Measures of Central Tendency

It is often convenient to use a central value to summarise a set of data. People frequently use a simple arithmetic average for this purpose. However, these are several different ways to find values around which a set of data tends to cluster. Such values are known as **measures of central tendency.**
In statistics, the three most commonly used measures of central tendency are the mean, median, and mode. Each of these measures has its particular advantages and disadvantages for a given set of data. Note that some distributions have outliers and you should pay a close attention.

Mean is defined as the sum of the values of a variable divided by the number of values and is calculated using

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}}$$

n - number of data values
x_i - each data value

Ex: Find the mean of 5, 7, 9, 8, 6, 5, 4, 10. $n = 8$

$$\therefore \bar{x} = \frac{5+7+9+8+6+5+4+10}{8} = \frac{54}{8} = 6.75$$

The **median** is the middle entry in an ordered list. There are as many data points above it as below it. To find the median,

a) If there is an odd number of data points, take the middle one (i.e. if there are 13 numbers, the median is the value of the 7th number when they are isted in ascending order).
b) If there is an even number of data points, the median is the average of the middle two numbers.

Ex: Find the median mark for each list of student grades.

a) 62, 64, 76, 89, 72, 54, 93

54, 62, 64, 72, 76, 89, 93

$\therefore$ Median = 72

b) 56, 84, 63, 67, 62, 98

First, list the numbers in ascending order.

56, 62, 63, 67, 84, 98

$\therefore$ Median $= \frac{63+67}{2} = 65$

The **mode** is the most frequent number in a data set. There can be no mode as well as more than one mode.

Ex: Find the mode(s) for each list of numbers.

a) 5, 7, 9, 8, 6, 5, 4, 10

$\therefore$ The mode is 5.

b) 25, 30, 32, 30, 25, 29

$\therefore$ The modes are 25 and 30.

c) 63, 57, 66, 83, 79, 72, 79, 69, 60, 63, 79, 85, 80

$\therefore$ The mode is 79.

A **weighted mean** gives a measure of central tendency that reflects the relative importance of the data:

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} w_i x_i}{\sum_{i=1}^{n} w_i}, \quad \text{OR } \bar{x} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}}$$

where x_i represents each data value and w_i represents its weight, or frequency.

When a set of data has been grouped into intervals, you can approximate the mean using the formula

$$\boxed{\bar{x} = \frac{\sum_{i=1}^{n} f_i m_i}{\sum_{i=1}^{n} f_i}, \quad \text{OR } \bar{x} = \frac{f_1 m_1 + f_2 m_2 + \dots + f_n m_n}{f_1 + f_2 + \dots + f_n}}$$

where f_i is the frequency for that interval and m_i is the midpoint value of an interval.

Ex (Finding the mean for grouped data: A sample of car owners were asked how old they were when they got their first car. Results are shown in the following table:

Age	Frequency, f	Midpoint, m	Frequency x Midpoint, f x m
16-20	10	18	10x18 = 180
21-25	18	23	18x23 = 414
26-30	12	28	12x28 = 336
31-35	8	33	8x33 = 264
36-40	2	38	2x38 = 76
		+	1270

a) Find the median and modal age.
b) Find the mean age.

Ans: $n = 50$

a) $\bar{x} = \dfrac{f_1 m_1 + f_2 m_2 + f_3 m_3 + f_4 m_4 + f_5 m_5}{f_1 + f_2 + f_3 + f_4 + f_5}$

$= \dfrac{10\times18 + 18\times23 + 12\times28 + 8\times33 + 2\times37}{10 + 18 + 12 + 8 + 2}$

$= \dfrac{180 + 414 + 336 + 264 + 76}{50}$

$= \dfrac{1270}{50}$

$= 25.4$

∴ The mean age is 25.4 years old.

b) Median $= \dfrac{t_{25} + t_{26}}{2}$

$= \dfrac{23 + 23}{2}$

$= 23$

Mode $= 23$

∴ Both, the median age and the modal age are 23 years old.

Ex: Find mean and median age of the data in previous example using Fathom.

Sol:
- Enter data into Fathom table. Create a Histogram
- Right click on the grapł and chose Plot value
- Type mean(Age)+median(Age)
- Change a max data value of 38 to 78.
- What do you notice? The mean increased from 25.4 to 26.2 and mean(Age)+median(Age) increased from 48.4 to 49.2. So, this shows that changing a data value has a greater effect on the mean than the median.

Homework: Pg 158: #1-3,5,6,8,10