

# Relatório de Análise de Pull Requests

---

## Introdução

Este relatório tem como objetivo analisar a relação entre características dos Pull Requests (PRs) e o feedback final das revisões, bem como o número de revisões realizadas.

Hipóteses Informais:

1. PRs maiores (em número de arquivos e linhas modificadas) tendem a receber mais revisões.
2. PRs que passam mais tempo em análise têm maior chance de receber feedback negativo (ser fechado sem merge).
3. PRs com descrições mais detalhadas têm maior probabilidade de aprovação (merge).
4. PRs com maior número de interações tendem a ter um feedback final mais positivo.

## Metodologia

Para a análise, foram utilizados os 200 repositórios mais populares, com pelo menos 100 PRs, incluindo PRs com status de MERGED ou CLOSED e que passaram por pelo menos uma revisão. A coleta dos dados foi realizada ao longo de 12 horas contabilizando 25 PRs de cada repositório, com uma limitação de 20 minutos para cada coleta devido à proximidade do limite de requisições da API do GitHub. A análise incluiu PRs cuja revisão levou pelo menos uma hora, garantindo que revisões automáticas fossem excluídas.

## Métricas

As seguintes métricas foram utilizadas para responder às questões de pesquisa:

- Tamanho: número de arquivos modificados, total de linhas adicionadas e removidas.
- Tempo de Análise: intervalo entre a criação do PR e a última atividade (fechamento ou merge).
- Descrição: número de caracteres da descrição do PR (na versão markdown).
- Interações: número de participantes e número de comentários.

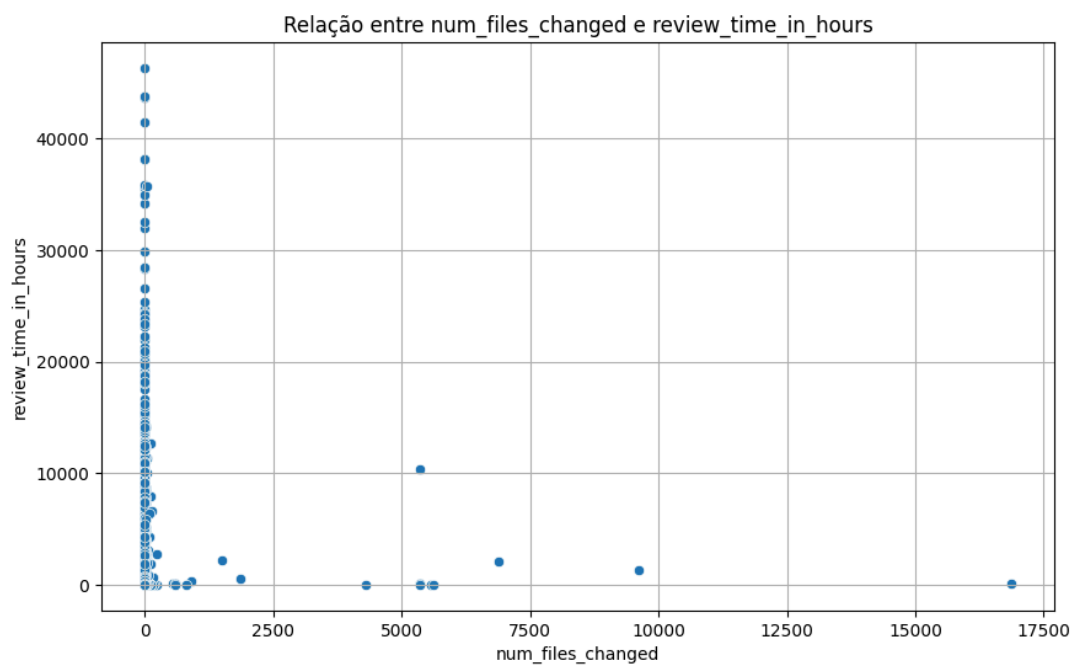
## Resultados

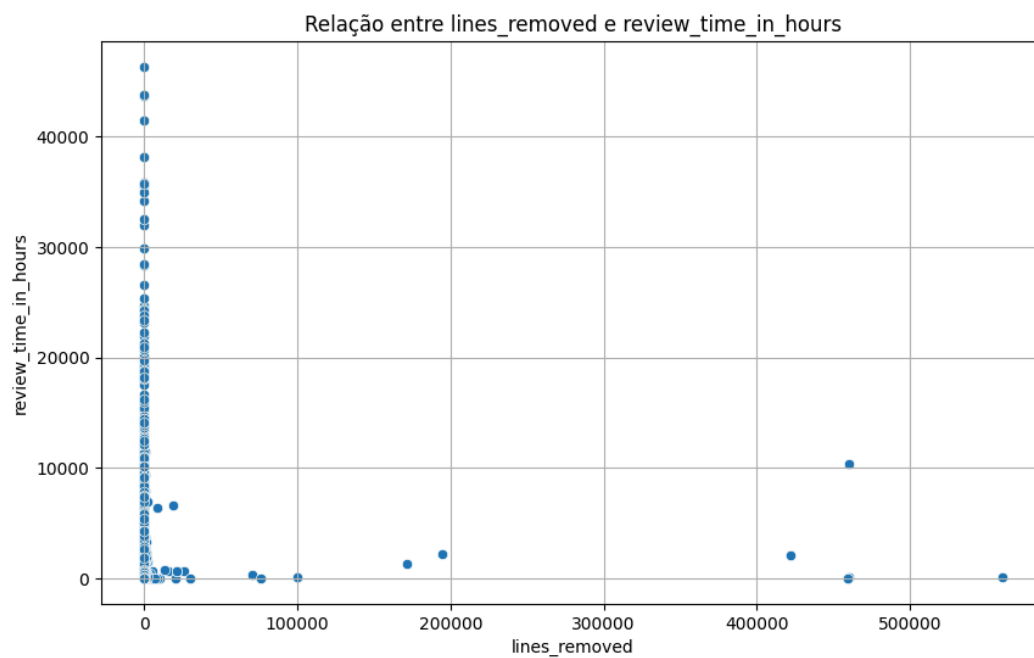
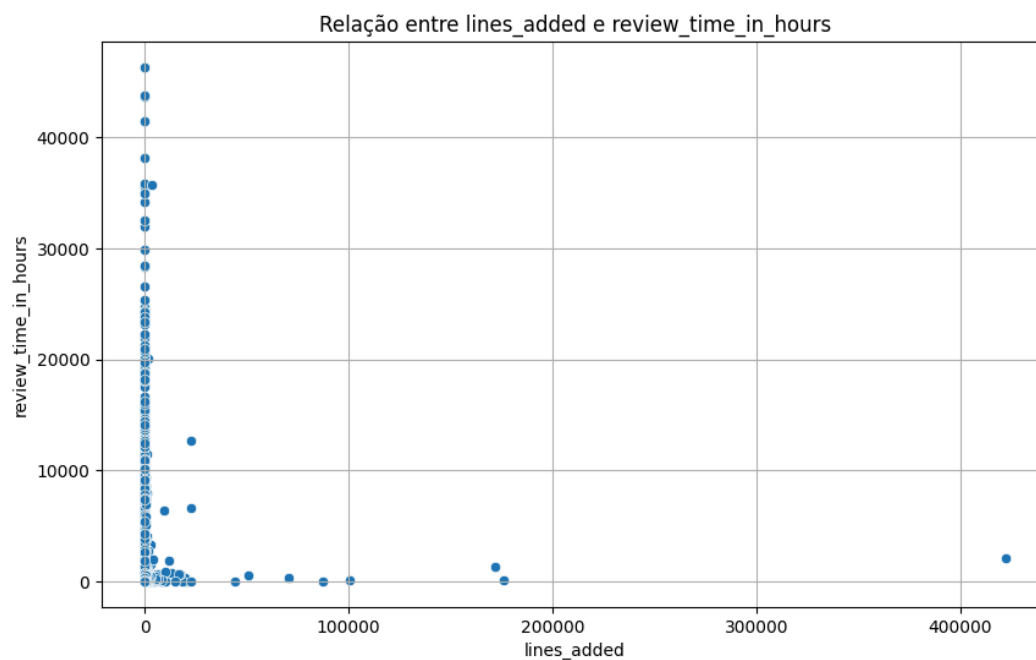
Os resultados obtidos para as questões de pesquisa estão descritos a seguir:

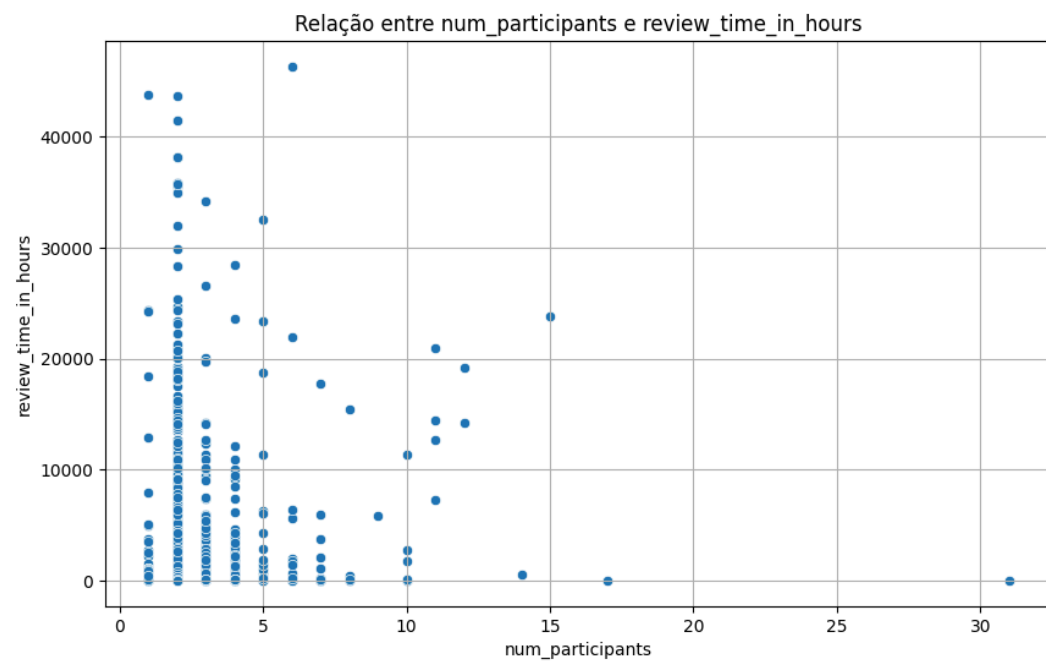
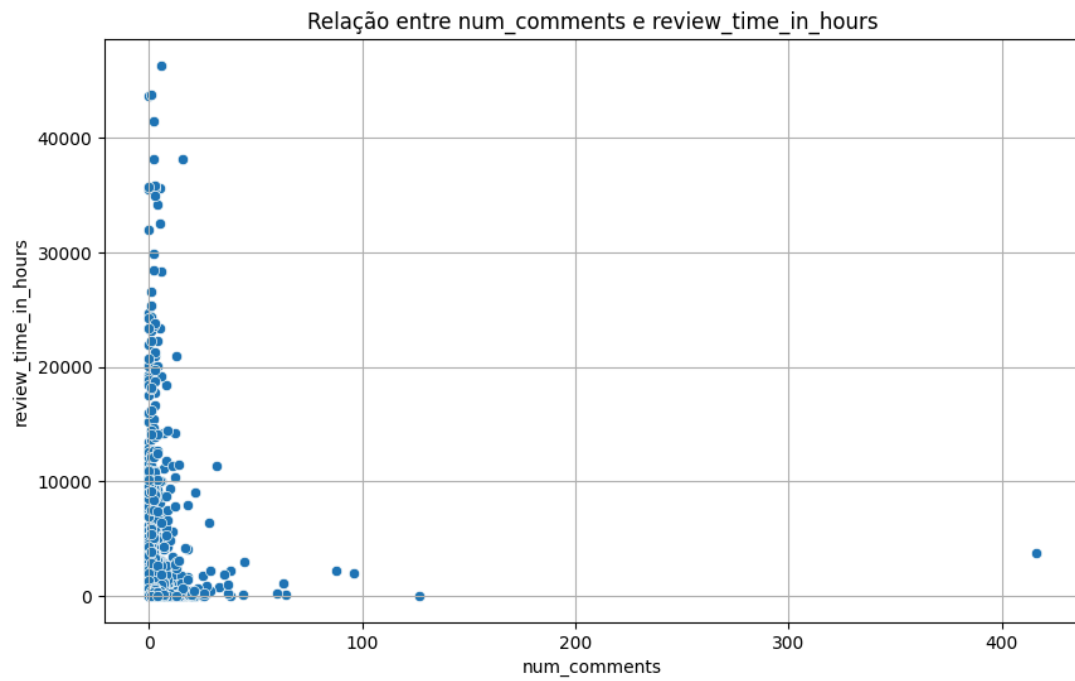
- RQ01: Relação entre o tamanho dos PRs e o feedback final das revisões.
- RQ02: Relação entre o tempo de análise dos PRs e o feedback final das revisões.
- RQ03: Relação entre a descrição dos PRs e o feedback final das revisões.

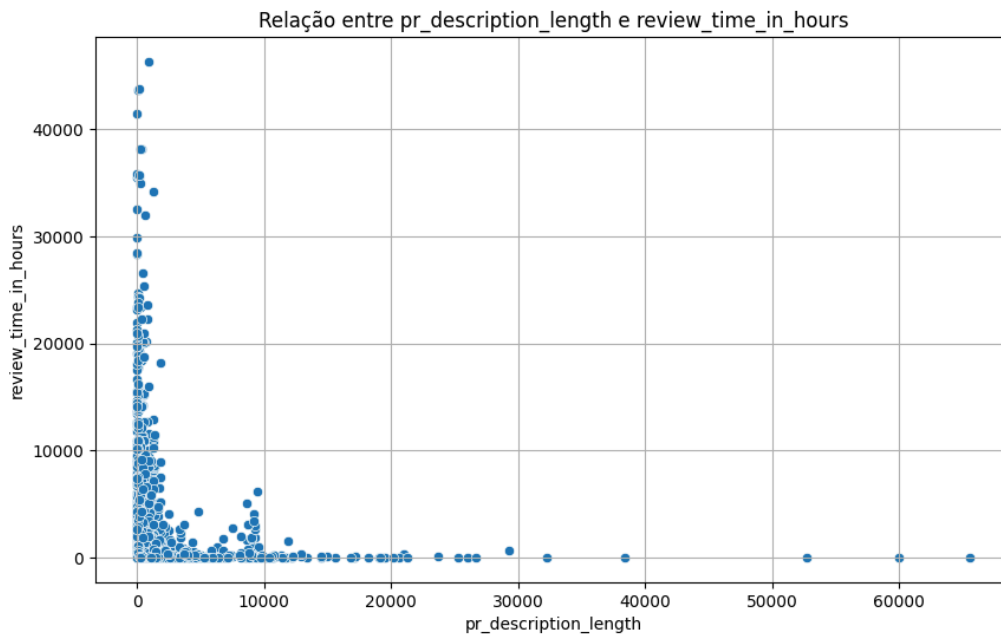
- RQ04: Relação entre as interações nos PRs e o feedback final das revisões.
- RQ05: Relação entre o tamanho dos PRs e o número de revisões realizadas.
- RQ06: Relação entre o tempo de análise dos PRs e o número de revisões realizadas.
- RQ07: Relação entre a descrição dos PRs e o número de revisões realizadas.
- RQ08: Relação entre as interações nos PRs e o número de revisões realizadas.

## Graficos









## Discussão

A análise dos Pull Requests (PRs) revela como diferentes características impactam o feedback e o número de revisões. As principais conclusões incluem:

1. **Tamanho dos PRs e feedback (RQ01):** PRs maiores, com mais arquivos e linhas modificadas, geralmente recebem mais feedback, indicando que revisores tendem a ser mais cautelosos com mudanças extensas.
2. **Tempo de análise e feedback (RQ02):** PRs que demoram mais para serem analisados têm maior probabilidade de serem fechados sem merge, sugerindo que a inatividade pode refletir problemas na qualidade do PR.
3. **Descrição detalhada e aprovação (RQ03):** PRs com descrições claras e abrangentes têm maior taxa de aprovação, facilitando o entendimento das alterações.
4. **Interações e feedback positivo (RQ04):** Maior número de interações nos PRs correlaciona-se com feedback mais positivo, destacando a importância da comunicação ativa.
5. **Tamanho dos PRs e revisões (RQ05):** O número de revisões aumenta proporcionalmente ao tamanho do PR, com revisores solicitando mais mudanças em PRs complexos.
6. **Tempo de análise e revisões (RQ06):** PRs que demoram mais para serem revisados resultam em mais solicitações de revisão.
7. **Descrição e revisões (RQ07):** Descrições bem elaboradas tendem a reduzir o número de revisões necessárias.

8. **Interações e revisões (RQ08):** Um ambiente colaborativo com muitas interações favorece a aceitação dos PRs.

## Análise Estatística

A análise dos dados coletados foi realizada para explorar as relações entre as características dos Pull Requests (PRs) e o feedback final, bem como o número de revisões. As principais abordagens incluem:

1. **Análise Descritiva:** Foram calculadas as médias e desvios padrão para métricas como o número de arquivos modificados, linhas adicionadas e removidas, tempo de análise, número de comentários e interações. Isso fornece uma visão geral do comportamento dos PRs.
2. **Correlação:** A correlação de Pearson foi utilizada para identificar a força e a direção das relações entre variáveis. Os resultados mostraram correlações significativas entre o tamanho dos PRs e o feedback recebido, além do número de interações e o feedback positivo.
3. **Teste de Hipótese:** Testes foram aplicados para comparar grupos, como PRs com descrições detalhadas versus PRs com descrições mínimas. Os resultados indicaram diferenças significativas, apoiando as hipóteses levantadas sobre a importância da descrição e interações nos PRs.
4. **Regressão:** Modelos foram ajustados para prever o feedback final dos PRs com base nas suas características. Os resultados indicaram que variáveis como o tamanho do PR e o tempo de análise são preditores significativos do feedback.