Section 3.2 – Measures of Central Tendency

MDM4U
Jensen

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?

The variable is the weekly wages for Americans, separated by gender

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?

The men's distribution is skewed to the right.

3. What is the most important difference between the distributions of weekly wages for men and for women?

The medians of the two distributions differ. Median measures the 50-50 point. The median for men's wages was larger than the median for women's wages

4. Would a few very large incomes pull the mean of a group of incomes up, down, or leave the mean unaffected?

A few very large incomes inflate the mean of a group of incomes. Hence, these very large incomes would pull the mean up.

5. What happens to the median of the set of data of salaries when the president decides to double his salary?

Increasing values at the extremes has no effect on the median. The middle most value is still in the middle.

- 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 center of the set of data
- The 3 measures of central tendency are:

1. Mean

2. Median

3. Mode

Part 2: The Mean

The mean: a measure of central tendency found by dividing the sum of all the data by the number of pieces of data. 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.

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

$$\bar{x} = \frac{\sum x}{n} = \frac{1+3+4+4+4+5+7+8+9}{9} = \frac{45}{9} = 5$$

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 5 pets. This is the 'fair share' value.

Part 2: Weighted Mean

Sometimes, certain data within a set are more significant 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 weighted mean gives a measure of central tendency that reflects the relative importance of the data.

Weighted mean formula:

$$\bar{x} = \frac{\sum xw}{\sum w}$$

Where x represents each data value and w represents its weight, or frequency.

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

$$\bar{x} = \frac{\sum xw}{\sum w} = \frac{4(2)+2(2)+5(3)+5(3)+4(1)}{2+2+3+3+1} = \frac{46}{11} = 4.2$$

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%

$$\bar{x} = \frac{\sum xw}{\sum w} = \frac{92(0.15) + 86(0.4) + 85(0.15) + 88(0.3)}{0.15 + 0.4 + 0.15 + 0.3}$$

$$= \frac{87.35}{1}$$

$$= 87.35\%$$

Part 3: Mean of Grouped Data

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

You no longer have actual data values, so you must then use the midpoint 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 interval midpoint as the data value.

Formula for mean of grouped data:

$$\bar{x} = \frac{\sum mf}{\sum f}$$

Where m is the midpoint of an interval and f is the frequency for that interval

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	18	180
21 – 25	18	23	414
26 – 30	12	28	336
31 – 35	8	33	264
36 – 40	2	38	76
	$\sum f = 50$		$\sum m \times f = 1270$

$$\bar{x} = \frac{\sum mf}{\sum f} = \frac{1270}{50} = 25.4$$

Part 4: Median

The median value is the middle data point in an ordered 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: 750, 800, 850, 1000, 1100, 1225, 3200

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

Downtown:

There are 7 elements in the set, so the median is the $\frac{7+1}{2} = 4^{th}$ element. The median is \$1000/month.

Suburbs:

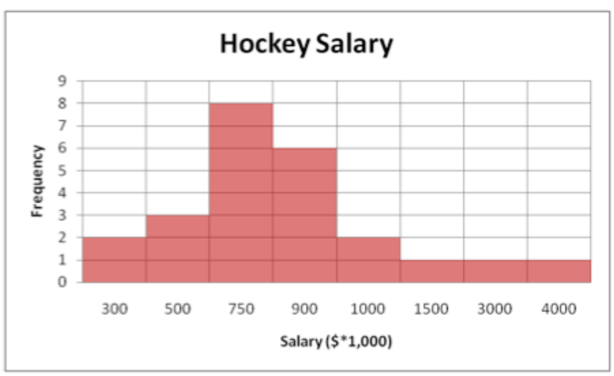
There are 8 elements in the set, so the median is halfway between the 4^{th} and 5^{th} element. Therefore the median is $\frac{585+625}{2} = \$605$/month.

Part 5: The mode

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

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

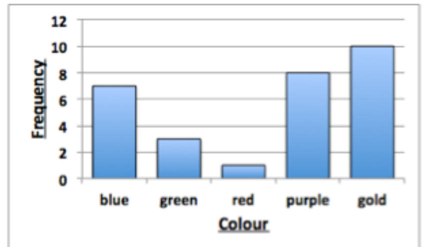
Salary (\$)	Number of Players
300 000	2
500 000	3
750 000	8
900 000	6
1 000 000	2
1 500 000	1
3 000 000	1
4 000 000	1



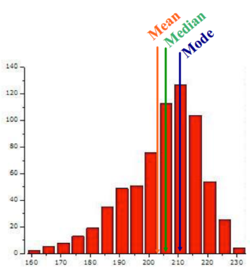
What is the mode of the salaries? \$750 000

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.



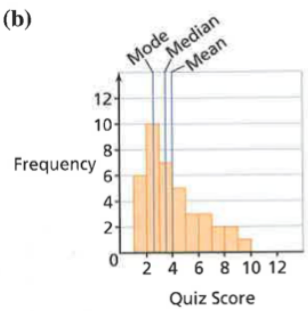
Skewed left:



The mean is <u>less</u> the median

Skewed right:

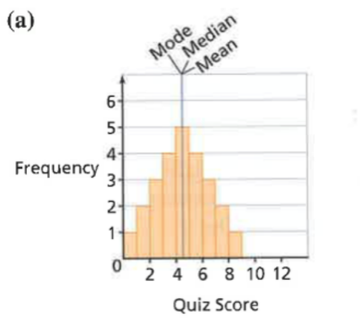
(b)



The mean is <u>greater</u> the median

Symmetrical:

(a)



The mean, median, and mode are <u>equal</u>

<u>Part 6: Application of Weighted Mean</u>

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%

$$\bar{x} = \frac{\sum xw}{\sum w}$$

$$85 = \frac{92(0.15) + 86(0.45) + 85(0.15) + x(0.3)}{0.15 + 0.4 + 0.15 + 0.3}$$

$$85 = 61 + x(0.3)$$

$$24 = x(0.3)$$

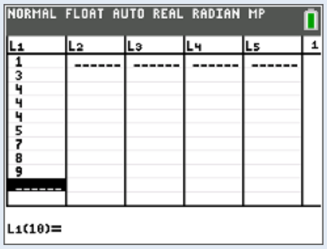
$$x = 80$$

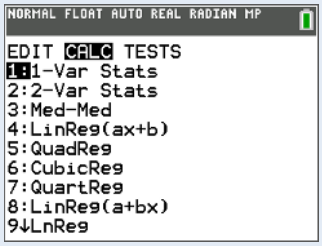
The students needs to get an 80% on the final exam.

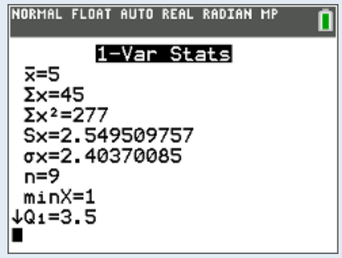
Using your Ti-84 to find the mean of a set of data

Example 1 done with Ti-84:

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







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