

APLICAÇÃO PARA ANÁLISE DE PRODUTOS E PREÇOS DE CONCORRENTES, PARA E-COMMERCE DE MATERIAL ESPORTIVO



Igor Gomes Silva
8º período do Curso de Sistemas de Informação do
Centro Universitário de Patos de Minas (UNIPAM).
igorgomes@unipam.edu.br



Eduardo Henrique Silva
Professor do Curso de Sistemas de Informação do
Centro Universitário de Patos de Minas (UNIPAM).
eduardohs@unipam.edu.br

RESUMO

Este artigo descreve o desenvolvimento de uma aplicação Business Intelligence (BI) para análise de produtos e preços de concorrentes. O objetivo do desenvolvimento dessa aplicação é proporcionar dados e informações para auxiliar em tomadas de decisões e planejamento estratégico. As ferramentas e softwares utilizados para o desenvolvimento do projeto foram: Microsoft Word; MySQL; Power BI; PyCharm; Scrapy e Scrapinghub. Portanto, o estudo mostra que a aplicação desenvolvida pode ser uma ferramenta de apoio à gestão, guiando a empresa com informações que ajudem o gestor em suas tomadas de decisões.

PALAVRAS-CHAVE: *Business Intelligence; Extração de Dados; Planejamento Estratégico; Power BI.*

ABSTRACT

This paper describes the development of a Business Intelligence (BI) application for an analysis of products and prices of competitors. The objective of the development of this application is to provide data and information to help managers in decisions and strategic planning. The tools and software used for the development of the project were Microsoft Word; MySQL; Power BI; PyCharm; Scrapy e Scrapinghub. Therefore, the study shows that the application developed can be a support tool to the management, by guiding the enterprise with information that helps the manager in his decisions.

KEYWORDS: *corn; productive chain; agribusiness*

APLICAÇÃO PARA ANÁLISE DE PRODUTOS E PREÇOS DE CONCORRENTES, PARA E-COMMERCE DE MATERIAL ESPORTIVO

1 INTRODUÇÃO

Uma estratégia competitiva está em conhecer muito bem as regras e os outros jogadores. No entanto, no ambiente competitivo real, as regras nem sempre são claras e mudam com muita velocidade, demandando respostas ágeis, o que, cada vez mais, dificulta o mapeamento dos movimentos dos principais concorrentes (CARVALHO, 2000).

Nos dias atuais, o papel dos sistemas de informação vai além do suporte às atividades operacionais das organizações. As empresas, atentas ao poder da informação, investem continuamente em novas tecnologias que agreguem valor ao negócio, permitindo uma gestão estratégica para a tomada de decisões nas organizações.

Nesse sentido, as empresas devem manter canais fortes de acompanhamento das mudanças no ambiente competitivo, monitorando as novas tendências. A elaboração de uma boa estratégia competitiva depende do mapeamento do ambiente e de seu monitoramento constante, para rapidamente capitalizar as oportunidades ou neutralizar as ameaças.

O Business Intelligence (BI) está abrindo suas asas para abraçar todos, desde pequenas e médias empresas até grandes organizações. Ferramentas de análise para funções muito específicas também estão entrando no mercado, e elas ajudam algumas empresas a adotar apenas análise em vez da implementação completa de BI baseada em Data Warehouse (TURBAN, 2009).

Nesse contexto, a proposta deste trabalho é descrever o desenvolvimento de uma aplicação para análise de produtos e preços de empresas concorrentes, de maneira que possibilite o monitoramento de produtos, apresentando gráficos com histórico de variações de preços, permitindo ao gestor identificar continuamente a oferta do concorrente, proporcionando uma análise do posicionamento do concorrente no mercado.

Espera-se, ao desenvolver esse sistema de análise, que, além de auxiliar em uma gestão estratégica, ele seja capaz de gerar informações essenciais, que possam servir para o planejamento de novas estratégias para a empresa.

2 REVISÃO DE LITERATURA

Esta seção contém informações que nortearam o desenvolvimento do projeto, por meio do embasamento da literatura publicada sobre o tema.

2.1 BUSINESS INTELLIGENCE

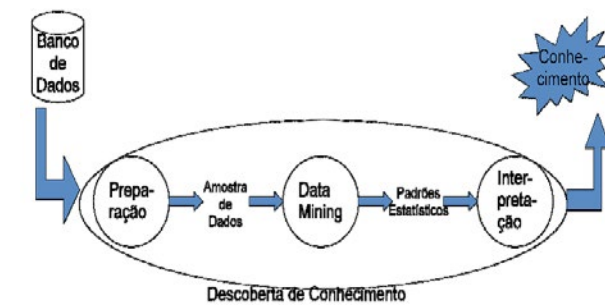
Conceitos e tecnologias de suporte à decisão foram implementados de forma crescente durante 30 anos, por muitos fornecedores que criaram uma diversidade de ferramentas e metodologias.

O conceito surgiu na década de 1990, e se refere aos processos de organização, coleta, análise, monitoramento e compartilhamento das informações que são a base da gestão de negócios, tais como informações sobre clientes, concorrentes, fornecedores e potenciais futuros clientes.

Business Intelligence (BI) é um termo "guarda-chuva" que inclui arquiteturas, ferramentas, bancos de dados, aplicações e metodologias. É uma expressão livre de conteúdo, portanto, significa coisas diferentes para pessoas diferentes. Parte da confusão relacionada ao BI é causada pela enxurrada de acrônimos e palavras da moda associadas a ele e a suas ferramentas como Business Performance Management - BPM (TURBAN, 2009).

Os principais objetivos do BI são permitir o acesso interativo aos dados, proporcionar a manipulação desses dados e fornecer aos gerentes e analistas de negócios a capacidade de realizar a análise adequada. Vários autores descrevem o processo geral de descoberta de conhecimento, conforme descrito na Figura 1. A entrada do processo é um banco de dados e a saída é um conjunto de conhecimentos. A etapa principal é a de mineração ou análise dos dados. Para tanto, é necessário ter antes uma etapa de preparação dos dados, a partir da base de entrada. Nessa etapa, os dados serão tratados (limpeza, integração, duplicidade) e amostras diferentes são geradas (LOH, 2014).

Figura 1. Processo geral de Descoberta de Conhecimento.



Fonte: LOH, (2014, P.28)

A etapa de análise tem como entrada uma amostra dos dados e gera como saídas padrões estatísticos. Mas esses padrões não são conhecimento. Eles precisam ser interpretados dentro do contexto da organização ou do domínio para se tornarem conhecimento. Os padrões estatísticos são, na maioria das vezes, hipóteses de causas, devendo ser mais bem avaliados, para realizar a preparação dos dados e depois a interpretação dos resultados. O processo do Business Intelligence baseia-se na transformação de dados em informações, depois em decisões e finalmente em ações. Os principais objetivos do Business Intelligence são permitir acesso interativo aos dados, proporcionar a manipulação desses dados e fornecer aos gerentes e analistas de negócios a capacidade de realizar a análise adequada (TURBAN, 2009).

2.1.1 DATA WAREHOUSE E DATA MARTS

Um Data Warehouse (DW) é um conjunto de dados produzido para oferecer suporte à tomada de decisões; é também um repositório de dados atuais e históricos de possível interesse aos gerentes de toda a organização. Os dados normalmente são estruturados de modo a estarem disponíveis em um formato pronto para as atividades de processamento analítico online (OLAP), com geração de relatórios e outras aplicações de suporte à decisão (TURBAN, 2009). Portanto, um Data Warehouse é uma coleção de dados orientada a assunto, integrada, variável no tempo e não-volátil, que proporciona suporte ao processo de tomada de decisões da gerência.

Um Data Warehouse une bancos de dados de toda uma empresa; já um Data Mart normalmente é menor e concentra-se em um assunto ou departamento específico. Um Data Mart é um subconjunto de um Data Warehouse, que normalmente consiste em uma única área temática,

como marketing, operações, vendas. Um Data Mart pode ser dependente ou independente.

Um Data Mart dependente é um subconjunto criado diretamente a partir do Data Warehouse. Ele tem a vantagem de usar um modelo de dados consistente e apresentar dados de qualidade. Os Data Marts dependentes suportam o conceito de um único modelo de dados em toda empresa, mas o Data Warehouse deve ser construído antes. Eles garantem que o usuário final visualize a mesma versão de dados acessada por todos os outros usuários do Data Warehouse.

O alto custo deste último, limita seu uso às grandes empresas. Como alternativa, muitas empresas usam uma versão de Data Warehouse reduzida em custo e escala, denominada Data Mart independente (TURBAN, 2009).

2.2 MINERAÇÃO DE DADOS

O processo de mineração corresponde à extração de minerais valiosos, como ouro ou pedras preciosas, a partir de uma mina. Uma característica importante desses materiais é que, embora não possam ser cultivados ou produzidos artificialmente, existem de maneira implícita e muitas vezes desconhecida em alguma fonte, podendo ser extraídos. Esse processo requer acesso à mina, o uso de ferramenta adequadas de mineração, a extração dos minérios propriamente dita e o seu posterior preparo para a comercialização (CASTRO, 2016).

O termo mineração de dados (MD) foi cunhado como alusão ao processo de mineração descrito anteriormente, uma vez que se explora uma base de dados (mina) usando-se algoritmos (ferramentas) adequados para se obter conhecimento (minerais preciosos). Os dados são símbolos ou signos não estruturados, sem significado, como valores em uma tabela, e a informação está contida nas descrições, agregando significado e utilidade aos dados. Por fim, o conhecimento é algo que permite uma tomada de decisão para agregação de valor, como, por exemplo, saber que vai chover no fim de semana pode influenciar sua decisão de viajar ou não para a praia (CASTRO, 2016).

As funcionalidades da mineração de dados são usadas para especificar os tipos de informações a serem obtidas nas tarefas de mineração. Em geral, essas tarefas podem ser classificadas em duas categorias: (1) descritivas: caracterizam as propriedades gerais dos dados; e (2) preditivas: fazem inferência a partir dos dados objetivando

predições. Em muitos casos, o usuário não tem ideia do tipo de conhecimento contido nos dados ou como usá-lo para gerar modelos preditivos, tornando importante a capacidade das ferramentas de mineração de dados.

Os algoritmos de aprendizagem de máquina são ferramentas poderosas para a descoberta de conhecimentos em bases de dados. Entretanto, uma etapa inicial do processo de mineração que não requer elevado nível de sofisticação é a análise descritiva dos dados, ou seja, o uso de ferramentas capazes de medir, explorar e descrever características intrínsecas aos dados. Especificamente, essas análises permitem investigar a distribuição de frequência, as medidas de centro e variação, e as medidas de posição relativa e associação dos dados. Além disso, técnicas elementares de visualização também são empregadas para um melhor entendimento de natureza e distribuição dos dados.

As análises descritivas permitem uma sumarização e compreensão dos objetos da base e de seus atributos, como qual o salário médio dos professores universitários brasileiros ou qual a distribuição salarial desses professores. Usando essas medidas, é possível saber, por exemplo, qual a posição relativa de um salário quando comparada à distribuição de salários disponível, o que permite identificar, por sua vez, se um salário está abaixo ou acima da média. Essas informações podem ser representadas por meio de gráficos do tipo torta, gráficos em barra, histogramas ou outras ferramentas equivalentes, cada uma capaz de explicar um conhecimento específico sobre os dados (CASTRO, 2016).

2.3 EXTRAÇÃO DE DADOS

Os avanços das tecnologias World Wide Web (www) possibilitam a geração de inúmeras fontes de dados. A Internet possui muitos dados de relevância, disponíveis em seus documentos, porém, seu modelo de publicação das informações permite aos usuários um modo informal de publicá-las, que remete à não garantia de consistência desses dados, dificultando sua filtragem para subsidiar diversas áreas de conhecimentos (PONTOLIO, 2015).

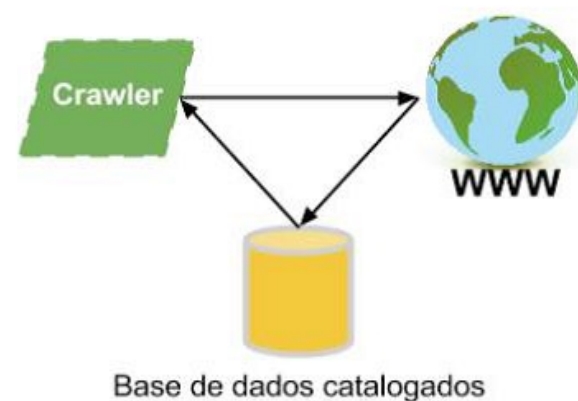
Segundo Mendonça (2003), fontes de informação são sistemas que respondem a consultas, retornando uma resposta apropriada para cada consulta submetida. No âmbito da Internet,

essas fontes possuem formatos textuais, que são manipulados pelo paradigma request-response implementado pelo protocolo Hypertext Transfer Protocol (HTTP), geralmente utilizando formatos semiestruturados.

Utilizando extração de dados para localizar, coletar e organizar dados de interesse apresentados nos formatos citados, é possível o enriquecimento de grandes bases de dados, permitindo a realização de consultas e o cruzamento de tais dados, o que não era possível devido a interfaces de consultas pré-estabelecidas nas fontes de informação, possibilitando maior integridade destas, já disponíveis, e facilitando a construção de agentes inteligentes.

Algumas das ferramentas de busca utilizam como forma de extrair informações programas capazes de percorrer a web automaticamente catalogando e classificando Universal Resource Locators (URLs), e as informações obtidas servem para o armazenamento em repositórios locais. Esses programas são chamados de Web Crawlers, representados na Figura 2. Porém, este tipo de software utiliza uma navegação entre as páginas do tipo "força bruta", isto é, não há especificações (filtros) que identifiquem os links que o software deverá acessar, pois ele se move automaticamente de site para site, seguindo os links embutidos em seus documentos, copiando tudo que lhe for possível (PONTOLIO, 2015).

Figura 2. Representação do funcionamento de um Crawler.

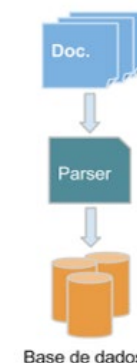


Fonte: PONTIOLLO, (2015, P.42)

Outro software capaz de realizar a extração de dados em documentos são os parsers, um programa que espera receber dados de uma maneira estruturada, de forma que informações consigam ser extraídas mecanicamente. Esse tipo

de abordagem oferece pouca inteligência, pois os parsers não são capazes de realizar a extração em ambientes não estruturados. Essa característica desta técnica pode ser observada na Figura 3.

Figura 3. Representação da funcionalidade de um Parse.



Fonte: PONTOLIO, (2015, P.43)

Sistemas de extração de dados visam localizar e extrair, de forma automática, informações relevantes em um documento ou coleção de documentos, contendo textos em linguagem natural, e estruturar tais informações para os padrões de saída, a fim de facilitar sua manipulação e análise. Nesse contexto, ressalte-se que o programa caracteriza esta ideia. O Data Scraping é também um software capaz de extrair dados da saída de um outro programa. Esse modelo é mais conhecido popularmente nos dias atuais como Web Scraping, um software capaz de extrair dados de documentos Web (PONTOLIO, 2015).

Web Scraping (conhecido também como Screen Scraping) é uma técnica de captura de dados a partir de sites. Está intimamente relacionado à indexação da web, indexando mais detalhes sobre ela, utilizando scripts ou bots (robôs de busca) e sendo, por fim, uma técnica mundial adotada pela maioria dos motores de busca.

3 METODOLOGIA

Foi realizado um estudo bibliográfico sobre Business Intelligence, que está sendo utilizado neste trabalho para obtenção de resultados. Também foram estudadas tecnologias para extração de dados, levando em consideração recursos, disponibilidade e custos para adoção do sistema.

O desenvolvimento da aplicação BI foi dividido em 4 fases: mapeamento das fontes dos dados, projeto, construção da solução BI e disponibilização aos usuários.

A fase do mapeamento das fontes dos dados envolveu a definição sobre em quais sites seria realizada a extração de dados. A extração de dados envolveu a criação de uma ferramenta Web Crawler utilizando o framework Scrapy. Foram criados crawlers denominados de spiders, e essas spiders percorrem os sites escolhidos extraíndo os dados (Scraping). De maneira estruturada, os dados podem ser utilizados para uma gama de aplicações. Os principais requisitos da ferramenta são:

- Percorrer e capturar elementos das páginas web;
- Apresentar os dados capturados em diferentes formatos JavaScript Object Notation (Json), Extensible Markup Language (XML), Comma-Separated Values (CSV);
- Oferecer portabilidade (execução em diversas plataformas).

A fase de projeto envolveu a definição das principais tecnologias a serem utilizadas na construção da aplicação. O Quadro 1 apresenta as tecnologias utilizadas no desenvolvimento da ferramenta.

Quadro 1. Ferramentas e Tecnologias Utilizadas

Nome	Descrição
Python	Linguagem de programação, utilizada no desenvolvimento da ferramenta.
PyCharm	Ambiente de desenvolvimento Python.
Scrapy	Framework para "crawlear" web sites e extrair dados estruturados.
Scrapinghub	Plataforma na nuvem para implantação e execução de rastreadores Web.
MySQL	Sistema de gerenciamento de banco de dados (SGBD), que utiliza a linguagem SQL (Structured Query Language) como interface.
Power BI	Framework coleção de serviços de software, aplicativos e conectores que trabalham juntos para transformar dados não relacionados em informações coerentes.
Visual Paradigm	Software utilizado para desenvolver os diagramas UML.

Fonte: Elaborado pelo Autor, 2017.

Na fase da construção da solução BI foi realizada a codificação de scripts na linguagem Python, utilizando o ambiente de desenvolvimento PyCharm e o framework Scrapy para extrair dados da Web e salvá-los no banco de dados. Foi utilizado também o framework Power BI na organização e manipulação desses dados.

Por último, a fase de disponibilização aos usuários, utilizando o Power BI para compartilhar relatórios por meio do serviço ou aplicativos móveis do Power BI.

4 DESENVOLVIMENTO E RESULTADOS

Para a composição deste projeto, foi empregado o framework Scrapy, uma biblioteca em Python open-source, que auxilia a criação de crawlers. O framework simula ações exercidas por um Browser, assim, a navegação consiste em "abrir" diferentes páginas de um site com o objetivo de coletar dados ou executar consultas. Com isto, é realizado downloads das páginas, com simples requisições HTTP, utilizando os métodos POST e GET.

O método GET é usado para recuperar qualquer informação referenciada por um URL, permitindo o primeiro reconhecimento dela. Mas o POST vai além: é usado para enviar um conjunto de dados para um serviço especificado por uma URL e recuperar as informações resultantes do processamento desses dados, o que pode ser compreendido como o envio de dados para consulta em um arquivo ou salvo diretamente no banco de dados.

O conceito central no Scrapy são as spiders, classes que herdam de scrapy.Spider, como mostrado na linha 7 da Figura 4. Na linha 11 da Figura 4, o start_urls, que define em quais URLs do site será realizada a extração de dados, o "name" define o nome da spider. Nesse caso foi escolhido o nome da loja virtual em que foi realizada a extração dos dados. Outro conceito importante no Scrapy são as settings (configurações). As settings oferecem uma maneira de configurar componentes do Scrapy, podendo ser iniciadas de várias maneiras, tanto via linha de comando, variáveis de ambiente em um arquivo settings.py quanto diretamente no spider. Na linha 17 da Figura 4, o download_delay é um componente das settings que define o intervalo entre cada requisição. Após alguns testes foi escolhido um delay de 2.0 segundos pelo motivo de não ser bloqueado pelos sites e não afetar a velocidade de download dos itens.

Figura 4. Trecho de código-fonte de uma Spider

```
7 class DafitiSpider(scrapy.Spider):
8     name = 'dafiti'
9
10
11     start_urls = ['https://www.dafiti.com.br/calçados/adidas/',
12                 'https://www.dafiti.com.br/calçados/asics/',
13                 'https://www.dafiti.com.br/calçados/mizuno/',
14                 'https://www.dafiti.com.br/calçados/new-balance/',
15                 'https://www.dafiti.com.br/calçados/nike/']
16
17     download_delay = 2.0
```

O método parse(), na linha 19 da Figura 5, recebe um objeto response, que representa uma resposta HTTP e é chamado de call-back. Os métodos call-backs no Scrapy são generators (ou retornam uma lista ou iterável) de objetos que podem ser dados extraídos ou novas requisições.

O motor do Scrapy itera sobre os objetos resultantes dos call-backs e os encaminha para o pipeline de dados ou para a fila de próximas requisições a serem feitas. Com esse esquema de requisições e call-backs que podem gerar novas requisições (como novos call-backs), pode-se programar a navegação por um site gerando requisições para os links a serem seguidos, até chegar nas páginas com os itens desejados para extração.

Na linha 23, da Figura 5, o dafiti_xpath define o caminho completo onde está o item a ser extraído ou apenas a class a que ele pertence. Por exemplo, existe uma lista de itens que pertencem a class "Product-box-detail", e a spider vai percorrer toda a página procurando apenas itens que estão dentro dessa class ignorando todas as outras.

Figura 5 – Trecho de código-fonte de uma Spider

```
19 def parse(self, response):
20     diaAtual = date.today()
21
22     dafiti_xpath = '//div[contains(@class, "product-box-detail")]'
```

Fonte: Elaborado pelo autor, 2017.

A marcação número 1 da Figura 6 representa o nome da spider, a ordem de execução e a quantidade de vezes que já foi executada. No caso da spider "netshoes", ela foi executada 36 vezes e é sempre a terceira spider a ser executada pela plataforma.

A marcação número 2 da Figura 6 apresenta a quantidade de itens que foram extraídos através da execução daquela spider. Por exemplo, a spider "dafiti" extraiu um total de 243 itens na sua execução. A quantidade de itens pode variar mesmo que não mude nada no código-fonte da spider, pelo fato de que o site alvo da extração pode sofrer alterações.

Figura 6. Printscreen Tela do histórico das spiders no Scrapinghub

Job Spider	Items	Requests	Errors	Log	Runtime	Status
1	121	6	0	17	0:01:26	2017-09-02 05:46:11 UTC
2	243	6	0	16	0:00:32	2017-09-02 05:49:10 UTC
3	41	3	0	16	0:00:40	2017-09-02 05:35:09 UTC

Fonte: Elaborado pelo autor, 2017.

O framework Power BI foi integrado ao projeto, um pacote de ferramentas de análise de negócios para analisar dados e compartilhar ideais desenvolvida pela Microsoft. A ferramenta conecta centenas de fontes de dados, simplificando a preparação dos dados, e conduz a análise.

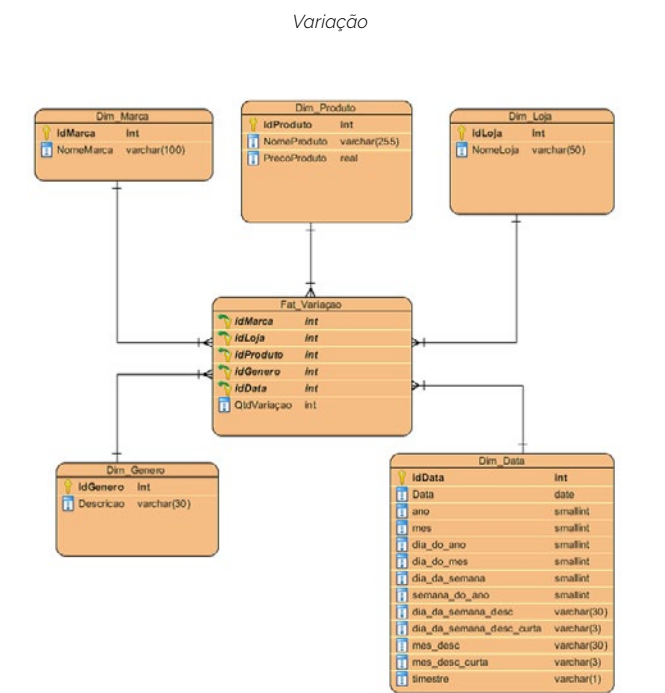
O framework pode unificar todos os dados, estando na nuvem ou localmente usando os gateways do Power BI, pode conectar bancos de dados SQL Server, modelos do Analysis Services e outras fontes de dados aos mesmos painéis no framework. A Figura 7 representa um Data Mart independente, focado nas variações de preços, com as seguintes dimensões: marca, gênero, produto, loja, data. Através desse fato, é possível fazer várias análises, gerando informações importantes para empresa.

Ele é a base para a descoberta de conhecimento. Nesse formato é possível identificar em quais lojas ocorre um volume maior de mudanças nos seus preços. Ao buscar padrões de quais dias da semana ou períodos do mês, geralmente ocorrem essas mudanças. Ao se fazer análise por datas, comparar datas específicas como Natal, Dia dos Namorados e Dia das Mães, pode-se assim reconhecer seu comportamento (quais foram os produtos e quais marcas tiveram queda de preço, quais não mudaram, mesmo no dia da promoção), analisar a porcentagem de desconto e verificar se este foi o mesmo em todos os produtos de uma determinada marca.

É importante ter o histórico completo de alterações, podendo-se, assim, analisar bons momentos para investir em marketing e promoções, buscar datas pouco exploradas pelos adversários, verificando se eles possuem poucos produtos com preços atrativos ao consumidor, podendo assim ser mais

efetivo em uma campanha de vendas.

Figura 7. Diagrama de Entidade e Relacionamento do Fato Quantidade



Fonte: Elaborado pelo autor, 2017.

A Figura 8 representa uma parte de um relatório em formato tabela de apenas um produto construído no Power BI. Nesse caso, a análise do produto foi feita do dia 9 ao dia 15 de setembro de 2017. O relatório contém o nome do produto, a data completa em que o produto foi extraído e adicionando à base de dados, o preço do produto e o nome da loja.

Com esses dados é possível analisar a variação de preços de um ou vários produtos do concorrente, identificar padrões e relações dessas variações, analisar o comportamento e a constância dessas variações perto de datas comemorativas que aquecem o mercado varejista, como Natal, Dia dos Namorados, Dia das Mães e muitas outras. Sabendo a oferta do seu concorrente periodicamente, é possível montar estratégias de marketing e promoções. O objetivo dos relatórios é entregar informações que ajudem na tomada de decisão dos gestores, dando a eles uma base de dados que ajude no planejamento de ações

baseadas nas informações.

Figura 8. Printscreen Relatório em formato tabela no Power BI

Nome	Ano	Trimestre	Mês	Dia	Preço	Loja
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	9	R\$384,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	10	R\$384,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	11	R\$384,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	12	R\$391,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	13	R\$391,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	14	R\$384,9	Netshoes
Tênis Adidas Springblade E-Force Fe...	2017	Tri 3	setembro	15	R\$384,9	Netshoes

Fonte: Elaborado pelo autor, 2017.

A Figura 9 apresenta o gráfico da variação de preço de apenas um produto. Existem várias formas de apresentar essa variação em um relatório, e a maneira de apresentar depende do gosto de cada gestor.

O Power BI proporciona criar vários modelos de relatórios dinâmicos que podem ser alterados ou customizados por qualquer pessoa com acesso àquela base de dados. Uma das grandes vantagens de se utilizar esse framework é a facilidade para compartilhar e publicar relatórios, pois ele salva passo a passo toda a lógica utilizada para se chegar ao resultado apresentando.

Figura 9. Printscreen Relatório em formato gráfico no Power BI



Fonte: Elaborado pelo autor, 2017.

5 CONCLUSÃO

O desenvolvimento da ferramenta teve como objetivo criar uma aplicação que permitisse a análise de produtos e preços de concorrentes, uma ferramenta que pode servir de apoio à tomada de decisão dos gestores, monitorando periodicamente a oferta dos seus concorrentes, emitindo diversos relatórios com análise de alteração de vários produtos.

O projeto está em desenvolvimento, algumas

funções estão sendo incrementadas, com o propósito de garantir tanto uma aplicação de qualidade quanto uma aplicação que agregue informações relevantes auxiliando em um planejamento estratégico.

Portanto, a aplicação irá possibilitar uma melhor gestão de conhecimento do mercado, pois os gestores vão ter acesso sempre que necessário a informações relevantes sobre produtos de seus concorrentes, conhecendo melhor a oferta de seus adversários, assim, auxiliando na elaboração de boas estratégias competitivas, mapeando o ambiente e o monitorando, para capitalizar oportunidades ou neutralizar possíveis ameaças.

É importante ressaltar que a aplicação é evolutiva: em determinado momento, uma análise descritiva dos dados satisfaz as necessidades da empresa, entretanto, o crescimento da base de dados possibilita a utilização de algoritmos de aprendizado de máquina, ferramentas poderosas para descoberta de conhecimento, que será incorporado ao projeto no momento adequado.

Futuramente pretende-se desenvolver um módulo para minerar opiniões de consumidores em redes sociais, opiniões sobre o próprio e-commerce e seus produtos, também sobre seus concorrentes. Essa técnica é conhecida como "mineração de opiniões em textos". Assim, a aplicação poderá atuar tanto na análise de produtos e preços, quanto na análise de opiniões de consumidores, identificando pontos positivos que devem ser mantidos e pontos negativos que apontam possíveis melhorias, auxiliando no planejamento de estratégias competitivas.

REFERÊNCIAS

CARVALHO, Marly Monteiro. Estratégia competitiva: dos conceitos à implementação. São Paulo: Editora Atlas, 2000.

CASTRO, L. N.; FERRARI, D. G. Introdução à Mineração de Dados: conceitos básicos, algoritmos e aplicações. São Paulo: Saraiva, 2016.

LOH, Stanley. BI na Era do Big Data para cientistas de dados. Porto Alegre: Amazon, 2014.

MENDONÇA, Eduardo. Extração Resiliente de Dados RDF a partir de Fontes Dinâmicas em Linguagem de Marcação. 2003. Disponível em:

<http://www.livrosgratis.com.br/arquivos_livros/cp108124.pdf>. Acessado em: maio de 2017.

PONTOLIO, Luan Silveira. Plataforma de Extração e Recuperação de Dados na Web no Contexto de Big Data. 2015. Disponível em: < <http://hdl.handle.net/11077/1006>>. Acesso em: maio de 2017.

TURBAN, Efraim et al. Business Intelligence: Um Enfoque Gerencial para a Inteligência do Negócio. Porto Alegre: Bookman Editora, 2009.