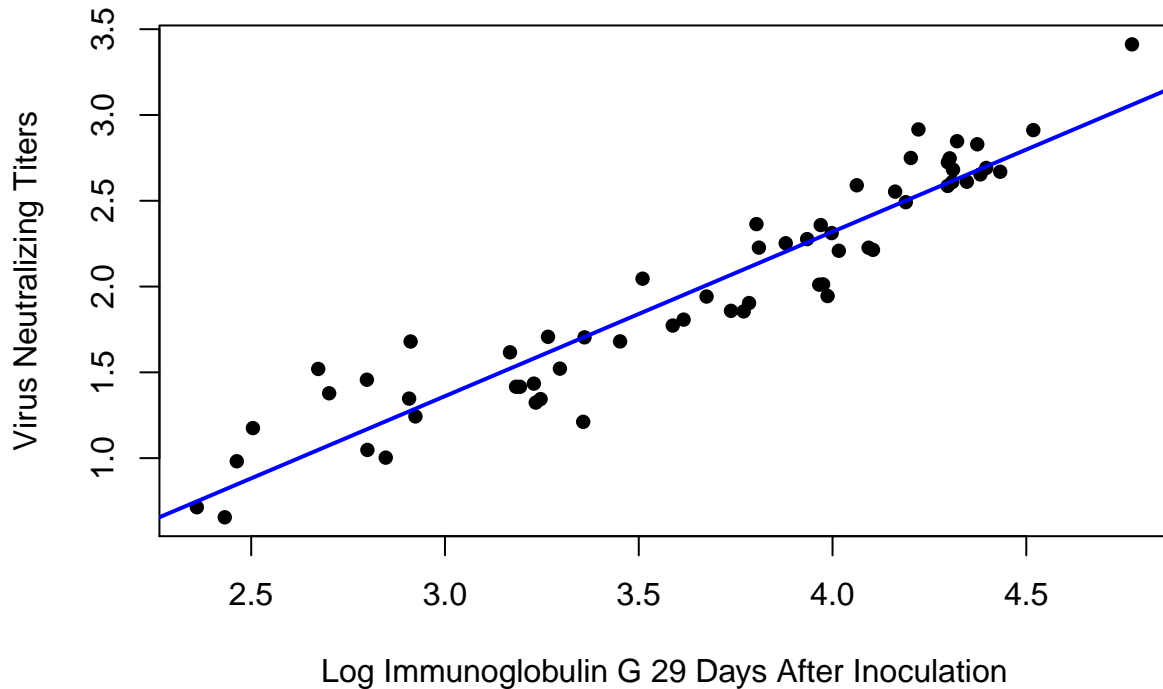
```
##           MEAN          SD          IQR          MEDIAN
## Dose Group: 1 1.5482880, 0.2316119, 0.2883617, 1.5255049
## Dose Group: 2 2.2738550, 0.3919578, 0.7718572, 2.2517141
## Dose Group: 3 2.3854777, 0.3149157, 0.4578352, 2.4250514
## Dose Group: 4 2.5756446, 0.4165343, 0.5004806, 2.6390034
## Dose Group: 5 1.1703048, 0.2953554, 0.4001061, 1.2091465
```

```
plot(immune_test$IgGpre, immune_test$IgG29, pch = 16, xlab = "Log Immunoglobulin G Prior to Inoculation",
     ylab = "Log Immunoglobulin G 29 Days After Inoculation")
abline(lm(immune_test$IgG29 ~ immune_test$IgGpre), lwd = 2, col = "firebrick")
```



```
plot(immune_test$IgG29, immune_test$VNT, pch = 16, ylab = "Virus Neutralizing Titers",
     xlab = "Log Immunoglobulin G 29 Days After Inoculation")
```

```
abline(lm(immune_test$VNT ~ immune_test$IgG29), lwd = 2, col = "blue")
```



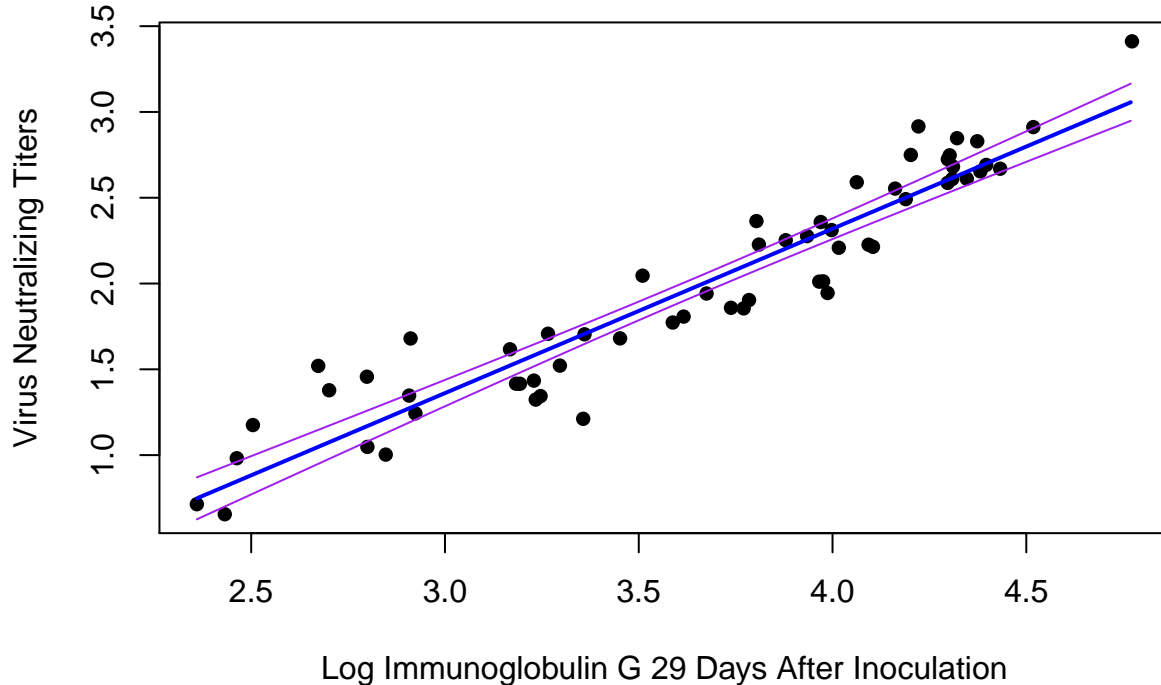
```
model = lm(data = immune_test, VNT ~ IgG29)
# Confidence Intervals
confint(model, level = .99)
```

```
##              0.5 %    99.5 %
## (Intercept) -1.9320169 -1.092745
## IgG29        0.8447733  1.071036
```

```
x = seq(min(immune_test$IgG29), max(immune_test$IgG29), .01)
mypoints = data.frame(IgG29 = x)
```

```
plot(immune_test$IgG29, immune_test$VNT, pch = 16, ylab = "Virus Neutralizing Titers",
     xlab = "Log Immunoglobulin G 29 Days After Inoculation",
     main = "VNT vs IgG29 With 95% Confidence Intervals")
myCI = predict(model, newdata = mypoints, interval = "confidence")
lines(x, myCI[, 2], col = "purple")
lines(x, myCI[, 3], col = "purple")
lines(x, myCI[, 1], col = "blue", lwd = 2)
```

## VNT vs IgG29 With 95% Confidence Intervals



```
m = aov(IgG29 ~ as.factor(dose), data = immune_test)
summary(m)
```

```
##              Df Sum Sq Mean Sq F value Pr(>F)
## as.factor(dose)  4 19.951   4.988   87.43 <2e-16 ***
## Residuals      55  3.138   0.057
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

*# We can conclude from the analysis of variance test that the coefficients of effects for different dose  
# on Virus Neutralizing Titers, are jointly significant. In other words, the means of VNT per each dose  
# group are significantly different from each other. We are told this information through our F Value,  
# which is essentially the ratio of the between-group variance against the within-group variance  
# in our model.*

```
confint(m, level = .95)
```

```
##              2.5 %      97.5 %
## (Intercept)    3.1210186  3.3973796
## as.factor(dose)2  0.4801546  0.8709881
## as.factor(dose)3  0.6718084  1.0626419
## as.factor(dose)4  0.7847145  1.1755480
## as.factor(dose)5 -0.7291440 -0.3383105
```

```
library(sqldf)
```

```
## Loading required package: gsubfn
```

```
## Loading required package: proto
```

```
## Warning in fun(libname, pkgname): couldn't connect to display ":0"
```

```
## Loading required package: RSQLite
```

```

ans = sqldf("
  SELECT dose, avg(IgG29)
  FROM immune_test
  GROUP BY dose
  ")
print(ans)

##   dose avg(IgG29)
## 1    1   3.259199
## 2    2   3.934771
## 3    3   4.126424
## 4    4   4.239330
## 5    5   2.725472

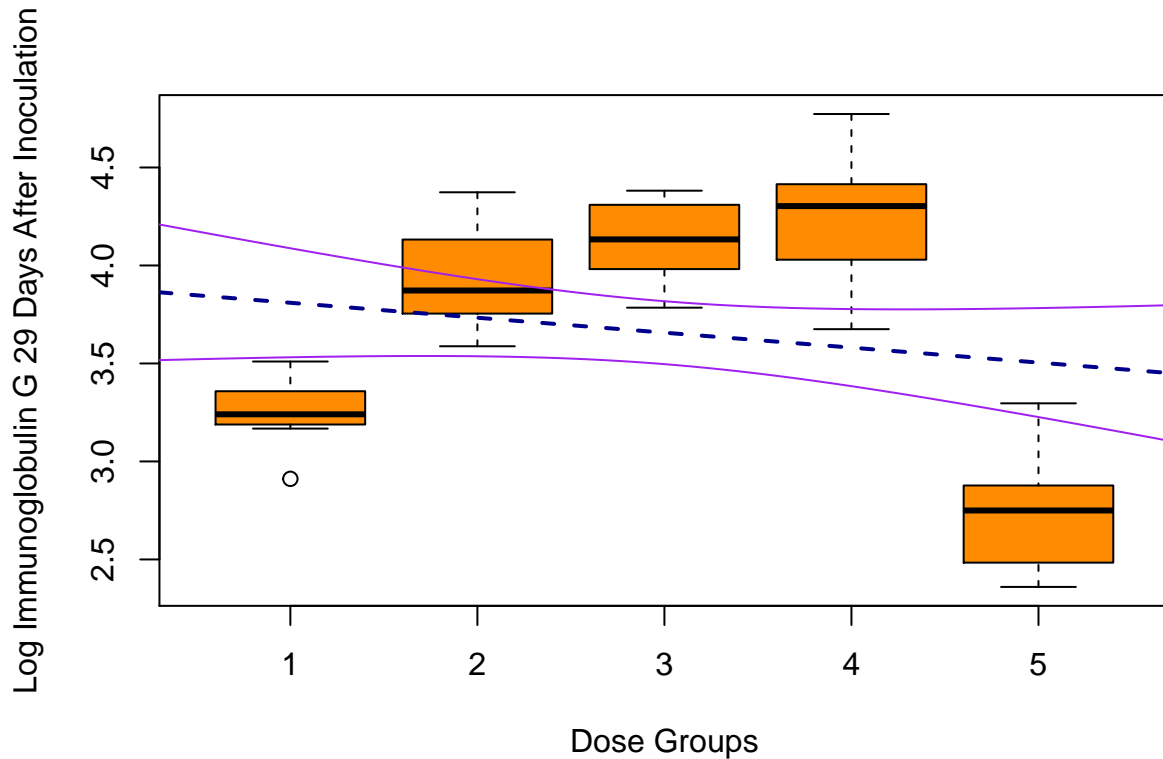
# The greatest log concentration seems to belong to dose group 4, at 4.239330.
# It's 95% Confidence Interval Values are (0.7847145, 1.1755480), with regards to its effect on
# Log Immunoglobulin G 29 Days After Inoculation.

# Running a boxplot regression with Confidence Intervals required that I converted "dose"
# to a numeric value.

immune_test$dose = as.numeric(immune_test$dose)
x = seq(0, 6, .01)
mypoints = data.frame(dose = x)
m = lm(IgG29 ~ dose, data = immune_test)

boxplot(IgG29 ~ dose, data = immune_test, col = "darkorange", xlab = "Dose Groups",
        ylab = "Log Immunoglobulin G 29 Days After Inoculation")
myCI = predict(m, newdata = mypoints, interval = "confidence")
abline(m, lty = 2, col = "darkblue", lwd = 2)
lines(x, myCI[, 2], col = "purple")
lines(x, myCI[, 3], col = "purple")

```



```
with(immune_test, t.test(IgG29[dose == 4], IgGpre[dose == 4]))
```

```
##
## Welch Two Sample t-test
##
## data: IgG29[dose == 4] and IgGpre[dose == 4]
## t = 22.965, df = 19.172, p-value = 2.065e-15
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  3.244909 3.895281
## sample estimates:
## mean of x mean of y
## 4.2393304 0.6692356
```

*# Yes it is significantly different.*

```
diff = sqldf("SELECT IgG29 - IgGpre
              FROM immune_test
              WHERE dose = 4")
```

*# Subtracting the baseline data from the IgPGpre data in the group 4 dose condition.*

```
diff
```

```
## IgG29 - IgGpre
## 1 4.150838
## 2 3.867206
## 3 3.731578
## 4 3.408268
## 5 3.262033
## 6 3.638150
## 7 3.150635
## 8 3.466389
```



```
## 9      3.463865
## 10     3.567617
## 11     3.535515
## 12     3.599043
```

```
mean(diff$`IgG29 - IgGpre`)
```

```
## [1] 3.570095
```

```
# The mean of the difference is 3.570095, the 95% CI, in the above t-test, is (3.244909, 3.895281).
```

```
names(diff) <- c("difference")
```

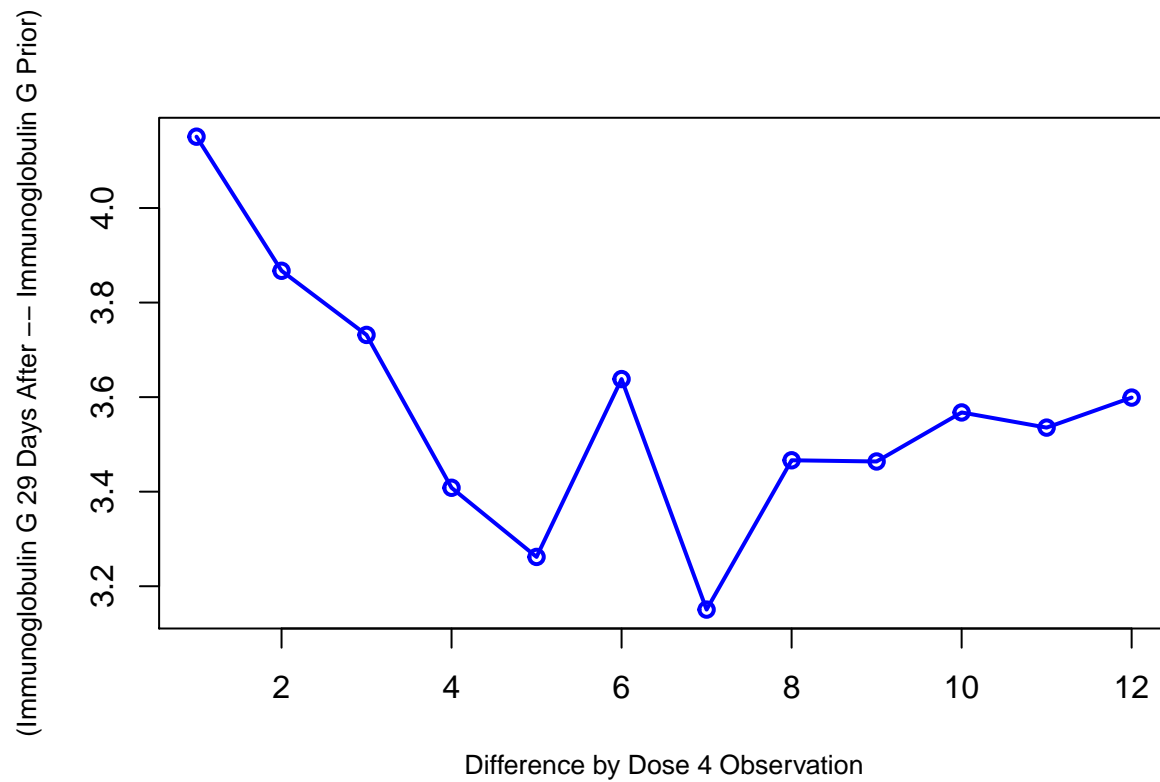
```
# Attaching Dose Value to the Data Frame, and using a rudimentary method to attach '4' to every row value
```

```
diff = sqldf("
  SELECT *,
  (CASE difference
  WHEN 100000 THEN 'No Comment'
  ELSE '4'
  END)
  AS 'Dose'
FROM diff
")
diff
```

```
##      difference Dose
## 1      4.150838    4
## 2      3.867206    4
## 3      3.731578    4
## 4      3.408268    4
## 5      3.262033    4
## 6      3.638150    4
## 7      3.150635    4
## 8      3.466389    4
## 9      3.463865    4
## 10     3.567617    4
## 11     3.535515    4
## 12     3.599043    4
```

```
# Here is the data frame, corresponding to all of the dose values and there respective 'IgG29 - IgGpre' values.
```

```
plot(diff$difference, type = "o", xlab = "Difference by Dose 4 Observation",
      ylab = "(Immunoglobulin G 29 Days After -- Immunoglobulin G Prior)", cex.lab = .8,
      col = "blue", lwd = 2)
```



```
expected_VNT = sqldf("
    SELECT avg(VNT) Expected_VNT
    FROM immune_test
    WHERE dose = 4
")
expected_VNT
```

```
## Expected_VNT
## 1 2.575645
```