

Homework 3

Jacob Plaza

3/2/2022

QUESTION 1

```
question1 = read.table(file.choose(),header = T,sep = ",")  
X = question1QUETY = question1SBP reg = lm(Y~X) reg  
residual = reg$residuals plot(residual) hist(residual)
```

QUESTION 3

```
question2 = read.table(file.choose(),header = T,sep = "") #a question2 = read.table(file.choose(),header = T,sep  
="")  
Armstrength = question2ARMSTRENGTHDynamiclift = question2DYNAMICLIFT  
plot(Armstrength, Dynamiclift)  
#b correlation=cor(Armstrength, Dynamiclift) correlation  
#c correlation_test = cor.test(Armstrength, Dynamiclift) correlation_test  
#d r_2 = correlation^2 r_2  
fit = lm(Armstrength ~ Dynamiclift) nova = anova(fit) nova SSY = nova [1,2] + nova[2,2] SSE = nova [2,2]  
r_2_comp = (SSY-SSE)/SSY r_2_comp
```

1) There are no patterns in the residual plot, so the assumption of constant variance is not violated. Also, from the histogram it appears that the data is normally distributed.

2)

	Degree Freedom	Sum of Sq	Mean Sq	F-Ratio
Regression	1	4.93598	4.93598	30.01
Residual	30	2.7237	0.1695	
Total	31	7.95968		

$$SSE = 0.09079(30) \\ = 2.7237$$

$$SST = 7.65968 - 2.7237 \\ = 4.93598$$

$$MST = 4.93598$$

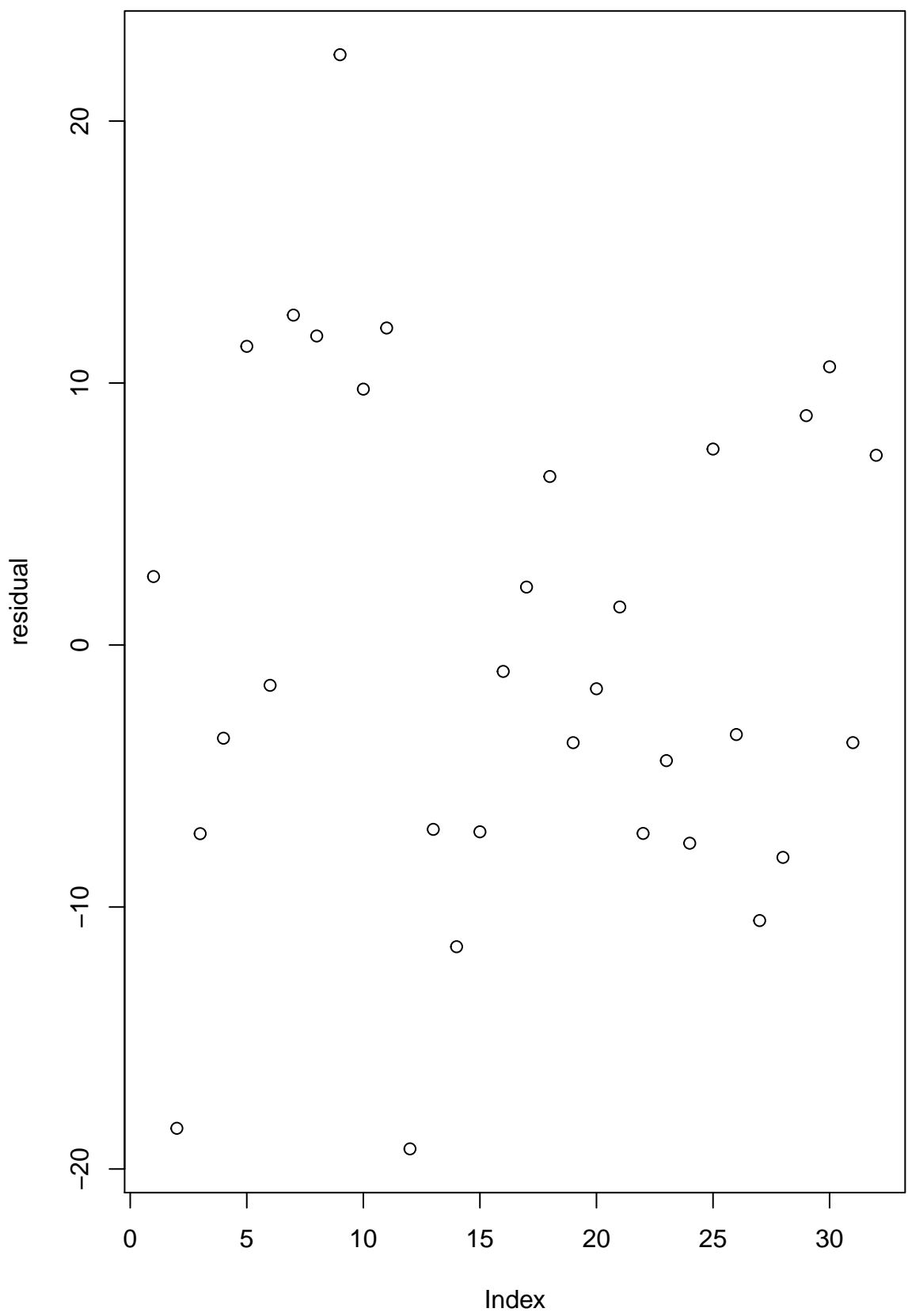
$$MSE = \frac{4.93598}{30}$$

$$F = \frac{4.93598}{0.1645} = 30.01$$

$$= 0.1645$$

3]

- a) The correlation coefficient will be low since the scatter plot does not suggest a linear relationship.
- b) The correlation coefficient is very small, 0.3726, indicating a weak linear relationship between Armstrength & Dynamic lift.
- c) P-value = 0.05217 since $0.05217 > 0.05$,
Sig = 0.05 we do not reject
null hypothesis.
There is no linear relationship.
- d) They yield the same results.
- e) the fit is bad. The R^2 shows that Armstrength only explains 13.4% of variation in Dynamic lift. Therefore, a linear model is not appropriate.



Histogram of residual

