



Diversidade e produtividade em times ágeis: percepções individuais *versus* dados.

Diversity and productivity in agile teams: individual perceptions versus data.

Taianne Valerie Alves Motta¹ , Gláucia Braga e Silva¹ 

¹Instituto de Ciências Exatas e Tecnológicas - Universidade Federal de Viçosa (UFV)
- Florestal – MG – Brazil

{taianne.mota, glaucia}@ufv.br

Abstract. *This paper investigates the influence of diversity on agile teams results, comparing individual perceptions of the developers with team's productivity and quality measurements. The study was carried out in an industry context, involving five teams with different diversity profiles. The work extends a previous research, including an investigation into the representativeness of diversity categories found in the teams and an analysis of soft skills based on the gender of the developer. Although other statistical analyzes can be performed, the results showed a positive influence of diversity on productivity and quality measurements of the evaluated teams. Regarding gender diversity, as expected, low representativeness was observed in the teams, but relatively promising results were found in terms of women's perception of the soft skills they possess. Finally, a certain discrepancy was observed when crossing metric results with the individual perceptions of the interviewed developers, indicating some communication failures within the evaluated teams.*

Keywords. *Diversity in Scrum teams; soft skills; communication in software engineering*

Resumo. *Este artigo investiga a influência da diversidade em times que adotam metodologias ágeis, comparando percepções individuais das pessoas desenvolvedoras com resultados de métricas de produtividade e qualidade aplicadas sobre dados de uma ferramenta de suporte. O estudo foi realizado no contexto da indústria, envolvendo cinco times com diferentes perfis de diversidade. O trabalho aprofunda uma pesquisa anterior, incluindo uma investigação sobre a representatividade das categorias de diversidade encontradas nos times e fazendo um recorte de gênero no que compete à presença de Soft Skills. Embora outras análises estatísticas possam ser realizadas, os resultados mostraram uma influência positiva da diversidade para os índices de produtividade e qualidade dos times avaliados. No que compete à diversidade de gênero,*

como já esperado, observou-se baixa representatividade nos times, mas resultados relativamente promissores em termos da percepção das mulheres quanto às Soft Skills que elas julgam possuir. Por fim, de forma geral, observou-se certa discrepância, ao se cruzar os resultados das métricas com as percepções individuais das pessoas desenvolvedoras entrevistadas, indicando falhas na comunicação dentro dos times avaliados.

Palavras-Chave. *Diversidade em times Scrum; soft skills; comunicação em Engenharia de Software*

1. Introdução

A Engenharia de Software Ágil evidencia a importância de times diversos dentro do desenvolvimento de software, pois possui um conjunto de métodos e práticas baseados nos valores e princípios expressos no Manifesto Ágil [Silveira and Prikładnicki 2019], como o planejamento adaptativo e o desenvolvimento com melhoria contínua. Times diversos são múltiplos e plurais, compostos por pessoas com várias características que se diferenciam e acrescentam, nas resoluções de problemas, gerando um conjunto mais amplo de habilidades, necessidades e vivências. Na busca pela qualidade do produto, torna-se fundamental melhorar o processo envolvido. Para isso, o ambiente de trabalho precisa ser um espaço livre, acolhedor e inclusivo, de forma a melhorar a produtividade e explorar o potencial criativo da equipe. Isso vai de encontro ao que defende o primeiro valor do Manifesto Ágil, definido por Kent Beck e outros dezesseis programadores [Beck et al. 2001], segundo o qual indivíduos e a interação entre eles são mais importantes do que processos e ferramentas.

A diversidade no contexto da Engenharia de Software (ES) tem sido objeto de estudo de vários trabalhos na literatura [Silveira and Prikładnicki 2019, Albusays et al. 2021, Rodríguez-Pérez et al. 2021], com um foco maior em questões de gênero. No entanto, pouco se conhece sobre o impacto de outras diversidades no contexto do desenvolvimento de software [Rodríguez-Pérez et al. 2021]. Além disso, existe também uma lacuna no que compete ao impacto da diversidade no contexto das metodologias ágeis [Silveira and Prikładnicki 2019].

Dessa forma, com o intuito de contribuir com novos estudos sobre a influência da diversidade, em especial, no contexto da Engenharia de Software Ágil, este trabalho será guiado pela investigação das seguintes questões de pesquisa:

- QP1: Quais são as categorias de diversidade mais/menos representativas nos times?
- QP2: Existe alguma influência na produtividade em função do índice de diversidade da equipe?
- QP3: Existe alguma influência na qualidade do produto em função do índice de diversidade da equipe?
- QP4: Existe alguma relação entre a presença de *Soft Skills* e o gênero da pessoa desenvolvedora?

Este trabalho propõe um estudo quali-quantitativo sobre a influência da diversidade em times ágeis, comparando percepções individuais com dados de produtividade e qualidade, por meio de uma pesquisa com desenvolvedores e da aplicação de métricas de produtividade e de qualidade de software no contexto de projetos reais da indústria. O estudo é conduzido em uma empresa de tecnologia e abrange a participação de cinco equipes com diferentes perfis de diversidade, as quais adotam metodologias ágeis em seus desenvolvimentos. Para medir o grau de diversidade das equipes avaliadas, o estudo irá contar com o índice de heterogeneidade de Blau [Blau 1977]. Neste artigo, resultados de um estudo anterior [Motta and Braga e Silva 2022] são ampliados, com investigação de duas novas questões de pesquisa, a saber a QP1 e a QP4. Além disso, aprofundam-se as análises das questões QP2 e QP3, cruzando-se os dados com as percepções individuais dos membros das equipes.

O restante deste artigo está organizado de acordo com as seções: seção 2, que apresenta conceitos sobre Metodologias Ágeis e Diversidade; seção 3, com os trabalhos relacionados; seção 4, em que se abordam materiais e métodos; seção 5, com os resultados obtidos; seção 6, com as ameaças à validade; e a seção 7, que traz as considerações finais.

2. Metodologias Ágeis e Diversidade

As metodologias ágeis surgiram para alinhar os valores e princípios do Manifesto Ágil [Beck et al. 2001], que sugere quatro conceitos fundamentais, sendo eles:

1. “Indivíduos e interação entre eles mais que processos e ferramentas.”
2. “Software em funcionamento mais que documentação abrangente.”
3. “Colaboração do cliente mais que negociação de contratos.”
4. “Responder a mudanças mais que seguir um plano.”

O primeiro conceito das Metodologias Ágeis demonstra o entendimento de que pessoas são as responsáveis por realizar o desenvolvimento de software e, portanto, a qualidade da interação entre elas pode resolver problemas crônicos de comunicação. A exemplo de que a comunicação é necessária para uma implementação bem-sucedida, e a criação de um entendimento compartilhado entre as partes interessadas constitui um fator importante para o sucesso do projeto [Diegmann and Rosenkranz 2017].

Na Engenharia de Software, a habilidade de comunicação é considerada uma *soft skill* imprescindível em um profissional da área. Essa se torna especialmente importante no contexto de algumas práticas do desenvolvimento ágil, com destaque para as atividades de requisitos, programação em pares (*Pair Programming*) e revisão em pares (*Pair Review*). No que diz respeito à atividade de levantamento de requisitos, a habilidade de comunicação em conjunto com um time mais diverso aumenta a qualidade das especificações, pois o time consegue entender e perceber as dificuldades e necessidades dos usuários mais de perto, [Barbosa et al. 2021].

Quando se tem um foco em pessoas e não em processos ou algoritmos, o ambiente de trabalho fica mais humanizado criando um contexto integrativo necessário para que a diversidade se torne o passo essencial, abrangendo e incluindo todas as pessoas. Por isso

que ao se considerar indivíduos e interações, o desempenho do trabalho deve ser feito em conjunto, abarcando as diferentes ideias do time, as pluralidades e diversidades, sejam elas culturais, físicas ou de gênero.

No contexto da Engenharia de Software, muito se fala sobre a importância da diversidade. Estudos mostram que equipes de software mais diversas podem aproveitar de informações mais amplas e possuem experiências e habilidades mais aprimoradas para resolução de problemas, tornando-se, portanto, mais eficazes [Vasilescu et al. 2015]. No entanto, em geral, as discussões têm se afunilado muito para questões de gênero, uma vez que o ambiente de Tecnologia da Informação (TI) é majoritariamente masculino. Entretanto, se faz necessário perceber o problema de maneira ampla, abordando outros aspectos sociais e outras características de diversidade que são igualmente importantes no contexto dos times [Rodríguez-Pérez et al. 2021].

3. Trabalhos Relacionados

No que compete à avaliação da produtividade de equipes ágeis, Melo et al. [Melo et al. 2011] acompanharam duas equipes para saber quais os conceitos e fatores de produtividade são utilizados em seus projetos, chegando à conclusão de que alguns princípios tradicionais de produtividade ainda impactam os times de desenvolvimento, mesmo com a adoção das práticas ágeis. Através de revisões literárias, alguns fatores de produtividade, como reuso e características de software, foram levantados. O trabalho também destacou a experiência e a capacidade de cumprir prazos como fatores pessoais de cada membro da equipe, e como fatores de projeto, comunicação. Em seu estudo Hernández et al. [Hernández et al. 2019], faz uma revisão de literatura sobre métricas que avaliam a produtividade no contexto de desenvolvimento ágil de software e discute sobre como as medidas são utilizadas majoritariamente de forma quantitativa.

Motta e Braga e Silva [Motta and Braga e Silva 2022] apresentaram um estudo quali-quantitativo sobre a influência da diversidade em times ágeis, por meio da aplicação de métricas de produtividade e qualidade sobre dados de uma ferramenta de suporte. As autoras também realizaram uma pesquisa com as pessoas desenvolvedoras envolvidas, de forma a complementar e entender melhor os dados.

No que compete à diversidade de gênero, a baixa participação feminina em projetos de desenvolvimento de *software* tem sido avaliada, principalmente, em termos da contabilização das contribuições em código e *pull request* aceitas [Rodríguez-Pérez et al. 2021, Canedo et al. 2019]. Zacchiroli [Zacchiroli 2021] analisou 1.6 milhões de *commits* e observou que cerca de 92% do código produzido nas plataformas de versionamento de código foi feito por homens. Catolino et al. [Catolino et al. 2019] realizaram um estudo empírico para analisar as relações entre o desequilíbrio de gênero e a ocorrência de *community smells*¹, concluindo que as mulheres são instrumentos fundamentais para redução destes *smells* em times de software, exercendo um importante papel em termos da comunicação.

O presente artigo aprofunda uma pesquisa anterior [Motta and Braga e Silva 2022], am-

¹Termo associado a *code smell*, que representam padrões de organização e comunicação abaixo do ideal em equipes de software.

pliando a amostra de dados e aprofundando as análises no que compete ao cruzamento dos resultados das métricas com as percepções dos envolvidos. Além disso, o trabalho faz uma análise das categorias de diversidades presentes nos times, destacando as sub-representações de minorias. Por fim, o trabalho faz um recorte de gênero para avaliar a presença de *Soft Skills* nas equipes, dentre as quais a habilidade de comunicação, de forma a investigar se os achados de Catolino et al. [Catolino et al. 2019] se manifestaram na amostra avaliada, embora não faça parte do escopo deste trabalho a avaliação de *community smells* em si.

4. Materiais e Métodos

Este estudo foi guiado por dois métodos de coleta de dados: um formulário para análise quali-quantitativa; e a extração de dados de uma ferramenta de apoio ao desenvolvimento, para aplicação de métricas de produtividade e qualidade. Os mesmos foram coletados no contexto de uma empresa de soluções tecnológicas e envolveu times que aplicassem *Scrum*² em seus projetos de software, uma vez que esta metodologia ágil valoriza a interação entre os indivíduos. Além disso, o *Scrum* trabalha com as *Sprints* (períodos curtos e fixos de tempo em que uma equipe trabalha para concluir uma quantidade definida de tarefas), que favorecem a aplicação das métricas. Os times foram escolhidos arbitrariamente. Ampliando a amostra do trabalho anterior [Motta and Braga e Silva 2022], cinco times participaram da pesquisa, nomeados aqui como: “Time A”, “Time B”, “Time C”, “Time D” e “Time E”.

Para se calcular o índice de diversidade dos times, foi utilizada a métrica conhecida como Índice de Blau [Collins 1979]. O índice analisa a medida da diversidade de espécies em um ecossistema com valores de zero (população heterogênea) a um máximo indicado pela quantidade de categorias. Neste estudo, foram consideradas as categorias: identidade de gênero, raça, Pessoas com Deficiência (PCD), orientação sexual e Estado.

Os resultados dependem do cálculo: $1 - (\sum_{x=0}^N p_i^2)$

Onde p_i é a porcentagem de indivíduos dentro da categoria e N é o número de categorias. O valor máximo de cada índice segue o seguinte padrão: 2 categorias com 35% = 0,65, 3 categorias com 30% = 0,70; 4 categorias com 25% = 0,75; 5 categorias com 20% = 0,8; E assim, sucessivamente.

4.1. Pesquisa com desenvolvedores

Para conhecer o perfil de diversidade dos times e também sua percepção quanto às *soft skills* e às medidas de produtividade e de qualidade do produto em seus projetos, foi elaborado um questionário, organizado em três dimensões:

1. **Perfil de Diversidade**, com perguntas sobre a identidade do respondente, incluindo sua identificação de gênero, orientação sexual, idade, raça, origem e se é pessoa com deficiência (PCD). Além de conter perguntas sobre sua posição/cargo dentro da empresa.
2. **Perfil Técnico**, com o objetivo de perceber como o respondente enxergava o contexto do seu time, incluindo prazos de *Sprints*, quantidades de bugs/débitos técnicos, retrabalhos, entre outros.

²<https://www.scrum.org/>

3. *Soft Skills*, com o objetivo de avaliar a qualidade das comunicações no contexto dos times ágeis e como o respondente considera sua percepção do time, realizando boas leituras emocionais e de impedimentos.

Os formulários foram elaborados no Google Forms³ e continham um total de 28 perguntas quali-quantitativas, sendo apenas uma aberta. As perguntas foram distribuídas em seções, sendo uma para cada dimensão: 11 perguntas, na seção de “Perfil de Diversidade”, 10 perguntas, na seção de “Perfil Técnico”, e 7 na seção de “*Soft Skill*”. As perguntas foram elaboradas com base nos trabalhos de Melo et al. [Melo et al. 2011], Silveira e Prikladnicki [Silveira and Prikladnicki 2018] e Kitchenham et al. [Kitchenham et al. 2009]. A maioria das perguntas emprega a escala de Likert utilizando 5 pontos, para resultados médios. A escolha dessa escala considerou o estudo de K. Vieira e M. Dalmoro [Kelmara M. Vieira 2008], que mostra a importância da mesma quando pondera escalas com formatos iguais, contínuo de direção e com a opção de neutro, necessária para a confortabilidade dos respondentes. O mesmo formulário foi aplicado para os diferentes times participantes de forma individual, com a finalidade de melhor organizar as respostas e cruzá-las com as métricas de cada time separadamente. A aplicação foi realizada remotamente, com cada respondente utilizando o próprio computador. O prazo para as respostas desses formulários foi de duas a três semanas e, no total, 24 indivíduos participaram.

4.2. Métricas de Produtividade e de Qualidade

Para avaliar a produtividade e a qualidade dos artefatos produzidos, de forma similar ao trabalho anterior [Motta and Braga e Silva 2022], foram utilizadas métricas que levaram em consideração os contextos dos times participantes e seus respectivos ambientes [Cesar 2017]. As métricas foram, em sua maioria, coletadas individualmente pelos times no ambiente da plataforma Microsoft Azure, o Visual Studio (VSTS)⁴. Nesse ambiente, é possível desenvolver e monitorar os produtos elaborados, em nuvem, localmente e, posteriormente, colocá-los em produção. Cada time escolheu um integrante para a retirada das métricas na ferramenta. Para isso, foram utilizadas consultas na plataforma, que selecionavam os dados separadamente, dentro dos seus contextos. A produtividade de cada time foi avaliada usando-se três métricas aplicadas sobre os dados do VSTS:

1. MP1 - Média de features por *Sprint*, calculada pela relação do número de trabalhos realizados (*features*) pelo número de *Sprints*;
2. MP2 - Média de *user stories* por *Sprint*, calculada pela relação do número de *user stories* pelo número de *Sprints*; e
3. MP3 - Percentual de tarefas cumpridas no prazo, calculada pela relação de tarefas que respeitaram as datas de entregas (*target dates*) dentre todas as tarefas previstas.

Para analisar a qualidade, outras duas métricas foram aplicadas no contexto dos times participantes e também de acordo com as possibilidades do VSTS:

1. MQ1- Média de bugs por *Sprint*, calculada pela relação do número de bugs pelo número de *Sprints*;

³<https://github.com/taiannemotta/Dados/blob/main/formularios.pdf>

⁴<https://azure.microsoft.com/>

2. MQ2 - Percentual de débitos técnicos, calculada pela relação do número de débitos técnicos registrados por features.

5. Análise e Discussão dos Resultados

Esta seção apresenta os resultados com a coleta e a análise de dados, de forma a responder as quatro questões de pesquisa endereçadas neste trabalho. As análises foram realizadas cruzando-se, os dados do formulário aplicado aos desenvolvedores com os resultados das métricas de produtividade (MPs) e de qualidade (MQs), de forma a verificar se as percepções individuais dos envolvidos estão alinhadas com os resultados.

Participaram do estudo, cinco times compostos por pessoas desenvolvedoras com diferentes níveis de diversidade, sendo seis pessoas no Time A (5 homens e 1 mulher); três pessoas (1 mulher e 2 homens) no Time B, três no time C (todos homens); oito pessoas no time D (2 mulheres e 6 homens) e quatro pessoas no time E (todos homens). Para se conhecer o perfil de diversidade de cada time, foram identificadas as categorias presentes e contabilizados os respectivos indivíduos em cada uma delas, segundo as respostas da primeira dimensão do formulário (Tabela 1). Para cada categoria, calculou-se o índice de Blau (Figura 1), sendo que quando um time apresenta este índice igual a 0,0, tem-se ausência de diversidade para aquela categoria.

Tabela 1. Indivíduos por Categoria de Diversidade

Quantidade	A (6)	B (3)	C (3)	D (8)	E (4)
Gênero	2	2	1	2	1
Raça	2	2	2	3	2
Orientação	3	2	2	3	1
PCD	1	1	1	2	1
Estado	5	2	2	4	1

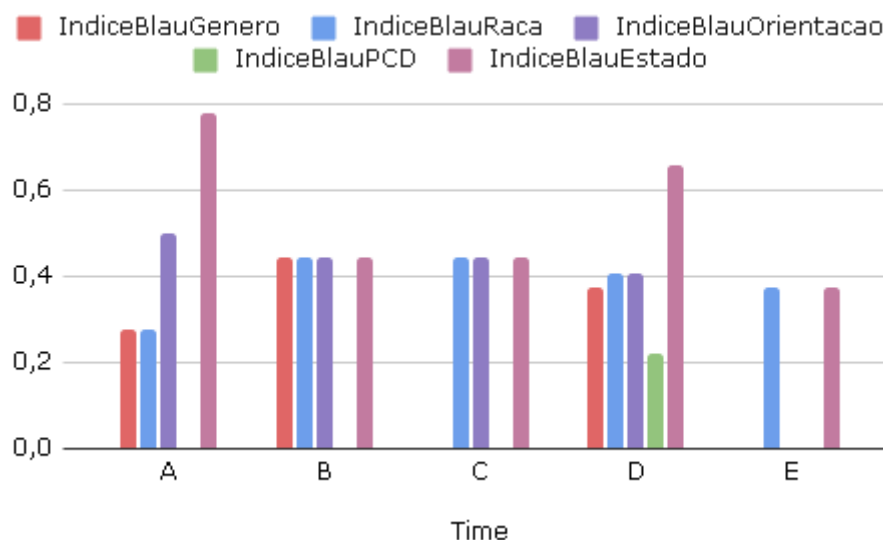


Figura 1. Índices de Blau calculados para cada time e por categorias.

Dentre os resultados mostrados no gráfico da Figura 1, é possível observar que todos os times participantes se mostram relativamente diversos, o que pode ser decorrente de políticas inclusivas e de diversidade da empresa envolvida. Na média, os times que podem ser considerados mais abrangentes em termos de diversidade são os Times B e D, possuindo níveis de diversidade próximos ao valor máximo e possuindo o maior número de diversidades. Os times que se apresentaram menos diversos considerando a mesma média, foram os times C e E. Além disso, o time D foi o único com diversidade para PCD, diferencial importante quando se considera a construção de sistemas mais inclusivos e com maior acessibilidade. No geral, confirmando outros achados na literatura [Burgstahler and Ladner 2007, Spingola 2018, Canedo et al. 2019], observa-se a baixa representatividade das minorias, em especial PCD e gênero feminino.

5.1. QP1: Quais são as categorias de diversidade mais/menos representativas nos times?

Com o intuito de conhecer o perfil de cada time participante e destacar as categorias de diversidade presentes, os dados do formulário foram analisados separadamente. De forma a complementar os resultados apresentados em [Motta and Braga e Silva 2022], o presente artigo analisa as categorias com maior ou menor representatividade nos times.

Inicialmente, foram analisados os dados relacionados a gênero. Embora o estudo não esteja restrito à diversidade de gênero, é importante ressaltar a problemática ainda presente sobre a baixa participação feminina no mercado de Tecnologia da Informação. Dos 24 respondentes, apenas 4 são mulheres (16%), percentual extremamente baixo que reforça os dados já conhecidos na literatura [Vedres and Vasarhelyi 2019] sobre a baixa representatividade de mulheres mesmo em empresas com políticas inclusivas. Conforme mostra a Figura 2. As 4 mulheres estão distribuídas em apenas 3 dos 5 times, com 1 ou 2 representantes, no máximo.

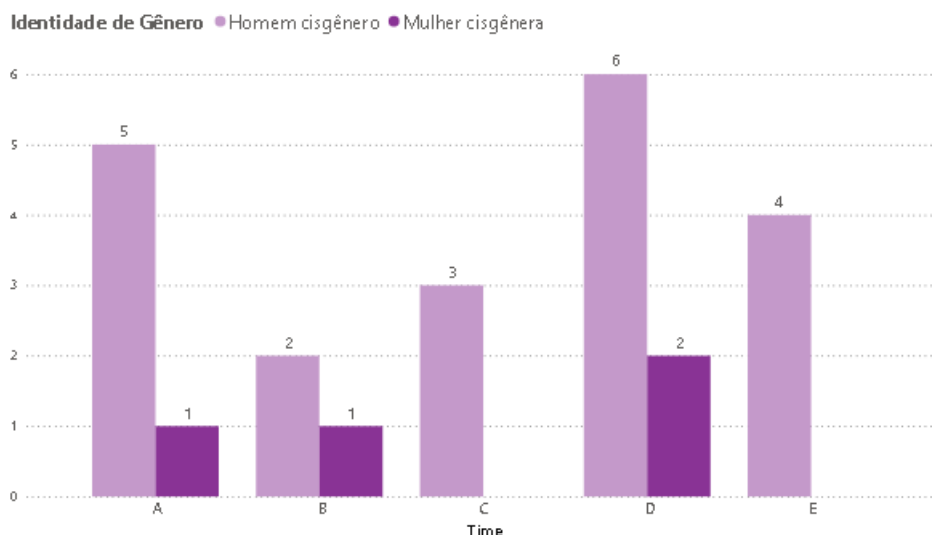


Figura 2. Participação de mulheres nos times.

Em relação à orientação sexual dos respondentes, 75% se declararam heterossexuais (18

de 24) e 6 indicaram outras orientações. A Figura 3 apresenta os dados detalhados, ilustrando a distribuição das orientações nos times analisados.

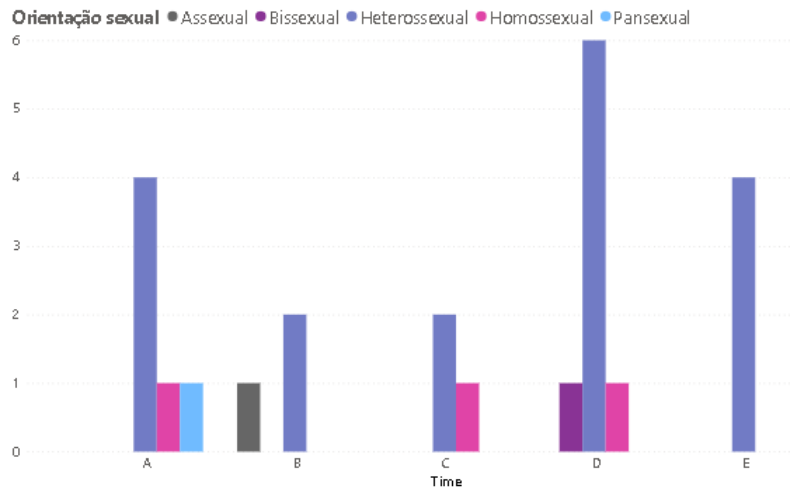


Figura 3. Diversidade em relação à orientação sexual dos times.

Em relação ao nível de diversidade de raça/cor (Figura 4), a maioria dos respondentes se declarou brancos (mais de 66%), aproximadamente 29% se declararam negros, sendo 25% como pardos e apenas uma pessoa se declarou preta (4,17%). Uma pessoa preferiu não se classificar (4,17%).

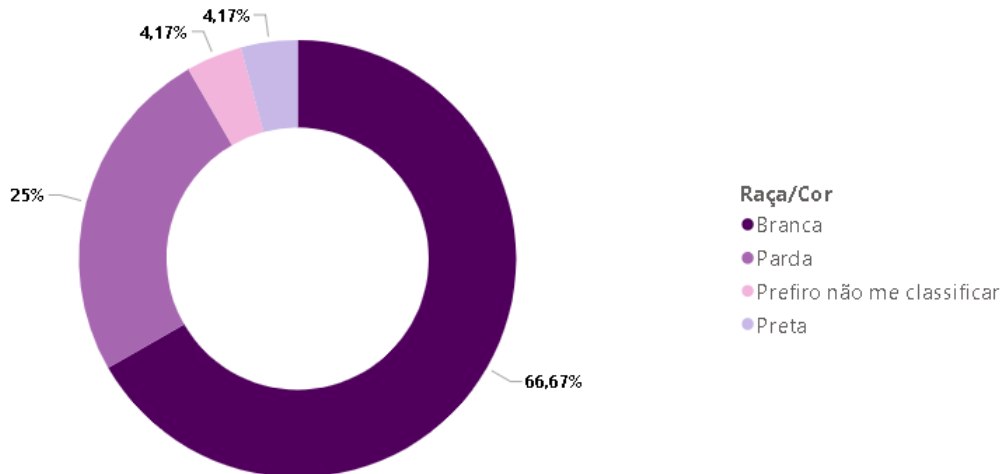


Figura 4. Diversidade em relação à raça/cor.

Em relação à diversidade de estado, a maioria dos participantes se encontram onde a empresa se localiza, conforme já esperado. Entretanto, existe uma participação diversa de pessoas de outros estados, conforme ilustra o gráfico da Figura 5. Acredita-se que esse comportamento pode estar associado ao trabalho remoto, que foi impulsionado pelas

restrições impostas pela pandemia da COVID-19, mas tem sido amplamente adotado em muitas das empresas de TI.

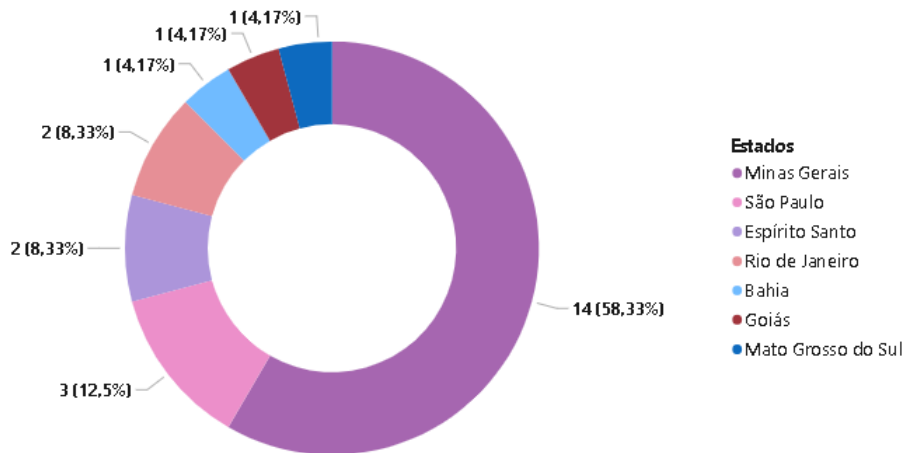


Figura 5. Diversidade de localização dos integrantes dos times.

Do ponto de vista técnico, os participantes foram perguntados sobre os papéis desempenhados no processo de software dentro dos times, tempo de participação e níveis de experiência. No que diz respeito ao papéis desempenhados (Figura 6), as respostas indicaram que a maioria dos respondentes tem atuado com desenvolvimento diretamente (Desenvolvedores Líderes/ Tech Lead e Desenvolvedores), representando 79,16%. Esse resultado era esperado uma vez que todos os times adotam uma metodologia ágil, que sugere uma divisão de papéis mais simplificada e tem foco maior em codificação.

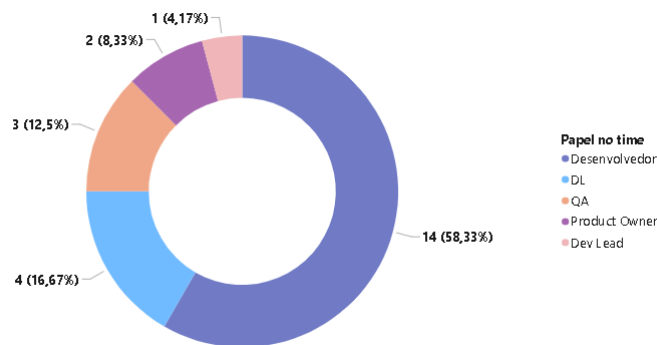


Figura 6. Papéis de atuação nos times.

Com relação ao tempo de participação (Figura 7), temos times relativamente “jovens”, com tempo variando de 3 a 17 meses. O time com o maior percentual de pessoas mais experientes durante o período da pesquisa foi o time C, que tem um integrante de 1 ano e 5 meses, e os outros dois integrantes com 5 e 6 meses respectivamente. Importante destacar que esse tempo se refere à participação do respondente no time avaliado e não

necessariamente na empresa. Fazendo um recorte para as mulheres, temos tempos de participação parecidos, sendo 2 integrantes, com 5 meses (times A e B) e 2 integrantes, com 4 meses (time D). Essa distinção é importante não só para as mulheres, mas para os novatos de forma geral, pois pode influenciar, por exemplo, na produtividade (curva de aprendizado) e na comunicação.

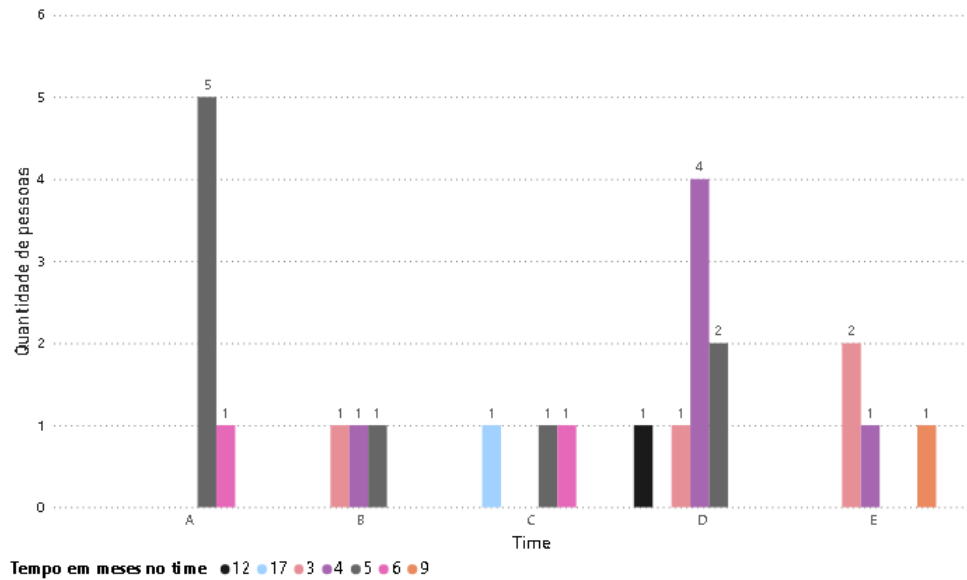


Figura 7. Tempo de participação no time.

Por fim, a Figura 8 apresenta as respostas relativas ao nível de experiência dos participantes, com a maioria deles nos níveis Pleno e Sênior. Esse conhecimento é importante porque pode ter impacto na produtividade dos times, na qualidade das produções e também na comunicação.

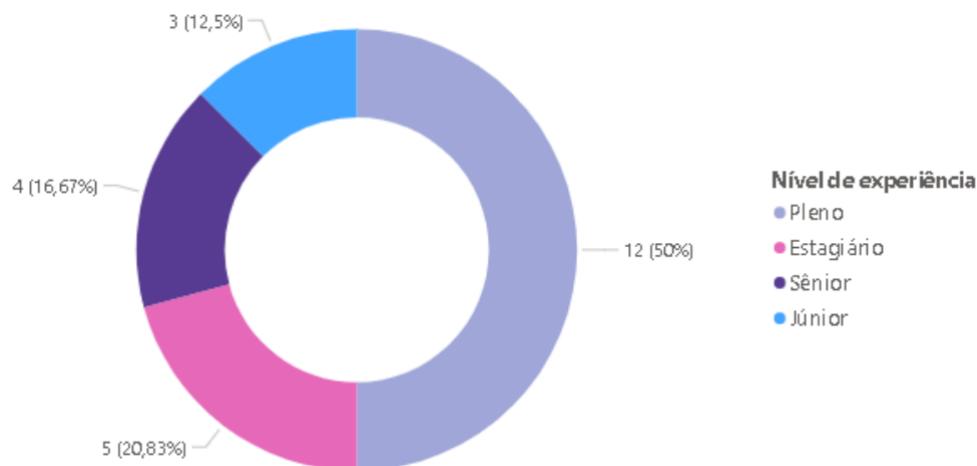


Figura 8. Níveis de experiência dos participantes.

5.2. QP2: Existe alguma influência na produtividade em função do índice de diversidade da equipe?

De forma similar ao trabalho de Motta e Braga e Silva [Motta and Braga e Silva 2022], a produtividade foi avaliada, a partir das três métricas de produtividades (MPs), mas agora ampliando a amostra para 5 times, conforme mostra a Figura 9. Importante destacar que as 3 métricas analisam produtividade em função das *Sprints*, sendo que os maiores resultados indicam melhores desempenhos.

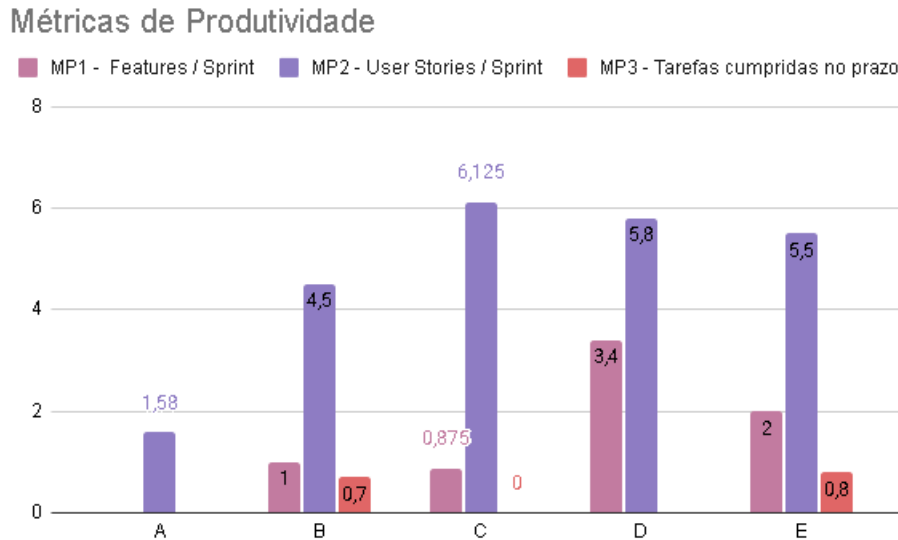


Figura 9. Resultados das métricas de produtividade para os 5 times.

Ressalta-se que o Time A não disponibilizou os cadastros das features do seu time, não sendo considerado na avaliação da MP1 e MP3, uma vez que as datas de entregas são utilizadas apenas nas features cadastradas dentro do VSTS. Na métrica MP1, o melhor resultado é do time D, com aproximadamente 3.5 features fechadas por *Sprint*. Apesar de ter três *Sprints* a menos do que os times B e C, é necessário ponderar que é o time com maior número de integrantes, o que também pode impactar na produtividade. Na métrica MP2, percebe-se que o time C com menor quantidade de integrantes (três) cadastrou a maior quantidade de User Stories, ficando próximo ao time D e E. Por fim, para a MP3, usada para medir as tarefas concluídas sem atrasos dentro de uma *Sprint*, apenas os times B, C e E foram considerados, porque os demais não faziam controle com datas de entregas (*target dates*). O time E tem uma média de 0.87 features entregues no prazo correto, próximo ao time B. Enquanto o time C não possui features sem atrasos.

A partir dos resultados, é possível observar que o time D apresentou os melhores valores de produtividade nas métricas MP1 e MP2, embora não tenha fornecido dados para o cálculo de MP3. Quando se considera os índices de diversidade, esse time também se destaca, pois é o único com diversidade em todas as categorias levantadas, com índices maiores em orientação sexual, cor/raça e, principalmente, estado.

Neste trabalho, com o intuito de cruzar os resultados da métrica MP3 com a percepção individual dos respondentes (aqui denominados Pessoas Desenvolvedoras (PDs)) quanto

à produtividade de seus times, as respostas da pergunta relacionada “Como você classifica o tempo das entregas de seu squad?” foram organizadas na Tabela 2. Ressalta-se que os times A e D não tinham o dado correspondente no VSTS (NA - Não Aplicável), pois como já informado anteriormente, não fizeram controle com datas de entregas (*target dates*) no VSTS. De qualquer forma, as percepções individuais dos integrantes desses times é bastante positiva quanto aos prazos das entregas. Acredita-se que a utilização desse recurso na ferramenta iria agregar em termos da coordenação das entregas do time.

Para os times B, C e E, existe alguma divergência na percepção dos times e os números no VSTS. O resultado comparativo é mais crítico para o time C que avalia de forma positiva as entregas (respostas maiores ou iguais a 3), apesar de ter 0% de tarefas entregues no prazo segundo os dados do VSTS. Para o time B, existe alguma coerência na média, mas avaliando-se as percepções individuais dos 3 integrantes, observa-se uma discrepância relevante em termos das notas que variam de 2 a 5. Ou seja, embora esse time tenha resultados razoáveis no VSTS (70% de entregas concluídas no prazo), a percepção de produtividade não é uniforme, o que pode indicar, por exemplo, falhas na comunicação do time. Os resultados são mais coerentes no contexto do time E, que tem 80% de tarefas cumpridas no VSTS e apresentou respostas 4 e 5 para a pergunta.

Tabela 2. Tarefas cumpridas no prazo - Percepção individual versus Dados.

Time	Como você classifica o tempo das entregas de seu squad?								VSTS (%)
	PD1	PD2	PD3	PD4	PD5	PD6	PD7	PD8	
A	4	3	5	4	4	4	-	-	NA
B	2	5	3	-	-	-	-	-	0,7
C	3	5	4	-	-	-	-	-	0
D	4	4	4	5	5	5	5	5	NA
E	4	5	4	4	-	-	-	-	0,8

Ainda com o intuito de avaliar a percepção dos indivíduos sobre a influência da diversidade na produtividade de seus times, foi usada a seguinte pergunta aberta no formulário: “Você percebe que a diversidade influencia a produtividade dentro do seu squad? Como?”. As respostas obtidas foram analisadas individualmente e utilizadas para gerar uma nuvem de palavras (Figura 10). Nessa nuvem, é notória a quantidade de “Não” para a responder à pergunta. Após análises das respostas, essa maior ocorrência revela que os respondentes parecem não reconhecer uma relação entre produtividade e diversidade. Essa percepção equivocada pode ser decorrente do fato de muitos não terem uma visão muito clara do que é a diversidade em si, valorizando mais as habilidades técnicas, ainda que trabalhem em um ambiente diverso. Esse achado já foi reportado em outro estudo na literatura, no âmbito da diversidade de gênero [Sarmiento et al. 2022], em que os participantes julgaram a diversidade menos relevante para mitigar *community smells* do que fatores técnicos tais como experiência e tamanho do time. Já os respondentes que reconhecem essa relação positiva, fizeram associações de como discussões sob diferentes visões podem mitigar riscos e encontrar soluções diversas para os problemas do desenvolvimento.



Figura 10. Nuvem de palavras com os termos mais frequentes das respostas.

5.3. QP3: Existe alguma influência na qualidade do produto em função do índice de diversidade da equipe?

Para avaliação da qualidade dos resultados dos times e suas possíveis relações com a diversidade, de forma similar ao trabalho de Motta e Braga e Silva [Motta and Braga e Silva 2022], foram utilizadas as duas métricas de qualidade (MQs), conforme ilustra a Figura 11. Importante destacar que para essas métricas, maiores resultados indicam um cenário menos satisfatório, já que estão sendo medidos número de *bugs* e débitos técnicos.

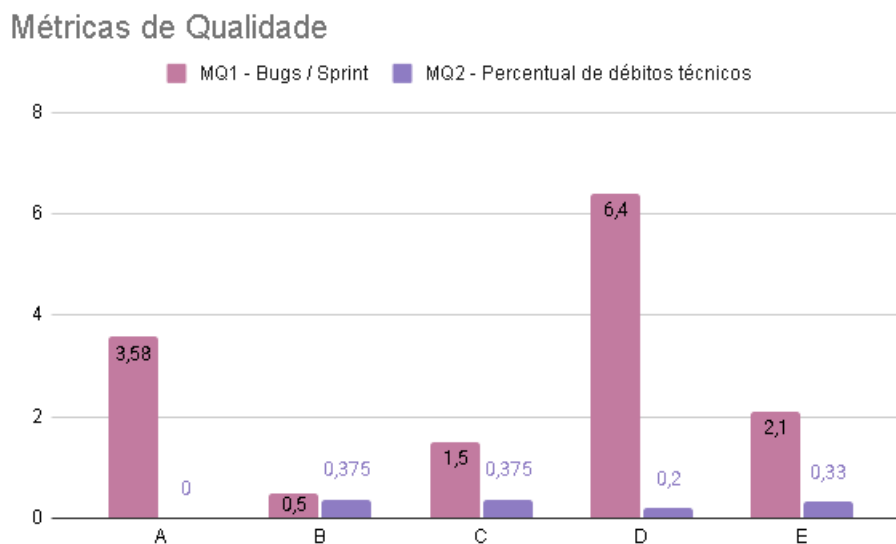


Figura 11. Resultados das métricas de qualidade para os 5 times.

Através da MQ1, foi possível observar que o número de bugs por *Sprint* dos times A e D se destaca em relação aos demais. O time B obteve a menor média de bugs por *Sprints* relatada, sendo 0.5 a cada ciclo da *Sprint*. Já o Time C, alcançou uma média de 1.5 bugs

por *Sprint*, seguido do time E com 2.1 bugs por *Sprints*. Em relação aos débitos técnicos (MQ2), todas as equipes alcançaram valores baixos, e o time A não obteve nenhum débito técnico. Nota-se que o time com a maior qualidade de produto é o time B, sendo um dos times mais diversos analisados no estudo, com diversidade de gênero, orientação, cor/raça e estado.

De forma a complementar as análises técnicas também para esta dimensão de qualidade, incluindo-se as percepções individuais dos integrantes dos times no que diz respeito ao percentual de débitos técnicos, os resultados da métrica MQ2 foram comparados com as respostas do formulário, conforme ilustra a Tabela 3.

Tabela 3. Débitos Técnicos - Percepção individual versus Dados.

Time	Como você classifica os débitos técnicos do seu squad?								VSTS (%)
	PD1	PD2	PD3	PD4	PD5	PD6	PD7	PD8	
A	5	3	5	3	4	4	-	-	0
B	2	5	2	-	-	-	-	-	0,375
C	3	4	4	-	-	-	-	-	0,375
D	4	4	3	5	4	4	3	5	0,2
E	4	5	4	3	-	-	-	-	0,33

Segundo os dados do VSTS, os times A e D apresentaram os melhores resultados, com 0 e 0.2% de débitos técnicos, mas existe uma percepção diferente dos participantes, que responderam com valores altos (maior ou igual a 3). Novamente aqui, parece haver alguma falha na comunicação e na divulgação dos dados da ferramenta para padronizar o conhecimento de todo o time.

Os times B, C e E apresentaram uma taxa de débitos técnicos em torno de 30 a 40%. Para os times C e E, parece haver uma coerência na percepção dos integrantes (PDs), que indicaram valores altos (maior ou igual a 3). Já para o time B, os resultados comparativos mostraram-se mais discrepantes, pois dos três membros, dois selecionaram a resposta 2 na escala Likert e um deles, a resposta 5. Embora, na média, o resultado seja próximo ao do VSTS, isso mostra uma percepção não uniforme entre os membros, que pode ser originada também por problemas de comunicação e não divulgação dos relatórios da ferramenta para todos os membros. De qualquer forma é importante observar que B tem o melhor resultado para a MQ2 e é um dos times mais diversos.

5.4. QP4: Existe alguma relação entre a presença de *Soft Skills* e o gênero da pessoa desenvolvedora?

Inicialmente, a investigação desta questão abrangeu algumas adaptações na replicação das análises anteriores [Motta and Braga e Silva 2022], com alguns ajustes e ampliação da amostra para 5 times. Para isso, foram utilizadas as seguintes perguntas da terceira dimensão do formulário, que abordavam aspectos como a capacidade de comunicação dos membros e a qualidade das interações:

1. Faço uma boa leitura emocional do meu time (observando mudanças comportamentais e expressivas que podem simbolizar algum desconforto).
2. Você considera satisfatória sua interação com seu squad?
3. Sinto que minhas ideias são recebidas e entendidas, pelo meu squad, da maneira que foram expressas.
4. Costumo compartilhar minhas dificuldades/impedimentos nas reuniões diárias (“*dailys*”) de forma honesta.
5. Em condições desfavoráveis ou desafiadoras, consigo manter meu bom desempenho.
6. Consigo perceber e ajudar se algum membro do time possui algum impedimento.

Segundo as percepções dos respondentes, todos os times consideraram ter boa interação com seus colegas de trabalho e saber compartilhar suas dificuldades e impedimentos nas reuniões diárias de forma honesta. Ainda com bons resultados, na pergunta sobre as condições desfavoráveis ou desafiadoras (5), a maioria considerou que consegue manter seu bom desempenho. Os times B, C e E consideram que conseguem perceber e ajudar se algum membro possui algum impedimento, com média 4, já os times A e D obtiveram uma média 5. Por fim, os times C e E, que apresentam o menor grau de diversidade (formados apenas por homens), também apresentaram a menor confiança em relação às *soft skills* avaliadas.

Acredita-se que a avaliação positiva dos times em relação à presença das *Soft Skills* pode ter alguma relação com o perfil de diversidade, uma vez que todos eles apresentaram em maior ou menor grau algum índice de diversidade em alguma das categorias consideradas. No entanto, alguns dos resultados obtidos para as métricas de produtividade (QP1) e qualidade (QP2) mostraram-se divergentes das percepções dos times nas dimensões **Perfil de Diversidade** e **Perfil Técnico** do formulário, indicando, por exemplo, uma falta de transparência no processo e até mesmo uma subutilização e/ou divulgação dos relatórios gerados pela ferramenta de apoio.

Com o intuito de aprofundar os achados preliminares e mais abrangentes sobre a relação entre *Soft Skills* e diversidade [Motta and Braga e Silva 2022], neste trabalho, essa questão é investigada mais pontualmente, a partir de um recorte de gênero. Para isso, as respostas para a pergunta “Dentre as habilidades não técnicas (*Soft Skills*) desejadas no contexto das metodologias ágeis, quais você considera possuir?” foram analisadas segundo o gênero do respondente, conforme mostra a Tabela 4.

A partir dos resultados, observa-se que as mulheres, no geral, não se consideram criativas (25% delas apenas) mas têm boas classificações em relação às outras habilidades, ficando com 50% dos votos apenas em “Comunicação” e “Liderança e influência”. Os resultados dos homens mostraram, em sua maioria, insegurança na habilidade de “Liderança e influência”, sendo essa *skill* a menos votada por eles com apenas 40% dos votos. A segunda habilidade com o menor número de votos masculinos foi a “Criatividade”. Essas *skills* são extremamente relevantes no contexto do trabalho em equipe, especialmente em

Tabela 4. Soft Skills versus Gênero

<i>Soft Skill</i>	#Mulheres	#Homens
Capacidade de lidar com conflitos	3/4 (75%)	11/20 (55%)
Autonomia e independência	3/4 (75%)	12/20 (60%)
Compreensão ética	4/4 (100%)	17/20 (85%)
Liderança e influência	2/4 (50%)	8/20 (40%)
Autogerenciamento	3/4 (75%)	12/20 (60%)
Comunicação	2/4 (50%)	15/20 (75%)
Facilidade de trabalhar em equipe	3/4 (75%)	20/20 (100%)
Aprendizagem contínua	3/4 (75%)	18/20 (90%)
Flexibilidade	3/4 (75%)	19/20 (95%)
Criatividade	1/4 (25%)	9/20 (45%)

times ágeis. O fato da “Comunicação” não estar com destaque (apenas 50% das mulheres e 75% dos homens) reflete, em alguns times, as inconsistências entre os resultados das métricas e as percepções individuais dos membros, já apontadas anteriormente.

6. Ameaças à Validade

Dentre as ameaças à validade, destacam-se: o tamanho da amostra de dados utilizada e a participação de times de uma única empresa; a não aplicação de algumas métricas por falta de dados de alguns times; limitações nas comparações entre os times, devido à ausência de times pouco/nada diversos, que pode ser explicado pelo estudo ter sido realizado no contexto de uma empresa com políticas inclusivas e o contexto de trabalho remoto presente na mesma; e falta de uma análise estatística para se avaliar correlações entre a diversidade dos times e os índices de produtividade e qualidade.

7. Considerações Finais

A falta de times diversos no contexto da Engenharia de Software é um débito social importante, que vem sendo discutido na literatura e trabalhado no âmbito de diversas empresas no mercado de TI. Embora para uma amostra relativamente pequena, extraída de uma única empresa, as análises apresentadas neste estudo também corroboram com a importância da diversidade em times de software, em especial, no contexto de metodologias ágeis. De forma geral, os resultados de produtividade e qualidade dos 5 times foram positivos e mostraram alguma relação com os bons índices de diversidade de todos.

De forma mais específica, após as análises guiadas pela investigação das quatro questões de pesquisa propostas, foi possível perceber que: as categorias de diversidade com menor representação nos times corresponderam, como esperado e já sabido, a grupos minoritários como mulheres e pessoas com deficiência - PCD (QP1); a produtividade variou em função do índice de diversidade da equipe (QP2), uma vez que os times que mais se destacaram, segundo as métricas aplicadas, foram também os mais diversos; em relação às métricas de qualidade (QP3), os resultados mostraram que o time com maior destaque também tinha o maior grau de diversidade; por fim, com relação à QP4, observou-se que apesar da pequena amostra e da baixa representatividade nos times, as

percepções das mulheres quanto às suas *Soft Skills* mostraram-se animadoras, ainda que elas tenham apresentado insegurança em termos de criatividade e liderança e influência. Além disso, as percepções individuais das mesmas sobre suas habilidades de comunicação mostraram-se piores que as dos homens. Esse resultado mostra-se curioso e precisa ser investigado para uma amostra mais representativa de times e cruzando-se com uma análise sobre dados de comunicação dos projetos, pois conforme já observado por Catolino et al. [Catolino et al. 2019], as mulheres são importantes mediadoras e facilitadoras no processo de comunicação. Por fim, um achado importante nas análises realizadas diz respeito a algumas falhas de comunicação dentro dos times que merece atenção, uma vez que houve divergência das percepções individuais dos resultados de produtividade e qualidade em comparação com os dados da ferramenta de suporte. Como em metodologias ágeis, tem-se uma abordagem orientada a pessoas, que valoriza fortemente as interações sociais, tem-se um gargalo importante que precisa ser gerenciado.

Para trabalhos futuros, é necessário: ampliar as amostras e realizar o estudo em diferentes empresas; considerar as características dos projetos e a complexidade dos produtos em desenvolvimento; aprofundar as análises estatísticas sobre os dados, de forma a avaliar a existência de possíveis correlações entre os índices de diversidade e os indicadores de produtividade e qualidade; e realizar análises com outros recortes para as categorias de diversidade consideradas.

8. Disponibilização dos Artefatos

Para fins de validação e replicação do estudo em outros contextos, os artefatos usados na condução deste estudo, incluindo o formulário com as perguntas e os dados analisados, encontram-se disponíveis em um repositório do GitHub⁵.

Referências

- [Albusays et al. 2021] Albusays, K., Bjorn, P., Dabbish, L., Ford, D., Murphy-Hill, E., Serebrenik, A., and Storey, M. (2021). The diversity crisis in software development. *IEEE Software*, 38(02):19–25.
- [Barbosa et al. 2021] Barbosa, M., Brilhante, I., Andrade, A., Conceição, J. L., Gomes, G., Conte, T., and Gadelha, B. (2021). Diversidade de gênero elicitación de requisitos: Uso do gendermag como estratégia de identificação de requisitos de usabilidade. In *Workshop On Requirements Engineering*, volume 24.
- [Beck et al. 2001] Beck, K., Beedle, M., Van Bennekum, A., Cockburn, A., Cunningham, W., Fowler, M., Grenning, J., Highsmith, J., Hunt, A., Jeffries, R., et al. (2001). Manifesto for agile software development.
- [Blau 1977] Blau, P. (1977). *Inequality and Heterogeneity: A Primitive Theory of Social Structure*, volume 7. Free Press.
- [Burgstahler and Ladner 2007] Burgstahler, S. and Ladner, R. (2007). Increasing the participation of people with disabilities in computing fields. *Computer (Long Beach, Calif.)*, 40(5):94–97.

⁵<https://github.com/taianemotta/Dados>

- [Canedo et al. 2019] Canedo, E., Tives, H., Bogo, M., Fagundes, F., and Siqueira de Cerqueira, J. (2019). Barriers faced by women in software development projects. *Information*, 10:309.
- [Catolino et al. 2019] Catolino, G., Palomba, F., Tamburri, D. A., Serebrenik, A., and Ferrucci, F. (2019). Gender diversity and women in software teams: How do they affect community smells? In *2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Society (ICSE-SEIS)*, pages 11–20, Montreal, QC, Canada. IEEE.
- [Cesar 2017] Cesar, E. (2017). *Fatores de influência na produtividade de desenvolvedores de organizações de software*. PhD thesis, Instituto de Computação / UFAM.
- [Collins 1979] Collins, R. (1979). Inequality and heterogeneity: A primitive theory of social structure. by peter m. blau. new york: Free press, 1977. *Social Forces*, 58(2):677–683.
- [Diegmann and Rosenkranz 2017] Diegmann, P. and Rosenkranz, C. (2017). Team diversity and performance-how agile practices and psychological safety interact. In *International Conference on Information Systems (ICIS)*.
- [Hernández et al. 2019] Hernández, G., Martinez, , Jimenez, R., and Jimenez, F. (2019). Productivity metrics for an agile software development team: A systematic review. *TecnoLógicas*, 22:63 – 81.
- [Kelmara M. Vieira 2008] Kelmara M. Vieira, M. D. (2008). Dilemas na construção de escalas tipo likert: o número de itens e a disposição influenciam nos resultados? v. 6 n. 3 (2013): *Edição Especial Epistemologia e Métodos de Pesquisa em Administração e Contabilidade*.
- [Kitchenham et al. 2009] Kitchenham, B., Pearl Brereton, O., Budgen, D., Turner, M., Bailey, J., and Linkman, S. (2009). Systematic literature reviews in software engineering – a systematic literature review. *Information and software technology*, 51(1):7–15.
- [Melo et al. 2011] Melo, C., Cruzes, D. S., Kon, F., and Conradi, R. (2011). Agile team perceptions of productivity factors. In *2011 Agile Conference*, pages 57–66.
- [Motta and Braga e Silva 2022] Motta, T. and Braga e Silva, G. (2022). Análise qualitativa sobre a influência da diversidade na produtividade de equipes ágeis: um estudo na indústria. In *Anais do VII Workshop sobre Aspectos Sociais, Humanos e Econômicos de Software*, pages 31–40, Porto Alegre, RS, Brasil. SBC.
- [Rodríguez-Pérez et al. 2021] Rodríguez-Pérez, G., Nadri, R., and Nagappan, M. (2021). Perceived diversity in software engineering: a systematic literature review. *Empirical Software Engineering*, 26(5):1–38.

- [Sarmiento et al. 2022] Sarmiento, C., Massoni, T., Serebrenik, A., Catolino, G., Tamburri, D., and Palomba, F. (2022). Gender diversity and community smells: A double-replication study on brazilian software teams. In *2022 IEEE International Conference on Software Analysis, Evolution and Reengineering (SANER)*, pages 273–283.
- [Silveira and Prikladnicki 2018] Silveira, K. K. and Prikladnicki, R. (2018). Perceptions on diversity in brazilian agile software development teams: A survey. In *Proceedings of the 1st International Workshop on Gender Equality in Software Engineering, GE '18*, page 37–40, New York, NY, USA. Association for Computing Machinery.
- [Silveira and Prikladnicki 2019] Silveira, K. K. and Prikladnicki, R. (2019). A systematic mapping study of diversity in software engineering: a perspective from the agile methodologies. In *Proceedings of the 12th International Workshop on Cooperative and Human Aspects of Software Engineering (CHASE), CHASE '19*, pages 7–10. IEEE, IEEE.
- [Spingola 2018] Spingola, E. M. (2018). Literature review on disability participation in the engineering field. In *2018 ASEE Annual Conference & Exposition, Salt Lake City, Utah*. ASEE Conferences. <https://peer.asee.org/30776>.
- [Vasilescu et al. 2015] Vasilescu, B., Posnett, D., Ray, B., van den Brand, M. G., Serebrenik, A., Devanbu, P., and Filkov, V. (2015). Gender and tenure diversity in github teams. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI '15*, page 3789–3798, New York, NY, USA. Association for Computing Machinery.
- [Vedres and Vasarhelyi 2019] Vedres, B. and Vasarhelyi, O. (2019). Gendered behavior as a disadvantage in open source software development. *EPJ Data Science*, 8.
- [Zacchiroli 2021] Zacchiroli, S. (2021). Gender differences in public code contributions: A 50-year perspective. *IEEE Software*, 38(2):45–50.