# AP Statistics Summer Assignment
Amity Regional High School

The **Urban Dictionary** defines AP Statistics as:

A division of math that requires much writing and analytical thought. Taken by AP nerds, those hell bent on going to the 'best' colleges, or by those with a masochistic streak. Can be taken in place of Calculus, or, for those who really enjoy self-punishment, concurrently. Mostly taken by seniors.
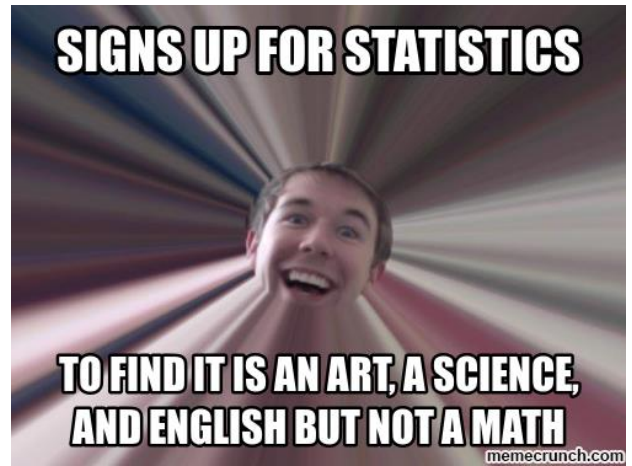
Welcome to AP Statistics! This course will be unlike any other math class you have ever taken! To get the most out of this course you need to be competent in basic algebra, be familiar with basic statistical measures, understand how to use a TI graphing calculator, and, most importantly, be willing to clearly and completely explain your answers in context.

Although this **_is_** a math course, AP Statistics concentrates **_much more heavily_** on thinking, reasoning, writing, and communicating than it does on number-crunching. Graphing calculators and computer software programs will do most of the tedious "busy work" for you.

This course is focused on four interrelated areas of concentration;

- **Exploring Data**. What patterns in a set of data do you see? What do these tell us about the data? What can we learn from them?
- **Sampling and Experimenting**. What plan would you implement to conduct a study? Can you effectively write a proposal for simulating a real-world situation?
- **Anticipating Patterns**. What can you surmise about random phenomena by using probability? How can you extrapolate your model into the future? What might you hope to have happen?
- **Statistical Inference**. How can you apply given parameters to your test hypothesis to see if it is valid?

A. Buy/Rent/Borrow a TI Graphing Calculator (preferably TI-84)*

It is mandatory that you have a charged **_TI_** calculator for every class day. Most of the examples and problems we will work on will use this calculator. If you cannot afford a TI calculator, contact Mrs. Shoop immediately upon your return to school and she will go over the process of borrowing a calculator for the year. I would stay away from any calculator other than a TI.

You will be expected to have some familiarity with the TI graphing calculator prior to taking AP Statistics. This knowledge includes entering data into lists, creating scatterplots, linear regression, the Y= equation editor, and window/zoom features. See part C of this document for more information.
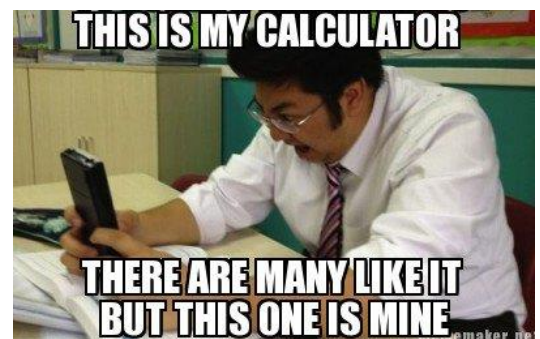
B. Understand what Statistics is All About*

What is Statistics:
Watch the videos

1. What is Statistics: Crash Course Statistics #1 https://www.youtube.com/watch?v=sxQaBpKfDRk
2. Prediction by the Numbers https://www.youtube.com/watch?v=9OIel5NUG7Q

Good resources for statistics articles/charts include:

- The Newark Star-Ledger (http://www.nj.com/)
- The New York Times (https://www.nytimes.com/)
- The Washington Post (https://www.washingtonpost.com/)
- Huffington Post (http://www.huffingtonpost.com/)
- Significance Magazine (http://www.significancemagazine.org/view/index.html)
- The Wall Street Journal's Numbers Guy (http://blogs.wsj.com/numbersguy/)
- BBC News Interactives and Graphics (http://www.bbc.co.uk/news/11628973)
- Nate Silver's 538 blog (http://fivethirtyeight.com/)
- Freakonomics (http://www.freakonomics.com/)
- WTF Visualizations (http://viz.wtf/)

C. Getting Ready AP Statistics - Expand your Lexicon and Practice

Reading and Vocabulary: You will use a free online statistical tutoring site that will give you information on variables and data displays. Follow the steps below:

a. Go to www.stattrek.com
b. Click on "AP Statistics" then "AP Tutorial"
c. Select the "Table of Contents" in the center of the page. Under each general topic are a list of subtopics. You will read the following subtopics to complete the vocabulary list.

General Topic:	Subtopics:
The Basics	Variables, Population vs. Sample, Central Tendency, Variability, Position
Charts & Graphs	Charts & Graphs, Patterns in Data, Dotplots, Histograms, Stemplots, Boxplots Scatterplots, Comparing Data Sets

You will be responsible for the following terms:
Categorical Variable, Quantitative Variable, Discrete Variable, Continuous Variable, Univariate Data, Bivariate Data, Population, Sample, Median, Mean, Outlier, Parameter, Statistic, Range, Center, Spread, Variance, Standard Deviation, Dotplot, Bar Chart, Histogram, Stemplot, Boxplot, Quartiles, Interquartile Range

Practice:

Types of Variables - Determine if the variables listed below are quantitative or categorical.

Time it takes to get to school	Number of people under 18 living in a household
Hair color	Temperature of a cup of coffee
Height	Amount of oil spilled
Age of Oscar winners	Jellybean flavors
Country of Origin	Number of pairs of shoes owned

Calculate Statistics

A statistic is a number calculator from a sample of data. Quantitative data has many different statistics that can be calculated. Determine the given statistics from the data below on the number of homeruns Mark McGuire hit in each season from 1982-2001.

70	52	22	49	3	32	58	39
39	65	42	29	9	32	9	33

Mean	
Minimum	
Maximum	
Median	
Q1	
Q3	
Range	
Interquartile Range (IQR)	

Dotplots

Directions: Determine each statistical measure for these dot plots. Grab your graphing calculator. Round to nearest hundredth where needed.

1.

Weights of Chickens in the Hen House
Ounces
MathBits.com

a) Describe the shape of the graph.

b) Mean

c) Median

d) Mode

e) Q₁ (first quartile)

f) Q₃ (third quartile)

g) IQR (interquartile range)

2.

Weights of Chickens in the Hen House
Ounces
MathBits.com

a) Describe the shape of the graph.

b) Mean

c) Median

d) Mode

e) Q₁
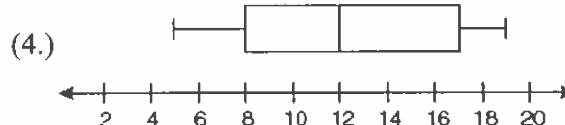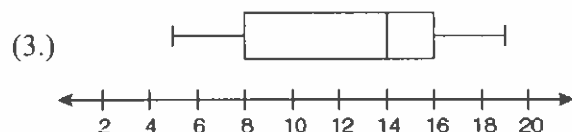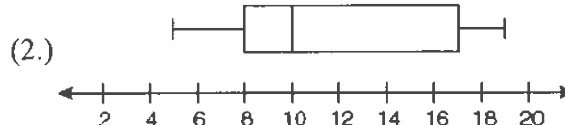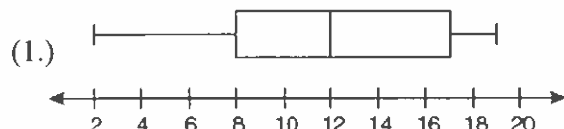
f) Q₃

g) IQR

Box Plots

Name_____

Directions: These questions deal with box and whisker plots.

1. Jason saves a portion of his salary from his part-time job in the hope of buying a used car. He recorded the number of dollars he was able to save over the past 15 weeks.

Dollars saved: 19, 12, 9, 7, 17, 10, 6, 18, 9, 14, 19, 8, 5, 17, 9

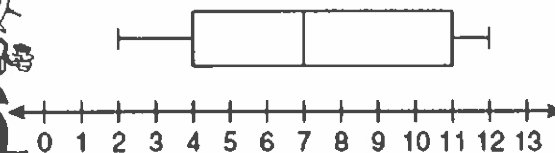Which box and whisker plot represents this data?

(1.)

(2.)

(3.)

(4.)

2. Examine the box plot:

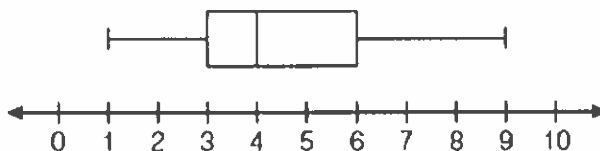(a) What is the value of the third quartile?

(b) What is the median score?

(c) What percent of the values are *less than 4*?

(d) What is the range of the data presented in the box plot?

3. A movie theater recorded the number of tickets sold daily for a popular vampire movie during the month of December. The box plot below shows the data for the number of tickets sold, in hundreds.

Determine whether the following statements are TRUE or FALSE regarding this box plot. Circle choice.

(a) 25% of the data is between 300 and 400.   T   F

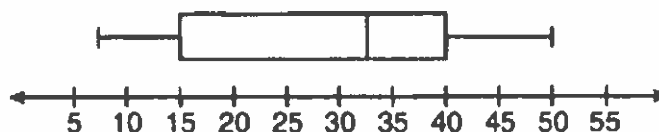(b) the second quartile is 600.   T   F

(c) the median is 400.   T   F
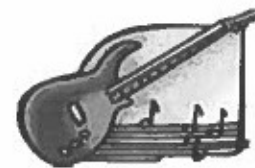
(d) the first quartile is 300.   T   F

(e) the range is 300.   T   F

(f) 50% of the data is between 300 and 600.   T   F

**4.** The box plot shows the ages of people attending a music concert.

(a) Which interval contains exactly 50% of the ages?

(b) What percentage of the ages are 15 or older?

(c) Which choice is the median age?   (1) 25   (2) 30   (3) 33   (4) 35

5. The box plot shows the lengths of cell phone calls made by three students in a one week period.

Which of the following statements is true?

(1.) All of Nikki's calls were longer than all of Paula's calls.

(2.) Paula made the fewest number of calls.

(3.) Mario made 50% more calls than Paula.

(4.) The median length of Mario's calls is equal to the median length of Nikki's calls.
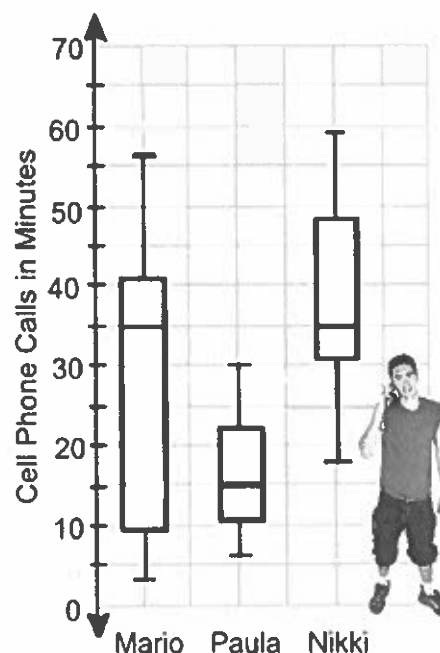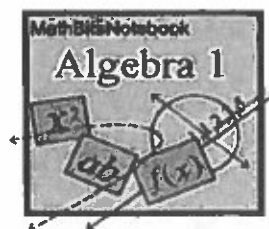

6. Examine the box plot:
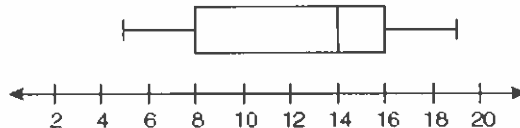
(a) Which interval represents the interquartile range, IQR?

(b) What percentage of the data lies in the IQR?

(c) Which interval could be described as being the most "clustered" section of the data?
     5-8,   8-14,   14-16,   16-19

Interpreting Graphs

Name_____

*Directions: Solve the following problems dealing with statistical graphs.*

1. Questions pertain to the histogram.

(a) How many students are in the Fencing Club?

(b) How many students' heights are greater than or equal to 66 inches tall?

(c) What percentage of the students have a height greater than or equal to 60 inches but less than 66 inches, to the *nearest tenth* of a percent?

Heights of Students in Fencing Club
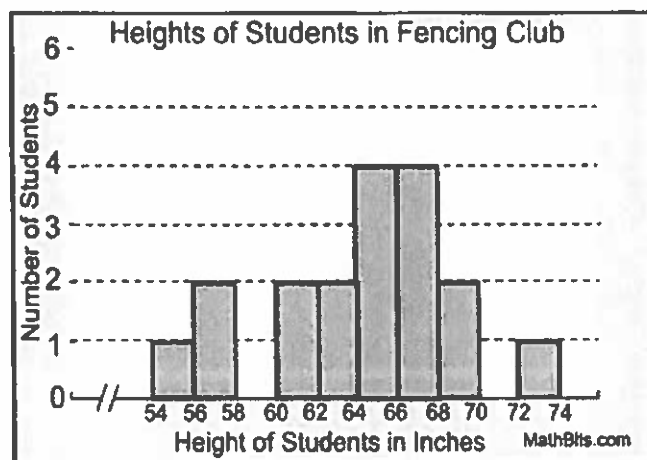
(d) In which interval, will you find the median student height?

2. The box and whisker plots represent the number of points scored in each game for the 2013 fall football season for two area high school teams.

(a) In general, which team scored more points per game?

(b) Which team's data shows more variability in the points scored?

(c) What is the difference between the median points scored by these two teams?

Points Scored
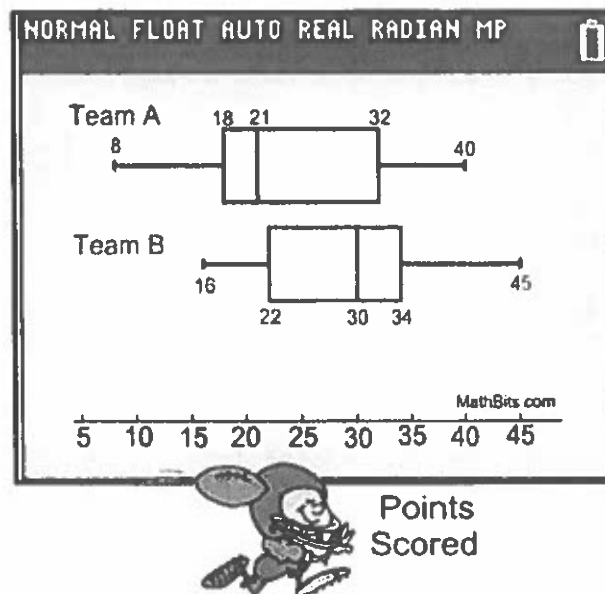
3. In the same morning, fifteen people spend money on $5 lottery tickets at a local gas station.
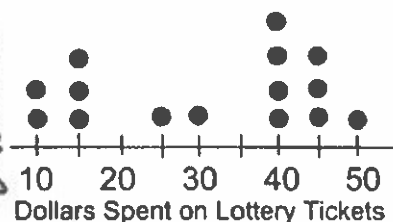The dot plot shows the amount spent by each customer on the tickets.

(a) What is the median of the dollars spent?

(b) What is the mean (average) of the dollars spent?

(c) What monetary value represents the first quartile of this data?

(d) What is the interquartile range?

~~**(e)** What is the population standard deviation for this data (to the *nearest hundredth*)?~~

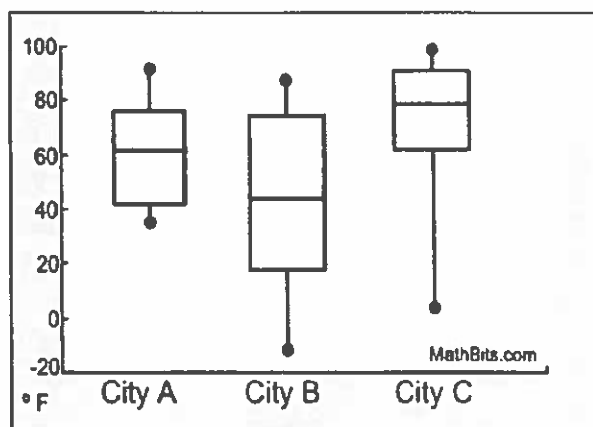**4.** The box and whisker plots represent monthly temperatures (° F) for three cities.

(a) Which city has temperatures most evenly spread over the greatest range?

(b) Which city has a potential outlier among its data?

(c) Which city has temperatures consistently above freezing (32° F)?

---

5. Examine the histograms at the right. Complete the chart by choosing the graph that best shows the stated condition. Place an ✗ under the correct heading for each condition.
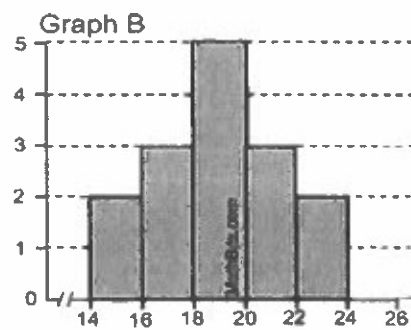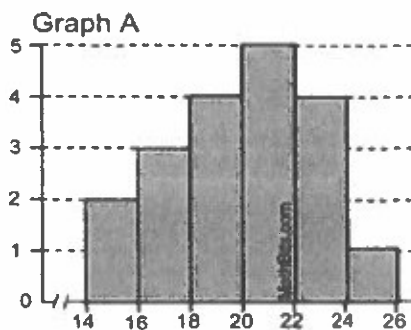
	Graph A	Graph B	Both Graphs
Larger Median			
Symmetric Graph			
Smaller Range			
Larger Data Count			
Interval Scale of 2			
Unimodal			

6. Two game players have been recording their best scores each day during a one week video tournament. Based on the box plots, determine if the following statements are **true** or **false**.

(a) The median score of Player A is greater than the median score of Player B.

(b) The highest score was obtained by Player B.

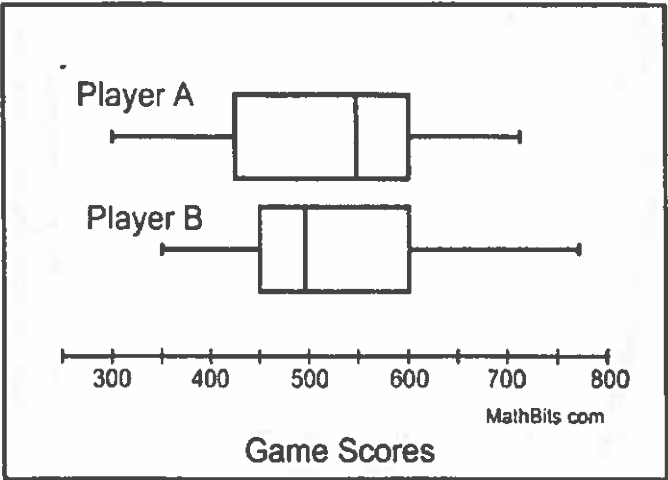**(c)** The interquartile range of the scores of Player B is less than the interquartile range of the scores of Player A.

7. Examine the dot plots. Complete the chart by choosing the best situation as it relates to each of the stated statistical measures. Place an ✗ under the correct heading for each statistic.
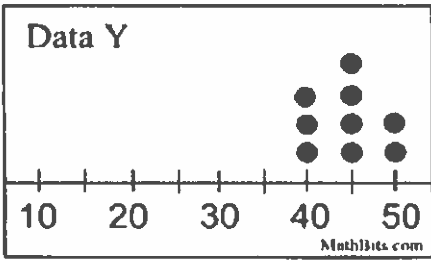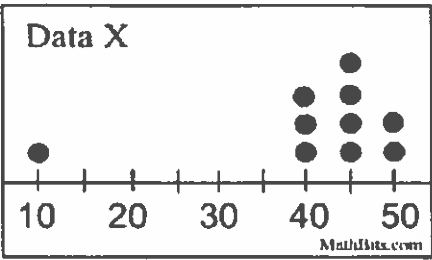
	Larger in Data X	Same in both data sets	Larger in Data Y
Median	\		
Range			
Mean			
Standard Deviation			
Mode			
Maximum			

You Are What You Eat
Student Activity

Name _____

Class _____

Problem 1 – Fast Food Hamburgers

Calories are a unit of energy. You may burn around 70 Calories just sitting at your desk and over 100 Calories while you sleep. The energy required to lift a book from the floor to the table is only about 0.003 Calories.
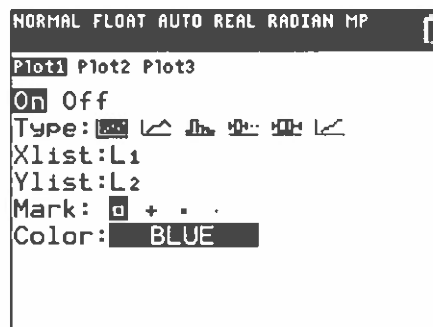
The table to the right shows the grams of fat and Calories for 11 hamburgers from three different fast food places. How many fast food hamburgers do you eat in a week?

Fat (g)	Calories
11	270
9	270
19	360
24	400
28	450
35	570
38	620
43	660
49	720
53	790
79	1060

1. Look at the data. What do you think is the relationship between the grams of fat and the Calories in a hamburger? Is it a positive or negative association?

2. In order to graph the data and set up an appropriate viewing window, consider the following.

   a. What is the independent variable? What is the domain?

   b. What is the dependent variable? What is the range of the data?

3. Estimate about how much the Calories increase for every increase of 1 gram of fat.

Plot the Calories vs. grams of fat as a scatter plot, and then find the line of best fit. Enter the data into your calculator by pressing [STAT] [ENTER] to open the ListEditor. Enter the grams of fat in L_1 and the Calories in L_2.

Plot the data as a scatter plot as shown to the right by pressing [2nd] [Y=] for [STAT PLOT]. Then press [WINDOW] and set up the window you decided upon in questions 2. The Xscl and Yscl sets up the tick marks. Press [GRAPH].

```
NORMAL FLOAT AUTO REAL RADIAN MP        []
Plot1 Plot2 Plot3
On  Off
Type: ▨ ⟋ ╨ ⊞ ⊞ ⟋
Xlist:L1
Ylist:L2
Mark: ▫ +  ·
Color:    BLUE
```

4. How does your answer for question 3 relate to the graph?
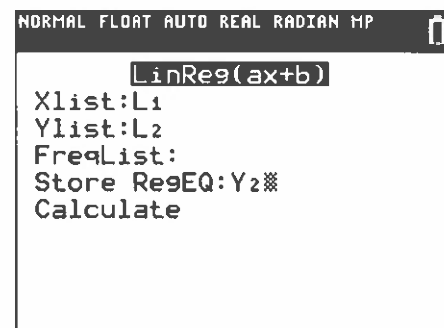
To create your own line of best fit press [STAT], arrow over to choose CALC, and up arrow to select Manuel-Fit. Arrow down to and select Calculate. Your equation will automatically be stored in Y_1. With the TI-84 Plus C you can press the soft key [GRAPH] to change the STYLE.

```
NORMAL FLOAT AUTO REAL RADIAN MP      []
            Manual-Fit
   Store EQ:
   Calculate
```

Move the cursor along to where the line of best fit appears to be and press [ENTER]. Move the cursor again to pick another point further away. Think of it as lining up your ruler along the data points. Press [ENTER] again. Use the arrows and type in a value to change the slope and y-intercept until you are satisfied with your line. To quit press [2nd] [MODE].

5. What is the equation of your line? Explain the meaning of your slope. Include units.
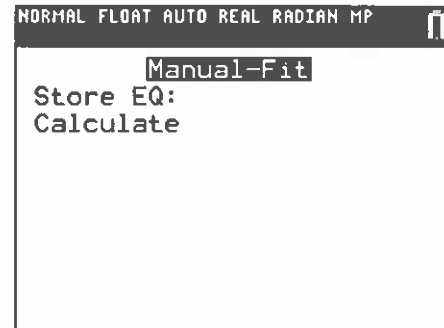
6. How does your equation compare with the linear regression? View the data and the two linear equations together. [STAT], CALC, LinReg (ax+b). Use Vars → Y Vars → Function to select Y_2 to store the regression equation in Y_2 as shown to the right and select Calculate.

```
NORMAL FLOAT AUTO REAL RADIAN MP      []
            LinReg(ax+b)
   Xlist:L1
   Ylist:L2
   FreqList:
   Store RegEQ:Y2▒
   Calculate
```

Use the regression equation to answer the following questions.

7. What does the y-intercept mean in terms of the number of hamburgers?

8. How many Calories are in a hamburger with 22 grams of fat?

9. If one of the fast food places created a triple burger with 1243 Calories, approximately how many grams of fat would it have?

Extension

10. Find nutritional data on other fast food hamburgers and compare them to your results.

11. Find nutritional data on various other types of food (ice cream, candy, chips, pop, etc.) and explore the relationship between grams of sugar and Calories.

TI-84+ Family Quick Reference Sheet Algebra 2

When in doubt, QUIT and go HOME.

To Find Intersection Pts:

1. Graph both equations.
2. Use CALC menu (2nd TRACE)
   Choose #5 Intersect
3. Move near the intersect location.
4. Simply press <ENTER> 3 times
   to reveal the answer.

If you are looking for more than one intersection point, repeat this process.

Check Inverse:

Enter your algebraic inverse in Y1. Graph. Use DRAW #8DrawInv to verify it is correct.

To see $\sqrt{-25} = 5i$, use $a + bi$ mode.

Logs and Exponents:

1. The LOG key is log base 10.
2. To enter: $\log_4 64$ use $\dfrac{\log 64}{\log 4}$
3. $27^{\frac{1}{3}}$ is 27^(1/3)   remember ( )

Summations:

Enter   sum(seq(2x+2, x, 2, 7, 1) $\sum_{k=2}^{7}(2k+2)$

- 2nd STAT(LIST) – MATH - #5 sum
- 2nd STAT(LIST) – OPS - #5 seq

The format for **seq**: *expression, variable, starting value, ending value, increment.*

Calculator ID #:
Choose 2nd MEM, #1 About
ID****-****-****

Diagnostics ON: must be ON to see correlation coefficient. *r*.
1. MODE – StatDiagnostics: ON
or 2. CATALOG. ALPHA D. DiagnosticOn. ENTER. ENTER

To Get Statistical Information:

1. Place data in Lists: STAT → EDIT
2. Engage 1-Variable Statistics: STAT → CALC #1 1-VAR STATS
3. On Home Screen indicate list containing the data: 1-VAR STATS L1

$\bar{X}$ = mean	Q_1 = data at the first quartile
Sx = the sample standard deviation	**med** = data at the median
σ_x = the population standard deviation	(second quartile)
n = the sample size (# of pieces of data)	Q_3 = data at the third quartile

To Get Scatter Plots and Regressions

(Linear, Quadratic, Exponential, Power, etc):
1. Place data in Lists: STAT → EDIT
2. Graph scatter plot: STAT PLOT #1 <ENTER> Choose ON.
   Choose the symbol for scatter plot, choose L1 , L2, choose mark
3. To graph, choose: ZOOM #9
4. To get regression equation: STAT → CALC #4 Lin Reg(ax+b)
   (or whichever regression is needed)
5. On Home Screen: LinReg(ax+b) L1 , L2 , Y1
6. to see graph – GRAPH

To get Y1 to appear, VARS → Y-VARS Choose FUNCTION, Y1 **OR ALPHA F4**

Functions:

$Y_1 = f(x)$ and $Y_2 = g(x)$

$(f + g)(x) \rightarrow Y_3 = Y_1 + Y_2$

$(f - g)(x) \rightarrow Y_3 = Y_1 - Y_2$

$(f \cdot g)(x) \rightarrow Y_3 = Y_1 Y_2$

$(f / g)(x) \rightarrow Y_3 = Y_1 / Y_2$

Composition:

$(f \circ g)(x) \rightarrow Y_3 = Y_1(Y_2)$

Normal Distributions   DISTR(2nd VARS)

1. normalcdf (lower, upper, mean, s.d.) *Finds prob. on cumulative interval.*
   - to enter ∞, use 10^99 or 1 EE 99.
2. normalpdf(x, mean, s.d.) *Graphs the normal distribution.*
   - Window: Xmin = mean – 3 s.d.; Xmax = mean + 3 s.d.; Xscl = s.d.
             Ymin = 0; Ymax = 1/(2 s.d.); Yscl = 0
3. ShadeNorm(lower, upper, mean, s.d.) *To see area and % under curve.*
   - must graph using normalpdf first, or you won't see your shading.

D. Understand Academic Integrity*

Academic integrity essentially means "intellectual honesty": honesty in the use of information, in formulating arguments, and in other activities related to the pursuit of knowledge and understanding. It is a core principle that underpins how we live and learn in a community of inquiry. As members of an academic community, we are entitled to a wide degree of freedom in the pursuit of scholarly interests. With that freedom, however, comes the responsibility to uphold the high ethical standards of academic conduct.

If I were to ask each of the AP Stats students their definition of "Academic Integrity", I would probably get different answers from everyone in the class.

A lot of the work you will be doing in AP Stats involves partners or small groups. In addition, much of the work is done at home. Because of the nature of the course, it is imperative that each and every one of us has the exact same understanding of what constitutes "Academic Integrity" as I do.

E. Beach Reading/Listening/Watching
If you have the time and want to delve a little further into the world of statistics, try out one or more of the following resources. ***These are not required at all.***

Books
- **Freakonomics: A Rogue Economist Explores the Hidden Side of Everything** by Levitt and Dubner
- **The Drunkard's Walk, How Randomness Rules Our Lives** by Leonard Mlodinow
- **Predictably Irrational** by Dan Ariely
- **Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists**, J. Best, University of California Press, 2001.
- **A Mathematician Reads the Newspaper**, J. A. Paulos, Basic Books, 1995.
- **200% of Nothing**, A. K. Dewdney, John Wiley and Sons, 1993.
- **Bringing Down the House: How Six Students Took Vegas for Millions**, B. Mezrich, Free Press, 2002
- **Moneyball: The Art of Winning an Unfair Game**, M. Lewis, Norton, 2003
- **Tainted Truth: The Manipulation of Fact In America**, C. Crossen
- **The Ghost Map: The Story of London's Most Terrifying Epidemic—and How It Changed Science, Cities, and the Modern World**, S. Johnson
- **The Tipping Point: How Little Things Can Make a Big Difference**, M. Gladwell, Little, Brown & Co., 2002

Videos (TED Talks)
- **The Freakononics of Crack Dealing:** http://www.ted.com/talks/steven_levitt_analyzes_crack_economics
- **Surprising Stats about Child Car Seats:** http://www.ted.com/talks/steven_levitt_on_child_carseats
- **What We Learned from 5,000,000 Books:** https://www.ted.com/talks/what_we_learned_from_5_million_books
- **How Juries are Fooled by Statistics:** http://www.ted.com/talks/peter_donnelly_shows_how_stats_fool_juries
- **Why Smart Statistics are the Key to Fighting Crime:** http://www.ted.com/talks/anne_milgram_why_smart_statistics_are_the_key_to_fighting_crime
- **Does Racism Affect How You Vote?:** http://www.ted.com/talks/nate_silver_on_race_and_politics
- **Flip Your Thinking About AIDS in Africa:** http://www.ted.com/talks/emily_oster_flips_our_thinking_on_aids_in_africa
- **Three Ways to Spot a Bad Statistic:** https://www.ted.com/talks/mona_chalabi_3_ways_to_spot_a_bad_statistic