

**UMUC Math-107**  
**Linear Modeling Project**  
**Brian Killough (Instructor)**

**Linear Data**

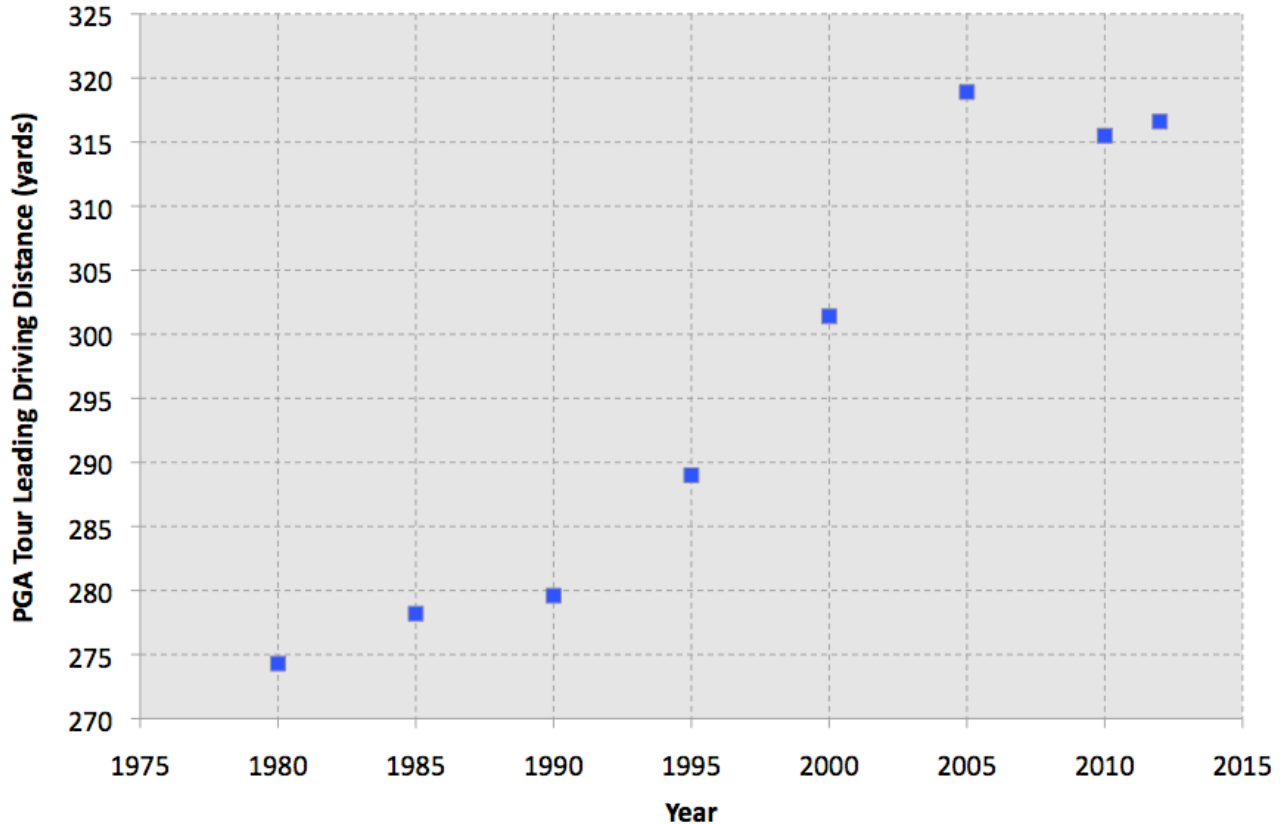
As a golfer, I decided that my project would examine the maximum driving distance of professional golfers over the past 30 years. I believe that the combination of enhanced equipment and improved physical fitness has created a linear increase in driving distance over 30 years. My data was obtained at the PGA Tour website ([www.pgatour.com](http://www.pgatour.com)).

Year	Distance
1980	274.3
1985	278.2
1990	279.6
1995	289.0
2000	301.4
2005	318.9
2010	315.5
2012	316.6



**Scatterplot**

**PGA Tour Driving Distance (1980 to 2012)**



The scatterplot above was created using EXCEL. I chose the YEAR on my x-axis and the leading (or maximum) driving distance on my y-axis. It is easy to see that that maximum driving distance is increasing over time, so we should see a POSITIVE slope for the linear model.

### **Linear Model**

I was able to use EXCEL to calculate the linear curve-fit or regression line for this data. The equation for this line is ...

$$Y = 1.545 (X) - 2790$$

The slope of my regression line is positive (1.545). This basically means that the average driving distance is rising 1.545 yards per year, or 15 yards every 10 years. If I use this model to forecast the future, in the year X=2100 the maximum PGA Tour driving distance will be  $Y = 1.545 (2100) - 2790 = 454$  yards !!! That is very hard to believe, but the data over the past 30 years has this trend.

### **Correlation Coefficient**

I used EXCEL to calculate my correlation coefficient of  $r=0.962$ . This result is positive due to the positive slope of the data and is a "very strong" relationship since it is very close to  $r=1.0$  (perfect linear fit). Below is a graph of the linear model (RED) and the raw data (BLUE squares).

