

## AP Statistics Summer Homework 2022

Dear Students,

Welcome to AP Statistics 2022-2023! To prepare for the AP Statistics class at Santa Clara High School you will need to spend approximately 18-20 hours completing the following summer assignments.

1. Join the Google Classroom to access a PDF of the textbook. Use code: **c2awibs**
2. Read and complete the questions for Chapter 1 (pages 1-76). Do not rush through and do multiple assignments on the same day. You need to absorb the concepts slowly and completely. Each one of these assignments is a typical representation of the workload for one day of an AP Statistics class. Note: If you do the reading guides and the assignments concurrently, it will be much easier.
3. After you've read each section, complete the exercises below:

### **Section 1.1**

Read pg. 1-6 and Do exercises 2, 4, 6, 7, & 8

Read pg. 7-11 and Do exercises 12, 14, 16, & 18

Read pg. 11-20 and Do exercises 20, 22, 24, 27-34

### **Section 1.2**

Read pg. 25-33 and Do exercises 38, 40, 42, 44, 46, 48

Read pg. 33-40 and Do exercises 54, 58, 59, 60, 66, 69-74

### **Section 1.3**

Read pg. 48-59 and Do exercises 80, 82, 84, 86, 87, 90, 92, 93, 95

Read pg. 60-68 and Do exercises 98, 100, 104, 106, 107-110

\*Note: I often assign even problems. Usually, there is a very similar odd problem to go with each even problem (for example, #1 is very similar to #2), and the answers to the odd problems are in the back of the book. Therefore, if you are having trouble with one of the even problems, check to see if there is a similar odd problem before it. Then, you can look and see how the book answers the odd problem and use it to help you with the even problem in your homework.

Acquire and familiarize yourself with the operation of a TI-84+ graphing calculator. The textbook will show you how to do several basic operations on the calculator. You can also get information from

<http://mathbits.com/MathBits/TISection/CachingPage.html>

**There will be a test during the first week of school.** The summer homework and reading guides will be due the day of the test also. If you have completed the all of the summer assignments, you should do very well on the first exam. Send me an email ([tburk@scusd.net](mailto:tburk@scusd.net)) if you have any questions. Good Luck and see you in August! --Ms. Burk

\*Additional online resources: If you are confused about ideas in the book, some additional resources that might be helpful are:

1. <http://www.statrek.com/ap-statistics-1/ap-statistics-intro.aspx>
2. [www.davidmlane.com/hyperstat/](http://www.davidmlane.com/hyperstat/)

These websites provide good overviews of Statistics. It is worth browsing through them even if you do understand the readings in the book.

<b>Topic: Ch. 1 Intro &amp; 1.1 Data Analysis: Making Sense of Data AND Analyzing Categorical Data</b> <ul style="list-style-type: none"> <li>- Bar Graphs and Pie Charts</li> <li>- Good and Bad Charts</li> <li>- Two-Way Tables and Marginal Distributions</li> <li>- Relationships Between Categorical Variables</li> <li>- Conditional Distributions</li> </ul>	<b>Name:</b>
<b>AP Topics</b> I. Exploring Data: describing patterns and departures from patterns E. Exploring Categorical Data <ol style="list-style-type: none"> <li>1. Frequency tables and bar charts</li> <li>2. Marginal and joint frequencies for two-way tables</li> <li>3. Conditional relative frequencies and association</li> <li>4. Comparing Distributions using bar charts</li> </ol>	<b>Period:</b>
	<b>Date:</b>
<b>Essential Questions:</b> <ul style="list-style-type: none"> <li>- What are the differences between categorical and quantitative variables?</li> <li>- What types of graphs are appropriate for presenting categorical data?</li> <li>- What makes some graphs of categorical data deceptive?</li> <li>- How do you present data for two categorical variables?</li> <li>- How do you describe the association between two categorical variables?</li> </ul>	

<b>Reading Guide</b>	<b>Read p. 1-6 and complete this reading guide</b>
	What is <b>data analysis</b> and why do we need it?
Define	<b>individuals</b> -
Define	<b>variable</b> -
	What are <b>categorical</b> and <b>quantitative variables</b> ?

	How can you distinguish between <b>categorical</b> and <b>quantitative variables</b> ?
	What are some examples of variables that take number values, but are not quantitative? How can you determine they are categorical instead of quantitative?
	Why is it important to distinguish between categorical and quantitative variables?
	<b><i>*The Example problems and Activities throughout the book are VERY IMPORTANT. Be sure to read through them. They will help you with the homework problems. For example, The Example on page 3-4 "Census at School" shows you what a complete answer would look like for problem #4 of your homework.</i></b>
Define	<b>distribution (of a variable) -</b>
Define	<b>inference -</b>

<b>Summary of Intro to Chapter 1 (In your own words, write a brief summary of what you have read in pages 1-6 here)</b>

**\*\*NOW Do p. 6-7 exercises 2, 4, 6, 7, & 8**

<b>Reading Guide</b>	<b>Read p. 7-11 and complete this reading guide</b>
	What is the difference between a <b>frequency table</b> and a <b>relative frequency table</b> ?
	What kind of data is usually presented in <b>frequency tables</b> and <b>relative frequency tables</b> ?
	What is <b>roundoff error</b> ?
	What are the most common graphical representations of categorical data?
	When is it appropriate to use a <b>pie chart</b> ?
	What is one advantage of a <b>bar graph</b> over a <b>pie chart</b> ?
	*Make sure you understand why a pie chart is NOT appropriate in the example on page 9-10 "Who Owns an MP3 Player"!!! People often try to use pie charts, even when they are not appropriate!
	What are the two problems that could arise with bar graphs illustrated by the example on p.10-11 "Who Buys iMacs"??

<b>Summary of p. 7-11</b>

**\*\*NOW Do p. 21-22 exercises 12, 14, 16, & 18**

<b>Reading Guide</b>	<b>Read p. 11-20 and complete this reading guide</b>
Define	<b>two-way table -</b>
	*Do you understand what <i>row variables</i> and <i>column variables</i> are?
	What is the first thing you should do when analyzing data presented in a two-way table?
Define	<b>marginal distribution (of one categorical variable) -</b>
	Describe a situation where you might use percents rather than counts.
Define	<b>conditional distribution (of a variable) -</b>
	How can you decide which conditional distributions to compare?
Define	<b>segmented bar graph -</b>
Define	<b>association (between two variables) -</b>
	If two variables have “no association,” what would be true about their conditional distributions? How would their segmented bar graphs look?

<b>Summary of p. 11-20</b>

**\*\*NOW Do p. 22-24 exercises 20, 22, 24, 27-34**

<b>Topic: 1.2 Displaying Quantitative Data with Graphs</b> <ul style="list-style-type: none"> <li>- Graphing Quantitative Data</li> <li>- Dotplots, Stemplots, &amp; Histograms</li> <li>- Describing Distributions of Quantitative Data</li> <li>- Comparing Distributions of Quantitative Data</li> <li>- Using Histograms Wisely</li> </ul>	<b>Name:</b>
<b>AP Topics</b> I. Exploring Data: describing patterns and departures from patterns <ul style="list-style-type: none"> <li>A. Constructing and interpreting graphical displays of distributions of univariate data (dotplot, stemplot, histogram, cumulative frequency plot) <ol style="list-style-type: none"> <li>1. Center and spread</li> <li>2. Clusters and gaps</li> <li>3. Outliers and unusual features</li> <li>4. Shape</li> </ol> </li> <li>C. Comparing distributions of univariate data (dotplots, back-to-back stemplots, parallel boxplots) <ol style="list-style-type: none"> <li>1. Comparing center and spread</li> <li>2. Comparing clusters and gaps</li> <li>3. Comparing outliers and unusual features</li> <li>4. Comparing shape</li> </ol> </li> </ul>	<b>Period:</b>  <b>Date:</b>
<b>Essential Questions:</b> <ul style="list-style-type: none"> <li>- What types of graphs can you use to picture the distributions of quantitative data?</li> <li>- When asked to describe a dotplot, stemplot, histogram, or boxplot, what are the key features you must include?</li> <li>- What is the difference between describing two distributions of quantitative data and comparing two distributions of quantitative data?</li> </ul>	

<b>Reading Guide</b>	<b>Read p. 25-33 and complete this reading guide</b>
Define	<b>dotplot -</b>
	What are the <u>3 steps</u> in making a <b>dotplot</b> ? <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> </ul>
	What is the purpose of graphing data?

<p><b>*THIS IS A KEY CONCEPT!!!</b> It applies to dotplots, stemplots, histograms, and boxplots.</p>	<p>What are the 4 key features you must address when asked to <b>describe the distribution of a quantitative variable</b>?</p> <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> <li>•</li> </ul>
<p>Define</p>	<p><b>mode –</b></p>
	<p>What should you look for when describing the <b>shape of a distribution</b>? What should you ignore?</p>
<p>Define</p>	<p><b>symmetric distribution -</b></p>
<p>Define</p>	<p><b>skewed to the right -</b></p>
<p>Define</p>	<p><b>skewed to the left -</b></p>
<p>Define</p>	<p><b>unimodal –</b></p> <p><b>bimodal –</b></p> <p><b>multimodal –</b></p>
<p><b>**AP EXAM TIP</b></p>	<p>What distinguishes <b>comparing</b> two distributions from <b>describing</b> two distributions? (Hint: Check out the AP Exam Tip on p. 30! These important and helpful tips are found throughout your book!)</p>



Define	<b>stemplot (also called a stem-and-leaf plot) -</b>
	What are the <u>4 steps</u> in making a <b>stemplot</b> ? <ul style="list-style-type: none"><li>•</li><li>•</li><li>•</li><li>•</li></ul>
	Why might you want to <b>split stems</b> ?
	Why would you use a <b>back-to-back stemplot</b> ?
	Tips to consider when using <b>stemplots</b> : <ul style="list-style-type: none"><li>•</li><li>•</li><li>•</li><li>•</li></ul>

**Summary of p. 25-33**

**\*\*NOW Do p. 41-43 exercises 38, 40, 42, 44, 46, 48**

<b>Reading Guide</b>	<b>Read p. 33-40 and complete this reading guide</b>
Define	<b>histogram -</b>
	What are the <u>3 steps</u> in making a <b>histogram</b> ? <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> </ul>
	How do you find the “center” or “midpoint” of a histogram?
	Tips to consider when making <b>histograms</b> : <ul style="list-style-type: none"> <li>•</li> <li>•</li> </ul>
	How are dotplots, stemplots, and histograms similar?

<b>Calculator Functions!!! Histograms</b>	*Try making a histogram for the data on foreign-born residents (see p. 33 for raw data) on your calculator following the instructions on p. 36-37 (If you have a TI-nspire, see <b>Appendix B</b> at the back of your book. It has all instructions for the TI-nspire.) <u>Note the AP Exam Tip at the end of the calculator instructions on p. 37.</u>
	List the three cautions (common errors) the book leaves you with regarding histograms? 1)  How are histograms different than bar graphs?  2)  3)

	<p>How do you choose whether to use a histogram, stemplot, or dotplot? What are the advantages and disadvantages of each?</p>
--	---

<b>Summary of p. 33-40</b>

**\*\*NOW Do p. 43-48 exercises 54, 58, 59, 60, 66, 69-74**

<b>Topic: 1.3 Describing Quantitative Data with Numbers</b> <ul style="list-style-type: none"> <li>- Calculating Measures of Center</li> <li>- Calculating and Interpreting Measures of Spread</li> <li>- Choosing the Most Appropriate Measures of Center and Spread</li> <li>- Identifying Outliers Using the 1.5 x IQR Rule</li> <li>- Making and Interpreting Boxplots</li> <li>- Using Appropriate Graphs and Numerical Summaries to Compare Distributions of Quantitative Data</li> </ul>	<b>Name:</b>
<b>AP Topics</b> <ul style="list-style-type: none"> <li>I. Exploring Data: describing patterns and departures from patterns</li> <li>B. Summarizing distributions of univariate data <ul style="list-style-type: none"> <li>1. Measuring center: median, mean</li> <li>2. Measuring spread: range, interquartile range, standard deviation</li> <li>3. Measuring position: quartiles, ...,</li> <li>4. Using boxplots</li> </ul> </li> </ul>	<b>Period:</b>  <b>Date:</b>
<b>Essential Questions:</b> <ul style="list-style-type: none"> <li>- How do you calculate median and mean?</li> <li>- How do you calculate and interpret range, interquartile range, and standard deviation?</li> <li>- How do you choose which measure of center and spread to use in describing quantitative data?</li> <li>- How do you identify outliers?</li> <li>- How do you make a boxplot?</li> <li>- What does a boxplot tell you about the data?</li> </ul>	

<b>Reading Guide</b>	<b>Read p. 48-59 and complete this reading guide</b>
	<b>Measures of Center</b>
Define Include formula	<b>mean (<math>\bar{x}</math> or <math>\mu</math>) -</b>
	What is the difference between $\bar{x}$ and $\mu$ ?

	What weakness of means does the Example “Travel Times to Work in North Carolina” illustrate?
Define	<b>Resistant measure -</b>
	What are two interpretations of mean the book mentions?
Define	<b>Median (M) -</b>
	How do you find the <b>median</b> of a distribution? 1)  2)  3)
	What is true about the median that is not true for the mean?
	When will the mean and median be exactly the same?  When will the mean and median be close together?  When will the mean lie to the right of the median?  When will the mean lie to the left of the median?

	<b>Measures of Spread</b>
	What are two things that must be included in a useful numerical description of data?
Define	<p><b>range -</b></p> <p>How is the statistical definition of range different than how the word “range” is often used in everyday language?</p>
Define	<p><b>quartile -</b></p> <p>What proportion/percent of the data lies within each quartile?</p>
	<p>How do you calculate quartiles <math>Q_1</math> and <math>Q_3</math>?</p> <ol style="list-style-type: none"> <li>1.</li> <li>2.</li> <li>3.</li> </ol>
Define *Note include the formulaic definition AND and the interpretation in words	<b>Interquartile range (IQR) -</b>
	Of range, $Q_1$ , $Q_3$ , and IQR, which is NOT a resistant measure of spread and why?
	*Make sure you understand how to interpret IQR as illustrated in the Example “Stuck in Traffic Again” on p.55!

	<b><u>Outliers</u></b>
Define	<b>The 1.5 x <i>IQR</i> rule for outliers -</b>
This is not explicitly listed step-by-step in the reading, but using the Example “Travel Times to Work in North Carolina” on p. 56, you should be able to work out steps.	List the steps for using 1.5 <i>IQR</i> rule for outliers to identify outliers? 1) 2) 3) 4) 5)
Define	<b>The five-number summary -</b>
Define	<b>boxplot (box-and-whisker plot) -</b>
	What are the <u>4 steps</u> in making a <b>boxplot</b> ? • • • •
	How are outliers identified in boxplots?
	Compare boxplots with histograms and stemplots?



	<p>When describing a <b>boxplot</b> what four features must you discuss?</p> <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> <li>•</li> </ul>
<p>Define Note: look up this definition online</p>	<p><b>parallel boxplot (side-by-side boxplot) -</b></p>
	<p>What do parallel boxplots allow you to do?</p>
<p><b>Calculator Functions!!! Boxplots</b></p>	<p>*Try making parallel boxplots for the data on travel time in North Carolina and New York (see p. 52 for both data sets) on your calculator following the instructions on p. 59 (If you have a TI-nspire, see <b>Appendix B</b> at the back of your book. It has all instructions for the TI-nspire.) <u>Note: ALWAYS use the boxplot function on the calculator that shows outliers separately!!!</u></p>

<p><b>Summary of p. 48-59</b></p>

**\*\*NOW Do p. 69-70 exercises 80, 82, 84, 86, 87, 90, 92, 93, 95**

<b>Reading Guide</b>	<b>Read p. 60-68 and complete this reading guide</b>
Define Include formula	<b>Standard deviation (<math>s_x</math>) -</b>
Define Include formula	<b>Variance (<math>s_x^2</math>)</b>
*You need to read the Example "How Many Pets?" on p. 60 to understand this.	Why do we need to square the deviations before adding them? (i.e. Why can't we just add deviations together?)
	List the steps to find the <b>standard deviation of <math>n</math> observations:</b> 1.  2.  3.
	What is the difference between $\sigma_x$ and $s_x$ ? Make sure to include when it is appropriate to use $\sigma_x$ vs $s_x$ .
	What does $s_x$ measure?
	Is $s_x$ ever less than zero?  When does it equal zero?
	How do you determine the units of measurement for $s_x$ ?

	Why isn't $s_x$ a resistant measure?
--	--------------------------------------

<b>Calculator Functions!!! Numerical Summaries (One-variable Stats)</b>	<p>*Try using your calculator to calculate numerical summaries for the data on travel time in North Carolina and New York (see p. 52 for both data sets) following the instructions on p. 63-64. This will save you a TON of time!!! (If you have a TI-nspire, see <b>Appendix B</b> at the back of your book. It has all instructions for the TI-nspire.) Note: at the end of the Technology Corner on p. 64, they introduce <b>minitab output</b>. It is VITAL that you understand how to read computer output for the AP EXAM!!! <u>They also explain potential differences in various types of computer output in the "Think About It" on p. 64.</u></p>
	<p>When should you use median and IQR?</p> <p>When should you use mean and <math>s_x</math>?</p>
	<p>What should you ALWAYS include along with numerical summaries? Why?</p>
	<p><b>Organizing a Statistics Problem</b> Copy the <b>four-step process</b> outlined by your book.</p> <ul style="list-style-type: none"> <li>•</li> <li>•</li> <li>•</li> <li>•</li> </ul>

<b>Summary of p. 60-68</b>

**\*\*NOW Do p. 71-73 exercises 98, 100, 104, 106, 107-110**