= key words

Definitions of Statistics chapters (1 2 3)

Chapter 1

- **Statistics:** is the science of conducting studies to collect, organize, summarizes, analyze, and draw conclusions from data.
- **A Variable:** is characteristic or attribute that can assume different values.
- **Data:** are the values (measurements or observations) that the variables can assume.
- **Random Variable:** variables whose determined by chance.
- **Data set:** Collection of data values.
- **Datum** Or a **data value:** Each value in the data set
- **A population:** consists of all subjects (human or otherwise) that are being studied.
- **A sample** : is a subset of the population (is a group of subjects selected from a population)
- **Descriptive statistic:** consists of the collection, organization, summarization, and presentation of data.
- **Inferential statistic**: consists of generalizing from samples to populations, performing estimations and hypothesis tests, determining relationships among variables, and making predictions. (and probabilities)
- **Probability**: the chance of event occurring.
- **Qualitative Variables:** are variables that have distinct categories , according to some characteristic or attribute.
- **Quantitative variables:** are variables that can be counted or measured.
- **Discrete Variables:** assume values that can be counted.
- **Continuous Variables:** assume an infinite number of values between any two specific values. They are obtained by measuring. They often include fractions and decimals
- **Nominal level:** classifies data into mutually exclusive, (non-overlapping) categories in which no order or ranking can be imposed on the data.
- **Ordinal level:** classifies data into categories can be ranked. However, precise differences between the ranks do not exist.

بالتوفيق.. استعين بالله.. وفالك الدرجة الكاملة
Go and get it 😊❤

= key words

- **Interval level :** ranks data and precise differences between units of measure do exist, however there is no meaningful zero.
- **Ratio level**: possesses all the characteristics of interval measurement and there exists a true zero. True ratios exist when the same variable is measured on two different subjects.
- **A random sample**: is a sample in which all members of the population have an equal chance of being selected.
- **Systematic sample**: is a sample obtained by selecting every $k\ th$ member of the population where k is a counting number.
- **Stratified sample**: is a sample obtained by dividing the population into subgroups or (strata) according to some characteristic relevant to the study. Then subjects are selected from each subgroup.
- **Cluster sample:** is obtained by dividing the population into sections or clusters and then selecting one or more clusters and using all members in the cluster(s) as the members of the sample.
- **Observational Study:** The researcher merely observes what is happening or what has happened in the past and tries to draw conclusions based on these observations.
- **Experimental Studies:** the researcher manipulates one of the variables and tries to determine how the manipulation influences other variables.
- when random assignment is not possible and a researcher uses intact groups, then he is performing a **quasi-experimental study**
- **Treatment group:** is a group that receives the special instruction/specific treatment.
- **Control group:** does not receive the special instruction/specific treatment.
- **Independent or Explanatory variable:** is the variable that is being manipulated by the researcher
- **Dependent or Outcome variable:** is the resultant variable.
- **A confounding variable**: is the variable that influences the dependent or outcome variable but was not be separated from the independent variable.

Chapter 2

بالتوفيق.. استعين بالله.. وفالك الدرجة الكاملة
Go and get it 😊❤

= key words

- **Raw data:** When data are collected in original form
- **Frequency distribution**: is the organization of raw data in table form, using classes and frequencies.
- **Class**: each raw data value is placed into a quantitative or qualitative category.
- **Frequency**: the number of data values contained in a specific class
- **Categorical Frequency Distribution:** Used for data that can be placed in specific categories, such as nominal- or ordinal-level data.
- **Grouped Frequency Distribution:** when the range of values in the data set is very large. The data must be grouped into classes that are more than one unit in width.
- **Class limits:** represent the smallest and largest data values that can be included in a class.
- **The class boundaries**: are used to separate the class so that there is no gap in frequency distribution.
- **Cumulative frequency**: the sum of the frequencies accumulated up the upper boundary of a class in the distortion.
- **Cumulative frequency distribution**: is a distribution that shows the number of data values less than or equal a specific value (usually an upper boundary)
- **The histogram**: is a graph that displays the data by using contiguous vertical bars (unless the frequency of a class is 0) of various heights to represent the frequencies of the classes.
- **The frequency polygon**: is a graph that displays the data by using lines that connect points plotted for the frequencies at the midpoints of the classes. The frequencies are represented by the heights of the points.
- **The ogive**: is a graph that represents the cumulative frequencies for the classes in a frequency distribution.
- **A bar graph**: represents the data by using vertical or horizontal bars whose heights or lengths represent the frequencies of the data .

بالتوفيق.. استعين بالله.. وفالك الدرجة الكاملة
Go and get it 😊❤

= key words

- **A Pareto chart**: is used to represent a frequency distribution for a categorical variable, and the frequencies are displayed by the heights of vertical bars, which are arranged in order from highest to lowest.
- **A time series graph (line chart)**: represents data that occur over a specific period of time.
- **A pie graph**: is a circle that is divided into sections or wedges according to the percentage of frequencies in each category of the distribution.
- **A dotplot**: is a statistical graph in which each data value is plotted as a point (dot) above the horizontal axis.
- **A stem and leaf plot:** is a data plot that uses part of a data value as the stem and part of the data value as the leaf to form groups or classes.

Chapter 3

- **A statistic**: Is a characteristic or measure obtained by using the data values from a sample.
- **A parameter:** Is a characteristic or measure obtained by using all the data values from a specific population.
- **The mean:** is the sum of the values, divided by the total number of values.
- **The median**: is the halfway point in a data set.
- **The median**: is the midpoint of the data array.
- **Data array**: is an ordered data set.
- **Mode**: The value that occurs most often in a data set
- **The midrange:** is defined as the sum of the lowest and highest values in the data set, divided by 2.
- **The range:** is the difference between the highest and lowest values in a data set.
- **The variance:** is the average of the squares of the distance each value is from the mean.
- **The standard deviation:** is the square root of the variance.

بالتوفيق.. استعين بالله.. وفالك الدرجة الكاملة
Go and get it 😊❤

= key words

- **The coefficient of variation** denoted by Cvar: is the standard deviation divided by the mean, expressed as a percentage.
- **Quartiles**: divide the data set into 4 equal groups.
- **The interquartile range (IQR):** is the difference between the third and first quartiles.
- **An outlier:** is an extremely high or an extremely low data value when compare with the rest of the data values

بالتوفيق.. استعين بالله.. وفالك الدرجة الكاملة
Go and get it 😊❤