

R1

Jeremiah Theisen

2024-09-14

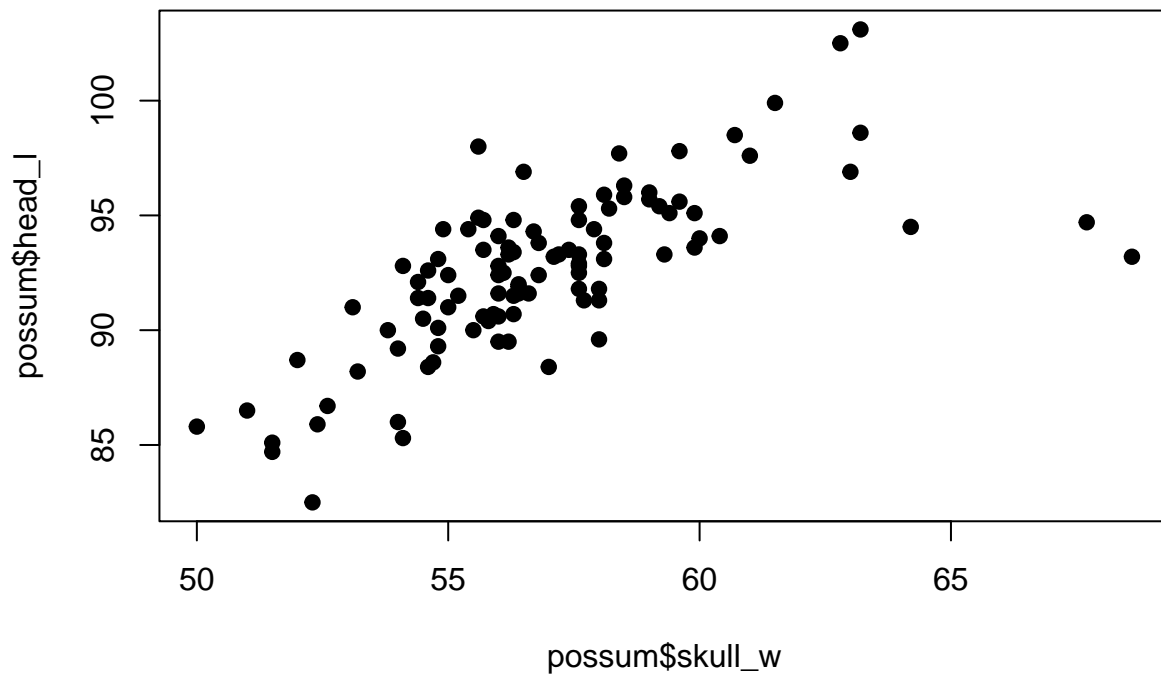
7.1.1

```
possum = read.csv("https://github.com/TienChih/tbil-stats/raw/main/data/possum.csv")  
names(possum)
```

```
## [1] "site"      "pop"       "sex"       "age"       "head_l"    "skull_w"   "total_l"  
## [8] "tail_l"
```

a.

```
plot(possum$skull_w, possum$head_l, pch=19)
```



b. The 62mm skull would be predicted to have a longer head

c. 53mm skull would be around 87mm long. 62mm skull would be around 96mm long. I would feel fairly confident that they would line up that way, but the exact numbers would be off by a bit.

d. 55mm. Very confident because there seems to be a large cluster at 55mm.

7.1.2

- a. c and f have strong linear relationships.

7.1.3

- a. $R = 0.35$
- b. $R = -0.97$
- c. $R = 0$
- d. $R = 0.65$
- e. $R = -0.38$
- f. $R = -0.85$
- g. $R = 0.99$
- h. the same graph as f.
- i. $R = 0.85$

7.1.4

I would say (b) is the best fit

7.1.5

- a. <https://www.desmos.com/calculator/lbg4ya4vwx>
- b. It fits a lot better than it did before.

7.1.6

- a. <https://www.desmos.com/calculator/iy0a1qf4e> This B_0 is smaller, and the B_1 is larger
- b. $r = 0.8906$. There is a strong but not perfect relationship.
- c. $r^2 = 0.7931$. There is a strong but not perfect relationship.
- d.

```
x1=c(1,2,3,4,4,5,6)
y1=c(1,3,2,4,5,4,7)
```

```
dummydata = data.frame(x1, y1)
```

- e.

```
dummymod=lm(y1~x1, data=dummydata)
print(dummymod)
```

```
##
## Call:
## lm(formula = y1 ~ x1, data = dummydata)
##
## Coefficients:
## (Intercept)      x1
##    0.05645    1.02419
```

The slope and intercept are the same.

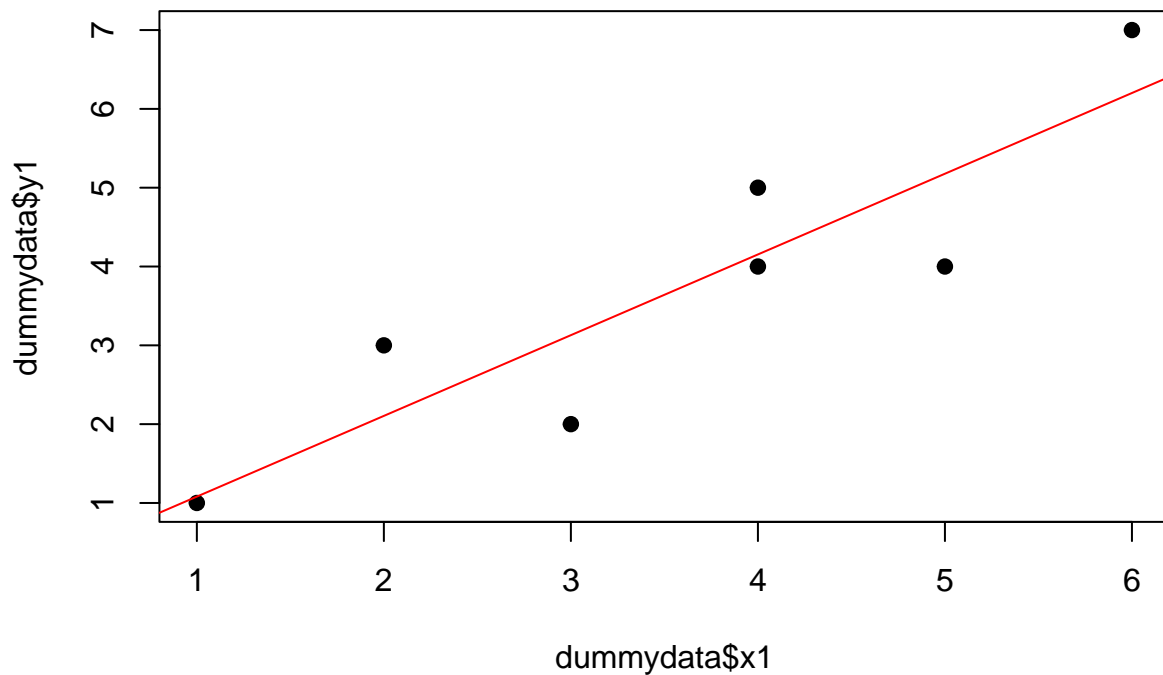
f.

```
cor(dummydata$x1, dummydata$y1)
```

```
## [1] 0.8905759
```

g.

```
plot(dummydata$x1, dummydata$y1, pch=19)
abline(dummymod, col="red")
```



7.1.7

a.

```
possummod=lm(head_l~skull_w, data=possum)
print(possummod)
```

```
##
## Call:
## lm(formula = head_l ~ skull_w, data = possum)
##
## Coefficients:
## (Intercept)      skull_w
##      46.1954       0.8158
```

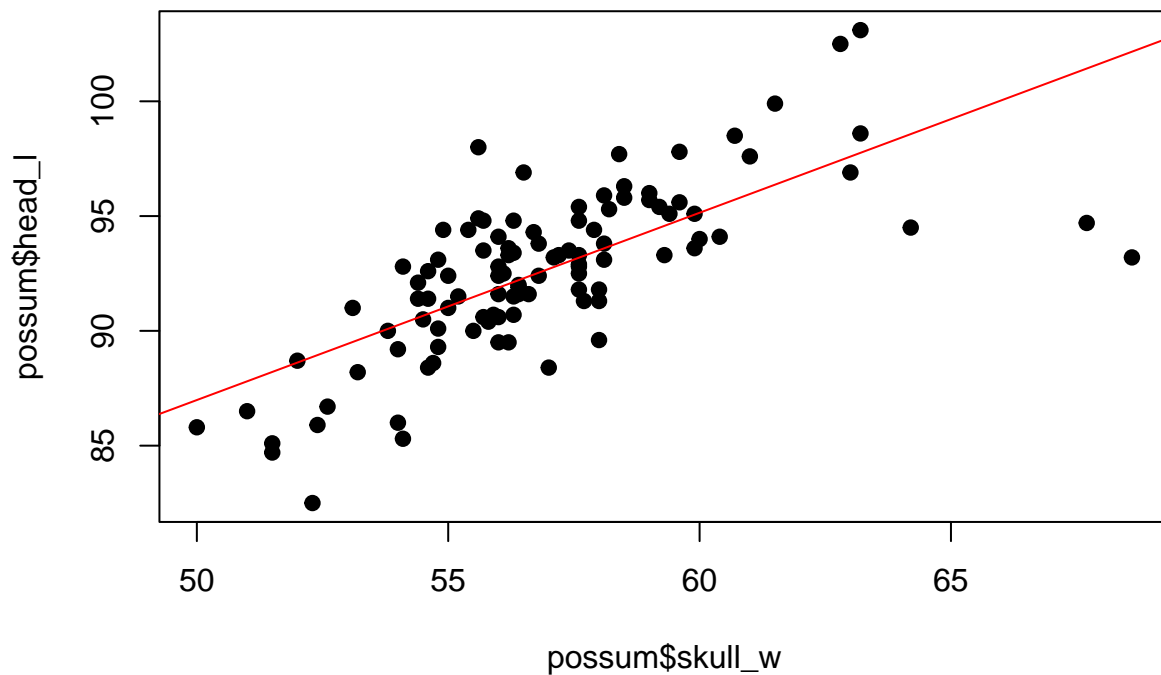
The slope is 0.8158, and intercept is 46.1954 b.

```
cor(possum$skull_w, possum$head_l)
```

```
## [1] 0.7108268
```

c. $r = 0.710$, $r^2 = 0.5041$. There is a weak correlation
d.

```
plot(possum$skull_w, possum$head_l, pch=19)
abline(possummod, col="red")
```



e.

```
summary(possummod)
```

```
##
## Call:
## lm(formula = head_l ~ skull_w, data = possum)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -8.9615 -1.3320  0.0588  1.6153  6.4444
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  46.19537    4.55355   10.14  <2e-16 ***
## skull_w       0.81583    0.07993   10.21  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.526 on 102 degrees of freedom
## Multiple R-squared:  0.5053, Adjusted R-squared:  0.5004
## F-statistic: 104.2 on 1 and 102 DF,  p-value: < 2.2e-16
```

f.

```
summary(possum)
```

```
##      site      pop      sex      age
## Min.   :1.000  Length:104  Length:104  Min.   :1.000
## 1st Qu.:1.000  Class :character  Class :character  1st Qu.:2.250
## Median :3.000  Mode  :character  Mode  :character  Median :3.000
## Mean   :3.625
## 3rd Qu.:6.000
## Max.   :7.000
##
##      head_l      skull_w      total_l      tail_l
## Min.   : 82.50  Min.   :50.00  Min.   :75.00  Min.   :32.00
## 1st Qu.: 90.67  1st Qu.:54.98  1st Qu.:84.00  1st Qu.:35.88
## Median : 92.80  Median :56.35  Median :88.00  Median :37.00
## Mean   : 92.60  Mean   :56.88  Mean   :87.09  Mean   :37.01
## 3rd Qu.: 94.72  3rd Qu.:58.10  3rd Qu.:90.00  3rd Qu.:38.00
## Max.   :103.10  Max.   :68.60  Max.   :96.50  Max.   :43.00
##
```

g.

```
possummod2=lm(total_l~tail_l, data=possum)
print(possummod2)
```

```
##
## Call:
## lm(formula = total_l ~ tail_l, data = possum)
##
```

```
## Coefficients:
## (Intercept)      tail_l
##      41.037      1.244
```

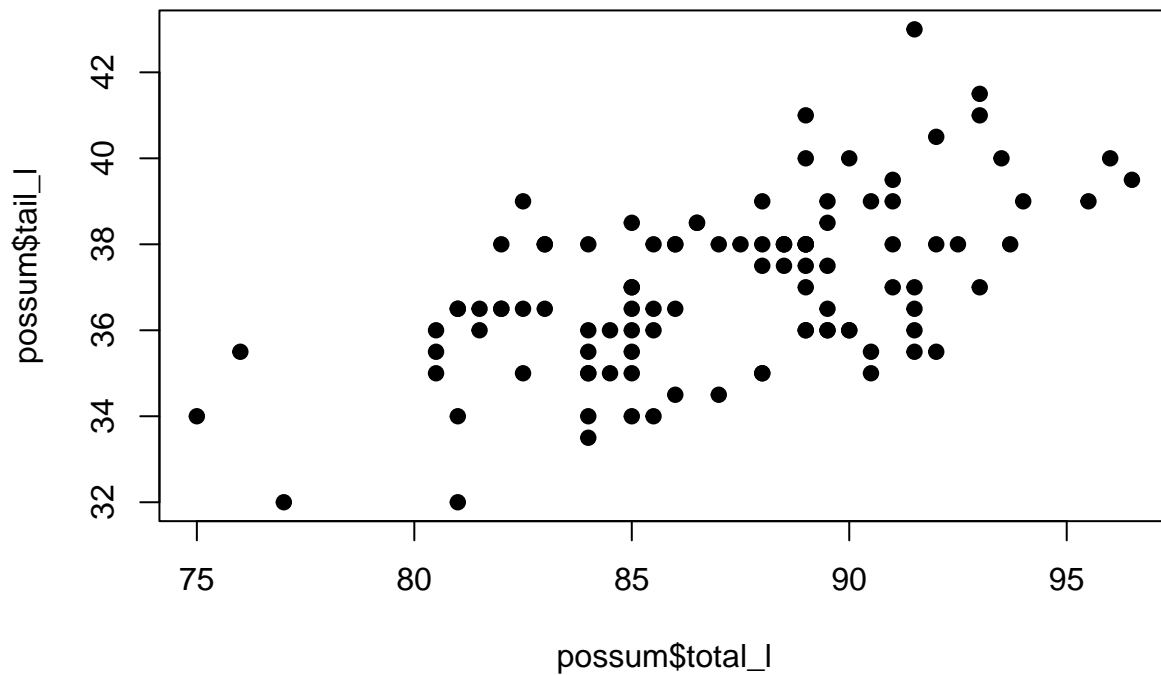
```
summary(possummod2)
```

```
##
## Call:
## lm(formula = total_l ~ tail_l, data = possum)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -9.2100 -2.3265  0.1792  2.7765  6.7900
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  41.0371     6.6568   6.165 1.43e-08 ***
## tail_l       1.2443     0.1796   6.927 3.94e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 3.572 on 102 degrees of freedom
## Multiple R-squared:  0.32, Adjusted R-squared:  0.3133
## F-statistic: 47.99 on 1 and 102 DF, p-value: 3.935e-10
```

```
cor(possum$total_l, possum$tail_l)
```

```
## [1] 0.5656455
```

```
plot(possum$total_l, possum$tail_l, pch=19)
abline(possummod2, col="red")
```



```
possummod3=lm(age~tail_l, data=possum)
print(possummod3)
```

```
##
## Call:
## lm(formula = age ~ tail_l, data = possum)
##
## Coefficients:
## (Intercept)      tail_l
##      -0.4276      0.1150
```

```
summary(possummod3)
```

```
##
## Call:
## lm(formula = age ~ tail_l, data = possum)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -3.0583 -1.2476 -0.5407  1.3874  5.1717
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -0.4276     3.5833  -0.119   0.905
```

```
## tail_1      0.1150      0.0966      1.191      0.237
##
## Residual standard error: 1.905 on 100 degrees of freedom
## (2 observations deleted due to missingness)
## Multiple R-squared:  0.01398,    Adjusted R-squared:  0.004121
## F-statistic: 1.418 on 1 and 100 DF,  p-value: 0.2366
```

```
cor(possum$age, possum$tail_l)
```

```
## [1] NA
```

```
plot(possum$age, possum$tail_l, pch=19)
abline(possummod3, col="red")
```

