



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO ACADÊMICO DO AGRESTE  
NÚCLEO DE TECNOLOGIA  
CURSO DE BACHARELADO EM ENGENHARIA DE PRODUÇÃO

BEATRIZ ROCHA FLORENCIO SILVA

**FALHAS NO SETOR DE MONTAGENS: UMA ANÁLISE ESTATÍSTICA PARA  
VERIFICAÇÃO DOS PRODUTOS, TURNOS E LINHAS MAIS IMPACTADOS**

Caruaru

2025

BEATRIZ ROCHA FLORENCIO SILVA

**FALHAS NO SETOR DE MONTAGENS: UMA ANÁLISE ESTATÍSTICA PARA  
VERIFICAÇÃO DOS PRODUTOS, TURNOS E LINHAS MAIS IMPACTADOS**

Trabalho de Conclusão de Curso apresentado ao Curso de Engenharia de Produção da Universidade Federal de Pernambuco - UFPE, na modalidade de monografia, como requisito parcial para a obtenção do título de Bacharel em Engenharia de Produção.

**Área de concentração:** Pesquisa Operacional.

Professor: Lucio Camara e Silva

Caruaru  
2025

Ficha de identificação da obra elaborada pelo autor,  
através do programa de geração automática do SIB/UFPE

Silva, Beatriz Rocha Florencio.

Falhas no setor de montagens: uma análise estatística para verificação dos produtos, turnos e linhas mais impactados / Beatriz Rocha Florencio Silva. - Caruaru, 2025.

64 : il., tab.

Orientador(a): Lúcio Camara e Silva

(Graduação) - Universidade Federal de Pernambuco, Centro Acadêmico do Agreste, Engenharia de Produção, 2025.

Inclui referências, apêndices, anexos.

1. Indústria de baterias. 2. Análise exploratória de dados. 3. Estatística descritiva. 4. Curva ABC. I. Silva, Lúcio Camara e. (Orientação). II. Título.

620 CDD (22.ed.)

BEATRIZ ROCHA FLORENCIO SILVA

**FALHAS NO SETOR DE MONTAGENS: UMA ANÁLISE ESTATÍSTICA  
PARA VERIFICAÇÃO DOS PRODUTOS, TURNOS E LINHAS MAIS  
IMPACTADOS.**

Trabalho de Conclusão de Curso  
apresentado à Coordenação do Curso de  
Engenharia de Produção do Campus  
Agreste da Universidade Federal de  
Pernambuco – UFPE, na modalidade de  
monografia, como requisito parcial para a  
obtenção do grau de bacharel em  
Engenharia de Produção.

Aprovada em: 14/08/2025

**BANCA EXAMINADORA**



Documento assinado digitalmente

**LUCIO CAMARA E SILVA**

Data: 19/08/2025 09:31:33-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Lúcio Camara e Silva (Orientador)  
Universidade Federal de Pernambuco



Documento assinado digitalmente

**LUCIMARIO GOIS DE OLIVEIRA SILVA**

Data: 19/08/2025 10:07:15-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Lucimário Gois de Oliveira Silva (Examinador Interno)  
Universidade Federal de Pernambuco



Documento assinado digitalmente

**THYAGO CELSO CAVALCANTE NEPOMUCENO**

Data: 20/08/2025 11:10:06-0300

Verifique em <https://validar.iti.gov.br>

---

Prof. Dr. Thyago Celso Cavalcante Nepomuceno (Examinador Externo)  
Universidade Federal de Pernambuco

## **AGRADECIMENTOS**

Primeiramente gostaria de agradecer aos meus Pais, Eronilson e Yana, meus maiores incentivadores da minha trajetória acadêmica, por todo esforço, suporte e investimento despendido propiciando conhecimento e persistência para finalizar esse trabalho.

Aos meus irmãos, Débora e Lucas, por sempre estarem presentes na minha vida e deixar essa trajetória mais leve.

Aos meus familiares, em nome de Mary e Patrícia, por sempre estarem me incentivando, apoiando e ajudando no meu desenvolvimento. A o meu orientador, Lúcio Camara, pela oportunidade de desenvolver este trabalho, por dar todo suporte, dedicar parte do seu tempo e conhecimento para que conseguisse concluir mais uma etapa.

A todo time da Acumuladores Moura do setor de produção da Montagem, controle de qualidade, gestão de suporte e demais áreas. Em especial, aos gestores Lorena Jordão, Eduardo Antonino, José Lins, Joseildo Cavalcante, Sandoval Santana e Alen Porto. Suas contribuições foram e serão um marco para o meu desenvolvimento profissional e pessoal.

Aos meus amigos que fiz desde o colégio e os da Universidade, que são muito importantes no meu desenvolvimento pessoal, profissional e deixaram a trajetória muito mais leve.

Aos meus mestres da Universidade, em nome de Maisa Mendonça, Marcele Elisa, Renata Maciel, Tharcylla Negreiros, Gilson Lima, Thalles Garcez, José Leão e Lucimário por todo conhecimento e ensinamentos compartilhados. Agradeço, também, a Jhonatha Joao Cavalcanti dos Santos pelo suporte nas atividades.

## RESUMO

Esta pesquisa tem como objetivo aplicar a Análise Exploratória de Dados (EDA - *Exploratory Data Analysis*), para identificação de produtos, turnos e linhas de produção mais críticos em uma unidade de montagem de baterias, a fim de subsidiar ações de melhoria da eficiência industrial. Metodologicamente, caracteriza-se como estudo de caso único, quantitativo e descritivo, com coleta de dados extraídos do sistema ERP (*Enterprise Resource Planning*) da empresa, referentes ao período de 2020 a 2024. Após a coleta, os dados foram tratados e visualizados em Python, possibilitando análises por Curva ABC, gráficos de Pareto, Heatmaps, Boxplots, séries temporais e clusterização. Os resultados revelaram padrões consistentes de concentração de falhas por tipo e turno, no qual um número reduzido de modelos de baterias concentra a maior parte das falhas. Verificou-se que os turnos A e B apresentaram os maiores índices de falhas, enquanto o turno D registrou volumes praticamente nulos, e que as linhas 2, 14, 3 e 12 concentraram mais de 80% das falhas acumuladas. Portanto, o trabalho contribui tanto de forma prática, oferecendo subsídios gerenciais para a empresa estudada, como na identificação dos itens considerados como críticos, quanto teórica, demonstrando o potencial da EDA como ferramenta estratégica para a excelência operacional na empresa.

**Palavras-chave:** Análise exploratória de dados; Eficiência operacional; Falhas; Indústria de baterias

## ABSTRACT

This research aims to apply Exploratory Data Analysis (EDA) to identify the most critical products, shifts, and production lines in a battery assembly unit, in order to support actions to improve industrial efficiency. Methodologically, it is characterized as a single-case, quantitative, and descriptive study, with data collected from the company's ERP (Enterprise Resource Planning) system for the period from 2020 to 2024. After collection, the data were processed and visualized in Python, enabling analyses through ABC Curve, Pareto charts, heatmaps, boxplots, time series, and clustering. The results revealed consistent patterns of failure concentration by type and shift, in which a small number of battery models accounted for the majority of failures. It was found that shifts A and B had the highest failure rates, while shift D recorded virtually zero volumes, and that lines 2, 14, 3, and 12 accounted for more than 80% of accumulated failures. Therefore, this work contributes both practically, by providing managerial insights for the studied company and identifying items considered critical and theoretically, by demonstrating the potential of EDA as a strategic tool for operational excellence in the company.

**Keywords:** Exploratory data analysis; Operational efficiency; Failure; Battery industry

## LISTA DE FIGURAS

|                                                                                        |    |
|----------------------------------------------------------------------------------------|----|
| Figura 1 – Série temporal da evolução das falhas                                       | 29 |
| Figura 2 – Top 10 modelos de bateria com mais falhas                                   | 31 |
| Figura 3 – Falhas por turno                                                            | 32 |
| Figura 4 – Falhas por linha de produção                                                | 33 |
| Figura 5 – Heatmap de falhas por linha e turno de produção                             | 34 |
| Figura 6 – Boxplot da distribuição das quantidades de falhas por modelo                | 35 |
| Figura 7 – Análise da quantidade de falhas por turno nos principais modelos de bateria | 37 |
| Figura 8 – Análise da quantidade de falhas por turno nas principais linhas de produção | 38 |
| Figura 9 – Gráfico de Pareto (Falhas por linha de produção)                            | 42 |
| Figura 10 – Heatmap de quantidade de falhas por tipo de defeito e turno                | 43 |
| Figura 11 – Clusterização dos modelos de bateria (turno + linha de produção)           | 44 |

## LISTA DE TABELAS

|                                                                    |    |
|--------------------------------------------------------------------|----|
| Tabela 1 – Falhas por modelo de bateria . . . . .                  | 24 |
| Tabela 2 – Distribuição das falhas por modelo de bateria . . . . . | 25 |
| Tabela 3 – Quantitativo das falhas . . . . .                       | 27 |
| Tabela 4 – Estatísticas descritivas das falhas . . . . .           | 28 |
| Tabela 5 – Curva ABC . . . . .                                     | 39 |

## **LISTA DE ABREVIATURAS E SIGLAS**

|     |                               |
|-----|-------------------------------|
| EDA | Análise Exploratória de Dados |
| ERP | Enterprise Resource Planning  |
| IA  | Inteligência Artificial       |

## SUMÁRIO

|          |                                                                                                    |           |
|----------|----------------------------------------------------------------------------------------------------|-----------|
| <b>1</b> | <b>INTRODUÇÃO.....</b>                                                                             | <b>12</b> |
| 1.1      | OBJETIVOS .....                                                                                    | 14        |
| 1.1.1    | Objetivo Geral .....                                                                               | 14        |
| 1.1.2    | Objetivos Específicos.....                                                                         | 14        |
| 1.2      | JUSTIFICATIVAS.....                                                                                | 15        |
| <b>2</b> | <b>REFERENCIAL TEÓRICO.....</b>                                                                    | <b>16</b> |
| 2.1      | EFICIÊNCIA OPERACIONAL E CONTROLE DE FALHAS NA INDÚS-<br>TRIA .....                                | 16        |
| 2.2      | CIÊNCIA DE DADOS E ANÁLISE DE FALHAS .....                                                         | 18        |
| 2.3      | A GESTÃO DA PRODUÇÃO ORIENTADA POR DADOS.....                                                      | 19        |
| <b>3</b> | <b>MÉTODO .....</b>                                                                                | <b>21</b> |
| 3.1      | CARACTERIZAÇÃO DA ORGANIZAÇÃO ESTUDADA .....                                                       | 21        |
| 3.2      | CLASSIFICAÇÃO METODOLÓGICA .....                                                                   | 21        |
| 3.3      | COLETA DE DADOS.....                                                                               | 22        |
| 3.4      | ANÁLISE DOS DADOS.....                                                                             | 23        |
| <b>4</b> | <b>RESULTADOS .....</b>                                                                            | <b>24</b> |
| 4.1      | ANÁLISE DAS ESTATÍSTICAS DESCRITIVAS DAS FALHAS (POR MO-<br>DELO DE BATERIA E TIPO DE FALHA) ..... | 24        |
| 4.2      | ANÁLISE DE SÉRIE TEMPORAL DE FALHAS (EVOLUÇÃO DAS FA-<br>LHAS AO LONGO DO TEMPO).....              | 29        |
| 4.3      | FALHAS POR PRODUTOS .....                                                                          | 30        |
| 4.4      | ANÁLISE DAS FALHAS POR TURNO .....                                                                 | 31        |

|      |                                                                                   |    |
|------|-----------------------------------------------------------------------------------|----|
| 4.5  | ANÁLISE DAS FALHAS POR LINHA DE PRODUÇÃO .....                                    | 32 |
| 4.6  | HEATMAP DE FALHAS POR LINHA E TURNO DE PRODUÇÃO .....                             | 33 |
| 4.7  | BOXPLOT DA DISTRIBUIÇÃO DAS FALHAS POR MODELO DE BATERIA.....                     | 35 |
| 4.8  | ANÁLISE DA QUANTIDADE DE FALHAS POR TURNO NOS PRINCIPAIS MODELOS DE BATERIA ..... | 36 |
| 4.9  | ANÁLISE DA QUANTIDADE DE FALHAS POR TURNO NAS PRINCIPAIS LINHAS DE PRODUÇÃO ..... | 38 |
| 4.10 | ANÁLISE DA CURVA ABC .....                                                        | 39 |
| 4.11 | GRÁFICO DE PARETO (FALHAS POR LINHA DE PRODUÇÃO) .....                            | 41 |
| 4.12 | HEATMAP – QUANTIDADE DE FALHAS POR TIPO DE DEFEITO E TURNO .....                  | 42 |
| 4.13 | CLUSTERIZAÇÃO DE MODELOS DE BATERIA (TURNO E LINHA DE PRODUÇÃO).....              | 44 |
| 4.14 | SUBSÍDIOS ANALÍTICOS PARA A DEFINIÇÃO DE ESTRATÉGIAS DE MELHORIA .....            | 45 |
| 5    | CONSIDERAÇÕES FINAIS.....                                                         | 47 |
|      | REFERÊNCIAS.....                                                                  | 49 |
|      | APÊNDICE A – ALGORITMO .....                                                      | 51 |

## 1 INTRODUÇÃO

A busca constante pela eficiência operacional constitui um dos principais desafios enfrentados pelas organizações industriais contemporâneas. Em mercados cada vez mais competitivos e dinâmicos, caracterizados pela elevada exigência dos consumidores, pelo acirramento da concorrência global e pela necessidade de inovação contínua, as empresas precisam desenvolver métodos eficazes para controlar a qualidade de seus processos e produtos (SLACK; CHAMBERS; JOHNSTON, 2009). No contexto industrial, falhas operacionais em linhas de produção representam um fator crítico que impacta diretamente a produtividade, os custos, a satisfação do cliente e, conseqüentemente, a posição competitiva da organização.

A identificação, análise e priorização de falhas operacionais emergem, portanto, como instrumentos estratégicos indispensáveis para a melhoria contínua dos processos produtivos. Conforme apontam Juran e Godfrey (1999), a gestão da qualidade requer uma abordagem sistemática para diagnosticar desvios e implementar ações corretivas fundamentadas em dados. No setor de fabricação de acumuladores de energia, a atenção a esses fatores se torna ainda mais vibrante, dado o rigor técnico exigido no desempenho e na segurança dos produtos. Qualquer inconsistência no processo de montagem pode comprometer a confiabilidade do produto final, acarretando conseqüências que extrapolam o âmbito interno da empresa, alcançando a reputação de marca e a fidelidade do cliente.

Nesse contexto, a transformação digital e a disponibilidade crescente de dados operacionais abrem novas possibilidades para uma gestão da qualidade baseada em evidências. A Análise Exploratória de Dados (EDA), como enfatiza Tukey (1977), surge como uma ferramenta poderosa para descobrir padrões, tendências e anomalias em grandes volumes de dados de forma intuitiva e visual. Ao aplicar técnicas de EDA no ambiente industrial, é possível obter insights relevantes sobre o comportamento das falhas operacionais, favorecendo uma tomada de decisão mais informada e assertiva.

Neste contexto, a utilização de linguagens de programação como o Python, amplamente difundido no meio acadêmico e empresarial, potencializa ainda mais essas análises ao proporcionar recursos avançados de tratamento, visualização e interpretação de dados (VANDERPLAS, 2016). Bibliotecas como *pandas*, *matplotlib* e *seaborn*, entre outras, permitem que grandes volumes de registros históricos sejam processados de maneira eficiente, revelando informações críticas que poderiam passar despercebidas em análises tradicionais.

Assim, a análise sistemática das falhas operacionais não apenas contribui para a melhoria da qualidade dos produtos, mas também fortalece a resiliência dos processos produtivos frente às variabilidades inevitáveis do ambiente industrial. Segundo Deming (1997), compreender as variações nos processos é o primeiro passo para controlá-los e melhorá-los continuamente. Aplicar a metodologia de EDA em bases de dados operacionais atende, portanto, à necessidade de mapear as principais fontes de variabilidade, identificar pontos críticos e orientar interven-

ções que promovam o aperfeiçoamento da eficiência produtiva.

A Indústria 4.0 caracteriza-se pela integração de tecnologias digitais como Internet das Coisas (IoT), Big Data e sistemas ciberfísicos, visando processos produtivos inteligentes e conectados. Nesse contexto, a Inteligência Artificial (IA) desempenha papel central ao possibilitar a análise de grandes volumes de dados, a automação inteligente, a manutenção preditiva e o controle de qualidade por meio de algoritmos de aprendizado e visão computacional. Dessa forma, a relação entre Indústria 4.0 e IA é complementar: enquanto a primeira fornece infraestrutura digital e conectividade, a segunda transforma dados em conhecimento aplicado, tornando a produção mais eficiente, flexível e sustentável (LEE; DAVARI; KAO, 2018).

Outro aspecto fundamental é a otimização das cadeias de suprimento, onde técnicas de IA contribuem para prever a demanda, planejar rotas logísticas e reduzir desperdícios (IVANOV et al., 2019). Do ponto de vista estratégico, a combinação entre Indústria 4.0 e IA permite não apenas maior eficiência e produtividade, mas também promove modelos de negócios baseados em personalização em massa, em que os produtos são adaptados às necessidades dos consumidores em larga escala.

Portanto, a relação entre Indústria 4.0 e IA pode ser compreendida como simbiótica: a infraestrutura digital da Indústria 4.0 fornece os dados e a conectividade necessários, enquanto a IA transforma esses dados em inteligência aplicada, resultando em processos produtivos mais ágeis, precisos e sustentáveis.

A indústria de baterias, objeto deste estudo, caracteriza-se por um elevado volume de produção seriada e por um alto grau de padronização dos processos. No entanto, mesmo em ambientes fortemente controlados, a ocorrência de falhas é inevitável. Em processos repetitivos e de larga escala, a análise minuciosa de falhas operacionais é essencial para reduzir perdas, otimizar recursos e elevar o nível de qualidade percebida pelo mercado. Em consonância com essa perspectiva, este trabalho propõe a realização de uma análise exploratória dos dados de falhas registrados no sistema ERP (Enterprise Resource Planning) de uma unidade de montagem de baterias.

Adotar uma abordagem orientada por dados é vital para superar a subjetividade nas decisões de melhoria operacional. Conforme defendem Provost e Fawcett (2013b), a ciência de dados oferece métodos rigorosos para transformar dados brutos em conhecimento acionável, capaz de fundamentar estratégias corporativas. No caso específico deste estudo, o tratamento e a visualização dos dados de falhas permitirão à organização compreender melhor a distribuição dos problemas, estabelecer prioridades de intervenção e, assim, otimizar seus esforços de gestão da produção.

Cabe ressaltar que a metodologia de análise adotada neste trabalho também se ancora no conceito de melhoria contínua proposto por Ishikawa (1985), para quem a qualidade não é um estado a ser atingido, mas sim um processo evolutivo baseado no monitoramento permanente dos processos e na correção sistemática de desvios. A partir da identificação dos modelos de baterias, turnos e linhas de produção mais críticos, pretende-se subsidiar a gestão da unidade

de montagem com informações que orientem o desenvolvimento de planos de ação focados, capazes de impactar positivamente os principais indicadores de desempenho.

Por fim, cabe destacar que embora a aplicação da EDA no contexto industrial esteja em crescimento, ainda são relativamente poucos os estudos que exploram o potencial dessa metodologia de forma prática e aplicada à gestão da produção de forma tão específica como a indústria de baterias. Assim, o trabalho pretende contribuir também para o fortalecimento do campo de estudos que articula, engenharia de produção e eficiência operacional.

Diante desse cenário, a condução de uma EDA sobre os históricos de falhas operacionais se mostra fundamental para a geração de valor tanto para a empresa estudada quanto para a comunidade acadêmica. O mapeamento dos produtos, turnos e linhas de produção mais impactados, bem como a identificação de padrões de falhas, será capaz de orientar a priorização de ações de melhoria, promover o uso inteligente de recursos organizacionais e impulsionar a trajetória da organização rumo à excelência operacional.

Neste caso, o presente trabalho foi dividido em cinco capítulos, sendo este o introdutório. Na sequência, é trago a fundamentação teórica que expõe e sustenta os conceitos basilares desta pesquisa. Adiante, é apresentado o método perseguido para realizar a presente pesquisa e que serve de disposição para futuras investigações e constatações. O quarto capítulo dialoga com os resultados encontrados. Por último, é dito as considerações finais do trabalho, delineando as conclusões, limitações e possíveis estudos futuros.

## 1.1 OBJETIVOS

Este tópico apresenta o objetivo geral, bem como os objetivos específicos deste estudo.

### 1.1.1 Objetivo Geral

Utilizar técnicas de análise de dados de falhas operacionais de uma unidade de montagem de baterias para identificar produtos, turnos e linhas de produção mais impactados, com o intuito de subsidiar ações de melhoria da eficiência industrial.

### 1.1.2 Objetivos Específicos

Para alcançar esse objetivo, delineou-se cinco objetivos específicos para atingir o objetivo maior que guia este estudo, sendo eles:

- Coletar e organizar os registros de falhas operacionais da unidade de montagem de baterias no período de 2020 a 2024, extraídos do sistema ERP da empresa.
- Tratar e preparar os dados para análise, incluindo limpeza, padronização e categorização das variáveis, com uso do Python.

- Aplicar técnicas de Análise Exploratória de Dados (EDA) para identificar padrões relacionados a produtos, turnos e linhas de produção.
- Identificar e priorizar os produtos, turnos e linhas de produção mais impactados pelas falhas operacionais.
- Propor recomendações gerenciais para ações corretivas e preventivas que visem à melhoria da eficiência operacional

## 1.2 JUSTIFICATIVAS

A relevância deste estudo está ancorada na necessidade crescente das indústrias em aprimorar seus processos produtivos por meio da utilização inteligente de dados. Em um ambiente industrial cada vez mais dinâmico e competitivo, caracterizado pela demanda por alta qualidade, baixos custos e agilidade operacional, a análise sistemática de falhas operacionais surge como uma estratégia essencial para o fortalecimento da eficiência e da vantagem competitiva.

No setor de produção de baterias, o rigor técnico exigido em relação à segurança, durabilidade e desempenho dos produtos torna a gestão das falhas ainda mais crítica. Pequenas inconsistências nos processos podem resultar em impactos significativos, não apenas nos custos internos, mas também na percepção de qualidade pelo mercado e na fidelização dos clientes. Assim, compreender a distribuição e a frequência das falhas operacionais permite que a organização direcione seus esforços de melhoria de forma mais estratégica e efetiva, otimizando recursos e minimizando desperdícios.

Dessa forma, a utilização de ferramentas de Análise Exploratória de Dados representa uma abordagem moderna, alinhada aos princípios da Indústria 4.0, promovendo a transformação de grandes volumes de dados brutos em conhecimento acionável. Além disso, fornece subsídios analíticos para a definição de estratégias de melhoria e tomada de decisão voltadas para identificação e resolução de falhas que geram maior impacto. A utilização de linguagens de programação Python e suas bibliotecas, como *pandas*, *matplotlib* e *seaborn*, proporcionou uma análise mais organizada e visual das informações, favorecendo a detecção de padrões e tendências que seriam difíceis de identificar por métodos tradicionais.

Do ponto de vista acadêmico, este trabalho também contribui para o fortalecimento da interface entre a Engenharia de Produção e pesquisas aplicadas que explorem a aplicação da EDA em problemas industriais, como no contexto específico da indústria de baterias.

## 2 REFERENCIAL TEÓRICO

### 2.1 EFICIÊNCIA OPERACIONAL E CONTROLE DE FALHAS NA INDÚSTRIA

A eficiência operacional tem sido, ao longo das últimas décadas, um dos principais pilares para a sobrevivência e o crescimento sustentado das organizações industriais. Em um cenário caracterizado por mercados altamente competitivos, margens de lucro cada vez mais estreitas e clientes cada vez mais exigentes, as empresas que não buscam incessantemente a excelência em seus processos correm o risco de perder competitividade e, em última instância, sua posição de mercado (SLACK; CHAMBERS; JOHNSTON, 2009).

De acordo com Chase, Jacobs e Aquilano (2006), a eficiência operacional está relacionada à capacidade de uma organização em maximizar seus outputs (produtos ou serviços) com a utilização racional de seus inputs (recursos humanos, materiais, financeiros e tecnológicos), minimizando desperdícios e perdas ao longo dos processos produtivos. Essa perspectiva vai ao encontro da filosofia da produção enxuta (lean manufacturing), que enfatiza a eliminação sistemática de desperdícios como forma de agregar valor ao cliente e aumentar a eficiência (WOMACK; JONES, 1996).

Segundo Campos (2004), a melhoria contínua só é possível quando há um monitoramento sistemático dos processos e das falhas, com uma atuação proativa da gestão. Organizações que adotam essa abordagem desenvolvem uma cultura de excelência, na qual a eficiência e a confiabilidade caminham juntas. Técnicas como FMEA (Failure Mode and Effects Analysis), RCA (Root Cause Analysis) e Análise de Árvore de Falhas (FTA) são amplamente empregadas para identificar causas raiz e evitar recorrências.

No contexto industrial, falhas operacionais constituem uma das principais fontes de desperdício e ineficiência. Cada falha representa não apenas uma interrupção do fluxo produtivo, mas também custos adicionais de retrabalho, refugo, manutenção corretiva, além de potenciais atrasos nas entregas e impactos negativos sobre a satisfação do cliente. Segundo Juran e Godfrey (1999), a não conformidade com padrões de qualidade planeja aquilo que denominam "custos da má qualidade", que englobam custos internos (detecção e correção de defeitos) e externos (reclamações de clientes, devoluções e garantias).

A prevenção de falhas e o controle rigoroso dos processos são, portanto, elementos fundamentais para a promoção da eficiência operacional. Nesse sentido, a gestão da qualidade, conforme proposto por Feigenbaum (1991), deve ser incorporada como uma responsabilidade de todos os níveis organizacionais e não apenas como uma função isolada em departamentos específicos. A qualidade deve ser construída ao longo de todo o processo produtivo e não apenas verificada ao final, o que demanda uma cultura organizacional orientada para a melhoria contínua.

As falhas operacionais podem ter origens variadas, como erros humanos, falhas de equi-

pamento, problemas de matéria-prima, inadequações no projeto de produto, deficiência em processos de manutenção, entre outros fatores. Cada tipo de falha traz consigo um impacto específico, variando desde pequenas interrupções na linha de produção até grandes paralisações que comprometem significativamente a capacidade de atendimento da empresa. Além dos impactos tangíveis, existem consequências intangíveis associadas às falhas, como a perda de confiança dos clientes, a desmotivação dos colaboradores e o desgaste da imagem corporativa no mercado. Esses efeitos, embora mais difíceis de quantificar, podem ser igualmente danosos ao longo do tempo.

Tradicionalmente, muitas organizações abordaram a gestão das falhas de maneira reativa, ou seja, adotando medidas corretivas somente após a ocorrência de problemas. Contudo, autores como Crosby (1979) e Garvin (1987) defendem que o verdadeiro caminho para a excelência operacional está na prevenção de falhas, e não apenas na sua correção. A abordagem preventiva implica a identificação precoce de desvios e a implementação de medidas sistemáticas para eliminar suas causas raízes. Isso requer o monitoramento contínuo dos processos produtivos, a capacitação constante das equipes, o investimento em manutenção preventiva e preditiva de equipamentos, e a construção de uma cultura organizacional orientada para a qualidade total.

O controle de falhas é um elemento indissociável da eficiência operacional, uma vez que as falhas não observadas ou não prevenidas podem comprometer a eficiência produtiva e elevar os custos (ANTUNES et al., 2008).

Segundo Ishikawa (1985), a qualidade é construída por todas as áreas da organização e deve ser considerada em todas as etapas, desde o planejamento do produto até o atendimento ao cliente. Assim, o controle de falhas não é um objetivo isolado, mas um componente integrante da estratégia empresarial.

Para que a gestão das falhas seja eficaz, é fundamental dispor de dados confiáveis e atualizados sobre o desempenho dos processos. Conforme destacado por Davenport e Harris (2007), organizações orientadas por dados tendem a tomar decisões mais rápidas, precisas e baseadas em evidências, o que lhes confere vantagem competitiva. A coleta e a análise sistemática dos dados de falhas possibilitam a construção de indicadores de desempenho, a identificação de padrões de ocorrência, a avaliação de tendências ao longo do tempo e o reconhecimento de áreas críticas que exigem atenção prioritária. Sem esse embasamento analítico, a tomada de decisão tende a ser baseada em percepções subjetivas e não em fatos concretos.

A digitalização dos processos e a análise avançada de dados, elementos centrais da Indústria 4.0, ampliam ainda mais essa integração, promovendo fábricas inteligentes, capazes de tomar decisões automatizadas com base no histórico e no comportamento dos ativos (WANG et al., 2018).

A evolução das tecnologias de informação e o advento dos ERP permitiram às organizações armazenar grandes volumes de dados operacionais. No entanto, como observam Provost e Fawcett (2013a), a simples posse dos dados não é suficiente: é necessário desenvolver a capacidade analítica para transformar dados em conhecimento e conhecimento em ação. A EDA,

nesse contexto, desponta como uma abordagem fundamental para extrair informações relevantes dos dados disponíveis, auxiliando no diagnóstico de falhas, na identificação de correlações relevantes e na construção de insights para a melhoria contínua dos processos.

No âmbito específico da produção de baterias, a análise cuidadosa dos dados de falhas operacionais assume um papel ainda mais relevante, dada a complexidade dos processos e a criticidade da qualidade do produto final. O presente estudo, ao propor a aplicação de técnicas de Análise Exploratória de Dados para a identificação e priorização de falhas, insere-se nesse esforço de promover uma gestão da produção mais eficiente, baseada em evidências e orientada para a excelência operacional.

## 2.2 CIÊNCIA DE DADOS E ANÁLISE DE FALHAS

A Análise Exploratória de Dados constitui uma etapa fundamental no processo de compreensão e interpretação de grandes volumes de dados. Introduzido por Tukey (1977), o conceito de EDA preconiza a necessidade de examinar os dados de maneira aberta e flexível, buscando identificar padrões, anomalias, tendências e estruturas ocultas que possam orientar hipóteses, tomadas de decisão e ações corretivas.

No ambiente industrial, a utilização da EDA tem ganhado destaque como uma poderosa ferramenta de apoio à gestão da produção e à melhoria contínua. Em setores como o de manufatura de baterias, onde a quantidade de informações geradas pelos processos produtivos é elevada, a capacidade de realizar uma análise exploratória eficiente torna-se um diferencial competitivo importante.

Conforme ressaltam Provost e Fawcett (2013b), o valor dos dados só é realizado plenamente quando eles são transformados em insights acionáveis. A EDA permite que gestores e engenheiros de produção entendam não apenas o "o quê" dos fenômenos observados, como a quantidade de falhas registradas em determinado período, mas também o "porquê" e o "como" esses eventos se distribuem, fornecendo uma base sólida para a formulação de estratégias de intervenção.

De acordo com Tukey (1977), a principal função da EDA é permitir que o analista se familiarize com os dados, identifique possíveis erros, detecte padrões significativos e formule hipóteses que possam ser posteriormente testadas. Diferentemente da análise confirmatória, que parte de hipóteses predefinidas, a análise exploratória assume uma postura aberta e investigativa, permitindo que os dados falem por si mesmos. Neste caso, a aplicação da EDA em ambientes industriais se justifica pela complexidade dos processos de produção e pela multiplicidade de variáveis envolvidas. Em fábricas de alta produção, como é o caso da indústria de acumuladores elétricos, dados são gerados continuamente por equipamentos, sensores, sistemas de controle de qualidade e registros operacionais.

Conforme destaca Montgomery (2020), a análise estatística dos dados de produção permite que as organizações monitorem seus processos em tempo real, detectem desvios rapidamente e implementem ações corretivas antes que problemas se tornem críticos. Assim sendo, a EDA

constitui um componente indispensável para a gestão moderna da produção industrial. Em um ambiente de elevada complexidade e intensa geração de dados, a capacidade de explorar e interpretar informações de maneira eficiente representa uma vantagem estratégica para as organizações.

Portanto, entende-se que a utilização de técnicas de EDA no controle de falhas operacionais permite uma visão mais clara e detalhada dos processos produtivos, possibilitando a priorização de ações corretivas, a identificação de oportunidades de melhoria e a construção de uma cultura organizacional orientada por dados. Ao adotar a EDA como parte integrante da gestão da organização, as empresas não apenas aumentam sua eficiência operacional, mas também fortalecem sua capacidade de inovação e adaptação em mercados cada vez mais exigentes.

### 2.3 A GESTÃO DA PRODUÇÃO ORIENTADA POR DADOS

A gestão da produção tem passado por transformações significativas nas últimas décadas, impulsionadas pela evolução tecnológica e pelo advento da Indústria 4.0. Nesse novo cenário, a capacidade de coletar, processar e analisar grandes volumes de dados tornou-se um diferencial competitivo essencial para as organizações industriais. A gestão da produção orientada por dados representa, assim, uma mudança de paradigma: de práticas baseadas em intuição e experiência acumulada para práticas fundamentadas em evidências quantitativas e análises sistemáticas (DAVENPORT; HARRIS, 2007).

De acordo com Provost e Fawcett (2013b), a ciência de dados aplicada à gestão empresarial possibilita a extração de informações relevantes a partir de grandes conjuntos de dados, permitindo que as decisões sejam tomadas com maior precisão, agilidade e fundamentação lógica. No contexto da produção, essa abordagem favorece a identificação de gargalos, a previsão de falhas, a otimização de recursos e o aprimoramento contínuo dos processos operacionais.

A produção orientada por dados busca transformar os registros operacionais, muitas vezes dispersos e subutilizados, em conhecimento estruturado e acionável. Essa transformação implica o desenvolvimento de uma cultura organizacional que valorize a coleta sistemática de dados, a integridade das informações, a análise crítica dos resultados e a disseminação do conhecimento gerado por essas análises. Conforme Slack, Chambers e Johnston (2009), o uso estratégico da informação é um dos fatores-chave para a construção de operações de excelência.

Entre as principais vantagens da gestão da produção orientada por dados, destaca-se a capacidade de detectar padrões de comportamento que não seriam perceptíveis por meio de análises tradicionais. A aplicação de técnicas de análise exploratória de dados, mineração de dados e aprendizado de máquina permite, por exemplo, antecipar tendências de falhas, prever a demanda por manutenção preventiva, estimar tempos de ciclo e ajustar parâmetros de processos de forma dinâmica. Essas práticas elevam a maturidade da gestão da produção, permitindo que a organização atue de maneira mais proativa e menos reativa frente aos desafios operacionais.

A adoção de gestão da produção orientada por dados traz diversos benefícios estratégicos. Segundo Marr (2016), as empresas que utilizam dados de forma eficaz obtêm melhorias mensuráveis em desempenho, como: redução de desperdícios e retrabalhos, aumento de produtividade por meio de automação inteligente, melhoria na tomada de decisões estratégicas, menor custo para a organização.

A análise de dados históricos e em tempo real permite ações preditivas em vez de reativas, promovendo uma gestão mais ágil e responsiva (WANG; WANG, 2020).

Além disso, o uso sistemático de dados contribui para a transparência organizacional e o alinhamento entre diferentes áreas da empresa. Ao trabalhar com indicadores objetivos e relatórios analíticos, as decisões se tornam menos suscetíveis a vieses pessoais e mais fundamentadas em fatos. Essa mudança é fundamental para consolidar práticas de melhoria contínua, que, segundo Deming (1997), dependem da capacidade da organização de medir seu desempenho de forma consistente e agir sobre as causas dos desvios detectados.

No ambiente industrial atual, marcado pela complexidade dos processos, pela variabilidade da demanda e pela necessidade de customização em massa, a gestão orientada por dados possibilita que as organizações se adaptem rapidamente às mudanças. Davenport e Harris (2007) apontam que empresas que investem em capacidades analíticas avançadas obtêm ganhos não apenas de eficiência, mas também de inovação, já que a análise de dados permite identificar oportunidades de melhoria e de desenvolvimento de novos produtos e processos.

Dessa forma, este trabalho se insere no esforço contemporâneo de reforçar a importância dos dados como base para a gestão industrial, ao demonstrar, através da análise exploratória de dados de falhas, como a informação estruturada e interpretada de maneira sistemática pode orientar ações corretivas e preventivas mais eficazes, promovendo a excelência dos processos produtivos.

,

### 3 MÉTODO

Este capítulo tem por finalidade apresentar detalhadamente os procedimentos metodológicos adotados para o desenvolvimento da presente pesquisa, incluindo a caracterização da organização estudada, a classificação metodológica, os instrumentos de coleta de dados utilizados, bem como os procedimentos empregados na análise dos dados. Conforme Lakatos e Marconi (2003), a metodologia é o conjunto de procedimentos intelectuais e técnicos adotados para atingir determinado fim, sendo essencial para a construção do conhecimento científico.

#### 3.1 CARACTERIZAÇÃO DA ORGANIZAÇÃO ESTUDADA

A pesquisa foi realizada em uma empresa de acumuladores de energia, localizada no município de Belo Jardim, estado de Pernambuco. Especificamente, a investigação foi conduzida na Unidade de Montagem de Baterias (UN01), setor onde são realizadas as etapas finais de montagem dos produtos fabricados pela organização. Esta unidade é caracterizada por um alto volume de operações repetitivas e controle rigoroso de qualidade, visto que a performance dos acumuladores de energia depende diretamente da precisão e confiabilidade de seu processo produtivo.

A discente autora desta pesquisa atua como supervisora no referido setor, o que lhe conferiu acesso facilitado às informações e à rotina operacional da unidade. Essa posição favoreceu a observação participante como estratégia complementar de coleta de dados, conforme destacado por Cervo, Bervian e Silva (2007), que apontam que o envolvimento direto do pesquisador com o ambiente investigado pode enriquecer a interpretação dos fenômenos observados, desde que mantida a objetividade científica.

A empresa de acumuladores de energia se destaca no cenário nacional pela sua capacidade produtiva, inovação tecnológica e compromisso com a sustentabilidade. Sua atuação abrange todo o território nacional e também mercados internacionais, o que a coloca entre as principais fabricantes de baterias da América Latina. A unidade investigada desempenha papel estratégico na cadeia de valor da organização, uma vez que é responsável por assegurar a conformidade dos produtos antes do envio ao mercado consumidor.

#### 3.2 CLASSIFICAÇÃO METODOLÓGICA

A presente pesquisa caracteriza-se como um estudo de caso único, conforme a conceituação de Yin (2015), uma vez que analisa em profundidade um único caso representativo, com o objetivo de compreender a ocorrência de falhas no processo de montagem de baterias. Estudos de caso são particularmente apropriados quando se busca examinar um fenômeno dentro de

seu contexto real, especialmente quando os limites entre o fenômeno e o contexto não são claramente definidos.

Quanto à abordagem, trata-se de uma pesquisa de cunho quantitativo, dado que se fundamenta na coleta, tratamento e análise de dados numéricos, com o intuito de identificar padrões e relações objetivas entre variáveis. Segundo Miguel et al. (2010), a abordagem quantitativa permite a aplicação de técnicas estatísticas e modelos matemáticos na exploração dos dados coletados, promovendo maior precisão e reprodutibilidade dos resultados.

A classificação da pesquisa é descritiva, pois seu objetivo central é descrever a localização das falhas no processo produtivo, identificando os pontos críticos e os tipos de problemas mais recorrentes. De acordo com Vergara (2007), pesquisas descritivas visam descrever as características de determinado fenômeno ou relações entre variáveis, sem necessariamente estabelecer relações de causa e efeito.

### 3.3 COLETA DE DADOS

A coleta de dados foi realizada a partir de fontes secundárias, extraídas diretamente do ERP da organização, complementadas por registros de observação participante da pesquisadora. Segundo Kauark Fabiane S.; Manhães (2010), a utilização de dados secundários permite o aproveitamento de informações já existentes, otimizando o tempo e os recursos da pesquisa, ao passo que a observação participante contribui com insights práticos sobre o cotidiano da organização.

A base de dados utilizada foi extraída do sistema ERP da empresa e abrangeu os registros de falhas ocorridas ao longo dos anos de 2020 a 2024, totalizando milhares de entradas associadas a diferentes modelos de baterias, turnos de produção e linhas produtivas. Após a limpeza e padronização dos dados, procedeu-se à construção de indicadores e gráficos capazes de revelar padrões e apoiar a identificação dos pontos mais críticos do processo produtivo.

Após a extração da base, realizou-se o procedimento de limpeza e organização dos dados com o apoio das bibliotecas do Python. Esse processo incluiu a exclusão de registros duplicados, padronização de nomenclaturas e validação da consistência dos dados. Miguel et al. (2010) ressaltam que a preparação adequada da base é uma etapa essencial para assegurar a qualidade e fidedignidade das análises posteriores.

Além disso, a pesquisadora utilizou-se da observação participante durante o período de coleta para complementar a compreensão dos dados. Por estar inserida no contexto organizacional e atuar diretamente na supervisão das atividades, foi possível captar detalhes operacionais e comportamentais que nem sempre estão evidentes nos registros eletrônicos, como aponta Cervo, Bervian e Silva (2007).

### 3.4 ANÁLISE DOS DADOS

A análise foi conduzida com o apoio da linguagem de programação Python, utilizando bibliotecas como *pandas*, *seaborn*, *matplotlib* e *numpy*, que permitiram a organização, tratamento, visualização e interpretação dos dados coletados.

Inicialmente, foi aplicada a Curva ABC para classificar os modelos de baterias de acordo com a quantidade acumulada de falhas. Essa análise, baseada no princípio de Pareto, revelou que um número relativamente pequeno de produtos responde pela maior parte dos desvios de qualidade, o que reforça a importância da gestão focada nos itens mais críticos.

Em seguida, foram construídos gráficos de barras e gráficos de Pareto para investigar a distribuição das falhas entre os turnos de produção. A análise também considerou a distribuição de falhas entre as linhas de produção da unidade UN01. Por meio de gráficos e visualizações comparativas, identificou-se as linhas mais afetadas por falhas. Linhas com menor volume de produção ou menor complexidade demonstraram índices significativamente mais baixos, sugerindo a existência de correlações entre o perfil da linha e a incidência de não conformidades.

Para aprofundar a análise, foi elaborado um heatmap cruzando linhas de produção e turnos, o que permitiu visualizar de forma integrada em quais turnos as linhas apresentaram maior quantidade de falhas. Outro recurso visual aplicado foi o boxplot, utilizado para avaliar a distribuição e a variabilidade das falhas entre os modelos de baterias mais críticos. Por fim, foi realizada a análise por agrupamentos (clusterização) e séries temporais, com o objetivo de detectar padrões sazonais e comportamentais ao longo do tempo.

De forma geral, os resultados da análise evidenciam que a utilização da EDA com Python foi eficaz para organizar, sintetizar e revelar os principais pontos de fragilidade do processo produtivo. A visualização clara e objetiva das falhas, distribuídas por modelo, linha e turno, permite à empresa direcionar seus esforços de melhoria contínua com maior precisão e efetividade. Ao transformar dados operacionais brutos em informações estruturadas, esta abordagem contribui diretamente para a construção de uma gestão da produção mais eficiente, inteligente e orientada por evidências.

## 4 RESULTADOS

Este capítulo fica a cargo de expor os resultados encontrados nesta pesquisa. Por questões de privacidade, a organização estudada não permite que os seus dados sejam divulgados, de modo que para descaracterizá-los usaremos codinomes ou nomenclaturas genéricas.

### 4.1 ANÁLISE DAS ESTATÍSTICAS DESCRITIVAS DAS FALHAS (POR MODELO DE BATERIA E TIPO DE FALHA)

Esta seção apresenta os resultados da análise estatística realizada sobre a base de dados de falhas da unidade UN01. O objetivo é identificar padrões de ocorrência associados aos modelos de baterias, aos tipos de falhas e à variabilidade dos registros, de forma a compreender quais elementos exercem maior impacto no processo produtivo. Para isso, além do volume total de falhas, foram aplicadas medidas descritivas como média, desvio padrão, valores mínimos e máximos, permitindo uma visão mais aprofundada sobre a intensidade e a dispersão dos problemas.

A Tabela 1 apresenta os dez modelos de baterias com maior incidência de falhas registradas no período analisado.

Tabela 1 – Falhas por modelo de bateria

| Modelo Bateria Anonimizado | Qtd Registros | Total Falhas |
|----------------------------|---------------|--------------|
| Modelo_110                 | 24.0410       | 166.8780     |
| Modelo_216                 | 14.7610       | 91.2670      |
| Modelo_107                 | 10.0210       | 55.8750      |
| Modelo_043                 | 4.9260        | 47.5070      |
| Modelo_044                 | 4.2210        | 38.5520      |
| Modelo_101                 | 5.4760        | 34.5900      |
| Modelo_062                 | 3.3470        | 33.0570      |
| Modelo_040                 | 3.4570        | 25.9760      |
| Modelo_059                 | 4.4890        | 24.4040      |
| Modelo_060                 | 1.9090        | 23.1800      |

Fonte: Esta pesquisa (2025).

O destaque absoluto é o Modelo\_110, responsável por 166.878 falhas, valor que representa o maior volume isolado dentre todos os modelos analisados. Esse resultado indica que este modelo, por si só, exerce impacto significativo na qualidade global do processo produtivo, demandando atenção imediata para a investigação de causas estruturais ou operacionais.

Na sequência, aparecem os Modelos\_216 e Modelo\_107, com 91.267 e 55.875 falhas, respectivamente, ambos figurando como itens de alto impacto no processo de montagem. Em conjunto, os três primeiros modelos somam mais de 313 mil falhas, o que equivale a aproximadamente 56% do total listado na tabela, evidenciando a forte concentração do problema em poucos itens.

Outros modelos também apresentam números expressivos. O Modelo\_043 e o Modelo\_044 registraram 47.507 e 38.552 falhas, configurando-se como produtos igualmente críticos, ainda que em patamares inferiores aos três principais. Já os Modelos\_101 e Modelo\_062 apresentaram 34.590 e 33.057 falhas, respectivamente, mantendo volumes relevantes e que não podem ser desconsiderados.

Na extremidade inferior da lista, mas ainda compondo o grupo dos dez mais problemáticos, encontram-se os Modelos\_040, 059 e 060, com valores entre 23 e 26 mil falhas cada. Embora em patamares mais baixos, esses modelos ainda representam um número significativo de não conformidades e merecem acompanhamento contínuo.

No que diz respeito, a Tabela 2, esta apresenta os principais parâmetros estatísticos relacionados à distribuição das falhas nos dez modelos de baterias com maior incidência de registros. Os indicadores de média, desvio padrão, valores mínimos e máximos permitem compreender não apenas o volume absoluto das falhas, mas também sua dispersão e variabilidade ao longo do período analisado.

Tabela 2 – Distribuição das falhas por modelo de bateria

| <b>Modelo Bateria Anonimizado</b> | <b>Média por Registro</b> | <b>Desvio Padrão</b> | <b>Mínimo</b> | <b>Máximo</b> |
|-----------------------------------|---------------------------|----------------------|---------------|---------------|
| Modelo_110                        | 5.951214                  | 11.876002            | 1.0           | 864.0         |
| Modelo_216                        | 6.182982                  | 16.876285            | 1.0           | 864.0         |
| Modelo_107                        | 5.575791                  | 14.910497            | 1.0           | 720.0         |
| Modelo_043                        | 9.644133                  | 22.245119            | 1.0           | 596.0         |
| Modelo_044                        | 9.133381                  | 16.761075            | 1.0           | 311.0         |
| Modelo_101                        | 6.316654                  | 20.956964            | 1.0           | 575.0         |
| Modelo_062                        | 9.876606                  | 37.148239            | 1.0           | 895.0         |
| Modelo_040                        | 7.541030                  | 19.515734            | 1.0           | 720.0         |
| Modelo_058                        | 5.436400                  | 15.878747            | 1.0           | 879.0         |
| Modelo_059                        | 12.142483                 | 55.335622            | 1.0           | 1248.0        |

Fonte: Esta pesquisa (2025).

Observa-se que os Modelos\_043, 044 e 062 apresentam as maiores médias de falhas por registro, todas próximas ou acima de 9 ocorrências, sugerindo que esses produtos, quando apresentam não conformidades, tendem a gerar falhas em maior intensidade. O Modelo\_059 se

destaca ainda mais, com média de 12,14 falhas por registro, configurando-se como o item mais crítico sob a ótica da gravidade dos eventos registrados.

Por outro lado, os Modelos\_110, 216 e 107, embora liderem em volume absoluto de falhas (conforme Tabela 1), apresentam médias por registro mais baixas, em torno de 5 a 6 falhas. Esse comportamento sugere que o elevado impacto desses modelos está associado mais à frequência de ocorrência do que à intensidade das falhas individuais.

Em relação ao desvio padrão, destaca-se novamente o Modelo\_059, com um valor de 55,33, indicando forte variabilidade na distribuição de falhas. Essa característica sugere a ocorrência de episódios pontuais de falhas em grande escala, possivelmente associados a lotes de produção com problemas específicos ou falhas críticas em componentes. O Modelo\_062 também apresenta alta dispersão (37,14), reforçando a hipótese de que determinadas situações ocasionam volumes anormais de falhas.

Quanto aos valores máximos, alguns modelos chamam a atenção por registrarem falhas de magnitude extremamente elevada em casos isolados. O Modelo\_059 apresentou o maior pico, com 1.248 falhas em um único registro, seguido pelos Modelos\_062 (895), 058 (879) e 110/216 (864). Esses picos, considerados outliers em relação ao comportamento médio, possuem alto potencial de comprometer o desempenho da linha de produção e indicam eventos de falhas críticas que merecem investigação detalhada.

No que se refere ao valor mínimo, todos os modelos apresentam registros unitários (1 falha), demonstrando que a base de dados contempla desde pequenas não conformidades até ocorrências de grande magnitude.

Em se tratando dos tipos de falhas, a Tabela 3 apresenta as falhas com maior incidência no processo produtivo, considerando o número de registros, o total de ocorrências e a média de falhas por registro. Essa análise permite identificar não apenas quais falhas são mais frequentes, mas também quais apresentam maior impacto em termos de intensidade, revelando diferentes perfis de criticidade.

Tabela 3 – Quantitativo das falhas

| Descrição Falha | Qtd Registros | Total Falhas | Média por Registro |
|-----------------|---------------|--------------|--------------------|
| Falha_013       | 23773.00      | 213856.00    | 9.00               |
| Falha_014       | 29552.00      | 200525.00    | 6.79               |
| Falha_015       | 40110.00      | 193838.00    | 4.83               |
| Falha_008       | 11598.00      | 70094.00     | 6.04               |
| Falha_006       | 13713.00      | 59058.14     | 4.31               |
| Falha_005       | 13503.00      | 55947.00     | 4.14               |
| Falha_024       | 3461.00       | 44067.00     | 12.73              |
| Falha_028       | 4534.00       | 25729.00     | 5.67               |
| Falha_025       | 1614.00       | 12595.00     | 7.80               |
| Falha_029       | 114.00        | 11880.00     | 104.21             |

Fonte: Esta pesquisa (2025).

Entre os códigos analisados, destacam-se as Falhas\_013, 014 e 015, que juntas somam mais de 608 mil ocorrências, configurando-se como os principais problemas enfrentados pela unidade UN01. A Falha\_013 lidera em volume total, com 213.856 registros e média de 9 falhas por ocorrência, o que evidencia sua relevância tanto pela frequência quanto pela gravidade. A Falha\_014, embora apresente média menor (6,79), acumula 200.525 falhas e se mantém como um dos pontos mais críticos. A Falha\_015, por sua vez, é a mais recorrente em termos de registros (40.110), mas sua média reduzida (4,83) sugere que, em geral, manifesta-se de forma menos intensa.

Outros códigos também demonstram grande relevância. As Falhas\_008, 006 e 005 apresentaram entre 55 mil e 70 mil ocorrências cada, confirmando que constituem fontes significativas de não conformidades, ainda que com médias por registro mais moderadas, variando de 4,14 a 6,04. Já a Falha\_024 se destaca pela intensidade, com uma média de 12,73 falhas por registro, mesmo apresentando menor frequência absoluta em comparação às anteriores. Esse comportamento sugere que, quando ocorre, trata-se de um evento concentrado e de alto impacto.

A análise ainda revela situações atípicas, como a Falha\_029, que apresenta média de 104,21 falhas por registro, a mais elevada entre todos os códigos. Apesar de ter sido registrada apenas 114 vezes, sua magnitude por evento a torna extremamente crítica, representando episódios de falhas massivas que podem comprometer de forma significativa a estabilidade do processo. Outro caso a ser destacado é a Falha\_025, com 12.595 falhas distribuídas em 1.614 registros, média de 7,80, configurando-se como um problema intermediário, mas de relevância.

Por último, a Tabela 4 apresenta os parâmetros estatísticos referentes à dispersão e amplitude das falhas, representados pelo desvio padrão, valor mínimo e valor máximo de ocorrências

registradas para cada tipo de falha. Esses indicadores fornecem uma visão complementar à análise de médias e totais (Tabela 3), permitindo compreender a variabilidade e a ocorrência de eventos extremos que impactam de forma significativa a estabilidade do processo produtivo.

Tabela 4 – Estatísticas descritivas das falhas

| <b>Descrição Falha</b> | <b>Desvio Padrão</b> | <b>Mínimo</b> | <b>Máximo</b> |
|------------------------|----------------------|---------------|---------------|
| Falha_013              | 21.264409            | 1.0           | 864.0         |
| Falha_014              | 12.683021            | 0.0           | 848.0         |
| Falha_015              | 10.951989            | 0.0           | 720.0         |
| Falha_008              | 16.979457            | 1.0           | 864.0         |
| Falha_006              | 9.366768             | 0.1           | 311.0         |
| Falha_005              | 34.903795            | 1.0           | 4000.0        |
| Falha_024              | 23.757690            | 1.0           | 340.0         |
| Falha_028              | 7.737805             | 1.0           | 100.0         |
| Falha_025              | 9.684827             | 1.0           | 102.0         |
| Falha_029              | 228.040459           | 1.0           | 1248.0        |

Fonte: Esta pesquisa (2025).

Observa-se que a Falha\_029 apresenta o maior desvio padrão, com 228,04, além de um valor máximo de 1.248 ocorrências em um único registro. Esses resultados confirmam o comportamento altamente atípico dessa falha, caracterizado por baixa frequência, mas intensidade extremamente elevada quando ocorre. Trata-se de um fenômeno crítico que deve ser investigado de forma prioritária, pois episódios dessa magnitude podem paralisar linhas inteiras e gerar perdas expressivas.

Entre as falhas mais recorrentes, a Falha\_013 apresenta desvio padrão de 21,26 e registros variando entre 1 e 864 falhas, evidenciando variabilidade significativa, ainda que em patamares controláveis em comparação à Falha\_029. A Falha\_008 mostra padrão semelhante, com desvio de 16,97 e máximo também de 864 falhas, o que sugere que ambas estão sujeitas a episódios de concentração crítica.

A Falha\_014 e a Falha\_015, apesar de liderarem em número absoluto de registros (Tabela 3), possuem desvios padrões mais moderados (12,68 e 10,95, respectivamente), indicando que tendem a se manifestar de forma mais homogênea ao longo do tempo. Isso reforça a hipótese de que seu impacto decorre da constância com que aparecem, e não de eventos pontuais de grande magnitude.

Casos específicos também merecem atenção, como a Falha\_005, que apresenta desvio padrão de 34,90 e máximo de 4.000 ocorrências em um único registro. Embora menos frequente em relação às falhas principais, sua variabilidade extrema sugere problemas localizados de

grande impacto, possivelmente vinculados a defeitos de lote ou falhas em processos críticos da produção. Já a Falha\_024, com desvio de 23,75, e a Falha\_025, com máximo de 102, configuram falhas intermediárias, com eventos que podem comprometer parcialmente a produção, ainda que em menor escala.

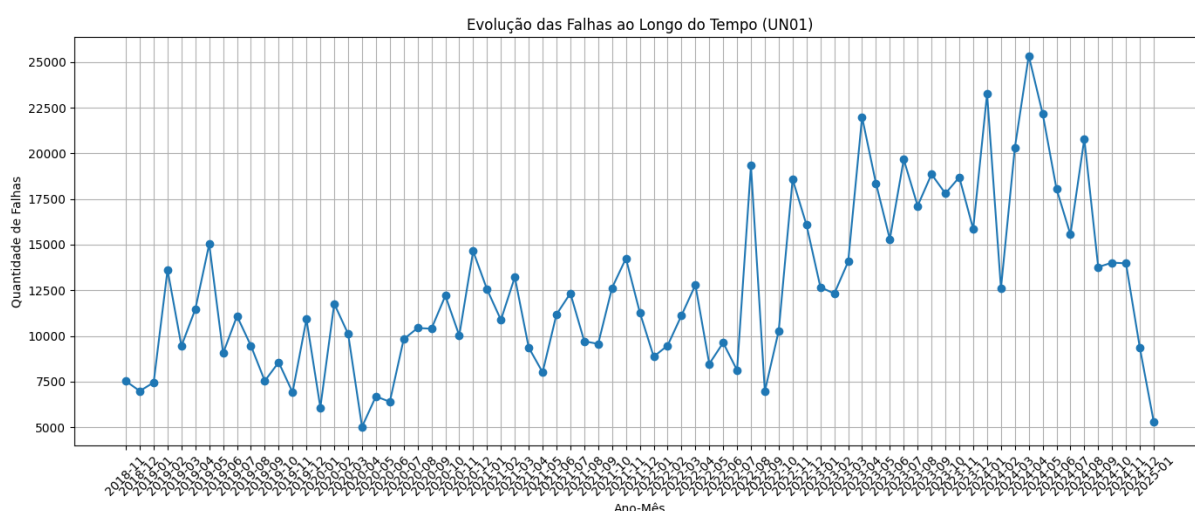
Em contrapartida, falhas como a Falha\_028 apresentam baixa variabilidade (desvio de 7,73) e amplitude restrita (máximo de 100 falhas), caracterizando um perfil de menor criticidade em termos de dispersão, embora não possam ser desconsideradas em função do volume acumulado.

De modo geral, a análise evidencia dois grupos distintos de falhas: aquelas com alta variabilidade e episódios extremos (como as Falhas\_029 e 005), que representam riscos imediatos e de grande escala para a produção; e aquelas de baixa variabilidade, mas alta recorrência (como as Falhas\_013, 014 e 015), que comprometem a qualidade de forma contínua e exigem monitoramento sistemático. Essa distinção reforça a necessidade de estratégias de controle diferenciadas: abordagens emergenciais e específicas para falhas de alta dispersão e políticas de prevenção e padronização para falhas de ocorrência frequente e homogênea.

#### 4.2 ANÁLISE DE SÉRIE TEMPORAL DE FALHAS (EVOLUÇÃO DAS FALHAS AO LONGO DO TEMPO)

A Figura 1 apresenta a evolução da quantidade de falhas registradas ao longo do tempo na unidade UN01, no período de novembro de 2018 a janeiro de 2025. A representação gráfica em linha permite visualizar as tendências, variações e padrões sazonais que caracterizaram o comportamento das falhas ao longo dos anos.

Figura 1 – Série temporal da evolução das falhas



Fonte: Esta pesquisa (2025).

No início do período analisado, entre 2018 e 2020, observa-se uma relativa estabilidade nos registros, com oscilações em torno de 6.000 a 15.000 falhas mensais. Apesar das variações pontuais, o volume de falhas manteve-se dentro de uma faixa relativamente constante. A partir

de 2021, contudo, verifica-se um aumento gradual na quantidade de falhas, com elevação mais expressiva ao longo dos meses subsequentes. Esse crescimento torna-se ainda mais acentuado em 2022, ano em que os registros atingem os maiores patamares da série, chegando a superar a marca de 25.000 falhas em determinados momentos.

O comportamento ao longo de 2023 é marcado por grande volatilidade, com variações bruscas na quantidade de falhas de um mês para outro. Essa instabilidade pode indicar a ocorrência de mudanças nos processos produtivos, alteração nos critérios de registro, ou ainda a presença de fatores externos que impactaram a operação da unidade. Já no final de 2024 e início de 2025, é possível identificar uma tendência de redução significativa nas falhas, sugerindo a efetividade de eventuais ações corretivas implementadas.

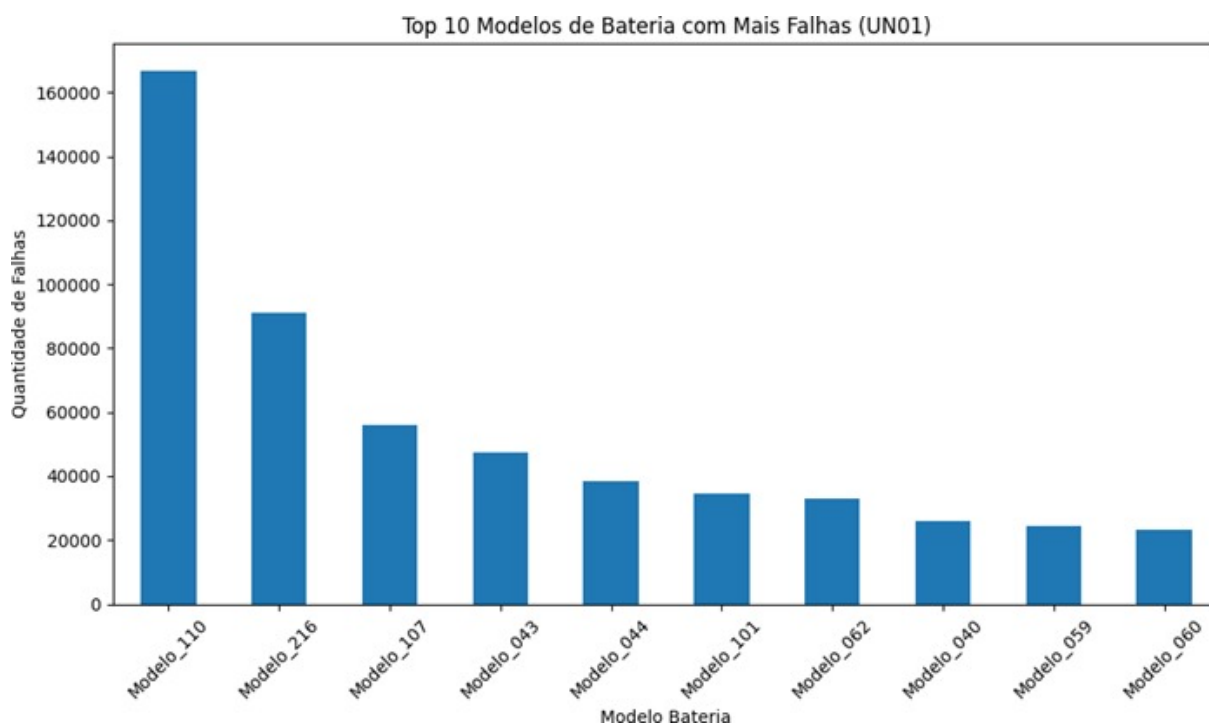
De maneira geral, a análise da série temporal revela uma tendência de crescimento no número de falhas ao longo do tempo, intercalada por períodos de oscilação e recente declínio. A persistência desse comportamento reforça a importância de monitorar continuamente os processos produtivos, aprofundar a investigação sobre as causas das falhas e consolidar as melhorias que começaram a ser observadas no período mais recente da série.

#### 4.3 FALHAS POR PRODUTOS

A primeira etapa da análise exploratória de dados consistiu na identificação dos modelos de baterias que apresentaram maior incidência de falhas no processo de montagem. Para isso, utilizou-se a base de dados extraída do sistema ERP da empresa, composta por registros de falhas referentes a 207 diferentes modelos de produtos.

A análise revelou que as falhas não se distribuem de maneira uniforme entre os modelos, concentrando-se em alguns produtos específicos. Os dez modelos que mais apresentaram falhas foram, em ordem decrescente de ocorrência Modelo\_110, Modelo\_216, Modelo\_107, Modelo\_43, Modelo\_44, Modelo\_101, Modelo\_62, Modelo\_40, Modelo\_101 e Modelo\_107, conforme pode-se ver na Figura 2.

Figura 2 – Top 10 modelos de bateria com mais falhas



Fonte: Esta pesquisa (2025).

A Figura 2 ilustra a quantidade de falhas registradas para cada um desses modelos, evidenciando que o modelo Modelo\_110 lidera o ranking com o maior número de falhas, enquanto o modelo Modelo\_107 aparece na décima posição, com um número relativamente menor.

O Modelo\_110 apresenta o maior número de falhas, mas é importante destacar que ele é o modelo mais produzido na unidade.

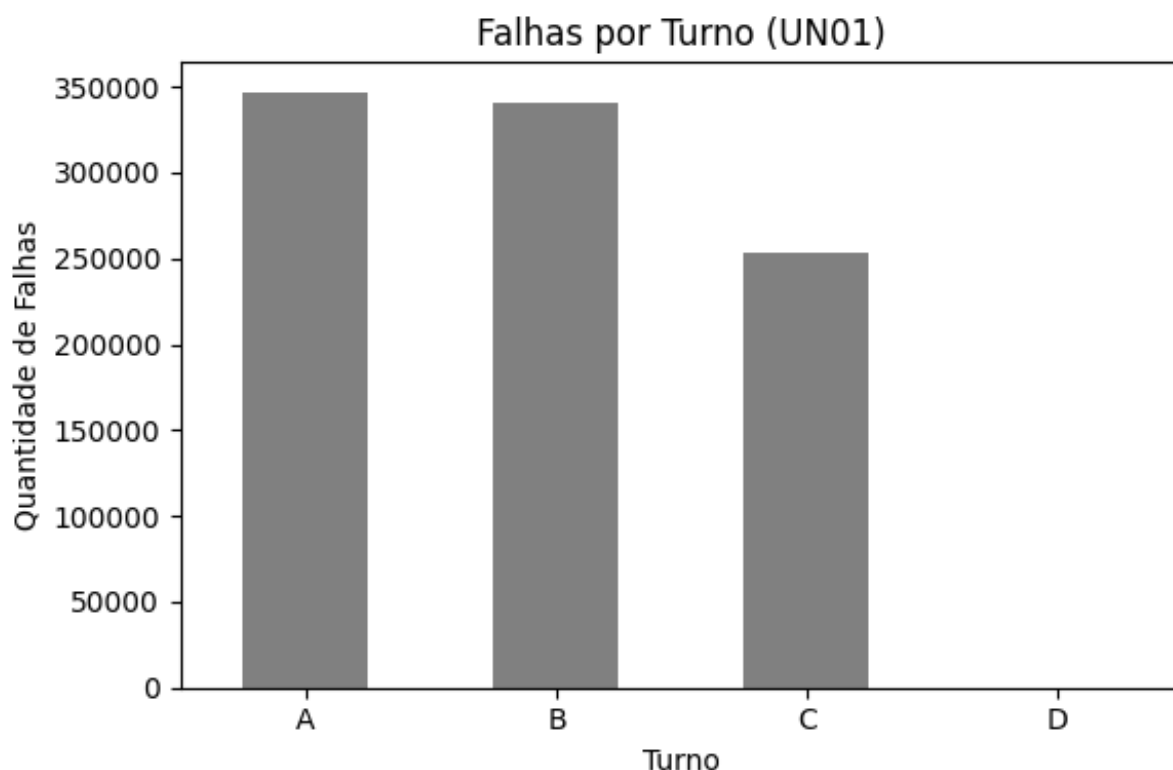
A concentração de falhas em determinados modelos sugere a necessidade de investigação aprofundada sobre características específicas de cada produto ou de seus respectivos processos de montagem, visando a identificação de possíveis causas estruturais ou operacionais que estejam contribuindo para a ocorrência de não conformidades.

#### 4.4 ANÁLISE DAS FALHAS POR TURNO

A segunda etapa da análise considerou a distribuição das falhas conforme o turno de trabalho da unidade fabril (UN01). A planta industrial opera em quatro turnos distintos, denominados A, B, C e D.

Conforme ilustrado na Figura 3, observou-se que os turnos A e B apresentaram as maiores quantidades de falhas registradas, ambos com valores próximos, seguidos pelo turno C, que apresentou uma quantidade significativamente menor. O turno D, por sua vez, demonstrou um número praticamente irrelevante de falhas em comparação aos demais por só ter produzido em um certo período de tempo, portanto é melhor desconsiderá-lo para comparações.

Figura 3 – Falhas por turno



Fonte: Esta pesquisa (2025).

Esses resultados indicam que os turnos A e B concentram a maior parte dos registros de falhas, o que pode estar relacionado a fatores como maior volume de produção, maior tempo de operação, ou mesmo aspectos específicos da gestão e da execução dos processos nesses períodos. A expressiva diferença observada entre os turnos sugere a necessidade de uma análise mais detalhada dos fatores operacionais e humanos envolvidos em cada um.

Com base no princípio de Pareto, torna-se evidente que os esforços de melhoria devem ser concentrados nos turnos A e B, uma vez que as ações corretivas nesses períodos têm potencial de impactar a maior parte das falhas identificadas. Estratégias como análise de causa raiz, revisões de procedimentos operacionais, reforço de treinamentos e revisões de processos específicos dos turnos A e B poderão gerar ganhos significativos de eficiência e qualidade.

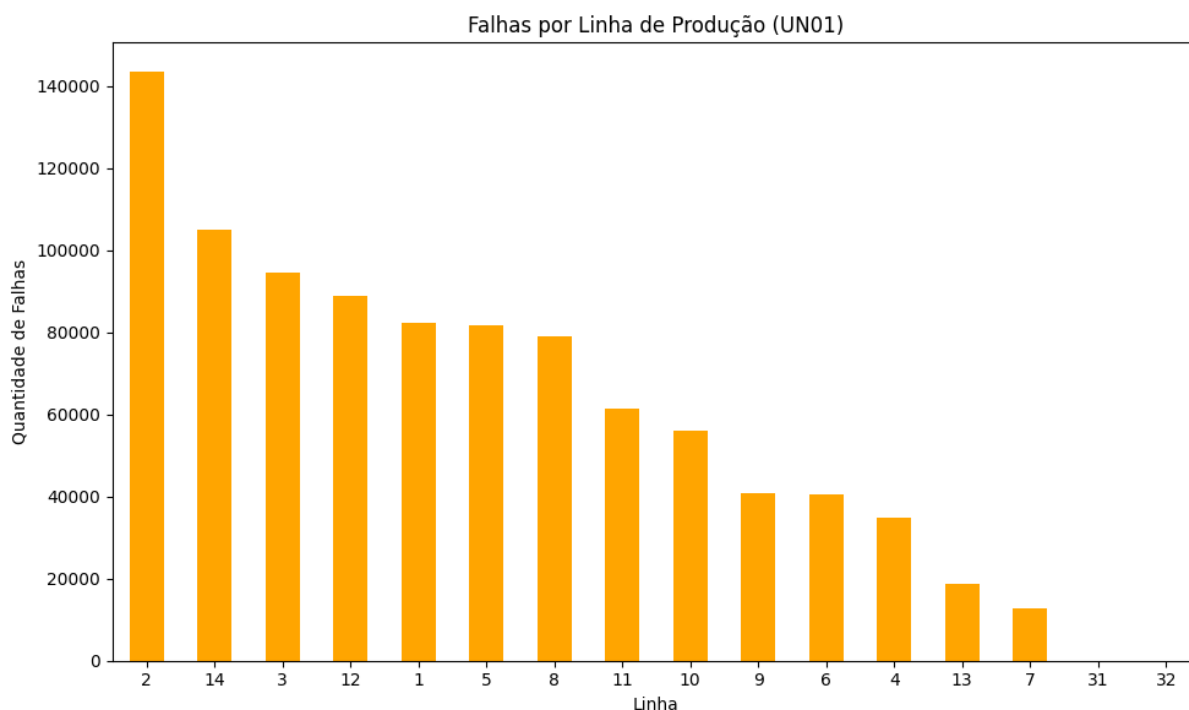
#### 4.5 ANÁLISE DAS FALHAS POR LINHA DE PRODUÇÃO

Além da análise por turno, também foi realizada uma avaliação da distribuição das falhas entre as linhas de produção da unidade UN01. Atualmente, a planta conta com 14 linhas ativas, numeradas conforme o sistema interno de identificação.

Conforme ilustrado na Figura 4, a linha 2 foi a que registrou a maior quantidade de falhas, seguida pelas linhas 14, 3 e 12. Na sequência, aparecem as linhas 1, 5, 8, 11 e 10, que também apresentaram volumes expressivos de falhas, ainda que inferiores às quatro primeiras. As linhas 9, 6 e 4 mostraram uma quantidade intermediária de registros, enquanto as linhas 13 e

7 apresentaram índices de falhas mais baixos. As linhas 31 e 32 foram aquelas que tiveram o menor número de falhas registradas.

Figura 4 – Falhas por linha de produção



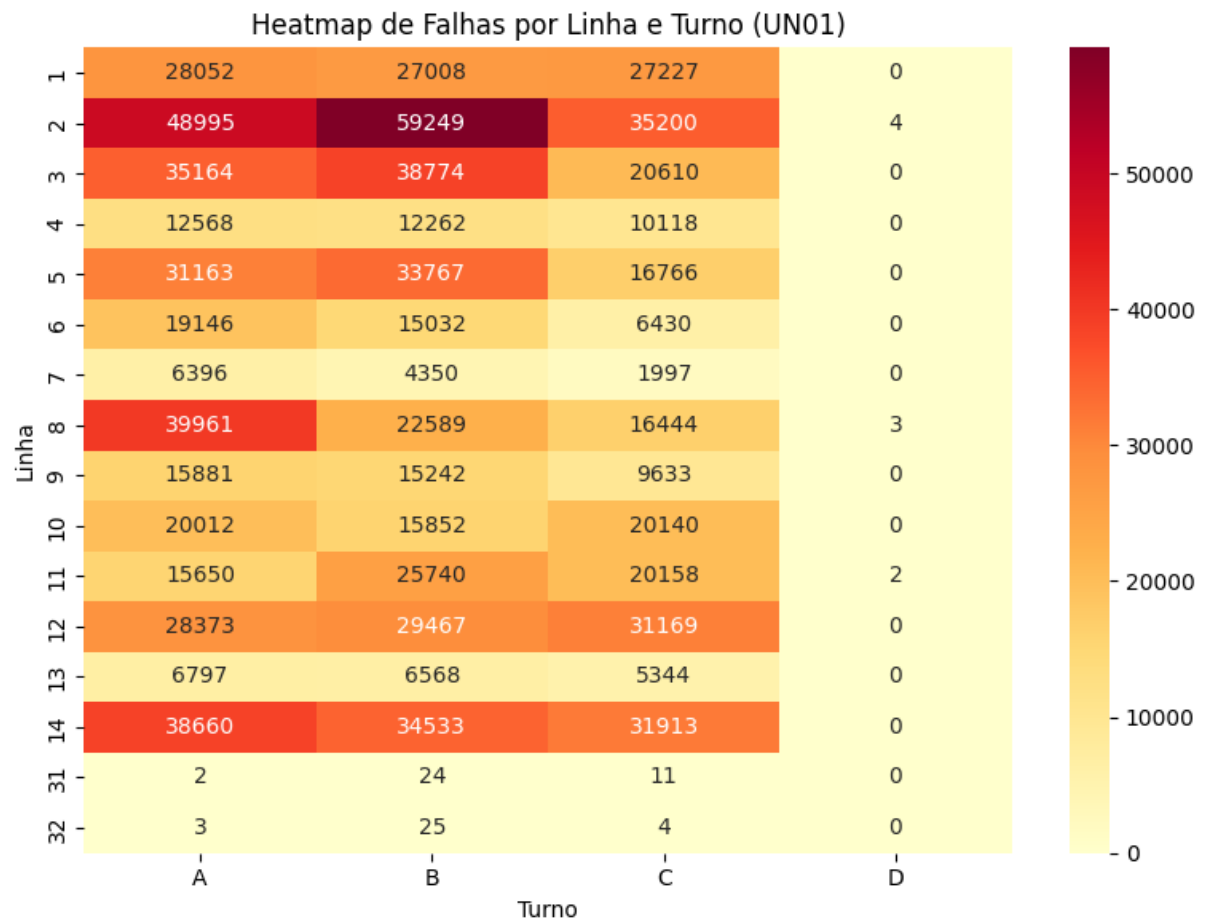
Fonte: Esta pesquisa (2025).

A expressiva diferença observada entre as linhas, especialmente o destaque da linha 2, indica a necessidade de uma investigação mais aprofundada para identificar fatores específicos que possam estar contribuindo para a maior incidência de falhas, como, por exemplo, o tipo de produto fabricado, o nível de automação, as condições de manutenção ou a experiência da equipe.

#### 4.6 HEATMAP DE FALHAS POR LINHA E TURNO DE PRODUÇÃO

Buscando um entendimento mais aprofundado dos dados, foi elaborado um heatmap, conforme pode-se ver na Figura 5, que cruza as informações de falhas entre as diferentes linhas de produção e os respectivos turnos de trabalho na unidade UN01. Essa análise possibilita visualizar não apenas quais linhas falharam mais, mas também em quais turnos essas falhas se concentraram, oferecendo uma perspectiva integrada e diferenciada.

Figura 5 – Heatmap de falhas por linha e turno de produção



Fonte: Esta pesquisa (2025).

O heatmap revela que a Linha 2 é a que apresenta o maior número de falhas, especialmente no turno B, seguido pelos turnos A e C. Esse padrão reforça o que já havia sido apontado nas análises anteriores, indicando que o turno B é particularmente crítico para esta linha. Também pode-se enxergar que a Linha 14 e Linha 3 também mantêm volumes elevados de falhas, novamente com predominância no turno B, sugerindo que esse turno pode apresentar fragilidades operacionais que merecem investigação.

Em linhas como a 1, 5, 8, 12 e 10, embora os números sejam menores que nas linhas principais, também se observa um comportamento semelhante: o turno B geralmente registra mais falhas do que os turnos A e C. A Linha 8 apresenta um comportamento interessante: apesar do volume total de falhas ser mais moderado, a discrepância entre o turno A (mais falhas) e os demais turnos é significativa, sugerindo uma possível particularidade na operação do turno A para essa linha.

Linhas com menor volume total de falhas, como a 6, 4, 13 e 7, apresentam uma distribuição mais homogênea, porém com volumes muito baixos, o que indica que essas linhas, embora funcionem com menos problemas, também devem ser monitoradas para garantir que mantenham esse desempenho. As linhas 31 e 32 praticamente não apresentaram falhas, estando consistentemente

temente zeradas nos turnos analisados. Isso pode indicar que essas linhas estão operando em níveis muito controlados ou que possuem uma produção muito reduzida no período.

Esses resultados mostram que não apenas a linha de produção, mas também o turno no qual as atividades são realizadas, exerce forte influência sobre o volume de falhas. O destaque do turno B em diversas linhas sugere que fatores como a equipe alocada, o tempo de operação das máquinas, o cansaço acumulado ou mesmo o ambiente físico podem estar impactando diretamente a eficiência produtiva.

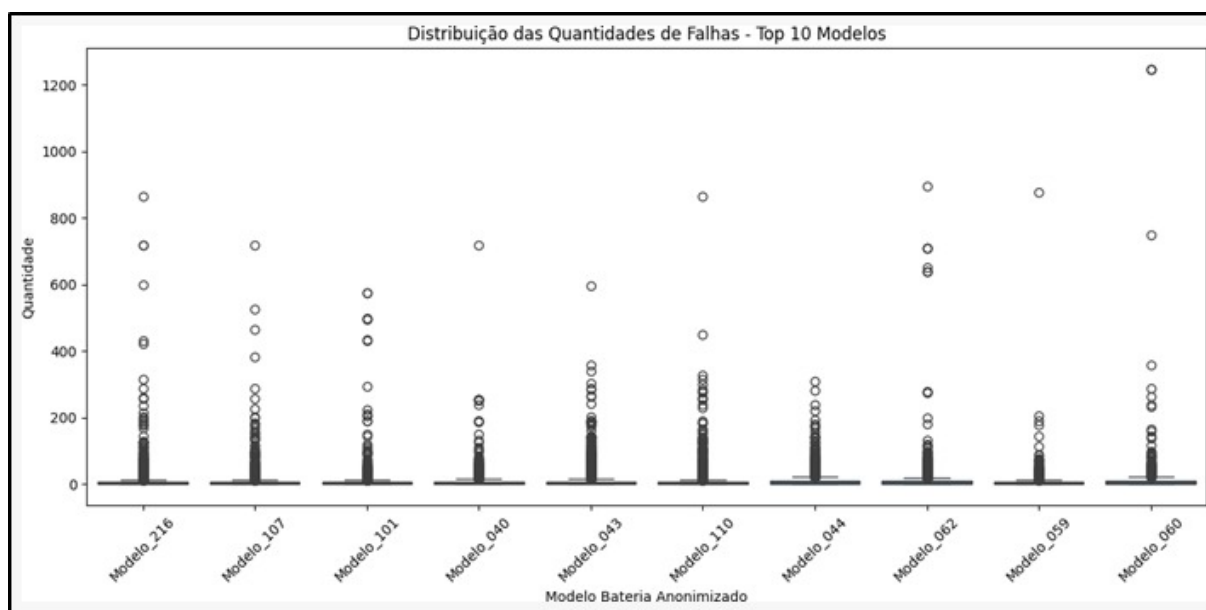
Portanto, recomenda-se que as próximas investigações analisem especificamente o que ocorre no turno B: rotinas de operação, treinamento dos operadores, manutenção preventiva de equipamentos, entre outros fatores que possam esclarecer o motivo das recorrências mais elevadas nesse período.

Essa abordagem integrada entre linha e turno amplia significativamente a compreensão sobre os padrões de falhas e permite direcionar ações corretivas de maneira mais estratégica e eficiente.

#### 4.7 BOXPLOT DA DISTRIBUIÇÃO DAS FALHAS POR MODELO DE BATERIA

Na sequência, buscou-se aprofundar a compreensão sobre o comportamento das falhas entre os diferentes modelos de baterias, e para isto foi elaborado o boxplot representado na Figura 6. Esse gráfico possibilita visualizar a distribuição da quantidade de falhas para os dez modelos que mais apresentaram problemas durante o período analisado.

Figura 6 – Boxplot da distribuição das quantidades de falhas por modelo



Fonte: Esta pesquisa (2025).

A análise do boxplot permite inferir que todos os modelos apresentam uma grande quantidade de outliers, evidenciando que, embora a maior parte das ocorrências de falhas se concentre

em valores mais baixos, há casos isolados de falhas extremamente elevadas. Modelos como o Modelo\_107, Modelo\_101 e Modelo\_62 se destacam por apresentarem valores máximos de falhas bastante superiores em relação aos demais, indicando que, embora sejam menos recorrentes, eventos de falhas críticas ocorreram para esses modelos.

A maioria dos modelos possui medianas muito próximas da base do boxplot, o que sugere que a maior parte das falhas ocorre em volumes baixos, sendo que as grandes quantidades de falhas são mais exceções do que regra. Modelos como Modelo\_40 e Modelo\_43 exibem uma distribuição mais compacta, indicando uma maior estabilidade nas quantidades de falhas observadas, ainda que também apresentem outliers importantes. Observa-se que, apesar da existência de casos críticos, a tendência geral para a maioria dos modelos é de um número relativamente baixo de falhas por ocorrência.

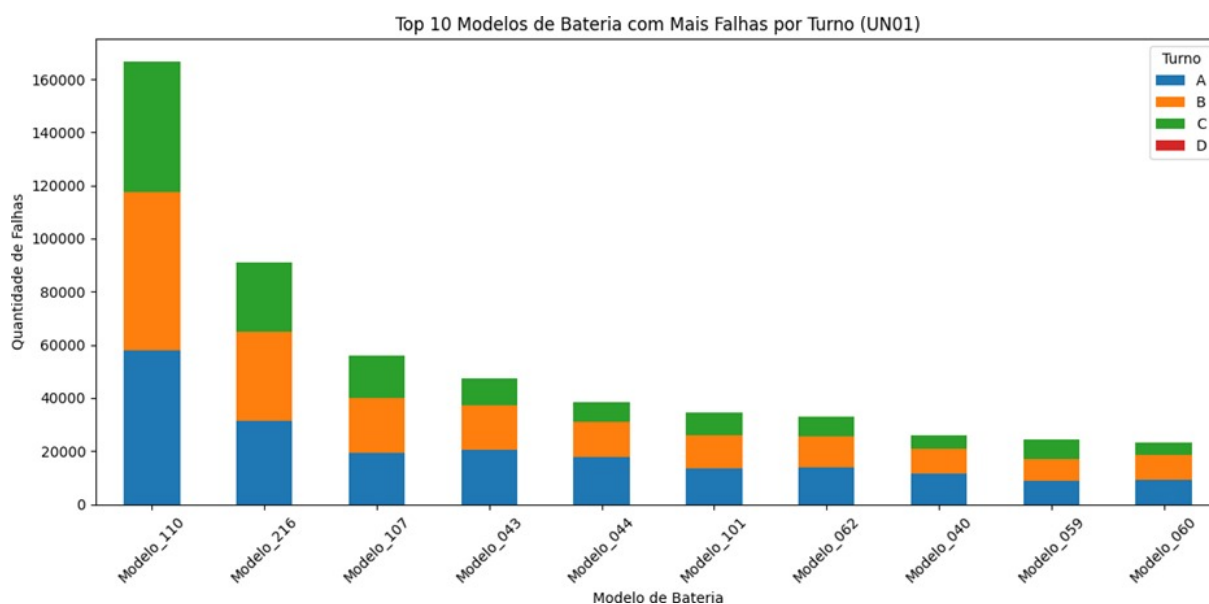
Em resumo, a análise do boxplot revela que os eventos de falhas em grande quantidade são pouco frequentes, mas quando ocorrem, atingem níveis bastante elevados, o que pode gerar impactos significativos na operação e nos índices de qualidade. Esses picos de falhas podem estar associados a fatores pontuais, como lotes de produção defeituosos, problemas na linha de montagem ou falhas específicas de componentes, e merecem investigação detalhada.

Assim, os resultados reforçam a necessidade de monitoramento constante e de implementação de ações preventivas voltadas para os modelos que apresentaram maior variabilidade e maior incidência de falhas críticas.

#### 4.8 ANÁLISE DA QUANTIDADE DE FALHAS POR TURNO NOS PRINCIPAIS MODELOS DE BATERIA

A Figura 7 apresenta o gráfico de barras empilhado, que ilustra a distribuição da quantidade de falhas por turno (A, B, C e D) nos dez modelos de bateria com maior incidência de falhas na unidade UN01.

Figura 7 – Análise da quantidade de falhas por turno nos principais modelos de bateria



Fonte: Esta pesquisa (2025).

A partir da análise do gráfico, é possível identificar que o modelo Modelo\_110 apresentou o maior número de falhas, com uma contribuição significativa de todos os turnos, especialmente dos turnos B e C. O modelo Modelo\_216 ocupa a segunda posição em quantidade de falhas, sendo também o turno B o principal responsável pela maior parte das ocorrências.

De maneira geral, para a maioria dos modelos analisados, os turnos A e B concentraram a maior parte das falhas, o que pode indicar fatores operacionais ou de processo que merecem uma investigação mais aprofundada. Os modelos Modelo\_43, Modelo\_44 e Modelo\_101 apresentam uma distribuição relativamente mais equilibrada das falhas entre os turnos A, B e C, embora, novamente, o turno B tenha uma contribuição levemente superior. No caso dos modelos Modelo\_62, Modelo\_40, Modelo\_101 e Modelo\_107, a predominância do turno A é mais evidente, sugerindo que a ocorrência de falhas nesses modelos possa estar mais associada a condições específicas desse turno.

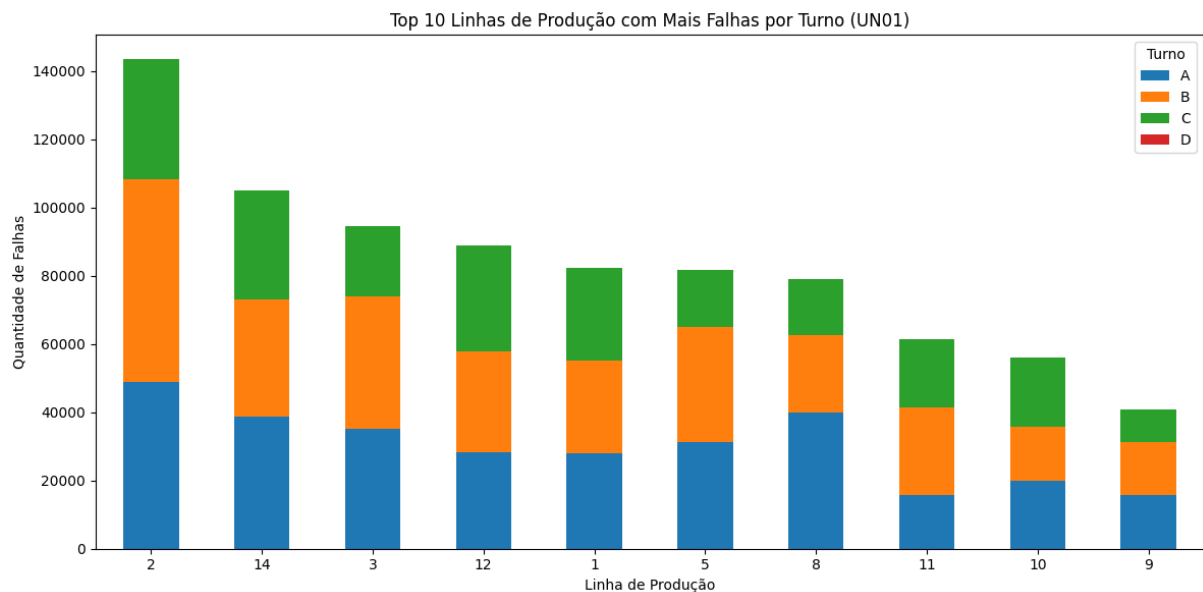
Essa distribuição das falhas por turno é um indicativo importante para a gestão da produção e da qualidade, pois aponta possíveis variações no desempenho das operações conforme o turno de trabalho. Fatores como troca de equipes, condições de manutenção dos equipamentos, ajustes de processo ou até aspectos ligados à rotina de inspeção e controle de qualidade podem impactar diretamente esses resultados.

Dessa forma, torna-se imprescindível analisar em detalhe os procedimentos e práticas de cada turno, visando identificar oportunidades de melhoria para reduzir a incidência de falhas, especialmente nos modelos mais críticos.

#### 4.9 ANÁLISE DA QUANTIDADE DE FALHAS POR TURNO NAS PRINCIPAIS LINHAS DE PRODUÇÃO

A Figura 8 exibe o gráfico de barras empilhado com a distribuição das falhas por turno (A, B, C e D) nas dez linhas de produção que mais registraram falhas na unidade UN01.

Figura 8 – Análise da quantidade de falhas por turno nas principais linhas de produção



Fonte: Esta pesquisa (2025).

A análise dos dados revela que a linha 2 é a que apresentou o maior número de falhas, destacando-se principalmente os turnos A, B e C, com volumes expressivos e relativamente próximos entre si. A linha 14 ocupa o segundo lugar em quantidade de falhas, com predominância dos turnos A e C, enquanto o turno B contribui de forma significativa, porém menor em comparação. As linhas 3 e 12 também apresentam números elevados de falhas, com forte participação dos turnos A e B, e menor contribuição do turno C. A distribuição das falhas nas linhas 1, 5, 8, 11, 10 e 9 mostra uma tendência semelhante, sendo os turnos A e B os principais responsáveis pela maioria das falhas.

A identificação das linhas de produção mais críticas é fundamental para direcionar ações de melhoria e planos de manutenção corretiva e preventiva. Além disso, o fato de os turnos A e B apresentarem as maiores incidências de falhas reforça a necessidade de avaliar, de maneira detalhada, as práticas operacionais e os fatores que podem estar influenciando o desempenho dessas linhas durante esses turnos.

Uma abordagem focada na análise de causa raiz, combinada com treinamentos específicos para as equipes de operação e manutenção, pode contribuir significativamente para a redução das falhas e melhoria da eficiência produtiva.

#### 4.10 ANÁLISE DA CURVA ABC

A análise dos dados operacionais da unidade UN01 no período compreendido entre os anos de 2020 a 2024 permitiu a construção de um diagnóstico detalhado sobre a incidência de falhas nos modelos de baterias nas linhas de produção do setor de montagem. Dentre os 207 modelos de baterias disponíveis no portfólio da organização, foi possível identificar um conjunto relativamente pequeno de itens que concentram a maior parte dos problemas reportados. Para essa identificação, foi empregada a ferramenta de gestão conhecida como Curva ABC, cuja proposta metodológica é classificar itens de acordo com sua relevância quantitativa, neste caso, considerando a frequência absoluta de falhas observadas.

A Curva ABC é amplamente utilizada na gestão de estoques, qualidade, produção e manutenção, por sua capacidade de apoiar a priorização de esforços em contextos organizacionais nos quais os recursos são limitados e a eficiência na alocação torna-se estratégica. Na abordagem adotada neste estudo, definiu-se que os itens da Classe A corresponderiam a aproximadamente 70% do total de falhas acumuladas, a Classe B corresponderia a 20% adicionais (alcançando 90%) e a Classe C os 10% restantes, compondo a totalidade dos registros. Neste sentido, os resultados obtidos com a aplicação da curva ABC, podem ser consultados na Tabela 01.

Tabela 5 – Curva ABC

| <b>Modelo Bateria Anonimizado</b> | <b>Falhas</b> | <b>% Acumulado</b> | <b>Classe ABC</b> |
|-----------------------------------|---------------|--------------------|-------------------|
| Modelo_110                        | 166878.0      | 17.473953          | A                 |
| Modelo_216                        | 91267.0       | 27.448272          | A                 |
| Modelo_107                        | 55875.0       | 33.389399          | A                 |
| Modelo_043                        | 47507.0       | 38.440766          | A                 |
| Modelo_044                        | 38552.0       | 42.539957          | A                 |
| Modelo_101                        | 34590.0       | 46.217874          | A                 |
| Modelo_062                        | 33057.0       | 49.732788          | A                 |
| Modelo_040                        | 25976.0       | 52.494787          | A                 |
| Modelo_059                        | 24404.0       | 55.089638          | A                 |
| Modelo_060                        | 23180.0       | 57.554341          | A                 |

Fonte: Esta pesquisa (2025).

O resultado da análise revelou uma concentração de falhas em apenas 10 modelos de baterias, os quais, somados, totalizam 554.966 ocorrências, o equivalente a 57,55% do total registrado no período analisado. Esses modelos compõem, portanto, a Classe A da Curva ABC e, de acordo com os critérios da metodologia, são os que demandam máxima atenção por parte da organização.

Destaca-se de forma evidente o Modelo\_110, responsável por 166.878 falhas, representando isoladamente 17,74% de todas as ocorrências. Esse modelo, por si só, compromete quase um quinto da qualidade do produto final entregue pelas linhas de montagem da UN01, sinalizando a urgência de ações corretivas e preventivas específicas. Os Modelo\_216 e Modelo\_107, com 91.267 e 55.875 falhas respectivamente, também figuram entre os mais críticos, acumulando juntos mais de 27% do total de falhas.

Ao expandir a análise até o Modelo\_60, observa-se que mesmo os itens na extremidade inferior da Classe A, como os Modelo\_59 e Modelo\_60, apresentam números expressivos de falhas, com mais de 23 mil ocorrências cada um.

Do ponto de vista gerencial, os achados da Curva ABC permitem uma reorientação estratégica no modo como a organização compreende e responde aos seus problemas de qualidade. A priorização dos modelos da Classe A torna possível desenvolver planos de ação focados, que maximizam o retorno dos investimentos em melhoria contínua. Isso significa, por exemplo, redirecionar esforços de auditoria interna, revisar parâmetros técnicos e processos produtivos especificamente associados a esses modelos, intensificar testes laboratoriais e, se necessário, renegociar padrões de fornecimento com parceiros externos.

Além disso, há implicações diretas na gestão do conhecimento e da memória organizacional. Ao identificar padrões históricos de falha e associá-los a modelos específicos de produto, a organização passa a construir um banco de dados estruturado que documenta não apenas as ocorrências, mas as decisões tomadas e os resultados dessas decisões. Este acervo se torna um ativo estratégico para a organização, permitindo que, em ciclos futuros, gestores possam consultar experiências anteriores e tomar decisões mais rápidas e fundamentadas, evitando a repetição de erros e promovendo o aprendizado institucionalizado.

Essa estruturação do conhecimento também favorece uma atuação mais sistêmica, pois as informações obtidas por meio da análise da Curva ABC podem ser compartilhadas entre áreas como engenharia, produção, suprimentos, logística e comercial. Isso favorece uma gestão integrada, que reconhece a complexidade dos problemas de qualidade como fenômenos multifatoriais que não podem ser tratados isoladamente.

Adicionalmente, os dados evidenciam a importância da gestão por indicadores. A mera percepção de que um modelo “parece apresentar mais defeitos” não é suficiente para embasar decisões em contextos complexos. A análise objetiva, baseada em dados históricos e metodologias consagradas, como a Curva ABC, eleva o nível de maturidade da gestão organizacional e consolida práticas gerenciais alinhadas à lógica da melhoria contínua e da excelência operacional.

A concentração de problemas em poucos modelos também oferece uma oportunidade de ganho financeiro. Reduzir significativamente as falhas nos modelos da Classe A pode resultar em impacto positivo direto sobre os indicadores de produtividade, retrabalho, devoluções, garantia e satisfação do cliente final. Tais indicadores não apenas afetam os custos operacionais, como também moldam a imagem da empresa no mercado, interferindo na sua competitividade

em um setor altamente sensível à qualidade, como é o de baterias automotivas.

No que diz respeito à unidade UN01, especificamente, os resultados servem como um alerta e ao mesmo tempo como uma oportunidade. A unidade pode se beneficiar enormemente da adoção de programas específicos de melhoria da qualidade voltados aos 10 modelos identificados. Isso pode incluir a criação de células de trabalho interdepartamentais, campanhas de treinamento técnico voltadas para as equipes que lidam com esses produtos, e até mesmo a aplicação de metodologias como Seis Sigma, FMEA ou Kaizen para atacar as causas raízes dos problemas identificados.

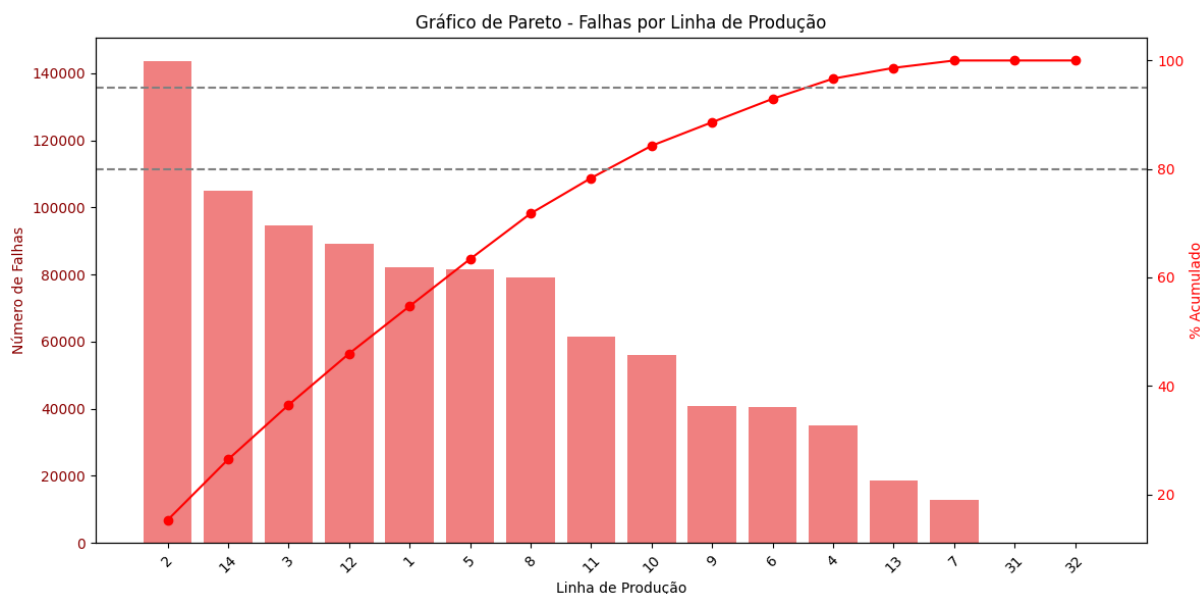
De maneira mais ampla, a organização como um todo pode utilizar este diagnóstico como base para decisões estratégicas relativas ao portfólio de produtos, política de desenvolvimento de novos modelos e até mesmo revisão de critérios de homologação técnica. A pesquisadora vislumbra que essa abordagem aqui adotada pode ser útil também para as demais unidades da organização. A centralização dos problemas em modelos específicos pode levar à descontinuidade de alguns itens com alta reincidência de falhas e baixo retorno econômico, ou à reformulação de seu projeto.

Em suma, a análise da Curva ABC revelou um retrato claro e inequívoco da distribuição das falhas ao longo do portfólio de baterias da organização. Mais do que uma simples categorização de dados, essa análise se apresenta como uma poderosa ferramenta de apoio à decisão, permitindo que a organização transforme dados brutos em ações gerenciais concretas, sustentadas por evidências e orientadas para a excelência operacional. Ao sistematizar o aprendizado obtido com essa análise e incorporá-lo à sua memória organizacional, a empresa dá um passo significativo na direção de uma gestão mais inteligente, resiliente e orientada por resultados.

#### 4.11 GRÁFICO DE PARETO (FALHAS POR LINHA DE PRODUÇÃO)

A Figura 9 apresenta o Gráfico de Pareto referente às falhas distribuídas pelas linhas de produção da unidade UN01. Assim como no gráfico anterior, este recurso visual facilita a identificação das linhas que mais contribuem para o volume de falhas, reforçando a lógica de priorização de esforços conforme o princípio de Pareto (80/20).

Figura 9 – Gráfico de Pareto (Falhas por linha de produção)



Fonte: Esta pesquisa (2025).

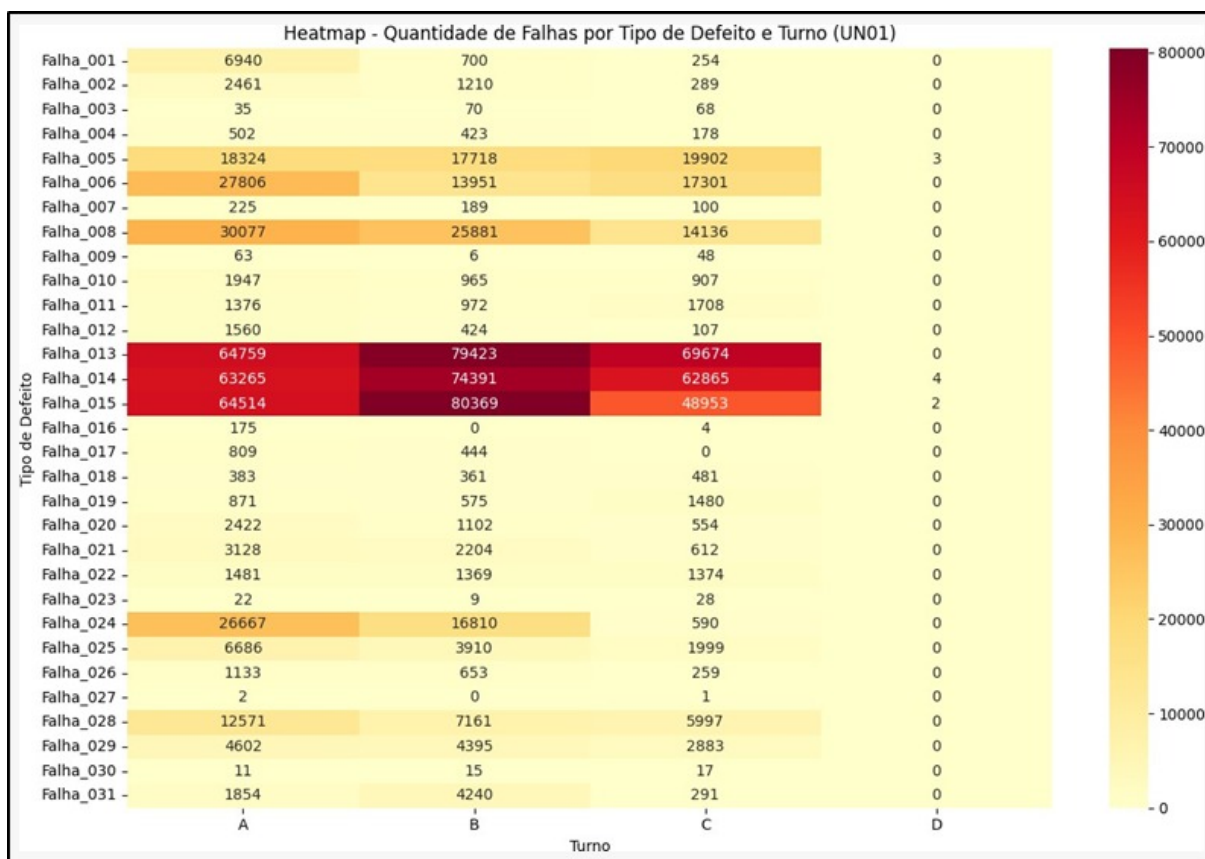
A interpretação dos dados revela que as Linhas 2, 14, 3 e 12 concentram os maiores volumes de falhas, destacando-se com números acima de 90.000 falhas cada uma. Essas quatro linhas, juntas, ultrapassam a marca de 80% das falhas acumuladas no processo, conforme evidenciado pela linha de % acumulado. As Linhas 1, 5, 8 e 11 também aparecem com volumes consideráveis, porém já contribuindo de forma mais moderada para o total de falhas. As demais linhas (10, 9, 6, 4, 13, 7, 31 e 32) apresentam falhas em volumes significativamente menores, mostrando uma tendência de baixa influência no total geral.

A análise aponta que ações de melhoria voltadas especialmente para as linhas 2, 14, 3 e 12 são prioritárias e podem gerar impacto direto na redução do número de falhas no processo como um todo. Recomenda-se a realização de diagnósticos específicos para essas linhas, abordando fatores como treinamento de operadores, manutenção de equipamentos, qualidade de insumos e ajustes de processos produtivos.

#### 4.12 HEATMAP – QUANTIDADE DE FALHAS POR TIPO DE DEFEITO E TURNO

A Figura 10 apresenta o heatmap da quantidade de falhas por tipo de defeito (identificados por códigos) e por turno de trabalho na unidade UN01. A utilização do recurso visual do heatmap facilita a identificação das concentrações mais elevadas de falhas, por meio da variação da intensidade de cor, permitindo uma análise mais imediata e comparativa entre os diferentes tipos de defeitos e turnos.

Figura 10 – Heatmap de quantidade de falhas por tipo de defeito e turno



Fonte: Esta pesquisa (2025).

Observa-se que as falhas Falha 13, Falha 14 e Falha 15 concentram os maiores volumes de ocorrências registrados, independentemente do turno. Estes três códigos se destacam de forma significativa, representando, isoladamente, os pontos mais críticos no processo produtivo. Especificamente, a Falha 13 apresenta valores elevados em todos os turnos analisados (A, B e C), atingindo a maior quantidade no turno B (79.423 falhas), seguido pelo turno A (64.759 falhas) e turno C (69.674 falhas).

As falhas Falha 14 e Falha 15 também exibem comportamento similar, com predominância de registros no turno B, reforçando a hipótese de que este turno apresenta maior vulnerabilidade ou exposição a condições de falha. Essa tendência pode ser atribuída a variáveis como aumento de volume de produção, diferença de experiência das equipes, variações nos processos de controle de qualidade ou fatores organizacionais específicos.

Além dos códigos mencionados, outras falhas que merecem destaque são: Falha 8, Falha 6 e Falha 24. Embora em patamares inferiores às três principais, também apresentam volumes expressivos, concentrados especialmente nos turnos A e B. O turno C, por sua vez, registra quantidades significativamente menores para a maioria dos tipos de falha, o que pode indicar menor carga de trabalho, melhor controle dos processos ou características específicas desse período de operação.

Outro ponto relevante é a quase inexistência de falhas no turno D, com registros prati-

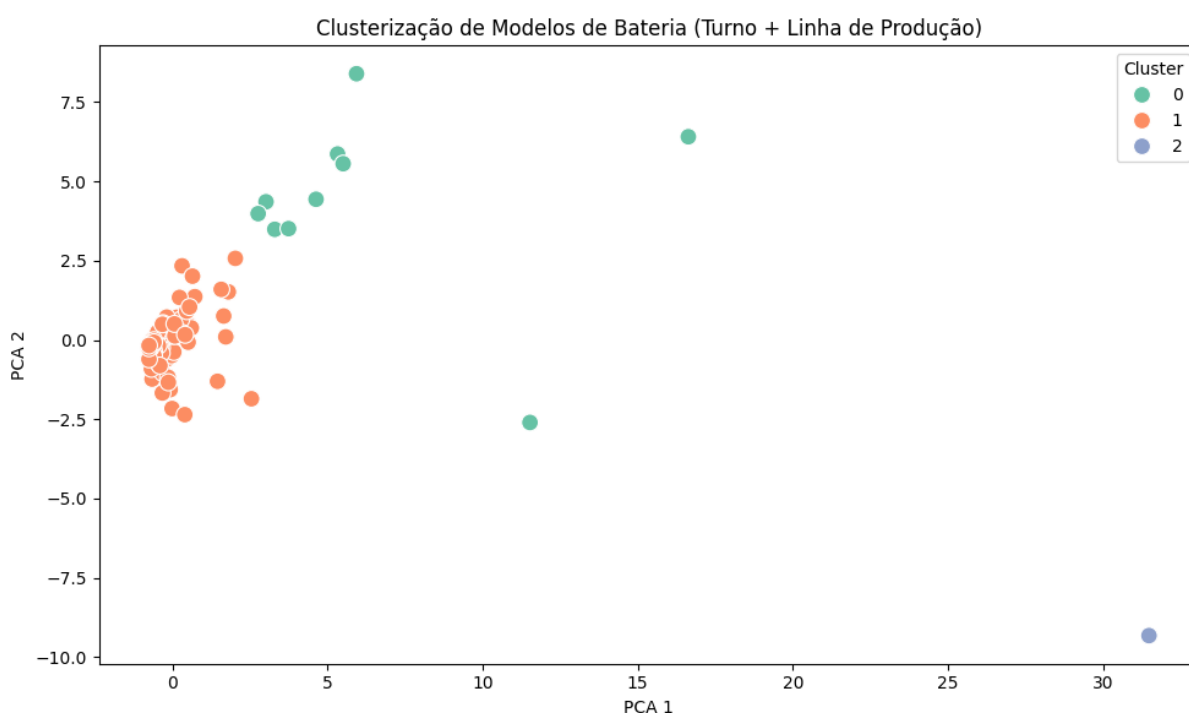
camente nulos para todos os códigos, o que pode indicar que este turno seja reservado para atividades de manutenção, setup ou que simplesmente não haja produção regular nesse horário.

De maneira geral, o heatmap confirma a concentração de falhas em alguns códigos específicos e destaca o turno B como o mais crítico em termos de volume de ocorrências. Esses resultados reforçam a necessidade de aprofundar a investigação sobre as causas raízes associadas às falhas nesses códigos e turnos, visando a implementação de ações corretivas direcionadas e eficazes.

#### 4.13 CLUSTERIZAÇÃO DE MODELOS DE BATERIA (TURNIO E LINHA DE PRODUÇÃO)

A Figura 11 apresenta o resultado da análise de clusterização dos modelos de bateria, considerando como variáveis o turno de produção e a linha de produção da unidade UN01. Para a visualização, foi aplicada a técnica de redução de dimensionalidade via Análise de Componentes Principais (PCA), permitindo representar a distribuição dos dados em duas dimensões (PCA 1 e PCA 2) sem perda significativa de informações.

Figura 11 – Clusterização dos modelos de bateria (turno + linha de produção)



Fonte: Esta pesquisa (2025).

Observa-se a formação de três clusters distintos. O Cluster 0 (representado em verde) agrupa modelos de bateria que apresentam maior dispersão em relação aos componentes principais, indicando características de produção mais heterogêneas ou variações significativas nos dados relacionados ao turno e à linha de produção. Esse grupo inclui modelos que, possivelmente, passam por diferentes condições operacionais ou regimes de produção.

O Cluster 1 (representado em laranja) reúne a maioria dos modelos de bateria analisados, formando um agrupamento denso e concentrado próximo à origem dos eixos de PCA. Essa configuração sugere forte similaridade entre os modelos dentro deste cluster, refletindo padrões de produção mais estáveis e homogêneos entre os turnos e linhas avaliados. Tal comportamento pode estar associado a linhas de produção mais consolidadas ou turnos com rotinas de fabricação padronizadas.

O Cluster 2 (representado em azul) é composto por poucos elementos, os quais se encontram fortemente afastados dos demais clusters. A distância significativa em relação aos outros grupos indica a presença de características muito distintas nos modelos aqui incluídos. Esses casos podem representar exceções no processo produtivo, como modelos de produção especial, processos experimentais, ou situações específicas de baixa demanda ou alta variabilidade.

De modo geral, a clusterização evidencia a existência de padrões de produção que podem ser úteis para a gestão da qualidade e para o planejamento de ações de melhoria contínua. A identificação de modelos pertencentes a clusters com maior dispersão ou diferenciação pode servir como ponto de partida para investigações específicas sobre práticas de produção, desempenho por turno e possíveis fontes de inconsistência no processo.

#### 4.14 SUBSÍDIOS ANALÍTICOS PARA A DEFINIÇÃO DE ESTRATÉGIAS DE MELHORIA

A identificação sistemática dos produtos, turnos e linhas de produção com maior incidência de falhas operacionais fornece uma base sólida para o direcionamento de estratégias de melhoria na unidade de montagem de baterias. A partir da análise exploratória dos dados históricos, foi possível mapear os principais pontos críticos do processo produtivo, possibilitando uma abordagem mais focada e eficiente para a implementação de ações corretivas e preventivas. Os dados analisados revelaram padrões recorrentes de falhas em determinados modelos de baterias, períodos produtivos específicos e linhas de montagem, permitindo a priorização de intervenções conforme o impacto sobre a eficiência operacional, os custos e a satisfação dos clientes.

Com base nesses resultados, torna-se pertinente a revisão dos procedimentos operacionais nas linhas de produção com maior concentração de falhas, com o objetivo de padronizar atividades e eliminar variabilidades que comprometem a qualidade do produto final. A capacitação direcionada das equipes de trabalho dos turnos mais afetados surge como outra estratégia necessária, voltada à disseminação de boas práticas de montagem e controle de qualidade, promovendo o desenvolvimento contínuo dos operadores e a redução das não conformidades.

Paralelamente, o aperfeiçoamento dos processos de inspeção e testes nas etapas críticas de fabricação deve ser considerado, de modo a garantir a identificação precoce de desvios e evitar a propagação de falhas ao longo da linha de produção. A intensificação da manutenção preditiva e corretiva nos equipamentos das linhas mais suscetíveis a falhas, com base nos dados de recorrência de problemas, também se apresenta como uma medida essencial para a estabilidade

operacional.

Ainda nesse contexto, destaca-se a importância da adoção de projetos de melhoria contínua, especificamente voltados aos modelos de baterias com maior índice de falhas, buscando a eliminação sistemática das causas-raiz dos problemas. A implementação de painéis de monitoramento em tempo real para acompanhamento dos principais indicadores de falhas por turno e por linha de produção pode contribuir significativamente para a agilidade na detecção e resposta a desvios.

Essas iniciativas, fundamentadas em uma análise exploratória robusta e orientada por dados, estão alinhadas às melhores práticas de gestão da qualidade e da produção. Como argumentam Deming (1997) e Ishikawa (1985), a melhoria contínua depende da compreensão profunda das causas de variação dos processos e da atuação sistemática para sua eliminação. Assim, ao aplicar os subsídios analíticos extraídos dos dados operacionais, a unidade de montagem de baterias poderá não apenas corrigir as falhas existentes, mas também fortalecer sua resiliência diante de novos desafios, promovendo um ambiente industrial mais eficiente, ágil e competitivo.

## 5 CONSIDERAÇÕES FINAIS

O presente estudo teve como objetivo analisar os dados de falhas operacionais de uma unidade de montagem de baterias com Python para identificar produtos, turnos e linhas de produção mais impactados, com o intuito de subsidiar ações de melhoria da eficiência industrial. Através da organização, visualização e interpretação dos dados históricos extraídos do sistema ERP da empresa, foi possível identificar, categorizar e priorizar os modelos de baterias, turnos e linhas de produção mais impactados por falhas.

Os resultados obtidos evidenciaram padrões relevantes que permitem compreender melhor a distribuição das falhas operacionais ao longo do processo produtivo. A clusterização aplicada aos dados possibilitou segmentar os produtos e áreas de produção mais críticos, fornecendo subsídios analíticos para a definição de estratégias de melhoria focadas. A análise revelou, ainda, que certos modelos de baterias e determinados turnos apresentaram índices de falhas significativamente superiores, sugerindo a necessidade de intervenções específicas, como revisões nos procedimentos operacionais, treinamentos direcionados para as equipes e reforço nos processos de inspeção e manutenção.

A análise do trabalho no geral, trouxe embasamento para direcionar a organização, em relação a que linha, modelo e turno focar seus trabalhos para obter resultados imediatos e significativos.

Além disso, o estudo reafirma a importância da cultura organizacional orientada por dados, evidenciando que a utilização de ferramentas de EDA, com o apoio da linguagem de programação Python e suas bibliotecas de análise, potencializa a capacidade da organização de detectar problemas, tomar decisões embasadas e promover melhorias contínuas nos processos industriais.

Entretanto, é importante reconhecer as limitações desta pesquisa. A análise foi realizada com base em dados históricos extraídos de um único período de tempo e de uma única unidade produtiva, o que restringe a abrangência dos resultados. Eventuais inconsistências nos registros do sistema ERP, como erros de categorização ou falhas de preenchimento, também podem ter impactado a qualidade dos dados analisados. Além disso, a pesquisa concentrou-se na análise descritiva dos dados, sem avançar para métodos preditivos ou de análise causal, que poderiam ampliar a profundidade dos insights gerados.

Diante dessas limitações, algumas sugestões para investigações futuras se tornam pertinentes. Estudos subsequentes podem incluir a aplicação de técnicas de modelagem preditiva, como algoritmos de machine learning, para antecipar a ocorrência de falhas e recomendar ações preventivas de forma automatizada. Recomenda-se também a ampliação do escopo temporal da análise, incluindo dados de períodos mais longos, para avaliar tendências sazonais ou evolutivas nos padrões de falhas. Outra possibilidade é a realização de estudos comparativos entre diferentes unidades produtivas da mesma empresa, permitindo uma análise mais ampla sobre a

eficácia dos processos em contextos variados.

Conclui-se, portanto, que a Análise Exploratória de Dados aplicada aos registros de falhas operacionais se mostrou uma abordagem eficaz para apoiar a gestão da qualidade e a melhoria da eficiência produtiva na indústria objeto deste estudo. A utilização de Python como ferramenta analítica reforçou a agilidade, a precisão e a profundidade da análise, contribuindo para a consolidação de práticas de gestão baseadas em evidências e alinhadas aos princípios da Indústria 4.0.

Destarte, com os resultados demonstrados, juntamente com as áreas de Produção e Qualidade, direcionaram os trabalhos de melhoria para uma das linhas modelo da produção a fim de reduzir o impacto das referidas falhas apontadas em objeto de estudo, explicitando-a com dados científicos.

## REFERÊNCIAS

- ANTUNES, J. et al. **Sistemas de produção: conceitos e práticas para projeto e gestão da produção enxuta**. Porto Alegre: Bookman, 2008.
- CAMPOS, V. F. **TQC – Controle da Qualidade Total**. Nova Lima: INDG, 2004.
- CERVO, A. L.; BERVIAN, P. A.; SILVA, R. d. **Metodologia científica**. São Paulo, SP: Pearson Prentice Hall, 2007.
- CHASE, R. B.; JACOBS, F. R.; AQUILANO, N. J. **Administração da produção para a vantagem competitiva**. Porto Alegre: [s.n.], 2006.
- CROSBY, P. B. **Quality Is Free: The Art of Making Quality Certain**. New York: McGraw-Hill, 1979.
- DAVENPORT, T. H.; HARRIS, J. G. **Competing on Analytics: The New Science of Winning**. Boston: Harvard Business School Press, 2007.
- DEMING, W. E. **A nova economia para a indústria, o governo e a educação**. Rio de Janeiro: Qualitymark, 1997.
- FEIGENBAUM, A. V. **Total Quality Control**. New York: [s.n.], 1991. v. 12.
- GARVIN, D. A. Competing on the eight dimensions of quality. 1987.
- ISHIKAWA, K. **What is Total Quality Control? The Japanese Way**. Englewood Cliffs: Prentice-Hall, 1985.
- IVANOV, D. et al. The impact of digital technologies on supply chains in industry 4.0. **International Journal of Production Research**, Taylor & Francis, v. 57, n. 6, p. 829–846, 2019.
- JURAN, J. M.; GODFREY, A. B. **Juran's Quality Handbook**. 5. ed. New York: McGraw-Hill, 1999.
- KAUARK FABIANE S.; MANHÃES, F. C. M. C. H. **Metodologia da pesquisa: um guia prático**. [S.l.: s.n.], 2010.
- LAKATOS, E. M.; MARCONI, M. d. A. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.
- LEE, J.; DAVARI, H.; KAO, H.-A. Recent advances and trends in predictive manufacturing systems in big data environment. **Manufacturing Letters**, Elsevier, v. 1, n. 1, p. 38–41, 2018.
- MARR, B. **Big Data in Practice: How 45 successful companies used big data analytics to deliver extraordinary results**. São Paulo: Alta Books, 2016.
- MIGUEL, P. A. C. et al. **Metodologia de Pesquisa em Engenharia de Produção e Gestão de Operações**. 2. ed. Rio de Janeiro: Elsevier Editora Ltda, 2010.
- MONTGOMERY, D. C. **Introduction to Statistical Quality Control**. [S.l.]: John Wiley & Sons, 2020.

PROVOST, F.; FAWCETT, T. Data science and its relationship to big data and data-driven decision making. **Big Data**, v. 1, n. 1, p. 51–59, 2013.

PROVOST, F.; FAWCETT, T. **Data Science for Business: What You Need to Know about Data Mining and Data-Analytic Thinking**. Sebastopol: O'Reilly Media, 2013.

SLACK, N.; CHAMBERS, S.; JOHNSTON, R. **Administração da Produção**. São Paulo: Atlas, 2009.

TUKEY, J. W. **Exploratory Data Analysis**. Reading, MA: Addison-Wesley, 1977.

VANDERPLAS, J. **Python Data Science Handbook: Essential Tools for Working with Data**. Sebastopol: O'Reilly Media, 2016.

VERGARA, S. C. **Projetos e relatórios de pesquisa em administração**. 9. ed. São Paulo: Atlas, 2007.

WANG, Q. et al. Battery manufacturing for electric vehicles: Advances and challenges. **Energy Storage Materials**, v. 17, p. 238–258, 2018.

WANG, Y.; WANG, Y. Applications of artificial intelligence in smart manufacturing: A review. **Engineering Applications of Artificial Intelligence**, v. 95, p. 103–135, 2020.

WOMACK, J. P.; JONES, D. T. Beyond toyota: how to root out waste and pursue perfection. **Harvard Business Review**, v. 74, p. 140–172, 1996.

YIN, R. K. **Estudo de Caso: Planejamento e métodos**. [S.l.]: Bookman Editora, 2015.

## APÊNDICE A – ALGORITMO

#Aplicação: Dados Baterias Moura

##Análise

por \*\*Beatriz Rocha\*\*

\*\*Engenharia de Produção / UFPE-CAA\*\*

### Importar pacotes utilizados

import numpy as np

import pandas as pd

import matplotlib.pyplot as plt

import seaborn as sns

import warnings

from sklearn.preprocessing import StandardScaler

from sklearn.cluster import KMeans

from sklearn.decomposition import PCA

warnings.filterwarnings('ignore')

pip install openpyxl

from google.colab import drive

drive.mount('/content/drive')

### 1. Coleta e preparação dos dados

Os dados trazem características sobre falhas em Baterias

# Caminho do arquivo

caminho\_arquivo = '/content/drive/MyDrive/Scrap Montagem 2024.xlsx'

# Leitura do arquivo Excel

df = pd.read\_excel(caminho\_arquivo)

# Visualizar as primeiras linhas do DataFrame

print(df.head())

df.head()

df\_un01 = df[df['Fábrica'] == 'UN01']

print(df\_un01.head())

## 2. Limpeza e tratamento dos dados

Toda análise será conduzida na Unidade de Montagem de Baterias (UN01)

# Limpeza de dados

```
df_un01_clean = df_un01.copy()
df_un01_clean = df_un01_clean[df_un01_clean['Modelo Bateria'].notna()]
df_un01_clean = df_un01_clean[df_un01_clean['Turma'].notna()]
df_un01_clean = df_un01_clean[df_un01_clean['Linha'].notna()]
df_un01_clean = df_un01_clean[df_un01_clean['Quantidade'].notna()]
df_un01_clean = df_un01_clean[df_un01_clean['Descrição Falha'].notna()]
# Converter coluna "Criado" para datetime
df_un01_clean['Criado'] = pd.to_datetime(df_un01_clean['Criado'], errors='coerce')
```

## 3. Análise exploratória (EDA)

# Criar um dicionário de mapeamento: nome original → nome genérico

```
modelos_unicos = df_un01_clean['Modelo Bateria'].dropna().unique()
map_modelos = {nome: f"Modelo_{str(i+1).zfill(3)}" for i, nome in enumerate(sorted(modelos_unicos))
}
```

# Aplicar a substituição

```
df_un01_clean['Modelo Bateria Anonimizado'] = df_un01_clean['Modelo Bateria'].map(map_modelos)
print(df_un01_clean.head())
```

# Verifica se a anonimização está correta linha a linha

```
verificacao = df_un01_clean['Modelo Bateria Anonimizado'] == df_un01_clean['Modelo Ba-
teria'].map(map_modelos)
```

# Verifica se todas as linhas são True (corretas)

```
print(verificacao.all())
```

# 1. Produtos com mais falhas

```
modelo_counts = df_un01_clean.groupby('Modelo Bateria Anonimizado')['Quantidade'].sum().
sort_values(ascending=False)
plt.figure(figsize=(10, 6)) modelo_counts.head(10).plot(kind='bar')
plt.title('Top 10 Modelos de Bateria com Mais Falhas (UN01)')
plt.ylabel('Quantidade de Falhas')
plt.xlabel('Modelo Bateria') plt.xticks(rotation=45) plt.tight_layout() plt.show()
```

# 2. Turnos mais críticos turno\_counts =

```
df_un01_clean.groupby('Turma')['Quantidade'].sum().sort_values(ascending=False)
```

```
plt.figure(figsize=(6, 4))
turno_counts.plot(kind='bar', color='gray')
plt.title('Falhas por Turno (UN01)')
plt.ylabel('Quantidade de Falhas')
plt.xlabel('Turno')
plt.xticks(rotation=0)
plt.tight_layout()
plt.show()
```

# 3. Linhas com maior tempo parado

```
linha_counts = df_un01_clean.groupby('Linha')['Quantidade'].sum().sort_values(ascending=False)
plt.figure(figsize=(10, 6))
linha_counts.plot(kind='bar', color='orange')
plt.title('Falhas por Linha de Produção (UN01)')
plt.ylabel('Quantidade de Falhas')
plt.xlabel('Linha') plt.xticks(rotation=0) plt.tight_layout() plt.show()
```

#4. Heatmap: Falhas por Linha e Turno

```
heatmap_data = df_un01_clean.pivot_table(
    index='Linha', columns='Turno', values='Quantidade', aggfunc='sum', fill_value=0)
plt.figure(figsize=(8, 6))
sns.heatmap(heatmap_data, annot=True, fmt='.0f', cmap='YlOrRd')
plt.title('Heatmap de Falhas por Linha e Turno (UN01)')
plt.xlabel('Turno')
plt.ylabel('Linha')
plt.tight_layout()
plt.show()
```

#5. Boxplot: Distribuição das quantidades de falhas por modelo

```
plt.figure(figsize=(12, 6))
top_modelos = modelo_counts.head(10).index
sns.boxplot(data=df_un01_clean[df_un01_clean['Modelo Bateria Anonimizado'].isin(top_modelos)],
    x='Modelo Bateria Anonimizado', y='Quantidade')
plt.title('Distribuição das Quantidades de Falhas - Top 10 Modelos')
plt.xticks(rotation=45)
plt.tight_layout()
plt.show()
falhas_modelo_turno = df_un01_clean.pivot_table(
    index='Modelo Bateria Anonimizado',
```

```

columns='Turma',
values='Quantidade',
aggfunc='sum',
fill_value=0
)

# Selecionar os 10 modelos com mais falhas totais (soma entre turnos)
top_modelos_falhas = falhas_modelo_turno.sum(axis=1).sort_values(ascending=False).head(10).index

# Filtrar a tabela para esses modelos
falhas_top = falhas_modelo_turno.loc[top_modelos_falhas]

# Criar o gráfico de barras empilhadas falhas_top.plot(kind='bar', stacked=True, figsize=(12,
6))

# Configurações do gráfico
plt.title('Top 10 Modelos de Bateria com Mais Falhas por Turno (UN01)')
plt.xlabel('Modelo de Bateria')
plt.ylabel('Quantidade de Falhas')
plt.xticks(rotation=45)
plt.legend(title='Turno')
plt.tight_layout()
plt.show()
falhas_linha_turno = df_un01_clean.pivot_table(
index='Linha',
columns='Turma',
values='Quantidade',
aggfunc='sum',
fill_value=0
)

# Selecionar as 10 linhas com mais falhas totais
top_linhas_falhas = falhas_linha_turno.sum(axis=1).sort_values(ascending=False).head(10).index
falhas_top_linhas = falhas_linha_turno.loc[top_linhas_falhas]

# Criar o gráfico de barras empilhadas falhas_top_linhas.plot(kind='bar', stacked=True, fig-
size=(12, 6))
plt.title('Top 10 Linhas de Produção com Mais Falhas por Turno (UN01)')
plt.xlabel('Linha de Produção')

```

```
plt.ylabel('Quantidade de Falhas')
plt.xticks(rotation=0)
plt.legend(title='Turno')
plt.tight_layout()
plt.show()
```

```
## 4. Curva ABC para Modelos de Bateria
```

```
# Agrupar os dados somando a quantidade de falhas por modelo abc_modelo = df_un01_clean.groupby
('Modelo Bateria
Anonimizado')['Quantidade'].sum().sort_values(ascending=False).reset_index()
```

```
# Renomear colunas
```

```
abc_modelo.columns = ['Modelo Bateria Anonimizado', 'Falhas']
```

```
# Calcular o percentual acumulado
```

```
abc_modelo['% Acumulado'] = 100 * abc_modelo['Falhas'].cumsum() / abc_modelo['Falhas'].sum()
```

```
# Função de classificação
```

```
ABC def classificar_abc(percentual):
```

```
if percentual <= 80:
```

```
return 'A'
```

```
elif percentual <= 95:
```

```
return 'B'
```

```
else:
```

```
return 'C'
```

```
# Aplicar a classificação
```

```
abc_modelo['Classe ABC'] = abc_modelo['%Acumulado'].apply(classificar_abc)
```

```
# Exibir as 10 primeiras linhas print(abc_modelo.head(17))
```

```
# Garantir que a tabela ABC está ordenada
```

```
abc_modelo_sorted = abc_modelo.sort_values(by='Falhas', ascending=False)
```

```
# Criar figura e eixos
```

```
fig, ax1 = plt.subplots(figsize=(12, 6))
```

```
# Barras: número de falhas por modelo
```

```
ax1.bar(abc_modelo_sorted['Modelo Bateria Anonimizado'], abc_modelo_sorted['Falhas'], co-
lor='skyblue')
```

```

ax1.set_xlabel('Modelo Bateria')
ax1.set_ylabel('Número de Falhas', color='blue')
ax1.tick_params(axis='y', labelcolor='blue')
ax1.set_xticklabels(abc_modelo_sorted['Modelo Bateria Anonimizado'], rotation=45) # (warning por causa do locator)
# Linha: percentual acumulado
ax2 = ax1.twinx()
ax2.plot(abc_modelo_sorted['% Acumulado'], color='red', marker='o') ax2.set_ylabel('% Acumulado', color='red')
ax2.tick_params(axis='y', labelcolor='red')

```

```

# Linhas horizontais para corte das classes ABC
ax2.axhline(80, color='gray', linestyle='--', linewidth=1)
ax2.axhline(95, color='gray', linestyle='--', linewidth=1)

```

```

# Título e ajustes
plt.title('Gráfico de Pareto - Falhas por Modelo de Bateria')
plt.tight_layout()
plt.show()

```

As barras azuis representam o número de falhas por modelo (ordenado do mais crítico ao menos).

A linha vermelha mostra o percentual acumulado, ajudando a visualizar os limites das classes A (até 80%) e B (até 95%).

### Gráfico de Pareto – Falhas por Turno

# Calcular total de falhas por turno

```

abc_turno = df_un01_clean.groupby('Turma')['Quantidade'].sum().sort_values(ascending=False).reset_index()

```

```

abc_turno.columns = ['Turno', 'Falhas']

```

# Calcular percentual acumulado

```

abc_turno['% Acumulado'] = 100 * abc_turno['Falhas'].cumsum() / abc_turno['Falhas'].sum()

```

# Classificação ABC

```

def classificar_abc(percentual): if percentual <= 80:

```

```

    return 'A'

```

```

    elif percentual <= 95: return 'B'

```

```

    else:

```

```
return 'C'
```

```
abc_turno['Classe ABC'] = abc_turno['%Acumulado'].apply(classificar_abc)
```

```
# Ordenar a tabela ABC de turnos
```

```
abc_turno_sorted = abc_turno.sort_values(by='Falhas', ascending=False)
```

```
# Gráfico de barras e linha acumulada
```

```
fig, ax1 = plt.subplots(figsize=(8, 5))
```

```
ax1.bar(abc_turno_sorted['Turno'], abc_turno_sorted['Falhas'], color='lightgreen')
```

```
ax1.set_xlabel('Turno')
```

```
ax1.set_ylabel('Número de Falhas', color='green')
```

```
ax1.tick_params(axis='y', labelcolor='green')
```

```
# Linha do % acumulado
```

```
ax2 = ax1.twinx()
```

```
ax2.plot(abc_turno_sorted['% Acumulado'], color='darkgreen', marker='o')
```

```
ax2.set_ylabel('% Acumulado', color='darkgreen')
```

```
ax2.tick_params(axis='y', labelcolor='darkgreen')
```

```
ax2.axhline(80, color='gray', linestyle='--')
```

```
ax2.axhline(95, color='gray', linestyle='--')
```

```
plt.title('Gráfico de Pareto - Falhas por Turno')
```

```
plt.tight_layout()
```

```
plt.show()
```

```
### Gráfico de Pareto – Falhas por Linha de Produção
```

```
# Calcular total de falhas por linha
```

```
abc_linha = df_un01_clean.groupby('Linha')['Quantidade'].sum().sort_values(ascending=False).reset_index()
```

```
abc_linha.columns = ['Linha', 'Falhas']
```

```
# Calcular percentual acumulado
```

```
abc_linha['% Acumulado'] = 100 * abc_linha['Falhas'].cumsum() / abc_linha['Falhas'].sum()
```

```
# Classificação ABC
```

```
def classificar_abc(percentual):
```

```
if percentual <= 80:
```

```
return 'A'
```

```
elif percentual <= 95: return 'B'
```

```
else:
```

```
return 'C'
```

```
abc_linha['Classe ABC'] = abc_linha['% Acumulado'].apply(classificar_abc)
```

```
# Ordenar a tabela ABC de linhas
```

```
abc_linha_sorted = abc_linha.sort_values(by='Falhas', ascending=False)
```

```
fig, ax1 = plt.subplots(figsize=(12, 6))
```

```
ax1.bar(abc_linha_sorted['Linha'].astype(str), abc_linha_sorted['Falhas'], color='lightcoral')
```

```
ax1.set_xlabel('Linha de Produção')
```

```
ax1.set_ylabel('Número de Falhas', color='darkred') ax1.tick_params(axis='y', labelcolor='darkred')
```

```
ax1.set_xticks(range(len(abc_linha_sorted)))
```

```
ax1.set_xticklabels(abc_linha_sorted['Linha'].astype(str), rotation=45)
```

```
# Linha do % acumulado
```

```
ax2 = ax1.twinx()
```

```
ax2.plot(abc_linha_sorted['% Acumulado'], color='red', marker='o')
```

```
ax2.set_ylabel('% Acumulado', color='red')
```

```
ax2.tick_params(axis='y', labelcolor='red')
```

```
ax2.axhline(80, color='gray', linestyle='--')
```

```
ax2.axhline(95, color='gray', linestyle='--')
```

```
plt.title('Gráfico de Pareto - Falhas por Linha de Produção')
```

```
plt.tight_layout()
```

```
plt.show()
```

```
#### Série Temporal de Falhas – Fábrica UN01
```

```
# Garantir que a coluna "Criado" está no formato datetime
```

```
df_un01_clean['Criado'] =
```

```
pd.to_datetime(df_un01_clean['Criado'], errors='coerce')
```

```
# Criar uma nova coluna com o formato Ano-Mês df_un01_clean['AnoMes'] = df_un01_clean['Criado']  
.dt.to_period('M')
```

```
# Agrupar por Ano-Mês e somar a quantidade de falhas
```

```
serie_temporal = df_un01_clean.groupby('AnoMes')['Quantidade'].sum().reset_index()
```

```
# Converter de Period para string para uso no eixo x
```

```
serie_temporal['AnoMes'] = serie_temporal['AnoMes'].astype(str)
```

```
# Criar o gráfico de linha (série temporal)
```

```

plt.figure(figsize=(14, 6))
plt.plot(serie_temporal['AnoMes'], serie_temporal['Quantidade'], marker='o', linestyle='-')
plt.title('Evolução das Falhas ao Longo do Tempo (UN01)')
plt.xlabel('Ano-Mês')
plt.ylabel('Quantidade de Falhas')
plt.xticks(rotation=45)
plt.grid(True)
plt.tight_layout()
plt.show()

### Série Temporal com Média Móvel (3 meses)
# (Re)garantir que 'Criado' está em datetime e criar coluna AnoMes
df_un01_clean['Criado'] = pd.to_datetime(df_un01_clean['Criado'], errors='coerce')
df_un01_clean['AnoMes'] = df_un01_clean['Criado'].dt.to_period('M')

# Agrupar falhas por Ano-Mês
serie_temporal = df_un01_clean.groupby('AnoMes')['Quantidade'].sum().reset_index()
serie_temporal['AnoMes'] = serie_temporal['AnoMes'].astype(str)

# Calcular média móvel com janela de 3 meses
serie_temporal['Média Móvel (3 meses)'] = serie_temporal['Quantidade'].rolling(window=3).mean()

# Plotar o gráfico
plt.figure(figsize=(14, 6))
plt.plot(serie_temporal['AnoMes'], serie_temporal['Quantidade'], marker='o', label='Falhas Mensais')
plt.plot(serie_temporal['AnoMes'], serie_temporal['Média Móvel (3 meses)'], color='red', linestyle='--', label='Média Móvel (3 meses)')
plt.title('Evolução das Falhas com Média Móvel (UN01)')
plt.xlabel('Ano-Mês')
plt.ylabel('Quantidade de Falhas')
plt.xticks(rotation=45)
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()

### Análise Cruzada da Quantidade de Falhas por tipo de defeito
# Agrupar por tipo de defeito e somar a quantidade de falhas
falhas_unicas = df_un01_clean['Descrição Falha'].dropna().unique()
map_falhas = nome: f"Falha_str(i+1).zfill(3)" for i, nome in enumerate(sorted(falhas_unicas))

```

```

print(falhas_unicas)
print(map_falhas)

# Aplicar a substituição
df_un01_clean['Descrição Falha Anonimizado'] = df_un01['Descrição Falha'].map(map_falhas)

# Agrupar por tipo de defeito e somar a quantidade de falhas
falhas_por_defeito = df_un01_clean.groupby('Descrição Falha Anonimizado')['Quantidade'].sum()
.sort_values(ascending=False).reset_index()

# Renomear as colunas para melhor leitura
falhas_por_defeito.columns = ['Tipo de Defeito', 'Total de Falhas']

# Exibir as 10 primeiras linhas (opcional)
print(falhas_por_defeito.head(10))
Os grupos FALHA 013, 014 E 015 concentram a maior parte das falhas.
# Verifica se a anonimização está correta linha a linha
verificacao2 = df_un01_clean['Descrição Falha Anonimizado'] == df_un01_clean['Descrição
Falha'].map(map_falhas)

# Verifica se todas as linhas foram corretamente mapeadas
print(verificacao2.all())
df_un01_clean[verificacao2]

#### Heatmap - Quantidade De Falhas Por Tipo De Defeito E Turno (UN01)
# Criar a tabela cruzada (pivot table): Tipo de Defeito x Turno
heatmap_defeito_turno = df_un01_clean.pivot_table(
index='Descrição Falha Anonimizado',
columns='Turma',
values='Quantidade',
aggfunc='sum',
fill_value=0
)

# Plotar o heatmap
plt.figure(figsize=(12, 8))
sns.heatmap(heatmap_defeito_turno, annot=True, fmt='.0f', cmap='YlOrRd')
plt.title('Heatmap - Quantidade de Falhas por Tipo de Defeito e Turno (UN01)')
plt.xlabel('Turno')

```

```
plt.ylabel('Tipo de Defeito')
```

```
plt.tight_layout()
```

```
plt.show()
```

```
### Clusterização de Modelos de Bateria
```

```
# Criar base com falhas por modelo de bateria x turno x linha de produção
```

```
modelo_turno = df_un01_clean.pivot_table(
```

```
index='Modelo Bateria Anonimizado', columns='Turma', values='Quantidade', aggfunc='sum',  
fill_value=0
```

```
)
```

```
modelo_linha = df_un01_clean.pivot_table(
```

```
index='Modelo Bateria Anonimizado', columns='Linha', values='Quantidade', aggfunc='sum',  
fill_value=0
```

```
)
```

```
# Garantir que todos os nomes das colunas sejam string
```

```
modelo_turno.columns = modelo_turno.columns.astype(str)
```

```
modelo_linha.columns = modelo_linha.columns.astype(str)
```

```
# Concatenar turnos e linhas em uma única base
```

```
base_cluster = pd.concat([modelo_turno, modelo_linha], axis=1)
```

```
# Padronizar os dados
```

```
scaler = StandardScaler()
```

```
X_scaled = scaler.fit_transform(base_cluster)
```

```
# Aplicar K-Means com 3 clusters
```

```
kmeans = KMeans(n_clusters=3, random_state=42, n_init=10)
```

```
clusters = kmeans.fit_predict(X_scaled)
```

```
# Reduzir para 2D com PCA para visualização
```

```
pca = PCA(n_components=2)
```

```
X_pca = pca.fit_transform(X_scaled)
```

```
# Criar DataFrame para visualização
```

```
cluster_df = pd.DataFrame(
```

```
'Modelo Bateria Anonimizado': base_cluster.index,
```

```
'Cluster': clusters,
```

```

'PCA1': X_pca[:, 0],
'PCA2': X_pca[:, 1]
)

# Plotar os clusters
plt.figure(figsize=(10, 6))
sns.scatterplot(data=cluster_df, x='PCA1', y='PCA2', hue='Cluster', palette='Set2', s=100)
plt.title('Clusterização de Modelos de Bateria (Turno + Linha de Produção)')
plt.xlabel('PCA 1')
plt.ylabel('PCA 2')
plt.legend(title='Cluster')
plt.tight_layout()
plt.show()

### Tabela de Modelos por Cluster
# Criar base com falhas por modelo de bateria x turno x linha de produção
modelo_turno = df_un01_clean.pivot_table(
index='Modelo Bateria Anonimizado', columns='Turma', values='Quantidade', aggfunc='sum',
fill_value=0
)

modelo_linha = df_un01_clean.pivot_table(
index='Modelo Bateria Anonimizado', columns='Linha', values='Quantidade', aggfunc='sum',
fill_value=0
)

# Garantir que nomes de colunas sejam strings
modelo_turno.columns = modelo_turno.columns.astype(str)
modelo_linha.columns = modelo_linha.columns.astype(str)

# Combinar os dados
base_cluster = pd.concat([modelo_turno, modelo_linha], axis=1)

# Padronizar os dados
scaler = StandardScaler()
X_scaled = scaler.fit_transform(base_cluster)

# Aplicar K-Means com 3 clusters
kmeans = KMeans(n_clusters=3, random_state=42, n_init=10)

```

```

clusters = kmeans.fit_predict(X_scaled)

# Adicionar resultados de cluster ao DataFrame
cluster_df_completo = base_cluster.copy()
cluster_df_completo['Modelo Bateria Anonimizado'] = base_cluster.index
cluster_df_completo['Cluster'] = clusters

# Agrupar modelos por cluster
modelos_por_cluster = cluster_df_completo[['Modelo Bateria Anonimizado', 'Cluster']].groupby(
('Cluster')['Modelo Bateria Anonimizado']).apply(list).reset_index()

# Exibir resultado print
(modelos_por_cluster)

# Estatísticas descritivas agrupadas
por Modelo
falhas_stats = (
df_un01_clean.groupby("Modelo Bateria Anonimizado")["Quantidade"]
.agg([
("Qtd Registros", "count"), # nº de vezes que aparece
("Total Falhas", "sum"), # soma das quantidades
("Média por Registro", "mean"),
("Desvio Padrão", "std"),
("Mínimo", "min"),
("Máximo", "max")
])
.reset_index()
.sort_values(by="Total Falhas", ascending=False)
)

# Exibir tabela ordenada
print(falhas_stats.head(10)) # top 10 falhas

# Estatísticas descritivas agrupadas por falha
falhas_stats = (
df_un01.groupby("Descrição Falha")["Quantidade"]
.agg([
("Qtd Registros", "count"), # nº de vezes que aparece
("Total Falhas", "sum"), # soma das quantidades

```

```
("Média por Registro", "mean"),  
("Desvio Padrão", "std"),  
("Mínimo", "min"),  
("Máximo", "max")  
])  
.reset_index()  
.sort_values(by="Total Falhas", ascending=False)  
)
```

```
# Exibir tabela ordenada
```

```
print(falhas_stats.head(10)) # top 10 falhas
```