

Research papers in Software Engineering

Form used to extract data on software engineering papers regarding its maturity in using i) statistics, ii) reproducibility and practical significance.

***Required**

1. What is the paper title? *

2. Where was it published? *

Mark only one oval.

- ☐ IST
- ☐ TSE
- ☐ EMSE
- ☐ TOSEM
- ☐ JSS

3. Of what type is the paper? *

Mark only one oval.

- ☐ Non-empirical paper *Stop filling out this form.*
- ☐ Non-primary study *Stop filling out this form.*
- ☐ Primary study

Primary study

4. Does it use QUALITATIVE data analysis? *

Tick all that apply.

- ☐ No
- ☐ Yes (Grounded theory)
- ☐ Yes (Thematic analysis)
- ☐ Yes (unknown/other type)

5. Does it use QUANTITATIVE data analysis (QDA)? *

Mark only one oval.

- ☐ No *After the last question in this section, stop filling in this form.*
- ☐ Yes
- ☐ Other: _____

6. Describe briefly the paper's goal *

Statistical techniques and approaches

7. Misc

Tick all that apply.

- ☐ Reasoning about data based on rel/abs diff, tables or raw data plotted
- ☐ Descriptive statistics (μ , σ , distributions, F-measure, graphs w/ distr/var)
- ☐ Bayesian inference
- ☐ Other: _____

8. Power analysis

Tick all that apply.

- ☐ A priori power analysis
- ☐ Post hoc power analysis
- ☐ Unclear

9. Tests for normality (claiming you have done it and don't report it is not ok!)

Tick all that apply.

- ☐ Kolmogorov–Smirnov test
- ☐ Shapiro–Wilk test
- ☐ Anderson–Darling test
- ☐ Pearson's chi-square test
- ☐ Cramér–von Mises criterion
- ☐ D'Agostino's K-squared test
- ☐ Jarque–Bera test
- ☐ Lilliefors test
- ☐ Normal probability plot(s)
- ☐ Other: _____

10. Parametric tests

Tick all that apply.

- ☐ z-test
- ☐ t-test
- ☐ F-test (simple anova)
- ☐ ANOVA, ANCOVA, MANOVA, MANCOVA
- ☐ Linear regression
- ☐ Factorial DOE
- ☐ Other: _____

11. Nonparametric tests

Tick all that apply.

- ☐ Fisher's exact test
- ☐ MWU
- ☐ Wilcoxon
- ☐ Kruskal-Wallis
- ☐ Chi²
- ☐ Friedman, Mood's median, Cochran-Mantel-Haenszel (X^2), and other *NOVA alternatives
- ☐ Non-linear regression (e.g. log)
- ☐ Other: _____

12. Correction for Type I errors (multiple testing)

Tick all that apply.

- ☐ Bonferroni
- ☐ Bonferroni-Holm
- ☐ Tukey HSD
- ☐ Fisher LSD
- ☐ Dunnett
- ☐ Hommel
- ☐ Benjamini-Hochberg
- ☐ Other: _____

13. Confidence intervals

Mark only one oval.

- ☐ Yes
- ☐ No
- ☐ Other: _____

14. Effect sizes

Tick all that apply.

- ☐ Common Language (CL)
- ☐ Cohen's d
- ☐ Glass' d
- ☐ Hedge's g
- ☐ Odds ratio
- ☐ Cliff's d
- ☐ Vargha-Delaney \hat{A}_{12}
- ☐ Spearman's ρ et al.
- ☐ Pearson's r et al. (incl. R^2 etc.)
- ☐ η^2
- ☐ Other: _____

15. Latent variable analysis

Tick all that apply.

- ☐ PCA
- ☐ Factor analysis (exploratory and/or confirmatory)
- ☐ Correspondence analysis
- ☐ Hierarchical clustering
- ☐ Other: _____

Practical significance

16. Is practical significance discussed?

Mark only one oval.

- ☐ Yes, it is explicitly discussed
- ☐ No, it is NOT discussed at all
- ☐ No, but it is implicitly discussed by being motivated by a practical problem and/or context

Reproducibility

17. Reanalysis: Is raw data available?

Mark only one oval.

- ☐ Available online
- ☐ Available online (dead link)
- ☐ Available in paper
- ☐ No

18. Repeatability

Tick all that apply.

- ☐ Is a repeatability package available, i.e., scripts for executing statistical tests and/or the experiment?
- ☐ Are artefacts made available explicitly by the authors, i.e., the sw/tools we evaluate?