

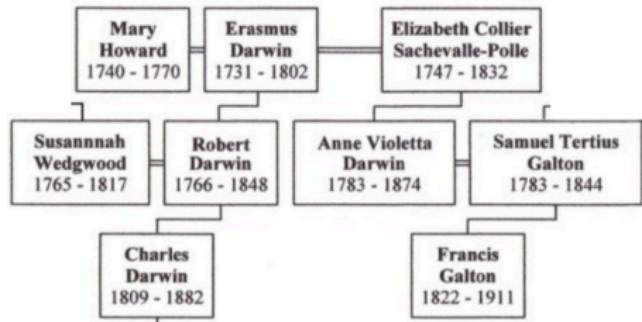
il modello lineare

Massimo Borelli

22 luglio 2024



.. una famiglia mica male!



246

Anthropological Miscellanea.

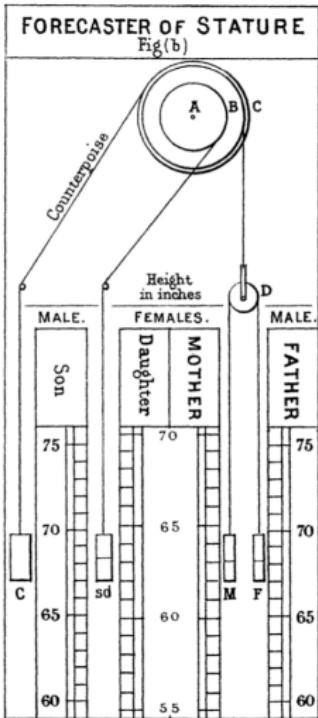
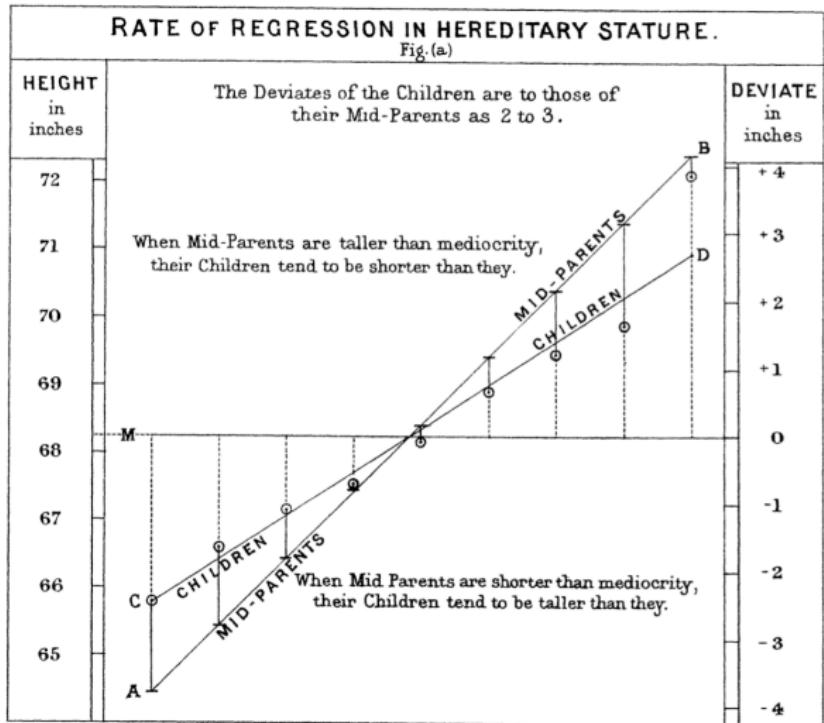
ANTHROPOLOGICAL MISCELLANEA.

REGRESSION *towards MEDIOCRITY in HEREDITARY STATURE.*

By FRANCIS GALTON, F.R.S., &c.

[WITH PLATES IX AND X.]

THIS memoir contains the data upon which the remarks on the Law of Regression were founded, that I made in my Presidential Address to Section H, at Aberdeen. That address, which will appear in due course in the Journal of the British Association, has already been published in "Nature," September 24th. I reproduce here



due strumenti basilari

- la **correlazione**
- la retta di **regressione**

la correlazione peso vs. altezza

Regression

Correlation

Variables

statura
peso

Partial out

Sample Correlation Coefficient

Pearson's r

Spearman's rho

Kendall's tau-b

Additional Options

Display pairwise

Report significance

Flag significant correlations

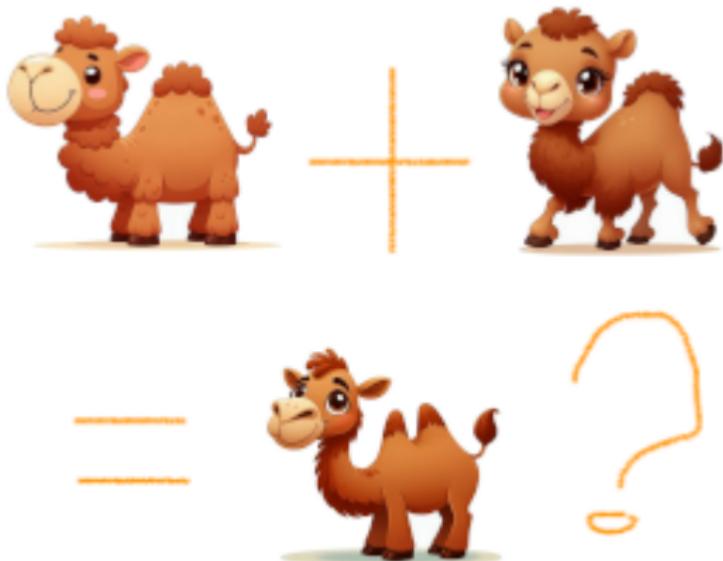
la correlazione peso vs. altezza

Tabella: Pearson's Correlations

		Pearson's r	p
statura	-	peso	0.744*** < .001

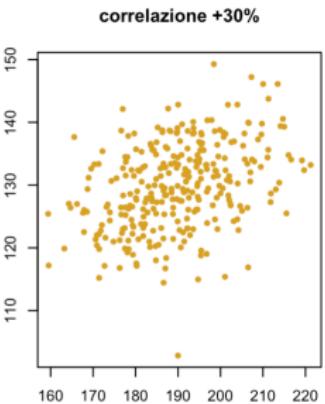
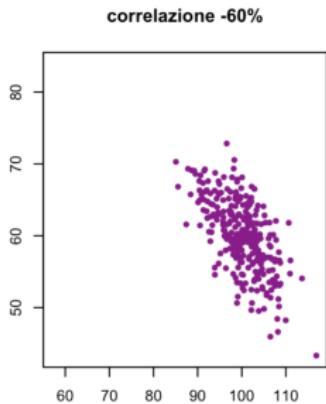
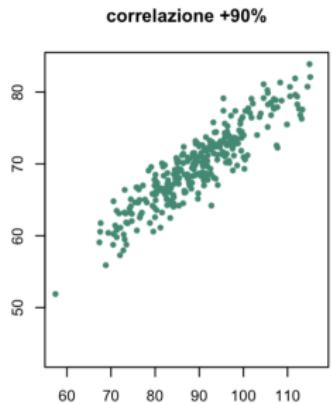
* p < .05, ** p < .01, *** p < .001

vi ricordate?



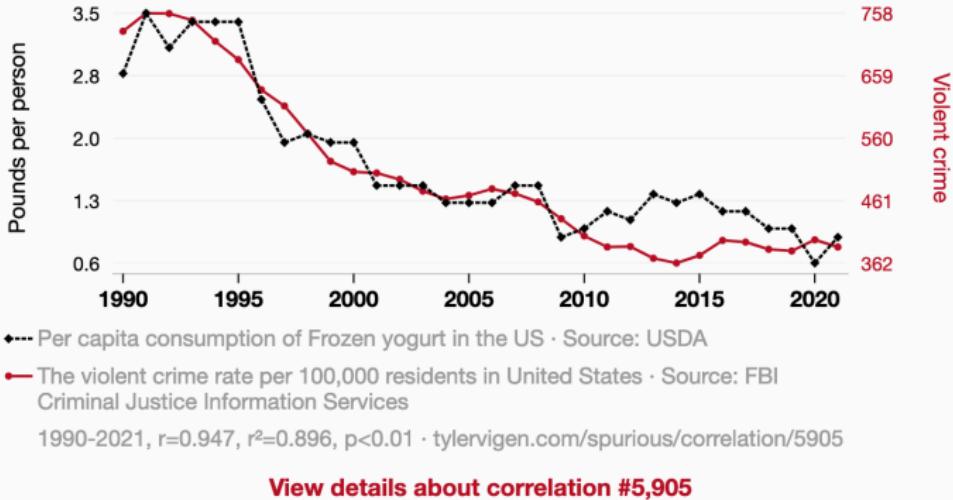
+90%, -60%, +30%

il coefficiente di correlazione ρ di Pearson



correlation is not causation

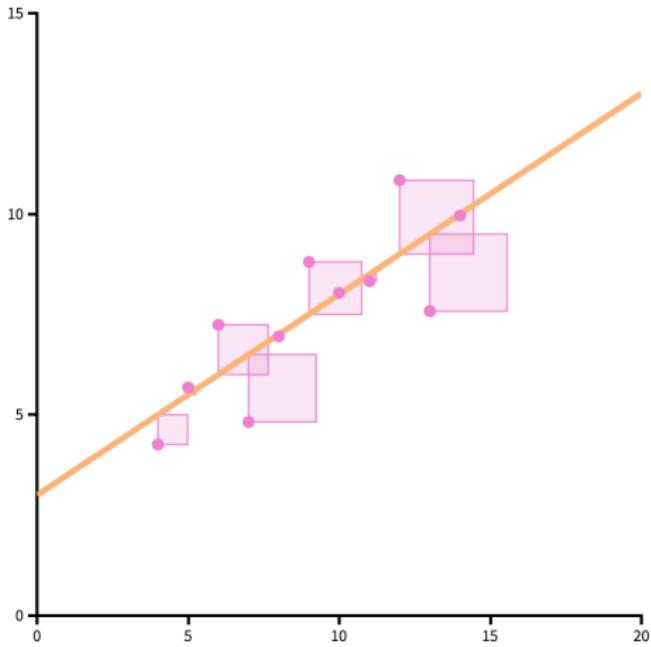
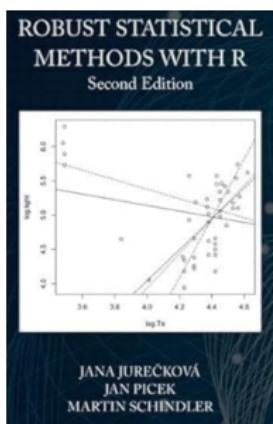
Frozen yogurt consumption correlates with Violent crime rates



<https://www.tylervigen.com/spurious-correlations>

il principio dei minimi quadrati

- Galton 1886 ?
- Legendre 1805 ?
- Gauss 1794 ?
- Boskovich 1757.



<https://seeing-theory.brown.edu/>

la retta di regressione

The screenshot shows a software interface for statistical analysis. At the top right, there is a icon of a scatter plot with a regression line and the word "Regression". Below it is a toolbar with several icons: a blue circle with "R", a black circle with a pencil, a green circle with a plus sign, a blue circle with an "i", and a red circle with a minus sign. On the left, a section titled "Linear Regression" is expanded, indicated by a green arrow pointing from the title above. This section contains a list of variables: "anno", "genero", "scarpe", "fumo", "sport", "frequenza", "massima", and "minima". To the right of this list are two buttons: a grey one with a downward arrow and a white one with a rightward arrow. Further down, there are settings for the regression: "Dependent Variable" set to "peso" (with a blue selection bar), "Method" set to "Enter" (with a dropdown arrow), and "Covariates" set to "statura".

Linear Regression

effetti casuali

Model Summary – peso

Model	R	R ²	Adjusted R ²	RMSE
H ₁	0.744	0.554	0.547	6.459

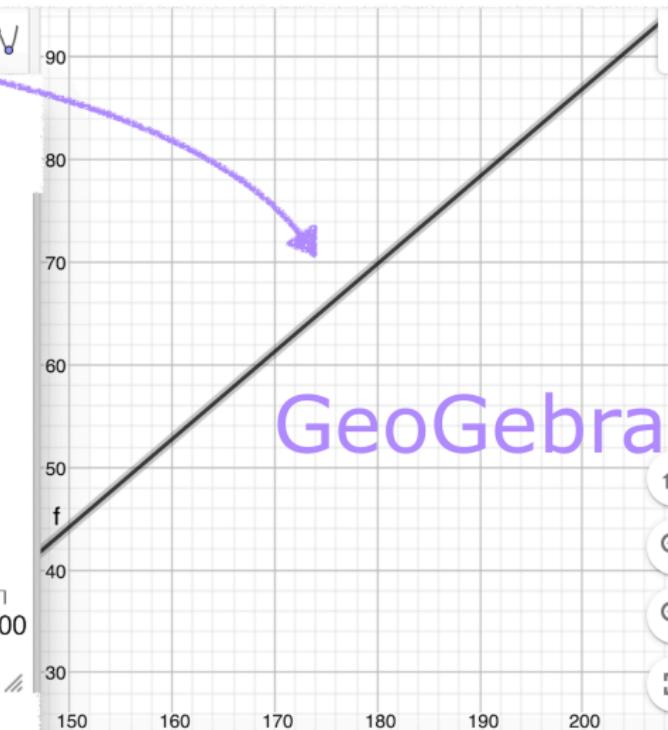
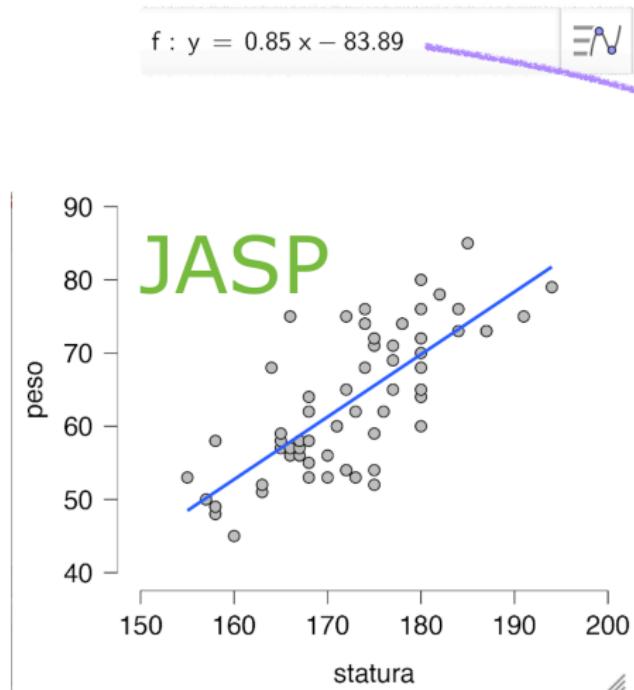


Coefficients

effetti fissi

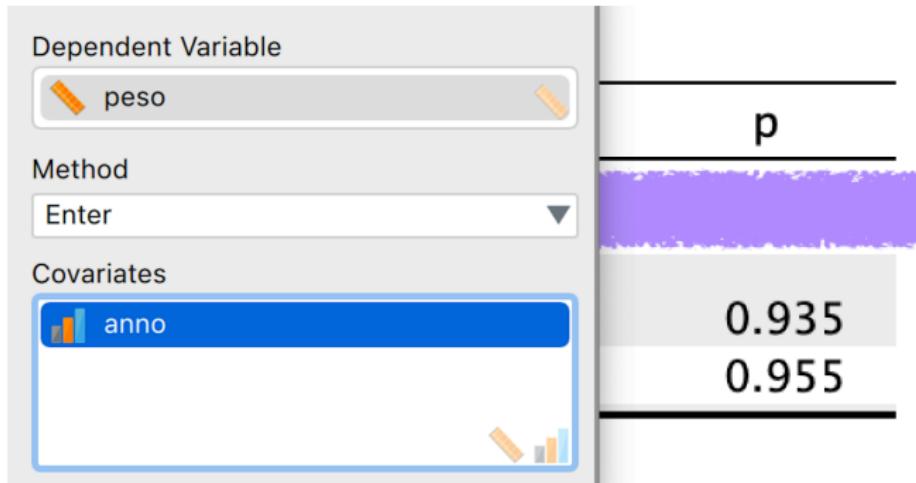
Model		Unstandardized	Standard Error	t	p
H ₁	(Intercept)	-83.891	16.677	-5.030	< .001
	statura	0.854	0.096	8.850	< .001

gli effetti fissi



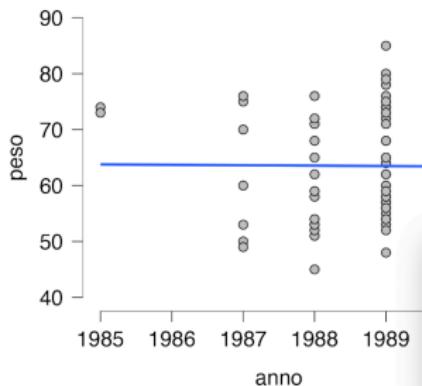
controprova, $p = \text{NS}$

Dimmi in che anno sei nata/o e ti dirò quanto pesi ..



spieghiamo il modello nullo H0

Dimmi in che anno sei nata/o e ti dirò quanto pesi ..



Descriptive Statistics

peso	
Mean	63.523
Std. Deviation	9.598

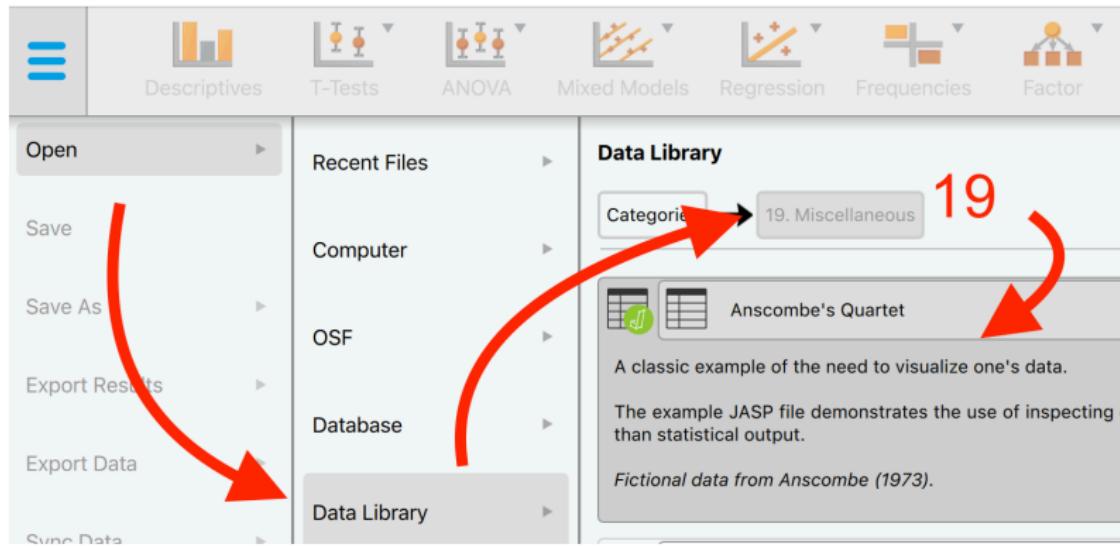
Model Summary – peso

Model	R	R ²	Adjusted R ²	RMSE
H ₀	0.000	0.000	0.000	9.598

il quartetto di Anscombe

https://it.wikipedia.org/wiki/Quartetto_di_Anscombe

diagnostica della regressione



diagnostica della regressione

▼ Plots

Residuals Plots

- Residuals vs. dependent
- Residuals vs. covariates
- Residuals vs. predicted
- Residuals histogram

Standardized residuals

Q-Q plot standardized residuals

United States
Environmental Protection
Agency

Office of Environmental
Information
Washington, DC 20460

EPA/240/B-06/003
February 2006

 **Data Quality Assessment:
Statistical Methods for
Practitioners**
EPA QA/G-9S