Stat*3240, Assignment 2
Graham Eckel
0679576

Q1

```
set.seed(2019-09-26)
dir = "E:\\...\\Applied Regression Analysis\\"
file1 = "3240_F19_A1_DDT.txt"
dfDDT = read.table(file=paste(dir,file1, sep=""), header=TRUE, sep=' ')

DDT = dfDDT$DDT

lmThick = lm(dfDDT$thickness~DDT)
summary(lmThick)
```

a)

A 95% confidence interval for the true mean egg thickness if DDT is 2000 ppm can be calculated by using the predict.lm function in R, as such:

```
dfDDT1 = data.frame(DDT=2000)
predict.lm(lmThick, newdata = dfDDT1, interval = "confidence")
```

```
      fit       lwr       upr
0.2806386 0.2464229 0.3148543
```

Which gives us a lower bound of 0.2464 and an upper bound of 0.3148.

b)

A 95% prediction interval for the true mean egg thickness if DDT is 2000 ppm can be calculated by using the predict.lm function in R, as such:

```
predict.lm(lmThick, newdata = dfDDT1, interval = "prediction")
```

```
      fit       lwr       upr
0.2806386 0.1693704 0.3919068
```

Which gives us a lower bound of 0.1694 and an upper bound of 0.3919.

c)

```
Residuals:
     Min       1Q   Median       3Q      Max 
-0.07255 -0.04945  0.01040  0.03718  0.08989 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept)  3.810e-01  2.412e-02  15.791   <2e-16 ***
DDT         -5.016e-05  1.840e-05  -2.726   0.0102 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.05204 on 33 degrees of freedom
Multiple R-squared:  0.1838,	Adjusted R-squared:  0.1591 
F-statistic: 7.432 on 1 and 33 DF,  p-value: 0.01018
```

From the summary table above, we can see that the R-squared value is 0.1838. R-squared can be viewed as the proportion of the variance in the response variable that be explained in the model. In this case, roughly 18% of the variance in eggshell thickness can be explained by the linear relationship with DDT.