

Faculdade de Ciências Gerenciais Alves Fortes

Cláudio Manoel Assis Soares

Como a Ciência de Dados auxilia na tomada de decisões

**Além Paraíba
2021**

Cláudio Manoel Assis Soares

Como a Ciência de Dados auxilia na tomada de decisões

Bacharelado em Administração

Faculdade de Ciências Gerenciais Alves Fortes

Professor(a) orientador(a): Allan Lima Ferreira

Além Paraíba
2021

ASSIS SOARES, Cláudio Manoel

Como a Ciência de Dados auxilia na tomada de decisões

Planejamento Estratégico

30 páginas

Bacharel em **Administração de Empresas** – Faculdade de Ciências
Gerencias Alves Fortes – FACE-ALFOR, mantida pela Fundação
Educacional de Além Paraíba – FEAP.

Coordenador: Allan Lima Ferreira

Professor(a) Orientador(a): Allan Lima Ferreira



Cláudio Manoel Assis Soares

Como a Ciência de Dados auxilia na tomada de decisões

Monografia apresentada a Faculdade de Ciências Gerenciais Alves Fortes – FACE ALFOR, mantida pela Fundação Educacional de Além Paraíba – FEAP, como requisito parcial à obtenção do título em Bacharel em Administração.

Banca examinadora:

Prof. Orientador(a): Allan Lima Ferreira

Prof. Convidado(a): Nome do Prof. Convidado(a)

Prof. Convidado(a): Nome do Prof. Convidado(a)

☐ Aprovado

☐ Aprovado com restrições

☐ Reprovado

Prof. Coordenador: Allan Lima Ferreira

Além Paraíba, 30 de Novembro de 2021

Dedicatória

Dedico este trabalho a meu herói, meu amado pai. Um homem incrível que sempre foi referência para mim. Eternamente vivo em nossas memórias e em nossos corações. Obrigado por tudo.

Agradecimentos

Agradeço a Deus por sempre me conceder força e sabedoria em momentos de incerteza, e por conduzir meus passos até aqui. Agradeço a minha mãe, parte fundamental desta conquista, irmão e irmã, cunhados e meus sobrinhos por me fazerem companhia nesta trajetória. Agradeço também a meu querido pai que olha lá do céu por todos nós. Obrigado Deus por minha família e por todo amparo que me deram, pois sem eles nada disso seria possível. Agradeço também ao meu orientador e coordenador pelas orientações que me deu para a construção da presente monografia.

Resumo

O presente trabalho busca demonstrar a relevância de Data Science para alavancar a probabilidade de sucesso no momento da tomada de decisões em uma organização, com o propósito de solucionar possíveis problemas de negócio ou de meramente realizar uma ação perante algum aspecto definido. O método utilizado para a realização deste trabalho foi primeiramente consultar literaturas da área, selecionar fragmentos de textos e discorrer sobre o tema com embasamento previamente estudado. Assim, resultou-se em um conteúdo de bastante veracidade e que aponta fatos no que tange à importância dos dados para as organizações. Empresas de diferentes portes precisam ter esse diferencial competitivo, que é a capacidade de analisar seus dados para tomar decisões com base neles, para que possam alavancar as chances de obterem um bom resultado. Para isso, precisam inicialmente implantar na cultura da organização a afirmação de que precisam se basear em dados a todo momento. Posteriormente, utilizar das técnicas necessárias para manusear corretamente tais dados a fim de transformá-los em informações, adquirir conhecimento a partir destas informações, e assim, poder fornecer aos gestores e tomadores de decisão *insights* para que possam tomar medidas de forma segura e com chances maiores de êxito.

Palavras Chaves: Dados. Resultado. Ciência de Dados. Inteligência de Negócios. Tomada de decisão.

Abstract

The present work seeks to demonstrate the relevance of Data Science to leverage the probability of success at the time of decision making in an organization, with the purpose of solving possible business problems or merely taking action in view of some defined aspect. The method used to perform this work was to first consult literatures in the area, select fragments of texts and discuss the theme with previously studied basis. Thus, it resulted in a content of very veracity and that points out facts regarding the importance of data for organizations. Companies of different sizes need to have this competitive differential, which is the ability to analyze their data to make decisions based on them, so that they can leverage the chances of getting a good result. To do this, they initially need to deploy to the organization's culture the statement that they need to be based on data at all times. Subsequently, use the necessary techniques to properly handle such data in order to transform it into information, acquire knowledge from this information, and thus be able to provide managers and decision makers *with insights so* that they can take action safely and with greater chances of success.

Key words: Data. Result. Data Science. Business Intelligence. Decision making.

Lista de Ilustração

Figura 01: Data Science no contexto da tomada de decisão	21
Figura 02: Exemplos de gráficos ineficientes	23
Figura 03: Exemplo de dashboard de desempenho	24
Figura 04: Processos, dados e pessoas: capital intelectual das empresas	27

Lista de Siglas

LGPD – Lei Geral de Proteção de Dados

IoT – Internet of Things (Internet das Coisas)

TI – Tecnologia da Informação

PCs – Personal Computer (Computador Pessoal)

DVDs – Digital Video Disk (Disco Digital de Vídeo)

DOD – Tomada de Decisão Orientada por Dados

SUMÁRIO

Introdução	9
Metodologia	10
1 – O que são dados	11
1.1 – Formato dos Dados	11
1.2 – Importância dos dados	12
1.3 – Importância da proteção de dados	13
1.4 – Big Data	14
1.4.1 – Os 5 V's do Big Data	16
1.4.1.1 – Variedade	16
1.4.1.2 – Velocidade	16
1.4.1.3 – Valor	17
1.4.1.4 – Volume	18
1.4.1.5 – Veracidade	18
2 – Ciência de Dados (<i>Data Science</i>)	20
2.1 – Processo Orientado a Dados	20
2.2 – Banco de Dados em Nuvem	22
2.3 – Business Intelligence	24
3 – Benefícios da análise dos dados para a tomada de decisões	28
Conclusão	31
Bibliografia	32

Introdução

Vivemos em um mundo cada vez mais cheio de novas tecnologias e novos métodos de administração. Com o advento da internet e sua ampla usabilidade, somados à necessidade de pessoas e empresas a utilizarem cada vez mais, a fim de se manterem atualizados e em constante inovação, inúmeros dados são gerados a todo momento nesta operação. E hoje, com o *Big Data*, novas tecnologias estão sendo estudadas para processar toda esta quantidade de dados, e cada vez mais eles precisam ser analisados e transformados em informação, para que assim, as organizações possam ser capazes de criar soluções para possíveis problemas de negócio. A Ciência de Dados (*Data Science*) é composta por um conjunto de conceitos provenientes de diversas áreas de conhecimento, capazes de conceder base teórica no momento da tomada de decisão. Assim, pode ser constatado o quão importante é a análise de dados antes de uma decisão estratégica ou solução de um problema corporativo.

Esta pesquisa tem relevância pois irá mostrar como as organizações podem utilizar os dados para auxiliar o gestor na tomada de decisões, a fim de tomá-las com mais ciência e aumentar suas chances de assertividade, não importando qual seja seu problema de negócio. Em algumas circunstâncias talvez nem seja um problema, apenas uma medida estratégica a ser tomada, ou um direcionamento de mercado relacionado à algumas das áreas que compõem a organização, enfim, com o intuito de melhorar algum setor da empresa. Os dados também podem dar direcionamento e embasamento teórico com relação a ações de campanhas de marketing, processos de produção, logística, dentre outras questões.

No decorrer do presente trabalho, que se inicia agora, será discorrido sobre a definição do que são dados, como eles são captados e armazenados, e porque são de extrema importância para as organizações. Também poderá ser observado no decorrer desta monografia como funciona a tecnologia de processamento de dados, computação em nuvem, dentre outros conceitos. E por fim, como os gestores podem se beneficiar das informações provenientes dos dados para suas possíveis ações estratégicas dentro do ambiente organizacional.

Metodologia

Primeiramente a presente monografia tem seu conteúdo baseado através de pesquisas realizadas inteiramente na internet, acessando conteúdos bibliográficos de autores da área, trabalhos acadêmicos também relacionados à área e capítulos de livros de estudiosos sobre o assunto, além de consultas a artigos e informativos publicados. Assim, o presente trabalho terá embasamento para comprovar a sua premissa.

Para chegar ao objetivo da presente monografia, foram identificados o tema e a área de estudo a serem retratados, o problema, hipótese, justificativa e os objetivos. Para isso, uma revisão literária foi consultada e retratada para complementar a veracidade de todo o texto.

1 – O que são dados

Os dados estão por toda parte, eles podem ser representados por imagens, números, caracteres, som, tabelas, vídeo, de diversas formas diferentes. Porém, dado é algo que não consegue transmitir uma mensagem clara, pois pode ser composto de diversos tipos de elementos a respeito de alguma coisa, algum assunto. Os dados constituem a matéria prima da informação. Ou seja, é através da coleta e interpretação de dados que se é capaz gerar mensagens claras (informações) a respeito de algo, e a partir destas informações interpretadas gerar conhecimento, que por sua vez é de extrema importância para pessoas e organizações.

De acordo com Cole Nussbaumer Knafllic (2019), uma boa visualização de dados pode significar a diferença entre o sucesso e o fracasso quando se vai comunicar a alguém as constatações da interpretação desses dados, e a capacidade de transmitir uma correta ideia com dados é uma habilidade muito importante em um mundo coberto de dados crescentes e da constante necessidade de tomadas de decisões orientadas por dados.

Hoje com as novas tecnologias que não param de evoluir, as pessoas – ou melhor, os consumidores – estão conectados quase o tempo inteiro, e com isso, as empresas acabam tendo a chance de também estarem conectadas a esses indivíduos. Segundo uma pesquisa da IBM com trinta mil consumidores em treze países mostrou que cerca de 78% a 84% dessas pessoas se baseavam nas redes sociais, no momento da pesquisa de produtos para consumo; 45% preferiam consultar amigos ou pessoas que já haviam consumido o mesmo produto anteriormente; e 18% levavam em conta informações dos produtores e vendedores para tomar sua decisão (Taurion, 2013). Isto indica que para as organizações se manterem conectadas aos possíveis clientes através dos dados, é necessário estar sempre atentas ao que acontece no mundo digital, pois a fonte de dados é imensa e eles estão em ampla circulação a todo momento.

1.1 – Formato dos Dados

Os dados podem apresentar formatos diferentes. Eles podem ser classificados como: Estruturados e Não Estruturados.

Dados Estruturados são dados que podem ser claramente definidos com padrões e são fáceis de serem encontrados através de pesquisas. Eles podem ser encontrados em resultados de pesquisas, medidas de controle de processos, e principalmente em Banco de Dados de sistemas internos de organizações. Portanto, os dados estruturados não aqueles que já estão

medidos, por assim dizer, através de alguma estatística ou qualquer outro tipo de mensuração ou tipagem.

Foi observado no parágrafo acima exemplos de dados estruturados provenientes de fonte de dados internos. Além deste, pode-se encontrar dados estruturados também de forma externa, como por exemplo em pontuações em sites de classificação e pela quantidade de reações em publicações de redes sociais, por exemplo.

Já os Dados Não Estruturados consistem em dados que geralmente não são tão fáceis de serem pesquisados, pois não estão classificados de forma ordenada, organizada, como os estruturados. Podem ser, por exemplo, mensagens de voz, postagens em rede social, e-mails, avaliação de funcionários, ilustrações e outras formas.

Os dados não estruturados possuem este nome justamente por não possuírem uma estrutura pré-definida, tampouco um tipo de modelo ou esquema organizado, podendo assim ter diversas formas (textual, não-textual) e podendo também serem gerados por humanos ou máquinas.

Pôde ser constatado anteriormente algumas formas de encontrar dados não estruturados de forma interna, e como exemplos de como encontrá-los em fonte de dados externa pode-se destacar conteúdos publicados em mídias sociais, comentários encontrados em fóruns online e na internet de modo geral, vídeos e imagens.

Existem ainda os Dados Semiestruturados, que são dados que não tem as características pré-definidas de um dado estruturado, mas ainda assim conseguem ter “tags”, espécie de marcadores que vão ser responsáveis por dar certa ordem a esses dados, separando-os por tipo, hierarquia, diferenciando um tipo de registro ou de medida, de outro.

1.2 – Importância dos dados

No âmbito das organizações, os dados trazem à tona aquilo que aconteceu/acontece dentro da empresa, portanto, nada mais correto do que se basear neles para decidir tudo o que for possível. Segundo Foster Provost e Tom Fawcett (2016), os últimos quinze anos testemunharam grandes investimentos em infraestrutura de negócios que têm melhorado a capacidade de coletar dados em toda a empresa. Praticamente todo modelo de negócio está propenso e aberto para a coleta de dados e, atualmente, são até instruídos a isso. Os dados podem ser úteis para o melhoramento das operações internas, gestão de *Supply Chain* (cadeia

de suprimentos), gestão e mensuração de satisfação e comportamento dos clientes, desempenho e iniciativas de marketing, captação de novos clientes e por aí vai.

Todos os aspectos do negócio possuem dados que podem ser coletados, analisados, interpretados e passados aos tomadores de decisão em forma de conhecimento, a fim de bem orientá-los na tomada de ações estratégicas. E, a melhor forma de extrair informação e conhecimento proveniente dos dados, é através do que se conhece por *Data Science* (Ciência de Dados). De acordo com Stanley Loh (2019), se os dados não estiverem disponíveis para análise, toda decisão se tornará um simples jogo de sorte ou azar. Por isto, dados e estatísticas são muito importantes e mudam drasticamente o direcionamento mercadológico e estratégico das organizações que fazem bom uso da análise de dados.

Quadro 1 – Dado x Informação x Conhecimento

Dado	Informação	Conhecimento
Conjunto de letras, números ou dígitos que não contém significado claro.	Dado trabalhado, útil, tratado, com valor significativo atribuído ou agregado a ele com um sentido natural e lógico.	Informação trabalhada por pessoas e pelos recursos computacionais, possibilitando a geração de cenários, simulações e oportunidades.

Fonte: Big Data: um estudo em gestão empresarial (2016).

1.3 – Importância da proteção de dados

A ideia de que dados são de extrema importância para pessoas e organizações já está bem implícita já é de amplo conhecimento ao redor do mundo. Devido a isso, alguns destes dados podem ser sigilosos e sua divulgação, além de não se fazer necessária, pode trazer conflitos indesejáveis. Conforme Taurion (2013), uma política de acesso e divulgação das informações devem ser instauradas e são essenciais para manter seguras as pessoas proprietárias destes dados, quando são de forma externa.

Com o desenvolvimento das tecnologias digitais, a privacidade acabou. Primeiro, há tecnologias para coleta e armazenamento de dados. Como já discutido antes, está todo mundo coletando dados sobre todos. A operadora de celular sabe por onde a gente anda e quando, qual o caminho que costumamos fazer, por onde costumamos andar em cada dia da semana e horário. [...] Aí alguém inventou a tecnologia de RFID, e ela está em cartões com chips, carros, produtos novos e vai estar em sacolas, carrinhos de supermercados, etc. Então não é só por celular. Os aplicativos e softwares que usamos

em celulares, tablets, notebooks e etc também estão avisando onde estamos, se estivermos conectados via Wifi, 3G ou 4G. (STANLEY LOH, 2019, p.145)

Portanto, há um certo monitoramento excessivo neste contexto. Por esse motivo, alguns países tem adotado medidas de proteção, por meio de legislações, com o intuito de diminuir a ampla “invasão” existente a dados de indivíduos e de empresas. Assim foram criando-se as políticas de privacidade, que tem como objetivo informar aos indivíduos quantos e quais dados estão ou vão ser coletados. No Brasil, a legislação que se refere é a Lei Geral de Proteção de Dados (LGPD). A LGPD (Lei nº 13.709, de 14/8/2018) entrou em vigor em 18 de Setembro de 2020 e atua perante o uso indevido, comercialização e vazamento irregular de dados, protegendo assim a privacidade da população.

A Lei obriga que empresas ou órgãos específicos deixem bem claro o interesse e comuniquem aos proprietários dos dados a finalidade para qual eles serão coletados, sendo estes dados pessoais, para então poderem analisá-los e tratá-los conforme informado anteriormente aos titulares. A criação constante de legislações, no Brasil e em países afora acerca da proteção de dados, reflete cada vez mais a importância dos mesmos na constante ascensão dos dados que cada vez mais impactam a vida de pessoas, empresas, sociedade, e também o mundo como um todo.

1.4 – Big Data

Sempre existiram dados disponíveis para análise no âmbito corporativo, advindos de sistemas internos de processos empresariais, mídias sociais, imagens, mídias tradicionais (jornais, televisão, rádio), além de vídeos na internet, sensores dos mais diversos tipos, mecanismos eletrônicos e etc. Neste sentido, atualmente a quantidade de dados que é gerada a cada instante aumentou exponencialmente e de forma gradativa, pois basicamente tudo aquilo que se conecta à internet pode gerar inúmeros dados, principalmente com o desenvolvimento de tecnologias como a IoT (*internet of things* ou internet das coisas), onde inúmeros aparelhos eletrônicos podem se conectar à internet e facilitar o cotidiano das pessoas. Os dados estão aumentando em volume, variedade e velocidade, portanto, é o *Big Data* que fornecerá a matéria prima para a Ciência de Dados.

Tudo aquilo que um indivíduo realiza no ambiente remoto gera dados que estão sendo armazenados. Sejam ações de cadastramento em sites, envio de e-mails, sistemas de pontuação de sites, comentários em fóruns, compras em sites de *e-commerce*, pesquisas, publicações em mídias sociais, tempo médio de consumo de internet, dentre muitas outras coisas, tudo isto

compreende um pouco do que o *Big Data* abrange. Se resume a uma imensa quantidade de dados desestruturados.

Segundo Foster Provost e Tom Fawcett (2016), essencialmente, o termo *Big Data* pode significar conjuntos de dados que são grandes demais para os sistemas tradicionais de processamento e, portanto, exigem novas tecnologias para processá-los.

A realização do vasto potencial do Big Data é viabilizada pela ubiquidade dos dispositivos de coleta e processamento de dados. O fato é que os sensores e os microprocessadores em breve serão onipresentes. Praticamente todos os dispositivos mecânicos ou eletrônicos podem deixar rastros que descrevem seu desempenho, localização e situação. Esses dispositivos e seus usuários se comunicam por meio da internet – o que gera outra fonte profusa de dados. Quando se combinam todos esses bits com os de outras mídias – telefonia com e sem fio, cabo, satélite e assim por diante – o futuro dos dados parece ainda mais amplo. (H. DAVENPORT, 2014, p.13)

Cada vez mais poderá ser observado que quatro forças ou tendências que ainda estão em desenvolvimento irão causar na humanidade revoluções significativas na indústria de TI e no uso da tecnologia em geral. Seriam elas *cloud computing*, mobilidade, *social business* e *Big Data* (Taurion, 2013). Estes aspectos caminham juntos, e precisam ser observados com bastante atenção pelas organizações, pois no futuro esses termos vão ser de grande responsabilidade e extremamente atuantes na forma como as empresas terão de pensar.

Big Data é algo que atualmente tem despertado interesse em muitos indivíduos e organizações, estudiosos e curiosos. Contudo, ainda é um conceito que demanda de poucas definições. Ainda segundo Taurion (2013), *Big Data* significa coisas diferentes para pessoas diferentes. A mídia, tecnológica e de negócios e fornecedores de tecnologia contribuem para confundir ainda mais as definições, pois quase sempre o *Big Data* é retratado conforme suas idealizações separadamente. Ou seja, podem levar o consumidor a achar que apenas adquirindo seus produtos, por exemplo, eles poderão ter uma solução rápida e extremamente fácil para as questões que envolvem *Big Data*, e assim, terão sanadas as necessidades cada vez mais evidentes de que dados são o futuro das organizações corporativas, e que com eles a probabilidade de se atingir o sucesso nas decisões é elevada.

“Os dados obtidos pelos sistemas transacionais, os visíveis comumente, representam uma parcela ínfima dos dados que circulam pelas empresas. Este imenso volume de “*shadow data*” [...], tem passado despercebido e não é utilizado para melhorar os processos de tomada de decisões” (Taurion, 2013). Isto significa que, o *Big Data* representará para as organizações

um vasto campo de possibilidades e de estratégias que antes não eram acessíveis, e que agora, serão a base do planejamento estratégico.

“[...] as análises são feitas com maiores quantidades de dados e com técnicas melhores. Modelos descritivos, como as técnicas de Data Mining e Business Intelligence, permitem entender o passado e encontrar padrões no histórico de dados. Uma vantagem do volume é guardar históricos mais antigos para análise (ex. mudanças climáticas). A análise permite identificar padrões mais exatos e ordem nos dados com amostras maiores. O ser humano é bom em identificar padrões, mas quando há muitos dados, isto gera confusão e conclusões erradas. (STANLEY LOH, 2019)

1.4.1 – Os 5 V's do Big Data

1.4.1.1 – Variedade

Quando se fala em variedade, dentro de *Big Data*, diz respeito à imensa quantidade de dados que provêm de diferentes fontes e em diferentes formatos. Dados que podem ser úteis em uma empresa são gerados em seus sistemas internos, como também podem ser gerados e captados através das mídias sociais via PCs, *smartphones*, *tablets* e *smartwatches*, pelos objetos com tecnologia de internet das coisas, como sensores e demais objetos inteligentes. “Variedade porque estamos tratando tanto de dados textuais estruturados quanto não estruturados como fotos, vídeos, e-mails e tuítes” Cezar Taurion (2013).

Como os dados podem assumir diversos formatos, fazendo assim serem variados em sua tipagem, variadas também são a forma como esses dados podem ser interpretados (sejam eles apresentados em gráficos, mapas, sons, imagens, redes, etc). Enfim, é uma diversidade relativamente grande que pode, por vezes, dificultar no processo de interpretação de tais dados, dependendo do problema a ser resolvido, ou da questão a ser resolvida a partir de uma ação que vai ser tomada.

1.4.1.2 – Velocidade

Devido ao avanço e à rapidez da tecnologia que atualmente pessoas e organizações tem acesso, os dados (estruturados e não estruturados) são gerados e estão sendo atualizados a todo instante. Essa é uma operação muito rápida, instantânea, e é necessário que se tenha um tempo de reação baixíssimo quanto a estes eventos, caso contrário, no ambiente corporativo

como exemplo, não seria possível tratar estes dados em volumes massivos com o objetivo de obter vantagens competitivas.

Na sociedade da informação é crucial saber tratar os dados com velocidade, pois quando não é possível fazê-lo, estes dados tornam-se inúteis estrategicamente para as empresas pois não geram informação atual. É estimado que cerca de 90% dos dados disponíveis digitalmente não estejam sendo aproveitados da maneira como poderiam (Taurion, 2013). De acordo com Stanley Loh (2019), a velocidade dos dados ajuda a tomar sábias decisões, e de forma rápida, e permite também com que eventos que acontecem em um lado do mundo, impacte rapidamente o outro lado, em minutos. As pessoas, no mundo eletrônico, disseminam tais informações de forma voraz, aumentando assim o índice de chegada de informações que antes não era acessível de forma tão rápida.

1.4.1.3 – Valor

Todo o conceito que foi apresentado sobre o que é *Big Data* só faz sentido para as empresas, do ponto de vista estratégico e da tomada de decisão, se o valor da análise desses dados compensar o custo que é para coletá-los, armazená-los e processá-los. Portanto, os tomadores de decisão precisam saber aplicar suas ações baseadas no conhecimento proveniente dos dados analisados, com o intuito de agir corretamente e com assertividade.

Segundo Taurion (2013), para se fazer bom uso do *Big Data*, empresas precisam investir em conhecimento e em novas tecnologias, além de promover mudanças na estrutura da empresa. Estrutura essa que abrange cultura organizacional e pensamento entre departamentos, por exemplo. Os colaboradores e os diferentes níveis gerenciais de uma empresa (horizontalmente e verticalmente) precisam comunicar-se bem e de forma igualitária. É preciso que seja instaurada na organização uma cultura de dados, haver esta mudança no *mindset* da empresa. De acordo com Stanley Loh (2019), pouco adianta uma organização ou indivíduo possuir determinados dados, se os mesmos não puderem ser transformados em informação,

depois adquirir conhecimento à partir das informações, e por fim servirem para solucionar questões administrativas utilizando junto inteligência e estratégia.

1.4.1.4 – Volume

De acordo com Stanley Loh (2019), a primeira característica do fenômeno pós-moderno conhecido como Big Data é seu grande volume, pois a quantidade de dados melhora o momento da ação/tomada de decisão.

Entretanto, é necessário manusear estes dados de forma assertiva e com muito cuidado, para que não haja erro proveniente de análises incorretas e falta de foco. Existe sim, uma dificuldade em manusear imensos volumes de dados, mas, ferramentas adequadas e mais modernas estão surgindo para tornar esta experiência de dados mais fácil para as organizações. Ainda segundo Stanley Loh (2019), o volume dos dados aumenta a cada ano pelas seguintes razões: o armazenamento de dados, atualmente, custa pouco. Existem diversas formas para armazenar dados, através de discos rígidos, DVDs, *pendrives*, dentre outras. Além disso, existem serviços grátis de armazenamento em nuvens, como Dropbox, Google Drive, One Drive e etc. Somado a isto, as pessoas estão constantemente se acostumando a utilizar desta e de outras tecnologias que, conseqüentemente, geram e armazenam mais e mais informações.

1.4.1.5 – Veracidade

Para que se possa ter certeza no momento de se basear em dados a fim de tomar uma medida, é preciso que se tenha ciência de que estes dados são verídicos, e é estritamente necessário que eles façam sentido. Do contrário, de nada adianta todo o trabalho e investimento na coleta, armazenamento, processamento e análise destes dados. Quando uma informação pode ser confiável? As fontes desses dados, são confiáveis? Das inúmeras informações disponíveis no mundo moderno de hoje, quantas delas são verídicas? Em um ambiente repleto de *fake news*, essas são perguntas recorrentes quando se trata de dados e nas vantagens que podem ser adquiridas se houver uma correta coleta, uma boa análise e bons *insights* gerados para facilitar a tomada de decisões do gestor de uma organização (Stanley Loh, 2019).

Os 5 V's do *Big Data* cada vez mais tem impactado a vida cotidiana da sociedade, mas acima de tudo, das organizações. Analisar dados para ter embasamento e aplicá-lo em uma correta tomada de decisão tem se tornado cada vez mais uma atividade complexa no meio

corporativo, uma vez que as características dos dados citadas acima estão cada vez mais presentes e atuantes.

2 – Ciência de Dados (*Data Science*)

Pode-se definir Ciência de Dados como um conjunto de técnicas, ferramentas e procedimentos provenientes de diferentes áreas do conhecimento a fim de analisar dados. Habilidades estas que, no contexto de *Data Science*, são utilizadas para analisar dados previamente coletados, interpretá-los e extrair conhecimento dos mesmos, a fim de entregar resultados que vão ajudar tomadores de decisão dos mais diversos segmentos.

Tomada de decisão está intimamente ligada a dados, uma vez que com os dados, a probabilidade de uma ação a ser tomada se tornar bem sucedida é muito maior do que tomá-las sem nenhum embasamento.

“Em um nível mais elevado, *Data Science* é um conjunto de princípios fundamentais que norteiam a extração de conhecimento a partir de dados”. (PROVOST e FAWCETT 2016, p. 34). Este trabalho busca salientar que o principal objetivo de *Data Science* é ser o ponto chave para que as empresas possam melhorar suas ações para obter vantagens de negócio, e como isso poderá ser feito.

No contexto organizacional, os dados precisam ser coletados da forma correta, pois toda essa operação custa tempo e dinheiro, recursos estes imprescindíveis para qualquer organização lucrativa. Não se deve contratar um profissional de dados e solicitar que sejam coletados dados sem nenhum tipo de classificação, pois além de demorar mais isso seria desnecessário.

Antes de tudo, é preciso ter definido o problema de negócio ou qualquer situação onde dados possam ser analisados para ajudar na tomada de decisão. Com o problema bem definido previamente é que se vai em busca da coleta dos dados que se referem à situação designada.

O principal objetivo da análise de dados é interpretar grandes quantidades de dados, a fim de extrair informações verídicas e organizadas dos mesmos, que serão transformados em conhecimento, que por sua vez irão ser necessários para uma boa tomada de decisão por parte dos executivos.

2.1 – Processo Orientado a Dados

Tudo se resume a ciência. Ciência é ter certeza daquilo que se está em mente. Portanto, toda organização seja ela lucrativa ou não lucrativa, necessita ter ciência daquilo que os rodeia para saber exatamente o que fazer, quais passos seguir. Por isso, os dados são valiosíssimos para que os indivíduos enxerguem claramente o ponto onde precisam estar no futuro. E neste

conceito, precisam ser o ponto de partida no momento dos processos organizacionais e das tomadas de decisão.

Tomada de decisão orientada por dados (DOD) refere-se à prática de basear as decisões na análise dos dados, em vez de apenas na intuição. Por exemplo, um negociante poderá selecionar anúncios baseados puramente em sua longa experiência na área e em sua intuição de que funcionará. Ou, pode basear sua escolha na análise dos dados sobre a forma como os consumidores reagem a diferentes anúncios. Ele também poderia utilizar uma combinação dessas abordagens. A DOD não é uma prática do tipo “tudo ou nada”, e diversas empresas a adotam em maior ou menor grau. (PROVOST e FAWCETT 2016, p.41)

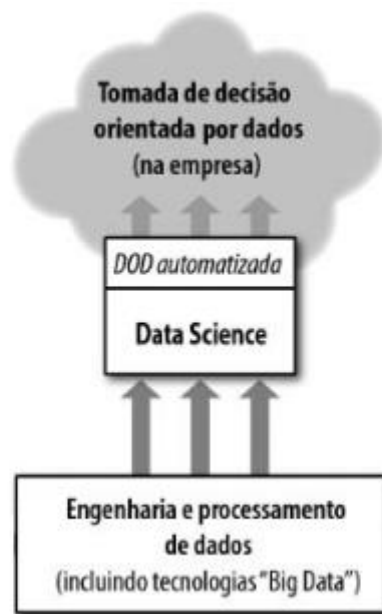
Ser uma organização que utiliza dados para sua tomada de decisão é algo que vem numa crescente exponencial dentre as organizações, visto que cada vez mais novas tecnologias de armazenamento e processamento de dados são desenvolvidas, empresas que estão migrando para o âmbito digital e o aumento de dados que estão sendo gerados a cada instante. Implementar um modelo de *data driven* ou se tornar uma empresa capaz de tomar decisões orientadas por dados, é algo que demanda de desafios.

Se faz necessário implementar uma cultura organizacional à base de dados, para que todos os setores da organização possam se comunicar adequadamente sem que um atrapalhe o outro, fazendo assim com que os processos internos possam continuar sendo executados da forma como devem.

Investimento em inteligência artificial e o advento da internet das coisas são alguns dos aspectos fundamentais que as empresas precisam estar de olho daqui para frente. Afinal, o

Big Data está longe de atingir todo o seu potencial, e num mercado cada vez mais competitivo, organizações precisam evoluir e se adaptar ao mercado e ao mundo, de forma geral.

Figura 1 - Data science no contexto da tomada de decisão



Fonte: Adaptado de Provost e Fawcett (2016).

Em um mundo cada vez mais competitivo e rápido, empresas precisam tomar suas decisões com base em dados concretos, e não apenas no achismo ou baseado em simples experiências relacionadas. Neste sentido, para o setor de marketing de uma organização por exemplo, é crucial ter uma visão 360° de um consumidor, olhando não somente o que ele comprou ou consumiu de determinada empresa, mas também o que ele pensa e diz sobre a organização pela internet à fora (Taurion, 2013).

2.2 – Banco de Dados em Nuvem

Para que se torne possível armazenar, processar e gerir a imensa quantidade de dados que uma empresa pode dispor, é de extrema importância ter capacidade de *hardware* e *software* para isto. A partir desta premissa, entra em contexto a computação em nuvem, que é responsável por uma mudança drástica no modelo de como as organizações lidam com seu ativo intelectual, sendo responsável também por uma revolução no mundo tecnológico das corporações.

Ainda é um tema que está em seu estágio inicial, contudo, é notável que já apresenta diversas vantagens para as organizações, como por exemplo um custo bastante baixo (ou até gratuito), indo de acordo com a demanda. A computação em nuvem está em constante

desenvolvimento, oferecendo serviços com cada vez mais segurança, efetividade e de certa forma, descomplicados. É uma verdadeira solução de TI para que empresas se modernizem e se atualizem conforme o mercado e as exigências que o mesmo as impõe.

A computação em nuvem e seus bancos de dados, inserida numa organização pode atingir todos os aspectos da mesma, facilitando o manuseio de suas informações, administrando o catálogo de clientes já existentes, buscando atingir clientes potenciais, potencializa o método de oferecer seus produtos ou serviços, facilita a compreensão de alcance que a empresa possui e relaciona seus diferenciais competitivos perante as demais empresas do mercado, enfim.

Portanto, empresas de pequeno porte agora podem ter as mesmas oportunidades que as empresas de grande porte, uma vez que esta tecnologia está a disposição de todos. E isto muda drasticamente as coisas, pois agora empresas pequenas podem competir com as maiores, pois tem a mesma capacidade e potência tecnológica que elas, sem que tenha sido preciso investir recursos em grandes infraestruturas físicas de TI.

Com a nuvem a disposição das empresas, alguns fatores se integram e são modificados e passam também a serem controlados. É necessário que se tenha conhecimento em TI e também na área de negócio para poder existir uma comunicação eficaz entre estes setores, integrando também todos os outros: produção, finanças, recursos humanos, etc.

Outro fator que é amplamente atingido são as questões de marketing e de divulgação de produtos e da marca, devido à facilidade de alcançar o público alvo e a computação em nuvem auxilia no direcionamento destas informações na rede. “E com a computação em nuvem, não é necessário dispor de um imenso parque computacional dentro de casa. Pode-se fazer esta análise em nuvens públicas e pagar-se apenas pelo consumo de recursos da análise.” (Taurion, 2013)

É primordial e impreterível reafirmar que o custo desta tecnologia não é excessivo, pois os mesmos são variáveis. Vão de acordo com a utilização de recursos que a empresa irá utilizar. Desse modo, micro e pequenas empresas podem ter acesso a esta tecnologia por preços

muito mais acessíveis do que grandes corporações, pois nas grandes sua utilização também é ampliada, gastando assim, obviamente, mais recursos para se ter a disposição.

“[...] ao possibilitar que todas as empresas, independentemente de seu porte, façam uso dos mesmos recursos tecnológicos e da mesma infraestrutura [...] a Computação em Nuvem possibilita que todas possam competir – o que não acontecia antigamente” [8]

2.3 – Business Intelligence

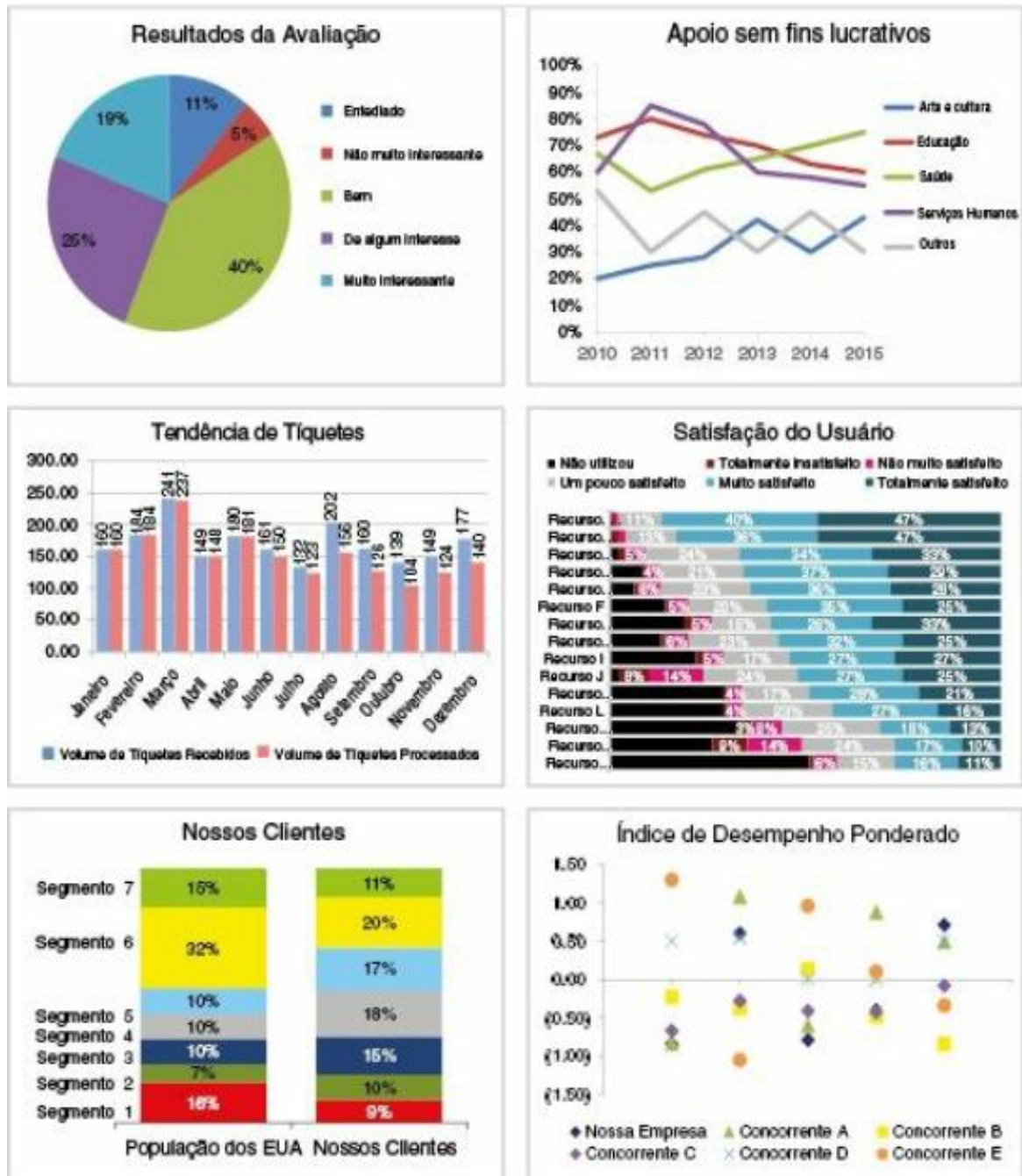
É primordial que se aplique *Big Data* nas empresas com o objetivo de obter vantagens competitivas, e para isso, é essencial coletar dados que a empresa já possui. Dados históricos, dados que podem ser localizados em fontes internas de armazenamento, nos sistemas de gerenciamento que constam dados de vendas, histórico de clientes, por exemplo. Somado a isso, com a ajuda de profissionais capacitados na área de ciência de dados, é preciso também coletar dados de fontes externas (onde é encontrado a maioria destes dados), quase sempre no formato não estruturado, que se é possível encontrar em quantidade muito maior do que os estruturados, devido a ampla usabilidade da rede por parte dos indivíduos, gerando inúmeros dados a todo instante. “[...] criar e tratar apenas de dados históricos [...], começa a se mostrar um processo lento demais para a velocidade com que os negócios precisam tomar decisões” Cezar Taurion (2013).

Mais importante que isso, as organizações precisam entender a importância e o valor de seus dados. E isso se alcança com inteligência de negócios e implantando uma cultura baseada na coleta e embasamento em dados, desde o topo da pirâmide organizacional até a base da mesma. *Business Intelligence* e *Big Data* estão interligados; portanto apenas coletar os dados e não saber trabalhar com eles significa ter uma organização incompleta do ponto de vista da inteligência corporativa. “Big Data tem uma abrangência muito maior que os projetos de BI que as empresas estão acostumadas a desenvolver. BI concentra-se na análise de dados gerados

pelos sistemas transacionais enquanto *Big Data* vai além, explorando fontes de dados externas [...]” (Taurion, 2013).

Neste contexto, *Big Data* e *Business Intelligence* precisam ser utilizados de forma complementar uma à outra. Assim, se encaixam ferramentas de BI para auxiliar a amostragem de insights coletados aos líderes e tomadores de decisão.

Figura 2 - Exemplos de gráficos ineficientes

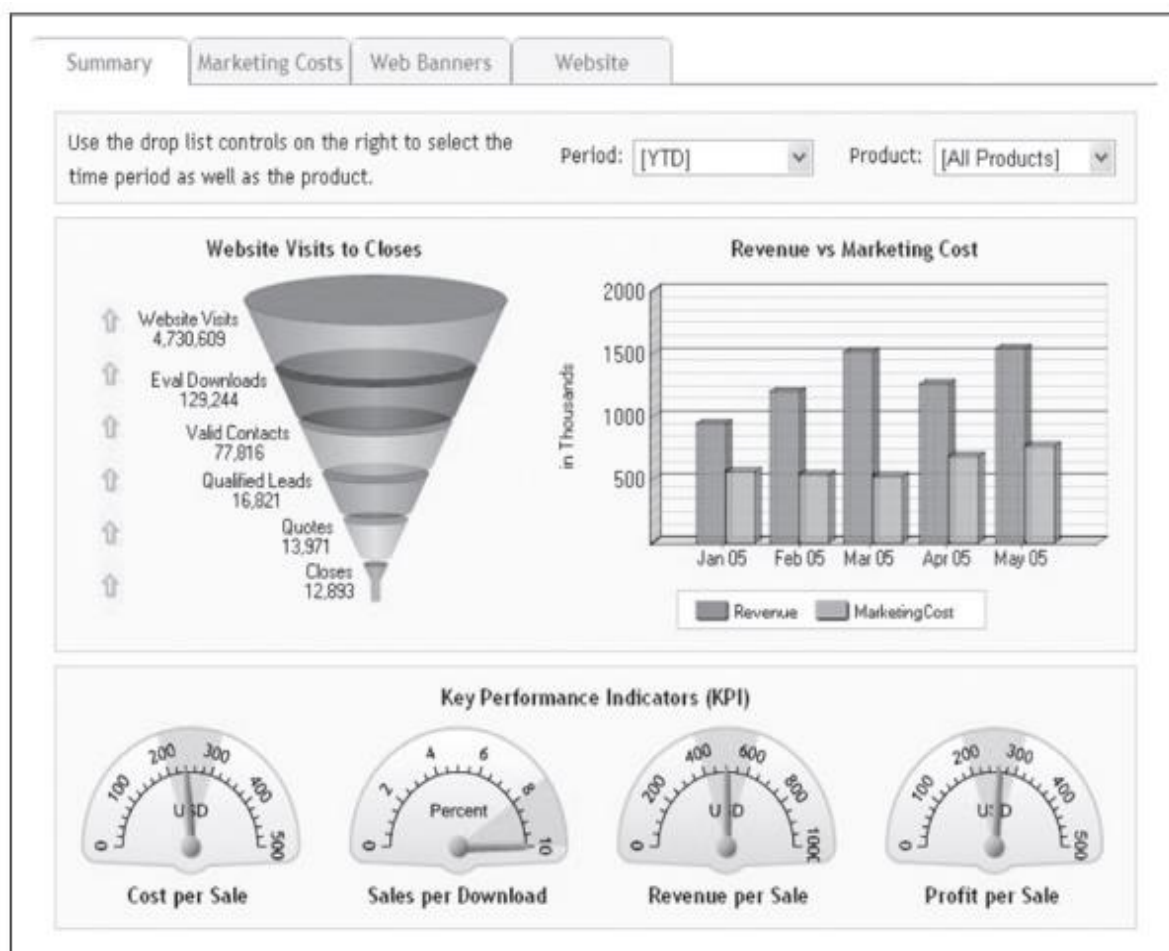


Fonte: Cole Nussbaumer Knaflie (2019).

Acima pode-se observar alguns exemplos de informações visuais que muitas vezes não auxiliam devidamente o processo de transformação desses dados em conhecimento por parte dos tomadores de decisão. Portanto, algumas outras técnicas mais atuais se tornam extremamente necessárias para ilustrar de forma mais coerente e facilitada de tudo aquilo que rodeia a organização.

E nisso entram formas de visualizações e de análise de negócios mais modernas provenientes de softwares mais bem elaborados para esta finalidade, que oferecem formatos de visualização interativa com recursos de *business intelligence* em uma interface de fácil compreensão, para que os próprios usuários possam analisar e elaborar seus relatórios e *dashboards*. Isto é chamado de Self-Service BI.

Figura 3 - Exemplo de dashboard de desempenho



Fonte: Efraim Turban (2009)

Portanto, este trabalho como um todo reflete em inteligência de negócios. Tudo está interligado. *Big Data* fornece material para a execução de *data science*, que por sua vez concede a tomadores de decisão e gestores de organizações embasamento teórico e verídico que

auxiliará em momentos de tomada de decisão. Tudo isto é *business intelligence*; planejamento, organização, direção e controle. BI é a chave para que empresas de variados portes possam obter sucesso em suas ações e em seus modos de administração.

3 – Benefícios da análise dos dados para a tomada de decisões

Partindo do princípio da questão, uma organização precisa passar pelo processo de identificação do problema de negócio ou a decisão a ser tomada. Este é o processo chamado de formulação, que é um dos principais componentes dentro do processo decisório (Thomas Davenport e Jin Ho-Kin, 2014). Um gestor pode identificar um problema ou uma questão a ser resolvida de variadas formas. Pelo cotidiano, apenas observando os processos; pelo pensamento analítico, analisando o mercado e os concorrentes, por exemplo. Estar sempre pensando no desenvolvimento da organização também traz à tona situações que demandam de atenção por parte da alta gerência.

Será necessário também, identificar os *stakeholders* deste processo decisório, pois claramente eles fazem parte do problema identificado, ou podem executar ações que vão influenciar na mudança de algum processo organizacional, que vão influenciar na ação a ser tomada. Com base no problema traçado e na identificação dos agentes que farão parte da solução da questão, a organização deve se encaminhar para a solução desta questão, empregando conceitos de *data science* além de habilidades quantitativas e analíticas.

Os dados também podem ser coletados e separados através de algoritmos preparados justamente para esse fim, utilizando técnicas como *machine learning* e *data mining*, então variáveis são formuladas e processadas por ferramentas de TI.

Até o lado mais escondido dos dados já está sendo, cada vez mais, alvo de interessantes aplicações. É o conceito de data mining (garimpagem de dados), que objetiva melhorar o uso desses gigantescos arsenais de informação através da identificação de padrões de correlação normalmente invisíveis em análises convencionais. (CARLOS BARBIERI, 2011)

Segundo Thomas Davenport e Jin Ho-Kin (2014), o último estágio deste procedimento é a comunicação dos resultados das análises aos *stakeholders* a fim de adotarem providências para sanar a questão ou o problema. É de extrema importância comunicar com eficiência os métodos e as prováveis soluções aos tomadores de decisão, sugerindo pensamentos e apresentando-lhes caminhos, dando embasamento teórico-prático para auxiliá-los no processo decisório.

Conforme Thomas Davenport e Jin Ho-Kin (2014) e Cole Nussbaumer Knaflitz (2019), é preciso contar histórias com estes dados, apresentando descobertas, mensurações hipóteses e possíveis ações, analisando tudo juntamente à questão abordada no início do processo.

O conceito de empresas que utilizam *big data* para desenvolver todas as técnicas e teorias que foram expostas no presente trabalho está aumentando, pois percebe-se no meio

corporativo uma necessidade de agir à base de dados, então é fundamental pensar sempre com inteligência corporativa e fazer com que empresas possam evoluir junto com as tecnologias e habilidades de negócios.

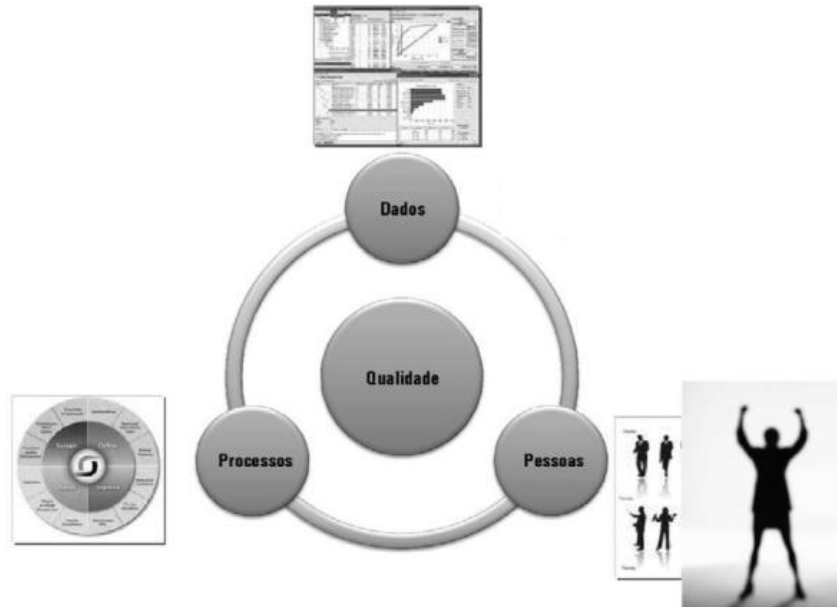
Dados de projetos, de blogs e de redes sociais, e de comportamento das pessoas nos seus mais variados papéis, vêm para se juntar ao mundo do varejo, onde tudo começou. O mundo do BI2 também falará de BA-Business Analytics, que misturará o BI em tempo real com mining e análises preditivas. Isso permitirá a criação de mecanismos inferenciais de negócios com latência quase zero e exigirá maior conteúdo qualitativo das informações. Como vimos, de início, a informática fez os dados. Depois transformou-os em informação. Agora o objetivo é usar conhecimentos, a partir daquelas matérias-primas (CARLOS BARBIERI, 2011).

O gerenciamento do conhecimento é uma modalidade que está sendo implantada para capturar e manusear todo conhecimento das organizações. Já estão sendo desenvolvidos métodos de acumulação e administração de conhecimento estruturado, conceito extremamente novo no âmbito organizacional, mas que as empresas da atualidade já praticam, mesmo que de forma inicial e muitas vezes ineficaz. Mas, de fato, as organizações precisam começar desde já a voltar seus olhares para dados, informação, conhecimento e capacidades analíticas para serem inovativas e capazes de se planejar e decidir daqui por diante.

De acordo com Carlos Barbieri (2011), o alicerce fundamental de uma organização bem sucedida, na era da informação, são os processos, dados e as pessoas, compondo o capital intelectual das organizações. O *peopleware* ou simplesmente recurso humano é um elemento-

chave, juntamente com a análise dos dados e os *insights* obtidos através desta análise para que uma corporação tenha qualidade em seus processos.

Figura 4 - Processos, dados e pessoas: capital intelectual das empresas.



Fonte: Adaptado de Barbieri (2011).

Diversas questões no meio corporativo podem ser levantadas para uma possível solução. Qual o produto/serviço que os clientes mais consomem? Qual o público alvo para determinado produto? Como atingir corretamente o cliente com as campanhas de marketing? Determinado processo em uma empresa, é um fator que atrai ou que repele clientes? Todas essas questões são exemplos de perguntas que podem ser respondidas através de análise de dados dentro do ambiente organizacional. Este é o benefício por trás da tomada de decisão a base de dados, pois estas decisões serão baseadas em detalhamentos de informações para entender o problema e solucioná-lo da maneira mais correta.

Conclusão

Uma organização empresarial não tem um cotidiano simplista e repetitivo. Cada dia é diferente do outro, situações que demandam de atenção surgem constantemente, e outras questões estão em constante mudança. É necessário enfrentar os desafios do mercado, sua competitividade e os obstáculos que surgem no decorrer dos processos organizacionais e manter-se em constante evolução, adquirindo novas tecnologias, pensamentos estratégicos e inovar para acompanhar a velocidade constante de mudanças no meio empresarial e adaptar-se às novas exigências de mercado para que se possa ter assertividade nas ações, fazendo assim com que organizações possam ter longevidade e se tornar referência no que tange a seu modelo de administração e seu modo de agir.

Por vezes as tomadas de decisão providas da alta gerência era feita unicamente levando em consideração experiências já vividas por estes indivíduos e com base também na intuição. De fato, pode sim ser possível acertar em uma ação sob esses aspectos. Mas, é muito mais fácil errar do que acertar, pois seriam decisões tomadas no achismo, sem ciência, simplesmente pelo viés do palpite. Devido a isto, essa forma de decidir está cada vez mais em desuso no segmento das empresas, por esses e inúmeros outros fatores.

A constante evolução da tecnologia permite que conhecimentos dos mais avançados cheguem rapidamente até as mais diversas áreas do planeta. Com isso, a rapidez de uma mudança neste pensamento é algo urgente para que empresas comecem a olhar para isto o mais rápido possível.

As pessoas estão tendo acesso a muita informação no ambiente digital, e também acabam tendo suas informações também captadas a todo instante. Cabe às organizações saber gerenciar estes dados da melhor forma possível, sem ferir a privacidade dos mesmos, podendo com isso se basear em fatores concretos no momento da tomada de uma decisão.

Na era da informação em que as empresas atuam nos dias de hoje, faz-se necessário minimizar os erros para se atingir os resultados desejados. E para isso, os olhos precisam estar voltados para a ciência de dados.

Bibliografia

[8] O QUE VOCÊ REALMENTE PRECISA SABER SOBRE COMPUTAÇÃO EM NUVEM
- São Paulo - Opus Software - 2015

BARBIERI, Carlos. **BI2 - Business Intelligence: Modelagem e Qualidade**. Rio de Janeiro: Elsevier, 2011.

DAVENPORT, Thomas H. & HO-KIM, Jin. **Dados Demais!: Como desenvolver habilidades analíticas para resolver problemas complexos, reduzir riscos e decidir melhor**. 1ª Edição. Rio de Janeiro: Elsevier, 2014.

KNAFLIC, Cole Nussbaumer. **Storytelling com dados: Um guia sobre visualização de dados para profissionais de negócios**. 1ª Edição. Rio de Janeiro: Alta Books, 2019.

LOH, Stanley. **Volume, velocidade, variedade, veracidade e valor: como os 5 Vs do big data estão impactando as organizações e a sociedade**. 1ª Edição. Porto Alegre, 2019.

PEREIRA, Vanessa Alves da Silva. **Big Data: Um estudo em gestão empresarial**. 2016. 17. Trabalho de Conclusão de Curso - Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2016.

PROVOST, Foster & FAWCETT, Tom. **Data Science para Negócios: O que você precisa saber sobre mineração de dados e pensamento analítico de dados**. 1ª Edição. Rio de Janeiro: Alta Books, 2016.

TAURION, Cezar. **Big Data**. Rio de Janeiro: Brasport, 2013.

TURBAN, Efraim. **Business Intelligence: Um enfoque gerencial para a inteligência do negócio**. 1ª Edição. Porto Alegre: Bookman, 2009.