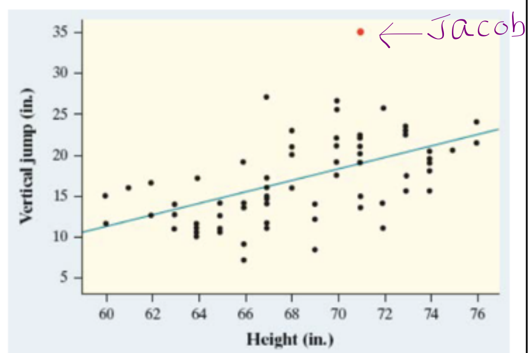
Pick up the Warm Up

- a review of influential points

Who's got hops? Haley, Jeff, and Nathan measured the height (in inches) and vertical jump (in inches) of 74 students at their school.

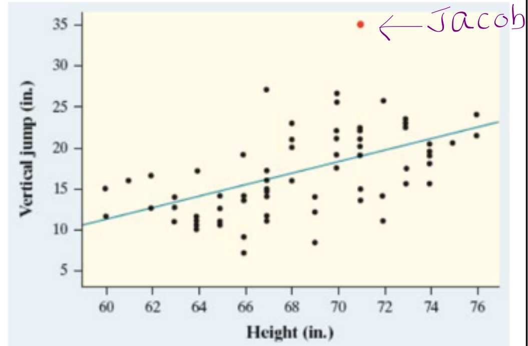
Here is a scatterplot of the data, along with the least-squares regression line. Jacob (highlighted in red) had a vertical jump of nearly 3 feet!



(a) Describe the influence that Jacob's point has on the equation of the least-squares regression line.

Who's got hops? Haley, Jeff, and Nathan measured the height (in inches) and vertical jump (in inches) of 74 students at their school.

Here is a scatterplot of the data, along with the least-squares regression line. Jacob (highlighted in red) had a vertical jump of nearly 3 feet!



(a) Describe the influence that Jacob's point has on the equation of the least-squares regression line.

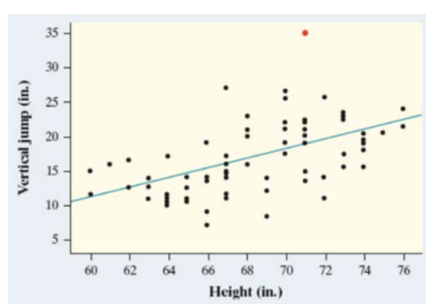
Because Jacob has an above-average height and an above-average vertical jump, his point increases the positive slope of the LSRL and decreases the y-intercept.

(b) Describe the influence that Jacob's point has on the standard deviation of the residuals and r^2.

Jacob's vertical jump is further from the LSRL than the other jumps. Because his point has such a large residual, it increases s.

The value of r^2 decreases

(c) Is Jacob's point an outlier? A high-leverage point? Both? Explain your answer.

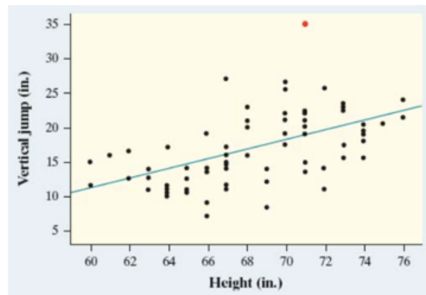


(b) Describe the influence that Jacob's point has on the standard deviation of the residuals and r^2.

Jacob's vertical jump is further from the LSRL than the other jumps. Because his point has such a large residual, it increases s.

The value of r^2 decreases

(c) Is Jacob's point an outlier? A high-leverage point? Both? Explain your answer.



(b) Describe the influence that Jacob's point has on the standard deviation of the residuals and r^2.

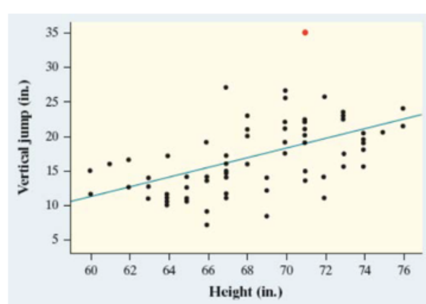
Jacob's vertical jump is further from the LSRL than the other jumps. Because his point has such a large residual, it increases s.

The value of r^2 decreases

(c) Is Jacob's point an outlier? A high-leverage point? Both? Explain your answer.

Jacob's point is an outlier because it has such a large residual (not in the pattern of data).

It is not a high leverage point because because his height is not larger (or smaller) than the heights of others.



Reminders

- You should have completed the Personal Progress Check [MCQ - Unit 1A]

~ you should have been given feedback on each question missed.

~ this first PPC will go in the Practice component of your grade.

Ch. 3 + 12.2 Test this Wednesday

- Submit HW then

- Please write the total score on the front sheet as well

$\frac{\quad}{32}$

Aim

Determine which of several transformations does a better job of producing a linear relationship.

What are the tools we have to help us choose a good prediction equation for a relationship between two quantitative variables?

~ Each team gets a car

~ How far the car moves depends on how far it is pulled back.

~ The challenge will be to get your team's car to go a certain distance.

~ Winner gets a prize, but....

You won't know the distance yet :(

So you will collect data and try to create the best model that you can use once you do find out the race distance. :)

Lesson 12.2 - Day 2

HOW CLOSE CAN YOU GET TO THE FINISH LINE?

definitely use stapplet to view the scatter plot, but

don't spend time sketching the scatter plots on your paper.

If you "see" a linear association, then focus on other things

- like residual plots
- if needed ..... r^2 and s

The Goal: Get your team's car to reach the finish line without going over.
The Catch: The distance will not be revealed until later.

Test drive your car for pull-backs of 2, 4, 6, and 8 inches. Measure the distance the car travels. Repeat this process 3 times. Fill in the table below.

Pull-back Distance (in.)	2	2	2	4	4	4	6	6	6	8	8	8
Distance Traveled (in.)												

You will create three different regression models for the data and will decide which is best. For each of the three models below:

a. Create a scatterplot
b. Calculate an LSRL
c. Analyze linearity using any strategies you choose. Show your work.

~ Each group collects own data.
~ Determines the best model
~ Once all groups have their model, the Race Distance will be announced.

Linear: Plot (x, y):

Quadratic: Plot (x^2, y):

Exponential: Plot (x, $\log y$)

Choosing the Best Regression

① Check the scatter plot for a linear pattern.

Choosing the Best Regression

(1) Check the scatter plot for a linear pattern.
(2) Check Residual plot for random scatter (no leftover pattern)

Choosing the Best Regression

(1) Check the scatter plot for a linear pattern.
(2) Check Residual plot for random scatter (no leftover pattern)

(3) If deciding between several good models, r^2 & S can be used to decide.

Life
insurance

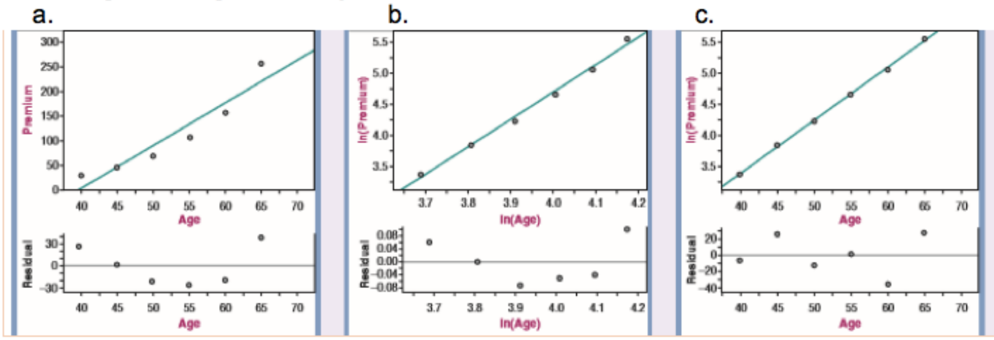
Check Your Understanding:

Age (years)	Monthly premium
40	$29
45	$46
50	$68
55	$106
60	$157
65	$257

Many adults plan ahead for their eventual passing by purchasing life insurance. Many different types of life insurance policies are available. Some provide coverage throughout an individual's life (whole life), while others last only for a specified number of years (term life). The policyholder makes regular payments (premiums) to the insurance company in return for the coverage. The table shows monthly premiums for a 10-year term-life insurance policy worth $1,000,000.

The output shows three possible models for predicting monthly premium from age. Option 1 is based on the original data, while Options 2 and 3 involve transformations of the original data. Each set of output includes a scatterplot with a least-squares regression line added and a residual plot. The equations are given below.

a. b. c.



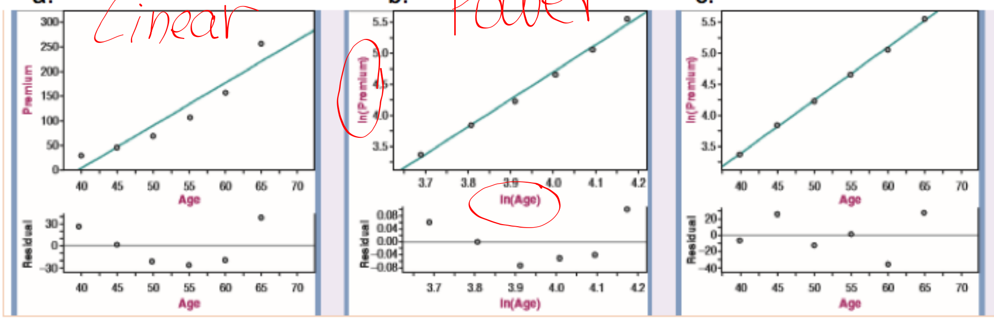
Many different types of life insurance policies are available. Some provide coverage throughout an individual's life (whole life), while others last only for a specified number of years (term life). The policyholder makes regular payments (premiums) to the insurance company in return for the coverage. The table shows monthly premiums for a 10-year term-life insurance policy worth $1,000,000.

40	$29
45	$46
50	$68
55	$106
60	$157
65	$257

ln(y) vs ln(x)

The output shows three possible models for predicting monthly premium from age. Option 1 is based on the original data, while Options 2 and 3 involve transformations of the original data. Each set of output includes a scatterplot with a least-squares regression line added and a residual plot. The equations are given below.

a. Linear b. Power c.



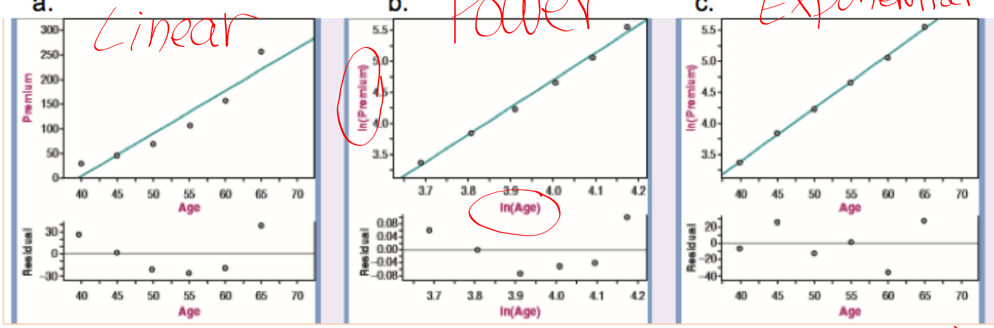
Many different types of life insurance policies are available. Some provide coverage throughout an individual's life (whole life), while others last only for a specified number of years (term life). The policyholder makes regular payments (premiums) to the insurance company in return for the coverage. The table shows monthly premiums for a 10-year term-life insurance policy worth $1,000,000.

40	$29
45	$46
50	$68
55	$106
60	$157
65	$257

ln(y) vs ln(x)

The output shows three possible models for predicting monthly premium from age. Option 1 is based on the original data, while Options 2 and 3 involve transformations of the original data. Each set of output includes a scatterplot with a least-squares regression line added and a residual plot. The equations are given below.

a. Linear b. Power c. Exponential



ln(y) vs. x

1. Use each model to predict how much a 58-year-old would pay for such a policy.

a. $\widehat{premium} = -343 + 8.63(age)$

b. $\ln(\widehat{premium}) = -12.98 + 4.416\ln(age)$

c. $\ln(\widehat{premium}) = -0.063 + 0.0859(age)$

1. Use each model to predict how much a 58-year-old would pay for such a policy.

a. $\widehat{premium} = -343 + 8.63(\underset{58}{age})$ $\approx \$157.54$

b. $\ln(\widehat{premium}) = -12.98 + 4.416\ln(\underset{58}{age})$

c. $\ln(\widehat{premium}) = -0.063 + 0.0859(\underset{58}{age})$

1. Use each model to predict how much a 58-year-old would pay for such a policy.

58

a. $\widehat{premium} = -343 + 8.63(age)$ $\approx \$157.54$

b. $\ln(\widehat{premium}) = -12.98 + 4.416 \ln(age)$

58

$\ln(\text{premium}) = 4.9509$

$\text{premium} = e^{4.9509} \approx \$141.17$

c. $\ln(\widehat{premium}) = -0.063 + 0.0859(age)$

58

1. Use each model to predict how much a 58-year-old would pay for such a policy.

58

a. $\widehat{premium} = -343 + 8.63(age)$ $\approx \$157.54$

b. $\ln(\widehat{premium}) = -12.98 + 4.416 \ln(age)$

58

$\ln(\text{premium}) = 4.9509$

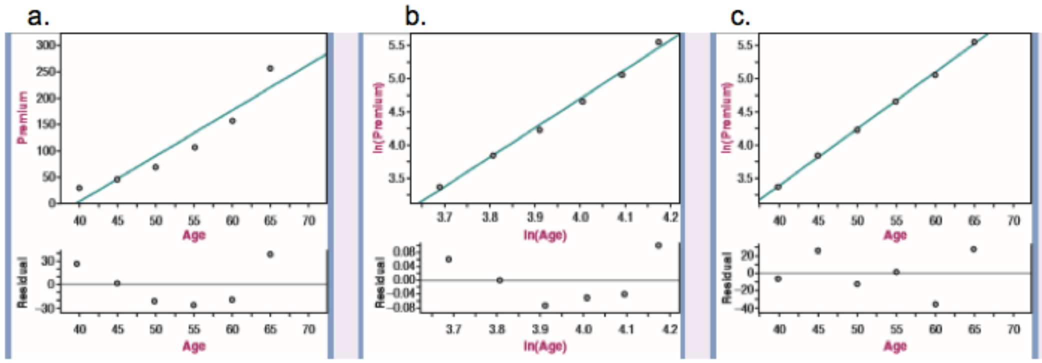
$\text{premium} = e^{4.9509} \approx \$141.17$

c. $\ln(\widehat{premium}) = -0.063 + 0.0859(age)$

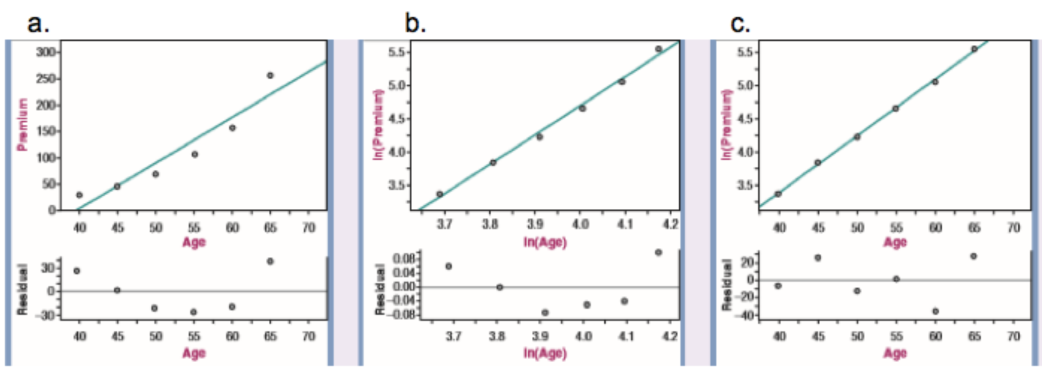
58

$\ln(\text{premium}) = 4.92$

$\text{premium} = e^{4.92} \approx \$137.00$



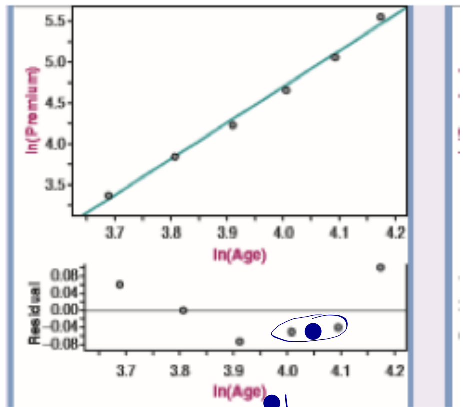
2. What type of function—linear, power, or exponential—best describes the relationship between age and monthly premium? Explain your answer.



2. What type of function—linear, power, or exponential—best describes the relationship between age and monthly premium? Explain your answer.

Exponential - The residual plot shows no leftover pattern.

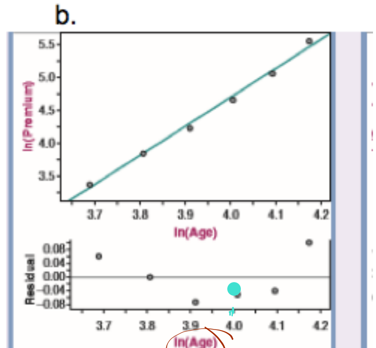
b.



4.06

3. If you used the **Power Model** (b), do you expect your prediction to be too large, too small, or about right. Justify your answer.

b.



3. If you used the **Power Model** (b), do you expect your prediction to be too large, too small, or about right. Justify your answer.

$\ln(58) = 4.0$

We would expect the predicted premium to be too high.

LCQ

Quick Lottery first

12.2.....43, 45, 47, 51-54

study pp. 803-810

and do the FRAPPY! to save time tomorrow.