1.3 – Linear Regression

(The Least Squares Line)

Math 111

Warnock - Class Notes

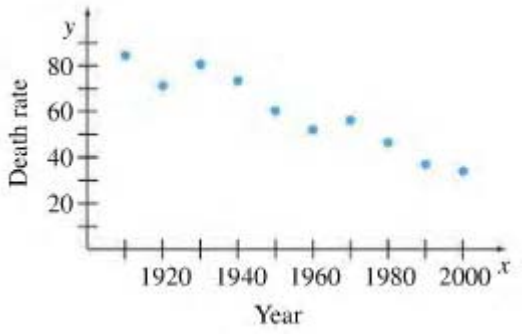
In this textbook, they talk about the "Least Squares Line" but we're going to cut to the chase and use _____ _____.

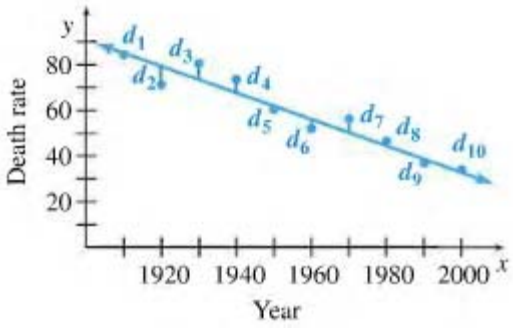
There are enough things for us to memorize this quarter, so we're not going to memorize or use the complicated Least Squares Formulas.

Here is an example of some data that we can model with a linear function.

(per 100,000 in the U.S.)

Accidental Death Rate	
Year	**Death Rate**
1910	84.4
1920	71.2
1930	80.5
1940	73.4
1950	60.3
1960	52.1
1970	56.2
1980	46.5
1990	36.9
2000	34.0

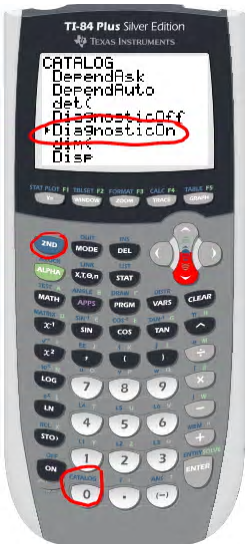




In your graphing calculator, you need to first turn DIAGNOSTICS to ON (so we see "r")

It is under CATALOG (2^{nd} - 0)

- The ALPHA D
- Or just scroll down to DiagnosticsOn and hit ENTER



Before we continue, **QUIT (2^{nd} MODE)** *(I don't like the terminology!) will exit you from a window you don't want to be in anymore.]*

1. **Enter** the **Data**
    - **STAT: EDIT**
      Enter the *x* data (be careful if dealing with years) in L_1 and the *y* data in L_2. Make sure your lists are of the same length, and the data points are correct.

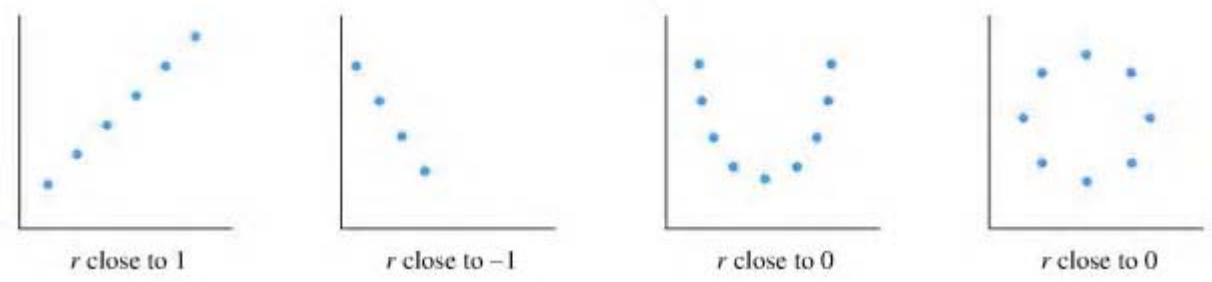
2. **Graph** the **Data** points

- **STAT PLOT** (**2nd Y=**)
  Turn on the first stat plot. (Make sure Xlist is L_1 and Ylist is L_2. It should already be set this way, and once you set it, you shouldn't have to do it again.)
- You can set the window *manually*, or
- **ZOOM: ZOOMSTAT** (menu item #9)
  This will give you exactly your points that you are plotting, so this is a good starting point. However, we will want to predict points beyond that, so now you can go to **WINDOW** and just expand beyond what has already been given to you.
- **GRAPH**

3. Finding/graphing a **Linear Regression** equation
    - In the main calculator window (**QUIT** from the graphing section)
    - **STAT → CALC: LinReg (ax+b)** (menu item #4)
    - After you select this command, type **Y1 –** this will store your equation in the Y1 slot for graphing.
        - **Y_1** is kind of interesting to find, so here it is
            - **VARS→Y-VARS: Function: Y_1**
    - If you have a newer version of the calculator, you're going to type Y1 after "Store RegEQ:"
    - At this point, you must hit **ENTER** to tell the calculator to actually take the Linear Regression or "Calculate" on a newer version.
    - Too see the linear regression equation – **Y=**
    - Too see the graph of the linear regression equation – **GRAPH**

4. Evaluating new data, using your Linear Regression Equation
    - **TRACE**
    - You must now hit up or down, so that you are tracing the linear regression, and not the points. Otherwise, it will exit you to the main screen if you try to type.
    - Once you see **Y_1** in the upper left corner, now you can type in your data to find the information you need. Be careful if you are dealing with years to know when $x = 0$ is.
    - If you get an error message, make sure that your window is large enough to see the points you are trying to look at.

The "r" value that you see is called the _____ _____. This measures the strength of the linear relationship between to two variables. The closer these values are to 1 or -1, the more linear the data is.



#1. The percent of female smokers 18 years and older has decreased from 1965 to 2002.

Years since 1960	**5**	**14**	**25**	**30**	**35**	**40**	**42**
Percent of Female Smokers	**33.9**	**32.1**	**27.9**	**22.8**	**22.6**	**21**	**19.1**

a) Use a linear regression to fit a line to the data in the table. Write the linear regression equation. Round to three decimal places.

b) Predict the percent at which female smokers are decreasing per year.

c) What is the Correlation Coefficient?

d) Predict the percent of female smokers in the year 2014 if the trend continues to decrease at the same rate. Round to the nearest tenth of a percent.

e) At what year will the percent of female smokers be zero? Is this reasonable?

#2. The Academy Awards is one of the most watched television shows. The table below gives the average cost of a 30-second commercial slot for the years 1998 through 2003.

Years since 1990	**8**	**9**	**10**	**11**	**12**	**13**
Academy Award Commercial Cost (in millions of dollars)	**0.9**	**1.0**	**1.2**	**1.35**	**1.25**	**1.4**

a) Find the linear regression equation for the given data. Round to three decimal places.

b) What is the correlation coefficient?

c) Why do you think there was a drop between 2001 and 2002?

d) If the rate continues to increase in the same manner, predict the cost of a 30-second commercial in the year 2014.

e) What was the actual cost of a 30-second commercial in 2014?

f) If the rate continues, what year will the cost of a 30-second commercial cost over $3,000,000? Do you expect this?