data.Normalization(clusterSim)

Types of variable normalization formulas

1	Selection of objects and variables	data matrix $[x_{ij}]$		
2	Variable scale level	ratio	ratio	interval
	Selection of variable normalization formula	n6 – quotient transformation (x/sd) n7 – quotient transformation (x/range) n8 – quotient transformation (x/max) n9 – quotient transformation (x/mean) n10 – quotient transformation (x/sum) n11 – quotient transformation x/sqrt(SSQ)	n1 – standardization n2 – Weber standardization n3 – unitization n4 – unitization with zero minimum n5 – normalization in range [–1, 1]	n1 – standardization n2 – Weber standardization n3 – unitization n4 – unitization with zero minimum n5 – normalization in range [–1 ,1]
	Transformed variable scale level	ratio	interval	interval

(n1) $$z_{ij} = s_j^{-1} x_{ij} - \bar{x}_j s_j^{-1},$$

(n2) $$z_{ij} = (x_{ij} - Me_j) \big/ 1{,}4826 \cdot MAD_j,$$

(n3) $$z_{ij} = r_j^{-1} x_{ij} - \bar{x}_j r_j^{-1},$$

(n4) $$z_{ij} = \left[x_{ij} - \min_i \{x_{ij}\} \right] \Big/ r_j$$

(n5) $$z_{ij} = (x_{ij} - \bar{x}_j) \Big/ \max_i \left| x_{ij} - \bar{x}_j \right|$$

(n6) $$x_{ij} \big/ s_j$$

(n7) $$x_{ij} \big/ r_j$$

n(8) $$x_{ij} \Big/ \max_i \{x_{ij}\}$$

(n9) $$x_{ij} \big/ \bar{x}_j$$

(n10) $$x_{ij} \Big/ \sum_{i=1}^{n} x_{ij}$$

(n11) $$x_{ij} \Big/ \sqrt{\sum_{i=1}^{n} x_{ij}^2}$$

where: $x_{ij}(z_{ij})$ – *i*-th observation on *j*-th variable (*i*-th normalized observation on *j*-th variable),

$\bar{x}_j, s_j$ – mean and standard deviation for *j*-th variable,

Me_j, MAD_j – median and median absolute deviation for *j*-th variable,

$r_j = \max_i \{x_{ij}\} - \min_i \{x_{ij}\}$.

References

Gatnar, E., Walesiak, M. (Eds.) (2004), *Metody statystycznej analizy wielowymiarowej w badaniach marketingowych [Multivariate statistical analysis methods in marketing research]*, Wydawnictwo AE, Wroclaw, pp. 35-38.

Jajuga, K., Walesiak, M. (2000), *Standardisation of data set under different measurement scales*, In: R. Decker, W. Gaul (Eds.), *Classification and information processing at the turn of the millennium*, Springer-Verlag, Berlin, Heidelberg, pp. 105-112.

Milligan, G.W., Cooper, M.C. (1988), *A study of standardization of variables in cluster analysis*, "Journal of Classification", vol. 5, pp. 181-204.