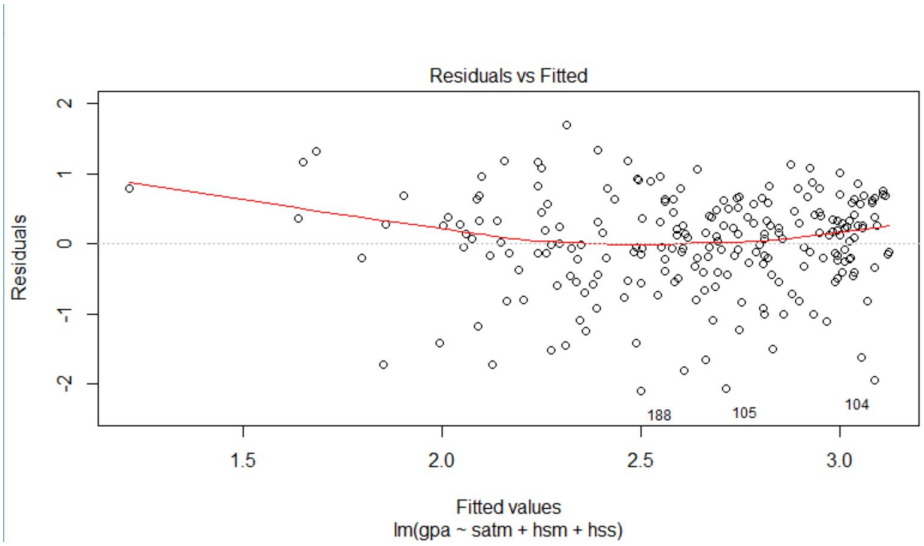
“I pledge my honor that I have abided by the Stevens Honor System” - Himanshu Rana

1)

a)



Based on this plot, the regression model seems adequate for the data

b)

```
Call:
lm(formula = gpa ~ satm + hsm + hss, data = csgpa)

Residuals:
     Min       1Q   Median       3Q      Max
-2.10954 -0.37657  0.08842  0.45121  1.68691

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 0.4843393  0.3601535   1.345    0.180
satm        0.0006383  0.0006092   1.048    0.296
hsm         0.1597422  0.0381288   4.190 4.05e-05 ***
hss         0.0545926  0.0337547   1.617    0.107
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7003 on 220 degrees of freedom
Multiple R-squared:  0.2036,    Adjusted R-squared:  0.1928
F-statistic: 18.75 on 3 and 220 DF,  p-value: 7.222e-11
```

Beta1 = 0.0006383 and std. Error 1 = 0.0006902

Beta2 = 0.1597422 and std. Error 2 = 0.0381288

Beta3 = 0.0545926 and std. Error 3 = 0.0337547

```
> anova(regression)
Analysis of Variance Table

Response: gpa
           Df  Sum Sq Mean Sq F value    Pr(>F)
satm        1   8.583  8.5829 17.5036 4.144e-05 ***
hsm         1  17.720 17.7198 36.1368 7.566e-09 ***
hss         1   1.283  1.2826  2.6158    0.1072
Residuals 220 107.877  0.4904
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

c) At alpha level = 5% = 0.05, the null hypothesis of the covariates satm and hsm should be rejected. While the hss should not because its p value is not less than alpha.

d) Hss has a better correlation to the first year gpa than SAT math and high school math scores.

2)

a)

SSM = 256.68

SSE = 91.14

SST = 347.82

b)

```
> wear1 <- c(9.1, 17.1, 20.8, 11.8)
> wear2 <- c(13.4, 20.3, 28.3, 16)
> wear1 <- c(9.1, 13.4, 15.6)
> wear2 <- c(17.1, 20.3, 24.6)
> wear3 <- c(20.8, 28.3, 23.7)
> wear4 <- c(11.8, 16, 16.2)
> combined <- data.frame(cbind(wear1, wear2, wear3, wear4))
> stacked <- stack(combined)
> stacked
> anova_results <- aov(values ~ ind, data = stacked)
> summary(anova_results)
            Df Sum Sq Mean Sq F value Pr(>F)
ind          3 256.68   85.56    7.51 0.0103 *
Residuals    8  91.14   11.39
---
Signif. codes:
0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
>
```

F = 7.51

P value = 0.0103

Since the p value is less than alpha we reject the null hypothesis. There is a significant difference of wear between the four different types.