

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS EXATAS E INFORMÁTICA
UNIDADE EDUCACIONAL PRAÇA DA LIBERDADE
Bacharelado em Engenharia de Software**

**Arthur Bicalho Lana Corrêa Fernandes
João Victor Tadeu Chaves Frois
Lucas Gabriel Padrão Rezende**

Trabalho laboratório de Experimentação de Software

Belo Horizonte
2022

Arthur Bicalho Lana Corrêa Fernandes
João Victor Tadeu Chaves Frois
Lucas Gabriel Padrão Rezende

Trabalho laboratório de Experimentação de Software

Trabalho apresentado na disciplina Laboratório
de Experimentação de Software.

Professor: José Laerte Pires Xavier

Belo Horizonte
2022

1 INTRODUÇÃO

O relatório apresentado tem como objetivo analisar as características de qualidade de sistemas Java utilizando a ferramenta de métricas de software CK. Essa atividade foi proposta na disciplina de Laboratório de Experimentação de Software.

1.1 Hipóteses

RQ 01. Qual a relação entre o tamanho dos PRs e o feedback final das revisões?

O tamanho dos PR's acabam influenciando no feedback final das revisões, visto que, dependendo da sua dimensão pode dificultar para análise do código desenvolvido, deixando assim passar falhas ou bugs. Dessa maneira quanto maior o PR maior a dificuldade de dar um feedback preciso e dele ser aceito.

RQ 02. Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?

O tempo de análise dos PR's acabam influenciando no feedback final das revisões, visto que, o tempo analisa o intervalo de criação do PR até a sua última atividade (sendo o fechamento dela ou o merge). Dessa maneira quanto maior o tempo de análise do PR, maior as chances do feedback ser falho, dado que, o tempo de análise recomendado é dentro de uma hora para caso tenha erros seja corrigido o mais rápido possível dando qualidade para o código. PR's com tempos de análise muito extensos têm chances menores de serem aceitos.

RQ 03. Qual a relação entre a descrição dos PRs e o feedback final das revisões?

A descrição dos PR's acabam influenciando no feedback final das revisões, visto que, a descrição analisa o número de caracteres do corpo de descrição do PR (sendo está na versão do markdown). Dessa maneira, quanto maior a descrição de um PR, isso será positivo, em virtude de, que foi uma descrição bem analisada e descrita e fazendo com que o PR tenha mais chances de ser aceito.

RQ 04. Qual a relação entre as interações nos PRs e o feedback final das revisões?

As interações de uma PR influenciam no feedback final das revisões, visto que, as interações são responsáveis por analisar o número de participantes e o número de comentários de uma PR. Dessa maneira, caso uma PR tenha muitas interações pode ser benéfico para o feedback final, visto que, muitas pessoas vão ter interagido com o PR evitando assim que falhas sejam aprovadas e tendo maiores chances de ser aceito.

RQ 05. Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?

O tamanho dos PR's acabam influenciando no número de revisões realizadas, visto que, dependendo da sua dimensão pode dificultar para análise do código desenvolvido, deixando assim passar falhas ou bugs. Dessa maneira quanto maior o PR será necessário realizar mais revisões devido ao seu tamanho, atrasando assim a sua aprovação.

RQ 06. Qual a relação entre o tempo de análise dos PRs e o número de revisões realizadas?

O tempo de análise dos PR's acabam influenciando no número de revisões realizadas, visto que, o tempo analisa o intervalo de criação do PR até a sua última atividade (sendo o fechamento dela ou o merge). Dessa maneira quanto maior o tempo de análise do PR, demonstra que foi necessário realizar diversas revisões, mostrando uma complexidade do código desenvolvido.

RQ 07. Qual a relação entre a descrição dos PRs e o número de revisões realizadas?

A descrição dos PR's acabam influenciando no número de revisões, visto que, a descrição analisa o número de caracteres do corpo de descrição do PR (sendo está na versão do markdown). Dessa maneira, quanto maior a descrição de um PR, isso será positivo, em virtude de que, foi uma descrição bem analisada e descrita e que as revisões foram eficientes.

RQ 08. Qual a relação entre as interações nos PRs e o número de revisões realizadas?

As interações de uma PR influenciam no número de revisões, visto que, as interações são responsáveis por analisar o número de participantes e o número de comentários de uma PR. Dessa maneira, caso uma PR tenha muitas interações pode ser benéfico para as revisões, demonstrando que a equipe está engajada na PR.

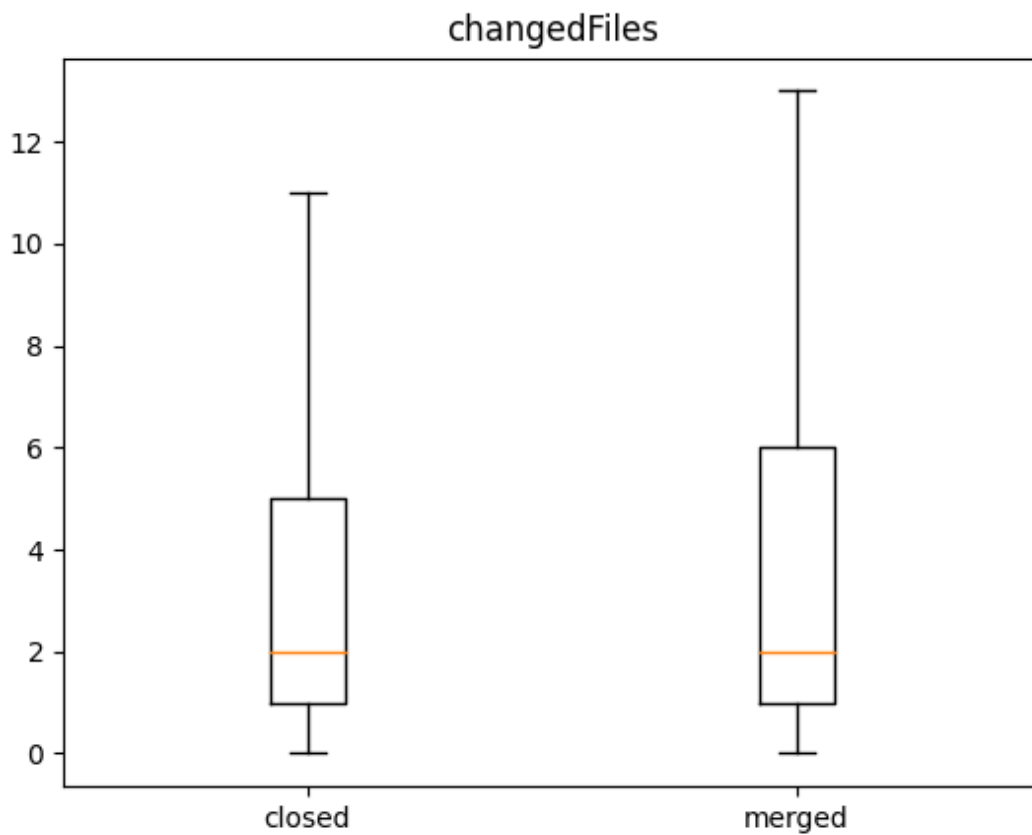
2. METODOLOGIA

O método utilizado para a análise deste trabalho foi o Code Review que tem como métrica a coleta de dados quantitativos e análises de estatísticas. Esta interpretação ocorre a partir da migração de PR's realizados nos repositórios mais populares do GitHub, que tem pelo menos cem PR's com status MERGED ou CLOSED. Este documento filtrou os PR's que foram revisados em pelo menos uma hora, isto é, o cálculo entre a data de criação e o seu encerramento deve ser maior que uma hora.

Para a análise estatística sobre o Feedback Final das Revisões foram gerados gráficos de caixa levando em consideração as medianas dos dados obtidos. Para verificar os Números de Revisões foram gerados gráficos de correlação calculando o coeficiente de Spearman entre os aspectos da questão de pesquisa.

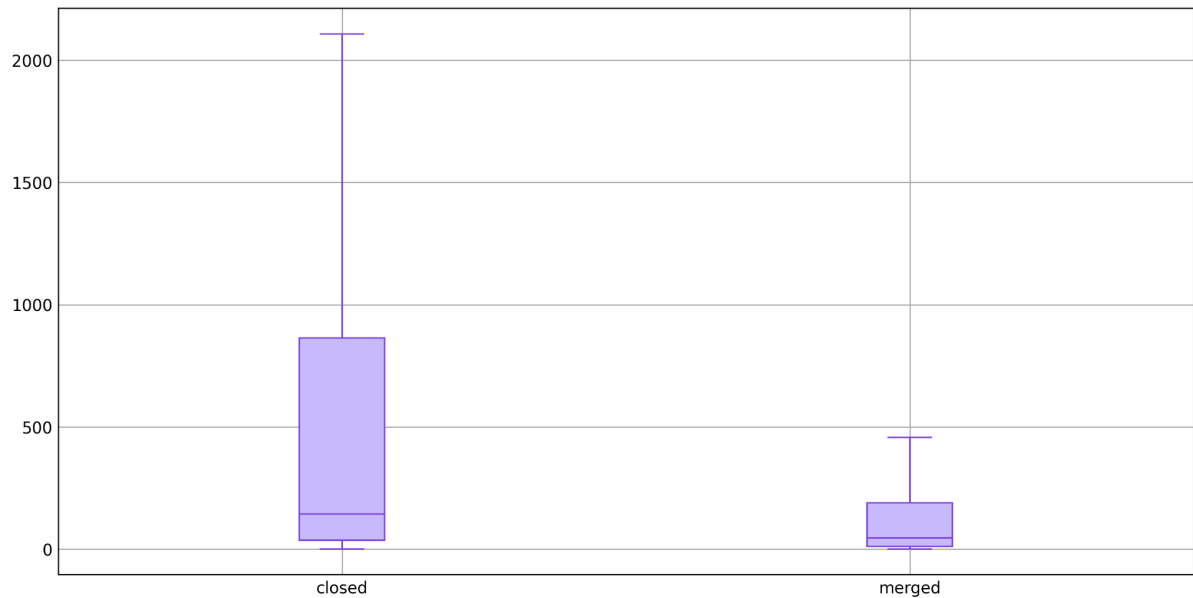
3. RESULTADOS

RQ 01. Qual a relação entre o tamanho dos PRs e o feedback final das revisões?



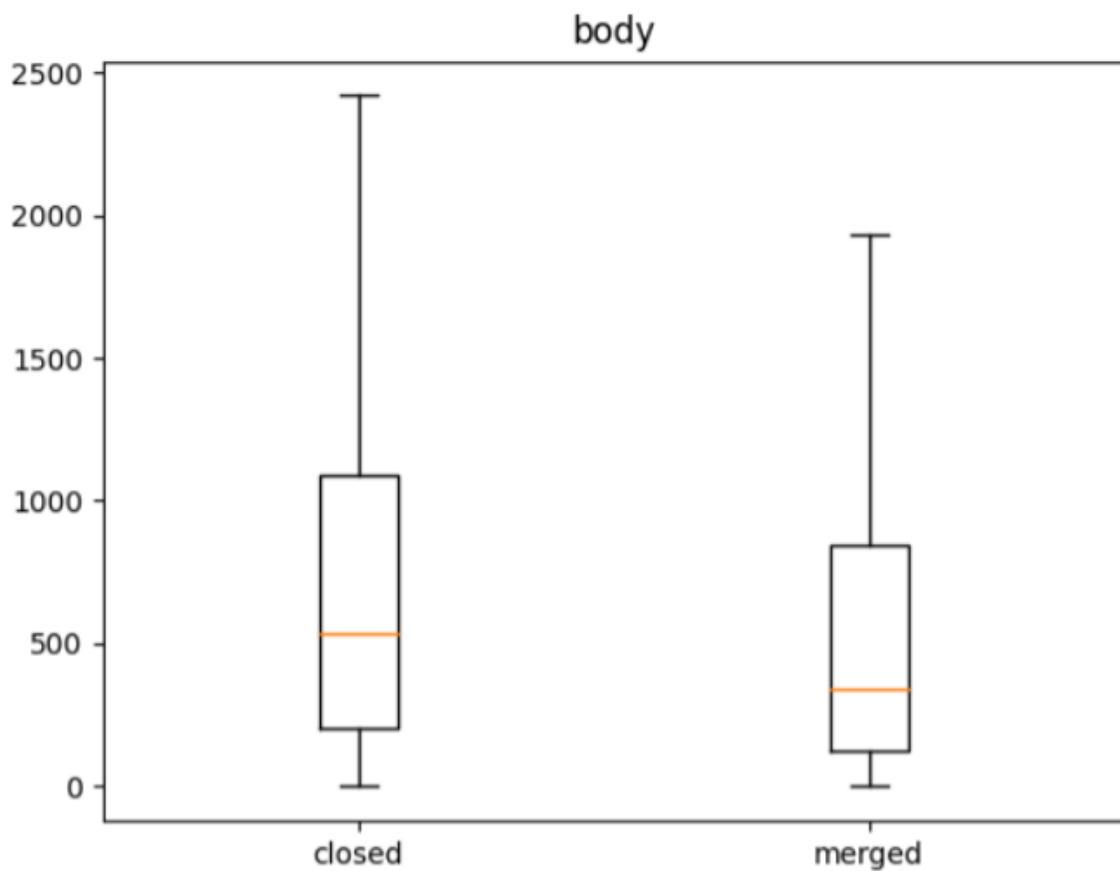
De acordo com o boxplot gerado, pôde-se observar que a quantidade de arquivos modificados no closed e merged são iguais, visto que, a sua mediana é igual e seus limites e distribuições são próximos. Dessa maneira, foi negada a ideia inicial de que os PR's com menos arquivos possuem mais chances de serem aceitos, visto que, não foi possível ver essa relação.

RQ 02. Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?



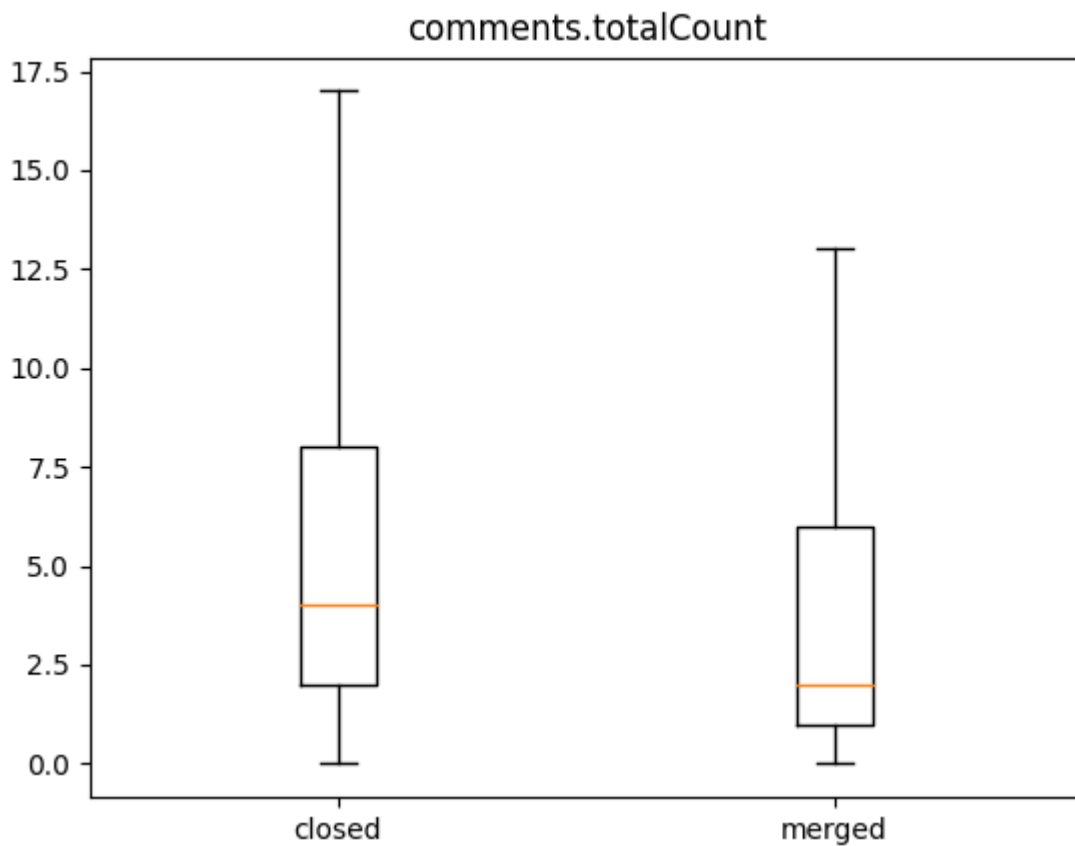
Analisando o boxplot gerado, pode-se ver que os PR's fechados e aceitos têm médias próximas. Entretanto, os PR's fechados tem um intervalo de variações de distribuições maior. Podendo assim confirmar a hipótese inicial, de que o PR que possui maior tempo de análise possui menos chance de ser aceito.

RQ 03. Qual a relação entre a descrição dos PRs e o feedback final das revisões?



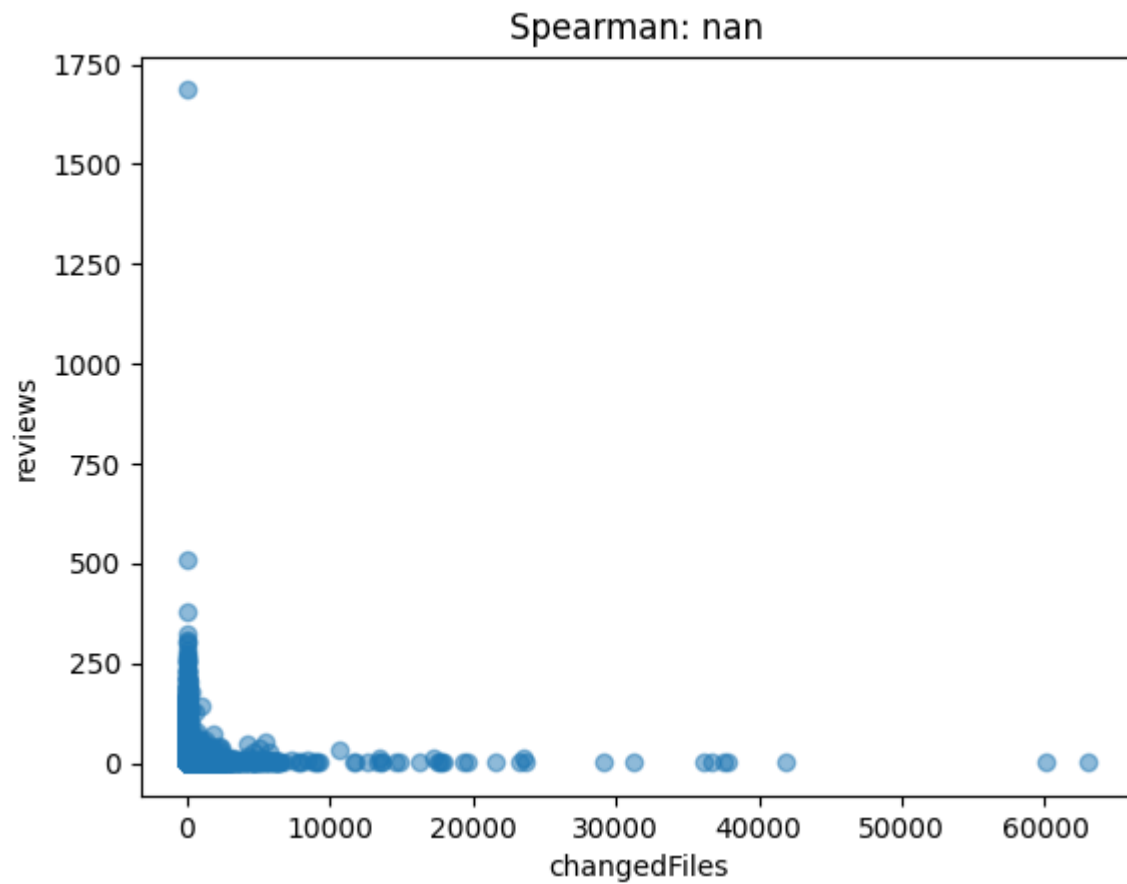
Os PR's fechados e aceitos mostrados no boxplot têm distribuições semelhantes e a mediana está bem próxima no quesito quantidade de caracteres da descrição, dando assim um aspecto indiferente no feedback da revisão. Refutando assim a hipótese inicial de que o PR com mais caracteres na descrição possuem mais chances de serem aceitos.

RQ 04. Qual a relação entre as interações nos PRs e o feedback final das revisões?



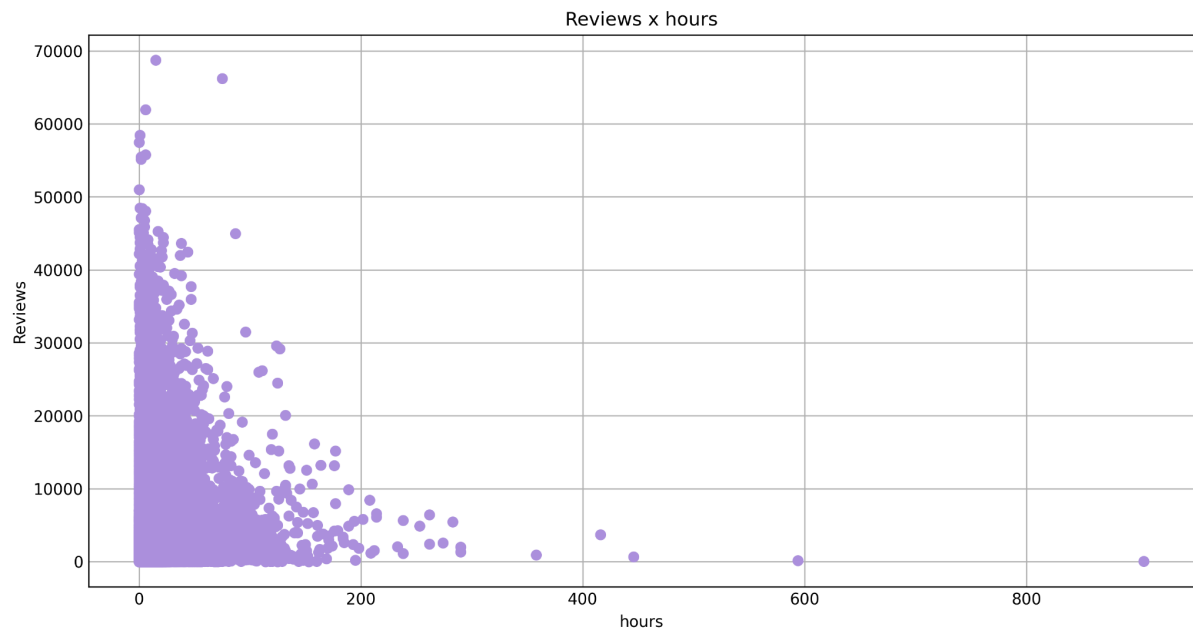
Pode-se observar no boxplot que os PR's fechados e aceitos tem uma variação semelhante. Entretanto, a mediana de interações em PR's aceitos é menor, dessa maneira a distribuição mais simétrica em PR's fechados. Dessa maneira, pode-se ver que os PR's aceitos tem uma quantidade de comentários significativos e não muito diferentes dos PR's fechados. Declinando assim a hipótese inicial de que PR's com muitas interações acabam tendo menos chances de serem aceitos.

RQ 05. Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?



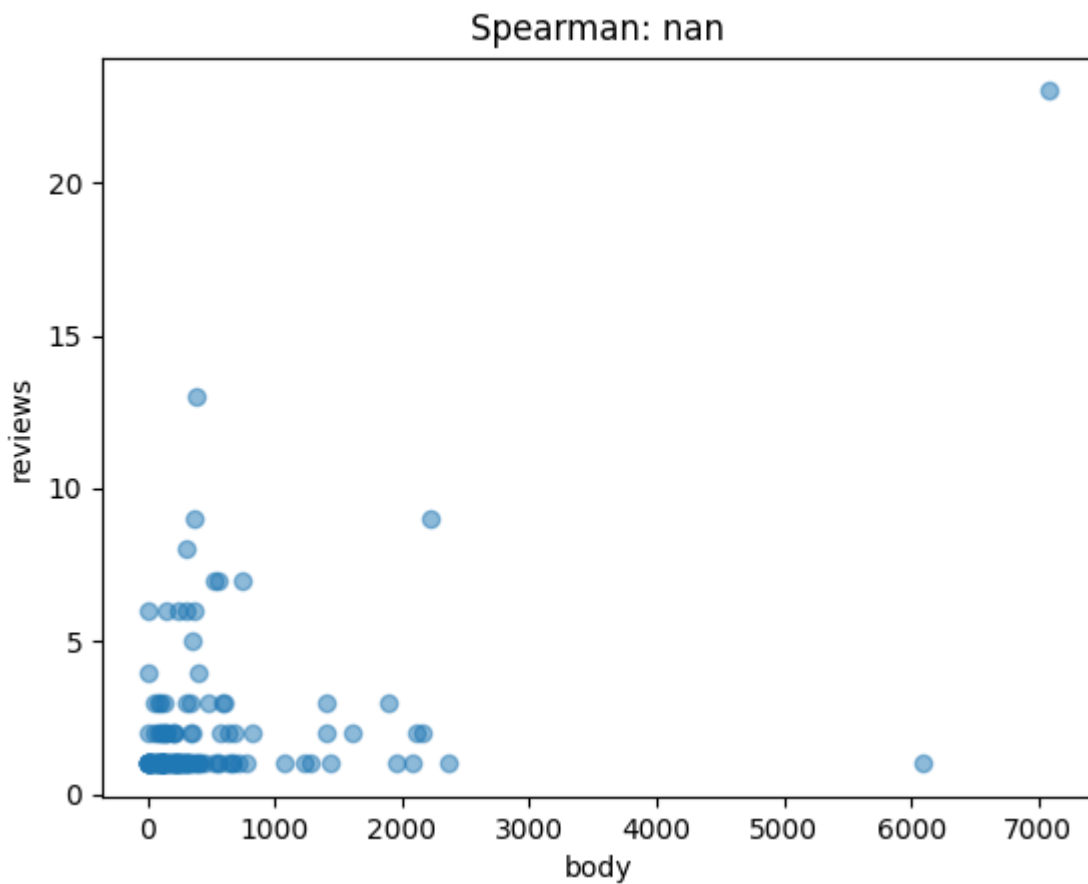
O tamanho do PR e o número de revisões que apresentou uma correlação positiva fraca entre os dados. Mesmo com os dados gerados fica inconclusivo se a quantidade de arquivos influencia na revisão final. Refutando assim a hipótese inicial de que quanto maior o pull request, maior vai ser o número de revisões realizadas.

RQ 06. Qual a relação entre o tempo de análise dos PRs e o número de revisões realizadas?



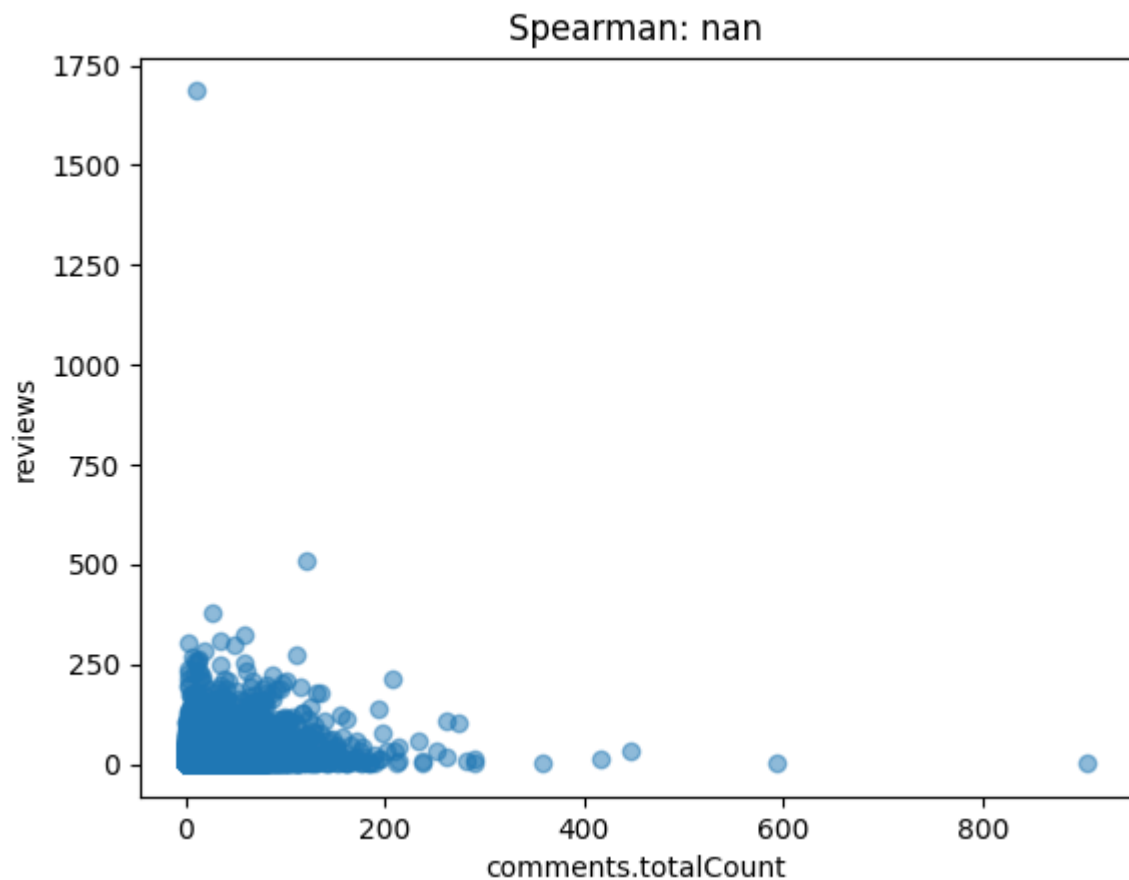
Existe uma pequena relação entre os dois aspectos. Visto que, o coeficiente de Spearman calculado mostra que existe uma correlação positiva entre os dados. Dessa maneira, a hipótese inicial pode ser confirmada de que quanto maior o tempo de análise, maior o número de revisões realizadas.

RQ 07. Qual a relação entre a descrição dos PRs e o número de revisões realizadas?



A descrição de PR's e o número de revisões apresentou uma relação positiva fraca. É possível ver uma pequena relação entre os dois aspectos, mas não é possível concluir se a quantidade de caracteres na descrição acaba influenciando na revisão. Podendo assim contrapor a hipótese inicial de que quanto maior a descrição do PR, maior o número de revisões realizadas.

RQ 08. Qual a relação entre as interações nos PRs e o número de revisões realizadas?



As interações nos PR's e o número de revisões apresentou uma correlação positiva considerável. Dessa forma, acabando tendo uma relação entre os aspectos, visto que, os PR's com menos comentários acabam sendo concentrados no eixo que apresenta menor número de revisões, confirmando assim a hipótese inicial, de que com mais interações possuem mais revisões.

4. CONCLUSÃO

A conclusão que o grupo chegou ao final deste trabalho é que os repositórios analisados mostraram um padrão nos seu resultado final, gerando assim empolgação no grupo, visto que, uma boa parte dos resultados encontrados condizem com as hipóteses planejadas pelo grupo.