Simple Linear Regression Practice - 1
Stat 3115, Spring 2015

1. The accompanying data was read from a graph that appeared in the article "Reactions on Painted Steel Under the Influence of Sodium Chloride, and Combinations Thereof" (*Ind. Engr: Chem. Prod. Res. Dev.,* 1985: 374-378). The independent variable is SO_2 deposition rate ($mg/m^2/d$), and the dependent variable is steel weight loss (g/m^2.)

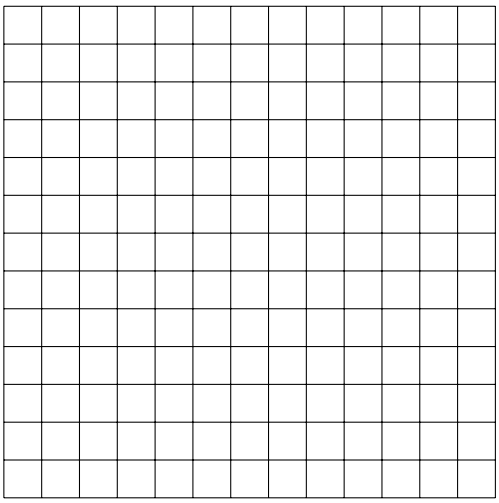
x	14	18	40	43	45	112
y	280	350	470	500	560	1200

   (a) Construct a scatter plot. Does the simple linear regression model appear to be reasonable in this situation?

   (b) Calculate the equation of the estimated regression line.

   (c) What percentage of observed variation in steel weight loss can be attributed to the model relationship in combination with variation in deposition rate?

   For your convenience, a grid has been supplied to help with scatter plot construction.



2. When anthropologists analyze human skeletal remains, an important piece of information is living stature. Often model skeletons are commonly based on statistical methods that utilize measurements on small bones. The following data was presented in the American Journal of Physical Anthropology to validate such a method.

Metacarpal Length (cm)	45	51	39	41	48	49	46	43	47
Stature (cm)	171	178	157	163	172	183	173	175	173

(a) Construct a scatter plot. Does the simple linear regression model appear to be reasonable in this situation?

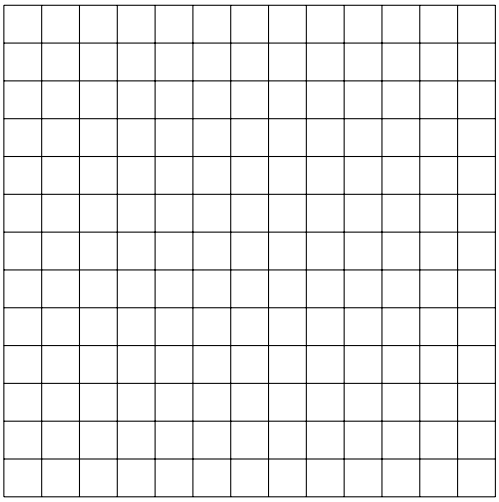
(b) Calculate the equation of the estimated regression line.

(c) What percentage of stature can be attributed to the model relationship in combination with the metacarpal length? In your opinion, is it fair to estimate living height via metacarpal length?

(d) At what point(s), if any, does the model fail to make physical sense?

(e) Give the predicted stature of an individual with metacarpal length of 30 cm.

For your convenience, a grid has been supplied to help with scatter plot construction.



3. Since World War II, plutonium for use in atomic weapons has been produced at an Atomic Energy Commission facility in Hanford, Washington. One of the major safety problems encountered there has been the storage of radioactive wastes. Over the years, significant quantities of these substances—including strontium 90 and cesium 137—have leaked from their open-pit storage areas into the nearby Columbia River, which flows along the Washington-Oregon border and eventually empties into the Pacific Ocean.

   To measure the health consequences of this contamination, and index of exposure was calculated for each of the nine Oregon counties having frontage either on the Columbia River or the Pacific Ocean. This particular index was based on several factors, including the county's stream distance from Hanford and the average distance of its population from any wanter frontage. The cancer mortality rate per 100,000 man-years was also determined for each of these same counties.

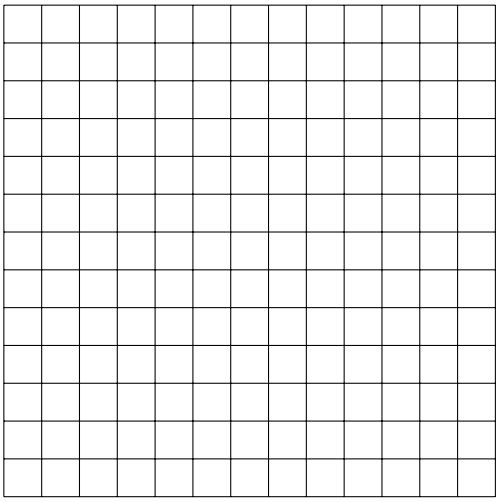
County	Exposure	Morality
Umatilla	2.49	147.1
Morrow	2.57	130.1
Gilliam	3.41	129.9
Sherman	1.25	113.5
Wasco	1.62	137.5
Hood River	3.83	162.3
Columbia	6.41	177.9
Clatsop	8.34	210
Portland	11.64	207.5

   (a) Construct a scatter plot. Does the simple linear regression model appear to be reasonable in this situation?

   (b) Calculate the equation of the estimated regression line.

   (c) What percentage of cancer mortality can be attributed to the model relationship in combination with the exposure level? In your opinion, is it likely that the radiation exposure increased cancer mortality among the population?

   For your convenience, a grid has been supplied to help with scatter plot construction.



4. The largest x values in both Problem 1 and 3 are significantly larger than the rest. These values may have been very influential in determining the equation of the estimated line. Delete these observations and recalculate the two equations. Do they appear substantially different from the originals? Consider predicted values.