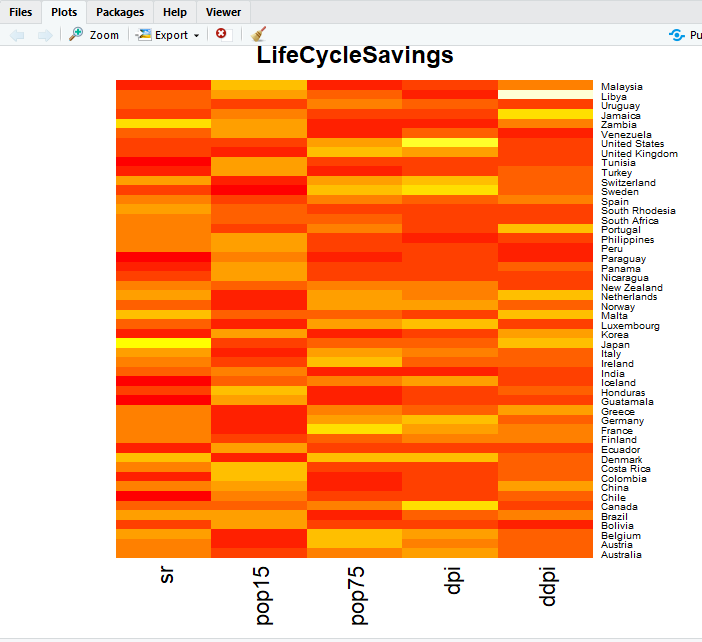
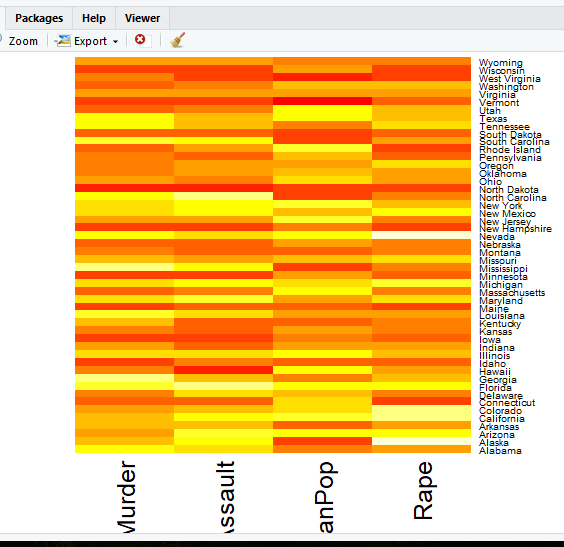
**W3\_Lab - Bruno Simione Beltrame – N01220860**

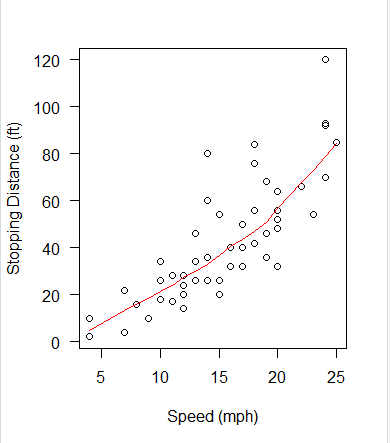
**Task1**



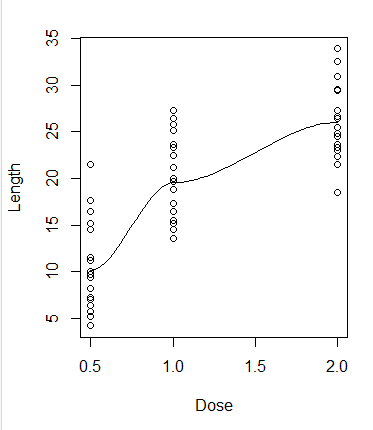
**Question 1**



**Task**

**Question 2**



**Question 3**

> cor(ToothGrowth$len, ToothGrowth$dose)

[1] 0.8026913

**Question 4**

> toothLm <- lm(ToothGrowth$len ~ ToothGrowth$dose)

> toothLm$coefficients

(Intercept) ToothGrowth$dose

7.422500 9.763571

**Question 5**

> summary(toothLm)

Call:

lm(formula = ToothGrowth$len ~ ToothGrowth$dose)

Residuals:

Min 1Q Median 3Q Max

-8.4496 -2.7406 -0.7452 2.8344 10.1139

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 7.4225 1.2601 5.89 2.06e-07 \*\*\*

ToothGrowth$dose 9.7636 0.9525 10.25 1.23e-14 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.601 on 58 degrees of freedom

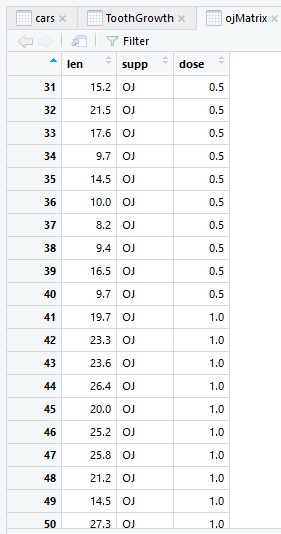
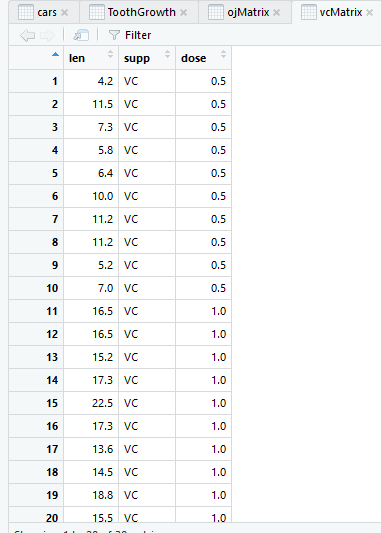
Multiple R-squared: 0.6443, Adjusted R-squared: 0.6382

F-statistic: 105.1 on 1 and 58 DF, p-value: 1.233e-14

Yes, it is a statistically significant model because the p-value (1.233e-14) is less than 0.05.

Around 64.43% of the dependent variable can be explained by the model.

**Question 6**

**Optional**

> ojLm <- lm(ojMatrix$len ~ ojMatrix$dose)

> vcLm <- lm(vcMatrix$len ~ vcMatrix$dose)

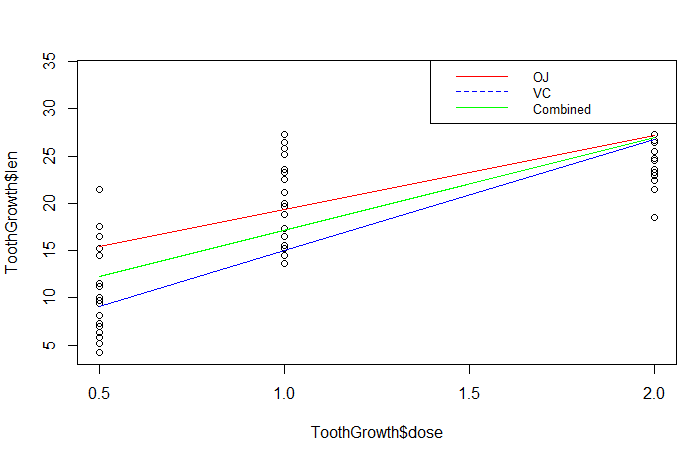
> toothLm <- lm(ToothGrowth$len ~ ToothGrowth$dose)

> plot(ToothGrowth$dose, ToothGrowth$len)

> lines(ojMatrix$dose, ojLm$fitted.values, col = "red")

> lines(vcMatrix$dose, vcLm$fitted.values, col = "blue")

> lines(ToothGrowth$dose, toothLm$fitted.values, col = "green")



**Question 7**

> summary(ojLm)

Call:

lm(formula = ojMatrix$len ~ ojMatrix$dose)

Residuals:

Min 1Q Median 3Q Max

-7.2557 -3.7979 -0.0643 3.3521 7.9386

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 11.550 1.722 6.708 2.79e-07 \*\*\*

ojMatrix$dose 7.811 1.302 6.001 1.82e-06 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.446 on 28 degrees of freedom

Multiple R-squared: 0.5626, Adjusted R-squared: 0.547

F-statistic: 36.01 on 1 and 28 DF, p-value: 1.825e-06

> summary(vcLm)

Call:

lm(formula = vcMatrix$len ~ vcMatrix$dose)

Residuals:

Min 1Q Median 3Q Max

-8.2264 -2.6029 0.0814 2.2288 7.4893

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 3.295 1.427 2.309 0.0285 \*

vcMatrix$dose 11.716 1.079 10.860 1.51e-11 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.685 on 28 degrees of freedom

Multiple R-squared: 0.8082, Adjusted R-squared: 0.8013

F-statistic: 117.9 on 1 and 28 DF, p-value: 1.509e-11

If we compare the R2 of both models, we can see that the VC Model (R2 = 80.82%) is said to be statistically more significant than the OJ Model (R2 = 56.26%). So, the linear model created for the OJ shows that we can explain 80.82% of the variations in the dependent variable using the model.