

Análise e diagnóstico das interrupções produtivas em uma Indústria Automotiva

Analysis and diagnosis of production interruptions in an Automotive Industry

Aíla Lima ^{1,2}  orcid.org/0000-0002-5343-4852

Joel Galdino ^{1,2}  orcid.org/0000-0002-0600-5967

Maria Gusmão ^{1,2}  orcid.org/0000-0002-0355-0978

Carmelo Bastos-Filho  orcid.org/0000-0002-0924-5341

¹ Escola Politécnica de Pernambuco, Universidade de Pernambuco, Recife, Brasil,

² Especialização em Ciência de Dados e Analytics, Escola Politécnica de Pernambuco, Pernambuco, Brasil,

E-mail: alm@ecomp.poli.br; galdino.joel@gmail.com; mlabg@ecomp.poli.br; carmelo.filho@upe.br

Resumo

Reduzir custos e aumentar a produtividade são objetivos comuns entre várias organizações e na indústria automotiva se tornou condição de sobrevivência. Para reduzir o desperdício de determinado recurso, primeiramente devemos identificar onde ele ocorre. O principal objetivo deste trabalho é identificar e dar visibilidade aos fatores relacionados ao tempo perdido com paradas produtivas na linha de produção de uma indústria automotiva. Para isto utilizamos ferramentas de mineração de dados e nos baseamos no CRISP DM. A atividade foi focada no diagnóstico e no levantamento dos perfis das paradas. Com a análise exploratória dos dados conseguimos identificar a alta ocorrência e a curta duração das interrupções na linha de produção. Além disso, ao tratar e transformar as justificativas automáticas do sistema de manufatura para um formato mais agrupado e fácil de se entender nos trouxe uma visão mais clara e permitiu que diversas ações focadas fossem realizadas pelos envolvidos. Em somente duas iniciativas projetou-se uma economia de mais de R\$: 75 mil reais anuais, e com perspectiva de ampliação para outras áreas da montagem. Para permitir que estas análises ocorram de forma mais tempestiva, desenvolvemos uma aplicação web para ser utilizada por supervisores e líderes de equipes na linha de produção.

Palavras-Chave: CRISP DM; Mineração de dados; Indústria Automotiva; Análise exploratória;

Abstract

Reducing costs and increasing productivity are common goals among many organizations, and in the automotive industry, it has become a condition for survival. To reduce the waste of a given resource, we must first identify where it occurs. The main objective of this work is to identify and give visibility to factors related to lost time with production stops in the production line of an automotive industry. For this, we use data mining tools and are based on CRISP DM. The activity was focused on diagnosing and surveying the profiles of the stops. With the exploratory analysis of the data, we were able to identify the high occurrence and short duration of interruptions in the production line. Furthermore, by treating and transforming the automatic justifications of the manufacturing system to a more grouped and easy-to-understand format, it brought us a clearer vision and allowed several focused actions to be taken by the stakeholders. In just two initiatives, savings of more than R\$ were projected: 75 thousand reais a year, and with the prospect of expansion to other areas of assembly. To allow these analyzes to take place more timely, we have developed a web application to be used by supervisors and team leaders on the production line.

Key-words: CRISP DM; Data mining; Automotive industry; Exploratory analysis.

1. Introdução

1.1. Contextualização

A primeira linha de montagem do mundo teve seu início no século XX na fábrica da Ford em Highland Park, Estados Unidos, onde a produção passou por uma revolução, pois, a partir daquele momento, os veículos chegavam até os operários, otimizando o processo produtivo e aumentando a quantidade de carros produzidos. Com o passar do tempo, o conceito de linha de produção foi sofrendo alterações e se tornando cada vez mais eficiente, saindo da fase de artesanato e chegando na indústria 4.0.

1.2. Descrição do Problema

Devido ao crescimento da demanda de veículos automotivos ao longo do último século, houve uma necessidade de elevar a produção de carros, reduzindo o tempo de fabricação. Entretanto, com o aumento da produção há também o aumento das perdas decorrentes desses processos. Considerando o aumento das perdas produtivas, o nível de capital investido para a produção de veículos cresce, ficando expresso no aumento dos valores que chegam ao consumidor.

1.3. Objetivo

O presente estudo tem como objetivo mapear as perdas produtivas na linha de produção de uma indústria automotiva, onde foram utilizados ferramentas e conhecimentos em ciência de dados para produção de um algoritmo para análise das paradas de linha.

1.4. Justificativa

As paradas de linhas impactam fortemente nos resultados, respondendo, conjuntamente, por, aproximadamente, 24% das perdas produtivas na Indústria Automotiva estudada.

Além de causarem diretamente perdas financeiras dentro do processo produtivo, as paradas de linha acarretam, também, a redução da produtividade e

podem, ainda, afetar negativamente a experiência do usuário final.

Esta realidade é comum a outras manufaturas, em que, de forma geral, atividades de montagem nas linhas de produção contribuem significativamente para o custo e a qualidade do produto final [1], compreendendo, em média, 40% do custo do produto e até 50% do custo total de fabricação [2, 3].

Neste contexto, é crucial identificar padrões e os principais fatores que influenciam na ocorrência de paradas de linha, e como é possível reduzi-los, gerando redução dos custos fabris, do retrabalho, aumento da qualidade e, conseqüentemente, da satisfação dos clientes.

2. Fundamentação Teórica

Visando explorar e extrair valor dos dados disponíveis há diversas abordagens e metodologias disponíveis. Dentre as quais podemos destacar: O Knowledge Discovery Databases (KDD), Sample, Explore, Modify, Model, Assess (SEMMA) e o Cross-Industry Standard Process for Data Mining (CRISP-DM) [4]. Devido a sua grande flexibilidade e independência em relação a ferramenta para sua aplicação, o CRISP DM foi escolhido como metodologia para apoiar a análise e mineração dos dados do nosso problema.

Este framework engloba o ciclo de vida básico de uma aplicação de Ciência de Dados e está apresentado na Figura 1. É dividido em 6 etapas [5]:

1. **Entendimento do Negócio:** Definição do problema, levantamento dos envolvidos, identificação das fontes de dados, levantamento dos riscos e planejamento.
2. **Entendimento dos Dados:** Integração das bases de dados, análise exploratória de dados e verificar a qualidade dos dados.
3. **Preparação dos dados:** Limpeza e transformação de dados, exploração adicional de dados, redução de dimensionalidade e Engenharia de atributos.
4. **Modelagem dos dados:** Seleção e desenho do modelo, construção do design de teste, otimização de hiperparâmetros e treinamento e validação dos modelos.

5. **Avaliação:** Teste do modelo e avaliação dos resultados.

6. **Produção:** Implantação do modelo em produção, plano de monitoramento e manutenção e relatórios finais e visualizações.

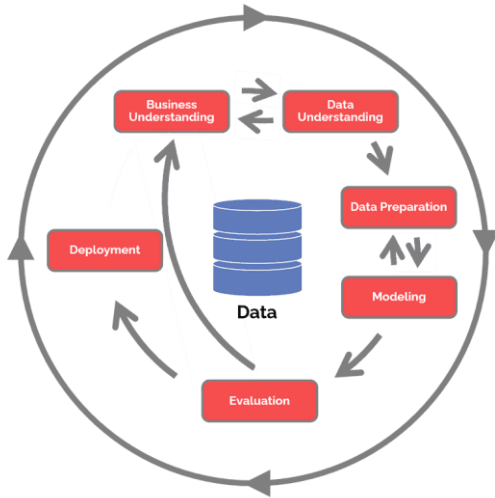


Figura 1: CRISP DM e suas 6 etapas no ciclo de vida de uma aplicação de ciência de dados. Fonte: <https://www.datascience-pm.com/crisp-dm-2/>

2.1. Área do negócio

A aquisição de vantagens competitivas deixou de ser um diferencial e hoje é condição de sobrevivência para as empresas. Principalmente em nichos tão competitivos quanto a indústria automobilística. Reduzir custos e aumentar a eficiência, este tem sido o mantra de todas as organizações. E é neste contexto que a inteligência artificial e a ciência de dados se tornaram ativos estratégicos. Vários subprocessos na cadeia produtiva automotiva já estão conectadas com essas tecnologias: Desenvolvimento, compras, logísticas, produção, marketing, vendas, pós-vendas e clientes conectados [6].

Em nosso trabalho abordamos a produção, mais especificamente a montagem. E neste contexto o principal objetivo da montagem é a entrega de veículos ao final da linha produtiva. Ação intimamente relacionada com a eficiência da linha de produção. Esta eficiência é calculada através do *Overall Production Efficiency* (OPE) que é a razão entre o número de veículos produzidos por hora na ponta da linha e a capacidade produtiva da planta.

Em linhas gerais quanto mais carros são produzidos em menos tempo maior é a OPE. Dessa forma, paradas ou interrupções produtivas reduzem este indicador e geram perdas financeiras para a planta. Permitir que os gestores e colaboradores na linha de produção possam atuar de forma mais tempestiva e assertiva nos fatores relacionados a estas interrupções é o principal objetivo deste trabalho.

2.2. Mineração de Dados

Mineração de dados pode ser entendido como o processo de exploração de grande massa de dados e posterior extração de informações úteis para o negócio [7]. Ela atua na busca por padrões e características que ajudem descrever ou explicar determinado fenômeno. Os seres humanos sempre se destacaram pela capacidade de encontrar padrões e avaliar comportamentos e tendências. No entanto, com o aumento exponencial da quantidade de dados disponíveis e na exigência de maior velocidade no processamento.

A avaliação manual na maioria das situações se torna ineficiente. A mineração de dados se vale de técnicas de aprendizado de máquina, inteligência artificial e estatística para transformar os dados em conhecimento. A mineração de dados pode ser dividida em três fases [8], observe a **Figura 2**.

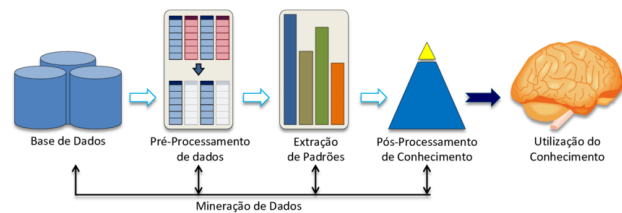


Figura 2: Fases do processo de mineração de dados. Fonte: [8]

A primeira fase conhecida como pré-processamento é responsável por adequar os dados brutos ao formato conveniente para a extração dos padrões. A depender da técnica escolhida e do estado atual dos dados, alguns ajustes são necessários. Isto é muito comum já que os dados podem ser oriundos das mais diversas fontes, seja o quantitativo de vendas de um produto ou a leitura de um manômetro de um vaso de pressão. Esta flexibilidade em relação

à natureza do negócio traz à tona outra função do pré-processamento que é o aprendizado sobre o domínio do negócio. Isto é fundamental para entender o que faz sentido e o significado dos dados utilizados.

Nesta fase são realizadas atividades como limpeza, preparação e transformação dos dados e atributos.

A segunda fase é responsável pela extração de padrões, e pode se subdividir outras etapas e aqui que a mineração de dados se vale de técnicas de aprendizado de máquina, de visualização e ferramentas estatísticas para automatizar e identificar retirada de conhecimento dos dados. É nesta etapa que pode ocorrer a definição dos hiperparâmetros de um modelo de aprendizado de máquina e a coleta das informações avaliadas por ele.

A terceira fase é chamada de pós-processamento do conhecimento. É nela que ocorre a avaliação e validação dos resultados. A mineração de dados deve ser um processo iterativo e interativo, isto fica bem evidenciado no ciclo do CRISP DM. Avaliar a qualidade das informações retiradas dos dados é fundamental para evitar um descolamento com a realidade do negócio. Muitas vezes se faz necessário várias rodadas de criação, avaliação e validação dos modelos para se chegar a resultados coerentes e confiáveis. *Outliers*, dados ruidosos, repetidos ou pobres semanticamente sempre podem estar presentes e devem ser tratados. Além disso, o conhecimento adquirido junto à mineração pode necessitar de uma adequação à linguagem do negócio. Processo de mineração que não é capaz de transmitir de forma clara e inteligível seus resultados, falha numa importante missão da mineração de dados que é apoiar as decisões estratégicas e de negócio.

2.3. Trabalhos Relacionados

A Indústria 4.0 é conhecida pelo frequente uso de sensores e outros dispositivos eletrônicos, conectados à rede, dentro dos processos produtivos a fim de aprimorar o produto entregue ao consumidor final. Atualmente já é possível coletar diversos dados sobre as máquinas e seu processo produtivo para proporcionar estudos que colaboram com a melhoria contínua. Muitas organizações ainda enfrentam dificuldades para diagnosticar as causas raízes dos defeitos, resultando em altos custos para retrabalhos e reparos [9].

Com isso, atingir a excelência operacional, manufaturando processos sem falhas, com o mínimo de desperdício de material, reduzindo constantemente os custos de produção é o principal propósito nas Indústrias atualmente [10].

Numerosos estudos têm sido realizados focando na implementação de mineração de dados em diferentes áreas da manufatura, como processos de produção, melhoria da qualidade, detecção de falhas, otimização do rendimento da fabricação, planejamento de requisitos de material e manutenção preditiva de máquinas [11, 12, 13].

Na detecção de falhas, a mineração de dados é usada para identificar padrões de defeitos, fatores que influenciam a falha de processos, tipos de defeitos e taxas de erro na fabricação [14, 15].

As quebras de máquinas no processo de produção são relatadas como um dos fatores mais importantes no diagnóstico de defeitos. Ciflikli e Kahya-zyirmidokuz em 2010 desenvolveram uma solução de mineração de dados para aumentar a produtividade da fabricação de carpetes que emprega análise de relevância de atributos, árvores de decisão e indução baseada em regras [16]. Os resultados indicaram que o isolamento das quebras da máquina no processo de produção e o modelo de árvore de decisão proposto produzem uma melhoria significativa de aproximadamente 73% na taxa de precisão.

Uma investigação mais aprofundada visa identificar problemas no procedimento que levam a defeitos. Os dados históricos das condições ambientais (umidade ou temperatura) e da máquina (pressão, tensão ou atual) no processo de produção são então avaliados para garantir que os padrões de qualidade foram encontrados. Um outro estudo realizado por Wang em 2013, focado em uma abordagem de mineração de dados, foi aplicada para um sistema zero de defeito de manufatura (ZDM) a fim de garantir a ausência de falhas nos produtos fabricados. A chave para alcançar o ZDM é se concentrar não apenas na qualidade dos produtos e nas condições do produto, mas também no estado do equipamento e na sua degradação [17].

Além disso, as matérias-primas de cada produto e quaisquer fatores relacionados às condições de trabalho, incluindo duração, turno e o nível de experiência dos trabalhadores também deve ser considerado na identificação de potenciais fontes dos

problemas que influenciam o sucesso ou fracasso do processo de produção [18, 19].

3. Materiais e Métodos

3.1. Descrição da base de dados

A nossa principal fonte de dados é o *Manufacturing Execution Systems* (MES). O sistema de execução de manufatura tem seu foco nas atividades de produção e faz a interface entre o planejamento (vendas) com o chão de fábrica. Devido à sua importância estratégica e por suas limitações no tráfego de dados, na realidade os dados foram retirados de um banco de dados espelho do MES. No entanto, iremos apenas nos referir como MES.

Na pesquisa por quais fatores impactam na parada produtiva na linha de produção buscamos diversas fontes que tivessem ligação com o fenômeno. Utilizamos os seguintes critérios para escolha dos dados:

- Disponibilidade
- Integridade
- Confiabilidade
- Importância para o problema.

Na **Figura 3** são apresentadas as fontes de dados utilizadas e as principais informações consideradas nas análises realizadas.



Figura 3: Fonte de dados e informações avaliadas. Fonte: Elaboração própria. Fonte: Autoria própria.

Após realizar integração e limpeza dos dados, condensamos todos os atributos numa única planilha no formato “.csv” e a partir dela realizamos a análise exploratória e a pesquisa por padrões. Na **Tabela 1** podemos observar o dicionário de dados construído. Devido aos critérios informados acima, o horizonte temporal de análise foi de janeiro de 2021 a abril de

2021. Onde tivemos um total de 150 mil amostras. Dividida entre veículos produzidos e paradas na linha.

Tabela 1: Dicionário de Dados utilizado na análise

Variável	Descrição	Tipo
ProductionOrder	Identificador do veículo	Texto
DifusionDate	Data de difusão do veículo	Timestamp
TransitDate_TRIM	Data de entrada na TRIM	Texto
Transit_Time_Time	Horário de entrada na TRIM	Texto
Turno_Transit	Turno de trânsito na TRIM	Texto
Ano_Transit	Ano de trânsito na TRIM	Texto
Mes_Transit	Mês de trânsito na TRIM	Texto
Dia_Semana_Transit	Dia da semana de trânsito na TRIM	Texto
Modelo	Modelo do veículo	Texto
Motor	Tipo de motorização do veículo	Número
Cat_Cambio	Tipo de câmbio do veículo	Texto
Traçao	Tração do veículo	Texto
CallStartTime	Data de início da parada de linha	Timestamp
StartTime_Time	Horário de início da parada de linha	Texto
Turno	Turno da parada de linha	Texto
Ano_StartTime	Ano da parada de linha	Texto
Dia_Semana_StartTime	Dia da semana da parada de linha	Texto
Mes_StartTime	Mês da parada de linha	Texto
CallEndTime	Data de finalização da parada de linha	Timestamp
Tempo_Parada(millseconds)	Duração da parada de linha	Texto
Domínio	Domínio da estação envolvida na parada	Texto
PDT	Estação de trabalho envolvida na parada	Texto
Workplace	Workplace envolvido na parada	Texto
Operation	Justificativa automática da parada	Texto
Taxa_abs	Taxa de absenteísmo do domínio	Float
Status_Producao	Status se o veículo parou ou não na linha	Texto

3.2. Análise descritiva dos dados

Os atributos relacionados na seção anterior podem ser organizados segundo os seis grupos presentes na **Figura 4**.



Figura 4: 6 Dimensões analisadas nos dados presentes. Fonte: Autoria própria.

Os grupos mês a mês, hora a hora, dia a dia e Turnos assumem a dimensão temporal. Se referem à distribuição das paradas produtivas ao longo do tempo. Visando encontrar padrões de ocorrência da parada. Inicialmente buscamos trabalhar com um período de tempo maior para avaliar possíveis efeitos sazonais tanto na produção quanto no número de paradas de linha ao longo dos meses. No entanto,

seguindo os critérios já mencionados, apenas 4 meses estavam aptos a serem utilizados.

Avaliamos o comportamento ao longo do dia, durante as 24h e os três turnos produtivos, turno A, Turno B e Turno C. Para os meses avaliados há presença produtiva em todos os dias da semana, de segunda a domingo.

Para a dimensão do produto temos presente algumas características produtivas dos veículos produzidos: modelo, câmbio, motorização e tração. O Modelo apresenta 3 opções, Modelo 1, Modelo 2 e Modelo 3.

O câmbio foi agrupado em três opções, AT6, AT9 e Manual. As duas primeiras são automáticas com seis e nove velocidades respectivamente e a última manual. A motorização apresenta 4 opções e a tração duas opções.

A última dimensão se refere à justificativa automática dada pelo MÊS à parada de linha. De forma crua ela se apresenta codificada e com muitas variações possíveis, mas foi trabalhada e traduzida em seis categorias. O que melhorou muito o entendimento sobre a parada e possibilitou encontrar padrões interessantes. A codificação foi substituída por operação manual, fixação, leitura, certificação, certificação de qualidade e diagnose elétrica.

3.3. Pré-processamento dos Dados

O pré-processamento de dados pode ser feito de diversas formas, a solução depende do conjunto de dados e do problema que se pretende resolver. Uma vez limpos, esses conjuntos podem ser gravados em uma nova estrutura de dados.

O pré-processamento dos dados consistiu em transformar os dados de entrada para um formato mais apropriado para análises posteriores. As etapas envolvidas no processamento incluíram a fusão de dados de múltiplas fontes e sua limpeza através do Pentaho Data Integration (PDI). A limpeza dos dados se deu a partir da seleção de variáveis, tratamento de dados faltantes e ajuste de dados duplicados. Inicialmente, foram excluídas colunas dos conjuntos de dados, essa estratégia foi utilizada devido ao fato de algumas variáveis não serem importantes para o resultado buscado.

Para o estudo foram compiladas, basicamente, 3 tabelas com informações de detalhes do veículo (motorização, câmbio, modelo etc.), detalhes da parada de linha (chassi, tempo de parada, turno, causas etc.) e detalhes de absenteísmo.

Considerando o tempo de parada, quantidade de parada e a sua causa como os principais *atributos* do estudo, elas foram agrupadas da seguinte forma:

- Grupos do tempo de parada: $\leq 1s$, 1s a 10s, 10s a 30s, 30s a 1min, 1min a 5min e $>5min$.
- Causa (Justificativa): Certificação, Certificação de Qualidade, Diagnose Elétrica, Fixação, Leitura e Operação Manual.

Escolhemos adicionar o tempo de parada com o intuito de visualizar padrões que poderiam ficar escondidos ao observar o tempo agregado. Outro motivo foi o de categorizar a variável tempo para permitir utilizar o qui quadrado ou K modes. Os valores inicialmente foram escolhidos de forma empírica e depois refinados junto aos colaboradores.

3.4. Metodologia experimental

Na **Figura 5** podemos observar um esquema resumido de como foram realizadas as análises.

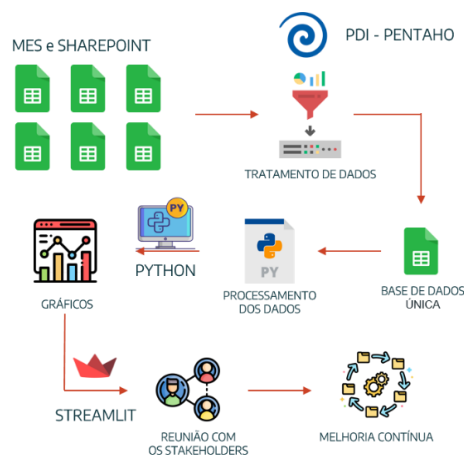


Figura 5: Resumo do processo de análise. Fonte: Autoria própria.

O processo se iniciava com a obtenção junto aos bancos de dados espelhos dos dados brutos. Então foi realizado o pré-processamento e limpeza desses dados de forma que eles estivessem mais amigáveis e no formato correto para a etapa de processamento.

Como exemplo desses ajustes foram a retirada de dados incoerentes, extração e criação de colunas novas a partir das originais. Para essas atividades foi utilizado o software Pentaho Data Integration (PDI).

A etapa de processamento dos dados é o coração do processo. Aqui ocorreu a tradução do Operation em códigos pouco amigáveis para agrupamentos mais claros e fáceis de compreender. Outras colunas foram criadas ou modificadas para facilitar a análise e permitir que a análise exploratória ocorresse. Aqui utilizamos o Collab do Google e o Jupyterlab versão 3.014, ferramenta integrada da plataforma Anaconda navigator versão 2.04.

Através das análises e dos gráficos construídos na etapa de processamento reuniões eram realizadas com supervisores, líderes de time e gestores para avaliar a veracidade e importância das informações conseguidas durante a análise exploratória. Essa troca permitiu estabelecer um processo de melhoria tanto para a análise quanto para a linha de produção que passou realizar intervenções apoiadas no conhecimento construído a partir das reuniões. Foram construídos gráficos de pareto, barras, histogramas, boxplot, violino e dispersão para realizar a análise descritiva das paradas de linha.

4. Análise e Discussão dos Resultados

Ao analisar o perfil das paradas de linha sob o prisma do tempo, do produto e das justificativas (Operation) algumas características se sobressaíram. Ao observar a **Figura 6** podemos perceber que a parada de linha é algo comum durante a rotina diária. Ao longo do período temos em média 1,42 paradas por veículo produzido e 66% dos veículos produzidos param em algum momento na linha. Muitos carros param e por múltiplas vezes. Esta característica pode indicar oportunidade de melhoria no processo ou na execução da montagem, já que eventos mais frequentes impactam mais no tempo de parada.

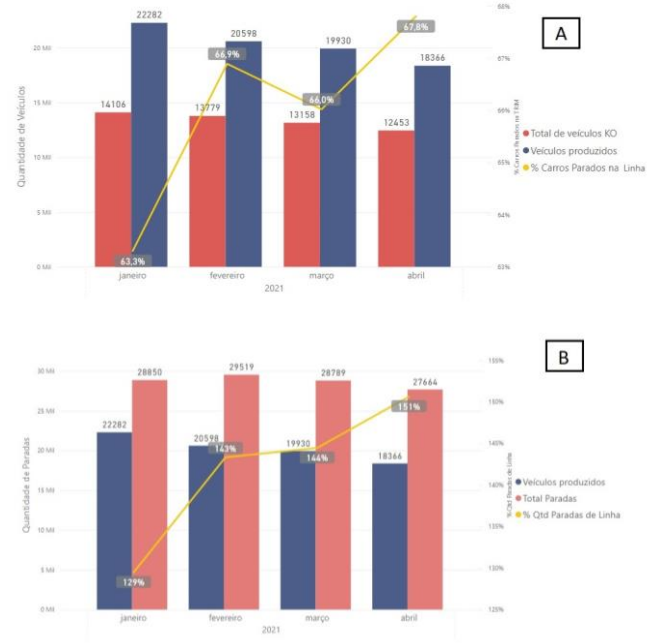


Figura 6: Gráfico contendo comparativo dos veículos produzidos vs quantidade de veículos parados (A) e vs quantidade de paradas de linha e consequentemente a taxa de parada por carro(B). Fonte: Autoria própria.

Outra característica é a baixa duração dos tempos de paradas. Temos paradas curtas mas com alta frequência, esse efeito se acumula ao longo do tempo gerando grande perda financeira. Mesmo quando observamos a distribuição do tempo de parada para uma justificativa específica numa faixa (categoria) de tempo as paradas tendem a se concentrar próximo ao primeiro quartil (**Figura 7**).

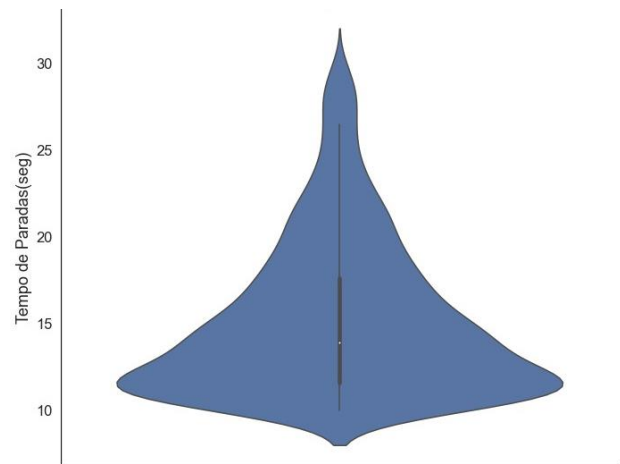


Figura 7: Gráfico de violino indicando a distribuição do tempo de parada para a justificativa de Certificação e tempo de parada entre 10s e 30s. Fonte: Autoria própria.

Se observarmos na faixa das paradas entre 10s e 30s há uma maior densidade próximo aos 10s. O que corrobora com o perfil de paradas curtas.

Ao confrontar o tempo de parada das principais estações responsáveis pelo maior tempo de parada acumulado no período. Além de permitir atuar diretamente nos locais mais críticos, também nos permitiu observar uma discrepância entre o tempo de parada das estações ao longo dos turnos.

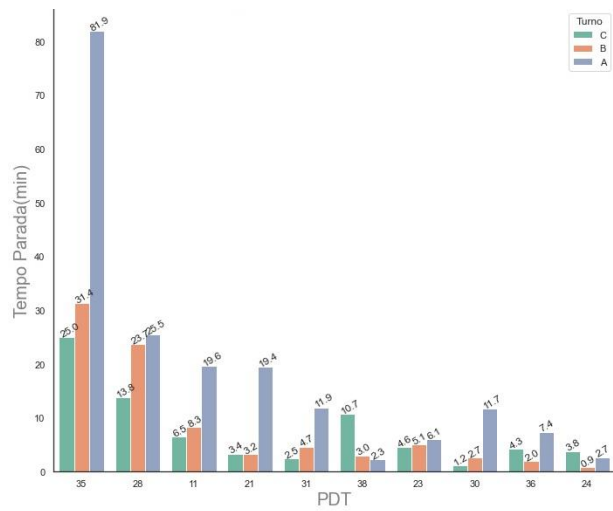


Figura 8: Gráfico contendo a estratificação do tempo de parada entre 1s e 10s para a justificativa de Certificação e distribuída por Turno. Fonte: Autoria própria.

Como pode ser visto na **Figura 8**, o turno A é responsável por 80% do tempo de parada para a estação de trabalho 35. Isto sinaliza um ponto de atenção. O que é feito neste turno, ou como são realizadas as atividades de montagem neste local e horário que podem explicar tamanha diferença?! Provavelmente a diferença deve estar em algum tipo de controle ou boa prática na gestão do ambiente laboral.

Com relação às justificativas automáticas observamos tanto a frequência de ocorrência quanto o total de tempo de parada acumulado para cada uma delas. Na **Figura 9** podemos observar o gráfico de pareto das justificativas (Operation) em relação ao somatório do tempo de parada.

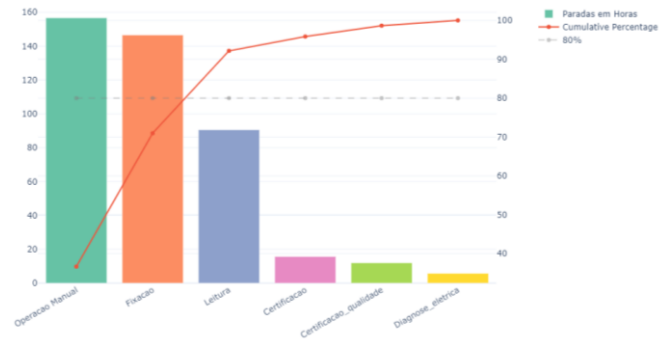


Figura 9: Gráfico diagrama de pareto do tempo de paradas por operation. Fonte: Autoria própria.

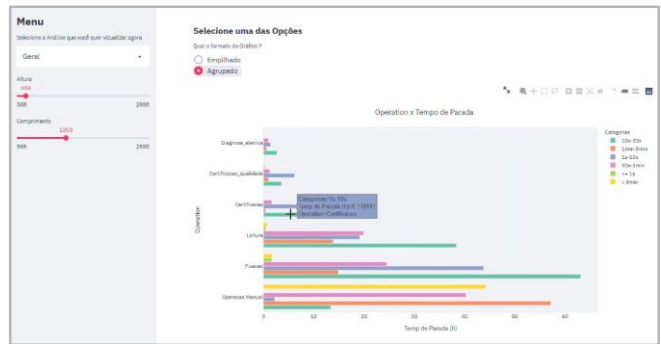


Figura 10: Aplicação web desenvolvida utilizando o framework open-source Streamlit. Fonte: Autoria própria.

Além disso, foi desenvolvido aplicativo web interativo, baseado nos dados previamente tratados, contendo todos os gráficos de análise diagnóstica para auxiliar nas tomadas de decisão dos Team Leaders (**Figura 10**).

5. Conclusões e Trabalhos Futuros

Algumas iniciativas foram implementadas após as reuniões visando reduzir o tempo perdido com paradas na linha de produção. Com isso, só na primeira semana de implementação houve um ganho de R\$ 86.418,15 reais devido a três atividades lançadas no sistema de projetos da empresa e foi iniciado uma expansão para as outras linhas produtivas, com o intuito de expandir essa análise. Este é um processo que deve ser contínuo e interativo entre o setor de análise e os envolvidos da linha de produção, visto que a linha de produção é um organismo vivo que tem seu comportamento alterado o tempo todo. Para o futuro se faz necessário permitir que estas análises ocorram de forma mais tempestiva

possível trazendo mais agilidade para os supervisores e líderes de equipes.

Além disso, avançar sobre os processos de mineração de dados como regras de associação e clusterização. Se fazem necessários para aumentar o entendimento sobre o fenômeno parada de linha. Permitindo, no futuro próximo, a construção de um modelo preditivo. Para que os envolvidos possam agir com antecedência. Adotando ações que mitiguem as perdas antes delas acontecerem. A adoção de uma política de dados mais clara e estruturada se faz necessária. Para permitir que mais informações possam ser analisadas e ampliar o espectro de atuação da companhia no aumento de sua eficiência produtiva. Sem dados não há informação e sem informação não há conhecimento.

Referências

- [1] KRUGH, MATTHEW ET AL. Prediction of defect propensity for the manual assembly of automotive electrical connectors. *Procedia Manufacturing*, v. 5, p. 144-157, 2016.
- [2] RÖHRDANZ, F., 1997. CAD method for industrial assembly, concurrent design of products, equipment and control systems. *Computer-Aided Design*, 29(10), p.737.
- [3] BI, ZHUMING M.; WANG, LIHUI; LANG, SHERMAN YT. Current status of reconfigurable assembly systems. *International Journal of Manufacturing Research*, v. 2, n. 3, p. 303-328, 2007.
- [4] U.SHAFIQUE AND H.QAISER. A Comparative Study of Data Mining Process Models (KDD, CRISP-DM and SEMMA). *International Journal of Innovation and Scientific Research*, ISSN 2351-8014 Vol. 12 No. 1 Nov. 2014, pp. 217-222.
- [5] What is CRISP DM? available at: <https://www.datascience-pm.com/crisp-dm-2/>. acessado em 18 de Agosto de 2021.
- [6] M.HOFMANN, F.NEUKART, T.BÄCK. Artificial Intelligence and Data Science in the Automotive Industry, arXiv preprint arXiv:1709.01989,(2017).
- [7] DATA ANALYTICS, BIG DATA, DATA SCIENCE. disponível em: <https://www.cetax.com.br/blog/data-mining/acessado-em-20-de-Agosto-2021>.
- [8] PARMEZAN, A. R. S. et al. Avaliação de Métodos para Seleção de Atributos Importantes para Aprendizado de Máquina Supervisionado no Processo de Mineração de Dados. 2012. Relatórios técnicos do laboratório de bioinformática, Universidade estadual do oeste do Paraná, Foz do Iguçu, 2012.
- [9] CHONGWATPOL, JONGSAWAS. Prognostic analysis of defects in manufacturing. *Industrial Management & Data Systems*, 2015.
- [10] STEVAN JUNIOR, S. L., LEME, M. O., SANTOS, M. M. D. Indústria 4.0: fundamentos, perspectivas e aplicações. São Paulo: Érica, 2018.
- [11] MACÊDO, LETICIA COSTA. Manutenção preditiva no contexto da indústria 4.0: um modelo preditivo em uma fábrica do ramo metalúrgico. 2020.
- [12] ZUEGE, TIAGO JASPER. Aplicação de técnicas de mineração de dados para detecção de perdas comerciais na distribuição de energia elétrica. 2018. Trabalho de Conclusão de Curso.
- [13] THIEDE, SEBASTIAN ET AL. Data mining in battery production chains towards multi-criterial quality prediction. *CIRP Annals*, v. 68, n. 1, p. 463-466, 2019.
- [14] KHAKIFIROOZ, MARZIEH; CHIEN, CHEN FU; CHEN, YING-JEN. Bayesian inference for mining semiconductor manufacturing big data for yield enhancement and smart production to empower industry 4.0. *Applied Soft Computing*, v. 68, p. 990-999, 2018.
- [15] HARDING, JENNY A.; SHAHBAZ, MUHAMMAD; KUSIAK, A. Data mining in manufacturing: a review. 2006.
- [16] ÇIFLIKLI, CEBRAİL; KAHYA-ÖZYIRMIDOKUZ, ESRA. Implementing a data mining solution for enhancing carpet manufacturing productivity. *Knowledge-Based Systems*, v. 23, n. 8, p. 783-788, 2010.
- [17] WANG, KE-SHENG. Towards zero-defect manufacturing (ZDM) — a data mining approach. *Advances in Manufacturing*, v. 1, n. 1, p. 62-74, 2013.
- [18] CHOUDHARY, ALOK KUMAR; HARDING, JENNY A.; TIWARI, MANOJ KUMAR. Data mining in manufacturing: a review based on the kind of knowledge. *Journal of Intelligent Manufacturing*, v. 20, n. 5, p. 501-521, 2009.
- [19] DEAN, JARED. Big data, data mining, and machine learning: value creation for business leaders and practitioners. John Wiley & Sons, 2014.