

# NOTES :

## CHAPTER 11 STATISTICS

### 1. MEASURES OF CENTRAL TENDENCIES

#### UNGROUPED DATA

##### MODE

The number that occurs most often in a set of numbers.

**HOW TO FIND THE MODE:**

- 1) put the numbers in order from smallest to largest
- 2) find the number that occurs the most

i.e. 2, 8, 12, 14, 19, 30, 30, 42, 51

**Special Cases**

- \* No mode - all numbers occur the same amount of times
- \* More than one mode - 1 or more numbers occur the most amount of times

##### MEDIAN

the middle number

**HOW TO FIND THE MEDIAN:**

- 1) put the numbers in order from smallest to largest.
- 2) cover up one number on each end until you get to the middle.

**If there are 2 numbers in the middle:** add them together and divide the answer by 2.

##### MEAN

the average value of numbers in a set.

**HOW TO FIND THE MEAN:**

- 1) add up all of the numbers
- 2) divide the answer by the number of values added together

i.e.  $6 + 7 + 9 + 2 \div 4$  (the number of values)  
= 6 (mean)

#### GROUPED DATA

**Estimated Mean:** *The mean of a grouped data that uses the midpoint to represent the class, thus resulting to an estimated mean.*

Class Interval	Mid-point	Frequency	Mid-point $\times$ Frequency
$140 \leq h < 150$	145	6	$145 \times 6 = 870$
$150 \leq h < 160$	155	16	$155 \times 16 = 2480$
$160 \leq h < 170$	165	21	$165 \times 21 = 3465$
$170 \leq h < 180$	175	8	$175 \times 8 = 1400$
<b>Totals</b>		<b>51</b>	<b>8215</b>

$$\begin{aligned} \text{Estimated Mean} &= \frac{8215}{51} \\ &= 161 \text{ (to the nearest cm)} \end{aligned}$$

**Median Class:** *The class that contains the middle-most value*

The median is the 26th value. In this case it lies in the  $160 \leq h < 170$  class interval.

**Modal Class:** *The class that has the most frequency.*

The modal class is  $160 \leq h < 170$  as it contains the most values.

3.

(a) CUMULATIVE FREQUENCY

A sub-total of frequencies from the beginning up to the current class.

LENGTH (CM)	FREQUENCY	CUMULATIVE FREQUENCY
20.5 < h ≤ 24.5	3	3
24.5 < h ≤ 28.5	7	(3+7)=10
28.5 < h ≤ 32.5	12	(3+7+12)=22
32.5 < h ≤ 36.5	6	(3+7+12+6)=28
36.5 < h ≤ 40.5	4	(3+7+12+6+4)=32

(b) CUMULATIVE FREQUENCY CURVE

Curve/Graphical representation of data using the cumulative frequencies

(c) ESTIMATING QUANTILES

To use the Cumulative Frequency Curve to estimate quartiles:

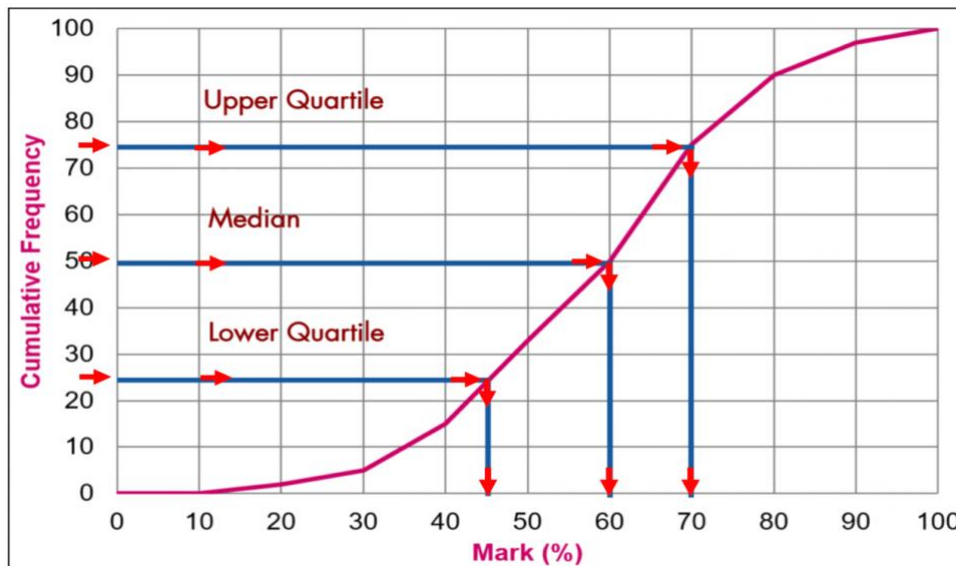
(1) Median: Take 50% of CF → → →

(2) Lower Quartile (Q1): Take 25% of CF → → →

(3) Upper Quartile (Q3): Take 75% of CF → → →

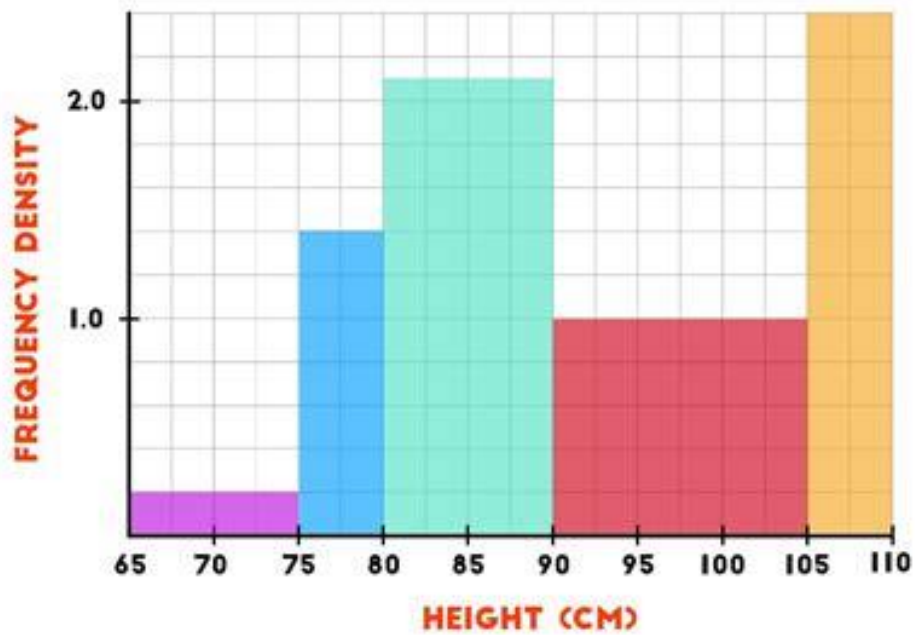
Find equivalent value on the **x-axis**

(Follow red arrows below!)



(4) To find the INTERQUARTILE RANGE (IQR) =  $Q_3 - Q_1$

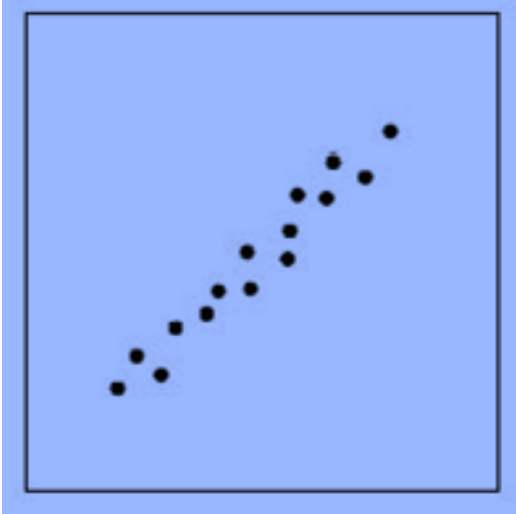

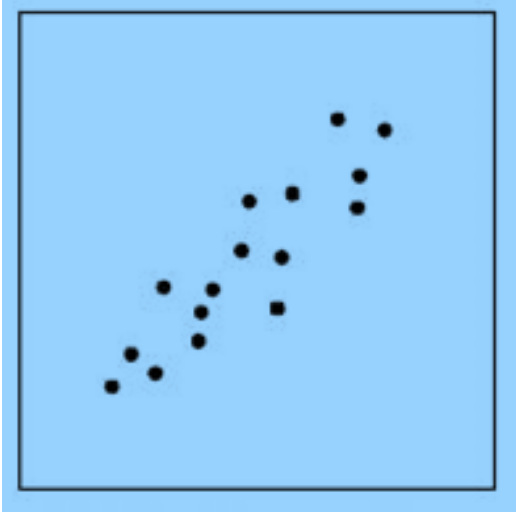
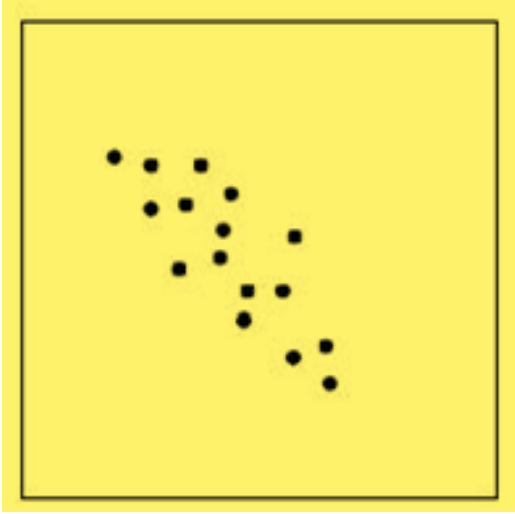
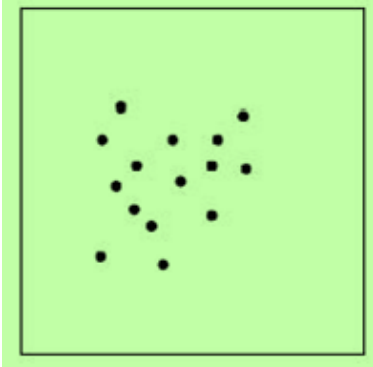
#### 4. FREQUENCY DENSITY



$$\text{Frequency Density} = \frac{\text{Frequency}}{\text{Class width}}$$

$$\text{Frequency} = \text{Frequency Density} \times \text{Class width}$$

4.  
(a) **CORRELATION**  
A Statistical measure that determines relationship involving two quantities. It helps identify whether one quantity strongly/weakly positive/negative affects the results of another quantity, or not all. A Scatter Diagram may be used to see the relationship of the two quantities involved.

STRONG – POSITIVE CORRELATION	STRONG – NEGATIVE
	
WEAK – POSITIVE CORRELATION	WEAK – NEGATIVE CORRELATION
	
<b>NO CORRELATION</b>	

## (b) LINE OF BEST FIT

### Line Of Best Fit - Investopedia

<https://www.investopedia.com/terms/l/line-of-best-fit.asp>

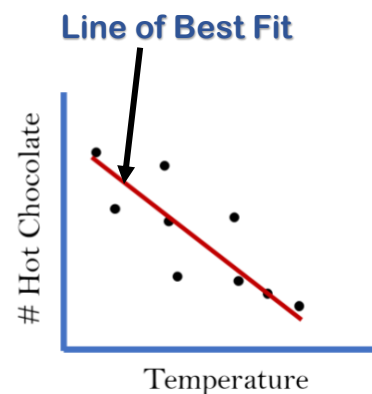
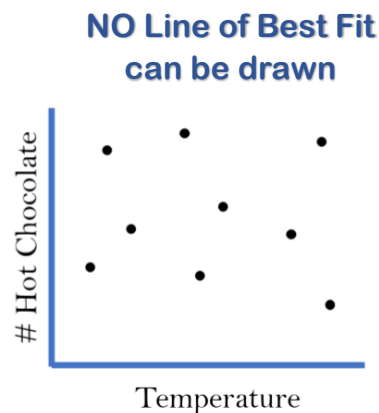
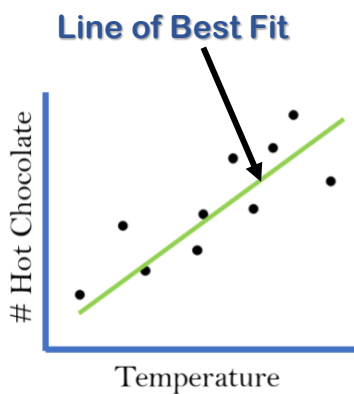
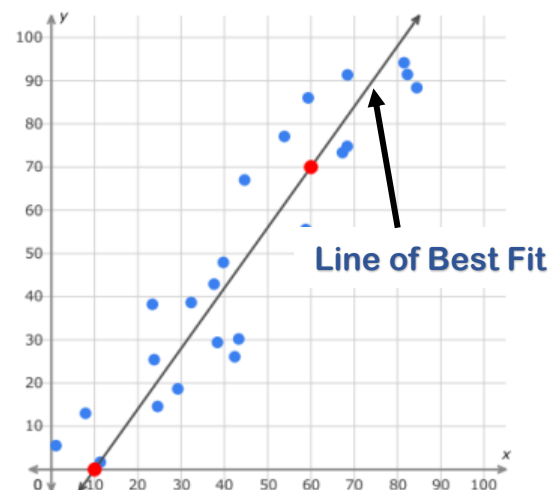
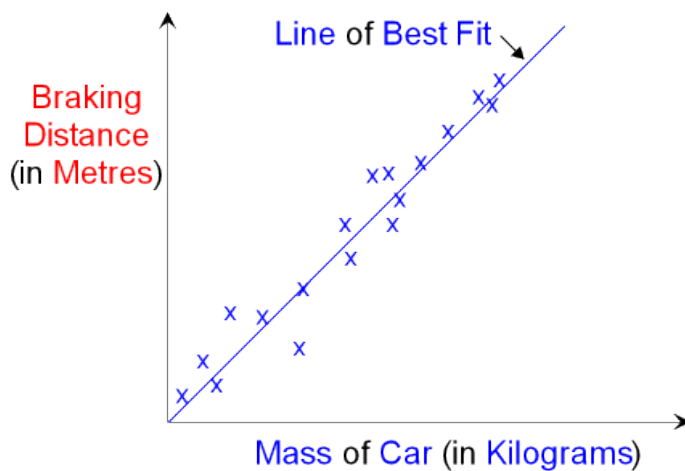
Line Of Best Fit: A line of best fit is a straight line drawn through the center of a group of data points plotted on a scatter plot. Scatter plots depict the results of gathering data on two

[https://www.varsitytutors.com/hotmath/hotmath\\_help/topics/line-of-best-fit-eyeball-met...](https://www.varsitytutors.com/hotmath/hotmath_help/topics/line-of-best-fit-eyeball-met...)

Line of Best Fit(Eyeball Method) A line of best fit is a straight line drawn through the maximum number of points on a scatter plot balancing about an equal number of points above and below the line. It is used to study the nature of relation between two variables.

**Important:**

**Line of Best Fit MUST pass through the point of averages of two quantities.**



**NO CORRELATION**