City Crime Rankings 2008-2009

Distribution Analysis

New to *City Crime Rankings*, this Distribution Analysis section presents charts depicting the distributions of the comparison scores as well as the individual and collective reported crime rates shown in this book to provide a mechanism of comparison beyond the rankings included in each analysis. The histograms illustrate the distribution of values for each analysis in the book. Along with each histogram, measures of central tendency such as median, mean, standard deviation, and minimum and maximum values are reported to provide further description of each distribution.

In each histogram (formatted as area charts for easier viewing), the values of the scores or rates are show along the bottom (x-axis) and the frequency of cases (i.e., metro areas or cities) are shown along the right (y-axis). The values along the bottom depict ranges for which the frequency of cases is totaled. These ranges and frequencies are different for each distribution (i.e., each histogram).

The median indicates the middle value of the distribution, which means that 50% of the metro areas or cities have scores or rates above and below that value. The mean is the average value of the distribution, and the standard deviation, described generally, is the measure of spread of all the values from the mean. The minimum and maximum values are the lowest and highest values of the distribution, respectively.

These statistics are based on a normal curve, so one standard deviation above and below the mean contains 68% of the distribution, two standard deviations above and below the mean contain 95% of the distribution, and three standard deviations above and below the mean contains 99.7% of the distribution. The use of these statistics is purely descriptive but helps the reader assess the distribution as a whole as well as illustrates where an individual value sits in terms of all the other values. For example, if a score is two or three standard deviations above or below the mean, it may be considered an outlier because it falls with only 5% or .3% of the values, respectively.

For example, Figure 1 depicts the comparison scores for metro areas in 2007. The median is -9.0, the mean is -5.4, the standard deviation is 38.0, the minimum value is -76.1, and the maximum value is 117.7. These statistics are interpreted as follows:

- The lowest comparison score for metro areas is -76.1.

- The highest comparison score for metro areas is 117.7.

- The range of scores (maximum minus minimum) is 193.8.

- 50% of the metro areas have comparison scores lower than -9.0, and 50% have scores higher than -9.0.

- The average comparison score for metro areas is -5.4 and the standard deviation is 38.0.

- 68% of the metro areas have scores between -43.4 and 32.6.

- 95% of the metro areas have scores between -81.4 and 70.6.

- 99.7% of the metro areas have scores between -119.4 and 108.6. (The fact that the lower end of this and the previous range are less than the minimum value of the distribution indicate the distribution is skewed.)

Assessing the score of -37.84 for York-Hanover, Pennsylvania, for example, reveals that it is in the lower 50% of all the scores (below the median) and falls within the first standard deviation of the mean with 68% of the other scores (between -43.4 and 32.6).

The remainder of this section presents charts and statistics for both metropolitan areas and cities in the categories listed here:
1. Comparison Score
2. Overall Crime
3. Violent Crime
4. Property Crime

A word of caution: These distribution analysis charts and statistics are provided to help understand the nature of the values within each analysis, but the analyses are still based on data that must be interpreted within the constraints noted earlier. These charts are only descriptions of the data and do not provide prediction or explanation of why these values are different.

Figure 1: Metropolitan Areas Comparison Score Distribution Analysis for 2007
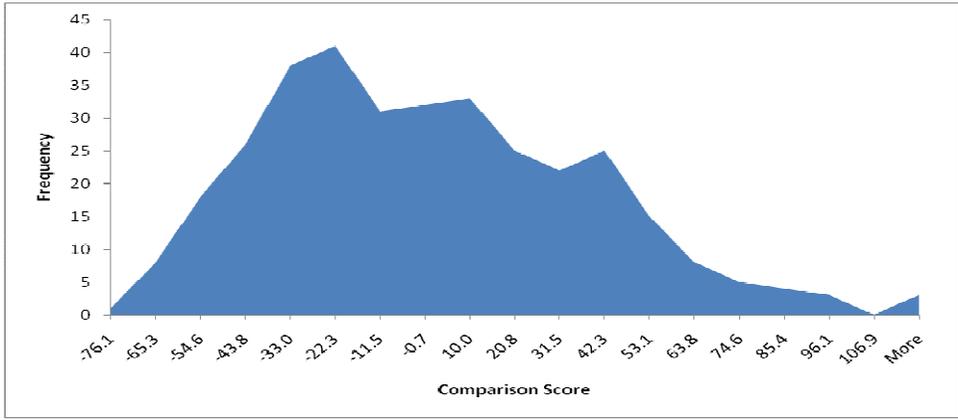
Median	-9.0
Mean	-5.4
Standard Deviation	38.0
Minimum	-76.1
Maximum	117.7

Figure 2: Cities Areas Comparison Score Distribution Analysis for 2007
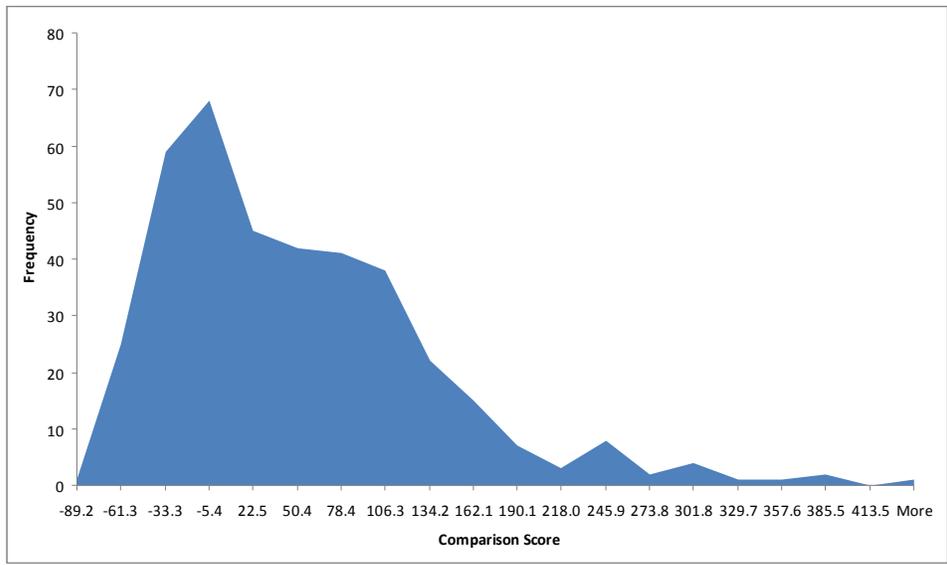
Median	20.1
Mean	37.1
Standard Deviation	87.1
Minimum	-89.2
Maximum	441.4