

Yellow highlight/orange text = not in GCSE (9-1) 1MA1			
HIGHER			
Ref	1ST0 (2017)	1MA1 (2015)	Notes
1 The collection of data			
(a) Planning			
1a.01	Know that a hypothesis can be tested only through the appropriate collection and analysis of data. Formal use of null hypothesis will not be required. [S]		
1a.02	Know the constraints that may be faced in designing an investigation to test a hypothesis including factors such as time, costs, ethical issues, confidentiality, convenience. [S]		
1a.03	Determine proactive strategies to mitigate issues that might arise during the statistical enquiry process. [S]		
(b) Types of data			
1b.01	Know and apply terms used to describe different types of data that can be collected for statistical analysis: raw data,[S] quantitative, qualitative, [U] categorical, ordinal, discrete, continuous, ungrouped, grouped, [S] bivariate [U] and multivariate. [B]	S2 interpret and construct ... for ungrouped discrete numerical data [S], ... S3 construct and interpret diagrams for grouped discrete data and continuous data, ... [B]	
1b.02	Know the advantages and implications of merging data into more general categories, and of grouping numerical data into class intervals. [S]	S3 construct and interpret diagrams for grouped discrete data and continuous data ...	
1b.03	Know and apply the terms explanatory (independent) variables and response (dependent) variables. [U]		Including appropriate choice of axes for scatter diagrams
1b.04	Know the difference between primary and secondary data. [S]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	Consideration of advantages/disadvantages and access issues. Implicit to 'knowing the limitations of sampling'.
(c) Population and sampling			
1c.01	Know the difference between population, sample frame and sample.	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
1c.02	Know that 'population' can have different meanings within a stated context.		
1c.03	Know reasons for employing judgement sampling or opportunity (convenience) sampling, and the associated risks of bias when these techniques are used.		
1c.04	1. Know appropriate sampling techniques in the context of the problem to avoid bias. 2. Understand random, systematic, and [S] quota [U] sampling. [S]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
1c.05	Know the key features of a simple random sample [S] and demonstrate understanding of how different techniques, both physical and electronic, are used to select random members from a population: including, but not limited to, dice, cards, random number lists, and calculator functions. [U]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
1c.06	Use stratification and know when this is appropriate before sampling takes place. [U]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
(d) Collecting data			
1d.01	a. Know that data can be collected from different sources: experimental (laboratory, field and natural), simulation, questionnaires, observation, reference, census, population and sampling. b. Know that sources of secondary data should be acknowledged. [S]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	

1d.02	Know the importance of reliability and validity with regard to collected data. [U]	S1 infer properties ... knowing the limitations of sampling [U]	Definitions of reliability and validity are given in guidance column.
1d.03	Determine factors that may lead to bias, including issues of sensitivity of the content matter [U] , level of control [B] and know how to minimise data distortion. [U]		
1d.04	Know the key features to be considered when planning data collection: leading questions, avoiding biased sources, time factors, open/closed questions, different types of interview technique. [S]		
1d.05	Know and demonstrate understanding of techniques used to deal with problems that may arise with collected data. [S]		
1d.06	Know why data may need to be 'cleaned' before further processing, including issues that arise on spreadsheets and apply techniques to clean data in context. [S]		
1d.07	Know the importance of identifying and controlling extraneous variable [U] and the use of control groups. [B]		
2 Processing, representing and analysing data			
(a) Tabulation, diagrams and representation			
2a.01	Represent data sets pictorially using calculated key values as necessary, and interpret and compare data sets displayed pictorially: tabulation, tally, pictogram, [S] pie chart, stem and leaf diagram, Venn diagram . [U]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: • appropriate graphical representation involving discrete, continuous and grouped data, [S] including box plots [B]	
2a.02	Interpret and compare data sets displayed pictorially: [S] population pyramid , choropleth map , [U] comparative pie chart , comparative 2D representations , comparative 3D representations [B]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: • appropriate graphical representation involving discrete, continuous and grouped data, [S] including box plots [B]	
2a.03	Represent data sets graphically using calculated key values as necessary, and interpret and compare data sets displayed graphically: bar charts, line graphs, time series, scatter diagrams, [S] bar line (vertical line) charts, frequency polygons, cumulative frequency (discrete and grouped) charts, histograms (equal class width), and box plots. [U]	S2 interpret and construct tables, charts and diagrams, including frequency tables, bar charts, pie charts and pictograms for categorical data, vertical line charts for ungrouped discrete numerical data [S], tables and line graphs for time series data [U] and know their appropriate use [S] S3 construct and interpret diagrams for grouped discrete data and continuous data, i.e. histograms with equal and unequal class intervals and cumulative frequency graphs, and know their appropriate use [B] S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: • appropriate graphical representation involving discrete, continuous and grouped data, [S] including box plots [B] ...	Note: the 'y' axis of histograms may be labelled 'frequency' where equal class widths are used.
2a.04	Calculate and use frequency density to draw histograms (unequal class width), and interpret and compare data sets displayed in histograms (unequal class width). [B]	S3 construct and interpret diagrams for grouped discrete data and continuous data, i.e. histograms with equal and unequal class intervals ... and know their appropriate use [B] S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: • appropriate graphical representation involving discrete, continuous and grouped data, [S] ...	

2a.05	Justify the appropriate format and produce accurate visualisation of data. [S]	S2 interpret and construct tables, charts and diagrams, including frequency tables, bar charts, pie charts and pictograms for categorical data, vertical line charts for ungrouped discrete numerical data [S], tables and line graphs for time series data [U] and know their appropriate use [S] S3 construct and interpret diagrams for grouped discrete data and continuous data, i.e. histograms with equal and unequal class intervals and cumulative frequency graphs, and know their appropriate use [B]	Awareness of capabilities of statistical software in producing graphs etc. Awareness of target audience.
2a.06	Recognise where errors in construction lead to graphical misrepresentation, including but not limited to incorrect scales, truncated axis, distorted sizing [S] or the misuse of formula when calculating the frequency densities of histograms. [B]		
2a.07	Extract and calculate corresponding values in order to compare data sets that have been presented in different formats and be able to present the same information in multiple formats. [S]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: • appropriate graphical representation involving discrete, continuous and grouped data, [S] including box plots [B] • appropriate measures of central tendency (median, mean, mode and modal class) and spread (range, including consideration of outliers [S], quartiles and inter-quartile range [B])	Includes extracting information from spreadsheets and graphs produced by software.
2a.08	Select [S] and justify [U] appropriate form of representation [S] with regard to the nature of data. [U]	S2 interpret and construct tables, charts and diagrams, including frequency tables, bar charts, pie charts and pictograms for categorical data, vertical line charts for ungrouped discrete numerical data [S], tables and line graphs for time series data [U] and know their appropriate use [S] S3 construct and interpret diagrams for grouped discrete data and continuous data, i.e. histograms with equal and unequal class intervals and cumulative frequency graphs, and know their appropriate use [B]	
2a.09	Determine skewness from data by inspection. [S] and by calculation. Use of: $\text{Skew} = 3(\text{mean} - \text{median})/\text{standard deviation}$ Formula will be given in the formulae sheet. [B]		
2a.10	Interpret a distribution of data in terms of skewness identified from inspection [S] or calculation. [B]		
(b) Measures of central tendency			
2b.01	Calculate averages for discrete and grouped data: mode, median, arithmetic mean [S], weighted mean, geometric mean, mean seasonal variation. [B] The term 'mean' should be understood to be 'arithmetic mean' [S] unless 'geometric mean' is stated. [B]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of central tendency (median, mean, mode and modal class) ... [S]	
2b.02	Justify the rationale for selecting appropriate types of average in context. [U]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of central tendency (median, mean, mode and modal class) ... [S]	
2b.03	Compare different data sets using appropriate calculated or given measure of central tendency: mode, modal class, median and mean. [S]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of central tendency (median, mean, mode and modal class) ... [S]	Includes e.g. selecting the appropriate values from those produced by statistical software.
(c) Measures of dispersion			

2c.01	Calculate different measures of spread: range, quartiles, [S] interquartile range (IQR), percentiles [U] interpercentile range, interdecile range and standard deviation. [B]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of ... spread (range, including ... quartiles and inter-quartile range [B]) ...	For standard deviation only the formulae for a set of values are given. Students will need to know how to apply these to grouped data. (see guidance column)
2c.02	Identify outliers by inspection [S] and using appropriate calculations. [B]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of ... spread (... including consideration of outliers [S]) ...	See guidance column for expected calculations.
2c.03	Comment on outliers with reference to the original data. [U]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of ... spread (... including consideration of outliers [S]) ...	Knowing that outliers may be genuine unusual values or may be the result of errors in recording data.
2c.04	Compare different data sets using appropriate calculated or given measure of spread: range, interquartile range (IQR), percentiles [S] and standard deviation. [B]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of ... spread (range, including ... quartiles and inter-quartile range [B]) ...	Including selecting the appropriate values from those produced by statistical software.
2c.05	Use calculated or given median and interquartile range (IQR) [U] or interpercentile range or interdecile range or mean and standard deviation [B] to compare data samples and to compare sample data with population data. [U]	S4 interpret, analyse and compare the distributions of data sets from univariate empirical distributions through: ... appropriate measures of ... spread (range, including ... quartiles and inter-quartile range [B]) ... S5 apply statistics to describe a population [S]	
2c.06	Use calculated or given means and standard deviation to standardise and interpret data collected in two comparable samples. Formulae for standard deviation will be given in the formulae sheet. [B]		Standardised score formula must be known
(d) Further summary statistics			
2d.01	Use different types of index [U] and weighted index [B] numbers in context, including but not limited to retail price index (RPI), consumer price index (CPI) and gross domestic product (GDP). [U]		R9 in 1MA1 includes working with percentage change.
2d.02	Interpret data related to rates of change over time (including, but not limited to, percentage change, births, deaths, house prices, and unemployment) when given in graphical form. [S] Calculate and interpret rates of change over time from tables using context specific formula. [U]		R9 in 1MA1 includes working with percentage change.
(e) Scatter diagrams and correlation			
2e.01	Know and apply vocabulary of correlation: positive, negative, zero, causation, association, interpolation and extrapolation. [S]	S6 use and interpret scatter graphs of bivariate data; recognise correlation [S] and know that it does not indicate causation; draw estimated lines of best fit; make predictions; interpolate and extrapolate apparent trends while knowing the dangers of so doing [U]	Including appropriate choice of axes for scatter diagrams
2e.02	Describe and make comparisons of correlation by inspection: strong or weak. [S]	S6 use and interpret scatter graphs of bivariate data; recognise correlation [S] ...	
2e.03	Know that correlation does not necessarily imply causation [U] and multiple factors may interact. [B]	S6 use and interpret scatter graphs of bivariate data; recognise correlation [S] and know that it does not indicate causation ... [U]	
2e.04	Determine line of best fit by eye, [S] by drawing through a calculated double mean point (\bar{x}, \bar{y}) [U] and by using the equation of the regression line. [B]	S6 use and interpret scatter graphs of bivariate data; ... [S] ... draw estimated lines of best fit ... [U]	
2e.05	Apply formula to determine Spearman's rank correlation coefficient. Values found using calculator functions will be permissible. [B]		Formula is given
2e.06	Interpret [U] calculated or [B] given Spearman's rank correlation coefficient in the context of the problem. [U]		
2e.07	Interpret given Pearson's product moment correlation coefficient (PMCC) in the context of the problem. [B]		The calculation of PMCC will not be required.

2e.08	Understand the distinction between Spearman's rank correlation coefficient and Pearson's product moment correlation coefficient (PMCC). [B]		How these relate to a scatter graph.
(f) Time series			
2f.01	Identify trends in data through inspection [S] and by calculation of 4 [U] or other determined appropriate [B] point moving averages. [U]	S2 interpret and construct tables, charts and diagrams, including ... [S], tables and line graphs for time series data [U] and know their appropriate use [S]	Including interpretation of the gradient of trend lines
2f.02	Interpret seasonal and cyclic trends in context. [U] Use such trends to make predictions. [B]		
(g) Quality assurance			
2g.01	Know that a set of sample means are more closely distributed than individual values from the same population. [B]		
2g.02	Use action and warning lines in quality assurance sampling applications. [B]		
(h) Estimation			
2h.01	Use calculated or given summary statistical data to make estimates of population characteristics. Use samples to estimate population mean. Use sample data to predict population proportions. [U]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U] S5 apply statistics to describe a population [S]	
2h.02	Apply Petersen capture recapture formula to calculate an estimate of the size of a population. [B]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
2h.03	Know that sample size has an impact on reliability and replication. [S]	S1 infer properties of populations or distributions from a sample, while knowing the limitations of sampling [U]	
3 Probability			
3p.01	Use collected data to calculate estimates of probabilities. [S]	P1 record, describe and analyse the frequency of outcomes of probability experiments using tables and frequency trees [S] P2 apply ideas of randomness, fairness and equally likely events to calculate expected outcomes of multiple future experiments [S] P3 relate relative expected frequencies to theoretical probability, using appropriate language and the 0-1 probability scale [S]	Use of relative frequency.
3p.02	Compare the probability of different possible outcomes using the 0-1 or 0-100% scale and statements of likelihood. [S]	P3 relate relative expected frequencies to theoretical probability, using appropriate language and the 0-1 probability scale [S]	
3p.03	Use probability values to calculate expected frequency of a specified characteristic within a sample or population. [S]	P2 apply ideas of randomness, fairness and equally likely events to calculate expected outcomes of multiple future experiments [S]	
3p.04	Use collected data and calculated probabilities to determine and interpret relative risks and absolute risks, and express in terms of expected frequencies in groups. [U]		
3p.05	Compare experimental data with theoretical predictions to identify possible bias within the experimental design. [S]	P5 understand that empirical unbiased samples tend towards theoretical probability distributions, with increasing sample size [U]	
3p.06	Recognise that experimental probability will tend towards theoretical probability as the number of trials increases when all variables are random. [S]	P5 understand that empirical unbiased samples tend towards theoretical probability distributions, with increasing sample size [U]	

3p.07	Use two-way tables, sample space diagrams, tree diagrams and Venn diagrams to represent all the different outcomes possible for at most three events. [U]	<p>P4 apply the property that the probabilities of an exhaustive set of outcomes sum to one; apply the property that the probabilities of an exhaustive set of mutually exclusive events sum to one [S]</p> <p>P6 enumerate sets and combinations of sets systematically, using tables, grids, Venn diagrams [S] and tree diagrams [U]</p> <p>P7 construct theoretical possibility spaces for single and combined experiments with equally likely outcomes and use these to calculate theoretical probabilities [S]</p> <p>P8 calculate the probability of independent and dependent combined events, including using tree diagrams and other representations, and know the underlying assumptions [U]</p> <p>P9 calculate and interpret conditional probabilities through representation using expected frequencies with two-way tables, tree diagrams and Venn diagrams [B]</p>	<p>1ST0 guidance column includes reference to exhaustive events.</p> <p>1ST0 requires general addition law.</p>
3p.08	Know and apply the formal notation for independent events. [U]	P8 calculate the probability of independent and dependent combined events, including using tree diagrams and other representations, and know the underlying assumptions [U]	
3p.09	Know and apply the formal notation for conditional probability. [U]	P9 calculate and interpret conditional probabilities through representation using expected frequencies with two-way tables, tree diagrams and Venn diagrams [B]	
3p.10	Comment on the differences between experimental and theoretical values in terms of possible bias. Formal tests of significance will not be required. [B]		
3p.11	Know and interpret the characteristics of a binomial distribution. [B]		
3p.12	Know and interpret the characteristics of a normal distribution. [B]		
3p.13	Know that, for a normal distribution, values more than three standard deviations from the mean are very unusual; know that approximately 95% of the data lie within two standard deviations of the mean and that 68% (just over two thirds) lie within one standard deviation of the mean. [B]		