George Smith

HW 1

IST 772

1.Using the material from this chapter and possibly other information that you look up,write a brief definition of these terms in your own words:

* + Mean – sum of numbers divided by the count of the numbers in a set of data values
  + Median – Middle number that separates higher half of data from the lower half of data in a set of data values.
  + Mode – number that appears most often in a set of data values.
  + Variance - Expectation of the squared deviation of a random variable from its mean, or how far a set of numbers is from their average value.
  + Standard Deviation – Similar to variance it is also a measure of how spread-out numbers are, however the output is not squared.
  + Histogram - Approximate representation of the distribution of numerical data.
  + Normal distribution - a function that represents the distribution of many random variables as a symmetrical bell-shaped graph.
  + Poisson distribution - a function that represents the distribution of many random variables as a symmetrical bell-shaped graph.

3. The data set that I selected was the AirPassengers data set. This Data set included The classic Box & Jenkins airline data. Monthly totals of international airline passengers, 1949 to 1960. When using the summary command on this dataset I received values of 265.5 and 280.3 for the median and mean respectively. In this instance the mean is the average amount of customers that flew on the Box & Jenkins airline from 1949 to 1960. The median represents the middle number of passengers that flew on the Box and Jenkins airline from 1949 to 1960 including all months.

4. The dataset I selected was the Nile dataset. The shape of the histogram appeared to be largest in the middle and dwindle down on both the left and right sides. The distribution this data fits most closely is the normal distribution. The Nile dataset fits this distribution because the dataset represents the average flow of Nile at Aswan from 1871 – 1970. I would not expect the flow of water to change drastically in a 100 year period unless there was drastic climate change in the area. As a result, the normal distribution appears logical.

