

STATICS NOTES: ①

CHAPTER NO : 1

Statisticsst Defination:

It is a science of conducting study about collection of data, organizing data, summarization of data, analysis of data and concluding / prediction / make a decision from or about data.

→ If you are certain about something than its unstatistics.

→ If uncertainty is involve than there will be statistics

Two Branches

→ Descriptive Statistics: Consist of the collection, organization, summarization and Presentation of data.

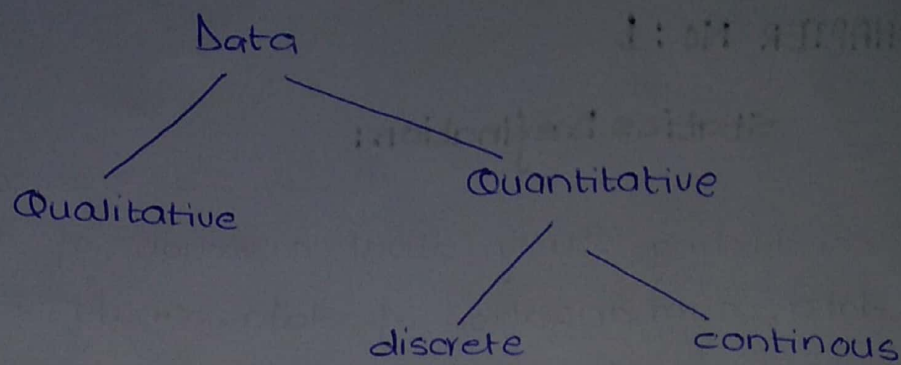
→ Inferential Statistics: Consists of generalizing from samples to population, performing estimation & hypothesis test, determining relationships among variables and making predictions.

VARIABLES: Variables is a characteristics or attribute that can have different values.

→ **Data** are the values (measurement or observation) that the variable can assume.

→ Variables whose values are determined by chance are called random variables.

② VARIABLES AND TYPES OF DATA:



Qualitative Variables: are variables that can be placed into distinct categories, according to some character or attribute.

For Example: If subjects are classified according to gender (male or female), then the variable **gender** is qualitative.

Other Examples are religious preference and geographic location.

Quantitative Variables: are numerical and can be ordered or ranked.

For Example: the variable age is numerical, and people can be ranked in order or according to their value of age.

Other Examples are height, weight & temperature of body.

→ **Discrete Variables:** assume values that can be counted. (whose values are countable)
e.g. no. of good books

→ **Continuous Variables:** Can assume an infinite number of values between any two specific values.

→ They are obtained by measuring

→ They often include fraction & decimal.

POPULATION:

A population consists of all subjects (human or otherwise) that are being studied.

SAMPLE:

It's a subset of population.

A sample is a group of subject selected from a population.

→ Consider a specific part of population.

MEASUREMENT OF SCALES:

1 → THE NOMINAL LEVEL OF MEASUREMENT

(No order & ranking can be imposed)

The measurement which based on language, colour etc.

2 → ORDINAL LEVEL OF MEASUREMENT.

(They can be ranked)

But the precise difference cannot be exist

3 → INTERVAL LEVEL OF MEASUREMENT:

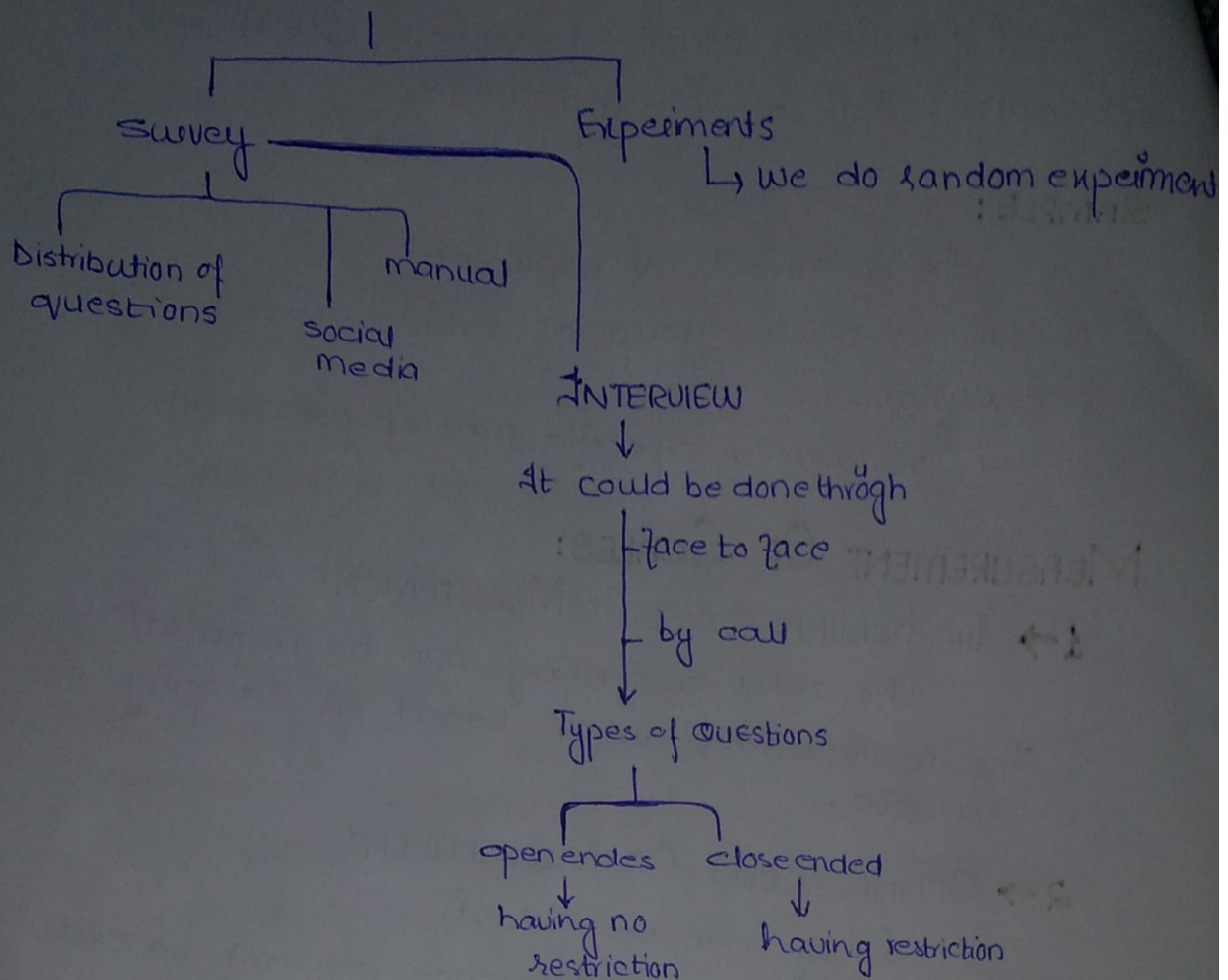
Rank data and precise difference b/w the units of measure do exist, there is no meaningful zero.

Zero temperature → non meaningful.
when classification based on temperature.

4 → RATIO LEVEL OF MEASUREMENT:

Contain all the characteristics of interval measurement and there exist a true zero. In addition, true ratio exist when the same variables is measured on two different member of the population

DATA COLLECTION:



- * The aim of questioner is should be very clear you must know
- * Remember → what will be the outcome of these questioners.

Once you collect the Data then the second Process is Sampling.

Sampling:

- ① Random Sampling
→ picking sample from population randomly
- ② Systematic Sampling
→ If we decide ^{that} we need a sample of specific size. It reduce the randomness
- ③ Stratified Sampling
→ If we have a population make its group (having each characteristic)

③

- ① Make a group of given data.
- ② Having all characteristics.
- ③ Then select random group.

④ CLUSTER SAMPLING

Already made group on the basis of something

APPLYING THE CONCEPT :

A study conducted at Manatee Community College revealed that student who attended class 95 to 100% of the time usually received an A in the class. Students who attended class 80 to 90% of the time usually received a B or C in the class. Student who attended class less than 80% of the time usually received a D or an F or eventually withdrew from the class.

Based on this information, attendance & grades are related. The more you attend class the more likely you will receive a higher grade. If you improve your attendance, your grades will probably improve. Many factors effect your grades in a course. One factor that you have considerable control over is attendance. You can increase your opportunities of learning.

1) What are the variable under study?

Answer! Grades & Attendance

(According to the definition of variable grades and attendance both have different values)

④ a) what are the data in study?

Answer: The data consist of specific grades and attendance

(Data are the measurement & observation that variables can assume)

3) Are descriptive, inferential or both types of statistics used?

Answer: Descriptive statistics.

4) What is the population under study?

Answer: Students of Manatee Community college.

5) Was a sample collected? If so, from where.

Answer: No, sample is not collected

6) From the information given, comment on the relationship between variables.

Answer: Better your attendance the higher your grade.

APPLYING THE CONCEPT - 2

Transportation Safety

The chart shows the number of job-related injuries for each of the transportation industries for 1998.

Industry	Number of injuries.
Railroad	4520
Intercity bus	5100
Subway	68500
Trucking	7144
Airline	9950

1) What are the variable under study?

Answer: Variables under study are Industry and number of injuries.

2) Categorize each variables as quantitative or qualitative.

Answer: The type of industries are qualitative variables and the number of injuries are quantitative.

3) Categorise each quantitative variable as discrete or continuous.

Answer: The number of injuries are discrete variables.

4) Identify the level of measurement for each variables.

Answer. The type of industry is nominal & number of injuries are ratio.

5) The railroad is shown as the safest transportation industry. Does that mean railroad have fewer accidents than the other industries? Explain.

Answer: The railroad is shown as the safest transportation because fewer people make accident by this transportation as shown from given data.

6) What factor other than safety influence a person's choice of transportation?

Answer: A person choice of transportation can be affected by cost, service, charges etc

7) From the information given, comment on the relationship b/w the variables?

From the given information we assumed railroad have less job related injuries as compare to the air line which have the most job-related injuries.

APPLYING THE CONCEPT - 3

AMERICAN CULTURE AND DRUG ABUSE

Assume you are the member of the family Research Council and have become increasingly concerned about the drug use by professional sports players. You set up a plan and conduct a survey on how people believe the American culture (television, movies, magazines and popular musics) influence illegal drug use. Your survey consist of 2250 adults and adolescents from around the country. A consumer group petition you for more information about your survey.

① What type of survey did you use (phone, mail or interview)?

Answer: I used a phone survey. The advantage to my survey method is that this was a relatively inexpensive survey method.

② What are the advantages and disadvantages of the surveying method you did not use?

Answer: A mail survey also would have been fairly inexpensive, but my response rate may have been much lower than what I got with my telephone. Interviewing would have allowed me to follow up questions and to clarify any questions of the respondent at time of the interview.

③ What type of scale did you use?

Answer: I used ordinal data on scale 1 to 5.

The scales were 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree.

(4) Did you use a random method for deciding who would be in your sample?

Answer: The random method that I used was random dialing method.

(5) Which of the methods (stratified, systematic, cluster or convenience) did you use?

Answer: To include people from each state, I used to stratified random sample, collecting data randomly from each of the area code and telephone exchange available.

(6) Why was that method more appropriate for this type of data collection?

Answer: This method allowed me to make sure that I had representation from each area in United States.

(7) If a convenience sample were obtained, consisting of only adolescents, how would the result of study be affected?

Answer: Convenience samples may not be representative of population & a convenience sample of adolescents would probably differ greatly from the general population with regards to influence of American culture on illegal drug use.