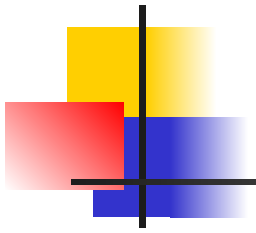
# Statistics for Managers using Microsoft Excel
## 3rd Edition

## Chapter 1
## Introduction and Data Collection



# Chapter Topics

- Why a manager needs to know about statistics
- The growth and development of modern statistics
- Key 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 citizens**

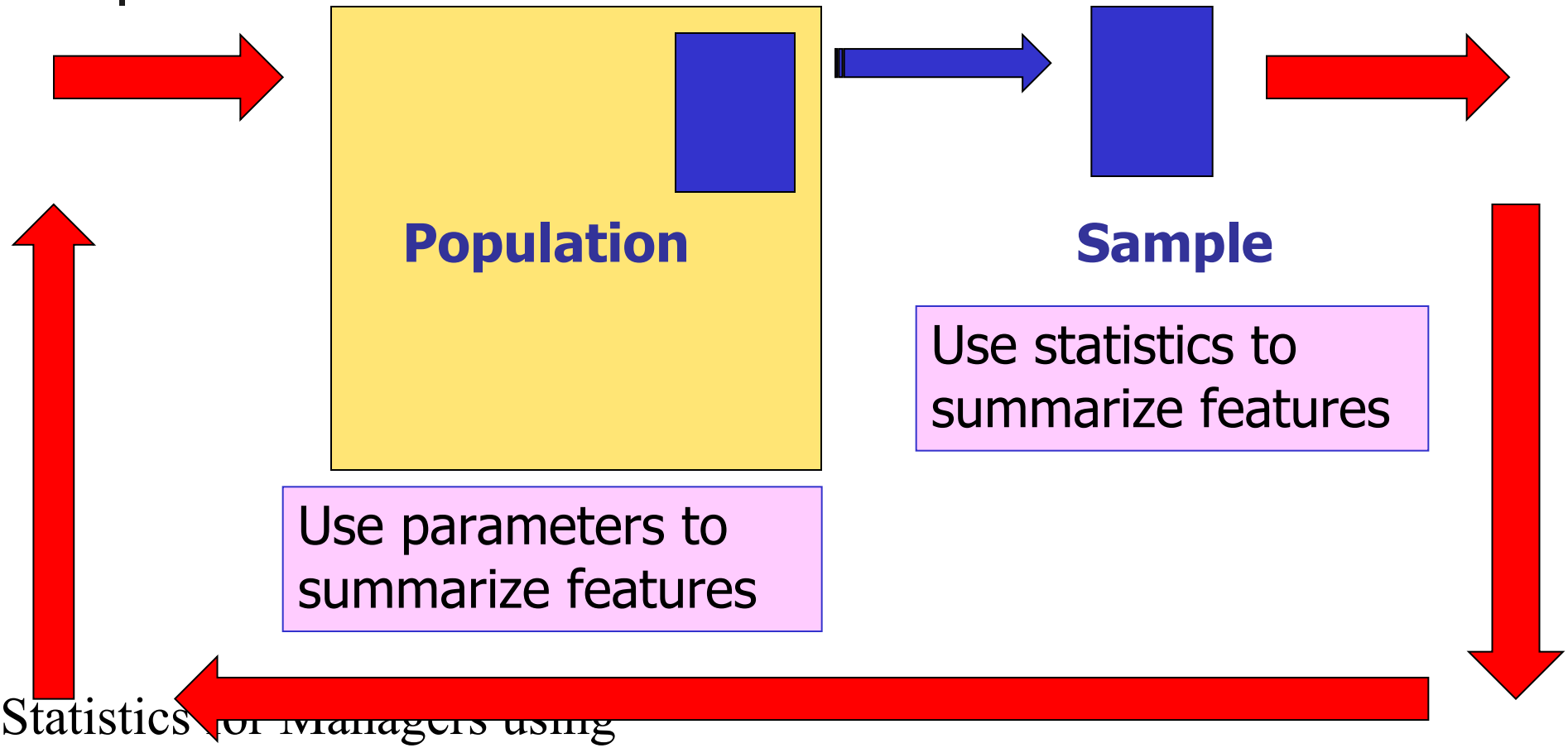
**The development of the mathematics of probability theory**

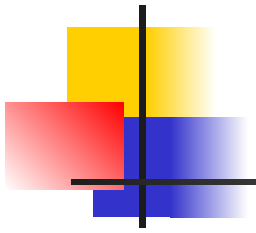
**The advent of the computer**

# Key Definitions

- A population (universe) is the 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



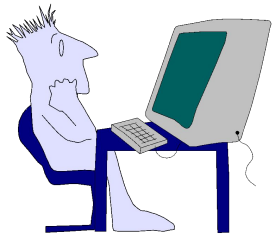


# Statistical Methods

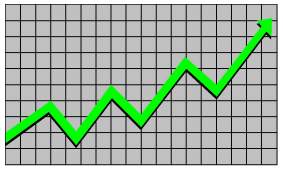
- Descriptive statistics
  - Collecting and describing data
- Inferential statistics
  - Drawing conclusions and/or making decisions concerning a population based only on sample data

# 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 120 pounds

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

# Why We Need Data

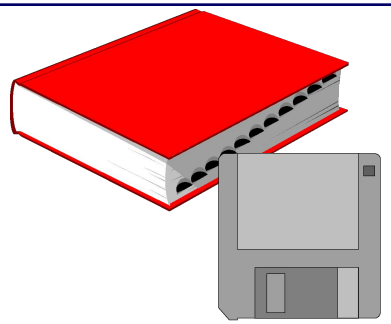
- To provide input to survey
- To provide input to study
- To measure performance of service or production process
- To evaluate conformance to standards
- To assist in formulating alternative courses of action
- To satisfy curiosity

# Data Sources

**Primary**
**Data Collection**

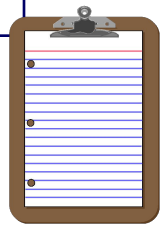
**Secondary**
**Data Compilation**

**Print or Electronic**

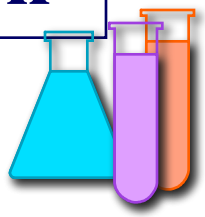


**Observation**

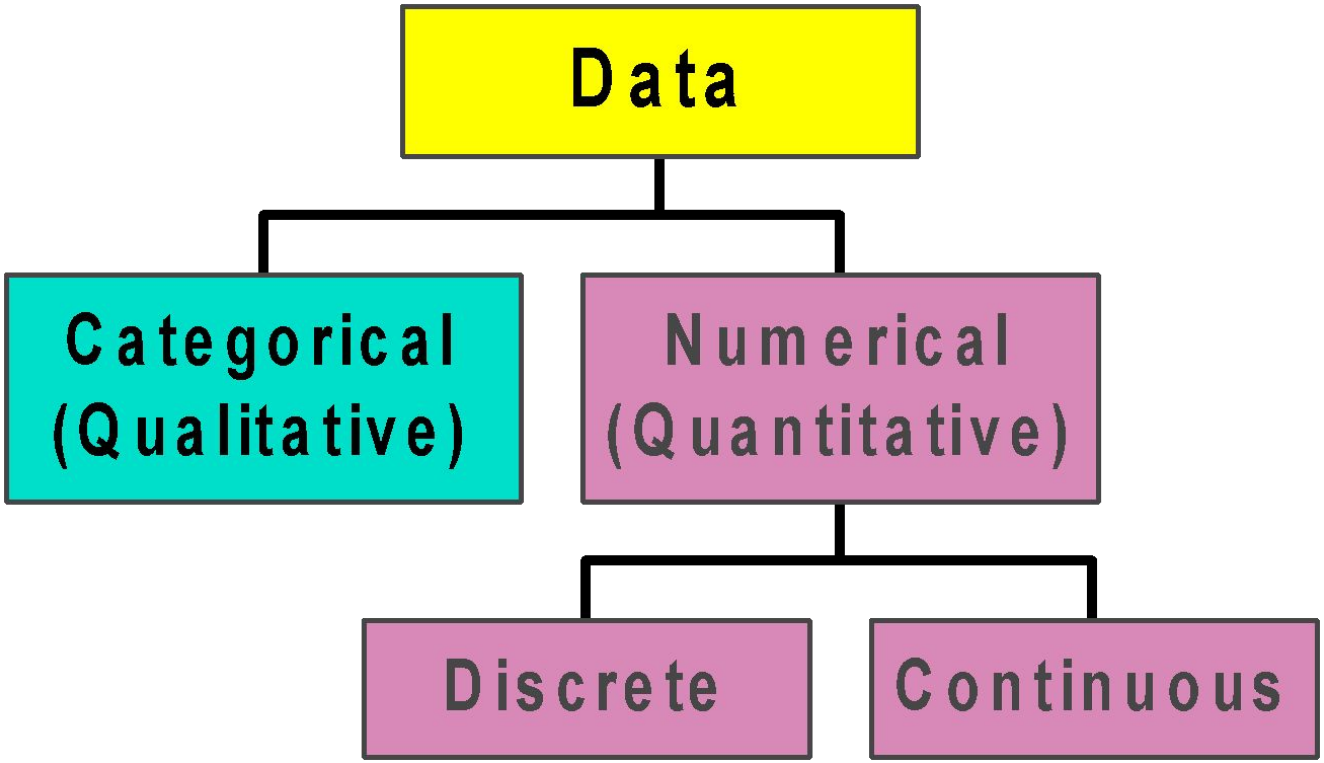
**Survey**



**Experimentation**



# Types of Data



# Design of Survey Research

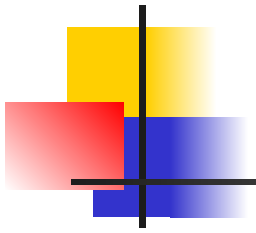
- Choose an appropriate mode of response
  - Reliable primary modes
    - Personal interview
    - Telephone interview
    - Mail survey
  - Less reliable self-selection modes (not appropriate for making inferences about the population)
    - Television survey
    - Internet survey
    - Printed survey on 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
  - Make questions clear and unambiguous. Use universally-accepted definitions
- Test the survey
  - Pilot test the survey on a small group of participants to assess clarity and length

# Design of Survey Research *(continued)*

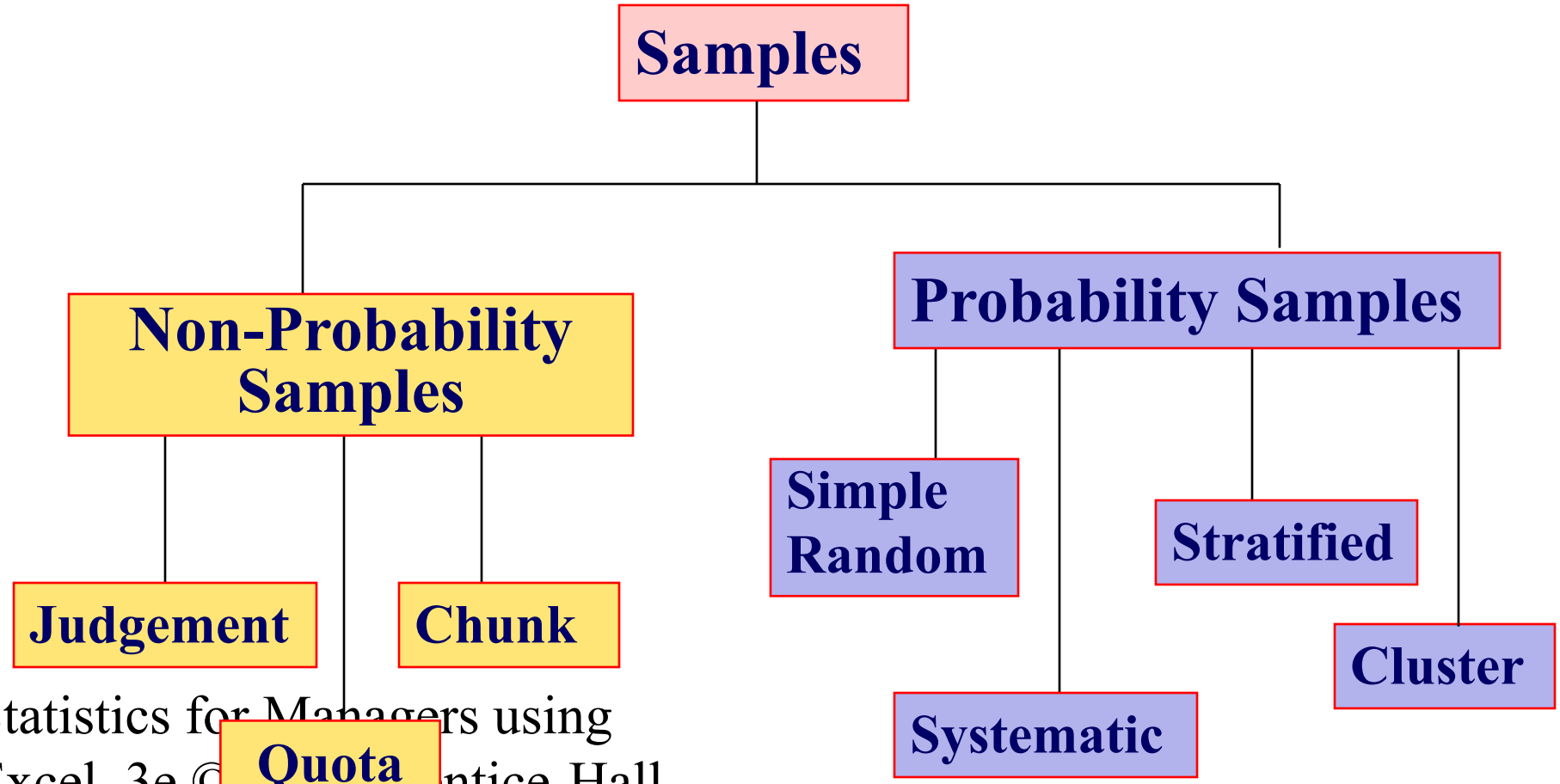
- Write a cover letter
  - State the goal and purpose of the survey
  - Explain the importance of a response
  - Provide assurance of respondent's anonymity
  - Offer incentive gift for respondent participation



# Reasons for Drawing a Sample

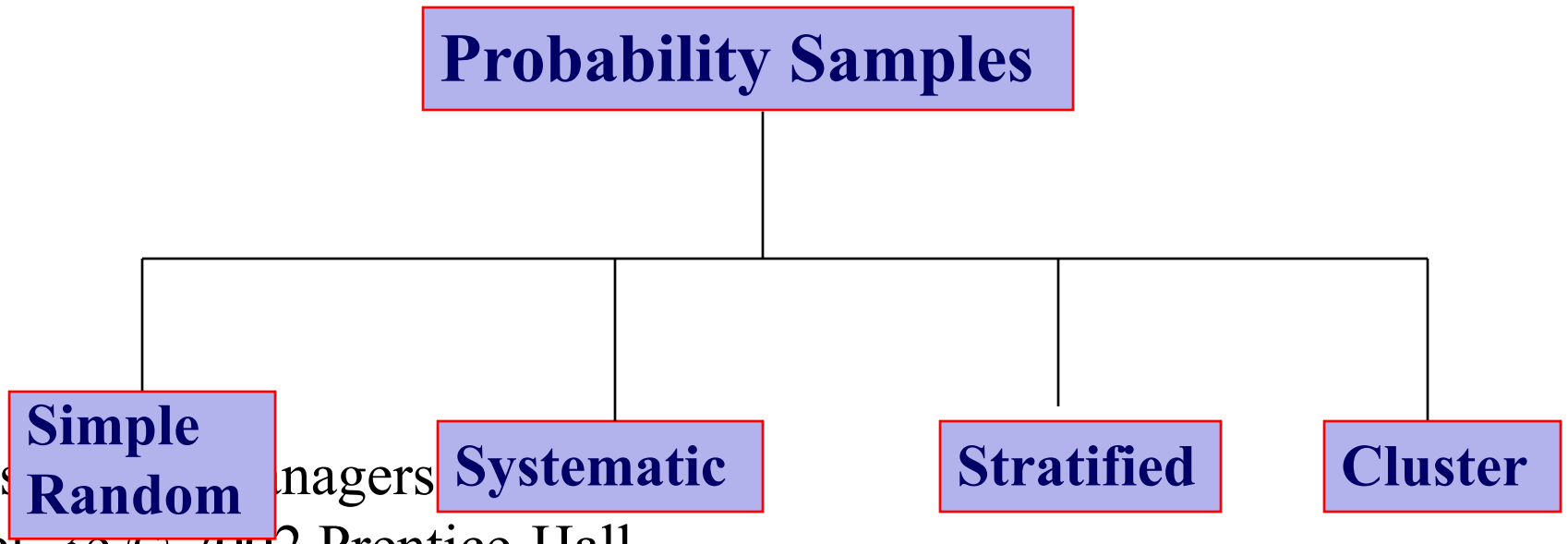
- Less time consuming than a census
- Less costly to administer than a census
- Less cumbersome and more practical to administer than a census of the targeted population

# Types of Sampling Methods



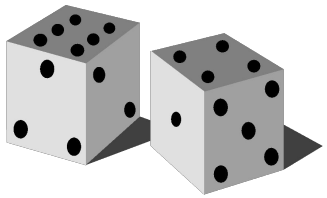
# Probability Sampling

- Subjects of the sample are chosen based on known probabilities



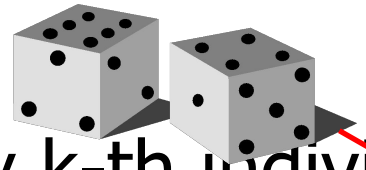
# Simple Random Samples

- Every individual or item from the frame has an equal chance of being selected
- Selection may be with replacement or without replacement
- Samples obtained from table of random numbers or computer random number generators



# Systematic Samples

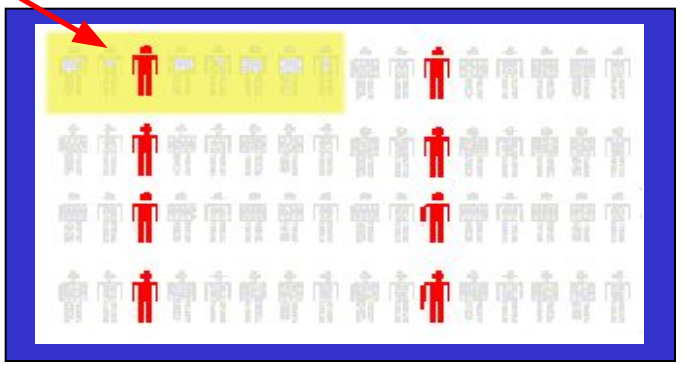
- Decide on sample size: n
- Divide frame of N individuals into groups of k individuals: k=n/n
- Randomly select one individual from the 1st group
- Select every k-th individual thereafter



N = 64

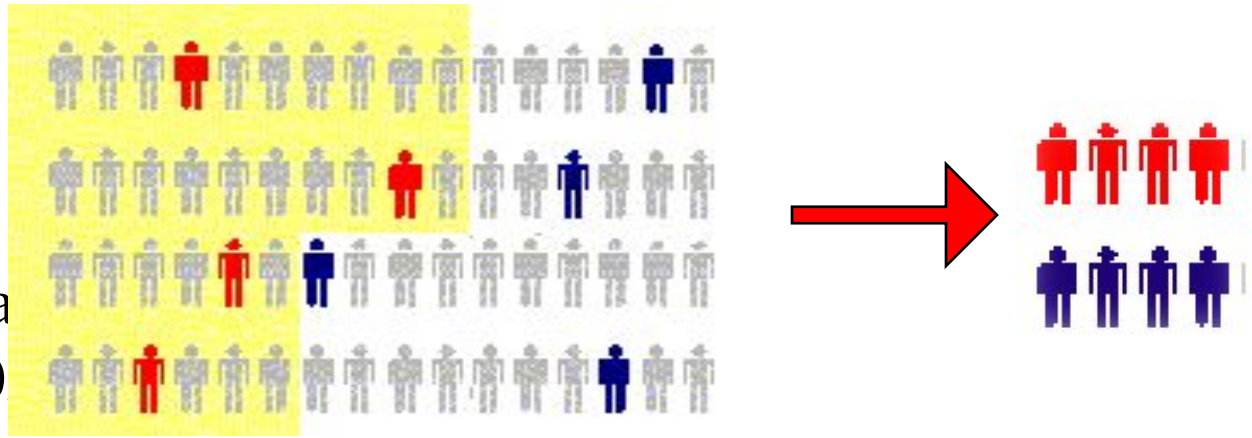
n = 8

k = 8



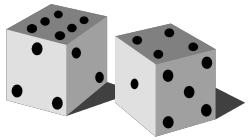
# Stratified Samples

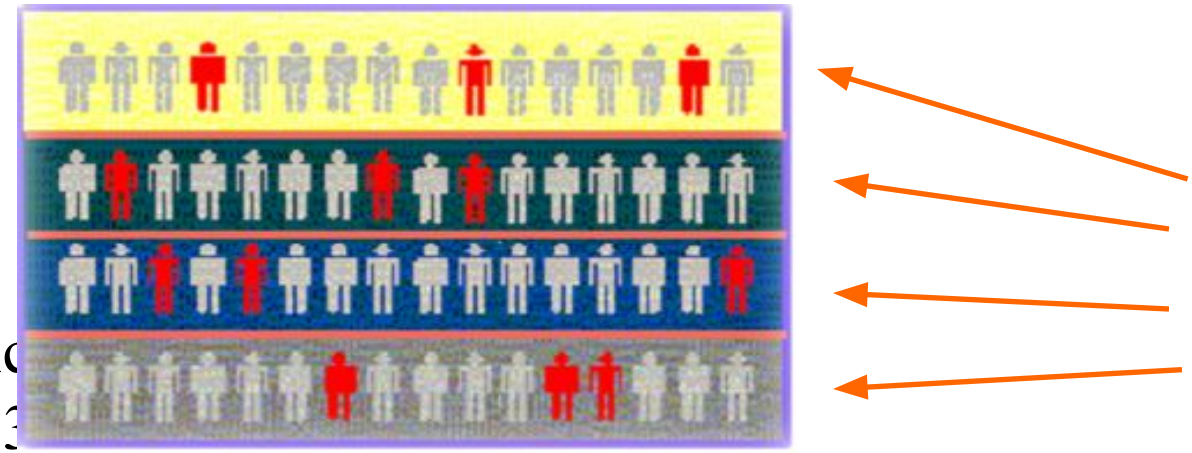
- Population divided into two or more groups according to some common characteristic
- Simple random sample selected from each group
- The two or more samples are combined into one



# Cluster Samples

- Population divided into several "clusters," each representative of the population
- Simple random sample selected from each
- The samples are combined into one

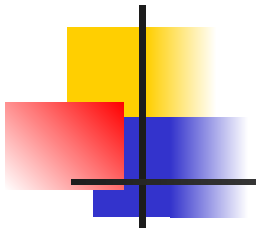




**Population divided into 4 clusters.**

# Advantages and Disadvantages

- Simple random sample and 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)

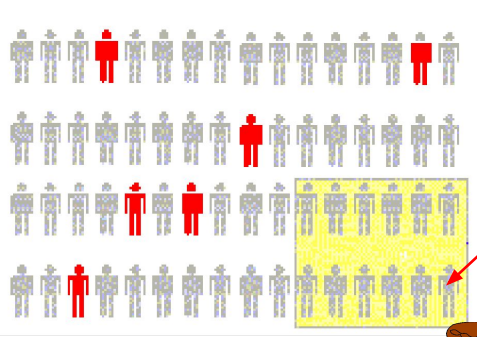


# Evaluating Survey Worthiness

- 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

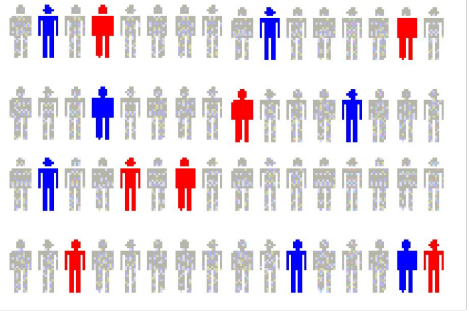
# Types of Survey Errors

- Coverage error
- Non response e
- Sampling error
- Measurement e



**Excluded from frame.**

**Follow up on non responses.**



**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 survey
- Discussed types of sampling methods
- Described different types of survey errors