**Best Fit Line**

You have learned how to find and write equations for lines of fit by hand. Many calculators use complex algorithms that find a more precise line of fit called the best-fit line.

One algorithm is called linear regression. We can find the linear regression.

STAT

To enter the data: EDIT L1 is independent variable; L2 is dependent variable

Enter

VARS

STAT

Calculator Steps: CALC 4: LinReg Y-VARS FUNCTION 1: Y1

Your calculator may also compute a number called the correlation coefficient. This number will tell you if your correlation is positive or negative and how closely the equation is modeling the data. The closer the correlation coefficient is to 1 or -1, the more closely the equation models the data.

Enter

0

2nd

To turn the correlation coefficient on: DiagnosticOn

* If the correlation coefficient is close to 1 or -1, the fit is .
* The farther away from 1 or -1, the the fit.
* If the scatterplot appears random, there is .
* If the correlation coefficient is positive, the slope will be .
* If the correlation coefficient is negative, the slope will be .

We will often need to interpret the slope and y-intercept in the context of the problem.

Slope can often follow this pattern:

* (Topic of data) (increases/decreases)
* (slope) (y-units) per (x-units).

The y-intercept is the starting value, or what the dependent variable is when the independent variable is 0.

**EXAMPLE:** The average lifespan of American women has been tracked, and the model for the data is *y* = 0.2*t* + 73, where *t* = 0 corresponds to years since 1960.

**INTERPRETATION of slope and y-intercept:**

**Real World Example 1: Box Office**

The table shows the amount of money made by movies in the United States. Use a graphing calculator to write an equation for the best-fit line for that data.



1. Enter the data into a list using the graphing calculator.

* Let x = the number of years after 2000.

2. Find the best fit line using the graphing calculator.

r = Describe the fit.

Interpret the slope.

Interpret the y-intercept.

3. EXTRAPOLATION: Use the equation and the table in the graphing calculator to predict what the box office income will be in 2013. State your answer as a complete sentence.

4. INTERPOLATION: Use the equation and the table in the graphing calculator to predict what the expected box office income was in 2008. How does the compare to the actual box office income given in the table? What is the difference?

In a college meal plan, you pay a membership fee; then all your meals are at a fixed price per meal.

1. If 30 meals cost $152.50 and 60 meals cost $250, find the membership fee and price per meal.
2. Write a formula for the cost of a meal plan, C, in terms of the number of meals, n.
3. Find the cost for 50 meals.
4. Determine the maximum number of meals you can buy on a budget of $300.

5. 6.

****

1. Equation
2. Interpret slope
3. Interpret y-int

**AFM Hwk: Regression**

**Put your answers on another sheet of paper.**

**The table below shows the number of people participating in high school athletics.**

1. Find the best-fit line.

2. Interpret the slope and y-intercept.

3. Predict the number of participants in 1988. State your answer as a complete sentence. Is this an interpolation or an extrapolation? Explain why.

4. Predict the number of athletes in 2013. State your answer as a complete sentence. Is this an interpolation or an extrapolation? Explain why.

**Interpret the slope and y-intercept for each real-world situation for 5 – 9.**

|  |  |
| --- | --- |
| **5.** . The function, $h\left(x\right)=2.90+0.79x$, models the cost of a hamburger with varying numbers of toppings.  | **6.** The height of a candle, in inches, as a function of time, in hours, when burning is modeled by $h\left(t\right)=-0.2t+12$. |
| **C:\Users\Jodie Kinkaid\AppData\Roaming\PixelMetrics\CaptureWiz\Temp\91.png7.**  | **C:\Users\Jodie Kinkaid\AppData\Roaming\PixelMetrics\CaptureWiz\Temp\92.png8.** Equation:Slope:Y-intercept:Given an example of an extrapolation |
| **9.** Margarita is hired by an accounting firm at a salary of $60,000 per year. Three years later her annual salary has increased to $70,500. Assume her salary increases linearly.  a) Find an equation that relates her annual salary S and the number of years, t, that she has worked on the firm. b) What do the slope and S-intercept of her salary equation represent? c) What will her salary be after 12 years with the firm? d) If her salary continues to grow linearly, in how many years would she have to work there to have an income of  $100,000? |

**I. Speed vs MPG**

An engineer collects data showing the speed of a Ford Taurus and its average miles per gallon.

|  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Speed | 30 | 35 | 40 | 40 | 45 | 50 | 55 | 60 | 65 | 65 | 70 |
| MPG | 18 | 20 | 23 | 25 | 25 | 28 | 30 | 29 | 26 | 25 | 25 |

1. Create a scatter plot on your calculator.
2. What type of relation exists (linear or quadratic)? Explain your answer with at least 2 explanations!
3. Based on your decision from number 2, write the equation of best fit for this data.
4. Use the function to predict mpg for a speed of 63 mph. Is this an extrapolation or interpolation? Explain.
5. What speed would get you 24 mpg?

**II. Exponential Function**:, where *a* is a nonzero constant, *b* is greater than 0 but not equal to 1, and *x* is a

 real number. The value of “*b*” determines whether it is a growth or decay. If 0 < *b* < 1, it

 is an exponential decay, if *b* > 1, then it is an exponential growth.

1. An equation that models the population of Washington D.C. since 1990 is. Explain the coefficients. If the trend continues, predict the population in 2010. When would the population be half of what it was in 1990?
* “a” value:
* “b” value:
* When:

The table below models the growth of the cable television industry. Let *x* represent the years since 1965. Use the table to find the best fitting function (linear or exponential) for the data. Answer the questions below.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Year | 1965 | 1970 | 1975 | 1980 | 1985 |
| Cable TV Subscribers (in millions) | 1.5 | 5.1 | 9.8 | 17.7 | 39.9 |

1. Linear Regression and Exponential Regression
2. Explain which function better models the data
3. Explain the coefficients from the best fit line.
4. According to the best fit model, how many subscribers should there be in 2000? Interpolation or extrapolation?
5. According to this model, when will there be 60 million subscribers?
6. Can we expect the growth in this industry to continue? Why or why not?

An additive to puppy food is shown to increase weight gain in underweight puppies when it is mixed with standard food. What mixture results in highest weight gain for the average puppy? Data are collected from a study of eight puppies fed with different percentages of additive.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Percentage additive (x) | .20 | .20 | .40 | .40 | .60 | .60 | .80 | .80 |
| Weight gain (kg) (y) | 4.1 | 6.2 | 6.5 | 7.3 | 3.1 | 4.8 | 0.5 | 1.2 |

1. Find quadratic and cubic models for these data.
2. Use each model to find the predicted percentage that produces the greatest weight gain.
3. How much difference is there in each of these predictions?
4. The table gives the number ***y*** (in millions) of cell phone subscribers from 1988 to 1997 where ***t*** is the number of years since 1987.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Years** | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
| **Cell phone subscribers (in millions)** | 1.6 | 2.7 | 4.4 | 6.4 | 8.9 | 13.1 | 19.3 | 28.2 | 32.8 | 48.7 |

1. Would linear or exponential be the best fit model? Why?
2. State best fit model for this data.
3. What do the values in your best fit model represent in context of this data? (You need two detailed sentences)
4. Predict the number of cell phone subscribers in the year 2003. Is this an extrapolation or interpolation?
5. In what year will there be 165 million subscribers?
6. **Projectile Motion Function**

The height of an object rising or falling under the influence of gravity is modeled by the function, , where *x* represents time in seconds, *y* represents the object’s height from the ground in meters or feet, *a* is half the downward acceleration due to gravity (on Earth, *a* is -4.9m/ss or -16 ft/ss), *v0* is the initial upward velocity of the object in meters per second or feet per second, and *s0* is the initial height of the object in meters or feet.

An object is projected upward, and the data is collected below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Time (s) | 1 | 2 | 3 | 4 | 5 | 6 |
| Height (m) | 120.1 | 205.4 | 280.9 | 346.6 | 402.4 | 448.4 |

1. Would a linear or quadratic model best fit this data? Why?
2. Explain what the coefficients mean in context of this data.
3. When does the object reach its maximum height?
4. What is the maximum height?
5. When does the object reach the ground?
6. When does the object reach 500 m? 600 m? Explain.
7. Give an example of an extrapolation.

**The data in the following table list natural gas consumption ( in quadrillion BTU) in the United States.**

**Natural Gas Consumption**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year** | **1960** | **1970** | **1980** | **1990** | **1997** |
| **Consumption** | **12.4** | **21.8** | **20.4** | **19.3** | **22.6** |

1. Which model best represents this data? Record this model.
2. Use this model to predict the natural gas consumption in 1982.

**Regression Practice**

1. The following data was obtained by throwing a rubber ball at a CBR.

|  |  |
| --- | --- |
| Time (sec) | Height (m) |
| 0.0000 | 1.03754 |
| 0.1080 | 1.40205 |
| 0.2150 | 1.63806 |
| 0.3225 | 1.77412 |
| 0.4300 | 1.80392 |
| 0.5375 | 1.71522 |
| 0.6450 | 1.50942 |
| 0.7525 | 1.21410 |
| 0.8600 | 0.83173 |

1. Use the data above to make a scatterplot, letting *x* represent the number of seconds elapsed.
2. Next, use a graphing calculator to find the model that best expresses the height and vertical velocity of the rubber ball. We can also use this model to predict the maximum height of the ball and its vertical velocity when it hits the face of the CBR.
3. Fit linear, quadratic, and exponential functions to the data. By comparing the values of, determine the function that best fits the data.
4. Graph the function of best fit with the scatterplot of the data.
5. Determine the maximum height of the ball (in meters).
6. With the model you selected in part (b), predict when the height of the ball is *at least* 1.5 meters.
7. *Projected Number of Alzheimer’s Patients: German psychiatrist Alois Alzheimer first described the disease, later called Alzheimer’s disease, in 1906. Since life expectancy has significantly increased in the last century, the number of Alzheimer’s patients has increased dramatically. The number of patients in the United States reached 4 million in 2000. The following table lists projected data regarding the number of Alzheimer’s patients in years beyond 2000.*

|  |  |
| --- | --- |
| Year, *x* | Projected Number of Alzheimer’s Patients in the United States (In millions) |
| 2000 | 4.0 |
| 2010 | 5.8 |
| 2020 | 6.8 |
| 2030 | 8.7 |
| 2040 | 11.8 |
| 2050 | 14.3 |

1. Draw a scatter plot of the data.
2. Fit linear, exponential, power, logistic and logarithmic functions to the data. By comparing the values of, determine the function that best fits the data.
3. Superimpose the regression curve on the scatter plot.
4. Use the regression model to estimate the number of Alzheimer’s patients in 2005, 2025, and 2100.
5. *Stopping Distance* A state highway patrol safety division collected the data on stopping distances in Table 2.16.
6. Draw a scatter plot of the data.
7. Fit linear, quadratic, and exponential, functions to the data. By comparing the values of, determine the function that best fits the data.
8. Superimpose the regression curve on the scatter plot.
9. Use the regression model to predict the stopping distance for a vehicle traveling at 25 mph.
10. Use the regression model to predict the speed of a car if the stopping distance is 300 ft.

Table 2.16 Highway Safety Division

|  |  |
| --- | --- |
| Speed (mph) | Stopping Distance (ft) |
| 10 | 15.1 |
| 20 | 39.9 |
| 30 | 75.2 |
| 40 | 120.5 |
| 50 | 175.9 |

1. *Home Schooling Growth* The estimated number of U.S. children that were home-schooled in the years from 1992 to 1997 were:

Table 1.13 Home Schooling

|  |  |
| --- | --- |
| Year | Number |
| 1992 | 703,000 |
| 1993 | 808,000 |
| 1994 | 929,000 |
| 1995 | 1,060,000 |
| 1996 | 1,220,000 |
| 1997 | 1,347,000 |

1. Produce a scatter plot of the number of children home-schooled in thousands (*y)* as a function of years since 1990 (*x*).
2. Find the linear regression equation. (Round the coefficients to the nearest 0.01.)
3. Does the value of  suggest that the linear model is appropriate?
4. Find the quadratic regression equation. (Round the coefficients to the nearest 0.01.)
5. Does the value of  suggest that a quadratic model is appropriate?
6. Use both curves to predict the number of U.S. children that are home-schooled in the year 2005. How different are the estimates?
7. *Writing to Learn*  Use the results of this exploration to explain why it is risky to use regression equations to predict *y-*values for *x* values that are not very close to the data points, even when the curves fit the data points very well.
8. In the years before the Civil War, the population of the United States grew rapidly, as shown in the following table from the U.S. Bureau of the Census.

|  |  |
| --- | --- |
| Year | Population in Millions |
| 1790 | 3.93 |
| 1800 | 5.31 |
| 1810 | 7.24 |
| 1820 | 9.64 |
| 1830 | 12.86 |
| 1840 | 17.07 |
| 1850 | 23.19 |
| 1860 | 31.44 |

1. Draw a scatter plot of the data.
2. Fit linear, quadratic, and exponential. By comparing the values of, determine the function that best fits the data.
3. Superimpose the best regression curve on the scatter plot.
4. Use the regression model to predict the population in 1870.
5. Use the regression model to predict the population in 1930. Explain why/why not you feel this prediction has validity. (Hint: you may want to complete this problem after you finish the problem dealing with Census records after the Civil War.)

AFM Hwk 9 – 3 Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

1. The table gives the number ***y*** (in millions) of cell phone subscribers from 1988 to 1997 where ***t*** is the number of years since 1987.

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Years** | 1988 | 1989 | 1990 | 1991 | 1992 | 1993 | 1994 | 1995 | 1996 | 1997 |
| **Cell phone subscribers (in millions)** | 1.6 | 2.7 | 4.4 | 6.4 | 8.9 | 13.1 | 19.3 | 28.2 | 32.8 | 48.7 |

1. Would linear or exponential be the best fit model? Why?
2. State best fit model for this data.
3. What do the values in your best fit model represent in context of this data? (You need two detailed sentences)
4. Predict the number of cell phone subscribers in the year 2003. Is this an extrapolation or interpolation?
5. In what year will there be 165 million subscribers?
6. **Projectile Motion Function**

The height of an object rising or falling under the influence of gravity is modeled by the function, , where *x* represents time in seconds, *y* represents the object’s height from the ground in meters or feet, *a* is half the downward acceleration due to gravity (on Earth, *a* is -4.9m/ss or -16 ft/ss), *v0* is the initial upward velocity of the object in meters per second or feet per second, and *s0* is the initial height of the object in meters or feet.

An object is projected upward, and the data is collected below.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Time (s) | 1 | 2 | 3 | 4 | 5 | 6 |
| Height (m) | 120.1 | 205.4 | 280.9 | 346.6 | 402.4 | 448.4 |

1. Would a linear or quadratic model best fit this data? Why?
2. Explain what the coefficients mean in context of this data.
3. When does the object reach its maximum height?
4. What is the maximum height?
5. When does the object reach the ground?
6. When does the object reach 500 m? 600 m? Explain.
7. Give an example of an extrapolation.

**The data in the following table list natural gas consumption ( in quadrillion BTU) in the United States.**

**Natural Gas Consumption**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Year** | **1960** | **1970** | **1980** | **1990** | **1997** |
| **Consumption** | **12.4** | **21.8** | **20.4** | **19.3** | **22.6** |

1. Which model best represents this data? Record this model.
2. Use this model to predict the natural gas consumption in 1982.

AFM Hwk 9 – 4 Power Function Model Name \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

|  |  |
| --- | --- |
| Diameter (mm) | Length (mm) |
| **17.6** | **159.9** |
| **26.0** | **206.9** |
| **31.9** | **236.8** |
| **38.9** | **269.9** |
| **45.8** | **300.6** |
| **51.2** | **323.6** |
| **58.1** | **351.7** |
| **64.7** | **377.6** |
| **66.7** | **384.1** |
| **80.8** | **437.2** |
| **82.9** | **444.7** |

**Data:** The data below represents the length and mid-shaft diameters of the humerus bones of African Antelopes.

The true antelopes are found only in Africa and Asia.  They range in size from 12" (30 cm. at the shoulder) pygmy antelopes to giant elands, which are over 6 feet tall (180 cm) at the shoulder.  Most antelopes are between 3 to 4 feet tall (90-120 cm) at the shoulder. The horns of antelopes, unlike the antlers of deer, are un-branched, are made of a shell with a bony core, and are not shed.  The majority of antelopes reside in Africa.



|  |  |
| --- | --- |
| **Task:** | Express answers to the *nearest hundredth*. |
|   | 1.) | Prepare a scatter plot of the data on your calculator. |
|   | 2.) | Determine a power regression model equation to represent this data. Record this. |
|   | 3.) | Decide whether the new equation is a "good fit" to represent this data. Explain. |
|   | 4.) | *Extrapolation:*  What length will correspond to a diameter of 84 mm? |
|   | 5.) | *Interpolation:* What length will correspond to a diameter of 47 mm? |
|   | 6.) | What mid-shaft diameter will correspond to a length of 305.7 mm? |

7. Write the equation of the power function that passes through the points (1, 2) and (5, 250).

8. Based on the function from #2, what is the value of x when y = 4?

**Sinusoidal Regression**