```
purpose: how perform experiments to evaluate methods, techniques and tools in SE
                                                                                               observations -> model
                                                            scientific method
                                                                                              formal theory -> observations
                                                            analytical method
                                             methods
                                                                                              solution changes are evaluated
                                                           engineering method
                                                                                              model proposed and evaluated through empirical studies
                                                           empirical method
                                                                                                            interpret a phenomenon based on explainations people bring to them
                                                                                   qualitative research
                                                                                                            subject is the person evaluating an object
                                                          research paradigms
                                                                                                            goal: understand why, how, ..
                                                                                                             quantify a relationship or compare 2 or more groups
                                                                                  quantitative research
                                                                                                             goal: identify a cause-effect relationship
                                                                                        questionnaire; typically retrospective
                                                                                                 descriptive
                                                                                                 explainatory (justify facts by analysing related variables)
                                                                            survey
                                                                                                 explorative (pre-study)
                                                                                        results are generalized to a larger population
                                                                                            observational study
                                                                                            track a specific attribute or enstablish relationship between attributes while monitoring projects,
                                                                                            activities, assignments
                                                                            case study
                                                                                                     compare methods/tools
                                                              strategies
                                                                                            goal
                                                                                                      how/why a certain phenomena occurs
                                                                                             controlled study
                                                                                             manipulate variables and control all other variables at fixed levels
                                                                                                      confirm theories / conventional wisdom
                                                                                                      explore relationships
                                                                            experiment
                                                                                                      evaluate accuracy of models
                                                                                                      validate measures
                                                                                             only quantitative
                                                                             risk/invasiveness: survey < experiment < case study
                                                                                                                           i.e. product, process, resource, model, metric theory
                                                                                 object of the study (WHAT)
                                                                                                                           i.e.evaluate impact of tecnique, caracterize process
                                                                                 purpose (WHY)
                                                                                                                          i.e. cost, reliability, effectivness
                                                                                 quality focust (which effect is studied)
                                                                  defintion
                                                                                                                           i.e. dev, manager, customer, researcher
                                                                                 perspective (whose view)
Experimentation in SE (book)
                                                                                                                             subjects & objects
                                                                                 context (WHERE)
                                                                                                                H0: no trends or patterns discovered
                                                                                                                H1..N: to reject H0
                                                                                    formal statement of HP
                                                                                                                          type I error: P (reject H0 | H0 true) -> false pattern detected
                                                                                                               error
                                                                                                                         type II error: P (not reject H0 | H0 false) -> pattern undetected
                                                                                       determine variables (input/output) + accepted values (measure scale)
                                                                                                                  population sampling
                                                                                                                  sample size
                                                                                                                                        randomization
                                                                                                                   design principles
                                                                                                                                        blocking
                                                                                                                                        balancing
                                                                                                                                                                     h0: same mean for output var
                                                                                                                                       1 factor / 2 treatments
                                                                                                                                                                      1 random subject on 1 random treatment
                                                                      planning
                                                                                                                                                                      1 subject on each treatment in random order
                                                                                           experiment design
                                                                                                                                                                        1 random subject on 1 random treatment
                                                                                                                                       decrease impact of variability between subjects
                                                                                                                                                                        1 subject on each treatment in random order
                                                                                                                                                                      h0: effect of T1 on Fa = effect of T2 on Fa
                                                                                                                  std design types
                                                                                                                                                                      h0: effect of Ti/Fj = 0 (for each i,j)
                                                                                                                                       2 factor / 2 treatments
                                                                                                                                                                      permutations (Fa=Ta1/Fb=Tb1, ..)
                                                                                                                                                                                                                    if Tb has different meaning depending on tratment of Fa
                                                                                                                                                                      nested permutations (Fa=Ta1/Fb=Tb1, ..)
                                                  experiment
                                                                                                                                                                      permutations (Fa=Ta1/Fb=Tb1/Fc=Tc1, ..)
                                                                                                                                       2+ factor / 2 treatments
                                                                                                                                                                                                                                partial permutations on factors that are weakly related
                                                                                                                                                                    fractional factorial design (A1, B1, C2 / A1, B2, C1 / .. )
                                                                                        instrumentation: objects, guidelines, measurment instruments
                                                                                                                                                                  factors: subject sampling, subject treatment, disturbing factors, ...
                                                                                                               horiz: causal relationship treatment - outcome?
                                                                                                  internal
                                                                                                                                                                  are experiment settings comparable with those existing in the generalized environment?
                                                                                                                horiz: is generalization possible?
                                                                                                  external
                                                                                                                 vert: relationship theory – observation? are experiment settings consistent with theory settings?
                                                                                                  construct
                                                                                      validity
                                                                                                                                                                   statistical tests, sample size, measurment
                                                                                                                  horiz: stat relationship treatment - outcome?
                                                                                                  conclusion
                                                                                                   internal > external > construct > conclusion
                                                                                                     obtain participants consent
                                                                                    preparation
                                                                                                     anonimity
                                                                                                     reward
                                                                    operation
                                                                                    execution
                                                                                    data validation
                                                                                                                                       central tendency: mean, median, mode, ..
                                                                                                   understand: descriptive analysis
                                                                                                                                       dispersion: std deviation, variance, coeff. of variation, ...
                                                                                                                                       dependency: Pearson coeff, Spearman coeff, Kendall coeff
                                                                                                               eliminate outliers generated by errors
                                                                                                   reduce
                                                                                                                  parametric: normally distributed data
                                                                    analysis & interpretation
                                                                                                                                            chi-2, binomial test (NP)
                                                                                                                  1 factor / 1 treatment
                                                                                                                                                                  t-test, F-test (P)
                                                                                                                  1 factor / 2 treatments (randomized design)
                                                                                                      HP test
                                                                                                                                                                  Mann-Whitney, Chi-2 (NP)
                                                                                                                                                                  paired t-test (P)
                                                                                                                  1 factor / 2 treatments (paired comparison)
                                                                                                                                                                  Wilcoxon, Sign test (NP)
                                                                                                                    introduction (objective, motivation, ..)
                                                                                                                    problem statement (experiment definition)
                                                                                                                   experiment planning
                                                                    presentation
                                                                                       documentation of results
                                                                                                                   experiment operation
                                                                                                                   data analysis
                                                                                                                   interpretation of result
                                                                                                                   discussion and conclusions
```

lab package for replication