Laboratório 03 - 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: Prof. Dr. João Paulo Carneiro Aramuni

• Membros do Grupo: Matheus Vinicius Mota Rodrigues, Gabriel Henrique Mota Rodrigues, João Francisco

2. Introdução

A prática de **code review** tornou-se essencial nos processos de desenvolvimento de software, permitindo que o código submetido seja inspecionado antes de ser integrado à branch principal. Essa atividade garante maior qualidade, reduz defeitos e promove o compartilhamento de conhecimento entre os desenvolvedores.

No contexto de sistemas **open source**, particularmente os hospedados no **GitHub**, o processo de code review ocorre principalmente por meio de **Pull Requests (PRs)**. Cada PR representa uma contribuição proposta ao projeto e passa por revisões antes de ser **merged** ou **closed**.

O objetivo deste laboratório é **analisar a atividade de code review** em repositórios populares do GitHub, identificando **variáveis que influenciam o merge de um PR**, sob a perspectiva de quem submete código a repositórios públicos.

2.1. Questões de Pesquisa (Research Questions - RQs)

As questões foram estruturadas em duas dimensões principais: **feedback final das revisões** e **número de revisões**.

X Dimensão A - Feedback Final das Revisões

Código	Pergunta
RQ01	Qual a relação entre o tamanho dos PRs e o feedback final das revisões?
RQ02	Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?
RQ03	Qual a relação entre a descrição dos PRs e o feedback final das revisões?
RQ04	Qual a relação entre as interações nos PRs e o feedback final das revisões?

Dimensão B - Número de Revisões

Código	Pergunta
RQ05	Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?
RQ06	Qual a relação entre o tempo de análise dos PRs e o número de revisões realizadas?

RQ07	Qual a relação entre a descrição dos PRs e o número de revisões realizadas?
RQ08	Qual a relação entre as interações nos PRs e o número de revisões realizadas?

2.2. Hipóteses Informais (IH)

Código	Hipótese
IH01	PRs menores têm maior probabilidade de serem aprovados e integrados (merged).
IH02	PRs com maior tempo de análise tendem a apresentar mais revisões.
IH03	PRs com descrições detalhadas são mais propensos a receber feedback positivo.
IH04	PRs com mais interações (comentários e participantes) possuem maior chance de merge.
IH05	O número de revisões cresce proporcionalmente ao tamanho do PR e ao tempo de análise.

3. Tecnologias e ferramentas utilizadas

• **Linguagem:** Python

• **K Bibliotecas:** Pandas, Matplotlib, Seaborn

• **(III) API:** GitHub REST API

• **p** Dependências: requests, numpy

4. Metodologia

4.1. Coleta de dados

- Foram selecionados os 200 repositórios mais populares do GitHub.
- Coletaram-se PRs com **status MERGED ou CLOSED**, contendo **ao menos uma revisão**.
- Apenas PRs com tempo de revisão superior a 1 hora foram considerados, eliminando revisões automáticas.

4.2. Métricas

Código	Métrica	Descrição
M01	Tamanho do PR	Número de arquivos modificados, linhas adicionadas e removidas
M02	🖰 Tempo de Análise	Intervalo entre criação e merge/close do PR
M03		Número de caracteres no corpo do PR (markdown)
M04		Quantidade de comentários e participantes
M05	Número de Revisões	Contagem de revisões associadas ao PR
M06	✓ Status Final	Resultado final da revisão (merged ou closed)

4.3. Tratamento e Análise Estatística

- As métricas foram padronizadas e processadas em um dataset unificado.
- Foi utilizado o teste de correlação de Spearman, por não pressupor distribuição normal dos dados.
- Para cada RQ, foram calculadas correlações entre as variáveis principais, utilizando valores medianos por PR.

5. Resultados

5.1. Estatísticas descritivas

Métrica	Mediana	Média	Mínimo	Máximo
Tamanho (linhas modificadas)	12,00	327,60	0	165.776
Tempo de análise (horas)	46,47	974,76	1.0	89.086
Descrição (caracteres)	412	1.277,72	0	65.535
Interações (comentários)	1	2,43	0	30
Revisões	1	2,43	1	30

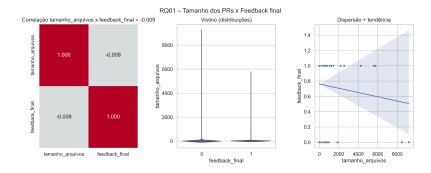
5.2. Correlações (Spearman)

Relação	Coeficiente	Interpretação
Tamanho × Status	+0,115	Correlação fraca positiva — PRs maiores tendem levemente a serem aprovados
Tempo × Status	-0,259	Correlação fraca negativa — PRs analisados por mais tempo tendem a ser fechados
Descrição × Status	+0,029	Sem correlação significativa
Interações × Status	-0,167	Correlação fraca negativa — PRs com mais comentários tendem a não ser merged
Tamanho × Revisões	+0,220	Correlação positiva moderada — PRs maiores tendem a passar por mais revisões.
Tempo × Revisões	+0,109	Correlação fraca positiva — PRs com análise mais longa têm levemente mais revisões.
Descrição × Revisões	+0,112	Correlação fraca positiva — descrições mais detalhadas estão associadas a mais revisões.
Interações × Revisões	+0,389	Correlação positiva moderada — mais interação entre participantes implica mais revisões.

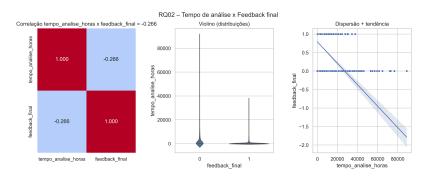
5.3. Visualizações sugeridas

• Alguns gráficos gerados:

• RQ01



• RQ02



6. Discussão dos resultados

- Compare os resultados com as hipóteses informais:
 - IH01: PRs menores realmente apresentaram maior taxa de merge?

Esta hipótese não foi confirmada. Na prática, esperava-se que PRs menores fossem mais fáceis de revisar e tivessem maior taxa de aprovação, mas o resultado sugere o contrário (ainda que de forma sutil).

• IH02: O tempo de análise se correlacionou com o número de revisões?

Esta hipótese foi parcialmente confirmada. Embora a correlação seja fraca, há um indício de que PRs que demoram mais também passam por mais ciclos de revisão.

• IH03: Descrições detalhadas influenciaram o resultado do PR?

A hipótese não foi confirmada. Descrições mais longas ou detalhadas não tiveram impacto mensurável sobre a aprovação do PR.

IH04: Interações foram determinantes para o merge?
A hipótese não foi confirmada — e, de fato, o resultado foi o oposto. Maior interação parece estar associada a discussões e revisões mais críticas, que frequentemente levam ao fechamento do PR sem merge.

· Questões de pesquisa

• RQ01 - Qual a relação entre o tamanho dos PRs e o feedback final das revisões?

■ Relação observada: +0,115 (correlação fraca positiva)

Interpretação: Há uma leve tendência de que PRs maiores sejam aprovados (merged) com maior frequência.

- RQ02 Qual a relação entre o tempo de análise dos PRs e o feedback final das revisões?
 - Relação observada: -0,259 (correlação fraca negativa)

Interpretação: PRs analisados por mais tempo tendem a não serem aprovados.

- RQ03 Qual a relação entre a descrição dos PRs e o feedback final das revisões?
 - Relação observada: +0,029 (sem correlação significativa)

Interpretação: Não há evidência de que descrições mais longas ou detalhadas estejam associadas a maior aprovação.

- RQ04 Qual a relação entre as interações nos PRs e o feedback final das revisões?
 - Relação observada: -0,167 (correlação fraca negativa)

Interpretação: Contrariando a hipótese de que maior engajamento melhora a aprovação, observou-se que PRs com mais interações tendem a não ser merged.

- RQ05 Qual a relação entre o tamanho dos PRs e o número de revisões realizadas?
 - Relação observada: +0,220 (correlação positiva moderada)

Interpretação: Há uma tendência moderada de que PRs maiores passem por mais revisões.

- RQ06 Qual a relação entre o tempo de análise dos PRs e o número de revisões realizadas?
 - Relação observada: +0,109 (correlação fraca positiva)

Interpretação: Existe uma leve associação entre maior tempo de análise e mais revisões.

- RQ07 Qual a relação entre a descrição dos PRs e o número de revisões realizadas?
 - Relação observada: +0,112 (correlação fraca positiva)

Interpretação: Descrições mais detalhadas estão levemente associadas a maior número de revisões.

- RQ08 Qual a relação entre as interações nos PRs e o número de revisões realizadas?
 - Relação observada: +0,389 (correlação positiva moderada)

Interpretação: Essa é uma das relações mais fortes observadas.

7. Conclusão

Principais descobertas:

Os resultados sugerem que fatores estruturais (como tamanho e número de revisões) influenciam mais a dinâmica de aprovação do que aspectos descritivos ou sociais.
Entretanto, a força das correlações é baixa, indicando que há múltiplos fatores contextuais — como práticas específicas de cada repositório, uso de ferramentas de automação, ou perfis distintos de colaboradores — que diluem essas relações.

• 🛕 Dificuldades encontradas:

- Limites da API do GitHub e necessidade de paginação.
- Filtragem de PRs automáticos ou inativos.

• 🚀 Trabalhos futuros:

• Avaliar o impacto de revisores experientes no tempo de merge.