File **masterTable.csv** contains the data used in the paper, as well as additional data not analyzed in the paper.

Every row of the table reports data from a primary study concerning one smell, one external software quality and one program. Hence, in general there are several rows concerning the same primary study, since most primary studies evaluated multiple smells against multiple qualities in multiple programs.

The meaning of each column is the following:

| Column | Meaning |
| --- | --- |
| PaperRef | The same reference to the primary study as used in the paper |
| Authors | The authors of the primary study |
| Year | The primary study's publication year |
| Title | The title of the primary study |
| SWquality | The software quality evaluated |
| LOCnormalized | Boolean: whether the effect of size was taken into account or not |
| CodeSmell | The code smell whose harmfulness was evaluated |
| Harmful | The harmfulness level resulting from the evaluation (1=harmful, 0= not harmful). See note below. |
| HowDetected | The type of method used to detect the existence of smell: either manually or via a tool |
| AnalyzedProject | The software project that provided the analyzed code. When the authors did not disclose the name of the analyzed projects, fantasy names are used, such as XYZAnon. |
| NumProjects | The number of projects the row refers to. Generally, every row gives the results of one project; however, in some cases the primary studies analyzed the code of several projects together. In these cases, NumProjects is greater than one. |
| ProjectNature | Industrial, OpenSource or Lab. 'Lab' indicates that the program was developed in a research environment, although with 'realistic' characteristics. |
| AnalyzedReleasesCommits | Number of project commits that were considered. |
| KLOC | Size of the analyzed product in KLOC. |
| NumClasses | The number of classes in the analyzed product(s). |
| Language | The implementation language of the analyzed product(s). |
| EvaluationType | How the evaluation of harmfulness was carried out: either formally or informally. |
| StatisticalTechnique | Statistical technique(s) used in the evaluation of harmfulness. |
| DefinitionUsed | Which definition of code smell(s) was adopted in the primary study |
| DetectionTool | What detection tool was used in the primary study to detect the presence of smells. |

Concerning the 'Harmful' column, in most cases the given value is either 0 (not harmful) or 1 (harmful). However, in just a few cases (7 out of 436), a value between 0 and 1 is given. Such values indicate that the corresponding primary studies do not provide definitive indications that the considered smell is dangerous with respect to the considered quality in the considered program.