



AP Statistics

Summer Math Packet

Course 1372

AP Statistics Summer Assignment

Welcome to AP Statistics future statisticians! The purpose of this assignment is to make you more comfortable exploring data analysis.

What can I expect from AP Statistics?

AP Statistics is a very unique and extremely rewarding class. While it is primarily a Math class, it also combines skills from English and Science classes. Clear communication skills are pivotal for this course; thus, you will find much more reading and writing in this class compared to what you are used to in a math class. On the AP exam, you will be expected to write descriptive paragraphs and draw conclusions from data, while explaining your reasoning for the methods you choose and the conclusions you draw.

What calculator do we use?

We use the TI-84 Plus CE graphing calculator. Every student enrolled in the course must have their own and is expected to bring it fully charged every day to class.

What is the purpose of this packet?

With the AP exam being in early May, it can be difficult to complete the required curriculum and include a much desired review by the end of April. This packet will address key prerequisite skills as well as instill statistical thinking into your minds. Additionally, this packet will introduce you to the capabilities of your graphing calculator and begin to familiarize you with the tools needed to succeed in AP Statistics.

Reminder:

AP Statistics, like any Advanced Placement class, is a college level course; the expectation is that you come to class **prepared**, **self-motivated**, and **hungry to learn**.

The summer assignment is composed of four parts.

1. Reading and Vocabulary: You will use a free online Statistical tutoring site that will give you information on variable and data displays. While reviewing the information on the site you will be completing a vocabulary list (See pages 3 and 4). You should print and do so in the spaces provided. Electronic versions will NOT be accepted. This is to be turned in on the first day of class. Follow the steps below:
 - a. Go to www.stattek.com
 - b. Click on "AP Statistics" then "AP Tutorial"
 - c. On the left side of the screen is a list of general topics. Under each general topic are a list of subtopics. You will read the following subtopics to complete the vocabulary list.

Topic: Exploring Data	
Subtopic: The basics	Variables
	Population vs sample
	Mean and median
	Variability
	Position

Subtopic: Charts and graphs	Patterns in data
	Dotplots
	Bar charts and histograms
	Stemplots
	Boxplots
	Cumulative plots
	Scatterplots
	Comparing datasets

2. Khan Academy: After familiarizing yourself with the vocabulary, you are to complete the Khan Academy course called “Get Ready for AP Statistics” using this link:
<https://www.khanacademy.org/join/FKKPZYQ7>

Keep in mind:

- Course is due by Monday, August 11, 2025 at 11:59pm
- If you don't link yourself to this course, you will not receive credit.
- Assessment on this material will be administered at the beginning of the school year.

3. Practice Problems: After completing #1 & 2, you should be able to complete the questions in the remaining pages of this packet. You should print and do so in the spaces provided. *Electronic versions will NOT be accepted.* This is to be turned in on the first day of class.

This packet should be printed... no electronic versions will be accepted. It should be completed by your return to school on **Tuesday, August 12, 2025.** You are expected to complete each part of each problem and to construct all data displays neatly. This assignment will be graded, and it will be assessed in the first grading period of the school year.

AP Statistics Summer Assignment

Part 1: Vocabulary List

Please define each of the following terms from the information on the [stattrek](http://stattrek.com) website. When asked to provide a UNIQUE example or sketch of the word, please provide one... One NOT given on the website and NOT the one your friends use.

1. Categorical Variables

Example:

2. Quantitative Variables

Example:

3. Discrete Variables

4. Continuous

5. Univariate Data

6. Bivariate Data

7. Population

Example:

8. Sample

Example:

9. Median

10. Mean

Formula:

11. Outlier

12. Parameter

13. Statistics

14. Range

15. Standard Score (z-score)

Formula:

16. Center

17. Spread

18. Variance

Formula:

19. Standard Deviation

Formula:

20. Symmetry

Sketch:

21. Unimodal

Sketch:

22. Bimodal

Sketch:

23. Skewness

Sketch Skewed left:

Sketch Skewed right:

24. Uniform

Sketch

25. Gaps

Sketch:

26. Outliers

Sketch:

27. Dot plots

28. Bar chart

29. Histogram

30. Difference between bar chart and histogram

31. Stemplots

32. Boxplots

33. Quartiles

34. Range

35. Interquartile Range

36. Four ways to describe data sets

37. Types of graphs that can be used for comparing data

AP Statistics Summer Assignment

Part 2: Practice Problems

CATEGORICAL OR QUANTITATIVE

Determine if the variables listed below are *quantitative* or *categorical*.

1. Time it takes to get to school
2. Number of people under 18 living in a household
3. Hair color
4. Temperature of a cup of coffee
5. Teacher salaries
6. Gender
7. Smoking
8. Height
9. Amount of oil spilled
10. Age of Oscar winners
11. Type of Depression medication
12. Jellybean flavors
13. Country of origin
14. Type of meat
15. Number of shoes owned

STATISTIC—WHAT IS THAT?

A statistic is a number calculated from 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	
IQR	

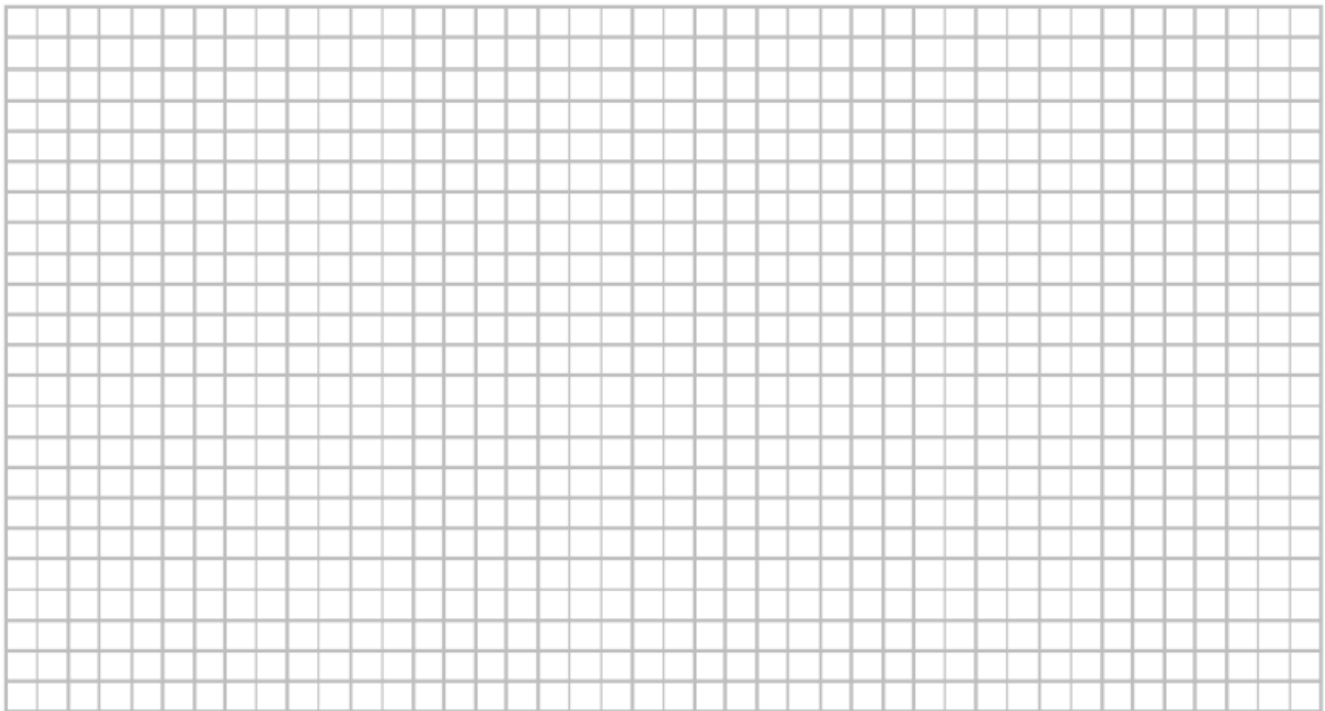
ACCIDENTAL DEATHS

In 1997 there were 92,353 deaths from accidents in the United States. Among these were 42,340 deaths from motor vehicle accidents, 11,858 from falls, 10,163 from poisoning, 4051 from drowning, and 3601 from fires. The rest were listed as “other” causes.

- a. Find the percent of accidental deaths from each of these causes, rounded to the nearest percent.

- b. What percent of accidental deaths were from “other” causes?

- c. NEATLY create a well-labeled **bar graph** of the distribution of causes of accidental deaths. Be sure to include an “other causes” bar.



- d. A pie chart is another graphical display used to show all the categories in a categorical variable relative to each other. Create a pie chart for the accidental death percentages. You may try using a software or internet source to make one and paste in the space below. (*Microsoft Excel works well*)

IT'S A TWISTA

The data below gives the number of hurricanes that happened each year from 1944 through 2000 as reported by *Science* magazine.

3	2	1	4	3	7	2	3	3	2	5	2	2	4	2	2	6	0	2	5	1	3	1	0
3	2	1	0	1	2	3	2	1	2	2	2	3	1	1	1	3	0	1	3	2	1	2	1
1	0	5	6	1	3	5	3																

- a. Make a dotplot to display these data. Make sure you include appropriate labels, title, and scale. The graph paper should help ensure you space your markings (you may use x's or dots) consistently.



SHOPPING SPREE!

A marketing consultant observed 50 consecutive shoppers at a supermarket. One variable of interest was how much each shopper spent in the store. Here are the data (round to the nearest dollar), arranged in increasing order:

3	9	9	11	13	14	15	16	17	17
18	18	19	20	20	20	21	22	23	24
25	25	26	26	28	28	28	28	32	35
36	39	39	41	43	44	45	45	47	49
50	53	55	59	61	70	83	86	86	93

a. Make a stemplot using tens of dollars as the stem and dollars as the leaves. Make sure you include appropriate labels, title and key.



SSHA SCORES

Here are the scores on the Survey of Study Habits and Attitudes (SSHA) for 18 first-year college women:

154 109 137 115 152 140 154 178 101 103 126 126 137 165 165
 129 200 148

and for 20 first-year college men:

108 140 114 91 180 115 126 92 169 146 109 132 75 88 113
 151 70 115 187 104

- a. Put the data values in order for each gender. Compute numeral summaries for each gender.

Women		Men	
Mean		Mean	
Minimum		Minimum	
Q1		Q1	
Median		Median	
Q3		Q3	
Maximum		Maximum	
Range		Range	
IQR		IQR	

- b. Using the minimum, Q1, Median, Q3, and Maximum from each gender, make parallel boxplots to compare the distributions.



ALGEBRA PAGE!

The prerequisite for AP Statistics is Algebra II. You will not find very much equation solving in this course, but some quick review of Algebra I and Algebra II content will be helpful.

To answer the following refer to the readings on www.stattek.com "Survey Sampling Methods".

The 7 types of sampling designs are:

- A. voluntary response B. convenience C. simple random D. stratified
E. cluster F. multistage G. systematic

1. The Maryland division of Weight Watchers is doing research to determine how many people on the Weight Watchers diet cheat at least once a week. They decide that anonymous surveys will give them an accurate representation but do not have time to get responses from ALL the Maryland Weight Watchers people. Read the scenarios below and determine which of the 7 sampling methods best describes it.

- _____ I. Randomly select 10 members from each of the WW centers in the Maryland division.
_____ II. Use an alphabetical listing of all Maryland division members. Randomly choose a starting person on the list. Then select every 20th person thereafter.
_____ III. Randomly select 2 or 3 branches of the Maryland division and survey every member of that center.
_____ IV. Send out the survey to every member of the Maryland division. Place drop boxes in each WW center. Anyone who returns a survey will be in the sample.
_____ V. The Maryland regional office is in Baltimore so they survey members at the WW center in Baltimore.
_____ VI. From a numbered list of all Maryland division members use a computer to randomly select 100 numbers and survey all members with those corresponding numbers.

2. What is the population of interest in the WW situation?

Here is a formula that is used often in AP Statistics: $z = \frac{x - \mu}{\sigma}$

1. If $z = 2.5$, $x = 102$, and $\mu = 100$, what is σ ? Show your work.

$$\sigma = \underline{\hspace{2cm}}$$

2. If $z = -3.35$, $x = 60$, and $\sigma = 4$, what is x ? Show your work.

$$x = \underline{\hspace{2cm}}$$

It is expected that you have a thorough understanding of linear functions and scatterplots.

1. The USDA reported that in 1990 each person in the United States consumed an average of 133 pounds of natural sweeteners. They also claim this amount has decreased by about 0.6 pounds each year.

- a. If 1990 could be considered “year 0”, which of the above numbers represents the slope and which represents the y-intercept?

- b. What is the equation of the line of best fit using the slope and y-intercept above?

- c. Predict the average consumption of sweeteners per person for the year 2005.

2. The following equation can be used to predict the average height of boys anywhere between birth and 15 years old: $y = 2.79x + 25.64$, where x is the age (in years) and y is the height (in inches).

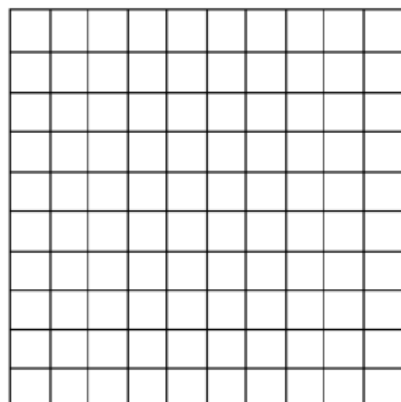
- a. What does the slope represent in this problem? Interpret it in the context of this problem/situation.

- b. What does the y-intercept represent in this problem? Interpret it in context.

3. Hilary wonders if people of similar heights tend to date each other. She measures herself, her dormitory roommate, and the women in the adjoining rooms; then she measures the next man each woman dates. Here are the data (heights in inches):

<u>Women:</u>	<u>66</u>	<u>64</u>	<u>66</u>	<u>65</u>	<u>70</u>	<u>65</u>
<u>Men:</u>	<u>72</u>	<u>68</u>	<u>70</u>	<u>68</u>	<u>74</u>	<u>69</u>

- A. Construct a scatterplot of the data.
- B. Describe the association between the heights of the women and the men they date.



You are expected to have a basic understanding of simple probability. If you find these problems “less than intuitive”, there are numerous sites available online that provide basic probability explanations.

1. A special lottery is to be held to select the student who will live in the only deluxe room in a dormitory. There are 100 seniors, 150 juniors, and 200 sophomores who applied. Each senior's name is placed in the lottery 3 times; each junior's name, 2 times; and each sophomore's name, 1 time. What is the probability that a senior's name will be chosen?

- A. $\frac{1}{8}$ B. $\frac{2}{9}$ C. $\frac{2}{7}$ D. $\frac{3}{8}$ E. $\frac{1}{2}$

2. Which of the following has a probability closest to 0.5?

- A. The sun will rise tomorrow.
B. It will rain tomorrow.
C. You will see a dog with only three legs when you leave the room.
D. A fair die will come up with a score of 6 four times in a row.
E. There will be a plane crash somewhere in the world within the next five minutes.

3. If a coin is tossed twice, what is the probability that on the first toss the coin lands heads and on the second toss the coin lands tails? (Hint: What are the possible outcomes when you toss a coin twice?)

- A. $\frac{1}{6}$
B. $\frac{1}{3}$
C. $\frac{1}{4}$
D. $\frac{1}{2}$
E. 1

4. If a coin is tossed twice what is the probability that it will land either heads both times or tails both times?

- A. $\frac{1}{8}$
B. $\frac{1}{6}$
C. $\frac{1}{4}$
D. $\frac{1}{2}$
E. 1

5. Calculate the following probabilities and arrange them in order from least to greatest.

- I. The probability that a fair die will produce an even number. _____
II. A random digit from 1 to 9 (inclusive) is chosen, with all digits being equally likely. The probability that when it's squared the answer will contain the digit 1. _____
III. The probability that a letter chosen from the alphabet will be a vowel. _____
IV. A random number between 1 and 20 (inclusive) is chosen. The probability that its square root will not be an integer. _____

ORDER: _____, _____, _____, _____

Problem Set 4: Use your TI-84+ to follow the directions and complete this problem set.

Histograms: A set of quantitative data is often separated into groups or intervals. For example, test scores may be separated into 90-99, 80-89, 70-79, and so on. The frequency refers to the count for each interval. A histogram is a visual representation of the frequencies plotted against the interval.

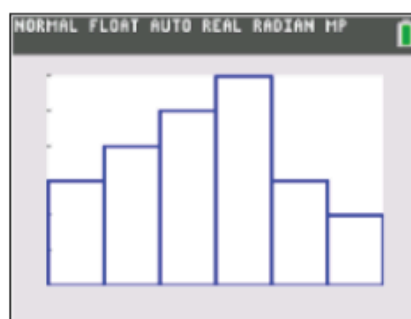
Exercise 1: Enter the following test scores into List 3 and create a histogram using intervals of width 10:

100, 100, 98, 95, 92, 88, 85, 84, 84, 81, 80, 77, 75, 72, 70, 70, 66, 65, 63, 60, 59, 55, 50

Procedure: First enter the scores into L3. You must now click on **STATPLOT** in the top left. (2nd → Y=) Choose Plot 1. Switch it ON. TYPE → choose the histogram. X LIST → L3 (2nd → 3).

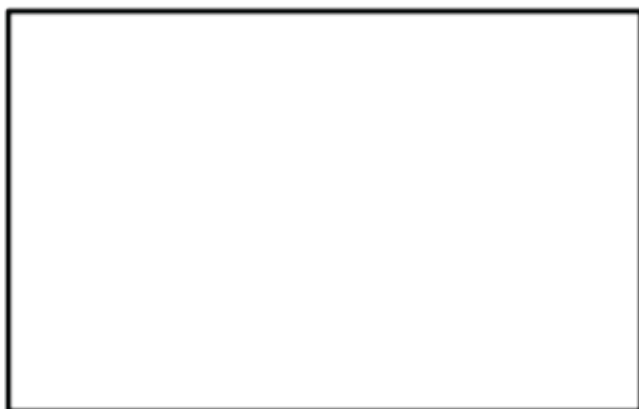
Next, go to **WINDOW**. Set the following: Xmin=50; Xmax=100; Xscl=10; Ymin=0; Ymax=6

The Xscl affects your bin width (the width of your bars).



Select **GRAPH** and you should see the histogram as shown on the right. Select **TRACE** to see frequencies of the intervals. When drawing a histogram, be sure to label both axes and scale reasonably.

1. Go back to **WINDOW** for this example and change the Xscl = 5 and **GRAPH**. Draw the histogram below and describe how this one differs from the first histogram. Which histogram would best describe the data? Explain below in full sentences.



Problem Set 6: Use your TI-84+ to complete the problem set. [use activity 3 notes]

Boxplots: A box plot is a graphical display for a set of quantitative data that only displays the five number summary: minimum, Q1, median, Q3, maximum

1. Using the same test scores from Set 6, create a box plot below. Label the axis and scale appropriately.

Go to **STATPLOT**, and you will notice two boxplot options. The first is the preferred option.



Use TRACE to get the values from the boxplot.

Suppose another class receives these test scores:

94, 93, 90, 83, 82, 82, 77, 76, 75, 70, 64, 60, 53, 52, 48, 44, 40

2. Create two box plots on the same set of axes to compare their performance with the original group.

3. Write a few sentences below comparing the centers and spreads of the two classes.