

Linear regression

Alizeh Azhar

9/20/2021

For problem 1:

- (A) the graph does not form a linear pattern as it has a strong relationship it would be best suited to model it with a curve instead of a line
- (B) this graph shows a strong relationship, linear model is not reasonable in this case and would be better to model it with a curve
- (C) this graph depicts a slight weaker relationship and the application of a linear model is reasonable as it forms a perfect linear pattern
- (D) the graph shows a weak relationship and you could fit a line thus linear model is reasonable
- (E) this graph shows a moderate relationship it is reasonable to fit a linear model
- (F) this graph shows a moderate linear relationship thus a linear model is reasonable

Problem 2: OLS Exercise 8.5

- (a) Exam 2 has the strongest correlation with the final exam. This is because the plot forms a better linear pattern as compared to Exam 1. Exam 1 has many outliers while majority of the points in Exam 2 linear on or closer to the fitted line as compared to exam 1.
- (B) Exam 2 is close to final exam, the student is more prepare for Exam2 than Exam 1 because it's the predictor of final performance.

problem 3

- 3A) There is strong linear relationship between husband and wife age. As husband's age increases woman's age rises.
- 3(B): there is a weak positive linear relationship between husband's and wives heights and they aren't any outliers
- 3(C): the plot comparing husband and wife's ages shows a stronger correlation
- 3D) :Conversion of units for heights from centimeters to inches does not influence the correlation between husband's and wife's height.

Problem 4: OLS Exercise 8.20

0.5 shows under estimation of the incidences of skin cancer. This is because the residual is positive.

Problem 5

5(A): please show that for the equation is written like this correction should be made $\text{height} = 106.23 + 0.61 \times \text{girth}$

5(B): a change in girth by 1cm increases height by 0.61cm and when shoulder girth is 0 the height is expected to be 106.23cm

5(C) Height prediction

```
height= 106.23 + 0.61*100
height
```

```
## [1] 167.23
```

5(D): Residual

```
Residual = 160 - 167.23
Residual
```

```
## [1] -7.23
```

The residual means that the student was 7.23 shorter than predicted

Problem 6 (A)

```
babies <- read.csv("https://www.openintro.org/data/csv/babies.csv")
model_babies= lm(bwt~gestation, data=babies)
model_babies
```

```
##
## Call:
## lm(formula = bwt ~ gestation, data = babies)
##
## Coefficients:
## (Intercept)    gestation
##    -10.0642      0.4643
```

6 (B) Write the equation of the line using mathematical notation (2 pts).

Birth weight = -10.06418 + 0.4642 X Gestation