GCSE MATHEMATICS · FORMULA & REVISION SHEET 3 OF 3

Statistics & Probability

Averages, charts, probability rules, tree diagrams and the data-handling cycle — Foundation core with Higher-tier extensions clearly marked. Works for AQA, Edexcel, OCR and WJEC.

FOUND + HIGHER needed by everyone HIGHER Higher tier only

1 Averages & Spread

The Three Averages — FOUND + HIGHER

Mean = $\frac{\textit{sum of values}}{\textit{number of values}}$

Median = *middle value when ordered*
for n values, position = (n+1)/2

Mode = *most frequent value*

Range = highest − lowest
(a measure of spread, not an average)

Averages from Tables — FOUND + HIGHER

Mean from frequency table = $\frac{\Sigma(fx)}{\Sigma f}$

Grouped data: use **midpoints** *for x → estimated mean*

Modal class = *group with highest frequency*

Quartiles & IQR — HIGHER

Lower quartile (Q1): $\frac{1}{4}$ *of the way through*

Upper quartile (Q3): $\frac{3}{4}$ *of the way through*

Interquartile range = Q3 − Q1
smaller IQR = more consistent data

Comparing Distributions — FOUND + HIGHER

Compare an **average** *(typical value)…*

…and a **measure of spread** *(consistency)*

Always write comparisons **in context** *of the question*

2 Charts, Tables & Diagrams

Common Charts — FOUND + HIGHER

Bar chart: *discrete data, equal-width bars*

Pictogram: *uses a key/symbol*

Pie chart: *360° ÷ total = degrees per item*

Line graph: *trends over time*

Scatter Graphs — FOUND + HIGHER

Correlation: *positive, negative, or none*

Line of best fit: *straight line through the trend*

Interpolation *(within data) is reliable;* **extrapolation** *(beyond) is not*

Frequency Diagrams — FOUND + HIGHER

Frequency polygon: *plot midpoints, join with lines*

Two-way table: *rows & columns, totals must agree*

Stem-and-leaf: *ordered, with a key*

Histograms — HIGHER

Unequal class widths → use **frequency density**

Frequency density = $\frac{\textit{frequency}}{\textit{class width}}$

Frequency = frequency density × class width = *area of bar*

Cumulative Frequency

HIGHER

Running total of frequencies

Plot against the **upper** *class boundary*

Read off median (½n), Q1 (¼n), Q3 (¾n)

Use to draw a **box plot**

Box Plots

HIGHER

Show: min, Q1, median, Q3, max

Box spans Q1 → Q3 (the IQR)

Great for comparing two data sets at a glance

3 Probability

Probability Basics

FOUND + HIGHER

$$\text{P(event)} = \frac{\textit{favourable outcomes}}{\textit{total outcomes}}$$

All probabilities lie between **0 and 1**

P(not A) = 1 − P(A)

Probabilities of all outcomes sum to 1

Expected Frequency

FOUND + HIGHER

Expected number = P(event) × number of trials

$$\text{Relative frequency} = \frac{\textit{times it happened}}{\textit{total trials}}$$

experimental probability

Combined Events

FOUND + HIGHER

AND rule (independent): *P(A and B) = P(A) × P(B)*

OR rule (mutually exclusive): *P(A or B) = P(A) + P(B)*

General OR rule (Higher): *P(A or B) = P(A) + P(B) − P(A and B)*

Use a **sample space diagram** *for two events*

Tree Diagrams

FOUND + HIGHER

Multiply along *the branches (AND)*

Add *the relevant end results (OR)*

Each set of branches must sum to 1

Higher: 'without replacement' changes the 2nd probabilities

Venn Diagrams & Set Notation

HIGHER

A ∩ **B** = *intersection (in both)*

A ∪ **B** = *union (in either)*

A′ = *complement (not in A)*

ξ = *the universal set (everything)*

Conditional Probability

HIGHER

Probability of B **given** *A has happened*

Reduce the total to only the outcomes where A occurred

Common in 'without replacement' problems

4 Sampling & the Data Handling Cycle

Sampling

FOUND + HIGHER

Population: *the whole group studied*

Sample: *a smaller part used to represent it*

Random sample: *everyone equally likely to be chosen*

Bias: *a sample that isn't representative gives misleading results*

Bigger, more random samples → more reliable conclusions

Data Types

FOUND + HIGHER

Qualitative: *words / categories*

Quantitative: *numerical*

Discrete: *counted (whole values)*

Continuous: *measured (any value in a range)*

Capture–Recapture

HIGHER

Estimate a population size:

$$\frac{\textit{total population}}{\textit{1st sample marked}} = \frac{\textit{2nd sample size}}{\textit{marked in 2nd sample}}$$

Assumes population unchanged & mixing is random

Stratified Sampling

HIGHER

Sample reflects group proportions

$$\textbf{From each group} = \frac{\textit{group size}}{\textit{total}} \times \textbf{sample size}$$

Fairer than simple random when groups differ in size

Exam technique — statistics: For "compare the data" questions you must mention *both* an average and the spread, and phrase it in context (not "the mean is higher" but "on average Class A scored higher"). On tree-diagram and "without replacement" questions, write the second-stage fractions with the new denominator — that single step is where most marks are lost.