

A Biplot View of Violent Crime in South Africa

N.J. le Roux¹, and B.W. Gurr²

¹Stellenbosch University, Stellenbosch, South Africa; njlr@sun.ac.za

²Stellenbosch University, Stellenbosch, South Africa; benjamingurr@outlook.com

Abstract

The violent crime and homicide rates of South Africa are a growing concern. The officially reported South African crime data for the period 2004 – 2013 are arranged in the form of per annum two-way tables where the rows depict twelve violent crimes and the columns represent the nine provinces of the country. The two-way tables are expressed in terms of frequencies and as rates per capita of the population in each province. Although mosaic plots allow for exploratory analysis of these two-way tables the focus is on the use of biplots for visualizing the inter-province changes in twelve violent crimes over the 2004 – 2013 study period.

Gower and others (2011, p. 300) give a general matrix formulation for a function of a contingency table. A weighted least-squares criterion is then defined leading to a singular value decomposition allowing for an inner product approximation as a basis for constructing several variants of correspondence analysis (CA) biplots. Each of these biplots shows the rows (columns) as points and the columns (rows) as calibrated axes approximating any one of Pearson residuals, deviations from independence, contingency ratios, chi-squared distances, correlations, row (column) profiles. A problem that occurs often in practice with CA biplots is that the displayed points might tend to be huddled together. Lambda-scaling is suggested by Gower and others (2011, p. 23) for addressing this problem in such a way that the underlying inner product approximation is not violated. Correspondence analysis (CA) biplots are constructed using Lambda scaling for the crime data and their usage are discussed in detail. It is shown how the occurrence of the various crimes changes differentially over time in the nine provinces.

In contrast with CA biplots it is argued that the crime data can be expressed in the form of compositional data and as such, allow for the construction of log-ratio biplots (Greenacre, 2010, p. 69). It is demonstrated that, in general, for the South African crime data the log-ratio biplots not only confirm the conclusions drawn from the CA biplots but have the ability to elucidate inter-provincial differences occurring over time that are not discernable in ordinary CA biplots.

References

Gower, J.C., Lubbe, S. and Le Roux, N.J. (2011). *Understanding biplots*. Chichester: Wiley.

Greenacre, M. (2010). *Biplots in practice*. Bilbao: Fundación BBVA.