

3 - Statistical Representation

Statistics

- Qualitative Data → Not a number, often descriptive eg. colour/make of car
- Quantitative Data → Data with numbers in it e.g. time, height, shoe size, age. Can be discrete or continuous
- Continuous Data → Data that can take on any value/range e.g. time/height
- Discrete Data → Data where there are no in-between values, specific. eg. how many children you have/Gender
- Bivariate Data → Combination of two variables (values) e.g. age + height
- Primary Data → Data you have collected yourself e.g. survey
- Secondary Data → Information from somebody else e.g. internet

Data Handling

- No overlaps
- Simple Understandable language/specific words
- Short questions with precise answers
- Tick Boxes
- Avoid open-ended questions
- Avoid leading questions/Unbiased question
- Time frames

Primary Data →

- can be expensive/time consuming
- trustworthy/ no bias
- better understanding of the data
- could have made mistakes
- less accuracy

Averages

- Mean = add all values and ÷ by quantity of values, total data ÷ total frequency

e.g. 1, 2, 2, 4, 6, 6, 7, 8

total = 36

quantity of values = 8

mean = $36/8 = 4.5$

- Median = align numbers in ascending order and find the middle number.
- Middle number = $\frac{1}{2}(N+1)$, N = total amount of data

if there is no middle number (even data), **+ the middle numbers then ÷ 2**

e.g. 1, 2, 2, 4, 6, 6, 7, 8 → median = $(4+6) \div 2 = 5$

- Mode = most common, highest frequency

e.g. 1, 3, 3, 4, 3, 6, 7 → mode = 3

- Set of Data →

$$\text{Mean} = \frac{\sum fx}{\sum f}$$

f = frequency

x = data

\sum = total

- Grouped data →

$$\text{Mean} = \frac{\sum fmp}{\sum f}$$

mp = midpoint

x	f	$c.f$
18	4	4
23	6	10
28	8	18
33	14	32
38	8	40
43	6	46
48	4	50
	50	

$$\text{Median} = \text{Size of } \frac{N}{2} \text{th value}$$

$$\text{Median} = \text{Size of } \frac{50}{2} = 25\text{th value}$$

f = frequency

Σ = total

Mode = group with highest frequency, modal group

Median = uses compound frequency

odd \rightarrow find what group $\Sigma \div 2$ th value would be in

even \rightarrow find what group $(\Sigma + 1) \div 2$ th value would be in

Average	Advantages	Disadvantages
Mode	<ul style="list-style-type: none"> • Quick and easy to determine • Actual value of data • Not affected by extreme values • Can be used for non-numerical data 	<ul style="list-style-type: none"> • May not exist • Does not use all values • Sometimes not very informative • Can change dramatically from sample to sample • Might be more than one
Median	<ul style="list-style-type: none"> • Resistant to outliers (extreme values) • Easy to find for ungrouped data 	<ul style="list-style-type: none"> • May not be so informative • Does not represent all data • Does not use all values • Often mis-understood
Mean	<ul style="list-style-type: none"> • Easy to find • Uses all values • Total number of values can be calculated from it 	<ul style="list-style-type: none"> • Extreme values can distort it • Has to be calculated

Range = Measure of spread

To compare consistency

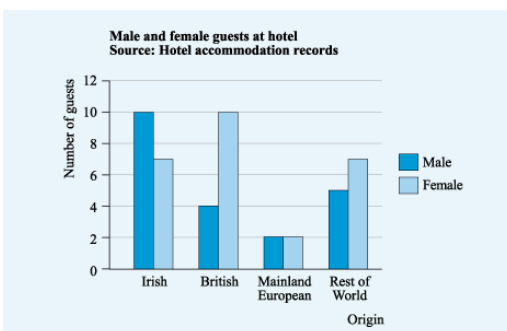
Lower the range, the higher the consistency

Helps us to make a conclusion

• Pie Charts \rightarrow

$$360/f = x = \text{degrees per } f$$

• Comparative Bar Chart:



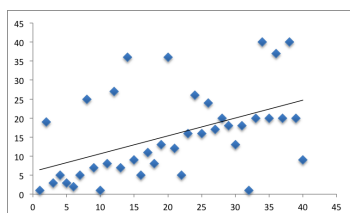
• Graphs \rightarrow

Axis Labels

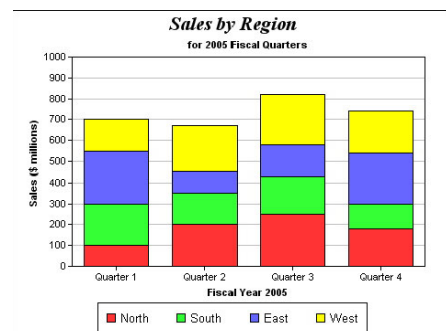
Key

Accurate Scale

Titles



Composite Bar Chart (stacked):



- Scatter Graph →

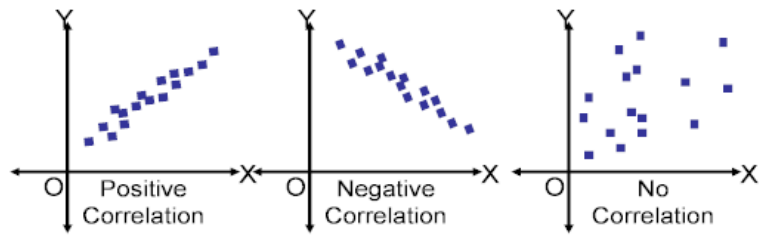
Compares 2 sets of data
 Comparison can be used to look for connections between data
 The connection is called correlation
 Line of best fit

Positive correlation = e.g. More illness + more absences

Negative correlation = e.g. More students + less vacancies in schools

No correlation = e.g. Pencil case colour + hair length

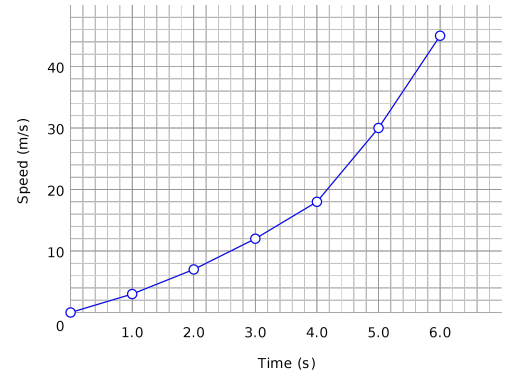
SCATTER PLOT EXAMPLES



- Line Graph →

Join each point individually with straight lines
 Used to show trend over a number of days or hours. Plotted as a series of points
 The end of the line graph does not have to join the axis

Trend → general direction of movement or change



- Stem and Leaf →

- include **key**

