THE ACADEMY OF ST NICHOLAS	Y12 Applied	CH04 2.1, 2.2, 2.3, 2.4	Statistics – Correlation	Lessons 2

Essential Knowledge Milestones

- Draw and interpret scatter diagrams for bivariate data
- Interpret correlation and understand that it does imply causation
- Interpret the coefficients of a regression line equation for bivariate data
- Understand when you can use a regression line to make predictions about the data

Success Criteria

- ❑ You can draw a scatter graph and describe the correlation between the two variables in context
- ❑ You can use 'common sense' to decide whether the correlation between two variables is 'causal'
- ❑ Given the equation of a regression line in the form y = a + bx you can describe what the values of a and b mean in context
- ❑ You can describe the situations when you can and cannot use your regression line to make predictions about the data using the key words explanatory variable, response variable, interpolation and extrapolation.

Assumed Prior Knowledge/ Links / Interleaving

GCSE (9-1) in Mathematics at Higher Tier

- S6 Use and interpret scatter graphs of bivariate data; recognise correlation 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

Teaching Points

- I don't think that the text books questions are particularly in depth so once again I have prepared a couple of packs that allow students to write in them. There are some slightly odd questions in there to try to drive home some key points.
- For bivariate data students should understand the terms explanatory and response variables and know where each is placed on the axes of a scatter diagram. This is particularly important as variables other than y and x could be used.
- The idea of causal relationships is really important and should be explored through class discussion. Can your students come up with variables that correlate but are not causal?
- As always, all interpretations must be given in context of the questions
- Students are not expected to know, calculate or understand the regression line formula. Students will need to understand the use of interpolation when using a regression line equation to make predictions within the range of values of the explanatory variable and they need to understand the dangers of extrapolation (predictions outside the range), again variables other than y and x could be used.
- Discussion of 'least squares' can help students to understand the principle of the regression. If students have done the standard deviation investigation sheet then they will understand the idea of 'squares' in this context.
- The interpretation of the coefficients a and b in $y= a+ bx$ is really important and will need lots of practice.
- Students will be expected to describe the correlation on a scatter diagram in terms of positive, negative or no correlation and strong or weak but no calculations need to be made. Values from calculations will not be given for interpretation.

Potential Barriers to Access /Misconceptions

- When drawing conclusions students should be careful to given full answers in context paying particular attention to units. E.g. $ (1000s) on the y axis.
- Students will interchange between independent/explanatory variable etc and probably incorrectly. This language needs to be modelled and encouraged from the outset.
- When interpreting coefficients in a regression line students will often interpret the gradient incorrectly. E.g. in g = 7.23 + 1.82w they will say 'as g goes up by 1 w goes up by 1.82', this is due to the idea of 1.82 being the coefficient of w. Try using the idea of substituting numbers into the equation to support them.

Opportunities for Reasoning/Problem Solving/Proofs

- There is opportunity for further use of the large data set here. Is there a correlation between daily hour's sunshine across 2 locations or years? What about rainfall in northern and southern hemisphere locations?
- Students will need strong reasoning skills when discussing 'causation' or the idea of the dangers of extrapolation.

Key Mathematical Vocabulary	scatter diagram, linear regression, explanatory (independent) variables, response (dependent) variables interpolation, extrapolation, product moment correlation coefficient (PMCC)

Personal Development	Notes	Resources
Pupils are taught that they must 'respect' each other's opinions and well-being when working collectively in class. Pupils to learn that mathematicians have 'ambition' to push boundaries when aiming to solve new problems **Resilience** – never giving up! Building confidence across the problem solving aspects of the course. **Ambition** – living life to the full – fulfilling dreams and aspirations – linking to future career and ambition plans. **Respect** – respect for others – the 9 protected characteristics **Personal Best** – First Work – Best Work every time		• Problem pack