

STAT 515 — Statistical Methods I

Lecture 1: Describing Data (Part 1)

Yen-Yi Ho

Department of Statistics
University of South Carolina

- What is Statistics
- Types of Data: Continuous, Binary, Categorical, Time-to Event
- Continuous Data
- Continuous Data: Visual Summary Measure

What is Statistics

The study of ...

- ... populations
- ... variation
- ... methods of the reduction of data

“The original meaning of the word ... suggests that it was the study of populations of human beings living in political union.”

Sir R.A. Fisher (the father of modern statistics)

What is Statistics

“...Statistics Science [is] the particular aspect of human progress which gives the 20th century its special character.... It is to the statistician that the present age turns for what is most essential in all its more important activities.” Sir R. A. Fisher

What is Statistics: Less Complimentary Views

- “Science is difficult. You need mathematics and statistics, which is dull like learning a language. ” Richard Gregory
- “There are three kinds of lies: lies, damned lies, and statistics.” Mark Twain, quoting Disraeli.

What is Statistics

- “Statistics is concerned with **METHODS** for **COLLECTING & DESCRIBING DATA** and then for **ASSESSING STRENGTH OF EVIDENCE IN DATA FOR/AGAINST SCIENTIFIC IDEAS!**”

Scott L. Zeger

Statistics Is the Science of Data...

- **Gathering Data:** Design of Experiments and Surveys, collect data online
- **Describing Data:** Descriptive Statistics
- **Using Data to Make Decisions:** Inference about Populations from Samples
- **The Theory Behind Analyzing Data:** Mathematical Statistics and Theoretical Probability

<https://www.youtube.com/watch?v=Ge9je05uYJ0>

FEATURES



Weighing the hype What poisonings data reveals about the weight loss drug boom

As new GLP-1 weight-loss injections like Ozempic, Wegovy and Mounjaro reshape medicine and culture, **Jordan G. Miller, Robert S. Miller, Shawn M. Varney** and **David Han** demonstrate how US poison centres are quietly documenting the growing pains of a pharmaceutical revolution

In 2021, the US Food and Drug

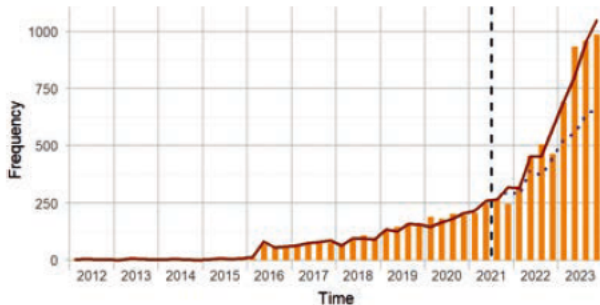
and #Wegovy racked up hundreds of millions

public and health-care providers responded to

Downloaded from academic.oup.com/jasip/article/23/2/10/8442597 by guest on 17 August 2026

Data Is Everywhere!

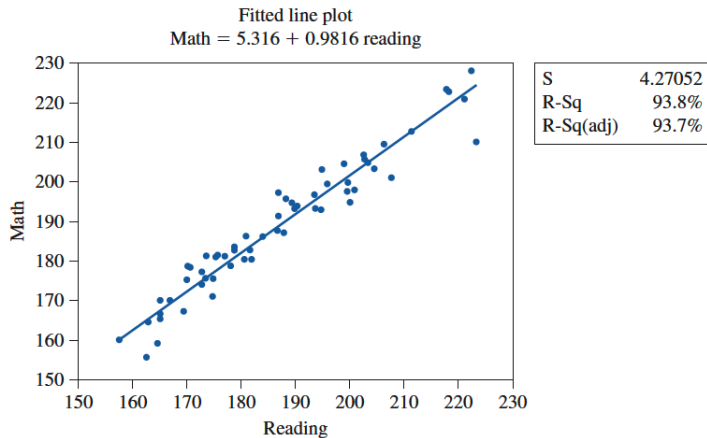
- Data is utilized and summarized frequently in magazine articles.
- Emergency physicians report a steady trickle of GLP-1-related visits: dehydration, abdominal pain, or anxiety over symptoms. In 2022 alone, poison-centre data indicated nearly a 100% increase in hospital evaluations tied to GLP-1 exposures compared with 2021.



Data Is Everywhere!

- Data is utilized and summarized frequently in popular media.
- From CNN.com, Monday July 8th, 2008:
“For the first time, an influential doctors group is recommending that some children as young as eight be given cholesterol-fighting drugs to ward off future heart problems ... With **one-third** of U.S children overweight and about **17 percent obese**, the new recommendations are important, ’ said Dr. Jennifer Li, a Duke University children’s heart specialist. ”

Data from 22 Medium-Sized Florida Elementary Schools



Data Provides Information

- Good data can be analyzed and summarized to provide useful information.
- Bad data can be analyzed and summarized to provide incorrect/harmful/non-informative information.

Steps in a Research Project

- Planning/design of study
- Data Collection
- Data Analysis
- Presentation
- Interpretation

Statistics CAN play a role in each of these steps! (but sometime is only called upon for the data analysis part)

To Do...

- Chapter 1 and Chapter 2 is either for self-study, or might have been seen in your previous course.
- The web page <http://people.stat.sc.edu/habing/515> contains four self-tests on the materials (with answers) and lecture slides.
- Another resource for studying these two chapters:
 - http://people.stat.sc.edu/hitchcock/stat515online/stat515_intro_video1.mp4
 - http://people.stat.sc.edu/hitchcock/stat515online/stat515_intro_video2.mp4
 - http://people.stat.sc.edu/hitchcock/stat515online/stat515_intro_video3.mp4
 - http://people.stat.sc.edu/hitchcock/stat515online/stat515_intro_video4.mp4
- The notes used in the videos are at:
 - http://people.stat.sc.edu/hitchcock/stat515prelimChap1_2notes.pdf
 - http://people.stat.sc.edu/hitchcock/stat515prelimChap2_part2notes.pdf

What We'll Learn in STAT 515

- Descriptive statistics: summarizing and visualizing data
- Probability: quantifying uncertainty
- Random variables and distributions
- Statistical inference: estimation and hypothesis testing

By the end, you'll read a claim backed by data — and know whether to believe it.

- What is Statistics
- Type of Data: Continuous, Binary, Categorical, Time-to Event
- Continuous Data
- Continuous Data: Visual Summary Measure

Binary (dichotomous) Data

- Do you have a facebook account: Yes/No
- Gender: Male/Female
- Cure: Yes/No
- Covid: Yes/No

Categorical Data

- Categorical Data (place observations in categories)

Categorical Data

- Categorical Data (place observations in categories)
- Nominal categorical data: no inherent order to categories
 - Eye color
 - Country of birth
 - Movie genre



Categorical Data

- Categorical Data (place observations in categories)
- Nominal categorical data: no inherent order to categories
 - Eye color
 - Country of birth
 - Movie genre



- Ordinal categorical data: order to categories
 - Degree of agreement, five categories from strongly disagree to strongly agree
 - Education Level: High school, bachelor's degree, master's degree, doctorate.
 - Shirt Sizes: Small, medium, large, extra large.
 - Income : brackets: Low income, middle income, high income.

Continuous Data

- Continuous Data (finer measurements)
 - Blood pressure, mmHg
 - Income level: dollars/year (Euro by year, etc.)
 - Age: years (months)
 - Weight: pounds (kilograms, ounces, etc.)



Different Methods for Different Data Types

- To compare the rating of two movies (thumbs up/ thumbs down), you could use
 - Fisher's Exact Test
 - Chi-Square Test
- To compare the blood pressures of students enrolled in two different sections of STAT 515, you could use
 - 2-Sample T-Test
 - Wilcoxon Rank Sum Test