

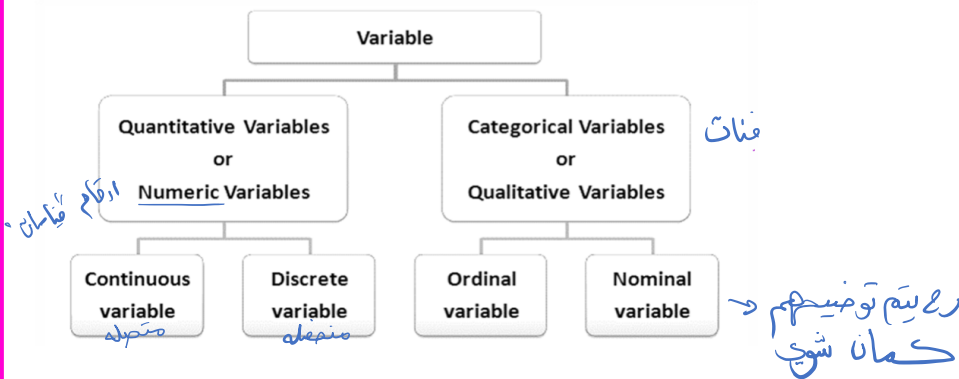
Done by : Farah Hasant

Lecture 2

# Variables

- 1) is an object, characteristic, or property that can have different values.
- 2) A quantitative variable can be measured in some way.
- 3. A qualitative variable is characterized by its inability to be measured but it can be sorted into categories.

لنوع لا يمكن قياسها بل يمكن تقسيمها إلى فئات



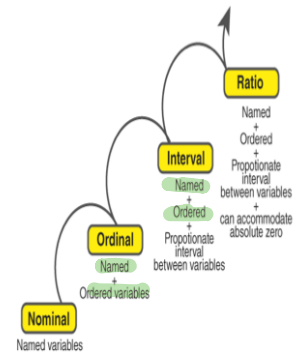
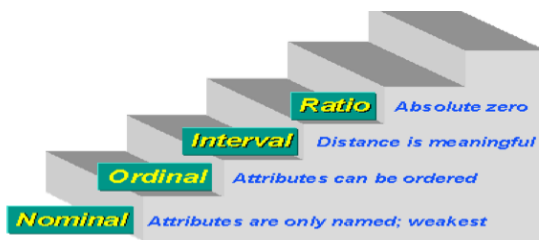
## Types of Variables

- Independent variable — the presumed cause (of a dependent variable)
- Dependent variable — the presumed effect (of an independent variable)

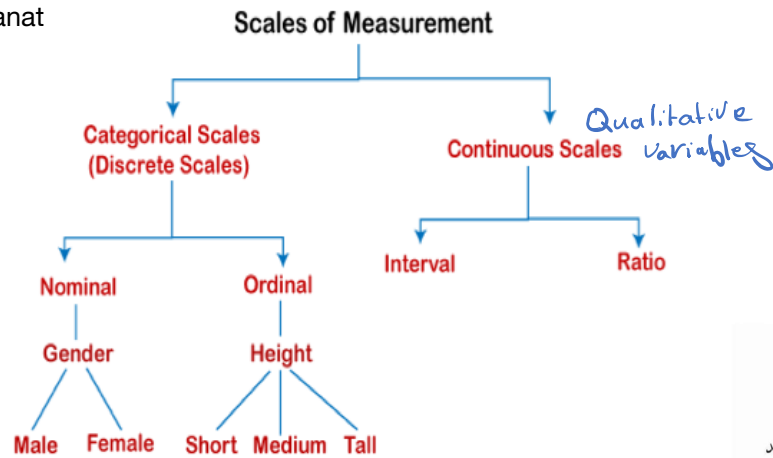
بغضب  
نتيجة تأثير independent  
بغضب انه لانه يكون له في depend ← يؤدي إلى حدوث effect

## Levels of Measurement: Nominal ,Ordinal , Interval and Ratio.

○



Done by : Farah Hasanat



## A. Nominal Level of Measurement :

- Categories that are distinct from each other such as gender, religion, marital status.
- They are symbols that have no quantitative value. → دين ما يعني كمي بس نومي
- Lowest level of measurement.
- Many characteristics can be measured on a nominal scale: race, marital status, and blood type.
  - Dichotomous.
- Appropriate statistics: mode, frequency
- We cannot use an average. It would be meaningless here.

### Examples of Nominal Data :

- Gender
- Hair color
- Housing style
- Marital status
- Ethnicity ( hispanic , asian )

© لا يمكن تحديد الفروق الدقيقة بين الرتب لأنها تشير إلى الطلب بدلاً من الكمية الدقيقة.

## B. Ordinal Level of Measurement → تصنيفاً حسب ارقام

- The exact differences between the ranks cannot be specified such as it indicates order rather than exact quantity.

- Involves using numbers to designate ordering on an attribute

بستخدم ارقام / عبارات تصف درجات لتصميم شيء ما

➤ Example: anxiety level: mild, moderate, severe. Statistics used involve frequency distributions and percentages.

➤ Appropriate statistics: same as those for nominal data, plus the median; but not the mean.

C. Interval level of Measurement : → فيها ارقام حقيقية Done by : Farah Hasanat

- They are real numbers and the difference between the ranks can be specified.
- Equal intervals, but no “true” zero.
- Involves assigning numbers that indicate both the ordering on an attribute, and the distance between score values on the attribute
- They are actual numbers on a scale of measurement.
- Example: body temperature on the Celsius thermometer as in 36.2, 37.2 etc. means there is a difference of 1.0 degree in body temperature.
- Appropriate statistics
- same as for nominal ◦ same as for ordinal plus, ◦ the mean

**Examples of Interval Scales**

**Example 1:**  
How likely are you to recommend the Santa Fe Grill to a friend? Definitely Will Not Recommend Definitely Will Recommend  
1 2 3 4 5 6 7

**Example 2:**  
Using a scale of 0–10, with “10” being Highly Satisfied and “0” being Not Satisfied At All, how satisfied are you with the banking services you currently receive from [read name of primary bank]?  
Answer: \_\_\_\_\_

**Example 3:**  
Please indicate how frequently you use different banking methods. For each of the banking methods listed below, circle the number that best describes the frequency you typically use each method.

Banking Methods	Never Use	Use Very Often									
Inside the bank	0	1	2	3	4	5	6	7	8	9	10
Drive-up window	0	1	2	3	4	5	6	7	8	9	10
24-hour ATM	0	1	2	3	4	5	6	7	8	9	10
Debit card	0	1	2	3	4	5	6	7	8	9	10
Bank by mail	0	1	2	3	4	5	6	7	8	9	10
Bank by phone	0	1	2	3	4	5	6	7	8	9	10
Bank by Internet	0	1	2	3	4	5	6	7	8	9	10

D. Ratio level of Measurement

- is the highest level of data where data can categorized, ranked, difference between ranks can be specified and a true or natural zero point can be identified.
- A zero point means that there is a total absence of the quantity being measured.
- All scales, whether they measure weight in kilograms or pounds, start at 0. The 0 means something and is not arbitrary (SUBJECTIVE).
- Example: total amount of money.

Question 3  
Which age group do you fall into?

15 - 25 years  A    26 - 35 years  B    36 - 50 years  C

51 - 65 years  D    Greater than 66  E



**What Type of Data To collect?**

- The goal of the researcher is to use the highest level of we measurement possible :
- A ) Is nominal, so the best we can get from this data are **frequencies**. asking (A)Do you smoke? !Yes !No
- (B) is ratio, so we can compute: **mean, median, mode, frequencies**. (B)How many cigarettes did you smoke in the last 3 days (72 hours)?

Done by : Farah Hasanat

## Parameter and Statistic

- Parameter** is a descriptive measure computed from the data of the **population**. **whole population**
- The population mean,  $\mu$ , and the population standard deviation,  $\sigma$ , are two examples of population parameters.
- If you want to determine the population parameters, you have to take a **census** of the entire population.
- Taking a census is very costly. → اجراء لتعداد مطلق للقائمة تعداد سكاني
- Parameters are numerical descriptive measures corresponding to populations.
- Since the population is not actually observed, the parameters are considered unknown constants. اذلا ما قدرنا اعداد لهكان يكون para غير ثابت "غير محدد"

**Statistic** is a descriptive measure computed from the data of the **sample**. **STATISTIC describes only a sample of the population.**

- For example, the sample mean,  $\bar{X}$ , and the standard deviation,  $s$ , are statistics.
- They are used to estimate the population parameters.

• *Statistics: It is a branch of applied mathematics that deals with collecting, organizing, & interpreting data using well defined procedures in order to make decisions.*

**Types of Statistics:** Descriptive Statistics, Inferential Statistics

**A) Descriptive Statistics.** It involves <sup>①</sup>organizing, <sup>②</sup>summarizing & <sup>③</sup>displaying data to make them more understandable.

**A. Measures of Location**

**Measures of Central Tendency:**

◦ 1) Mean ◦ 2) Median ◦ 3) Mode

◦ **Measures of noncentral Tendency- Quantiles:**

◦ Quartiles ◦ Quintiles ◦ Percentiles. نسبة مئوية



Done by : Farah Hasanat

## B. Measure of Dispersion (Variability):

- Range
- Interquartile range → **بين ربعي**
- Variance
- Standard Deviation
- Coefficient of variation

## c . Measures of Shape:

- Mean > Median-positive or right Skewness **انحراف**
- Mean = Median- symmetric or zero Skewness
- Mean < Median-Negative of left Skewness

**حرف الكبر**  
**وذن الكبر**  
**... لثقل لثما بلبي**



## ○ **B) Inferential Statistics.** **احصاء استنتاجي**

- It reports the degree of confidence of the sample statistic that predicts the value of the **population parameter**
- Is the procedure used to reach a conclusion about a population based on the information derived from a sample that has been drawn from that population.
- Are used to test **hypotheses** (prediction) about relationship between variables in the population. A relationship is a bond or association between variables.
- An important aspect of statistical inference involves **reporting the likely accuracy**, or of **confidence** of the sample statistic that predicts the value of the population parameter.
-

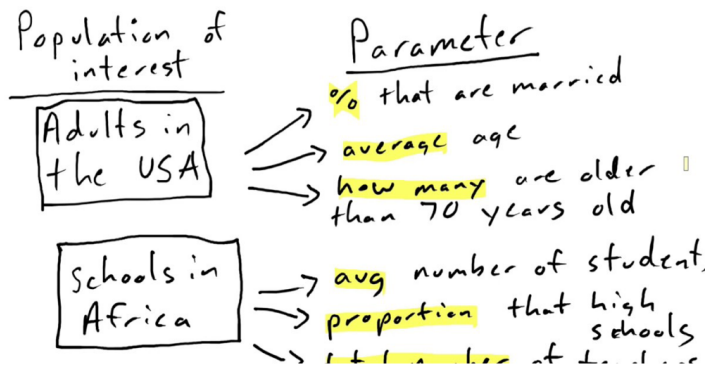
Done by : Farah Hasanat

## Bivariate Parametric Tests: ثنائي المتغير

- One Sample t test (t)
- Two Sample t test (t)
- Analysis of Variance (ANOVA (F).) ارتباط
- Pearson's Product Moment Correlations (r).
- Nonparametric statistical tests: Nominal Data:
- Chi-Square Goodness-of-Fit Test
- Chi-Square Test of Independence
- Nonparametric statistical tests: Ordinal Data:
- Mann Whitney U Test (U)
- Kruskal Wallis Test (H)



## Examples of Parameters



## Research Hypothesis

- A tentative prediction or explanation of the relationship between two or more variables. توقع أو تفسير مؤقت للعلاقة بين متغيرين
- It's a translation of research question into a precise prediction of the expected outcomes ترجمة سؤال البحث إلى توقع دقيق للنتائج المتوقعة
- In some way it's a proposal for solution/s يقترح حل
- In qualitative research, there is NO hypothesis نوعي ماهي

Done by : Farah Hasanat

## ○ Research Hypothesis (شروطه)

- ➤ States a prediction
- ➤ Must always involve at least **two** variables
- ➤ Must suggest a predicted relationship between the independent variable and the dependent variable
- ➤ Must contain terms that indicate a relationship (e.g., more than, different from, associated with).

مصادر لفرقيات

## ○ Hypotheses Criteria


لكتابة كل شكل اعلاوي

- ■ Written in a declarative form.
- ■ Written in present tense.
- ■ Contain the population ✓
- ■ Contain variables. ✓
- ■ Reflects problem statement or purpose statement.
- ■ Empirically testable. → قابل للاختبار تجريبيا

مكتوبه بصيغة الحاضر

اللهم صل وسلم على نبينا محمد

يعكس  
بيان مشكلة  
بيان غرض

**Note**  : A hypothesis is made about the value of a parameter, but the only facts available to estimate the true parameter are those provided by the sample. If the statistic differs (and of course it will) from the hypothesis stated about the parameter, a decision must be made as to whether or not this difference is significant. If it is, the hypothesis is rejected. If not, it cannot be rejected.

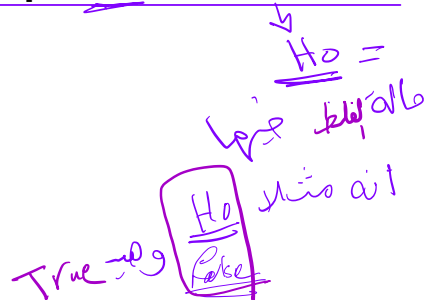
يتم وضع فرضية حول قيمة parameter، ولكن الحقائق الوحيدة المتاحة لتقدير parameter الحقيقية هي تلك التي قدمتها sample. إذا statistic اختلف (وسيختلف بالطبع) عن لفرضية المذكورة حول parameter، فيجب اتخاذ قرار بشأن ما إذا كان هذا الاختلاف مهماً أم لا. إذا كان الأمر كذلك، يتم رفض الفرضية. إذا لم يكن كذلك، فلا يمكن رفضه.

- $H_0$ : The null hypothesis. This contains the hypothesized parameter value which will be compared with the sample value.
- $H_1$ : The alternative hypothesis. This will be “accepted” only if  $H_0$  is rejected. ✓

Technically speaking, we never accept  $H_0$  What we actually say is that we do not have the evidence to reject it.

Two types of errors may occur:

- $\alpha$  (alpha) and  $\beta$  (beta). The  $\alpha$  error is often referred to as a Type I error and  $\beta$  error as a Type II error.
- You are guilty of an alpha error if you **reject  $H_0$**  when it really is true.   
 *H<sub>1</sub> =*   
 *← انه انما يري H<sub>1</sub> انما يمزج حاسب H<sub>0</sub> True*
- You commit a beta error if you “accept”  $H_0$  when it is false.





Done by : Farah Hasanat

		STATE OF NATURE	
		H <sub>0</sub> Is True	H <sub>0</sub> Is False
DECISION	Do Not Reject H <sub>0</sub> <i>Handwritten: Ho ✓</i>	GOOD	β Error (Type II Error)
	Reject H <sub>0</sub> <i>Handwritten: Hi ✓</i>	α Error (Type I Error)	GOOD

# Types of Errors

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ  
 اللهم صل وسلم على نبينا محمد

If You.....	When the Null Hypothesis is...	Then You Have.....
Reject the null hypothesis <i>Handwritten: Hi ✓</i>	True (there really are no difference)	Made a Type I Error
Reject the null hypothesis <i>Handwritten: Hi</i>	False (there really are difference)	☺
Accept the null hypothesis <i>Handwritten: Ho</i>	False (there really are difference)	Made Type II Error
Accept the null hypothesis <i>Handwritten: Ho</i>	True (there really are no difference)	☺

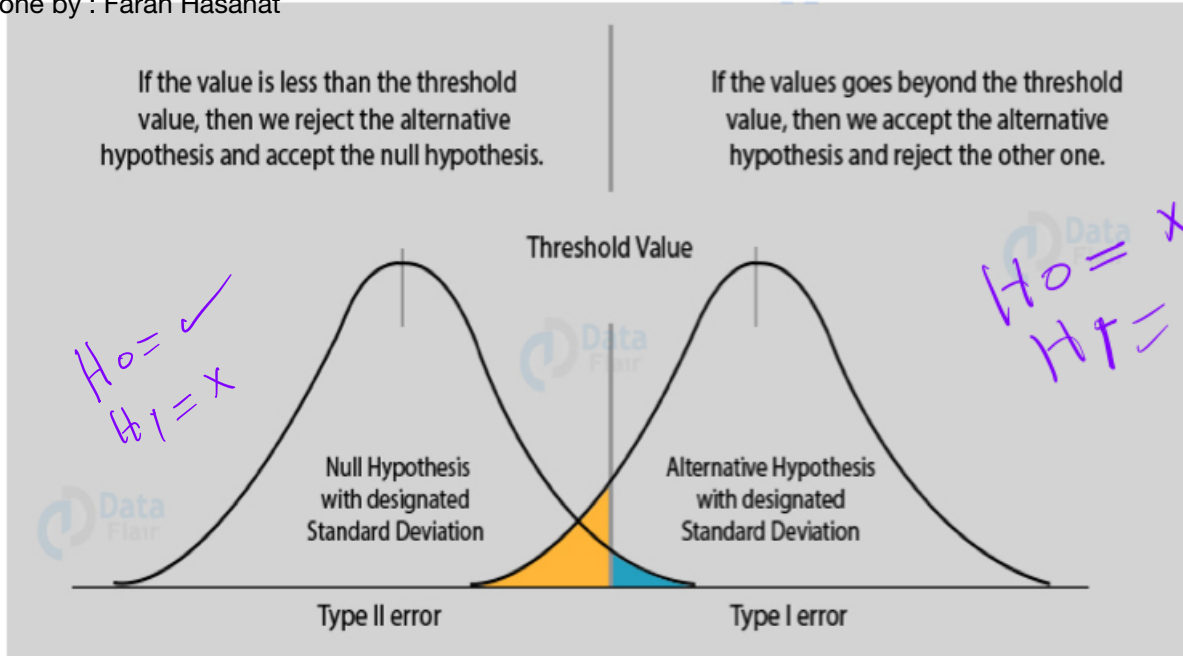
	H <sub>0</sub> rejected <i>Handwritten: Hi ✓</i>	Fail to reject H <sub>0</sub> = H <sub>0</sub>
H <sub>0</sub> false	Correct	Type II error
H <sub>0</sub> true	Type I error	correct

Alpha ( $\alpha$ ) = Prob (Type I error)

Beta ( $\beta$ ) = Prob (Type II error)

Power =  $1 - \beta$

Done by : Farah Hasanat



1. Formulate  $H_0$  and  $H_1$ .  $H_0$  is the null hypothesis, a hypothesis about the value of a parameter, and  $H_1$  is an alternative hypothesis.  
e.g.,  $H_0: \mu=12.7$  years;  $H_1: \mu \neq 12.7$  years
2. Specify the level of significance ( $\alpha$ ) to be used. This level of significance tells you the probability of rejecting  $H_0$  when it is, in fact, true. (Normally, significance level of 0.05 or 0.01 are used)
3. Select the test statistic: e.g., Z, t, F, etc.
4. Establish the critical value or values of the test statistic needed to reject  $H_0$ . DRAW A PICTURE!
5. Determine the actual value (computed value) of the test statistic.
6. Make a decision: **Reject  $H_0$**  or **Do Not Reject  $H_0$** .

## Steps in Hypothesis Testing

### HYPOTHESIS TESTING



اللهم صل وسلم على نبينا محمد