لجان الدُّفعات

DOPAMINE ACADEMY

Statistics

Subject: Lecture 2

25-7-2024

▸ Subject : .....

Types of Data

- **Qualitative** نوعي مافي ارقام
  - **Nominal** "Name"
    ex: eye color
    female / male
    UnOrdered
    "Just description"
    * يعني وصف بدون ترتيب
  - **Ordinal**
    ex: • degree of pain
    • grade of malignancy
    . درجة خطورة الورم
    Ordered
    * مافي ارقام بس في ترتيب
- **Quantitative** كمي في ارقام
  - **continuous**
    Data with fraction
    يعني الوزن من هذول
    يكون عدد صحيح يعني ممكن
    يكون 47.5 يعني في اعشار
    No gaps
    there are gaps
  - **Discrete**
    Data without fraction
    Integer
    يعني عدد الطلاب عن
    المحاضرة عدد صحيح
    طالب مثلاً .

Nominal data presentation :-

1. Counts in each category حسب
2. Histogram رسم بياني
3. Relative frequency
4. Relative frequency diagram
5. pie Chart رسم بياني دائري

• Relative frequency = $\frac{N}{\text{total}}$

• Σ Relative frequency = 1
مجموع

Ordinal data presentation :-

1. tools of nominal data
2. Cumulative relative frequency
3. Cumulative relative frequency polygon
4. percentiles

	N	Relative freq	Cumulative Relative freq
	68	0.34	0.34
		x / =	
	92	0.46	x

Subject : ..... / /

Percentiles : divide a distribution into equal or ordered parts.

Median : 50th percentile divides a population into 2 equal size parts.

لو حكيتلك إنه median لعلاماتكم 20 يعني 50% جابوا أعلى من 20 و 50% جابوا أقل من 20.

لو حكيتلك $P_{20} = 18$ يعني 20% جابوا أقل من 18 و 80% جابوا أعلى من 18

لو حكيتلك $P_{75} = 23$ يعني 75% جابوا أقل من 23 و 25% جابوا أعلى من 23

Descrete data presentation :-

1. tools of ordinal data

2. mean = sum of observations / total N ← هاد القانون لبس بيكون عندك raw data

ex: 5 , 6 , 7.2 , 10.4 , ... $\frac{5+6+7.2+10.4}{4} = \ldots$

"Risk factor"

# of RF	N
0	10
1	15
2	25

table → mean = $\frac{(0 \times 10)+(1 \times 15)+(2 \times 25)}{50}$

σ σ^2

* Indicator of Dispersion :- Standard deviation , variation

Median "central tendency" : 1. order observation (smallest → largest)

A) 1 2 3 4 5 (N)

5 , 10 , (15) , 20 , 30 odd 2. odd : $(N+1)/2$ → median الرقم الناتج هو موقع

$\frac{N+1}{2} = \frac{6}{2} = 3$ موقع median

even : ! /2 ← القيمتين اللي بالوسط

B)

5 , 10 , 15 , 30 , 40 , 20 even $\frac{15+20}{2} = \frac{35}{2} = 17.5$

5 10 15 20 30 40 ✓

+

2

▶ Subject : .....

عندك Data معينة وانت قاعد تعملها خريطة بـ ان حين راح تأثر ؟

الأصل 4.02 2.03 3.08

أحوال حين !

اكبر ال mean ⑪ لأنه بتحتد على مجموع

التل 40.2 2.03 3.08

أما ال median ترتيب بحيث راح يتأثر

اشلك !

mean $\frac{4.02 + 2.03 + 3.08}{3} = \boxed{3.04}$ $\frac{40.2 + 2.03 + 3.08}{3} = \boxed{15.1}$

median 2.03 (3.08) 4.02 2.03 (3.08) 40.2

$\frac{N+1}{2} = \frac{4}{2} = 2$ $\frac{N+1}{2} = \frac{4}{2} = 2$

Continuous data presentation :-

1. Central tendency (median)
2. dispersion
   - Range — largest value - Smallest value — • حسب الترتيب
   - Interquartile range — $Q_3 - Q_1 = P_{75} - P_{25}$
   - Standard deviation — square root of the variance
     (∓) ناتج الجذر التربيعي
     • sum of the Squared deviation $(x - mean)^2$ $(N-1)$

Importance of ordered array

1. determine quickly the value of smallest measurement
2. the value of the largest measurement.

Interval width = Range / K → number of interval

mid interval = $\frac{\text{largest} + \text{smallest}}{2}$ المنتصف الوحدة