

Lab 2, Checkpoint 2 - Uma visualização para publicação

1. Introdução

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Quisque tempus condimentum massa ut posuere. Curabitur eget maximus metus, in tincidunt odio. Fusce eleifend mauris sit amet neque finibus, non ornare lorem laoreet. Mauris ultrices sollicitudin enim, dapibus tempor ipsum sollicitudin at. Fusce pulvinar eros in felis eleifend, quis vulputate purus luctus.

2. Dataset

Os dados utilizados na análise foram extraídos por meio Travis Torrent. Neles estão presentes informações sobre builds, commits, casos de testes e etc de diferentes projetos desenvolvidos sob as linguagens de programação Java ou Ruby hospedados no GitHub. Com base nos dados, foi investigada a seguinte pergunta:

- ***Maiores projetos (linhas de código) tendem a possuir uma densidade de testes maior? A linguagem utilizada pode influenciar nesse aspecto?***

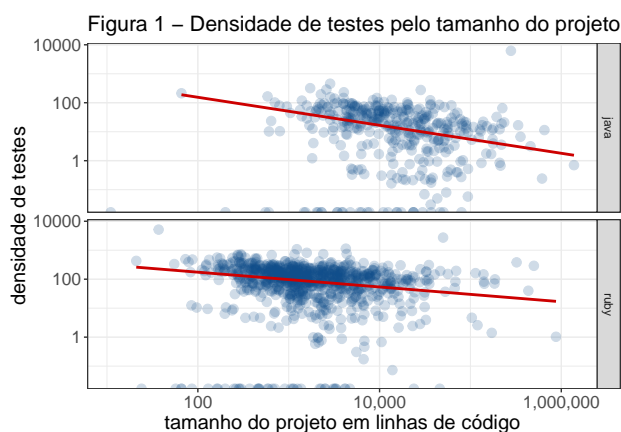
Para esse estudo foram descartados alguns projetos da linguagem de programação Javascript existentes nos dados, devido a análise concentrar-se na integração contínua de testes em projetos Java e Ruby.

No que se refere a investigação da pergunta proposta, foram utilizadas as seguintes variáveis:

- ***tamanho do projeto em linhas de código:*** maior tamanho do projeto considerando linhas de código encontrado nos dados;
- ***densidade de testes:*** densidade máxima de testes encontrada nos dados considerando cada mil linhas de código do projeto.
- ***linguagem:*** principal linguagem de programação utilizada no projeto, variando entre Java, Ruby e Javascript;

3. Densidade de testes e tamanho do projeto

Assumindo a densidade máxima de testes criados pelos desenvolvedores para cada mil linhas de código como a densidade de testes do projeto e o tamanho do projeto como o número máximo de linhas de código, podemos verificar a pergunta investigada. Assim, a fim de analisar a relação entre o tamanho do projeto e a densidade testes, a Figura 1 é apresentada.



Com base na Figura 1, é notada a existência de pequena tendência (fraca correlação) entre os projetos onde quanto maior o projeto menor sua densidade de testes em ambas as linguagens. No entanto, é perceptível a existência de uma grande dispersão nos projetos Java e Ruby, onde é possível encontrar projetos com densidade de testes abaixo de 1 e com tamanho superior a 10 mil linhas de código. Em outro extremo podemos identificar projetos com tamanho superior a 10 mil linhas de código e com densidade de testes superior a 100, apresentando o oposto da tendência. Assim, é possível concluir que não existe correlação forte e válida para essa análise entre o tamanho do projeto e a sua densidade de testes, logo a pergunta proposta nesse estudo não pode ser confirmada.

Por fim, foi calculado a correlação entre essas variáveis por meio do coeficiente de correlação de Pearson, resultado no valor $p = 0.01302925$, confirmando a inexistência de uma correlação.