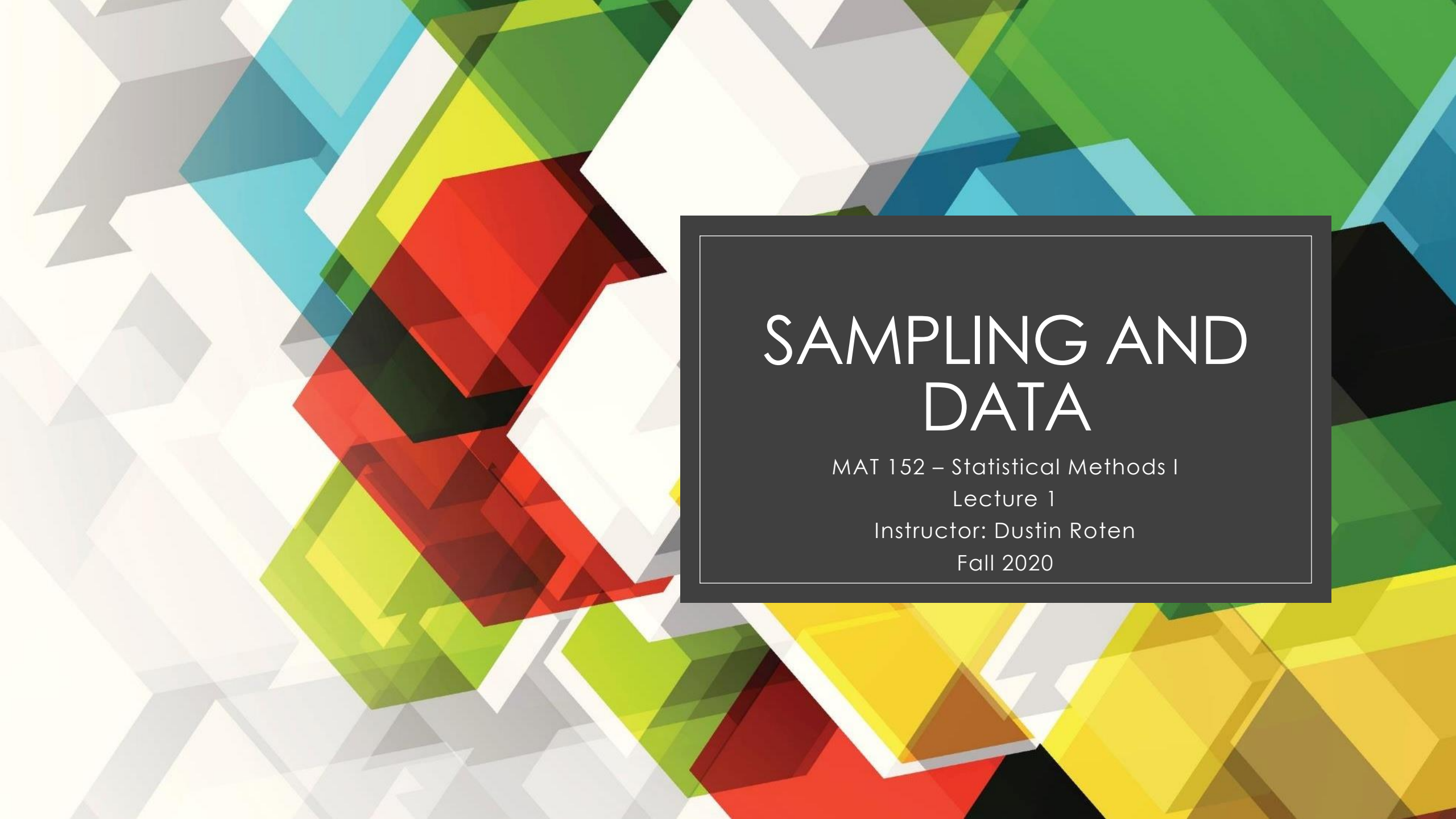


SAMPLING AND DATA

MAT 152 – Statistical Methods I

Lecture 1

Instructor: Dustin Roten

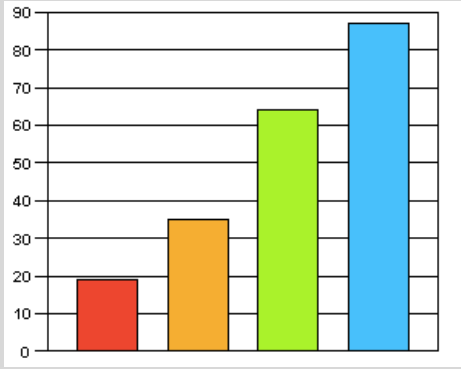
Fall 2020

Statistics is a process

Collecting Data

Analyzing Data

Interpreting Data



Presenting Data

What is *statistics*?

- Statistics is:
  - The collection of data
    - What question(s) are being asked?
    - What information needs to be collected?
      - Demographics? Grades? Rainfall totals?
    - How can this information be collected?
      - Surveys? Observations? Instrumentation?
  - The Analysis of data
    - What mathematical tools do we need?
  - The interpretation of data
    - What do our results tell us?
    - Was our hypothesis correct?
  - The presentation of data
    - How can we communicate our findings to others?

Statistics is used to answer questions in almost every field:

Medicine

Economics

Engineering

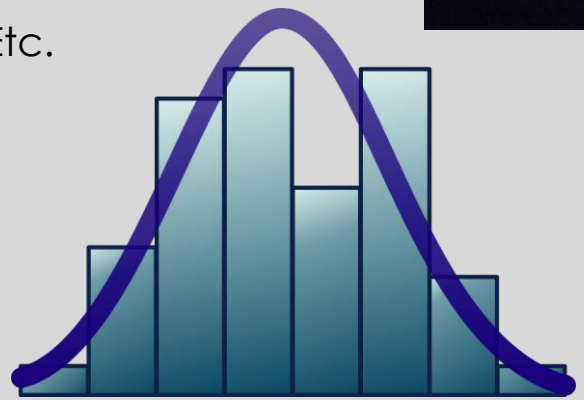
Meteorology

Astronomy

Sports (Sabermetrics)

Construction

Etc.



Terminology

- Population
- Parameter
- Sample
- Statistic
- Data
- Variable

Exploring terms with an example:

Suppose a statistics instructor at Forsyth Tech wishes to determine the average GPA of an FT student.

Population:
A collection of persons, things, or objects under study:
All students at Forsyth Tech

Parameter:
A numerical characteristic of the whole population:
The average GPA of all FT students (??? Unknown parameters)

Data:
The actual values of a variable

Data can be continuous or discrete

Sample:
A representative sample of the population:
1000 students selected at random
The sample MUST be reflective of the population

Statistic:
A number that represents a property of the sample:
The average GPA of the 1000 selected students

Variable:
A characteristic or measurement that can be determined for each member of the population:
$X = \text{Student GPA}$

Example

(Textbook: Chapter 1, #42)

A fitness center is interested in the average amount of time a client exercises in the center each week.

What is the population?

All clients registered at the fitness center

What is the sample?

A subset (group) of the clients at the center

What is the parameter?

The average amount of time clients exercise in the center

What is the statistic?

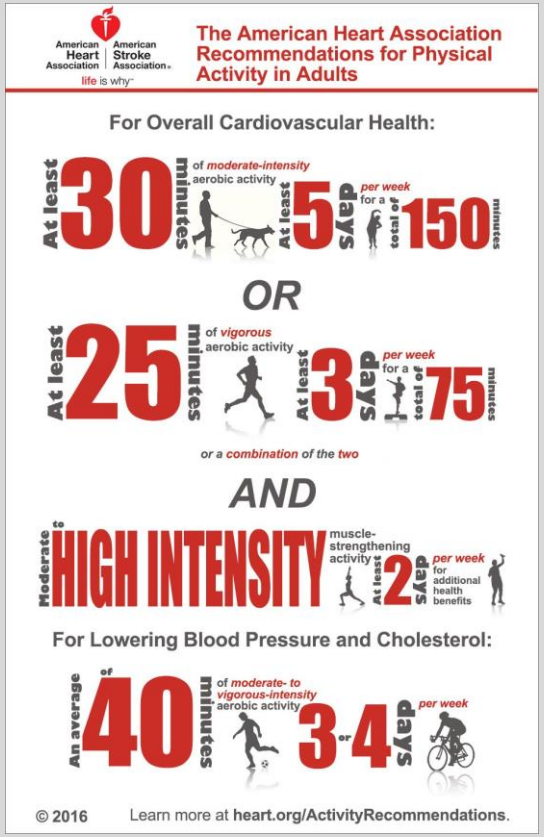
The average amount of time the sample (group) of clients exercises in the center

What is the variable?

$X =$ The amount of time spent exercising by a single client

What is the data?

Values of X from the sample (20min, 30min, 45min, 60min, etc.)



Different types of variables

- Categorical variables – place the person or thing into a category (discrete data)
- Numerical variables – values with equal units (numerical data)

Data can be **qualitative** or **quantitative**.

Qualitative/Categorical Data – Describing attributes of a population

Examples: Color, shape, sex, political party, etc.

Quantitative Data – Always numbers

- Quantitative *discrete* data - All data that are the result of counting; only certain numerical values are allowed

  Examples: Number of insurance policies sold, a sports team's wins/losses, the number of pairs of shoes you own (whole numbers)

- Quantitative *continuous* data – Data that are not only made up of counting numbers, but that may include fractions, decimals, or irrational numbers

  Examples: The distance from a university's dorms to the middle of the campus, the wingspans of blue birds, the weights of Forsyth Tech students' backpacks

Examples

(Example 1.9 from textbook) Indicate whether quantitative data are continuous or discrete

- The number of pairs of shoes you own
  - Quantitative discrete
- The type of car you drive
  - Qualitative
- The distance it is from your home to the nearest grocery store
  - Quantitative continuous
- The number of classes you take per school year
  - Quantitative discrete
- The type of calculator you use
  - Qualitative
- Weights of sumo wrestlers
  - Quantitative continuous
- Number of correct answers on a quiz
  - Quantitative discrete
- IQ scores
  - Quantitative continuous

Qualitative Data

The type of car you drive
The type of calculator you use

Quantitative Discrete Data

The number of pairs of shoes you own
The number of classes you take per school year
Number of correct answers on a quiz

Quantitative Continuous Data

The distance it is from your home to the nearest grocery store
Weights of sumo wrestlers
IQ scores

A Quick Review

- Population/Parameter – a collection of objects being studied
- Sample/Statistic – a subset of the objects being studied
- Data/Variable – values of a characteristic from a sample

- Categorical variables – characteristics used to group objects
- Continuous variables – numerical values

- Qualitative data – characteristic(s) of the objects in the sample
- Quantitative discrete data – data resulting from counting, only certain values are allowed
- Quantitative continuous data – numerical values; a full range of values are possible