

Statistical methods for the detection of falsified data by interviewers and application survey data in Africa

Souleymane Diakité ENSAE-Senegal, e-mail : souleymanediakite2@yahoo.fr

Abstract

Data quality has a significant impact on the results of analyzes. The concern for quality is all the more justified, if those responsible for the collection are not the professional trade. According to the director of the Institute of Statistics of Mali, 80% of people working in the field of statistics in Mali are not statisticians.

In this work, we applied several methods to detect falsified data. Including law Benford, hierarchical and mixed ascending classification or analyze discriminant. Indicators used: the percentage of extreme values, the percentage of missing values, the percentage jump so the percentage of modality "Other." The results show that the classification seems to be better compared to the application of Benford's law or analyze discriminant. Also the best indicators for the detection of falsified data are ratios of extreme values and missing values. These ratios are much lower in the falsifiers.

Key Words: Falsifiers, Interviewers, Benford laws, classification, analyze discriminant.