

City University of New York (CUNY)

## CUNY Academic Works

---

Open Educational Resources

Queensborough Community College

---

2020

### Clear-Sighted Statistics: Module 1: Introduction to Statistics (slides)

Edward Volchok

*CUNY Queensborough Community College*

[How does access to this work benefit you? Let us know!](#)

More information about this work at: [https://academicworks.cuny.edu/qb\\_oers/144](https://academicworks.cuny.edu/qb_oers/144)

Discover additional works at: <https://academicworks.cuny.edu>

---

This work is made publicly available by the City University of New York (CUNY).

Contact: [AcademicWorks@cuny.edu](mailto:AcademicWorks@cuny.edu)

# Introduction to Statistics

## Module 1



0



The goal of this course is to help you

**SEE:**

**S**elect the  
right technique

**E**xecute it  
properly

**E**xplain your  
findings



1

## Lecture Objectives

Explain the discipline of statistics and some of the common misconceptions about it

Describe what you need to succeed in an introductory statistics class

Trace the origins of the science of statistics

Discuss how the study of statistics should lead to healthy skepticism

Why learning statistics is important for your future



2

**What is (are) Statistics?**



3

Clear-Sighted Statistics

## Statistics has 2 distinct meanings

The discipline of analyzing data & making appropriate decisions based on the analysis

↔

Data or summarized data derived from a sample

4

Clear-Sighted Statistics

## First meaning of Statistics

**Statistics:** The science of collecting, organizing, presenting, analyzing, and interpreting data used in making decisions

Statistics has a utilitarian purposes

σ   ρ   Σ   μ   α   β

5

Clear-Sighted Statistics

## Second meaning of Statistics

Statistic	Parameter
<ul style="list-style-type: none"> <li>A random variable or datum sourced from a sample</li> <li>Often symbolized with Latin letters</li> </ul>	<ul style="list-style-type: none"> <li>A random variable or datum sourced from a population</li> <li>Often symbolized with Greek letters</li> </ul>

6

Clear-Sighted Statistics

## Populations vs. Samples

Population	Sample
<ul style="list-style-type: none"> <li>A complete set of the objects or subjects of interest</li> </ul>	<ul style="list-style-type: none"> <li>A subset of a population chosen in such a way that it represents the population</li> </ul>

7

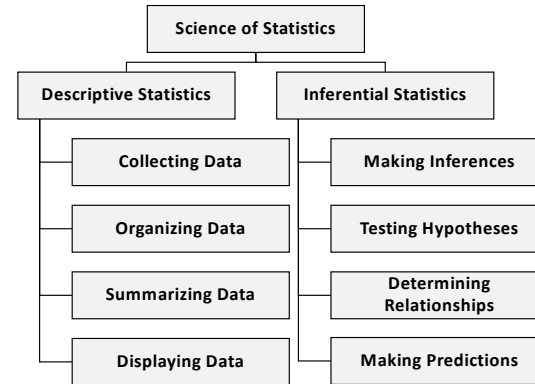
## Tukey: Statistics ≠ Mathematics

“Statistics is a science in my opinion, and it is no more a branch of mathematics than are physics, chemistry, and economics; for if its methods fail the test of experience—not the test of logic—they are discarded.”\*



John W. Tukey  
1915 - 2000

\*L. V. Jones (ed.) The Collected Works of John W. Tukey: Philosophy and Principles of Data Analysis 1949-1964, Volume III. (New York: Chapman & Hall, 1986), p. 66.



## Descriptive Statistics: Example #1

Only 33.33% of community college students referred to remedial math complete this sequence\*

This means only 1/3 of the students who need remedial math complete this requirement

\*<http://ccrc.tc.columbia.edu/Community-College-FAQs.html>



## Descriptive Statistics: Example #2

45% of community college graduates earn more than \$45,000 a year\*

This statistic describes 45 of 100 community college graduates

\*<http://salarysurfer.cccco.edu/SalarySurfer.aspx>



## Inferential Statistics deal with...

Decisions, estimates, predictions, or generalizations about a population, based on a sample

A *Population*\* is a collection of all possible individuals, objects, or measurements of interest

A *Sample* is a portion, or part, of the population of interest

\*Sometimes called a Universe

12

## Statistics and Empiricism

Empiricism: All knowledge is derived from sense-perception

Empirical knowledge is *a posteriori* (from the latter). It comes from experience or inductive reasoning

Logic and mathematics rely on *a priori* (from the former), or deductive reasoning

Knowledge is not based on conjecture, abstractions, or appeals to authority

13

## What you need to succeed in an introductory statistics class

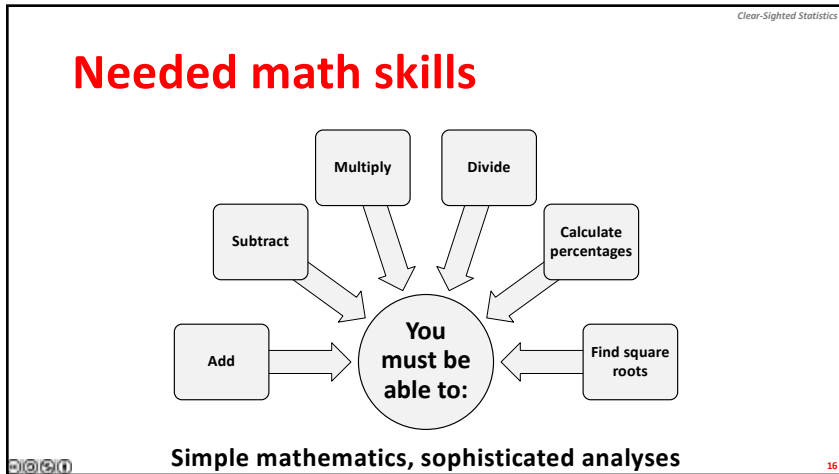
14

## First: A Chinese Proverb

Not hearing is not as good as hearing.  
Hearing is not as good as seeing.  
Seeing is not as good as knowing.  
Knowing is not as good as acting.  
True learning continues until it is put into action.

**Learning statistics requires effort**

15



16


Clear-Sighted Statistics

## Three Components of Statistical Literacy\*

Familiarity with using statistical thinking to deal with problems containing data.

Appreciating the role statistics plays in decision-making by going through the process of collecting, displaying, and analyzing data.

Being able to read data resources, data analyses, and summarized information critically."



Jun, Li. "Statistics Education for Junior High Schools in China." Paper in the Proceedings of the 2004 Roundtable of the International Association for Statistics Education

17

17

Clear-Sighted Statistics

## The Rise of Statistical Science

18

18

- Clear-Sighted Statistics
- ## 3 factors led to the rise of modern statistics\*
- Development of probability theory in the 17<sup>th</sup> century
- Rise of the modern nation state
- Interest in the scientific study of human behavior
- \*Walker, Helen M., *Studies in the History of Statistical Method: With Special Reference to Certain Educational Problems*, (Philadelphia: Williams & Wilkins, 1929)
- 19

19

Clear-Sighted Statistics

## Statistic is the Science of the State


**Discipline started in the Enlightenment**  
(18<sup>th</sup> century Europe)

**Political Arithmetic**

**Etymology:** (Origin of the word Statistics)  
**Latin:** *Statisticum* ("of the state")  
**German:** *Statistik* ("data relating to the state")  
**Italian:** *Statistica* ("statesman")

20

Clear-Sighted Statistics




## Quételet Exemplifies the Rise of Modern Statistics

Belgian astronomer & proto-sociologist

Among the first to use statistics to study social phenomena (Social Physics\*)

Focused on the "average man"

First to apply "bell curve" to social matters




Adolphe Quételet  
1796-1874

\*1835: *Sur l'homme et le développement de ses facultés, ou Essai de physique sociale*

21

Clear-Sighted Statistics



## Auguste Comte used the term "Social Physics" before Quételet

Comte disagreed with Quételet's method for collecting data

Eventually called Social Physics **Sociology**


Auguste Comte  
1789 - 1857

22

Clear-Sighted Statistics

**"[Statistics is] the most important science in the whole world: for upon it depends the practical application of every other [science]."\***

**– Florence Nightingale**




**Florence Nightingale**  
(1820 – 1910)

Nightingale's annotations to Adolphe Quételet's *Physique Sociale* cited in "Florence Nightingale," I. Bernard Cohen, *Scientific American*, 250(3): p. 136, March 1984.

23

Clear-Sighted Statistics

**“Statistics is the grammar of science.”**



**Karl Pearson**  
(1857 – 1936)

Pearson, Karl. *The Grammar of Science*, (New York: Cosimo, 2007). Originally published in 1892.

24

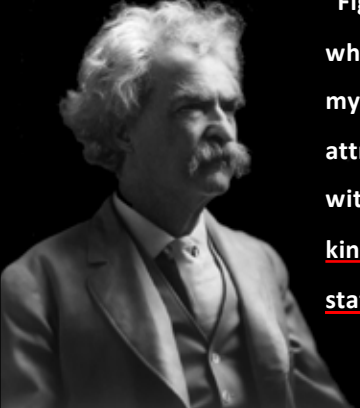
24

Clear-Sighted Statistics

## Statistics and Lies

25

25



**“Figures often beguile me, particularly when I have the arranging of them myself; in which case the remark attributed to Disraeli would often apply with justice and force: ‘There are three kinds of lies: lies, damned lies and statistics.’”**

*Autobiography of Mark Twain*

**Mark Twain**  
1835 - 1910


26

26

Clear-Sighted Statistics

## Statistics and lies

“...may I quote a remark I once heard: ‘There are three kinds of lies: white lies, which are justifiable; common lies—these have no justification; and statistics.’ Our meaning is similar when we say: ‘Anything can be proven by figures’; or, modifying a well-known quotation from Goethe, with numbers ‘all men may content their charming systems to defend.’”\*



**Richard von Mises**  
1883 - 1953

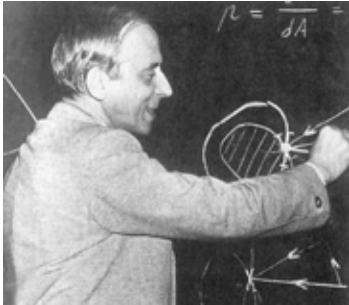
\*Richard von Mises, *Probability, Statistics and Truth*, (New York: Dover Publications, 1981), p. 1.

27

27



## Statistics and “truthful” conclusions



**Richard von Mises**  
1883 - 1953

“...a great deal of meaningless and unfounded talk is presented to the public in the name of statistics....my purpose is to show that, starting from statistical observations and applying to them a clear and precise concept of probability it is possible to arrive at conclusions that are as reliable and ‘truthful’ as those obtained in any other exact science.”\*

\*Richard von Mises, *Probability, Statistics and Truth*. (New York: Dover Publications, 1981), p. 1.

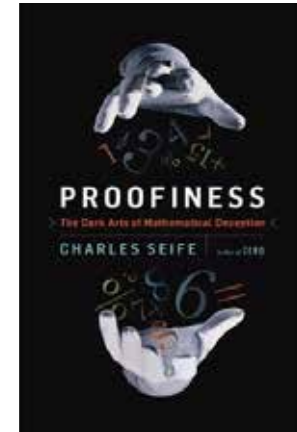
28

28

## The dark art of Proofiness...

“[Proofiness] is...using numbers to prove what you know in your heart is true, even when you know it’s not. Numbers have a particular ability to fool us. It’s using that ability to turn nonsense into something that is believable with numbers.”

Charles Seife  
New York University  
Arthur L. Carter Journalism Institute  
<http://well.blogs.nytimes.com/2010/10/29/the-dark-art-of-statistical-deception/>



29

29

“Statistics are like bikinis.

What they reveal is suggestive,

but what they conceal is vital.”\*

Prof. Aaron Levenstein  
Baruch College/CUNY

\*Cited in Alan Murray (2010), *The Wall Street Journal Essential Guide of Management*. (New York: HarperCollins, 2010). Chapter 9, location 1862.

30

30

## The Importance of Skepticism

31

31

# A brief moment for classical Greek philosophy

## Cynicism and Skepticism



## Cynicism

- Society is incompatible with happiness
- Renounced wealth and physical comfort
- Disregarded all social conventions
- There's no "truth"
- People are self-interested and greedy



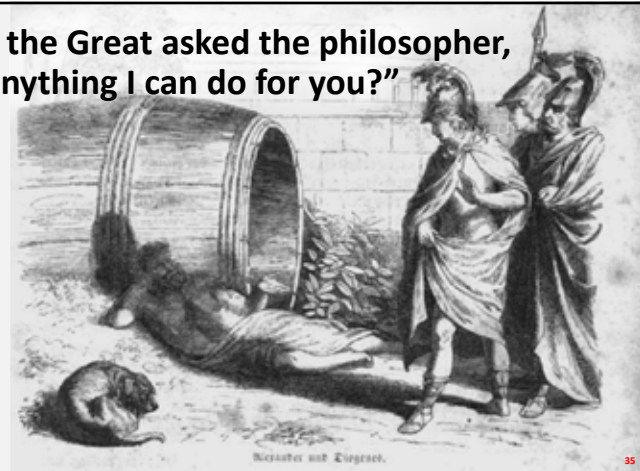
Diogenes  
c. 412 BCE – 323 BCE  
"I am looking for an honest man"

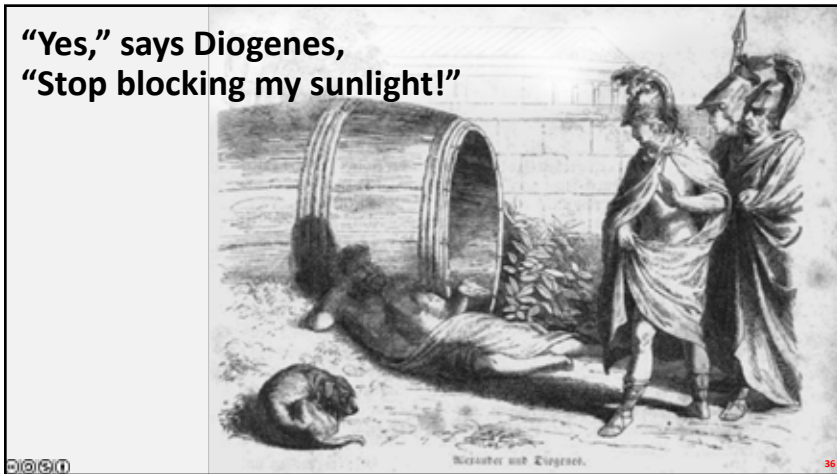


## Diogenes meets Alexander the Great



Alexander the Great asked the philosopher, "Is there anything I can do for you?"





36

**Skepticism**

- Truth is uncertain
- Must apply reason *and* critical thinking
- Skeptics doubt and investigate

Pyrrho  
c. 360 BCE – 270 BCE

37

37

**Data Scientist Cathy O’Neal:  
On Being a Data Skeptic\***

“A skeptic is someone who maintains a consistently inquisitive attitude toward facts, opinions, or (especially) beliefs stated as facts. A skeptic asks questions when confronted with a claim that has been taken for granted. That’s not to say a skeptic brow-beats someone for their beliefs, but rather that they set up reasonable experiments to test those beliefs.”

Cathy O’Neal  
\*On Being a Data Skeptic

38

38

**Statistics leads to  
Skepticism not Cynicism**

39

39

# The Skeptic's Starting Point: The 5 "W" Questions



**Who?**  
**What?**  
**When?**  
**Where?**  
**Why?**



# Why learning statistics is important for your future



## Who uses statistics?

Government Officials	Marketers	Marketing Researchers	Data Scientists
Financial Analysts	Quality Control Experts	Insurance Executives	Social Scientists
Educators	Politicians	Pollsters	Health Care Professionals



# Jobs, Salaries, and Statistics

	Job Title	Statistics Needed	Starting Salary
1	Community Organizer	No	\$14,954
2	Corrections Officer	No	\$27,000
3	Criminologist (Counseling)	No	\$23,963
4	Criminologist (Research)	Yes	\$31,500
5	Human Resources	No	\$27,040
6	Marketing Research Analyst	Yes	\$32,000
7	Marketing Research Interviewer	No	\$12,000
8	Private Detective	No	\$26,750
9	Probation Officer	No	\$23,594
10	Public Opinion Researcher (w/Statistics)	Yes	\$27,101
11	Public Opinion Researcher (No Statistics)	No	\$16,500
12	Public Relations Specialist	No	\$21,325
13	Residential Counselor	No	\$19,680
14	Social Worker	No	\$23,147

Source: *Chronicle Guidance Publications*, The Bureau of Labor Statistics and the Nation Association of Colleges and Employers. Salary are from 1996 to 2003. Pamela Paxton, "Dollars and Sense," *Teaching Sociology*, Vol. 34, January 2006, pp. 65-70.



# How do we interpret this table?

Jobs listed in alphabetical order

Change order: Order by salary

	Job Title	Statistics Needed	Starting Salary
1	Community Organizer	No	\$14,954
2	Corrections Officer	No	\$27,000
3	Criminologist (Counseling)	No	\$23,963
4	Criminologist (Research)	Yes	\$31,500
5	Human Resources	No	\$27,040
6	Marketing Research Analyst	Yes	\$32,000
7	Marketing Research Interviewer	No	\$12,000
8	Private Detective	No	\$26,750
9	Probation Officer	No	\$23,594
10	Public Opinion Researcher (w/Statistics)	Yes	\$27,101
11	Public Opinion Researcher (No Statistics)	No	\$16,500
12	Public Relations Specialist	No	\$21,325
13	Residential Counselor	No	\$19,680
14	Social Worker	No	\$23,147



# Reordered Table

	Job Title	Statistics Needed	Starting Salary
1	Marketing Research Analyst	Yes	\$32,000
2	Criminologist (Research)	Yes	\$31,500
3	Public Opinion Researcher (w/Statistics)	Yes	\$27,101
4	Human Resources	No	\$27,040
5	Corrections Officer	No	\$27,000
6	Private Detective	No	\$26,750
7	Criminologist (Counseling)	No	\$23,963
8	Probation Officer	No	\$23,594
9	Social Worker	No	\$23,147
10	Public Relations Specialist	No	\$21,325
11	Residential Counselor	No	\$19,680
12	Public Opinion Researcher (No Statistics)	No	\$16,500
13	Community Organizer	No	\$14,954
14	Marketing Research Interviewer	No	\$12,000



# Stat Jobs on average pay 41% more than Non-Stat Jobs

Av. Stat job pays \$30,200

Average Non-Stat job pays \$21,450

Difference = \$8,750

	Job Title	Statistics Needed	Starting Salary
1	Marketing Research Analyst	Yes	\$32,000
2	Criminologist (Research)	Yes	\$31,500
3	Public Opinion Researcher (w/Statistics)	Yes	\$27,101
4	Human Resources	No	\$27,040
5	Corrections Officer	No	\$27,000
6	Private Detective	No	\$26,750
7	Criminologist (Counseling)	No	\$23,963
8	Probation Officer	No	\$23,594
9	Social Worker	No	\$23,147
10	Public Relations Specialist	No	\$21,325
11	Residential Counselor	No	\$19,680
12	Public Opinion Researcher (No Statistics)	No	\$16,500
13	Community Organizer	No	\$14,954
14	Marketing Research Interviewer	No	\$12,000



Clear-Sighted Statistics

**Learn Statistics if you want more money**




48



49

Clear-Sighted Statistics

Except where otherwise noted *Clear-Sighted Statistics* is licensed under a Creative Commons License. You are free to share derivatives of this work for non-commercial purposes only. Please attribute this work to Edward Volchok.



50