

Bibliography

Abramowitz, M. and I. Stegun (1965). *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*. Applied Mathematics Series. Dover Publications.

Anaconda (2016, November). Anaconda Software Distribution. Computer Software. V. 2-2.4.0. <https://www.anaconda.com>.

Anscombe, F. J. (1973). Graphs in statistical analysis. *The American Statistician* 27(1), 17–21.

Arfken, G., H. Weber, and F. Harris (2011). *Mathematical Methods for Physicists: A Comprehensive Guide*. Elsevier Science.

Bayes, T. (1763). An essay towards solving a problem in the doctrine of chances. *Philosophical Transactions* 53, 370–418.

Berkowitz, B. D. (2018). *Playfair: The True Story of the British Secret Agent Who Changed How We See the World*. George Mason University Press.

Bernoulli, J., J. Bernoulli, and E. D. Sylla (2006). *The Art of*

Conjecturing, Together with Letter to a Friend on Sets in Court Tennis. Johns Hopkins University Press.

Bertin, J. and M. Barbut (1967). *Sémiologie graphique: Les diagrammes, les réseaux, les cartes.* Gauthier-Villars.

Bokeh Development Team (2018). Bokeh: Python library for interactive visualization. <https://bokeh.pydata.org/en/latest/>.

Bostridge, M. (2015). *Florence Nightingale: The Woman and Her Legend.* Penguin Books Limited.

Bowley, A. L. (1928). The Standard Deviation of the Correlation Coefficient. *Journal of the American Statistical Association* 23(161), 31–34.

Bradley, M. *Charles Dupin (1784 - 1873) and His Influence on France.* Cambria Press.

Brown, M. B. and A. B. Forsythe (1974). Robust tests for the equality of variances. *Journal of the American Statistical Association* 69(346), 364–367.

Buckmaster, D. (1974). The Incan Quipu and the Jacobsen hypothesis. *Journal of Accounting Research* 12(1), 178–181.

Busby, M. (2020). Cambridge college to remove window commemorating eugenicist. www.theguardian.com/education/2020/jun/27/cambridge-gonville-caius-college-eugenicist-window-ronald-fisher. Accessed: 2021-02-14.

D'Agostino, R. and E. S. Pearson (1973). Tests for Departure from Normality. Empirical Results for the Distributions of b_2 and $\sqrt{b_1}$. *Biometrika* 60(3), 613–622.

- D'Agostino, R. B. (1971, 08). An omnibus test of normality for moderate and large size samples. *Biometrika* 58(2), 341–348.
- D'Agostino, R. B. (1970, 12). Transformation to normality of the null distribution of g_1 . *Biometrika* 57(3), 679–681.
- de Heinzelin, J. (1962, Jun). Ishango. *Scientific American* (206:6), 105–116.
- Díaz Díaz, R. (2006). Apuntes sobre la aritmética maya. *Educere* 10(35), 621–627.
- Fisher, R. (1963). *Statistical Methods for Research Workers*. Biological monographs and manuals. Hafner Publishing Company.
- Franks, B. (2020). *97 Things About Ethics Everyone in Data Science Should Know*. O'Reilly Media.
- Gartner (2017). Gartner Says within Five Years, Organizations Will Be Valued on Their Information Portfolios. www.gartner.com/en/newsroom/press-releases/2017-02-08-gartner-says-within-five-years-organizations-will-be-valued-on-their-information-portfolios. Accessed: 2021-01-04.
- Gauss, C. F. (1823). *Theoria combinationis observationum erroribus minimis obnoxiae*. Number V. 2 in *Commentationes Societatis Regiae Scientiarum Gottingensis recentiores: Classis Mathematicae*. H. Dieterich.
- Good, I. J. (1986). Some statistical applications of Poisson's work. *Statistical Science* 1, 157–170.
- Hand, D. J. (2020). *Dark Data: Why What You Don't Know Matters*. Princeton University Press.

Hassig, R. (2013). El tributo en la economía prehispánica.

Arqueología Mexicana 21(124), 32–39.

Heath, T. L. (2017). *Euclid's Elements (The Thirteen Books)*.

Digireads.com Publishing.

Heffelfinger, T. and G. Flom (2004). Abacus: Mystery of the

Bead. <http://totton.idirect.com>. Accessed: 2021-02-03.

Henderson, H. V. and P. F. Velleman (1981). Building multiple

regression models interactively. *Biometrics* 37(2), 391–411.

Howard, L. (1960). Robust tests for equality of variances.

In I. Olkin and H. Hotelling (Eds.), *Contributions to Probability and Statistics: Essays in Honor of Harold Hotelling*, pp. 278–292. Stanford University Press.

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment.

Computing in Science & Engineering 9(3), 90–95.

Ioannidis, Y. (2003). The history of histograms (abridged).

www.vldb.org/conf/2003/papers/So2Po1.pdf. Accessed: 2021-02-14.

Jones, E., T. Oliphant, P. Peterson, et al. (2001–). SciPy: Open

source scientific tools for Python. <http://www.scipy.org/>.

Kolmogorov, A. (1933). Sulla determinazione empirica di una

legge di distribuzione. *Inst. Ital. Attuari, Giorn.* 4, 83–91.

Laplace, P. S. and A. Dale (2012). *Pierre-Simon Laplace*

Philosophical Essay on Probabilities: Translated from the Fifth French Edition of 1825 with Notes by the Translator. Sources in the History of Mathematics and Physical Sciences. Springer New York.

Lichtheim, M. (2019). *Ancient Egyptian Literature*. University of California Press.

Loupart, F. (2017). Data is giving rise to a new economy. *The Economist* - www.economist.com/briefing/2017/05/06/data-is-giving-rise-to-a-new-economy. Accessed: 2021-01-02.

Mackinlay, J. (1986, Apr). Automating the Design of Graphical Presentations of Relational Information. *ACM Trans. Graph.* 5(2), 110–141.

Mann, H. B. and D. R. Whitney (1947, Mar). On a test of whether one of two random variables is stochastically larger than the other. *Ann. Math. Statist.* 18(1), 50–60.

McCann, L. I. (2015). Introducing students to single photon detection with a reverse-biased LED in avalanche mode. In E. B. M. Eblen-Zayas and J. Kozminski (Eds.), *BFY Proceedings*. American Association of Physics Teachers.

McKinney, W. (2012). *Python for Data Analysis: Data Wrangling with Pandas, NumPy, and IPython*. O'Reilly Media.

McKinsey & Co. (2014). Using customer analytics to boost corporate performance.

Melissinos, A. C. and J. Napolitano (2003). *Experiments in Modern Physics*. Gulf Professional Publishing.

P. Hlobil (2021). Pandas Bokeh. <https://github.com/PatrikHlobil/Pandas-Bokeh>.

PA Media (2020). UCL renames three facilities that honoured prominent eugenicists. www.theguardian.com/education/

2020/jun/19/ucl-renames-three-facilities -that-honoured-prominent-eugenicists. Accessed: 2021-02-14.

Pearson, E. S. (1931, 05). I. Note on tests for normality.

Biometrika 22(3-4), 423-424.

Pearson, K. (1920). Notes on the history of correlation.

Biometrika 13(1), 25-45.

Pearson, K. (1968). *Tables of the Incomplete Beta-Function: With a New Introduction*. Cambridge University Press.

Pletser, V. (2012). Does the Ishango Bone Indicate Knowledge of the Base 12? An Interpretation of a Prehistoric Discovery, the First Mathematical Tool of Humankind.

arXiv math.HO 1204.1019.

Plotly Technologies Inc (2015). Collaborative data science.

<https://plot.ly>.

R. A. Fisher (1918). The correlation between relatives on the supposition of mendelian inheritance. *Philos. Trans. R. Soc.*

Edinb. 52, 399-433.

Richterich, A. (2018). *The Big Data Agenda: Data Ethics and*

Critical Data Studies. Critical, Digital and Social Media

Studies. University of Westminster Press.

Rogel-Salazar, J. (2014). *Essential MATLAB and Octave*. CRC

Press.

Rogel-Salazar, J. (2016, Mar). Motor Trend Car Road Tests.

<https://doi.org/10.6084/m9.figshare.3122005.v1>.

- Rogel-Salazar, J. (2020). *Advanced Data Science and Analytics with Python*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press.
- Rogel-Salazar, J. (2021a, May). GLA World Cities 2016. <https://doi.org/10.6084/m9.figshare.14657391.v1>.
- Rogel-Salazar, J. (2021b, Dec). Normal and Skewed Example Data. <https://doi.org/10.6084/m9.figshare.17306285.v1>.
- Rogel-Salazar, J. (2022a, Feb). Anscombe's Quartet. <https://doi.org/10.6084/m9.figshare.19221720.v3>.
- Rogel-Salazar, J. (2022b, Feb). Jackalope Dataset. <https://doi.org/10.6084/m9.figshare.19221666.v3>.
- Rogel-Salazar, J. (2022c, Feb). Python Study Scores. <https://doi.org/10.6084/m9.figshare.19208676.v1>.
- Rogel-Salazar, J. (2022d, Jan). Starfleet Headache Treatment - Example Data for Repeated ANOVA. <https://doi.org/10.6084/m9.figshare.19089896.v1>.
- Rogel-Salazar, J. (2017). *Data Science and Analytics with Python*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. CRC Press.
- Rothamsted Research (2020). Statement on R. A. Fisher. www.rothamsted.ac.uk/news/statement-r-fisher. Accessed: 2021-02-14.
- Samueli, J.-J. (2010). Legendre et la méthode des moindres carrés. Bibnum journals.openedition.org/bibnum/580. Accessed: 2021-02-14.

Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. *Biometrics Bulletin* 2(6), 110–114.

Scheinerman, E. A. (2012). *Mathematics: A Discrete Introduction*. Cengage Learning.

Scientific Computing Tools for Python (2013). NumPy. <http://www.numpy.org>.

Seaborn, J. B. (2013). *Hypergeometric Functions and Their Applications*. Texts in Applied Mathematics. Springer New York.

Shapiro, S. S. and Wilk, M. B. (1965, 12). An analysis of variance test for normality (complete samples)†. *Biometrika* 52(3-4), 591–611.

Short, J. E. and S. Todd (2017). What’s your data worth? MIT Sloan Management Review, sloanreview.mit.edu/article/whats-your-data-worth/. Accessed: 2021-01-08.

SINTEF (2013). Big Data, for better or worse: 90% of world’s data generated over last two years. www.sciencedaily.com/releases/2013/05/130522085217.htm. Accessed: 2021-01-01.

Smirnov, N. V. (1939). Estimate of deviation between empirical distribution functions in two independent samples. *Bulletin Moscow University* 2(2), 3–16.

Stahl, S. (2006). The evolution of the normal distribution. *Mathematics Magazine* 79(2), 96.

- Statista (2020). The 100 largest companies in the world by market capitalization in 2020. www.statista.com/statistics/263264/top-companies-in-the-world-by-market-capitalization. Accessed: 2021-01-03.
- Strang, G. (2006). *Linear Algebra and Its Applications*. Thomson, Brooks/Cole.
- Student (1908). The probable error of a mean. *Biometrika* 6(1), 1–25.
- The Event Horizon Telescope Collaboration (2019). First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole. *ApJL* 875(L1), 1–17.
- Three Dragons, David Lock Associates, Traderisks, Opinion Research Services, and J. Coles (2016). Lessons from Higher Density Development. Report to the GLA.
- Tufte, E. (2022). The Work of Edward Tufte and Graphics Press. <https://www.edwardtufte.com/tufte/>. Accessed: 2022-23-02.
- Tukey, J. W. (1977). *Exploratory Data Analysis*. Number v. 2 in Addison-Wesley Series in Behavioral Science. Addison-Wesley Publishing Company.
- van der Zande, J. (2010). Statistik and history in the German enlightenment. *Journal of the History of Ideas* 71(3), 411–432.
- Vinten-Johansen, P., H. Brody, N. Paneth, S. Rachman, M. Rip, and D. Zuck (2003). *Cholera, Chloroform, and the Science of Medicine: A Life of John Snow*. Oxford University Press.

Waskom, M. L. (2021). Seaborn: Statistical data visualization. *Journal of Open Source Software* 6(60), 3021.

Wasserstein, R. L. and N. A. Lazar (2016). The ASA statement on p-values: Context, process, and purpose. *The American Statistician* 70(2), 129–133.

Welch, B. L. (1947, 01). The generalization of 'Student's' problem when several different population variances are involved. *Biometrika* 34(1-2), 28–35.

Wilcoxon, F. (1945). Individual comparisons by ranking methods. *Biometrics Bulletin* 1(6), 80–83.

William H. Kruskal and W. Allen Wallis (1952). Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association* 47(260), 583–621.