

Mapping to AQA 2009 Statistics specification

Higher tier objectives appear in bold text.

Whole lessons or sections of lessons are given to match each objective as appropriate.

3.1.1 Planning a strategy	Lessons
Hypothesis	
Specifying a hypothesis to be tested.	Types of data – Planning a statistical enquiry
Planning an Investigation	
Determining the data needed to address hypotheses and selecting an appropriate method for obtaining the data. Justifying the choice of method by comparing it with possible alternatives.	Types of data – Planning a statistical enquiry
Specifying a research question to be investigated and breaking it down into sub-questions as necessary.	Types of data – Planning a statistical enquiry
Deciding between survey/experiment	
Awareness of possible problems including: <ul style="list-style-type: none"> • identifying the population; • questionnaire distribution and collection; • non-response; • errors in recording answers; • missing data. 	Population and sampling

3.1.2 Data Collection	Lessons
Types of Data	
Raw data. Primary and secondary data sources. Qualitative and quantitative variables. Categorical data. Discrete and continuous data. Grouped and ungrouped data. Bivariate data. Implications of grouping for loss of accuracy in presentation and calculation.	Types of data
Classification of data; class limits and intervals. Implications of grouping for loss of accuracy in presentation and calculation.	Types of data – Scales of measurement Calculating the mean – Grouped frequency and the mean

Obtaining Data	
<p>Obtaining data by counting or measuring; accuracy of such measures. Design and use of efficient methods of recording data, appropriate to the purpose for which it will be used.</p>	<p>Types of data – Planning a statistical enquiry Population and sampling – Questionnaires</p>
Census Data	
<p>Obtaining information from well-defined populations.</p>	<p>Types of data – Planning a statistical enquiry Types of data – Secondary data</p>
Sampling	
<p>Purpose of sampling; variability between samples.</p> <p>Randomness. Random numbers from tables, calculators and computers.</p> <p>Sampling from a well-defined population. Sample frame. Simple random sampling; the condition that all members of the population are equally likely to be included in the sample. Use of stratification in sample design using a single category. Awareness of the dangers of convenience sampling. Stratified sampling with no more than two sets of categories. Cluster sampling and quota sampling with particular reference to its use in conducting large scale opinion polls. An awareness of multi stage sampling. Strengths and weaknesses of the various sampling methods, including the dangers of convenience sampling. The criteria used for selecting sample members in national opinion polls: geographical area, sex, age group, social and economic backgrounds. Associated sources of bias.</p> <p>Biased samples arising from sampling from a wrong population or non-random choice of individual elements.</p>	<p>Population and sampling – Methods of sampling</p> <p>Population and sampling – Methods of sampling</p> <p>Population and sampling – Methods of sampling</p> <p>Population and sampling – Methods of sampling</p>
Conducting a survey/experiment	
<p>Obtaining primary data by questionnaire. Use and reasons for pilot studies and pre-testing. Problems of design, wording, biased questions, definitions, obtaining truthful answers. The advantages and disadvantages of closed and open questions. The use of opinion scales. The technique of random response, in its</p>	<p>Population and sampling – Questionnaires</p> <p>Population and sampling – Sensitive questions</p>

<p>simplest form, for obtaining truthful answers to sensitive questions.</p> <p>Obtaining data by interview. Advantages and disadvantages of interviews compared with written questionnaires.</p> <p>Simulation. Use of, for example, dice, random number tables, ICT.</p> <p>Using secondary data; sources, reliability, accuracy, relevance and bias. Difference between sample and census data.</p> <p>Designing and obtaining data from simple statistical experiments. Obtaining data from observation or experiments (laboratory, field or natural experiments) Examples of extraneous variables. Issues of inter-observer bias. Explanatory and response variables; identification of the variables to be investigated. Use of a control group; use of random allocation to experimental and control groups. Matched pairs of groups; “before and after” experiments. Identification of extraneous variables and methods of controlling them: the need to hold extraneous variables constant for both groups.</p> <p>Surveys.</p>	<p>Population and sampling – Questionnaires</p> <p>Population and sampling – Methods of sampling</p> <p>Types of data – Secondary data Types of data – Planning a statistical enquiry</p> <p>Population and sampling – Experimental data</p> <p>Types of data – Planning a statistical enquiry</p>
---	--

3.1.3 Tabulation and Representation	Lessons
<p>Tabulation</p> <p>Construction of frequency tables by tallying raw data. Use of five bar gates expected.</p> <p>Class intervals. Open-ended classes.</p> <p>Simplifying tables by combining categories and reducing the number of significant figures; resulting effects on readability: identifying or masking of patterns/trends; loss of detail. Problems of under and over simplification</p>	<p>Continuous data and frequency polygons – Raw data and frequency tables</p> <p>Continuous data and frequency polygons – Grouping continuous data</p> <p>Continuous data and frequency polygons – Grouping continuous data</p>

<p>resulting from unsuitable choice of group size or number of significant figures.</p> <p>Reading and interpreting data presented in tabular form.</p> <p>Design of tables to summarise data effectively. Design and use of appropriate two-way tables.</p>	<p>Cumulative frequency – Cumulative frequency graphs</p> <p>Continuous data and frequency polygons – Raw data and frequency tables</p>
<p>Diagrammatic Representation</p>	
<p>Qualitative data: bar and pie charts, pictograms. Multiple and composite bar charts. Dot plots for small data sets. Comparative pie charts (area proportional to total frequency).</p> <p>Discrete data; vertical line graphs. Cumulative frequency step polygons.</p> <p>Continuous data; grouped frequency diagrams, including histograms, with equal class intervals. Frequency polygons. Cumulative frequency graphs. Population pyramids. Histograms with equal or unequal class intervals.</p> <p>Output gap charts.</p> <p>Stem and leaf diagrams.</p> <p>Choropleth maps</p> <p>Transforming data presentation from one form to another.</p> <p>The shapes and simple properties of frequency distributions; symmetrical, positive and negative skew. The shape and simple properties of the normal frequency distribution.</p> <p>Bivariate data: scatter diagrams.</p> <p>Time series: line graphs.</p> <p>Other diagrammatic representations for comparisons of data using length. Use of area and volume. Comparison of the</p>	<p>Presenting discrete data Presenting continuous data Bar charts and histograms</p> <p>Time series and quality assurance – Line graphs Cumulative frequency – Cumulative frequency step polygons</p> <p>Cumulative frequency – Cumulative frequency graphs Continuous data and frequency polygons – Frequency polygons Presenting continuous data – Population pyramids Bar charts and histograms – Histograms</p> <p>Bar charts and histograms – Output gap charts Presenting discrete data – Stem and leaf diagrams Presenting continuous data – Choropleth maps Presenting discrete data Presenting continuous data</p> <p>Skewness and standardized scores – Skewness</p> <p>Probability distributions – The normal distribution</p> <p>Scatter graphs</p> <p>Time series and quality assurance</p> <p>Presenting discrete data Presenting continuous data</p>

<p>various diagrammatic representations using area or volume, including their advantages and disadvantages.</p> <p>Visual misrepresentation: misuse or omission of origin or scale. Broken, incorrect or changed scales. Incomplete definitions and labelling. Simple misuse of area and volume (calculations not expected at Tier F) Misuse of length, area and volume in pictorial comparison.</p> <p>Read or interpret information presented in diagrammatic form; distinction between well and poorly presented data. Spotting possible errors in a data set by recognising outliers that do not fit a general pattern. Diagrams drawn from the media and from Government and other statistical sources may be used. Where these are not of the types named in the specification, the interpretation required will be at an appropriate level for Foundation or Higher tier.</p>	<p>Presenting continuous data – Misleading charts</p> <p>Presenting discrete data Presenting continuous data Range and quartiles – Outliers</p>
---	---

3.1.4 Data Analysis	Lessons
<p>Measures of Location</p>	
<p>Mean, median and mode for raw data. Use of change of origin when calculating the mean. Effect on the mean of linear transformations.</p> <p>Mean, median and mode for discrete frequency distributions. Modal class for grouped frequency distributions. Median for grouped frequency distributions. Mean for grouped frequency distributions. Use of Σfx notation</p> <p>Advantages and disadvantages of each of the three measures of location in a given situation. Reasoned choice of a measure of location appropriate to the nature of the data and the purpose of the analysis.</p> <p>Geometric mean.</p>	<p>Averages Calculating the mean – Linear transformations and the mean</p> <p>Calculating the mean – Grouped frequency and the mean</p> <p>Averages – Which average?</p> <p>Index numbers and geometric mean – Geometric mean</p>

Measures of Spread	
<p>Range.</p> <p>Quartiles for discrete data. Quartiles and percentiles, for grouped frequency distributions.</p> <p>Deciles.</p> <p>Interquartile range for discrete and continuous data. Interpercentile ranges. Interdecile range.</p> <p>Variance and standard deviation. Candidates should know how to use the formula for finding the variance and standard deviation. The formula for the variance and standard deviation will be given.</p> <p>Advantages and disadvantages of each of these measures of spread.</p> <p>Construction of box and whisker plots. Use of box and whisker plots to identify outliers.</p> <p>Calculation, interpretation and use of standardised scores.</p> <p>Use of tabulated data, diagrams, measures of location, measures of spread and skew to compare data sets. Calculation, interpretation and use of measures of skewness. Candidates should know how to use Pearson's measure. The formula for this will be given.</p>	<p>Range and quartiles – Range</p> <p>Range and quartiles – Interquartile range Cumulative frequency – Cumulative frequency step polygons</p> <p>Range and quartiles – Interquartile range Cumulative frequency – Cumulative frequency step polygons</p> <p>Standard deviation – Standard deviation</p> <p>Standard deviation – Choosing the best measure</p> <p>Presenting discrete data – Box plots</p> <p>Skewness and standardized scores – Standardized scores</p> <p>Skewness and standardized scores – Skewness</p>
Other Summary Statistics	
<p>Simple index numbers. Weighted index numbers. Chain base numbers. General Index of Retail Prices (RPI). General Index of Consumer Price (CPI). Indices to measure GDP and Retail sales</p> <p>Crude rates. Standardised rates. Candidates are not expected to calculate these rates but should be able to understand and interpret such measures.</p>	<p>Index numbers and geometric mean</p> <p>Skewness and standardized scores – Crude and standardized rates</p>

Time Series	
<p>Drawing a trend line by eye and using it for prediction. Evaluating and plotting appropriately chosen moving averages. Trend line based on moving averages.</p> <p>Identification of seasonal variation. Seasonal effect at a given data point. Average seasonal effect. Prediction of future values.</p> <p>'Z' charts</p>	<p style="text-align: center;">Scatter graphs – Lines of best fit Time series and quality assurance – Time series graphs</p> <p style="text-align: center;">Time series and quality assurance – Seasonal effects</p> <p style="text-align: center;">Time series and quality assurance – Z charts</p>
Quality Assurance	
<p>Plotting sample means, medians or ranges over time to view consistency and accuracy against a target value.</p>	<p style="text-align: center;">Time series and quality assurance – Quality assurance</p>
Correlation and Regression	
<p>Scatter diagrams. Recognition by eye of positive correlation, negative correlation, no correlation.</p> <p>The distinction between correlation and causality.</p> <p>Interpret values of Spearman's correlation coefficient in the context of a problem. Spearman's rank correlation coefficient as a measure of agreement; its calculation and limitation in interpretation.</p> <p>Interpret bivariate data presented in the form of a scatter diagram. Comparison of the degree of correlation between two or more pairs of data sets with reference to scatter diagrams. Product Moment correlation coefficient, and its interpretation.</p> <p>Fitting a straight line of best fit by eye through (x, y) to the plotted points on a scatter diagram. Obtaining the regression equation in the form $y = mx + c$; the interpretation of the regression coefficients m and c. Non-linear data.</p> <p>Interpolation and extrapolation.</p>	<p style="text-align: center;">Scatter graphs – Scatter graphs</p> <p style="text-align: center;">Scatter graphs – Cause and effect</p> <p style="text-align: center;">Correlation – Measuring correlation</p> <p style="text-align: center;">Correlation – Measuring correlation</p> <p style="text-align: center;">Scatter graphs – Lines of best fit Correlation – Non-linear relationships</p> <p style="text-align: center;">Scatter graphs – Lines of best fit</p>

Estimation	
<p>Estimation of population mean from a sample. Estimation of a population proportion from a sample; the use of this method of estimation in opinion polls. Variability in estimates from different samples and the effect of sample size. Estimation of population size based on the capture/recapture method. Conditions for this method to be appropriate. An elementary quantitative appreciation of appropriate sample size. Understanding what affects the accuracy of the estimates.</p>	<p>Population and sampling – Methods of sampling</p> <p>Population and sampling – Capture–recapture method</p>

3.1.5 Probability	Lessons
<p>Probability</p>	
<p>Probability of an event, impossible events, certain events. Use of words such as possible, likely. Putting events into order of probability. Probability on a scale from 0 to 1.</p> <p>Probability as the limit of relative frequency as the number of observations increases. Equally likely events.</p> <p>Sample space: pictorial representation; probability by counting. Use of Venn diagrams, tables and Cartesian grids.</p> <p>Exhaustive events.</p> <p>Mutually exclusive events, the addition law. The general addition law.</p> <p>Independent events, the multiplication law. The general multiplication law.</p> <p>Tree diagrams. Two stage only. Independent or ‘with replacement’ only</p> <p>An intuitive approach to conditional probability e.g using two-way tables or Venn diagrams.</p>	<p>Simple probability and independent events – Introduction to probability</p> <p>Diagrams, dependent events and experimental probability – Probability from experiments</p> <p>Simple probability and independent events – Combined events</p> <p>Simple probability and independent events – Combined events</p> <p>Simple probability and independent events – Combined events</p> <p>Simple probability and independent events – Combined events</p> <p>Simple probability and independent events – Tree diagrams for independent events Diagrams, dependent events and experimental probability – Dependent events</p> <p>Diagrams, dependent events and experimental probability – Dependent events</p>

<p>Expected frequencies. Comparison of actual frequencies with expected frequencies.</p>	<p>Diagrams, dependent events and experimental probability – Probability from experiments</p>
--	---

<p>3.1.6 Data Interpretation</p>	<p>Lessons</p>
<p>(an appreciation of limitations and conclusions)</p>	<p><i>Throughout</i></p>