ASSIGNMENT NO.1

1). A manufacturer says the Z-Phone smartphone has a mean consumer life of 42 months with a standard deviation of 8 months. Assuming a normal distribution, what is the probability that a given random Z-Phone will last between 20 and 30 months?

2). An experiment to investigate the survival time in hours of an electronic component consists of placing the parts in a test cell and running them for 100 hours under elevated temperature conditions. (This is called an "accelerated" life test.) Eight components were tested with the following resulting failure times:

75, 63, 100+, 36, 51, 45, 80, 90

The observation 100+ indicates that the unit still functioned at 100 hours. Is there any meaningful measure of location that can be calculated for these data? What is its numerical value?

3). On average, do people gain weight as they age? Based on a dataset of 250 samples, some summary statistics for both age (x) and weight (y) are:

$$\sum_{i=1}^{n} x_i = 11211.00 \qquad \sum_{i=1}^{n} y_i = 44520.80 \qquad \sum_{i=1}^{n} x_i y_i = 1996904.15$$

$$\sum_{i=1}^{n} x_i^2 = 543503.00 \qquad \sum_{i=1}^{n} y_i^2 = 8110405.02$$

Assume that the two variables are related according to the simple linear regression model.
a. Calculate the least squares estimates of the slope and intercept.

b. Use the equation of the fitted line to predict the weight that would be observed, on average, for a man who is 25 years old.

c. Suppose that the observed weight of a 25-year-old man is 170 lbs. Find the residual for that observation.

d. Was the prediction for the 25-year-old in part (c) an overestimate or underestimate? Explain briefly.

4). A gasoline manufacturer is investigating the "cold start ignition time" of an automobile engine. The following times (in seconds) were obtained for a test vehicle: 1.75, 1.92, 2.62, 2.35, 3.09, 3.15, 2.53, 1.91. Calculate the sample mean, sample variance, and sample standard deviation. Construct a box plot of the data. A second formulation of the gasoline was tested in the same vehicle, with the following times (in seconds): 1.83, 1.99, 3.13, 3.29, 2.65, 2.87, 3.40, 2.46, 1.89, and 3.35. Use these new data, along with the cold start times reported in the previous exercise, to construct comparative box plots. Write an interpretation of the information that you see in these plots.

5). Assume a linear model and add 0-mean Gaussian noise to generate 100 samples.

a. Divide your sample into training and testing sets (80:20).

b. Use linear regression (from the sklearn package) for the training half. Compute the mean squared error (MSE) on the testing set.

c. Plot the fitted model along with the data.

d. Repeat the same for polynomials of degrees 2 and 3 as well.