EDP 613 Fall 2020

Chapter 2 Slides

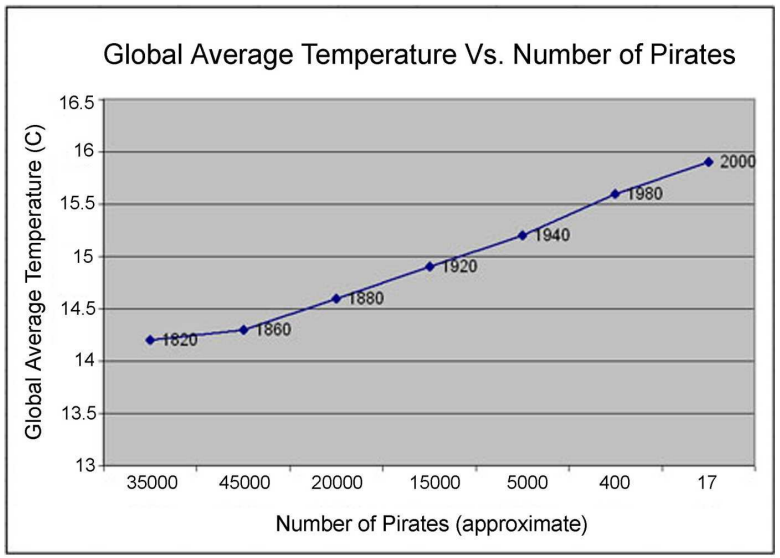
Abhik Roy

Abhik.Roy@mail.wvu.edu

West Virginia University

Always Remember!

STOP GLOBAL WARMING: BECOME A PIRATE



WWW.VENGANZA.ORG

CORRELATION DOES NOT PROVE CAUSATION!

Example

Construct a frequency table for the numbers of newspapers sold at a local shop over the last 10 days:

$$22, 20, 18, 23, 20, 25, 22, 20, 18, 20$$

Example

Construct a relative frequency table for the numbers of newspapers sold at a local shop over the last 10 days:

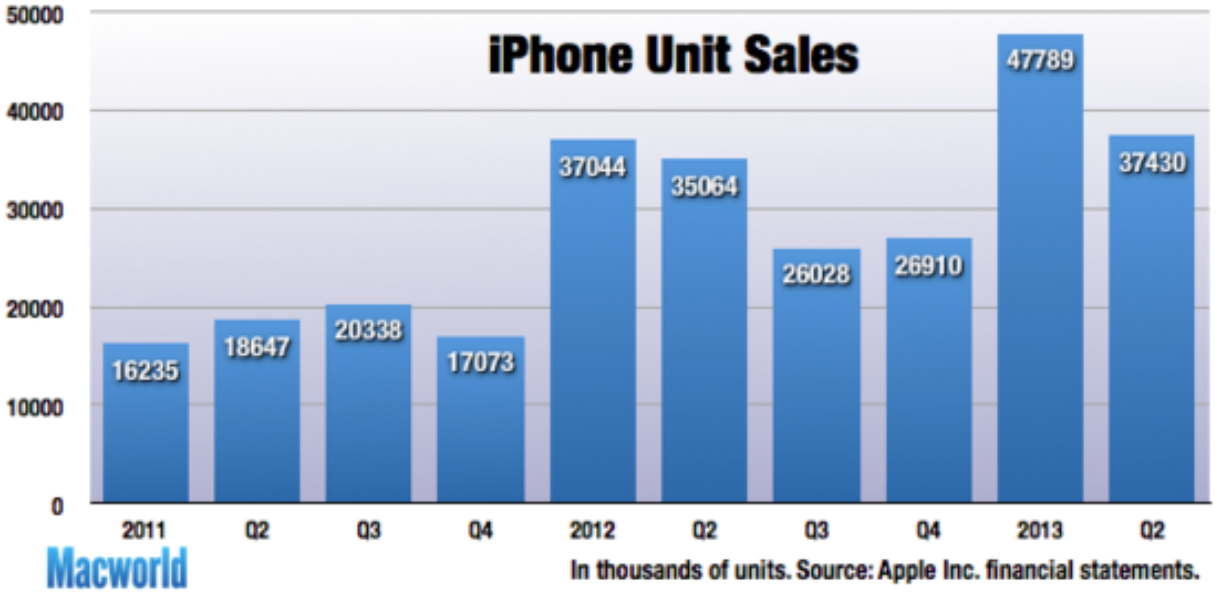
$$22, 20, 18, 23, 20, 25, 22, 20, 18, 20$$

On Your Own

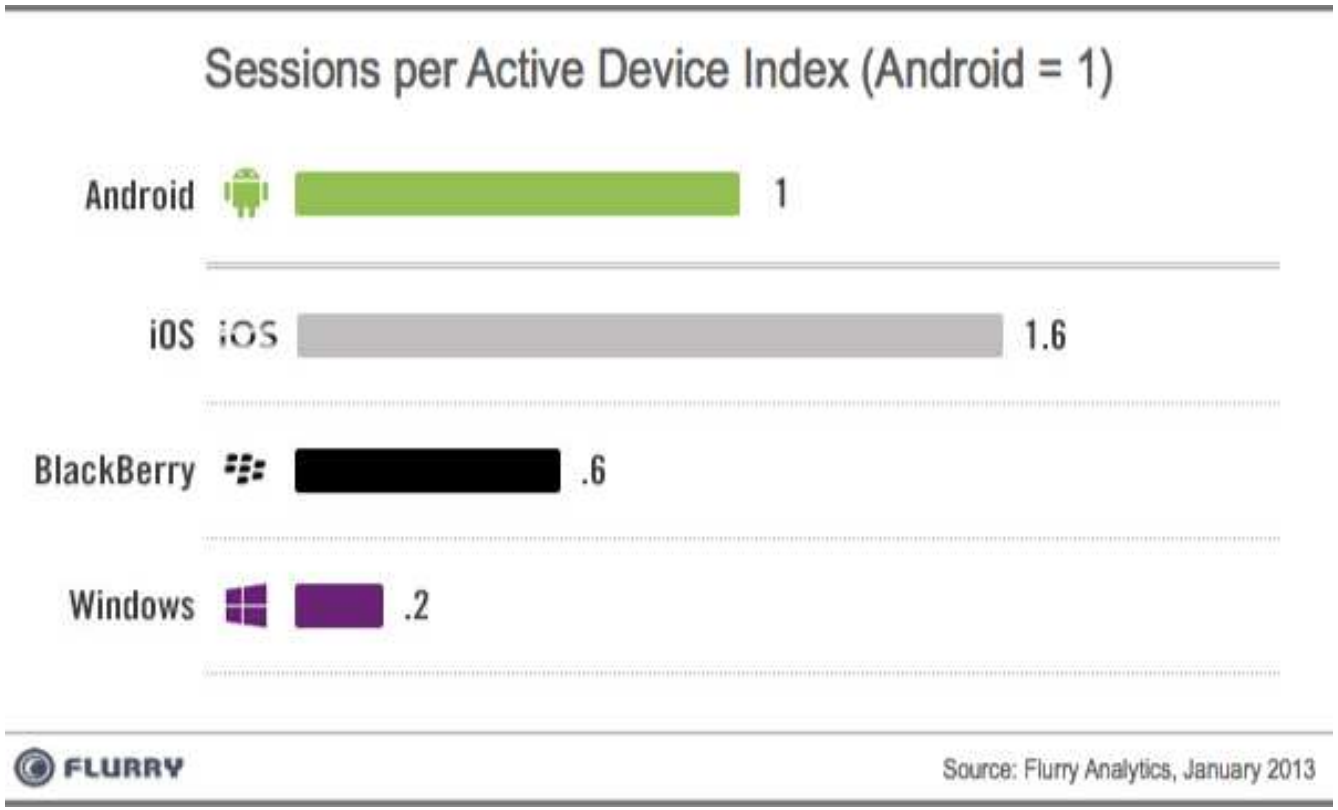
Construct a frequency table for the data below that sampled students and asked the total number of children in their family (including themselves).

2	2	2	4	5	3	3	3	3
2	1	2	3	5	3	4	3	1
2	3	5	3	2	1	3	2	

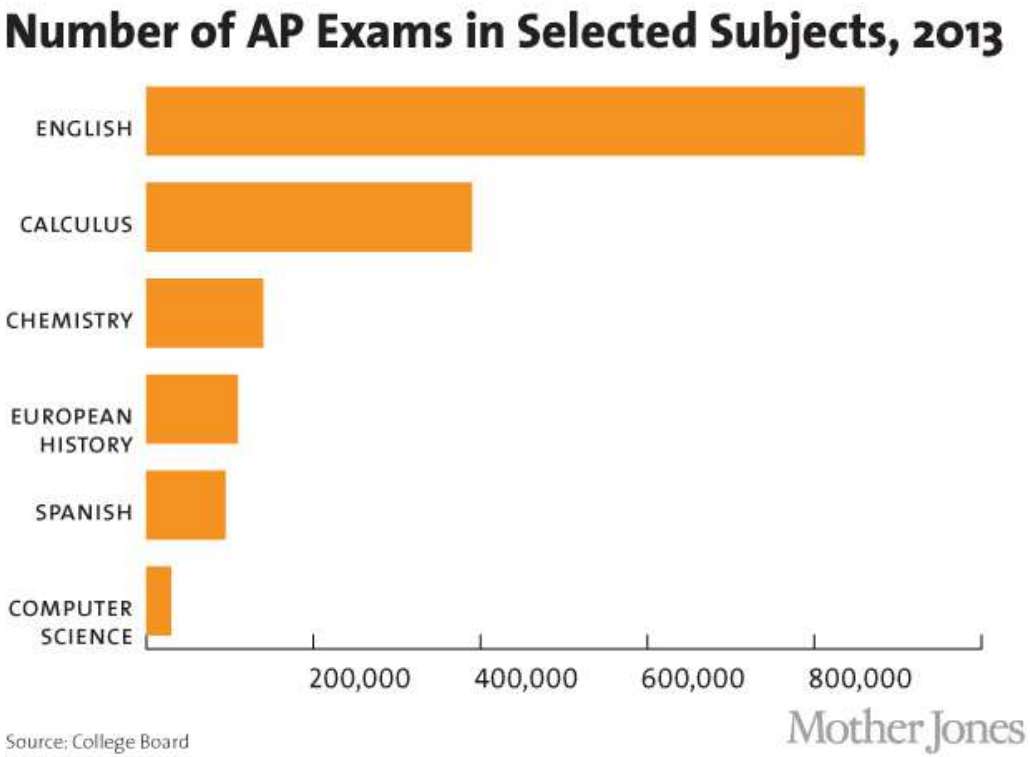
Bar Graph (vertical)



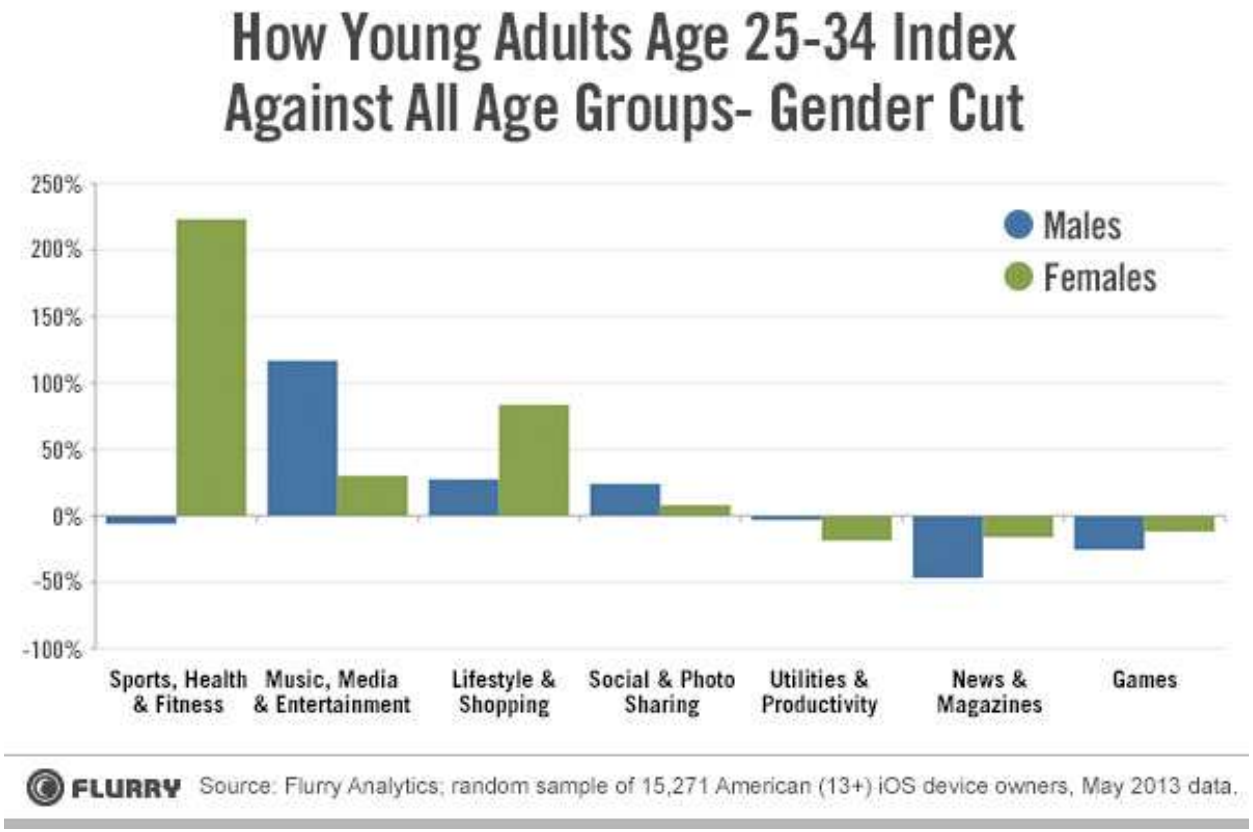
Bar Graph (horizontal)



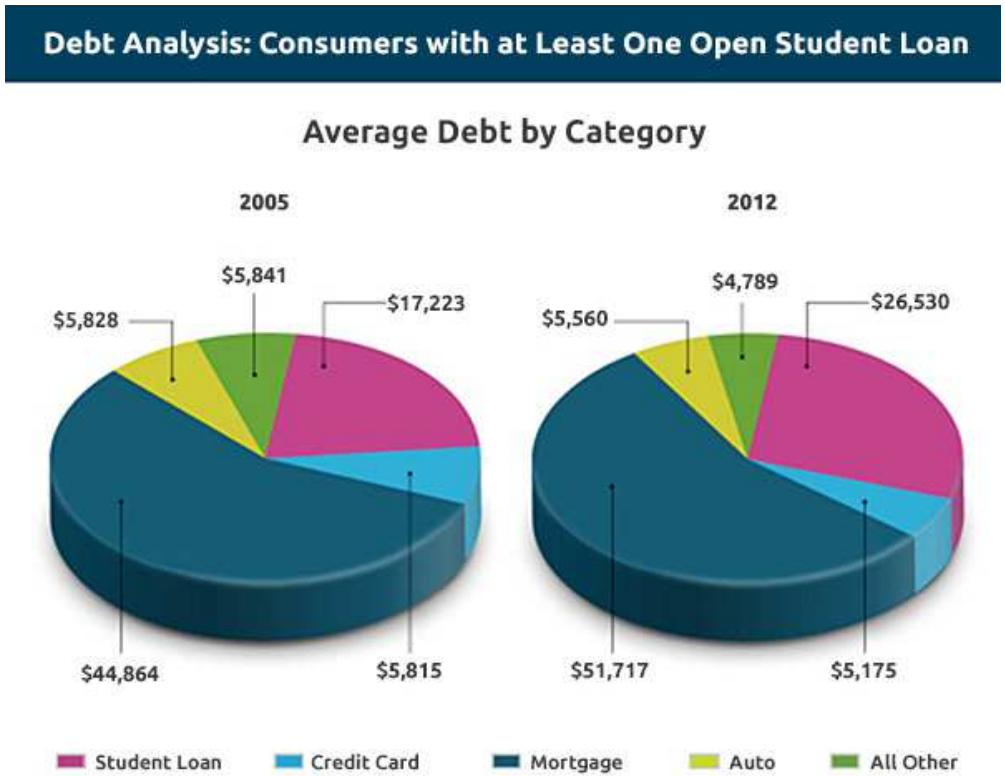
Pareto chart



Side-by-side Bar Graph



Pie Chart


Debt Analysis: Consumers with at Least One Open Student Loan
Average Debt by Category
2005
2012
$5,828
$5,841
$17,223
$44,864
$5,815
$5,560
$4,789
$26,530
$51,717
$5,175
Student Loan
Credit Card
Mortgage
Auto
All Other
Source: FICO™ Banking Analytics Blog. © 2012 Fair Isaac Corporation

Example

Make a frequency distribution for the following data, using 5 classes:

5	10	7	19	25	12	15	7	6	8
17	17	22	21	7	7	24	5	6	5

On Your Own

The following table presents the purchase totals (in dollars) of a random sample of gasoline purchases at a convenience store. Construct a frequency distribution using a class width of 10, and using 0 as the lower class limit for the first class.

76.59	48.55	93.66	60.17	39.10
93.28	65.43	34.12	80.41	77.16
80.07	93.46	39.19	43.84	44.70
68.74	89.98	6.97	52.86	68.93

Example

Construct a frequency histogram using the following data using only 5 classes:

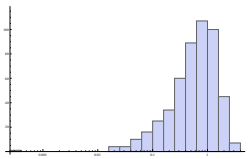
5	10	7	19	25	12	15	7	6	8
17	17	22	21	7	7	24	5	6	5

On Your Own

The following table presents the purchase totals (in dollars) of a random sample of gasoline purchases at a convenience store. Construct a frequency histogram using a class width of 10, and using 0 as the lower class limit for the first class.

76.59	48.55	93.66	60.17	39.10
93.28	65.43	34.12	80.41	77.16
80.07	93.46	39.19	43.84	44.70
68.74	89.98	6.97	52.86	68.93

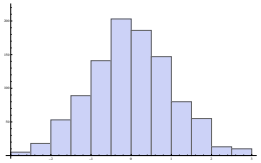
Shapes of Histograms



Skewed left

Negatively skewed

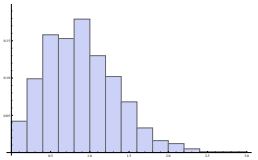
Long left tail



Approximately symmetric

Normal

Basically symmetric



Skewed right

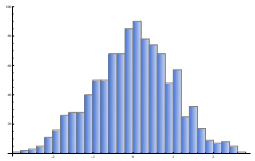
Positively skewed

Long right tail

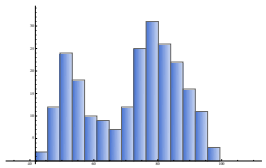
Modes

A histogram is

- *unimodal* if it has one mode.



- *bimodal* if it has two clearly distinct modes.



Example

Construct a frequency and a cumulative frequency plot in the following data set:

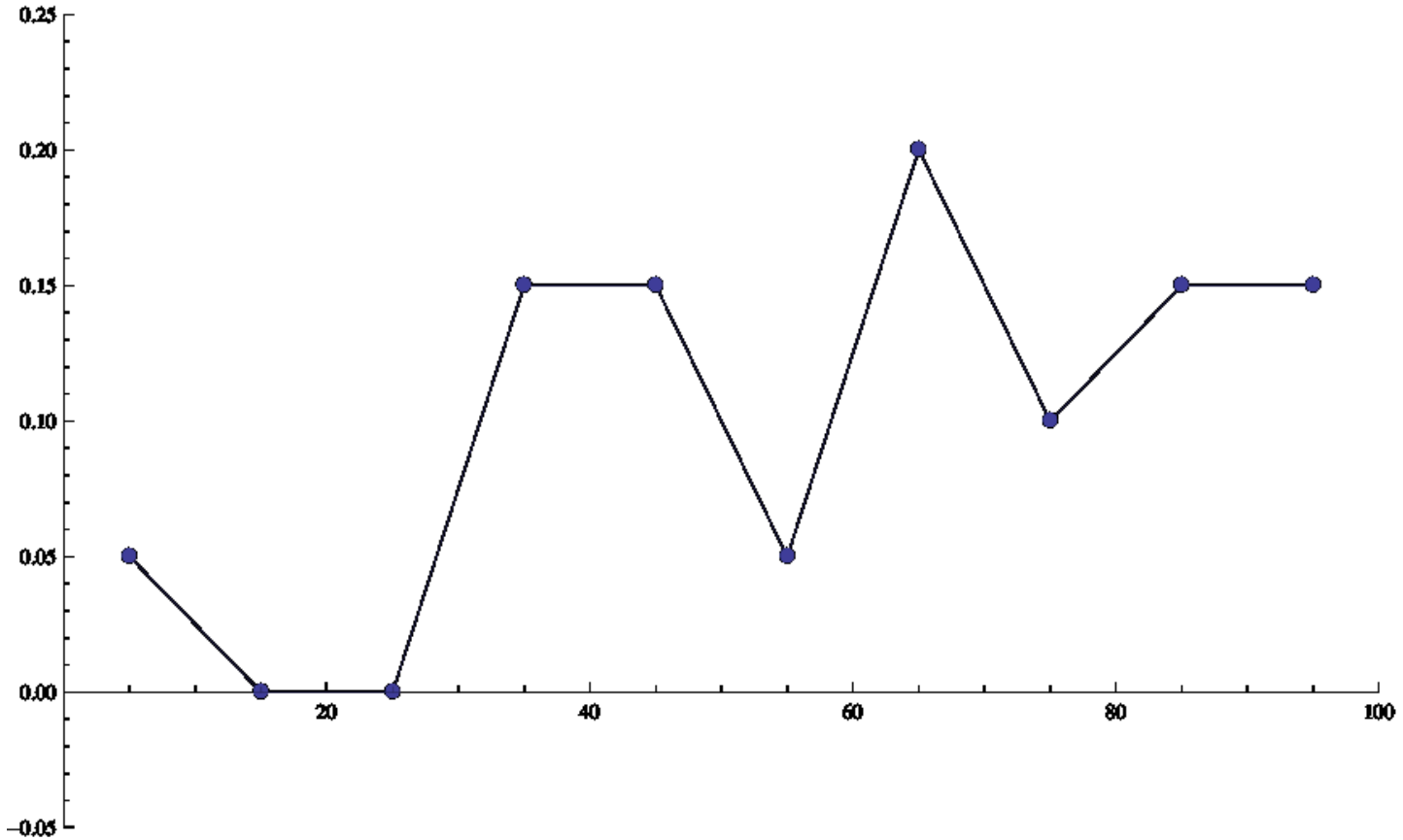
5	10	7	19	25	12	15	7	6	8
17	17	22	21	7	7	24	5	6	5

On Your Own

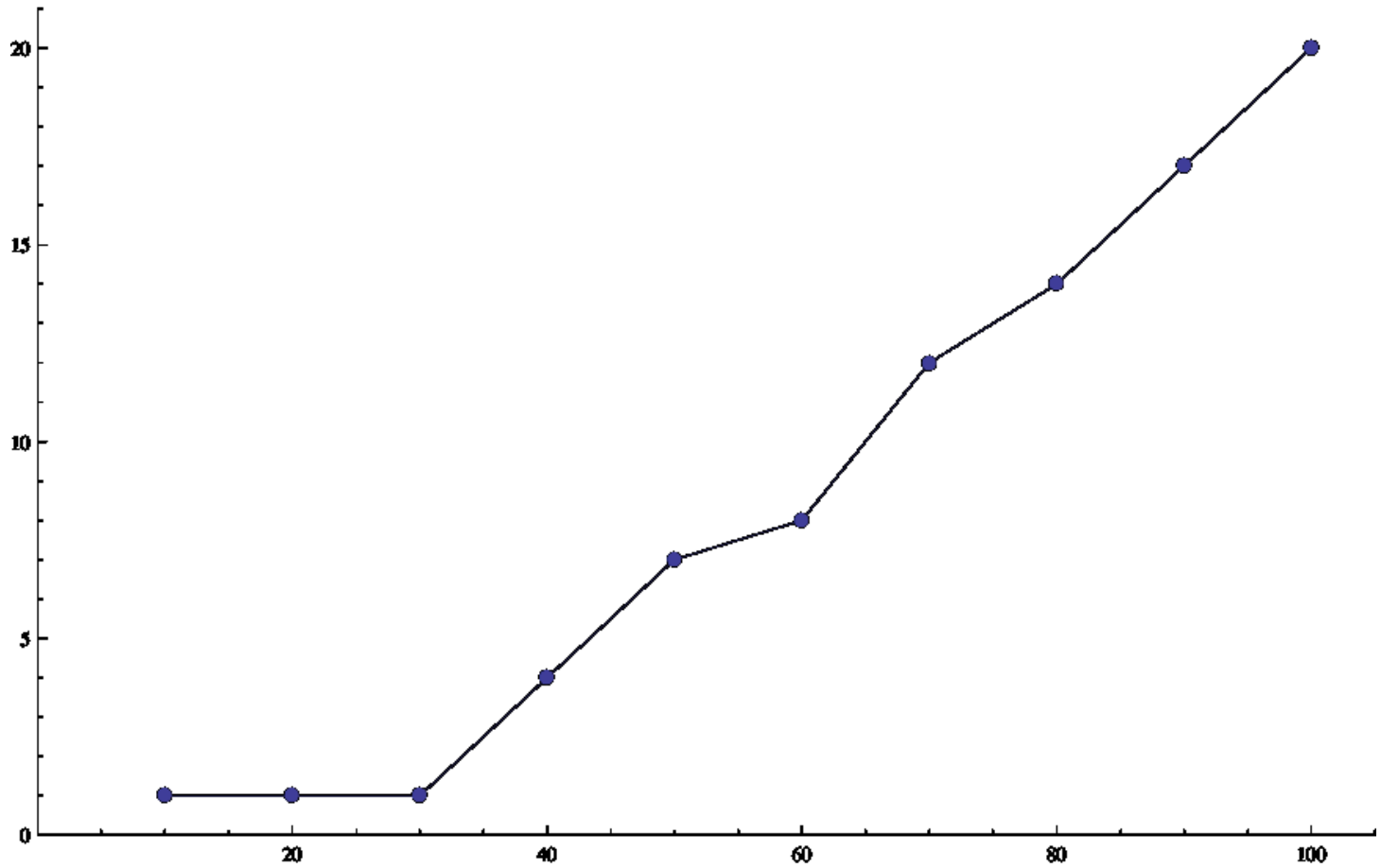
The following table presents the purchase totals (in dollars) of a random sample of gasoline purchases at a convenience store. Construct a frequency plot and a cumulative frequency plot for the following data set: using a class width of 10, and using 0 as the lower class limit for the first class.

76.59	48.55	93.66	60.17	39.10
93.28	65.43	34.12	80.41	77.16
80.07	93.46	39.19	43.84	44.70
68.74	89.98	6.97	52.86	68.93

Frequency Plot



Cumulative Frequency Plot



Example

This data set gives number of staff in an organization. Draw a time-series plot of the data.

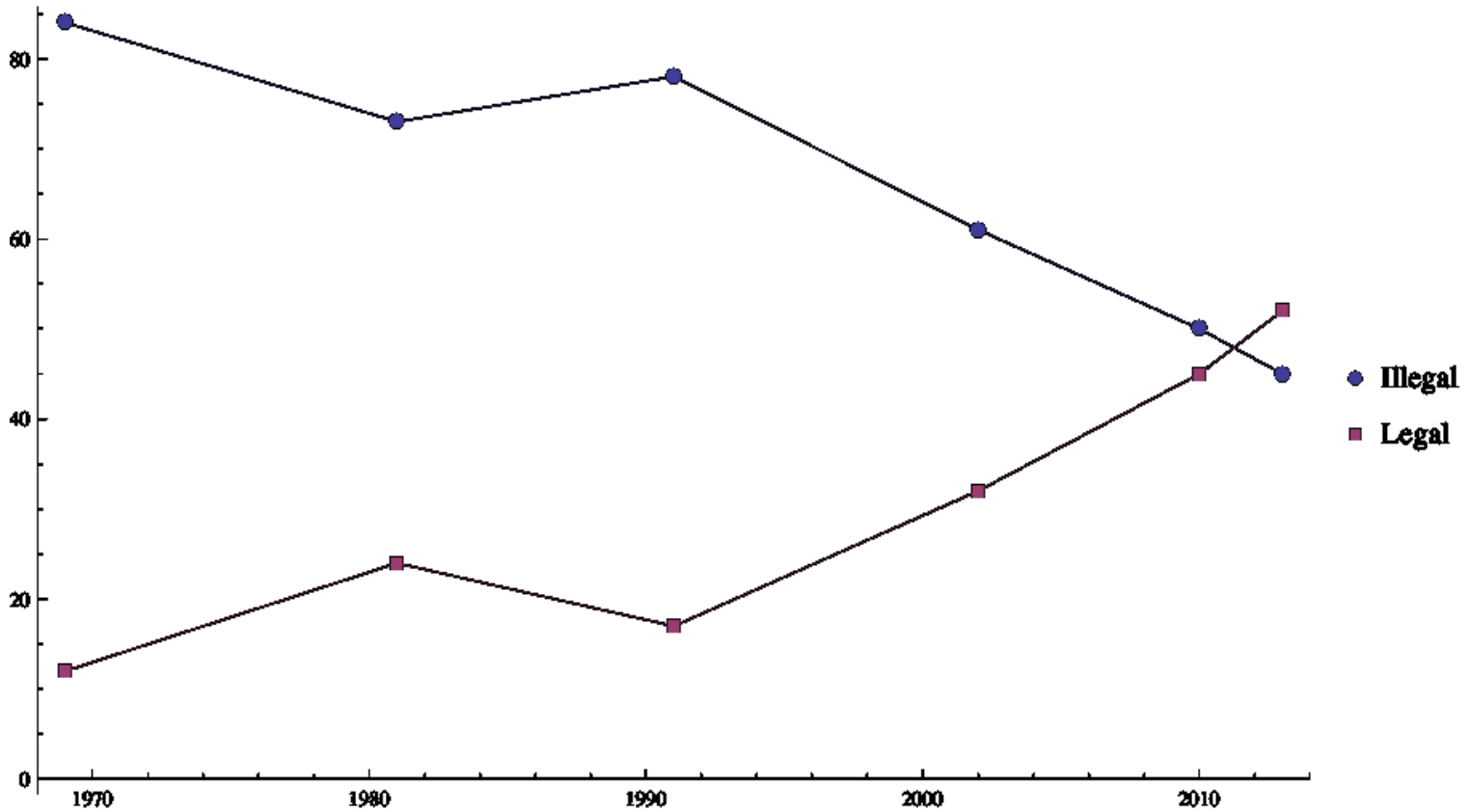
Year	1	2	3	4	5	6
Total	10	25	40	55	60	65

On Your Own

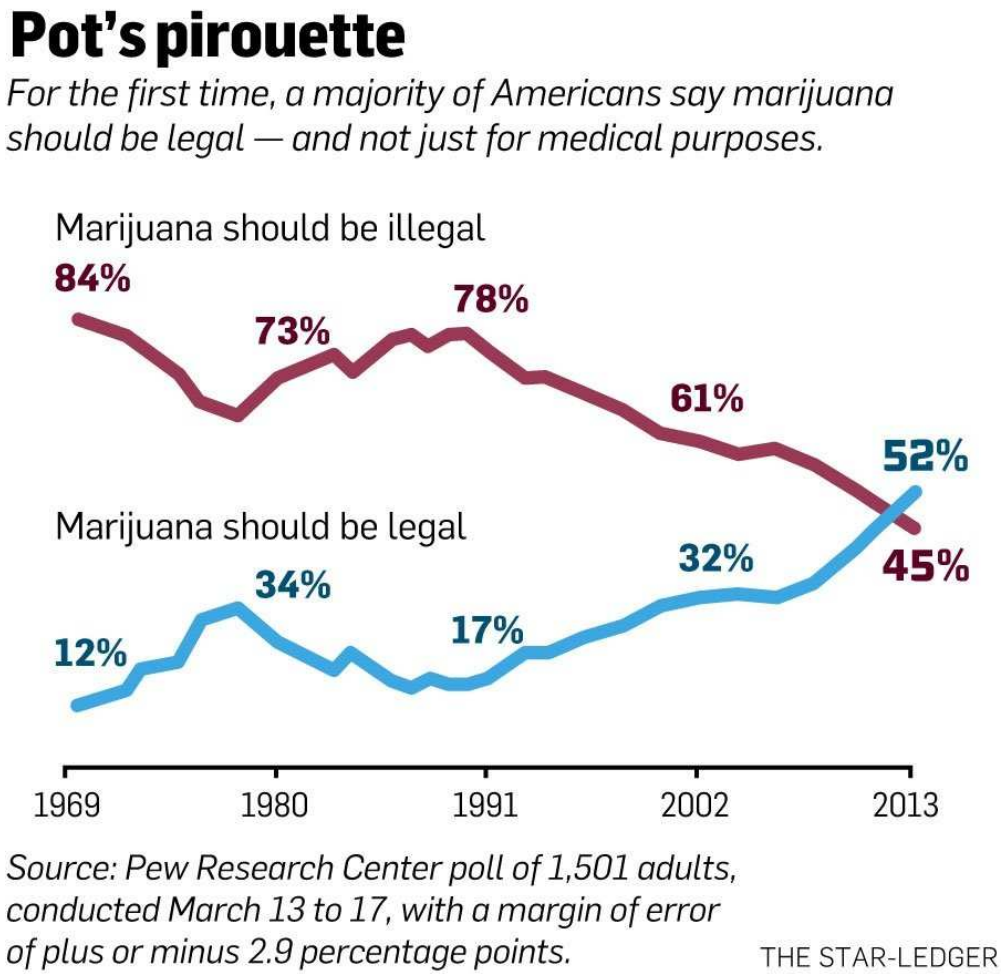
This data set provides percentage of the US populous and their views on the legalization of marijuana between the years of 1969-2013. Draw a time-series plot of the data.

Year	1969	1981	1991	2002	2010	2013
Illegal	84	73	78	61	50	45
Legal	12	24	17	32	45	52

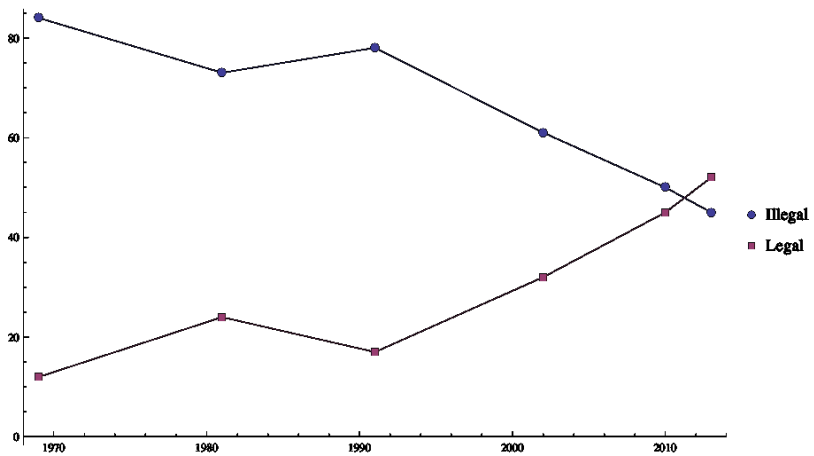
Actual Time Series Plot

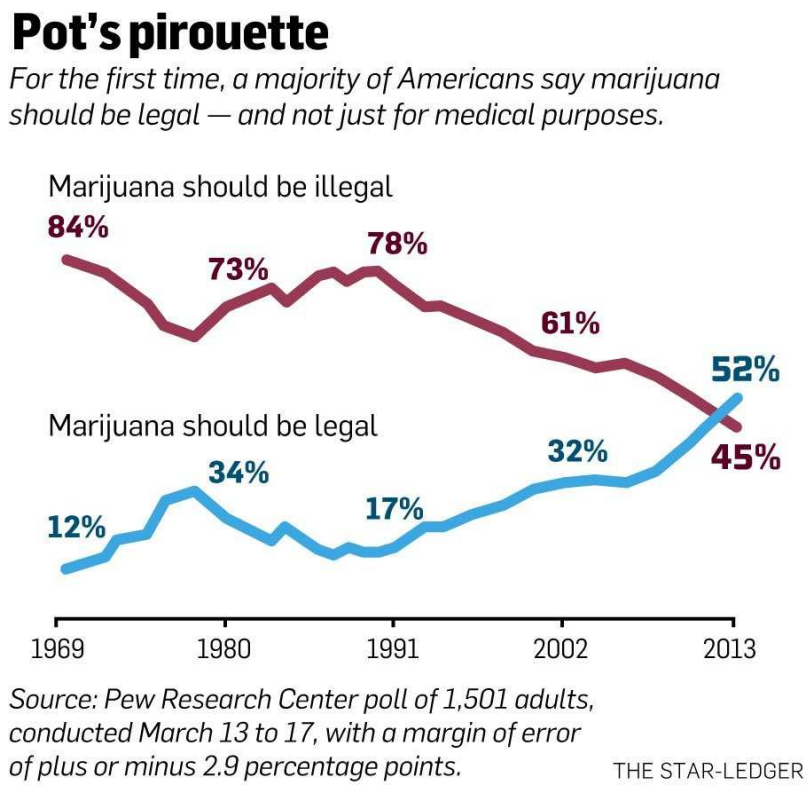


Actual Time Series Plot



Comparison


80
60
40
20
0
1970
1980
1990
2000
2010
Illegal
Legal


Pot's pirouette
For the first time, a majority of Americans say marijuana should be legal — and not just for medical purposes.
Marijuana should be illegal
84%
73%
78%
61%
52%
Marijuana should be legal
12%
34%
17%
32%
45%
1969
1980
1991
2002
2013
Source: Pew Research Center poll of 1,501 adults, conducted March 13 to 17, with a margin of error of plus or minus 2.9 percentage points.
THE STAR-LEDGER

Section 2.4: Graphs Can Be Misleading

That's it for the problem session! Let's take a 10 minute break!