# Marking key Statistical Investigation 1 – Unit 3

|  |  |
| --- | --- |
| **Introduction** | |
| I have been asked to decide which, if any, is the better predictor of a person’s height; his/her arm span or the length of his/her right foot. In order to make an informed decision, I will compare the strength of the association between the pairs of variables; that is, the association between height and arm span and the association between height and the length of the right foot. | |
| **Specific behaviours** | **Marks** |
| Provides a simple introduction of the question | 1 |
| Restates question in his/her own words | 1 |
| Mentions comparison of association between variables | 1 |
| **Total** | **/3** |

|  |  |
| --- | --- |
| **Numerical analysis** | |
| Of the data, there are four outliers  Height Arm span Right foot  170 70 23  55 55 23  147 27 24  158 80 25  These have been removed for the graphical analysis | |
| **Specific behaviours** | **Marks** |
| Identifies one outlier | 1 |
| Identifies at least two outliers | 1 |
| Identifies all outliers | 1 |
| Removes outliers | 1 |
| **Total** | **/4** |

| **Graphical analysis** | |
| --- | --- |
|  | |
|  | |
|  | |
| **Specific behaviours** | **Marks** |
| Constructs a single scatterplot | 1 |
| Shows comparative scatterplots | 1 |
| Uses a similar scale for the horizontal axis (height) | 1 |
| Presents correct graphs, including labelling | 2 |
| **Total** | **/5** |

|  |  |
| --- | --- |
| **Interpretation** | |
| Discussion of sample interpretation:  Outliers have been removed from four data points as these were probably due to incorrect measurement and would therefore have not given a true picture of any association between the variables.  The scatterplot for height and arm span is positive and linear. The association appears to be strong.  The scatterplot for height and right foot length is not linear. The association appears to be weak. | |
| **Specific behaviours** | **Marks** |
| Interpretation linked to numerical and graphical data | 1 |
| Makes an appropriate statement about the outliers | 1 |
| Gives a reason for removing the outliers | 1 |
| Compares the shape of each scatterplot | 2 |
| States the linear relationship between height and arm span | 1 |
| States the non-linear relationship between height and length of right foot | 1 |
| **Total** | **/7** |

|  |  |
| --- | --- |
| **Conclusion** | |
| Short statement outlining summary of findings  Sample conclusion:  The removal of four outliers from the data has reduced the sample to 26 students.  Analysis of data from these 26 students has shown that the linear relationship between height and arm span is a better predictor of height as there is a strong association between the two variables. | |
| **Specific behaviours** | **Marks** |
| Makes a valid statement about the results | 1 |
| Relates conclusion back to the original question | 1 |
| Refers to strong relationship/association between height and arm span | 1 |
| Provides a concise and coherent summary of the analysis | 1 |
| **Total** | **/4** |