# P01 Instructor & Reference Guide – Descriptive Statistics and Visualisation

## Theory Brief

Descriptive statistics provide a summary of data through measures such as central tendency (mean, median, mode) and dispersion (variance, standard deviation). A histogram visualises the distribution of a single numeric variable, revealing patterns like skewness or modality.

## Worked Example

Below is a snapshot of the first few rows of the synthetic dataset and summary statistics:

Student\_ID Mathematics Physics Chemistry Computer Age  
 1 55.1 61.2 47.9 80.2 18  
 2 61.3 53.3 60.9 54.1 19  
 3 77.9 47.6 62.8 70.2 20  
 4 66.9 65.7 66.7 65.5 21  
 5 74.2 49.8 62.7 81.0 18

### Basic Statistics

* Student\_ID\_mean: 75.50
* Student\_ID\_median: 75.50
* Student\_ID\_mode: 1.00
* Student\_ID\_var: 1887.50
* Student\_ID\_std: 43.45
* Mathematics\_mean: 66.26
* Mathematics\_median: 65.45
* Mathematics\_mode: 58.60
* Mathematics\_var: 85.24
* Mathematics\_std: 9.23
* Physics\_mean: 59.44
* Physics\_median: 61.00
* Physics\_mode: 61.80
* Physics\_var: 146.50
* Physics\_std: 12.10
* Chemistry\_mean: 61.53
* Chemistry\_median: 62.10
* Chemistry\_mode: 58.60
* Chemistry\_var: 130.96
* Chemistry\_std: 11.44

### Correlation Matrix

Student\_ID Mathematics Physics Chemistry  
Student\_ID 1.000000 0.056925 -0.071927 -0.079921  
Mathematics 0.056925 1.000000 -0.112505 0.053850  
Physics -0.071927 -0.112505 1.000000 -0.034815  
Chemistry -0.079921 0.053850 -0.034815 1.000000

### Visualisations

An example plot is saved in the results folder as P01\_plot1.png.

## Evaluation Rubric

* Correctness of statistical calculations (50%)- Quality of histogram and visual interpretation (30%)- Clarity of written analysis and report (20%)

## Common Pitfalls

Forgetting to handle missing values, misinterpreting the histogram scale, or confusing variance with standard deviation.