**IE 340**
**PROCESS IMPROVEMENT**
**THROUGH PLANNED EXPERIMENTATION**

**IE406**
**Simulation**

# Introduction and Data Collection

Dr. Xueping Li
University of Tennessee

Based on Business Basic Statistics (9th Edition)

# Chapter Topics

- Why a Manager Needs to Know About Statistics
- The Growth and Development of Modern Statistics
- Some Important Definitions
- Descriptive Versus Inferential Statistics

# Chapter Topics

*(continued)*

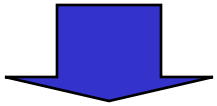
- Why Data are Needed
- Types of Data and Their Sources
- Design of Survey Research
- Types of Sampling Methods
- Types of Survey Errors

# Why a Manager Needs to Know About Statistics

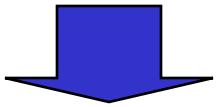
- To Know How to Properly Present Information
- To Know How to Draw Conclusions about Populations Based on Sample Information
- To Know How to Improve Processes
- To Know How to Obtain Reliable Forecasts

# The Growth and Development of Modern Statistics

**Needs of government to collect data on its citizenry**



**The development of the mathematics of probability theory**

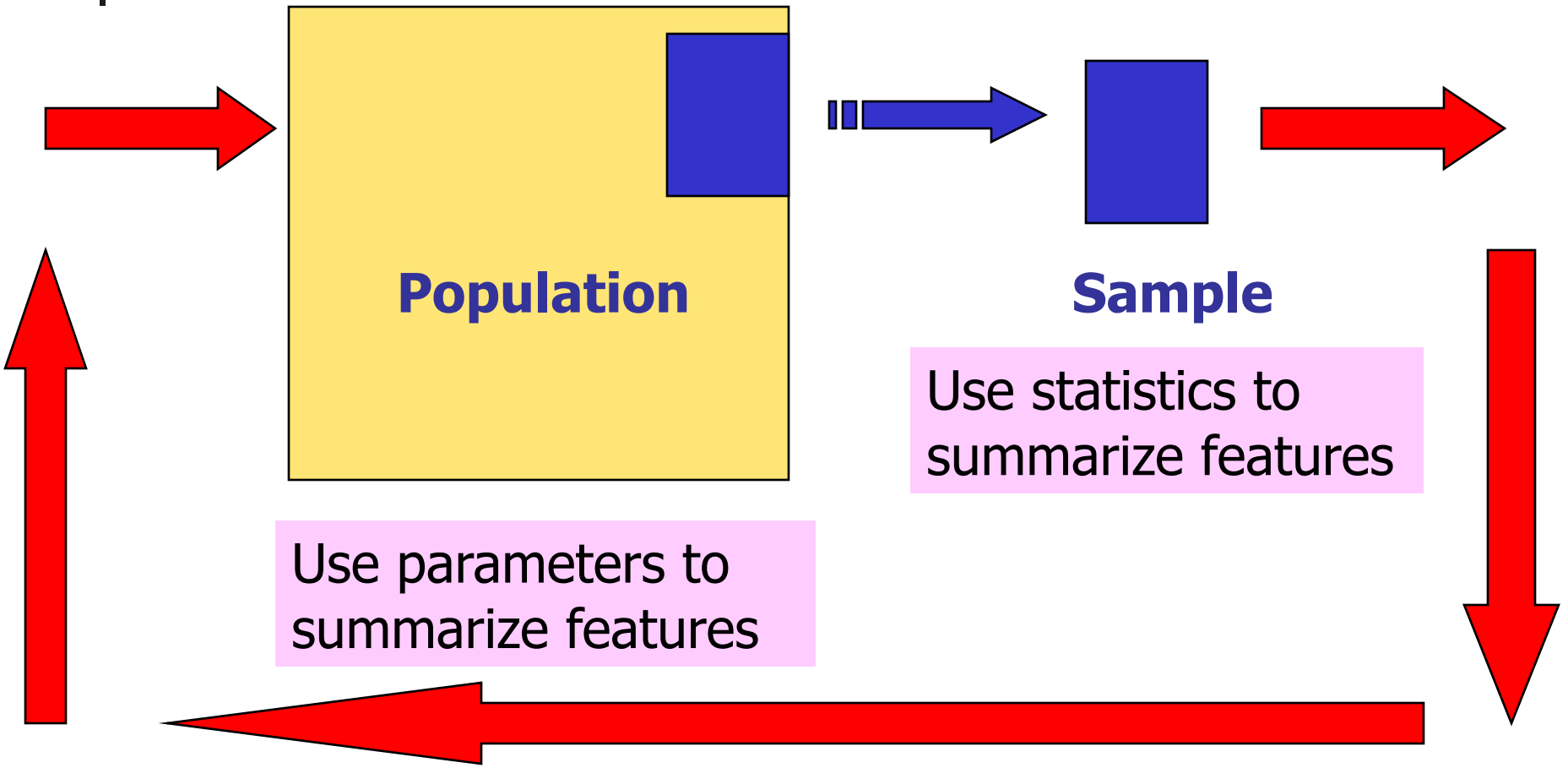


**The advent of the computer**

# Some Important Definitions

- A **Population** (Universe) is the Whole Collection of Things Under Consideration
- A **Sample** is a Portion of the Population Selected for Analysis
- A **Parameter** is a Summary Measure Computed to Describe a Characteristic of the Population
- A **Statistic** is a Summary Measure Computed to Describe a Characteristic of the Sample

# Population and Sample

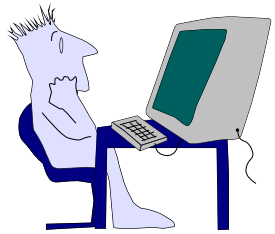


# Statistical Methods

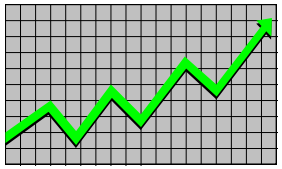
- Descriptive Statistics
  - Collecting and describing data
- Inferential Statistics
  - Drawing conclusions and/or making decisions concerning a population based only on sample data

# Descriptive Statistics

- Collect Data
  - E.g., Survey



- Present Data
  - E.g., Tables and graphs



- Characterize Data
  - E.g., Sample Mean = $\frac{\sum X_i}{n}$

# Inferential Statistics

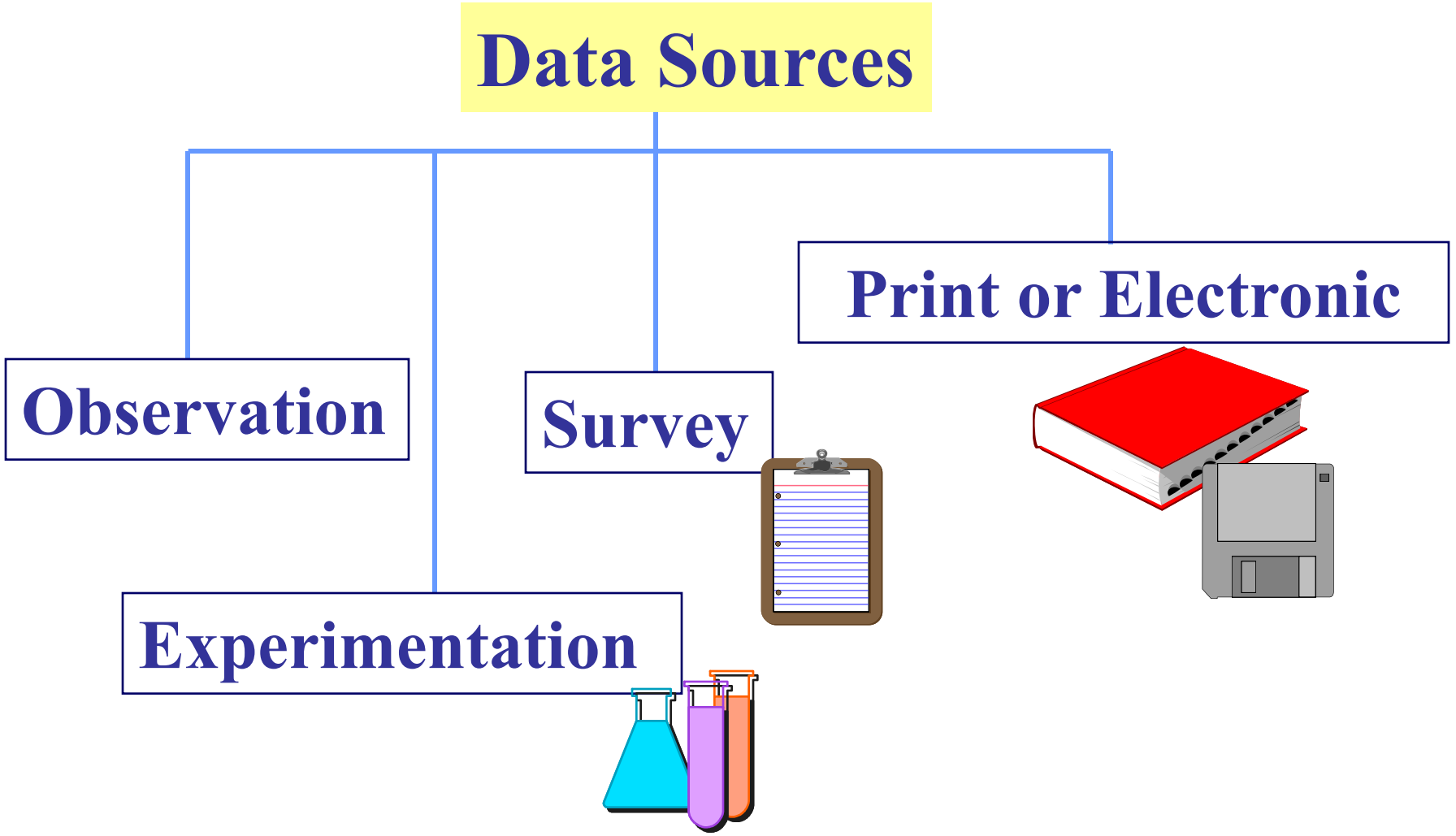
- Estimation
  - E.g., Estimate the population mean weight using the sample mean weight
- Hypothesis Testing
  - E.g., Test the claim that the population mean weight is

**Drawing conclusions and/or making decisions concerning a population based on sample results.**

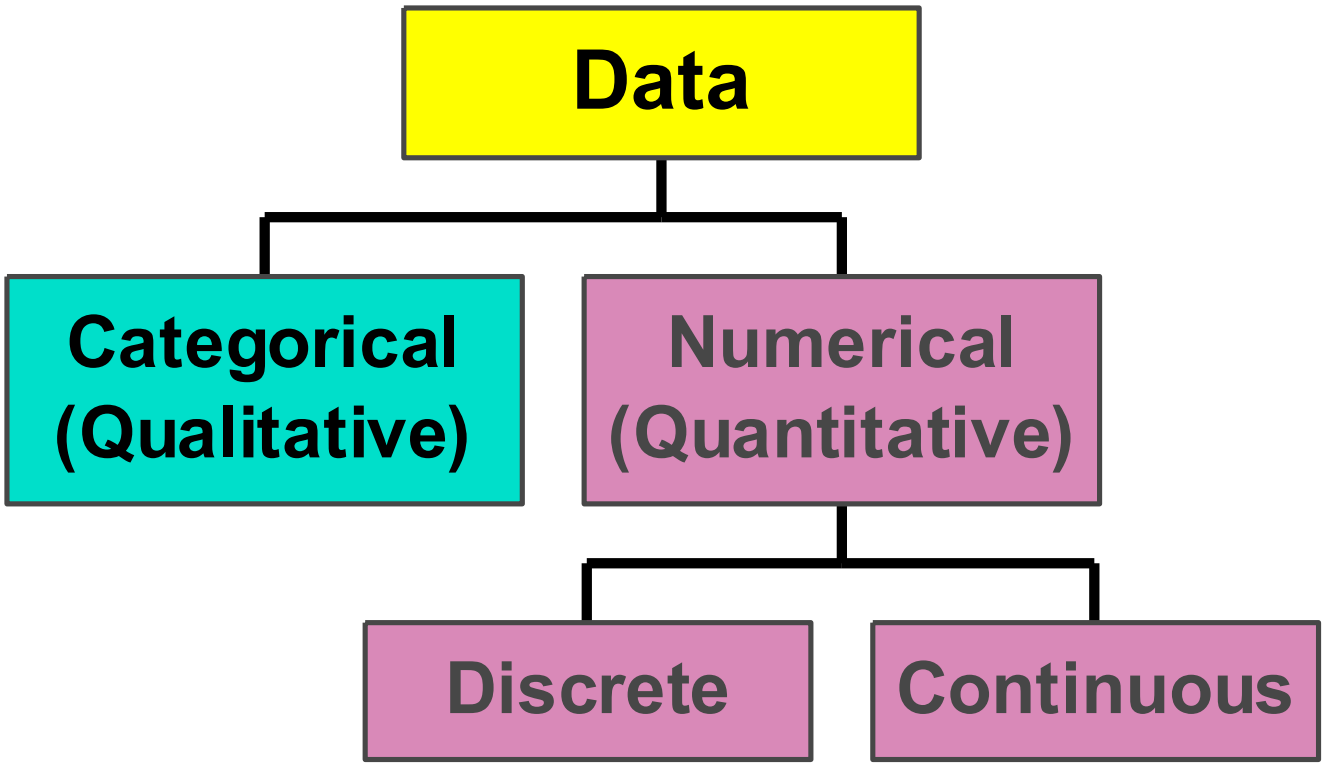
# Why We Need Data

- To Provide Input to Survey
- To Provide Input to Study
- To Measure Performance of Ongoing Service or Production Process
- To Evaluate Conformance to Standards
- To Assist in Formulating Alternative Courses of Action
- To Satisfy Curiosity

# Data Sources



# Types of Data



# Design of Survey Research

- Choose an Appropriate Mode of Response
  - Reliable primary modes
    - Personal interview
    - Telephone interview
    - Mail survey
  - Less reliable self-selection modes (not appropriate for making inferences about the population)
    - Television survey
    - Internet survey
    - Printed survey in newspapers and magazines
    - Product or service questionnaires

# Design of Survey Research *(continued)*

- Identify Broad Categories
  - List complete and non-overlapping categories that reflect the theme
- Formulate Accurate Questions
  - Clear and unambiguous questions use clear operational definitions – universally accepted definitions
- Test the Survey
  - Pilot test on a small group of participants to assess clarity and length

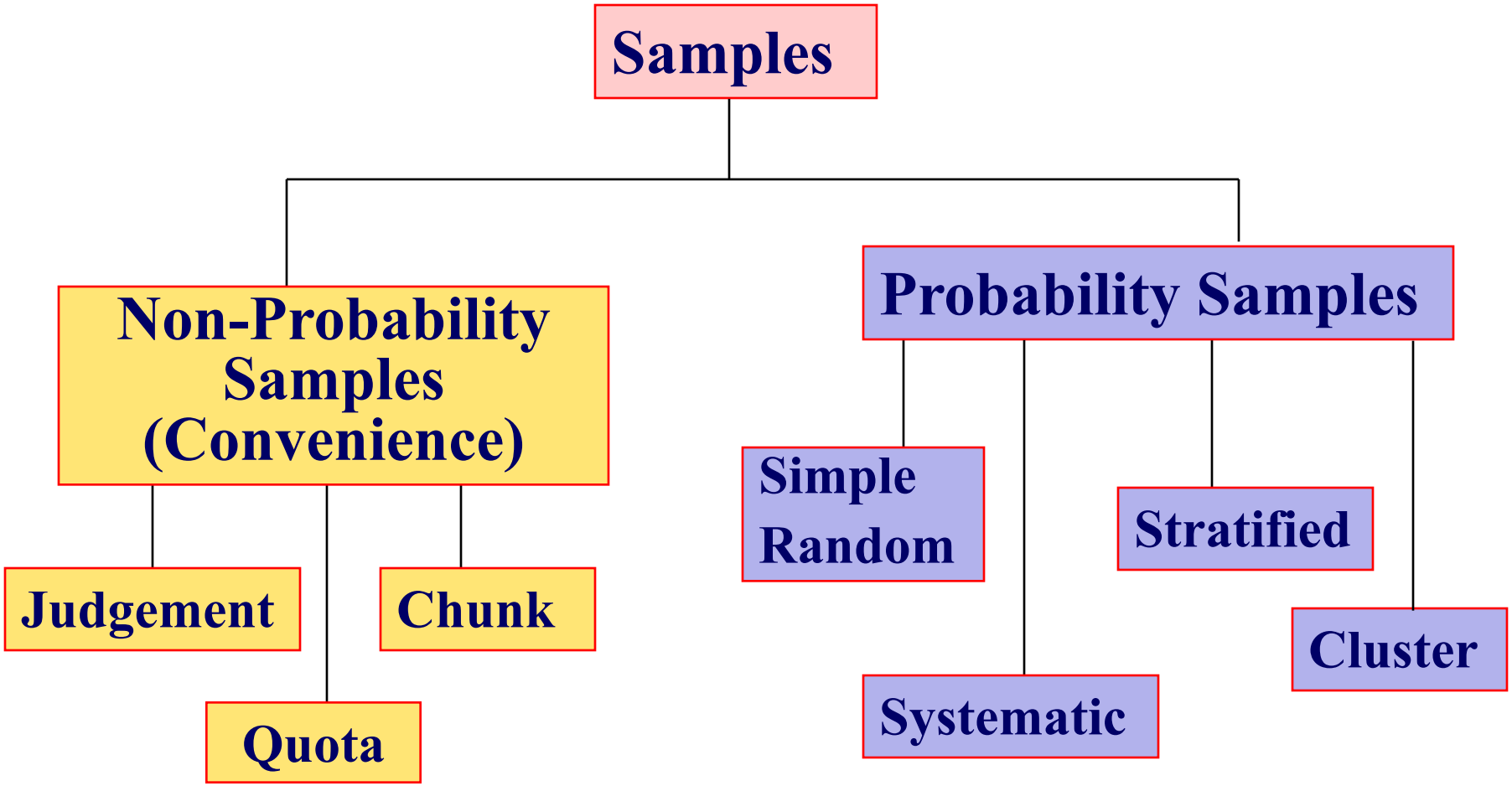
# Design of Survey Research *(continued)*

- Write a Cover Letter
  - State the goal and purpose of the survey
  - Explain the importance of a response
  - Provide assurance of respondent anonymity
  - Offer incentive gift for respondent participation

# Reasons for Drawing a Sample

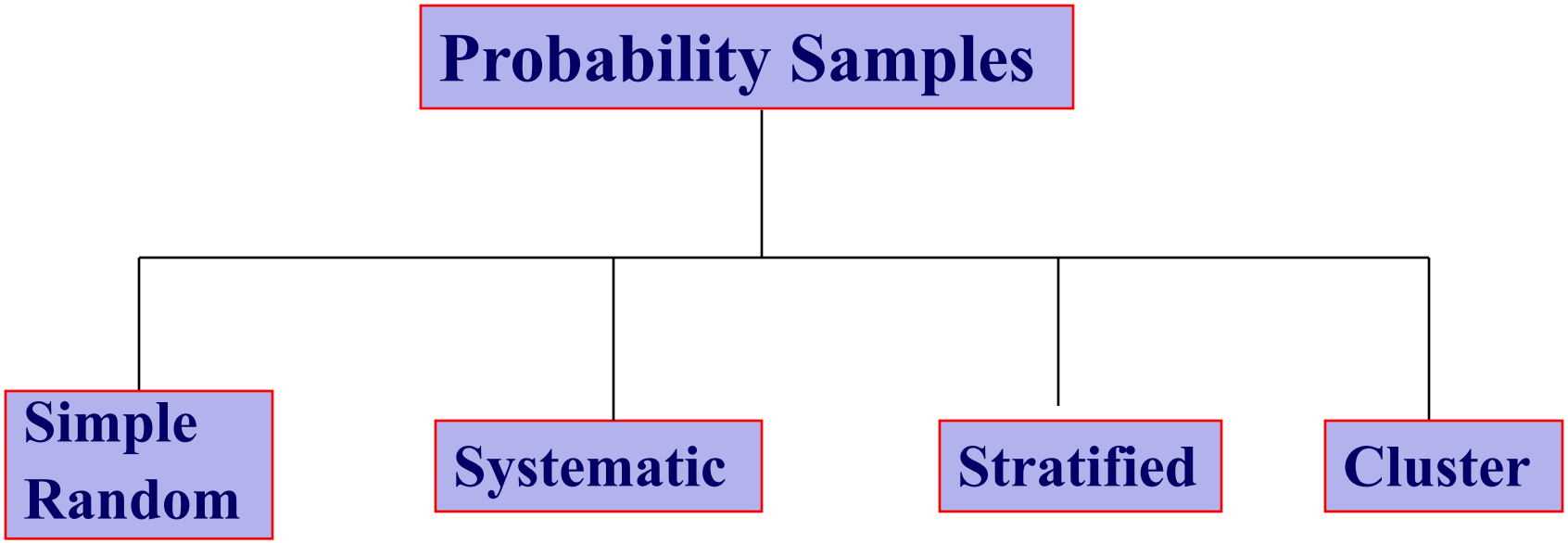
- Less Time Consuming Than a Census
- Less Costly to Administer Than a Census
- Less Cumbersome and More Practical to Administer Than a Census of the Targeted Population

# Types of Sampling Methods



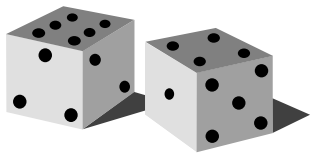
# Probability Sampling

- Subjects of the Sample are Chosen Based on Known Probabilities



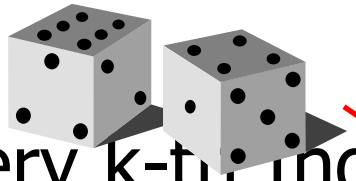
# Simple Random Samples

- Every Individual or Item from the Frame Has an Equal Chance of Being Selected
- Selection May Be With Replacement or Without Replacement
- One May Use Table of Random Numbers or Computer Random Number Generators to Obtain Samples



# Systematic Samples

- Decide on Sample Size: n
- Divide Frame of N individuals into Groups of k Individuals: k=N/n
- Randomly Select One Individual from the 1st Group
- Select Every k-th Individual Thereafter

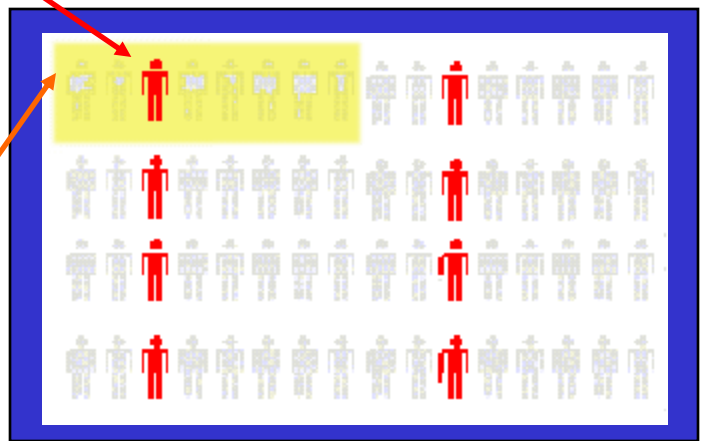


N = 64

n = 8

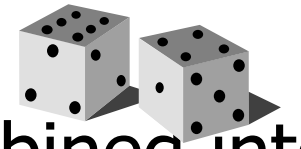
k = 8

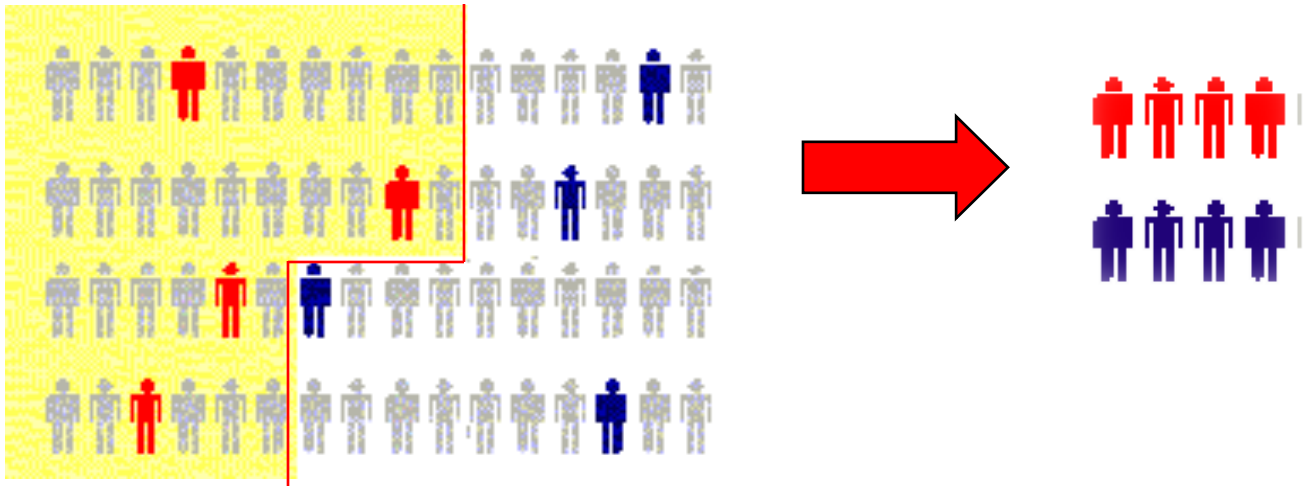
First Group



# Stratified Samples

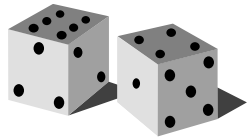
- Population Divided into 2 or More Groups According to Some Common Characteristic
- Simple Random Sample Selected from Each Group
- The Two or More Samples are Combined into One



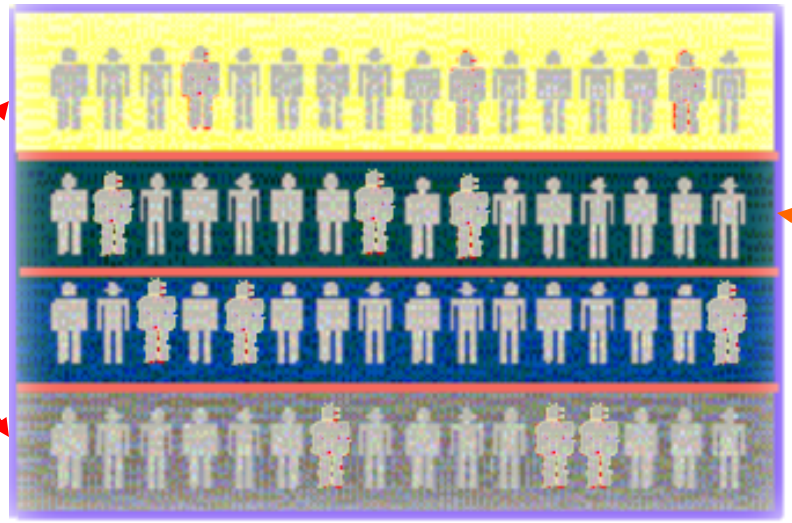


# Cluster Samples

- Population Divided into Several "Clusters," Each Representative of the Population
- A Random Sampling of Clusters is Taken
- All Items in the Selected Clusters are Studied



**Randomly selected 2 clusters**



**Population divided into 4 clusters**

# Advantages and Disadvantages

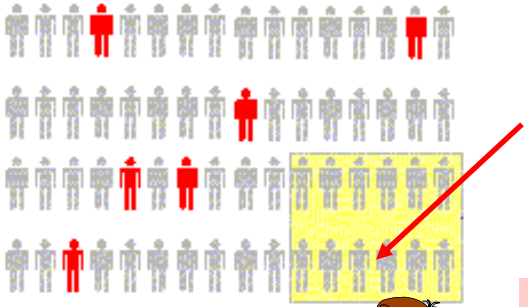
- Simple Random Sample & Systematic Sample
  - Simple to use
  - May not be a good representation of the population's underlying characteristics
- Stratified Sample
  - Ensures representation of individuals across the entire population
- Cluster Sample
  - More cost effective
  - Less efficient (need larger sample to acquire the same level of precision)

# Evaluating Survey Worthiness

- What is the Purpose of the Survey?
- Is the Survey Based on a Probability Sample?
- Coverage Error – Appropriate Frame
- Nonresponse Error – Follow up
- Measurement Error – Good Questions Elicit Good Responses
- Sampling Error – Always Exists

# Types of Survey Errors

- Coverage Error
- Nonresponse Error
- Sampling Error
- Measurement Error



**Excluded from frame**

**Follow up on nonresponses**

**Chance differences from sample to sample**

**Bad Question!**

# Chapter Summary

- Addressed Why a Manager Needs to Know about Statistics
- Discussed the Growth and Development of Modern Statistics
- Addressed the Notion of Descriptive Versus Inferential Statistics
- Discussed the Importance of Data

# Chapter Summary

*(continued)*

- Defined and Described the Different Types of Data and Sources
- Discussed the Design of Surveys
- Discussed Types of Sampling Methods
- Described Different Types of Survey Errors