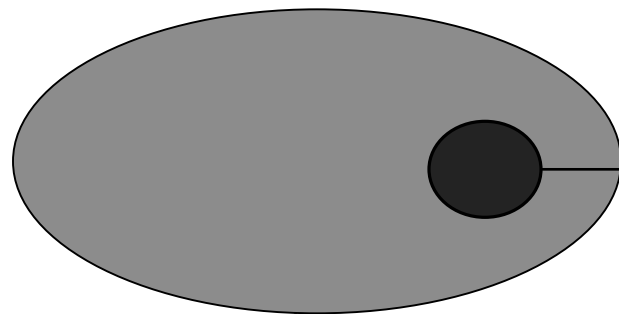


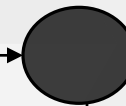
Basic Statistical Terms

Population:
Complete
collection to
be studied



Sampling
Process

Sample: Part
of population



Parameter

Inference

Statistic

Characteristic of
a population

Characteristic
of a sample

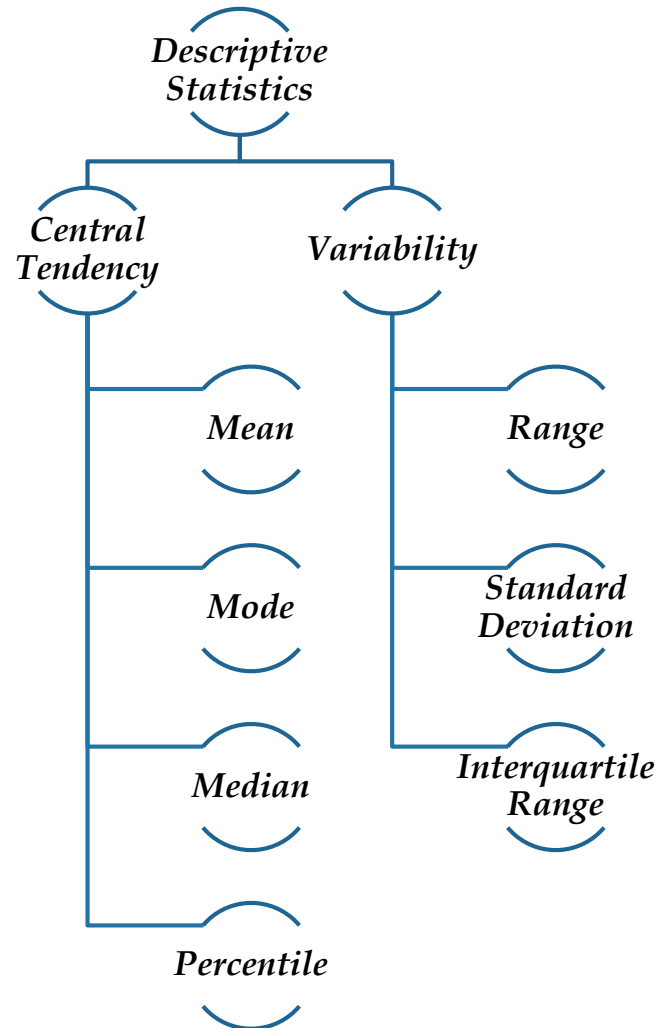
N number of members
 μ mean
 σ standard deviation

n
 \bar{x}
 s

Notations

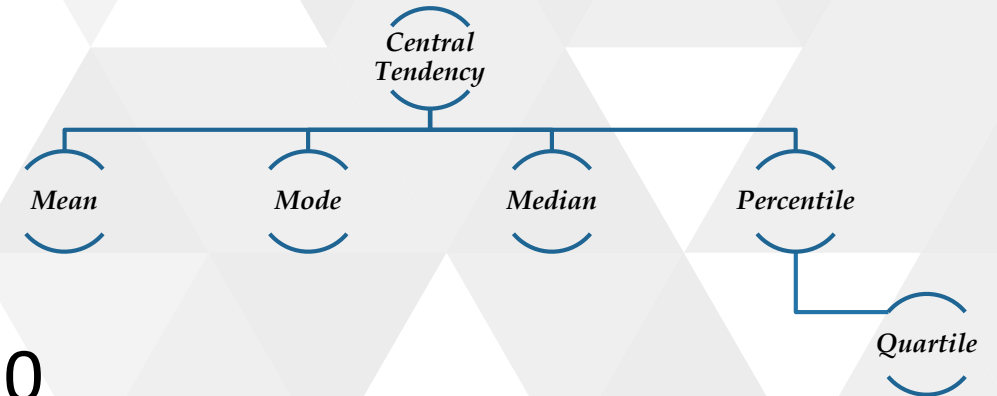
	Population Parameters	Sample Statistics
Mean	μ	\bar{x}
Standard Deviation	σ	s
Variance	σ^2	s^2
Proportion of population having an attribute	P	p
Proportion of population not having an attribute	Q (=1-P)	q (=1-p)
Correlation coefficient	ρ	r
Number of elements	N	n

Descriptive Statistics



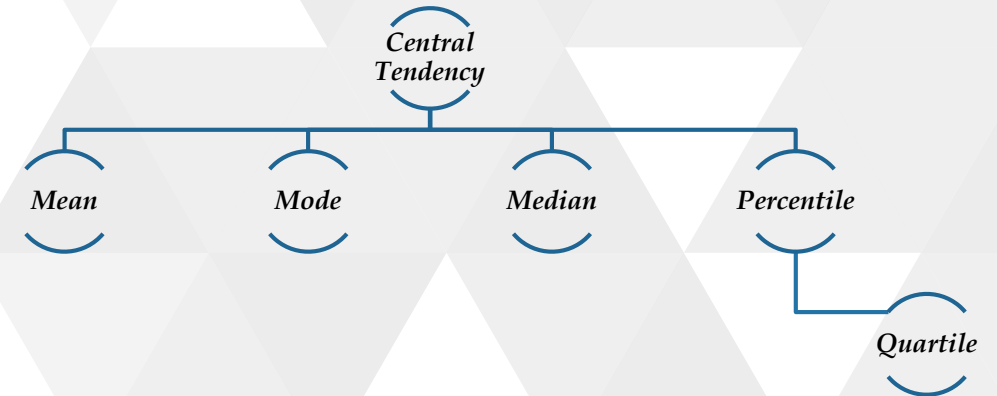
Mean

- ❖ Also known as Average
- ❖ Affected by extreme values
- ❖ Example: 10, 11, 14, 9, 6
- ❖ $\text{Mean} = (10+11+14+9+6)/5 = 50/5 = 10$



Mode

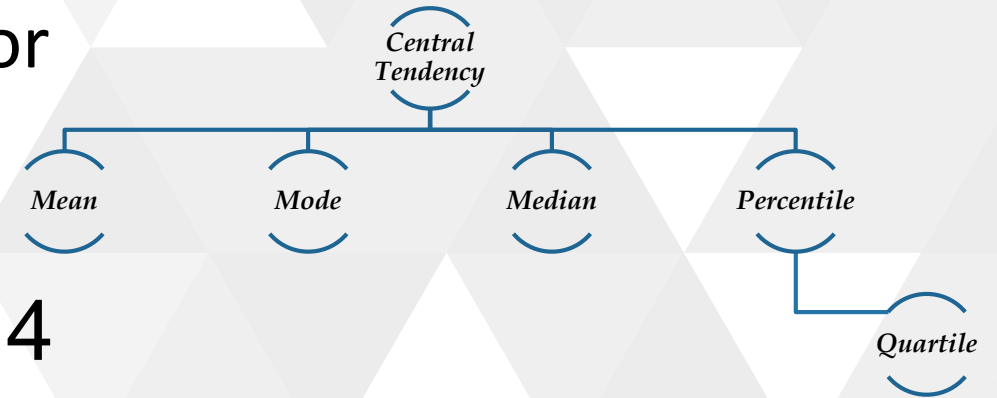
- ❖ Most occurring item
- ❖ Example: 10, 11, 14, 9, 6, 10
- ❖ Mode = 10



Median

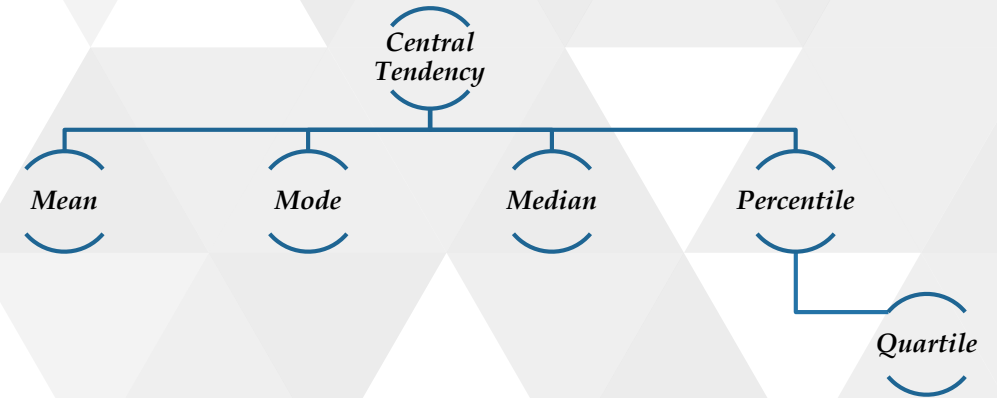
- ❖ Middle value when put in ascending or descending order.
- ❖ Example: 10, 11, 14, 9, 6
- ❖ In ascending order - 6,9,10,11,14
- ❖ Median = 10

- ❖ Example: 10, 11, 14, 9, 6, 11
- ❖ In order - 6,9,10,11, 11,14
- ❖ Median = 10.5



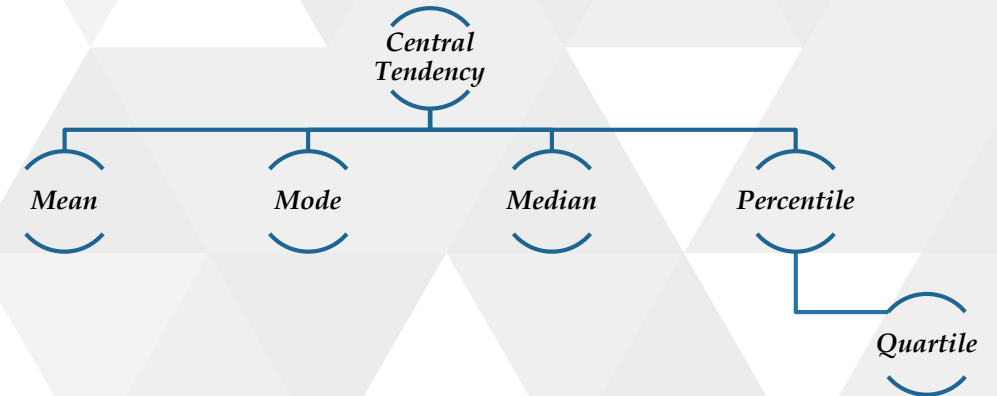
Percentile

- ❖ Median divides the data in two equal parts when arranged in ascending or descending order
- ❖ Percentile divides data in 99 parts
- ❖ Quartile divides data in 4 parts
- ❖ Example: 6,9,10,11, 11,14
- ❖ $Q1=9$, $Q2=10.5$, $Q3=11$

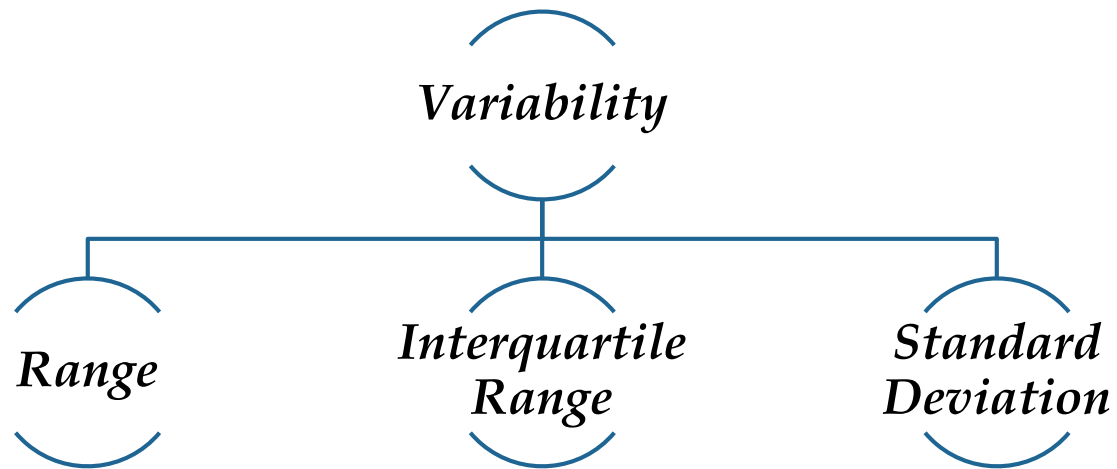


Percentile/Quartile Steps

- ❖ Arrange in ascending or descending order
- ❖ Calculate location(i) = $P.(n)/100$
- ❖ P=percentile, n=numbers in data set
- ❖ If i is whole number – Percentile is average of (i)th and (i+1)th location
- ❖ If i is “not” a whole number – Percentile is located at (i+1)th whole-num.
- ❖ Example: 6,9,10,11, 11,14
- ❖ Q1=9, Q2=10.5, Q3=11

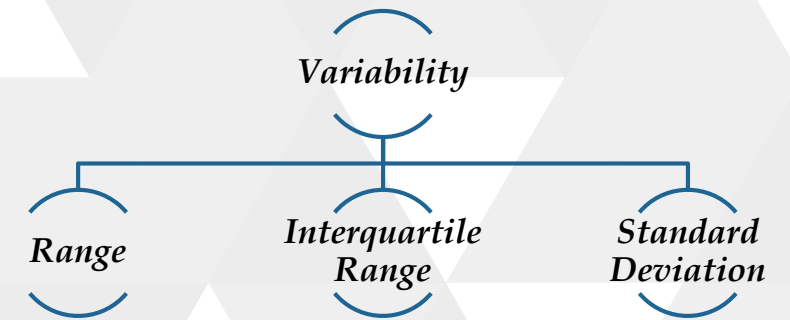


Descriptive Statistics



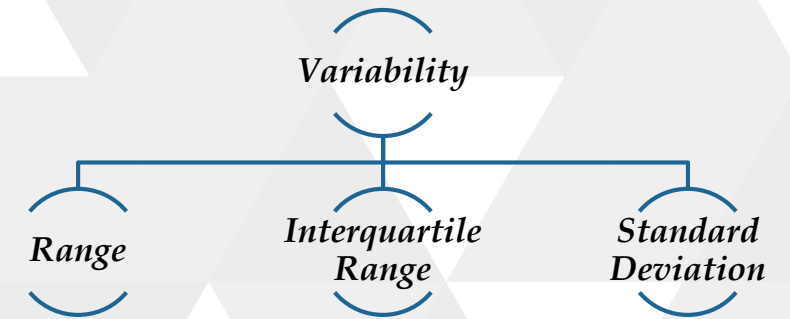
Range

- ❖ Difference between lowest and the highest value.
- ❖ Example: 6,9,10,11, 11,14
- ❖ $\text{Range} = 14 - 6 = 8$



Interquartile Range

- ❖ Range of middle 50% data
- ❖ $IQR = Q3 - Q1$
- ❖ Example: 6, 9, 10, 11, 11, 14
- ❖ $Q1 = 9$, $Q2 = 10.5$, $Q3 = 11$
- ❖ $IQR = 11 - 9 = 2$
- ❖ Box-and-Whisker Plot

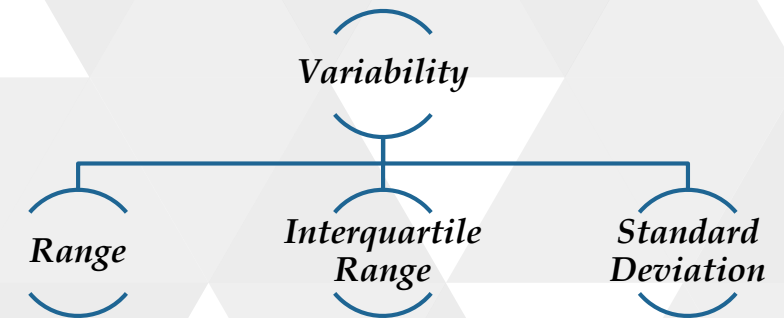


Standard Deviation

- ❖ Variance = average of squared deviation about the arithmetic mean.
- ❖ Square root of variance is standard deviation

$$s = \sqrt{\frac{\sum (x - \bar{x})^2}{n - 1}}$$

$$\sigma = \sqrt{\frac{\sum (x - \mu)^2}{N}}$$



Standard Deviation

x	$x - \bar{x}$	$(x - \bar{x})^2$
100	0	0
101	1	1
99	-1	1
102	2	4
98	-2	4
100	0	0
$\bar{x} = 100$	$\sum(x - \bar{x}) = 0$	$\sum(x - \bar{x})^2 = 10$

$$s^2 = \frac{\sum(x - \bar{x})^2}{n-1}$$

$$s^2 = 10/5 = 2$$

$$s = \sqrt{2} = 1.414$$