Stat 50 Quiz #1 Name: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**All work and answers should be on separate paper, not on this quiz.**

**Show All Your Work!!!!!**

**Late submissions will not be graded.**

**The quiz will be open book and open notes (only the textbook for this class and the notes from this class), and you may use a basic scientific calculator. No statistical functions. No internet. No outside help from anyone.**

**This is an individual quiz, so working together is not allowed.**

**Problems are expected to be completed using the methods and procedures presented in class.**

1. Consider the following data set:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 30 | 16 | 22 | 23 | 18 | 18 |
| 20 | 24 | 19 | 13 | 9 | 28 |

1. Compute the mean.
2. Compute the standard deviation.
3. Consider the above data to be the . Now, consider the transformation . Compute the mean, variance, and standard deviation of the transformed data **without actually constructing the transformed data set.**
4. The U.S. National Center for Health Statistics compiles data on the length of stay by patients in short-term hospitals. A random sample of 21 patients yielded the following data on length of stay, in days:

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 1 | 1 | 3 | 3 | 3 | 4 | 5 |
| 5 | 5 | 5 | 7 | 9 | 9 | 10 |
| 12 | 13 | 13 | 15 | 18 | 42 | 55 |

1. Obtain and interpret and.
2. Determine and interpret the interquartile range.
3. List the five-number summary.
4. Identify potential outliers, if any.
5. Construct a boxplot.

3. The following data represent scores of 25 students on a calculus test. Use the data to construct a relative frequency histogram. Use class widths of 40 - 49, 50 – 59, 60 – 69, etc.

72 72 93 70 59 78 74 96 73 80

91 67 72 88 83 76 74 75 68 67

74 76 79 49 84

4. During the early part of the 1994 baseball season, many sports fans and baseball players noticed that the number of home runs being hit seemed to be unusually large. Here are the data on the number of home runs hit by American and National League teams:

American League 35, 40, 44, 49, 51, 54, 57, 58, 58, 63, 68, 68, 75, 77

National League 29, 31, 45, 46, 47, 47, 48, 53, 54, 55, 55, 63, 63, 67

Construct back to back stem and leaf plots for the two data sets.

5. A large population has a mean of 27 and a standard deviation of 3.2. A sample of size 50 is taken from the population. The sample mean is 26.6 and the sample standard deviation is 3.5. The values of the population mean and standard deviation are called\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, while the sample mean and sample standard deviation are called \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_?

Cause and effect can sometimes be established in a \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_, but generally not in an \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_. Explain.

6. Define a Simple Random Sample (SRS) and give an example of a sampling scheme which does not produce an SRS.

For questions 7 and 8, state whether the investigation is an observational study or a designed experiment. Explain your answer.

7. Researchers randomly assigned 39,876 initially healthy women 45 years of age or older to 100 mg of aspirin or placebo on alternative days and then monitored them for 10 years for a first major cardiovascular event (i.e. nonfatal myocardial infarction, non-fatal stroke, or death from cardiovascular causes).

8. A study involving 3800 people over the age of 65 in East Boston. The study took place over a six year period and compared the age at which death occurred for right handed people vs. left handed people among the 3800 people who died during the six year period. The study concluded there was no difference.