Statistics

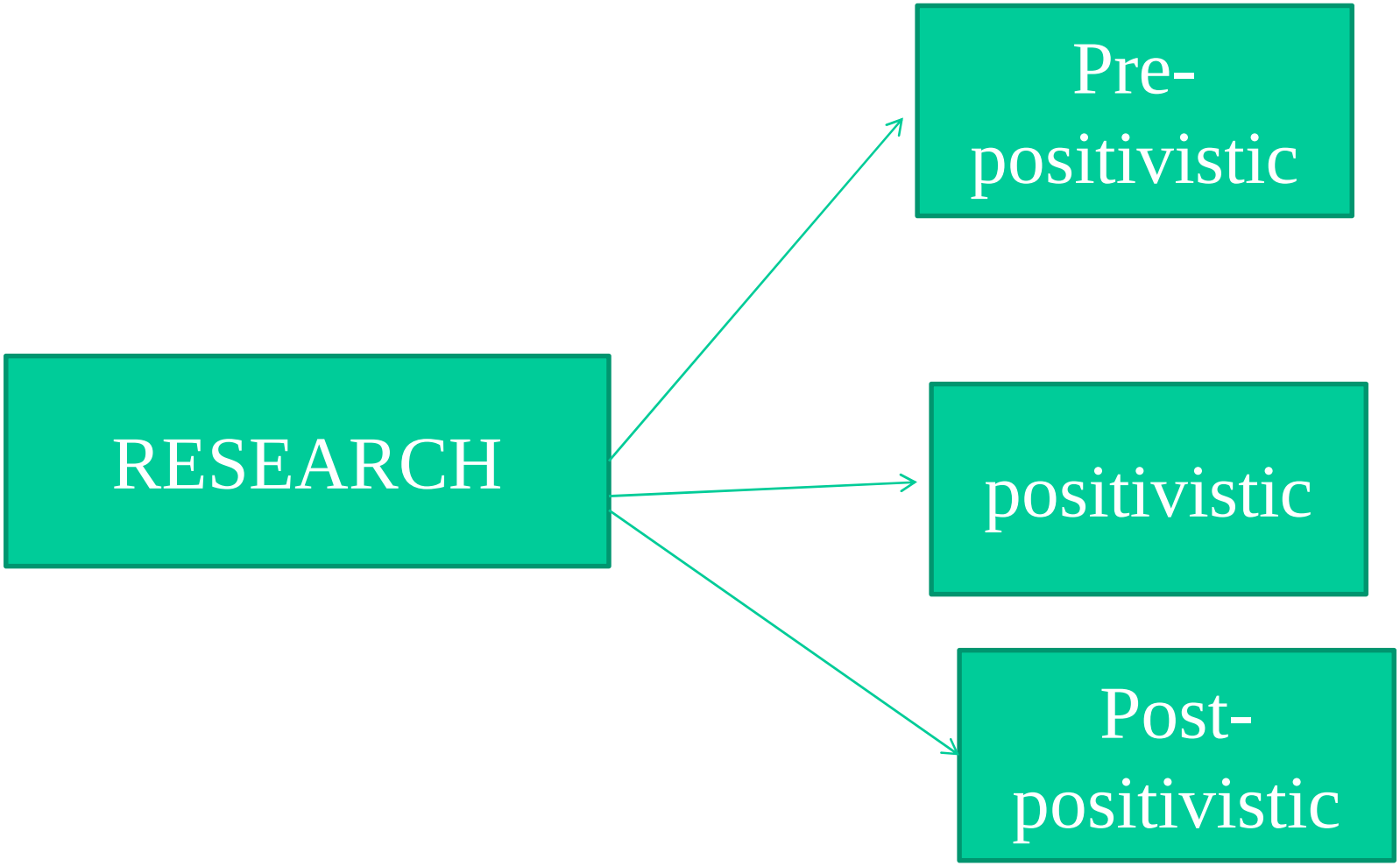
Nur Hidayanto PSP
English Education Dept.

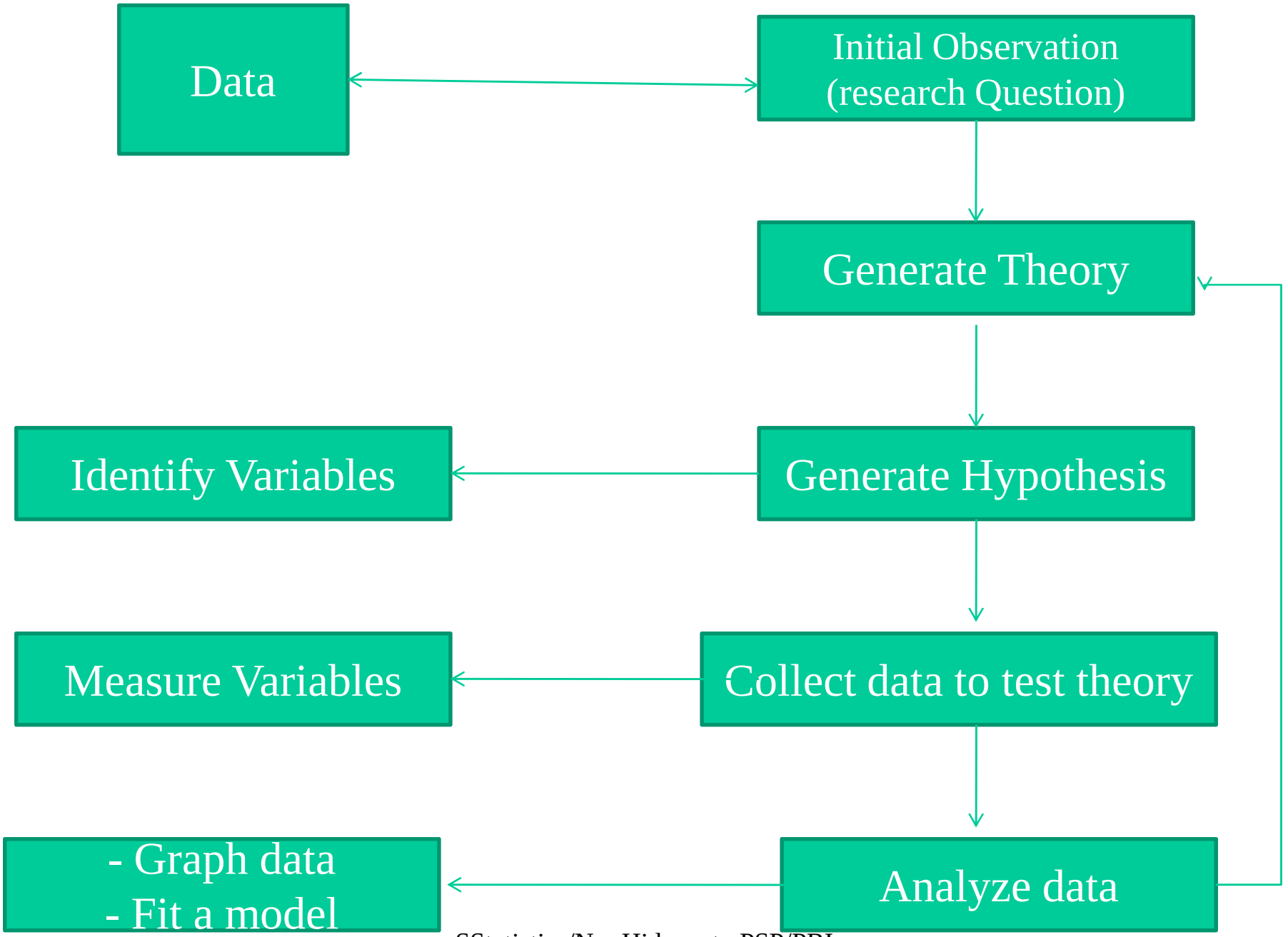
RESEARCH

• STATISTICS

WHAT’S THE RELATIONSHIP?

RESEARCH


RESEARCH
Pre-
positivistic
positivistic
Post-
positivistic


Data
Initial Observation
(research Question)
Generate Theory
Identify Variables
Generate Hypothesis
Measure Variables
Collect data to test theory
- Graph data
- Fit a model
Analyze data

Educational Statistics:
Statistical Terms and Vocabulary

- *Statistics*: a set of methods, procedures and rules for organizing, summarizing, and interpreting information.
  - This is a *general* definition.
  - Later, a distinction between *statistics* and *parameters* will be made.
  - Here, it would be better to speak of statistical *methods*.

Educational Statistics:
Statistical Terms and Vocabulary

Population vs *Sample*

- *Population:* all members of a particular group (e.g., all Appstate freshman, all males over the age of 21, all of the schools in NC).
- *Sample:* a subgroup of a population that is usually assumed to be *representative* of the population (e.g., 10 Appstate freshman selected at random).

Educational Statistics:
Statistical Terms and Vocabulary

Variable: any characteristic that can vary across individuals, groups, or objects. For example:

- Weight
- Occupation
- Grade-point average
- Level of test anxiety

Later we will look at various *types* of variables

Educational Statistics:
Statistical Terms and Vocabulary

Values: the numerical value of a particular *realization* of a variable.

For instance if the variable is weight and Mortimer weighs 147 lbs. Then the *value* of the *variable* for Mortimer is *147*.

Make sure you can distinguish between variables and values

Educational Statistics:
Statistical Terms and Vocabulary

Parameters and Statistics

- *Parameter:* the value of a variable in a population.
- *Statistic:* the value of a variable in a sample.
- Statistics are often used to *estimate* or draw *inferences* about parameters.

Educational Statistics:
Statistical Terms and Vocabulary

Types of statistical methods:

- Descriptive statistics: methods used to summarize, organize, and simplify data.
- Inferential statistics: methods that allow us to make generalizations about populations based on data obtained from samples.

Educational Statistics:
Statistical Terms and Vocabulary

The *correlational* method of research.

- Examines relationships among two or more *variables*.
- For example: What is the relationship between hours of TV watched per day and the number of calories consumed per day?
  - Note that there no cause-effect relationship is postulated.
  - Correlation does not imply causation.

Educational Statistics:
Statistical Terms and Vocabulary

The *experimental* method is used when the researchers wants to establish a cause and effect relationship.

- The researcher *manipulates* one variable (the *independent*) variable, and
- Observes (or *measures*) what happens to a second variable (the *dependent variable*),
- while *Controlling* for all other variables (*extraneous variables*).

Educational Statistics:
Statistical Terms and Vocabulary

A *quasi-experiment* is similar to a (true) experiment except that here the independent variable is *not* manipulated by the researcher.

For example, in studying the effects of sex on mathematics achievement a researcher compares boys and girls (the independent variable).

Educational Statistics:
Statistical Terms and Vocabulary

Types of measurement.

- The *type* of measurement scale has implications for the type of statistical procedure employed.
- Some statistical procedures assume a certain level of measurement.
- Three tpes of measurement can be distinguished: *nominal, ordinal,* and *scale.*

Educational Statistics:
Statistical Terms and Vocabulary

Types of measurement: *nominal*.

- Coarse level of measurement used for identification purposes.
- Substitutes numbers for other categorical labels.
- No order of magnitude is implied.
- Examples: sex (male or female), student classification (freshman, sophomore, junior, senior), etc.

Educational Statistics:
Statistical Terms and Vocabulary

Types of measurement: *ordinal*.

- Objects measured on an ordinal scale differ from each other in terms of magnitude, but the units of magnitude are not equal.
- The objects can be ordered in terms of their magnitude (more or less of an attribute.
- Examples: class rank, seeding in golf or tennis, percentiles, level of motivation.
- Do not allow common mathematical opperations.

What about grades or GPA?

Educational Statistics:
Statistical Terms and Vocabulary

Types of measurement: *scale*.

- On an interval scale, objects are not only ordered by magnitude, but the distance between any two adjacent units is equal to the distance between any other two adjacent units.
- Examples: SAT scores, Celsius and Fahrenheit scales, developmental scale scores (e.g., EOG/EOC).
- Allow common mathematical opperations.

Educational Statistics:
Statistical Terms and Vocabulary

Discrete and Continuous variables

- Variables can also be described in terms of the types of *values* they can be assigned.
  - *Discrete* variables are categorical. No values between two adjacent values are permissible.
  - *Continuous* variables can (theoretically) have an infinite number of values.

Descriptive Statistics

2nd meeting

VALIDATING YOUR INSTRUMENTS

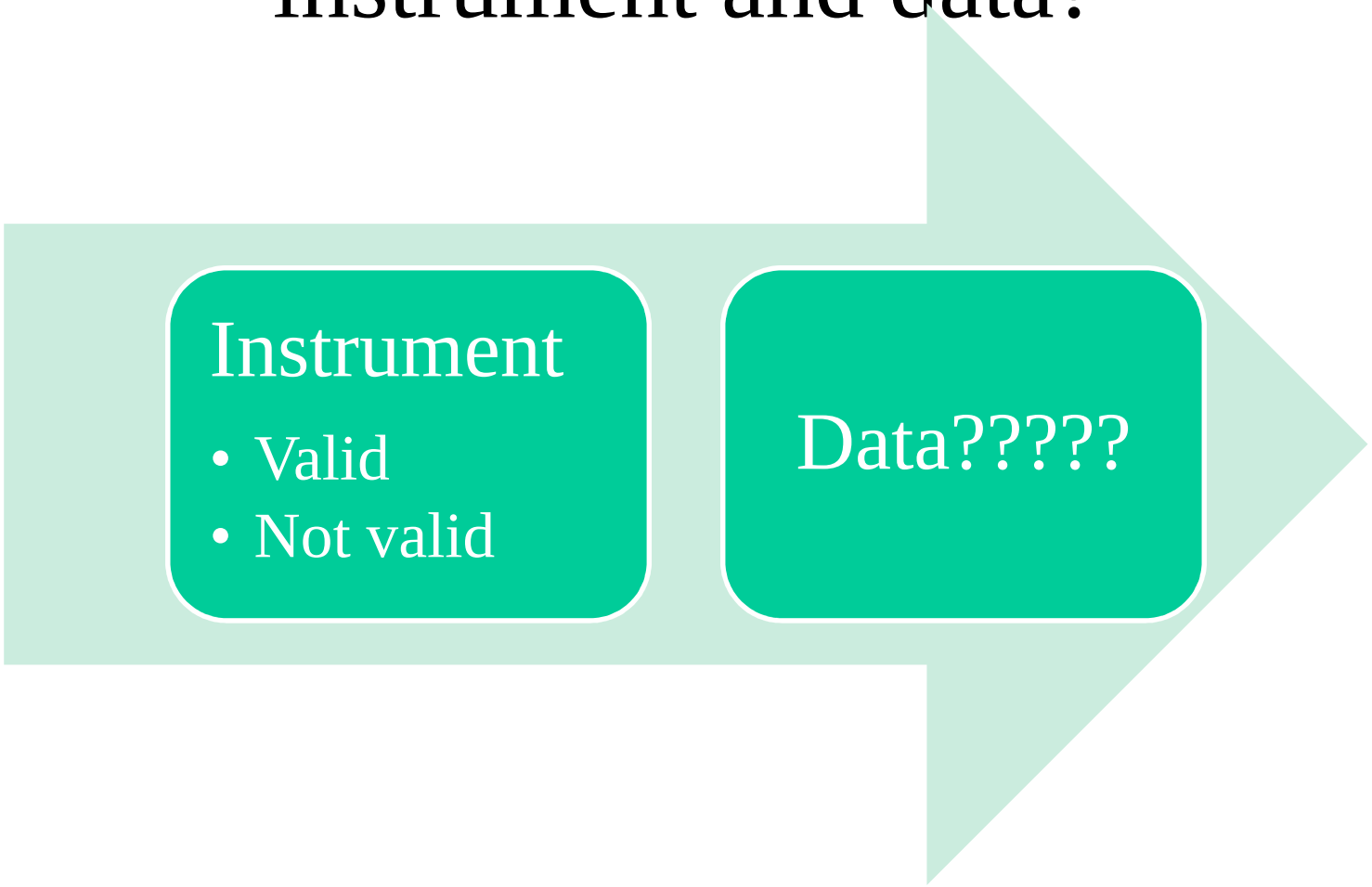
QUESTIONNAIRE

AND

TESTS

SStatistics/Nur Hidayanto PSP/PBI

What is the relationship between you instrument and data?



VALIDITY

What is validity?

If your instrument is valid, so what?

How do you validate your instruments?

- Expert judgment
- Using computer software
  - Questionnaire: items: SPSS
  - Tests : item difficulty, discrimination index, distractors: iteman

So?

- Do all the “theses” in the library show the reality that happens in the field?

- GARBAGE IN – GARBAGE OUT

COMMON MISTAKES

- NO TRY OUT--→ INVALID INTRUMENT ---→ INVALID DATA ---→ WRONG CONCLUSION

Go to the library

Find a thesis which used a questionnaire as one of the instruments in collecting the data.

a. Did the researcher conduct a try out to validate the questionnaire?

b. Are all the items in the questionnaire good?

c. What is the implication?

Review:

Statistics has two major chapters:

- Descriptive Statistics
- Inferential statistics

Statistics

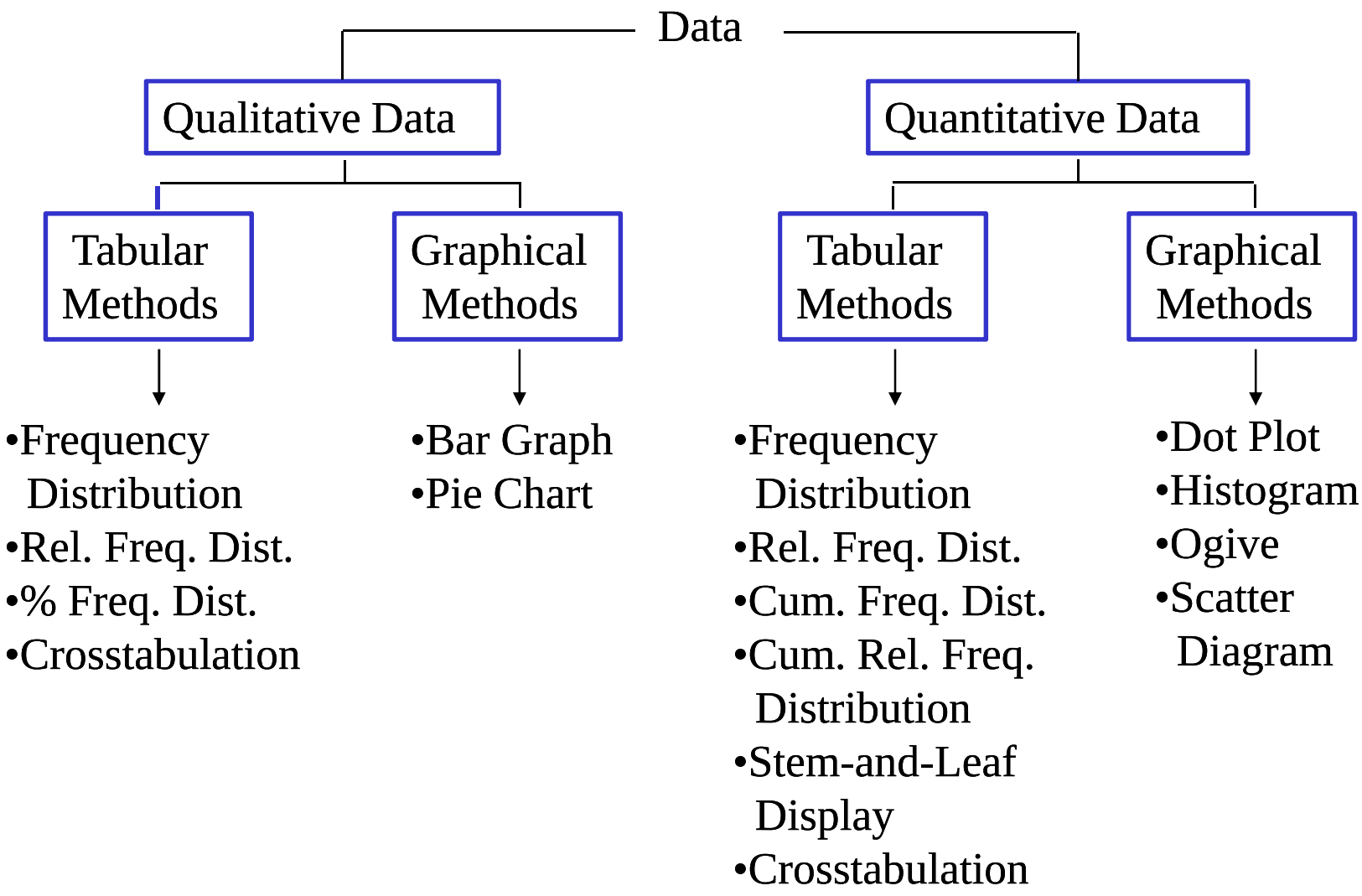
Descriptive Statistics

- Gives numerical and graphic procedures to summarize a collection of data in a clear and understandable way

Inferential Statistics

- Provides procedures to draw inferences about a population from a sample

TABULAR DAN GRAFIS

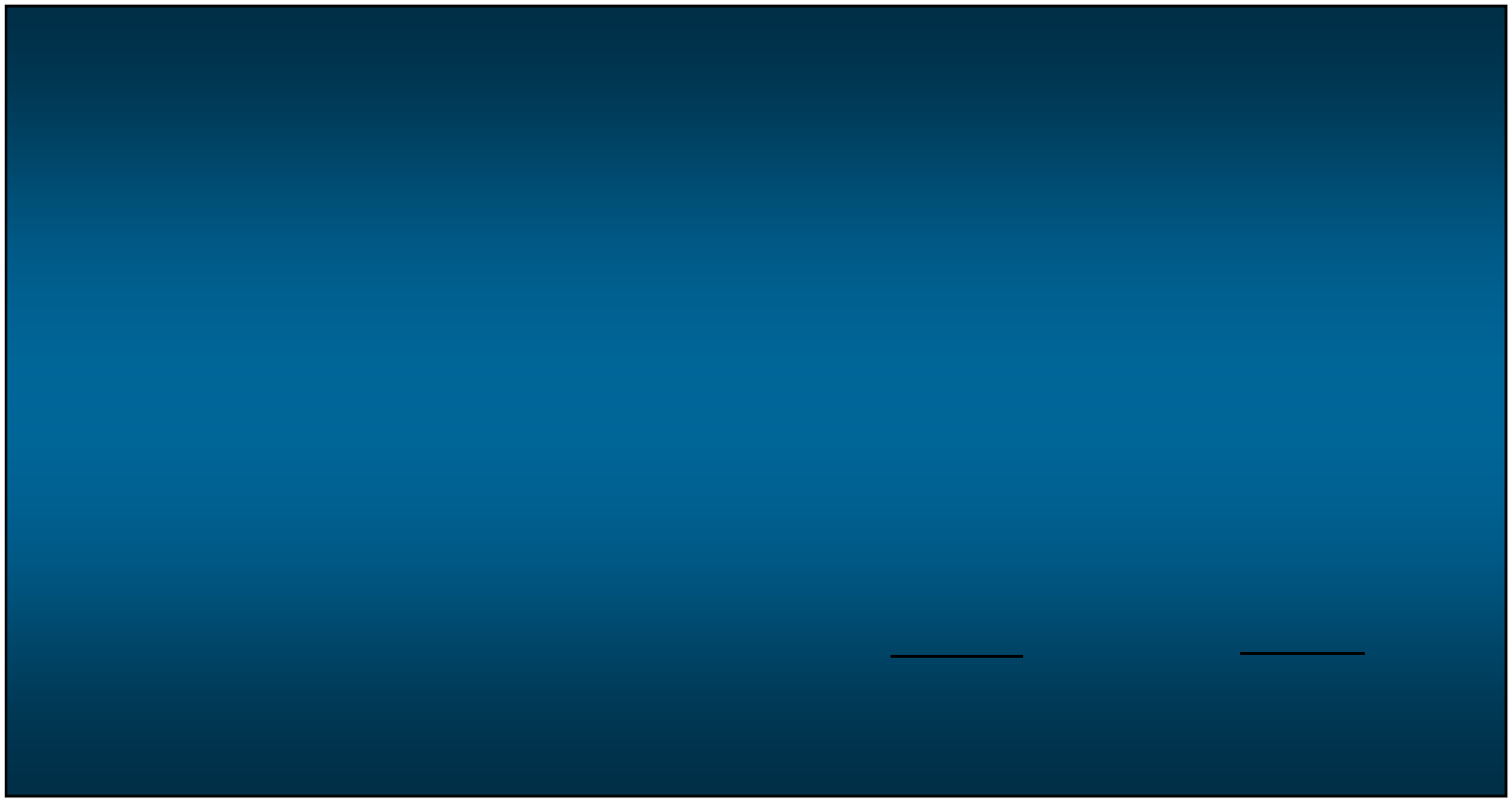


QUALITATIVE DATA

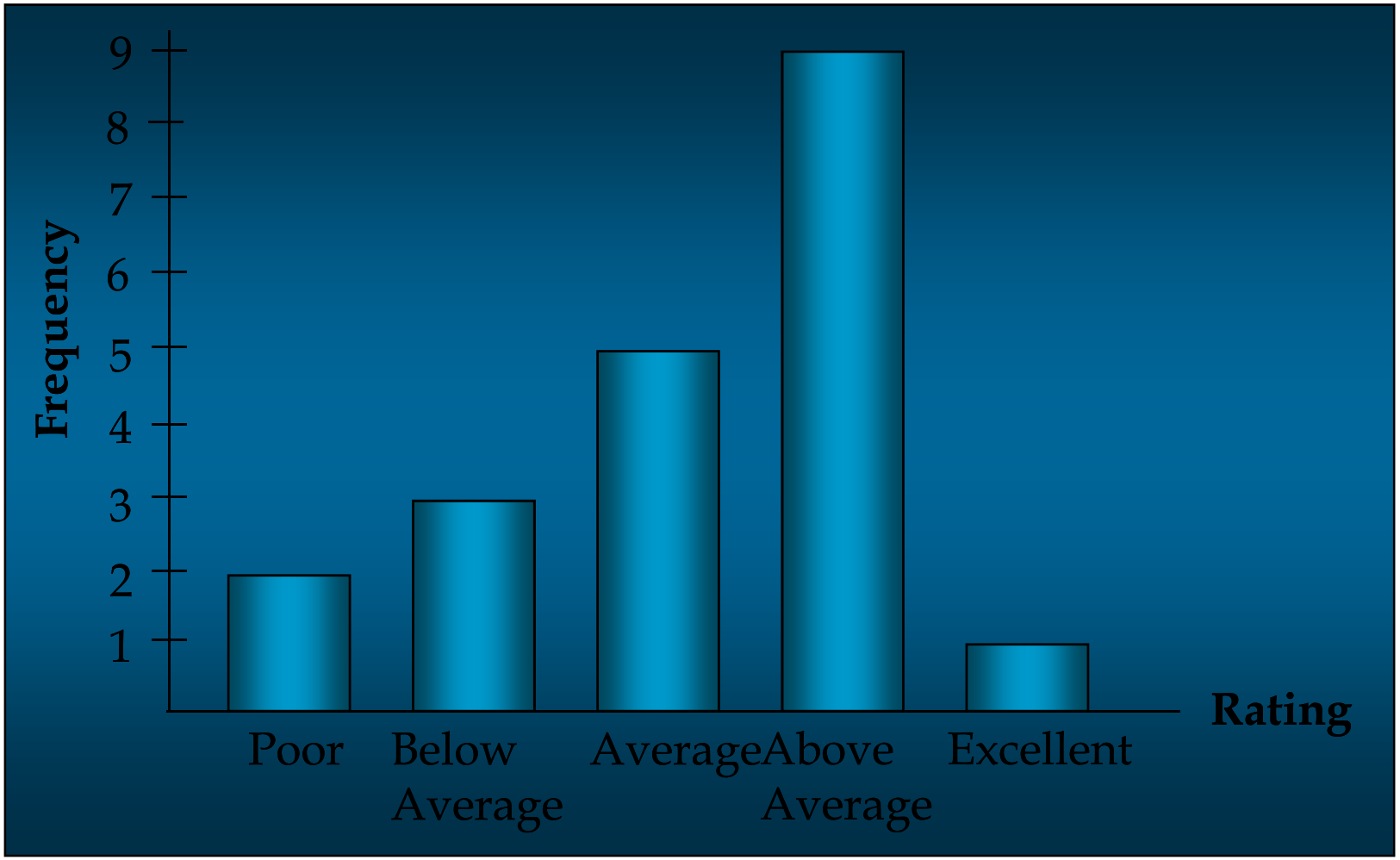
The students of Mr. Enjoy class are expected to judge the use of new teaching technique “CAT” by choosing :***excellent***, ***above average***, ***average***, ***below average***, or ***poor***. The following data show the answers of 20 different students in his class.

Below Average	Average	Above Average
Above Average	Above Average	Above Average
Above Average	Below Average	Below Average
Average	Poor	Poor
Above Average	Excellent	Above Average
Average	Above Average	Average
Above Average	Average	

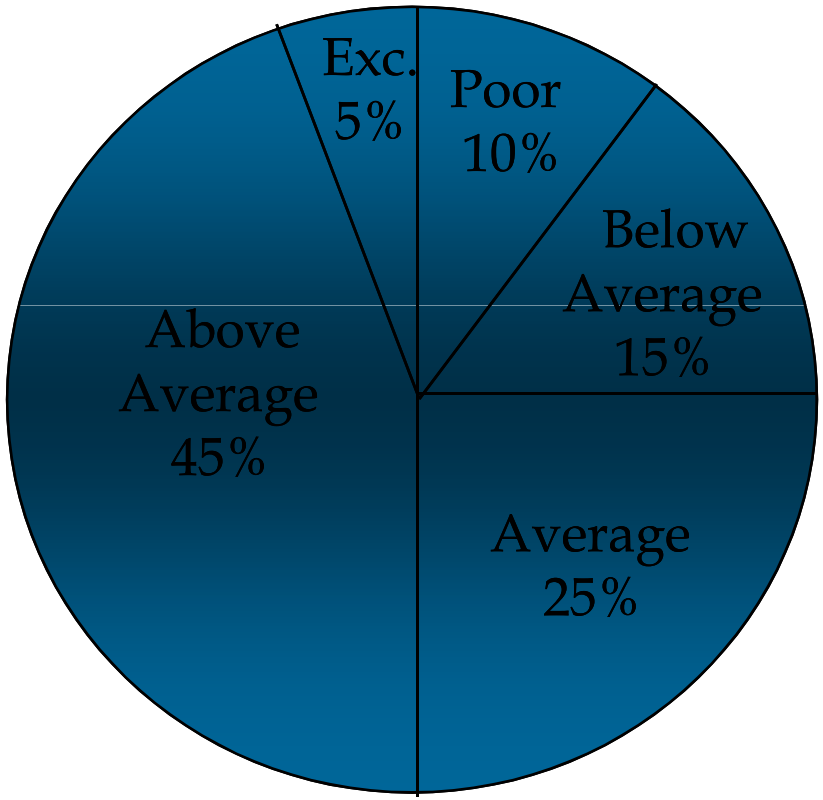
DISTRIBUTION TABLE



BAR GRAPH


Frequency
9
8
7
6
5
4
3
2
1
Poor
Below Average
Average
Above Average
Excellent
Rating

PIE CHART



Quality Ratings

SStatistics/Nur Hidayanto PSP/PBI

QUANTITATIVE DATA
Descriptive Measures

- ***Central Tendency measures***. They are computed to give a "center" around which the measurements in the data are distributed.

- ***Variation or Variability measures***. They describe "data spread" or how far away the measurements are from the center.

- ***Relative Standing measures***. They describe the relative position of specific measurements in the data.

Measures of Central Tendency

- **Mean**:

  Sum of all measurements divided by the number of measurements.

- **Median**:

  A number such that at most half of the measurements are below it and at most half of the measurements are above it.

- **Mode**:

  The most frequent measurement in the data.

Example of Mean

Name	score	Name	acore
A	65	P	55
B	68	Q	70
C	70	R	65
D	85	S	80
E	90	T	95
F	87	U	70
G	95	V	85
H	76	W	85
I	70	X	80
J	90	Y	76
K	90	Z	60
L	80	A1	78
M	90	B1	68
N	75	C1	90
O	60	D1	70

- **An English teachers conducted a mid-semester test (max 100) and the table shows the students' scores.**
- What does the table show you?
- What treatment should teachers give to the students?

Example of Median

Name	score	Name	acore
A	65	P	55
B	68	Q	70
C	70	R	65
D	85	S	80
E	90	T	95
F	87	U	70
G	95	V	85
H	76	W	85
I	70	X	80
J	90	Y	76
K	90	Z	60
L	80	A1	78
M	90	B1	68
N	75	C1	90
O	60	D1	70

- What is the median?

Example of Mode

Name	score	Name	acore
A	65	P	55
B	68	Q	70
C	70	R	65
D	85	S	80
E	90	T	95
F	87	U	70
G	95	V	85
H	76	W	85
I	70	X	80
J	90	Y	76
K	90	Z	60
L	80	A1	78
M	90	B1	68
N	75	C1	90
O	60	D1	70

What is the mode?

Variance (for a sample)

- **Steps:**
  - Compute each deviation
  - Square each deviation
  - Sum all the squares
  - Divide by the data size (sample size) minus one: n-1

Example of Variance

Measurements	Deviations	Square of deviations
x	x - mean	
3	-1	1
5	1	1
5	1	1
1	-3	9
7	3	9
2	-2	4
6	2	4
7	3	9
0	-4	16
4	0	0
40	0	54

- Variance = 54/9 = 6
- It is a measure of "spread".
- Notice that the larger the deviations (positive or negative) the larger the variance

The standard deviation

- It is defines as the square root of the variance
- In the previous example
- Variance = 6
- Standard deviation = Square root of the variance = Square root of 6 = 2.45

Further Notes

- When the Mean is greater than the Median the data distribution is skewed to the Right.
- When the Median is greater than the Mean the data distribution is skewed to the Left.
- When Mean and Median are very close to each other the data distribution is approximately symmetric.