Did you know: Half of all Americans earn less than the median salary!

Prob/Stat/Discrete
Unit 4B Guided Notes

Name_____

Section 4.3: Measures of Central Tendencies

Objectives:
1. Can you determine the mean, median, and mode of a population and of a sample?
2. Can you determine the weighted mean of a data set and the mean of a frequency distribution?
3. Can you describe the shape of a distribution as symmetric, uniform, or skewed and compare the mean and median for each?

Measures of Central Tendency:
- A value that represents a typical, or central, entry of a __data__ __set__.
- Most common measures of central tendency:
  - __mean__
  - __median__
  - __mode__

Mean (_average_)
- The __sum__ of all the data entries divided by the number of __entries__.

- Sigma notation: $\Sigma x = $ __sum__ all of the data entries (x) in the data set.

μ • Population mean: Using every entry from the entire population. $\mu = \dfrac{\Sigma x}{N}$

$\bar{X}$ • Sample mean: Using data collected from a sample. $\bar{x} = \dfrac{\Sigma x}{n}$

Median: The value that lies in the __middle__ of the data when the data set is **ordered**.
- Measures the center of an ordered data set by dividing it into two __equal__ parts.
- If the data set has an
  - odd number of entries: median is the middle data entry.
  - even number of entries: median is the __average__ of the two middle data entries.

Mode: The data entry that occurs with the __most__ frequency.

- If no entry is repeated the data set has __no__ __mode__.
- If two entries occur with the same greatest frequency, each entry is a mode (it's __bimodal__).

Example 1: The prices (in dollars) for a sample of roundtrip flights from Chicago, Illinois to Cancun, Mexico are listed. Find the mean, median, and mode price of the flights.
872 432 397 427 388 782 397

Mean $= \dfrac{sum}{7} = \dfrac{3695}{7} \approx 527.9$

Median: 388   397 397 / 427 / 432 782 872
   Arrange small → large, find middle
2

Mode : 397 (occurs twice & that's the most any of them occurs)

Example 2: At a political debate a sample of audience members was asked to name the political party to which they belong. Their responses are shown in the table. What is the mode of the responses?

Political Party	Frequency, f
Democrat	34
Republican	56
Other	21
Did not respond	9

most frequent

"Republican" occurs the most (56 times) so it is the mode
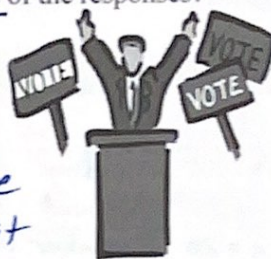
How could the mean of the above data be found? This is called a __weighted__ __mean__ .

Weighted Mean
- The mean of a data set whose entries have varying __weights__ . (like tests are weighted more than homework)

- $$\bar{x} = \frac{\Sigma(x \cdot w)}{\Sigma w}$$ where w is the weight of each entry x.

  ⤷ given as %, convert to decimal

Example 4: You are taking a class in which your grade is determined from four sources: 65% from your test scores, 10% from your assignments, 20% from your final exam, and 5% from your binder grade. Your scores are 86 (test average), 96 (assignment average), 82 (final exam), and 98 (binder). →partic. What is the weighted mean of your scores? If the minimum average for an A is 90, did you get an A?

$$\frac{x}{} \quad \frac{w}{} \quad \frac{x \cdot w}{}$$

Tests $86 \times .65 = 55.9$

Assignments $96 \times .10 = 9.6$

Final $82 \times .20 = 16.4$

Participation $98 \times .05 = 4.9$

$\Sigma w = 1.00 \quad \Sigma x \cdot w = 86.8$

$\bar{x} = \frac{\Sigma(x \cdot w)}{\Sigma w}$

$= \frac{86.8}{1.00}$

$= \boxed{86.8}$

Comparing the Mean, Median, and Mode
- All three measures describe a typical entry of a data set.
- Advantage of using the mean:
  - The mean is a reliable measure because it takes into account __every__ __entry__ of a data set.

- Disadvantage of using the mean:
  - Greatly affected by __outliers__ (a data entry that is far removed from the other entries in the data set). Do **NOT** use the mean as a "typical entry" if there exists one or more extreme outliers!

3

Example 5: Find the mean, median, and mode of the sample ages of a college class shown. Which measure of central tendency best describes a typical entry of this data set? Are there any outliers?

Ages in a class						
20	20	20	20	20	20	21
21	21	21	22	22	22	23
23	23	23	24	24	65	

$n = 20$ (number of entries)

mean, $\bar{X} = \frac{sum}{n} = \frac{475}{20} \approx \boxed{23.8 \text{ years old}}$

median: between 21 & 22 $= \frac{21+22}{2} = \boxed{21.5 \text{ years old}}$

Mode: $\boxed{20 \text{ years old}}$ (most frequent — 6 of them)
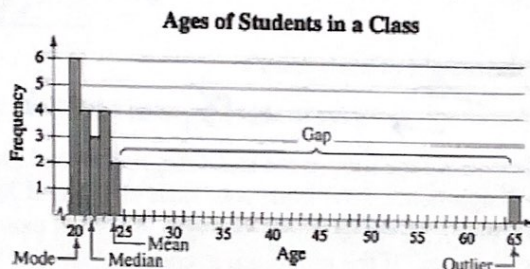
Things to consider…
1) Is there an outlier? If so, do not use the ___mean___ . Yes: 65 is an outlier
2) Is the mode a "typical central" entry? No — it's the youngest age
3) The median is a good choice if the mean cannot be used.

Best: MEDIAN

Sometimes a graphical approach can help you decide which entry is best to use.

Ages of Students in a Class

Shapes of Distributions
- o Symmetric
- o Uniform
- o Skewed Left
- o Skewed Right

Symmetric Distribution
- A ___vertical___ line can be drawn through the middle of a graph of the distribution and the resulting halves are approximately ___mirror___ images. The mean, median, and mode are the ___SAME___ values.
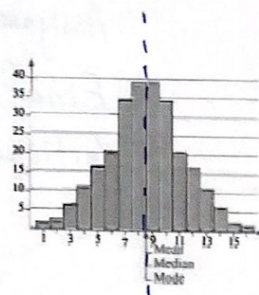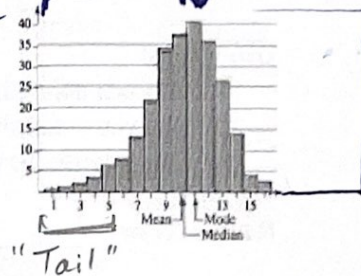

Uniform Distribution (___rectangular___)
- ___All___ entries or classes in the distribution have equal or *approximately* equal frequencies.
- Symmetric (and so the mean, median, and mode are the same values.)

4
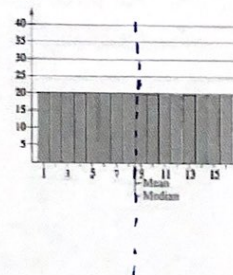
The graph is more heavily piled up on the LEFT   Mode (highest peak)

Skewed Left Distribution (negatively skewed)
- The "___tail___" of the graph elongates more to the left.
- The ___mean___ is to the left of the median.

$$Mean < median < mode$$

"Tail"

More weight on the RIGHT

Skewed Right Distribution (positively skewed)
- The "___tail___" of the graph elongates more to the right.
- The ___mean___ is to the right of the median.

$$mode < median < mean$$

"tail"
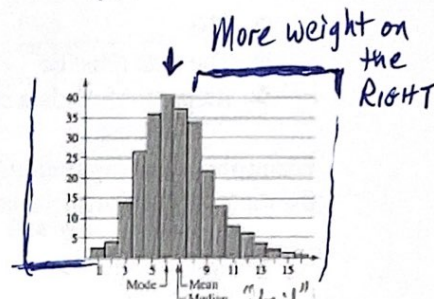
Using a calculator to find Mean and Median ☺
Step 1: Input your data.

Button	Comments
STAT	
EDIT	(or hit ENTER)
Arrow up to L1 and hit CLEAR	To erase any previous entries
Type in each data entry, and hit ENTER after each one.	Note: The data does NOT have to be ordered.
Make sure you don't forget to clear out L1… there could be "hidden" entries that impact your calculations!	

Step 2: Perform calculations with your data.

Button	Comments
	Note: L1 must be entered!!!
STAT	
CALCULATE	Use the right arrow button to select this.
ENTER	To select 1-Var Stats (one variable statistics)
ENTER ENTER ENTER	To activate the command. L1 is the default… you can pick another list by typing in a new one (hit 2nd first.)
You can easily see the mean, and if you scroll down, you will find the median (M.) Many other statistical values are also given that we will discuss this semester. → Med =	

→ $\bar{x}$

Example 6: Use your calculator to find the mean and median of the following list of monthly fees, in dollars, for car insurance for teenagers:

247, 130, 290, 150, 160, 221, 205, 187, 199, 220, 263, 144, 172, 261, 220, 201, 189, 240, 156

mean: $\bar{x} = \$202.89$

median: Med = $\$201.00$

more piled up on the right

5

Note: mean is to the right of median
⟹ Right-Skewed Distrib

130   170   210   250   290