

data before any analysis takes place so that the extreme values associated with the outliers do not skew the results.

Data

Data > US Violent Crime
by State

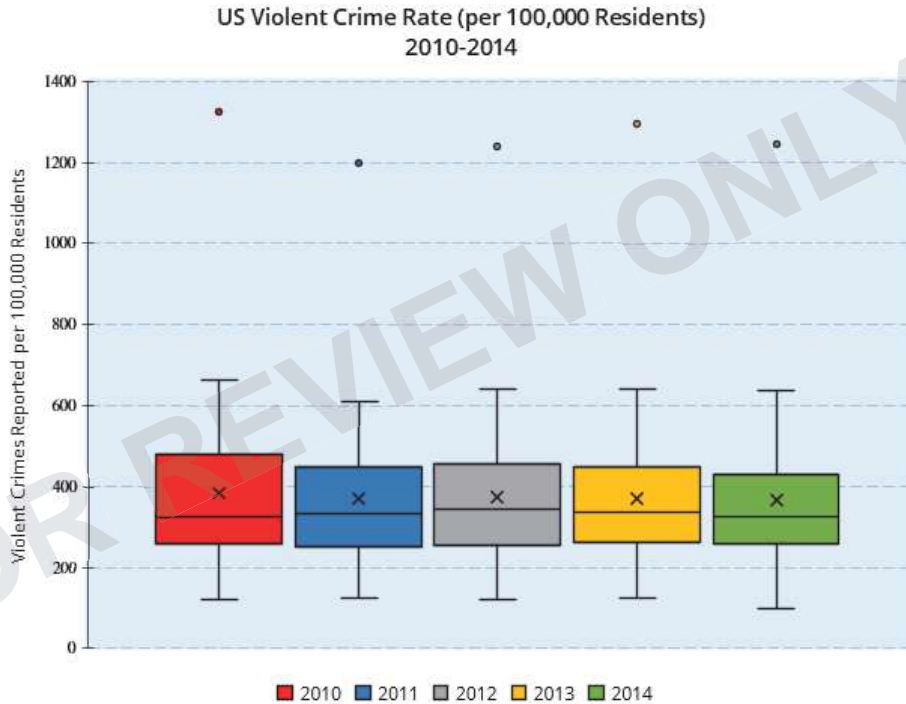


Figure 4.3.5

z-Scores

The **z-score** is a standardized measure of relative position with respect to the mean and variability (as measured by the standard deviation) of a data set.

z-Score

The **z-score** transforms a data value into the number of standard deviations that value is from the mean.

$$z = \frac{x - \mu}{\sigma}$$

FORMULA

Describing a data value by its number of standard deviations from the mean is a fundamental concept in statistics that is found throughout this book. It is used as a standardization technique, a yardstick, to describe properties of data sets and to compare the relative values of data from different data sets.

Example 4.3.4

Suppose you scored an 86 on your biology test and a 94 on your psychology test. The mean and standard deviations of the two tests are given.

Test Scores		
Course	Mean	Standard Deviation
Biology	74	10

Test Scores		
Course	Mean	Standard Deviation
Psychology	82	11

What are the z -scores for your two tests? On which of the tests did you perform relatively better?

Solution

The z -score for the biology test is $z = \frac{86 - 74}{10} = 1.20$.

The z -score for the psychology test is $z = \frac{94 - 82}{11} = 1.09$.

On the biology test you scored 1.2 standard deviations above the mean, compared to only 1.09 standard deviations above the mean for the psychology test. Even though the raw score on the psychology test is larger than the raw score on the biology test, relative to the means and variability in the data sets, the performance on the biology test was slightly better. Once again, changing the scale of the data has beneficial effects. It enables the comparison of two measurements that are drawn from different populations.

Rounding Rule

When calculating a z -score, round to two decimal places.

Properties of a z -Score

- If a z -score is negative, the corresponding data value is less than the mean.
- Conversely, if a z -score is positive, the corresponding data value is greater than the mean.
- The z -score is a unit-free measure. That is, regardless of the original units of measurement (centimeters, meters, or kilometers), an observation's z -score will be the same.

PROPERTIES