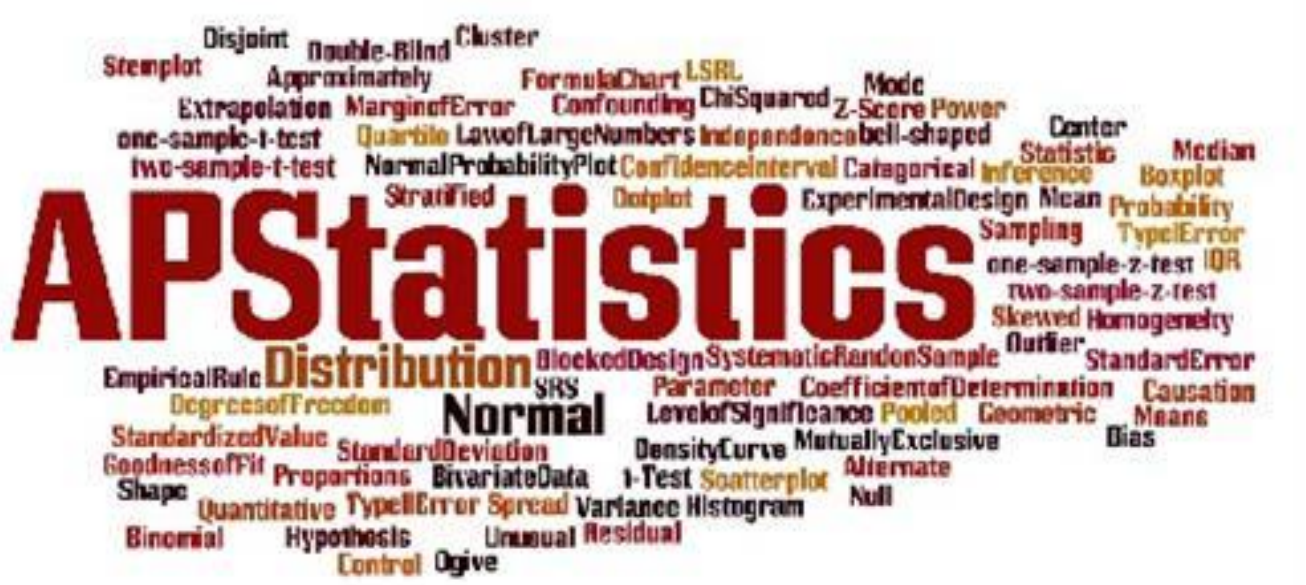
LAWRENCE TOWNSHIP
PUBLIC SCHOOLS
www.ltps.org • 2565 Princeton Pike, Lawrenceville, NJ 08648 • 609-671-5500

# Lawrence High School's
# AP Statistics
# 2021Summer Jumpstart Assignment




Lawrence High School
2525 Princeton Pike, Lawrenceville NJ 08648
609-671-5510 Fax 671-3411

LAWRENCE TOWNSHIP
PUBLIC SCHOOLS
www.ltps.org • 2565 Princeton Pike, Lawrenceville, NJ 08648 • 609-671-5500

To incoming AP Statistics students,

To be best prepared for your AP Statistics course, you will complete this packet, which you will **submit shortly after returning to school in September**. Your instructor will announce the exact due date. This assignment will be graded and included in the first marking period earn grade. You will be able to access this assignment through ltps.org and/or classroom.google.com**. Please be sure to sign up for the correct instructor's Google classroom.**

**Google Classroom**:
This Google group has been established in order to provide you, the students, with support as a group. If you are struggling with concepts/material, you can post questions to one another and work together just as if you were working in groups in school. In addition, you have the educational resources listed below for assistance.

**Google Classroom Codes:**

- Go to www.classroom.google.com
- Sign into your school provided **.info account.**
- **Mr. Krajunus'** Access code hzfzv3v
- **Mr. Shive's** Access code: a2uwmvg
- The instructors might monitor the Google Classroom throughout the summer to provide guidance. Use this Google Classroom to reach out to the instructor.

**Directions:**

- Complete all questions/examples.
- All work must be your original work.
- Use separate pages to show your work, be sure each additional work page has your name on it. Work should be organized and neatly labeled.
- Any problem with no work shown will receive 0 points.
- When asked to explain or justify, you must write in full sentences.

**Scoring/Grading:** The AP Statistics Summer Assignment is a 32 point assignment.

- The first page (probability) is a total of 8 points.
- The summary statistics and two graphing data sections are each 8 points.
- **This will count as your first grade of the marking period and will be entered into the grading system.**
- **The assignment is due on the first day of the first full week of school. Your instructor will announce the exact date.**

**Resources:**
For additional examples and support you can reference any of the sites listed below and search the skill/concept.

- stattrek.com
- KhanAcademy.com
- You Tube.com or Teacher Tube.com
- MathIsPower4u.com
- IXL.com
- Math 8th Seminar materials

Good luck with the assignment and enjoy your summer!!!!
Mr. Krajunus and Mr. Shive

Lawrence High School
2525 Princeton Pike, Lawrenceville NJ 08648
609-671-5510 Fax 671-3411

Suppose you toss one coin and roll one six-sided die.

(a) List all the possible outcomes.

(b) Find the probability of getting a head.

(c) Find the probability of getting a 1, 2, or 3 on the die.

(d) Find the probability of getting a head or a five.

1. In a statistics class there are 18 juniors and 10 seniors; 5 of the seniors are females, and 10 of the juniors are males. Complete the table below:

| | Male | Female | Total |
|---|---|---|---|
| Junior | | | |
| Senior | | | |
| Total | | | |

If a student is selected at random, find the probability of selecting

(a) a junior or a female

(b) not a junior male

LAWRENCE TOWNSHIP PUBLIC SCHOOLS
www.ltps.org • 2565 Princeton Pike, Lawrenceville, NJ 08648 • 609-671-5500

# Summary Statistics – Center and Spread

A *statistic* is a number that gives information about a set of data. Common examples include mean, median, mode (which we won't worry about in AP Stat), range, standard deviation, and more!

**SYMBOLOGY**

In statistics, we use a variety of *symbols* to represent statistics. Sometimes, the symbol used depends on whether we are talking about a **population** or a **sample** (select members of a given population)

| | Mean | Standard Deviation | Median | Number of data points |
|---|---|---|---|---|
| Population | **μ** ("myu") | **σ** ("sigma") | *No symbol* (Often abbreviated "**Med**.") | **n** |
| Sample | $\bar{x}$ ("x-bar") | **s** | | |

**Measures of CENTER**

The *center* of a data set lets us understand the "average" or "typical" value of a number in that data set. There are two main measures of center: mean and median.

**MEAN**

Add up all data points, then divide by the number of data points.

- **μ (or $\bar{x}$)** = $\frac{\sum x}{n}$
- "Sum of data points over number of data points"

***Example 1:*** *Science grades of a sample of 15 juniors: 91, 87, 66, 74, 85, 98, 43, 88, 77, 62, 83, 91, 89, 52, 100*

This is a SAMPLE, so $\bar{x} = \frac{\sum x}{n} = \frac{1186}{15}$ = **79.07**

---

***Example 2:*** *Heights of all 6 people in a family (inches): 47, 58, 61, 65, 68, 70*

This is the POPULATION, so **μ** = $\frac{\sum x}{n} = \frac{369}{6}$ = **61.5**

**MEDIAN**

The *middle number* of the data set, assuming that the data points are **in order** (smallest to largest)

- If there are 2 numbers in the middle, find the *mean* of those two numbers!
- A nice **trick** for finding the *position* of the median is to use $\frac{n+1}{2}$

***Example 1:*** *Science grades of a sample of 15 juniors: 91, 87, 66, 74, 85, 98, 43, 88, 77, 62, 83, 91, 89, 52, 100*

$\frac{n+1}{2} = \frac{15+1}{2}$ = **8.** Median is the **8th** number (IN ORDER)

43, 52, 62, 66, 74, 77, 83, **85**, 87, 88, 89, 91, 91, 98, 100

---

***Example 2:*** *Heights of all 6 people in a family (inches): 47, 58, 61, 65, 68, 70*

$\frac{n+1}{2} = \frac{6+1}{2}$ = **3.5.** Median is **between the 3rd & 4th** number

47, 58, 61, 65, 68, 70; Average = $\frac{61+65}{2}$ = **63**

**Measures of SPREAD**

The *spread* of a data set tells us whether the data points are far apart or clustered together. The most important measure of spread is **standard deviation,** which is the **typical distance of the data points from the mean**. Other measures of spread, such as Range and IQR, will be discussed in Part 5.

The formulas for Standard Deviation are as follows. Note that they are slightly different for a population and a sample (the sample one will be slightly larger to account for the fact that the sample doesn't include all members of a population)

Population: $\sigma = \sqrt{\frac{\Sigma(x_i-\mu)^2}{n}}$ Sample: $s = \sqrt{\frac{\Sigma(x_i-\bar{x})^2}{n-1}}$

***You will NOT have to calculate Standard Deviation by hand in this course!***

What you *will* have to do, however, is be able to *interpret* and *compare* the Standard Deviations of different data sets:

- ***Larger* Standard Deviation:** The data is more spread out (points are typically further from the mean)
- ***Smaller* Standard Deviation:** The data is closer together (points are typically closer to the mean)

**Example:**

**Data Set 1:** 1, 2, 3, 17, 18, 19; μ = 10, σ = 8.04 **Data Set 2:** 7, 8, 9, 11, 12, 13; μ = 10, σ = 2.16

*Notice how Data Set 1 is more spread out, while Data Set 2 is closer together. This is reflected in the fact that Set 1's Standard Deviation (8.04) is higher than Set 2's Standard Deviation (2.16)*

**Practice Problems**

1. Find the mean and median of the following data set. **Show work** when appropriate!
   **Teaching experience of all LHS math teachers (n = 15):** 1, 3, 3, 3, 4, 4, 5, 5, 5, 6, 7, 7, 18, 23, 26
   *Symbol for mean:* _____ *Value of mean:* _____ *Position of Median:* _____ *Value of Median:* _____

2. Find the mean and median of the following data set. **Show work** when appropriate!
   **Weights of 8 randomly-selected chickens on a farm (in pounds):** 5.4, 5.7, 6.2, 6.9, 7.2, 7.2, 8.1, 9.0

   *Symbol for mean:* _____ *Value of mean:* _____ *Position of Median:* _____ *Value of Median:* _____

3. Find the mean and median of the following data set. **Show work** when appropriate!
   **Temperature readings on <u>all</u> thermostats in an office building:** 71, 72, 72, 74, 68, 74, 71, 72, 69, 76

   *Symbol for mean:* _____ *Value of mean:* _____ *Position of Median:* _____ *Value of Median:* _____

4. List the following data sets in order from *least spread out* to *most spread out*. Then, **write** 1-2 sentences explaining how you could tell.

   **LHS Math Teachers:** σ = 7.02 **Chickens:** s = 1.21 **Thermostats:** σ = 2.26

LAWRENCE TOWNSHIP PUBLIC SCHOOLS
www.ltps.org • 2565 Princeton Pike, Lawrenceville, NJ 08648 • 609-671-5500

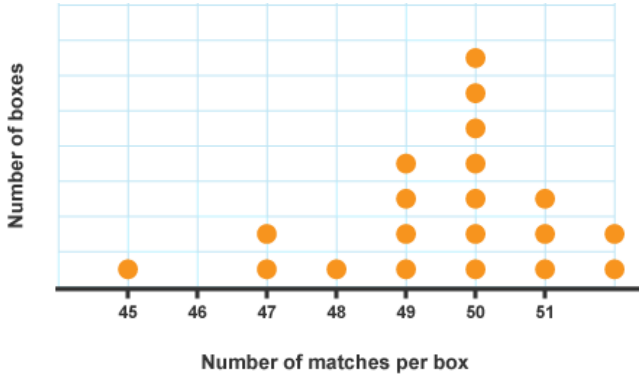
# Graphing Data – Dotplots and Stem-and-Leaf Plots

Statistics such as mean, median, and standard deviation are very useful in summarizing data and giving overall trends. But they don't tell the full story. By making a *graph* of the data, we can go *beyond* the numbers and see *shapes* and *patterns* in the data. Shown below are two common ways in which to graph data

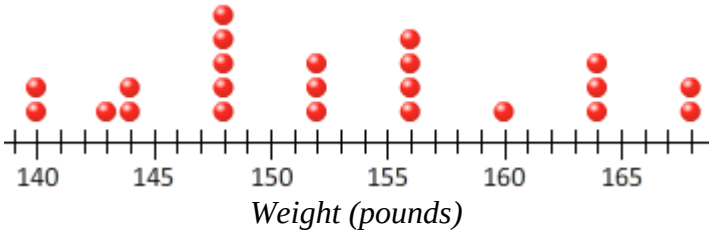
### *Dotplots*

- Make an AXIS on the bottom (you can go by 1s, 2s, 5s, 10s…whatever makes sense for the data!)
- Put one dot for each data point on the axis. If there is more than one data point for a given value, *stack* the dots!

**Example:** *Number of matches in 20 randomly-selected boxes.*
45, 47, 47, 48, 49, 49, 49, 49, 50, 50, 50, 50, 50, 50, 50, 51, 51, 51, 52, 52



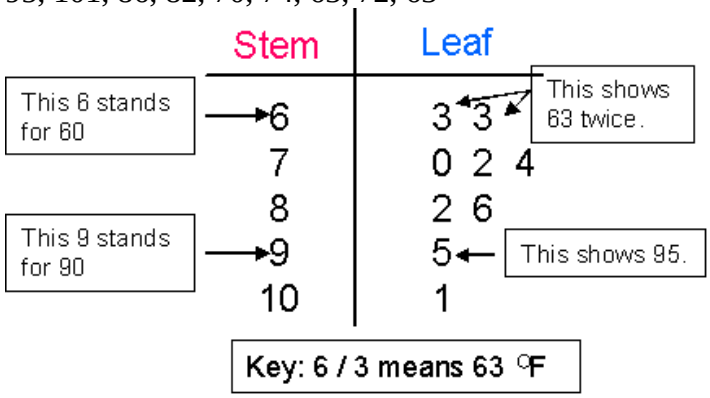

**Example:** *Weights of players on a high school baseball team*
140, 140, 143, 144, 144, 148, 148, 148, 148, 148, 152, 152, 152, 156, 156, 156, 156, 160, 164, 164, 164, 168, 168




### *Stem-and-Leaf Plots (also called Stemplots)*

- Use a KEY to determine what the stems and leaves are worth
- DO NOT SKIP STEMS. If there are no data points for that stem, just keep the stem there and put no leaves after it. Skipping the stem will alter what the stemplot looks like.

**Example:** *Temperatures at OU football games, 2009*
95, 101, 86, 82, 70, 74, 63, 72, 63




**Example:** *Gross National Product (per capita) of West African countries*
180, 240, 260, 270, 310, 330, 360, 370, 390, 410, 480, 500, 710, 730, 890

Lawrence High School
2525 Princeton Pike, Lawrenceville NJ 08648
609-671-5510 Fax 671-3411

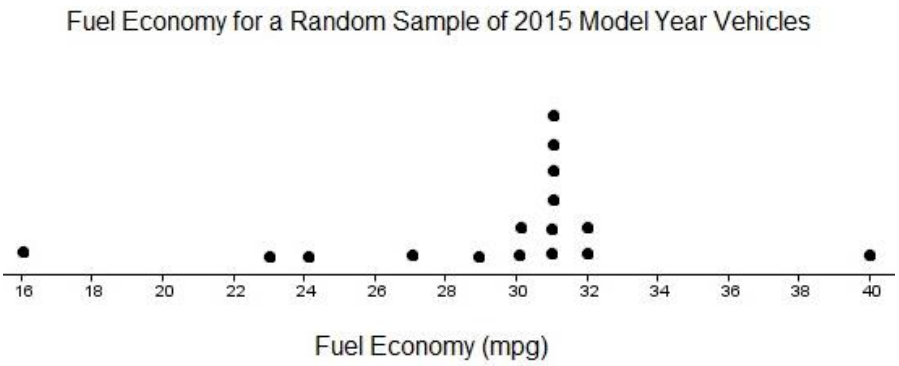
| Stem | Leaf |
|---|---|
| 1 | 80 |
| 2 | 40 60 70 |
| 3 | 10 30 60 70 90 |
| 4 | 10 80 |
| 5 | 00 |
| 6 | |
| 7 | 10 30 |
| 8 | 90 |

Notice how the 6 is still here even though there are no data points in this interval

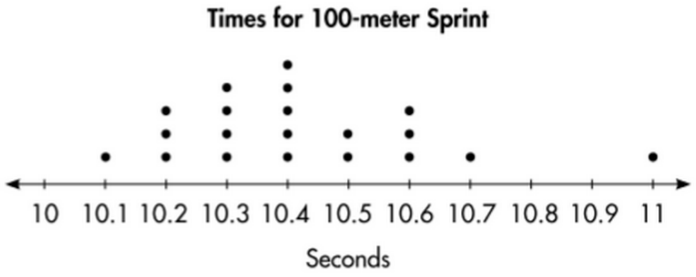
*Key: 4 | 80 means an income of $480.*

**Practice Problems**

1. List all the data points in the dotplot:




2. Using the dotplot shown, find the **mean** and **median** 100-meter sprint time




3. The following stemplot shows the final exam scores of a class with 10 students. List the scores of each student:

| stem | leaf |
|---|---|
| 6 | 9 |
| 7 | |
| 8 | 7 8 8 9 |
| 9 | 0 6 7 7 |
| 10 | 0 |

Key: 6 | 8 means 68

4. Find the **mean** and **median** of the data shown in the stemplot.
***NOTE:*** *Look at the key carefully!*

| Stem | Leaf |
|---|---|
| 2 | 0 2 3 6 |
| 3 | 2 3 5 6 7 |
| 4 | 6 8 9 |
| 5 | 4 7 |
| 6 | 2 |
| 7 | 3 |

**KEY:** 4|6 = 4.6

# Graphing Data – Box and Whisker Plots

**Example:** Ages of 9 employees in an office

**37, 24, 51, 46, 62, 28, 35, 49, 55**

**Finding MEDIAN:* (9 people. $\frac{9+1}{2}$ = 5, it's the 5th number. **Remember that numbers must be <u>in order</u>!**)

24, 28, 35, 37, **46,** 49 51, 55, 62

**Median = 46**

**Finding QUARTILES:* There are 9 data points, so *split* the data set in half! ***NOTE:** If there is one number that serves as the median, as with this data set, **it is not included in <u>either</u> half!***

*Q1*
24, **28, 35**, 37 — *Average: 31.5*

46 (not included)

*Q3*
49, **51, 55,** 62 — *Average: 53*

***BOXPLOT:** Min = 24, Q1 = 31.5, Med = 46, Q3 = 53, Max = 62


**RANGE:** 62 – 24 = **<u>38</u>**
**IQR:** 53 – 31.5 = **<u>21.5</u>**

**Practice Problems**

1. Construct a box and whisker plot for the following data set. **Be sure to include an <u>axis</u>** like the ones in the examples!

17, 21, 24, 26, 31, 33, 36, 37, 41, 48