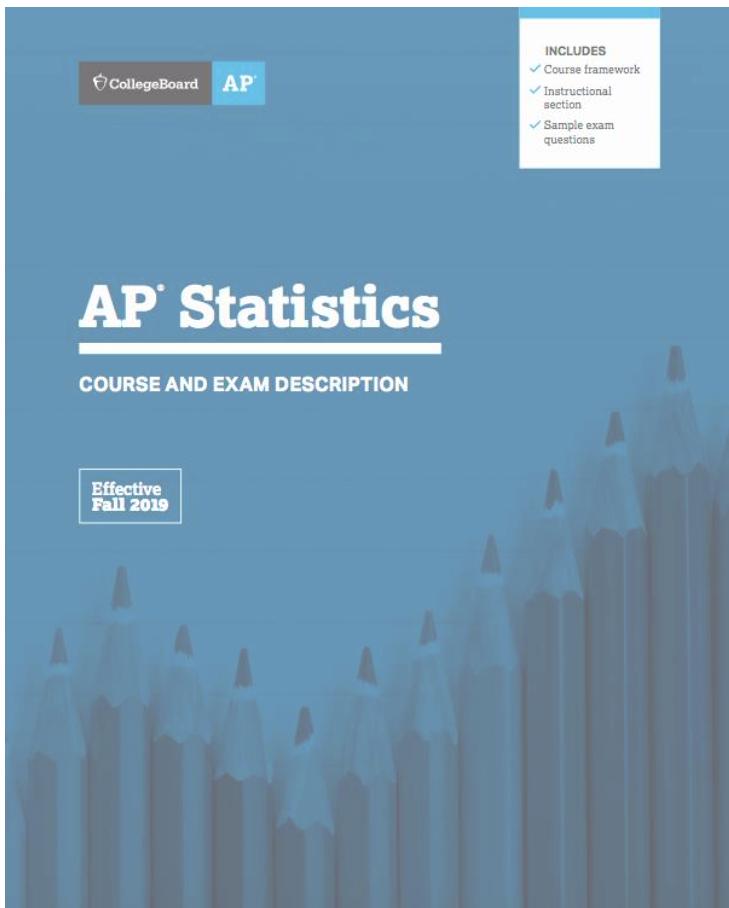


AP Statistics Summer Institute 2019

E-mail: dave_ferris@nobl.k12.in.us

Web site: noblestatman.com



1. What are the big ideas of AP Statistics (major concepts)?

2. What are the skills students need to “master?”

The Course and Exam Description (CED) Binder:

1. Tells you **WHAT** to teach and
2. What will be on the **EXAM**...
3. It does **NOT** tell you **HOW** to teach!

That's the professionally fun part!

DIRECTIONS: Complete the table below by filling in the last column with 1-2 examples of EKs, LOs, EUs, and Big Ideas for your chosen course topic. If it helps, you might consider starting at the bottom of the table, then working your way up.

Course Topic: _____

Component	Ask yourself:	Examples
Big Ideas	What are the recurring <u>themes</u> ?	
Enduring Understanding (EU)	What will you come to <u>learn over time</u> ?	
Learning Objective (LO)	What must you <u>do</u> to be successful?	
Essential Knowledge (EK)	What must you <u>know</u> to get started?	

2017 AP Statistics Exam

2. The manager of a local fast-food restaurant is concerned about customers who ask for a water cup when placing an order but fill the cup with a soft drink from the beverage fountain instead of filling the cup with water. The manager selected a random sample of 80 customers who asked for a water cup when placing an order and found that 23 of those customers filled the cup with a soft drink from the beverage fountain.
- (a) Construct and interpret a 95 percent confidence interval for the proportion of all customers who, having asked for a water cup when placing an order, will fill the cup with a soft drink from the beverage fountain.
- (b) The manager estimates that each customer who asks for a water cup but fills it with a soft drink costs the restaurant \$0.25. Suppose that in the month of June 3,000 customers ask for a water cup when placing an order. Use the confidence interval constructed in part (a) to give an interval estimate for the cost to the restaurant for the month of June from the customers who ask for a water cup but fill the cup with a soft drink.

1. List what this question requires students to DO (not what they know).

2. Read the titles of the four Skills categories. What do you notice compared to what students need to demonstrate in the FRQs.

Treasure Map

Checkpoint 1

What do the three-letter abbreviations above each colored box represent in the Course at a Glance?

Checkpoint 2

How many units are there for the entire year? Why do you think they are sequenced in this order?

Checkpoint 3

Which unit has the most topics?
Which unit has the least?

Checkpoint 4

What is the information at the bottom of each column telling you?

Multiple-choice: ~35 questions

Free-response: 2 questions

- 1 Exploring Data question
- 1 Investigative Task question

Checkpoint 7

How is the spiraling of skills displayed across the units?

Checkpoint 6

What is the range of class periods shown for Unit 5?
Why do you think these are given as a range?

Checkpoint 5

What do you think is measured on the Personal Progress Checks (PPC)?

Checkpoint 8

Under each unit title, there are percentages:
What are they percentages of?

15–23%



Checkpoint 9

What is the relationship between the Unit at a Glance table and the Course at a Glance?

Checkpoint 10

Where do the colors of the Course Skills show up in the Unit at a Glance?

INSTRUCTIONAL APPROACHES

Lesson & Instructional Strategy Planning Template

Topic Number:			
Topic Title:			
Instructional Strategy:			
UBD Components*	Big Idea:		
	Enduring Understanding (EU):		
	Learning Objective (LO):		
	Essential Knowledge (EK):		
Course Skills*:			
Materials:			
Time Allotted:			
Basic Procedures:			

Use coding to save space, e.g., Skill **2.C** or Learning Objective **UNC-1.C**

Web Site List

CED RESOURCES:

TPS supplements, correlations: go.bfpub.com/ap-course-updates-ap-statistics.html

RESOURCES, HANDOUTS, ACTIVITIES:

noblestatman.com (“**ehandouts**” and other resources)

AP Central (official **documents** and AP **Exam problems**)

apstatsmonkey.com (**clearinghouse** for many useful resources, including Best Practices
Night at the AP Reading, FRAPPY’s, etc.)

amstat.org (American Statistical Association)

(STEW **lesson plans** for activities and mini-projects; GAISE report: guidelines for
statistical instruction)

TOOLS:

stapplet.com (online “calculator” for all computations and inference procedures)

Against All Odds statistics **videos** (can stream for free--learner.org)

amstat.org (Census At School: **survey** and **student data**)

Rossmann Chance **applets** (many good simulation applets)

onlinestatbook.com/stat_sim/sampling_dist/index.html (great sampling distribution
demonstration **applet**)

getkahoot.com, quizlet.com, quizizz.com (engaging online review **games**)

StatCrunch (Stats **Software**: teacher account is free, students pay small fee)

StatKey (**simulation** website app)

tuvalabs.com (online **tool** for analyzing distributions and scatterplots)

gapminder.org (amazing online analysis **tool** of United Nations data)

Classifying Statistics Problems (lccconline—**practice** at choosing the correct inference
procedure)

ARTICLES and NEWS:

fivethirtyeight.com (great current **articles** with a statistical slant; engaging graphs)

www.causeweb.org/sbi (Simulation Based Inference discussions/**blog**. This is a “trending”
topic among high school and college statistics teachers.)

thisisstatistics.org (engaging information on statistics as a **career**)

tylervigen.com/spurious-correlations (funny, non-causation relationships)

M&M's Activity:

My Guess: _____

Guesses:

Actual:



Describe the distribution of the guesses:

Compare and contrast the distributions of guesses vs. actual:

FIRST DAY OF SCHOOL IDEAS:

- Register and give CensusAtSchool survey: amstat.org/censusatschool
(Your students' data can be downloaded as csv file for later use.)

The screenshot shows the homepage of the Census at School website. At the top, there's a large logo for "CENSUS at SCHOOL" with a checkmark icon. To the right is a map of the United States with the American flag pattern overlaid. Below the logo is a yellow banner with the text "Welcome to Census at School - United States". On the left, there's a vertical navigation menu with options: Home, Student Section, Teacher Section, Random Sampler, and International. The main content area features a news section with the headline "Census at School is an international classroom project that engages students in grades 4-12 in statistical problem-solving. Students complete a brief online survey, analyze their class census results, and compare their class with random samples of students in the United States and other countries. [More](#)". Below this is a "What's New?" section with a sub-headline about the American Statistical Association and Population Association of America seeking champions to expand U.S. Census at School nationally. To the right, there are links for "About Census at School", "Privacy Statement", and "Resources", along with a small thumbnail image of students raising their hands.

- Fun Size M&M's: guess # in bag, # of colors, mean of distributions, etc.
- Show thisisstatistics.org
- Memory Game Activity (see noblestatman.com or TPS5e Ch. 4 resources)
- Show interesting/famous/current graphs, data stories, articles about stats and surveys

FiveThirtyEight.com has great statistics stories and analysis:

The screenshot shows the homepage of FiveThirtyEight. At the top, there's a navigation bar with categories: Politics, Sports, Science & Health, Economics, and Culture. Below the navigation is a large image of a woman clapping in front of flags, with the caption "2017 FRENCH ELECTION". The main content area is divided into two columns. The left column, titled "THE LATEST", contains three articles: "Rajon Rondo Got Hurt Just When He Was Starting To Matter" (NBA), "The Utah Jazz Got Good While You Weren't Watching" (Sports), and "The French Election Is Way Too Close To Call" (Politics). The right column, titled "INTERACTIVES", contains two data visualizations: "NBA Predictions" (updated 14 hours ago) showing the chance of winning the NBA Finals for four teams, and "How Popular Is Donald Trump?" (updated 1 hour ago) showing approval and disapproval percentages over time. Both visualizations include a "See all NBA teams and games" or "See all polls" link.

- Look at some cool data sets: TuvaLabs.com, Census at School, gapminder.org, Bodyfat dataset, McDonald's menu, Arby's Menu, etc.

US Census at School Measurement Questions

Use Safari. Class ID: _____ Password: _____

The following questions require measurements. Please fill these out prior to taking the online survey.

4. How tall are you without your shoes on? Answer to the nearest centimeter.

5. What is the length of your right foot (without your shoe on)? Answer to the nearest centimeter.

6. What is your arm span? (Open arms wide and measure distance across your back from tip of right hand middle finger to tip of left hand middle finger.) Answer to the nearest centimeter.

9. How long does it usually take you to travel to school? Answer to the nearest minute.

14. What is the length of your left foot (without your shoe on)? Answer to the nearest centimeter.

16. What is the length of your index finger (finger next to your thumb) on your left hand? Answer to the nearest centimeter.

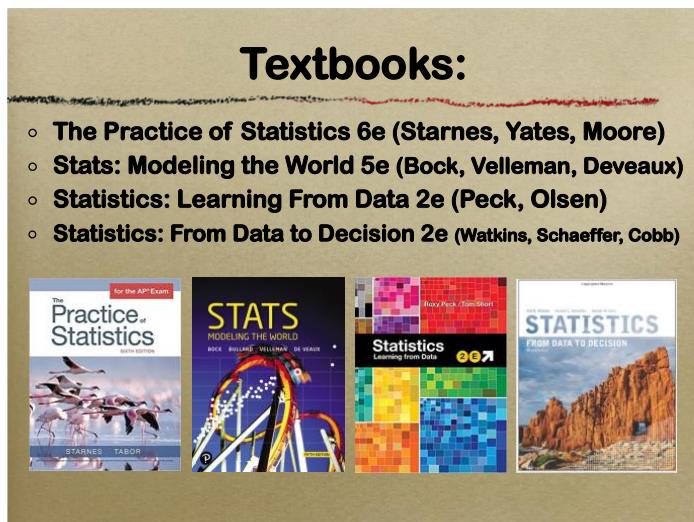
17. What is the length of your ring finger? (located between your middle finger and little finger) on your left hand? Answer to the nearest millimeter (there are 10 millimeters in one centimeter).

26. How many hours of sleep do you usually get when you have school the next day?

27. How many hours of sleep do you usually get when you don't have school the next day?

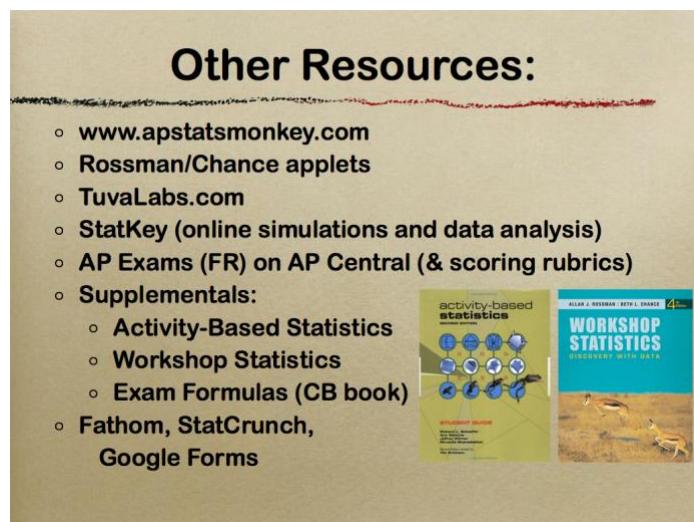
Essential resources from apcentral.collegeboard.com:

- **Course Overview**
- **Full Course Description** (including AP StatsTopic Outline—later in this handout)
 - “Statistics is a discipline in which clear and complete communication is an essential skill.”
 - “formulate cogent answers”
- **AP Statistics Teacher’s Guide** (BIG pdf file)
- **AP Teacher Community** and e-mail discussion group (see green sidebar)
- **Audit information**
- **Special Focus:** Sampling Dist’ns, Inference, Planning and Conducting a Study
 - Excellent “extra” information about these three important topics



Textbooks:

- **The Practice of Statistics 6e** (Starnes, Yates, Moore)
- **Stats: Modeling the World 5e** (Bock, Velleman, DeVeaux)
- **Statistics: Learning From Data 2e** (Peck, Olsen)
- **Statistics: From Data to Decision 2e** (Watkins, Schaeffer, Cobb)



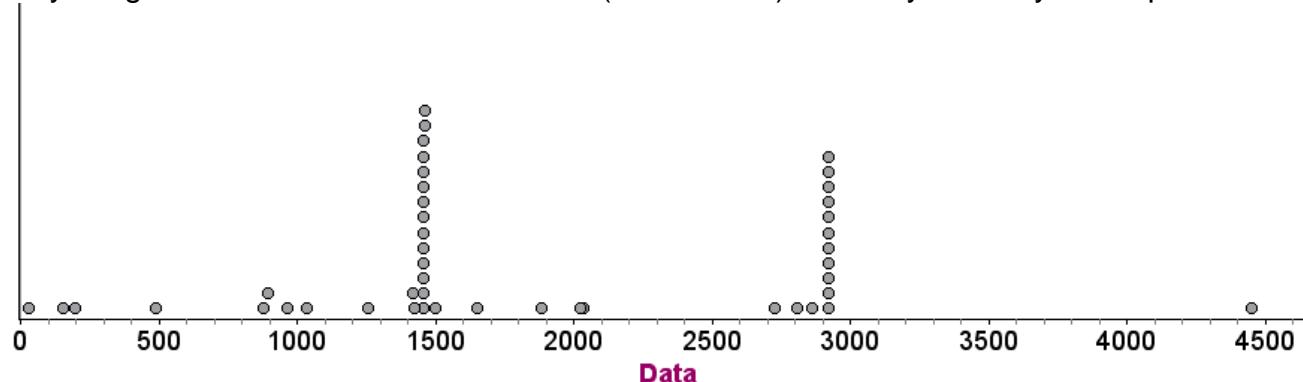
Other Resources:

- www.apstatsmonkey.com
- Rossman/Chance applets
- TuvaLabs.com
- StatKey (online simulations and data analysis)
- AP Exams (FR) on AP Central (& scoring rubrics)
- Supplements:
 - Activity-Based Statistics
 - Workshop Statistics
 - Exam Formulas (CB book)
- Fathom, StatCrunch,
Google Forms

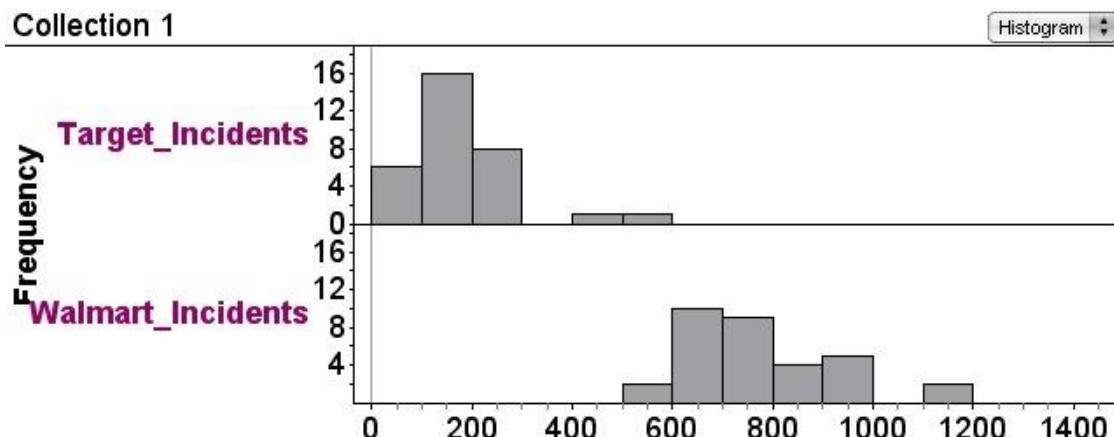
News, Tweets, AP Stat e-mails: great class openers!

Outliers and Inliers

Can you figure out the context of this data? (Hint: $n = 45$) You may ask 20 yes/no questions.

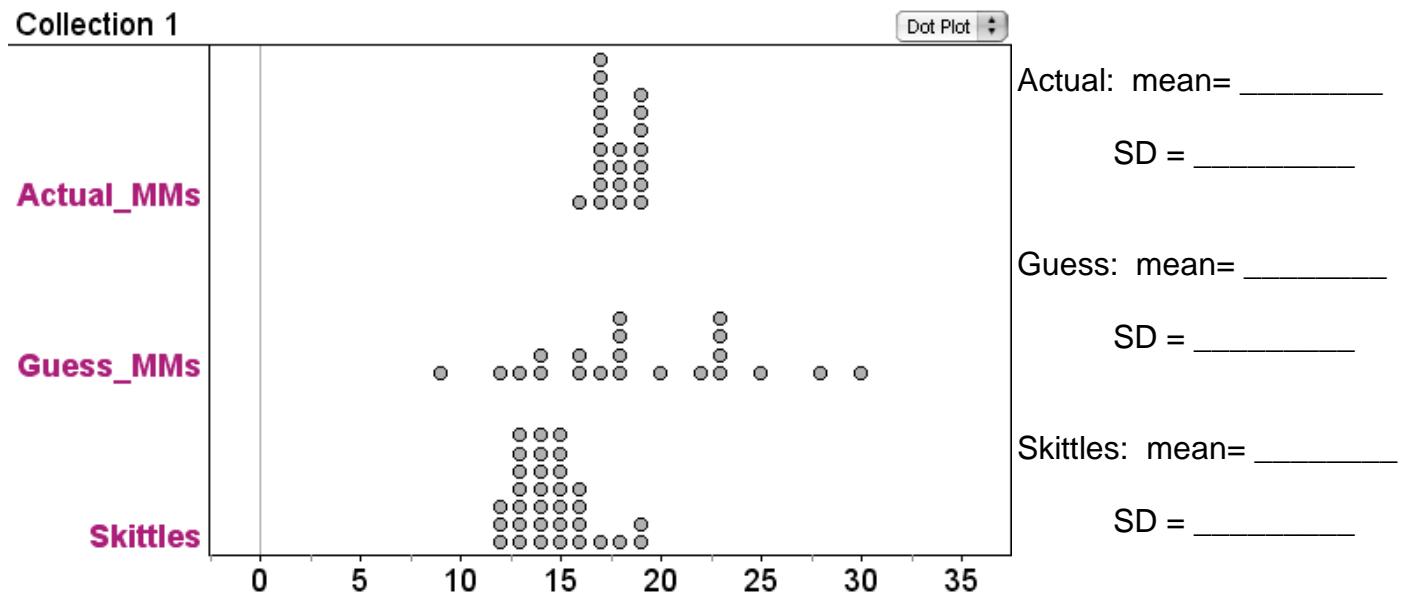


Is Wal-Mart Safe?¹



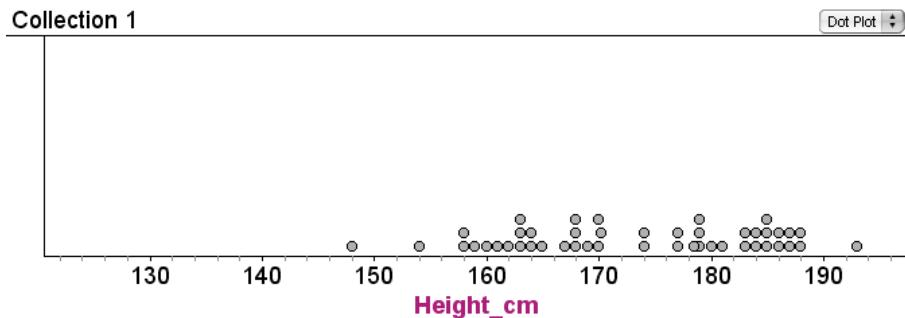
¹ See notes at the end of this handout for details on how the data was collected.

Standard Deviation:

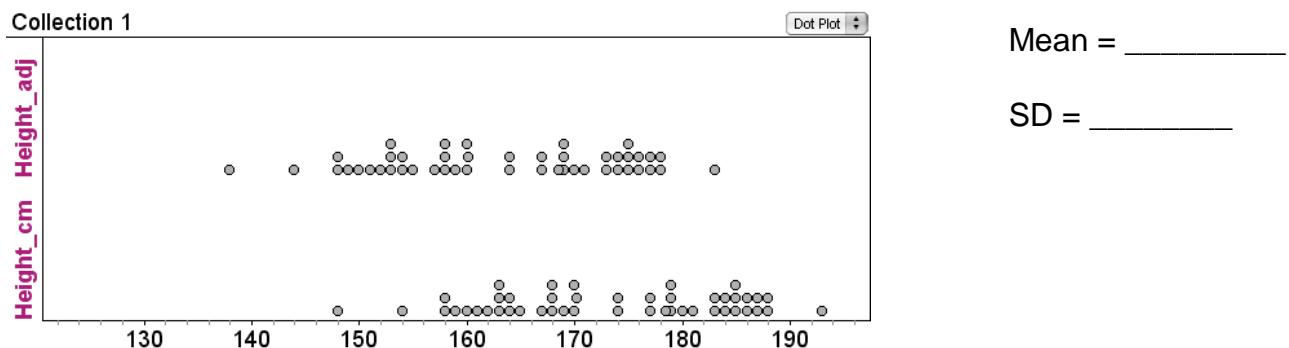


Transforming Data

Below is the distribution of student heights in cm.



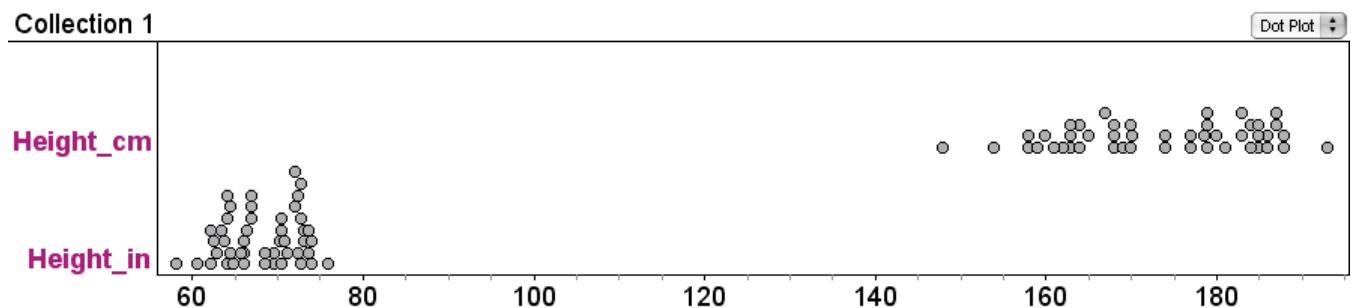
1. What if we discovered that the tape measure used to measure height had 10cm cut off (i.e. it started at 10 instead of 0)? We need to subtract 10cm from each student's height. How would this transformation affect the mean and SD?



2. What if we had to convert each height to inches and recalculate the mean and SD? What is the new mean and SD?

Mean = _____

SD = _____



Another application of transformations:

8. A company sells concrete in batches of 5 cubic yards. The probability distribution of X , the number of cubic yards sold in a single order for concrete from this company, is shown in the table below.

$X = \text{the number of cubic yards}$	10	15	20	25	30
Probability	0.15	0.25	0.25	0.30	0.05

The expected value of the probability distribution of X is 19.25 and the standard deviation is 5.76. There is a fixed cost to deliver the concrete. The profit Y , in dollars, for a particular order can be described by $Y = 75X - 100$. What is the standard deviation of Y ?

- (A) \$332.00
- (B) \$432.00
- (C) \$532.00
- (D) \$1,343.75
- (E) \$1,400.00

Categorical Data

Below are some data and questions from the Titanic disaster.

<http://lib.stat.cmu.edu/S/Harrell/data/descriptions/titanic.html>
<http://www.encyclopedia-titanica.org/>

1. Give the marginal distribution of class (in %'s).

		Titanic Passengers		Row Summary
		No	Yes	
Class	1st	129	193	322
	2nd	161	119	280
	3rd	574	137	711
Column Summary		864	449	1313
$S1 = \text{count}()$				

2. Give the conditional distribution of survival by class (in percent).

3. Of the first class passengers, what percent survived? $P(\text{survived} | \text{1st class}) = \underline{\hspace{2cm}}$

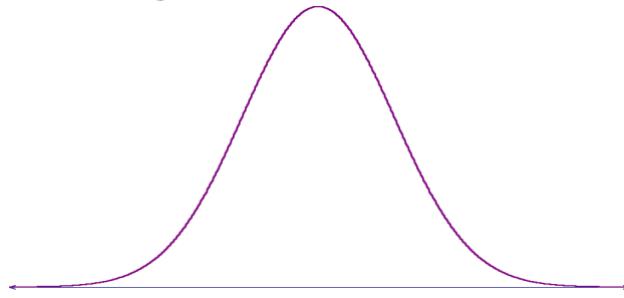
4. Of the survivors, what percent were 1st class? $P(\text{1st class} | \text{survived}) = \underline{\hspace{2cm}}$

5. What percent of passengers were 1st class? $\underline{\hspace{2cm}}$

6. What percent of passengers were either first class or survived? $\underline{\hspace{2cm}}$

7. Do survivability and class appear to be independent? Explain.

Common exercises using the normal model:



1. Find the z-score for a student who scored 660 on the SAT Verbal, where the mean is 505 and the standard deviation is 110.
2. What percentile did this student score?
3. What SAT Verbal score represents the first (lower) quartile?
4. Approximately what percent of students scored between 450 and 560?
5. What score would a student have to earn to be in the top 5% of all SAT Verbal scorers?
6. Approximately what percent of students scored between 395 and 615?
7. Approximately what percent of student scored within two standard deviations of 505?

8. Which is more extreme, a foot length of 32 cm or a height of 190 cm?

Heightcm	173.597 cm 11.36 cm
Footcm	26.6964 cm 3.37966 cm

S1 = `mean()`
S2 = `stdDev()`

Age Guessing Activity

1. Guess the ages of the following people, and put your guess in this column:

Name:	Actual Age:	Your Guess:
Donald Trump	_____	_____
Nate Silver	_____	_____
Bill Gates	_____	_____
Johnny Depp	_____	_____
Adele	_____	_____
Alex Trebek	_____	_____
Daisy Ridley	_____	_____
Miley Cyrus	_____	_____
Tom Brady	_____	_____
J. K. Rowling	_____	_____
Mick Jagger	_____	_____
Mark Zuckerberg	_____	_____

2. Put the actual ages of each person in the first column.

3. Type both lists into your calculator. (Actual ages in L1, your guesses in L2.)

4. Make a scatterplot for these two lists. (x-axis is L1, y-axis is L2)

5. Calculate r: _____

6. Describe below what you discovered about your age guesses:

Commonly Asked Regression Questions

(as seen on AP Statistics exams...can be used with any bivariate data set)

1. Describe the association in context.

2. Is a linear model appropriate to describe this relationship? Explain.

3. Write the equation for the linear model on this data.

4. Explain the meaning of the slope in this linear model

5. Explain the meaning of the y-intercept in this linear model

6. Find the value and explain the meaning of the correlation coefficient.

7. Find the value of and interpret r-squared

8. a. Using the linear model, predict _____ when _____ is _____.
b. Is the residual for this data point positive or negative? Is the model over- or underestimating? Explain.

9. Comment on any outliers present. Fully describe their effect on the analysis, if any.

10. Interpret regression and model information from a computer printout.

Starburst Grab Activity

NAME _____ Guess # Starbursts you can grab: _____

During Homecoming Week, your teacher is planning to give away a handful of Starburst candies to one lucky student in each class. To know how many candies to buy, it might be good to find out how many candies the typical student can grab with one hand. Naturally, we should collect some data! Each student will grab a handful of Starbursts and record the number grabbed.

1. Come to the container and grab one handful of Starbursts. You must shake your handful twice to allow any loose candies to fall out. After two shakes, this is your official sample. Count how many you grabbed and tally on the board and in a computer/calculator.
2. Based on the class sample, how many Starbursts should be bought per class so to have enough for the winner to grab? Explain your choice.
3. There is some variation in the number of Starbursts grabbed.
 - a) Name some statistics that could be calculated to measure the variation in this data.
 - b) Which one is the best for this data? Explain.
4. One measure of the total amount of variation is called Sum of Squares Total (SST). This is calculated by measuring the distances from each data point to the mean, squaring these distances, and adding them all up: $SST = \sum (x_i - \bar{x})^2$ Calculate the SST for this data.

$$SST = \underline{\hspace{10cm}}$$

5. The amount of total variation in this data makes it difficult to predict the number of Starbursts for a particular student. So now we will consider how improve our ability to predict. Perhaps the winning students should be picked first, before the teacher goes to Wal-Mart to buy Starbursts. What could be measured on each student to help accurately estimate the total amount of Starbursts needed? Pick three possible measurements, and explain each choice.

#1:

#2:

#3:

6. Decide as a class on the one variable that we should use to help with our prediction. Then create an appropriate model that relates this new variable to the # of Starbursts grabbed.

After the model is created, calculate the amount of variation that “still remains.” In other words, the total amount of variation remaining between the actual # of Starbursts grabbed and the predicted # of Starbursts grabbed according to this new model. This statistic is called the Sum of the Squares of the Errors (Residuals), or SSE.

$$\text{SSE} = \underline{\hspace{10cm}}$$

7. Go back and find the SST from #4. Notice that the SSE is smaller. In other words, by using a model on a second (explanatory) variable, we have been able to eliminate or account for some of the variation in the # of Starbursts grabbed. Calculate the percent of the SST that was “accounted for.”

The % of variation in the number of Starbursts that has been accounted for by the

$$\underline{\hspace{10cm}} \text{model on } \underline{\hspace{10cm}} = \underline{\hspace{2cm}} \%$$

This number is called _____. (Yes, it really is _____ ^ _____.)

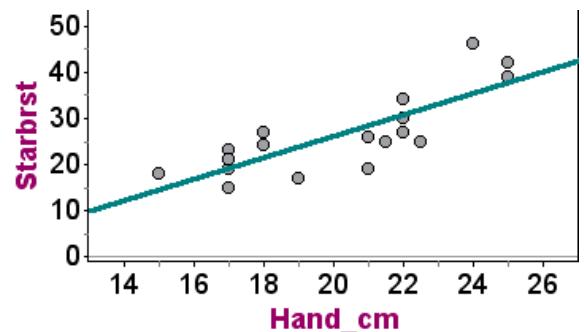
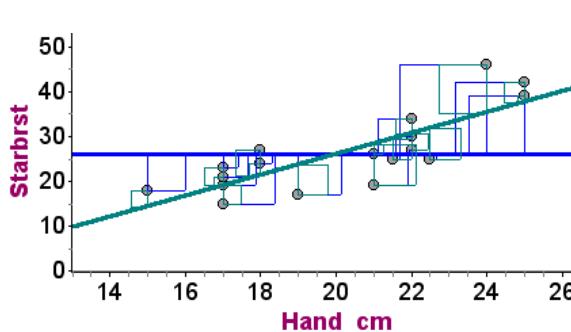
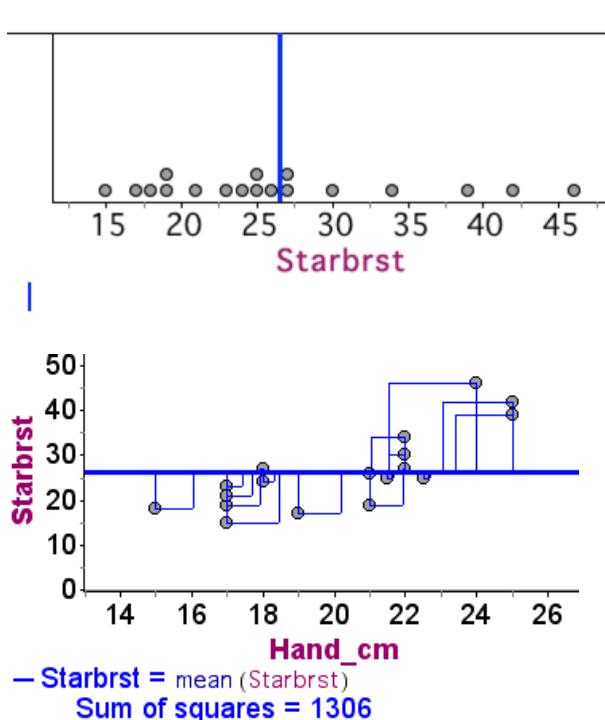
Starburst Grab Activity Teacher Notes:

1. Have students guess how many Starbursts they can grab with one hand: _____
2. Then tell the story on the sheet @ Homecoming...

Sample Data:

Mean = 26.5 candies, SST = 1306, SSE = 451.9, r-squared = 0.65

Hand_cm	# Starbst
22.0	34
21.0	19
21.5	25
17.0	23
15.0	18
18.0	27
22.0	27
18.0	24
25.0	39
17.0	19
22.0	30
21.0	26
24.0	46
22.5	25
25.0	42
17.0	15
17.0	21
19.0	17

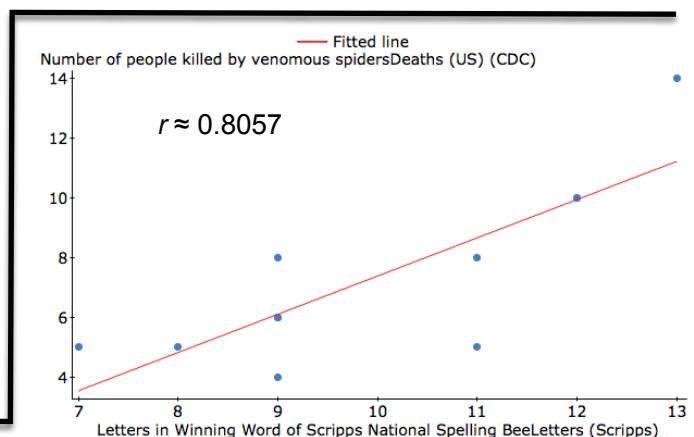


Age Guessing Key:

Name:	Actual BD:	AGE:
Donald Trump	6/14/46	_____
Nate Silver	1/13/78	_____
Bill Gates	10/28/55	_____
Johnny Depp	6/9/63	_____
Adele	5/5/88	_____
Alex Trebek	7/22/40	_____
Daisy Ridley	4/10/92	_____
Miley Cyrus	11/23/92	_____
Tom Brady	8/3/77	_____
J. K. Rowling	7/31/65	_____
Mick Jagger	7/26/43	_____
Mark Zuckerberg	5/14/84	_____

Wal-Mart vs. Target:

This phase of the study focused on the 50 Wal-Mart stores out of the 460 analyzed stores that experienced the “highest rate” of reported police incidents in 2004. Target stores chosen for the comparative analysis were within a 10-mile radius of the 50 “high incident” Wal-Mart stores. Of these 50 “high incident” Wal-Mart stores, three stores did not have a Target within 10 miles, leaving a sample of 47 Wal-Mart stores for further analysis. Because of further data restrictions, the sample for comparison was limited to 32 Wal-Mart stores and 30 nearby Target stores.

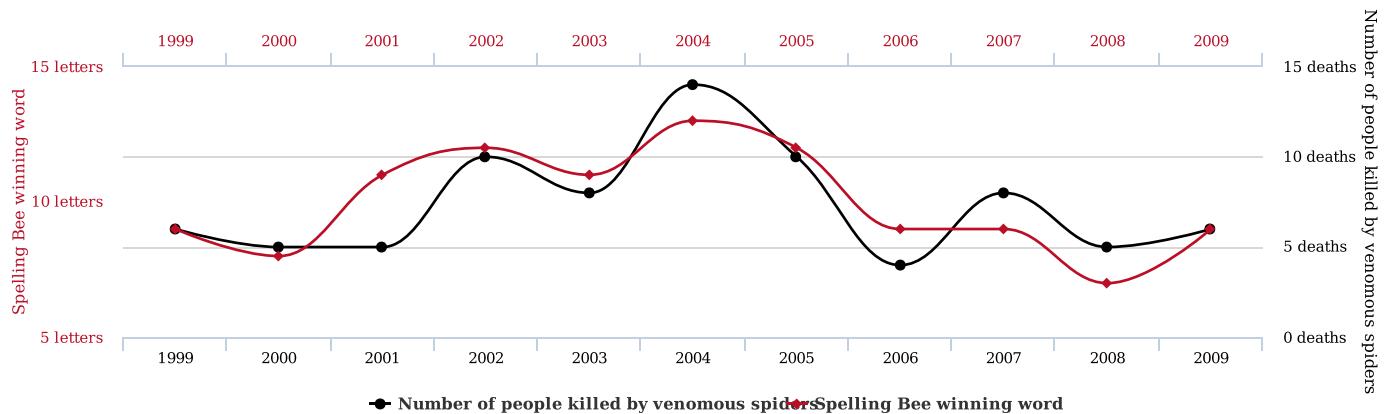


Letters in Winning Word of Scripps National Spelling Bee

correlates with

Number of people killed by venomous spiders

($r \approx 0.8057$)

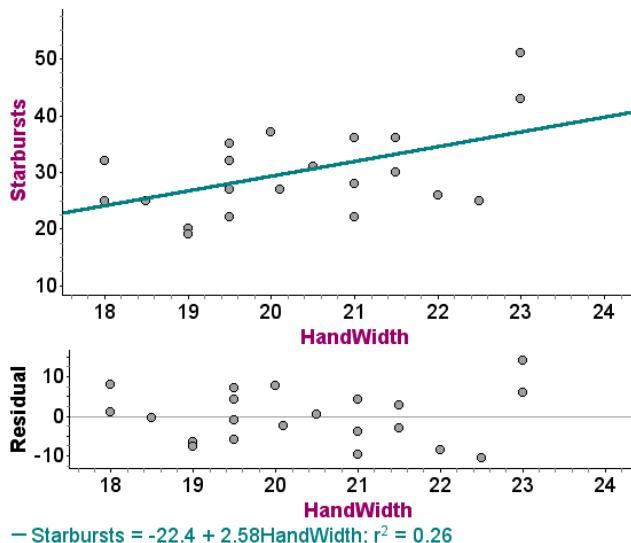


Graphs above show two ways of graphing the same data: one uses time and two different y-axes. The other shows only the two variables of interest. From tylervigen.com/spurious-correlations

Sample Data Sets for Regression (from real students):

Starburst Grab Activity:

Hand Width	Starbursts
21.5	30
21.0	28
19.5	22
19.5	32
20.0	37
19.0	20
22.0	26
19.0	19
21.0	22
23.0	51
23.0	43
18.0	25
20.1	27
21.0	36
19.5	27
19.5	32
21.5	36
18.0	32
22.5	25
18.5	25
20.5	31
19.5	35



Counting Steps:

LegLength	LongWalk
95.00	228
111.26	187
108.00	199
94.00	235
98.00	231
111.00	196
103.00	236
119.00	187
107.00	192
113.00	183
93.00	213
107.00	217
91.00	252
108.00	217
100.00	248
95.00	246
100.00	254
116.00	214
83.00	240
96.00	203
113.00	217

