

## Section 3.2 – Measures of Central Tendency

MDM4U

David Chen

### Part 1: Video

This video will review shape of distributions and introduce measures of central tendency. Answer the following questions while watching.

<http://www.learner.org/courses/againstallodds/unitpages/unit04.html>

1. What variable is examined in comparing men and women workers at the beginning of the video?
2. Would you describe the shape of the distribution of men's weekly wages as symmetric, skewed to the left or skewed to the right?
3. What is the most important difference between the distributions of weekly wages for men and for women?
4. Would a few very large incomes pull the mean of a group of incomes up, down, or leave the mean unaffected?
5. What happens to the median of the set of data of salaries when the president decides to double his salary?

- In this section, you will learn how to describe a set of numeric data using a single value
- The value you calculate will describe the \_\_\_\_\_ of the set of data
- The 3 measures of central tendency are:

## **Part 2: The Mean**

**The mean:** a measure of central tendency found by dividing the \_\_\_\_\_ by the \_\_\_\_\_. In statistics, it is important to distinguish between the mean of a population and the mean of a sample of that population. The sample mean will approximate the actual mean of the population, but the two means could have different values. Different symbols are used to distinguish the two kinds of means.

<b>Population Mean</b>	<b>Sample Mean</b>
$\mu = \frac{\sum x}{N}$	$\bar{x} = \frac{\sum x}{n}$
$\mu$ – mu; population mean	$\bar{x}$ – x-bar; sample mean
$\sum$ --- sigma; the sum of	$\sum$ --- sigma; the sum of
$N$ --- number of values in the population	$n$ --- number of values in the sample

**Example 1:** A group of elementary school children were asked how many pets they have. Here are their responses, arranged from lowest to highest:

1      3      4      4      4      5      7      8      9

What is the mean number of pets for this group of children?

What does this number tell us? One way to think of it would be that if every child in the group had the same number of pets, each would have \_\_\_\_\_ pets. This is the ‘fair share’ value.

## Part 2: Weighted Mean

Sometimes, certain data within a set are more \_\_\_\_\_ than others. For example, the mark on a final exam is often considered to be more important than the mark on a term test for determining an overall grade for a course. A \_\_\_\_\_ gives a measure of central tendency that reflects the relative importance of the data.

### **Weighted mean formula:**

**Example 2:** The personnel manager for a marketing company considers five criteria when interviewing a job applicant. The manager gives each applicant a score between 1 and 5 in each category, with 5 as the highest score. Each category has a weighting between 1 and 3. The following table lists a recent applicant's scores and the company's weighting factors. Determine the weighted mean score for this job applicant.

Criterion	Score, $x_i$	Weighting Factor, $w_i$
Education	4	2
Job experience	2	2
Interpersonal skills	5	3
Communication skills	5	3
References	4	1

**Example 3:** The table below shows a student's performance in an MDM4U class. What would their final mark be based on the weightings shown?

Category	Mark, $x$	Weighting, $w$
Assignments	92	15%
Tests	86	40%
ISU	85	15%
Exam	88	30%

## Part 3: Mean of Grouped Data

Supposed your data have already been organized into a frequency table with \_\_\_\_\_.

You no longer have actual data values, so you must then use the \_\_\_\_\_ of each class to estimate a mean weighted by the frequency.

Finding the average (mean) of grouped data is the same as finding a weighted average; except that you have to use the \_\_\_\_\_ as the data value.

Formula for mean of grouped data:

**Example 4:** A sample of car owners was asked how old they were when they got their first car. The results were then reported in a frequency distribution. Calculate the mean.

Age	16 --- 20	21 --- 25	26 --- 30	31 --- 35	36 --- 40
Frequency	10	18	12	8	2

**Solution:**

Age	Frequency, $f$	Midpoint of age, $m$	$f \times m$
16 – 20	10		
21 – 25	18		
26 – 30	12		
31 – 35	8		
36 – 40	2		

#### **Part 4: Median**

The median value is the \_\_\_\_\_ data point in an \_\_\_\_\_ set, dividing the set into two sets of equal size.

To find the median of a distribution:

1. Arrange all observations in order of size, from smallest to largest
2. If the number of observations  $n$  is odd, the median is the middle most observation in the ordered list
3. If the number of observations  $n$  is even, the median is the average of the two middle most observations in the ordered list

**Tip:** The middle most piece of data is in the  $\frac{n+1}{2}$  position

**Example 5:** Monthly rents downtown and in the suburbs are collected from the classified section of a newspaper. Calculate the median rent in each district

Downtown: 850, 750, 1225, 1000, 800, 1100, 3200

Suburbs: 750, 550, 900, 585, 220, 625, 500, 800

Start by ordering the sets of data...

Downtown:

Suburbs:

**Downtown:**

There are \_\_\_\_ elements in the set, so the median is the \_\_\_\_ element. The median is \_\_\_\_ /month.

**Suburbs:**

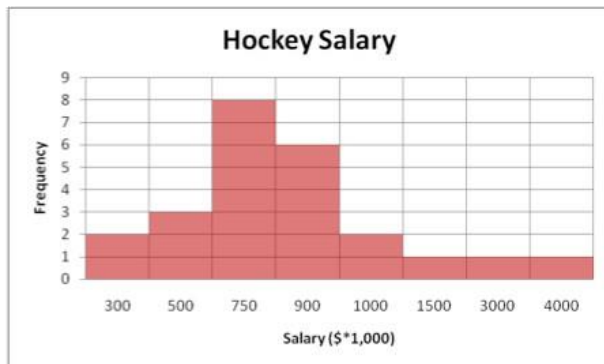
There are \_\_\_\_ elements in the set, so the median is halfway between the \_\_\_\_ and \_\_\_\_ element. Therefore the median is \_\_\_\_ /month.

### **Part 5: The mode**

The mode is simply the \_\_\_\_ value or range of values in a data set. It is easy to determine the mode from a histogram as it is the \_\_\_\_ column.

**Example 6:** The table and histogram show the current salaries for a hockey team.

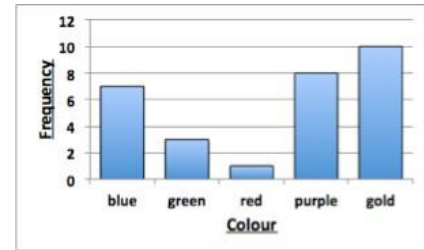
Salary (\$)	Number of Players
300 000	2
500 000	3
750 000	8
900 000	6
1 000 000	2
1 500 000	1
3 000 000	1
4 000 000	1



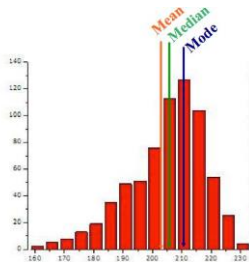
What is the mode of the salaries?

**Note:** If no measurement is repeated, the data set has no mode. If it has two measurements that occur most often, it is called bimodal and has two modes.

**Note:** If you have categorical data, the mode is the only appropriate measure of central tendency to use.

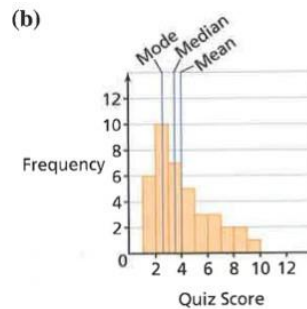


**Skewed left:**



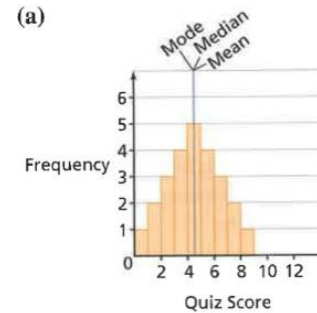
The mean is \_\_\_\_\_ the median

**Skewed right:**



The mean is \_\_\_\_\_ the median

**Symmetrical:**



The mean, median, and mode are \_\_\_\_\_

### Part 6: Application of Weighted Mean

Before you start your homework, use the table below to determine what mark the student needs on the final exam to earn a final grade of 85%.

Category	Mark, $x$	Weighting, $w$
Assignments	92	15%
Tests	86	45%
ISU	85	10%
Exam	?	30%

### Using your Ti-84 to find the mean of a set of data

### Example 1 done with Ti---84:

- Input data in to L1: STAT → EDIT
- Calculate mean: STAT → CALC → 1---VARSTATS → List: L1 → CALCULATE

NORMAL FLOAT AUTO REAL RADIAN MP					
L1	L2	L3	L4	L5	1
1	-----	-----	-----	-----	
3					
4					
4					
5					
7					
8					
9					
-----					
L1(10)=					

NORMAL FLOAT AUTO REAL RADIAN MP

EDIT **Calc** TESTS

1:1-Var Stats

2:2-Var Stats

3:Med-Med

4:LinReg(ax+b)

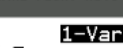
5:QuadReg

6:CubicReg

7:QuartReg

8:LinReg(a+bx)

9:LnReg



NORMAL FLOAT AUTO REAL RADIAN MP

**1-Var Stats**

$\bar{x}=5$   
 $\Sigma x=45$   
 $\Sigma x^2=277$   
 $Sx=2.549509757$   
 $\sigma x=2.40370085$   
 $n=9$   
 $\min X=1$   
 $\downarrow Q1=3.5$

### Example 2 done with Ti---84

- Input scores in to L1 and frequencies in L2: STAT → EDIT
- Calculate mean: STAT → CALC → 1---VARSTATS → List: L1, Frequencies: L2 → CALCULATE

[illegible]

NORMAL FLOAT AUTO REAL RADIAN MP

EDIT **CALC** TESTS

1:1-Var Stats

2:2-Var Stats

3:Med-Med

4:LinReg(ax+b)

5:QuadReg

6:CubicReg

7:QuartReg

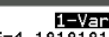
8:LinReg(a+bx)

9:LnReg

NORMAL FLOAT AUTO REAL RADIAN MP

**1-Var Stats**

List:L1  
FreqList:L2  
Calculate



1-Var Stats

$\bar{x}=4.181818182$

$\Sigma x=46$

$\Sigma x^2=206$

$Sx=1.167748416$

$\sigma x=1.113404429$

$n=11$

$\min X=2$

$Q_1=4$