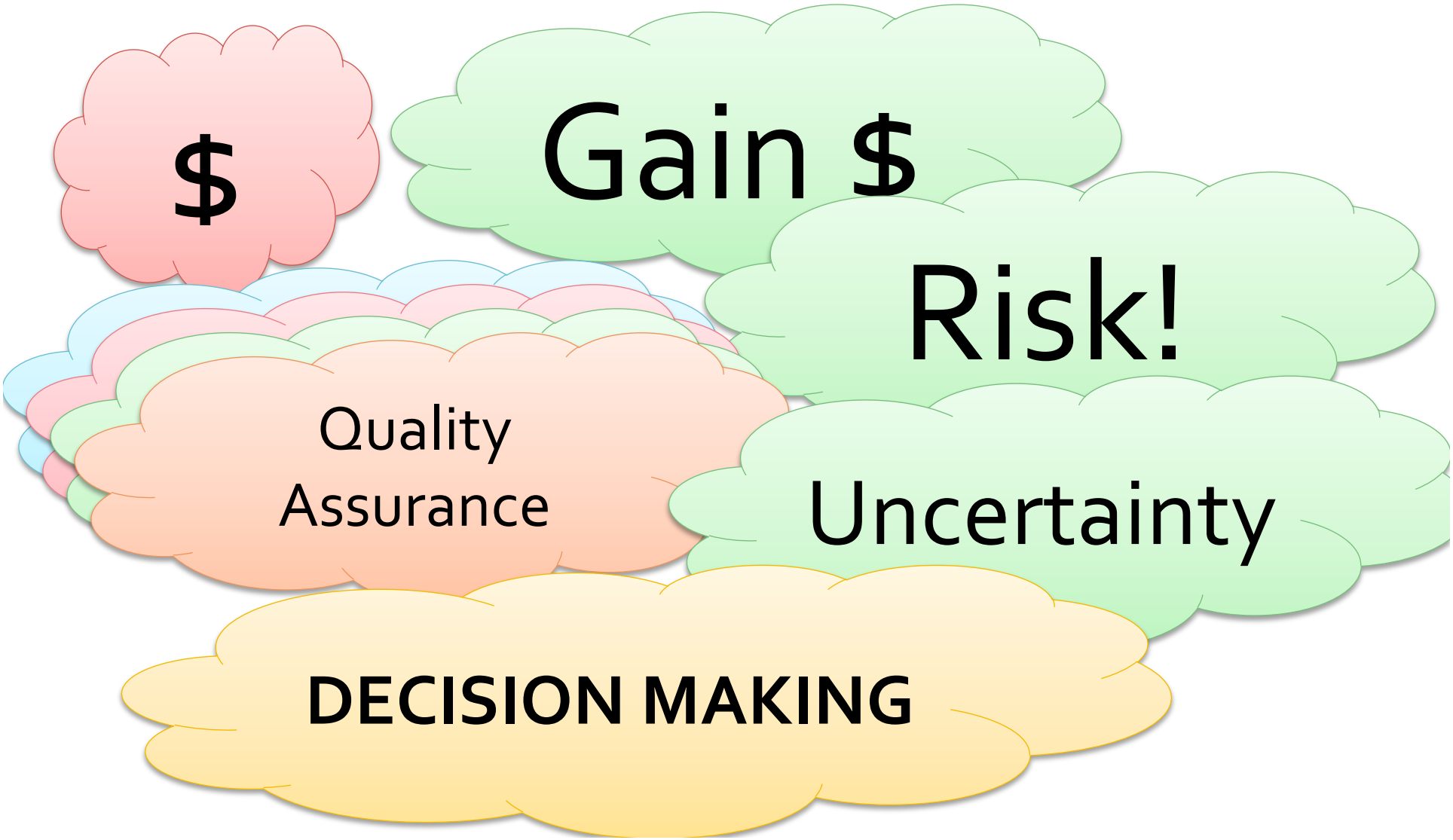
Chapter 1- Data and Statistics

Statistics in Business



Objectives

- Understand how and why research samples are collected
- Explain what is meant by the term random sample and explain how a random sample may be generated
- Describe the four levels of measurements

Objectives

- Statistical applications in business
- Understand important terms used in statistics
- Understand different types of data and its sources
- Be aware of how errors can arise in data
- Understand the role of descriptive statistics and statistical inferences.
- Be aware of ethical guidelines for statistical practice.

Statistical Applications

- Business needs a record of its past history
  - with respect to sales, costs, sources of materials, market facilities, etc.
- Statistics are used to measure progress, financial standing, and economic growth.
  - A record of business changes- of its rise and decline and of the sequence of forces influencing it- it necessary for estimating future developments.

Statistical Applications

- Our behavior in the marketplace help companies make decisions on products to be retained, dropped, or modified.

Business Statistics - Topics

Basic Concepts

Algebra

Probabilities

Ch1 to 3

Normal Distributions
Statistics Tools

Ch 4 to 7

Research Techniques
Hypothesis Testing

Ch8,11,12,14

Data Collection

- Tell a story

Student Name	Gender	Status	University Year	Age Bracket	Hours of Sleep in a day
Sandra	F	Full Time	1	20-23	9-11
Eric	M	Full Time	2	20-23	6-8
Brad	M	Part Time	1	27+	3-5

Elements

Variables

Scales of Measurement

- Scales of Measurement
  - Nominal scale
  - Ordinal scale
  - Interval scale
  - Ratio scale

Scales of Measurement

- Nominal scale
  - A variable consist of labels or names used to identify an attribute of the element
  - 1: Male, 2: Female
  - 1:Full-Time, 2:Part-Time

Scales of Measurement

Nominal + Rank

- Ordinal Scale
  - The data exhibits the properties of nominal data
  - The order or rank of the data is meaningful.

[5]	[4]	[3]	[2]	[1]
Strongly agree	**Agree**	**No Opinion**	**Disagree**	**Strongly disagree**

Scales of Measurement

No Absolute Zero

- Interval Scale
  - All the properties of ordinal data
  - The interval between values is expressed in terms of a fixed unit of measure.
  - Interval data are always numeric.
    - E.g: GPA scores, Temperature

[4.00]	[3.00]	[2.00]	[1.00]	[0.00]
A	**B**	**C**	**D**	**Fail**

Scales of Measurement

Absolute Zero

- Ratio scale
  - All the properties of interval data
  - The ratio of two values is meaningful.
    - E.g: distance, height, weight .
  - Bob is 140lbs, Mary is 70lbs.
  - Bob is _____ heavier than Mary.

Scales of Measurement

Eric has a GPA is 3.00, Sam has a GPA of 1.50.

Can you say Eric is two-times smarter than Sam?

Problem # 1.1

For each of the following, indicate the scale of measurement that best describes the information.

- In 2008, Dell corporation had approximately 78,000 employees.

ANSWER: Ratio scale; there is an absolute zero point associated with the number of employees.

Source: Fortune, May 4, 2009, p.F-48

Problem # 1.1

For each of the following, indicate the scale of measurement that best describes the information.

- USA Today reports that the previous day's highest temperature in the United States was 105 degrees in Death Valley, California.

ANSWER: Interval scale; there is no absolute zero point for temperature.

- Source: USA Today, June 19, 2009, p.12A

Problem # 1.1

For each of the following, indicate the scale of measurement that best describes the information.

- An individual respondent answers “yes” when asked if TV contributes to violence in Canada.

ANSWER: Nominal scale; we could use "1" to identify yes and "0" for no.

Problem # 1.1

For each of the following, indicate the scale of measurement that best describes the information.

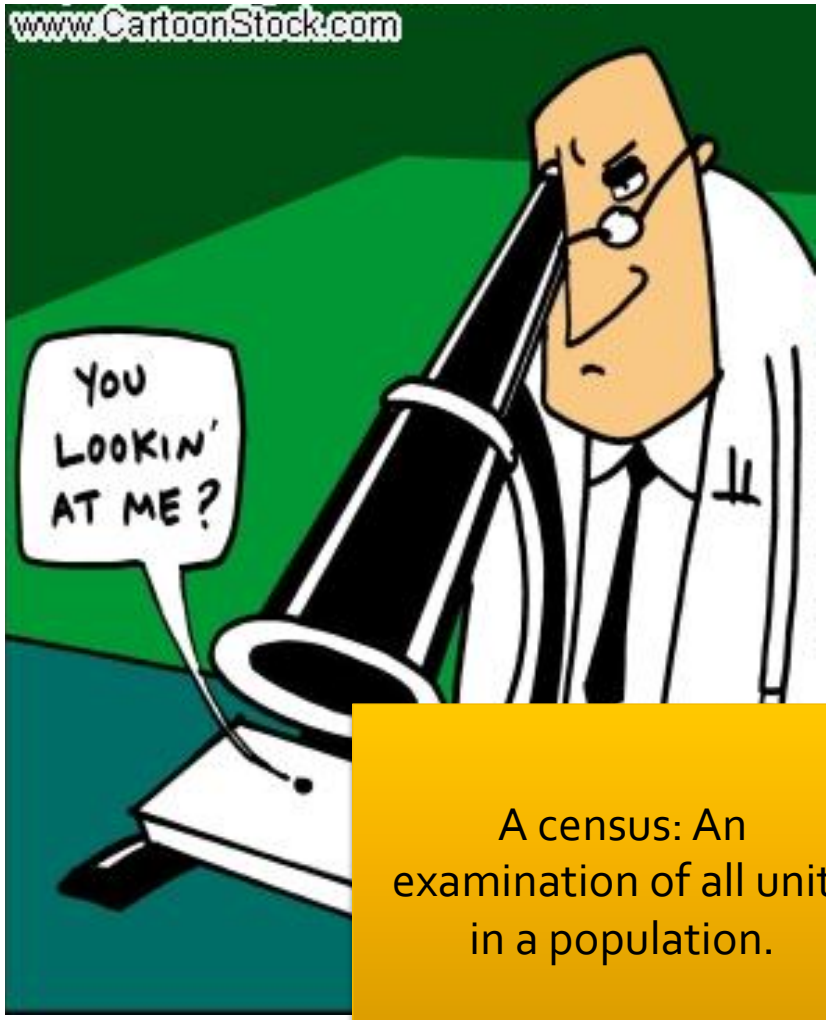
- In a comparison test of family sedans, a magazine rates the Toyota Camry higher than the VW Passat.

ANSWER: Ordinal scale; the cars are ranked but there is no measure for the distance between them.

Research

Population:
All students who have taken COMM 215

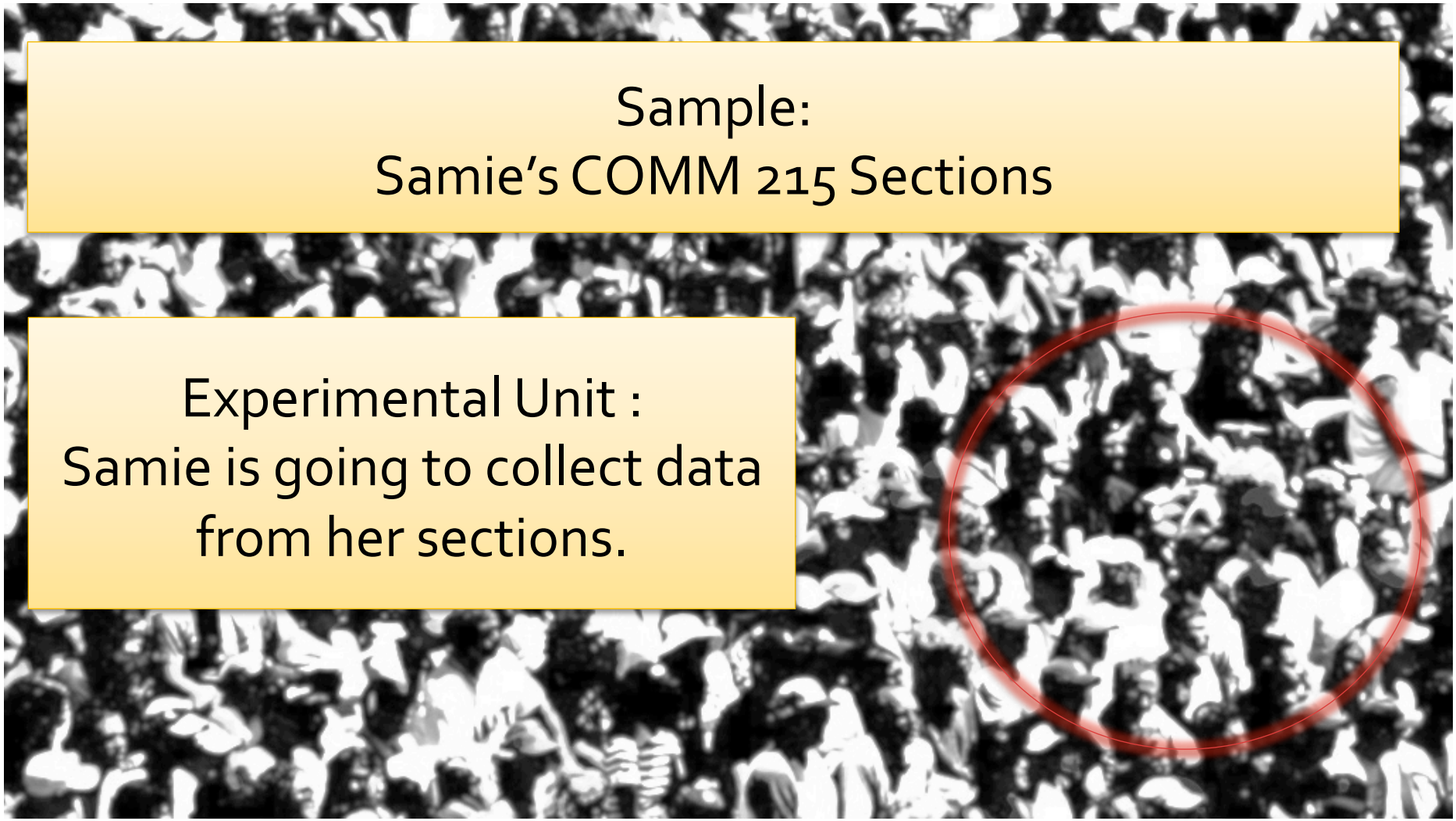
A **population** is a set of units (usually people, objects, transactions, or events) that we are interested in studying.
E.g. All students who have taken COMM 215.



A census: An examination of all units in a population.

http://www.cartoonstock.com/newscartoons/cartoonists/hkh/lowres/hkhn17l.jpg

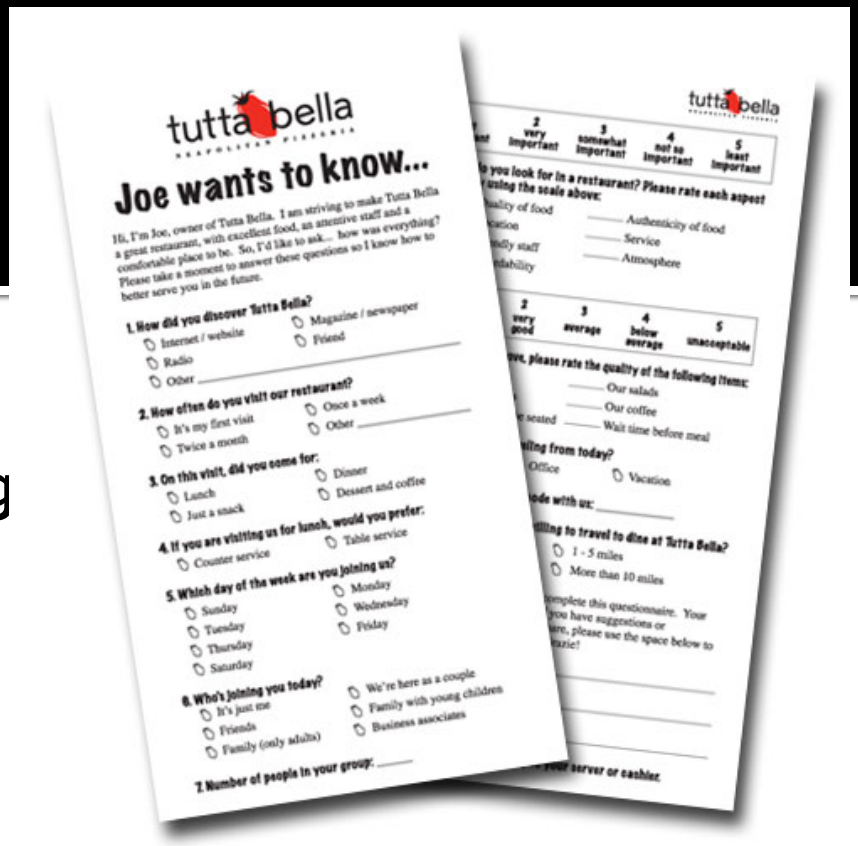
Research



Sample:
Samie's COMM 215 Sections

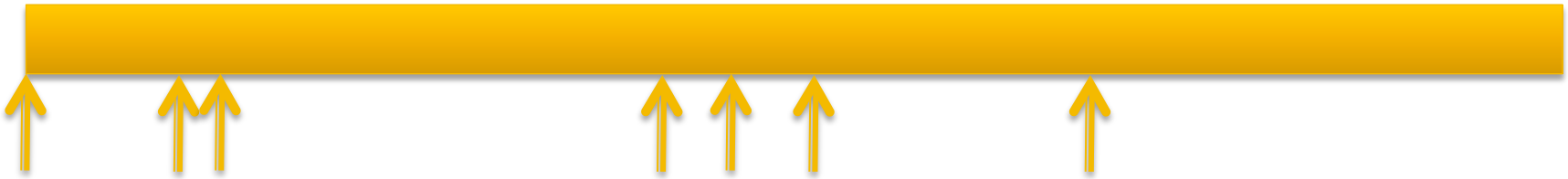
Experimental Unit :
Samie is going to collect data from her sections.

Sampling



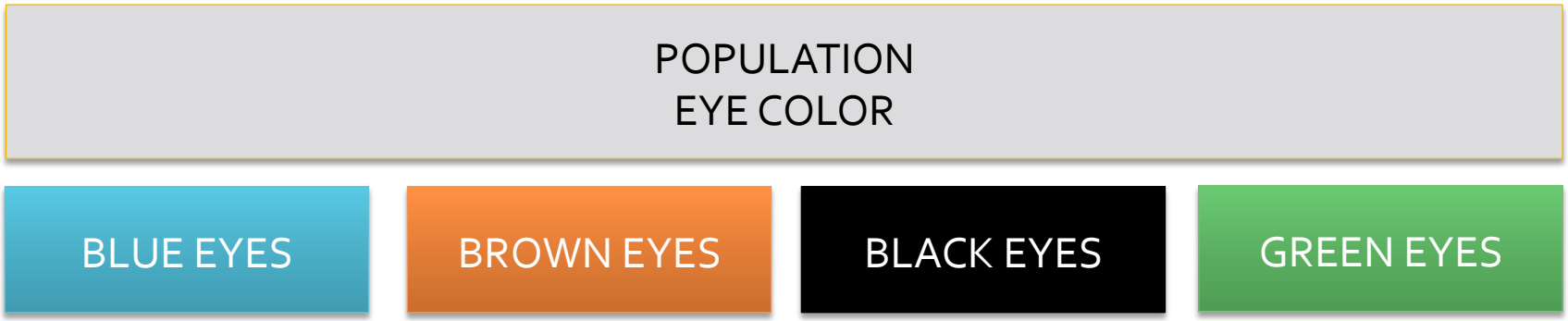
Voluntary response sampling

Simple Random sampling



Sampling

Stratified random sampling



Systematic sampling

Types of Data

- Categorical Data vs Quantitative Data
- Cross-Sectional vs Time Series

Types of Data-Categorical

- Categorical Data
  - Grouped by specific categories.
  - Categorical data are obtained using either the **Nominal** or **Ordinal** scale of measurement.

Types of Data-Quantitative

- Quantitative Data
  - Numeric values
  - Quantitative data are obtained using either the **Ratio** or **Interval** scale of measurement.

Types of Data

- Cross-Sectional Data- same point in time
  - Today, I will be collecting data from all COMM 215 students at the same time.
- Time series data- different times
  - Through out the semester, I will be collecting data every week from a few specific students.

Data Sources

- Existing Sources
  - Statistics Canada (www.statcan.gc.ca)
- Data Acquisition Errors
  - Making errors during data collection
  - Writing 24-year-old as 42 year-old
  - Asking ambiguous questions
  - Inconsistency
  - Spotting outliers
  - Selection bias

Descriptive Statistics

- **Descriptive Statistics**
  utilizes numerical graphical methods
  - to look for patterns in a data set
  - to summarize the information revealed in a data set
  - to present the information in a convenient form.

Statistical Inference

- **Inferential Statistics** utilizes sample data
  - to make estimates, decisions, predictions, or other generalizations about a larger set of data.

Sample of 500 Students in COMM215	
Year	COMM215 Failure Rate
2009	12.7%
2010	12.3%
2011	12.1%
2012	????

Problem # 1.2
Key elements of Statistical Problem

Cola wars is the popular term for the intense competition between Coca-Cola and Pepsi displayed in their marketing campaigns. Their campaigns have featured movie and television stars, rock videos, athletic endorsements, and claims of consumer preferences based on taste tests. Suppose, as part of a Pepsi marketing campaign, 1,000 cola consumers are given a blind taste test. Each consumer is asked to state a preference for brand A or brand B.

- Describe the population.
- Describe the variable of interest.
- Describe the sample.
- Describe the inference.

Problem # 1.2
Key elements of Statistical Problem

- **Population of interest**: all cola consumers
- **Variable of interest**: cola preference
- **Sample**: 1,000 cola consumers selected
- **Inference**: generalization of the cola preference of 1,000 sampled consumers to the population of all cola consumers.

Ethical Guidelines for Statistical Practice

- Statistical thinking
  - Know how the data was collected- Is the data from a reliable source?
  - Is it a random sample? Or "Self-Selected"?
  - Is it possible? Does the data make sense?
- If the biased sample was intentional, with the sole purpose to mislead the public the researchers would be guilty of **unethical statistical practice**.

Misleading Statistics

- " Research Study Reveals…"

As a statistics student who wishes to avoid accepting biased results, what single question should be foremost in your mind as you begin reading the article?

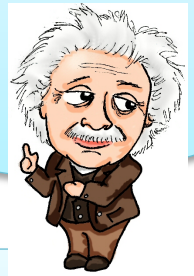
- You should ask yourself what benefits the person or company stands to gain from the conclusions reached by the study.

Misleading Statistics

- *Forensic DNA Databases and Race: Issues, Abuses and Actions held June 19-20, 2008, at New York University. To link to this paper, visit www.gene-watch.org.*

Misleading Statistics

'One in several billion.'



'Oh, about 1 in 100.'

'Assuming that the defendant did not commit this crime, what is the probability that the defendant and the culprit having identical fingerprints?'

'Let me ask you a different question. What is the probability that a fingerprint lifted from a crime scene would be wrongly identified as belonging to someone who wasn't there?'

Misleading Statistics

- It's all about the question asked.
- The defendant's fingerprints had been incorrectly identified as being the same as the ones lifted from the scene.
- It is not a fact, it is a science, and is governed by probabilities.

References

- Statistics for Business and Economics
- Keller, G. (2012). *Statistics for management and economics.* Mason: Cengage Learning.
- McClave, J. T., Benson, G. P., & Sincich, T. (2008). *Statistics for Business and Economics.* New Jersey: Prentice Hall.
- Weiers, R. M. (2011). *Introduction to Business Statistics.* Mason: Cengage Learning.

New Years Resolution

- "Research has shown that, after 6 months, fewer than half the people who make New Year's resolutions have stuck with them, and, after a year, that number declines to around 10%."

- Psychology Today, Jim Taylor, Ph.D

Life: New Year's Resolutions: Why They Don't Stick