1 Definitions

$$variance = s_y^2 = \frac{\sum(Y_i - \overline{Y})^2}{N-1}$$

$$standard\ deviation = s_y = \sqrt{variance}$$

$$covariance\ of\ X\ and\ Y = s_{yx} = \frac{\sum(Y_i - \overline{Y})(X_i - \overline{X})}{N-1}$$

2 Deriving the Formula of the Slope

Estimatation of the slope (b_X) is straightforward. It's simply the covariance of X and Y divided by the variance of X (Knoke, et. al: 174–175). The formulas for both covariance and variance are both divided by $N-1$, which factors out in Eq 4. Equation 4, then, equals the formula for the slope provided by Charles.

$$\begin{aligned}
Slope\ of\ X = b_X &= \frac{Covariance\ of\ X\ and\ Y}{Variance\ of\ X} && (1)\\
&= \frac{s_{yx}}{s_x^2} && (2)\\
&= \frac{\frac{\sum(Y_i-\overline{Y})(X_i-\overline{X})}{N-1}}{\frac{\sum(X_i-\overline{X})^2}{N-1}} && (3)\\
&= \frac{\sum(Y_i - \overline{Y})(X_i - \overline{X})}{\sum(X_i - \overline{X})^2} && (4)
\end{aligned}$$

3 Deriving the Formula of r

Although a bit more complicated, deriving the formula for r is also relatively straighforward. The formula for computing r is the covariance of X and Y divided by the product of the standard deviations of X and Y. In Eq 7, I substitute the square root of the variances for s_y and s_x. In Eq 8, I reach back into my high school algebra, remembering[1] that you can multiply terms with equal exponents (i.e., $(x^a)(y^a) = (xy)^a$). That permits me to place the entire denominator under the root. Then, in Eq 9, I'm able to split the root over both the numerator and denominator portions of the lower half of the equation because $(\frac{x}{y})^a = \frac{x^a}{y^a}$. In Eq 9, then, we have $\sqrt{(N-1)\times(N-1)}$ which reduces to $(N-1)$ in Eq 10. In Eq 11, the $N-1$'s cancel out which equals the formula for

[1] This is a lie. I had to look it up. The GRE has a pretty decent math review available at http://www.gre.org/pracmats.html#gentest

r provided by Charles. (Except that Charles forgot to include the summation signs in the denominator.)

$$
\begin{aligned}
correlation\ coef = r_{yx} &= \frac{Covariance\ of\ X\ and\ Y}{(Std\ Dev\ of\ y)(Std\ Dev\ of\ x)} && (5) \\
&= \frac{s_{yx}}{(s_y)(s_x)} && (6) \\
&= \frac{\frac{\sum (Y_i-\overline{Y})(X_i-\overline{X})}{N-1}}{\sqrt{\frac{\sum (Y_i-\overline{Y})^2}{N-1}}\sqrt{\frac{\sum (X_i-\overline{X})^2}{N-1}}} && (7) \\
&= \frac{\frac{\sum (Y_i-\overline{Y})(X_i-\overline{X})}{N-1}}{\sqrt{\frac{\sum (Y_i-\overline{Y})^2}{N-1} \times \frac{\sum (X_i-\overline{X})^2}{N-1}}} && (8) \\
&= \frac{\frac{\sum (Y_i-\overline{Y})(X_i-\overline{X})}{N-1}}{\frac{\sqrt{\sum (Y_i-\overline{Y})^2 \times \sum (X_i-\overline{X})^2}}{\sqrt{(N-1)\times(N-1)}}} && (9) \\
&= \frac{\frac{\sum (Y_i-\overline{Y})(X_i-\overline{X})}{N-1}}{\frac{\sqrt{\sum (Y_i-\overline{Y})^2 \times \sum (X_i-\overline{X})^2}}{N-1}} && (10) \\
&= \frac{\sum (Y_i-\overline{Y})(X_i-\overline{X})}{\sqrt{\sum (Y_i-\overline{Y})^2 \times \sum (X_i-\overline{X})^2}} && (11)
\end{aligned}
$$