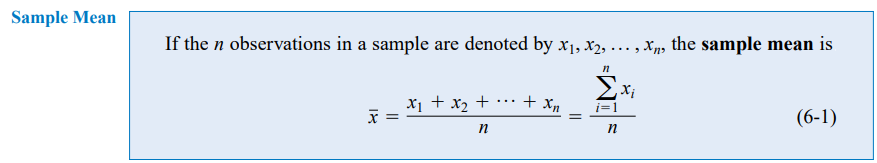
**Chapter 6: Descriptive Statistics**

1. **Numerical summaries of data**

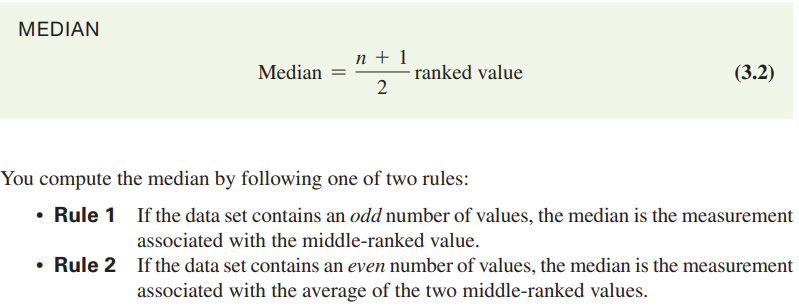
**Sample Mean** = sum of values divided by the number of values.



**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 - 2**

**Sample Median:** is the “middle” number (50% above, 50% below).



1. Rank the data set in increasing order.

2. Median position = position in the ordered data.

n: sample size

* If the number of values is odd, the median is the middle number.
* If the number of values is even, the median is the average of the two middle numbers.

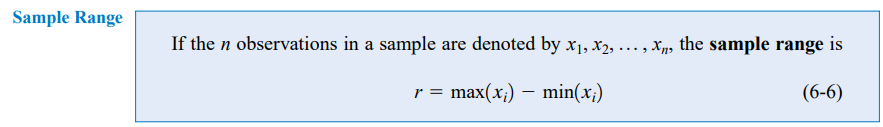
**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**Sample Mode:**

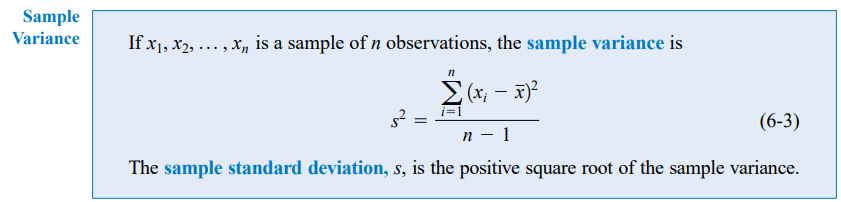
Value that occurs most often.

There may be **no mode** OR there may be **several modes**.

**Sample Range**



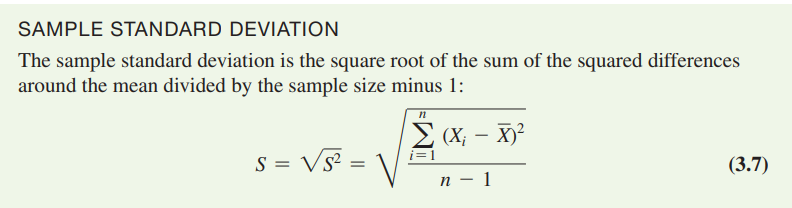
**Sample Variance:** Average (approximately) of squared deviations of values from the mean.

****

**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 - 4 - ^2**

**Sample Standard Deviation:** Is the square root of the variance.

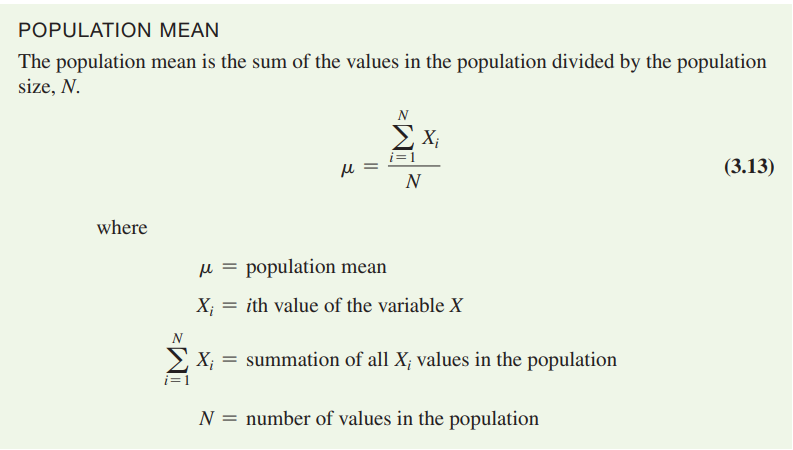
****

**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 - 4**

1. **Numerical Descriptive Measures for a Population**

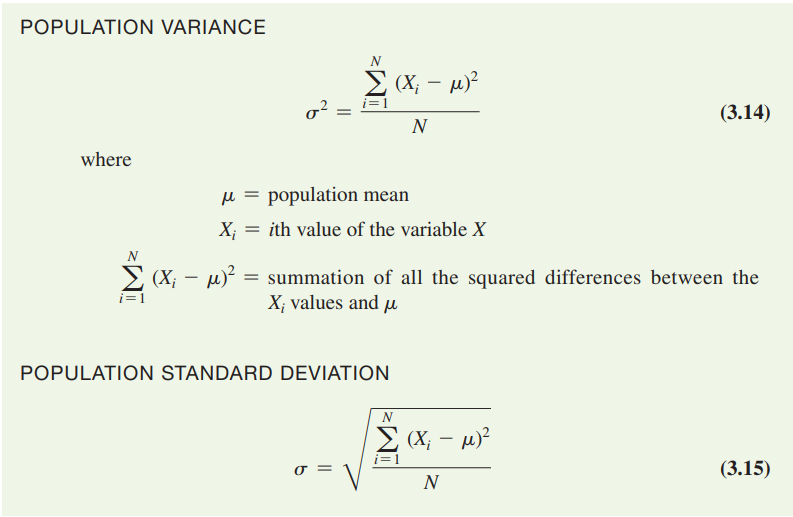
**Population Mean**

****

**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 - 2**

**Population Variance & Population Standard Deviation**

****

**Population Variance**

**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

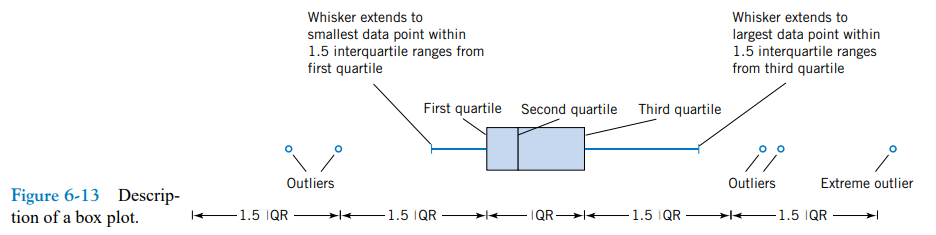
**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 - 3 - ^2**

**Population Standard Deviation**

**580VN: Menu - 6 - 1 - Data - AC - OPTN - 2**

**570VN: Mode - 3 - 1 - Data - AC - Shift - 1 - 4 – 3**

1. **Chart & Diagram & Table**

**3.1 Box Plots **

**Quartiles** split the ranked data into **4 segments** with an **equal number of values** per segment.

The first quartile, **Q1**, is the value for which 25% of the values are smaller and 75% are larger.

**Q2** is the same as the median (50% of the values are smaller and 50% are larger).

Only 25% of the values are greater than **the third quartile**.

**Find a quartile** by determining the value in the appropriate position in the ranked data, where:

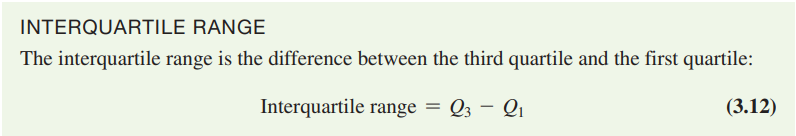
First quartile position: ranked value.

Second quartile position: ranked value.

Third quartile position: ranked value.

where n is the number of observed values.

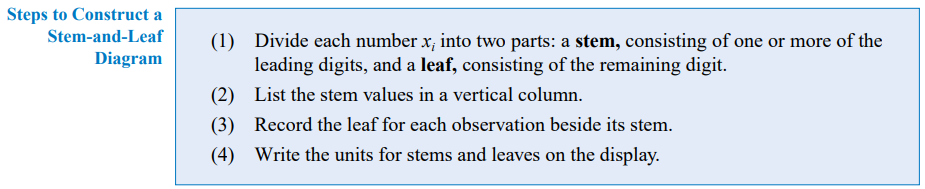
The **Interquartile Range** (**IQR**) is **Q3 – Q1** and measures the spread in the **middle 50% of the data**.

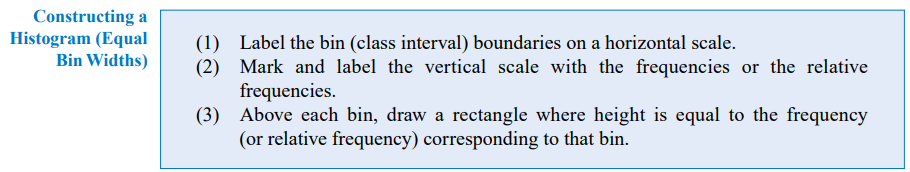


An **outlier** is an observation that is numerically distant from the rest of the data

(< Q1 - 1.5 \* IQR **OR** > Q3 + 1.5 \* IQR).

**3.2 Stem and leaf diagram**

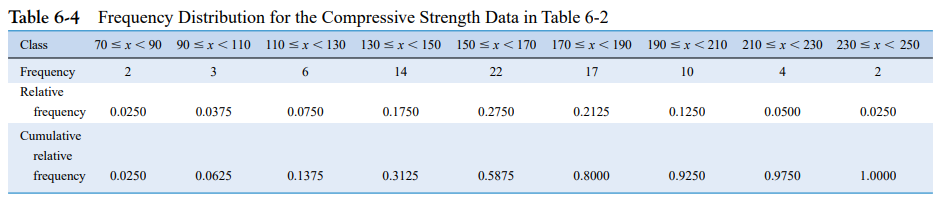
****

**3.3 Histogram **

**3.4 Frequency distribution**

**n: sample size**

**class/bin**

****

population, sample

probability distribution table → population

liệt kê → 90% sample