AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak

1. What is statistics?
2. What is variability?

The study of variability.
Differences, how things differ. There is variability everywhere. We all look different, act different, have different preferences. Statisticians look at these differences.

Inferential and Descriptive
Tell me what you got! Describe to me the data that you collected, use pictures or summaries like mean, median, range, etc.
5. What are INFERENTIAL STATS?

Look at your data, and use that to say stuff about the BIG PICTURE. Like tasting soup, a little sample can tell you a lot about the big pot of soup (the population).
6. Compare Descriptive and Inferential Descriptive explains you about the STATS

## 7. What is data?

8. What is a population?
data that you have, inference uses that data you have to try to say something about an entire population.

Any collected information. Generally, each little measurement. Like, if it is a survey about liking porridge: the data might be "yes, yes, no, yes, yes." If it is the number of saltines someone can eat in 30 seconds: the data might be "3, 1, 2, 1, 4, 3, 3, 4."

The group you're interested in. Sometimes it's big, like "all teenagers in the US." Other times it is small, like "all AP Stats students in my school."
9. What is a sample?

AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak
10. Compare population to sample
11. Compare data to statistics
13. What is a parameter?

## 12. Compare data to parameters

A numerical summary of a population. Like a mean, median, range, etc. of a population.
14. What is a statistic?

A numerical summary of a sample. Like a mean, median, range, etc. of a sample.

AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak
15. We are curious about the average wait time at a Dunkin Donuts drive through in your neighborhood. You randomly sample cars one afterYou randomly sample cars one after- never know. The statistic is " 3.2 minnoon and find the average wait time is 3.2 minutes. What is the population parameter? What is the statistic? What is the parameter of interest? What is the data?

The parameter is the true average wait time at that Dunkin Donuts. This is a number you don't have and will utes." It is the average of the data you collected. The parameter of interest is the same thing as the population parameter. In this case, it is the true average wait time of all cars. The data is the wait time of each individual car, so that would be like " 3.8 min , 2.2 min , $0.8 \mathrm{~min}, 3 \mathrm{~min}$. " You take that data and find the average, that average is called a "statistic," and you use that to make an inference about the true parameter.
16. Compare DATA-STATISTIC-PARAMETER using a categorical example

Data are individual measures, like meal preference: "taco, taco, pasta, taco, burger, burger, tacos." Statistics and Parameters are summaries. A statistic would be "42\% of sample preferred tacos" and a parameter would be "42\% of population preferred tacos."

## 17. Compare DATA-STATISTIC-PARAMETER using a quantitative example

Data are individual measures, like how long a person can hold their breath: " $45 \mathrm{sec}, 64 \mathrm{sec}, 32 \mathrm{sec}, 68$ sec." That is the raw data. Statistics and parameters are summaries like "the average breath holding time in the sample was 52.4 seconds" and a parameter would be "the average breath holding time in the population was 52.4 seconds."
18. What is a census?

AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak
19. Does a census make sense?
20. What is the difference between a parameter and a statistic?
21. If I take a random sample 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them, and one of them had 9 pickles, then the number 9 from that burger would be called $\qquad$ ?
22. If I take a random sample 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them, and the average number of pickles was 9.5 , then 9.5 is considered a $\qquad$ ?
23. If I take a random sample 20 hamburgers from FIVE GUYS and count the number of pickles on a bunch of them... and I do this because I want to know the true average number of pickles on a burger at FIVE GUYS, the true average number of pickles is considered a $\qquad$ ?
24. What is the difference between a sample and a census?

Like a sample of the entire population, you get information from every member of the population.

A census is ok for small populations (like Mr. Martinez's students) but impossible if you want to survey "all US teens."

BOTH ARE A SINGLE NUMBER
SUMMARIZING A LARGER GROUP OF NUMBERS. But parameter come from populations. Statistics come from statistics.
a datum, or a data value.
statistic
parameter, a one number summary of the population. The truth. AKA the parameter of interest.

With a sample, you get information from a small part of the population.

AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak
25. Use the following words in one sentence: population, parameter, census, sample, data, statistics, inference, population of interest.
26. If you are tasting soup.. Then the flavor of each individual thing in the spoon is the $\qquad$ , the entire spoon is a $\qquad$ . The flavor of all of that stuff together is like the and you use that to about the flavor of the entire pot of soup, which would be the $\qquad$ -
27. What are random variables?
28. What is the difference between quantitative and categorical variables?

In a census, you get information from the entire population. You can get a parameter from a census, but only a statistic from a sample.

I was curious about a population parameter, but a census was too costly so I decided to choose a sample, collect some data, calculate a statistic and use that statistic to make an inference about the population parameter (aka the parameter of interest).

If you are tasting soup.. Then the flavor of each individual thing in the spoon is the DATA, the entire spoon is a SAMPLE. The flavor of all of that stuff together is like the STATISTIC and you use that to MAKE AN INFERENCE about the flavor of the entire pot of soup, which would be the PARAMETER. Notice you are interested in the parameter to begin with, that is why you took a sample.

If you randomly choose people from a list, then their hair color, height, weight and any other data collected from them can be considered random variables.

Quantitative variables are numerical measures, like height and IQ. Categorical are categories, like eye color and music preference.

[^0]AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak
29. What is the difference between quantitative and categorical data?
30. What is the difference between discrete and continuous variables?
31. What is a quantitative variable?
32. What is a categorical variable?
33. What do we sometimes call a categorical variable?
34. What is quantitative data?
35. What is categorical data?

The data is the actual gathered measurements. So, if it is eye color, then the data would look like this: "blue, brown, brown, brown, blue, green, blue, brown, etc." The data from categorical variables are usually words. If it was weight, the the data would be quantitative like: "125, 155, 223, 178, 222, etc." The data from quantitative variables are numbers.

Discrete can be counted, like "number of cars sold" they are generally integers (you wouldn't sell 9.3 cars), while continuous would be something like weight of a mouse (4.344 oz).

Quantitative variables are numeric like: Height, age, number of cars sold, SAT score.

Qualitative variables are like categories: Blonde, Listens to Hip Hop, Female, yes, no... etc.
qualitative

The actual numbers gathered from each subject: 211 pounds or 67 beats per minute.

The actual individual category from a subject, like "blue" or "female" or "sophomore"

AP Stats Summer Vocab
Study online at https://quizlet.com/_bl3nak

| Categorical data |  |
| :--- | :---: |
| Color | Number <br> of toys |
| Brown | 2 |
| Yellow | 5 |
| Red | 4 |
| Blue | 3 |
| Green | 6 |

36. What is a random sample?
37. What is frequency?
38. data or datum?
39. What is a frequency distribution?
40. What is meant by relative frequency?

When you choose a sample by rolling dice, choosing names from a hat, or other RANDOMLY generated sample. Humans can't really do this well without the help of a calculator, cards, dice, or slips of paper.

How often something comes up.
Datum is singular. Like "hey dude, come see this datum I got from this rat!" Data is the plural. Like "hey look at all that data Edgar got from those chipmunks over there!"

A table, or a chart, that shows how often certain values or categories occur in a data set.

| Class (Rs.) | Tally Marks | Frequency <br> students |
| :---: | :--- | :---: |
| $20-30$ | HH | 5 |
| $30-40$ | HH III | 8 |
| $40-50$ | HH III | 9 |
| $50-60$ | HH HH | 10 |
| $60-70$ | HH I | 6 |
| $70-80$ | II | 2 |
| Total |  | $\mathbf{4 0}$ |

The PERCENT of time something comes up (frequency/total).
41. How do you find relative frequency? Just divide frequency by TOTAL.
42. What is meant by cumulative frequency?

ADD up the frequencies as you go. Suppose you are selling 25 pieces of

Study online at https://quizlet.com/_bl3nak candy. You sell 10 the first hour, 5 the
second, 3 the third and 7 in the last hour, the cumulative frequency would be $10,15,18,25$.
43. Make a guess as to what relative cu- It is the ADDAD up PERCENTAGES. mulative frequency is.

An example is selling candy, 25 pieces sold overall, with 10 the first hour, 5 the second, 3 the third, and 7 the fourth hour. We'd take the cumulative frequencies, $10,15,18$, and 25 and divide by the total giving cumulative percentages: $0.40,0.60$. 0.72 , and 1.00 . Relative cumulative frequencies always end at 100 percent.
44. What is the difference between a bar Bar charts are for categorical data chart and a histogram? (bars don't touch) and histograms are for quantitative data (bars touch).

45. What is the mean?

The old average we used to calculate. It is the balancing point of the histogram.
46. What is the difference between a population mean and a sample mean?

Population mean is the mean of a population, it is a parameter. Sample mean is a mean of a sample, so it is a statistic. We use sample statistics to make inferences about population parameters.
47. What symbols do we use for popula- Mu for population mean (parameter), tion mean and sample mean?
x-bar for sample mean (statistic).
Population Mean $=\mu$
Sample Mean $=\bar{X}$

How can you think about the mean and median to remember the difference when looking at a histogram?

Mean is balancing point of histogram, median splits the area of the histogram in half.
49. What is the median?

The middlest number, it splits area in half (always in the POSITION $(n+1) / 2)$.

Find the median of 5, 8, 4, 1, 2

50. What is the mode?
51. When do we often use mode?
52. Why don't we always use the mean, It is not RESISTANT, it is impacted by we've been calculating it all of our life?
53. When we say "the average teenager" It depends, if we are talking height, it 54. are we talking about mean, median or mode?

## AP Stats Summer Vocab

Study online at https://quizlet.com/_bl3nak
m al income, we'd probably use the
i median, if we were talking about
g music preference, we'd proba- bly
h use the mode to talk about the
t average teenager.
b
e
t
h
e
m
e
a
n

What is a clear example of where the mean would change but median wouldn't? (this would show how the median is resistant)
55. How are mean, median and mode positioned in a skewed left histogram?
56. How are mean, median, and mode positioned in a skewed right histogram?

Imagine if we asked eight people how much money they had in their wallet. We found they had (1, 2, 2, 5, 5, 8, $8,9)$. The mean of this set is 5 , and the median is also 5 . You might say "the average person this group had 5 bucks." But imagine if one of them just got back from the casino, and instead it was (1, 2, 2, 5, 5, 8, 8, 9000), in this case, the median would still be 5 , but the mean goes up to over 1000 . Which number better describes the amount of money the average person in the group carries, 5 bucks or 1000 bucks? I think 5 is a better description of the average person in this group and the 9000 is simply an outlier.

Goes in that order from left to right. Mean-median-mode.


Goes in the opposite order. Mode-me-dian-mean.


The mean chases the tail, the mean chases the tail, high-ho the derry-oh the mean chases the tail... and outliers.


[^0]:    A categorical variable takes on values

