The Data Driven Manager

2

Data and Measurement

Learning Objectives

- Discern between qualitative and quantitative data, continuous and discrete data
- Compare/contrast measurement and underlying characteristics
- Distinguish between nominal and ordinal scales
- Distinguish between interval, ratio and absolute scales

Learning Objectives

- Identify the measurement level of data given background information
- Describe the process of measurement
- Demonstrate and recall the five (5) aspects of data
- Discern between population and sample statistics

“When you can measure what you are speaking about, and express it in numbers, you know something about it

…but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind.”

Lord Kelvin (1883)

Data and Measurement

How Are They Related?

Data Costs Money

We must make data both

- Efficient and
- Effective

How Do We Make Data Efficient and Effective?

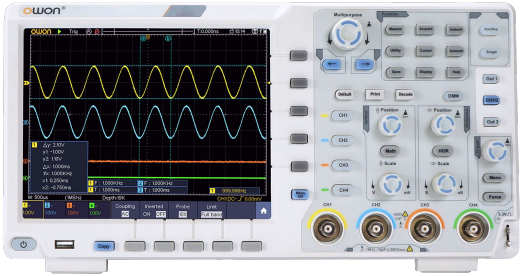
Make certain that the measurement process itself is effective:

- Capable
- Acceptable

Measurement and Data

How do we study, record and communicate an event? We assign numbers.

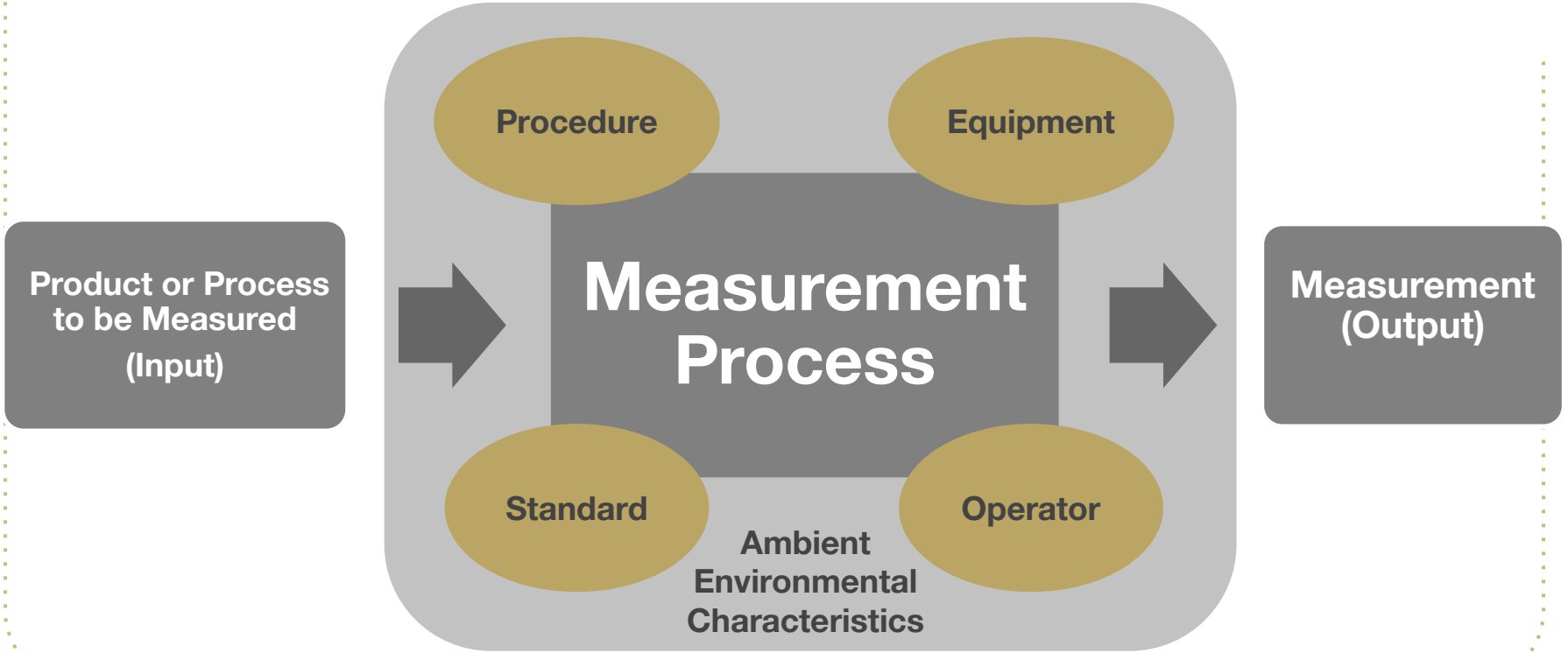
- Measurement is the **process**



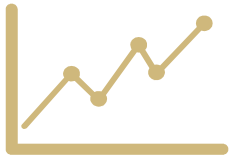
- Data is the **output**

1.5	1.3	1.6	1.5
1.6	1.4	1.2	1.3
1.2	1.7	1.3	1.2
1.4	1.4	1.1	1.7

Measurement as a Process



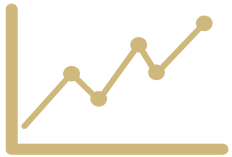
Measurement as a Process



As in any process, measurement systems must demonstrate:

- Stability through time, or control
- The ability to generate reliable, or repeatable and reproducible measures
- The ability to generate valid measures

Measurement as a Process



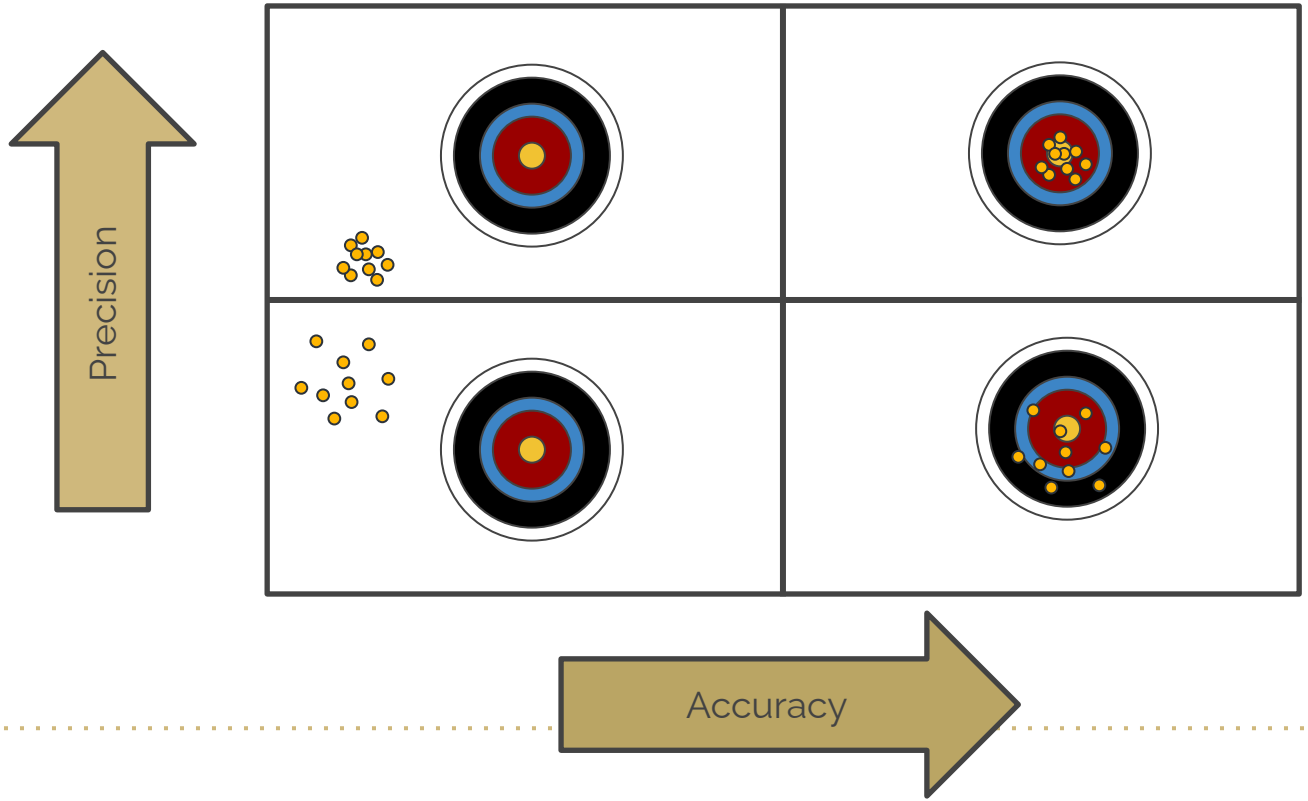
Reliability in measurement is a measure of the

- Precision of the device / method

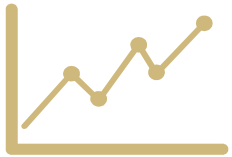
Validity in measurement is a measure of the

- Accuracy of the device / methods

Precision vs. Accuracy

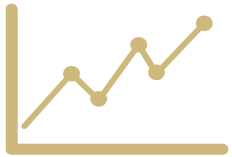


Understanding Data



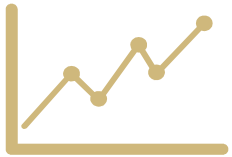
Once the measurement process is found to be capable and acceptable, we can begin to analyze the generated data

Quantitative Data



- **Quantitative** data are data measured along a numerical scale.
  - Often referred to as **continuous**.

Qualitative Data

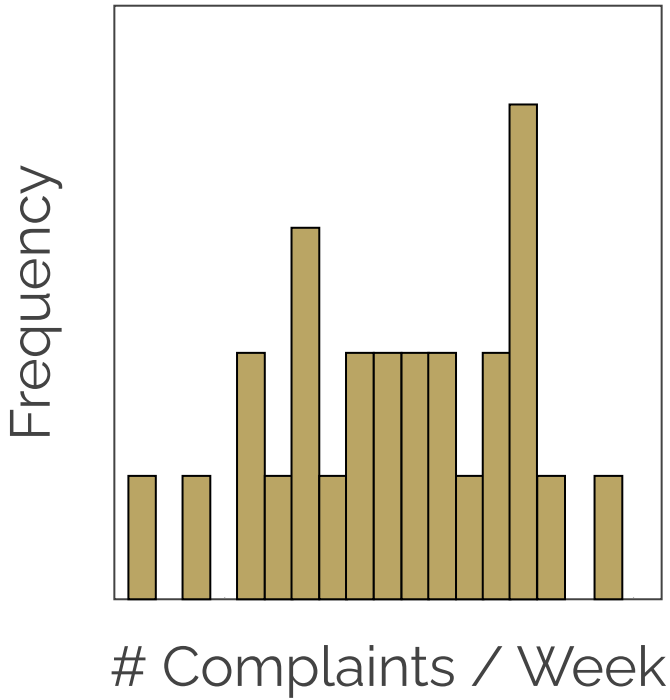


- Qualitative data are descriptions that fall into categories.
  - Often referred to as discrete.
  - Frequencies, proportions, or rates.

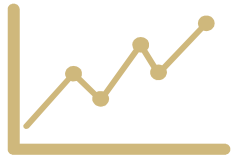
Discrete Data

Discrete Data:

- Items/Units we **count**



Discrete Examples

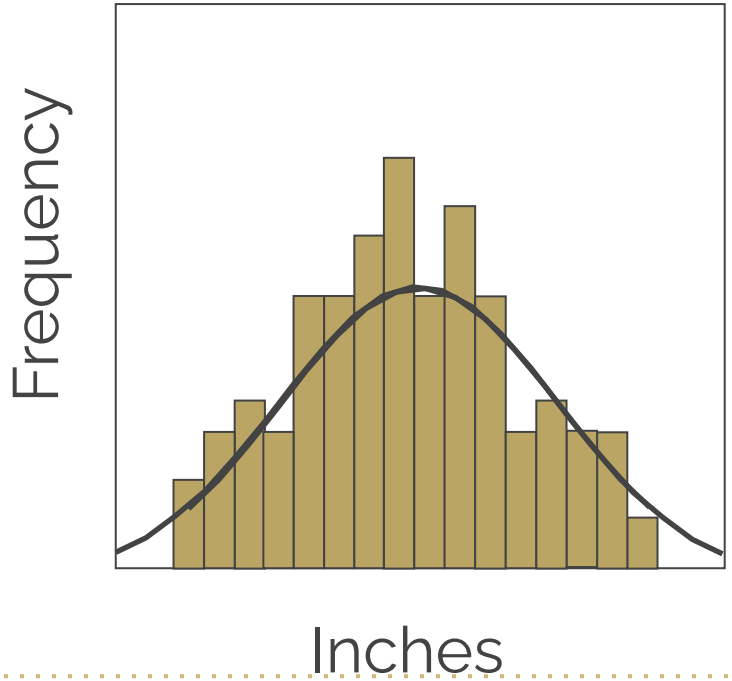


- Examples of discrete data in business:
  - Complaints per sales period
  - Number of defects per unit
  - Percent defective units
  - Number of orders shipped on time

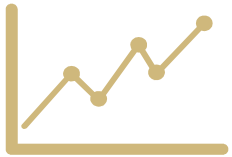
Continuous Data

Continuous Data:

- Items/Units we **measure**



Continuous Examples



- Examples of continuous data in business:
  - Dimensions (height, length, width)
  - Temperature
  - Speed
  - Volume of sales

Data

Underlying Property

Operational Definition

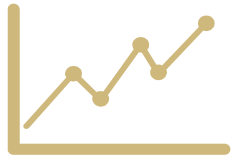
Criterion Measure

Data

Measurement

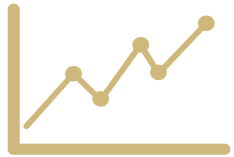
Measurement Scales

Measurement



Measurement is the assignment of numbers or other symbols to an underlying attribute, characteristic or property.

Measurement Scale

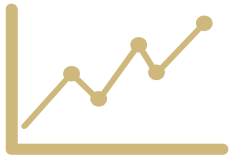


The numbers, or symbols, are assigned such that the relationships amongst the numbers or symbols reflect relationships in the attribute studied.

Measurement Scale

Measurement Scale

Measurement Scale



- Measurements are not the same as the attribute studied

- To draw conclusions, we must consider how the measurement maps to the attribute

Data

Underlying Property

Operational Definition

Criterion Measure

Data

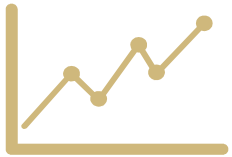
Data

Underlying Property ⟷ Criterion Measure

Operational Definition

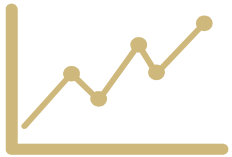
Data

Measurement Scale



- The level of measurement restricts:
  - Transformations that may be performed
  - Statistics that may be calculated
  - Statistical procedures employed

Measurement Scales



- Nominal
- Ordinal
- Interval
- Ratio
- Absolute or Ratio Discrete

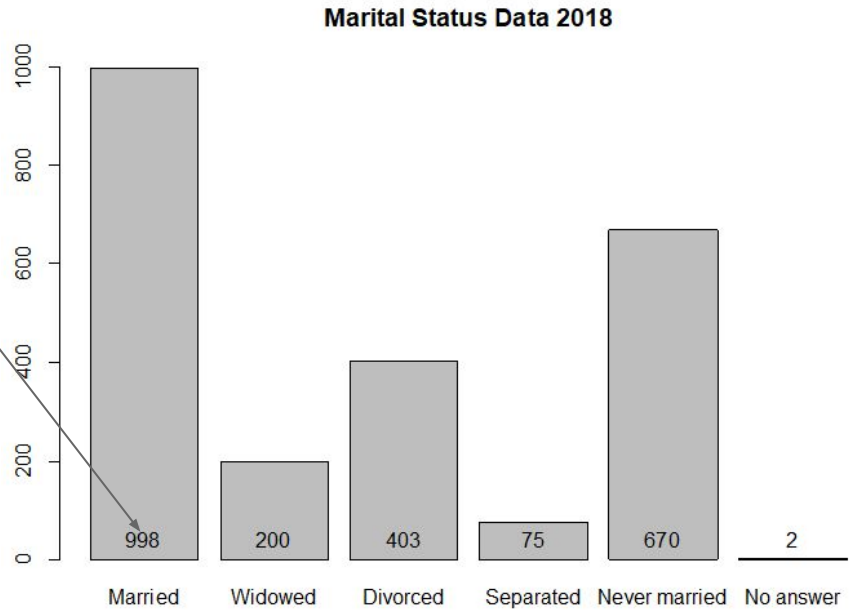
Nominal Scale

Numbers are assigned to **categorize**, identify or name attributes

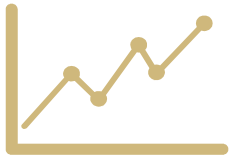
- Zip codes
- Area codes
- Numbers assigned to types of nonconformity in products
- Numbers assigned to presence or absence of an attribute (e.g., 0, 1)
- Numbers assigned to sales territories

Nominal Scale

Frequency = Count

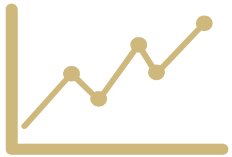


Nominal Scale



- Nominal Scale values can only be used to indicate = or ≠
- Analysis is restricted to **frequency** or **proportion** of values in the assigned categories

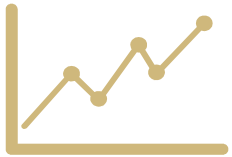
Ordinal Scale



Numbers are assigned to observations, such that the order of the numbers corresponds to the **order** of the underlying property studied

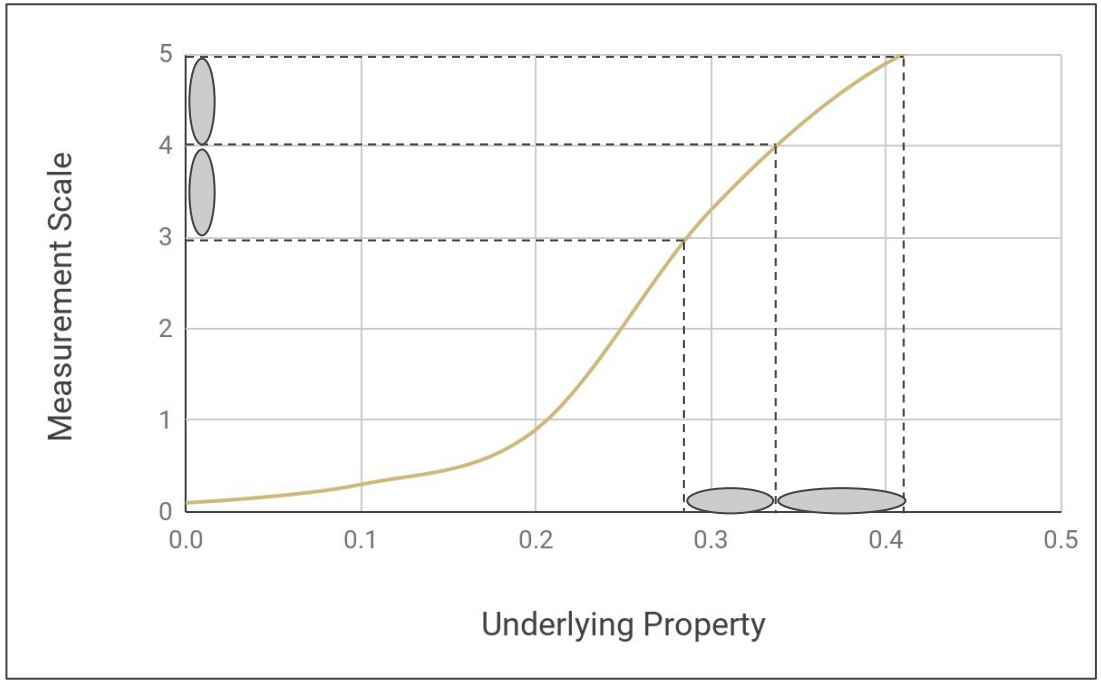
- 5-point scale to measure customer satisfaction
- Letter Grades for Academic performance
- Sound intensity measured in decibels
- Socio-economic status
- Project priority numbers

Ordinal Scale

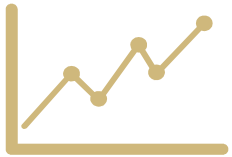


- Ratings and Rankings are ordinal scale
  - **Ratings**: assign a subjective score on a scale
  - **Rankings**: result from sorting items, assigning a unique number to each item.

Ordinal Scale



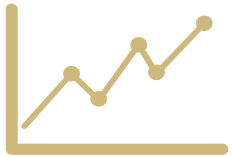
Ordinal Scale



- The median and mode may be used for “center” and range or interquartile range for dispersion

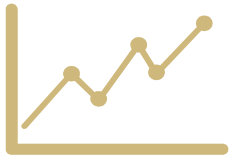
- Ordinal Scale values can be used to determine = or ≠ and > or < as well, but NOT magnitude

Measurement Scales



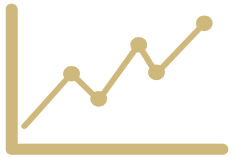
- Nominal
- Ordinal
- Interval
- Ratio
- Absolute or Ratio Discrete

Interval Scale



- Numbers are assigned to observations such that differences between any two numbers (the interval) correspond to proportional differences in the underlying property studied, and there are equal intervals along the scale.

Interval Scale

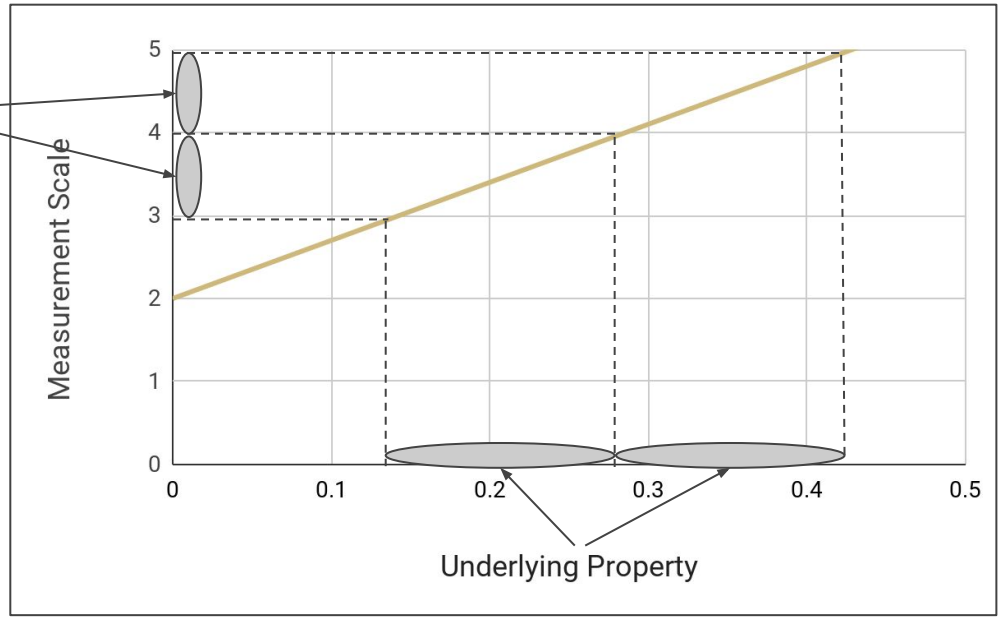


- Examples
  - Temperature measured in degrees Fahrenheit
  - Directional distance from a reference point
  - Calendar date
  - Height above sea level

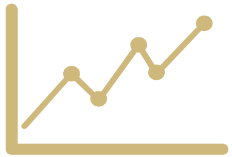
- Zero is a value on the continuum, hence negative values are possible

Interval Scale

Note the proportionality of the differences (interval) in the values on the measurement scale are the same as for the underlying property.



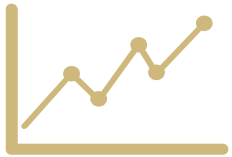
Interval Scale



- Statistics such as the mean and standard deviation may be used as well as median, mode, and range

- Interval Scale values can be used to determine = or ≠, > or <, and you can use sums (+) and differences (–) with meaning

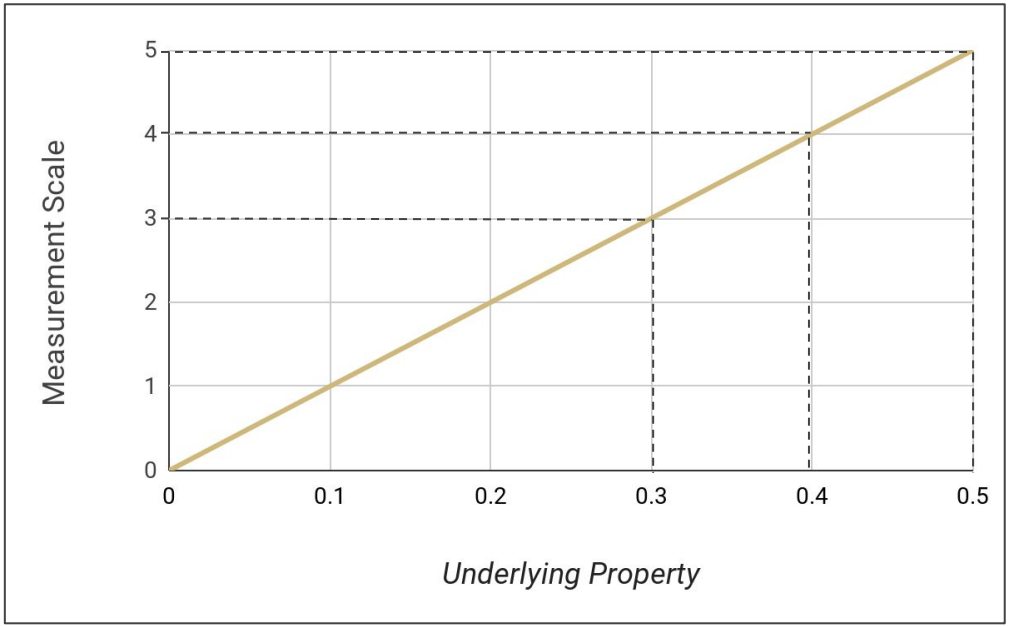
Ratio Scale



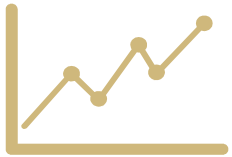
- Numbers are assigned to observations such that an interval scale has been achieved, and there is a zero point which corresponds to a zero, null state, or absence of the underlying property
  - Length, height, width, distance
  - Volume, weight
  - Cycle-time and time-to-repair

Ratio Scale

Same as Interval Scale but **zero** represents an **absence** or zero amount of the underlying Property



Ratio Scale



- Ratio Scale values can be used to determine = or ≠, > or <, and you can use sums (+) and differences (-) with meaning as well as compute ratios and products (× and ÷) with meaning

Ratio vs Interval Scale

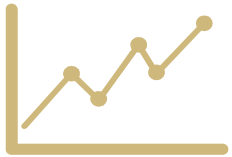
- Consider a bathroom scale:
- Two people are weighed on the same bathroom scale, one weighs 200 pounds and the other weighs 100 pounds, is the first twice as heavy as the second?

Ratio vs Interval Scale

- What if I told you the scale was NOT set at zero to begin with, but at 50 pounds. Then what?
- What are the corresponding Scales of Measurement for each situation?

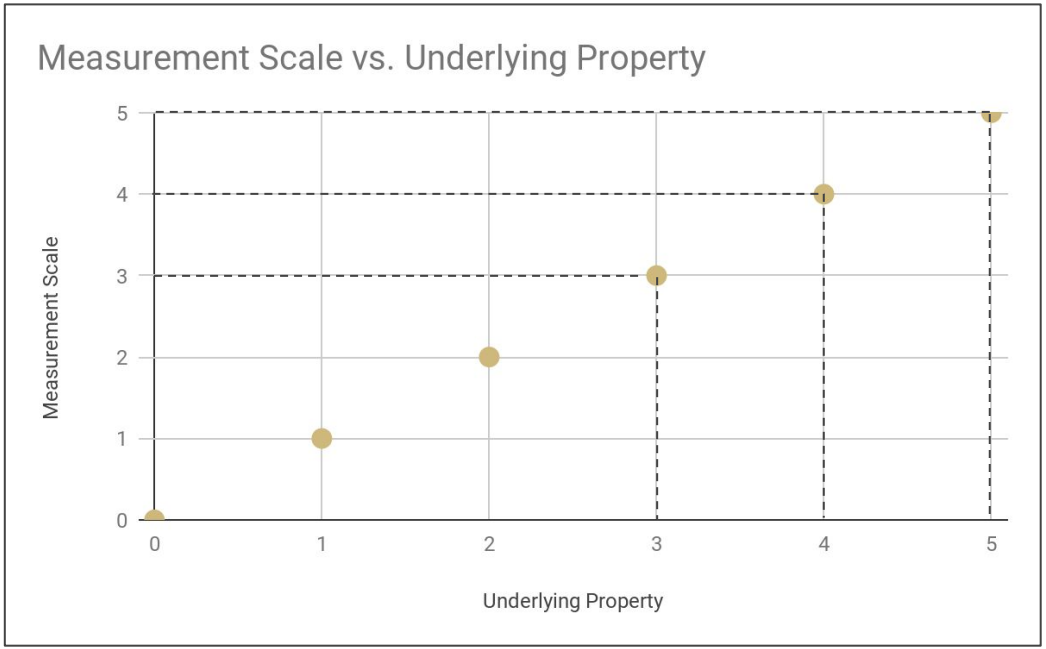
Absolute Scale

Count Data



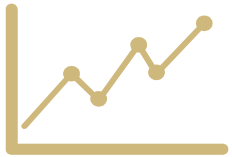
- Numbers are assigned to observations such that the numbers directly correspond to the underlying property being studied
  - The number of defects
  - The number of scratches observed
  - The number of parts made
  - The number of safety accidents
  - The number of customer complaints

Absolute Scale



Absolute Scale

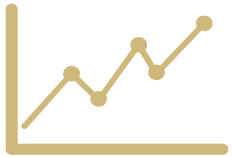
Count Data



- Some transformations are permissible, particularly a one-to-one transformation.
- Data on an absolute scale have some of the properties of ratio data, hence it may be called Ratio Discrete.
- The mean, median, and mode may be used.

Absolute Scale

Count Data



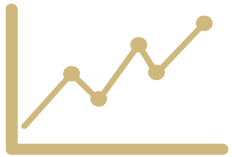
- Various statistical procedures may be used depending on resolution.
- In many cases, standard parametric methods may be used, in other cases nonparametric methods must be employed.
- For Absolute scales, the scale value **IS** the underlying characteristic.

Characteristics of Measurement Scales

Scale	Relationships	Interpretation	Statistics
Nominal	= ≠	Equivalence or not, No order and no magnitude	Frequencies, Mode, Proportions
Ordinal	= ≠ < >	Order, but no magnitude	Median, Range, IQR, plus those above
Interval	= ≠ < > plus equivalence of intervals	Magnitude, meaningful differences of intervals. Zero is a point on the scale.	Mean, Standard Deviation plus those above
Ratio	= ≠ < > plus equivalence of intervals and a true zero	Interval scale with a True Zero (Absence of the Characteristic)	Geometric Mean, Coefficient of Variation, plus those above
Absolute	All of the above apply	Is Discrete, but has all the Properties of Ratio Scale	All of the above apply

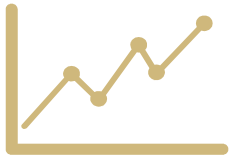
The Big 5 Aspects of Data

Big 5 Aspects of Data



- Location or Central Tendency
- Spread or Dispersion (Variability)
- Shape
- Time Sequence
- Relationship

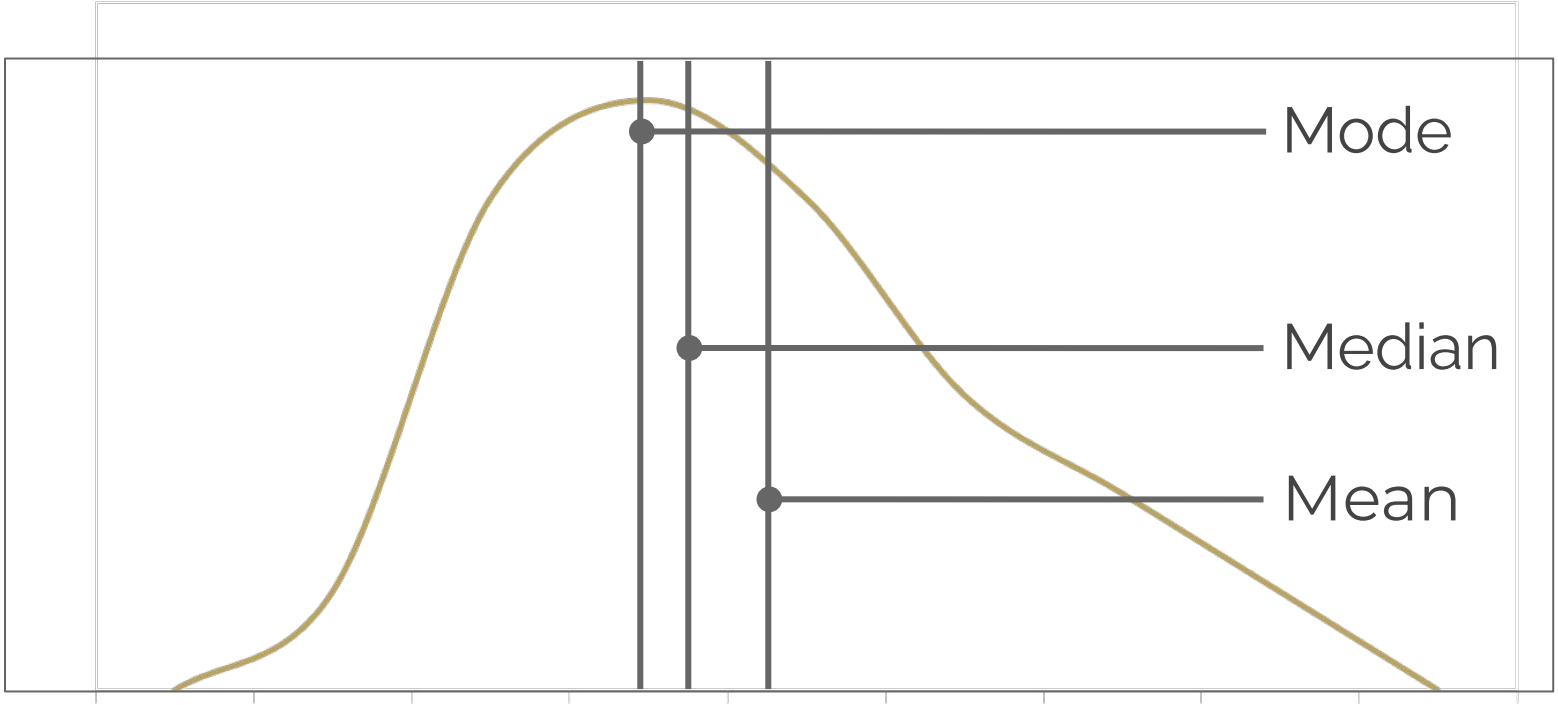
Measures of Location



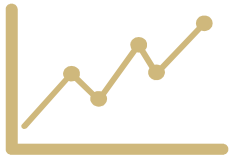
Measures of location, sometimes called measures of central tendency, describe a middle or central point or tendency of a distribution.

- Mean, Median, Mode

Measures of Location



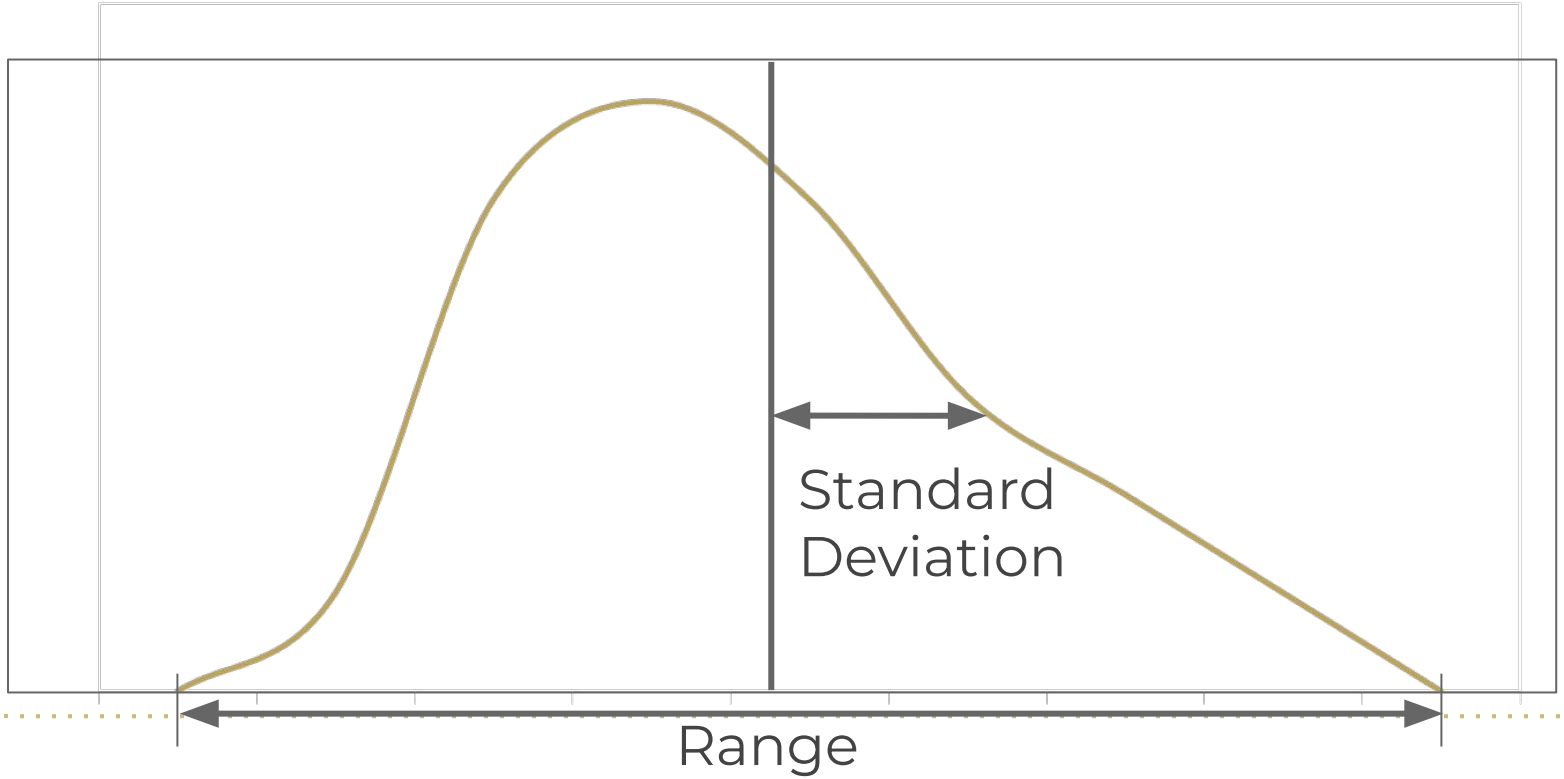
Measures of Spread



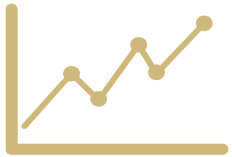
Measures of dispersion reflect the variation or spread in a data set or distribution. Some of the common measures of dispersion are:

- Range, Standard Deviation, Variance

Measures of Spread



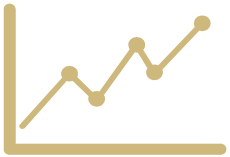
Measures of Shape



Measures of shape reflect the type of distribution sampled. There are two measures:

- Skewness
- Kurtosis

Skewness



Skewness is concerned with the symmetrical nature of the distribution, and is the degree of departure from symmetry of a distribution.

Skewness

Basically measures “lopsidedness.”

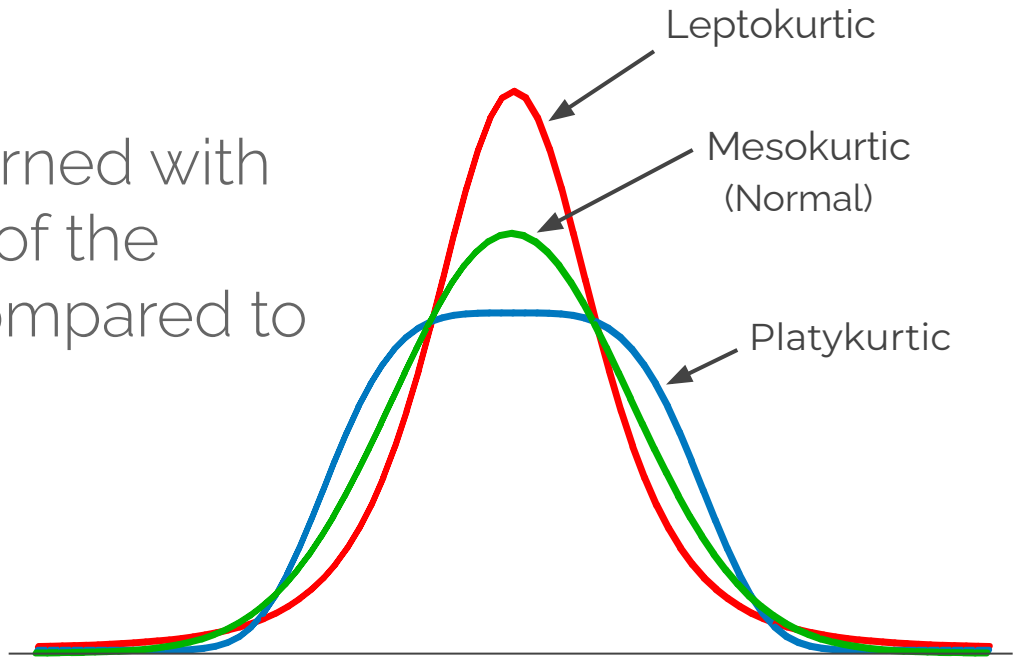
Symmetric distributions have zero skewness.

Positively skewed

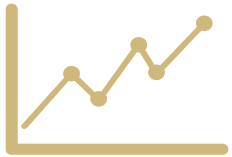
Negatively skewed

Kurtosis

Kurtosis is concerned with the peakedness of the distribution as compared to the tails

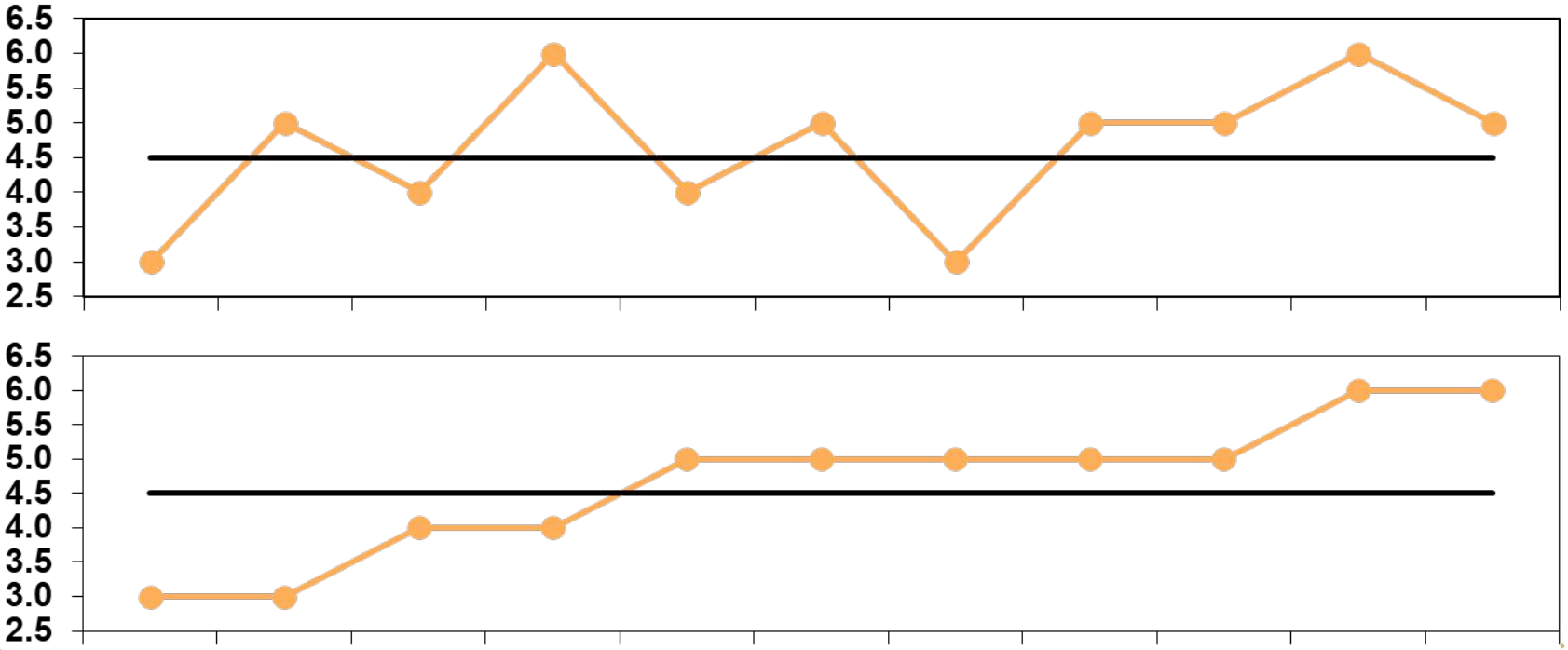


Time Sequence

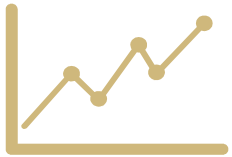


Time ordered data indicates the stability of the process through time.

Time Sequence



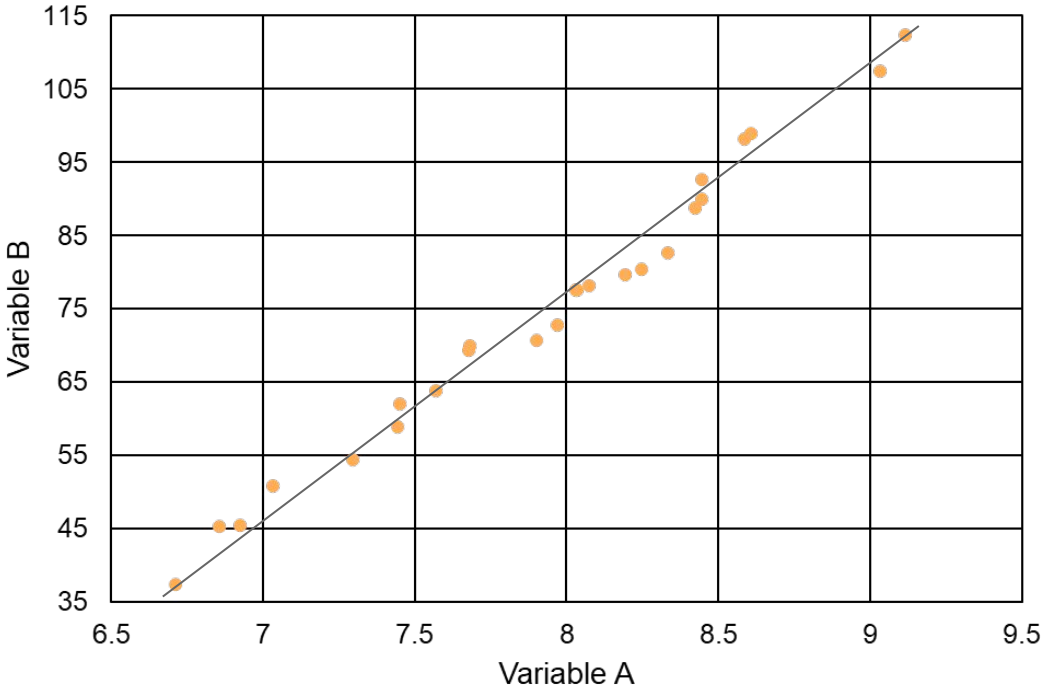
Measures of Relationship



Measures of relationship quantify the “strength” of the relationship between two variables.

- Correlation
- Association

Correlation Example



$Y=-161.8151 + 29.8373X$
Linear: $r^2 = 0.9852$

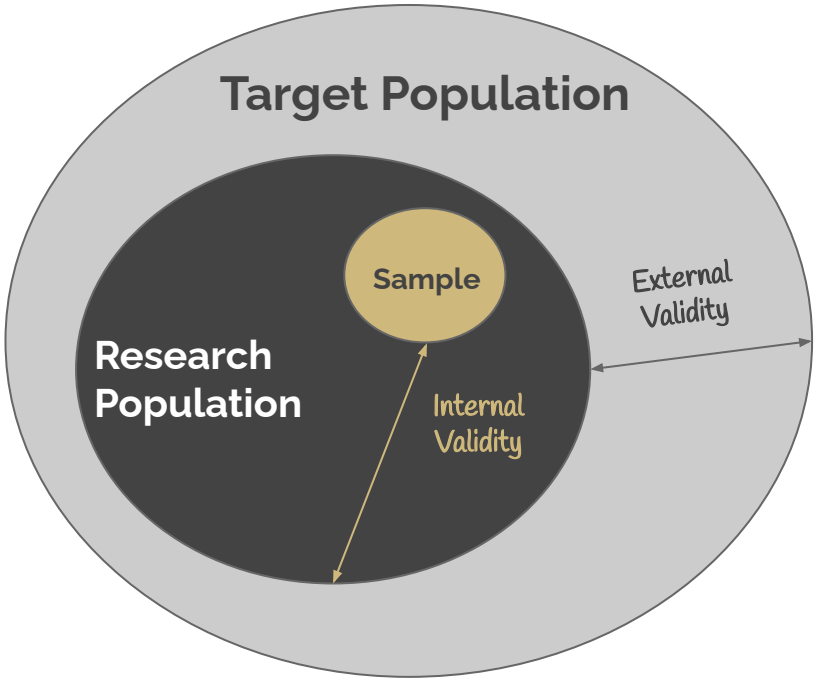
Sampling Concepts

Population vs Sample

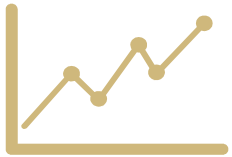
Population: Group of **all** items possessing a common characteristic of interest to a manager

Sample: A representative **portion** of a population that is used to reach conclusions about the population it represents

Population vs Sample

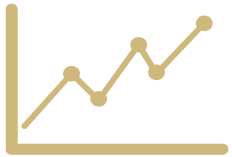


Random Sampling



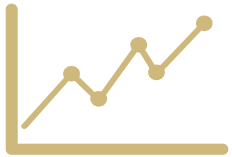
- Every possible sample of size n has an equal chance of being selected
  - Representative
  - Crucial for making generalizations from the sample to the population

Statistics



- A **statistic** is a measure calculated from sample data that may be used to make inferences about a population
  - The average is a “statistic”
  - The range is another “statistic”
  - There are many more...

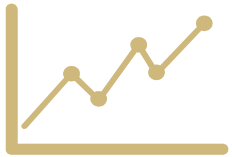
Statistics



Descriptive Statistics

- Describe a characteristic of a sample
- Frequently used to make inferences about population parameters
- Represented by letters in English

Parameters



Population parameters

- Describes a characteristic of the population
- Represented by Greek letters (with few exceptions)

Statistics & Parameters

Sample Statistics	Population Parameters	Description
$\overline{X}$	μ	Mean
$\tilde{X}$	M	Median
s	σ	Standard Deviation
s^2	σ^2	Variance
R	NT'	Range / Natural Tolerance
g_3	γ_3	Skewness
g_4	γ_4	Kurtosis

Statistics & Parameters

Sample Statistics	Population Parameters	Description
$\overline{c}$	**λ**	Items / Unit (Avg. Count)
Q1	Q1'	1st Quartile
Q3	Q3'	3rd Quartile
IQR	IQR'	Interquartile Range
Mode	N/A	Mode
p	π	Proportion