Shannon McNish
MSAN 502 - Homework 3
7/30/17

I wrote a function called bestfit that takes data and finds the best fit line in the sense of least squares. The data that is fed to this function must have the independent and dependent columns labeled as x and y. The function plot the x and y variables to check that there is a linear relationship. It also print the value of the slope and intercept values using statsmodel.ols to check my model against. The return values of the function are intercept of the best fit line, slope of the best fit line, and the sum of the squares of the distances from the observed y-values to the y-values on the best fit line.

I used the three provided text files, as well as one I found on my own which plots the mean average winter temp against years.

Here are the results:

Data	**X and Y**	**Intercept**	**Slope**	**Sum of Squares**
TV life	Life Expectancy (y) as a function of Televisions per Thousand People (x)	57.34	0.032	1417.89
Population	National Population (y) as a function of Year (x)	-3,740,767,381.46	2,002,587.09	5,282,560,399,260,044.0
NBA	Team Winning Percentage (WinPct) (y) as a function of PM (x)	0.50	0.032	0.046
NY average winter temperature	Average winter temperature (y) as a function of year (x)	-22.29	0.023	901.76