**DATA ANALYTICS FOR** **UNIVARIATE DISTRIBUTIONS**

**Key Descriptive Statistical Indices**

***--Show All Work--***

Once data are collected statisticians are typically initially interested in three fundamental statistical properties of **UNIVARIATE DISTRIBUTIONS**: (1) **Central Tendency** (or Central Location), (2) **Dispersion** (or Variability), and (3) the **Shape** (or Form) of the data distribution. There are multiple measures of central tendency, dispersion and shape. We will follow the standard convention of using the level of measurement (nominal, ordinal, interval/ratio or categorical vs. quantitative) of the variable as the major decision criterion.

The *purposes* of this assignment are to introduce you to the *computation* and *interpretation* of some of the fundamental statistical techniques used to describe *univariate* ("uni"= one; "variate"= variable) distributions. Although these techniques should be familiar to you, it is important to review them because they provide the foundation for the more sophisticated procedures which we will be discussing.

**Sociological Research Question:** Suppose you are an educational sociologist and the superintendent of the local school district asks you to analyze some quantitative data the office has collected on teachers. The superintendent wants to be informative for the upcoming collective bargaining meeting with the teachers’ union.

Using the data set on teachers, your tasks are to:

***A.  Measures of Central Tendency or Central Location***

1. Compute and interpret the following Measures of Central Tendency for the salary variable for teachers 29-41:

Arithmetic Mean

Median

Mode

Mid-range

***B.  Measures of Dispersion or Varia*tion**

1. Compute and interpret the following Measures of Dispersion (Variability) for the salary variable for teachers 29-41:

Range

Standard Deviation

Variance

Sum of Squares (technically not)

Interquartile Range

**NOTE**:  The *sum of squares* is one of the most conceptually important and widely used computations in statistics.  Be sure you understand its computation and interpretation and its relationship to the *three-step method* for computing the standard deviation.  It will be used again in subsequent assignments.

***C. “Measures” of Shape***

1. Construct and interpret the following Graphic Techniques for the salary variable for teachers 29-41:

Boxplot

Frequency Polygon

Percentage Histogram

a. Create a frequency and percentage univariate table and a bar graph for the qualitative variable “Religious Preference” for teachers 29-41.

* Using your *sociological* and *statistical imaginations,* what do you convey to the superintendent?

***D. The Normal Curve (NC) as a Probability Distribution & Standard (or z) Scores***

4. Assume you are a sociology graduate student specializing in social stratification. Government census data indicate that the mean household income of Americans is **$74,580** and the standard deviation is $8,443. Using the empirical rule regarding areas under the *NC*, determine the percent of Americans who earn between +1, +2 and +3 & -1, -2, and -3 standard deviations from the mean.

a. What would be the z-score (standard score) and percentile rank for a salary of:

(1) $39,332

(2) $74,580

(3) $85,145

**🡪 Now using SPSS calculate the above statistical problems to double-check your computations and submit your computer outputs. Are there any discrepancies? If so, why?**