Final Exam Review

May 3, 2016

Final Exam Review

May 3, 2016

Final Exam Information

- **DATE:** Tuesday, May 10

- **TIME**: 7:40 pm – 9:40 pm (ET)*

  *Students living outside the 6 New England states will have a 24 hour window in which to complete the exam. You must arrange, with a proctor, any two hour period between May 10, 7:40 pm and May 11, 7:40 pm (ET) in which to complete your exam.

- **LOCATION**:
  - If you live within the 6 New England states: **Maxwell-Dworkin G115**
  - If you live outside the 6 New England states: You must arrange for a proctor:

- http://www.extension.harvard.edu/resources-policies/exams-grades-transcripts/exams-online-courses

- Proctor questions should be directed to: distance_exams@dcemail.harvard.edu or call (617) 495-0977 Monday through Friday, 9 am to 5 pm eastern time.

Preliminary Matters

The final exam will be mostly multiple choice.

Topics you should know:

Percents
Summary Statistics (mean, median, mode)
Confidence Intervals
Hypothesis Testing
Linear Growth
Scatterplots
Regression and Correlation
Exponential Growth and Decay

The exam will not cover:

coins
dice
Megabucks
1st type of Standard Deviation Formula
Z-scores
Excel

Agenda

- Summary Stats: Mean, Median, Mode
- Confidence Intervals
- Hypothesis Testing
- Q&A
- **Break**
- Linear Growth
- Regression
- Exponential Growth and Decay
- Q&A

Measures of Location
Mean, Median, Mode

Also known as location descriptors, measures of center, measures of central tendency:

- Mean: the average of all data values.
- Median: middle data value when data values are in numerical order.
- Mode: most frequently occurring data value.

The Mean

$$\text{Mean} = \frac{\text{Sum of all data values}}{\text{Total Number of Data}}$$

An Example for mean

- A running club made up of seventeen members has been training together for the Boston marathon. They tend to run in groups according to their ability. On the day of the marathon, their times were recorded in the chart that follows.
- What is the mean time for the running club in the marathon?

Example: Mean, Median, Mode

Time in marathon in hours	Number of runners
2.4	1
3.2	2
3.6	3
3.8	5
4.0	3
4.2	3

Mean Marathon Time

$$\text{Mean} = \frac{\text{Sum of all data values}}{\text{Total Number of Data}}$$

$$\text{Mean} = 1\frac{(2.4)+2(3.2)+3(3.6)+5(3.8)+3(4.0)+3(4.2)}{17}$$

$$\frac{2.4+6.4+10.8+19+12+12.6}{17}$$

$$\frac{63.2}{17} = 3.7 \text{ hours}$$

Median

- The median is the data value(s) in the middle when data is sorted in either ascending or descending order
- If you have an odd amount of data, there is one median value
- If there is an even amount of data, there are two values in the middle which must be averaged to get the median

Median

If I have 5 pieces of data as follows:

2 6 9 12 15

There is one middle value

2 6 9 12 15

The median is 9

Median

If I have 6 pieces of data as follows:

3 7 10 12 16 19

There are two middle values

3 7 10 12 16 19

Average these to get the median

$\frac{10+12}{2} = 11$ is the median

Median using running club

Median Value = middle value

Make sure data is in order from smallest to largest, or largest or smallest

Odd amount of data =17

Because it is odd, there is one middle value

Divide 17 by 2=8.5

Round up to 9 –look at the 9^{th} piece of data

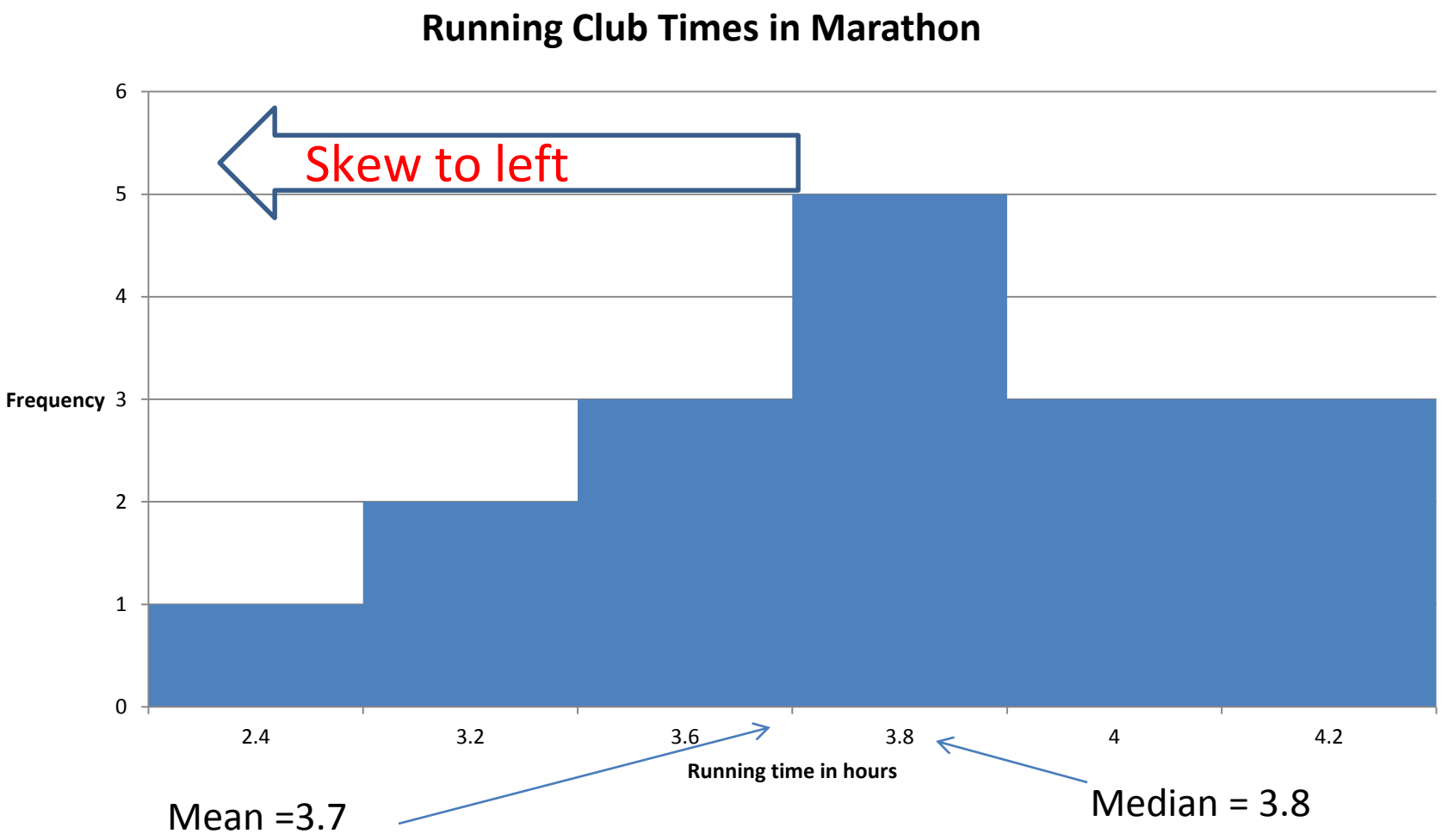
Running club example: Median

Hours to complete marathon	Number of runners
2.4	1
3.2	2
3.6	3
3.8	5
4.0	3
4.2	3

Median

- 3.8 is the median
- It also happens to be the mode as it is the most frequently occurring
- So to recap:
- Mean = 3.7
- Median=3.8
- Mode =3.8

Histogram of Running Club



Skew to left

- Mean lower than median
- Pulled in direction of the skew

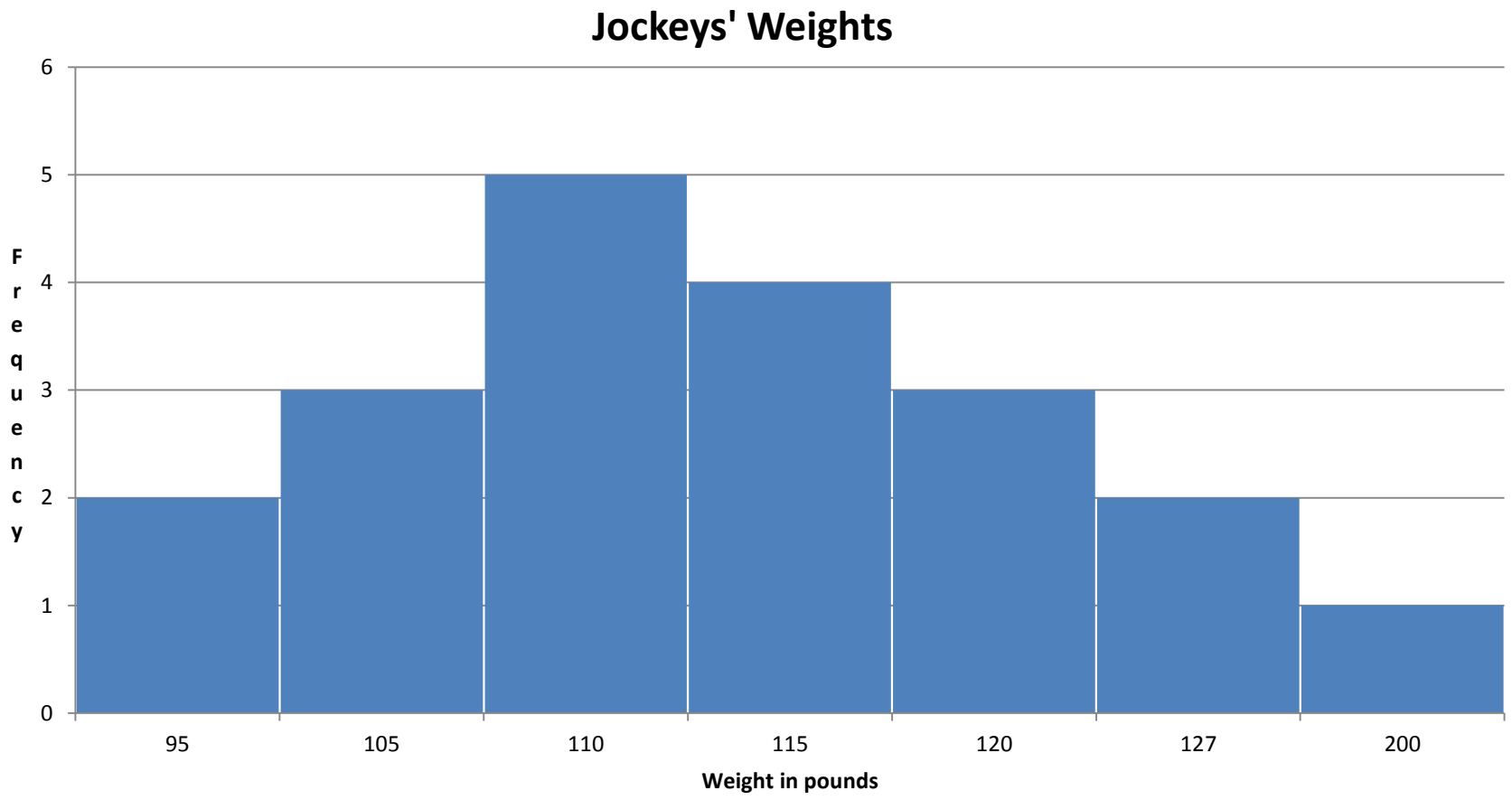
Another Example

Recently in the running of the Kentucky Derby, the jockeys had to weigh in. Normally, the jockeys are very light weight, but one unfortunately gained a lot of weight prior to the race. The data on the jockeys' weights follow.

Example: Mean, Median, Mode

Jockey weight in pounds	Number of jockeys
95	2
105	3
110	5
115	4
120	3
127	2
200	1

Kentucky Derby example



Jockey example

Before we do any calculations, do you think the mean is less than, equal to, or greater than the median?

Jockey example

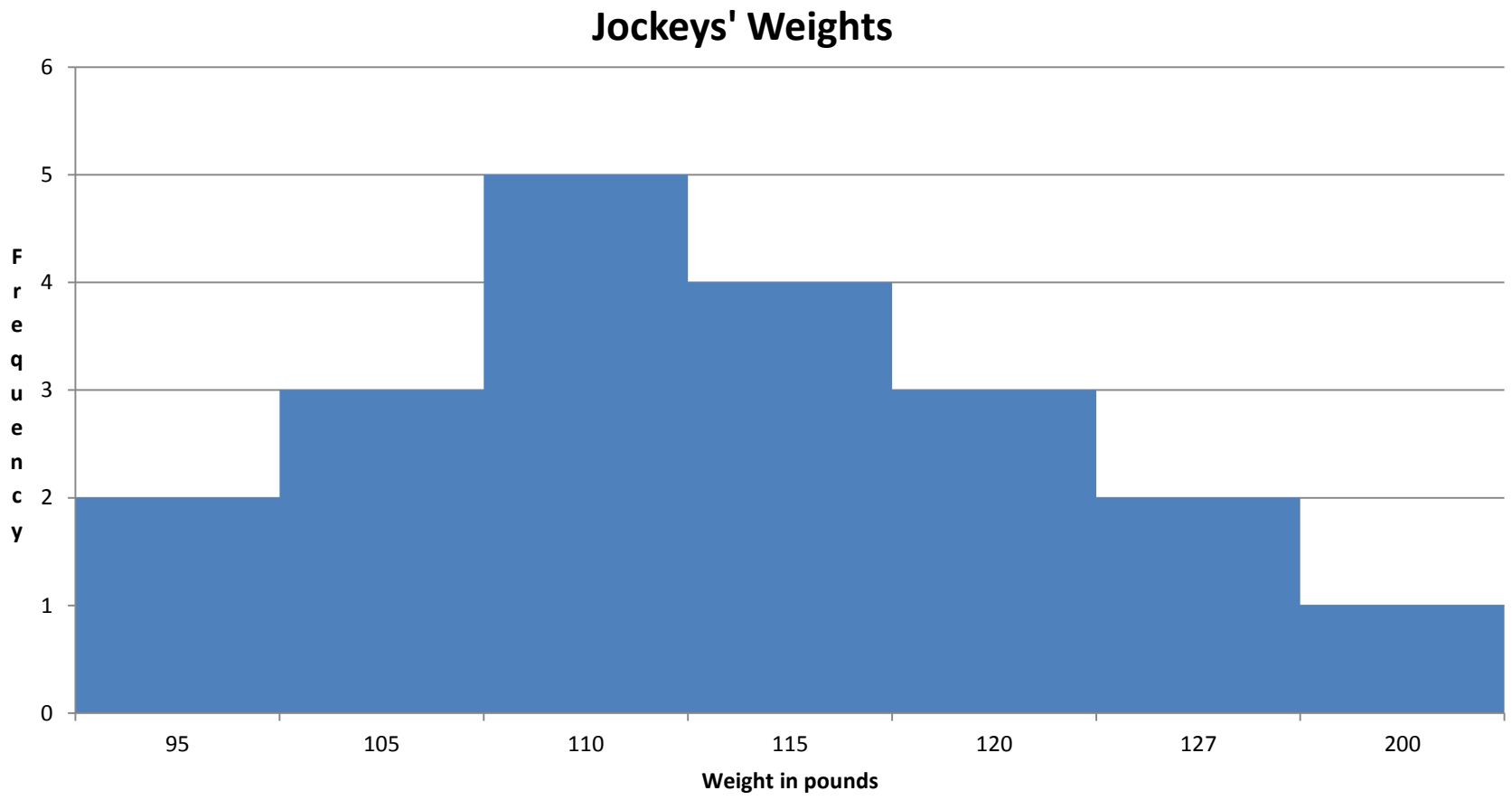
Before we do any calculations, do you think the mean is less than, equal to, or greater than the median?

It's greater.

Why?

Is this skewed right, skewed left, or is it normal?

Kentucky Derby example



Jockey example

Before we do any calculations, do you think the mean is less than, equal to, or greater than the median?

It’ s greater.

Why?

Is this skewed right, skewed left, or is it normal?

Skewed right

Jockey example

Let's do the math:

Mean:

$$\frac{2(95)+3(105)+5(110)+4(115)+3(120)+2(127)+1(200)}{20} =$$

$$\frac{190+315+550+460+360+254+200}{20}$$

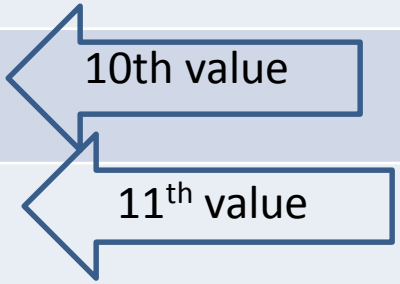
$$\frac{2329}{20} = 116.45 \text{ or } 116 \text{ to nearest whole pound}$$

Jockey example

- Median is the data value(s) in the middle
- This set of data has 20 values
- That means there are two middle values
- Which ones are in the middle?
- Divide 20 by 2 and you get 10
- Average 10^{th} value with the one above it, the 11^{th}, and you will have the median value

Example: Mean, Median, Mode

Jockey weight in pounds	Number of jockeys
95	2
105	3
110	5
115	4
120	3
127	2
200	1



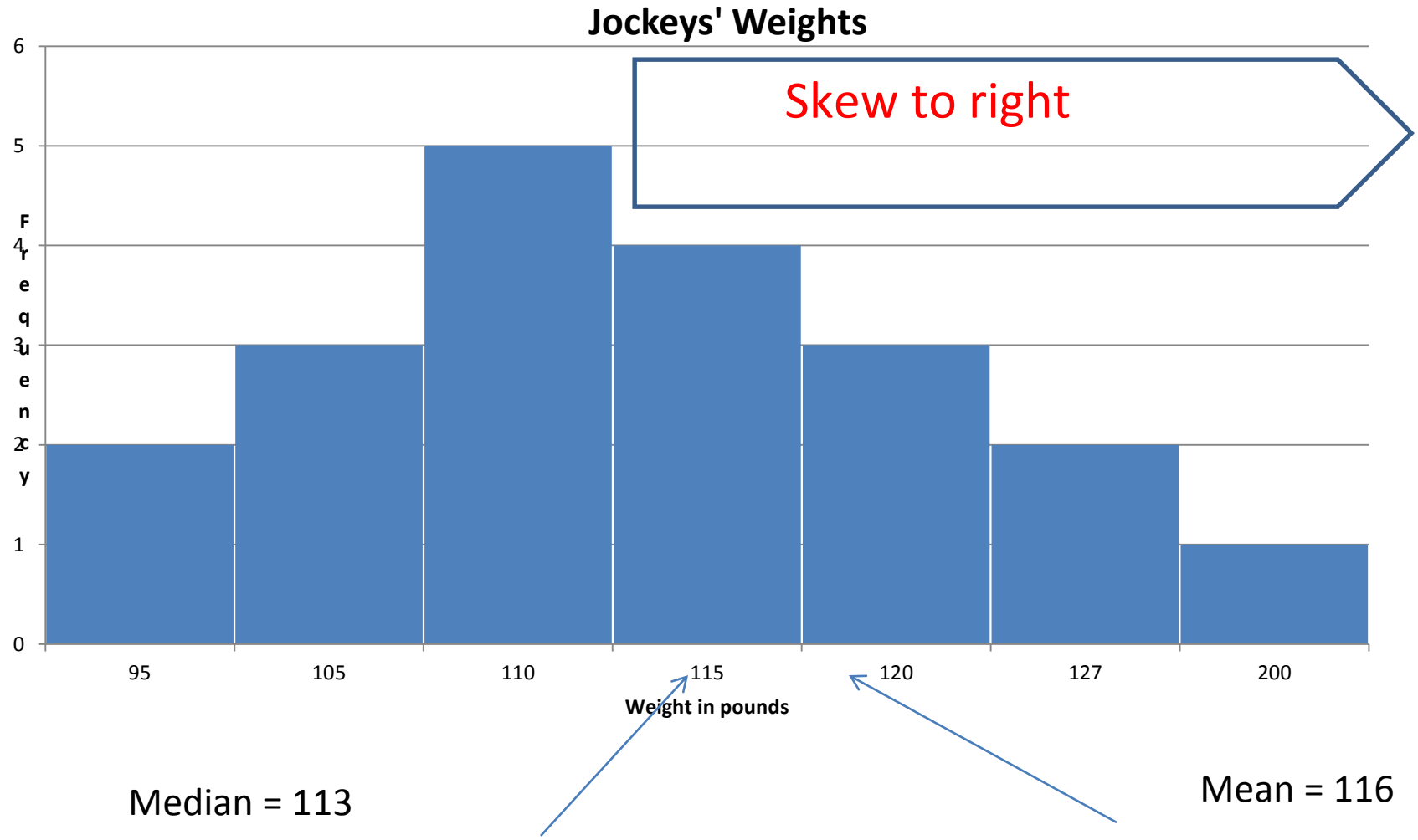
So median is…

The median is:

$$\frac{110 + 115}{2} = \frac{225}{2} = 112.5 \text{ or } 113$$

So our mean, 116, is greater than the median because the mean has been pulled in the direction of the skew

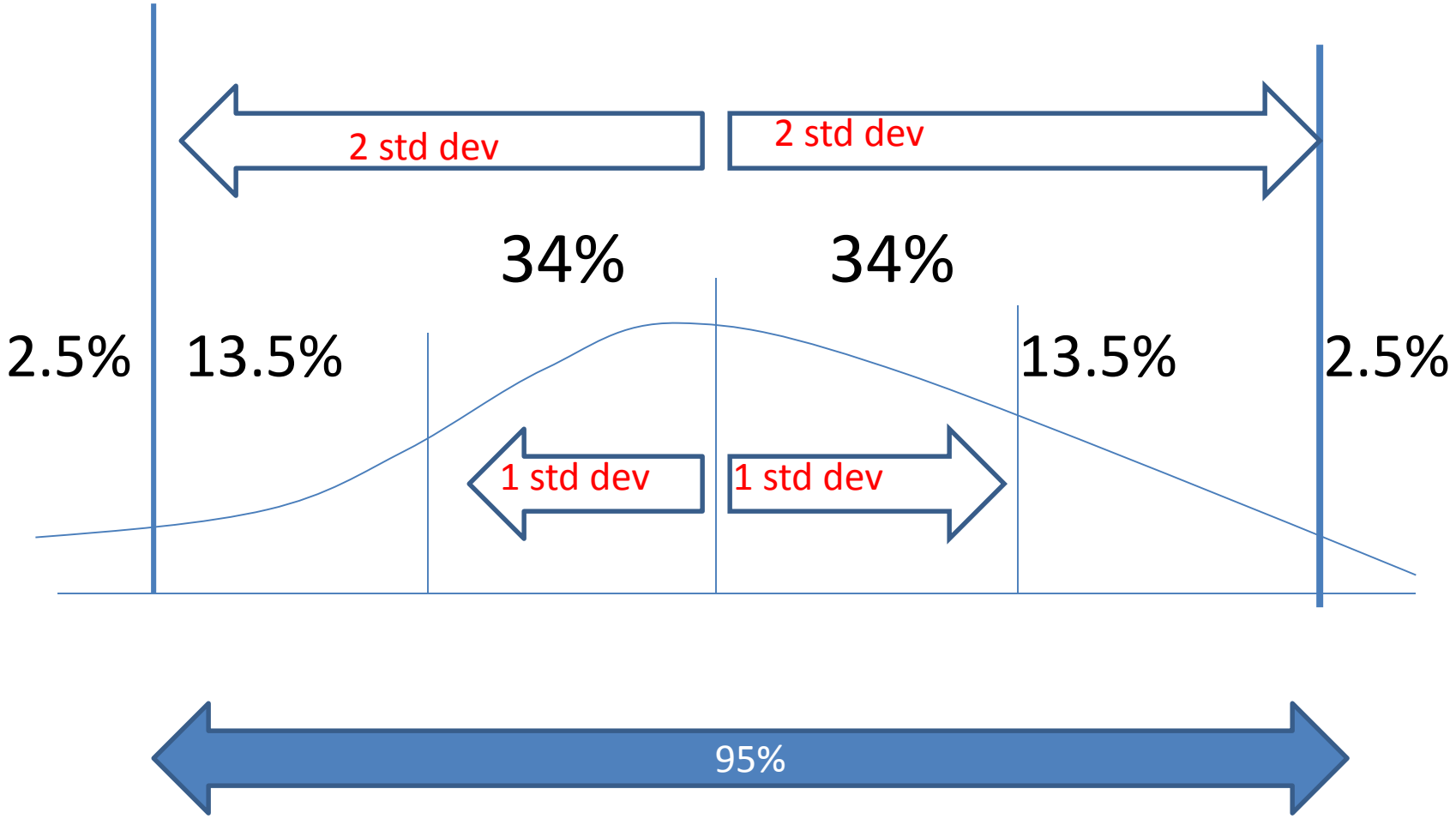
Kentucky Derby example



Confidence Intervals

- It all started with the Normal Distribution Curve….
- Rough and ready rules…

Normal distribution curve



Rough and ready rules

- 68% of the data is within 1 standard deviation from the mean
  - 34 % of each side
- 95% of the data is within 2 standard deviations
  - 47.5 on each side
- 100% of the data is under the curve
  - 50% on each side

95% Confidence Interval

- The 95% confidence interval is:
- The mean ± 2 standard deviations
- We use *p* to represent the mean and σ to represent our standard deviation so our formula to get the 95% confidence interval is: $p \pm 2\sigma$

A New Standard Deviation Formula

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

Used for problems involving binary data.

Standard deviation formula

- If you have binary data involving only two choices
- *p* is the percentage
- n is the sample size
- Let’s do an example

Survey of number of lefties

- We took a random sample of people, and asked how many were left-handed:
- Found that out of a sample of 850 people, 79 were left-handed
- Make this a percentage: 79/850 =9.3%
- Note: Can only use new standard deviation formula with %
- Now let's find the true proportion of people who are left-handed

Calculate 95% confidence interval

- Calculate standard deviation (round to 1 dp):
- Change 9.3% to the decimal form: .093
- Get your sample size = 850
- Put these values into the standard deviation formula

Left hand example

$$\sigma = \sqrt{\frac{.093(1-.093)}{850}}$$

Lefties

- You will get this:
- σ = .00996175
- Change to a percent:
- .99175%
- Round appropriately to 1.0%

Calculate 95% confidence interval

- Get margin of error =
- 2 * 1.0% = 2.0%
- Our confidence interval =
- 9.3% ± 2.0%, or
- 7.3%  11.3%

Conclusion

The true proportion of left-handed people is 9.3% ± a margin of error of 2.0%

More examples

- What if you were happy with a wider margin of error such that your 95% confidence interval was 9.3 ± 3.0% error instead of 2%?
- How do you calculate this? You know the margin or error, and the mean, but not the sample size?
- Use the formula for standard deviation, and solve for n

Solving for n

- Since the margin or error is 3.0%, our standard deviation is half that or 1.5%
- So change this to a decimal, .015, and put the values you know into your formula:
- Solve for n using algebra, or by guessing what n is and calculating how close you can get to .015

Solving for different Standard Deviation

.015 is std dev associated with a 3 % margin of error

$$.015 = \sqrt{\frac{.093(1-.093)}{n}}$$

Solving for new std dev

- .093 *(1-.093)=.093*.907=.084351
- Now you have:

New std deviation

$$.015 = \sqrt{\frac{.084351}{n}}$$

Guessing

850 gave you a margin of error of 2%, so try 500

This will give you 1.3%

Still too low, you want 1.5%

Try 400

This will give you 1.45%

Try 375

This will give you 1.5% (or pretty close)

1.4997867%

Use algebra to get this

This is where you get to once you multiply .093*(1-.093)

$$.015 = \sqrt{\frac{.084351}{n}}$$

Now square both sides of the equation

Algebra

Squaring both sides gets rid of the square root sign and you are left with:

$.015^2 = \frac{.084351}{n}$

$.000225 = \frac{.084351}{n}$

$.000225n = .084351$

$N = 374.89$ or

375

Tips for test if you get one of these

- If you get one of these and it is multiple choice, plug in the answers to the formula in the denominator, and solve for new standard deviation
- Remember the new std deviation will be ½ of the new margin of error

Hypothesis Testing

- We are still using the same formula to calculate standard deviation
- We still are concerned with $p \pm 2\sigma$ which is now our likely region
- But, we have a claim percentage and an observed percentage
- The claim percentage may be plainly stated, and sometimes it's a little trickier

Hypothesis Testing

- A car company claims only 2% of their cars needs repairs in the first 2 years of operation. You test it and find that the 35 out 1600 need service.
- What is the claim percentage?
- What is the observed percentage?

Hypothesis Testing

- A car company claims only 2% of their cars needs repairs in the first 2 years of operation. You test it and find that the 35 out 1600 need service.
- What is the claim percentage? 2%
- What is the observed percentage?

$$\frac{35}{1600} = 2.2\%$$

Hypothesis Testing

- Sometimes you need to do a little bit of thinking to get to the claim percentage
- Let’s do an example

Hypothesis Testing Example

A director of a charitable organization wants to sell Christmas trees to raise money for a project.

She has observed that most people seem to like Douglas Firs the best of the five common varieties of Christmas trees.

She decides to ask one of her volunteers who has taken Math E-3 to conduct a hypothesis test to test the popularity of the 5 trees.

Hypothesis Test

- The volunteer sets up the hypothesis test by formulating his Null Hypothesis.
- Then he picks a random sample of Christmas tree purchasers, and asks them to choose their favorite from among the five different types of Christmas trees.

Volunteer Survey

He does a survey of 600 Christmas tree enthusiasts and gets the following results:

Tree	# who prefer	% of people who prefer
Douglas Firs	150	150/600 =25.0%
Scotch Pine	125	125/600=20.8%
Blue Spruce	125	125/600=20.8%
Arizona Cypress	95	95/600=17.5%
Eastern Red Cedar	105	105/600=17.5%

Christmas Tree Hypothesis Test

What is his null hypothesis?

a)People prefer Douglas Firs 25 % of the time.

b)People enjoy all different varieties equally and would choose them equally, meaning 25% per tree.

c)People would choose Douglas Firs 50% or more of the time.

d) People would choose all trees equally at 20 % per tree.

Christmas Tree Hypothesis Test

What is his null hypothesis?

a)People prefer Douglas Firs 25 % of the time.

b)People enjoy all different varieties equally and would choose them equally, meaning 25% per tree.

c)People would choose Douglas Firs 50% or more of the time.

d) People would choose all trees equally at 20 % per tree.

Why is it d?

Well, it's not:

- a) People prefer Douglas Firs 25 % of the time.
- 25% is an observed percentage
- The Null Hypothesis should contain your claim percentage
- The observed percentage is used to compare against your likely region
- The observed percentage would never be your claim percentage

Why d???

Well it's not:

b) People enjoy all different varieties equally and would choose them equally, meaning 25% per tree.

And, it's not c

c) People would choose Douglas Firs 50% or more of the time.

Why d?

AND THE WINNER IS:

d) People would choose all trees equally at 20 % per tree.

There are 5 trees

If you are trying to see if the varieties are equally popular, you would have an equal percentage of those surveyed choosing them

100%/5 choices =20% apiece

Hypothesis Testing

- Step 1: State your Null Hypothesis
  - Important to have a claim percentage
- Step 2: Calculate your standard deviation using new formula
  - Important to use the claim percentage in your standard deviation calculation
- Step 3: Draw your normal curve with your claim percentage as the mean

Hypothesis Testing

- Step 4: See if your observed percentage is inside the likely region
- Step 5: State your conclusion in proper statistical language
  - If your observed result is **outside p±2σ**, we can **reject** the NH at a 5% los (level of significance)
  - If your observed result is inside p±2σ, we cannot reject the NH at a 5% los

Hypothesis Testing

- Step 6: Explain it to me like I’m a five year old
- No seriously, put in your own words

Hypothesis Testing

Important stuff:

Never use the word accept –it is cannot reject

Use your claim percentage when calculating your standard deviation

Our standard deviation for the Christmas Tree Test would use what?

Hypothesis Testing

- 20%
- So, calculate the standard deviation using our formula:

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

Hypothesis Test

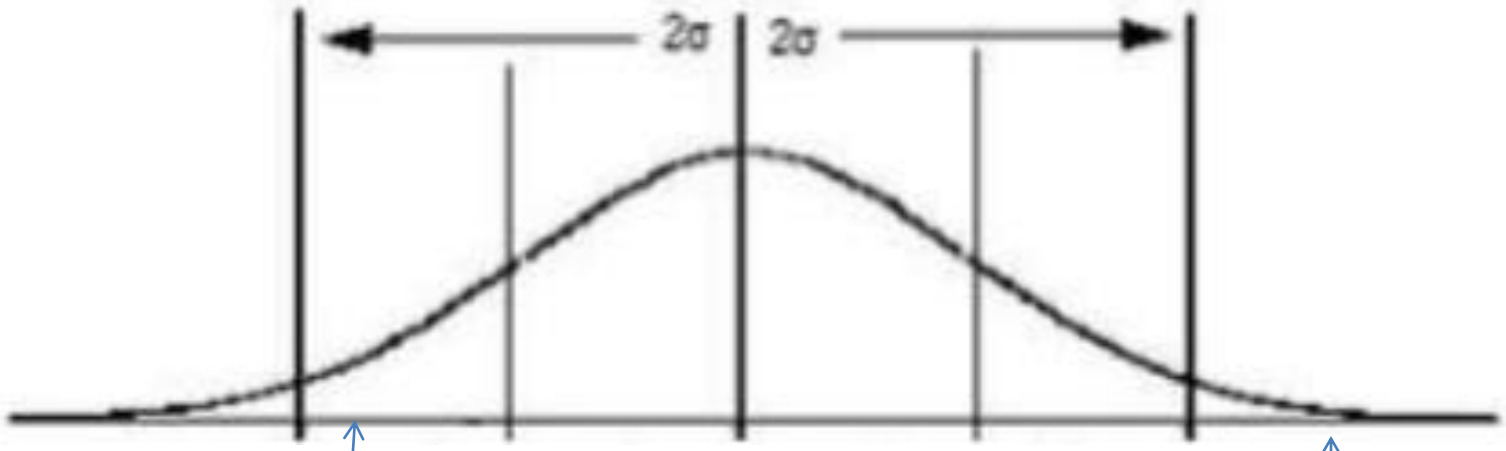
- Here it is:

$$\alpha = \sqrt{\frac{.20(1-.20)}{600}}$$

- .01633 =1.633% =1.6%
- Change this to a percentage always
- Round to 1 dp unless otherwise instructed
- Always round last – in other words, after you convert to a percentage

Draw the curve

Draw curve:



15.8%
Red Cedar

17.5%
Arizona Cypress

20.8%
Scotch Pine and Blue Spruce

25.0%
Douglas Firs

So answer this…

Which conclusion is correct?

a. We can reject the NH at a 5% los because our observed % is inside $p \pm 2\sigma$
b. We cannot reject the NH at a 5 % los because our observed % is outside the $p \pm 2\sigma$
c. We can reject the NH at a 5% los because our observed % is outside the $p \pm 2\sigma$
d. We accept the NH at a 5% los because our observed % is inside the $p \pm 2\sigma$

So answer this…

Which conclusion is correct?

a. We can reject the NH at a 5% los because our observed % is inside $p \pm 2\sigma$
b. We cannot reject the NH at a 5 % los because our observed % is outside the $p \pm 2\sigma$
c. We can reject the NH at a 5% los because our observed % is outside the $p \pm 2\sigma$
d. We accept the NH at a 5% los because our observed % is inside the $p \pm 2\sigma$

Break time

- Up next:
- Linear Growth
- Regression
- Exponential Growth and Decay

Linear Growth

The equation of a line is:

$Y=mx+b$

What does this mean?

Let's break it down

x and y are two variables that have a relationship to each other that is characterized by a constant rate of change that we represent as m

Y=mX+b

- So, what does the b represent?
- It is where the line is positioned, specifically where the line crosses the y axis
- So if we have an equation of a line like y=4x+3, we know the slope is 4
- That means as x goes up by 1, y will go up by 4

Y=4x+3

- We know the y intercept is 3
- What does that mean?
- It means the line crosses the y axis at 3, or at the coordinate (0,3)
- So, let's look at how to graph this

How do we get the points?

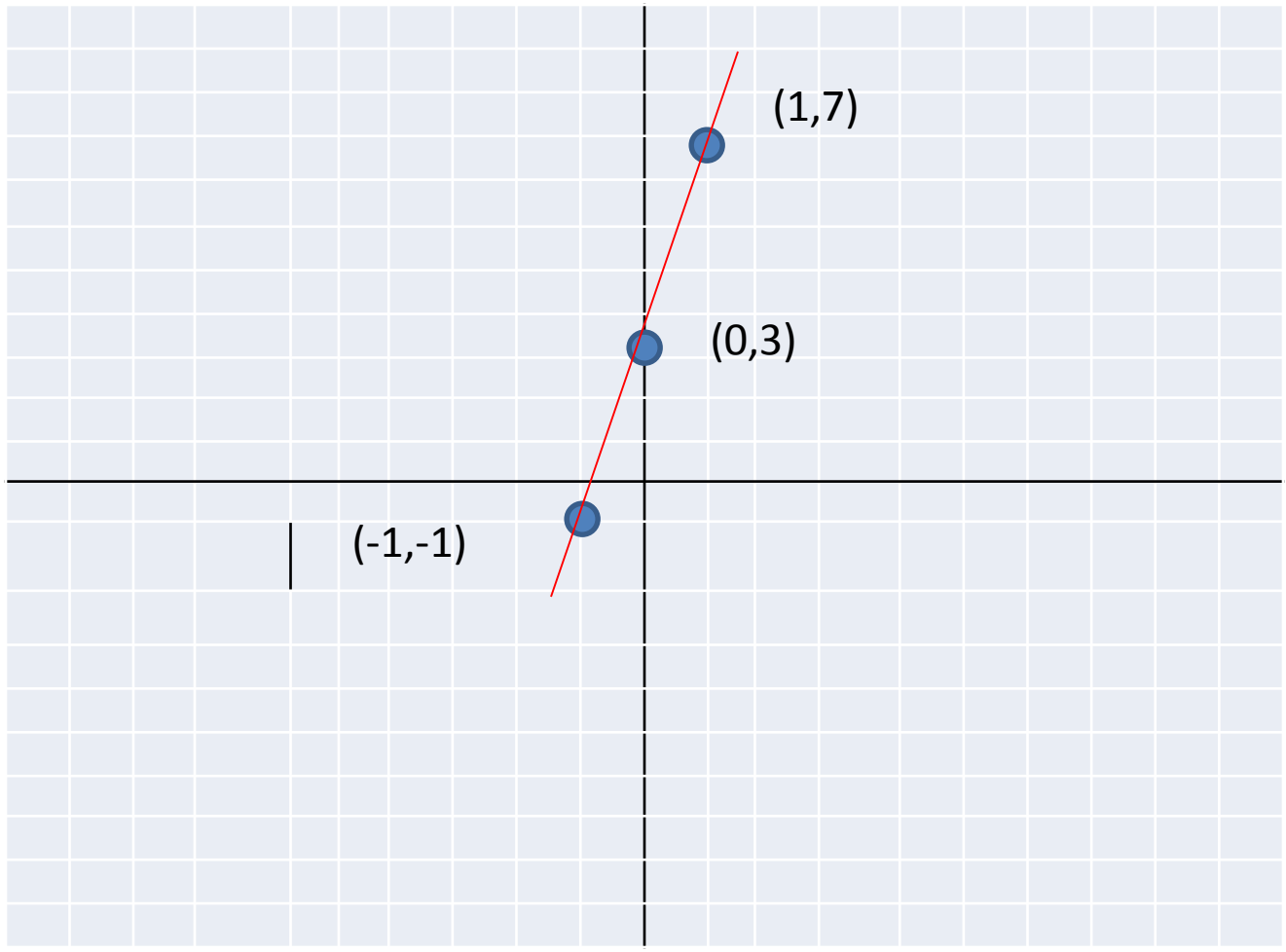
- Make a table
- Pick a value for x
- Calculate Y using the formula, and substituting your values for x

The Table

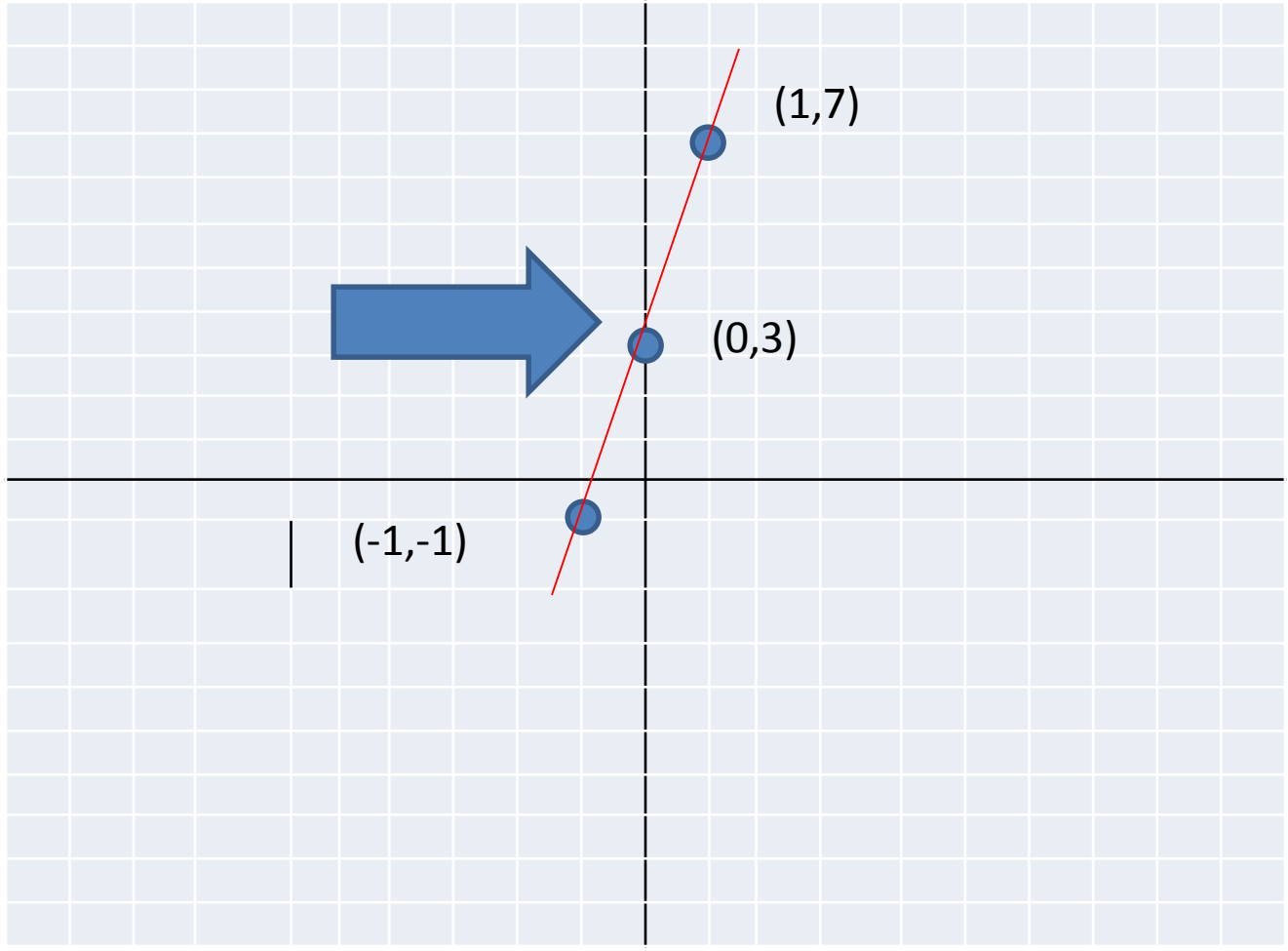
Y= 4(x) + 3

x	y
0	If x=0, then y=4(0)+ 3= 3
1	If x=1, then y=4(1)+ 3= 7
-1	If x=-1, then y=4(-1)+3=-1

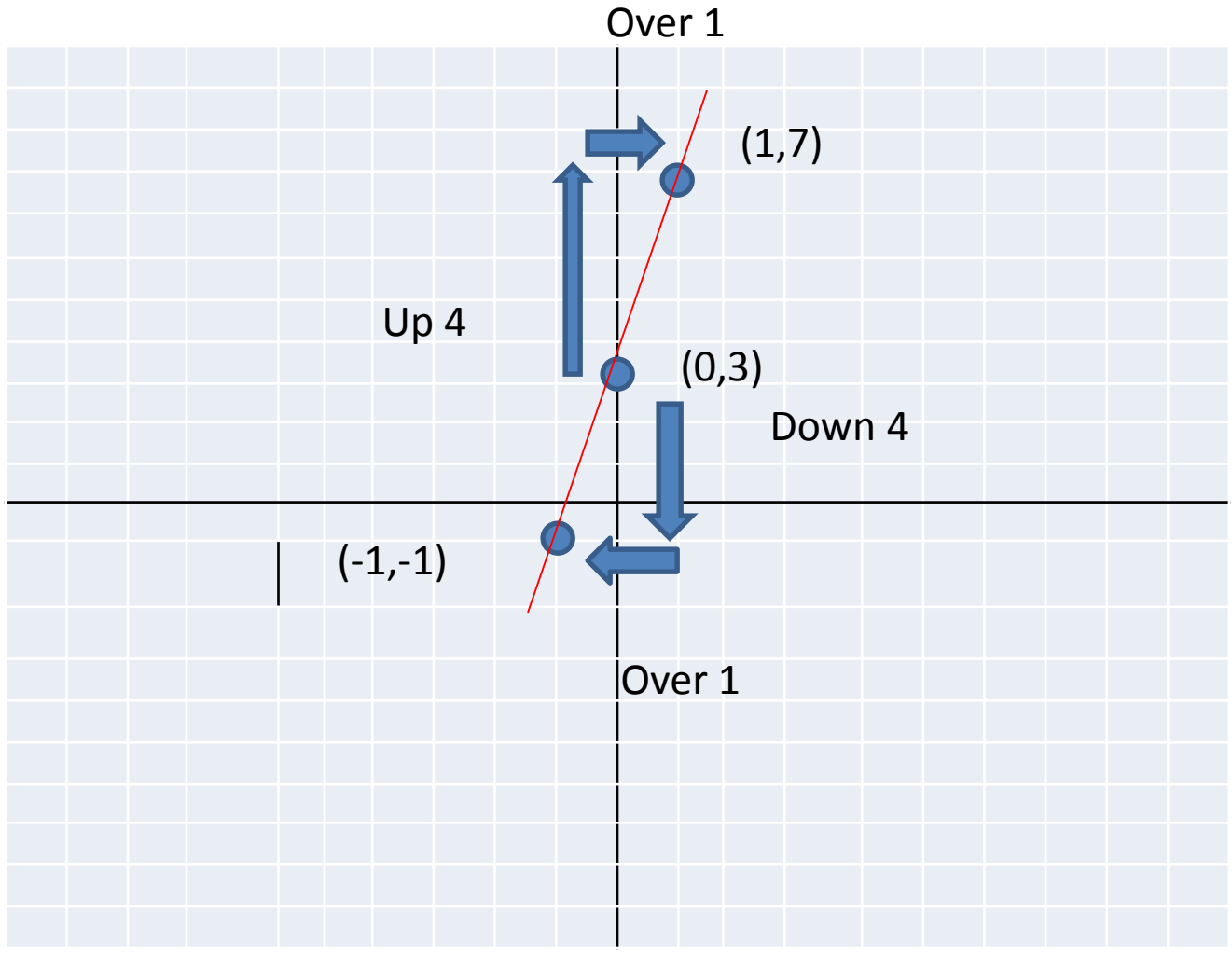
Graph of Y=4x+3



The y intercept shows you where the line crosses the y axis



The slope shows you where to go



Characteristics of y=4x+3

- The slope is positive
- The y intercept is positive
- So, it y intercept is above the origin (0,0)
- And the line slants upward as you look from left to right

Lines and Linear growth problem

If you have a line with two points (2,8) and (-1,2), what is the slope of the line?

a. 2
b. 8
c. -1
d. 1/2

Lines and Linear growth problem

If you have a line with two points (2,8) and (-1,2), what is the slope of the line?

a. 2
b. 8
c. -1
d. 1/2

Line Basics

- From the points, you can calculate the slope
- Two points: (2,8) (-1,2)
- Slope = $\frac{Y_2 - Y_1}{X_2 - X_1} = \frac{8 - 2}{2 - (-1)} = \frac{6}{3} = 2$

Linear Growth Problem

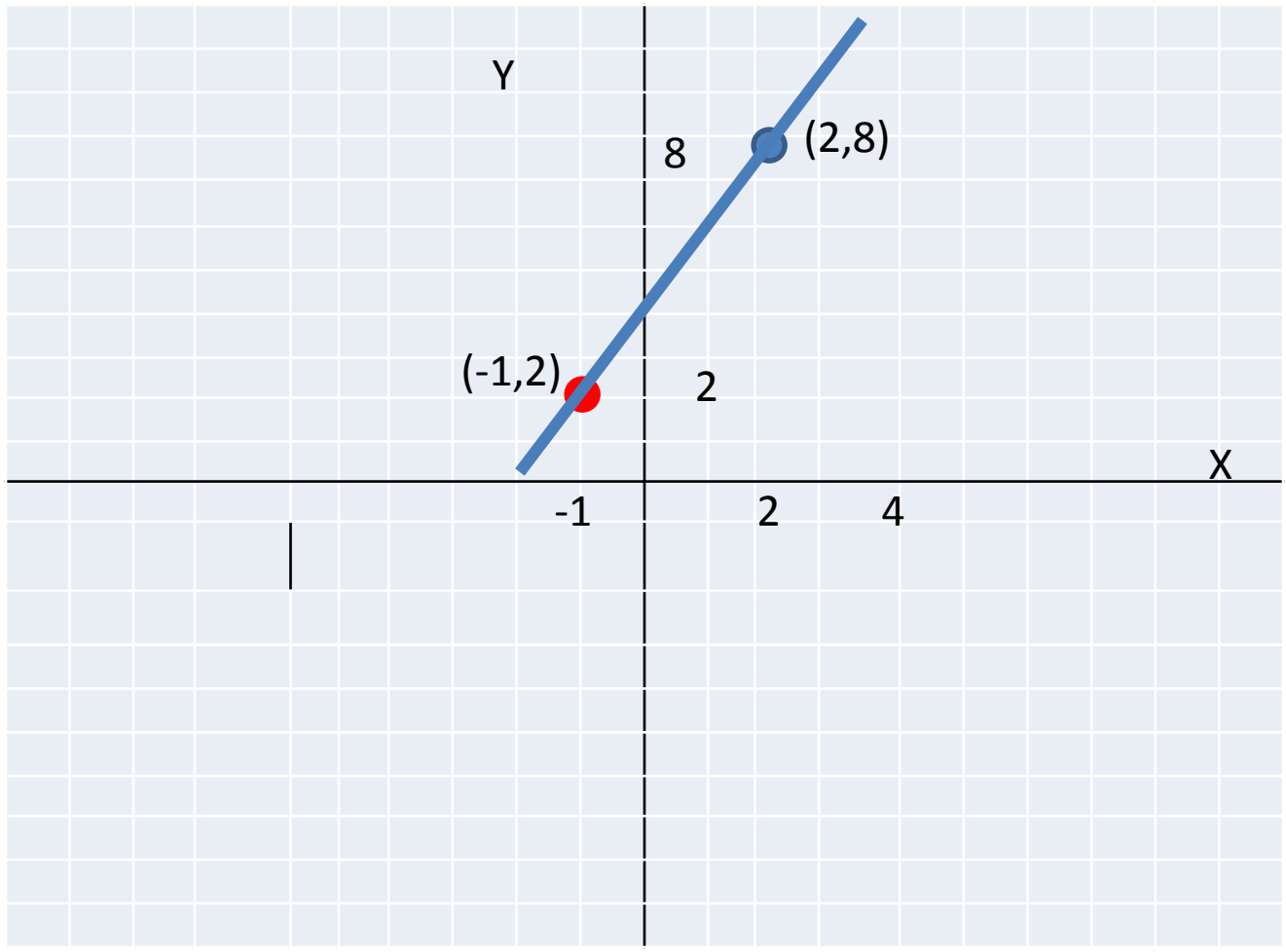
What is the equation of this line that contains (2,8) and (-1, 2)?

You already have slope, (it is 2), but how would you get y intercept?

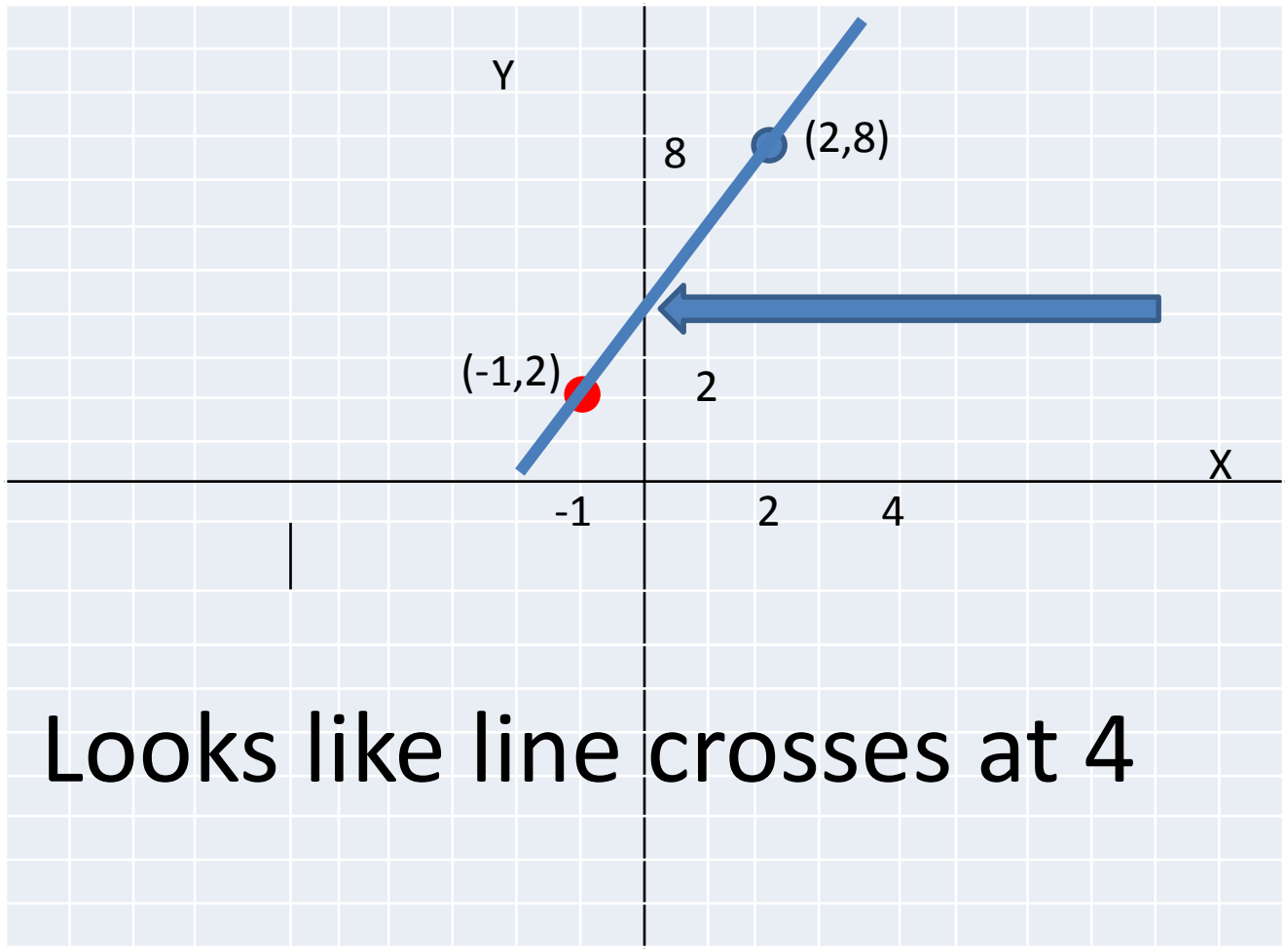
Either graph the points, draw the line, and see where the line crosses the y axis, or

Calculate algebraically

Draw a line



See where the line crosses the y-axis



Looks like line crosses at 4

Line Basics

- From the graph of the line, you can see where the line crosses the y intercept
- From these two pieces of information, you can get the equation of the line you created
- The formula for a line is y=mx+b where m is the slope and b is the y intercept (where the line crosses the y axis)

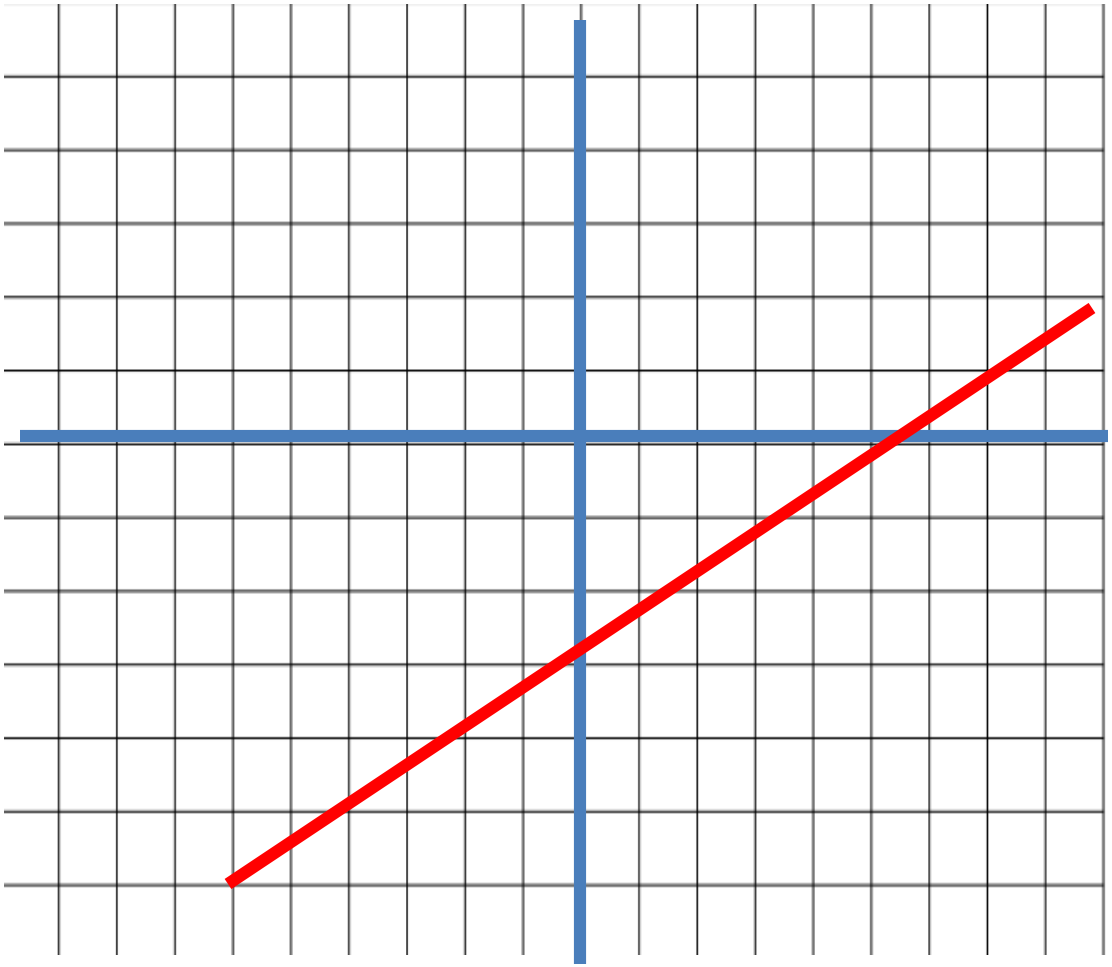
Line Basics

- So, we have the slope =2, and it looks like the y intercept is 4,
- So plugging in to y=mx+b, the equation of our line is:
- Y=2x+4
- We could get the y intercept algebraically
- Method to follow but you don't have to use it

Getting the y intercept algebraically

- Two points: (2,8) (-1,2)
- Get your slope which we did and got 2
- Your equation so far is:
- Y=2x+b
- Put in one of your points for x and y and solve for b
- Using (2,8) : 8= 2(2)+b
- 8=4+b, therefore b=4

Let’s look at some other lines



Possible equations

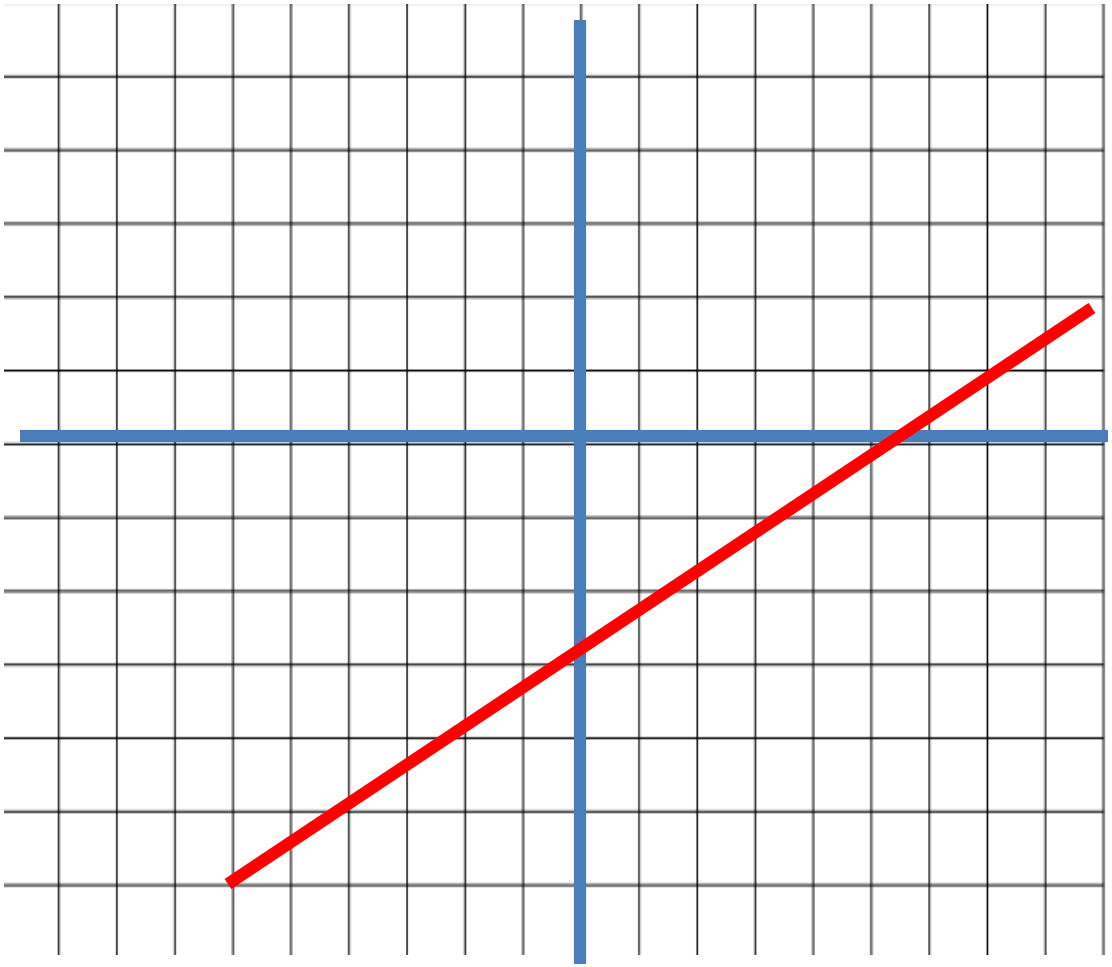
a. Y= - 2x+3

b. Y= x + 4

c. Y= -3x +1

d. Y= 2x -3

Let’s look at some other lines



Possible equations

a. Y= - 2x+3

b. Y= x + 4

c. Y= -3x +1

d. Y= 2x -3

Why is this the only equation which could possibly apply?

The slope must be positive, and the y intercept must be negative

a. Y= - 2x+3

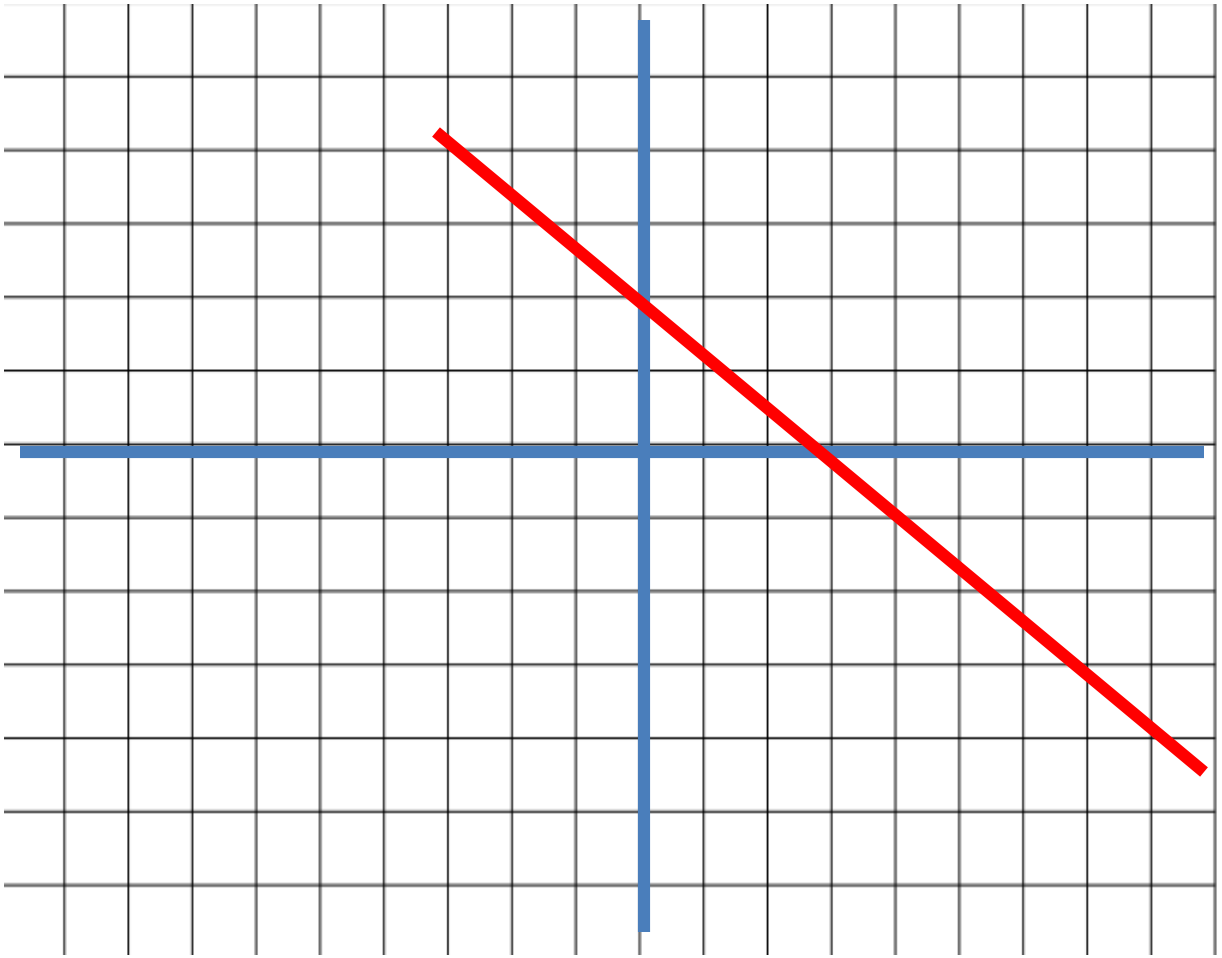
b. Y= x + 4

c. Y= -3x +1

d. Y= 2x -3

Choice d is the only one that had positive slope and a negative y intercept

How about another one?



Possible equations

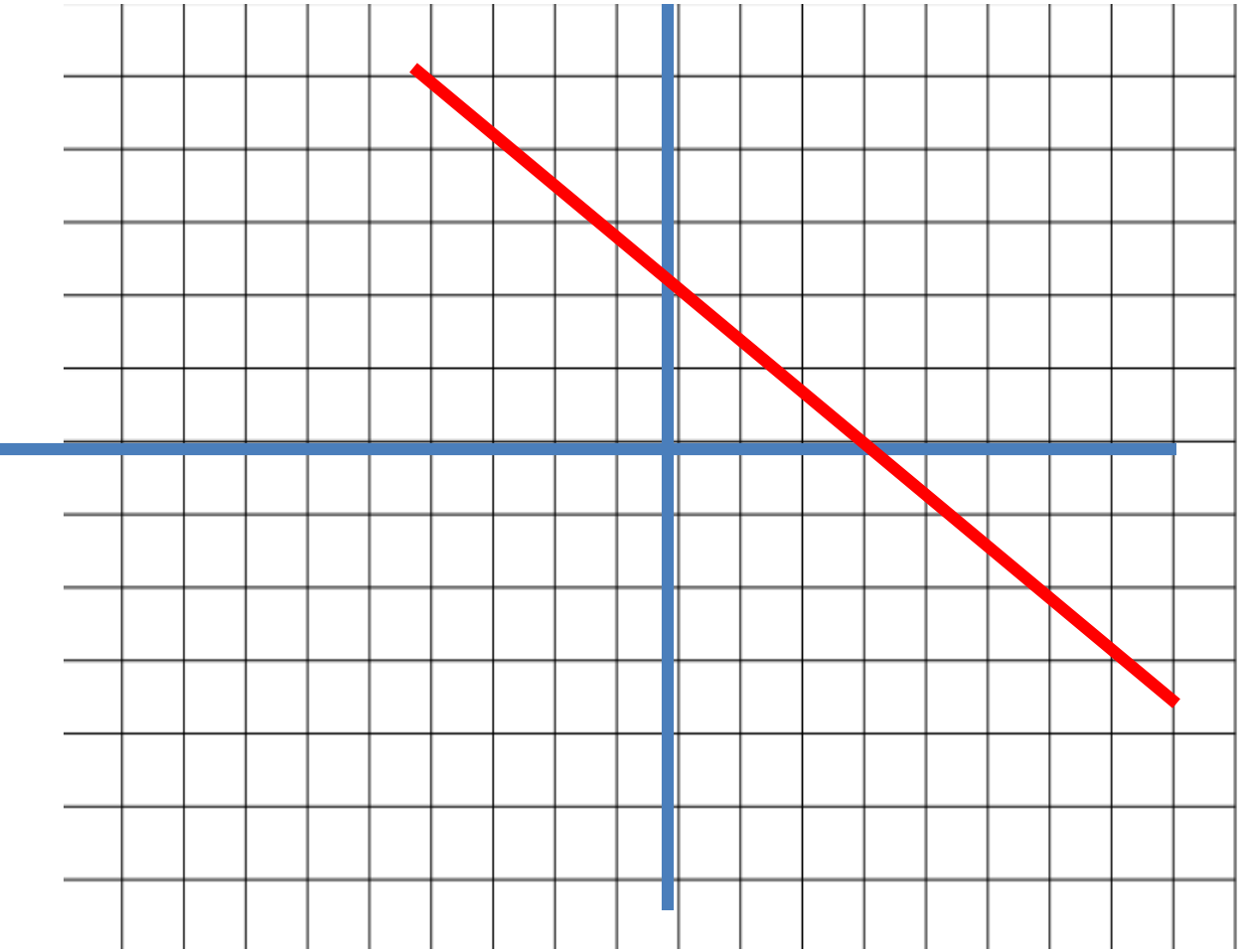
a. Y= 2/3 X +2

b. Y= - 2/3+2

c. Y= 2/3x-3

d. Y= -2/3x-2

How about another one?



Possible equations

a. Y= 2/3 X +2

b. Y=-2/3+2

c. Y= 2/3x-3

d. Y= -2/3x-2

Why is that the right choice?

The slope must be negative and the y intercept must be positive

a. Y= 2/3 X +2

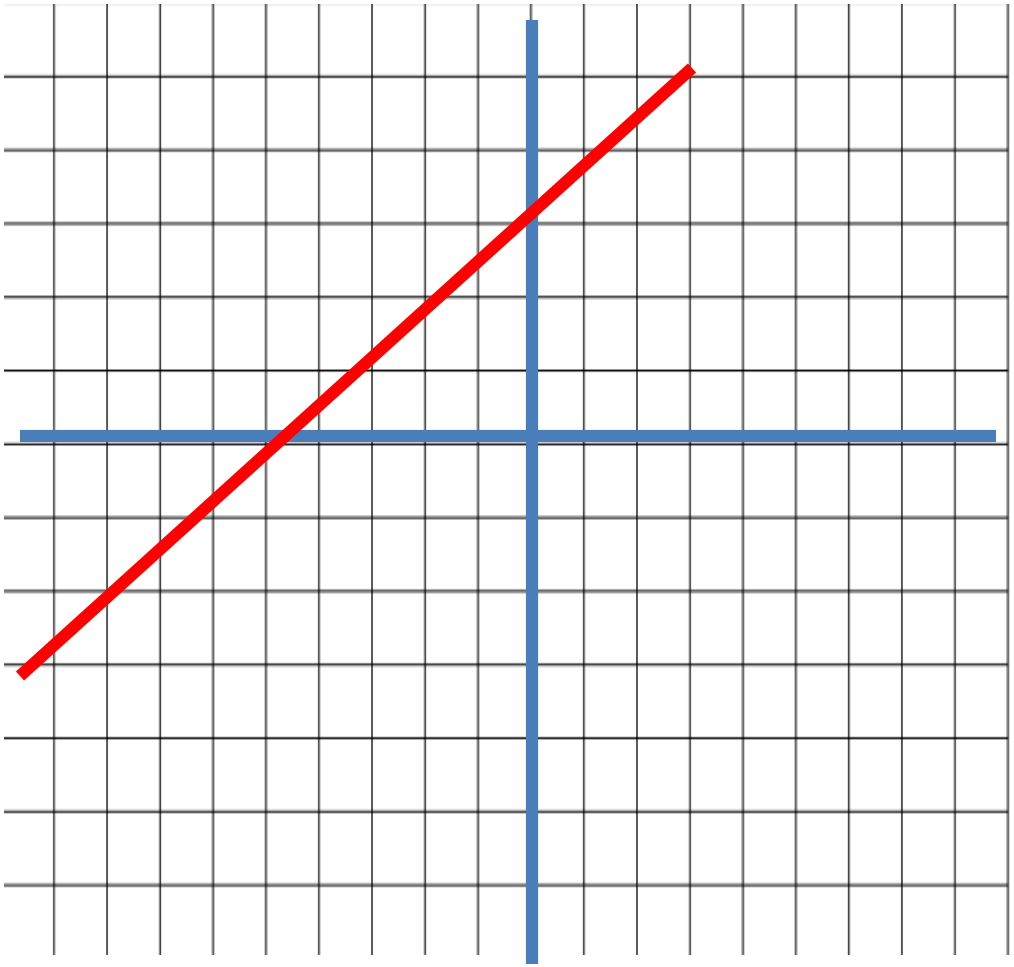
b. Y=-2/3+2

c. Y= 2/3x-3

d. Y= -2/3x-2

Choice b is the only one that has a negative slope and a positive y intercept

You could do this all day, right?



Possible equations

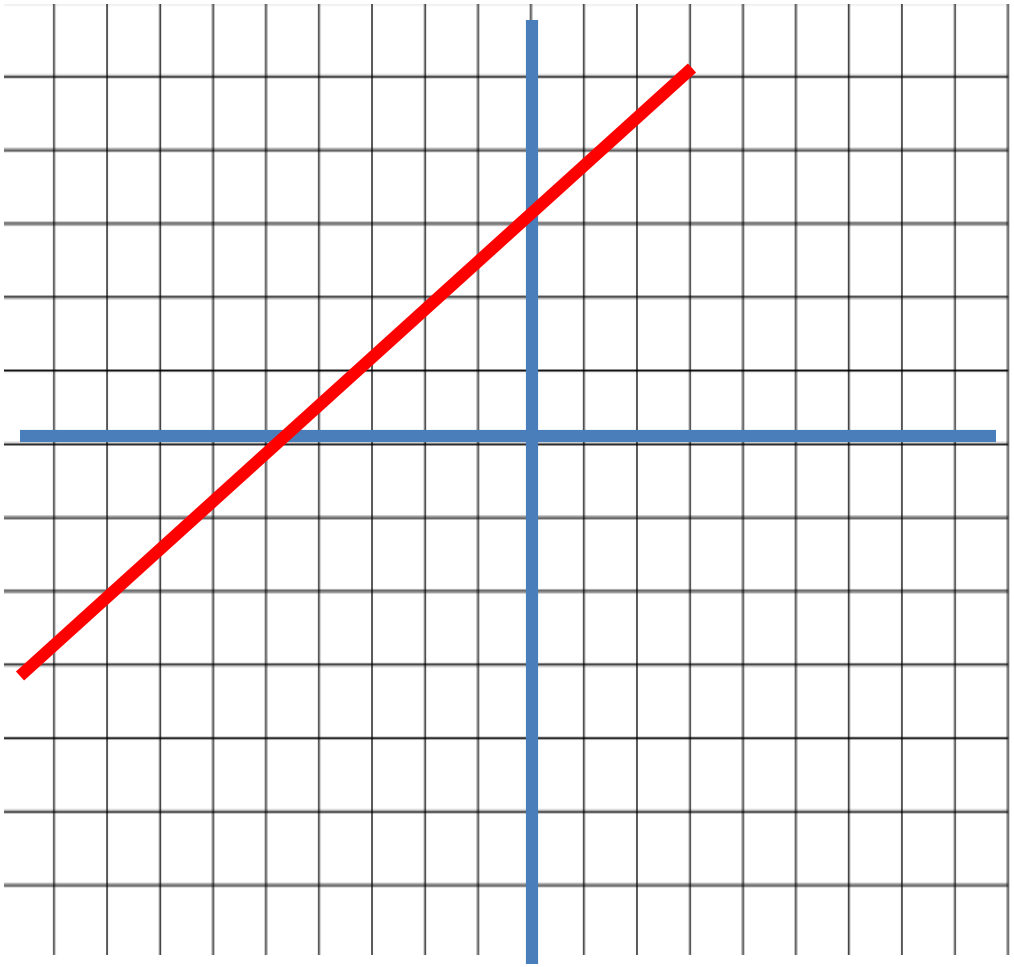
a. Y= 3/4x- 4

b. Y= -3/4x +3

c. y= 3/4x+3

d. Y= -3/4x-4

You could do this all day, right?



Possible equations

a. Y= 3/4x- 4

b. Y=-3/4x +3

c. y= 3/4x+3

d. Y=-3/4x-4

Alright, why?

It must have a positive slope as the line moves upward as you look from left to right, and it must have a positive y intercept

a. Y= 3/4x- 4

b. Y=-3/4x +3

c. y= 3/4x+3

d. Y=-3/4x-4

Choice c is the only one that has a positive slope and a positive y intercept

Regression

- Regression…
- Just a fancy name for taking a bunch of points, and estimating a line
- We still need Y=mx+b

Regression steps

- You get a table of data
- You plot that data on a graph
- You put whatever the independent variable is on the x or horizontal axis
- And you put the dependent variable on the y or vertical axis
- Let’s do an example

Regression

A high school teacher suspected that her students' grades for the term were adversely affected by the number of days they were absent from class. She went onto her online system and found the following data:

Regression

# of days absent	Term average
3	85
6	80
10	60
2	90
1	70
4	95
4	56
5	69
15	45
1	92

How to analyze the data

Draw a scatter plot:

What goes on the x axis?

Days absent or term average?

Why?

How to analyze the data

Draw a scatter plot:

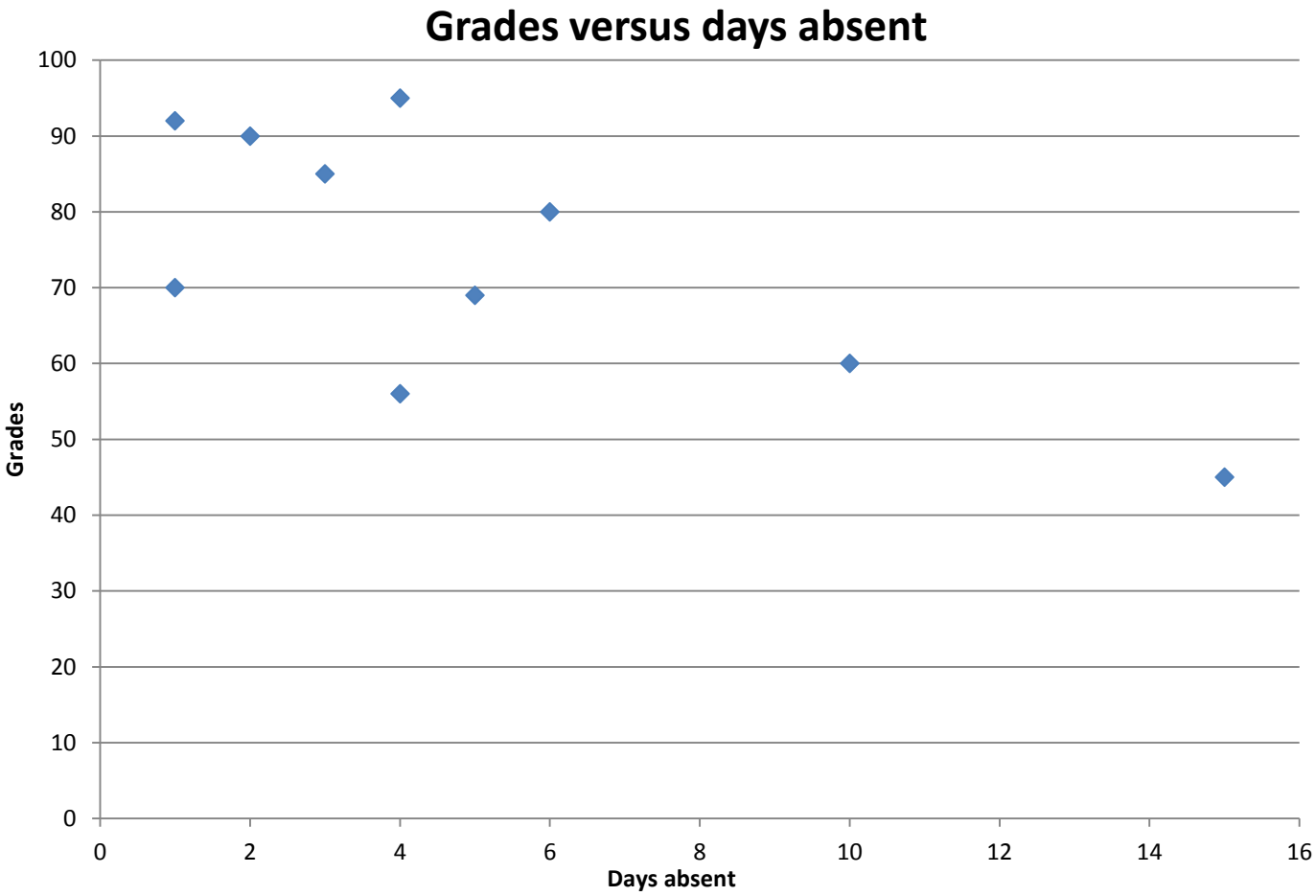
What goes on the x axis?

Days absent

Why?

It is the independent variable.

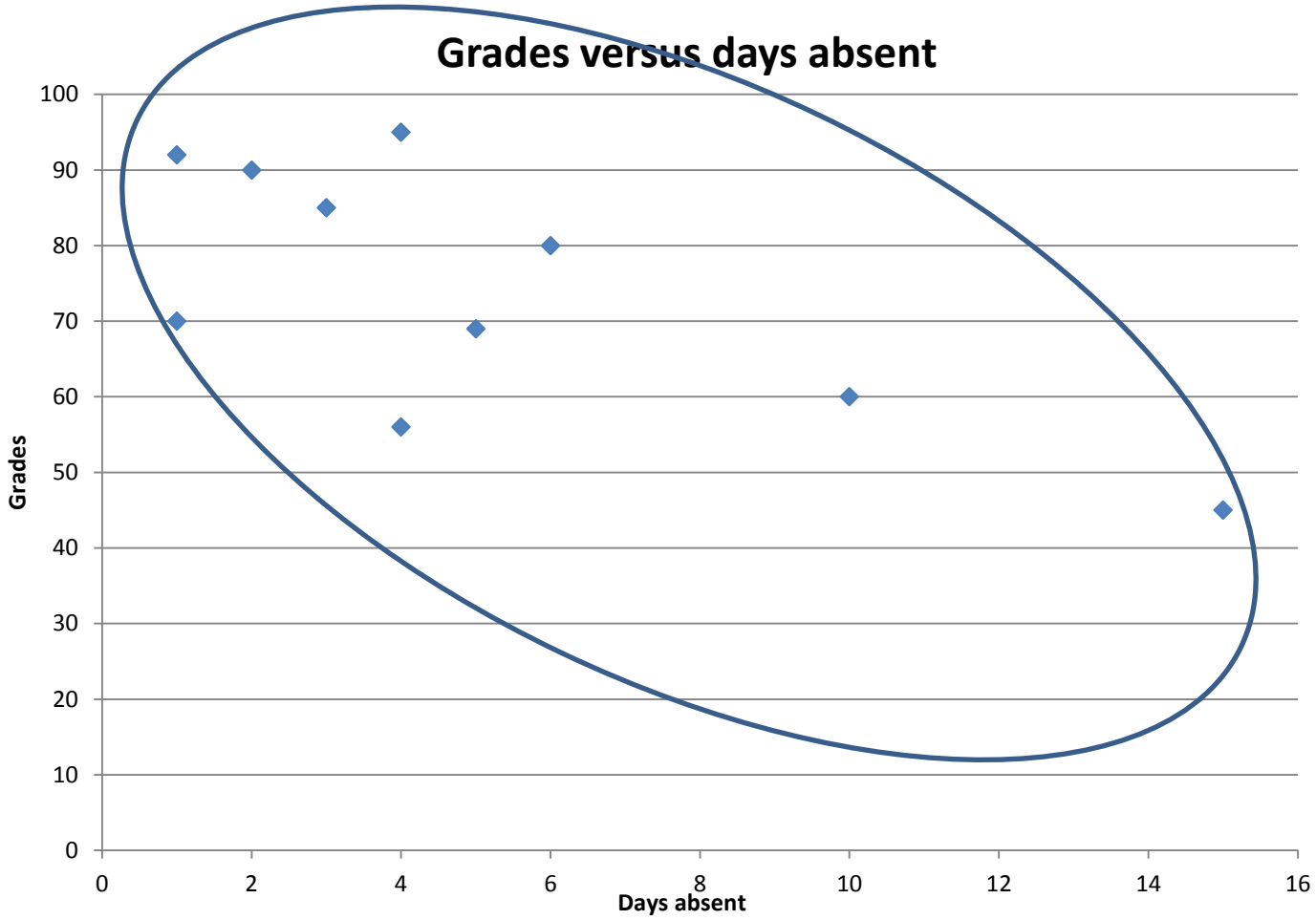
Scatter plot



Now that you have your scatter plot

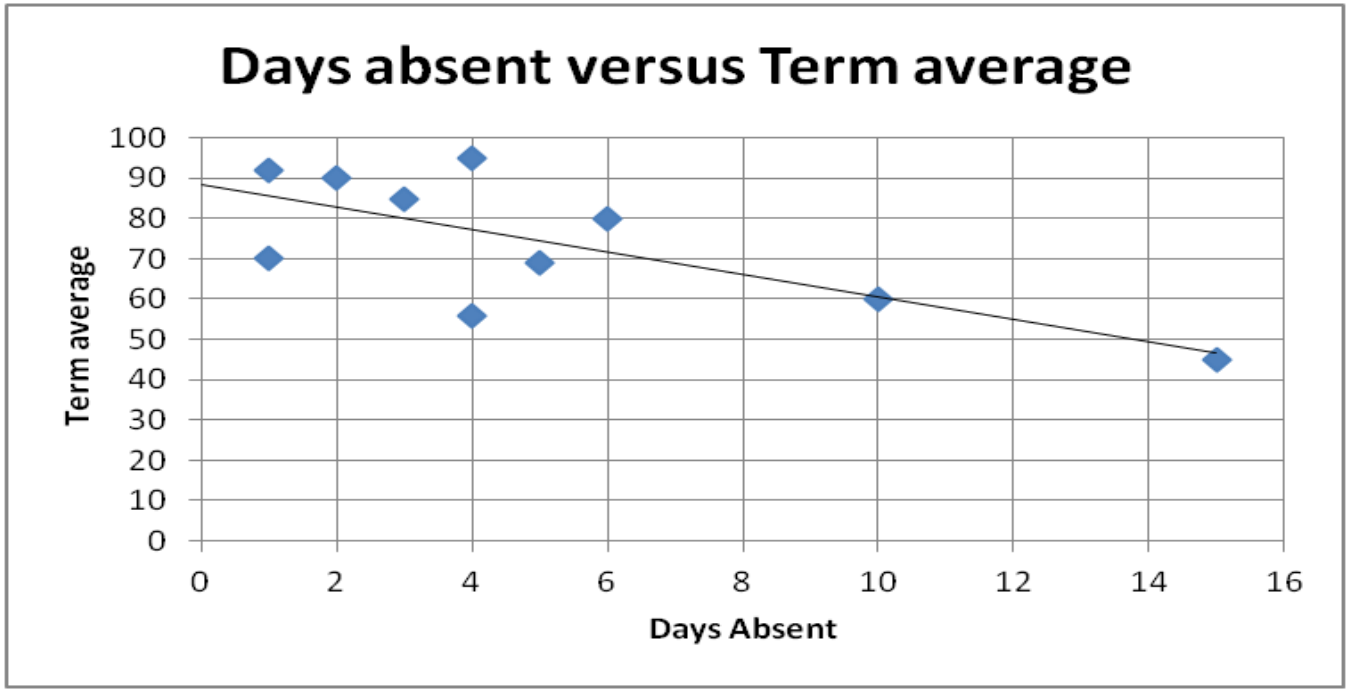
- Draw a line that represents to you the best fit for the data
- Calculate the slope of the line
- Calculate or draw the y –intercept
- Get your equation using Y= mx+b
- Make predictions based on your line

Here's the Scatter plot



Regression

Draw a regression line that represents the trend of the data on your graph.



Find the equation of your line

Use two points on your line to calculate slope:

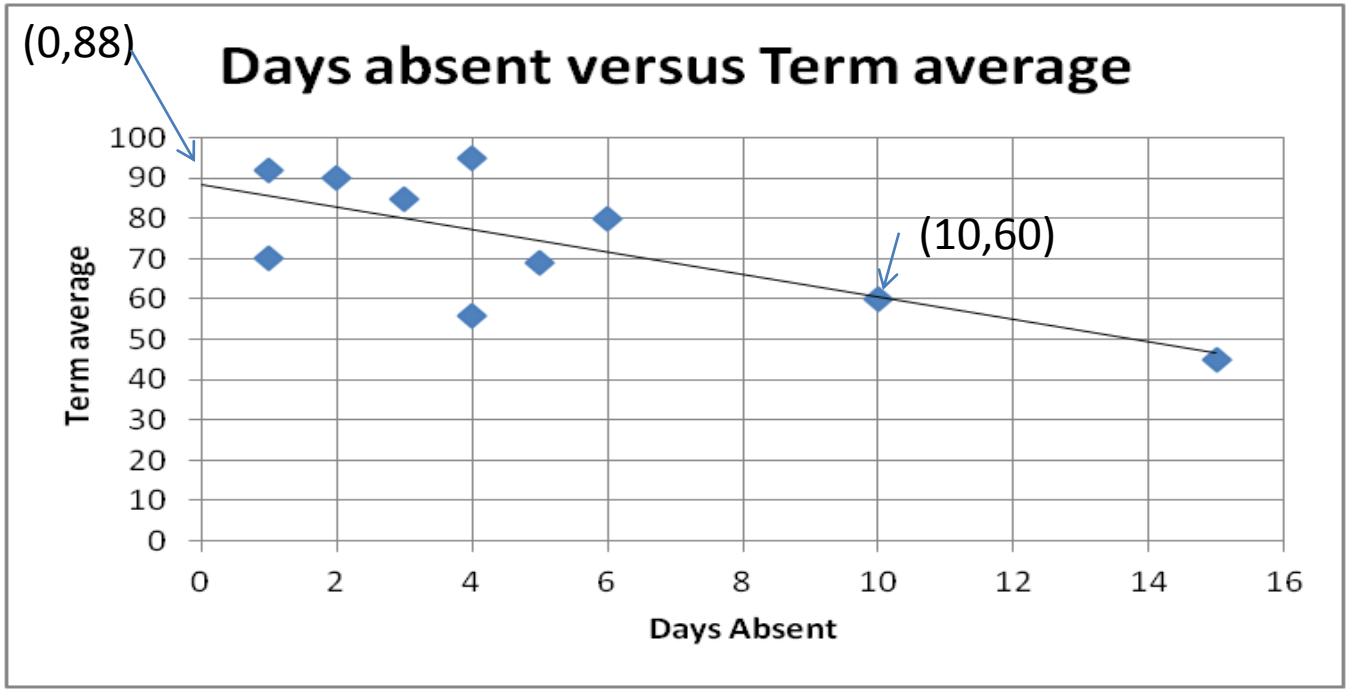
- Mine are (0,88) and (10,60).
- Change in y/change in x=

  $\frac{88-60}{0-10} = \frac{28}{-10} = -2.8$

- My y-intercept from my graph is 88, so my equation is:
- Y= -2.8X + 88

Regression

Getting the slope



What does my equation mean???

My equation: $Y = -2.8X + 88$

Explain in words what your slope means.

Key questions:

- What is x and what is y?
- What is slope?

What does my slope mean?

Key questions:

- What is x and what is y?
  - ✓X is the days absent
  - ✓Y is the term average measured in points
- What is slope?
  - ✓Slope is the change in y over the change in x, or
  - ✓The change in y for every 1 unit change in x

So, how do I put this together?

- My equation: Y= -2.8X + 88
- So, the term average (y) will decrease by 2.8 points for every 1 day (x) increase in absences.
- What about the y-intercept?
- What does that mean?
- The y intercept is the value of y when x=0
- So again, what is x and what is y?

Y intercept

- My equation: Y= -2.8X + 88
- X is days absent, y is term average
- So, the y-intercept means….

Y-intercept deep meaning revealed!!

- If there are no absences, the term average will be an 88%
- Let’s use our equation to make predictions

Predictions based on your equation

My equation: : $Y = -2.8X + 88$

Use your equation to predict a student's grade if they were absent for:

- 7 days
- 14 days
- 20 days

Predictions

- We are trying to find y when x (days absent) equals 7,14, or 20 days
- So use your formula and put these values in for x
- Y=-2.8(7)+88
- Y=-2.8(14)+88
- Y=-2.8(20)+88

Predictions

- We are trying to find y when x (days absent) equals 7,14, or 20 days
- So use your formula and put these values in for x
- Y=-2.8(7)+88=68.4
- Y=-2.8(14)+88=48.8
- Y=-2.8(20)+88=32

How confident should we feel about our predictions?

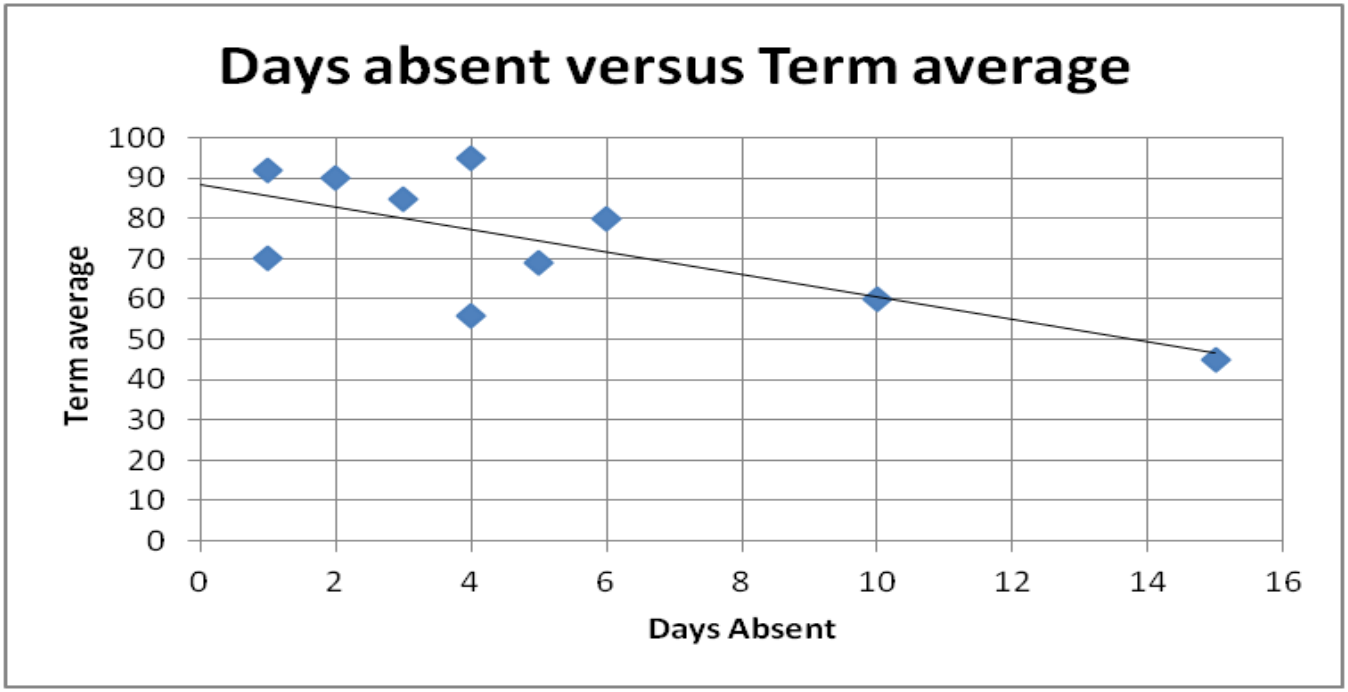
- That depends on whether we are interpolating or extrapolating
- Interpolating: Making predictions when we are inside known data values
- Extrapolating: making predictions when are outside known data values
- Which is more reliable?

Interpolation is more reliable

- So which of my predictions is/are interpolation, which is/are extrapolation?
- Y=-2.8(7)+88=68.4
- Y=-2.8(14)+88=48.8
- Y=-2.8(20)+88=32
- You need the graph again, you say?

Regression

Interpolation versus Extrapolation



7 days

14 days

20 days

Well, now we know

- The predictions for 7 and 14 are more reliable because we are interpolating between known data values, while
- The prediction for 20 days is less reliable because we are extrapolating beyond known data values

Regression

From your line, describe the correlation between the two variables:

Key questions:

Is the correlation positive or negative?

Is it strong, moderate, or weak?

Correlation

From your line, describe the correlation between the two variables:

Key questions:

Is the correlation positive or negative?

Negative

Is it strong, moderate, or weak?

Moderately strong

Does correlation mean causation?

- No!
- Could be other factors at work here

Pearson’s r value

- Number between -1 and 1
- Represents the correlations
- Circle the number that estimates Pearson's r value:

-.6 1.5 .1 .7 1 -1

Pearson’s r value

- Number between -1 and 1
- Represents the correlations
- Circle the number that estimates Pearson's r value:

-.6 1.5 .1 .7 1 -1

- Why?
- It’s negative, and it’s closer to -1 than 0 and we said it was moderately strong.

Exponential Growth and Decay

Linear Growth vs Exponential Growth

Characterized by a constant number.	**Characterized by a constant percent growth.**
Growth rate is called slope.	**Growth rate is expressed as a percent, and added to 100%. For decay rate, the decay rate would be subtracted from 100%**
Incremental change is added to the initial amount.	**The growth factor is multiplied by the initial amount.**

Linear Growth vs. Exponential Growth

Linear	Exponential
Equation looks like: y= (growth rate) X + initial amount	**Equation looks like: Y= initial amount (growth factor)^{n}**
$Y= mx +b$	$Y=A(1+/-r)^{n}$

Methodology for doing problems

Make a table

- Time in years, minutes, seconds whatever the unit is on the left
- Amount of whatever is increasing decreasing on right

Example of Population Growth

- A population is increasing at an annual rate of 4% per year. The population is now 40,000.
- What will be population in 10 years?
- What will be population in 25 years?
- How long until the population doubles?

Population problem Step 1

- Identify Growth rate: 4%
- Identify Growth factor: 104% or 1.04
- Set up table:

Set up population problem

Time in years	Population
0	40,000
1	$40{,}000(1.04)^1$
2	$40{,}000(1.04)^2$
n	$40000(1.04)^n$

Population problem

- Identify general equation:
- $Y = 40000(1.04)^n$
- Answer questions:
- Population in 10 years:
- $Y= 40,000(1.04)^{10} =$
- 59,209.77 rounds to 59,210

Answering questions

- Population in 25 years:
- $40{,}000(1.04)^{25}$ =
- 106,633.4533 =106,634
- How long until population doubles?
- Use the rule of 70 to help you get close

Rule of 70

The rule of 70 can help you get close to the number of years it will take to double when you are dealing with growth, or how long it will take to halve when you are dealing with decay.

Just take 70 the growth/decay rate expressed as a number, not a decimal

So for our example: $70 \div 4 = 17.5$

Rule of 70

- Let's plug our estimate of 17.5 years into our formula:
- $Y=40000(1.04)^{17.5}=79459.06$
- That's pretty close to 80000, right?
- The answer is actually 17.67 years, but approximately 17.5 years would be a good answer here

Decreasing population

- Population currently at 60,000
- Decreasing at rate of 4.5%
- What will the population be in 10 years?
- What will population be in 25 years?
- How long until population halves?

Population decreasing steps 1 & 2

- Identify Decay Rate:
- 4.5%
- Identify Decay Factor:
- How?
- 100%-4.5% =
- 95.5%
- What is this as a decimal?
- .955

Step 3 Set up table

Time in years	Population
0	60000
1	$60000(.955)^1$ =57300
2	$60000(.955)^2$ =54721
n	$60000(.955)^n$

Step 4&5

- What is our general equation:
- $Y= 60000(.955)^n$
- Answer questions
- What will population be in 10 years?
- $Y=60000\ (.955)^{10} = 37,860$
- What will population be in 25 years:
- $Y= 60000(.955)^{25} = 18977$

Answering questions population decrease

- How long until population halves?

- Let's use Rule of 70 to get close
- 70÷4.5=15.55…

- So let's plug in 15.5 to our equation:
- $Y=60000(.955)^{15.5}=29390$

Halving

- Close, but we've gone a little too far

- Try 15
- $Y=60000(.955)^{15}=30074$

- About 15 years is a good answer

Compounding Interest

What's our formula?

$Q=A(1 + r/n)^{nt}$

Q is the your ending amount

A is your original amount

r is your growth rate

Compounding interest

- n = number of compounding periods
- r/n is the growth rate adjusted for n compounding periods
- So, 1 + r/n is your growth factor
- t= time in years

Compounding

- What is our number periods (n) if we are compounding:
- Annually?
- 1
- Semi-annually
- 2
- Quarterly
- 4

Compounding

The beauty of interest

earning interest …

What if we start earning money

from our job in biomedical

research, and we invest $1000 in a very safe investment and it earns 2.1% annually.

Compounding

- What if that investment compounds interest annually. How much would we have after 4 years?
- What is our growth rate?
- What is this as a decimal?

Compounding

- What if that investment compounds interest annually. How much would we have after 4 years?

- What our growth rate?
- 2.1%
- What is this as a decimal
- .021

Compounding

- What amount would you have after 4 years and an annual rate of 2.1% investing $1000?
- Annual compounding?
- $Q=1000(1+.021/1)^{(4*1)}=1000(1.021)^4$
- $1086.68
- Quarterly compounding?
- $Q=1000(1+.021/4)^{(4*4)}=1000(1+.00525)^{16}$
- $1087.39

Comparison of interest over 4 years

Initial amount invested	Compounding period	Formula	Amount after 4 years
$1000	Annual	$1000(1+.021/1)^{(1*4)}$	$1086.68
$1000	Semi-annual	$1000(1+.021/2)^{(2*4)}$	$1087.15
$1000	Quarterly	$1000(1+.021/4)^{(4*4)}$	$1087.39
$1000	Monthly	$1000(1+.021/12)^{(12*4)}$	$1087.55
$1000	Daily	$1000(1+.021/365)^{(365*4)}$	$1087.63

Formulas to know

- Mean: $\frac{\text{Sum of all data values}}{\text{Total Number of Data}}$
- Median = data value in the middle when data is ascending/descending order
  - One value when there is an odd amount of data
  - Two values averaged when there is an even amount of data
- Mode = most frequently occurring data value

Formulas to know

- Confidence Interval: $p \pm 2\sigma$
- Equation of a line: $Y = mx+b$
  - Calculation of slope:

$$\frac{(Y_2-Y_1)}{(X_2-X_1)}$$

- Exponential Growth and Decay: $Y=A(1\pm r)^n$
- Compounding formula: $Y=A(1 + r/n)^{nt}$

Formulas to know

Standard Deviation Formula

$$\sigma = \sqrt{\frac{p(1-p)}{n}}$$

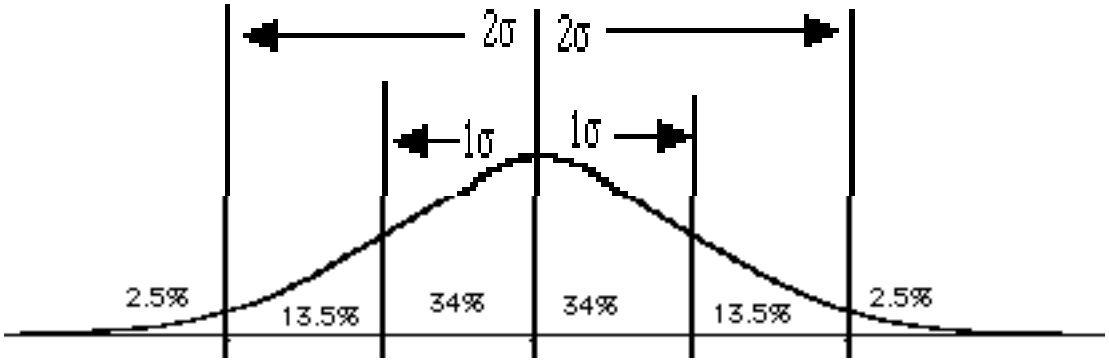
Formulas to know

Percent change formula:

$$\frac{\text{New} - \text{Old}}{\text{Old}} \times 100$$

Also…..

- Know the Rough and Ready rules:



Suggestions for Exam Prep

Review:

- your notes
- homework assignments and solution sets
- lecture videos, if needed
- final exam review video and slides

Practice writing and reciting essential formulas

Do the Practice Test to be posted under “Final Exam” module

On test day…

1. **The Final Exam is a closed book exam – no notes allowed.**

2. **Don't forget to bring your calculator, a few sharpened pencils, a straightedge and your photo I.D.**

   DISTANCE STUDENTS w/Proctor: be sure to PRESS HARD when writing your work and answers to ensure they are legible after scanning by the proctor.

3. **Scratch paper will be provided.**

Good Luck on the exam!!!

Professor Bird

Sue Jessica

Nina Lori

http://www.lz95.org/classes/mrsgibsonsfirstgradeclass_bgibson/