

# Innovative R tools supporting market survey quality

Ágnes Konkoly

University of Debrecen, Faculty of Informatics  
agi@konkoly.org

## Abstract

The aim of this work is to examine and develop the quality of surveys. To achieve these goals we consider the following topics: sentiment analysis [3], wordcloud, robot detection and other factors that has an effect on the survey quality.

Sentiment analysis is an expert tool for evaluating text fields of surveys, which process can be hard to automatize. It helps us to identify those clients who are satisfied or dissatisfied with us. Here can be mentioned another useful tool, the wordcloud. Wordcloud draws in a cloud the most frequent words of our documents.

After working with texts, I tried to increase the reliability with robot detection. Here I developed scripts which aim to recognise robots based on the answers and the response time of the respondent. Another attempt to increase reliability was to detect the possible duplications or similarities [2],[1] in the survey. I also tried to find those questions on which an enormous amount of participants left the survey. With this we can easily find if we made some mistakes while creating the survey.

*Keywords:* surveys, R programming language, text mining, robot detection

*MSC:* 62D99,68U15

## References

- [1] Hamming, R. W. (1950). *Error detecting and error correcting codes*, Bell System Technical Journal, Vol. 29 (2): 147–160
- [2] Levenshtein, V. I. (1966). *Binary codes capable of correcting deletions, insertions, and reversals*, Soviet Physics Doklady, Vol. 10 (8): 707–710.
- [3] Miner, G. (2012). *Practical Text Mining And Statistical Analysis For Non Structured Text Data Applications*, Academic Press