

Correlation

	Fórmula	Legenda	
Pearson's r	$r = \frac{cov(x,y)}{s_x \cdot s_y} \quad -1 \leq r \leq 1$	cov(x,y): covariância de x e y S _x : desvio padrão de x	Correlation coefficient (r)
Coefficient of determination	r ²		r ² é a % de variação de y que é explicada pela variação de x
Correlation of Population	ρ		Verdadeira correlação da população que se estudo através da amostra H ₀ : ρ = 0
Hypothesis Testing	$t = \frac{r\sqrt{N-2}}{\sqrt{1-r^2}} \quad df = N - 2$		

Regression

	Fórmula	Legenda	
			Melhor função de regressão é a que minimiza a soma dos quadrados dos resíduos
Regression Line	$\hat{y} = ax + b$	ŷ: valor esperado y: valor observado a: slope	Linha de Regressão sempre passará pelo ponto (X _{média} , y _{média})
	$\hat{y} = \beta_1 x + \beta_0$		Linha de Regressão para a população
Resíduo	(y - ŷ)		Erro entre o valor observado(real) e o esperado
Slope	$a = r \frac{s_y}{s_x}$	r: Pearson's r s _y : desvio padrão de y s _x : desvio padrão de x	Indica quanto y varia, variando x em 1 unidade
Y-Intercept	$b = \bar{y} - r \frac{s_y}{s_x} \bar{x}$	x: valor médio de x y: valor médio de y	
Standard Error	$SE = \sqrt{\frac{\sum (y - \hat{y})^2}{N - 2}}$	N: número de pontos	Standard Error of estimate
Hypothesis Testing			Realiza-se o Hypothesis Test para o slope para confirmar se existe de fato correlação do atributo x com o atributo y (H ₀ : β ₁ = 0)

Links

	Link
T-table	https://s3.amazonaws.com/udacity-hosted-downloads/t-table.jpg
GraphPad	https://www.graphpad.com/quickcalcs/