


**MINERAÇÃO DE PROCESSOS EM CONCESSIONARIAS DA INDÚSTRIA AUTOMOTIVA:
ESTUDO DE CASO, ANÁLISE E NEGÓCIO**

Wilson Portela Junior^A, Márcio Fontana Catapan^B, Fernando Deschamps^C



ARTICLE INFO	RESUMO
<p>Article history:</p> <p>Received 26 August 2022</p> <p>Accepted 18 November 2022</p>	<p>Objetivo: Como objetivo, busca-se aplicar uma metodologia de mineração de processos nas informações de pós-venda de concessionárias de automóveis, de uma montadora brasileira, analisando como negócio.</p>
<p>Palavras-chave:</p> <p>Minação de processos; Concessionárias; Pós-venda; Indústria automotiva.</p> <div data-bbox="172 943 480 1182">  </div>	<p>Referencial teórico: A utilização da mineração de processos na indústria automotiva proporciona diversos benefícios gerenciais, que são refletidos na melhora da qualidade do produto. Quando se trata de concessionária de automóveis e seus serviços de pós-venda, satisfação e/ou reclamações de cliente, evidencia uma escassez na literatura que reforçam o uso da técnica de mineração de processos. Dessa forma, para cobrir tal lacuna, há clara a necessidade desta pesquisa.</p> <p>Metodologia/abordagem: Primeiramente, técnicas de estruturação de dados foram desenvolvidas para a adequação dos dados brutos cedidos pela empresa estudada, para um padrão adequado à mineração de processos. Busca-se assim, obter informações, tendências e padrões em dados de clientes através do sistema de pós-venda destas concessionárias, durante um período e, apresentar o resultado destas análises.</p> <p>Resultados: Como resultados, destaca-se a criação de gráficos e tabelas que mostram a evolução de cada problema, relatados pelos clientes, em quilometragens de revisão e qual foi encontrado pelas concessionárias, baseando-se em palavras-chave citadas pelos clientes no momento da reclamação.</p> <p>Pesquisa, implicações práticas e sociais: As implicações desta pesquisa culminam na melhora da análise e utilização dos dados pela empresa estudada, exaltam a importância da mineração de processos, para encontrar informações escondidas na grande quantidade de dados, gerada por uma empresa multinacional do setor automotível.</p> <p>Originalidade: Nesta pesquisa foi detectada para o setor de concessionárias de automóveis há uma escassez de pesquisas voltadas a mineração de processos. Logo, esta pesquisa, teve como objetivo propor esta metodologia.</p> <p>Doi: https://doi.org/10.26668/businessreview/2022.v7i4.963</p>

^A Mestrando em Engenharia de Manufatura pela Universidade Federal do Paraná (UFPR). Av. Cel. Francisco H. dos Santos, 100, Jardim das Américas, Curitiba - PR, CEP: 81530-000. E-mail: wilson.portela@ufpr.br
Orcid: <https://orcid.org/0000-0003-1973-7187>

^B Doutorado em Engenharia Mecânica pela Universidade Federal do Paraná (UFPR). Av. Cel. Francisco H. dos Santos, 100, Jardim das Américas, Curitiba - PR, CEP: 81530-000. E-mail: marciocatapan@ufpr.br
Orcid: <https://orcid.org/0000-0003-1039-3939>

^C Doutorado em Engenharia de Produção pela Universidade Federal do Paraná (UFPR). Av. Cel. Francisco H. dos Santos, 100, Jardim das Américas, Curitiba - PR, CEP: 81530-000. E-mail: fernando.deschamps@ufpr.br
Orcid: <https://orcid.org/0000-0002-5269-3721>

PROCESS MINING IN AUTOMOTIVE DEALERSHIPS: CASE STUDY, ANALYSIS AND BUSINESS**ABSTRACT**

Objective: As an objective, it is sought to apply a process mining methodology in the after-sales information of car dealerships, of a Brazilian automaker, analyzing it as a business.

Theoretical framework: The use of process mining in the automotive industry provides several managerial benefits, which are reflected in the improvement of product quality. When it comes to car dealerships and their after-sales services, customer satisfaction and/or complaints, there is a scarcity in the literature that reinforces the use of the process mining technique. Thus, to cover such a gap, there is a clear need for this research.

Methodology/approach: First, data structuring techniques were developed to adjust the raw data provided by the company studied to a standard suitable for process mining. Thus, it is sought to obtain information, trends and patterns in customer data through the after-sales system of these dealerships, during a period and, present the result of these analyses.

Results: As results, we highlight the creation of graphs and tables that show the evolution of each problem, reported by customers, in review mileage and which was found by the dealerships, based on keywords cited by customers at the time of complaint.

Research, practical and social implications: The implications of this research culminate in the improvement of the analysis and use of data by the company studied, exalting the importance of process mining, to find hidden information in the large amount of data, generated by a multinational company in the automotive sector.

Originality: In this research it was detected that for the sector of car dealerships there is a scarcity of researches focused on process mining. Therefore, this research aimed to propose this methodology.

Keywords: Process mining, Dealerships, After sales, Automotive industry.

LA MINERÍA DE PROCESOS EN LOS CONCESIONARIOS DE AUTOMÓVILES: ESTUDIO DE CASOS, ANÁLISIS Y NEGOCIO**RESUMEN**

Objetivo: Como objetivo, se busca aplicar una metodología de minería de procesos en la información de posventa de las concesionarias de automóviles, de una automotriz brasileña, analizando como negocio.

Marco teórico: El uso de la minería de procesos en la industria del automóvil proporciona varios beneficios de gestión, que se reflejan en la mejora de la calidad del producto. Cuando se trata de los concesionarios de automóviles y sus servicios posventa, la satisfacción de los clientes y/o las reclamaciones, se evidencia una escasez en la literatura que refuerza el uso de la técnica de minería de procesos. Por lo tanto, para cubrir esta laguna, existe una clara necesidad de esta investigación.

Metodología/enfoque: En primer lugar, se desarrollaron técnicas de estructuración de datos para adaptar los datos brutos proporcionados por la empresa estudiada, a un estándar adecuado para la minería de procesos. Así, se busca obtener información, tendencias y patrones en los datos de los clientes a través del sistema de posventa de estos concesionarios, durante un periodo y, presentar el resultado de estos análisis.

Resultados: Como resultados, destaca la creación de gráficos y tablas que muestran la evolución de cada problema, reportado por los clientes, en el kilometraje de revisión y que fue encontrado por los concesionarios, en base a las palabras clave citadas por los clientes en el momento de la reclamación.

Implicaciones de la investigación, prácticas y sociales: Las implicaciones de esta investigación culminan en la mejora del análisis y uso de los datos por parte de la empresa estudiada, exaltando la importancia de la minería de procesos, para encontrar información oculta en la gran cantidad de datos, generados por una empresa multinacional del sector del automóvil.

Originalidad: En esta investigación se ha detectado que para el sector de los concesionarios de automóviles hay una escasez de investigaciones centradas en la minería de procesos. Por lo tanto, esta investigación tenía como objetivo proponer esta metodología.

Palabras clave: Mineração de processos, Concessionárias, Postventa, Industria del automóvil.

INTRODUÇÃO

O contexto contemporâneo empresarial é resultado de inúmeras transformações, que fizeram com que as empresas repensassem a forma de como elas entregam seus produtos ou serviços aos clientes. Assim, nesse atual cenário, as empresas são convocadas a adotarem melhores práticas, caso queiram se diferenciar em um mercado, que é marcado pela competitividade acirrada (MARTINS; SANTOS; CARVALHO, 2010).

Ahn e Sohn (2009) citam que a percepção do pós-venda, como sendo uma oportunidade e vantagem competitiva para o negócio. Ela exige uma mudança de visão, deixando de lado a perspectiva que é tão comum no mercado centrada no produto, tratando o pós-venda como espécie de mal necessário, para uma visão mais centrada no cliente. Ainda, esses mesmos autores destacam que o universo de informações existentes quanto ao pós-venda é muito amplo, o que exige a adoção de técnicas que facilitem a análise de informações.

Desta maneira, o problema em questão a ser tratado é: como aplicar uma metodologia de mineração de processos que seja factível a área de serviço de pós-venda de concessionária de veículos? Embora o processo de mineração trate do contexto organizacional dos processos de negócios, poucas pesquisas têm sido realizadas na análise de processos de negócios na perspectiva organizacional (SONG e VAN DER AALST, 2008).

Neste contexto, Suriadi et al. (2013) e Van Eck, Leemans, Van der Aalst. (2015), relatam que a carência de metodologias de mineração de processos, específicas da área de pós-venda, faz com que a curva de aprendizado para que se possa desenvolver um projeto nessa área, torne-se muito complexa e trabalhosa. Além disso, poucas das metodologias para aplicação geral da mineração de processos encorajam, explicitamente, a análise iterativa e/ou a associação com o conhecimento tácito contido nas pessoas envolvidas no processo a ser estudado.

Sendo assim, objetiva-se nesta pesquisa o desenvolvimento e a aplicação de uma metodologia de mineração de processos, de maneiras sistêmica, no setor de pós-venda de concessionárias de veículos, que auxilie gestores na tomada de decisão.

REFERENCIAL TEÓRICO

Este capítulo tem como objetivo fornecer conceitos para o melhor entendimento da pesquisa, introduzindo os conceitos-chave e seus desmembramentos específicos.

Mineração de Processos

A mineração de processos é uma técnica que bem serve aos esforços da empresa de tratamento e análise dos dados de que dispõe. Van Der Alst (2016) define que é a área de pesquisa, na qual o principal objetivo é a extração de informações sobre o comportamento dos processos de negócios do log de eventos (registro completo de um período ou um evento (VAN ECK ET AL, 2015)). Esses são registrados pelos sistemas de informação dentro da organização. Esses podem ser armazenados tanto em tabelas de banco de dados como em logs de mensagens, arquivos de e-mail, entre outras fontes.

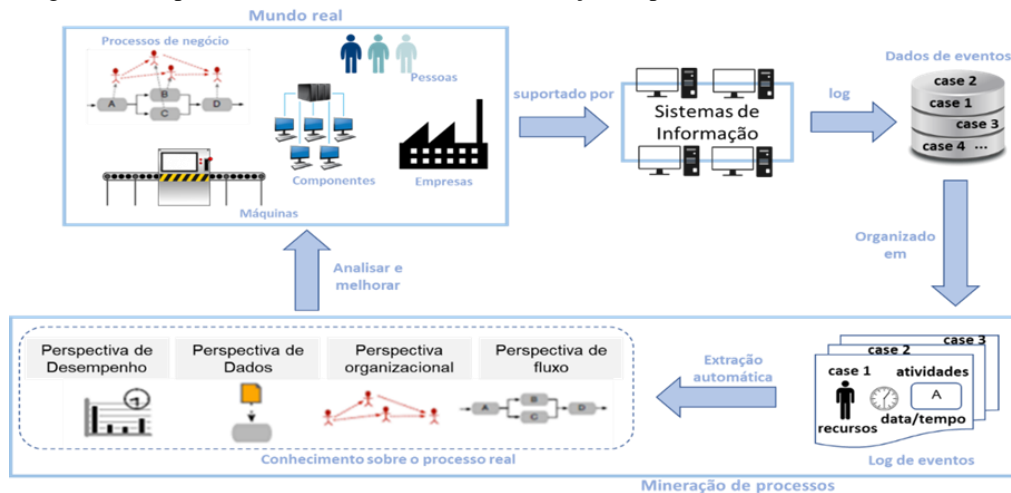
A mineração de processos (do inglês: process mining), é uma disciplina localizada entre o aprendizado de máquina (do inglês: machine learning) e mineração de dados (do inglês: data mining), e a modelagem de processos e análise. Esta técnica busca descobrir, monitorar e otimizar processos através dos dados extraídos dos sistemas atuais.

Segundo Van Der Alst (2011), a mineração de processos pode ser classificada em três grupos. A saber:

- **Descoberta:** é a técnica leva em consideração um registro de eventos e produz um modelo, sem que haja a necessidade de qualquer informação adicional;
- **Conformidade:** onde o modelo do processo existente é comparado com um registro de eventos. Assim, através da verificação de conformidade pode-se analisar se os acontecimentos estão de acordo com os registros armazenados, sendo possível detectar, localizar e explicar os desvios existentes no processo;
- **Aprimoramento:** a qual esta técnica busca estender ou aperfeiçoar o modelo do processo, através das informações geradas no decorrer dos eventos.

É possível ainda abordar a mineração de processos através de perspectivas segundo Van der Aalst (2015). A Figura 1 apresenta as quatro perspectivas obtidas a partir do uso da mineração de processos.

Figura 1. Perspectivas obtidas com o uso da mineração de processos. Santos e Neumann (2019)



Verifica-se na Figura 1 uma visão geral do ciclo de aplicação da mineração de processos, os quais sistemas existentes no mundo real geram e registram logs nos sistemas de informação. Esses logs incluem informações sobre as instâncias dos processos, atividades realizadas, tempos, recursos (pessoas), dentre outras. Ainda, segundo Van der Aalst (2016), o uso da mineração de processos então se utiliza desses logs para extrair as quatro perspectivas descritas a seguir:

Perspectiva de fluxo, a qual se concentra no processo e na ordenação das atividades. O propósito desta perspectiva é encontrar uma boa caracterização de todos os caminhos possíveis, por exemplo, expressando-os em termos de uma rede de Petri ou alguma outra notação, tal como o Business Process Model Notation (BPMN) ou o Event-driven Process Chain (EPC). Esta perspectiva permite à organização identificar não somente fluxos padrões, geralmente descobertos através do mapeamento tradicional, mas também fluxos raros existentes na prática e difíceis de serem identificados com técnicas manuais de mapeamento.

Perspectiva organizacional, que foca em informações sobre os recursos presentes no log, ou seja, quais atores (pessoas, sistemas, funções ou departamentos) estão envolvidos e como eles se relacionam. O propósito é estruturar a organização por meio da classificação de pessoas em termos de funções e unidades organizacionais ou para mostrar a rede social (handover de trabalhos, por exemplo).

Perspectiva de dados, concentra-se nas propriedades ou atributos associados as instâncias (cases) do processo. Uma instância pode ser caracterizada pelo caminho no processo ou pelos “atores” que trabalham nele. No entanto, cases pode também ser caracterizados pelos valores dos seus elementos de dados. Por exemplo, se um case representa um pedido de

reposição, pode ser interessante conhecer o fornecedor ou o número de produtos encomendados.

Perspectiva de tempo, que está preocupada com o momento e a frequência dos eventos. Quando os eventos registram timestamps é possível identificar gargalos, medir os níveis de serviço, monitorar a utilização dos recursos, e prever o tempo restante das instâncias em execução.

Tais informações podem ser usadas para auxiliar no desenvolvimento e implementação de novos sistemas. Esses suportam a execução de processos de negócios, servindo também, para apoiar melhorias, a auditoria ou análise de processos existentes, bem como para eliminar desperdícios em um ambiente da organização (VAN ECK, LEEMANS, VAN DER AALST., 2015).

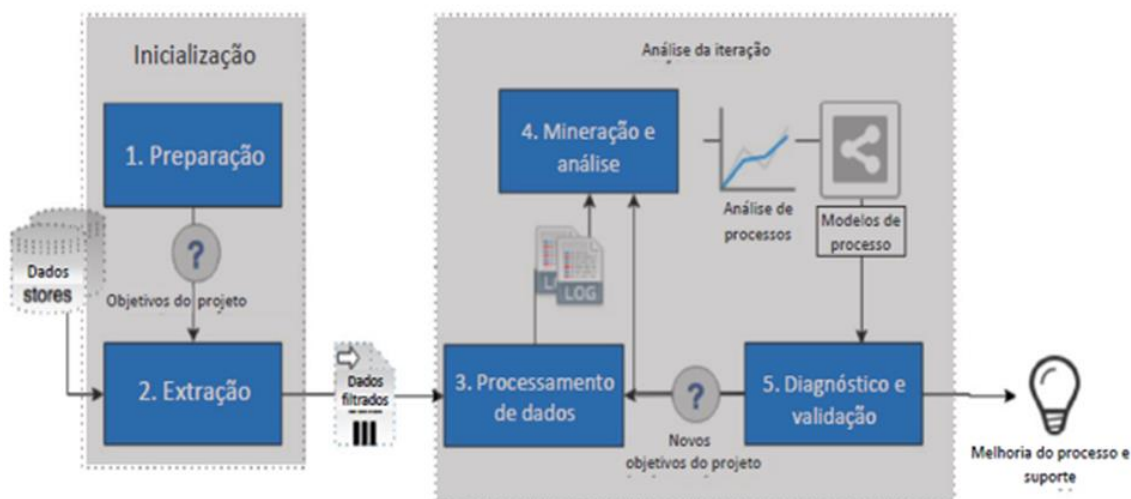
METODOLOGIA DA PESQUISA

Para esta etapa denominada desenvolvimento, as características definidas como sendo ou não desejáveis para uma metodologia, que seja compatível com o contexto da área de pós-venda de concessionárias de veículos, são analisadas. Como primeira fase, busca-se selecionar a metodologia de mineração de processo adequada.

Seleção da metodologia

Esta seção discorre sobre a seleção do processo da mineração, propriamente dito, sendo utilizada a metodologia PM² desenvolvida por Van Eck, Leemans, Van der Aalst (2015). Maita et al (2018) já destacavam que a opção por essa metodologia é pelo fato por ser a mais utilizada e mais bem consolidada dentre as metodologias disponíveis. O orientador dessa obra, Wil Van der Aalst, publicou mais de 250 artigos em periódicos, 22 livros (como autor ou editor), 550 publicações em conferências e 80 capítulos de livros relacionados aos temas mineração de processos, redes de Petri e gerenciamento de processos de negócios. Isso demonstra a importância do autor na área. Cabe-se ressaltar que o índice “h” deste pesquisador é maior que 100.

O PM² busca guiar organizações para desenvolver projetos de mineração de processos, de forma a melhorar a performance e/ou a conformidade dos processos no que tange às regras e regulamentações. O desenvolvimento do PM² ocorre em seis etapas, ilustradas na Figura 2, como sendo: (1) Preparação, (2) Extração, (3) Processamento de dados, (4) Mineração e análise, (5) Diagnóstico e validação (VAN ECK, LEEMANS, VAN DER AALST., 2015).

Figura 2. Visão geral do PM². Adaptado de Van Eck, Leemans, Van der Aalst. (2015)

Conforme ilustrado na Figura 2, os dois estágios iniciais são (1) Preparação e (2) Extração e compõem a inicialização do processo, onde serão definidas as questões de pesquisa e ocorrerá a extração dos dados.

Os três próximos estágios: (3) Processamento de dados, (4) Mineração e análise e (5) Diagnóstico e validação, compreendem uma etapa iterativa de análise. Nessa etapa, ocorre uma ou mais instâncias de cada uma delas. Essa etapa busca responder às questões de pesquisa previamente colocadas. Esse processo não possui um tempo padrão de análise, uma vez que depende da complexidade e tratamento dos dados. Uma vez que se chega a um resultado, este pode ser utilizado para o desenvolvimento da etapa (6) Melhoria do processo e suporte.

Uma descrição do objetivo, das atividades, bem como suas entradas e saídas em cada etapa, segundo Van Eck, Leemans, Van der Aalst. (2015), encontram-se a seguir:

- 1) Preparação – Busca organizar o projeto e determinar as questões de pesquisa para ao fim, poder melhorar a performance do processo ou analisar a conformidade com regras e regulamentos. Possui três atividades (i) Identificação das questões de pesquisa, (ii) Seleção do processo e (iii) Composição do time do projeto (pessoas envolvidas no processo e profissionais de TI);
- 2) Extração – Busca extrair os dados de evento e, quando possível, os modelos de processo. Possui três atividades: (i) Determinação do escopo, (ii) Extração dos dados do sistema e (iii) Transferência do conhecimento do processo (extração de conhecimento tácito dos envolvidos no processo);

- 3) Processamento de dados – Busca criar os logs de eventos como visualizações diferentes daquelas obtidas pelos dados de eventos, bem como processar os logs de modo que estes fiquem tratados para serem inseridos na mineração e análise.
- 4) Mineração e análise: Busca aplicar as técnicas de mineração de processos nos logs de eventos para responder as questões de pesquisa e obter abstrações sobre a performance a conformidade do processo.
- 5) Diagnostico e validação: Busca relacionar os resultados das análises a ideias de melhorias de modo a atingir o objetivo do projeto.

Com base no que foi exposto até aqui, uma vez definido o protocolo de pesquisa e consequente revisão bibliográfica da literatura, é visível a carência de artigos e publicações que mostrem a aplicação da técnica de mineração de processos no pós-venda. Em relação à definição da metodologia de mineração de processos, optou-se pela utilização da metodologia PM², sendo esta já devidamente consolidada e de larga aplicação. Desta maneira, o próximo capítulo trará a aplicação desta metodologia em um estudo de caso.

ESTUDO DE CASO

O estudo de caso foi realizada em uma indústria automobilística localizada na cidade de São José dos Pinhais – PR. A empresa vem investindo em sistemas e metodologias que buscam otimizar a grande quantidade de dados gerados advindos tanto de clientes como das áreas de projeto e manufatura.

A montadora dispõe de uma base de dados onde são catalogadas todas as reclamações de clientes em todas as concessionárias do Brasil. Nela, além das reclamações, são listados os modelos, o tipo do motor, o chassi, a quilometragem em que o problema foi encontrado, o diagnóstico técnico da concessionária, a classe de problema que foi encontrada a falha, a especificação da falha, além das datas de reclamação, fabricação e compra dos veículos, entre outros atributos. Porém, não são realizadas o tipo de análise ou predição baseada nesses dados, servindo apenas para a exposição de resultados do setor.

O pós-venda da empresa sofre com um alto custo de reposição de peças em concessionárias. Essas, muitas vezes desnecessárias e sem um plano robusto de predição de problemas nos veículos, resultando em falta de velocidade no momento de resolver um problema relatado por um cliente. Por esta razão, foi proposta a presente análise.

Preparação

O setor de qualidade da empresa acompanhada de forma sistemática o índice de defeitos em seus veículos, mas esses não são diretamente reportados a este ou a outro setor da fábrica. Essas reclamações e identificações de defeitos vêm através das reclamações dos clientes da rede de concessionárias de automóveis, onde os clientes compraram seus veículos. Para a realização de um estudo exploratório para identificar os principais defeitos de seus veículos, definiu-se o escopo do projeto, a identificação dos modelos que concentram o maior número de problemas, e a partir deles obter uma visão geral dos defeitos. Ficou definido que o sistema de informações que poderia fornecer os dados para o projeto seria o sistema de registros de pedidos de reparo que a concessionária utiliza. Essa tomada de decisão veio de Especialistas em Negócios da própria empresa estudada.

Para este estudo, a empresa disponibilizou logs de sua cadeia de concessionárias de automóveis, onde a partir do número do chassi, é possível identificar e rastrear os eventos que ocorrem em um veículo. Esses registros também possuem códigos de problemas, usados pela empresa, quantos quilômetros um veículo têm em uma ordem de reparo e o modelo do carro.

As possíveis ocorrências são fabricação de veículos, registro, entrega de veículos, data de venda do veículo e ordem de reparo. O registro BIM e BVM (etapas onde o veículo passa por checagem de qualidade na fábrica) não serão considerados por serem registros internos da empresa e não têm influência no processo. As informações disponíveis para análise são as seguintes:

- i) A fabricação de veículos é o evento em que foi concluída a sua fabricação.
- ii) O veículo entregue é o evento quando ele é concedido da fábrica para a concessionária.
- iii) Veículo vendido é o evento em que ele é entregue ao cliente.
- iv) Ordem de reparo é quando um técnico identifica um problema e gera uma ordem para reparar p veículo.

A Figura 3 mostra o exemplo de um registro, onde o número do chassi foi borrado devido à privacidade dos dados da empresa.

Figura 3. Exemplo de registro de eventos

Chassis	Event Date	Event	Problem Code	Km	Model
93YHSR6P5*****	18/08/2018	Vehicle Manufactured			
93YHSR6P5*****	18/08/2018	BVM Registration			
93YHSR6P5*****	18/08/2018	Vehicle Delivered			
93YHSR6P5*****	18/08/2018	BIM Registration			
93YHSR6P5*****	18/08/2018	Vehicle Sold			
93YGSR6R3*****	07/12/2019	Repair Order	GSFA-V01-CAR BODY	40543	E5
93YHSR2L6*****	18/08/2018	Repair Order	GSFA-V13-TRIMS	41534	E6
93YHSR2L6*****	18/08/2018	Repair Order	GSFA-P14-CAR BODY	41042	E7
93YHSR2L6*****	18/08/2018	Repair Order	GSFA-V14-SEATS	41259	E8
93YHSR2L6*****	18/08/2018	Repair Order	GSFA-V14-INTERIOR	41042	E9

Seleção de processos

Para a seleção dos processos a serem estudados dentro da área de pós-venda e, para verificar quais dados e registro de logs de eventos estariam disponíveis, foi realizado durante um mês, diversas reuniões foram realizadas. Os participantes dessa foram o gerente de qualidade de serviços da rede de concessionárias, o diretor de qualidade da fábrica de automóveis, diversos analistas e/ou engenheiros de qualidade, durabilidade e incidentologia. Nessas reuniões foram apresentadas imagens com exemplos de resultados e tabelas contendo dados que deveriam ser extraídos do sistema.

Desta forma, ficou definido que os dados seriam disponibilizados pela área de durabilidade. Esses abrangeriam os processos de tratamento de defeitos e reclamações, reportados pela rede de concessionárias, através da disponibilização dos logs de eventos.

Identificação das questões de pesquisa

O projeto aplicado na empresa foi estruturado de modo a ter como questão de pesquisa inicial: “Como é o fluxo de tratamento de reclamações e defeitos pela empresa em questão?”

A opção por uma questão de pesquisa inicial mais abstrata vai ao encontro das sugestões da metodologia mencionadas. Tal prática abre mais espaço para refinamento posterior através da análise exploratória, resultando em ideias concretas de melhoria e insights valiosos (VAN ECK, LEEMANS, VAN DER AALST., 2015).

Identificação das questões de pesquisa

Os profissionais envolvidos no projeto foram selecionados de modo a cobrir a esfera gerencial e técnica. Sendo assim, foram selecionados os seguintes cargos:

- i. Gestor do negócio: Gerente de qualidade pós-venda;
- ii. Especialista no negócio: pessoa responsável pela gestão dos dados da área de incidentologia (visão mais voltada ao processo);

iii. Especialista em sistemas: pessoa responsável pela análise de dados e sistemas (visão mais voltada aos dados, com formação na área de computação);

iv. Analista de processos: o próprio pesquisador desta dissertação;

Um ponto de atenção identificado na composição do time, foi que apenas o especialista no negócio possuía conhecimento sobre mineração de processos. Com a seleção do processo, das questões de pesquisa e da equipe do projeto, é possível passar para a etapa de extração dos dados.

Extração

Nesta fase, o especialista no negócio extraiu os dados de uma concessionária de automóveis, na qual julgou que teria dados maduros o suficiente e que gerariam uma boa amostra.

Determinação do escopo

Para a determinação do escopo, foi discutido junto ao gerente de qualidade de pós-venda e ao especialista do negócio, quais os elementos estavam disponíveis no sistema e como deveriam ser selecionados. Nesse ponto foi identificado um problema: o sistema não permite que haja grande manipulação de dados na extração propriamente dita. Dessa forma, apenas o período de extração pôde ser parametrizado e todas as outras colunas foram extraídas para tratamento posterior. Foi determinado o período de 01/01/2017 a 31/12/2019 para ser extraído do sistema, pois eram informações mais consolidadas e eram de anos completos.

Extração dos dados do sistema

Na fase de extração, pretende-se adquirir informações cadastradas em sistemas de informação. Os dados extraídos tiveram influência direta no escopo do projeto, que foi definido na etapa anterior. Em busca de defeitos relatados pelos clientes, onde além de definir quais dados seriam extraídos, foi determinado o período a ser extraído.

Adequação dos dados

Para a realização da adequação, foi necessário acessar o banco de dados responsável pelo armazenamento das informações do pós-venda. Foi possível extrair as informações contidas no período de 2019 a 2020. A planilha resultante foi gerada em MS Excel, com uma dimensão de 57 colunas e 154.563 linhas.

Entretanto, os dados brutos extraídos do banco de dados da empresa não estavam preparados para uma mineração de processos, pois havia campos vazios, datas não preenchidas, além da impossibilidade da inserção no software DISCO®, visto a falta das informações importantes nas planilhas fornecidas. Embora, já foi possível determinar que o chassi seria o evento do processo, onde as reclamações (que muitas vezes se repetem) giram em torno da identificação dos carros nas concessionárias.

Para solucionar os problemas citados, foi desenvolvido primeiramente um programa no Visual Basic for Applications (VBA). O objetivo foi de transformar as sete datas integrantes das planilhas (catalogadas como colunas) em linhas, para servirem como atividades para o processo.

Com isso, pôde-se visualizar as sete atividades do processo, que são: fabricação do veículo, compra, geração da nota fiscal, reclamação do cliente, faturamento do concerto para a concessionária, faturamento do concerto para a montadora e data base de incidentes (data em que são catalogadas as reclamações no banco de dados).

Para cada linha criada, foi então realocada a data referente a atividade em uma coluna adicional, que serve para o atributo de “timestamp” da mineração de processos no software DISCO®.

Porém, notou-se ainda que das sete atividades, três seriam fixas e não necessitariam aparecer mais de uma vez, mesmo que o cliente fosse a concessionária fazer mais de uma reclamação. Caso contrário, geraria loops de mineração de processos inexistentes. São elas: Fabricação do veículo, compra e geração da nota fiscal. Portanto, foi desenvolvido outro programa em VBA para excluir as linhas em que essas três datas fossem repetidas (quando um chassi aparecesse mais de uma vez na planilha). A Figura 4 mostra como ficou a tabela desenvolvida.

Figura 4. Reclamações de um veículo (único chassis)

Activity	Fabrication	Invoice	Purchase	Billing Dealer	Billing Company	Base Date	Reclaim	Chassis	Time stamp
Fabrication	19/05/2018							93YHSR6P5CJ717875	19/05/2018
Invoice		18/08/2018						93YHSR6P5CJ717875	18/08/2018
Purchase			25/09/2018					93YHSR6P5CJ717875	25/09/2018
Billing Dealer				05/07/2019				93YHSR6P5CJ717875	05/07/2019
Billing Company					19/07/2019			93YHSR6P5CJ717875	19/07/2019
Base Date						19/07/2019		93YHSR6P5CJ717875	19/07/2019
Reclaim							05/07/2019	93YHSR6P5CJ717875	05/07/2019
Fabrication	17/05/2018							93YHSR6P5CJ717875	17/05/2018
Invoice		18/08/2018						93YHSR6P5CJ717875	18/08/2018
Purchase			26/09/2018					93YHSR6P5CJ717875	26/09/2018
Billing Dealer				03/07/2019				93YHSR6P5CJ717875	03/07/2019
Billing Company					19/07/2019			93YHSR6P5CJ717875	19/07/2019
Base Date						19/07/2019		93YHSR6P5CJ717875	19/07/2019
Reclaim							03/07/2019	93YHSR6P5CJ717875	03/07/2019

Com essa modificação, os chassis que foram às concessionárias mais de uma vez não tiveram os dados de fabricação, compra e nota fiscal repetidos. Com isso, apenas os dados da nova reclamação foram contabilizados, o que, aliado às modificações anteriores, possibilitaria uma análise mais concreta e sem falhas no software DISCO®. Os logs de eventos são então importados para o software de mineração. Como já mencionado anteriormente, neste projeto o software utilizado foi o DISCO®.

Transferência de conhecimento do processo

A transferência do conhecimento iniciou-se antes do processo de Extração propriamente dito e estendeu-se até a etapa de Mineração e Análise. Ela se deu inicialmente por meio de visitas in loco em 3 concessionárias guiadas pelo gerente de qualidade de pós-venda, com inclusão dos responsáveis por cada seção.

Em seguida, o especialista no negócio e o especialista em sistemas tiveram participação direta em diversos momentos, por meio de reuniões e e-mails, para auxiliar na interpretação dos dados e sanar dúvidas pontuais de interpretação.

Um ponto de atenção apresentado nessa etapa foi a necessidade de alinhamento dos termos utilizados por cada um dos envolvidos. Tanto do lado dos responsáveis pelo processo laboratorial quanto do analista de processos.

Processamento de dados

Os dados recebidos estão preparados para se tornarem logs de eventos. Tais dados são parametrizados para um formato que pode ser processado no software de mineração. Muitas vezes os dados vêm com mais informações do que o necessário, e que não seriam relevantes para a análise, por isso nesta etapa também está definido quais dados têm relevância.

Na Figura 5, pode identificar as estatísticas geradas no software DISCO®, dentro dos atributos do modelo do veículo e código do problema, onde é possível identificar que existem 57 códigos de problema e 10 modelos de veículos, que estão cadastrados no sistema. A figura mostra que existe um predomínio do modelo A1, sendo responsável por mais de 60% das solicitações de reparação. No código do problema também é possível ter uma ideia de quais são os problemas mais constantes.

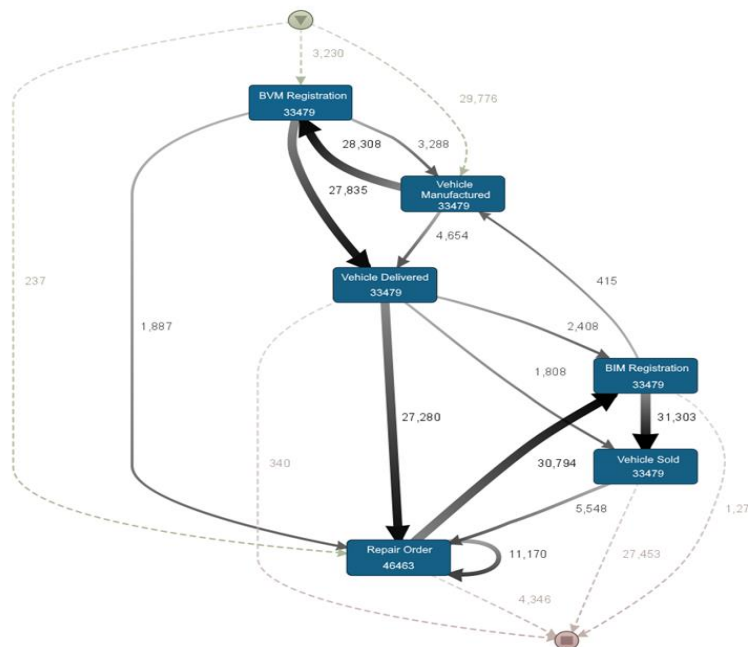
Figura 5. Estatística geradas pelo software Disco®

Problem Code Other attribute			Model Other attribute		
All values (57)			All values (10)		
Value	Frequency	Relative frequency	Value	Frequency	Relative frequency
GSFA-S22-PEDAL - PARKING BRAKE - JACK	6,808	14.65 %	A1	28,938	62.28 %
GSFA-M07-LOWER BASE ENGINE - CLUTCH	4,151	8.93 %	B2	5,429	11.68 %
GSFA-P05-ACCESS - VEHICLE PROTECTION	3,301	7.1 %	C3	4,543	9.78 %
GSFA-V14-SEATS	3,121	6.72 %	D4	2,919	6.28 %
GSFA-V07-EXTERIOR ACCESSORIES	2,446	5.26 %	E5	2,891	6.22 %
GSFA-S34-SUSPENSIONS - FILTRATION	2,096	4.51 %	F6	1,199	2.58 %
GSFA-V03-SEALING	2,018	4.34 %	G7	331	0.71 %
GSFA-P11-AIR CONDITIONNING	1,701	3.66 %	H8	164	0.35 %
GSFA-M05-CYL. BOCK- ENGINE COVERS - BR...	1,690	3.64 %	I9	27	0.06 %
GSFA-P20-COOLING	1,548	3.33 %	J10	22	0.05 %
GSFA-V10-DASHBOARD	1,466	3.16 %			
GSFA-P21-FUEL CIRCUIT	1,364	2.94 %			
GSFA-V04-MANEUVER OPENING PANEL	1,205	2.59 %			

Mineração e análise dos dados

Após a importação dos registros de eventos, o software de mineração gerou diversos mapas de processos. A Figura 6 é o mapa do processo em relação à frequência das ocorrências, considerando a frequência absoluta. Através do mapa do processo é possível identificar que 33.449 veículos foram cadastrados no sistema de informações, e que havia 46.463 ordens de reparo, nas quais 11.171 veículos possuíam mais de uma ordem de reparo.

Figura 6. Mapa do processo (frequência única)

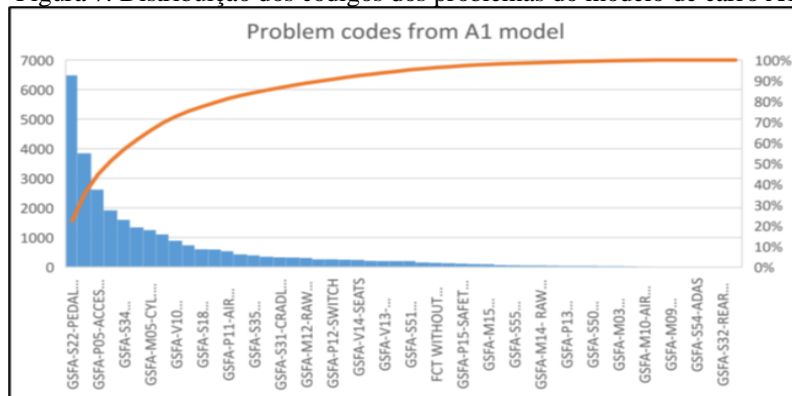


Utilizando as ferramentas de filtro do software de mineração, foram isolados os três modelos que apresentaram mais defeitos. Juntos a esses três modelos representam mais de 82% do total de defeitos. Após serem isolados, novos logs de eventos foram gerados. Alguns filtros

também foram usados para medir as mudanças nos atributos do quilômetro e o tempo que leva entre a venda de um veículo e um pedido de conserto.

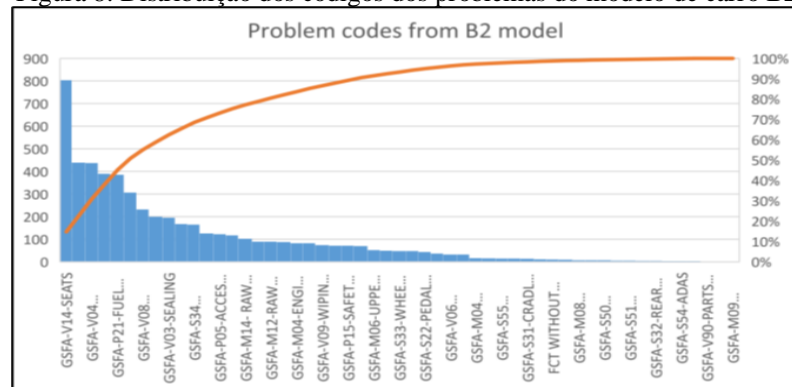
A partir dos registros de eventos fornecidos na etapa anterior, foram gerados diagramas de Pareto para melhor compreender a distribuição dos defeitos nos modelos de carros A1, B2 e C3, que concentraram a maioria dos defeitos. Na figura 7, do modelo A1 o defeito mais constante é o código GSFA-S22 (22,4%) seguido por GSFA-M07 (13,3%) e GSFA-P05 (9,06%).

Figura 7. Distribuição dos códigos dos problemas do modelo de carro A1



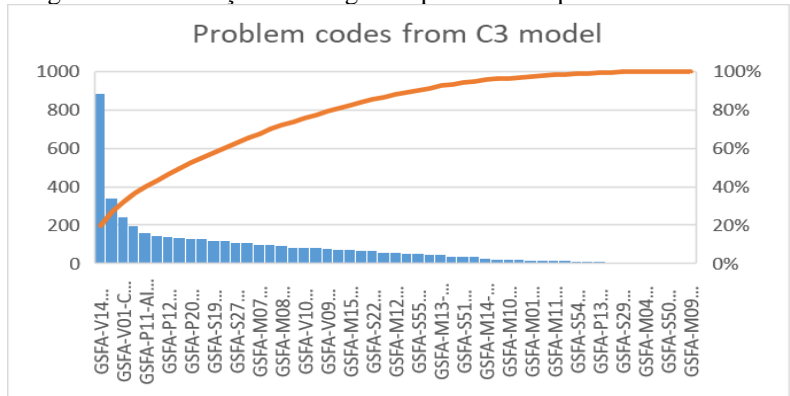
Na Figura 8, no modelo de carro B2, o defeito que mais ocorreu foi o GSFA-V14 (14,81%). Na sequência, os dois próximos códigos de defeitos têm valor muito próximo, GSFA-P11 (8,09%) e GSFA-V04 (8,05%). Neste modelo os defeitos são mais distribuídos do que no modelo A1.

Figura 8: Distribuição dos códigos dos problemas do modelo de carro B2



A Figura 9 mostra o diagrama do modelo C3, que é o terceiro modelo mais defeituoso. O código de problemas mais frequente é GSFA-V14 (19,46%), significativamente abaixo dos códigos GSFA-M06 (7,48%) e GSFA-V01 (5,35%).

Figura 9: Distribuição de códigos de problemas a partir do modelo C3



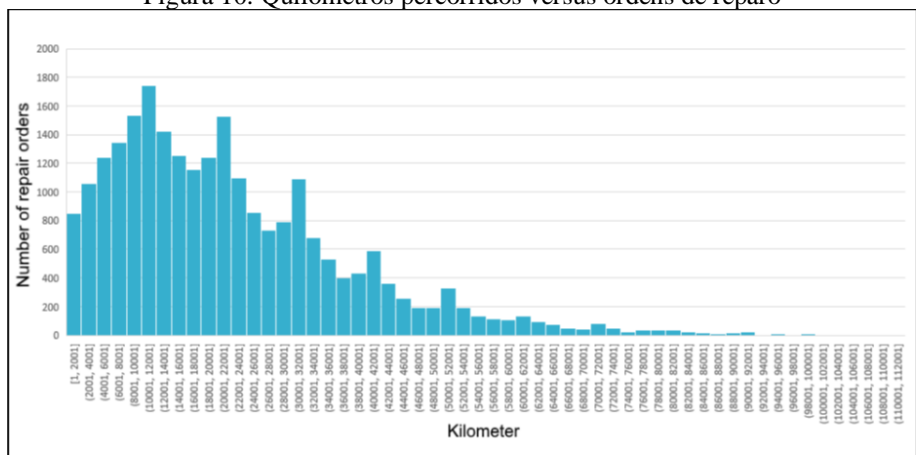
Uma análise foi feita focada nos seguintes tópicos:

O tempo que um veículo vendido leva para voltar para a concessionária;

- Reparo necessário, que pode ser uma primeira ordem de reparo ou uma segunda ordem de reparo;
- O veículo já foi previamente reparado.
- A análise revelou que, em ambos os casos, os defeitos ocorrem principalmente antes de completar um ano.

Outra análise realizada foi o número de quilômetros percorridos pelo veículo, quando foi feita uma ordem de reparo. O histograma na Figura 10 mostra as faixas de quilômetro e o número de ordens de reparo que foram feitas, em cada faixa. Através deste histograma é possível reforçar a análise do boxplot, em que os defeitos estão ocorrendo nos primeiros meses de uso, devido a estes estarem mais concentrados nas faixas iniciais de rodagem dos veículos.

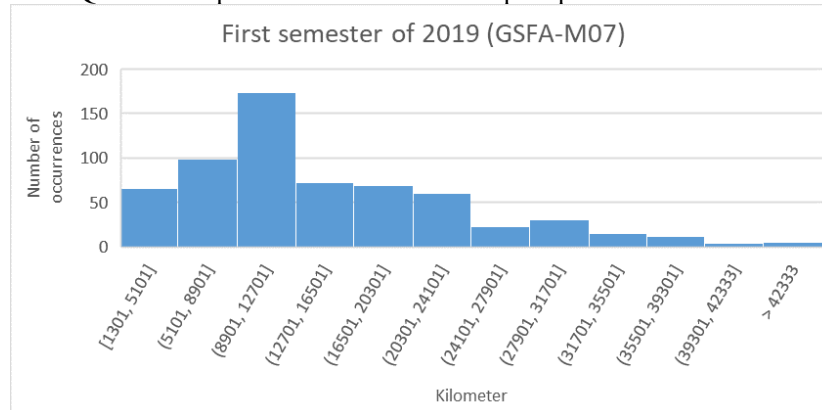
Figura 10: Quilômetros percorridos versus ordens de reparo



Também, com foco nos quilômetros percorridos em que uma ordem de reparo é iniciada, foi realizada uma análise utilizando um filtro para isolar os códigos de problemas em uma determinada faixa de tempo. Isso porque, a data mais antiga do pedido de reparo é janeiro de 2019 e os dados foram extraídos em janeiro de 2020.

Desta forma, foi considerado para análise para dividir o intervalo de tempo em semestres de 2019. Primeiro, foram analisados os dois códigos de defeito com as maiores ocorrências, ou seja, o GSFA-S22 e o GSFA-M07, representando juntos mais de 23% do total de defeitos. No primeiro semestre, esse defeito teve 1314 ocorrências, com maior incidência entre 7412Km e 11111Km. No segundo semestre, o código de defeito GSFA-S22 aumentou de 1314 ocorrências para 5313 ocorrências. Esse aumento representa pouco mais de 400%. Além do aumento visível é possível notar que o ápice do problema ocorreu em um número maior de quilômetros percorridos. Como mostrado na Figura 11.

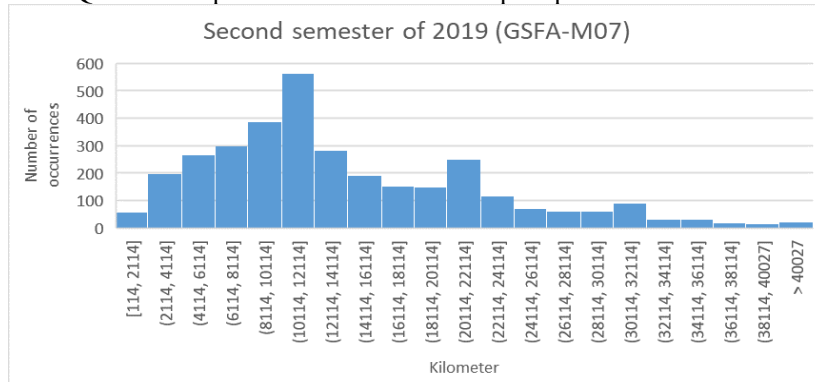
Figura 11: Quilômetros percorridos na ordem de reparo para GSFA-M07 no I semestre



Analisando o código de defeito GSFA-M07, foram encontradas 627 ocorrências no primeiro semestre de 2019, com maior incidência entre 8901Km e 12701Km.

No segundo semestre do ano de 2019, o código de defeito GSFA-M07 aumentou de 627 para 3.304 ocorrências, o que representa um aumento de pouco mais de 520%. Como no caso do defeito GSFA-S22, o pico de maior ocorrência foi em maior número de quilômetros percorridos. Como é possível visualizar na Figura 12.

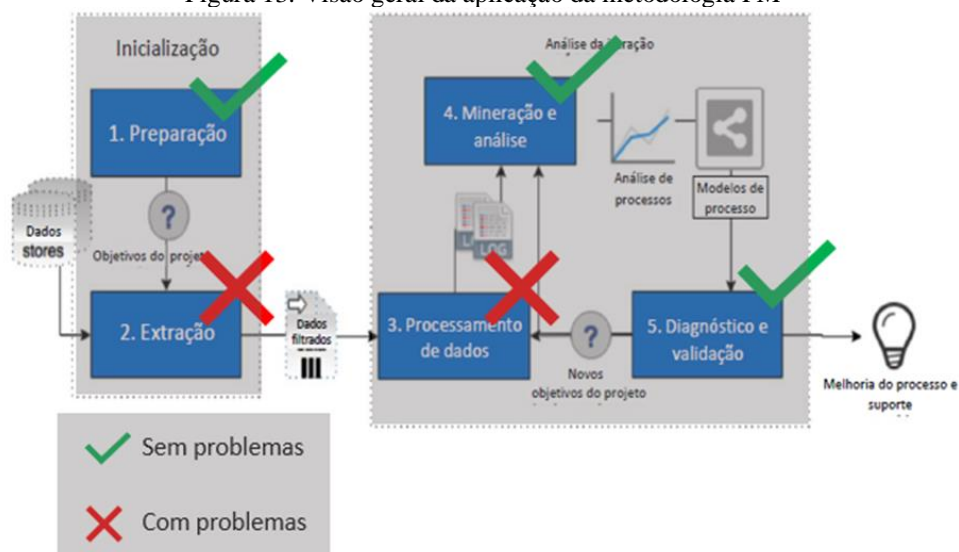
Figura 12: Quilômetros percorridos na ordem de reparo para GSFA-M07 no II semestre



Diagnóstico e validação

A avaliação desta proposta feita foi realizado com base na visão geral do estudo de caso sobre o processo de suporte de *hardware* da IBM, presente em Van Eck, Leemans, Van der Aalst. (2015). Dessa forma, a visão geral da mineração de processos a metodologia PM² está presente no fluxo da Figura 13.

Figura 13: Visão geral da aplicação da metodologia PM²



Primeiramente, na etapa de preparação, notou-se que, a determinação dos objetivos, recursos necessários, custos envolvidos, comunicação com o setor de pós-venda e incidentologia, usando a base de dados foi realizada de forma satisfatória. A realização da análise dos dados foi desenvolvida após a aprovação de gestores do setor de qualidade de pós-venda e de qualidade, que confiaram no projeto e compreenderam seus objetivos. Após isso, a realização do planejamento ocorreu sem problemas e os resultados atenderam aos objetivos propostos.

A dificuldade nas etapas de extração e processamento dos dados se deu em função da busca pelas modificações ideais dos dados brutos. O desenvolvimento de diversas bases de dados, em VBA, para cada planilha de veículos, mostrou-se um processo demorado, não apenas pela programação em si, mas também pela demora na execução dos programas pelos computadores da empresa.

Certos programas VBA necessitaram de mais de 48 horas de execução, o que restringiu de certa forma o avanço do projeto. Talvez, possa-se dizer que esta etapa seja a mais crítica da aplicação das técnicas propostas, visto que demanda conhecimentos prévios em técnicas de estruturação de dados, linguagens de programação ou avançados em manipulação de planilhas. Além disso, necessitam de computadores com um grau de processamento compatível para a utilização dos softwares, já que em projetos com uma maior quantidade de dados, a etapa de adequação demandará um alto número de modificações e, conseqüentemente, um tempo maior para execução dos programas.

Destaca-se a velocidade do processo de mineração. Atualmente, os responsáveis da empresa envolvidos na análise de reclamações das concessionárias, levam em torno de 2 a 3 dias para formatar a base de dados e gerar os gráficos e análises necessárias, através dos recursos existentes como o software MS-Excel®. Através do uso do software de mineração DISCO® e feitas as devidas adequações na base de dados, a geração de gráficos e estatísticas levou em torno de 4 horas. Ou seja, uma redução aproximada de 80% do tempo atual.

Assim, pode-se dizer que a aplicação da metodologia proposta no setor de concessionárias automobilística, tem grandes chances de utilização e ganhos futuros.

RESULTADOS E CONCLUSÕES

A aplicação da metodologia em uma indústria automobilística destacou os passos necessários para a realização de uma mineração de processos no pós-venda. Ela mostra que, embora ainda não utilizada, a mineração de processos pode sim ser aplicada no pós-venda e seus benefícios são reais.

Quanto à etapa de extração dos dados, as dificuldades se concentraram na busca pela base ideal. Na empresa onde a abordagem foi aplicada, existem diversas bases de dados que envolvem informações de pós-venda, cada uma com características específicas, como: avaliação de clientes em redes sociais, avaliação de atendimento em concessionárias realizadas pela ANFAEVA (Associação Nacional dos Fabricantes de Veículos Automotores), dados retirados de um aplicativo para celulares para clientes da empresa, entre outros. Incluindo a base de dados com informações relacionadas à solução de problemas realizados nas

concessionárias, a qual foi utilizada nesse estudo. A escolha da base ideal se deu ao fato de que a melhor se adaptaria à mineração de processos, compondo os conceitos básicos da mineração (instâncias, atividades e timestamp), pois sem essas características, ficaria inviável a aplicação. Tendo escolhida a base de dados, não houve problema na extração dos dados e a aplicação dos métodos, pôde seguir seus demais passos.

A realização da mineração de processos em si revelou informações importantes para a empresa. A sequência padrão dos processos do pós-venda, os intervalos de tempo médio e máximo entre os processos, características dos clientes, como os produtos se comportaram na mão dos clientes e quantas vezes cada veículo precisou de reparos em concessionárias, foram pontos importantes.

Os dados obtidos foram suficientes para a realização do projeto e notou-se que a mineração de processos pôde, de fato, auxiliar na busca e resolução dos problemas de gerenciamento de dados do pós-venda da empresa. Clarificando assim, os processos e identificando o desempenho de cada método dentro do setor, como por exemplo, o fato das reclamações começarem em média dez meses após a compra do veículo por parte do cliente. A observação de que em média a segunda reclamação ocorre aproximadamente cinco meses após a primeira, sendo esse fato evidenciado também nos mapas de processo gerados.

A utilização do software DISCO® provou-se bem-sucedida, através da geração de gráficos e tabelas de fácil compreensão, além da grande quantidade de informação produzida. Os dados modificados também mostraram que as estruturações realizadas foram adequadas para as análises no software DISCO® e fundamentais para o encontro dos resultados.

A etapa de aprofundamento dos dados demandou estudos adicionais realizados pelo próprio software DISCO®, onde através de filtros nos dados, foi possível alcançar as análises adicionais propostas na etapa de preparação. As informações encontradas puderam esclarecer questões relacionadas à problemas encontrados pelos clientes e características específicas de revisões de veículos. Essa etapa permite que além das informações resultantes da mineração de processos, novas características sejam descobertas e estudadas em relação com o pós-venda.

Por fim, avaliou-se que o projeto de aplicação atendeu aos objetivos propostos na fase de planejamento. Alguns pontos se destacam em relação das abordagens: a empresa pôde melhorar o tratamento dos dados de pós-venda gerados; foi possível obter um maior conhecimento sobre os processos que envolvem o pós-venda da empresa, inclusive os processos envolvendo clientes; foi possível avaliar o desempenho de cada processo, analisando suas frequências e durações; foi possível também possibilitar uma melhor preparação para o

atendimento ao cliente, já que os resultados da aplicação conferem as concessionárias informações chave para a solução de problemas encontrados.

De maneira geral, o projeto de aplicação de uma metodologia de mineração de processos, no pós-venda da indústria automobilística, mostrou que muitas informações podem ser descobertas e a abordagem proposta, pode ser de grande ajuda para o gerenciamento do setor. Através das análises realizadas, pode-se perceber as características chave dos processos do pós-venda na empresa. Bem como, informações relevantes em relação à procura dos clientes às concessionárias e à solução dos problemas por parte dos técnicos do pós-venda.

Pode-se citar ainda que a abordagem proposta, embora com pequenos problemas em sua execução, mostrou-se factível não só para o setor de pós-venda automobilístico, mas para qualquer tipo de indústria. Sobretudo, é necessário apenas decidir previamente quais características adicionais serão estudadas, bem como a análise das bases de dados disponíveis e a adequação de dados para a aplicação de mineração de processos.

A metodologia proposta mostrou-se também de fácil aplicação, tendo como maior dificuldade a etapa de estruturação dos dados, que demandará conhecimentos técnicos de programação para sua concretização. E por fim, a metodologia proposta mostrou-se útil, já que os resultados encontrados pela aplicação foram de grande valia para a empresa, que agora possui mais uma ferramenta para atender às reclamações de clientes. Assim, agiliza a rapidez do atendimento e a qualidade dos serviços, além de uma preparação prévia, por parte da empresa, em relação à logística de peças para reposição em concessionárias.

REFERÊNCIAS

AHN, Jin Sook; SOHN, So Young. Customer pattern search for after-sales service in manufacturing. *Expert Systems with Applications*, v. 36, n. 3, p. 5371-5375, 2009.

MAITA, A., CORRÊA MARTINS, L., PAZ, C., RAFFERTY, L., HUNG, P., PERES, S., & FANTINATO, M. (2018). A systematic mapping study of process mining. *Enterprise Information Systems*, 12, 505–549.

SONG, M. S.; VAN DER AALST, Wil MP. Towards comprehensive support for organizational mining. *Decision Support Systems*, 2008.

VAN DER AALST, W. M. P. Process mining in the large: a tutorial. In: *Business Intelligence*. Springer International Publishing, 2014. p. 33-76.

VAN DER AALST, W. M. P. *Process mining: data science in action*. Springer, 2016.

VAN DER AALST, W. M. P., WEIJTERS, Ton; MARUSTER, Laura. Workflow mining: Discovering process models from event logs. *IEEE Transactions on Knowledge and Data Engineering*, v. 16, n. 9, p. 1128-1142, 2004.

VAN ECK, M. L.; LU, X.; LEEMANS, S. J. J.; VAN DER AALST, W. M. P. PM2: A process mining project methodology. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, v. 9097, p. 297–313, 2015.

VAN ECK, Maikel L. et al. A Process Mining Project Methodology. In: *International Conference on Advanced Information Systems Engineering*. Springer, Cham, 2015. p. 297-313.