

# Chapter 7: Scatterplots and Correlation

1. We use the word 'correlation' to refer to what kind of association?

**LINEAR**

2. T/F: If two variables have a high correlation, then one must cause the other.

**FALSE**

Data Set:

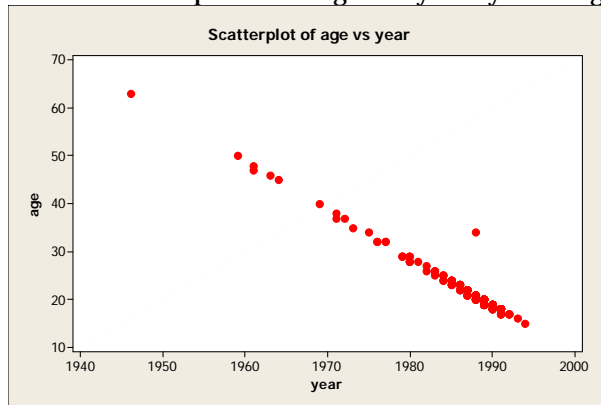
- Fall 2009 COC Math 140 Survey Results: height column fixed (suspicious observations deleted)
- <http://www.canyons.edu/faculty/morrowa/140/datasets/>
- Note: Survey questions are online at <http://classes.thegradekeeper.com/math140.php> (if you need a reminder).

3. Age and Year

a) What relationship would we expect to be between age and year of birth?

**VERY STRONG LINEAR (NEGATIVE)**

b) Examine the relationship between age and year by creating a scatterplot.



c) The plot doesn't meet my expectation--why?

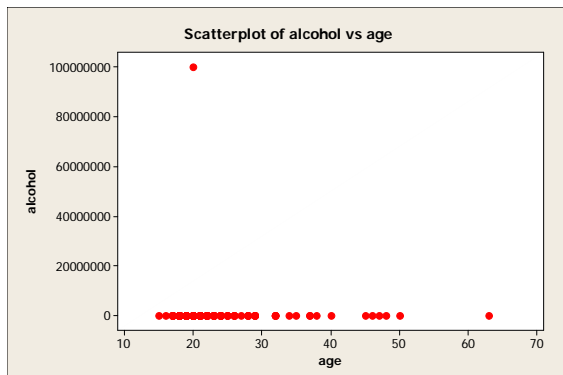
**THERE IS ONE POINT THAT DOES NOT FALL ON THE LINE. THEREFORE, THERE MUST BE A MISTAKE ON DATA ENTRY.**

d) Hold your mouse over the outlier in the graph to discover which observation/row it corresponds to, and then remove that point.

**OBSERVATION 176 (YEAR 1988, AGE 34)**

4. Age and Alcohol

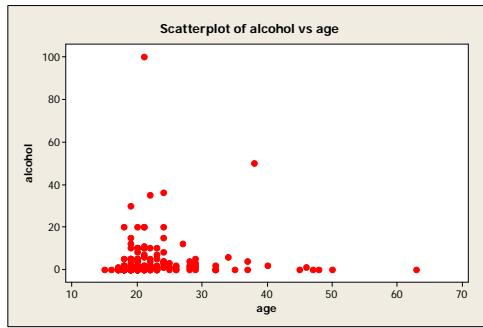
a) Upon examining a scatterplot for age and alcohol, a non-statistician claims that the two variables are correlated. Provide the scatterplot and explain why his conclusion was wrong.



**THE APPEARANCE OF A LINEAR RELATIONSHIP ISN'T REAL. THE VALUE AT 100,000,000 IS AN OUTLIER. WE MUST EXAMINE THE DATASET WITHOUT IT.**

**NOT ONLY IS THERE AN OUTLIER, THE APPARENT LINEAR RELATIONSHIP IS NOT, IN FACT, A RELATIONSHIP (ALCOHOL DOES NOT CHANGE AS AGE CHANGES).**

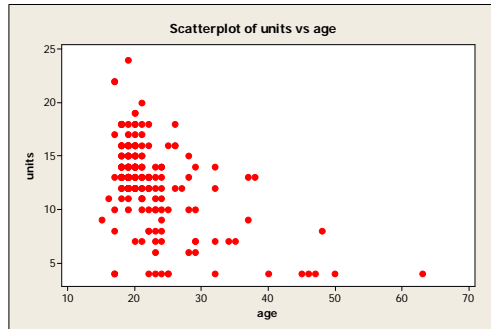
b) Now use the same procedure as above to remove the suspicious observation and redraw the scatterplot. Now what do you conclude?



NOW THERE DEFINITELY DOESN'T APPEAR TO BE A LINEAR RELATIONSHIP

5. Age and Units

a) Create a scatterplot for age and units. Describe the relationship.



NO APPARENT RELATIONSHIP

b) is it appropriate to find r? why or why not?

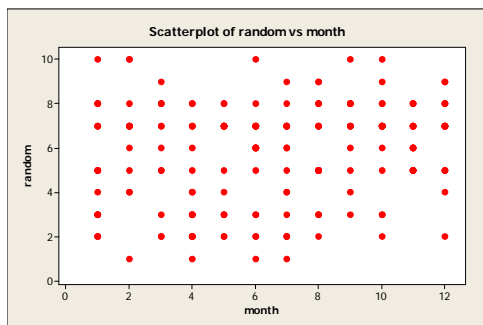
NO, NO LINEAR RELATIONSHIP

6. Month and Random

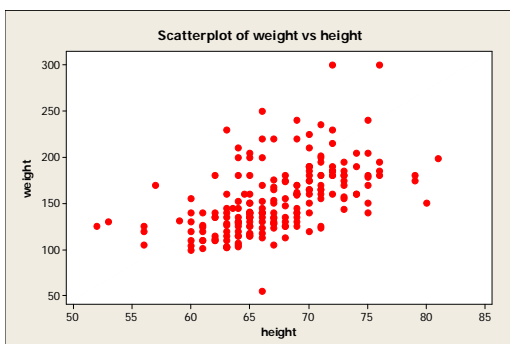
a) Even before we create the scatterplot, what would you expect it to look like?

TOTALLY SCATTERED – THERE SHOULD BE NO RELATIONSHIP BETWEEN THE MONTH SOMEONE IS BORN IN AND THE NUMBER THEY CHOSE

b) Now check your guess by creating the scatterplot.



7. Examine the relationship between height and weight. Find r.



THERE APPEARS TO BE A LOOSE POSITIVE LINEAR RELATIONSHIP. THERE DOES NOT APPEAR TO BE ANY QUESTIONABLE OUTLIERS.

R = 0.540