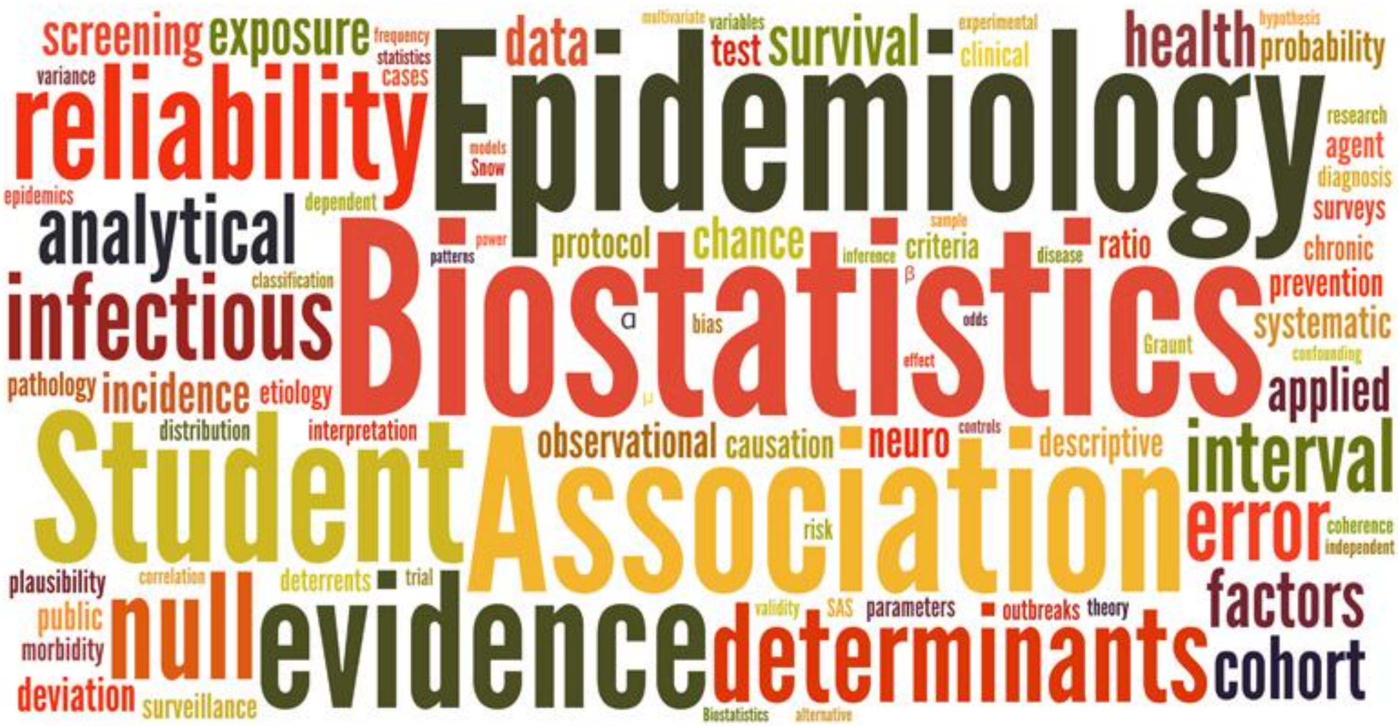
University of Jordan
Faculty of Medicine



☑ Sheet (Notes) ☐ Slide ☐ Handout

Number: 10

Done By: Dana Rida

Subject: Organizing and Displaying Data

Doctor: Mohammad Nassar

Organizing and Displaying Data

Lec 2

Levels of measurements

The researcher chooses his own level to measure his own variables in his study, for example if he is doing a questionnaire to measure the quality of life for patients, what level of measurement will he use? This is what this lecture all about.

Four levels (from lowest to the highest)

1-nominal : the numbers given to the characteristics are not given for their numerical value but as a code for the variable. *not in book nor slides* We have 2 types on nominal measurements, first, nominal dichotomous, the variables given 1or 2 codes , 0 or 1, 2 or 3 , so what is important is that I only use 2 codes (e.g male=0, female=1 it doesn't mean females are more by one) numbers don’t mean anything just for statistical purposes, , if coding exceeds 2 variables it is then nominal categorical level (e.g medical diagnosis; give codes for patients come from medical, surgical, neuro so on, codes are given to them as 1,2,3, the numbers don’t mean their selves they are just codes they neither point to a status , so 3 doesn’t mean has a worse problem than 1.

2-ordinal: the numbers here have a little bit meaning, e.g at the end of the semester we evaluate the doctors, the questionnaire would be from 1-5 (from rare to always) so the numbers here have a certain meaning but not the exact value of the number (so if I choose 4 it is not the double of 2) so these numbers we only use their ordering meaning (e.g pain intensity, when you ask your patient to state their level of pain, but we cannot get the exact difference between patients) . The whole point of coding data is to make the statistical procedure easier, dealing with numbers easier than words.

3-interval: numbers should be equal they have their real meaning , but there is no absolute zero, e.g temperature even if I got a zero that doesn’t mean I don’t have heat, so it's not a real zero

4-ratio, here there could be an absolute zero. All lab results are considered ratio. (weight, hemoglobin etc..)

I can't find age *zero* or weight *zero,* pulse ..etc but we use interval and ratio to strengthen our results, but usually it is used in chemistry, physics and math since they are theoretical.

We use something called robust statistics or parametric statistics, the main assumption to use it is to utilize the interval and ratio level of measurements. Interval and ratio almost considered the same, named as continuous or scale level , both mainly used in experimental research, because they need highest level of statistics.

Subjects	Nominal	Ordinal	Interval	Ratio
1	2	4	70	180
2	1	1	0	110
3	2	3	55	165
4	1	2	20	130

we have here 4 subjects, and in the ratio column is their weight, their weight is considered a ratio level (since there is an absolute zero) when converting them to a lower level we used a cutoff point, <= 130 we gave them code 1, >=130 code 2 , which this is a nominal level and we can consider it a dichotomous type (non-parametric), because I only used 2 codes. in nominal it is harder to compare between patients, imagine if had 400 patients and most of them were above 130 it will be hard to compare between them since they all have the same code. You are free to choose your cut off point it is just an assumption, but in this example we chose 130. Now to convert it from ratio level to ordinal, simply ranking them, so number 110 took rank 1 (an ascending ranking, starting from the least, we can do the opposite) 130>>2 and so on. This ranking is important later on to get the median. Now to convert it to interval , X-(observed weight)= interval, in this example x was 110 (a criterion ,it was the lowest weight) >> 110-(observed weight which is 110) = 0 , and so on to the other weights.

So from here we get that when you are at the highest level of measurement you can change yourself to lower levels, but not the opposite way around, so when collecting data try your best to collect at the highest level.

Other examples .. when I say 39% people have insurance, this is a nominal dichotomous because I can only get from this if they are insured or not. At the lowest level of measurements you have got a general idea, to get to specific ideas you need higher levels.

An important example

Table (1): Factors associated with the development of hematoma. (n = 66)

Factors	*Hematoma* Yes(n=6)	*Hematoma* No(n=60)	*P value*
Prophylactic anticoagulants (Heparin)			
Yes	6	31	
No	0	29	0.031*
Smoking			
Yes	0	14	0.335
No	6	46	
Medical illness			
Present	1	15	0.548
Not present	5	45	
Menstrual phase			
phase I (periovulatory)	1	26	0.38
phase II (perimenstrual)	4	28	0.42
phase III (amenorrhea)	1	6	0.45

*

The prophylactic coagulant at which level the researcher measured it ?

Since here I can see that collecting this data depended on 2 answers yes or no(1 or 0), (forget about the numbers here concentrate on the variables) so this is a nominal dichotomous. Also smoking and medical illness both are nominal dichotomous.

what about the menstrual phase? Since it has more than 3 answers and it has nothing to do with ordering (it doesn't show measure of danger or pain) it is a nominal categorical. If the menstrual phase is a crucial thing or a crucial variable related to hematoma formation, like phase 3 higher risk in hematoma than 1, then it will be an ordinal level.

Any data presented in frequencies and percentages are nominal and those in standard deviation and means they are at highest level of measurement, like age and its relation with hematoma, since age is written in years and the numbers give the actual meaning then they are ratio or better called continuous, and also total mass(in grams) breast mass in grams, hospital stay in days etc, all the info in the next table are considered continuous(the doctor prefers to say continuous rather than ratio or interval, because it's a little bit questionable to choose between ratio or interval)

Variables	*Hematoma Mean (SD) (n=6)*	*No Hematoma Mean (SD) (n=60)*	*P Value*
Age (Years)	*39.83 (6.8)*	*34.75(8.8)*	*0.175* [a]
Total mass (gm)	*3186.67(932.4)*	*1829(969.6)*	*0.002** [a]
Left breast mass (gm)	*1585(433.3)*	*868.1(449.4)*	*< 0.001** [a]
Right breast mass (gm)	*1601.6(502)*	*940.3(469.4)*	*0.002** [a]
Hospital stay(day)	*6.3(1.2)*	*4.7(1.6)*	*0.02** [a]
Duration of operation(hr)	*3.4(0.66)*	*2.8(0.66)*	*0.05** [a]
Estimated blood loss(ml)	*566.6(87.5)*	*327.5(118.7)*	*< 0.001** [a]
Low SBP first period (mmHg)	*86.1(5.49)*	*88.6(8.5)*	*0.48* [a]
Low SBP second period (mmHg)	*81.3(5.3)*	*89.8(9.5)*	*0.03** [a]
Mean BP first period (mmHg)	*63.6(1.2)*	*71.2(5.8)*	*0.002** [b]
Mean BP second period (mmHg)	*62.3(2.05)*	*75.1(6.9)*	*< 0.001** [b]
Peak SBP first period (mmHg)	*97.1(4.02)*	*107.8(10)*	*0.012** [a]
Peak SBP second period (mmHg)	*90.8(3.7)*	*115.9(13.6)*	*< 0.001** [b]
BMI (4 International Categories)	*50.83 (mean rank)*	*31.7 (mean rank)*	*0.018** [b]
Hemoglobin preoperative (gm/dl)	*12.6(0.79)*	*12.4(1.17)*	*0.663* [a]
Hemoglobin postoperative (gm/dl)	*9.3(1.16)*	*10.3(1.1)*	*0.03** [a]