

CONSTRUCTION OF GROUPED FREQUENCY DISTRIBUTION

Example:

The following data represents the ages of 30 students in a statistics class. Construct a frequency distribution that has five classes.

Ages of Students

18	20	21	27	29	20
19	30	32	19	34	19
24	29	18	37	38	22
30	39	32	44	33	46
54	49	18	51	21	21

STEPS FOR CONSTRUCTION

STEP 1: Number of classes = $m = 5$ (given)

STEP 2: Find range

Minimum value = 18

Maximum value = 54

Range = $54 - 18 = 36$

STEP 3: Find Class width

$$h = \frac{\text{Range}}{m} = \frac{36}{5} = 7.2 \cong 8$$

STEP 4: Decide class intervals:

The minimum data entry¹⁸ may be used for the lower limit of the first class. To find the lower class limits of the remaining classes, add the width (8) to each lower limit.

STEP 5:

Construct the table having columns for class intervals, tally marks, frequency, class boundaries, midpoints, relative frequency, cumulative frequency (as needed).

Note: Make a tally mark for each data entry in the appropriate class. The number of tally marks for a class is the **frequency** for that class.

Example continued:

Ages

Ages of Students

Number of students

Class	Tally	Frequency, f
18 – 25		13
26 – 33		8
34 – 41		4
42 – 49		3
50 – 57		2
		$\sum f = 30$

Check that the sum equals the number in the sample.

STEP 6: Find Class Boundaries

Find the class boundaries for the “Ages of Students” frequency distribution.

Ages of Students

Class	Frequency, f	Class Boundaries
18 – 25	13	17.5 – 25.5
26 – 33	8	25.5 – 33.5
34 – 41	4	33.5 – 41.5
42 – 49	3	41.5 – 49.5
50 – 57	2	49.5 – 57.5
	$\sum f = 30$	

The distance from the upper limit of the first class to the lower limit of the second class is 1.

Half this distance is 0.5

$$\text{i.e., } \frac{26 - 25}{2} = 0.5$$

STEP 7: Find midpoint:

The **midpoint** of a class is the sum of the lower and upper limits of the class divided by two.

$$\text{Midpoint} = \frac{\text{upper limit} + \text{lower limit}}{2} = \frac{1+4}{2} = 2.5$$

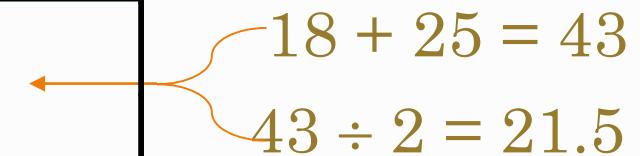
Class	Frequency, f	Midpoint
1 – 4	4	2.5

Example:

Find the midpoints for the “Ages of Students” frequency distribution.

Ages of Students

Class	Frequency, f	Midpoint
18 – 25	13	21.5
26 – 33	8	29.5
34 – 41	4	37.5
42 – 49	3	45.5
50 – 57	2	53.5
	$\sum f = 30$	

$$\begin{aligned}18 + 25 &= 43 \\43 \div 2 &= 21.5\end{aligned}$$
An orange curly brace groups the numbers 18 and 25, with an arrow pointing to the equation 18 + 25 = 43. Another orange curly brace groups the result 43 and the divisor 2, with an arrow pointing to the equation 43 ÷ 2 = 21.5.

Relative Frequency

The **relative frequency** of a class is the portion or percentage of the data that falls in that class.

To find the relative frequency of a class, divide the frequency f by the sample size n .

$$\text{Relative frequency} = \frac{\text{Class frequency}}{\text{Sample size}} = \frac{f}{n}$$

Class	Frequency, f	Relative Frequency
1 – 4	4	0.222

$$\sum f = 18$$

$$\text{Relative frequency} = \frac{f}{n} = \frac{4}{18} \approx 0.222$$

Percentage Frequency

The **percentage frequency** of a class is the percentage of the data that falls in that class.

To find the %age frequency of a class, multiply the relative frequency f by 100.

$$\text{\%age frequency} = \frac{(\text{class frequency}) * 100}{\text{Sample size}}$$

or

$$\text{\%age } f = (\text{relative } f) \times 100$$

Example continued..

Find the relative frequencies for the “Ages of Students” frequency distribution.

Class	Frequency, f	Relative Frequency
18 – 25	13	0.433
26 – 33	8	0.267
34 – 41	4	0.133
42 – 49	3	0.1
50 – 57	2	0.067
	$\sum f = 30$	$\sum \frac{f}{n} = 1$

Portion of students

$$\frac{f}{n} = \frac{13}{30} \approx 0.433$$

Cumulative Frequency

The **cumulative frequency** of a class is the sum of the frequency for that class and all the previous classes.

Ages of Students

Class	Frequency, f	Cumulative Frequency
18 – 25	13	13
26 – 33	+ 8	21
34 – 41	+ 4	25
42 – 49	+ 3	28
50 – 57	+ 2	30
	$\sum f = 30$	

Total number
of students

GRAPHICAL REPRESENTATION OF GROUPED FREQUENCY DISTRIBUTION

Frequency Histogram

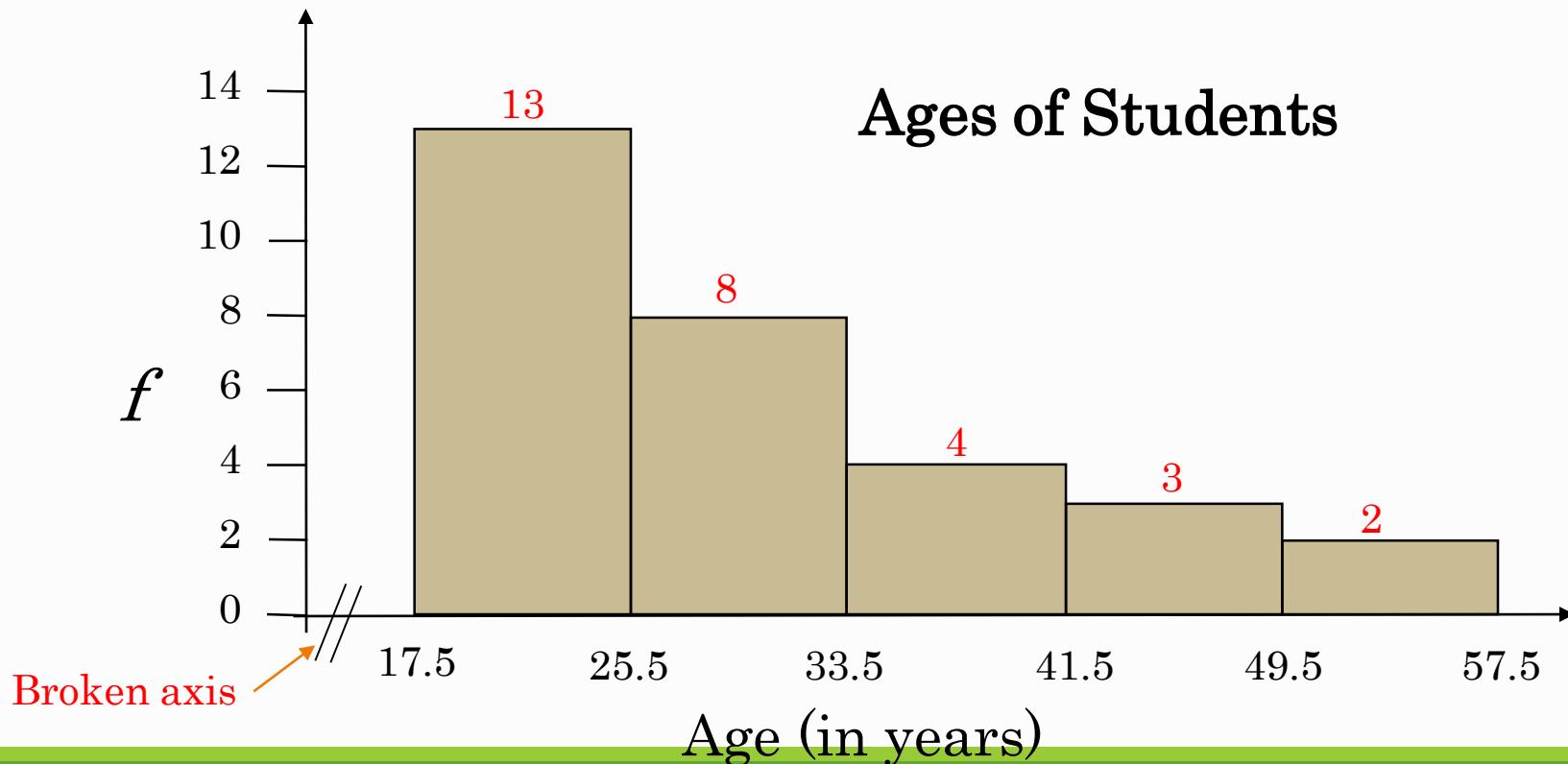
A **frequency histogram** is a bar graph that represents the frequency distribution of a data set.

1. The horizontal scale is quantitative and measures the data values.
2. The vertical scale measures the frequencies of the classes.
3. Consecutive bars must touch.

Simply Histogram is the bar diagram between class boundaries and the frequencies.

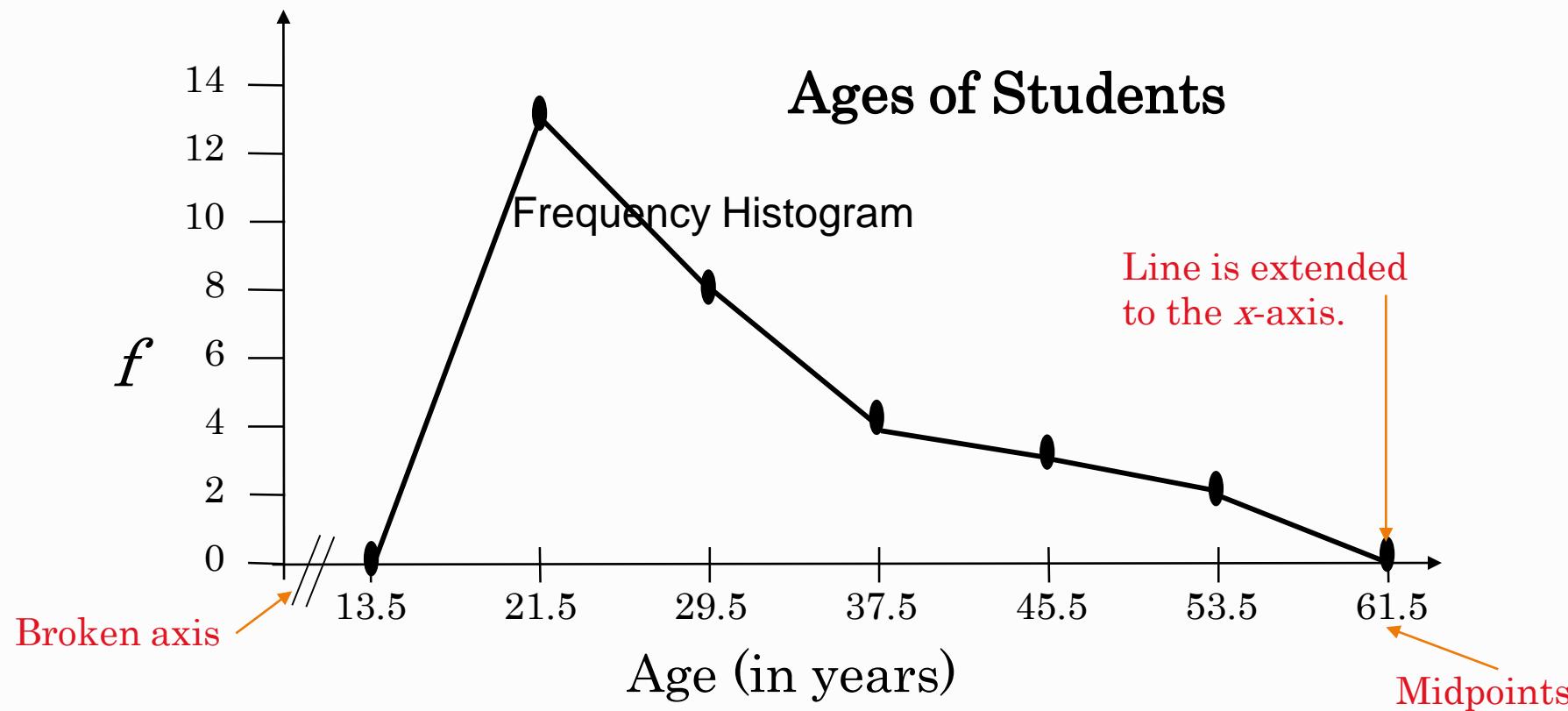
Example continue:

Draw a frequency histogram for the “Ages of Students” frequency distribution. Use the class boundaries.



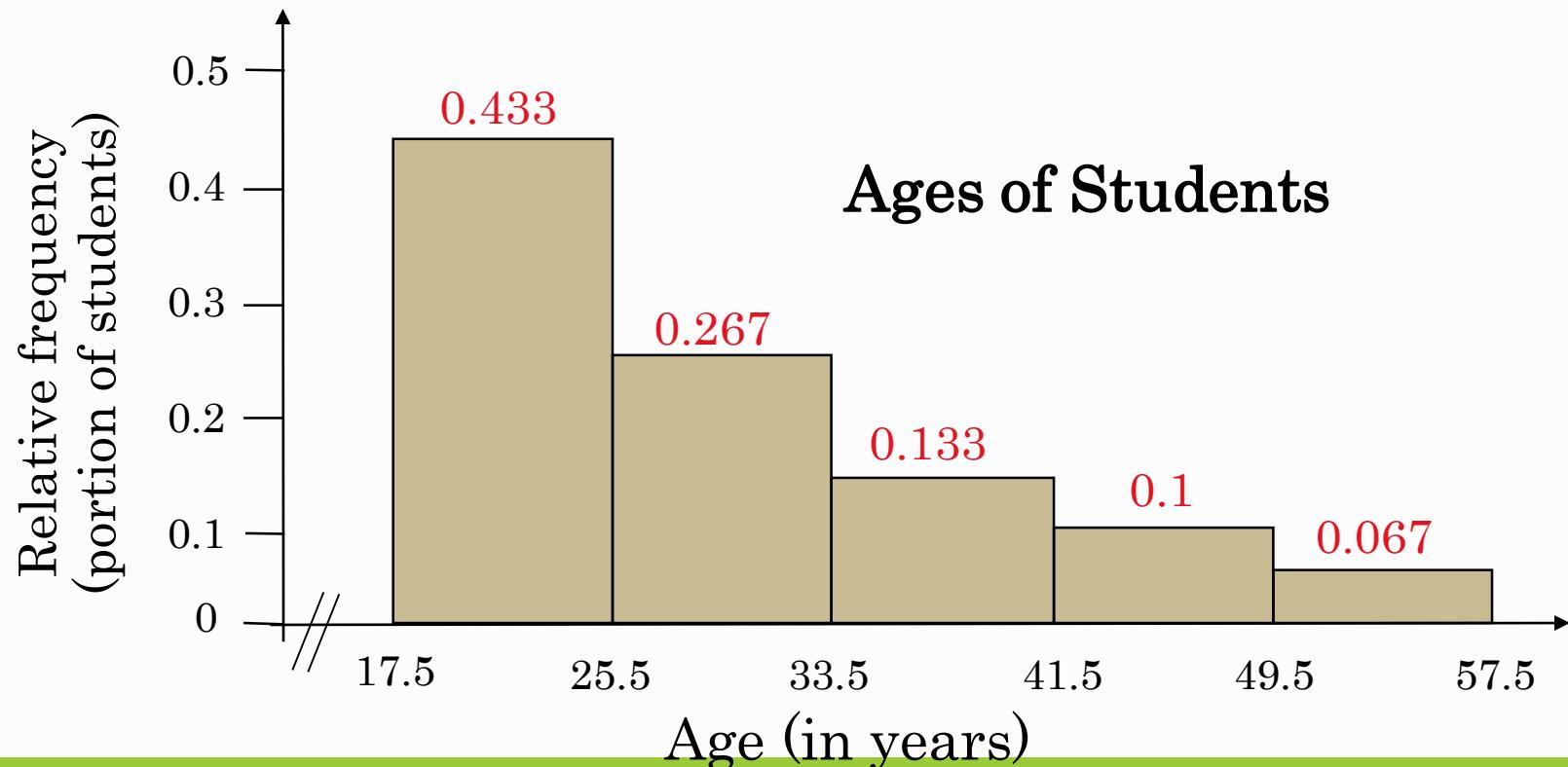
Frequency Polygon

A **frequency polygon** is a line graph that emphasizes the continuous change in frequencies. (Graph between midpoints and frequencies)



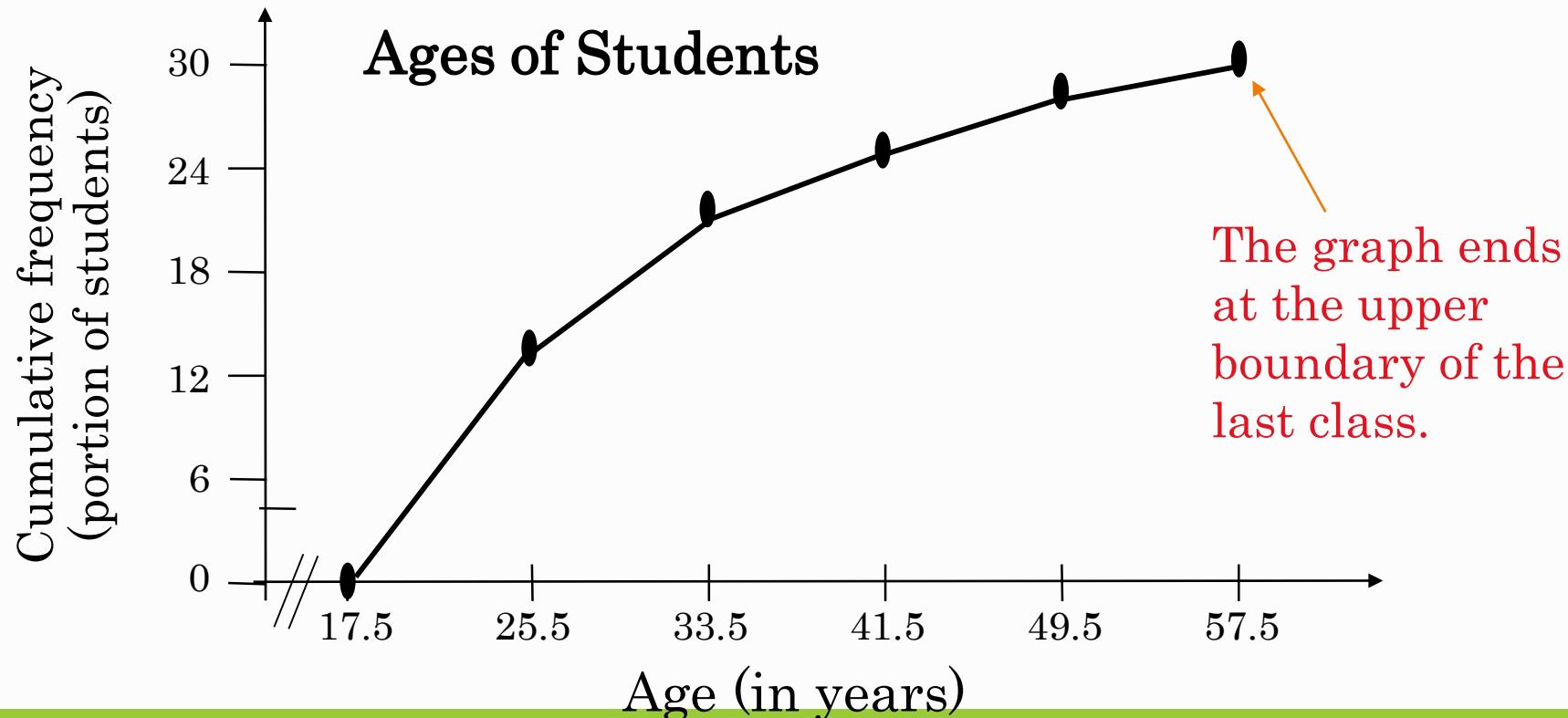
Relative Frequency Histogram

A **relative frequency histogram** has the same shape and the same horizontal scale as the corresponding frequency histogram.



Cumulative Frequency Polygons (Ogives)

A **cumulative frequency polygons or ogive**, is a line graph that displays the cumulative frequency of each class at its upper class boundary.

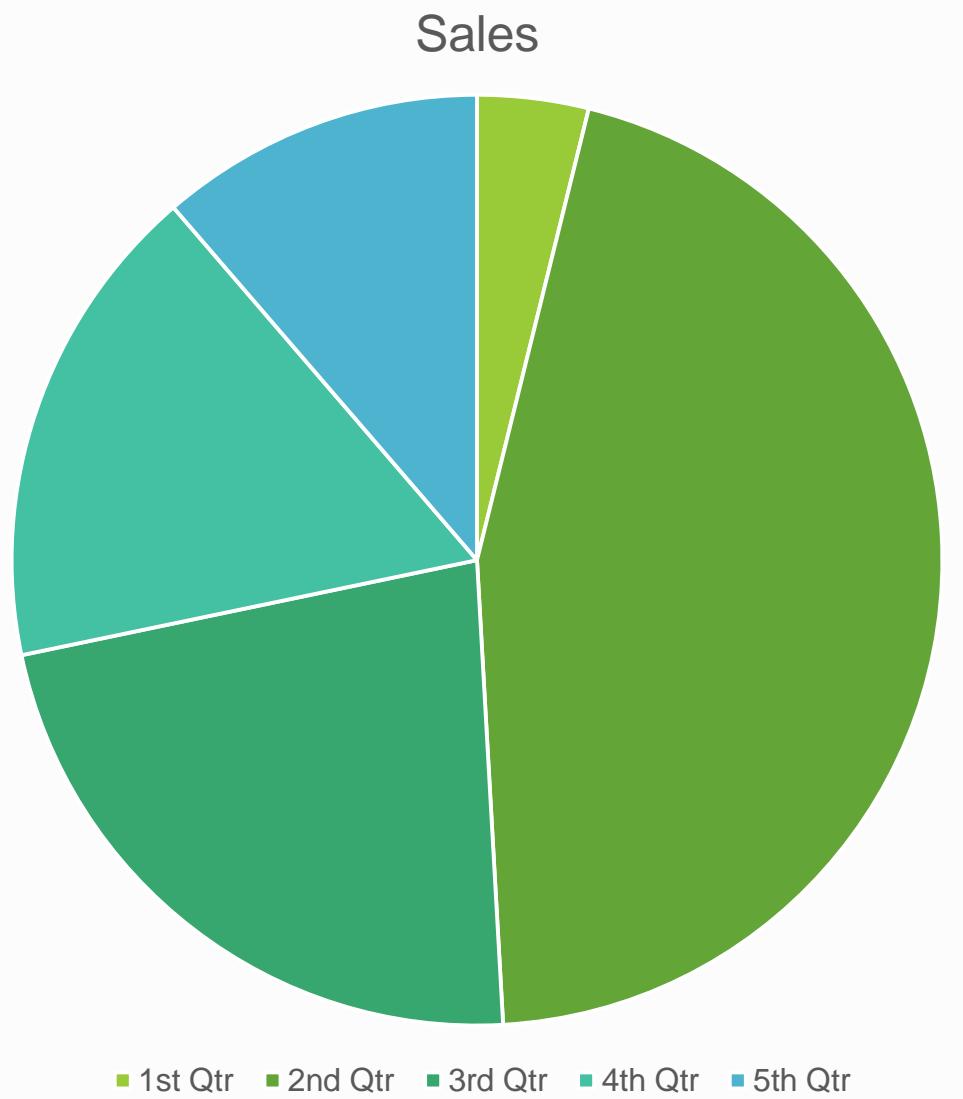


Pie Chart

A **Pie Chart or Pie diagram** is a graphical device of circular shape divided into sub-divisions in terms of angles, called sectors.

$$\text{sector angle} = \frac{\text{component value}}{\text{total value}} \times 360^{\circ}$$

frequency	Sector angle
13	156
8	96
4	48
3	36
2	24
$\Sigma=30$	$\Sigma=360$



Task:

Pick a data of your choice either from book or some data on covid 19 from google, construct a frequency distribution for it and draw all the graphical representations for it which we have learnt in the class.