

ARE YOU READY FOR STATISTICS?

The first unit of Statistics/AP Statistics begins with the basics of descriptive statistics. Many of these concepts have been covered in previous math and science courses. Over the summer you will review these topics and complete this packet. Writing and reading comprehension are important components to the course. In this packet are three tasks that require you to justify your reasons. Listed below is some Internet sites that contain information about descriptive statistics.

www.mste.uiuc.edu/hill/dstat/dstat.htm

www.chelt.ac.uk/acadres/stats/descript.htm

www.habermas.org/stat2f98.htm

The following packet is due at the first meeting of Statistics/AP Statistics. This will be graded as the first homework assignment. There will two parts to the grade:

1. consideration of correctness of answer.
2. completeness of work.

There will be a quiz on this material.

I. Measuring Central Tendency

- a. Find the mean, median, and mode of the following collection.

15, 11, 19, 15, 14, 13, 17, 11, 12, 17, 15, 14, 15

To begin, order the fourteen numbers.

11, 11, 12, 13, 14, 14, 14, 15, 15, 15, 17, 17, 19

To find the **mean**, divide the sum of the numbers by 14.

$$\text{mean} = \frac{2(11) + 12 + 13 + 3(14) + 4(15) + 2(17) + 19}{14} \approx 14.4$$

The **mode** is 15 because that is the number that occurs the most frequently.

The **median** is the average of the two middle numbers for even data list.

$$\text{Median} = \frac{14 + 15}{2} = 14.5$$

For odd data list, the median is the middle number.

- b. Find the quartiles of the collection in Example a. Then, sketch a box-and-whisker plot of the data

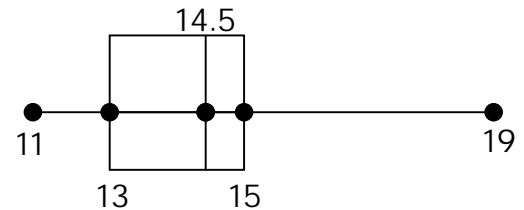
11, 11, 12, 13, 14, 14, 14, 15, 15, 15, 15, 17, 17, 19
← Lower half Upper half →

The **first quartile** is 13 (the median of the lower half)

The **second quartile** is 14.5 (the median)

The **third quartile** is 15 (the median of the upper half).

A **box-and-whisker plot** for the data labels the endpoints of the data and marks the quartiles. It is shown at the right.



The **range** is the difference of the highest and lowest data points. That is, the range is $19 - 11 = 8$.

Exercises

- Find the mean, median, mode, and range of the following collection of scores on a test.
32, 72, 81, 95, 98, 58, 77, 75, 83, 97, 45, 89, 93, 57,
82, 97, 52, 75, 79, 78, 99, 98, 54, 75, 85, 61, 55, 86
- Find the first, second, and third quartiles of the collection of data in Exercise 1.
- Construct a box-and-whisker plot of the collection of data in Exercise 1.
- Complete #1-3 for the following set of data. The weights (in pounds) of eleven children are as follows: 39, 52, 40, 45, 46, 55, 48, 40, 43, 47, 44

II. Organizing Data

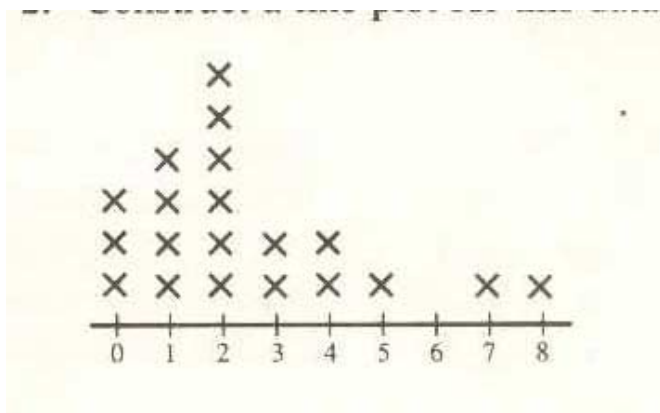
At a car dealership, the number of new cars sold in a week by each salesperson was as follows:

5, 8, 2, 0, 2, 4, 7, 4, 1, 1, 2, 2, 0, 1, 2, 0, 1, 3, 3, 2.

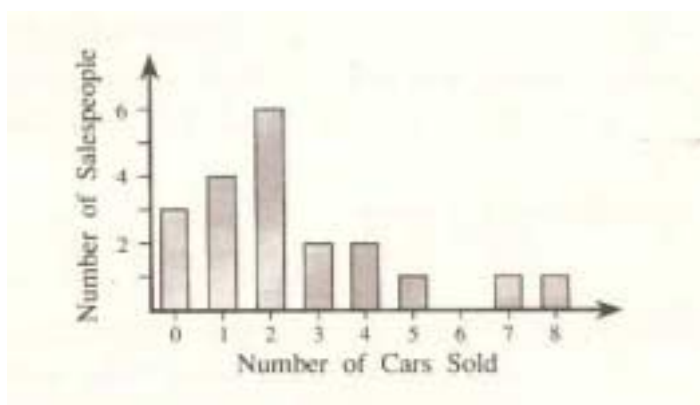
- a. Construct a frequency distribution for this data.

Number	Tally	Frequency
8		1
7		1
6		0
5		1
4		2
3		2
2		6
1		4
0		3

- b. Construct a line plot for this data



- c. Construct a bar graph that shows the number of salespeople who sold 0-8 cars.



2. Each of the members of a recent eighth class was asked to name his/her favorite among these subjects: English, foreign language, history, mathematics, science. The results are shown in the table. Construct a bar graph that shows these results.

English	62
Foreign Language	40
History	40
Mathematics	18
Science	33

III. Constructing Stem-and-Leaf Plots and Histograms

I. Construct a stem-and-leaf plot for the data

Unordered Data

63, 52, 84, 83,
51, 32, 58, 35,
45, 41, 65, 75,
59, 67, 25, 46

Stem-and-leaf Plot

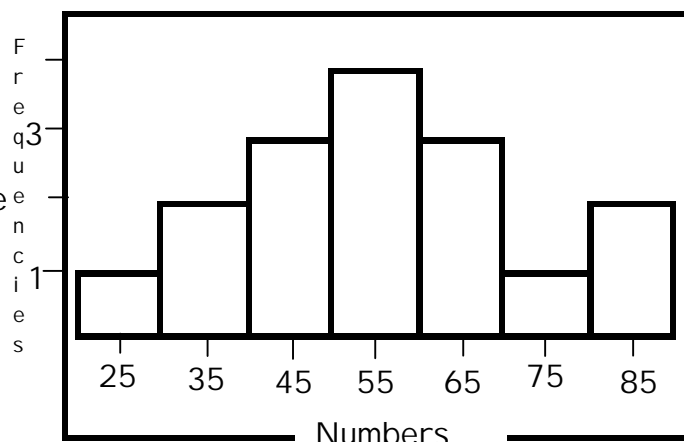
2	5
3	2 5
4	1 5 6
5	1 2 8 9
6	3 5 7
7	5
8	3 4

Leaves should be in increasing order.

A stem-and-leaf plot orders data in increasing or decreasing order.

II. Histograms

1. Construct and label a horizontal number line that is scaled to contain all of the values of the variable of interest.
2. Construct and label a vertical axis so that the greatest frequency can be represented.
3. Construct the bars of equal width that are centered above each value. The heights of the bars represent the frequencies of the values.

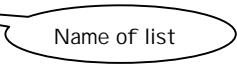


Exercises

1. Construct a stem-and-leaf plot for the data
15, 59, 66, 42, 48, 23, 70, 81, 35, 51, 68, 29, 77, 92,
85, 16, 37, 59, 61, 76, 40, 25, 86, 11
2. Construct a histogram for the above data.

IV. Using the TI-83 Calculator

1. Entering data into a list: **STAT**→ **1:Edit**→ enter data in appropriate list
2. To clear a list: **highlight the list name and hit clear**. The list is empty when you see the dashes.
3. To sort a list in ascending order: **STAT**→ **2:SortA(L₁)**
4. To find the mean of a list: **2nd STAT**→ **MATH** →**3:mean(L₁)**
5. To find the median of a list: **2nd STAT**→ **MATH** →**4:median(L₁)**
6. To find the sum of a list: **2nd STAT**→ **MATH** →**5:sum(L₁)**



Name of list

Exercises

1. Enter the above data from Exercise 1 in a list.
2. Sort the data in ascending order and find the mode.
3. Find the mean of the data.
4. Find the median of the data.
5. Find the sum of the data.

Task 1: Write two specific survey questions that you would ask voters in the next senatorial election in your state. Choose the type of question and response (yes/no, scale of 1 to 5, numerical responses, etc.) that would be most appropriate for the issues involved. What relationships would be expected when the responses are analyzed?

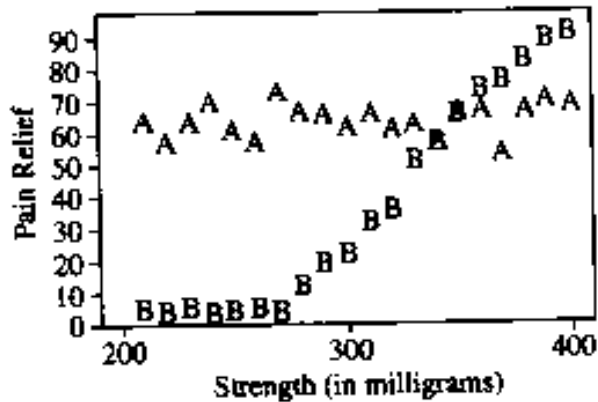
Task 2: Find a set of numbers that will satisfy the following conditions:

- The median of a set of 20 numbers is 24.
- The range is 42.
- To the nearest whole number the mean is 24.
- No more than three numbers are the same.

Show your strategy.

Task 3: Two pain relievers, A and B, are being compared for relief of postsurgical pain. Twenty different strengths (doses in milligrams) of each drug were tested. Eight hundred postsurgical patients were randomly divided into 40 different groups. Twenty groups were given drug A. Each group was given a different strength. Similarly, the other twenty groups were given different strengths of drug B. Strengths used ranged from 210 to 400 milligrams. Thirty minutes after receiving the drug, each patient was asked to describe his or her pain relief on a scale of 0 (no decrease in pain) to 100 (pain totally gone).

The strength of the drug given in milligrams and the average pain rating for each group are shown in the scatterplot below. Drug A is indicated with A's and drug B with B's.



- Based on the scatterplot, describe the effect of drug A and how it is related to strength in milligrams.
- Based on the scatterplot, describe the effect of drug B and how it is related to strength in milligrams.
- Which drug would you give and at what strength, if the goal is to get pain relief of at least 50 at the lowest possible strength? Justify your answer based on the scatterplot.