25th ANNIVERSARY EDITORIAL

(Methods and Data)

## Introduction

One of the sections of the 25th Anniversary Editorial of the International Journal of Social Research Methodology (IJSRM), presents the thematic trends in published contributions, for the whole period of 25 years of the journal’s life. Investigating the thematic trends in published contributions was not an easy task, not only because of the huge number of published papers, but also because of various technical details (for example, the published papers were not accompanied by keywords in the first volumes).

Coding and charting the thematic trends of published papers proved to be a very laborious task and the workload was shared between four researchers. The aim of this document is to help interested readers understand the nature and the structure of the dataset. This could be useful to those interested in extending our own analysis, which had to be confined in the limited space of an Editorial.

## Methodology

It is important for prospective users of the dataset to understand how the published content was coded and how reliable the coding was.

Three coders worked in parallel for three weeks, under the supervision of an experienced researcher. They coded each of the contributions published by IJSRM, not only in the first 25 issues, but also in the ‘latest papers’ section of the journal’s web page, which includes papers which have not yet been assigned to specific volumes/issues.

The three coders and the experienced researcher, developed a coding scheme, which included all the necessary variables to be coded in an Excel file. Through long online meeting, the group discussed the aims of the coding exercise, the structure and content of the coding scheme etc. The group coded a number of common papers to confirm that they interpreted the coding scheme in the same way. Regular online meetings and email exchanges were necessary to discuss various issues which emerged and to keep the coders in sync. To make sure that the coders did not ‘drift’ over time, they were instructed to ‘blindly’ re-code 5%-10% of each other’s excel file incrementally (every few days). The coders were in communication all the time and they exchanged emails where they would update each other about coding difficulties in order to remain in sync. As a result of this procedure, various issues came to the surface (e.g. there were many papers which could not be easily categorized as Qualitative, Quantitative or Mixed, so a new category was created; more information later).

When all the coding was completed, the experienced researcher re-coded blindly 50 random papers - around 5% of the total number of papers in the database - but no major discrepancies were detected (for example, in one case, the number of views was mis keyed as ‘867’ instead of ‘861’) .

Overall, there is no reason to believe that there is widespread bias or errors in the data. We expect the dataset to give a fair interpretation of what has been published in the journal in the last 25 years of its life.

## Variables in dataset

The dataset includes the following variables:

|  |  |
| --- | --- |
| Vol | Volume |
| No | Issue number |
| Title | The title of the paper (no coding, it was just copied and pasted) |
| Abstract | The abstract of the paper (no coding, it was just copied and pasted) |
| Keywords | The keywords of the paper (no coding, it was just copied and pasted) |
| Paradigm | Main research paradigm. Takes four values: Qualitative, Quantitative, Mixed Methods, General/Other. Note: The General/Other category refers to papers which cannot be described accurately by the three other codes (Qualitative, Quantitative, Mixed Methods) |
| Views | Number of views (as reported on the journal’s web page) |
| CrossRef | Number of CrossRef citations (as reported on the journal’s web page) |
| Altmetric | Altmetric count (as reported on the journal’s web page) |

##

## Data filtering

The original dataset consisted of 1043 records, but book reviews, editorials and other small items were removed, resulting to a ‘clean’ dataset of 924 published papers (including ‘Research Notes’).

## Dataset format

The dataset is provided as an R data frame, with the name ‘EditorialData.Rda’.