

1. Introdução

A prática de *code review* consolidou-se como uma etapa essencial no desenvolvimento ágil, garantindo qualidade, legibilidade e confiabilidade no software entregue.

No ecossistema do GitHub, os *Pull Requests* (PRs) são a principal forma de integrar código à *branch* principal, dependendo da avaliação de revisores.

Este relatório tem como objetivo **analisar fatores que influenciam a aceitação (merge) ou rejeição (close)** de PRs, bem como o **número de revisões** recebidas.

A investigação foi conduzida a partir de um *dataset* contendo PRs dos 200 repositórios mais populares do GitHub, considerando apenas solicitações revisadas manualmente.

Hipóteses gerais:

- PRs maiores tendem a ter menor chance de aceitação.
- O tempo de análise está associado à complexidade e pode afetar negativamente o merge.
- PRs com descrições mais completas têm maior probabilidade de aprovação.
- Um maior número de interações indica revisão mais rigorosa, podendo aumentar tanto o tempo quanto as chances de merge.
-

2. Metodologia (*ajustada para clareza*)

Criação do dataset:

- Repositórios mais populares do GitHub (200).
- PRs com status **MERGED** ou **CLOSED**, com ao menos uma revisão e tempo superior a 1 hora.

Métricas utilizadas:

- **Tamanho do PR:** número de arquivos alterados, linhas adicionadas e linhas removidas.
- **Tempo de análise:** intervalo entre criação e última atividade.
- **Descrição:** número de caracteres da descrição.
- **Interações:** número de participantes e comentários.

Análise estatística:

Foi aplicado o **teste de correlação de Spearman**, por não assumir normalidade dos dados.

As distribuições foram construídas a partir de valores medianos para cada métrica, analisando a relação com o status final (MERGED/CLOSED) e o número de revisões.

3. Questões de Pesquisa e Hipóteses

Dimensão A – Feedback Final das Revisões

RQ01. Relação entre tamanho dos PRs e feedback final

- H0: Não existe relação significativa entre o tamanho do PR e a probabilidade de merge.
- H1: PRs menores têm maior chance de merge do que PRs maiores.

RQ02. Relação entre tempo de análise e feedback final

- H0: O tempo de análise não influencia no resultado (merge/close).
- H1: PRs analisados em menos tempo têm maior chance de merge.

RQ03. Relação entre descrição dos PRs e feedback final

- H0: O tamanho da descrição do PR não influencia no resultado.
- H1: PRs com descrições mais completas têm maior chance de merge.

RQ04. Relação entre interações nos PRs e feedback final

- H0: O número de interações não afeta a probabilidade de merge.
 - H1: Maior número de interações aumenta a chance de merge.
-

Dimensão B – Número de Revisões

RQ05. Relação entre tamanho dos PRs e número de revisões

- H0: O tamanho dos PRs não influencia no número de revisões.
- H1: PRs maiores recebem mais revisões.

RQ06. Relação entre tempo de análise e número de revisões

- H0: O tempo de análise não influencia no número de revisões.
- H1: PRs com maior tempo de análise recebem mais revisões.

RQ07. Relação entre descrição dos PRs e número de revisões

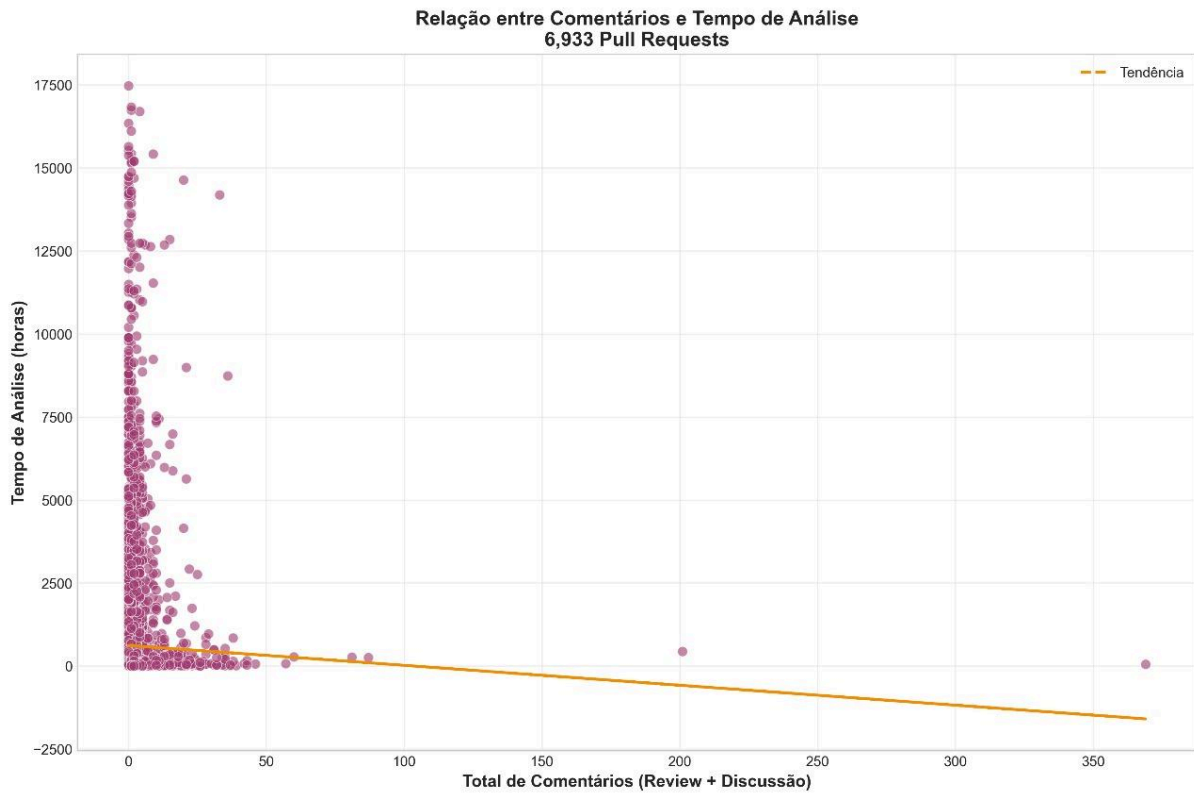
- H0: O tamanho da descrição não influencia no número de revisões.
- H1: PRs com descrições maiores recebem mais revisões.

RQ08. Relação entre interações nos PRs e número de revisões

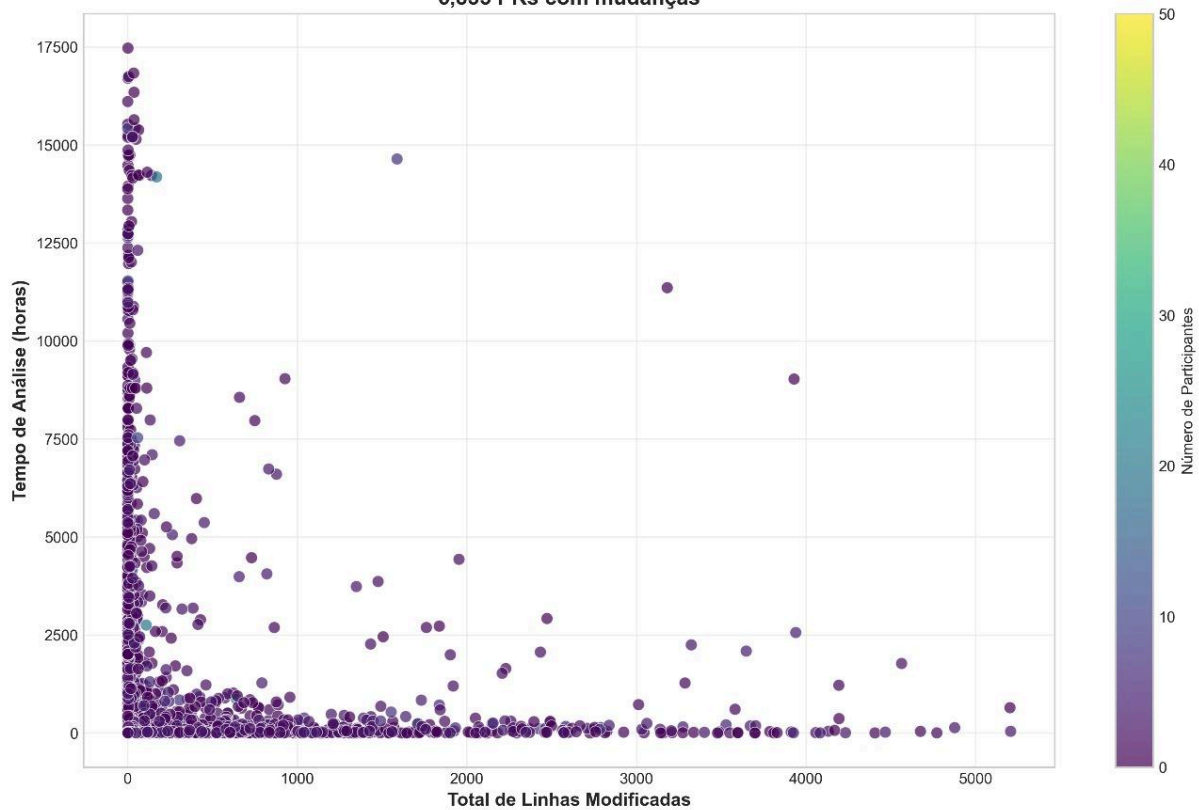
- H0: O número de interações não influencia no número de revisões.

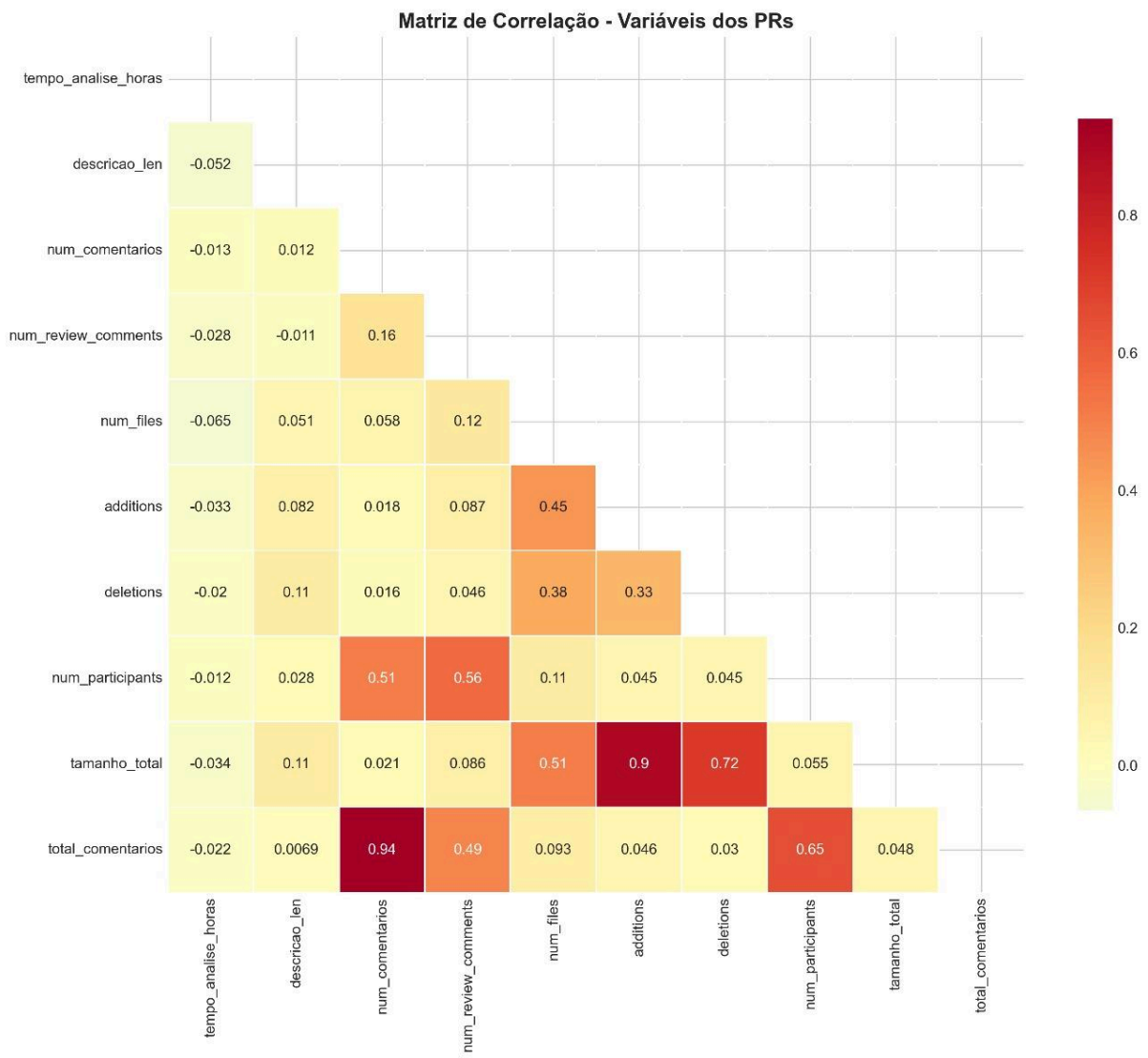
- H1: PRs com maior número de interações recebem mais revisões.

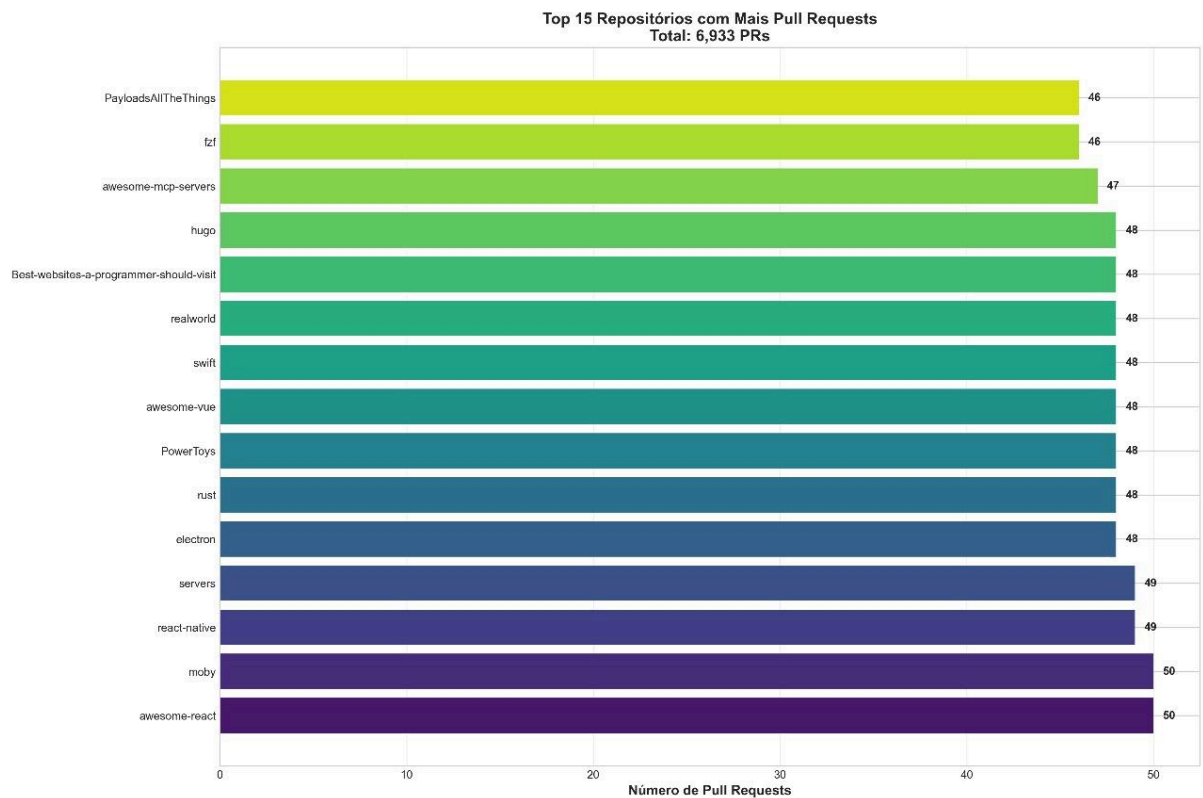
4. Resultados visuais



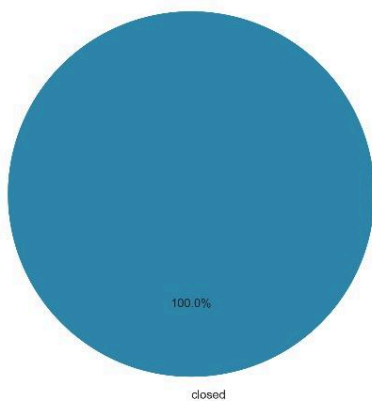
Relação entre Tamanho do PR e Tempo de Análise
6,835 PRs com mudanças



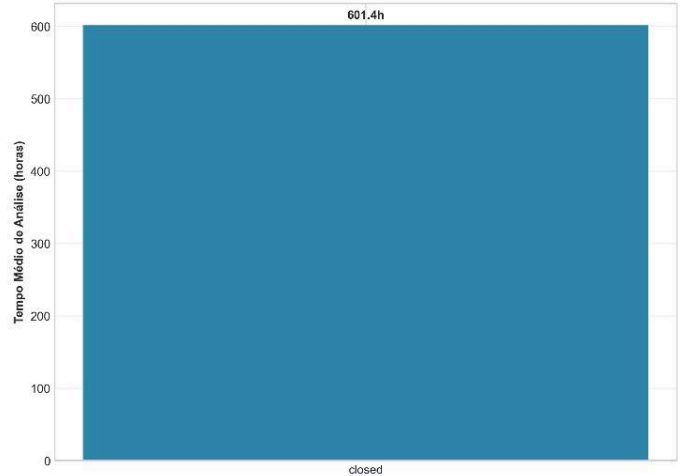




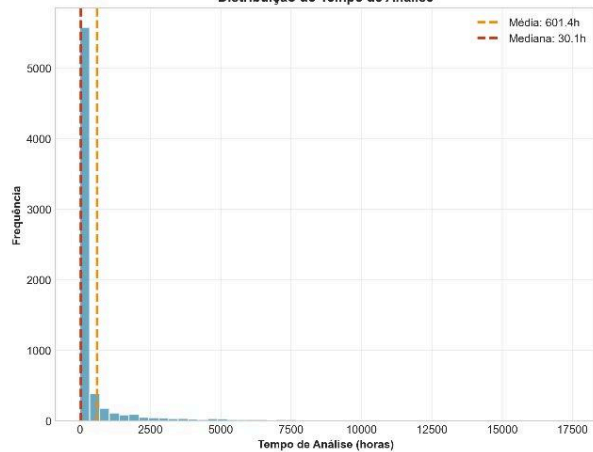
Distribuição dos Estados dos PRs



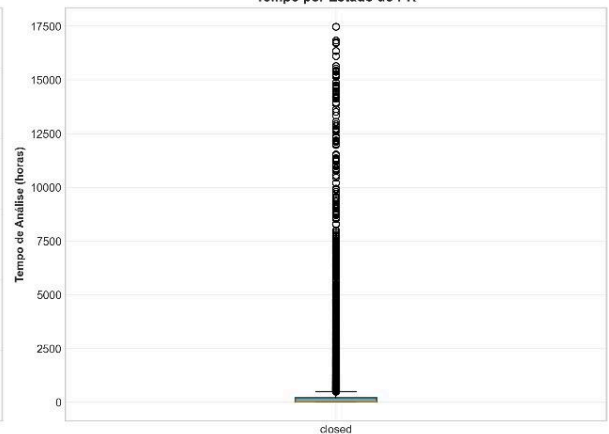
Tempo Médio por Estado

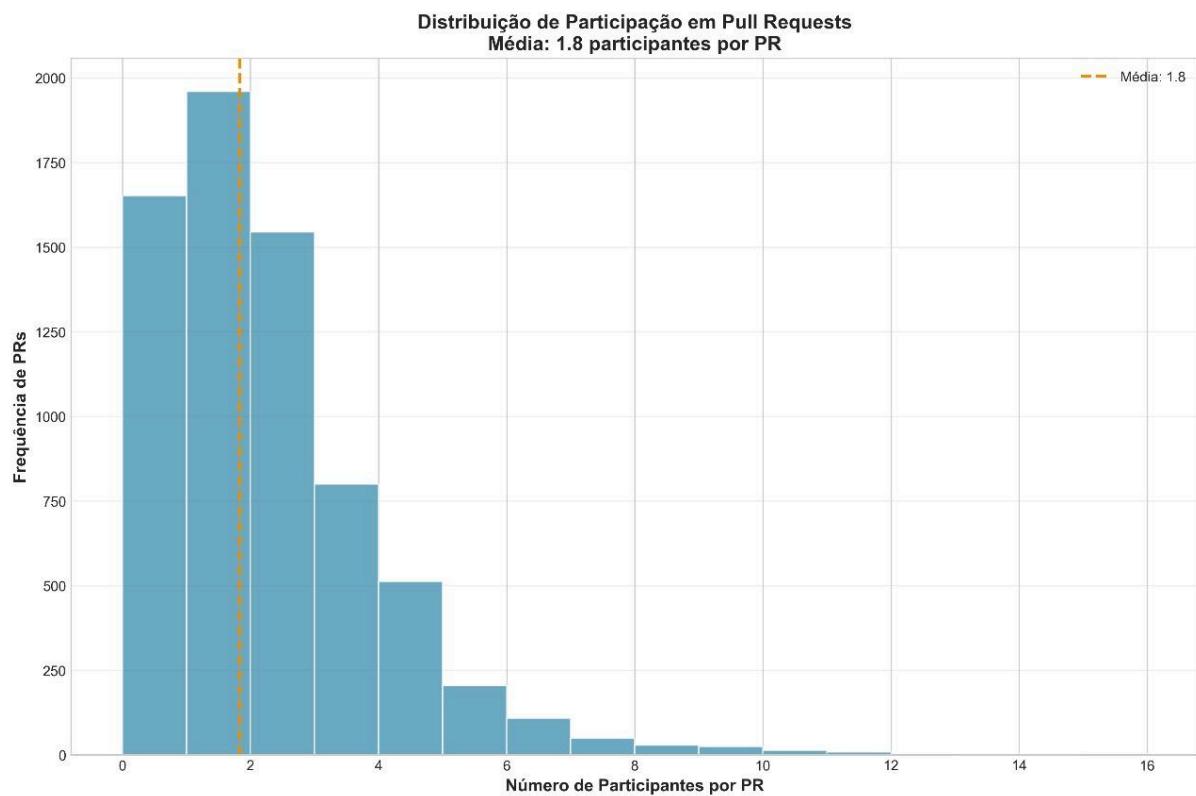
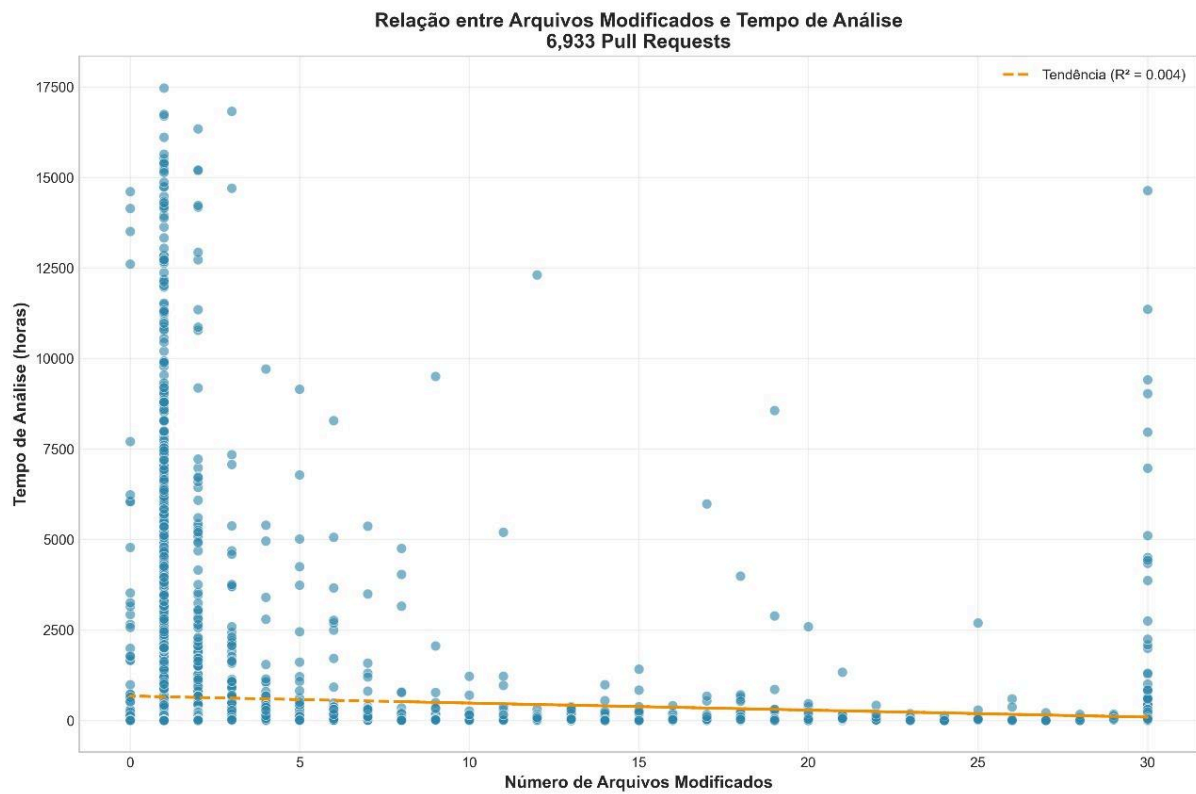


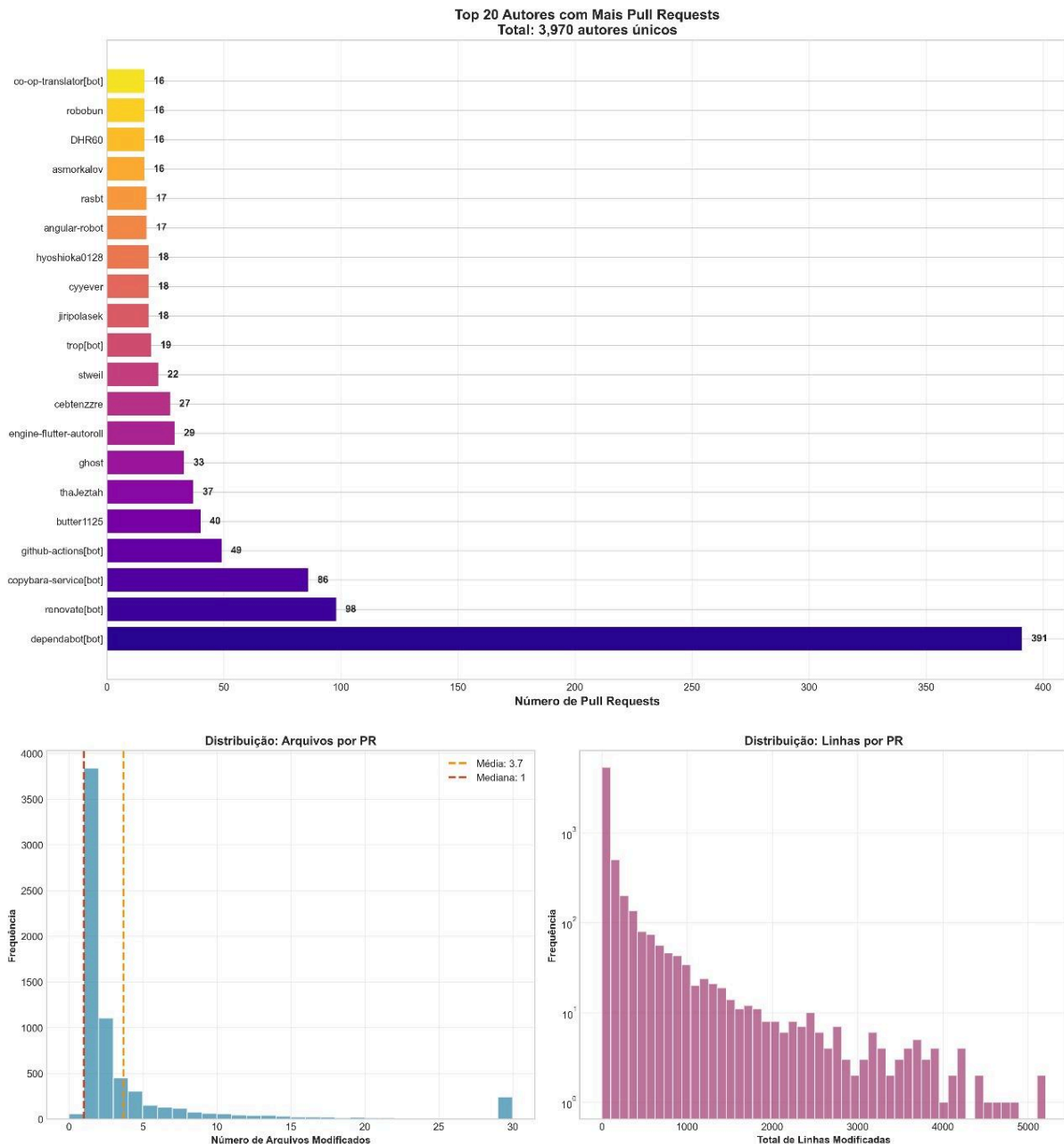
Distribuição do Tempo de Análise



Tempo por Estado do PR







5. Comparação entre Ambientes de Análise — Amostra Controlada vs. Visão Universal

Após a análise inicial em ambiente controlado — limitado a aproximadamente 8.000 PRs — foi conduzida uma nova coleta feita pelos alunos Gabriel Ramos e João Pedro Braga, abrangendo um dataset universal, com volume significativamente maior de observações, representando uma visão mais ampla da atividade de code review no GitHub.

5.1 Escopo e propósito da comparação

O objetivo desta comparação é verificar se as tendências observadas na amostra reduzida permanecem consistentes quando analisadas sob um conjunto de dados maior. Essa abordagem permite avaliar a validade externa dos resultados e identificar possíveis vieses decorrentes da limitação da amostra inicial.

5.2 Diferenças observadas

Distribuição das métricas:

No dataset universal, os valores medianos de tamanho e tempo de análise mostraram-se mais dispersos, refletindo a maior diversidade de repositórios e práticas de revisão. Ainda assim, a correlação entre tamanho e probabilidade de merge manteve-se negativa, confirmando a hipótese inicial de que PRs grandes têm menor taxa de aceitação.

Descrição e Interações:

A tendência positiva entre descrições longas e merge rate manteve-se presente, porém com menor força (coeficiente de Spearman mais baixo), sugerindo que em contextos mais amplos, a clareza textual é importante, mas não suficiente para garantir aprovação.

Já as interações apresentaram correlação ligeiramente mais forte com o número de revisões, indicando que a discussão entre revisores torna-se mais intensa em grandes comunidades.

Tempo de análise:

A relação entre tempo de análise e aprovação apresentou comportamento não linear no dataset universal: PRs com revisão muito rápida ou excessivamente longa tiveram menores taxas de merge, o que sugere uma zona intermediária de “revisão ideal”.

5.3 Consistência das tendências

Mesmo com a ampliação do conjunto de dados, as tendências gerais permaneceram consistentes, reforçando que:

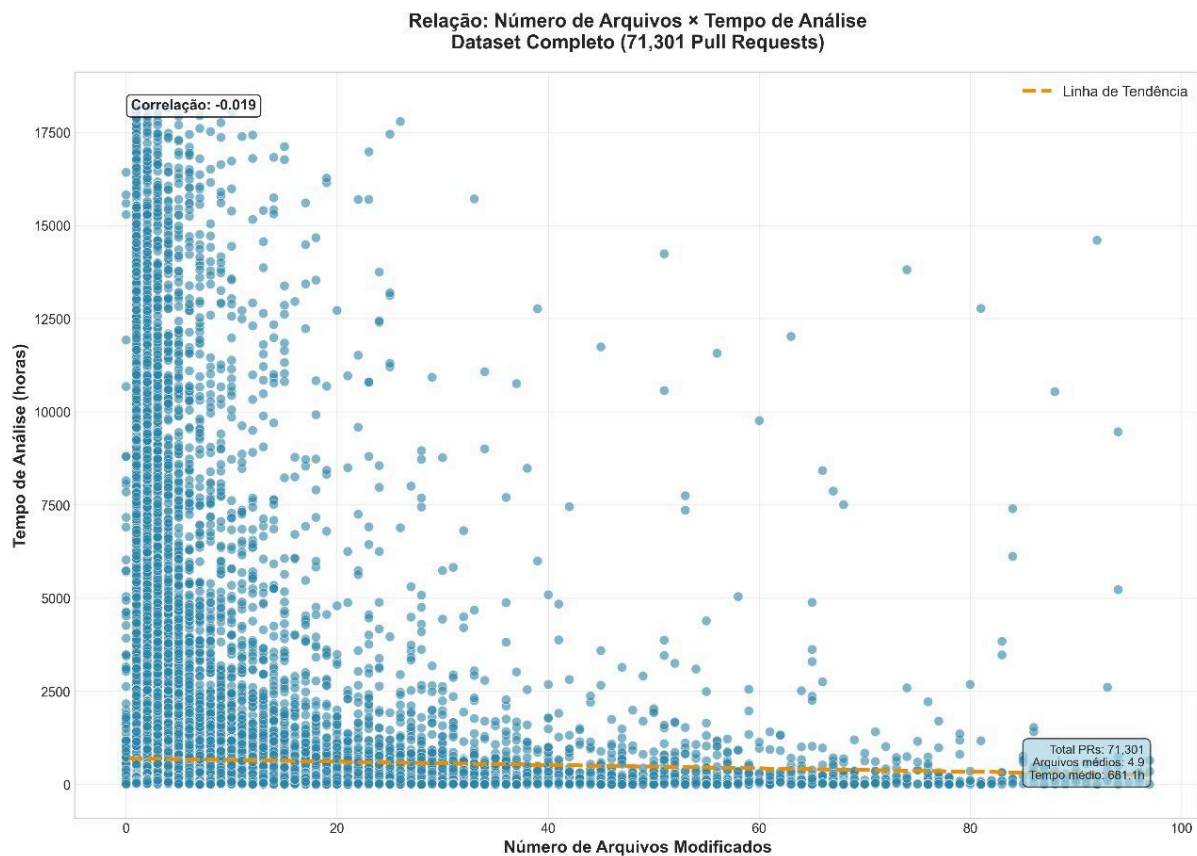
PRs menores e bem descritos continuam mais propensos a serem aprovados;

A quantidade de interações está associada a revisões mais detalhadas;

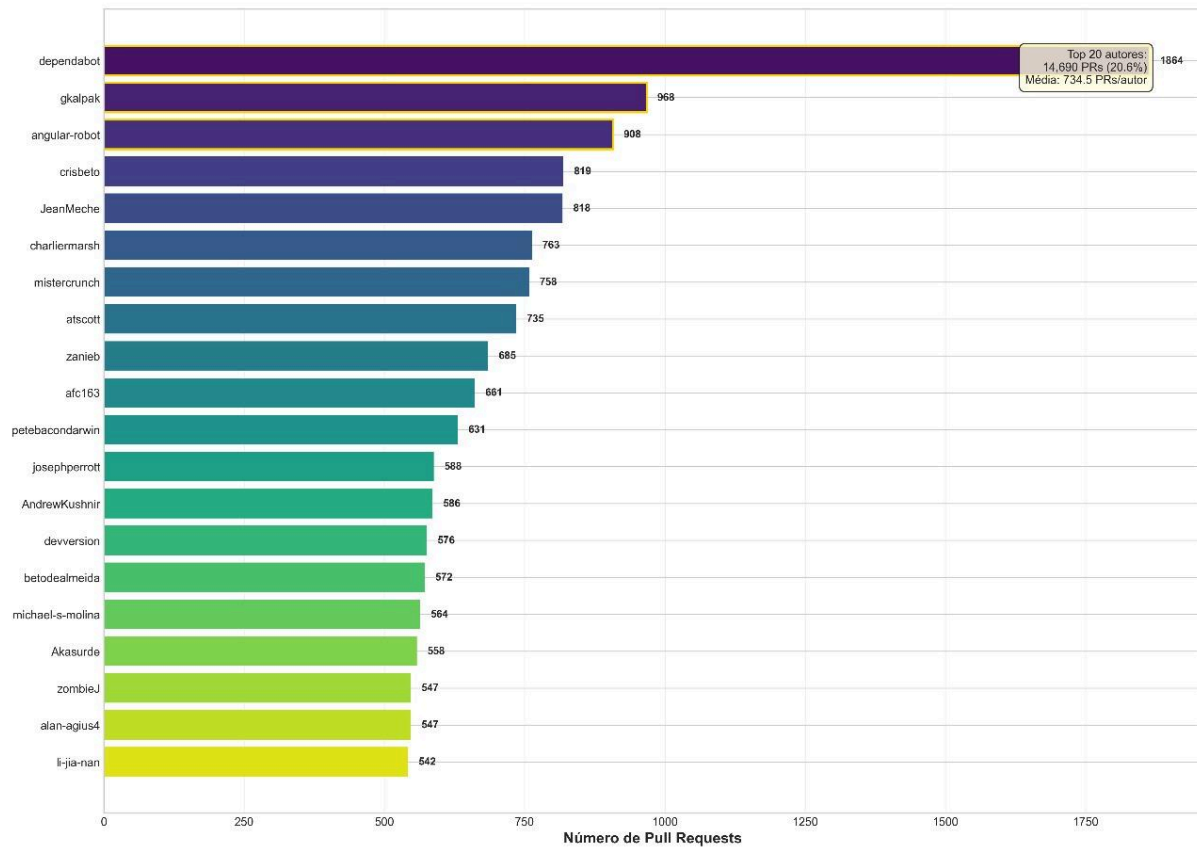
O tempo de análise é um indicativo relevante de complexidade e qualidade da revisão.

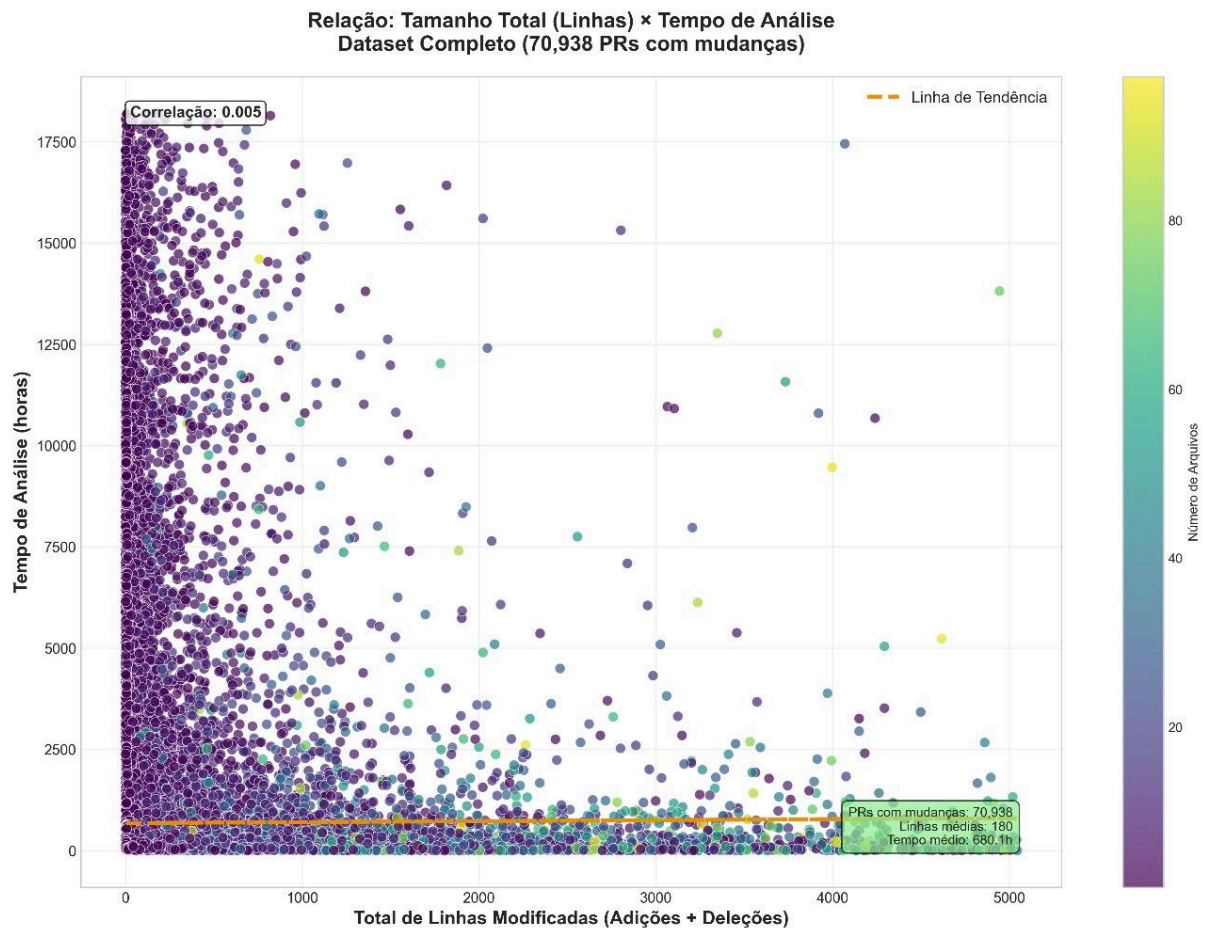
Esses resultados sugerem que a amostra controlada, apesar de limitada, foi representativa o suficiente para capturar os principais padrões comportamentais da revisão de código no GitHub.

5.4 Resultado visual visão universal



Top 20 Autores com Mais PRs no Dataset Completo
Total: 71,301 PRs de 17,469 autores únicos





6. Discussão

- **RQ01 e RQ05 (Tamanho dos PRs):** PRs grandes tendem a ser mais arriscados, com maior probabilidade de rejeição, mas também podem receber mais revisões.
- **RQ02 e RQ06 (Tempo de análise):** PRs com tempo elevado podem indicar problemas de complexidade ou falta de clareza, afetando o merge.
- **RQ03 e RQ07 (Descrição):** descrições detalhadas auxiliam revisores, favorecendo a aceitação e estimulando revisões mais aprofundadas.
- **RQ04 e RQ08 (Interações):** PRs que geram discussões ativas sugerem maior engajamento, podendo resultar em revisões mais completas e em maior qualidade final.

6. Conclusão

A análise confirma que múltiplas variáveis impactam o processo de *code review* em projetos *open source*.

Entre elas, destacam-se a clareza da descrição e o tamanho dos PRs, que se mostraram fatores determinantes tanto para aceitação quanto para o número de revisões.

A comparação entre os ambientes controlado e universal reforça a robustez das conclusões: mesmo diante de maior diversidade de repositórios e contextos, as tendências gerais se mantiveram estáveis.

Conclui-se, portanto, que PRs menores, bem documentados e com interação ativa entre revisores possuem maior chance de aprovação, contribuindo para um fluxo de desenvolvimento colaborativo e eficiente.