

- \* Quantitative
- Numerical Data - Two Types (discrete - continuous)
  - Discrete - (counting) → <sup>عددي</sup> ← <sup>متردد</sup> <sup>متقطع</sup>
  - Continuous - (measurement) → <sup>قياسي</sup> ← <sup>متصل</sup> <sup>مستمر</sup>
- discrete → <sup>متقطع</sup> 8 cats  
8.5 cats X
- Continuous → <sup>مستمر</sup> 5.56 m  
distance - speed - weight

- \* Qualitative
- Descriptive data based on observations  
use word to describe this data → <sup>وصفي</sup> <sup>كلمات</sup>
  - Involves 5 senses
  - see, feel, taste, hear, smell

Color / Soft / hard / low / high

### \* Nominal scale data

1. Qualitative / Categorical
2. Names, colors, labels, Gender, etc.
3. order does not matter

بيانات اسمية لا ترتب

1. Red → 5 50%
2. Blue → 3 30%
3. Green → 2 20%

### \* ordinal scale data

1. Ranking / Placement
2. The order matters
3. Differences cannot be measured

بيانات الترتيب

لا يمكن قياس الاختلافات  
لأنها ليست بالبيانات

- |     |       |
|-----|-------|
| 1st | 41.53 |
| 2nd | 41.55 |
| 3rd | 41.02 |



بيانات القياس الكمي

# \* Interval Scale data

1. The order matters
2. Differences can be measured (except ratios)
3. No True "0" starting Point

30 °F → 30 °F  
60 °F  
90 °F

# \* Ratio Scale Data

بيانات القياس الكمي النسبي

1. The order matters
2. Differences are measurable (including ratios)
3. Contains a "0" starting Point

30, 56, 70, 82, 90

$$\frac{90}{30} = 3$$

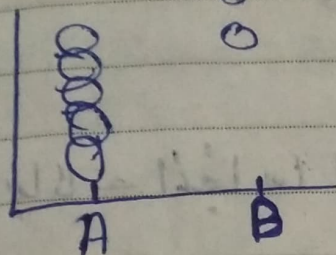
نقطة الصفر  
نسبة

Data	Nominal	ordinal	Interval	Ratio
labeled	✓	✓	✓	✓
meaningful order	X	✓	✓	✓
measurable Difference	X	X	✓	✓
true zero starting Point	X	X	X	✓



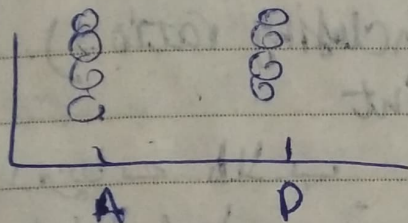
## Hypothesis testing and Null Hypothesis

عندما دوا A ودوا B فعند انهاء علاج ٢ اشخاص دوا A و ٢ اشخاص دوا B ونقوى فرضيات الوقت الى هبطوا فيه فدائيه ولو يردو ما عينا فتره



There is ~~no~~ difference in recovery time

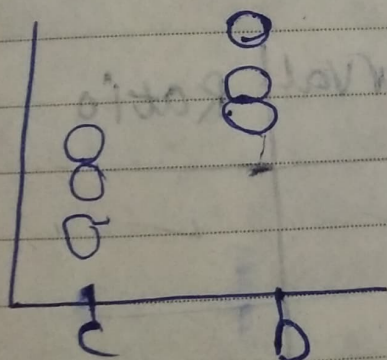
لاي فرق بين هذين المجموعتين كبيره من الناحية



نقر بانه نفس الوقت  
Null Hypothesis

## Alternative Hypothesis

فرضيات بديله



مما يعني اننا قد نرفض  
reject or fail to reject the  
Null Hypothesis

A statistical Test  
needs 3 things

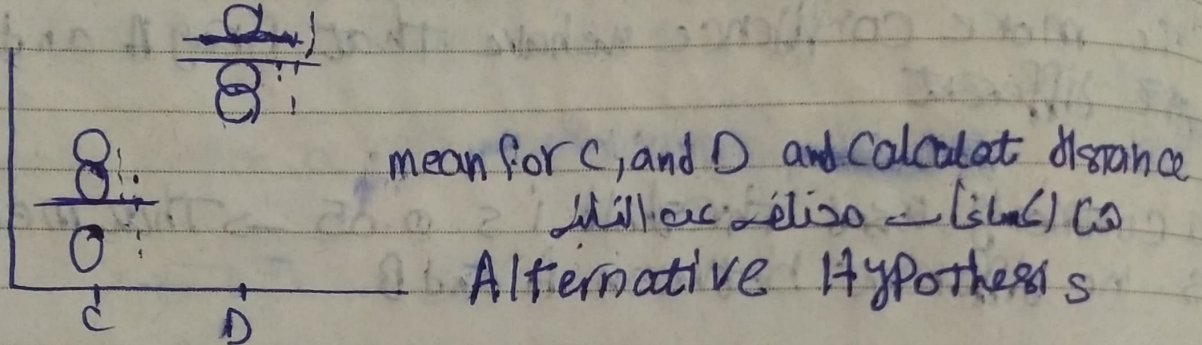
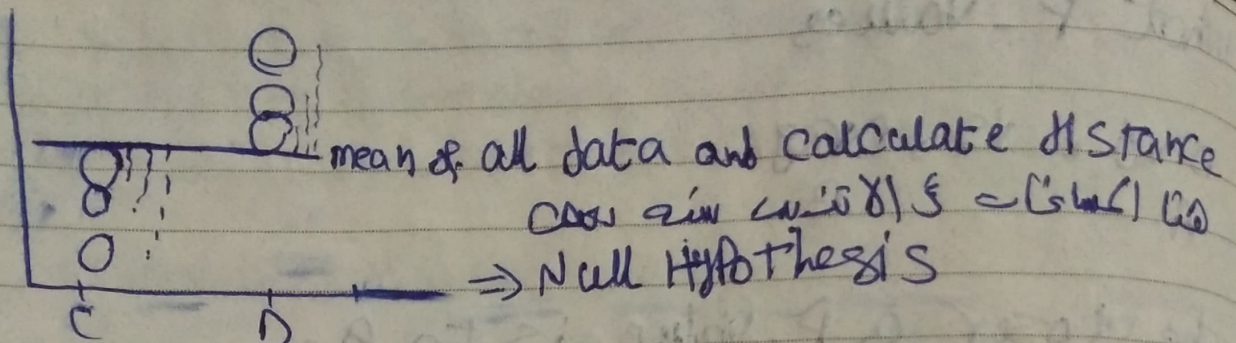
الاشيا

- 1) it's need the data.
- 2) it needs a null, or primary Hypothesis (i.e. it needs something to reject or fail to reject)
- 3) it needs Alternative Hypothesis

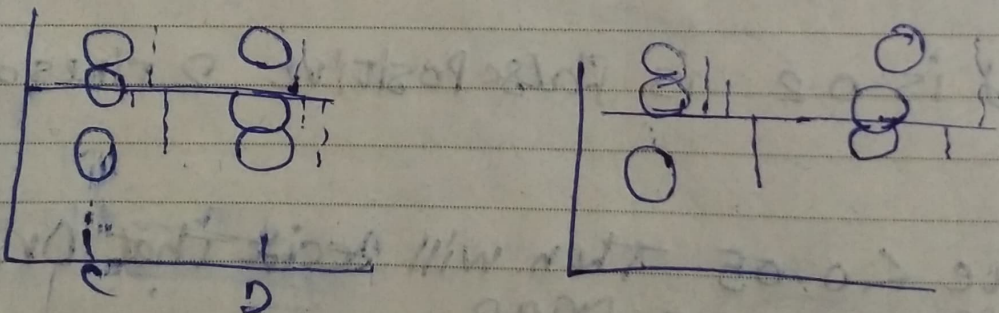
Alternative Hypothesis: There is a difference in recovery times between Drug C and Drug D

Null Hypothesis: There is no difference in recovery time between Drug C and Drug D

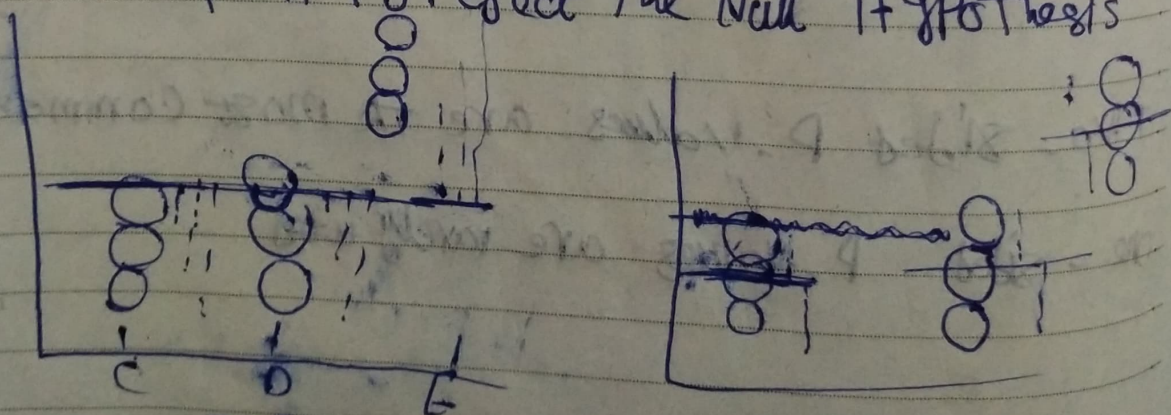




reject the Null Hypothesis or mix



fail to reject the Null Hypothesis



Drug C, D are the same

Drug E doing it's own thing  
 reject or fail to reject the Null Hypothesis



## What P-Values

Drug A  $\longleftrightarrow$  Drug B

The closer a P-value is to 0, the more confidence we have that Drug A and B are different

a commonly used threshold is 0.05  $\rightarrow$  That means no difference between A and B

getting a small P-value when there is no difference is called a False Positive

threshold is 0.2 is False Positive 2 times out of 10

P-value  $< 0.05$  then will decide that Drug A is different from Drug B

Two-sided P-values are the most common

one-sided P-values are rarely used





P-value is Composed of Three Parts.

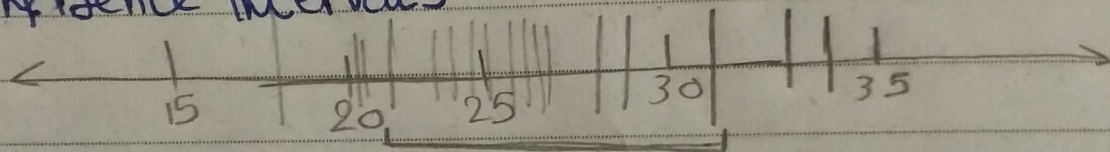
- ① The Probability random chance would result in the observation
- ② The Probability of observing something else that is equally rare
- ③ The Probability of observing something rarer or more extreme

$$P\text{-value for 2 Heads} = 0.25 + 0.25 + 0 = 0.5$$

We only reject a hypothesis if the P-value is less than 0.05

$0.5 > 0.05$ , we fail to reject the hypothesis

### \* Confidence Intervals



A 95% Confidence Interval is

just an interval that covers

95% of the means

We know that anything outside of it occurs less than 5% of the time.

That is to say, the P-value of anything outside of the confidence interval is  $< 0.05$



regression analysis

Winston's Categorical, Categorical, Continuous

variables

variables