

ANÁLISE DE DADOS PARA APOIO À TOMADA DE DECISÃO: ESTUDO EM UMA PEQUENA EMPRESA DO RAMO FARMACÊUTICO

ILTON LUIZ ASSI

UNIVERSIDADE PRESBITERIANA MACKENZIE (MACKENZIE)

ALMIR MARTINS VIEIRA

UNIVERSIDADE PRESBITERIANA MACKENZIE (MACKENZIE)

FERNANDO NASCIMENTO ZATTA

HMZ CONSULTORIA DE NEGÓCIOS, GOVERNANÇA, AUDITORIA E CONTROLES

ANÁLISE DE DADOS PARA APOIO À TOMADA DE DECISÃO: ESTUDO EM UMA PEQUENA EMPRESA DO RAMO FARMACÊUTICO

INTRODUÇÃO

Dentre os Objetivos de Desenvolvimento Global Sustentável para 2030 da Organização das Nações Unidas (ONU), a 4ª diretriz recomenda “Educação de Qualidade”, e a 17ª diretriz recomenda “Parcerias e Meios de Implementação”. Tais diretrizes são replicadas no relatório do Report GEM Brazil 2019 (ONozato *et al.*, 2019).

No mesmo contexto, o relatório do Sebrae – SP de 2018, destaca que as Pequenas e Médias Empresas (PMEs) contribuem com mais de R\$ 635 bilhões por ano à economia estadual, apresenta crescimento de 5,2% em relação aos resultados de 2016 referentes aos últimos dados publicados pela instituição. Esse montante de valor econômico agregado é distribuído nos seguintes setores: serviços (41%), comércio (37%), indústria (12%), construção civil (7%) e agropecuária (3%) (Duchateau; Moreira, 2018). As PMEs no estado de São Paulo (SP) constituem 98% das empresas do estado, geram 50% dos empregos, são responsáveis por 39% dos salários pagos e 27% do PIB estadual conforme dados do Sebrae – SP (Duchateau; Moreira, 2018).

A taxa de sobrevivência das PMEs, considerando o setor do comércio no estado de SP é de 76,3%, ou seja 1 em 4 empresas fecha antes de completar 2 anos de existência, assim é premente a necessidade de melhor gestão e maior aprendizagem organizacional meio à alta competitividade e incertezas mercadológicas atuais, objetivando a diminuição dessa alta taxa de mortalidade organizacional (Duchateau; Moreira, 2018).

Os estudos sobre Desenvolvimento Organizacional foram impulsionados na década de 1990, suportados pelo processo de globalização, exponencial crescimento tecnológico, que demandaram novas estratégias de gestão em um ambiente de alta e constante mudança (Hiatt; Creasey, 2012). Os conceitos de *Business Intelligence (BI)* e de *Análise Preditiva de Dados (APD)* são potencializados pelo aumento sem precedentes na criação de dados, avanços no desenvolvimento de softwares e consequente informatização da organização, maior competitividade e necessidade de negócios sustentáveis no longo prazo (Siegel, 2017).

No princípio, tais conceitos eram acessíveis somente para grandes organizações, pois o custo de implementação e de conhecimento das técnicas necessárias eram fator limitante, fato este superado atualmente com a ampla e capilarizada divulgação dos conceitos de *BI* e *APD* (Siegel, 2017; Turban *et al.*, 2008). Assim é selecionada uma pequena empresa do setor de comércio sendo uma farmácia situada em uma cidade no interior de SP como fonte de obtenção de dados e oportunidade de aplicação.

O objetivo deste estudo é contribuir para o aumento da competência dos gestores na identificação de melhores estratégias comerciais para racionalizar os processos de tomadas de decisão baseadas nos resultados da pesquisa. A unidade de estudo considera o contexto organizacional constituindo por uma empresa familiar de pequeno porte, a qual tem como fator limitador de atuação, proposições e ações que sejam determinadas pelo desejo, motivação, compreensão e cultura organizacional criada por seu fundador (Schein, 2017).

Assim, assume-se a necessidade de profundo alinhamento ao cliente, o que requer saber o que fazer, quem envolver, como engajar, até onde alterar, quanto tempo o processo de mudança deve perdurar, qual intensidade reforçar o processo para institucionalizar as mudanças propostas (Hiatt; Creasey, 2012; Kotler; Pfoertsch; Spohnholz, 2021). Diante deste contexto, este estudo foi iniciado com realização de análises para geração de *insights* importantes, submetidos ao fundador para a sua possível aprovação, e assim exploradas e implementadas de maneira gradual (Knafllic, 2019).

É necessário que o início do negócio seja desenhado como um trabalho de designer, definindo claramente os desejos do “cliente” a partir dados, e o que ele faria com tais dados para proposição de uma função prática; minimizar as barreiras de compreensão impostas pelo ERP, ou seja, construir uma visualização amigável e compreensível para “ele”, valorização de uma forma viável de interpretação, tal como a apresentação dos dados, gerando lógicas, com resultados práticos que proporcionem uma via de aceitação à mudança planejada e alteração do status quo organizacional (KNAFLIC, 2019).

Para a execução da mudança planejada, demanda-se consciência, desejo, conhecimento, habilidades e institucionalização de novos processos. É crucial que esse ciclo seja realizado de forma sequencial pois, as fases são interdependentes, sendo que a não realização do antecedente desabilita automaticamente a implementação da proposição consequente, conceito importante para a interpretação das relações causais do MBA (HIATT; CREASEY, 2012; TAN; STEINBACH; KUMAR, 2009). Importante ainda é considerar o estilo de liderança exercido na organização, no caso em particular em que há predominância da liderança transacional, ou seja, a prática do comando e controle é muito presente (BASS; BASS, 2008).

Para a implementação do projeto é fundamental a análise do exercício da liderança na organização, ou seja, os fatores relacionais entre líder e liderados, a consequente aceitação da distribuição da liderança em situações complexas, níveis de autonomia outorgada e se esta existe, distribuição e compartilhamento da liderança, o comportamento e cultura da organização (LICHTENSTEIN *et al.*, 2006; ROBBINS; JUDGE; SOBRAL, 2011). O papel do “consultor”, neste caso, é compreender esse ambiente com o intuito de contribuição com menor interferência possível, ou seja, o desejo de inovação técnica ou processual do consultor não deve sobrepor o desejo do proprietário ou fundador da organização na implementação de qualquer mudança. Tal afirmação é desafiadora pois, psicologicamente nosso comportamento habitual, em muitos casos é irracional, mas acaba sempre com uma justificativa, esses comportamentos são amplamente determinados por forças inconscientes, forjadas por nosso passado, características comportamentais que geralmente guiam nossos processos decisórios enfatizando intuições e mitigando processos racionais e sistemáticos, o autoconhecimento fundamental para qualquer consultor ou estudioso quando em um ambiente, pois este ambiente é de seu total controle (BAZERMAN; MOORE, 2014; KETS DE VRIES; BALAZS, 2011).

Para um processo decisório ser gerido com maior racionalidade, as seguintes etapas devem ser cumpridas: (i) Definição do problema; (ii) Identificação de critérios relevantes para o processo de decisão; (iii) Ponderar os critérios estabelecidos anteriormente de acordo com seu valor relativo à decisão; (iv) Gerar alternativas ou cursos de ação; (v) Gerar uma classificação de cada alternativa segundo cada critério e por fim identificar a solução ideal e aplicá-la. O que a princípio até parece fácil e lógico, mas raramente é aplicado em sua totalidade (BAZERMAN; MOORE, 2014).

Somos influenciados por nossas emoções, subjetividades que são carregadas por nossos vieses que são dominados por forças inconscientes, assim construímos heurísticas, que resultam em julgamentos de baixa qualidade pois, utilizamos quantidade insuficiente de dados e informações concretas. Possuímos conscientização limitada, e invertemos preferências, tudo isso ofusca ainda mais nossa capacidade cognitiva, gerando decisões equivocadas na maioria das vezes (BAZERMAN; MOORE, 2014; KETS DE VRIES; BALAZS, 2011). Assim, o cenário deste estudo demanda a utilização de duas ferramentas de uma lista conceitual e de práticas de BI e APD, provocando mudança do “*status quo*” organizacional, sem grande impacto no comportamento organizacional, ou seja, ser eficiente e eficaz.

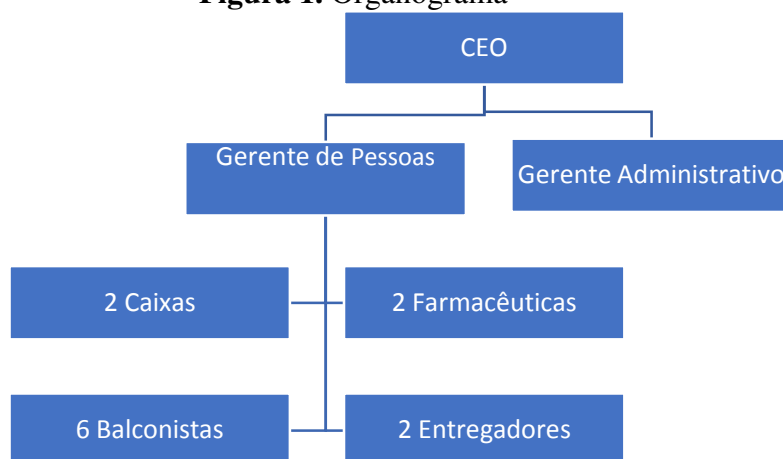
CONTEXTO INVESTIGADO

O proprietário é ex-funcionário do ramo farmacêutico com carreira iniciada na função de balconista, tornando-se empreendedor em 1984, com 18 anos de idade, auxiliado financeiramente pelo seu pai na aquisição da farmácia. Possui formação educacional com segundo grau incompleto. Verifica-se que sua experiência profissional, dedicação, resiliência e relacionamento são fatores críticos de sucesso para a sustentabilidade e longevidade da sua organização. O desenvolvimento organizacional ocorrido a partir de 1990 vem demandando liderança, maiores e melhores conhecimentos, “*skills*”, e competências na difícil tarefa de posicionamento organizacional no mercado.

No tocante ao tamanho e aos desafios, tais como, problemas econômicos, mercado altamente regulamentado, globalização e aumento exponencial da exigência por melhor gestão e liderança, levou o proprietário à decisão de se integrar a uma rede associativa de farmácias em 2010, a qual se mantém nos dias atuais. A partir dessa associação foram oportunizadas novas formas de gestão mais modernas, padronizadas e efetivas ao negócio. O cerne da união de pequenos negócios e redes associativas conta atualmente com 270 associados, cujo maior objetivo é gerar maior força de negociação na relação com o mercado distribuidor de medicamentos, e ainda oferecer um maior leque de serviços. Como resultado procura inserir seu negócio em um novo patamar de atuação e ações inovadoras (CHRISTENSEN; RAYNOR; MCDONALD, 2015).

Tal processo disruptivo demanda nova fase evolutiva pessoal e profissional de seu proprietário, pois o mesmo necessita entender, implementar e adaptar sua organização às mudanças requeridas e as exigidas pela rede, para estar inserido em um novo ambiente de negócios. Com faturamento atual de aproximadamente R\$ 480.000,00 mês, com mais de 6.000 itens cadastrados, constituídos por medicamentos éticos, genéricos e similares, acessórios, produtos de higiene e alguns alimentos, a organização gera uma rentabilidade bruta de média de 29%. A composição de recursos humanos da organização contém 15 pessoas, conforme organograma apresentado na figura 1.

Figura 1. Organograma



Fonte: elaborado pelos autores.

O estilo de liderança é centralizado no proprietário que como causa gera sobrecarga de estudo e intensa interação de alto nível sobre conflito entre proprietário, gerente e funcionários. Este estilo de gestão depende de comunicação e capacidade de escuta ativa como fator crítico de sucesso. Requer uma explanação ampla e detalhada, sobre o que será feito, como será feito, qual o nível de interferência sobre as tarefas diárias, e resultados esperados para que a

implementação da mudança planejada seja bem sucedida (HIATT; CREASEY, 2012; KOTLER; PFOERTSCH; SPONHOLZ, 2021).

DIAGNÓSTICO DA SITUAÇÃO-PROBLEMA

Dados são fatos que registram o passado e o presente das organizações, atualmente com menor custo e maior rapidez. Dada a acessibilidade às ferramentas computacionais, os dados podem inferir ações preditivas quando tratados e modelados adequadamente, e consequentemente compreendidos pelo gestor. Entretanto, para cada aplicação são necessárias duas definições: (1) O que será predito; e (2) O que será feito com isso (SIEGEL, 2017).

Certamente cada gestor possui diferentes capacidades ou competências, estilos de gestão e de compressão dos *insights* gerados pela aplicação de APD e BI, que de forma causal gerarão mudanças significativas ou não para a organização. Assim, essa mudança necessita ser iniciada no íntimo do gestor, ou seja, necessitando entender com profundidade o que será mudado e qual será o benefício gerado para o negócio.

Em relação ao estudo realizado na farmácia, optou-se pela utilização de um sistema de análise de dados utilizando um banco de dados SQL PBI (TURBAN *et al.*, 2008), e posteriormente uma ferramenta de associação *MBA*(*Market Basket Analysis*) (LANZ, 2015), utilizando dados do banco de dados do ERP da farmácia.

A construção da base relacional de dados

A definição de iniciar o estudo por esse processo foi determinada por três motivos: (i) utilizar de forma mais efetiva e ágil o sistema de ERP da farmácia; (ii) causar um impacto positivo na compreensão dos dados, utilizando as formas gráficas do PBI; e (iii) aprender como os dados estavam dispostos, como extraí-los e utilizá-los fora do ERP.

A partir desse momento, foi apresentado ao proprietário o que seria realizado e que benefício essa base lhe traria. Após a criação do *dashboard*, a apresentação ao proprietário foi acompanhada de grande sucesso, pois o motivou a dedicar mais tempo do seu dia a dia para maior compreensão e interpretação dos gráficos combinados e devidamente relacionados em apenas uma tela, visualmente mais agradável, e principalmente sem ter que “navegar” pelo sistema ERP de maneira individualizada, o que provoca maior tempo do gestor.

Com a “simples” construção relacional do banco de dados da farmácia com a ferramenta PBI como, vendas, compras, formas de pagamentos e recebimentos, margem de lucratividade geral e por grupo de produtos, foram gerados *insights* sobre o que é prioridade estratégica para a organização de forma lógica e factual, com apenas alguns clicks.(TUCCI, 2017; WOODY *et al.*, 2016).

Simple dashboard

Utilizando dos relatórios contidos no ERP, o primeiro objetivo foi alcançado com a criação de motivação do proprietário com *insights* práticos no dia a dia da gestão, sendo que para tal, o mesmo, prontamente, ajustou sua agenda, reservando maior tempo para o processo de aprendizagem sobre a interpretação dos dados combinados em apenas uma tela, coesa, com informações precisas e úteis (SIEGEL, 2017; TUCCI, 2017; WOODY *et al.*, 2016).

O principal fator motivacional reside no fato de que o PBI possibilita acesso com facilidade, sendo atualizado de forma automática, com as informações desejadas, selecionadas previamente e combinadas, e sem restrições para que estas sejam vistas de formas individualizadas, caso necessário, para objetivos estratégicos específicos (TUCCI, 2017).

Com os resultados apresentados, o projeto recebeu autorização de andamento, uma vez que os benefícios até então apresentados eram fatos relevantes ao negócio como um todo, facilitando a criação de valor ao aprendizado organizacional e ao gestor, pois amplia a visão organizacional de um estágio micro para macro, ou seja, de uma gestão processual para uma gestão de aplicações conceituais.

INTERVENÇÃO PROPOSTA

Um passo à frente – Regras de Associação (MBA)

A etapa para delineamento e organização dos dados foi complexa, pois o ERP da organização não continha relatórios que contemplassem a formatação dos dados necessários para a aplicação do algoritmo MBA – APRIORI, de forma que nem os pesquisadores, bem como o proprietário tinham conhecimento de como obter os dados do sistema. Para tal, foi obtido auxílio do departamento de Tecnologia da Informação (TI) da rede, que foi contemplada com explicação detalhada de minha necessidade e consequente orientação da extração do banco de dados contido no ERP para posterior modelagem. Mesmo assim, o formato do banco de dados que foi extraído e recebido do departamento de TI, não continha o formato ideal o que demandou muito estudo de modelação o para sua correta utilização.

Em relação à composição do banco de dados, foi estabelecido para este estudo um corte temporal de 4 meses, relativo aos meses de agosto a novembro de 2022, no qual foi gerado um conteúdo que proporcionou uma tabela com mais de 28 mil linhas de transações, para que fossem contemplados todos os cupons fiscais do período determinado, contendo as seguintes informações: data, número do cupom fiscal e produtos neles registrados. A modelagem para a aplicação do algoritmo de associação MBA - APRIORI constituiu em: (i) relacionar cada item registrado (produto) com seu número correspondente do cupom fiscal; (ii) exclusão de todos os caracteres especiais (+, /, -, “,”, &); (iii) espaços e linhas em branco; (iv) produtos repetidos em um mesmo ticket; informações sem sentido como, por exemplo, registro de uma letra M como produto. Tal modelagem gerou um Data Frame => df = 9.366 linhas e 19 colunas, representando, 9.366 transações com até 19 produtos na maior lista.

Desta forma, concluiu-se com o delineamento do estudo, relacionando os bancos de dados (PBI), e posteriormente definindo temporalmente o período de análise do (MBA), avaliando o conteúdo e o modelando para a aplicação correta do algoritmo. A próxima seção destaca os resultados e *insights* obtidos.

RESULTADOS OBTIDOS

Resultados Práticos do *Dashboard*

Na construção da base de dados relacionais, o estudo apresenta constatações práticas que podem proporcionar ao proprietário uma visão condensada e mais adequada dos principais resultados financeiros da empresa com a seguinte estrutura:

1. Vendas
 1. Resultados diário, mensal, anual
 2. Resultados por grupo de produtos
 3. Rentabilidade por grupo, por mês e por ano
 4. Forma de pagamento
2. Clientes
 1. Percentual de fidelização
 2. Índice de Inadimplência

3. Compras
 1. Valores de compras por mês, grupo de produto
 2. Distribuição de itens comprados
 3. Ranking dos fornecedores
4. Cartões informativos
 1. Vendas totais
 2. Custos Totais
 3. Rentabilidade Bruta

Insights

O processo de relacionamento do banco de dados em um *dashboard*, fornece e estimula a melhora do processo de tomada de decisão potencializada por meio de interpretações combinadas utilizando apresentações gráficas de agrupamentos (produtos, clientes com cartão fidelidade, rentabilidade por grupo) e ajustes de foco de gestão em busca de maior e melhor efetividade e resultado financeiro organizacional. Também cria novas formas de análises, como, por exemplo, na análise das formas de pagamento. O PBI também expõe uma falha de sistema que apresentou 50% dos dados como não identificados no sistema (ERP). Essa falta de ocorrência pode ser um problema na identificação das formas de pagamento no sistema ou causada por erro humano.

O PBI ainda gerou uma proposta, que não foi acatada, a qual foi a inserção de uma tabela de fluxo de caixa no *dashboard*. Esta rejeição foi feita por o proprietário possuir tabela (Excel®), utilizada em separado, a qual ele mesmo atualiza sua posição diariamente, para manter o controle individual dessa informação.

A visualização dos grupos de produtos e percentual de lucratividade, por meio da utilização da função DAX do PBI, apresenta uma oportunidade de melhoria estratégica nas vendas de balcão, tratando-se de uma mudança a ser implementada acompanhada de processo de comunicação para a adequação de procedimentos de venda interna com foco sobre qual grupo deve ser imputada maior energia. Outro ponto de melhoria apresentado com a construção do *dashboard* foi a gestão dos cartões de fidelidade. Este item apresentou aderência de 50% dos clientes, o que gera oportunidade de ampliação e melhoria de gestão.

Em relação às compras, essa operação passou a ser realizada de forma mais adequada, pois o maior índice de compra é do fornecedor “rede”, que proporciona os maiores índices de desconto e prazos para pagamento. Por fim, a primeira etapa da proposta de análise de dados da farmácia foi cumprida com sucesso e de acordo com a entrega objetivada. A construção de um *storytelling* com os dados pode valorizar o estudo e motivar o proprietário no processo de aprendizagem analítica de dados (KNAFLIC, 2019). A seguir, tem-se um processo preditivo de análise.

Aplicação do Algoritmo Apriori (MBA)

O MBA é um algoritmo de importância fundamental para construção de recomendações de compras, indicando a possibilidade de realização de vendas ativas, alteração geográfica dos produtos no interior da loja e a realização de promoções de produtos em conjunto de forma física e online (LANZ, 2015; TURBAN *et al.*, 2008). E ainda, estabelecendo regras de combinação, venda de “pacotes” de produtos que podem ser impulsionadas na organização contemplando necessidades específicas da organização com base em padrões de compras.

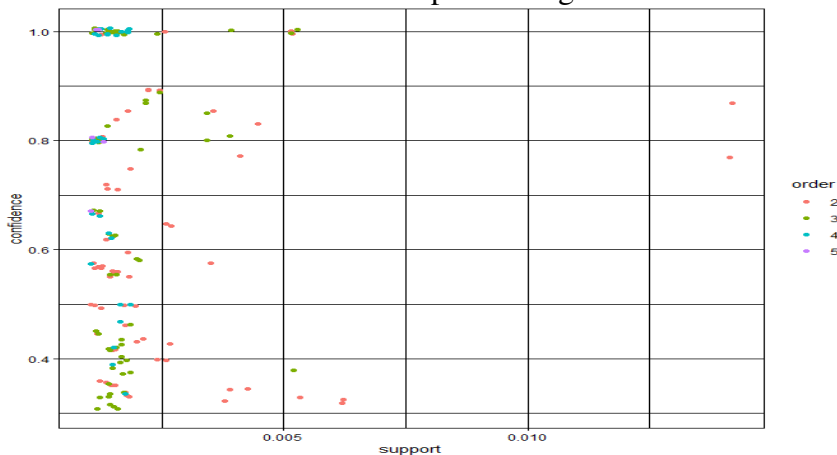
Primeira rodada do MBA

Essa rodada foi realizada como teste utilizando o software RStudio®, considerando apenas o mês de agosto de 2022. O objetivo foi a verificação de associações de alta frequência e consequentemente, com pouca significância estratégica para os especialistas. Como exemplo, uma relação de alta frequência de associação foi a “taxa de entrega”, item que é cobrado em vendas efetuadas por via eletrônica.

Mesmo essa relação sendo de conhecimento dos especialistas, isso proporciona importante informação, pois o item se revela uma importante fonte de receita e uma excelente forma de relação empresa cliente. Foi constatada sua viabilidade financeira, uma vez que os dados proporcionam a primeira decisão baseada em evidências e proporciona o cancelamento do projeto de terceirização das entregas. Porém, em vez do cancelamento, foi decidido que será realizada a troca das duas motos e fomentar contatos de entrega por meios eletrônicos para ampliação do serviço.

Por outro lado, das 169 regras criadas também constavam relações entre produtos injetáveis e materiais descartáveis obrigatórios para a tarefa, como seringas, agulhas escalpes entre outros tidos como materiais de consumo interno informação que não agrega valor. Assim como a taxa de entrega, esse tipo de material foi retirado do banco de dados após essa primeira análise. Com a finalidade de ilustrar as relações obtidas nesta fase, tem-se o gráfico 1 e a tabela 1, para apresentação das condições para sua criação (suporte), determinado em 0,001 e confiança mínima em 0,3.

Gráfico 1. Dispersão Regras



Fonte: Base de dados da organização.

O resultado da análise foi essencial para uma tomada de decisão baseada em evidências, na manutenção da frota de entrega em domicílio.

Tabela 1. Associações (MBA)

	lhs	rhs	support	confidence	coverage	lift	count
[1]	{FLUVIRAL DIA C 4 COMP}	=> {FLUVIRAL NOITE C 4 COMP}	0.001172333	0.8000000	0.001465416	545.920000	4
[2]	{FLUVIRAL NOITE C 4 COMP}	=> {FLUVIRAL DIA C 4 COMP}	0.001172333	0.8000000	0.001465416	545.920000	4
[3]	{FR HUGGIES SUPREMECARE 6 C 20}	=> {TAXA DE ENTREGA}	0.001172333	0.8000000	0.001465416	4.618613	4
[4]	{HUMULIN N NPH 100UI ML 10ML}	=> {GLIFAGE XR 500MG C 30 COMP}	0.001172333	1.0000000	0.001172333	29.929825	4
[5]	{TORAGESIC SL 10MG C 10 COMP}	=> {TAXA DE ENTREGA}	0.001172333	0.8000000	0.001465416	4.618613	4
[6]	{CONFORT MASTER ABSORVENTE GER C 20}	=> {TAXA DE ENTREGA}	0.001172333	0.8000000	0.001465416	4.618613	4
[7]	{IBUPRIL 300MG C 20 COMP}	=> {TAXA DE ENTREGA}	0.001172333	0.5714286	0.002051583	3.299009	4
[8]	{CATETER I.V.C AGULHA 16G}	=> {CATETER I.V.C AGULHA 18G}	0.001465416	0.8333333	0.001758499	406.190476	5
[9]	{CATETER I.V.C AGULHA 18G}	=> {CATETER I.V.C AGULHA 16G}	0.001465416	0.7142857	0.002051583	406.190476	5
[10]	{CATETER I.V.C AGULHA 16G}	=> {TAXA DE ENTREGA}	0.001172333	0.6666667	0.001758499	3.848844	4
[11]	{CELEBRIN XPE 120ML}	=> {TAXA DE ENTREGA}	0.001758499	0.6000000	0.002930832	3.463959	6
[12]	{STREPSILS C 24 PAST}	=> {TAXA DE ENTREGA}	0.001172333	0.5000000	0.002344666	2.886633	4
[13]	{SEKI XPE 120ML}	=> {TAXA DE ENTREGA}	0.001172333	0.4444444	0.002637749	2.565896	4
[14]	{LEITE APTAMIL 1 PREMIUM 800G}	=> {TAXA DE ENTREGA}	0.001758499	0.5454545	0.003223916	3.149054	6
[15]	{CATETER I.V.C AGULHA 18G}	=> {TAXA DE ENTREGA}	0.001172333	0.5714286	0.002051583	3.299009	4
[16]	{IBUPROFENO 50MG GTS 30ML}	=> {TAXA DE ENTREGA}	0.001465416	0.5555556	0.002637749	3.207370	5
[17]	{CAPTOPRIL 25MG C 30 COMP}	=> {HIDROCLOROTIAZIDA 25MG C 30 COMP}	0.001758499	0.5000000	0.003516999	25.462687	6
[18]	{DECONGEX PLUS C 12 COMP}	=> {TAXA DE ENTREGA}	0.001172333	0.5000000	0.002344666	2.886633	4
[19]	{LUFTAL GTS 15ML}	=> {LACTO PURGA C 6 COMP}	0.001465416	0.4166667	0.003516999	52.654321	5
[20]	{LABIRIN 16MG C 30 COMP}	=> {TAXA DE ENTREGA}	0.001465416	1.0000000	0.001465416	5.773266	5

Fonte: Base de dados da organização.

Segunda Rodada do MBA

Após o primeiro teste, foi utilizada a base de dados total com 28.413 registros considerando os 4 meses de coleta de dados. Nova modelagem foi necessária com a retirada dos seguintes itens, ainda, da tabela de Excel®. Foram retirados 201 registros com identificação M (sem significado para a análise), 2.051 registros de taxa de entrega, 556 registros identificados como agulhas e cateteres, retiradas conforme as justificativas apresentadas anteriormente.

Com a limpeza realizada, e a tabela no formato CSV, esta foi importada para o ambiente Jupyter para execução do algoritmo. Esse processo resultou em uma base de dados limpa com 9.366 linhas com 4.210 itens, com apenas 1 ticket com 19 produtos, conforme demonstrado na Figura 2.

Figura 2. Data Frame

```
In [3]: df.shape
```

```
Out[3]: (9366, 19)
```

Fonte: Base de dados da organização.

O Data Frame foi construído de acordo com o número de itens em cada linha, em que cada linha é uma lista com X itens (TAN; STEINBACH; KUMAR, 2009). Ainda no processo de configurações do `df = farmaci`, foram substituídos os NaN por zero(0) utilizando a função `replace` e `inplace` para que as alterações fossem realizadas no mesmo Data frame. É criada a *List Of Transaction*, e em seguida `row.values.tolist` e por fim, para eliminação de todos os zeros do Data Frame a função `filter (lambda(x: x!=0))`. (HAIR *et al.*, 2009).

Com a base de dados organizada na plataforma, utilizouse o pacote de *MLXTEND*, que contém o *MBA (APRIORI)*, e assim foi criada uma matriz para a aplicação do algoritmo. A aplicação do algoritmo apresenta as associações transacionais com as quantidades de itens em cada uma delas e seus respectivos e percentuais de frequência. São 3.343 transações com 1 item, e 1 transação com 19 itens.

Tabela 2. Quantidade por transações

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	19
33	24	15	97	44	28	147	76	39	30	19	10	6	8	5	3	1
43	61	21	1	3	3											
35	26	16	10	4	3	1,5	0,8	0,4	0,3	0,2	0,11	0,06	0,09	0,05	0,03	0,01
%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%

Fonte: Base de dados da organização.

A tabela 3, a seguir, apresenta as associações, tomando por base a frequência dos itens.

Tabela 3. Frequência MBA

```
In [16]: frequent_itemsets = apriori(df, min_support = 0.001, use_colnames = True)
frequent_itemsets.sort_values(by=['support'], ascending = False)

Out[16]:
```

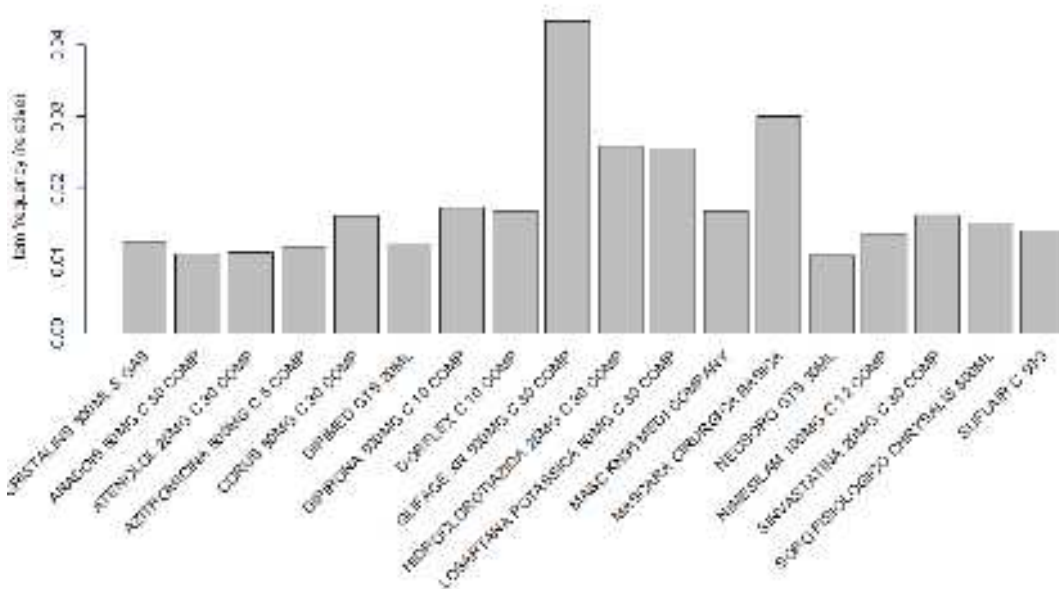
	support	itemsets
281	0.043242	(GLIFAGE XR 500MG C 30 COMP)
364	0.030322	(MASCARA CIRURGICA BASICA)
292	0.025945	(HIDROCLOROTIAZIDA 25MG C 30 COMP)
345	0.025625	(LOSARTANA POTASSICA 50MG C 30 COMP)
201	0.017403	(DIPIRONA 500MG C 10 COMP)
...
296	0.001068	(HUMECTOL D C 20 COMP)
301	0.001068	(IBUPRIL 300MG C 20 COMP)
308	0.001068	(JANUMET 50 1000MG C 56 COMP)
311	0.001068	(KIT PANT HIDROCAUTE SH 350ML COND 175ML)
596	0.001068	(SINVASTATINA 20MG C 30 COMP, GLIFAGE XR 500MG...

597 rows x 2 columns

Fonte: Base de dados da organização.

São relacionadas 597 linhas com 2 colunas nas quais o produto Glifage XR 500mg, representa 4,32% das associações considerando o suporte. O gráfico 2 retrata a informação anterior, seguido por venda de máscaras descartáveis seguido por dois produtos com a mesma frequência, Hidroclorotiazida e Losartana. Com as regras criadas e os limites de suporte e confiança determinados, optou-se por analisar as 39 regras pelo índice de LIFT. Este índice é de grande importância por revelar o peso de cada transação no conjunto de regras que forma gerado, conseqüentemente uma melhor avaliação estratégica do comportamento de associações do banco de dados (TAN; STEINBACH; KUMAR, 2009).

Gráfico 2. Frequência/Produtos



Fonte: Base de dados da organização.

O panorama demonstra que o output das regras (fator de prioridade, apresentado em ordem decrescente - strings na Figura 3), ordena as 39 regras no índice do Lift e o classifica do maior para o menor, conforme se apresenta na tabela 4.

Figura 3. String Lift Ranking (decrecente)

```
In [19]: rules.sort_values(by=['lift'], ascending = False)
Out[19]:
```

Fonte: Base de dados da organização.

A tabela 4 considera leva em conta três elementos fundamentais para o caso organizacional analisado neste estudo: informação de maior impacto, informação para ação e associações raras.

Informação de maior impacto

De acordo com a análise do especialista (proprietário), 85% dos produtos listados com grande poder associativo, tendem a ser relacionados ao programa governamental “farmácia popular”: tal informação tem grande importância pois, neste grupo de produtos, possui características especiais sendo: (1) social – é um programa de medicamentos de uso contínuo subsidiado pelo governo; (2) financeira - são medicamentos de custo muito baixo, exemplo (uma caixa de sinvastatina custa R\$ 3,00 para o cliente); (3) giro – são medicamentos que pela combinação das características anteriores apresentam alto giro comercial e podem gerar fidelização.

Contudo, esse tipo de produto é de pouca rentabilidade em relação ao número de transações que são realizadas, ou seja, servem como produto “chamariz” ou de entrada para outras transações, quando isso não ocorre, não produzem valor à organização.

Informação para ação

Em mãos de dados consistentes como as apresentadas na tabela 4 (p. 12), os processos decisórios passam a ser assertivos, efetivos, com maior alinhamento estratégico, ou seja, tornam-se racionais, eficientes e eficazes.

Assim, com o output do sistema de associação em mãos, o proprietário da farmácia já busca uma nova análise, sendo: quais desses produtos são efetivamente produtos do “grupo farmácia popular” ou produtos similares e ou genéricos? Apurando tal questão, torna-se factível a realização de ajustes nas vendas sem prejuízo ao cliente, mas com maior lucratividade para a organização.

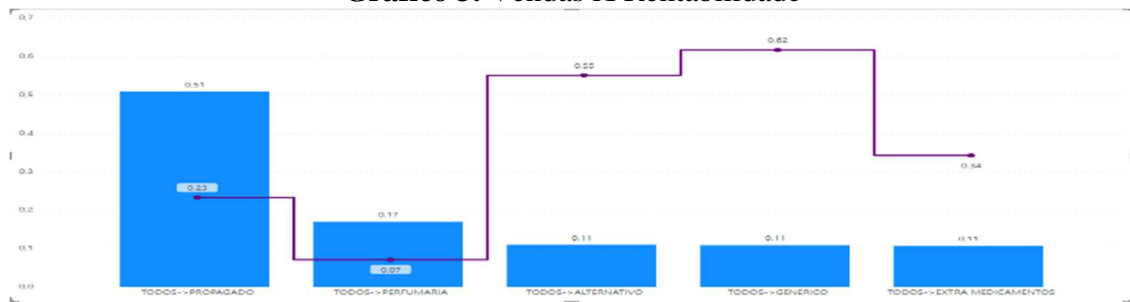
Associações raras

Com grande experiência profissional e com base nas associações apresentadas ao proprietário, nenhuma das 39 associações se mostrou como um fator de ruptura prática (rara), ou seja, o banco de dados não apresentou padrões fora dos que já era intuitivamente conhecido. Esse fato, certamente não minimiza a significância do estudo e dos insights proporcionados anteriormente, pelo contrário proporciona um “gostinho de quero mais”, como por exemplo uma nova modelagem dos dados, a utilização de outros algoritmos, a ampliação temporal. São propostas novas possibilidades que antes nem eram cogitadas para melhorar e ampliar as competências analíticas da organização.

O novo contexto interpretativo a partir da avaliação dos dados de associação MBA, instigamos ações preditivas, com execução factível em nível e grau que o especialista do negócio definirá visando manter o seu excelente atendimento aos seus clientes, mas com o objetivo claro e assertivo de potencializar suas vendas com produtos com maior valor agregado.

Informações geradas a princípio com o PBI, são reforçadas com a aplicação do MBA conforme dados constantes no gráfico 3 (a seguir) e na tabela 4 (p. 12).

Gráfico 3. Vendas X Rentabilidade



Fonte: Base de dados da organização.

Conforme o panorama apresentado, tem-se duas formas de análise que apontam para a mesma direção estratégica: a organização necessita de melhor gestão de suas vendas.

Tabela 4. Ranking Lifit

antecedentes	consequente	suporte	confiança	lift	alavancagem	convicção
11 (EQUIPO P NUTRIÇÃO ENTERAL BIOSANI)	(FRASCO P ALIMENTAÇÃO 300ML)	0,128%	75,0%	390.250.000	0,001278	3.992.313
12 (FRASCO P ALIMENTAÇÃO 300ML)	(EQUIPO P NUTRIÇÃO ENTERAL BIOSANI)	0,128%	66,7%	390.250.000	0,001278	2.994.875
0 (CLENIL HFA 200MCG 200 DOSES)	(AEROLIN SPRAY 200 DOSES)	0,128%	42,9%	64.741.935	0,001261	1.738.416
10 (DIMETILIV GTS 15ML)	(LACTO PURGA C 6 COMP)	0,150%	38,9%	40.925.094	0,001458	1.620.814
25 (LACTO PURGA C 6 COMP)	(SIMETICONA 75MG GTS 15ML)	0,288%	30,3%	35.078.652	0,002801	1.423.069
24 (SIMETICONA 75MG GTS 15ML)	(LACTO PURGA C 6 COMP)	0,288%	33,3%	35.078.652	0,002801	1.485.746
23 (LUFTAL GTS 15ML)	(LACTO PURGA C 6 COMP)	0,150%	29,8%	31.346.880	0,001447	1.410.709
33 (SINASTATINA 20MG C 30 COMP, HIDROCLOROTIAZID,,,	(CORUS 50MG C 30 COMP)	0,160%	44,1%	27.184.598	0,001543	1.760.432
31 (GLIFAGE XR 500MG C 30 COMP, CORUS 50MG C 30 C,,,	(SINASTATINA 20MG C 30 COMP)	0,139%	37,1%	22.589.610	0,001327	1.564.751
26 (GLIFAGE XR 500MG C 30 COMP, CORUS 50MG C 30 C,,,	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,182%	48,6%	18.720.988	0,001718	1.893.996
34 (CORUS 50MG C 30 COMP, HIDROCLOROTIAZIDA 25MG ,,,	(SINASTATINA 20MG C 30 COMP)	0,160%	29,4%	17.887.701	0,001512	1.393.373
37 (GLIFAGE XR 500MG C 30 COMP, LOSARTANA POTASSI,,,	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,246%	46,0%	17.729.877	0,002317	1.803.806
32 (SINASTATINA 20MG C 30 COMP, CORUS 50MG C 30 ,,,	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,160%	45,5%	17.519.641	0,00151	1.785.768
4 (CAPTOPRIL 25MG C 30 COMP)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,214%	43,5%	16.757.917	0,002008	1.723.328
27 (GLIFAGE XR 500MG C 30 COMP, HIDROCLOROTIAZIDA,,,	(CORUS 50MG C 30 COMP)	0,182%	27,0%	16.627.193	0,001706	1.347.339
29 (SINASTATINA 20MG C 30 COMP, GLIFAGE XR 500MG,,,	(CORUS 50MG C 30 COMP)	0,139%	26,0%	16.020.789	0,001301	1.329.420
13 (GLIBENCLAMIDA 5MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,342%	62,7%	14.510.385	0,003181	2.568.141
35 (GLIFAGE XR 500MG C 30 COMP, HIDROCLOROTIAZIDA,,,	(LOSARTANA POTASSICA 50MG C 30 COMP)	0,246%	36,5%	14.247.222	0,002283	1.534.641
16 (HUMULIN N NPH 100UI ML 10ML)	(GLIFAGE XR 500MG C 30 COMP)	0,117%	57,9%	13.388.694	0,001087	2.272.301
9 (CORUS 50MG C 30 COMP)	(SINASTATINA 20MG C 30 COMP)	0,352%	21,7%	13.203.947	0,003257	1.256.309
8 (SINASTATINA 20MG C 30 COMP)	(CORUS 50MG C 30 COMP)	0,352%	21,4%	13.203.947	0,003257	1.252.072
20 (HIDROCLOROTIAZIDA 25MG C 30 COMP)	(LOSARTANA POTASSICA 50MG C 30 COMP)	0,865%	33,3%	13.008.333	0,007983	1.461.563
21 (LOSARTANA POTASSICA 50MG C 30 COMP)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,865%	33,8%	13.008.333	0,007983	1.470.272
6 (CORUS 50MG C 30 COMP)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,545%	33,6%	12.932.261	0,005024	1.465.905
7 (HIDROCLOROTIAZIDA 25MG C 30 COMP)	(CORUS 50MG C 30 COMP)	0,545%	21,0%	12.932.261	0,005024	1.245.085

14	(GLIBENCLAMIDA 5MG C 30 COMP)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,139%	25,5%	9.824.740	0,001247	1.307.284
30	(SINVASTATINA 20MG C 30 COMP, CORUS 50MG C 30 ...)	(GLIFAGE XR 500MG C 30 COMP)	0,139%	39,4%	9.110.213	0,001236	1.578.652
22	(SINVASTATINA 20MG C 30 COMP)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,363%	22,1%	8.509.540	0,003204	1.250.037
3	(ATENOLOL 25MG C 30 COMP)	(LOSARTANA POTASSICA 50MG C 30 COMP)	0,235%	21,0%	8.176.667	0,002062	1.232.644
38	(SINVASTATINA 20MG C 30 COMP, GLIFAGE XR 500MG,,)	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	0,107%	20,0%	7.708.642	0,000929	1.217.569
28	(CORUS 50MG C 30 COMP, HIDROCLOROTIAZIDA 25MG ,,,)	(GLIFAGE XR 500MG C 30 COMP)	0,182%	33,3%	7.708.642	0,00158	1.435.138
19	(SINVASTATINA 20MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,534%	32,5%	7.508.418	0,004627	1.416.739
39	(SINVASTATINA 20MG C 30 COMP, HIDROCLOROTIAZID,,)	(GLIFAGE XR 500MG C 30 COMP)	0,107%	29,4%	6.801.743	0,000911	1.355.408
36	(LOSARTANA POTASSICA 50MG C 30 COMP, HIDROCLOR,,)	(GLIFAGE XR 500MG C 30 COMP)	0,246%	28,4%	6.566.621	0,002082	1.336.163
18	(MAL,ENALAPRIL 10MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,192%	26,9%	6.212.935	0,001613	1.308.221
2	(ATENOLOL 25MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,299%	26,7%	6.166.914	0,002505	1.304.671
1	(ARADOIS 50MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,288%	26,2%	6.062.136	0,002407	1.296.660
15	(HIDROCLOROTIAZIDA 25MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,673%	25,9%	5.995.610	0,005605	1.291.624
5	(CORUS 50MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,374%	23,0%	5.325.049	0,003035	1.242.968
17	(LOSARTANA POTASSICA 50MG C 30 COMP)	(GLIFAGE XR 500MG C 30 COMP)	0,534%	20,8%	4.817.901	0,00423	1.208.537

Fonte: Base de dados da organização.

CONTRIBUIÇÃO TECNOLÓGICA-SOCIAL

O presente estudo de análise e mineração de dados, considerado pequeno e aplicado a uma pequena empresa, apresenta um novo campo de prestação de serviço e aprendizagem, confirmado de forma empírica que independentemente do tamanho da organização podendo utilizar ferramentas de BI e APD para obter *insights* importantes para o desenvolvimento organizacional, potencializando o processo de aprendizagem individual e organizacional que abarca: (a) os aspectos técnicos para alunos de disciplinas de análise de dados, para gestores de organizações no foco do objeto da pesquisa, na ampliação de sua visão analítica de seu negócio, (c) na elucidação de possíveis ações estratégicas mais assertivas sobre as transações de sua organização; (d) a evolução do processo de decisão podem ser mais racionais e baseados em evidências e por fim (e), a afirmação que não há necessidade de alto volume de investimentos financeiros para a realização deste tipo de análise.

Para gestores, de forma visual entendendo que um *dashboard* proporciona pode proporcionar, “ascende”, impulsionando o desejo de maior compreensão dos fatos que ocorrem em sua organização, mesmo que em fase inicial o utilize apenas como uma ferramenta descritiva.

O exercício com o algoritmo Apriori (MBA), pode proporcionar tanto ao autor bem como ao proprietário, um salto qualitativo em relação a análise de dados fornecida para a tomada de decisão, uma vez que na modelagem dos dados, o aprendizado sobre a organização é exponencial, sendo que sua aplicação prediz condições e ou padrões na vida cotidianas da organização não observados de outra forma. Além disso oportuniza a melhora da qualidade e na quantidade de análise de dados gerados pela organização, melhor visão operacional da organização e melhores decisões.

Como limitação, é importante ressaltar que este estudo e seus achados, principalmente no tocante ao perfil do negócio, pode possibilitar pesquisas futuras, sugerindo assim a aplicação de outros algoritmos, como clusterização em um nível para além do PBI, a implementação de redes autônomas de comunicação com clientes entre muitas outras possibilidades.

Referências

BASS, M Bernard; BASS, Ruth. **The Bass Handbook of Leadership - Theory, Research and Managerial Applications**. 4ªed. New York: The Free Press, 2008.

BAZERMAN, Max H.; MOORE, Don. **O Processo Decisório: Tradução Daniel Vieira**. 8. ed. Rio de Janeiro: Elsevier/Campus, 2014.

CHRISTENSEN, Clayton M.; RAYNOR, Michael; MCDONALD, Rory. What is disruptive innovation?. **Harvard Business Review**, [s. l.], n. December, p. 1–10, 2015.

DUCHATEAU, Philippe Vedolim; MOREIRA, Marcelo. **Panorama dos Pequenos Negócios 2018 - Sebrae -SP**. [S. l.: s. n.], 2018. Disponível em: https://www.sebrae.com.br/Sebrae/Portal/Sebrae/UFs/SP/Pesquisas/Panorama_dos_Pequenos_Negocios_2018_AF.pdf. .

HAIR, Joseph F. Jr *et al.* **Análise Multivariada de Dados -Tradução Adonai Schlup Sant’Anna**. 6. ed. Porto Alegre RS: bookman, 2009.

HIATT, Jeffrey M; CREASEY, Timothy J. **Advance praise for Change Management : The People Side**. 2. ed. Loveland, Colorado USA: Prosci Research, 2012.

KETS DE VRIES, Manfred; BALAZS, Katharina. The shadow side of leadership. In: BRYMAN ALAN; COLLINSON DAVID; GRINT KEITH; JACKSON BRAD; UHL-BIEN MARY (org.). **Leadership Today**. Londres: SAGE Publications, 2011. p. 380–392.

KNAFLIC, Clole Nussbaumer. **Storytelling com dados: Um guia sobre visualização de dados para profissionais de negócios; Traduzido por João Tortello**. Rio de Janeiro: Alta Books, 2019.

KOTLER, Philip; PFOERTSCH, Waldemar; SPONHOLZ, Uwe. **H2H Marketing**. Cham: Springer International Publishing, 2021. *E-book*. Disponível em: <https://link.springer.com/10.1007/978-3-030-59531-9>.

LANZ, Brett. **Machine Learning with R**. Singapore: Springer Singapore, 2015. *E-book*. Disponível em: <http://link.springer.com/10.1007/978-981-10-6808-9>.

LICHTENSTEIN, Benyamin B. *et al.* Complexity leadership theory: An interactive perspective on leading in complex adaptive systems. **Emergence: Complexity and Organization**, [s. l.], v. 8, n. 4, p. 2–12, 2006.

ONOZATO, Erika *et al.* **Report GEM Brazil**. [S. l.: s. n.], 2019.

ROBBINS, Stephen P.; JUDGE, Timoty A.; SOBRAL, Felipe. **Comportamento Organizacional - Teoria e Prática no Contexto Brasileiro , Tradução Rita de Cássia Gomes**. 14. ed. São Paulo: Pearson Prentice Hall, 2011.

SCHEIN, Edgard H. **Cultura Organizacional e Liderança; Tradução Ailton Bomfim Brandão**. São Paulo: Atlas Ltda, 2017.

SIEGEL, Eric. **Análise Preditiva**. Rio de Janeiro: Alta Books, 2017.

TAN, Pang-Ning; STEINBACH, Michael; KUMAR, Vipin. **Introdução ao DATAMINING: Mineração de dados**. Rio de Janeiro: Editora Ciência Moderna Ltda., 2009.

TUCCI, Pierstefano. **Power BI Succinctly**. Morrisville USA: SyncFusion Inc., 2017. *E-book*. Disponível em: www.sunconfusion.com.

TURBAN, Efraim *et al.* **Business Intelligence: Um Enfoque Gerencial para a Inteligência do Negócio**. Porto Alegre RS: Artmed Editora SA, 2008.

WOODY, Buck *et al.* **Data Science with Microsoft SQL Server 2016**. [S. l.]: Microsoft Press, 2016.