Simple Statistics in AP Psychology

Or how to use a “common language” so we can all understand the results

TOPIC 0.6 Statistical Reasoning in Everyday Life		
0.6-1 Describe descriptive statistics. **0.6-2** Explain how we describe data using three measures of central tendency and percentile rank. **0.6-3** Explain the relative usefulness ofthe two measures of variation. **0.6-4** Describe inferential statistics. **0.6-5** Explain how we determine whether an observed difference can be generalized to other populations.	• *Descriptive Statistics* • *Histogram* • **Measure of Central Tendency** • **Mode** • **Mean** • **Median** • **Percentile Rank** • *Skewed Distribution*	• **Measures of Variation** • **Range** • **Standard Deviation** • **Normal Curve** • *Inferential Statistics* • **Meta-analysis** • **Statistical Significance** • **Effect Size**

CONGRATULATIONS!
YOU'VE COLLECTED THE DATA

- Now what?
- Yup, Statistics…
- Don't freak out… it's not that bad… in fact, you know a lot of it already…
- You need to know the difference between:

DESCRIPTIVE STATISTICS

and

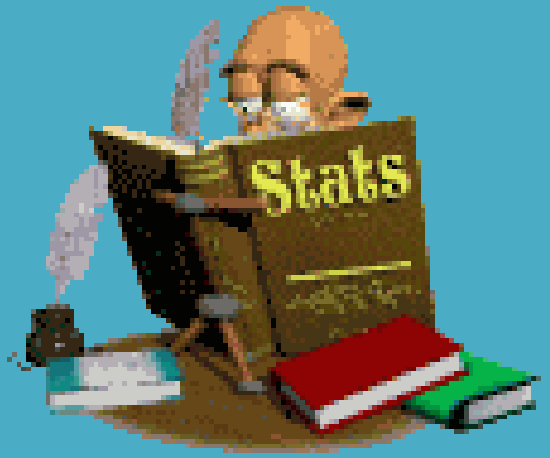
INFERENTIAL STATISTICS

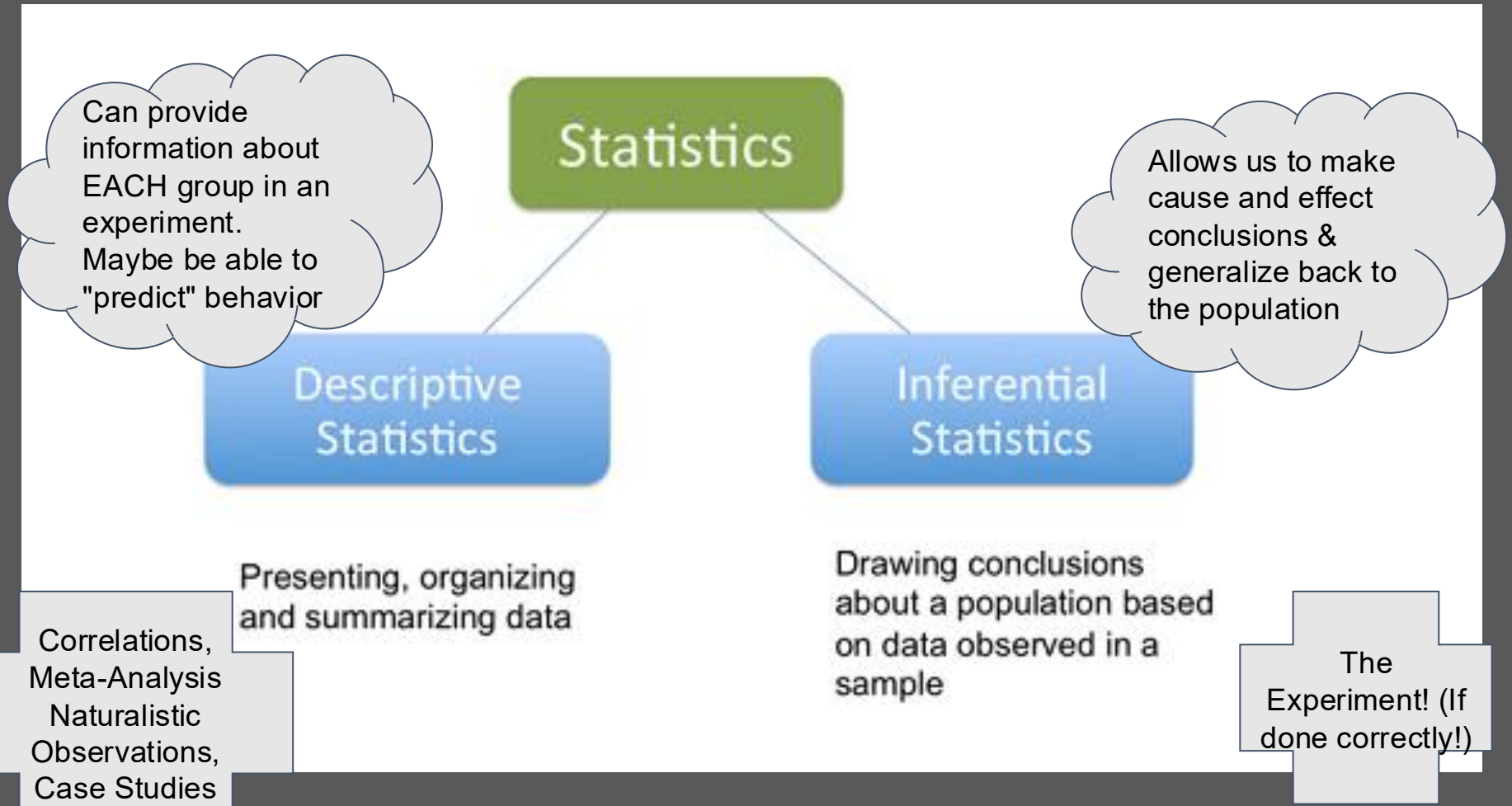
Module 0.6 Statistical Reasoning in Everyday Life

- STATISTICS!!!! Myers 4ed pages 0.40 to 0.46
- Don't barf… I promise it won't be that bad…
- And you only have one day to endure the pain…. Think about all those AP STATS students who must spend a WHOLE YEAR learning this stuff
- But seriously, this is important stuff to understand, so stick with me…
- Take a large gulp of coffee and let's get this party started!

Statistics

- Recording the results from our studies.
- Must use a common language so we all know what we are talking about.
- Statistical illiteracy creates needless health scares such as when it was reported that women had 100% increase in strokes from taking "the pill".
- Actual stat: Chances went from 1 in 7,000 to 2 in 7,000 (and probably resulted in 13,000 additional abortions).

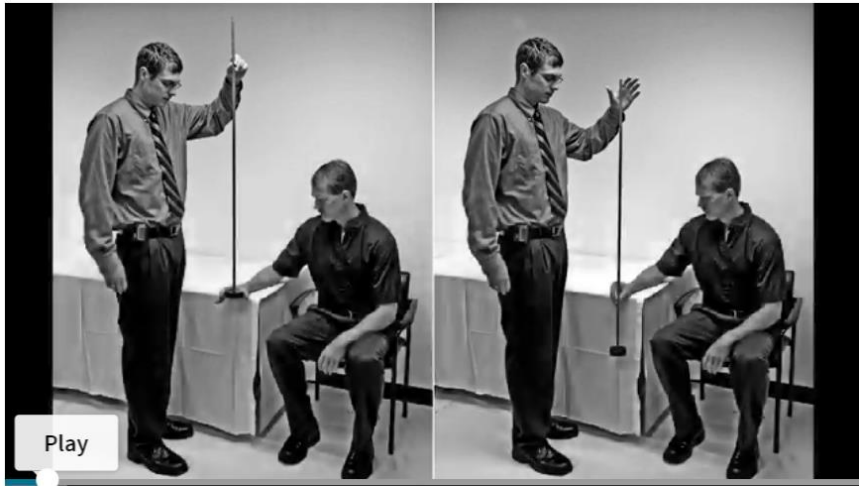



Statistics
Descriptive Statistics
Inferential Statistics
Can provide information about EACH group in an experiment. Maybe be able to "predict" behavior
Allows us to make cause and effect conclusions & generalize back to the population
Presenting, organizing and summarizing data
Drawing conclusions about a population based on data observed in a sample
Correlations, Meta-Analysis Naturalistic Observations, Case Studies
The Experiment! (If done correctly!)

DESCRIPTIVE STATISTICS

Remember, descriptive statistics DESCRIBES the data of your research.
In other words, it only summarizes the information about the SAMPLE you studied.

- Allows you to possibly make a **PREDICTION of what MAY happen,** but you cannot draw any conclusions about a population
- You **CANNOT** infer or determine causality of the results to a larger population from which the sample was selected.



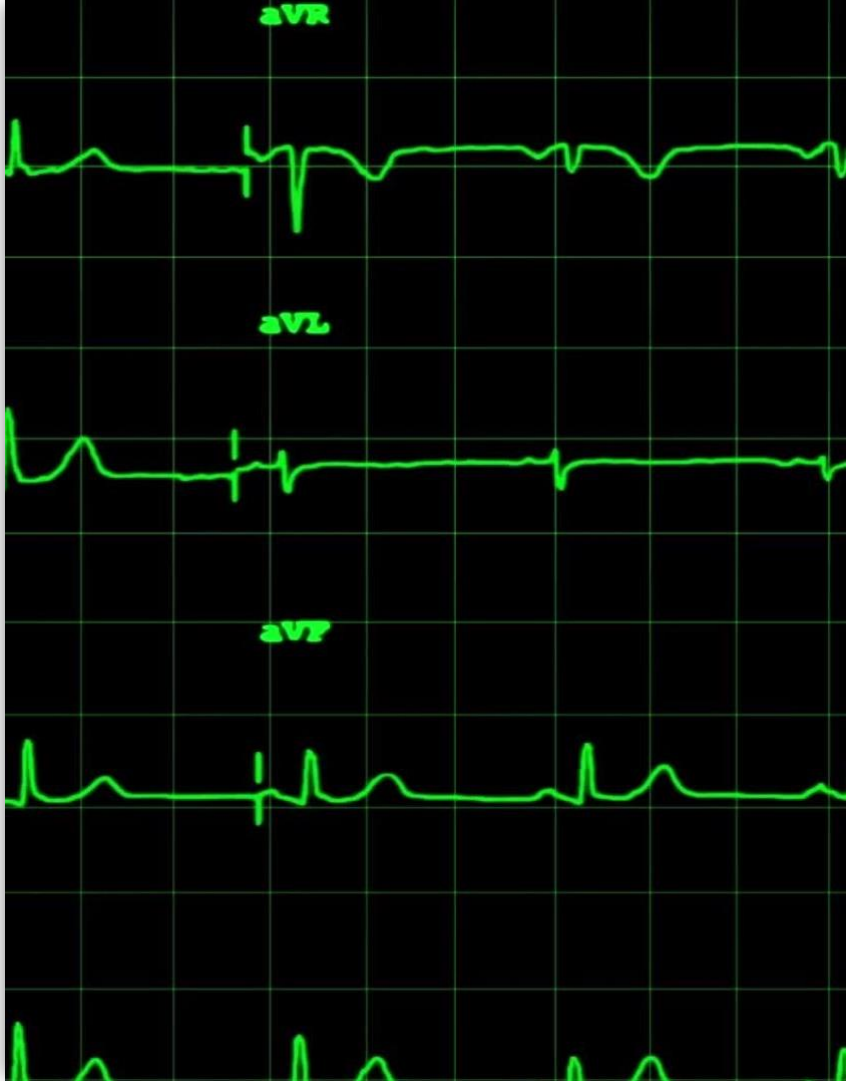
Example of how Descriptive Stats are used in Psych

- When research psychologists run a study, they collect a set of measurements as numbers.
- In cognitive psychology experiments, for example, a common measurement is the participant's reaction time, measured in milliseconds.
- Each participant's reaction time is referred to as their raw score.

TYPES OF DESCRIPTIVE STATISTICS

- **Include:**

  1. Correlations (see previous slides, been there, done that!)
  2. Histograms and Frequency Polygons (Bar graphs and line graphs)
  3. Measures of Central Tendency (Mean, Median and Mode. You know it, but I'll review it)
  4. Measures of Variability (Range, Variance and Standard Deviation, easy peasy)



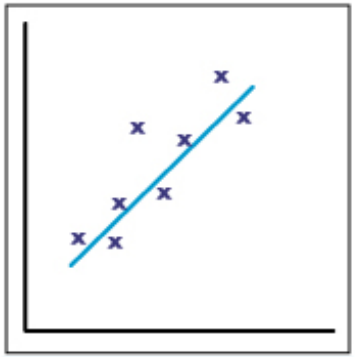
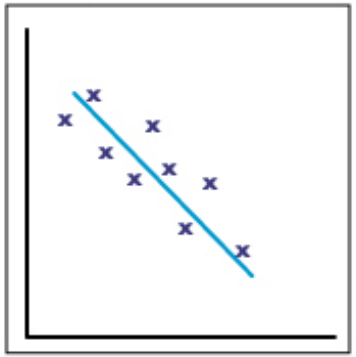
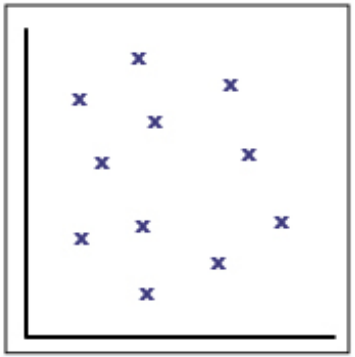
1. Correlation Coefficient

- The correlation "r" measures the strength of the linear relationship between two quantitative variables.
- Range is from -1 to +1
- The relationship gets weaker the closer you get to zero.

Which is a stronger correlation?

- r-.13 or r+.38
- r-.72 or r+.59
- r-.91 or r+.04

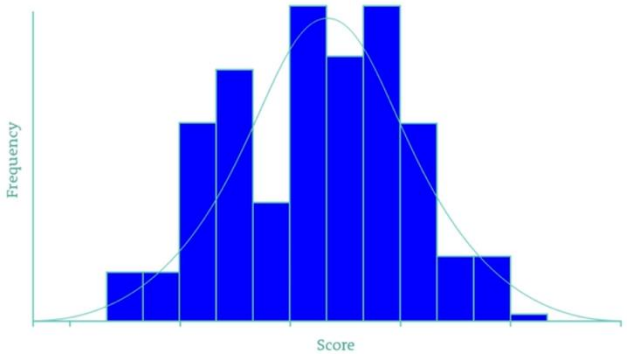
More on correlations...

Positive correlation	Negative correlation	No correlation
		
The points lie close to a straight line, which has a positive gradient.	The points lie close to a straight line, which has a negative gradient.	There is no pattern to the points.
This shows that as one variable **increases** the other **increases**.	This shows that as one variable **increases**, the other **decreases**.	This shows that there is **no connection** between the two variables.

Use "r" to signify a correlation

Example of how Stats are used in Psych

To understand these data and identify patterns, researchers often arrange the data in a **histogram plot**, from lowest to highest values. The histogram allows them to see, at a glance, how the scores are grouped, which helps them determine the results of the study.



However, when the data sets contain hundreds or even thousands of raw scores, psychologists calculate numbers, called descriptive statistics, that allow them to summarize important characteristics of the distribution.

These statistics allows us to describe the data in terms of (1) the center of the data and (2) the spread of the data. Picking the right descriptive statistic is key to interpreting data correctly.

Center of the data = use measures of central tendency

Spread of the data = use measures of variation

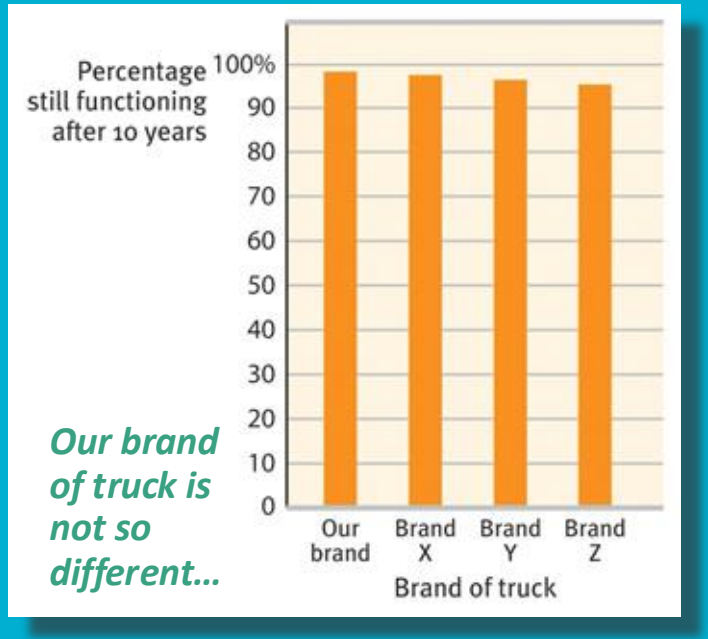
2. Histograms or Bar Graph

- Histograms **and** bar graphs
- Depicts frequency of a distribution
- Y axis is always vertical and represents the frequency you are **measuring**. (aka the *dependent variable*)
- X axis is horizontal and represents what you **are manipulating**. (aka *the independent variable*)

Tools for Describing Data

The bar graph is one simple display method but even this tool can be manipulated.

Is the data for both charts quantitative or qualitative?



Why is there a difference in the apparent result?

3. Measures of Central Tendency

IMPORTANT: **Measures of Central Tendency attempt to mark the CENTER of a distribution**. How do they do this? By using:

- Mean
- Median
- Mode

THINK about this… the measures of central tendency all look at how to measure the middle of a bunch of numbers… just three different ways to do it.

Measures of central tendency

Are you looking for just ONE NUMBER to describe a population's income, height, or age?

Options:

Mode

- the most common level/number/ score

Note—Bimodal means two modes

Mean

(arithmetic "average")

- the sum of the scores, divided by the number of scores

Median

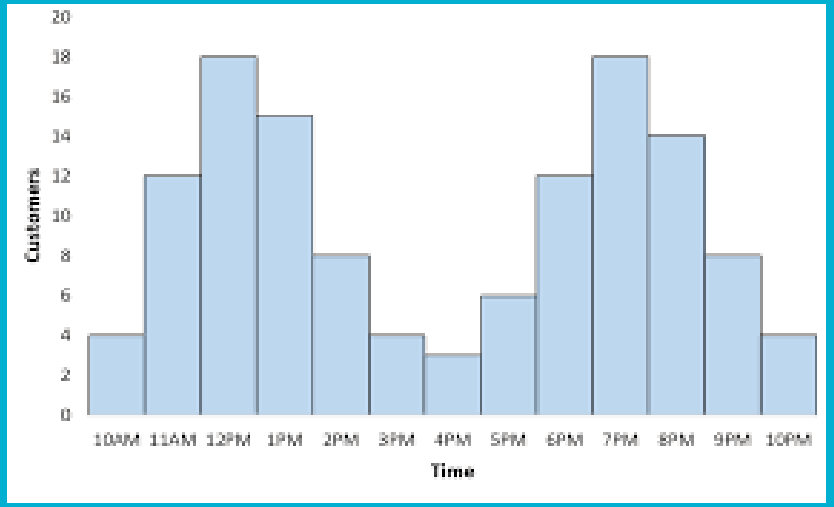
(middle person's score, or 50th percentile)

- the number/level that half of people scored above and half of them below

Bimodal distributions

- A set of scores with two peaks or modes around which values tend to cluster, such that the frequencies at first increase and then decrease around each peak.

- For example, when graphing the heights of a sample of adolescents, one would obtain a bimodal distribution if most people were either 5'7" or 5'9" tall.



4. Percentile Rank: the percentage of scores that are LOWER than a given score.

READ PERCENTILE SCORES FROM LOW TO HIGH!

If you line up all the scores and you fall in the 79th percentile, that means your score is HIGHER than 79 percent of your peers

	A	B	C
1	Products	Sales	
2	Product 1	$ 48,947	
3	Product 2	$ 42,781	85th percentile
4	Product 3	$ 42,306	
5	Product 4	$ 41,471	
6	Product 5	$ 40,556	66th - 85th percentile
7	Product 6	$ 38,802	
8	Product 7	$ 34,725	
9	Product 8	$ 27,916	33rd - 66th percentile
10	Product 9	$ 22,309	
11	Product 10	$ 18,576	
12	Product 11	$ 18,043	20th - 33rd percentile
13	Product 12	$ 14,430	
14	Product 13	$ 13,804	
15	Product 14	$ 13,651	20th percentile
16	Product 15	$ 13,469	
17			

5. Measures of Variability

IMPORTANT: Unlike measures of central tendency, measures of variability attempt to depict the DIVERSITY of the distribution. How do they do this? By using:

- Range
- Standard Deviation

THINK ABOUT THIS: measures of variability look at how far apart the numbers lay apart from each other

RANGE

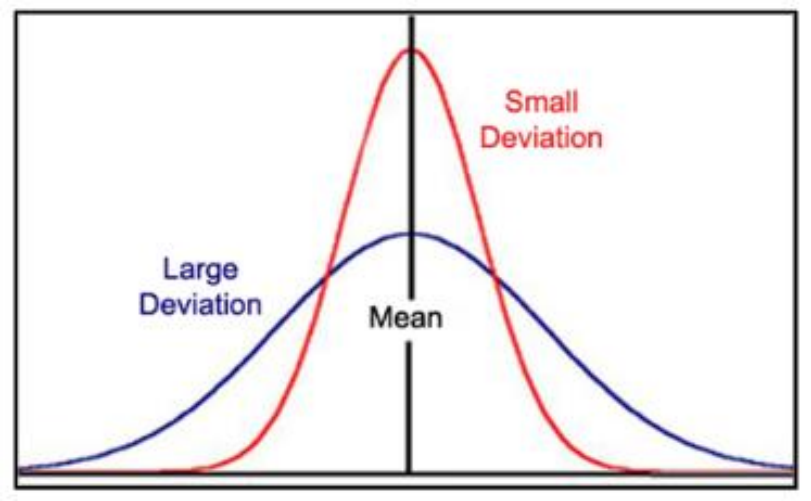
- Range is simply the distance between the highest and lowest score in a distribution

Example: 3, 5, 9, 15, 20 and 23.

The Range is 20 (23-3)

Standard Deviation

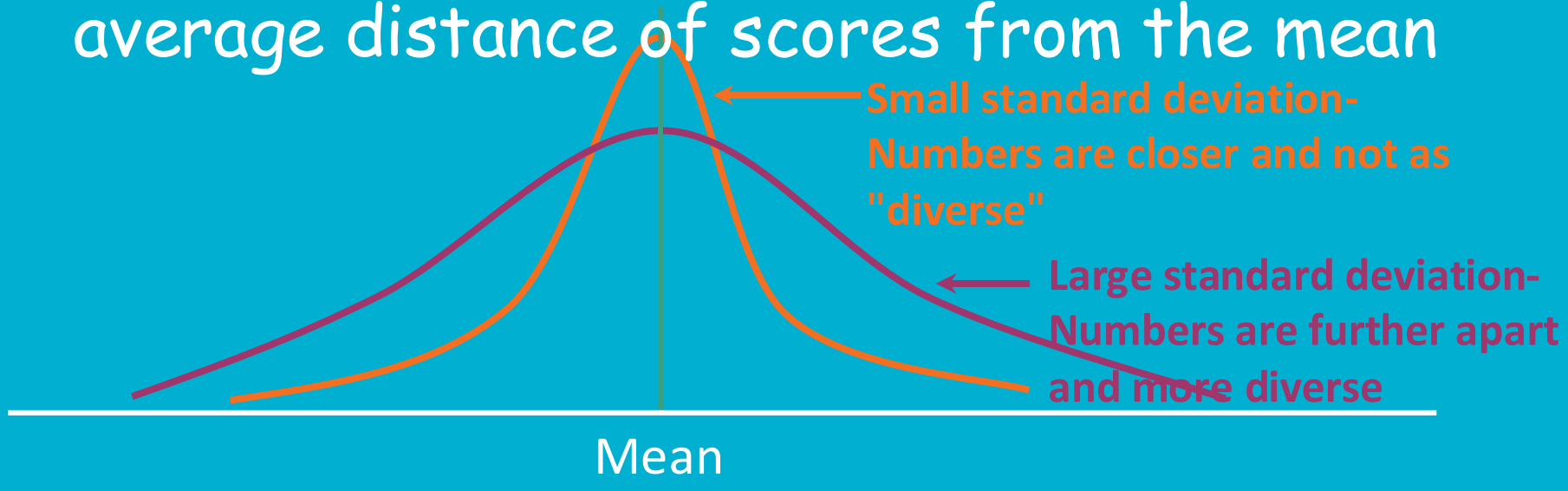
A computed measure of how much the scores vary around the mean. 0 means every score was the mean. The larger my std. Dev., the further apart the spread of my scores from the mean/avg.



Read scores from the CENTER moving out to see how far out the bell curve is—is it narrow or flat?

Measures of variation:
how spread out are the scores?

- **Range:** the difference between the highest and lowest scores in a distribution
- **Standard deviation:** a calculation of the average distance of scores from the mean



So…..quick review

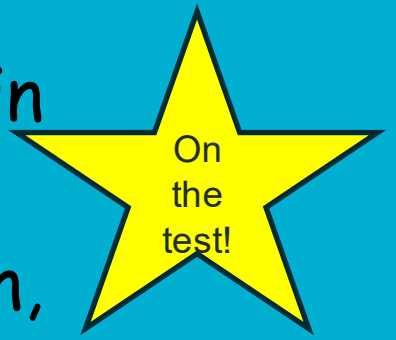
Measures of central tendency try to show how closely the numbers are clustered together (put your hands together for central tendency)

While…

Measures of variability try to show how diverse or far apart the numbers are (put your hands apart for measures of variability)

THE BELL CURVE

- A "normal" curve is a theoretical bell-shaped curve.
- We typically use the normal bell curve in intelligence testing.
- In this theoretical bell shape, the Mean, Median & Mode all fall in the center.
- The center is referenced to as the "50th percentile" or a Standard Deviation of 0



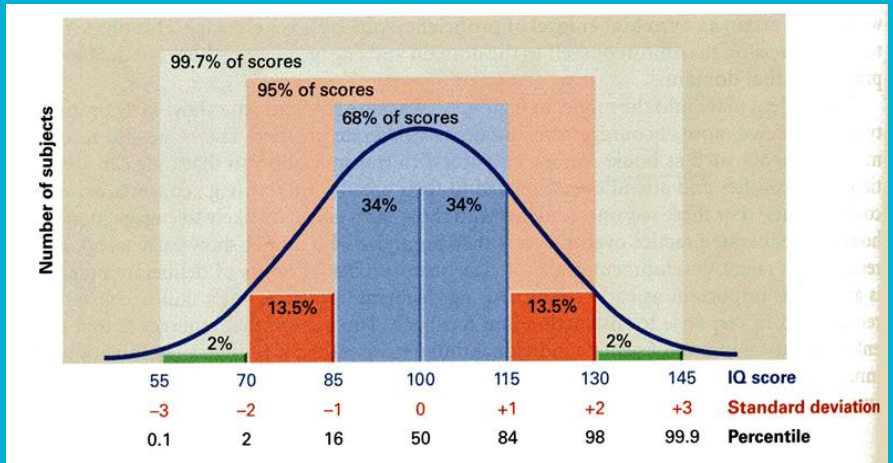
How to read the Normal Curve

Approximately 68% of scores in a normal distribution will fall within one standard deviation of the mean.

Approximately 95% of the scores fall within two standard deviations of the mean

Almost 99% of scores fall within three standard deviations of the mean.

On the test!

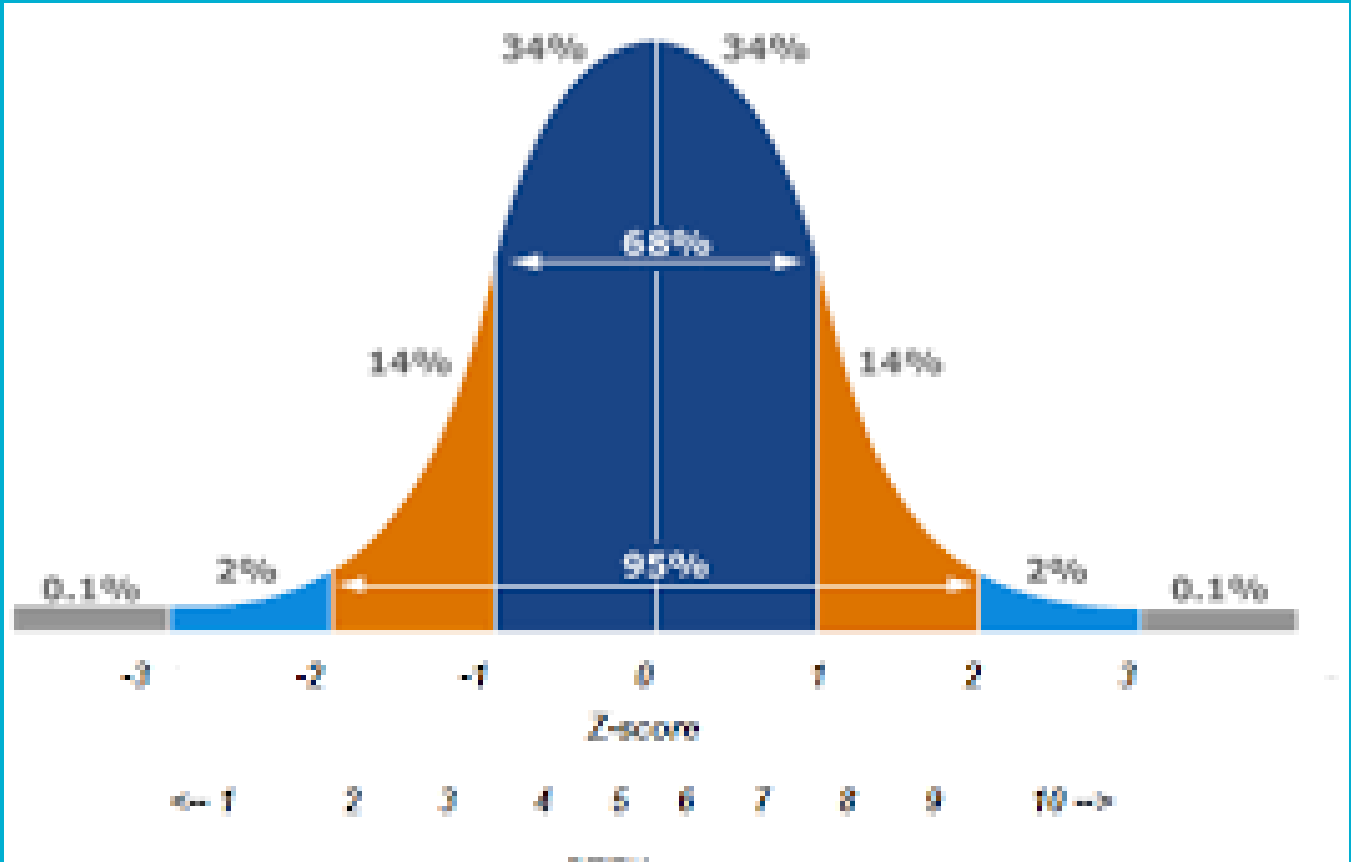


MEMORIZE THIS:

68-95-99

A 68-year-old man was driving along I-95 going 99 miles per hour

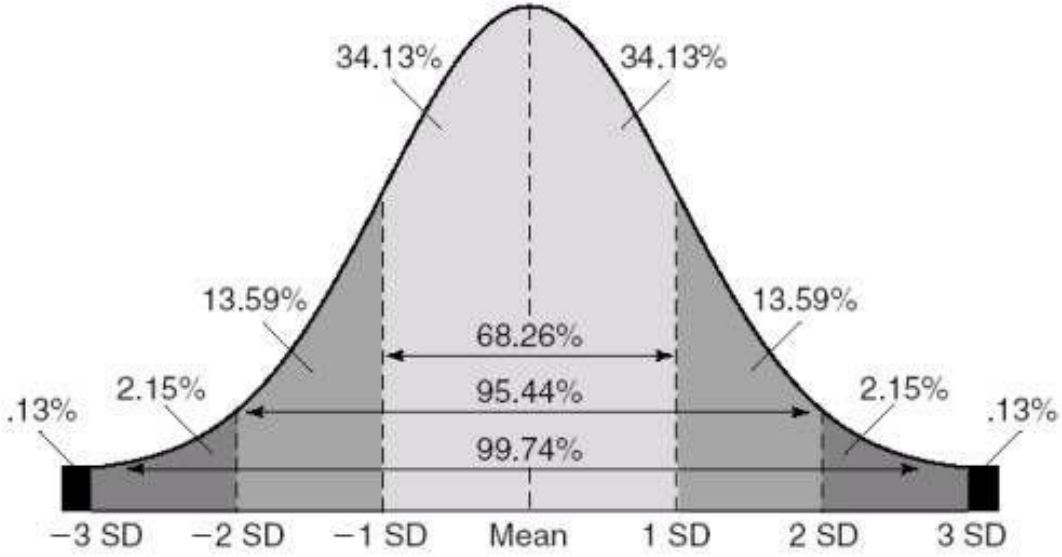
Normal Distribution



The Normal Curve and Comparing Scores

NOTICE that the SD scores move from the center to the sides

NOTICE that the percentile scores move from left to right

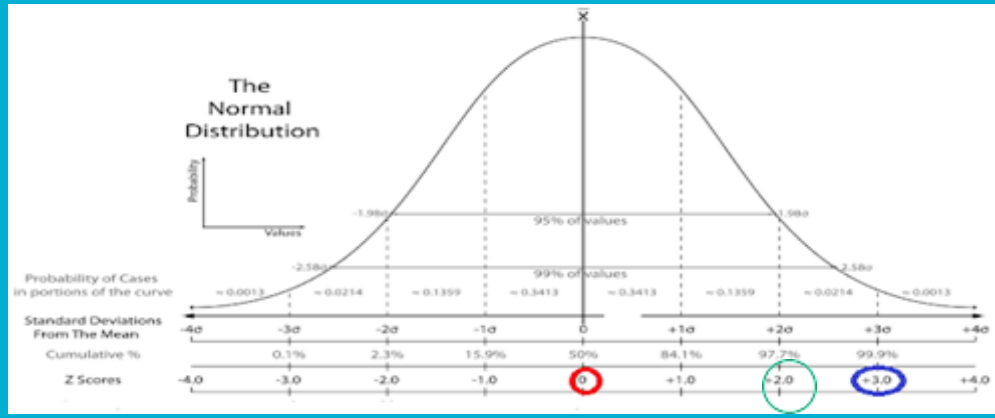


z-scores	−3	−2	−1	0	1	2	3
Percentile ranks	0.1	2	16	50	84	98	99.9
IQ	55	70	85	100	115	130	145
SAT	200	300	400	500	600	700	800

FIGURE 15.8 Percentile ranks and standard scores in relation to the normal curve. SD = standard deviation.

How to find the Standard Deviation (S.D.)

Knowing the normal curve is symmetrical and the three numbers (68-95-99) allows you to calculate the approximate percentages of scores falling between any given S.D. scores.



For example, 47.5 percent (95/2) of scores fall between the S.D. scores of 0 and +2 (and the converse is true for 0 and -2)

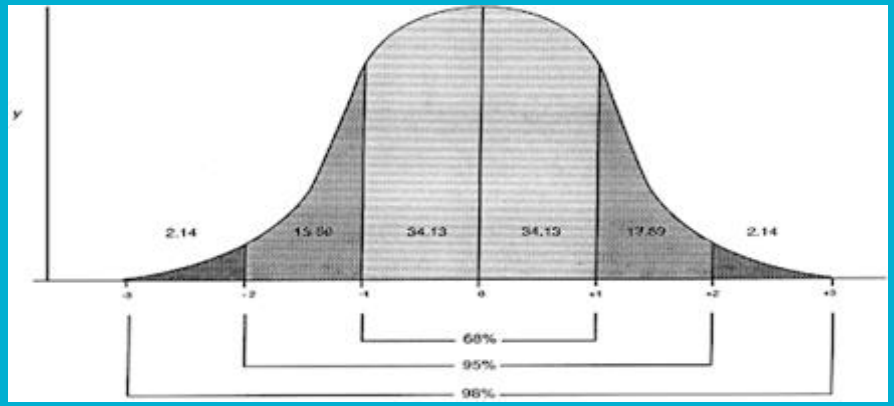
49.5 percent (99/2) of scores fall between the S.D. scores of 0 and +3 (and the converse is true for 0 and -3)

Examples for S.D. scores and the Bell Curve

I give a quiz and the mean (average) is 80 and the standard deviation is an 8.

A score of 72 would be a -1 S.D. score

A score of 84 would be a +0.5 S.D. score



Comparing S.D. scores and percentiles

While S.D. scores measures the distance away from the mean, percentiles indicate the distance of a score from 0 (like measuring on a ruler).

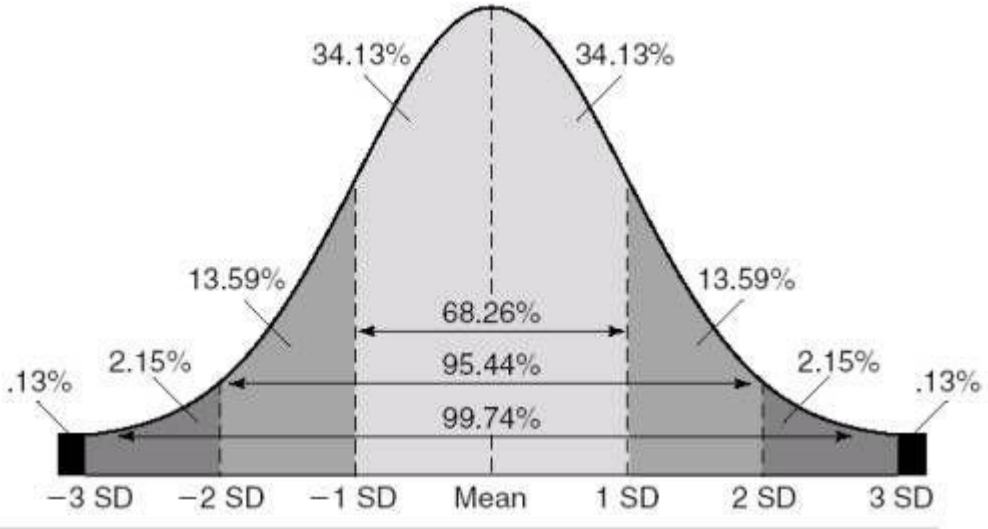
Someone who scores in the 90th percentile on a test has scored better than 90% of their peers.

S.D. scores count from the center point of the bell graph (usually range from 0 to 145 with 100 at center point)

Percentiles start all the way to the left or (at 0 and go up to 100 with 50 being at center point).

Comparing percentiles with S.D.scores

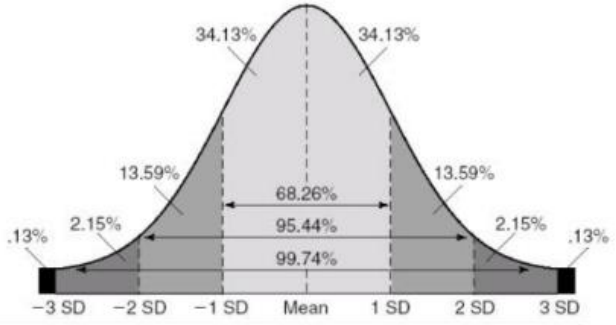
There is a relationship between percentiles and S.D.scores. Someone who scores at the 50th percentile has a S.D. score of 0 and someone at the 99th percentile has an approximate S.D.score of +3.



z-scores	−3	−2	−1	0	1	2	3
Percentile ranks	0.1	2	16	50	84	98	99.9
IQ	55	70	85	100	115	130	145
SAT	200	300	400	500	600	700	800

FIGURE 15.8 Percentile ranks and standard scores in relation to the normal curve. SD = standard deviation.

measurement, the concept of standard deviation is not relevant.



z-scores	−3	−2	−1	0	1	2	3
Percentile ranks	0.1	2	16	50	84	98	99.9
IQ	55	70	85	100	115	130	145
SAT	200	300	400	500	600	700	800

FIGURE 15.8 Percentile ranks and standard scores in relation to the normal curve. SD = standard deviation.

Understanding Test Scores*

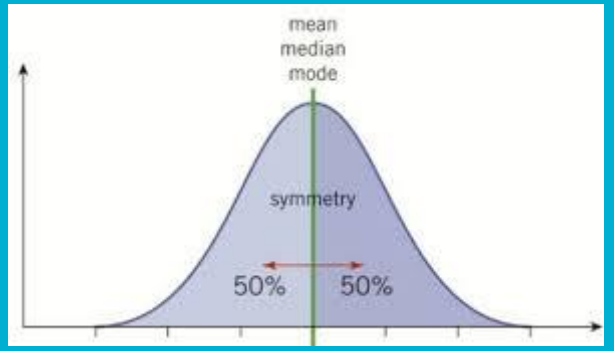
Table of Standard Scores, Percentiles, Z-Scores, T-Scores, Stanines, and Descriptions

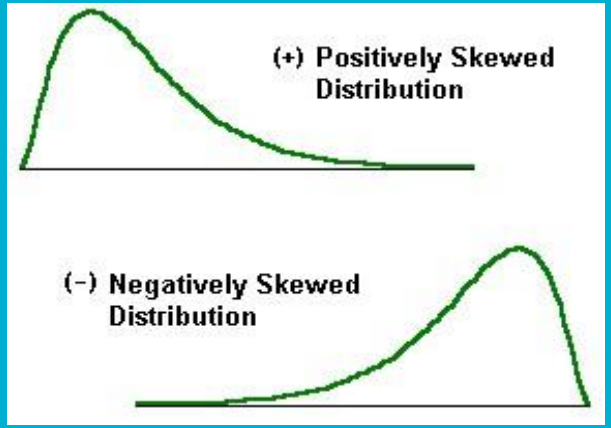
Standard Score Tests with average of 100 and SD 15	Percentile Rank	Z-Scores	T-Scores	Stanine approximate	Description
140	99.6	2.67	77	9	Very Superior
135	99	2.33	73	9	Very Superior
130	98	2.0	70	9	Superior
125	95	1.67	67	8	Superior
120	91	1.33	63	7	High Average
115	84	1.0	60	7	High Average
110	75	.67	57	6	Average
105	63	.33	53	6	Average
100	50	0	50	5	Average
95	37	-0.33	47	4	Average
90	25	-0.67	43	4	Low Average
85	16	-1	40	3	Low Average
80	9	-1.33	37	3	Borderline
75	5	-1.67	33	2	Borderline
70	2	-2.0	30	1	Low

What is meant by a positive S.D. score?
What does a negative S.D. score indicate?
What S.D. score would a person receive who scores at the 50th percentile?
What is the S.D. score for someone whose standard IQ score is 130
If someone's S.D. score is 1.67, what is their percentile score? Their z-score?
How many standard deviations is the S.D. score away from a standard score of 100?
Which score is higher? A standard score of 120? A percentile rank of 95? A z-score of 2.0

Skewed Distributions

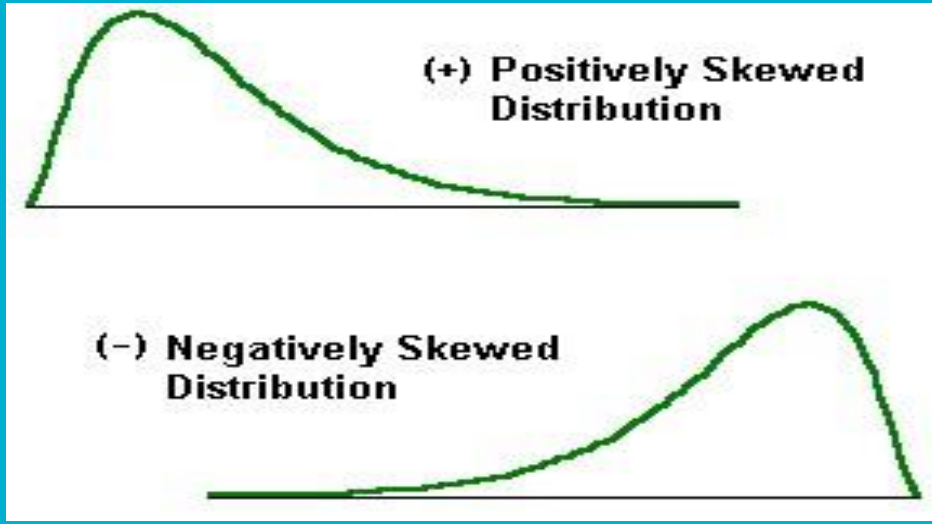
- A "normal bell curve" is symmetrical
- A "normal bell curve" has mean, median and mode all lined up in the center
- But sometimes, we have a few very high or very low numbers in a distribution.
- When we have a few very high or very low numbers, it will "distort" the normal bell curve.
- Those very high or very low numbers are called "outliers"
- The distortion of the bell curve creates "a skewed distribution"





Skewed Distributions

- Outliers skew distributions.
- If a group has one high score, the curve has a positive skew (contains more low scores and mean is located to the right of the median)
- If a group has a low outlier, the curve has a negative skew (contains more high scores and the mean is located to the left of the median)



HINT: Ask yourself-"Where is the tail?". That is how you determine whether the skew is positive or negative.

Skewed Distributions

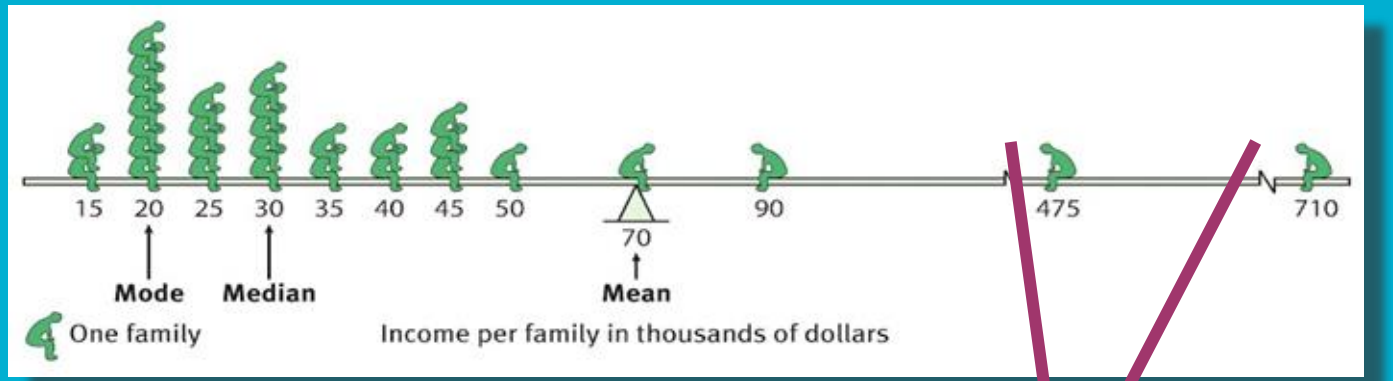
A distorted distribution; NOT evenly distributed around the mean.

Positively skewed means more people have LOW scores.

Negatively skewed means more people have HIGH scores

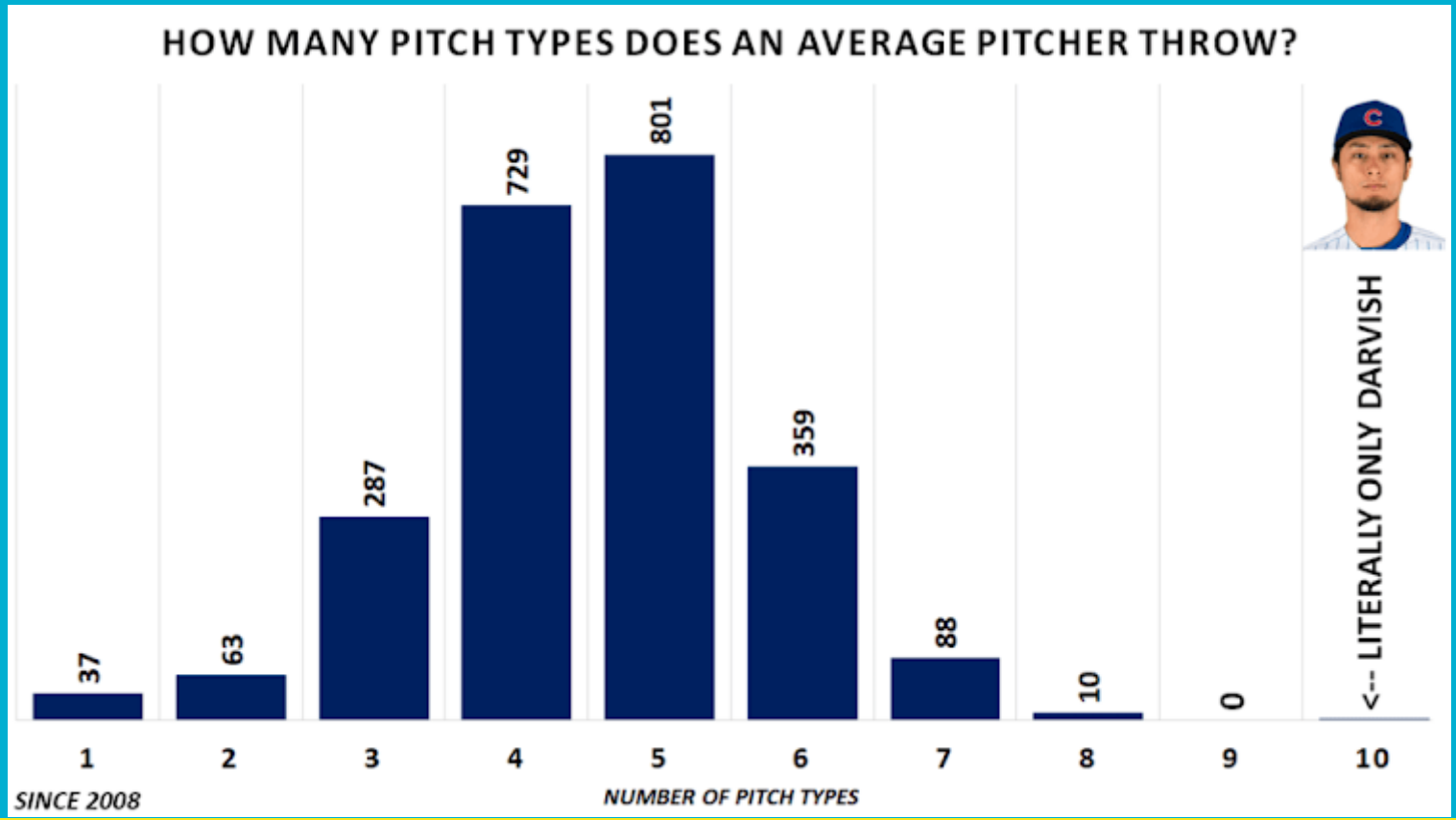
Measures of central tendency

Here is the mode, median, and mean of a family income distribution. Note that this is a **skewed** distribution; a few families greatly raise the mean score.

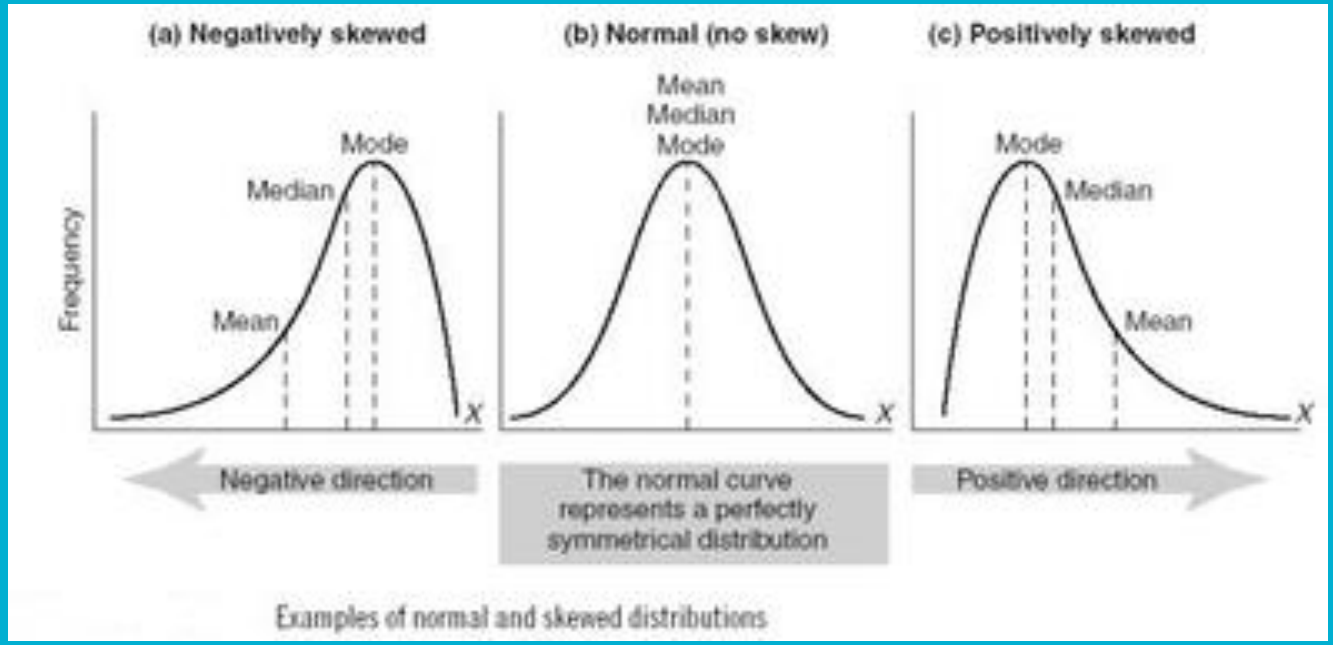


Why does this seesaw balance?
Notice these gaps?

Based on this curve, what kind of skew is it?



Notice location of mean, median and mode with a skewed curve



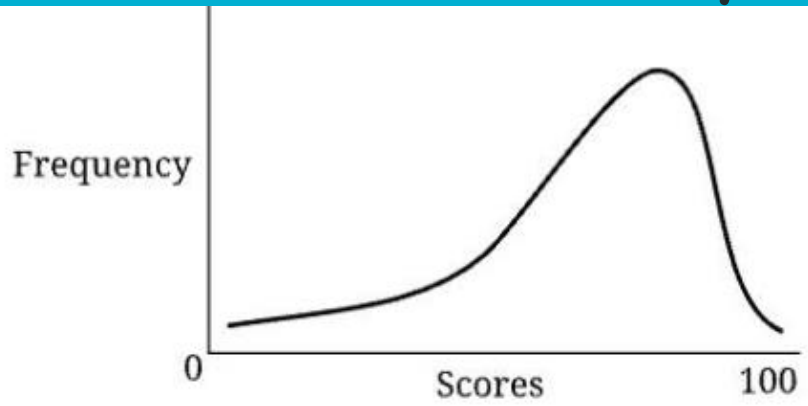
The mean and skewness

If you see a situation with a skewed curve, do NOT use the mean.

It is MEANINGLESS! (at least for a skew)

The outliers pull the mean toward the tail, so the Mean is being "mean" by pulling the tail… use the MEDIAN instead (it's nicer!)

Measures of Central Tendency



The graph represents scores on an exam about the influence of groups on individual behaviors.

Which of the following is a measure of central tendency that best represents the type of distribution represented in the graph?

a. Mode
b. Mean
c. Standard Deviation
d. Median

The Median in a skew is the best measure of central tendency!

- It IS a skewed distribution, but that's not what they're asking.
- The student would have to recognize the type of distribution first, and then be able to identify what measure of central tendency is best when assessing data from this type of distribution - and the answer is median.
- The low outliers would drastically affect the mean, and not affect the mode at all. So median is the best Measure of Central Tendency
- The median is the measure of central tendency that is least affected by outliers in the data and therefore best represents the data.

More practice . . .

Which of the following is the measure of central tendency that would be most affected by a few extreme scores?.

- A mean
- B range
- c) median
- d) mode
- e) random sample

The four families on your block all have annual household incomes of $25,000. If a new family with an annual income of $75.000 moved in, which measure of central tendency would be least affected?

- A mean
- B median
- C mode
- D standard deviation

Regression toward the mean

In statistics, regression toward the mean is the phenomenon that arises if a sample point of a random variable is extreme, a future point is likely to be closer to the mean or average

It's just a fancy way of saying that in any event where luck is involved, extreme outcomes are followed by more moderate ones.

Descriptive Statistics: Group work

01E Understanding Descriptive versus Inferential Stats and Statistical Significance

1. Write down today's phone usage (in minutes) of each person at your table and at least one other table.
   - Settings
   - Screen time
   - Convert to minutes

2. Sort the numbers: arrange them in order from lowest to highest

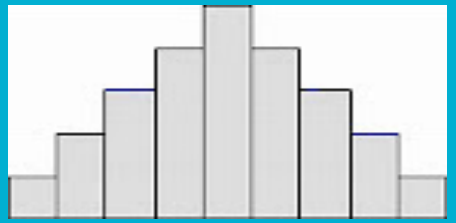
3. Create a frequency histogram

X axis = "students"

Y axis = "cell phone usage"

# of bins = columns you will need for y axis

https://statscharts.com/bar/histogram



Descriptive Statistics: Group Work

Measures of Central Tendency

- Is the data positively or negatively skewed?
- What is the mode of your distribution?
- What is the median of your distribution?
- What is the mean of your distribution?
- Which measure of central tendency would be most useful in the analysis? Why?

Measures of Variability

- What is the range of your distribution?
- What is the standard deviation?
- What does this standard deviation tell you about phone usage among our peers?

STANDARD DEVIATION CALCULATOR

But first a warm-up . . .

- Your favorite!
- Let's think statistically!

The Living Distribution

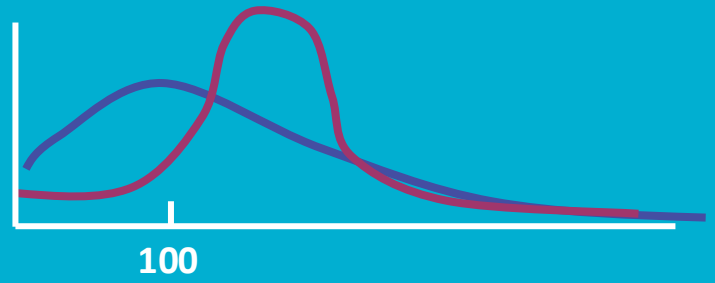
- Group activity!
- Split into "short" and "tall students"

Applying the concepts

Try, with the help of this rough drawing below, to describe intelligence test scores at a high school and at a college using the concepts of range and standard deviation.

Intelligence test scores at a high school

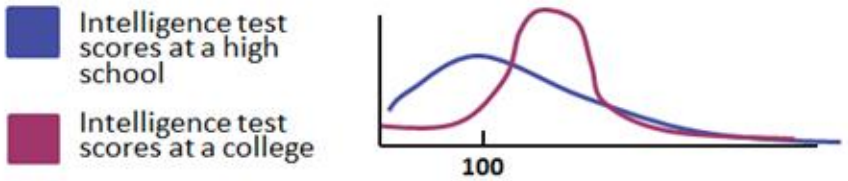
Intelligence test scores at a college



Notice that in this fictional example, the range is the same, but the mean is different. More importantly, the standard deviation is smaller at a college. Possible explanation: there is likely to be a narrower range of intelligence test scores at a college than at a high school, because at a given college, people with lower intelligence test scores might not have the SAT/ACT scores and grades to be accepted, and people with higher intelligence test scores might have the SAT/ACT scores to apply to a college with a higher median student ability level.

Applying the concepts

Try, with the help of this rough drawing below, to describe intelligence test scores at a high school and at a college using the concepts of range and standard deviation.



Why different intelligence scores at a high school and college?

High School should have a normal bell curve as it should reflect the general population

But colleges have an admission program where they SELECT their population. The better the school, the more skewed their population will be… which way will the population be skewed?

Negatively skewed….

Descriptive Analysis

Creates reports and graphs that provides information that **describes or summaries that data.**

Inferential Analysis

Collects data from a sample and **draws conclusions** about the population from the sample.

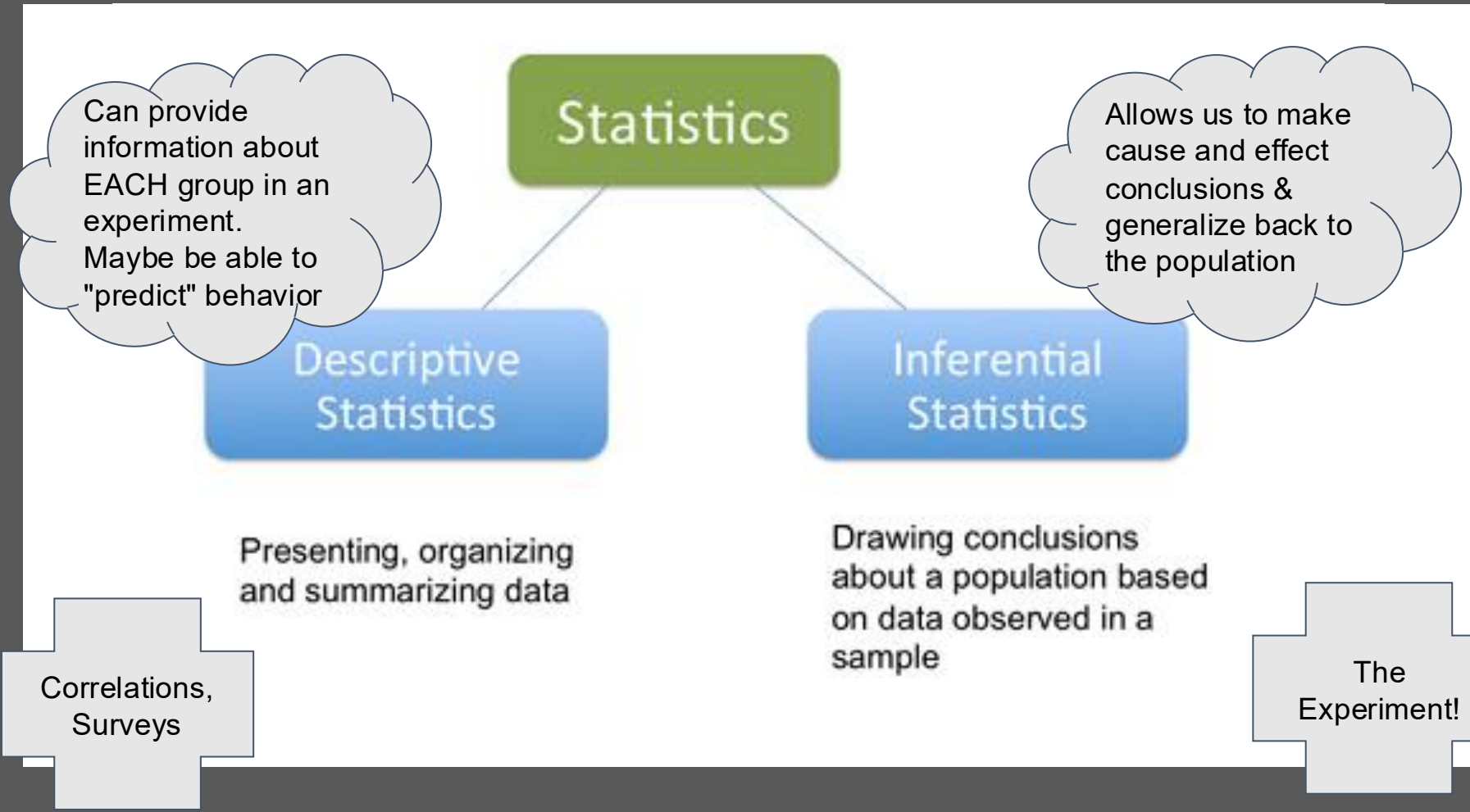
INFERENTIAL STATISTICS!!!

Descriptive Statistics (measures of central tendency and measures of variability) provides a way to summarize (or describe) the information about the sample studied…

Inferential Statistics is to determine whether or not the findings can be applied back to the larger population from which the sample was selected to draw a conclusion.

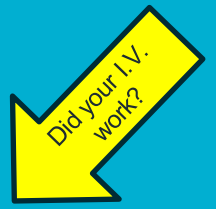
In other words, *inferential statistics allows you to* apply the results of your research (Causation)

TEXTBOOK: numerical data that allow one to generalize—to infer from sample data the probability of something being true of a population


Statistics
Can provide information about EACH group in an experiment. Maybe be able to "predict" behavior
Allows us to make cause and effect conclusions & generalize back to the population
Descriptive Statistics
Inferential Statistics
Presenting, organizing and summarizing data
Drawing conclusions about a population based on data observed in a sample
Correlations, Surveys
The Experiment!

INFERENTIAL STATISTICS

- Applying the results of your research to the larger population from which the sample was collected.
- These statistics take into account:
  - (1) the magnitude of the difference found between your control and experimental groups and
  - (2) the size of the sample.



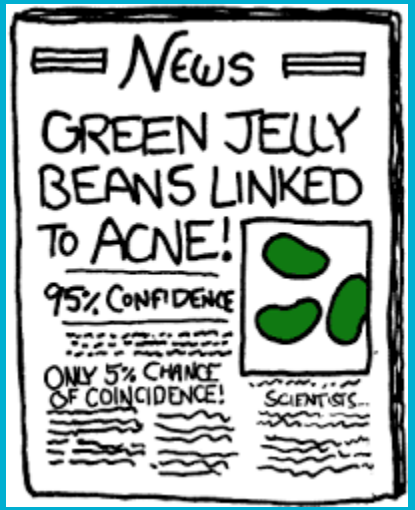
NOTE—you will sometimes be able to tell when you are in inferential stats because you will see a "p-value". If the p value is .05 OR LESS, it is considered statistically significant. It means there is a 5% chance or less that the results of your study (from the manipulation of the independent variable) were due to chance. Due to chance is another way to say "it was just luck"

VERY IMPORTANT SLIDE!

Statistical Significance

- A statistical statement of how likely it is that a result, such as a difference between the samples occurred by chance alone, assuming there is no difference between the populations being studied (which is "bad"- meaning the research did not work/theory is wrong).

- We want the differences to be the result of the manipulation of the I.V., not by chance (which is good, meaning your research worked/theory is correct!).



Statistical Significance (Memorize and Understand this slide!!!)

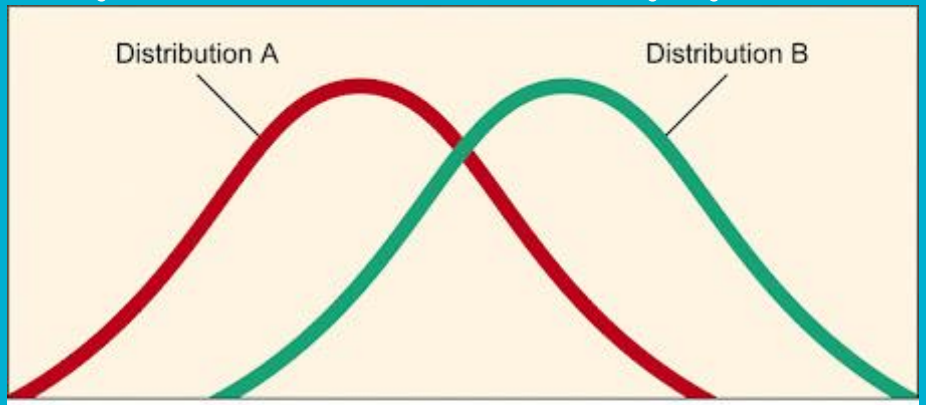
- To determine whether a research finding is statistically significant, **researchers <u>compare</u> the control group to the experimental group.**
- **When the observed difference between these two groups is large, it is statistically significant.**
- **Thus, the results occurred due to your manipulation of the independent variable and NOT a confounding variable!**
- What helps to make a finding statistically significant more likely is a LARGE SAMPLE SIZE! It allows you to GENERALIZE the results back to the population you claimed to be studying!

Significance Difference

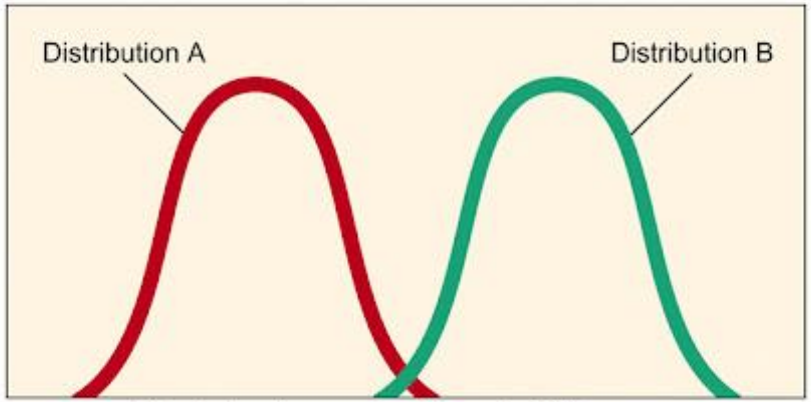
Group A:
Control
Group B:
Experimental

Just by looking at these two groups, there is a lot "in common"; they seem to share a lot of the same results in the overlap.

When you look at the results of these two groups, there is hardly any overlap. They don't share much in common. Your manipulation of the I.V. looks like it worked



Not likely to be a significant difference



Likely to be a significant difference

RESULTS **NOT** STATISTICALLY SIGNIFICANT.
Overlap between 2 groups is very large.
P value greater than .05

RESULTS ARE STATISTICALLY SIGNIFICANT.
Overlap between 2 groups is very small.
P value less than .05

How can we "generalize" the results from our experimental group to the population?

Inferential Stats allow us to determine if the observed differences in our experimental and control group are reliable if there are three things. ..

1. **A representative sample:** The selected sample should look like the population. It should not be biased (unrepresented) nor a convenience sample—AAQ!!!
2. **A large sample:** Averages based on many cases are more precise than based on a few and make it more replicable to study (and hopefully find a similar estimate next time) Think of the crayons
3. **More estimates better than fewer:** Conduct multiple studies and combine them using meta-analysis

Generalizability:

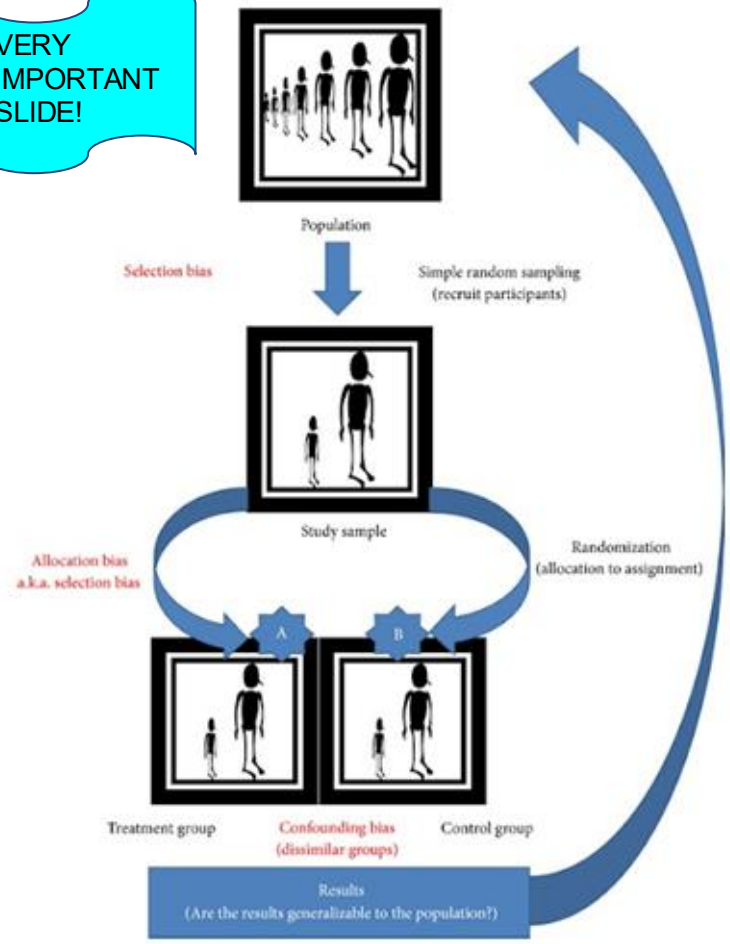
Can the results of your study be applied to a population larger than or different from the particular sample?

For a well-designed experiment with a LARGE ENOUGH Random Sample we can make an inference about the population.

With Random assignment, we can infer there is cause and effect (showing there is a large difference in outcomes between the control and experimental group)

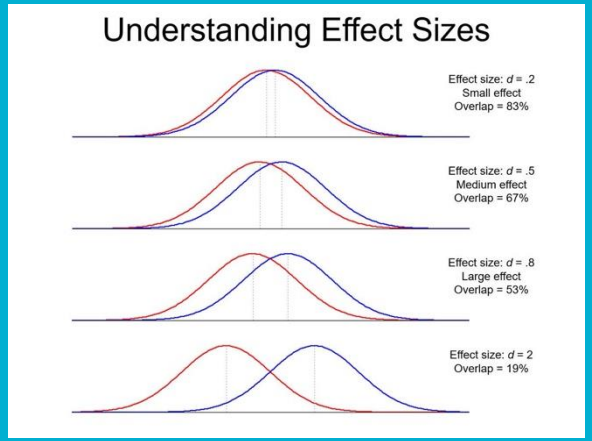
YES!

VERY IMPORTANT SLIDE!



Effect Size

- Just because you have "statistical significance", does not mean you have "practical significances"
- EFFECT SIZE is the strength of the relationship between the two variables. The larger the effect size, the more one variable can be explained by the other.
- Example: researchers found birth order to be statistically significant for intelligence. But the difference in intelligence scores was just 1.5 I.Q. points, so IQ is probably determined by factors other than just birth order.



Inferential Statistics: Statistical Significance

01E Understanding Descriptive versus Inferential Stats and Statistical Significance

A statistically significant difference is unlikely to be due to chance.
Three factors contribute to whether a difference between groups will be statistically significant:

- the size of the difference between the group means
- the size of the sample
- the variance within the groups
  (small variance within each group is better)

Statistical Significance

Melissa is running a study to see if girls and boys average different amounts of participation in classrooms. She hypothesizes that girls participate more than boys and plans to observe students in various classes and record how often they raise their hands to answer the teachers' questions.

Assignment

01E Understanding Descriptive versus Inferential Stats and Statistical Significance

Statistical Significance

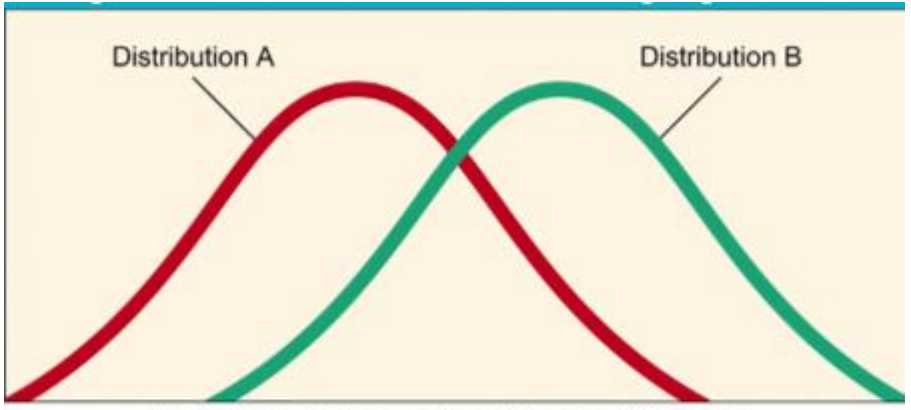
Why does the small sample size in the first example increase the likelihood that Melissa's results are due to chance?

1. For the first part of her data collection, Melissa selects two boys and two girls to study. She finds that the girls raise their hands an average of 4.7 times per week while the boys only raise their hands an average of 1.3 times in the same classes. Do you think this difference is likely to be statistically significant? Why or why not?

Statistical Significance

Why does the small difference between boys and girls in the second example increase the likelihood that Melissa's results are due to chance?

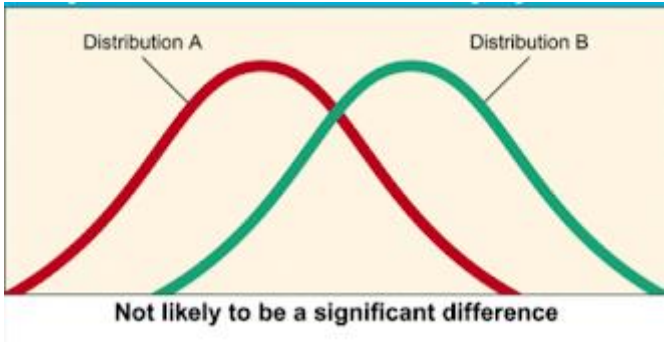
2. Next, Melissa expands her study to observe 50 boys and 50 girls. She finds that girls raise their hands an average of 3.1 times per week and boys raise their hands an average of 3.0 times per week in their social studies classes. Do you think this difference is likely to be statistically significant? Why or why not?



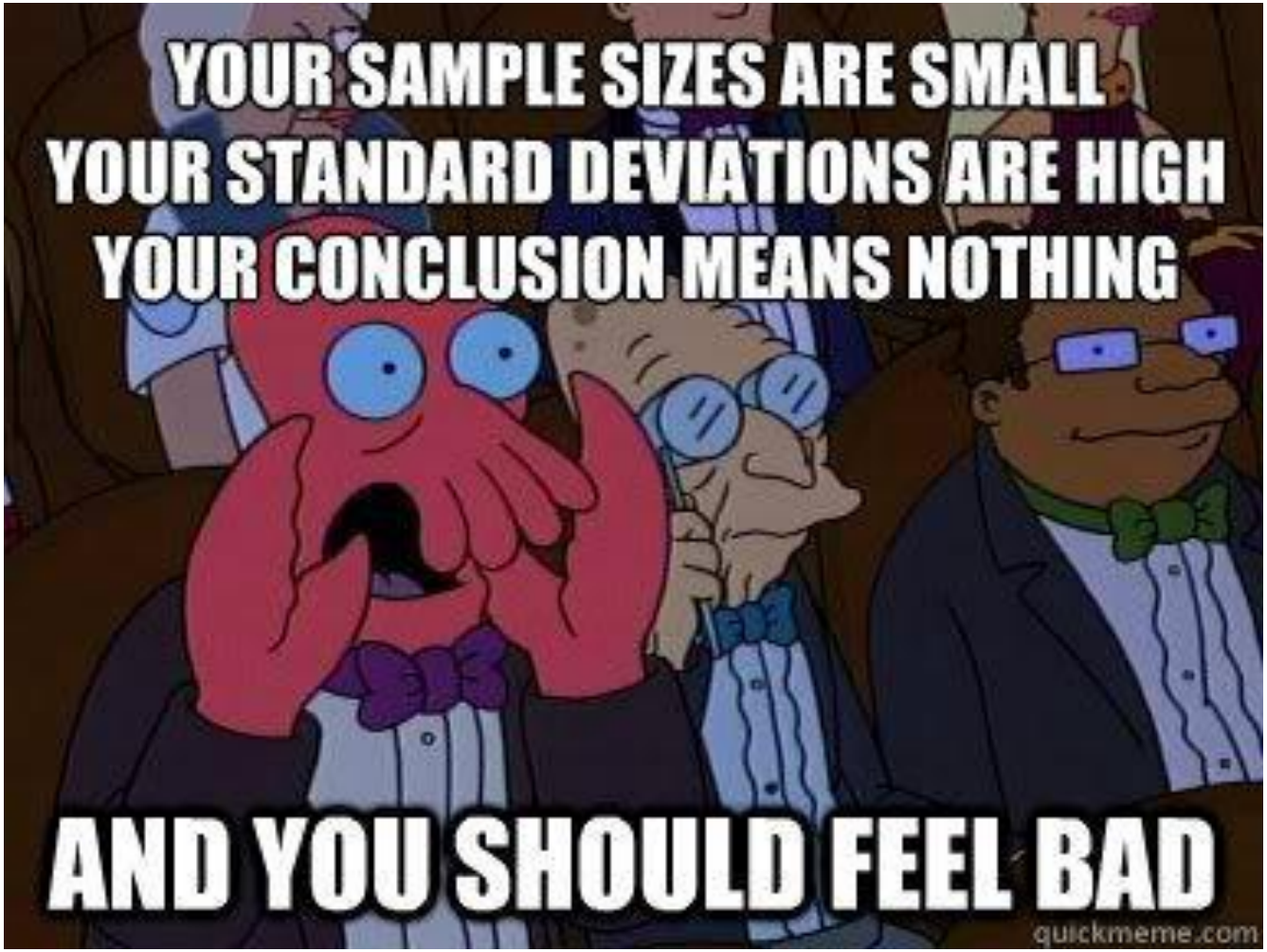
Statistical Significance

Why does greater variability among students' responses increase the likelihood that Melissa's results are due to chance?

3. In a third data collection, Melissa studies another 50 boys and another 50 girls. This time, she finds that girls raise their hands an average of 3.2 times per week and boys only raise their hands an average of 2.4 times in their math classes. Some students—both boys and girls—never raise their hands while others—again both boys and girls—raise their hands more than 15 times per week. Does this affect the likelihood of the difference between girls' and boys' participation to be statistically significant?



🔑01E Answer Key: Understanding Descriptive versus Inferential Stats and Statistical Significance


YOUR SAMPLE SIZES ARE SMALL
YOUR STANDARD DEVIATIONS ARE HIGH
YOUR CONCLUSION MEANS NOTHING
AND YOU SHOULD FEEL BAD
quickmeme.com