Common Problems Associated with Data Collection and Statistical Analyses

- **Bias**

  Certain outcomes are systematically favored as a result of:
  - the sampling method
  - the wording of questions (in questionnaires)
  - the measurement instrument

- **Confounding**

  Spurious association between two variables with no cause-effect relationship caused by a third hidden variable

Ballot Initiative

Initiative 9999 will cut wasteful government spending and provide a much needed tax relief for millions of citizens.

☐ yes ☐ no

Initiative 9999 will eliminate 500 jobs for policemen and firefighters and jeopardize the safety of millions of citizens.

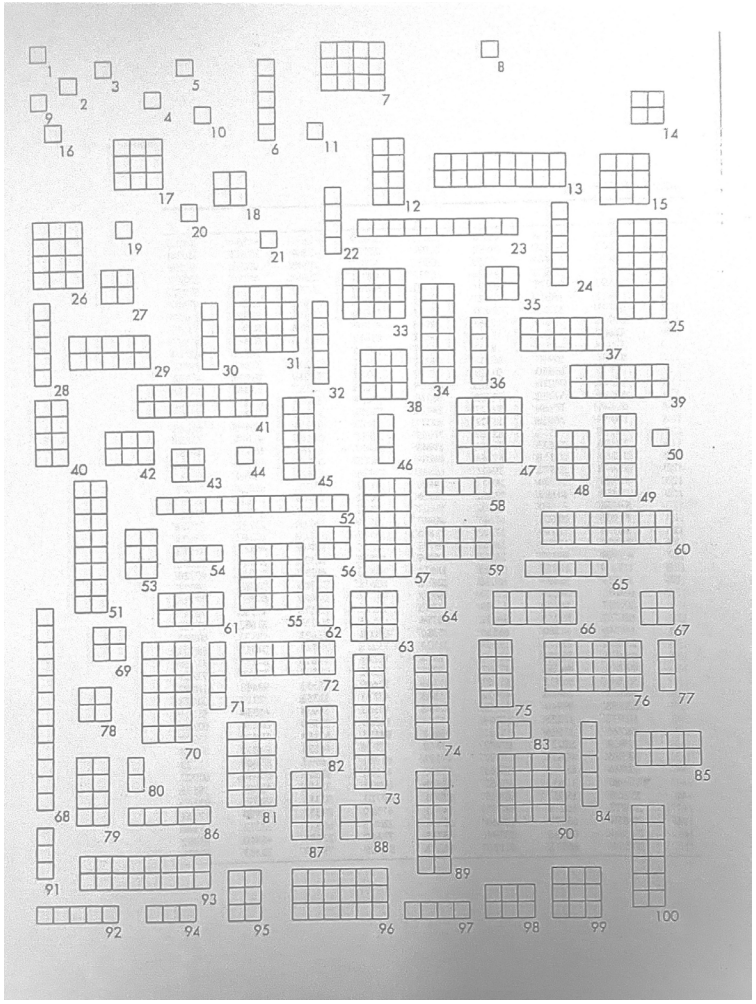
yes ☐

no ☐

Technology Fee

- During the 1998/99 academic year at EWU, the student association conducted a survey among students to measure the support of implementing a $35/quarter Technology Fee.
- Questionnaires were distributed among students in the PUB computer lab and in the dormitories.

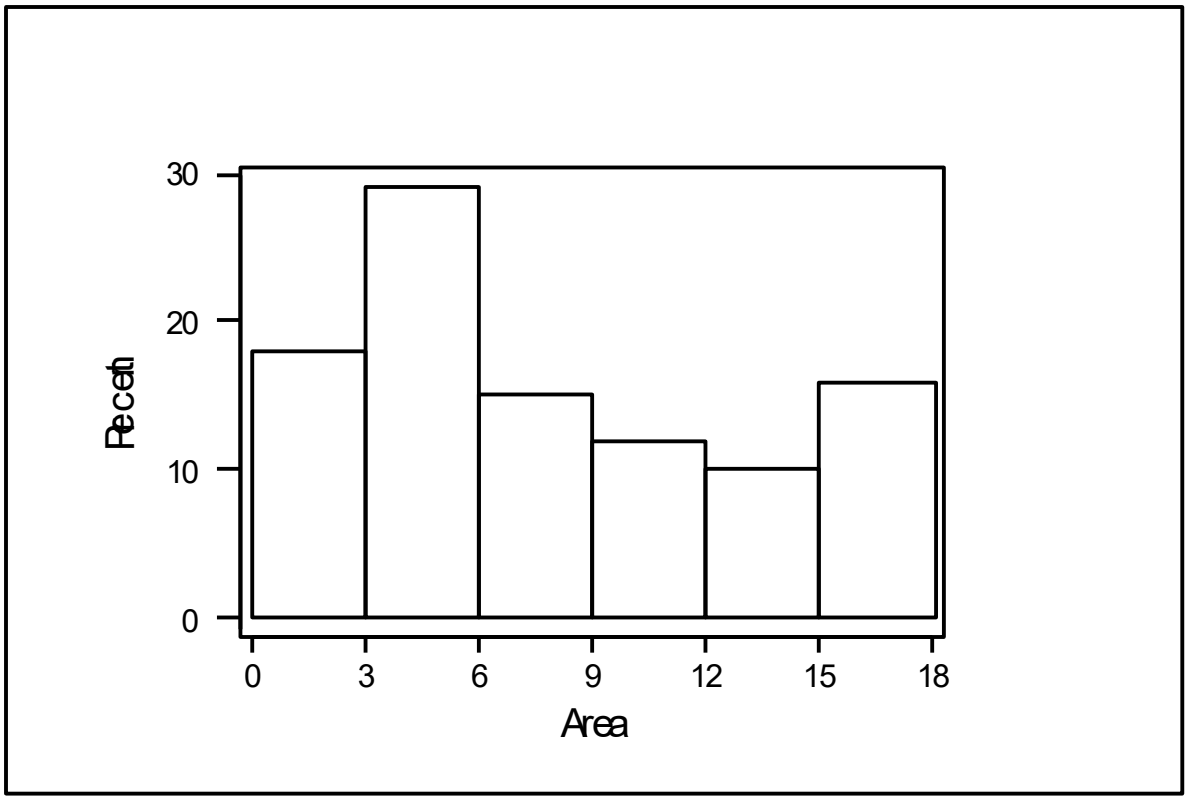
CASE: Random Rectangles



CASE: Random Rectangles

Leaf	Stem	Leaf
	1	
	2	
	3	
	4	
	5	
	6	
	7	
	8	
	9	
	10	
	11	
	12	
	13	
	14	
	15	
	16	
	17	

CASE: Random Rectangles

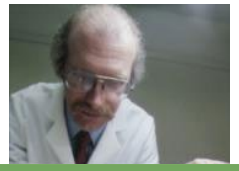


Population Parameters:

Mean μ = 7.46

Standard Deviation σ = 5.22

Experiments



PRINCIPLES OF EXPERIMENTAL DESIGN

The basic principles of statistical design of experiments are

1. Compare two or more treatments. This will control the effects of lurking variables on the response.

2. Randomize—use impersonal chance to assign experimental units to treatments.

3. Repeat each treatment on many units to reduce chance variation in the results.

Experiments

Terminology

- Treatments are imposed on experimental units or subjects.
- Each treatment is a combination of levels of the explanatory variables called factors.
- Outcomes are the measured variables that are used to compare the treatments (response variables).

Experiments

BLOCK DESIGN

A **block** is a group of experimental units or subjects that are known before the experiment to be similar in some way that is expected to affect the response to the treatments. In a **block design,** the random assignment of units to treatments is carried out separately within each block.

Example : Comparing Treatment Regimens for Lowering Blood Pressure

- Four different treatment regimens (3 drugs and a placebo) are to be evaluated for their impact on diastolic blood pressure. 32 patients (16 men and 16 women) with high blood pressure are selected randomly for the experiment. The subjects are randomly assigned to the 4 treatments (4 subjects of each gender to each treatment). The difference in diastolic blood pressure (before and after treatment) is measure for each subject.
- Is there evidence of a difference between the different treatments? between men and women's reaction to the treatments?

Diastolic Blood Pressure Data

Subjects randomly assigned

	Regimen 1	Regimen 2	Regimen 3	Regimen 4
Male	6	9	8	7
	12	1	14	16
	2	13	5	3
	15	4	11	10
Female	13	15	3	7
	11	8	1	4
	6	10	16	12
	5	2	9	14

Decreases in diastolic blood pressure for 16 male and 16 female patients, randomly assigned to each of four different treatment regimens.

Diastolic Blood Pressure Data

	Regimen 1	Regimen 2	Regimen 3	Regimen 4
Male	5	6	7	6
	5	6	8	5
	2	4	6	4
	4	4	7	5
Female	2	3	4	4
	5	5	5	4
	3	3	6	3
	5	4	7	6

[Change in diastolic blood pressure]

Decreases in diastolic blood pressure for 16 male and 16 female patients, randomly assigned to each of four different treatment regimens.