GRACE
CHRISTIAN SCHOOL

AP Statistics
2019 Summer Assignment
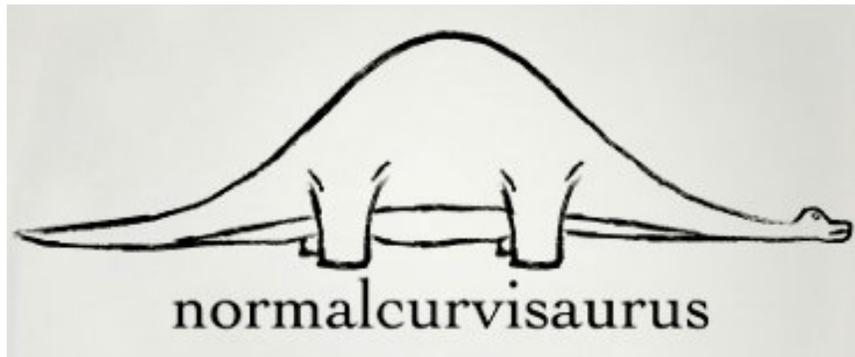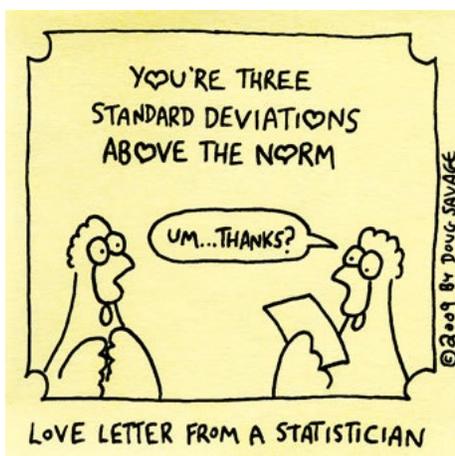
Meagan Stone
mstone@gracechristian.net

Adapted from Bob Cashill's (Verona High School) AP Statistics Summer Assignment

AP Statistics 2019-2020

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-84 or TI-84 Plus, and, most importantly, be willing to explain your answers, not just simply get the correct answer. The textbook for this class is *The Practice of Statistics* (fifth edition) by Starnes, Tabor, Yates, and Moore.We will also heavily utilize JMP, statistical software that will allow us to visualize data and make statistical discoveries.

This course is a college-level math course that moves at a fast pace. A considerable amount of reading and writing is required. 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 calculations 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?

During the second semester, you all will participate in a semester-long a project that requires you to incorporate these areas of concentration. You will work as a statistician for a charity or non-profit of your choice. You will brainstorm with representatives of the organization to generate questions about their organization that they want answered using statistics. You will then gather data and perform the appropriate tests. The project will culminate with you all completing a written research report and presenting your findings formally to the class and to your organization.

AP Statistics 2019 Summer Assignment

This summer assignment is meant to get you started on introductory topics, such as how statistics is useful course of study, displaying data, and reading displays so that we can spend more time on the difficult material during the school year.

Please feel free to email me over the summer if you have questions or run into difficulties with completing the work.

A. **Buy a Graphing Calculator**

It is strongly recommended that you buy a ***TI-84 Plus*** calculator if you do not already have one. All of the examples and problems we will work on will be based on this calculator. You will be using your calculator extremely often and familiarity with it is easily worth several points on the AP Exam. Take a picture of you with your calculator and submit it.

B. **Visit CollegeBoard**

Become familiar with the AP Statistics resources on **the CollegeBoard website**.

C. **Watch (and Take Notes on) The Joy of Stats**

The Joy of Stats is a 59 minute long documentary which takes viewers on a rollercoaster ride through the wonderful world of statistics to explore the remarkable power they have to change our understanding of the world. It is presented by Professor Hans Rosling, whose eye-opening, mind-expanding and funny online lectures have made him an international internet legend.

Rosling's presentations are grounded in solid statistics illustrated by the visualization software he developed. The animations transform development statistics into moving bubbles and flowing curves that make global trends clear, intuitive, and even playful. During his legendary presentations, Rosling takes this one step farther, narrating the animations with a sportscaster's flair.

You can view the video **here**.

Please complete **Attachment A** as you watch the video.

D. **Create a Real World Statistics Portfolio**

The beauty of statistics is that it is all around us. We see examples of good and bad stats every single day in newspapers and magazines, on newscasts, during sporting events, and on a host of web sites--especially those dealing with politics, the economy, and the government.

To really see how statistics impacts our lives, a major part of your summer assignment is to develop a **Real World Statistics Portfolio**.

Begin to collect newspaper, magazine, video or internet articles/stories/advertisements that include statistical concepts--both good and bad. These may include things like graphs, charts or tables. They may also report conclusions made as a result of looking at data—e.g., newspaper/internet reports on drug/social/psychology studies. Depending on what kinds of articles you all choose, we may use them throughout the year as examples.

One goal of this assignment is to encourage you to read an article objectively and draw conclusions based on what the statistics are actually telling you or to notice gaps or shortcomings in data. The

sample article in **Attachment B** is about a topic that is often very emotionally charged. The goal is to read this without letting your emotions interpret the statistics.

For each article/web page/etc., highlight the statistics mentioned and answer the following questions:

a. What is the source of the data? When and where was it published?
b. What was the purpose of the article/statistic/data? Why was it written/produced?
c. Were any conclusions stated? If so, what were they?
d. Are the article/statistics/data convincing? Do you believe the stated results? Explain.

You will submit three articles over the summer. You must submit each article and the responses to the questions via Talon by the due date. Each article must be current— that is, it must have appeared in the news on or after May 1st, 2019.

Good resources for statistics articles/charts include:
- The Newark Star-Ledger
- The New York Times
- The Washington Post
- 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/)

Attachment B is a sample of an article and responses. Each article must be emailed to me by the due date with the questions and your answers to the questions attached.

E. **Chapter 1 Reading Guide**

Read Chapter 1 and complete the Chapter 1 Reading Guide (See **Attachment C**). You should complete each section of the study guide and submit it via email by its due date. You may type the answers to the reading guide. Remember that this is to be your work only. You should not collaborate with other students. The answers should be located in the reading. (Check out a book for the summer.)

AP Statistics 2019 Checklist

By submitting these assignments, you are guaranteeing that all work was completed using appropriate resources (i.e., your instructor, the textbook, searching the internet for terms you don't understand) and is your own. You are pledging that you have not copied assignments from other students or searched for exact solutions online.

Assignment	Due Dates
A. Buy/Own a Graphing Calculator. (Take a picture of you with it and submit!) We will have a calculator check on the first day of school.	Monday, July 8
B. Visit the AP Statistics page on Talon. Under Lessons, choose Chapter 1 Notes and Resources and explore the resources available there.	Monday, July 8
C. Watch "The Joy of Statistics" and turn in Attachment A.	Monday, July 15
D. Create a Real World Statistics Portfolio (three articles total). (See Attachment B for example.)	Monday, July 29
E. Read Chapter 1 in *The Practice of Statistics, ed. 5* and complete the Chapter 1 reading guide. (See Attachment C.)	Monday, July 29

Attachment A — "The Joy of Statistics"

Hans Rosling — http://www.gapminder.org/videos/the-joy-of-stats/

As you watch the video, please fill in the blanks in the following statements.

1. Sweden has _____ boats, which means one out of _____ people own a boat.

2. The average human sleeps _____ years of their life.

3. Hans's official area of expertise is _____.

4. The _____ and _____ supplies a large amount of the data that Hans uses in his presentations.

5. Chris Vien is the _____ for San Francisco.

6. The word statistics comes from the word _____.

7. Statistics were started in _____ in 1749 when they started recording information on _____ and _____.

8. These data were published in the _____.

9. Sweden estimated they had a population of _____ million until they started collecting data and found they only had a population of _____ million.

10. It took _____ years before the other countries in Europe started collecting and using statistics.

11. The early term for statistics was _____.

12. David Speigelhalter talked about the _____ Society of London & was intrigued that in 1842 they kept statistics on _____ rates.

13. He noted that _____ had the highest rates.

14. Babbage wrote Tennyson about a line in his poem "_____ and _____" indicating that the correct number should be _____.

15. In Hans's 200 year analysis, he uses _____ and _____ as his two variables.

16. Hans was born in _____.

17. For Sweden residents, the average number of legs is _____ legs/person.

18. Francis Galton came up the concept/ realization of the _____.

19. The Poisson distribution describes the number of Prussian Officers _____ by their horses each year.

20. Florence Nightingale started collecting data on _____ and _____ at age nine.

21. During the war she collected mortality data & graphed the data using _____.

22. By comparing diets of Americans to the Japanese & French, the joke goes that heart attacks don't kill you, _____ kills you.

23. In the 1950's, Sir Richard Doll first showed the correlation between _____ and _____.

24. Data is the _____ of science.

25. _____ kicked off the first data explosion.

26. A single letter is about _____ of data.

27. It was estimated that the internet will contain _____ of data by 2010.

28. Peter Norvig is the _____ at Google.

29. Hans is very intrigued and interested in Google's _____.

30. Sweden's goal in the Olympics is not to win, but to _____.

31. The Sloan Sky Surveyor is located at an observatory in _____.

32. Sep Kamvar of Stanford is working on a _____ database.

33. Women feel _____ more, but also feel _____.

34. Men feel _____ more, but feel more _____.

35. Young people associate happy with _____.

36. Older people associate happy with _____.

<u>**Attachment B — Real World Statistics Portfolio**</u>

Marijuana and race: ACLU finds blacks in New Jersey arrested at 3 times the rate of whites
By Star-Ledger Staff on June 04, 2013 at 10:20 AM, updated June 05, 2013 at 10:35 AM

By Seth Augenstein and James Queally

Black people in the U.S. are nearly four times more likely to be arrested for marijuana possession than white people - despite comparable usage rates, according to an American Civil Liberties Union report released today.

Blacks are nearly three times more likely to be arrested for marijuana possession than whites in New Jersey, and roughly four times as likely nationwide, according to a report released today by the American Civil Liberties Union that showed a deep racial divide in arrests in several counties throughout the state.

While the two races use marijuana at roughly the same rate, the report said, blacks were up to 30 times more likely to be arrested in some parts of the country. In two Alabama counties, for example, 100 percent of those arrested for marijuana possession in 2010 were black, the report said.

Nationwide, blacks were arrested at a rate of 716 per 100,000 in 2010 — the most recent year studied — up from 537 in 2001, according to the report. Whites were arrested at a rate of 192 per 100,000 in 2010, nearly the same as in 2001. The report also said marijuana arrests accounted for more than half of all drug arrests in the country, and roughly 43 percent in New Jersey.

The report was compiled using data from FBI's Uniform Crime Reports and the U.S. Census. It does not specify how police came to make the arrests or if the arrests were made in conjunction with other crimes. However, the ACLU said the arrests were not connected to violent crime.

Alexander Shalom, policy counsel for the state chapter of the ACLU, said the racial divide could be deeper if adjusted for the Hispanic population, which is counted as white in FBI reports.

Still, the data is clear evidence that police across the country unfairly target minorities in order to pump up arrest totals, according to one of the report's authors.

"The war on marijuana has disproportionately been a war on people of color," said Ezekiel Edwards, director of the ACLU's Criminal Law Reform Project.

According the report, blacks in New Jersey were 2.8 times more likely to be arrested for marijuana possession than whites, below the national average of 3.73.
Several counties reported deep disparities in arrest rates.
In Hunterdon County, where more than 90 percent of the population is white, blacks were five times as likely to be arrested for marijuana possession, the report said. Five other counties — Ocean, Monmouth, Mercer, Warren and Salem counties — had disparities higher than the national average, the report said.
Law enforcement officials were quick to note the limitations of the study.

Toms River Police Chief Michael Mastronardy said it would be difficult to determine racism as a motive for the disparity without knowing how police came to make those arrests. Marijuana arrests are often secondary to motor vehicle stops or domestic disturbances, situations in which officers are dispatched without knowing the race of the people involved.

"If you're stopping somebody at night, you don't know (the race)," he said, adding the same is true " if you're called to a disturbance and you start running warrants."

John Kuczynski, chief of detectives for the Hunterdon County Prosecutor's Office, shot down any suggestion police in his county were making arrests based on race. He said the bulk of marijuana arrests in the county are the result of motor-vehicle stops, but added that he is willing to discuss the report's findings with the ACLU.

"It's not based upon race," he said. "It's based upon offense whoever is trafficking."

The state Attorney General's Office declined to comment. Calls to law enforcement leaders in Monmouth, Salem and Warren counties were not returned.

A graphic from the ACLU report "The War on Marijuana in Black and White" shows how black people are 2.8 times likelier to be arrested for marijuana than white people in New Jersey. Some of the counties with the higher disparities are rural areas, like Hunterdon, Warren and Salem counties.

Peter Moskos, a criminologist at the John Jay School of Criminal Justice in New York, said the data was a product of "racism without racists." Drug task forces formed to combat the crack and heroin boom decades ago have shifted their focus to marijuana in order to maintain funding from national sources, he said.

"Once you have the apparatus, you're not going to care what the drug is," he said.

A 2012 Star-Ledger report showed heroin, not marijuana, was the problematic drug in the three New Jersey counties with the worst arrest disparities — Hunterdon, Ocean and Monmouth.

The report also found states spent an estimated $3.61 billion enforcing marijuana possession laws in 2010 alone. In New Jersey, that figure was $127.3 million per year, a number which drew the ire of pro-marijuana advocates.

"You could build a lot of schools with that," said Sen. Nicholas Scutari (D-Union) who introduced legislation to de-criminalize marijuana earlier this year.

Roseanne Scotti, the director of the state chapter of the Drug Policy Alliance, called the report's findings "morally appalling and fiscally ridiculous."
"I think when you look at some of the counties like that, it's hard to imagine or find any way to reason your way out of the fact that it could be anything other than intentional discrimination," she said. "The numbers don't lie."

---

a. **What is the source of the data? When and where was it published?**

NJ.com; New Jersey on 6/4/10.

b. **What was the purpose of the article/statistic/data? Why was it written/produced?**

The purpose of the article was to draw attention to the disparate rates of marijuana arrests in New Jersey between blacks and whites. The study purported to show that blacks are more likely than whites in many parts of the country to be singled out for arrest.

c. **Were any conclusions stated? If so, what were they?**

Primary conclusion is that race plays a significant part in marijuana arrests.

d. **Is the article/statistics/data convincing? Do you believe the stated results? Explain.**

It would be nice to know what percentage of marijuana arrests occur after traffic stops or disturbances when the officer didn't initially know the race. The proportion of blacks that are arrested for possession of marijuana after a traffic stop or disturbance versus the proportion of whites that are arrested for possession of marijuana after a traffic stop or disturbance would also affect the way that I view this article. (There are many more conclusions we could draw from this article. This is just a sample response. You should include as much as you can about the statistics in the article. Be as specific as possible. You don't have to absolutely believe or not believe the conclusions stated, but give reasons for any hesitations you have about the statistics in the article.)

Chapter 1: Exploring Data

Key Vocabulary:

- individual
- variable
- frequency table
- relative frequency table
- distribution
- pie chart
- bar graph
- two-way table
- marginal distributions
- conditional distributions
- side-by-side bar graph

- association
- dotplot
- stemplot
- histogram
- SOCS
- outlier
- symmetric
- Σ
- $\overline{x}$
- spread
- variability

- median
- quartiles
- Q_1, Q_3
- IQR
- five-number summary
- minimum
- maximum
- boxplot
- resistant
- standard deviation
- variance

Data Analysis: Making Sense of Data (pp.2-6)

1. *Individuals* are…

2. A *variable* is …

3. When you first meet a new data set, ask yourself:

    • Who…

    • What…

    • Why, When, Where and How…

4. Explain the difference between a *categorical* variable and a *quantitative* variable. Give an example of each.

5. Give an example of a categorical variable that has number values.

6. Define *distribution*:

7. What are the four steps to *exploring data*?

   • Begin by….

   • Study relationships…

   • Start with a …

   • Then add…

8. Answer the two questions for the *Check Your Understanding* on page 4*:*

9. Define *inference.*

Analyzing Categorical Data (pp.8-22)

1. A *frequency* table displays…

2. A *relative frequency table* displays…

3. What type of data are *pie charts* and *bar graphs* used for?

4. *Categories* in a bar graph are represented by _____ and the *bar heights* give the category _____.

5. What is a *two-way table*?

6. Define *marginal distribution*.

7. What are the two steps in examining a marginal distribution?

8. Answer the two questions for the *Check Your Understanding* on page 14.

9. What is a *conditional distribution*? Give an example demonstrating how to calculate one set of conditional distributions in a two-way table.

10. What is the purpose of using a *segmented bar graph*?

11. Answer question one for the *Check Your Understanding* on page 18.

12. Describe the four steps to organizing a statistical problem:

    •   State…
    •   Plan…

    •   Do…

    •   Conclude…

13. Explain what it meant by an *association* between two variables.

Analyzing Categorical Data (pp.27-42)

1. What is a *dotplot*? Draw an example.

2. When examining a distribution, you can describe the overall pattern by its

   S_____   O_____   C_____   S_____

3. If a distribution is *symmetric*, what does it look like?

4. If a distribution is *skewed to the right*, what does it look like?

5. If a distribution is *skewed to the left*, what does it look like?

6. Describe and illustrate the following distributions:

   a. Unimodal

   b. Bimodal

   c. Multimodal

7. Answer questions 1-4 for the *Check Your Understanding* on page 31.

8. How are a *stemplot* and a *histogram* similar?

---

9. When is it beneficial to *split the stems* on a stemplot?

10. When is it best to use a *back-to-back stemplot*?

11. List the three steps involved in making a histogram.

12. Why is it advantageous to use a relative frequency histogram instead of a frequency histogram?

13. Answer questions 2-4 for the *Check Your Understanding* on page 35.

Analyzing Categorical Data (pp.50-67)

1. What is the most common *measure of center*?

2. Explain how to calculate the *mean,* $\bar{x}$.

3. What is the meaning of $\sum$?

4. Explain the difference between $\bar{x}$ and μ.

5. Define *resistant measure*.

6. Explain why the mean is not a resistant measure of center.

7. What is the *median* of a distribution? Explain how to find it.

8. Explain why the median is a resistant measure of center?

9. How does the shape of the distribution affect the mean and median?

10. What is the *range*?

11. Is the range a resistant measure of spread? Explain.

12. How do you find *first quartile* Q_1 and *third quartile* Q_3?

13. What is the *Interquartile Range* (IQR)?

14. Is the IQR and the quartiles a resistant measure of spread? Explain.

15. How is the IQR used to identify *outliers*?

16. What is the *five-number summary* of a distribution?

17. Explain how to use the five-number summary to make a *boxplot*.

18. What does the *standard deviation* measure? How do we calculate it?

19. What is the relationship between *variance* and *standard deviation*?

20. What are the *properties* of the standard deviation as explained on page 64?

21. How should one go about choosing measures of center and spread?