Yiqiao Jin
UID: 305107551
Section 1C

4.2
A. Number of square feet in the home seem to be more related because by observation, there is a stronger positive linear relationship in Plot B than Plot A

B. I will make a better prediction with number of square feet. Since there's a stronger linear correlation between Total value of the house and Number of square feet, we can predicted a more reliable total value through linear regression.

4.6
Professors in 1985 earned the most during the first employment and professors in 2007 earns the least. The negative trend shows that over time, the salary received by professors during their first employment decreases.

4.10
The graph shows very little trend. Points are scattered randomly on the plot, and there is not a sign of linear correlation.

4.14
I may find the correlation, but the correlation is non-linear. The highest fertility rate occurs at the age of about 28.

4.18
A 0.374
B -0.903
C 0.777

4.30
A.
Independent Variable: Mother's Height
Dependent Variable: Daughter's Height
B.
The height is about 62 inches
C.
Height = 29.92+0.5417*60 = 62.42 inches
D.
The slope indicates that for each inch increase in mother's height, we expect daughter's height to increase by 0.5417 inches.
E.
Possible factors: mother's living habits, the average time mothers spend on exercises, the duration of pregnancy

4.32

A. About 330 thousand dollars

B. Predicted price = -11.77+3000*0.1146 = 332.03 thousand dollars

4.34

A. Foot Length = Hand Length*0.998 + 5.67

B. s=0.948*(1.230/1.168)=0.998

C. a=23.318-17.682*0.998 = 5.671

D. Foot Length = 18*0.998 + 5.67 = 23.634 cm

4.40

The correlation is approximately -1 because there is a strong negative linear correlation between the two variables

4.42

	Predictor(x)	Response(y)
a	Weight of gold nuggets	Value of gold nuggets
b	Time since Chlorine is added	Concentration of Chlorine
c	Circumference of oak trees	Age of trees

4.54

A.

Extrapolation occurs when we try to predict the y value with an out-of-range x value using the regression equation

B.

Correlation of determination is the square of correlation efficient.

It shows what percentage of variation in y is explained by x

C.

No. The mere fact that two variables are correlated does not mean that there exists causation between them.

4.68

A.

The trend is positive. It means that as people per household increases, the amount of trash increases.

B.

$r=\sqrt{0.769} = 0.877$

C.

Slope=11.30

It means for each additional person in the house, there are 11.30 additional pounds of trash produced.

D.

The intercept is meaningless. x=0 exceeds the domain of our sample. It's obvious that 0 pounds of trash is produced when there are 0 people in the house.

4.70

The regression is inappropriate because the graph does not indicate that there exist a linear correlation between the two variables. Since the regression line cannot be derived according to the samples, I don't think I can raise the life expectancy by buying more TVs. Correlation does not mean causation.