

STAT 100 Lecture 1: Hello, and Welcome to Statistics

Nate Strawn

<http://www.math.umd.edu/~nstrawn/>

Today's Agenda

- 1 Introductions and course syllabus
- 2 Why are you taking this class, and what is Statistics?
 - Sections 1.1-1.7
- 3 Descriptive statistics
 - Sections 2.1-2.3

Introductions

- My name is Nathaniel (Nate) Strawn.
- I was born and raised in Texas.
- I'm getting a Ph.D. in pure math.

Why are you taking this class?

- To exploit the structure of reality?
 - Consequences of causality
 - Totally deep stuff, man
- To understand statistical arguments and evidence?
 - Calorie restriction has a *significant*, positive effect on the longevity of mice.¹
 - A nation's emphasis on religion is *significantly*, positively correlated with it's homicide rate.²
- To fulfill your degree requirements?
 - Where is your next meal coming from?
 - Everyone sells out eventually!

¹R. Weindruch, R. Walford, S. Fligel, and D. Guthrie. *The Retardation of Aging in Mice by Dietary Restriction: Longevity, Cancer, Immunity and Lifetime Energy Intake*. Journal of Nutrition Vol. 116 No. 4, pp. 641-654, 1986

²G. Jensen. *Religious Cosmologies and Homicide Rates among Nations*. Journal of Religion and Society No. 8, pp. 1-14, 2006.

Chapter 1: What is Statistics?

Definition

Statistics is a mathematical science that focuses on methodologies for collecting, analyzing, interpreting, explaining, and presenting *Data*.

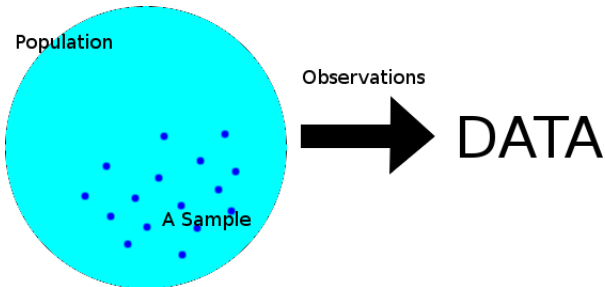
“The best thing about being a statistician is that you get to play in everybody else’s backyard”-John Tukey

“There are three kinds of lies: lies, damned lies, and statistics.”-Benjamin Disraeli

Populations and Samples

Definition

A **Population** is a set of objects from which observations or measurements are drawn. **Data** is a set of observations or measurements. A **Sample** is any subset of the population.

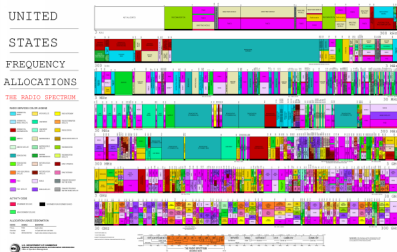


Variables and an Example

Definition

A *Variable* is an observable or measurable quantity whose value may change from one member of a population to another. Note that variables can be *discrete* or *continuous*

Consider the US frequency allocation chart:



What is a population that this information describes? What is a sample? What can be observed or measured? What could be a variable? Is this variable discrete or continuous?

Another Example: NFL Stats 2007

Rank	Team	Games	Pts/Game	Tot Pts	Scrm Plys	Yds/G	Yds/P
1	New England Patriots	16	36.8	589	1058	411.2	6.2
2	Green Bay Packers	16	27.2	435	985	370.7	6
3	Dallas Cowboys	16	28.4	455	975	365.7	6
4	New Orleans Saints	16	23.7	379	1060	361.2	5.5
5	Indianapolis Colts	16	28.1	450	1020	358.7	5.6
6	Philadelphia Eagles	16	21	336	1047	358.1	5.5
7	Jacksonville Jaguars	16	25.7	411	1022	357.4	5.6
8	Cleveland Browns	16	25.1	402	1004	351.3	5.6
9	Seattle Seahawks	16	24.6	393	1056	348.9	5.3
10	Cincinnati Bengals	16	23.8	380	1008	348	5.5
11	Denver Broncos	16	20	320	976	346.3	5.7
12	Arizona Cardinals	16	25.2	404	1016	344.1	5.4
13	Minnesota Vikings	16	22.8	365	964	336.2	5.6
14	Houston Texans	16	23.7	379	968	333.6	5.5
15	Washington Redskins	16	20.9	334	1052	333.4	5.1
16	New York Giants	16	23.3	373	1041	331.4	5.1
17	Pittsburgh Steelers	16	24.6	393	1000	327.4	5.2
18	Tampa Bay Buccaneers	16	20.9	334	975	326.8	5.4
19	Detroit Lions	16	21.6	346	965	322.9	5.4
20	San Diego Chargers	16	25.8	412	980	315.2	5.1
21	Tennessee Titans	16	18.8	301	1037	311.7	4.8
22	Baltimore Ravens	16	17.2	275	1042	302	4.6
23	Atlanta Falcons	16	16.2	259	987	301	4.9
24	St. Louis Rams	16	16.4	263	1026	297.5	4.6
25	Oakland Raiders	16	17.7	283	1000	294.8	4.7
26	New York Jets	16	16.8	268	1011	294.7	4.7
27	Chicago Bears	16	20.9	334	1035	293.2	4.5
28	Miami Dolphins	16	16.7	267	989	287.5	4.7
29	Carolina Panthers	16	16.7	267	989	284.9	4.6
30	Buffalo Bills	16	15.8	252	919	277.1	4.8
31	Kansas City Chiefs	16	14.1	226	1001	276.8	4.4
32	San Francisco 49ers	16	13.7	219	925	237.3	4.1

Two Branches of Statistics

Definition

Inferential Statistics deals with inferences, or implications and generalizations based upon data.

Definition

Descriptive Statistics deals with organizing, displaying, and measuring data.

Chapter 2: Organization and Description of Data

Data needs to be organized so that we can comprehend trends. Since most people are visual learners, graphs are crucial for data presentation.

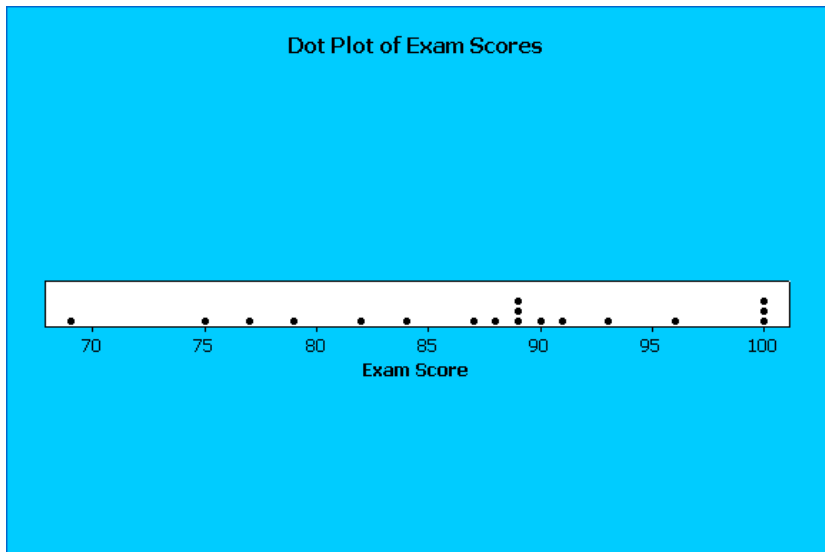
- Frequency distributions
- Stem-and-leaf display
- Dot diagram
- Histogram
- Pie chart

Stem-and-Leaf Display Example

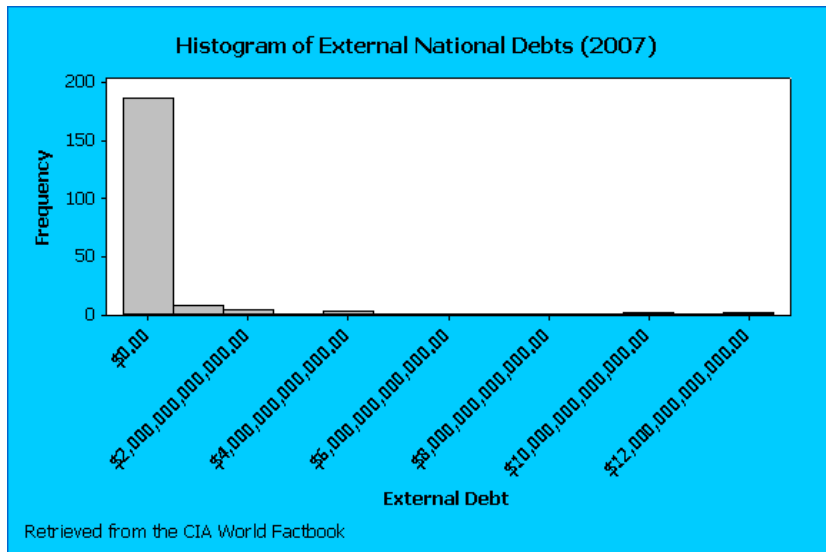
Suppose the exam scores after a statistics exam are 69, 75, 77, 79, 82, 84, 87, 88, 89, 89, 89, 90, 91, 93, 96, 100, 100, and 100. Then you would have a stem-and-leaf display as follows:

Stem	Leaf
6	9
7	579
8	2478999
9	0136
10	000

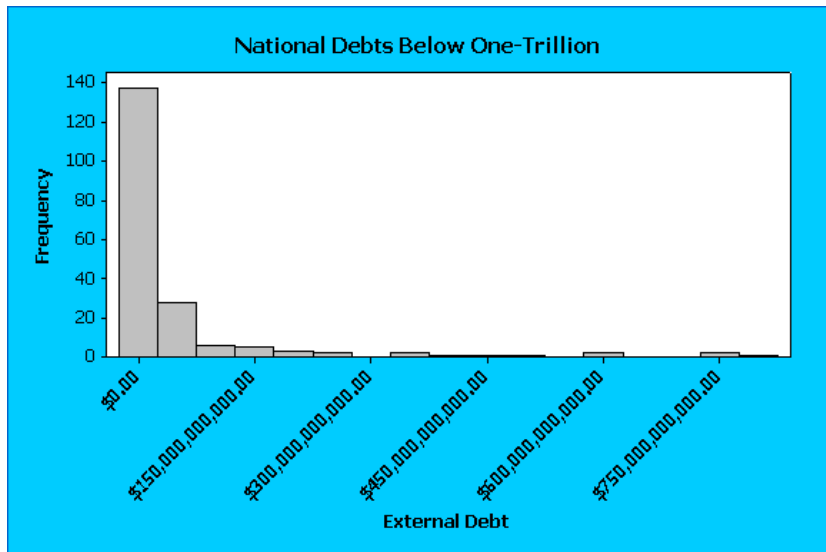
Dot Diagram Example



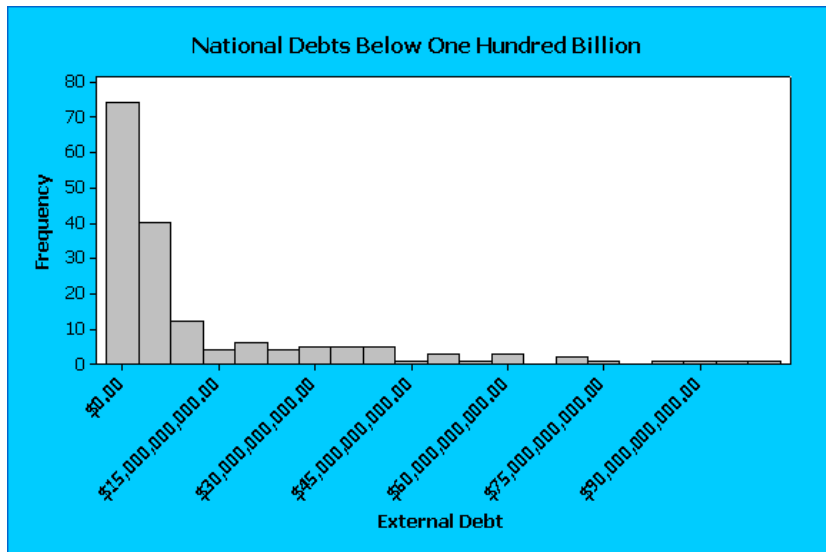
Histogram Example



Histogram Example

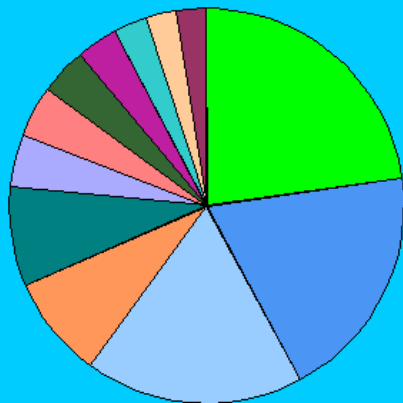


Histogram Example



Pie chart Example

Proportion of Global Debt



Retrieved from the CIA World Factbook

For Next Time

- Read Section 1.1-1.7 and 2.1-2.3 from Johnson and Bhattacharyya
- Online homework 2.3: 2.3, 2.5, 2.15, 2.19, 2.23
 - Section 0101:
<http://edugen.wiley.com/edugen/class/cls73726/>
 - Section 0201:
<http://edugen.wiley.com/edugen/class/cls73727/>