

Bibliography

- Cattell, R.B. (1966), “The scree test for the number of factors.” *Multivariate Behavioral Research*, 1, 629–637.
- Cohen, Mike X. (2022), *Practical Linear Algebra for Data Science*, 1 edition. O’Reilly Media, CA: Sebastopol.
- Ferré, L. (1995), “Selection of components in principal component analysis: A comparison of methods.” *Computational Statistics & Data Analysis*, 19, 669–682.
- Fisher, R. A. (1988), “Iris.” UCI Machine Learning Repository. <https://doi.org/10.24432/C56C76>.
- Gwet, K. L. (2020), *Beginner’s Guide to Principal Components: Applications with Microsoft Excel*, 1 edition. AgreeStat Analytics, MD: AgreeStat Analytics.
- Hotelling, H. (1933), “Analysis of a complex of statistical variables into principal components.” *Journal of Educational Psychology*, 24, 417–441, 498–520.
- Jackson, D. (1993), “Stopping rules in principal components analysis: a comparison of heuristical and statistical approaches.” *Ecology*, 74, 2204–2214.

- Jolliffe, I.T. (1986), *Principal Component Analysis*. Springer-Verlag, New York.
- Jolliffe, I.T. (2002), *Principal Component Analysis*. Springer, New York.
- Kaiser, H.F. (1960), “The application of electronic computers to factor analysis.” *Educational and Psychological Measurement*, 20, 141–151.
- Lukjancenko, O., T.M. Wassenaar, and D.W. Ussery (2010), “Comparison of 61 sequenced escherichia coli genomes.” *Microbial Ecology*, 60, 708–720.
- McCabe, G. P. (1984), “Principal variables.” *Technometrics*, 26, 137–144.
- Pearson, K. (1901), “On lines and planes of closest fit to systems of points in space.” *Phil. Mag.*, 6, 559–572.
- Singh, K. (2014), *Linear Algebra: Step by Step*, 1 edition. Oxford University Press, Great Clarendon Street, Oxford, OX2 6DP.
- Xu, D. and Y. Tian (2015), “A comprehensive survey of clustering algorithms.” *Annals of Data Science*, 2, 165–193. <https://doi.org/10.1007/s40745-015-0040-1>.
- Zwick, W. and W. Velicer (1986), “Comparison of five rules for determining the number of components to retain.” *Psychological Bulletin*, 99, 432–446.

Author Index

Cattell, R.B., 42

Cohen, MX, 9

Ferré, L., 41

Gwet, KL, 9

Hotelling, H., v

Jackson, D., 41

Jolliffe, I.T., 41, 42, 50, 51, 53

Kaiser, H.F., 41

Lukjancenko, O., 69

McCabe, G.P., 50

Pearson, K., v

Singh, K, 9

Tian, Y., 68

Ussery, D.W., 69

Wassenaar, T.M., 69

Xu, D., 68

Zwick W., 41

Subject Index

- Between-Cluster Sum of Squares, 106
- Centering, 18
- Centroid, 74
- Cluster Analysis, 63
- Clustering, 63
- Composite Score, 4, 23
 - Indicator, 4
 - Variable, 4
- Correlation Coefficient, 5
- Correlation Matrix, 21
- Data Table, 83
- Distance, 71
- Eigenvalue, 22
- Eigenvector, 22
- Elbow Method, 42
- Elimination Method, 50
- Euclidean Distance, 71
- Factors, 24
- Inclusion Method, 50
- Loading, 43, 50
- Loadings, 24
- Named Range, 116
- Named Ranges, 88
- Named Variables, 88
- Outliers, vii, 38, 53
- Principal Component, 12
 - Scores, 13
- Principal Component Scores, 4
- Principal Components, 12, 22
- Scree Graph, 41
- Scree Plot, 41
- Standardization, 17
- Sum of Squared Errors, 106
- Total Sum of Squares, 107
- Within-Cluster Distance, 106

Within-Cluster Sum of Squares,

106

