



RECURSOS PARA A AVALIAÇÃO DE EFICÁCIA DE UM PROGRAMA KAIZEN DIRECIONADO À REDUÇÃO DO LEAD TIME

RESOURCES FOR THE EVALUATION OF THE EFFECTIVENESS OF A KAIZEN PROGRAM AIMED AT REDUCING LEAD TIME

RECURSOS PARA LA EVALUACIÓN DE LA EFICACIA DE UN PROGRAMA KAIZEN DIRIGIDO A LA REDUCCIÓN DEL LEAD TIME

Thiago Loretti Arico Martins^{1*} & Marilia Macorin de Azevedo²

^{1,2}Centro Estadual de Educação Tecnológica Paula Souza

¹thiago.martins@cpspos.sp.gov.br ²marilia.azevedo@cpspos.sp.gov.br

ARTIGO INFO.

Recebido: 30.03.2025

Aprovado: 20.05.2025

Disponibilizado: 06.06.2025

PALAVRAS-CHAVE: Kaizen. Mann-Whitney U. Lead time. Teste de normalidade. Histograma.

KEYWORDS: Kaizen. Mann-Whitney U. Lead time. Normality test. Histogram.

PALABRAS CLAVE: Kaizen. Prueba U de Mann-Whitney. Lead time. Prueba de normalidad. Histograma.

***Autor Correspondente:** Martins, T. L. A.

RESUMO

Este artigo propõe recursos para a avaliação da eficácia de um programa Kaizen na redução do lead time em um processo de confecção de uniformes profissionais. Por meio de um estudo de caso realizado na empresa Alpha, foram comparados os períodos anteriores e posteriores à adoção do programa Kaizen, com o objetivo de mensurar seus impactos no desempenho produtivo. A análise concentrou-se em dados não paramétricos, utilizando ferramentas estatísticas como estatísticas descritivas, histogramas, gráficos de séries temporais, a verificação da normalidade dos dados com o teste Anderson-Darling, bem como a aplicação do teste não paramétrico de Mann-Whitney para investigar potenciais diferenças entre os cenários avaliados. Os resultados apontaram uma redução no lead time e uma tendência decrescente ao longo do tempo, embora a análise estatística não tenha evidenciado diferenças significativas. Diante da inexistência de critérios padronizados para a mensuração objetiva da eficácia do Kaizen, o estudo propõe um conjunto de instrumentos analíticos que possibilitam avaliar de forma sistemática as melhorias em processos produtivos. Dessa forma, contribui-se para o avanço das pesquisas sobre melhoria contínua, oferecendo subsídios teóricos e práticos aplicáveis em contextos industriais semelhantes.

ABSTRACT

This article proposes resources for evaluating the effectiveness of a Kaizen program in reducing lead time in a process for manufacturing professional uniforms. Through a case study conducted at Alpha, the periods before and after the adoption of the Kaizen program were compared,

with the aim of measuring its impacts on production performance. The analysis focused on nonparametric data, using statistical tools such as descriptive statistics, histograms, time series graphs, verification of data normality with the Anderson-Darling test, as well as the application of the nonparametric Mann-Whitney test to investigate potential differences between the scenarios evaluated. The results indicated a reduction in lead time and a decreasing trend over time, although the statistical analysis did not reveal significant differences. Given the lack of standardized criteria for the objective measurement of the effectiveness of Kaizen, the study proposes a set of analytical instruments that allow for the systematic evaluation of improvements in production processes. In this way, it contributes to the advancement of research on continuous improvement, offering theoretical and practical support applicable in similar industrial contexts.

RESUMEN

Este artículo propone recursos para evaluar la efectividad de un programa Kaizen en la reducción del tiempo de entrega en un proceso de fabricación de uniformes profesionales. A través de un estudio de caso realizado en Alpha, se compararon los períodos previos y posteriores a la adopción del programa Kaizen, con el objetivo de medir su impacto en el rendimiento de la producción. El análisis se centró en datos no paramétricos, utilizando herramientas estadísticas como estadística descriptiva, histogramas, gráficos de series temporales, verificación de la normalidad de los datos mediante la prueba de Anderson-Darling, así como la aplicación de la prueba no paramétrica de Mann-Whitney para investigar las posibles diferencias entre los escenarios evaluados. Los resultados indicaron una reducción del tiempo de entrega y una tendencia decreciente a lo largo del tiempo, aunque el análisis estadístico no reveló diferencias significativas. Dada la falta de criterios estandarizados para la medición objetiva de la efectividad de Kaizen, el estudio propone un conjunto de instrumentos analíticos que permiten la evaluación sistemática de las mejoras en los procesos de producción. De esta manera, contribuye al avance de la investigación sobre mejora continua, ofreciendo soporte teórico y práctico aplicable en contextos industriales similares.



INTRODUÇÃO

A busca por eficiência e otimização de processos é essencial para a competitividade na indústria de vestuário, onde agilidade na entrega, controle de custos e conformidade com o mercado visam a sustentação das empresas. No Brasil, o setor têxtil e de confecção tem grande relevância econômica, sendo o segundo maior gerador de empregos no Brasil (Associação Brasileira da Indústria Têxtil e de Confecção [ABIT], 2023).

Diante desse cenário, o sistema *lean manufacturing*, especialmente a ferramenta *Kaizen*, tem se mostrado eficaz na redução de desperdícios e aumento da eficiência (Womack, Jones, & Roos, 2004). O *Kaizen*, que significa “mudar para melhor”, surgiu na Toyota e se consolidou como método de melhoria contínua. Sua aplicação inclui ajustes incrementais em processos produtivos, *Just in Time* (JIT), automação e uso eficiente de recursos.

Contudo, medir os ganhos do *Kaizen* nem sempre é uma atividade direta, podendo abranger redução do tempo de ciclo, aumento da produtividade e diminuição de defeitos (Martins & Azevedo, 2023). Martins & Azevedo (2023) verificaram em uma revisão sistemática da literatura não padronização no que diz respeito à avaliação da eficácia de *Kaizens* e, em especial, a avaliação quanto à redução de *lead time*. No estudo de caso na empresa Alpha, um programa *Kaizen* recém-implementado no processo de costura resultou na redução do *lead time*, levantando uma questão fundamental: Quais ferramentas podem ser usadas para comparar dados de *lead time* antes e depois da implementação de um programa *Kaizen* voltado à sua redução?

Para responder a essa questão, definiu-se o seguinte objetivo geral: selecionar ferramentas que permitam inferir sobre o comportamento do *lead time* antes e depois do *Kaizen*. Como suporte, estabeleceu-se o objetivo específico de avaliar a distribuição dos dados de *lead time* por peça, possibilitando a escolha de ferramentas adequadas.

FUNDAMENTAÇÃO

P-VALOR

O P-Valor mede a compatibilidade dos dados com a hipótese nula, indicando a probabilidade de obter os resultados observados sob essa premissa (Cohen, 2011; Goodman, 1999). Valores baixos sugerem inconsistência com a hipótese nula, enquanto valores altos indicam compatibilidade. Goodman (1999) define o P-Valor como a chance de observar um resultado tão extremo quanto, ou ainda mais, do que aquele obtido, assumindo que a hipótese nula seja verdadeira. Ele destaca que um P-Valor de 0,05 não significa que há 5% de chance de a hipótese nula ser falsa, sendo essa uma interpretação equivocada. Além disso, o P-Valor não considera o tamanho do efeito observado, levando à preferência por intervalos de confiança.

Cohen (2011) ressalta que o P-Valor é apenas uma estimativa, sujeita a erros, pois modelos estatísticos possuem pressupostos que raramente são atendidos integralmente. Goodman (1999) apresenta essa relação por meio da curva normal, onde α (nível de significância) é definido antes do experimento, enquanto o P-Valor surge após a análise dos dados.

TESTE DE NORMALIDADE DE ANDERSON-DARLING

Engmann e Cosineau (2011) apresentam o teste de Anderson-Darling como uma alternativa para detectar desvios das distribuições amostrais da normalidade, sendo inicialmente direcionado para aplicações de engenharia. A estatística do teste de Anderson-Darling para



CC BY 4.0
DEED
Attribution 4.0
International

uma amostra não segue uma direção específica. No teste Anderson-Darling (AD), a estatística calculada é comparada a valores críticos para um nível de significância (α). Se AD exceder o valor crítico, há evidências para rejeitar a hipótese nula, indicando que a amostra não segue a distribuição teórica esperada.

TESTE DE MANN-WHITNEY U

O teste de Mann-Whitney U é um método não paramétrico usado para comparar duas amostras independentes de dados contínuos ou ordinais, avaliando se há diferença significativa entre elas (Oti, Olusola, & Esemokumo, 2021). Aplicado quando as suposições da ANOVA não são atendidas (Siegel & Castellan, 1988), segue as etapas (Mann & Whitney, 1947): coleta de amostras independentes, ordenação dos valores, cálculo da estatística U, obtenção do teste U, determinação do P-Valor e tomada de decisão (se P-Valor $< \alpha$, há diferença significativa).

Útil para distribuições não normais, Mann e Whitney (1947) demonstraram sua aplicabilidade comparando o tempo de sobrevivência de ratos tratados e não tratados com um medicamento, reforçando seu uso em diversas áreas.

METODOLOGIA

Esta pesquisa utilizou o método de estudo de caso, reconhecido por sua capacidade de investigar fenômenos complexos em contextos reais (Yin, 2001; Stake, 1995). As etapas propostas por Yin (2001) são:

- Projeto do estudo de caso: trata da criação de questões gerais e critérios para julgar a qualidade dos projetos de pesquisa.
- Preparação para coleta de dados: envolve a obtenção de dados de múltiplas fontes, como entrevistas, documentos, observações e artefatos. Isso inclui a definição do protocolo de pesquisa, que deve conter uma visão geral do projeto do estudo de caso, procedimentos de campo, questões do estudo de caso e guia para o relatório.
- Coleta de evidências: explora os princípios e fontes para coletar evidências.
- Análise das evidências do estudo de caso: aborda estratégias e métodos de análise.
- Composição de um relatório de estudo de caso: cobre a estrutura e o procedimento para criar um relatório.

Todo o roteiro proposto por Yin (2001) foi adotado para esta pesquisa quanto ao protocolo para estudo de caso, pois esse protocolo garante rigor metodológico, abordando desde a concepção do estudo até a análise e redação final (Yin, 2001). A coleta de dados é destacada como etapa central, exigindo acuracidade em campo e uma análise criteriosa das evidências.

PROJETO DO ESTUDO DE CASO

O estudo se baseou na questão de pesquisa: Quais ferramentas podem ser utilizadas para comparar os dados de *lead time* antes e após a implementação de um programa *Kaizen*?

A coleta de dados foi planejada a ocorrer diretamente no *Enterprise Resource Planning* (ERP) Microsoft Dynamics AX® da empresa Alpha, garantindo confiabilidade interna. Para evitar distorções, os pesquisadores planejaram verificação de inconsistências na base de dados, decidindo pela exclusão ou manutenção de registros em conjunto com a gestão do processo de costura. Os modelos de dados planejados incluíram produto, fornecedor, datas e horários de início e término, e quantidade.

Entre as cinco fábricas da empresa Alpha, escolheu-se a unidade de Jandira/SP devido à recente implementação do *Kaizen* e à proximidade dos pesquisadores. A unidade opera com 32 máquinas de costura, com 16 costureiras e 1 líder de costura. A coleta de dados foi planejada a ocorrer semanalmente no ERP Microsoft Dynamics AX®, de 2 de outubro a 4 de dezembro de 2023, totalizando 10 semanas, com participação dos pesquisadores e da líder do processo.

O projeto de pesquisa iniciou-se com a formulação da questão de pesquisa e a definição dos objetivos gerais e específicos. Em seguida, foram estabelecidos os critérios para a coleta de dados, incluindo a seleção das categorias relevantes de acordo com o protocolo estabelecido na seção Metodologia, a validação junto à gestão do processo, a verificação de inconsistências e o uso de dados do ERP para garantir a validade dos resultados. A unidade de costura em Jandira/SP foi inserida no protocolo de estudo de caso.

A análise planejada envolveu a verificação da distribuição e normalidade, além da classificação dos dados quanto ao tipo, como discretos, ordinais, contínuos e binários. Por fim, foi planejada a seleção das ferramentas analíticas adequadas e testes estatísticos, culminando na elaboração do relatório final. Tais macroestágios apoiam-se na proposta de Yin (2001) para sua concepção.

PREPARAÇÃO PARA COLETA DE DADOS

No planejamento da coleta de dados, previu-se a extração direta do ERP da empresa Alpha, com a exportação dos dados em formato .xls para posterior seleção e limpeza no Microsoft Excel®. A análise estatística foi planejada para ser conduzido no Minitab®, escolhido devido à familiaridade dos pesquisadores com a ferramenta e sua integração com o Excel.

Definiu-se que a primeira coleta abrangeeria o mês de setembro, representando o período pré-*Kaizen*, permitindo comparações com os dados pós-implementação. Conforme o protocolo estabelecido, a extração dos dados ocorreria às segundas-feiras, excluindo registros do mesmo dia e considerando a semana anterior (segunda a sexta), com exceção de feriados. Para garantir robustez estatística, estabeleceu-se que a coleta seria realizada por um período de 10 semanas, formando um conjunto de dados significativo para a comparação dos cenários antes e após o *Kaizen*. Planejou-se que a distribuição dos *leads times* seria analisada utilizando ferramentas adequadas para inferência sobre seu comportamento.

Considerou-se que todos os recursos necessários para a coleta e análise dos dados estariam disponíveis, e foram previstas medidas para garantir a ética do processo. Avaliou-se que a realização de uma coleta piloto não seria necessária, dado que o procedimento poderia ser replicado de forma confiável. Tal decisão partiu da premissa de que, por se tratar de uma coleta de dados direta no ERP da empresa, a replicabilidade da coleta poderia ser feita a qualquer tempo.

COLETA DE EVIDÊNCIAS

A coleta de dados começou na primeira segunda-feira de outubro de 2023, abrangendo registros de 1º a 29 de setembro de 2023. Foram analisados possíveis erros de digitação e falhas de apontamento, resultando na exclusão de 13 registros com campos vazios ou caracteres inválidos.

Além disso, identificou-se que alguns registros apresentavam tempos irrealisticamente curtos entre o início e o fim da produção. Em consenso com a gestão do processo, foram excluídos



20 registros com tempo inferior a 0,01 hora, considerado inviável para a realização de uma operação e entendeu-se como falha de apontamento.

No total, 33 registros de produção foram descartados por inconsistências. No mês de setembro/2023, 777 registros foram considerados válidos para a análise.

As coletas subsequentes ocorreram em 09/10, 16/10, 23/10, 30/10, 06/11, 13/11, 20/11, 27/11 e 04/12. Nessas coletas, 129 registros foram descartados por dados inválidos ou falhas de apontamento. Com a finalização da fase de coleta, iniciou-se a análise das evidências do estudo de caso (Quadro 1).

Quadro 1. Compilação das ideias /projetos aprovados e concluídos de outubro a novembro de 2023

Ideia	Situação encontrada	Situação desejada	Aprovado?	Concluído
1 - Substituição dos cabides de arremate por mesa	Alta fadiga das colaboradoras, o que influí na produtividade durante o dia	Menor fadiga e maior produtividade durante o dia	Sim	Outubro
2 - Retreinamento da equipe quanto a 5S's e interpretação de fichas técnicas	Falhas quanto aos 5S's e atrasos/ retrabalhos devido dificuldades de interpretação das fichas técnicas	5S's adotados. Facilidade de interpretação de fichas técnicas	Sim	Novembro
3 - Elaboração de ficha para acompanhamento da produção	Irrastreabilidade de quem executou a operação, o que torna difícil orientar a costureira em caso de erro.	Possibilidade de rastreio quanto à realização da operação	Sim	Novembro
4 - Caixa para organização das matrizes de fixação de botões	Demora na localização das matrizes corretas/ perda de matrizes de fixação de botões	Localização padronizada e corretamente identificada das matrizes de fixação de botões	Sim	Outubro
5 - Adoção de esquadro ajustável para posicionamento de patch	Método com esquadro padrão leva tempo demais e maior possibilidade de erros	Método de marcação de posição de patch mais rápido e mais preciso	Sim	Outubro
6 - Identificação visual das produções a serem realizadas no dia ou urgentes	A identificação de produções de outros dias permite que as costureiras adotem um ritmo mais lento, caso percebam que estão adiantadas	A identificação somente das produções que devem ser produzidas no dia permite que a liderança tenha maior controle sobre as entregas do dia	Sim	Outubro
7 - Criação de <i>kit</i> de ferramentas da costureira	Costureiras utilizam ferramentas compartilhadas e com localização não-padronizada, o que gera perdas e atrasos para localização de ferramentas.	Cada costureira com seu <i>kit</i> , permitindo ajustes e set-ups rápidos de máquina	Sim	Novembro
8 - Adoção de gerenciamento visual de metas e resultados	Não há gerenciamento, report à equipe ou cobrança baseada em metas	Gerenciamento e report à equipe baseado em metas/ maior comprometimento da equipe com resultados	Sim	Novembro
9 - Organizador para os gabaritos de passadoria	Desorganização quanto aos gabaritos de passadoria, o que causa atrasos para localização e perdas	Gabaritos em posições padronizadas, permitindo a rápida e fácil localização	Sim	Novembro

Fonte: Autores (2025).

Com os projetos *Kaizen* implementados, pôde-se partir para a análise de evidências, respeitando-se assim o protocolo estabelecido na seção Metodologia.

ANÁLISE DAS EVIDÊNCIAS

Com o sucesso da coleta de dados, que ocorreu conforme o planejado, pôde-se partir para a fase de análise das evidências, onde pretendeu-se analisar a normalidade dos dados e as estatísticas descritivas de média, moda, mediana e desvio padrão. Para auxílio no trabalho com os dados obtidos, utilizou-se o *software* Minitab®, conforme planejado.

ANÁLISE DAS EVIDÊNCIAS DO MÊS DE SETEMBRO

Tratando-se primeiramente do mês de setembro, calculou-se as estatísticas descritivas planejadas (Tabela 1).

Tabela 1. Estatísticas descritivas para os dados de hora por peça do mês de setembro/2023

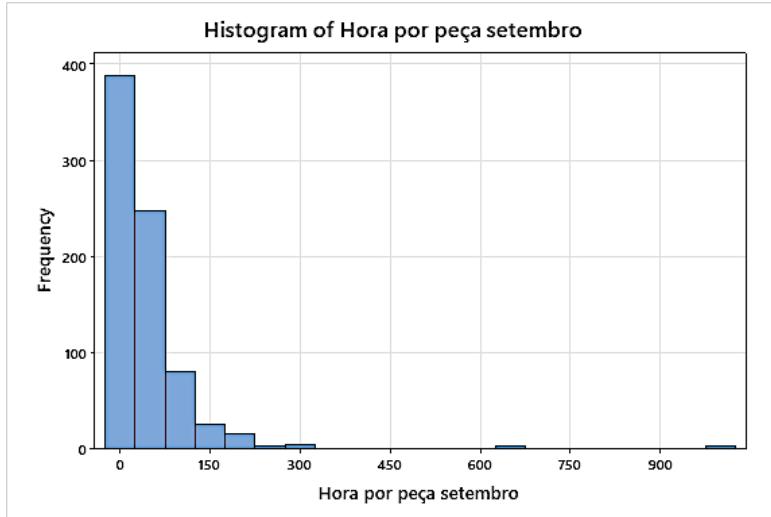
Variável	N	N*	Média	Erro padrão da média	Desvio Padrão	Mínimo	Q1	Mediana	Q3	Máximo
Horas por peça em setembro	777	0	50,92	3,34	92,99	0,07	9,8	25,34	57,3	1008

Fonte: os autores (2025).

Apresenta-se a análise da hora por peça para setembro, com N = 777 registros contínuos. Os principais valores estatísticos são: média = 50,92 h, erro padrão = 3,34 h, desvio padrão = 92,99 h, mínimo = 0,07 h, máximo = 1.008 h, 1º quartil = 9,80 h, mediana = 25,35 h e 3º quartil = 57,30 h.

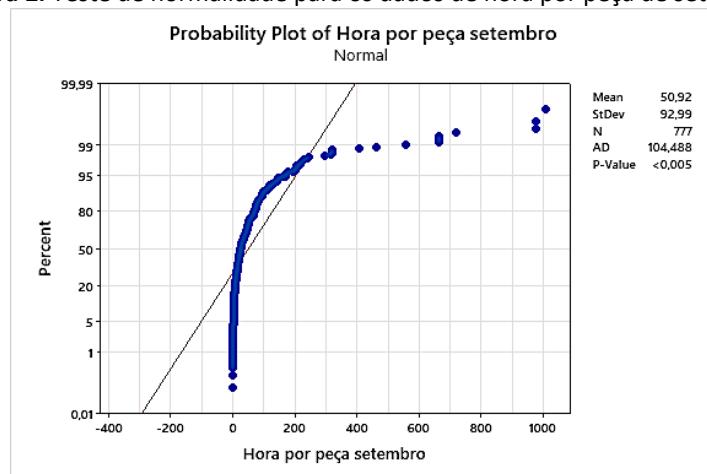
A grande dispersão dos dados sugere uma distribuição ampla, reforçada pelo alto desvio padrão. Como os dados são contínuos, utilizou-se o Histograma, recomendado por Coelho, Silva e Manicoba (2016), para representar graficamente a distribuição, centralidade e dispersão do *lead time*. O Histograma exibe intervalos de valores na base e frequência na altura dos retângulos, permitindo identificar a distribuição e sua relação com as especificações do estudo (Figura 1).

Figura 1. Distribuição de frequências de horas por peça do mês de setembro



Fonte: Autores (2025).

A Figura 1 demonstra que 743 registros (95,62%) possuem *lead times* entre 0 e 150 horas, enquanto 34 registros (4,37%) estão acima desse intervalo. Para verificar a hipótese de não normalidade, utilizou-se o teste de Anderson-Darling, recomendado por Engmann e Cosineau (2011), como ferramenta para identificar discrepâncias na distribuição dos dados. Originalmente aplicado à engenharia, o teste se mostrou adequado para a análise do *lead time* neste estudo (Figura 2).

Figura 2. Teste de normalidade para os dados de hora por peça de setembro

Fonte: Autores (2025).

Adotando-se uma significância de 5% e com um P-Valor obtido menor que 0,005, sugere-se a não normalidade dos dados coletados no mês de setembro/2023.

ANÁLISE DAS EVIDÊNCIAS DO MÊS DE OUTUBRO

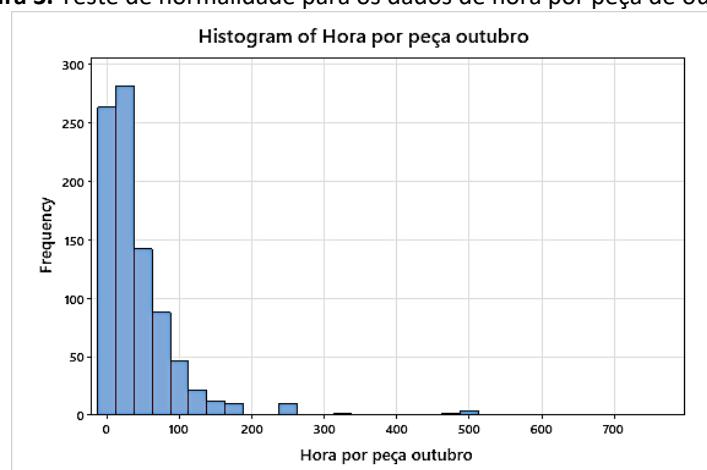
Foram calculadas as estatísticas descritivas com o auxílio do software Minitab® para o mês de outubro/2023, já com leve interferência do *Kaizen* iniciado em 02/10/2023 (Tabela 2).

Tabela 2. Estatísticas descritivas para os dados de hora por peça do mês de outubro/2023

Variável	N	N*	Média	Erro padrão da média	Desvio Padrão	Mínimo	Q1	Mediana	Q3	Máximo
Horas por peça em outubro	889	0	46,39	2,3	68,61	0,01	9,53	26,06	55,66	766,45

Fonte: Autores (2025).

A Tabela 2 apresenta a análise da hora por peça para outubro/2023, com N = 889 registros. Estatísticas principais: média = 46,39 h, erro padrão = 2,30 h, desvio padrão = 68,61 h, mínimo = 0,01 h, máximo = 766,45 h, 1º quartil = 9,53 h, mediana = 26,06 h e 3º quartil = 55,66 h. A dispersão ainda é alta, mas o desvio padrão menor (68,61 h) sugere uma leve redução em relação a setembro (92,99 h). A média caiu 4,53 h, enquanto a mediana subiu 0,72 h, indicando mudanças na distribuição dos dados (Figura 3).

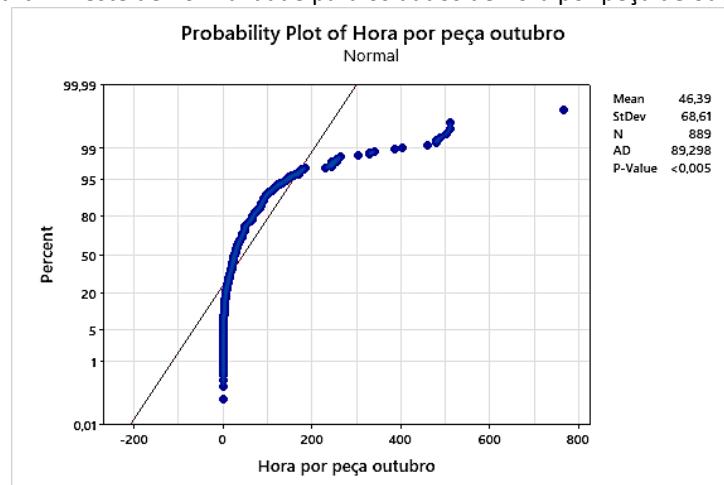
Figura 3. Teste de normalidade para os dados de hora por peça de outubro

Fonte: Autores (2025).

Mostra-se que 820 registros (92,24%) possuem *lead times* entre 0 e 100 horas, enquanto 69 registros (7,76%) estão acima desse intervalo. Não houve ocorrências acima de 600 horas,

diferentemente de setembro/2023. Para confirmar a hipótese de não normalidade, aplicou-se o teste de Anderson-Darling, conforme Yazici e Yolacan (2007) (Figura 4).

Figura 4. Teste de normalidade para os dados de hora por peça de outubro



Fonte: Autores (2025).

Com base no resultado acima, pode-se concluir que, para um P-Valor apresentado menor que 0,05 e para uma significância de 95%, assim como para o mês de setembro/2023, os dados de outubro/2023 também não apresentam normalidade.

ANÁLISE DAS EVIDÊNCIAS DO MÊS DE NOVEMBRO

Foram calculadas então, as estatísticas descritivas com o auxílio do *software* Minitab® novembro/2023, também com interferência do *Kaizen* iniciado em 02/10/2023 (Tabela 3).

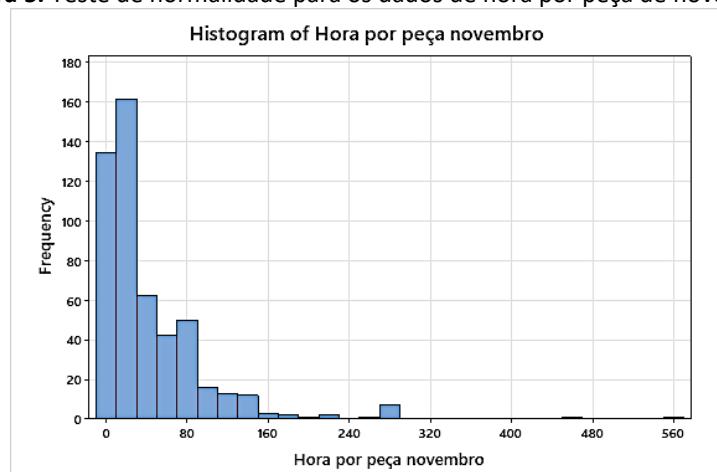
Tabela 3. Estatísticas descritivas para os dados de hora por peça do mês de novembro/2023

Variável	N	N*	Média	Erro padrão da média	Desvio Padrão	Mínimo	Q1	Mediana	Q3	Máximo
Horas por peça em outubro	508	0	43,91	2,55	57,5	0,03	9,02	23,87	60,31	557,81

Fonte: Autores (2025).

Apresenta-se a análise da hora por peça para novembro/2023, com N = 508 registros. Estatísticas principais: média = 43,91 h, erro padrão = 2,55 h, desvio padrão = 57,50 h, mínimo = 0,03 h, máximo = 557,81 h, 1º quartil = 9,02 h, mediana = 23,87 h e 3º quartil = 60,31 h. Embora o intervalo entre mínimo e máximo ainda seja amplo, o valor máximo reduziu 208,64 h em relação ao mês anterior. A redução do desvio padrão em 11,11 h sugere menor variabilidade no processo (Figura 5).

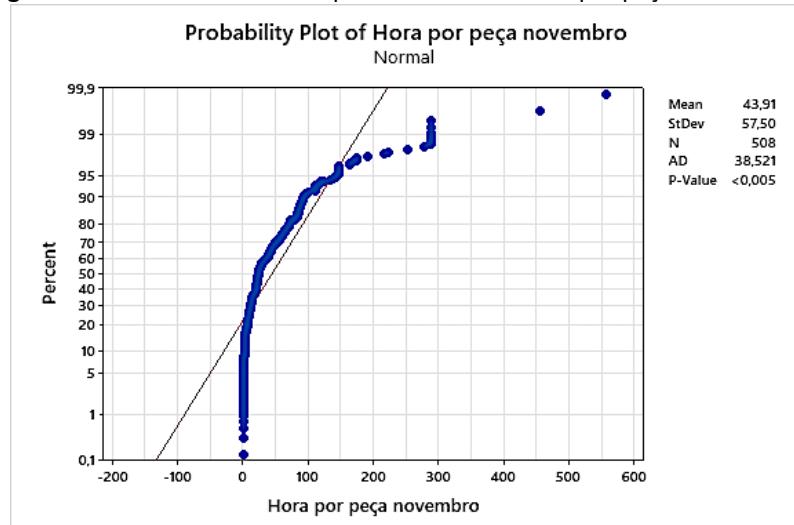
Figura 5. Teste de normalidade para os dados de hora por peça de novembro



Fonte: Autores (2025).

Apresenta-se que 449 registros (88,39%) possuem *lead times* entre 0 e 80 horas, enquanto 59 registros (11,61%) estão acima desse intervalo. Essa redução sugere a influência do *Kaizen* no processo. Para confirmar a hipótese da não normalidade, aplicou-se o teste de Anderson-Darling, conforme Yazici e Yolacan (2007) (Figura 6).

Figura 6. Teste de normalidade para os dados de hora por peça de novembro



Fonte: Autores (2025).

Com base no resultado acima, pode-se inferir que, para um P-Valor apresentado menor que 0,05 (obtido <0,005), assim como para o mês de setembro/2023 e outubro/2023, os dados de novembro/2023 também não apresentam normalidade.

COMPARATIVOS DOS DADOS DOS MESES DE SETEMBRO, OUTUBRO E NOVEMBRO DE 2023

Baseando-se nos Histogramas apresentados, observa-se nesses 3 meses analisados um estreitamento da faixa onde concentra-se a maior parte dos apontamentos (Tabela 4).

Tabela 4. Maiores faixas de ocorrência de hora por peça de setembro a novembro

Meses/2023	Faixa de maior ocorrência de <i>lead times</i> por peça
Setembro	0 a 150 horas por peça
Outubro	0 a 100 horas por peça
Novembro	0 a 80 horas por peça

Fonte: Autores (2025).

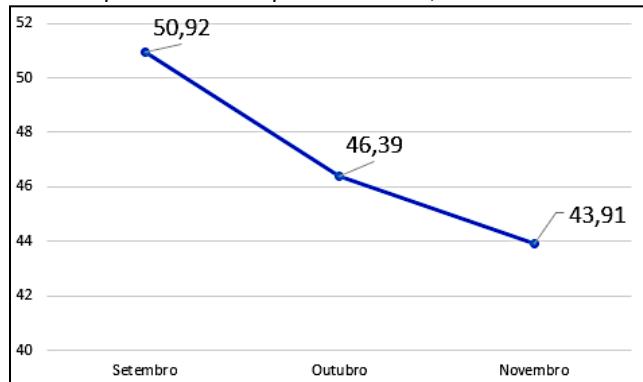
Sugere-se que há um aumento de concentração de apontamentos de produção com *lead times* menores, mês a mês, desde o início do *Kaizen*. Avaliando-se ainda os demais dados de estatística descritiva (Tabela 5), podendo sugerir que com o avanço do *Kaizen* outros ganhos podem ser observados, como a redução do desvio padrão.

Tabela 5. Demonstrativo das estatísticas descritivas de setembro a novembro

Meses/2023	População	Média	Erro padrão	Desvio padrão	Mínimo	Q1	Mediana	Q3	Máximo
Setembro	777	50,92	3,34	92,99	0,07	9,8	25,34	57,3	1008
Outubro	889	46,39	2,3	68,61	0,01	9,53	26,06	55,66	766,45
Novembro	508	43,91	2,55	57,5	0,03	9,02	23,87	60,31	557,81

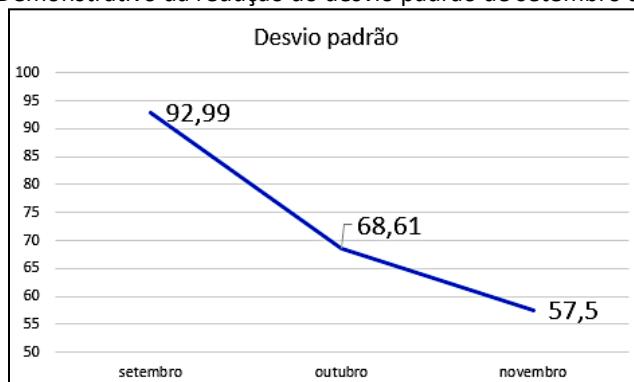
Fonte: Autores (2025).

Antunes (2015) define o gráfico de séries temporais como uma ferramenta para organizar informações quantitativas ao longo do tempo. Para este estudo, ele será utilizado na análise da média, desvio padrão, mínimo, máximo e acompanhamento da média semanal (Figura 7).

Figura 7. Série temporal da média para setembro, outubro e novembro de 2023

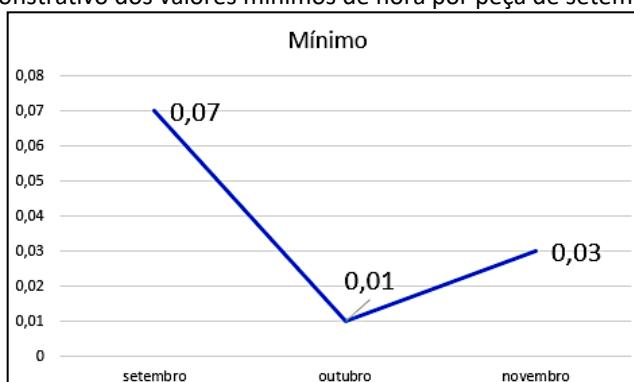
Fonte: Autores (2025).

Por se tratar de dados que compreendem toda a população, não foram calculados intervalos de confiança e utilizou-se a média direta. Pode-se concluir, assim, que houve uma redução no valor médio de *lead times* de setembro a novembro de 2023. Para o desvio padrão, pode-se também observar uma redução ao longo dos meses de pesquisa (Figura 8).

Figura 8. Demonstrativo da redução do desvio padrão de setembro a novembro

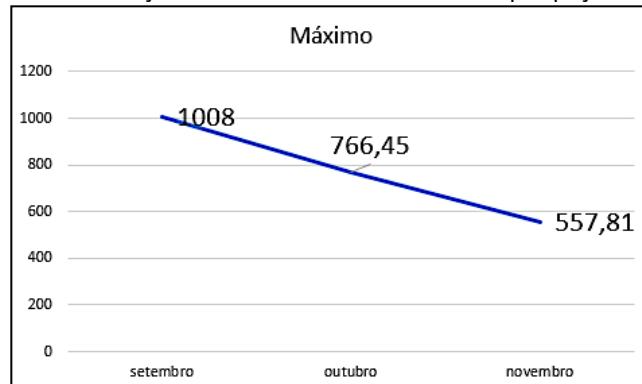
Fonte: Autores (2025).

Pode-se concluir, com base nos histogramas apresentados, que houve redução da variabilidade do processo no que tange ao *lead time* por peça observando-se os dados de setembro a outubro. Os valores de *lead time* mínimo também apresentaram redução comparando-se os meses de setembro a outubro de 2023 (Figura 9).

Figura 9. Demonstrativo dos valores mínimos de hora por peça de setembro a novembro

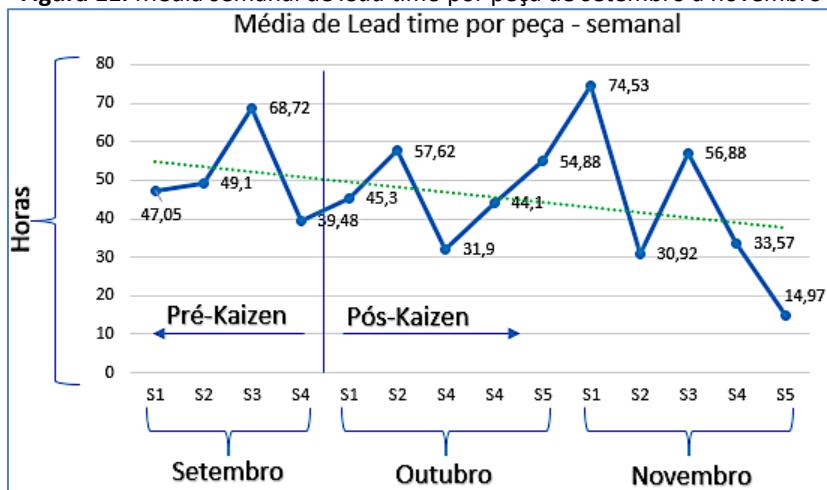
Fonte: Autores (2025).

Devido a proximidade de valores obtidos nos meses de setembro a novembro de 2023, sendo 0,07 h, 0,01 h e 0,03 h, respectivamente, entendeu-se para este estudo de caso como não significantes para julgamento. Analisando-se os valores máximos obtidos mensalmente, de setembro a novembro de 2023 pode-se verificar queda nos *leads times* (Figura 10).

Figura 10. Demonstrativo de redução dos valores máximos de hora por peça de setembro a novembro

Fonte: Autores (2025).

Pode-se verificar que desde a implementação do *Kaizen* os valores máximos de *lead time* por peça tiveram uma redução de 1.008 horas por peça em setembro para 766,45 em outubro e 557,81 em novembro. Analisando-se o período de 01 de setembro a 30 de novembro de 2023, pôde-se verificar as médias das *leads times* por peça para o processo de costura semanalmente (Figura 11).

Figura 11. Média semanal de *lead time* por peça de setembro a novembro

Fonte: Autores (2025).

Apesar da redução no *lead time*, ainda há grande variabilidade entre as semanas (S1 a S5, eixo x) em relação às horas (eixo y). A linha de tendência (pontilhada), gerada pelo Microsoft Excel, segue um modelo de regressão linear simples, que busca relacionar variáveis quantitativas (Martins, 2019). Dado que os dados não apresentam normalidade (Figuras 5, 7 e 9), foi necessário aplicar um teste estatístico adequado para comparar os cenários pré e pós-*Kaizen*.

APLICAÇÃO DO TESTE DE MANN-WHITNEY U

Considerando que os dados pré e pós-*Kaizen* atendem às premissas de não normalidade e independência, adotou-se o teste de Mann-Whitney U para verificar diferenças significativas entre os cenários analisados (Tabela 6).

Tabela 6. Estatísticas descritivas do teste de Mann-Whitney U

Estatísticas descritivas		
Amostra	N	Mediana
Pré-Kaizen	777	25,3407
Pós-Kaizen	1397	25,0119

Fonte: Autores (2025).

Mostra-se que o período pré-*Kaizen* teve 777 registros (setembro/2023), com mediana de 25,3407 h. Já o pós-*Kaizen* teve 1.397 registros (soma de outubro e novembro/2023), com mediana de 25,0119 h. O teste, realizado no Minitab® com 95% de confiabilidade, gerou o P-Valor (Tabela 7).

Tabela 7. Resultados de P-Valor para o teste de Mann-Whitney U

Teste de Mann-Whitney U		
Hipótese nula	$H_0: n_1 - n_2 = 0$	
Hipótese alternativa	$H_0: n_1 - n_2 \neq 0$	
Método	W-Valor	P-Valor
Não ajustado para empates	849833.00	0.730
Ajustado para empates	849833.00	0.730

Fonte: Autores (2025).

Indica-se que a hipótese nula é assumida quando não há diferença significativa entre as medianas de n_1 (pré-*Kaizen*) e n_2 (pós-*Kaizen*), enquanto a hipótese alternativa sugere uma diferença estatística entre os períodos. O Minitab® apresenta os resultados “não ajustado para empates” e “ajustado para empates”, mas como o P-Valor foi idêntico em ambos, conclui-se que valores repetidos não afetaram significativamente o teste. O P-Valor de 0,730, superior ao esperado ($<0,05$) para rejeição da hipótese nula, indica que não há diferença estatística significativa entre as medianas. Assim, a hipótese alternativa foi rejeitada, mantendo-se a hipótese nula.

SELEÇÃO DE FERRAMENTAS E TESTES

Verifica-se a existência de outros testes estatísticos além do teste de Mann-Whitney U, a saber, o teste t de Student (Zimmerman, 1987), a ANOVA (Boisgontier & Cheval, 2016), o teste de Wilcoxon (Barros & Mazucheli, 2005) e o teste de Kruskal-Wallis (Pontes, 2000), e a partir disso, procedeu-se à análise da aderência de cada um dos testes ao caso (Quadro 2).

Quadro 2. Comparação dos testes estatísticos, quando usar e motivo de exclusão

Teste Estatístico	Tipo de Teste	Quando Usar	Motivo de Exclusão
Teste t de Student	Paramétrico	Comparação de médias entre dois grupos com distribuição normal	Os dados de <i>lead time</i> não seguem uma distribuição normal
ANOVA	Paramétrico	Comparação de médias entre três ou mais grupos com distribuição normal	Requer mais de dois grupos para análise
Teste de Mann-Whitney U	Não Paramétrico	Comparação de medianas entre dois grupos quando a distribuição não é normal	Adequado para este estudo, pois os dados não são normais e trabalha-se com dois grupos independentes
Teste de Wilcoxon	Não Paramétrico	Comparação de pares de dados dependentes (antes e depois da intervenção)	Os dados são independentes, enquanto esse teste é para amostras pareadas
Teste de Kruskal-Wallis	Não Paramétrico	Comparação de três ou mais grupos quando a distribuição não é normal	Requer mais de dois grupos para análise

Fonte: Autores (2025).

Na análise da eficácia do *Kaizen*, constatou-se que algumas ferramentas da qualidade não foram relevantes para o estudo. O Histograma (Souza Neto et al., 2017) permitiu visualizar a distribuição dos *leads times*, e o Gráfico de Séries Temporais (Antunes, 2015) evidenciou tendências ao longo do tempo. No entanto, ferramentas como Diagrama de Pareto (Juran & Defeo, 2010), Diagrama de Causa e Efeito (Bezerra et al., 2015), Gráfico de Dispersão (Solanki et al., 2021), Folha de Verificação (Carvalho et al., 2015), Gráfico de Controle (Suman;



Prajapati, 2018) e Fluxograma (Martins & Azevedo, 2023) foram descartadas, pois o foco do estudo estava na mensuração estatística da eficácia do *Kaizen*, e não na priorização de causas ou no mapeamento do processo (Quadros 3 e 4).

Quadro 3. Comparação dos testes estatísticos

Ferramenta da Qualidade	Finalidade	Motivo de Uso ou Exclusão
Histograma	Analizar a distribuição de frequência dos <i>leads times</i> e visualizar padrões	Utilizado para observar a concentração dos <i>leads times</i> em diferentes faixas
Diagrama de Pareto	Priorizar os principais fatores que contribuem para um problema	Não utilizado , pois o foco não era priorizar causas de problemas, mas avaliar estatisticamente a eficácia do <i>Kaizen</i>
Diagrama de Causa e Efeito (Ishikawa ou Espinha de Peixe)	Identificar causas potenciais que impactam a variação do <i>lead time</i>	Não utilizado , pois o foco era medir estatisticamente a redução do <i>lead time</i> e não identificar causas raiz
Gráfico de Dispersão	Analizar a relação entre duas variáveis	Não utilizado , pois não havia necessidade de verificar a correlação entre variáveis distintas
Folha de Verificação (Check sheet)	Coletar e organizar dados de forma estruturada	Não utilizado , pois os dados foram extraídos diretamente do ERP da empresa, sem necessidade de registro manual
Gráfico de Controle (CEP - Controle Estatístico de Processo)	Monitorar a estabilidade do processo ao longo do tempo	Não utilizado , pois o objetivo do estudo foi comparar períodos distintos e não verificar se o processo estava sob controle
Fluxograma	Representar visualmente o fluxo do processo e suas etapas	Não utilizado , pois a pesquisa não tinha o objetivo de mapear o processo produtivo, mas sim de analisar a eficácia estatística do <i>Kaizen</i>
Gráfico de Séries Temporais	Acompanhar a tendência da média do <i>lead time</i> ao longo do tempo	Utilizado para identificar tendências e variações ao longo dos meses

Fonte: Autores (2025).

Quadro 4. Compilação das ferramentas selecionadas

Recurso selecionado	Base científica	Útil para a análise de cenários?	O que foi apresentado
Histograma	Coelho, Silva & Maniçoba (2016)	Sim	Aumento da frequência de ocorrência de apontamentos de produção em faixas de menor <i>lead time</i>
Teste de normalidade de Anderson-Darling	Engmann & Cosineau (2011)	Sim	Não-normalidade dos dados pré e pós- <i>Kaizen</i>
Série temporal	Antunes (2015)	Sim	Média, desvio padrão, mínimo, máximo e acompanhamento da média semanal
Análise de tendência	Martins (2019)	Sim	Tendência de redução do <i>lead time</i>
Teste de Mann-Whitney U	Ruxton (2006); Milenovic (2011); Oti et al., (2021); Tai et al., (2022)	Sim	Para período analisado, não assume hipótese alternativa, de que existe diferença estatisticamente significativa para 2 períodos

Fonte: Autores (2025).

Pode-se verificar que são apresentadas as ferramentas e testes utilizados visando-se responder à questão desta pesquisa, as referências científicas utilizadas, uma consideração dos autores quanto à utilidade das ferramentas para o propósito central e o que a ferramenta pôde apresentar para esta pesquisa.

CONSIDERAÇÕES

Este estudo teve como propósito avaliar a eficácia de um programa *Kaizen* na redução do *lead time* em um processo de confecção de uniformes profissionais, por meio da identificação e

aplicação de ferramentas estatísticas apropriadas para a análise comparativa dos cenários pré e pós-implementação.

A avaliação do *lead time* antes e depois do *Kaizen* foi conduzida com base em um conjunto de testes e ferramentas estatísticas, incluindo Histogramas, Teste de Normalidade de Anderson-Darling, Séries Temporais, Análise de Tendência e o Teste de Mann-Whitney U. A escolha dessas ferramentas foi pautada na necessidade de analisar um conjunto de dados não paramétrico e de verificar a presença de diferenças estatisticamente significativas entre os períodos analisados.

Os resultados apontaram uma redução progressiva da média do *lead time* ao longo dos meses subsequentes à implementação do *Kaizen*. O estreitamento da dispersão dos dados e a tendência decrescente na variabilidade indicam melhorias no desempenho do processo produtivo. No entanto, a análise do teste de Mann-Whitney U revelou que, para o período de análise considerado, não houve diferença estatisticamente significativa entre os cenários pré e pós-*Kaizen*, sugerindo que a estabilidade do processo ainda estava em construção ou que o impacto do *Kaizen* pode demandar um período mais longo para se tornar estatisticamente detectável.

A presente pesquisa contribui para a literatura ao demonstrar uma abordagem estruturada para a mensuração dos impactos do *Kaizen* em processos produtivos. Ao selecionar e aplicar ferramentas estatísticas adequadas, este estudo fornece um referencial para futuras análises sobre a eficácia do *Kaizen* na redução do *lead time*, especialmente em ambientes produtivos com características semelhantes ao estudado.

CONCLUSÕES

O objetivo específico deste estudo foi avaliar a distribuição dos dados de *lead time* por peça, possibilitando a escolha de ferramentas adequadas, o que foi realizado por meio da análise de normalidade. Os resultados dessa análise levaram à escolha de abordagens estatísticas não paramétricas, garantindo que os métodos aplicados fossem adequados à natureza dos dados. Com isso, o objetivo geral de selecionar ferramentas que permitam inferir sobre o comportamento do *lead time* antes e depois do *Kaizen* foi atendido, julgando-se os recursos úteis e não úteis à análise por meio dos Quadros 2 e 3. Para responder à questão de pesquisa – "Quais ferramentas podem ser usadas para comparar dados de *lead time* antes e depois da implementação de um programa *Kaizen* voltado a sua redução?" – foram aplicados os recursos apresentados no Quadro 4, propondo-se uma padronização para a avaliação da eficácia de *Kaizens* na redução do *lead time*.

Os resultados apontam para uma redução progressiva do *lead time* e da dispersão dos dados, sugerindo melhorias no processo produtivo. No entanto, o teste de Mann-Whitney U não encontrou diferença estatisticamente significativa entre os cenários analisados, o que indica que um período mais longo de observação pode ser necessário para captar os efeitos do *Kaizen* de forma mais robusta.

Para estudos futuros, recomenda-se:

- Acompanhamento do *lead time* por um período mais longo para verificar a consolidação dos efeitos do *Kaizen*;
- Avaliação de outros indicadores, como taxa de retrabalho, produtividade e custos operacionais;
- Aplicação de testes estatísticos adicionais, como Wilcoxon ou modelos de regressão;

- Comparação entre diferentes unidades produtivas para avaliar a replicabilidade dos efeitos do *Kaizen*.

Este estudo contribui para a mensuração da eficácia do *Kaizen* na redução do *lead time* e fornece um referencial para futuras investigações sobre melhoria contínua em processos produtivos.

REFERÊNCIAS

- ABIT Associação Brasileira da Indústria Têxtil e de Confecção. (2023). Perfil do setor. Recuperado de: <https://www.abit.org.br/cont/perfil-do-setor>
- Antunes, J. L. F. & Cardoso, M. R. A. (2015). Aplicações da epidemiologia: Uso da análise de séries temporais em estudos epidemiológicos. *Epidemiologia e Serviços de Saúde*, 24(3), 565-576. <https://doi.org/10.5123/S1679-49742015000300024>
- Barros, E. A. C. & Mazucheli, J. (2005). Um estudo sobre o tamanho e poder dos testes t-Student e Wilcoxon. *Acta Scientiarum Technology*, 27(1), 23-32.
- Bezerra, M. J. S., et al. (2015). Melhoria de processos no setor de planejamento de uma equipe de construção e montagem industrial: Um estudo de caso. *Anais do XXXV Encontro Nacional de Engenharia de Produção*.
- Boisgontier, M. P. & Cheval, B. (2016). The ANOVA to mixed model transition. *Neuroscience & Biobehavioral Reviews*, 68, 1004-1005.
- Carvalho, W. J. S., et al. (2015). Análise e aplicabilidade de ferramentas básicas da qualidade como auxílio na melhoria do processo produtivo: Estudo de caso em uma indústria de confecção. *Anais do XXXV Encontro Nacional de Engenharia de Produção*.
- Coelho, F. P. S., Silva, A. M., & Maniçoba, R. F. (2016). Aplicação das ferramentas da qualidade: Estudo de caso em pequena empresa de pintura. *REFAS*, 3(1), 1-10.
- Cohen, H. W. (2011). P values: Use and misuse in medical literature. *American Journal of Hypertension*, 24(1), 18–23. <https://doi.org/10.1038/ajh.2010.205>
- Engmann, S. & Cousineau, D. (2011). Comparing distributions: The two-sample Anderson-Darling test as an alternative to the Kolmogorov-Smirnov test. *Journal of Applied Quantitative Methods*, 6(3), 1-10.
- Goodman, S. N. (1999). Toward evidence-based medical statistics. 1: The p-value fallacy. *Annals of Internal Medicine*, 130, 995-1004.
- Imai, M. (1988). *Kaizen: A estratégia para o sucesso competitivo* (C. Fagnani Lucca, Trad.). IMAM.
- Juran, J. M. & DeFeo, J. (2010). *Juran's Quality Handbook: The complete guide to performance Excellence* (6^a ed.). McGraw-Hill Education.
- Mann, H. B. & Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), 50-60.
- Martins, M. E. G. (2019). Regressão linear simples. *Revista de Ciência Elementar*, 7(3), 1-10.
- Martins, T. & Azevedo, M. (2023). Contribuições do *Kaizen* nas indústrias de confecção: Uma revisão sistemática da literatura. *Anais do XXX Simpósio de Engenharia de Produção*, Bauru, SP.
- Milenovic, Z. M. (2011). Application of Mann-Whitney U test in research of professional training of primary school teachers. *Metodicki obzori*, 11(6), 73.
- Oti, E. U., Olusola, M. O., & Esemokumo, P. A. (2021). Statistical analysis of the median test and the Mann-Whitney U test. *International Journal of Advanced Academic Research*, 7(9), 44.
- Ruxton, G. D. (2006). The unequal variance t-test is an underused alternative to Student's t-test and the Mann-Whitney U test. *Behavioral Ecology*, 17(3), 1-10. <https://doi.org/10.1093/beheco/ark016>
- Siegel, S. & Castellan, N. J. Jr. (1988). Nonparametric statistics for the behavioral sciences (2^a ed.). McGraw-Hill.
- Souza Neto, R. M., Galdino, D. D. E., Dantas, S. M., Santos, M. W. L., & Neto, J. M. S. (2017). Aplicação das sete ferramentas da qualidade em uma fábrica de blocos standard de gesso. *Anais do XXXV Encontro Nacional de Engenharia de Produção*, Joinville.
- Stake, R. E. (1995). The art of case study research. Sage Publications.
- Solanki, A. B., et al. (2021). Application of quality tools to reduce rejection of ice cream cones: A case in ice cream cone making company. *Technology Innovation in Mechanical Engineering-2021*, Elsevier Ltd.
- Suman, G. & Prajapati, D. R. (2018). Control chart applications in healthcare: A literature review. *Punjab Engineering College*, 12(1), 1-10.
- Tai, K. Y., Dhaliwal, J., & Balasubramaniam, V. (2022). Leveraging Mann-Whitney U test on large-scale genetic variation data for analysing malaria genetic markers. *Malaria Journal*, 21(79), 1-10. <https://doi.org/10.1186/s12936-022-04104-x>
- Womack, J. P., Jones, D. T., & Roos, D. (2004). A máquina que mudou o mundo (3^a ed.). Elsevier.
- Yazici, B. & Yolacan, S. (2007). A comparison of various tests of normality. *Journal of Statistical Computation and Simulation*, 77(2), 175-183. <https://doi.org/10.1080/10629360600678310>
- Yin, R. K. (2001). Case study research: Design and methods (3^a ed.). Sage Publications.
- Zimmerman, D. W. (1987). Comparative power of Student t test and Mann-Whitney U test for unequal sample sizes and variances. *The Journal of Experimental Education*, 55(3), 171-174.