MATH E-3: Lecture 4

Quantitative Reasoning: Practical Math

Announcements

February 16, 2016

https://upload.wikimedia.org/wikipedia/commons/a/ac/Widener_library_2009.JPG

Homework

- Assignment 1 with grader comments available
- Graded Assignment 2 will be available by 2/20; compare your graded assignment with posted solutions
- Assignment 3 is due 2/20 by the upload deadline
- Assignment 4 will be posted tomorrow, 2/17

Homework Assignment Submission
"Homework Help Center"

CHECKLIST

- ☑ single PDF file
- ☑ 4mb maximum
- ☑ file name (example: albrigo.assign1)
- ☑ work must be neat and legible (flag your final answer) and scanned to the appropriate drop box (pages in order and no upside down or sideways pages)
- ☑ review your assignment after submitting
- ☑ upload by Saturday, 11:59 a.m. (ET) deadline

Review carefully the Homework Policies section outlined in the syllabus.


QUESTIONS

Australia's population hits 24 million people, ABS clock shows

The number of people in Australia has surpassed 24 million for the first time, according to the Bureau of Statistics' (ABS) population clock.

When did the population reach 24 million? According to the ABS population clock, it reached 24 million about 12:50am (AEDT) on February 16, 2016.

Who was the 24 millionth person? There is no official 24 millionth person. It could be a newborn baby or a person moving to Australia for work.

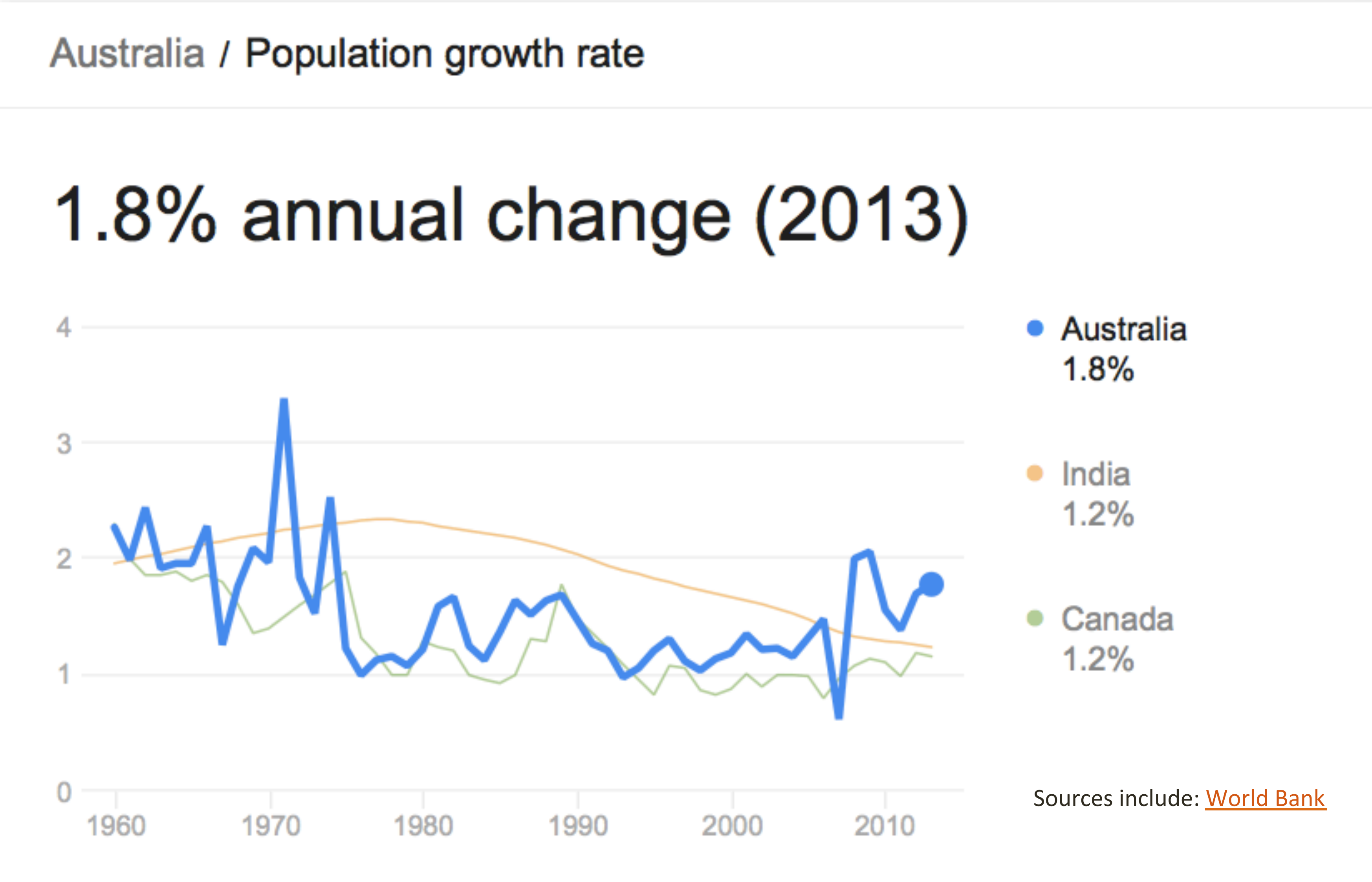
How long has it been since we reached 23 million? It has been about three years since Australia reached 23 million (in the March quarter of 2013).

When will the population reach 25 million? It is projected that Australia will reach 25 million in 2018 and will keep increasing by a million persons every two to three years.

And what happens beyond 25 million? According to the ABS, it will take Australia 66 years to double its population from 24 million to 48 million. Australia is projected to reach 50 million in 2089.

The population clock is an indication of the current population based on a projection calculated using births and deaths data from the ABS, and migration figures from the Department of Immigration and Border Protection.

And population growth rates . . .



Sources include: [World Bank]

More data . . .

	1901 (Federation)	**1968** (when Australia was approximately half today's population)	**2015**
Total Australian population	3,788,123	12,008,635	23,781,169
Population growth rate	1.5%	1.8%	1.4%
Number of births	102,945	240,906	303,965
Total Fertility Rate (TFR) *(Births per woman)*	Not available	2.9	1.8
Sex ratio *(Males per 100 females)*	110.1	101.3	99
Life expectancy *(years)*	55.2 (males), 58.8 (females) *(1901-1910)*	67.6 (males), 74.2 (females), *(1965-1967)*	80.3 (males), 84.4 (females), *(2012-2014)*
Median age *(The age at which half the population is older and half is younger)*	22.5	27.8	37.4
Born in Australia	2,908,303	9,419,542	16,890,250
Born in Australia (%)	76.8%	81.2%	72.0%
Born in UK *(includes Ireland in 1901)*	679,159	870,548	1,221,260
Born in UK *(includes Ireland in 1901)* (%)	17.9%	7.5%	5.2%

Some reactions . . .

Migration pushing Australia to 'third-world style population growth rate'

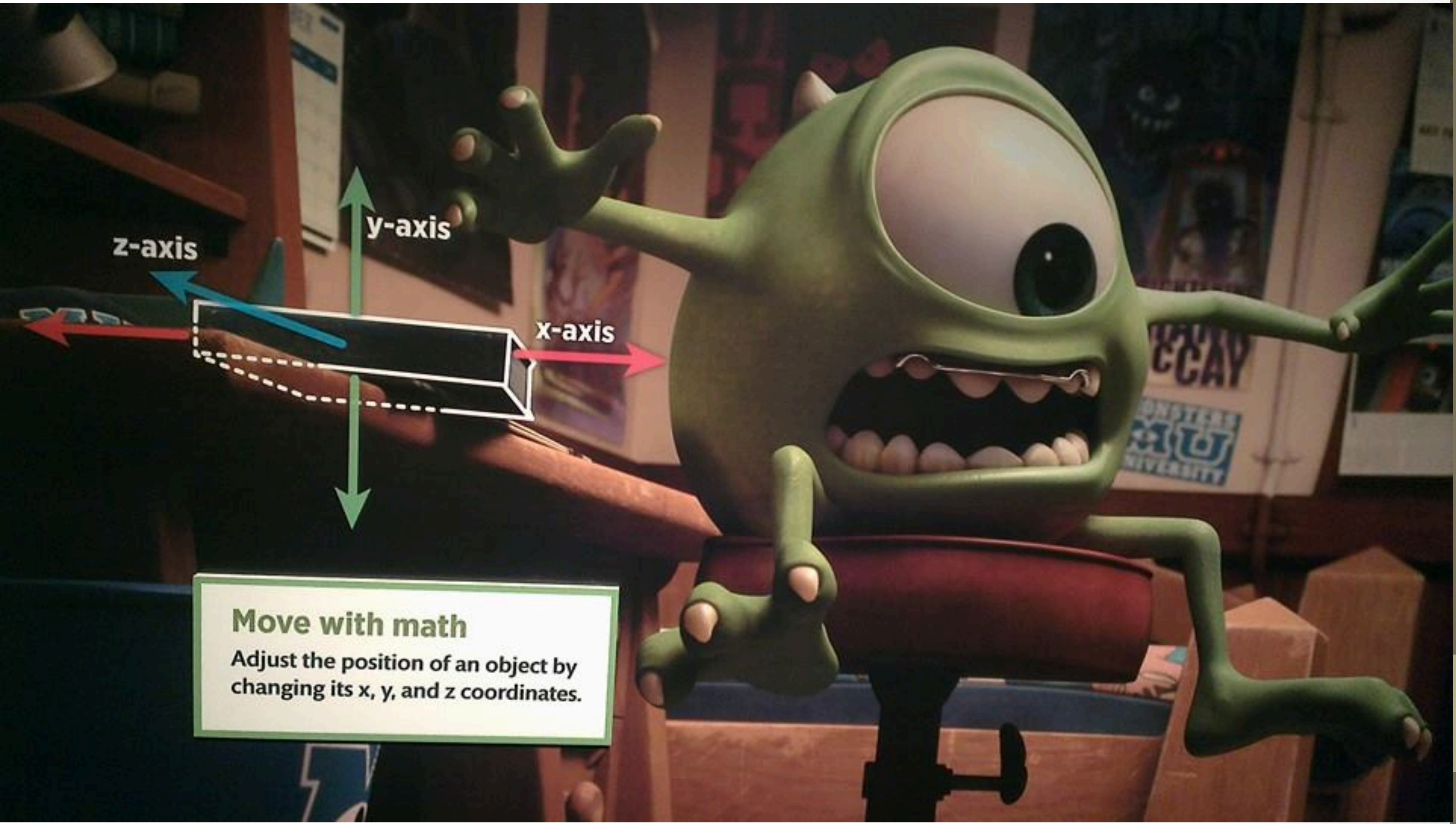
http://www.theguardian.com/australia-news/2016/feb/16/migration-pushing-australia-to-third-world-style-population-growth-rate

Former Labor foreign minister and NSW premier Bob Carr says Australia's 'hugely over-ambitious' immigration rate must be cut by up to 50 per cent.

http://www.sbs.com.au/news/article/2016/02/16/bob-carr-urges-50-cut-immigration?cid=cxenseab_b&cx_navSource=related-side-cx#cxrecs_s

On a lighter note . . . Math and Pixar

(Boston Museum of Science)



3-d geometry: cylinder, sphere and cone

Percents in the news . . .

CEO DEFENDS 5,000% DRUG PRICE HIKE

A drug commonly used to treat life-threatening parasitic infections has jumped in price to $750 per tablet, from $13.50, after being acquired by a pharmaceutical startup run by a former hedge-fund manager. The 62-year-old drug, Daraprim, was acquired by Turing Pharmaceuticals this year. Turing's CEO Martin Shkreli defended the price increase in an interview with Bloomberg Monday. “At the end of the day, the price per course of treatment — to save your life — was only $1,000," Shkreli said. Daraprim is still underpriced relative to its peers." He also explained that the company simply needed to make a profit on the drug. Biotech stocks plummeted Monday after Hillary Clinton announced that she would release a plan to combat high prescription-drug costs in response to the Turing incident.

READ IT AT THE NEW YORK TIMES >

emails@thedailybeast.com

Let's check the math . . .

% change = (new – old)/old

Let’s check the math . . .

% change = (new – old)/old

% change = (750 – 13.50)/13.50

Let's check the math . . .

% change = (new – old)/old

% change = (750 – 13.50)/13.50

% change = 736.50/13.50

Let's check the math . . .

% change = (new – old)/old

% change = (750 – 13.50)/13.50

% change = 736.50/13.50

% change = 54.55555 . . .

Let’s check the math . . .

% change = (new – old)/old

% change = (750 – 13.50)/13.50

% change = 736.50/13.50

% change = 54.55555 . . .

% change = 5,455.56%

So the story understated it a bit . . .

Presenting and Analyzing Data

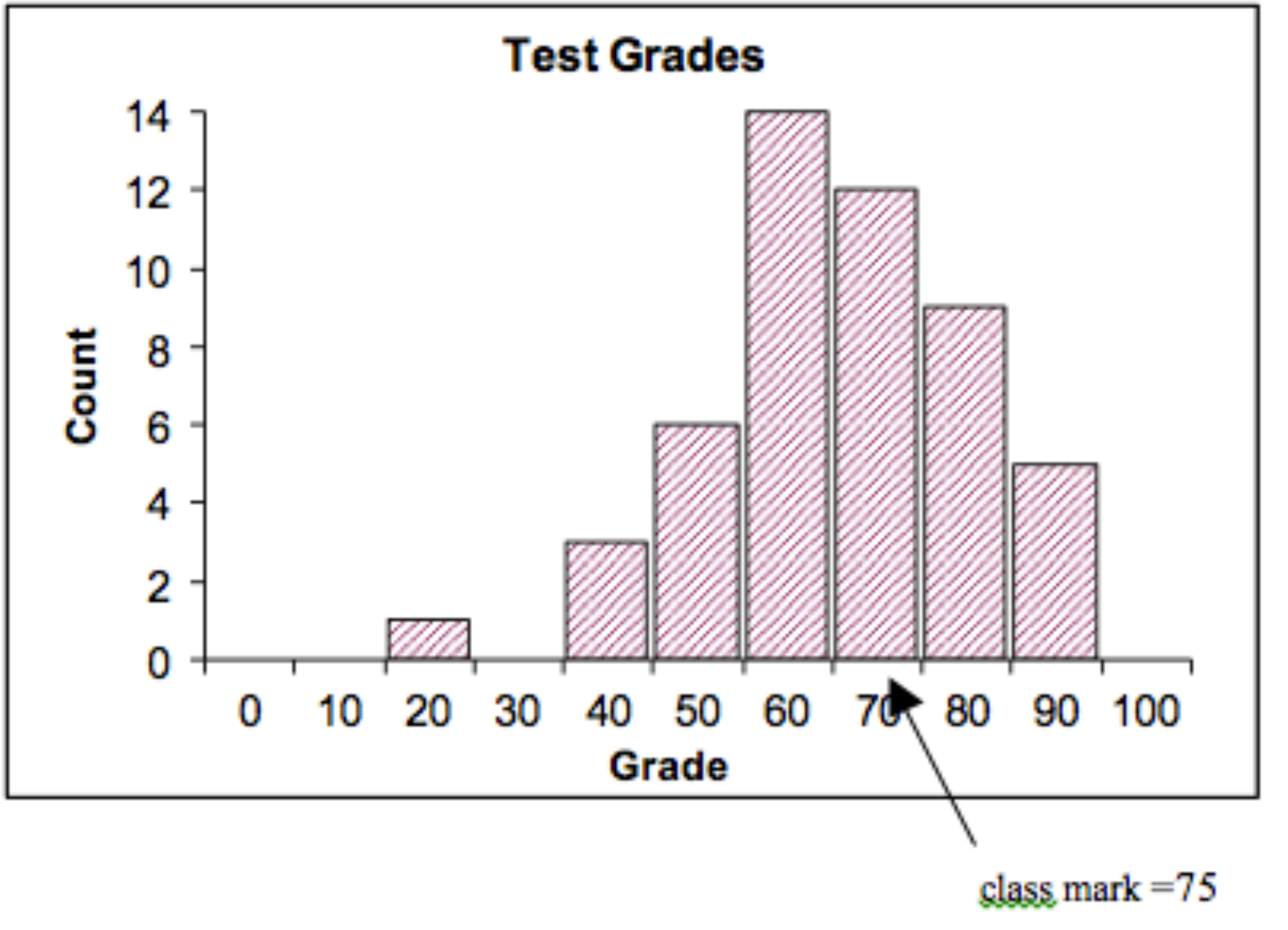
- Histograms
- Stem and Leaf Displays
- Shapes of Histograms
- Measure of Location
- Measures of Dispersion

Grades received in class:
(from lecture 1)

73	47	71	65	85	61	74	80	65	67
62	84	71	99	70	88	69	83	81	71
80	68	26	58	71	90	58	64	70	95
91	67	48	41	62	75	66	77	78	66
50	52	83	58	50	87	91	60	78	67

* Information and Graphs on the next few slides are taken from the Core Curriculum Quantitative Reasoning Requirement Data Text, © President and Fellows of Harvard College, 1991.

Histogram


Test Grades
14
12
10
8
6
4
2
0
Count
0
10
20
30
40
50
60
70
80
90
100
Grade
class mark =75

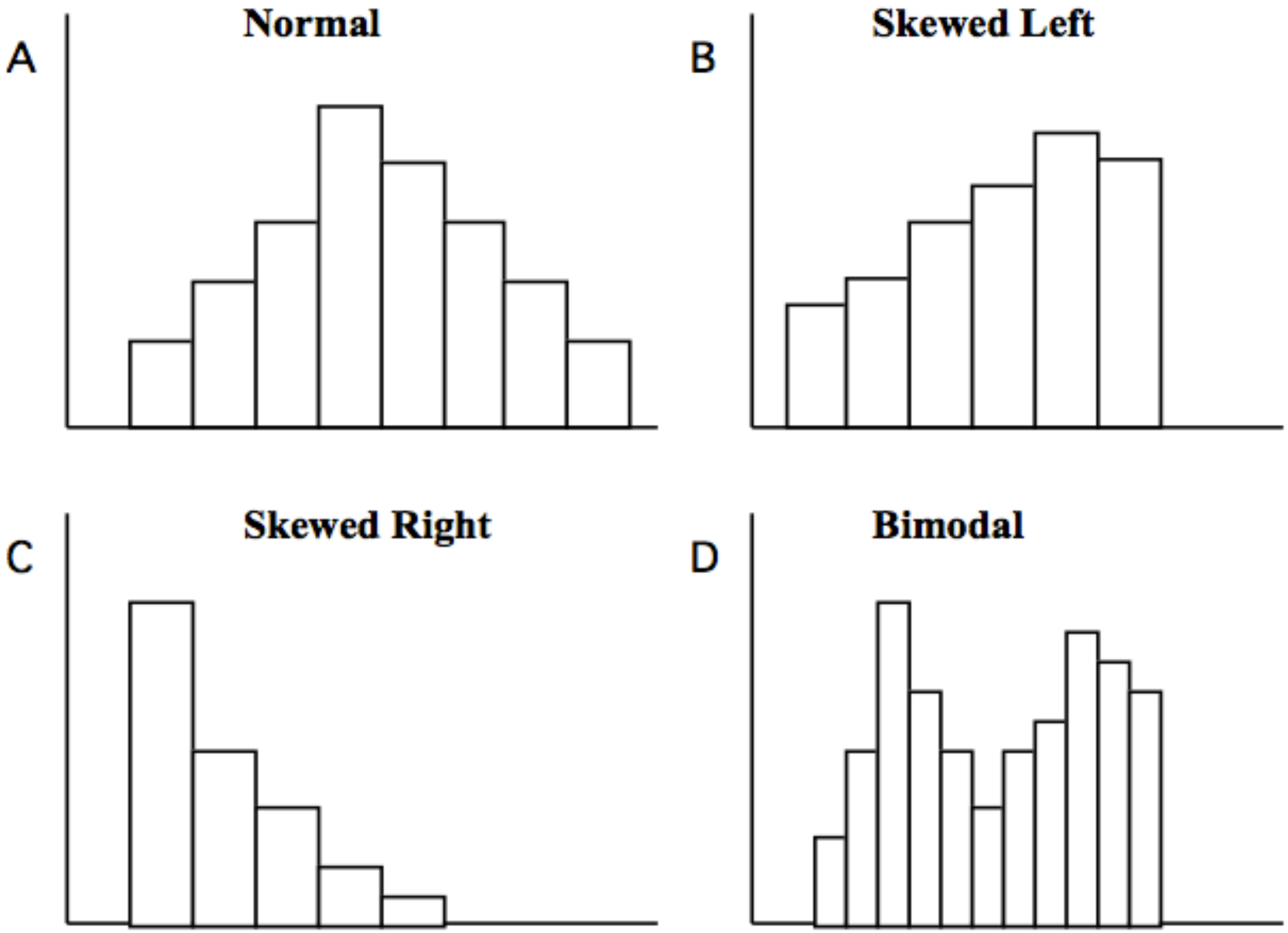
Stem and Leaf Display, part 1

Stem	**Leaf** (1st Pass)
2	6
3	--
4	7 8 1
5	0 2 8 8 0 8
6	2 8 7 5 2 1 9 6 4 0 5 7 6 7
7	3 1 1 0 1 5 4 7 0 8 8 1
8	0 4 3 5 8 7 0 3 1
9	1 9 0 1 5

Stem and Leaf Display, part 2

Count	**Stem**	**Leaf** (2nd Pass – numerical Order)
1	2	6
0	3	--
3	4	1 7 8
6	5	0 0 2 8 8 8
14	6	0 1 2 2 4 5 5 6 6 7 7 7 8 9
12	7	0 0 1 1 1 1 3 4 5 7 8 8
9	8	0 0 1 3 3 4 5 7 8
5	9	0 1 1 5 9
Total 50		

Shapes of Histograms



Measures of Location #1: the Mean

Mean – The mean is the ***Average*** of all data values.

i.e. add data and divide by how many there are.

$$\textbf{Mean} = \frac{\text{Sum of all data values}}{\text{total number of data}} = \frac{\sum x}{n} = \bar{x}$$

We denote 'how many' or the frequency as 'n'

We denote the mean by an x with a bar, $\bar{\mathbf{x}}$

Example of finding the mean

Example 1: A class of 15 students in English 101 received the following grades on a test. Find the class average or **mean**.

100, 60, 90, 90, 70, 100, 100, 90, 90, 80, 70, 100, 80, 90, 80

$$\bar{x} = \frac{\sum x}{n} = \frac{100+60+90+90+70+100+100+90+90+80+70+100+80+90+80}{15}$$

$$= \frac{1290}{15} = \mathbf{86}$$

$\bar{x} = \mathbf{86}$ Note: $\mathbf{n = 15}$

Using a frequency table

Table 1: *A Frequency Table for English 101*

Data Value	How Many
60	1
70	2
80	3
90	5
100	4

Finding the mean from the freq. table

$$\textbf{Mean} = \frac{60(1)+70(2)+80(3)+90(5)+100(4)}{15}$$

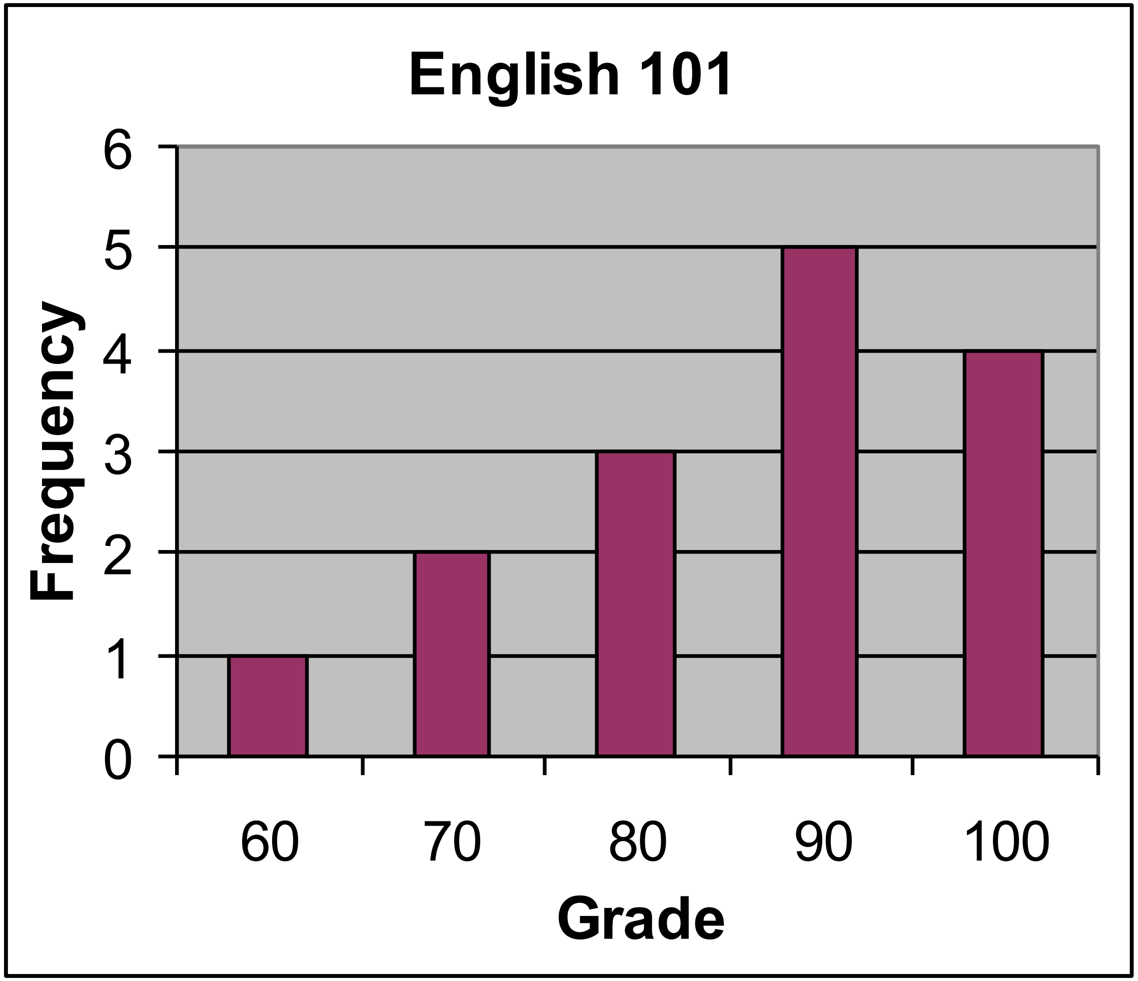
$$= \frac{60+140+240+450+400}{15}$$

$$= \frac{1290}{15}$$

$$= \mathbf{86}$$

Histogram of grade data

NOTE: columns should be <u>touching</u>



Another example . . .

Example 2: The following is a table showing the weights (in kg) of **24** prize pumpkins. (The weights have been rounded to the nearest whole pound.) Find the mean (average) weight and draw the histogram associated with these data. The raw data have been put into a frequency table, which facilitates our calculations and drawing.

28	35	42	14	28	42
42	14	21	35	42	35
21	42	28	42	21	35
35	21	42	28	35	21

Note: **n = 24** (total number of pumpkins.)

Putting the data into a frequency table

28	35	42	14	28	42
42	14	21	35	42	35
21	42	28	42	21	35
35	21	42	28	35	21

Table 2: *Prized Pumpkin Weights (kg)*

Weight	**Frequency**
14	2
21	5
28	4
35	6
42	7

Finding the mean

Mean $= \dfrac{2(14)+5(21)+4(28)+6(35)+7(42)}{24}$

$= \dfrac{28+105+112+210+294}{24}$

$= \dfrac{749}{24} = 31.2$

$\bar{x}$ = 31.2 kg The mean weight of this group of pumpkins is 31.2 kg.

The Median

- (NB: "median" ≠ "medium")

- **The Median is the *Middle*** data value – after the data values have been sorted into numerical order.

NOTE: It is important to determine whether there are an **even** or an **odd** number of data.

Example of finding the median

(Previous data) A class of 15 students in English 101 received the following grades on a test.

Find the class **median**.

100, 60, 90, 90, 70, 100, 100, 90, 90, 80, 70, 100, 80, 90, 80

- i) are the data sorted?
- ii) are there an odd or even number of data?

Median continued . . .

There are 15 data values – an odd number.

And the data are not sorted . . . yet.

Median continued . . .

There are 15 data values – an odd number.

And the data are not sorted . . . yet.

Data Value	How Many
60	1
70	2
80	3
90	5
100	4

Now they are!

Median continued…

For an odd number of data, the position of the median is n/2. Here the position of the Median = 15/2 = 7.5

You have to round 7.5 up to 8.

So the Median will be in the 8^{th} Position.

This is not the same saying that the Median is 8, which it obviously is not! (why?)

Median continued …

<u>Finding the actual MEDIAN data value</u>. In this case, we are finding the MEDIAN score on the English 101 test. So we need to find the 8th score – now that the scores are in order:

Data Value	**How Many**	**positions**
60	1	1st
70	2	2nd and 3rd
80	3	4th, 5th, 6th
90	5	7th - 11th
100	4	12th - 15th

Clearly the 8th score is a 90. So the median score is 90.

Seeing the Median visually

60, 70, 70, 80, 80, 80, 90, 90, 90, 90, 90, 100, 100, 100, 100

this is the **8th spot.**

Notice the same number of data are to the left and right of the median.

Thus the median is the center of the data.

Another example of the median . . .
(remember those pumpkins?)

28	35	42	14	28	42
42	14	21	35	42	35
21	42	28	42	21	35
35	21	42	28	35	21

Table 2: *Prized Pumpkin Weights (kg)*

Weight	**Frequency**
14	2
21	5
28	4
35	6
42	7

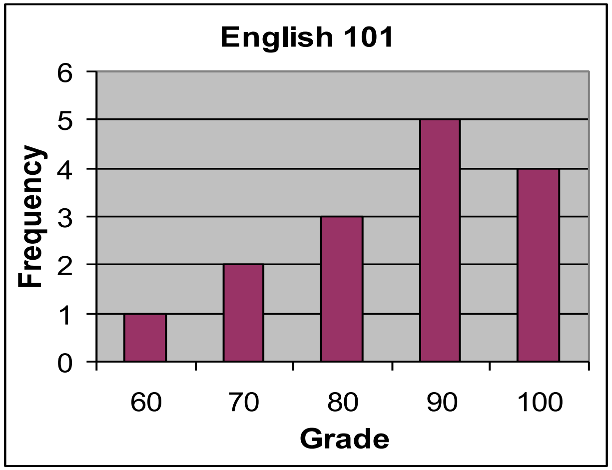
Pumpkin median . . .

Remember this time there are an <u>even</u> number of data; therefore there are 2 middle values. Calculate n/2, which = 12. The median is the average of the 12^{th} and the 13^{th}, i.e. the average of 35 and 35, so 35.

Weight	**Frequency**	**Positions**
14	2	1st and 2nd
21	5	3rd - 7th
28	4	8th - 11th
35	6	12th - 17th
42	7	18th - 24th

The Mode…
Note: columns should be touching

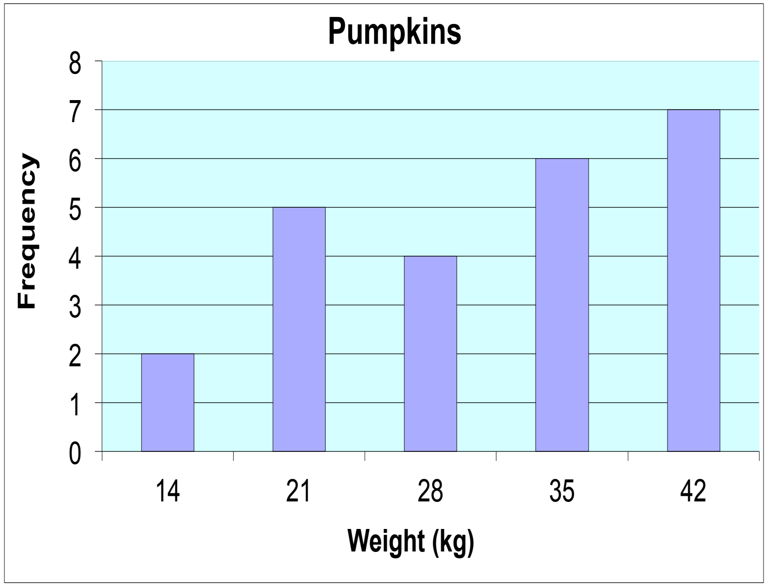
Mode - ***Most*** frequently occurring data value. This is usually the easiest one to "calculate."



The highest bar is over the 90, so the mode is 90.

Another example of the mode
Note: columns should be touching

- The highest bar is over the 42, so the mode is 42 kg.



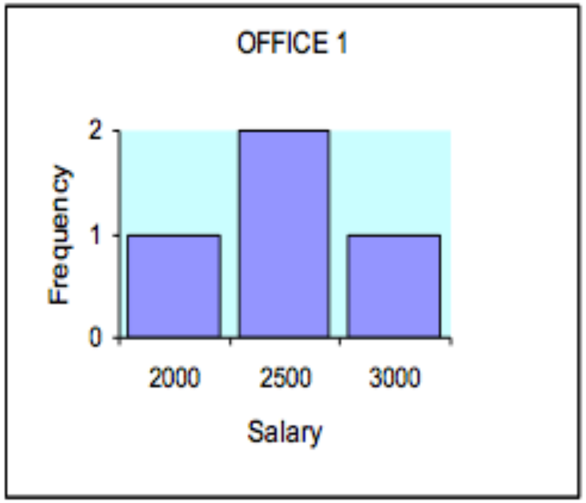
Comparing the mean and median

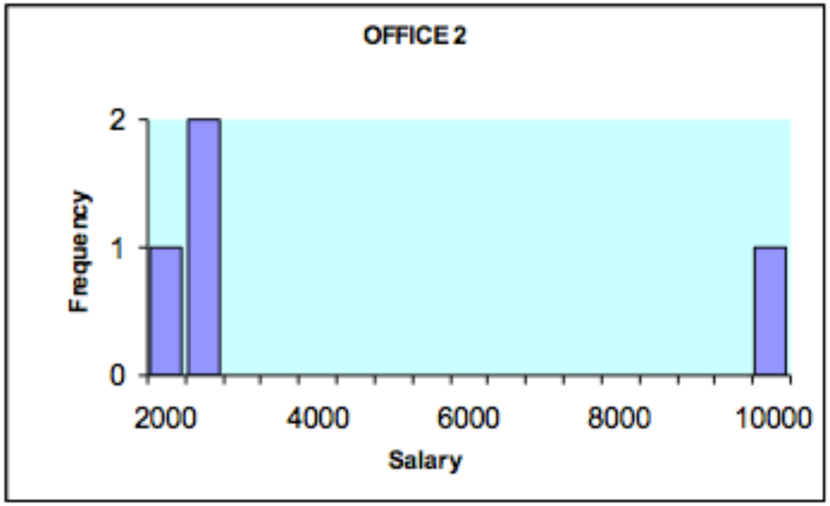
OFFICE 1

Name	*Salary*
Jane	$2,000
Rae	2,500
Rita	2,500
Pat	3,000

OFFICE 2

Name	*Salary*
Rose	$2,000
Carol	2,500
Connie	2,500
Ann	10,000





mean salary = $2,500
median salary = 2,500
mode = 2,500

mean salary = $4,250
median salary = 2,500
mode = 2,500

Measures of dispersion: the range

Range = Largest minus smallest data value.

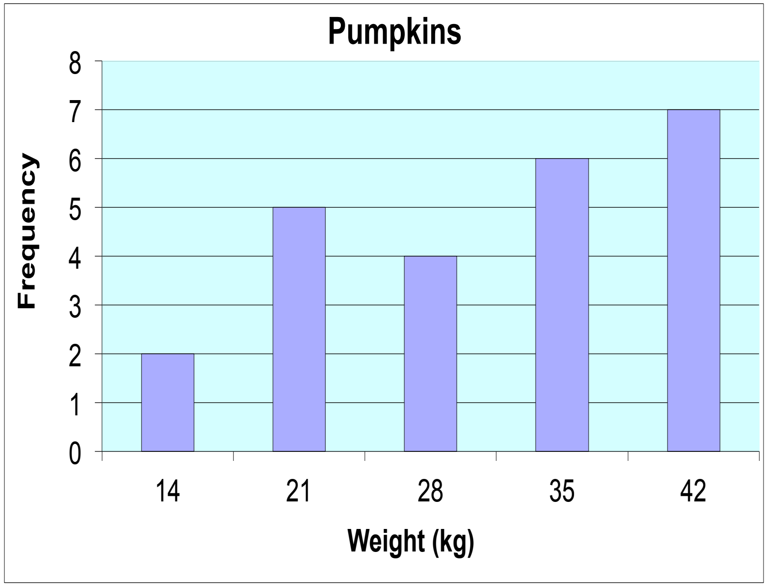
Example 1. Data from Example 1.

Data Value	**How Many**
60	1
70	2
80	3
90	5
100	4

Range = 100 – 60 = **40**

Another range example
Note: columns should be touching

- Range = largest – smallest = 42 – 14 = 28 kg.



The standard deviation

(why? I hear you ask…)

Standard Deviation: The average of the deviations of each data value from the mean.

The formula is $\sigma = \sqrt{\dfrac{\Sigma(x - \bar{x})^2}{n}}$

Score data one more time . . .

Data Value	How Many
60	1
70	2
80	3
90	5
100	4

Imagining the standard deviation . . .

(remember that song by John Lennon?)

1. First find the mean – since we're interested in seeing how much the data "deviate" from the mean
2. Then see how far from the mean each value is
3. Then, since some of these distances are positive and some are negative . . . "do something" to make the negative signs "go away" (apart from just ignoring them . . .). You actually <u>square</u> each of the deviations
4. Then if there are more than one of a particular value, multiply the deviation by the frequency of that value
5. Total up all the "squared deviations"
6. Take the mean of this total
7. Take the square root of this mean
8. What is left should be the standard deviation. Pretty cool, huh?

Here it is in action . . .

Data Value A	Deviation B	Deviation^2= (Column B)^2 C	Frequency D	Column C*D E
x	$x - \bar{x}$	$(x - \bar{x})^2$	freq	$(x - \bar{x})^2 * f$
60	60-86= –26	$(-26)^2$=676	1	676
70	70–86=–16	$(-16)^2$=256	2	512
80	80–86=–6	36	3	108
90	90–86=4	16	5	80
100	100–86= 14	196	4	784
			n=15	Σ = 2160

Last three steps . . .

To finish it off, steps 6, 7, and 8:

Step 6: Sum the last column - this is symbolized by the Σ so Σ=2160

Step 7: Divide this total by the total frequency (the sum of Column D = **n**) $\frac{2160}{15}$

Step 8: Take the Square Root of this answer. Written in the formula it looks like this:

$$\sigma = \sqrt{\frac{2160}{15}} = \sqrt{144} = \mathbf{12}$$

Standard deviation of the prize pumpkins
(sounds like the title of a horror movie)

We already calculated the mean weight of these pumpkins; it is 31.2 kg. (I rounded the mean to 1 decimal place for this example.)

Data Value A	Deviation B	Deviation^2= (Column B)^2 C	Frequency D	Column C*D E
x	$x - \bar{x}$	$(x - \bar{x})^2$	freq	$(x - \bar{x})^2 * f$
14	14–31.2= –17.2	$(-17.2)^2$=295.84	2	295.84*2=591.68
21	21–31.2= –10.2	$(-10.2)^2$=104.04	5	104.04*5=520.20
28	28–31.2= –3.2	10.24	4	40.96
35	35–31.2= 3.8	14.44	6	86.64
42	42–31.2= 10.8	116.64	7	816.48
			n=24	Σ = 2055.96

Last few steps . . .

$$\sigma = \sqrt{\frac{2055.96}{24}} = \sqrt{85.665} = \mathbf{9.2553...} \qquad \textbf{Rounds to 9.3}$$

So the standard deviation of the prize pumpkin weights is 9.3 kg.

A way of using the standard deviation . . .

- Sometimes it is useful to know how much of the data is "within one standard deviation of the mean."

- Let's try that here: what % of the score data (first example) is within one s.d. of the mean?
- Mean = 86, s.d. = 12.
- "Within 1 s.d. of the mean" means between 86-12 and 86+12, i.e. between 74 and 98. So how many scores are between 74 and 98?
- Answer: The 3 80s and the 5 90s, making 8 scores out of 15, which equals <u>53.3%</u>.

You have just been initiated into the secret mysteries of statistics!!

But also, this is probably the most tricky thing we do in this course, so you may take some comfort from that.

You will need to be able to a problem like this on the first test, so do the practice examples, as well as the problems on the assignment.

Statistics is actually very useful, since people are always interested in counting things, finding averages, discovering whether data are spread out or tightly clustered together, and especially whether a certain occurrence is expected or "out of the ordinary." More to come on this . . .