

Relatório Final - Laboratório 03

**Gabriel Alejandro Figueiro Galindo¹, Marcelo Aguilar Araújo D’Almeida¹,
Philippe Roberto Dutra Chaves Vieira¹**

¹Instituto de Ciências Exatas e Informática
Pontifícia Universidade de Minas Gerais (PUC Minas)
Belo Horizonte – MG – Brasil

1. Introdução e Hipóteses

A prática de code review envolve a interação entre desenvolvedores e revisores para inspecionar o código produzido, garantindo sua integração à base principal sem riscos de introduzir novos erros à aplicação. Em sistemas open source, especialmente nos desenvolvidos no GitHub, esse processo ocorre por meio da avaliação de contribuições submetidas via Pull Requests (PRs). Para que um código seja incorporado à branch principal, é necessário que um colaborador do projeto analise e discuta a solicitação, decidindo por aprovar ou rejeitar o merge.

Neste relatório, o objetivo é analisar a atividade de code review em repositórios populares do GitHub, identificando variáveis que influenciam a aprovação de um PR sob a perspectiva dos desenvolvedores que contribuem para os projetos selecionados. Com base nas questões de pesquisa propostas, foram formuladas as seguintes hipóteses informais:

- RQ 01: PRs grandes têm mais chances de serem rejeitados, pois são mais difíceis de revisar e podem acabar introduzindo mais problemas no repositório. PRs pequenos costumam ser mais bem recebidos e aprovados com mais facilidade.
- RQ 02: PRs que levam mais tempo para serem analisados tendem a ser rejeitados, pois o atraso pode indicar dúvidas ou problemas detectados pelo revisor.
- RQ 03: PRs com descrições mais longas e detalhadas tendem a ser aprovados, pois facilitam o entendimento do propósito da mudança e demonstram cuidado por parte do autor.
- RQ 04: PRs com muitas interações (comentários e participantes) tendem a ser rejeitados, pois isso pode indicar divergência entre os participantes, em vez de um consenso.
- RQ 05: PRs maiores provavelmente passam por mais revisões, já que o volume de código aumenta a complexidade da análise, o que pode levar a um número maior de ajustes.
- RQ 06: PRs que levam mais tempo para serem analisados tendem a ter mais revisões, pois o tempo está sendo usado para discutir ou refinar as mudanças propostas.
- RQ 07: PRs com descrições curtas tendem a passar por mais revisões, pois os revisores precisarão de mais iterações para entender o que foi feito. Por outro lado, descrições mais longas podem reduzir o número de revisões necessárias.
- RQ 08: Um maior número de interações tende a resultar em mais revisões, já que cada troca de mensagens pode levar a novas versões do código até que todas as sugestões ou objeções sejam resolvidas.

Após a coleta de dados, essas hipóteses foram analisadas e discutidas para verificar a sua validade.

2. Metodologia

Para a realização do trabalho, foi desenvolvido um script em Python responsável pela coleta de dados por meio da API GraphQL do GitHub. O código tem como objetivo recuperar os 200 repositórios mais populares da plataforma que possuam, no mínimo, 100 pull requests (PRs), a fim de construir uma base de dados ampla e variada. Durante o processo, PRs com tempo de revisão inferior a uma hora são desconsideradas, pois o foco está apenas naquelas revisadas por humanos. Para a análise das hipóteses informais propostas, foram coletadas as seguintes métricas de cada PR:

- repository: Nome do repositório onde o PR foi gerado;
- number: Número do PR em relação à seu repositório;
- title: Título do PR;
- created at: Data de criação do PR;
- closed at: Data de fechamento do PR;
- review count: Número de reviews realizadas neste PR;
- description length: Tamanho da descrição (em caracteres);
- file count: Número de arquivos com mudanças;
- additions: Número de linhas adicionadas pelo PR;
- deletions: Número de linhas deletadas pelo PR;
- comments count: Número de comentários no PR;
- participants count: Número de participantes do PR;
- pr state: Status do PR (MERGED ou CLOSED).

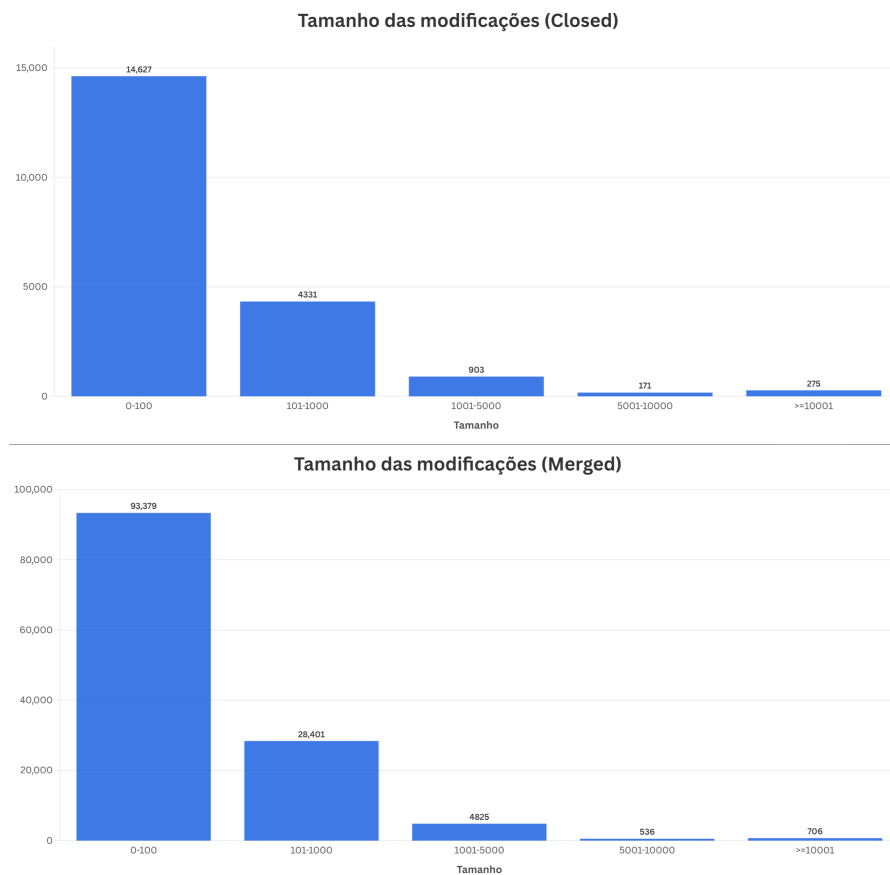
3. Resultados Obtidos

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

A análise do tamanho dos Pull Requests (PRs), considerando o número total de linhas adicionadas e removidas, revelou as seguintes medianas por status:

- PRs **CLOSED** (rejeitados) apresentaram uma mediana de **24** linhas modificadas.
- PRs **MERGED** (aprovados) apresentaram uma mediana de **26** linhas modificadas.

Esses resultados mostram que a maioria dos PRs, tanto aprovados quanto rejeitados, é relativamente pequena em tamanho. Apesar da média dos PRs rejeitados ser maior (indicando a presença de alguns PRs muito grandes entre os rejeitados), a mediana semelhante sugere que, **na maioria dos casos, o tamanho do PR não é o principal fator determinante para o seu feedback final**. Em outras palavras, **PRs grandes podem ter mais riscos associados, mas a rejeição não é diretamente proporcional ao tamanho das mudanças**.

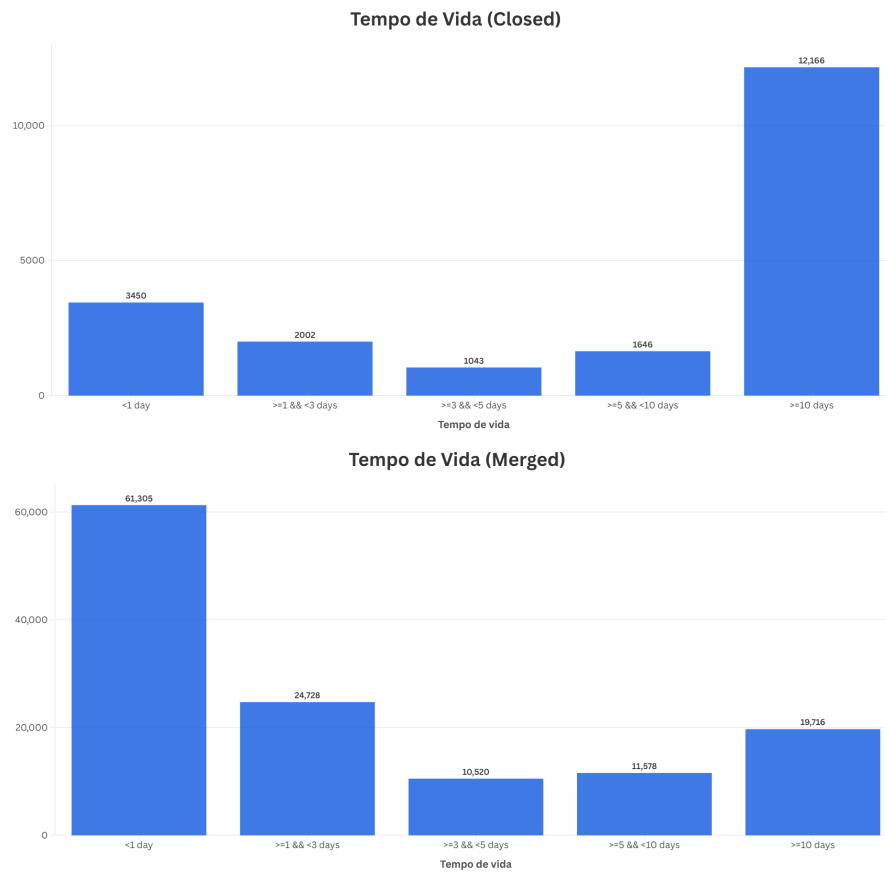


3.2. RQ 02: Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?

A análise do tempo de análise dos Pull Requests (PRs), medido em minutos entre a criação e a decisão final, revelou as seguintes medianas por status:

- PRs **CLOSED** (rejeitados) apresentaram uma mediana de **6.287** minutos (cerca de 4 dias e 9 horas).
- PRs **MERGED** (aprovados) apresentaram uma mediana de **2.553** minutos (cerca de 1 dia e 18 horas).

Esses resultados indicam que PRs rejeitados tendem a demorar mais tempo para serem analisados em comparação com os PRs aprovados. Isso sugere que, **quando um PR leva muito tempo para ser revisado, é mais provável que existam dúvidas, divergências ou problemas que dificultam sua aprovação**. Assim, um tempo de análise maior pode estar associado a um feedback final mais negativo.

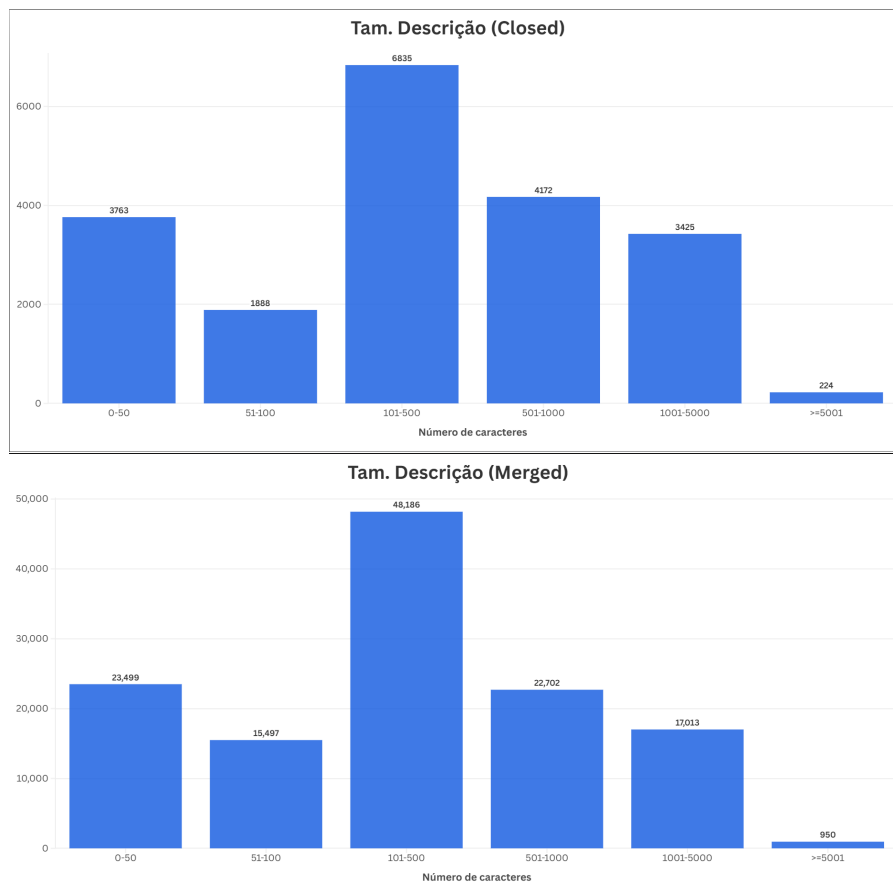


3.3. RQ 03: Qual a relação entre a descrição dos PRs e o feedback final das revisões?

A análise do tamanho das descrições dos Pull Requests (PRs), considerando o número de caracteres na descrição, revelou as seguintes medianas por status:

- PRs **CLOSED** (rejeitados) apresentaram uma mediana de **223** caracteres na descrição.
- PRs **MERGED** (aprovados) apresentaram uma mediana de **319** caracteres na descrição.

Os dados mostram que PRs aprovados tendem a ter descrições mais longas e detalhadas do que PRs rejeitados. Isso sugere que **descrições mais completas facilitam o entendimento das mudanças propostas e aumentam as chances de aprovação**. Por outro lado, PRs com descrições curtas ou inexistentes podem gerar dúvidas nos revisores, o que pode contribuir para a sua rejeição.

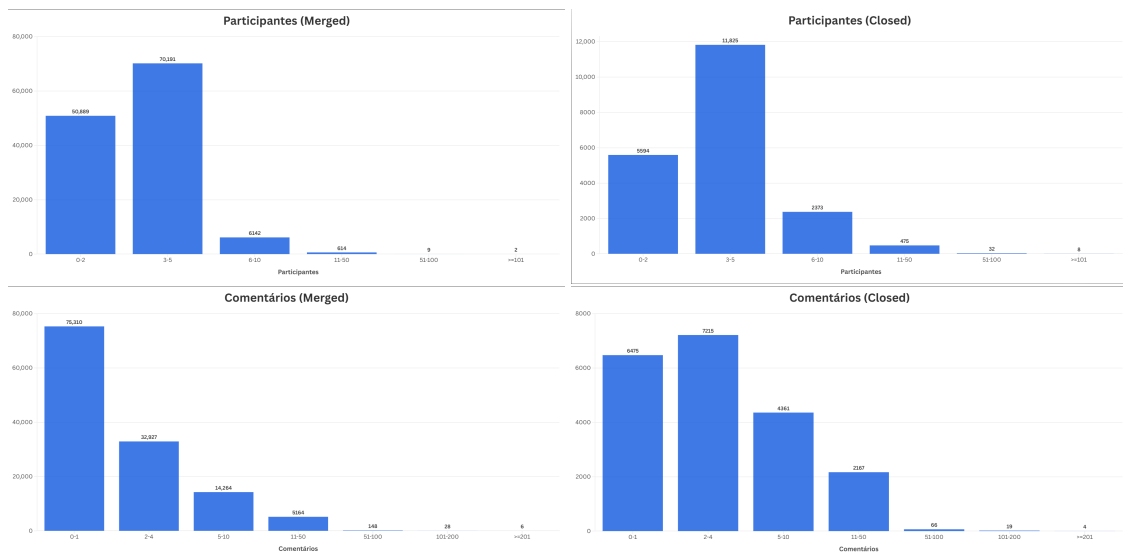


3.4. RQ 04: Qual a relação entre as interações nos PRs e o feedback final das revisões?

A análise das interações nos Pull Requests (PRs), considerando o número de participantes e o número de comentários, revelou as seguintes medianas por status:

- PRs **CLOSED** (rejeitados) por participantes apresentaram uma mediana igual a **3**.
- PRs **MERGED** (aprovados) por participantes apresentaram uma mediana igual a **3**.
- PRs **CLOSED** (rejeitados) por comentários apresentaram uma mediana igual a **3**.
- PRs **MERGED** (aprovados) por comentários apresentaram uma mediana igual a **1**.

Os dados mostram que PRs aprovados, tendem a ter poucos comentários (50% da amostra com 1 único comentários) e poucas pessoas trabalhando nele (30% com até 2 pessoas e 80% com até 5 participantes).

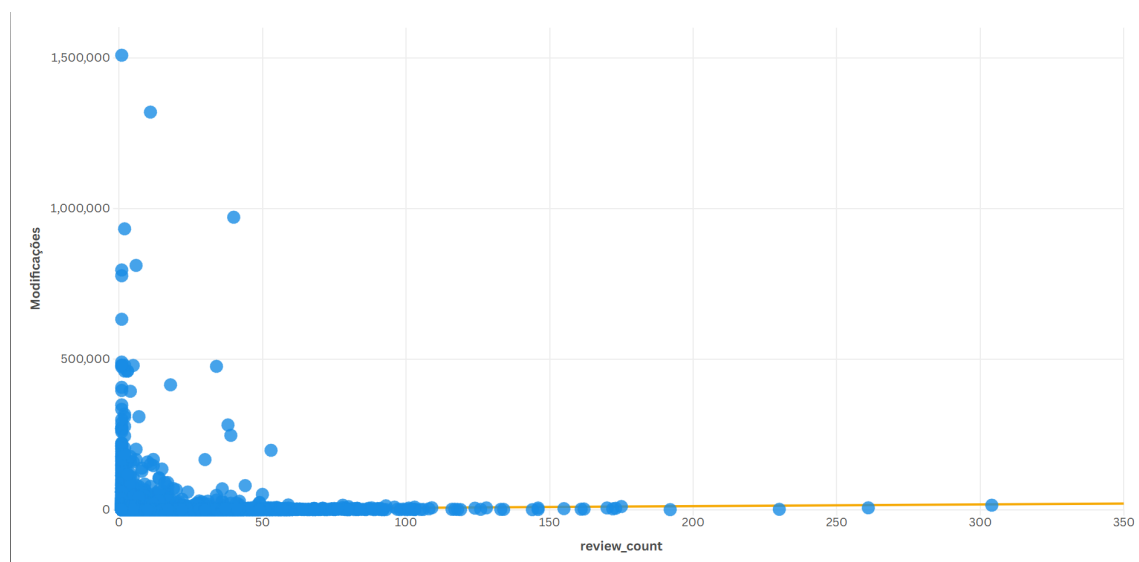


3.5. RQ 05: Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?

A análise do tamanho (número de modificações) dos Pull Requests (PRs), considerando as informações adicionadas e deletadas, em comparação ao número de reviews do PR, revelou o seguinte:

- A Mediana (independente do STATUS) foi de **25** itens modificados.
- A Correlação entre os 2 dados (modificações X reviews) foi de $\simeq 0,030$.

Com o valor da Correlação, é possível visualizar que o número de itens modificados por PR não tem tanta influência no número de reviews destes mesmos PRs.



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

A análise do tempo de vida dos Pull Requests (PRs), considerando a diferença entre a data de fechamento e a data de criação, em comparação ao número de reviews do PR, revelou o seguinte:

- A Mediana (independente do STATUS) foi de **1dia e 11horas** de tempo de vida dos PRs.
- A Correlação entre os 2 dados (tempo de vida X reviews) foi de $\simeq 0,050$.

Com o valor da Correlação, foi possível identificar que estes dois dados (tempo de vida e número de reviews) não afetam uns aos outros. Com a mediana, ficou fácil de identificar que os PRs são analisados e finalizados em um tempo relativamente rápido (não passando de 2 dias na maioria dos casos).

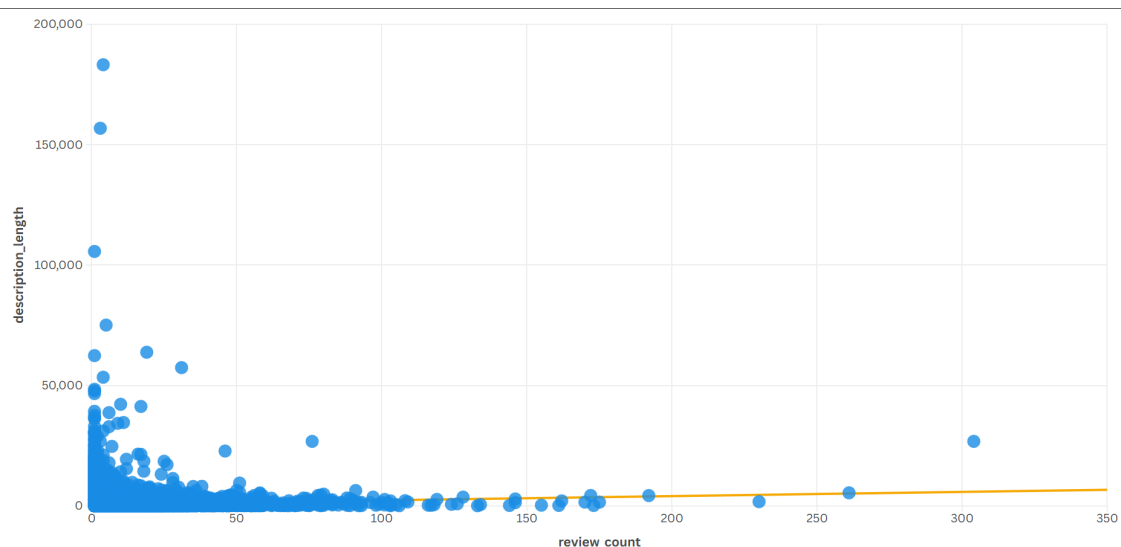
Por conta do tipo do dado gerado pela diferença entre as duas datas (duração em segundos, sendo então formatada para "dia.Horas:Minutos:Segundos") não foi possível criar uma visualização gráfica que não apresentasse dados ou muito grandes ou uma representação distorcida dos dados formatados.

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

A análise do tamanho das descrições dos Pull Requests (PRs), considerando o número de caracteres na descrição, em comparação ao número de reviews do PR, revelou o seguinte:

- A Mediana (independente do STATUS) foi de **253** caracteres na descrição.
- A Correlação entre os 2 dados (caracteres na descrição X reviews) foi de $\simeq 0,070$.

A análise da Correlação demonstra pouca relação entre os dois tipos de dados observados, porém, igual ao definido na RQ03, uma descrição com um certo número de caracteres é ideal para poder descrever o que está sendo feito no PR, assim não requerer desencadeando tantos reviews.



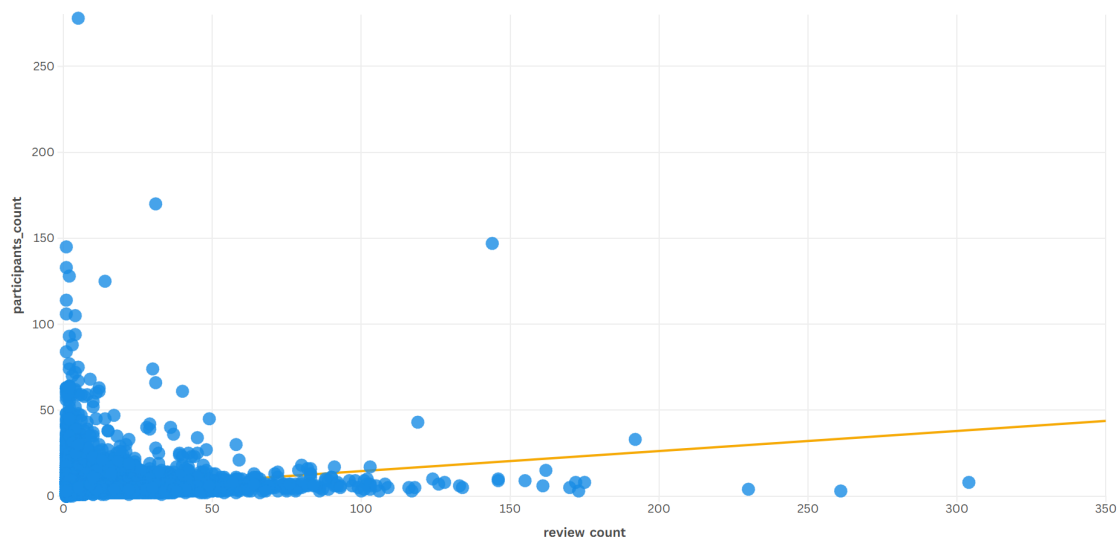
3.8. RQ 08: Qual a relação entre as interações nos PRs e o número de revisões realizadas?

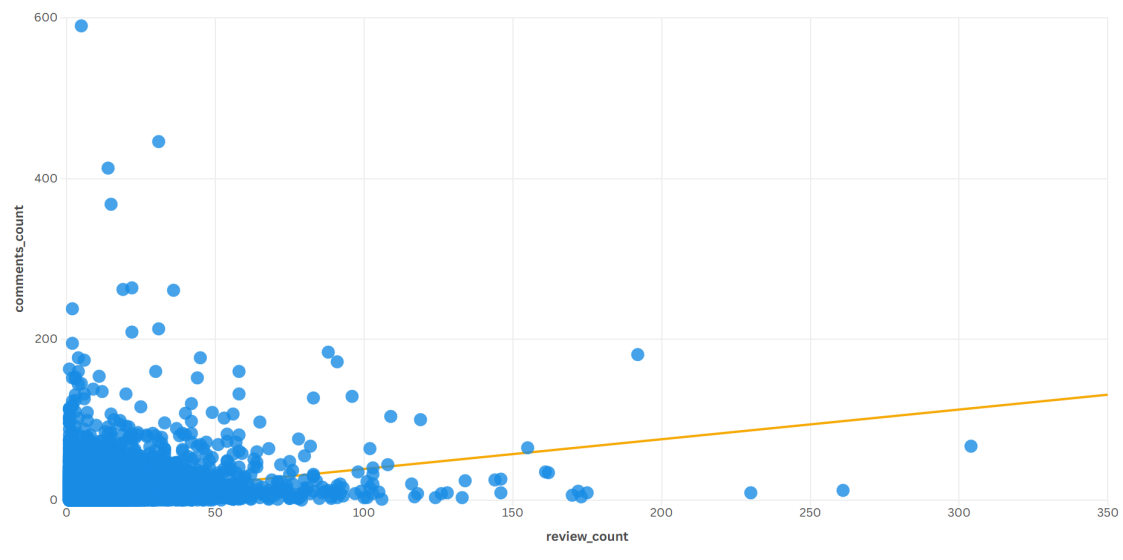
A análise das interações nos Pull Requests (PRs), considerando o número de participantes e o número de comentários, em comparação ao número de reviews do PR, revelou o seguinte:

- A Mediana (independente do STATUS), do nº de participantes, foi de **3** por PRs.
- A Mediana (independente do STATUS), do nº de comentários, foi de **1** por PR.
- A Correlação entre os 2 dados (nº de participantes X reviews) foi de $\simeq 0,247$.
- A Correlação entre os 2 dados (nº de comentários X reviews) foi de $\simeq 0,315$.

A análise das Correlações foram as seguintes:

- Os dados nº de participantes e nº de reviews tem uma boa correlação entre eles. Foi observado que, geralmente, quanto maior o número de participantes no PR menor é o nº de reviews deste PR antes de este ser aceito. Isto se deve ao fato que as modificações no código foram feitas e validadas por diferentes indivíduos antes do PR ser iniciado, assim, diminuindo a chance de erros (consequentemente, diminuindo a necessidade de review).
- Os dados nº de comentários e nº de reviews tem uma boa correlação entre si. Foi observado que, geralmente, quanto maior o número de comentários no PR menor é o nº de reviews deste PR antes de este ser aceito. Isso se deve ao fato que foi realizada uma discussão entre diversos indivíduos neste PR resultando no esclarecimento de certas dúvidas, mal-entendidos ou até a identificação de erros, assim, gerando uma baixa necessidade para reviews no mesmo PR.





4. Conclusão

Como conclusão deste trabalho tiramos que atualmente os Pull Requestes (PRs) são bem regularizados, passando por algumas etapas e pela visão de diferentes pessoas antes de ter seu "destino" definido (ser aceito ou recusado; MERGED ou CLOSED). É bem normal que os PRs tenham bastantes comentários, participantes e uma descrição bem completa e, apesar de serem definidos em um tempo relativamente rápido, têm tremenda importância no processo de desenvolvimento e recebem a devida atenção.

Por este trabalho, ficou evidente para o grupo que a utilização de PR e, consequentemente, sua análise, é algo essencial para todos repositórios populares; uma vez que PR é uma forma de se dividir o trabalho no sistema para outras pessoas e, sua análise, é algo essencial para manter a qualidade do sistema e prevenir o surgimento de novos erros.