Bivariate Datasets

**Formula Sheet**

|  |  |  |
| --- | --- | --- |
| **Name** | **Definition** | **Formula** |
| Joint Relative Frequency | The ratio of frequency in a particular category to the total number of data points. (Two-way frequency table) | $$\frac{joint frequency}{total surveyed}$$ |
| Marginal Relative Frequency | The ratio of a sum of joint relative frequencies to the total number of data points. (Two-way frequency table) | $$\frac{category total}{total surveyed}$$ |
| Conditional Relative Frequency | Finding the ratio of frequency in a particular category to the total number of data points in a particular category. (Two-way frequency table) | $$\frac{joint frequency}{category total}$$ |
| Least Squares Regression | A method of fitting a curve to a set of points representing statistical data in such a way that the sum of the squares of the distances of the points from the curve is a minimum. | Use GeoGebra to find the precise equation for the line of best fit.  |
| Residual | The vertical distance between a data point and the line of best fit; the result of subtracting the predicted *y-*value from the actual *y-*value. | Residual = $$y−\hat{y}$$$$y=actual value $$$$\hat{y}=predicted value$$ |
| Correlation Coefficient | A number (usually represented with the variable *r*) in the range of *−1 ≤ r ≤ 1*that indicates how accurately an equation represents the data being examined. | Use GeoGebra to find the precise correlation coefficient.  |