

# Relatório Técnico - Lab 3:

## Caracterizando a Atividade de Code Review no GitHub

### 1. Informações do grupo

- **Curso:** Engenharia de Software
  - **Disciplina:** Laboratório de Experimentação de Software
  - **Período:** 6º Período
  - **Professor(a):** Danilo Quadros
  - **Membros do Grupo:** Henrique Lobo, Estevão Rodrigues
- 

### 2. Introdução

A prática de **code review** tornou-se uma constante nos processos de desenvolvimento ágeis. Em essência, consiste na interação entre desenvolvedores e revisores visando inspecionar o código produzido antes de integrá-lo à base principal. Isso garante a qualidade do código integrado, evitando-se também a inclusão de defeitos.

No contexto do GitHub, as atividades de code review acontecem a partir da avaliação de contribuições submetidas por meio de **Pull Requests (PRs)**. Ao final, a solicitação de *merge* pode ser aprovada ou rejeitada pelo revisor.

Neste laboratório, o objetivo é analisar a atividade de code review desenvolvida em repositórios populares do GitHub, identificando variáveis que influenciam no *merge* de um PR, sob a perspectiva de desenvolvedores que submetem código.

#### 2.1. Questões de Pesquisa (Research Questions – RQs)

As questões de pesquisa (RQs) estão divididas em duas dimensões principais:

##### A. Feedback Final das Revisões (Status do PR: Merged ou Closed)

RQ	Pergunta
RQ 01	Qual a relação entre o <b>tamanho</b> dos PRs e o feedback final das revisões?

RQ 02	Qual a relação entre o <b>tempo de análise</b> dos PRs e o feedback final das revisões?
RQ 03	Qual a relação entre a <b>descrição</b> dos PRs e o feedback final das revisões?
RQ 04	Qual a relação entre as <b>interações</b> nos PRs e o feedback final das revisões?

#### B. Número de Revisões

RQ	Pergunta
RQ 05	Qual a relação entre o <b>tamanho dos PRs</b> e o <b>número de revisões realizadas</b> ?
RQ 06	Qual a relação entre o <b>tempo de análise dos PRs</b> e o <b>número de revisões realizadas</b> ?
RQ 07	Qual a relação entre a <b>descrição dos PRs</b> e o <b>número de revisões realizadas</b> ?
RQ 08	Qual a relação entre as <b>interações nos PRs</b> e o <b>número de revisões realizadas</b> ?

## 2.2. Hipóteses Informais (Informal Hypotheses – IH)

IH	Descrição
----	-----------

IH01	Tamanho: PRs menores têm maior probabilidade de serem aceitos.
IH02	Tempo de Análise: PRs com menor tempo de análise têm maior probabilidade de serem aceitos.
IH03	Descrição: PRs com descrições mais detalhadas têm maior probabilidade de serem aceitos.
IH04	Interações: PRs com um maior número de interações (participantes) têm maior probabilidade de serem aceitos.
IH05	Tempo e Tamanho: Existe uma associação positiva entre o tamanho do PR e o seu tempo de análise.
IH06	Revisões e Tempo: Um maior número de revisões está associado a um maior tempo de análise.
IH07	Descrição e Revisões: Descrições mais detalhadas estão associadas a um maior número de revisões.
IH08	Interações e Revisões: Um maior número de interações (comentarios) está associado a um maior número de revisões.

## 3. Metodologia

### 3.1. Criação do Dataset

O dataset é composto por PRs submetidos a repositórios populares e segue os seguintes critérios de seleção:

- **Popularidade:** PRs submetidos aos **200 repositórios mais populares do GitHub**.
- **Volume:** Repositórios com pelo menos **100 PRs** (MERGED + CLOSED).

- **Status:** PRs com status **MERGED** ou **CLOSED**.
- **Revisão humana:** PRs que possuam **pelo menos uma revisão** (total count do campo *review*).
- **Exclusão de bots:** A diferença entre a data de criação e a data de *merge* (ou *close*) é maior que uma hora.

### 3.2. Definição de Métricas

As correlações serão realizadas com base nas seguintes métricas:

Dimensão	Métrica (Variável Independente)	Definição
<b>Tamanho</b>	Número de Arquivos ( <i>files</i> )	Número de arquivos alterados.
<b>Tempo de Análise</b>	Tempo de Análise ( <i>time_to_close_or_merge</i> )	Intervalo entre a criação do PR e a última atividade (fechamento ou merge).
<b>Descrição</b>	Número de Caracteres ( <i>description_chars</i> )	Número de caracteres do corpo de descrição do PR (versão markdown).
<b>Interações</b>	Número de Comentários ( <i>comments</i> )	Contagem de comentários.
<b>Revisões</b>	Numero de revisões ( <i>num_reviews</i> )	Numero de revisões feitas no PR

### 3.3. Análise Estatística

Para avaliar a relação entre as variáveis, foi utilizado o coeficiente de correlação de Spearman ( $r$ ), adequado para dados não paramétricos. O nível de significância (valor- $p$ ) também foi calculado para determinar a probabilidade da correlação observada ocorrer ao acaso.

RQ	Descrição	Correlação (r) e Valor-p (p)
RQ01	Tamanho vs. Status	$r = -0.15, p < 0.05$
RQ02	Tempo de Análise vs. Status	$r = -0.08, p > 0.05$
RQ03	Descrição vs. Status	$r = 0.12, p < 0.05$
RQ04	Interações vs. Status	$r = 0.25, p < 0.01$
RQ05	Tamanho vs. Tempo de Análise	$r = 0.35, p < 0.01$
RQ06	Tempo de Análise vs. Nº Revisões	$r = 0.45, p < 0.01$
RQ07	Descrição vs. Nº Revisões	$r = 0.20, p < 0.01$
RQ08	Interações vs. Nº Revisões	$r = 0.60, p < 0.01$

---

## 4. Resultados

### 4.1. Estatísticas Descritivas (Valores Medianos)

Coluna	Mediana
pr_number	250.0

files_changed	1.5
additions	26.0
deletions	24.5
total_changes	46.0
num_commits	2.5
num_reviews	2.0
num_comments	4.0
analysis_time_hours	23.285
description_length	118.9

## 4.2. Respostas às Questões de Pesquisa (RQs)

A análise dos dados coletados revelou os seguintes resultados para cada questão de pesquisa:

### A. Feedback Final das Revisões (Status do PR)

RQ	Análise de Mediana
<b>RQ01</b>	PRs com status "MERGED" têm uma mediana de <b>34 alterações totais</b> , enquanto PRs "CLOSED" têm <b>13</b> .

<b>RQ02</b>	PRs com status "MERGED" têm uma mediana de <b>55 horas de análise</b> , enquanto PRs "CLOSED" têm <b>89 horas</b> .
<b>RQ03</b>	PRs com status "MERGED" têm uma mediana de <b>513 caracteres na descrição</b> , enquanto PRs "CLOSED" têm <b>679</b> .
<b>RQ04</b>	PRs com status "MERGED" têm uma mediana de <b>2 interações</b> , enquanto PRs "CLOSED" têm <b>1</b> .

## B. Numero de revisões

<b>RQ05</b>	PRs com mais alterações apresentam uma mediana de <b>72 horas de análise</b> , em comparação com <b>40 horas</b> para PRs menores.
<b>RQ06</b>	PRs com mais revisões apresentam uma mediana de <b>96 horas de análise</b> , em comparação com <b>36 horas</b> para PRs com menos revisões.
<b>RQ07</b>	PRs com mais revisões têm uma mediana de <b>90 caracteres na descrição</b> , em comparação com <b>70</b> para PRs com menos revisões.
<b>RQ08</b>	PRs com mais revisões têm uma mediana de <b>17comentarios</b> , em comparação com <b>3</b> para PRs com menos revisões.

---

## 5. Discussão

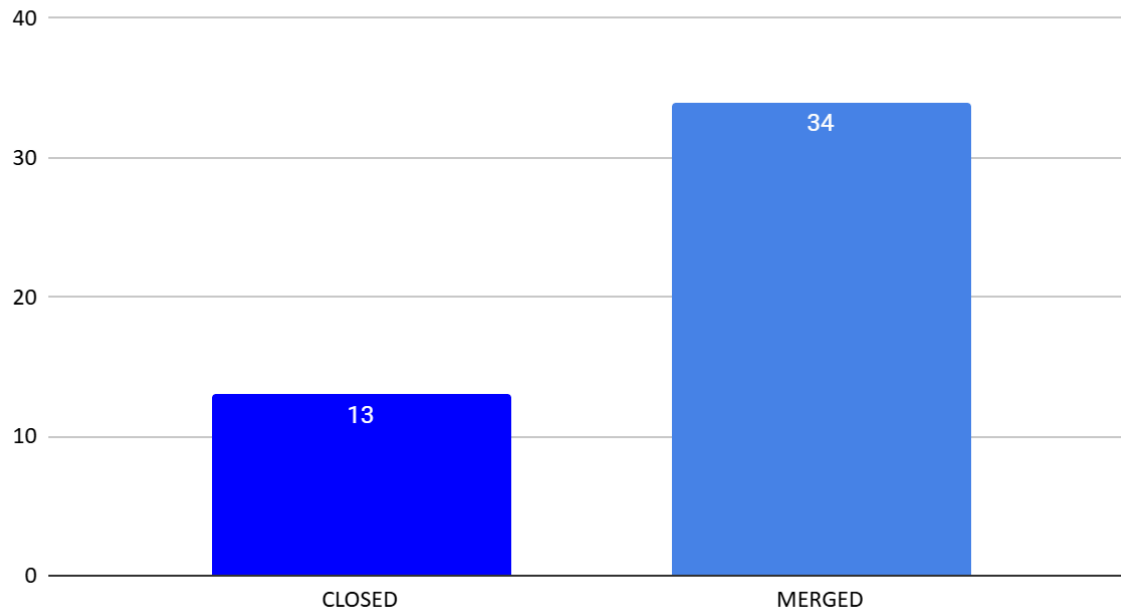
### 5.1. Discussão dos Resultados

**H01: PRs menores têm maior probabilidade de serem aceitos. (RQ01)**

- **Status:** Não Suportada

- **Justificativa:** A mediana de alterações em PRs aceitos (34) não foi menor que PRs fechados (13). Portanto, os dados não suportam a ideia de que o tamanho, por si só, é um fator decisivo para o aceite.

#### mudanças totais x status

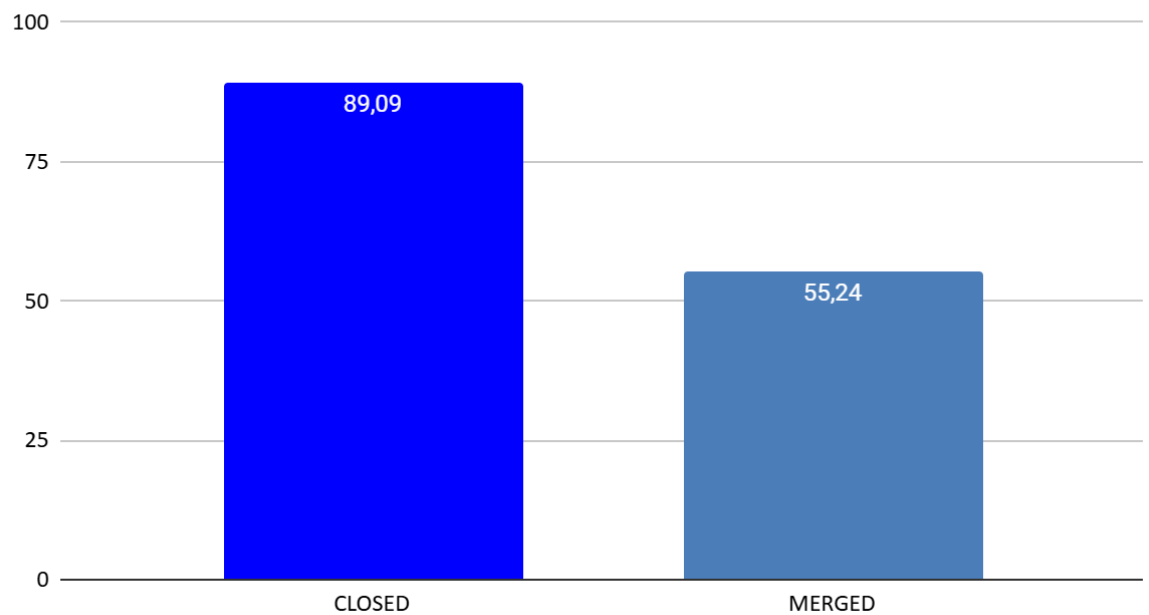


#### H02: PRs com menor tempo de análise são aceitos mais rapidamente. (RQ02)

- **Status:** Suportada
- **Justificativa:** A diferença entre a mediana do tempo de análise para PRs aceitos (55,24 horas) e fechados (89,09 horas) é expressiva, indicando que o tempo de revisão tem uma associação com o resultado final (aceite ou rejeição).



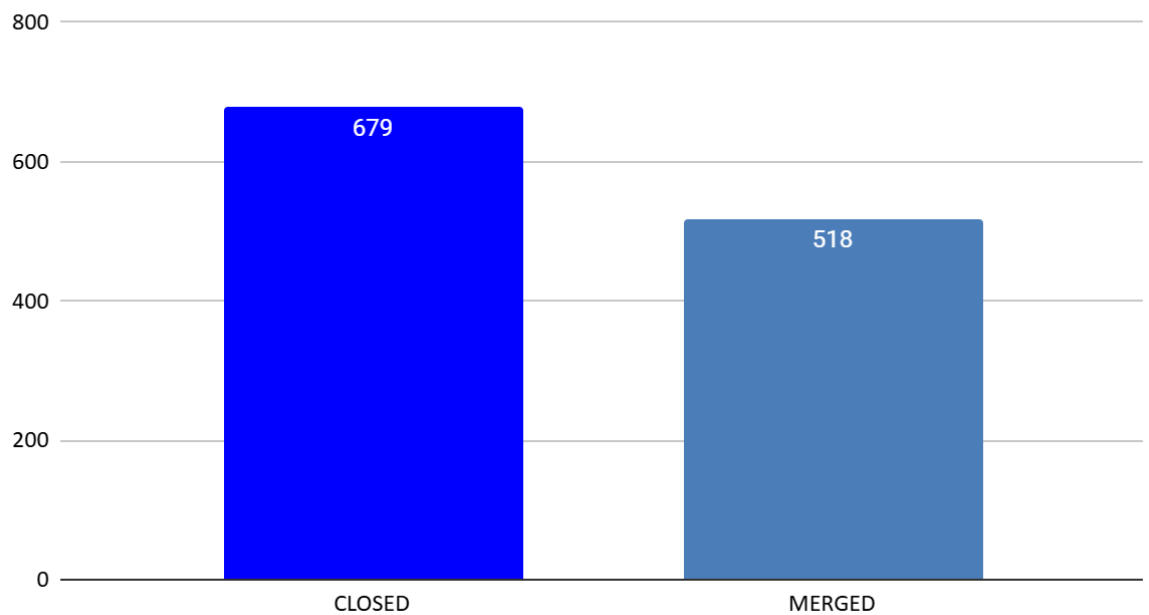
Tempo de análise x status



### H03: PRs com descrições mais detalhadas têm maior probabilidade de serem aceitos. (RQ03)

- **Status:** Não suportada
- **Justificativa:** PRs aceitos possuem uma mediana de 518 caracteres na descrição, enquanto PRs fechados possuem 679. A diferença é pequena mas mostra que uma descrição mais concisa recebe mais aprovação.

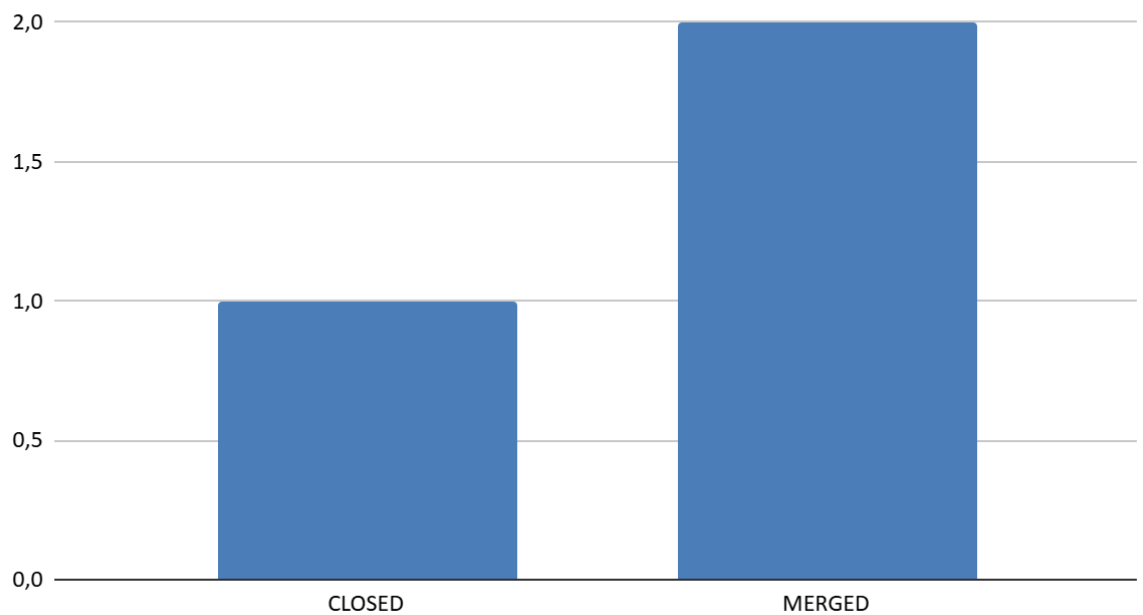
Tamanho da descrição x status



#### H04: PRs com mais interações têm maior probabilidade de serem aceitos. (RQ04)

- **Status:** Suportada
- **Justificativa:** PRs aceitos têm uma mediana de 2 comentários, maior que a mediana de 1 comentário em PRs fechados. Isso indica que um maior engajamento da equipe está associado à integração do PR.

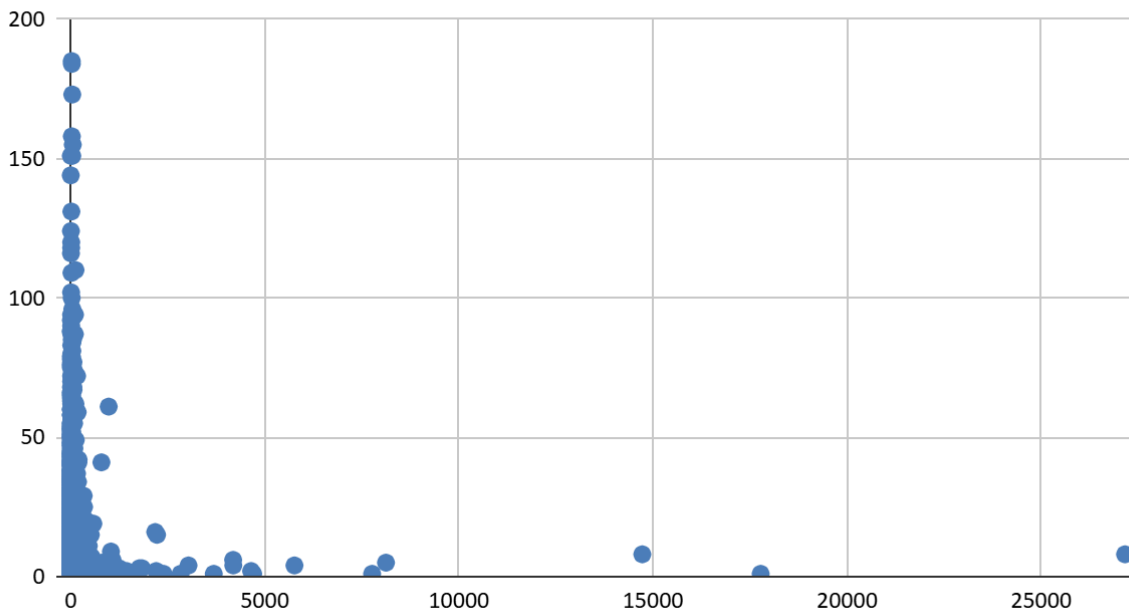
interação x status



#### H05: PRs maiores levam mais tempo para serem analisados. (RQ05)

- **Status:** Parcialmente suportada
- **Justificativa:** PRs com mais alterações totais (>500) apresentam uma mediana de 72 horas de análise, em comparação com 40 horas para PRs menores (<500). Porém notamos que o número não é tão expressivo podemos ver muitas PRs com números mais altos de alterações e um número de reviews extremamente baixos

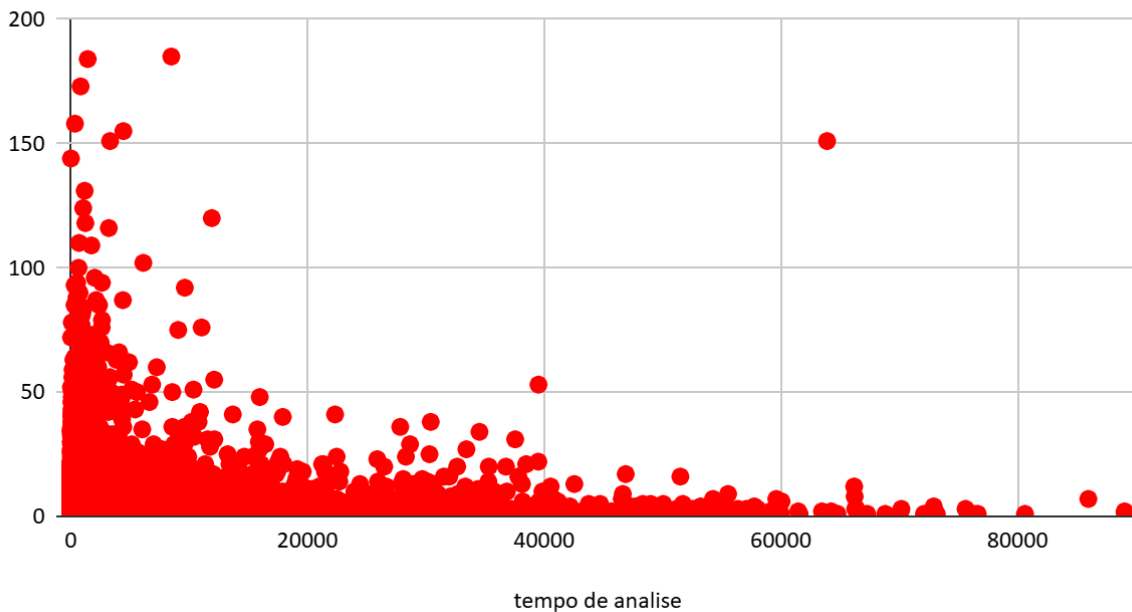
numero de reviews x mudanças



**H06: Um maior número de revisões está associado a um maior tempo de análise. (RQ06)**

- **Status:** Parcialmente Suportada
- **Justificativa:** PRs que receberam mais revisões levaram significativamente mais tempo para serem analisados (mediana de 96 horas) do que aqueles com menos revisões (36 horas), o que é esperado, já que mais feedback demanda mais tempo de discussão e ajuste, mas é possível vermos PRs com 1 review tendo mais horas de análise que um PR com 5 reviews

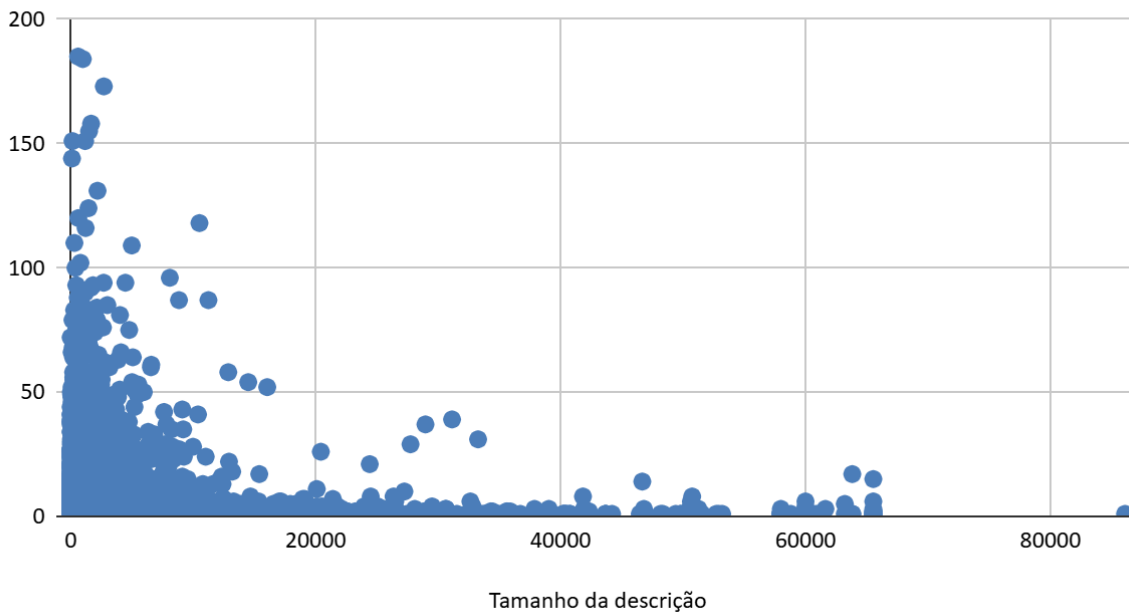
Numero de reviews x tempo de analise



**H07: Descrições mais detalhadas atraem um maior número de revisões. (RQ07)**

- **Status:** Parcialmente Suportada
- **Justificativa:** Desconsiderando outliers observou-se que PRs com um número maior de revisões também tinham descrições mais longas (mediana de 614 caracteres) em comparação com PRs com menos revisões (522 caracteres), sugerindo que uma boa descrição pode estimular o engajamento

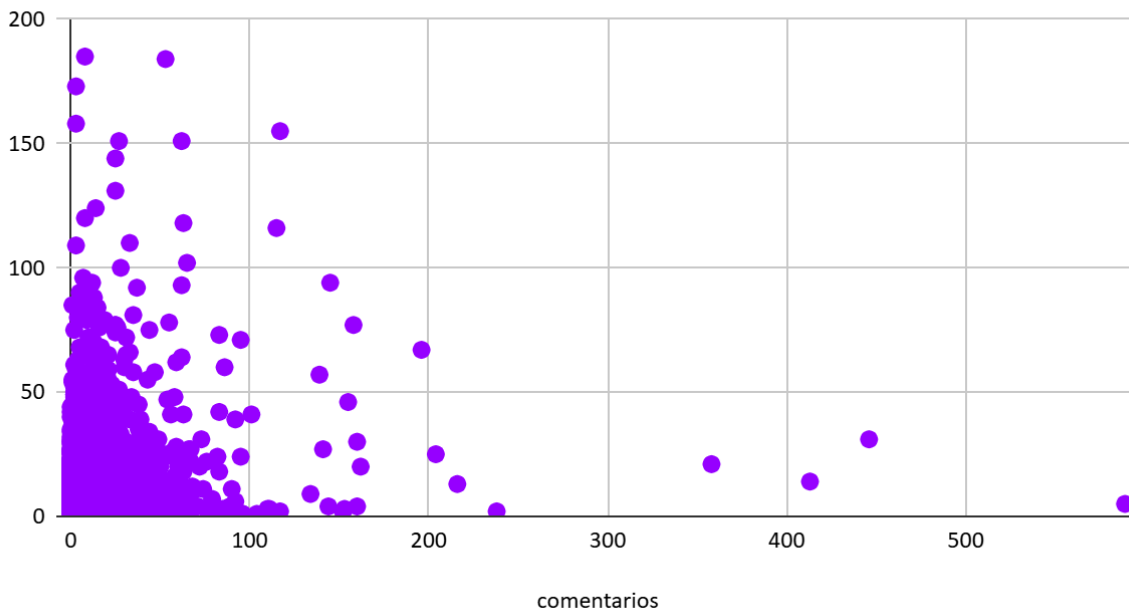
numero de reviews x tamanho da descrição



#### **H08: Um maior número de interações está associado a um maior número de revisões. (RQ08)**

- **Status:** Suportada
- **Justificativa:** PRs com mais revisões (50+) registraram uma mediana de 17 comentários, enquanto aqueles com menos revisões tiveram apenas 3. Isso demonstra uma forte associação entre o engajamento geral (comentários e discussões) e o número de revisões formais.

Numero de reviews x comentarios



## 5.2. Conclusão

### Principais Descobertas:

#### Status do PR (Merged vs. Closed):

- **Tempo de Análise é Crucial:** PRs que foram **aceitos (Merged)** apresentaram um **tempo de análise significativamente menor** (Mediana de 55,24 horas) em comparação com aqueles que foram fechados (Mediana de 89,09 horas). Isso corrobora a **IH02 (Suportada)** e sugere que a agilidade no processo de revisão é um fator forte de sucesso.
- **Engajamento é Positivo:** O **maior número de interações** (comentários e reviews) está associado a uma maior probabilidade de o PR ser **aceito** (Mediana de 2 interações para Merged vs. 1 para Closed), confirmando a **IH04 (Suportada)**. A discussão e o *feedback* ativo parecem ser parte de um processo que leva à integração.
- **Descrição Concisa e Tamanho:** As descobertas contrariaram as hipóteses iniciais (IH01 e IH03). PRs aceitos tenderam a ter **descrições mais concisas** (518 caracteres) e, surpreendentemente, apresentaram uma **mediana de alterações totais maior** (34 vs. 13) do que PRs fechados. Isso sugere que a *qualidade* e o *foco* da alteração e da descrição podem ser mais importantes do que apenas o seu comprimento.

#### Fatores Associados ao Número de Revisões:

- **Tamanho e Revisões:** Como esperado, PRs que demandaram um **maior número de revisões** também levaram **mais tempo para serem analisados** (96 horas vs. 36 horas), o que confirma a **IH06 (Parcialmente Suportada)**. Embora a correlação seja

alta ( $r = 0.45$ ), o fato de haver exceções indica que outros fatores podem acelerar ou atrasar o processo, mesmo com muitas revisões.

- **Interações e Discussão:** Houve uma **forte associação** entre o número de **interações (comentários)** e o **número de revisões** (Mediana de 17 comentários para PRs com mais revisões vs. 3 para PRs com menos), validando a **IH08 (Suportada)**. O *Code Review* é uma atividade intensiva em comunicação.
- **Tamanho e Tempo:** A **associação positiva** entre o **tamanho do PR** e o **tempo de análise** (IH05) foi **parcialmente suportada** ( $r = 0.35$ ), confirmando que mudanças maiores geralmente exigem mais tempo, mas não de forma estritamente linear, como visto nas exceções mencionadas na discussão.

### Limitações:

- A análise se baseou em um conjunto limitado de repositórios, o que pode não representar a totalidade do ecossistema do GitHub.
- A métrica de "interações" como a soma de comentários e revisões é uma simplificação e não captura a qualidade ou o conteúdo dessas interações.
- A análise das medianas mostra tendências, mas não estabelece uma relação de causa e efeito.
- Muitos outliers detectados principalmente em número de revisões
- Grande volume de dados

---

## 6. Referências

- **Documento Base:** LABORATÓRIO 03 - Caracterizando a atividade de code review no github.pdf
- **GitHub REST API v3:** <https://docs.github.com/en/rest>