PART A

1. Regression model with weekly wages as the response and years of education and experience as predictors.

```
##
## Call:
## lm(formula = wage ~ educ + exper, data = uswages2)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -1014.7  -235.2   -52.1   150.1  7249.2
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -239.1146    50.7111  -4.715 2.58e-06 ***
## educ          51.8654     3.3423  15.518  < 2e-16 ***
## exper          9.3287     0.7602  12.271  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 426.8 on 1964 degrees of freedom
## Multiple R-squared:  0.1348, Adjusted R-squared:  0.1339
## F-statistic:   153 on 2 and 1964 DF,  p-value: < 2.2e-16
```

2. The *percentage of variation* (R^2) in the response explained by the model is **13.51%**.

```
## [1] 0.134793
```

3. The *case number* is **1550**, and the *residual value* belonging to this case number is **7249.174**.

```r
#Case number
which.max(lmodsum$residuals)
```

```
## 15387
##  1550
```

```r
#Max residual
max(lmodsum$residuals)
```

```
## [1] 7249.174
```

4. The mean is close to zero, whereas the median is close to -52.14. This leads us to conclude that the distribution of the residuals is **skewed right** given the mean is bigger than the medians. In addition, in general incomes tend to have a right skewed distribution, where the majority make little/average money, and a few ones make very high profits. Also, one of the assumptions when using RSS is that the mean of the errors (residuals) is zero, which is apporximately correct in here too.

```
## [1] -1.381535e-15
```

```
## [1] -52.14337
```

5. For two people with the same education and one year difference in experience, the model suggests a difference in predicted weekly wages of **$9.33**.

6. The *correlation* of fitted values is approximately **zero**.

```
#Correlation
cor(lmod$residuals, lmod$fitted.values)
```

```
## [1] 6.35678e-17
```

6. Geometrically speaking, a correlation of zero hints at **orthogonality**. Orthogonality simply means that the two vecotrs are perpendicular to one another (form a 90 degree angle). In Figure one, we are able to see the plot of the residuals vs the fitted. Next to it (right), we are able to see the theoretical geometric reppresentation and indeed, our correlation of zero, is able to prove that the fitted values (y^) are orthogonal with the residuals (e^).

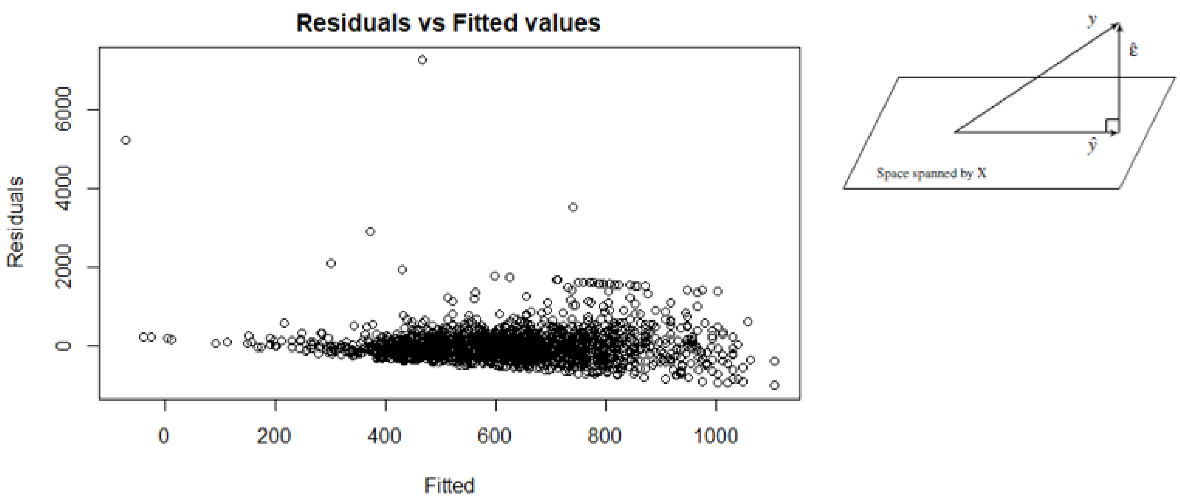


Figure 1: 'Geometric Reppresentation of Residuals vs Fitted'

PART B

```
## [1] "X matrix"

##       [,1] [,2] [,3]
##  [1,]    1    2   -2
##  [2,]    1   -1   -2
##  [3,]    1    3   -2
##  [4,]    1    3    3
##  [5,]    1    2    3
##  [6,]    1    1    3
##  [7,]    1    0    0
##  [8,]    1    0    0
##  [9,]    1   -1    0
## [10,]    1    0    1
```

```
## [1] "B matrix"

##      [,1]
## [1,]    1
## [2,]   -1
## [3,]    2
```

```
## [1] "e matrix"

##              [,1]
##  [1,]  0.36807403
##  [2,] -0.71390603
##  [3,]  0.05006762
##  [4,]  2.35495303
##  [5,]  0.29459138
##  [6,] -0.61626467
##  [7,]  0.14338742
##  [8,] -0.44539615
##  [9,] -0.94897086
## [10,]  0.02089350
```

```
## [1] "Y matrix"

##             [,1]
##  [1,] -4.6319260
##  [2,] -2.7139060
##  [3,] -5.9499324
##  [4,]  6.3549530
##  [5,]  5.2945914
##  [6,]  5.3837353
##  [7,]  1.1433874
##  [8,]  0.5546039
##  [9,]  1.0510291
## [10,]  3.0208935
```

Figure 2: 'Matrices X, B, e, and Y'

1. Beta^ coefficient esimtates

```
#B^= (XtX)^-1XtY
xtxi <- solve(t(X) %*% X)
b = xtxi %*% t(X) %*% Y
```

```
##            [,1]
## [1,]  1.1021718
## [2,] -0.7820157
## [3,]  1.8277382
```

2. True variance of sigma^2

```
#var(B) = (XtX)^-1sigma^2 , now sigma is var(e)= sigma^2*I, where sigma is varaince of
#e (1^2) & I is the identity matrix
B_var= xtxi %*% diag(3)
```

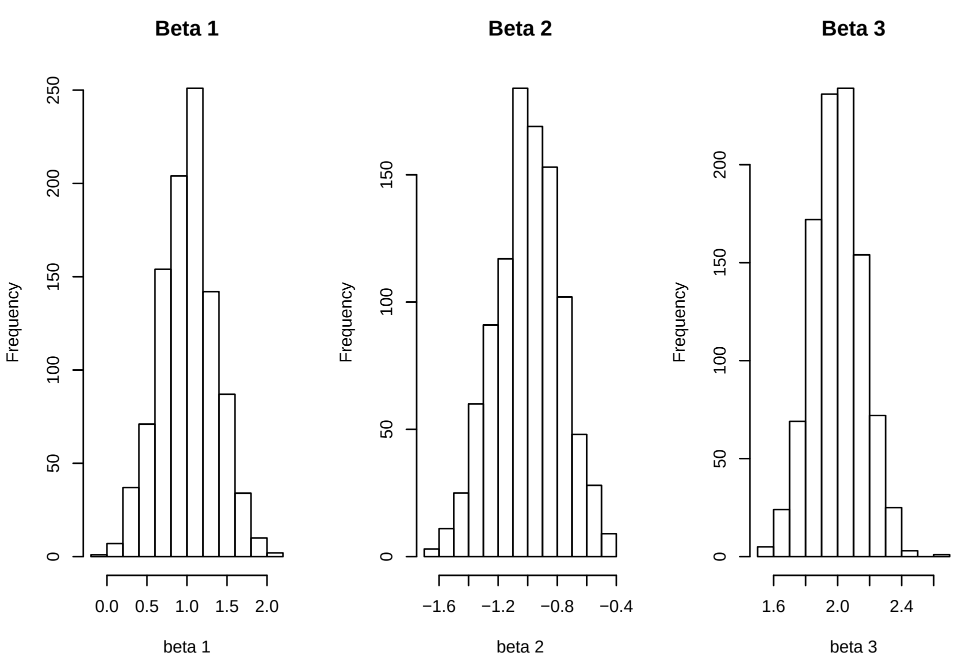
```
##              [,1]         [,2]         [,3]
## [1,]  0.139180672 -0.042016807 -0.003413866
## [2,] -0.042016807  0.050420168 -0.008403361
## [3,] -0.003413866 -0.008403361  0.027442227
```

3. Estimate of sigma^2 throguh the residuals

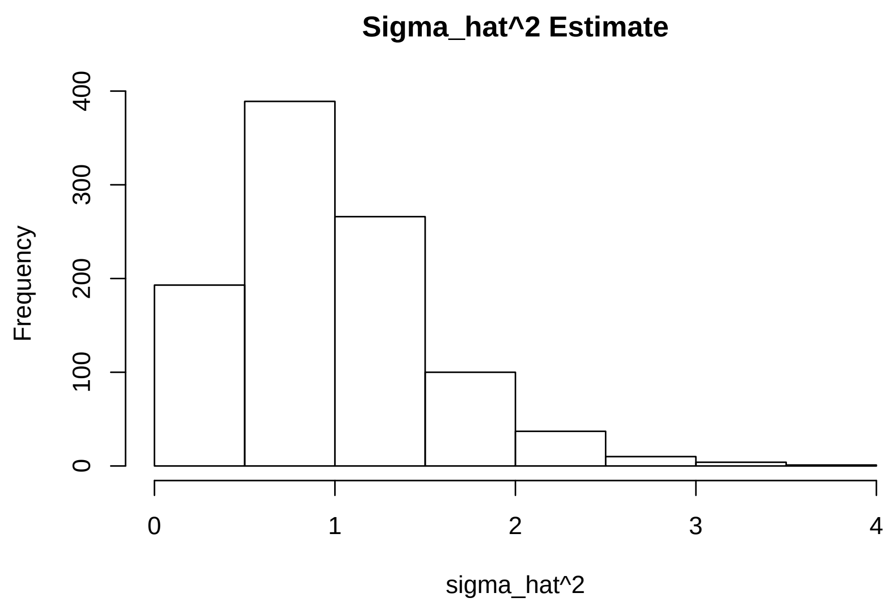
```
#sigma_hat^2 = e^t*e^/ (n-p) , RSS/df.residuals
#e= y-y^, y^ = xb^y => x(x^tx)^-1xty
e = sum((Y - X%*%b)^2) / (10-3)
```

```
## [1] 0.5720924
```

4. Yes, the current estimates match with the coefficients of beta (1,-1,2), and their variances are very similar to the ones in number 2. The variances being: **beta1**=0.1318, **beta2**=0.0498, **beta3=0.0265**.



5. Sigma_hatˆ2, with a **mean** of about 1.0030, does indeed provide a **good** estimate for the actual sigmaˆ2 parameter which is defined as 1. The main reasons for such reality is the use of an e with **normal** distribution (mean=0, sd=1) and RSS for the actual estimation of sigmaˆ2.



6. The distribution used was the **uniform** one, which has expectation of zero, and variance of one. All of the three Betas (1,2,3) share similar variances to previous results, but with a slight increase in the variance. Same for the sigmaˆ2 it is quite similar, with the main difference being in the distribution, which is less skewed and almost normal.