



ARTIGO CONVIDADO



REFLEXÕES PÓS-PANDEMIA SOBRE DESAFIOS E OPORTUNIDADES PARA PESQUISA DE MARKETING NO SÉCULO XXI

Joe F. Hair Jr.¹ | Dana E. Harrison² | Jeffrey J. Risher³

¹University of South Alabama, Alabama, EUA.

²East Tennessee State University, Tennessee, EUA.

³Southeastern Oklahoma State University, Durant, EUA.

Corresponding Author:

Joe F. Hair Jr.

E-mail: jhair@southalabama.edu

Editor in Chief

PhD. Eliana Andréa Severo
Universidade Federal de Pernambuco (UFPE)

Editor Executivo

Altieres de Oliveira Silva
Alumni.In Editors - UNICURITIBA

Tradutora

DSc. Alessandra Yula Tutida
UNIVALI - ABRATES (Associação Brasileira de Tradutores Intérpretes)

O artigo foi convidado pelo Editor Executivo - Altieres de Oliveira Silva - ao autor Joe F. Hair Jr. sobre o estudo inédito.

How to cite this article:

Hair Jr., J. F., Harrison, D. E. & Risher, J. J. (2022). Post-Pandemic Reflections on Challenges and Opportunities for Marketing Research in the 21st Century. *Iberoamerican Journal of competitive intelligence*, 12(1), e0412.

RESUMO

O papel do marketing está evoluindo rapidamente, e os métodos de design e análise usados pelos pesquisadores de marketing também estão mudando. Essas mudanças estão surgindo de transformações nas habilidades de gestão, inovações tecnológicas, evolução contínua do comportamento do cliente e, mais recentemente, da pandemia de Covid-19. Mas talvez o impulsionador mais substancial dessas mudanças seja o surgimento da big data e os métodos analíticos usados para examinar e entender os dados. Para continuar sendo relevante, a pesquisa de marketing deve permanecer tão dinâmica quanto os próprios mercados e se adaptar de acordo com o seguinte: os dados continuarão aumentando exponencialmente; a qualidade dos dados melhorará; a análise será mais poderosa, mais fácil de usar e mais amplamente utilizada; as decisões de gestão e clientes serão cada vez mais baseadas no conhecimento; questões e desafios de privacidade serão um problema e uma oportunidade à medida que as organizações desenvolvem suas habilidades analíticas; a análise de dados se estabelecerá firmemente como uma vantagem competitiva, tanto no setor de pesquisa de marketing quanto no meio acadêmico; e no futuro próximo, a demanda por cientistas de dados altamente treinados excederá a oferta.

Palavras-chave: big data, análise de dados, qualidade de dados, análise de marketing, investigação de marketing.

DOI: <https://doi.org/10.24883/IberoamericanIC.v12i.2022.e0411>

This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Iberoamerican Journal of Competitive Intelligence* published by Alumni.In.



POST-PANDEMIC REFLECTIONS ON CHALLENGES AND OPPORTUNITIES FOR MARKETING RESEARCH IN THE 21ST CENTURY

ABSTRACT

The role of marketing is evolving rapidly, and design and analysis methods used by marketing researchers are also changing. These changes are emerging from transformations in management skills, technological innovations, continuously evolving customer behavior, and most recently the Covid-19 pandemic. But perhaps the most substantial driver of these changes is the emergence of big data and the analytical methods used to examine and understand the data. To continue being relevant, marketing research must remain as dynamic as the markets themselves and adapt accordingly to the following: data will continue increasing exponentially; data quality will improve; analytics will be more powerful, easier to use, and more widely used; management and customer decisions will increasingly be knowledge-based; privacy issues and challenges will be both a problem and an opportunity as organizations develop their analytics skills; data analytics will become firmly established as a competitive advantage, both in the marketing research industry and in academics; and for the foreseeable future, the demand for highly trained data scientists will exceed the supply.

Keywords: big data, data analytics, data quality, marketing analytics, marketing research.

1 | INTRODUÇÃO

O papel do marketing para consumidores e empresas está evoluindo rapidamente (Ferrell, Hair, Marshall & Tamlia 2015). De fato, a função de marketing está mudando fundamentalmente como resultado da transformação digital, dados, análise de dados e, mais recentemente, dispositivos móveis pessoais, como software de infusão, que rastreiam o tempo enquanto os indivíduos trabalham ou se divertem, podem pesquisar empresas ou clientes locais enquanto estão em uma cafeteria, enviar uma fatura ou cotação do escritório de um cliente e muito mais (Shah et al 2014). Os métodos de design e análise de pesquisa de marketing também estão mudando rapidamente (Hair, Black, Babin & Anderson 2019). Essas mudanças estão surgindo de transformações nas habilidades de gestão (Henke, Levine & McInerney 2018; Davenport

2018), inovação tecnológica, muitas vezes no espaço de marketing digital (Davenport 2018), e evolução contínua do comportamento do cliente, principalmente resultante de mudanças nas necessidades e expectativas (Wedel & Kannan 2016), mais recentemente impactado pela pandemia de Covid-19. Para continuar sendo relevante e eficaz, a pesquisa de marketing deve permanecer tão dinâmica quanto os próprios mercados.

Ao mesmo tempo, a pesquisa acadêmica está aplicando cada vez mais ferramentas e algoritmos analíticos avançados (Erevelles, Fukawa & Swayne 2016; Syam & Sharma 2017, Wedel & Kannan 2016; Petrescu & Krishen, 2017), e até mesmo inteligência artificial em combinação com redes neurais (Ababukar et al. 2017). Assim como a análise avançada está fornecendo vantagens competitivas importantes nas práticas do setor, também está forçando mudanças rápidas na pesquisa acadêmica de ciências sociais (Krishen & Petrescu 2018).

Direções para pesquisas, particularmente marketing e ciências sociais relacionadas, incluem aplicações de técnicas e tecnologias de análise avançada. Wedel e Kannan (2016), por exemplo, sugerem como métodos de *machine learning* e tecnologias de computação cognitiva podem ser aplicados para entender melhor os problemas e oportunidades de marketing. Da mesma forma, Erevelles, Fukawa e Swayne (2016) identificam proposições que examinam soluções de big data para diversas atividades de marketing, bem como como essas descobertas podem ser aplicadas para desenvolver vantagens competitivas sustentáveis. Syam e Sharma (2017) destacam a importância de entender como o *machine learning* (ML) e a *artificial intelligence* (AI) podem avançar na pesquisa de marketing, e (Abadular et al. 2017) propõem maneiras de aplicar AI e redes neurais. Assim, à medida que esses e outros desenvolvimentos paralelos são aplicados para expandir nosso conhecimento sobre tendências de marketing e comportamento do consumidor, as empresas podem atender com mais eficácia às necessidades dos consumidores e reduzir a sobrecarga de informações.

O que não foi antecipado quando inicialmente compartilhamos nosso pensamento sobre o papel emergente da pesquisa de marketing no século 21 (Hair, Harrison & Risher, 2018) foi a influência subsequente da pandemia de Covid 19 e os desenvolvimentos econômicos e sociais relacionados. Esta versão atualizada do nosso artigo anterior inclui comentários sobre desenvolvimentos mais recentes na pesquisa de marketing, bem como impressões iniciais de como a pandemia está afetando a pesquisa de marketing.

O objetivo deste artigo, portanto, é sugerir como as tendências de mercado emergentes, e particularmente as de marketing, provavelmente impactarão a pesquisa de marketing. Primeiro, exploramos o papel emergente da função de marketing que está se desenvolvendo a partir de novos modelos de negócios. Em segundo lugar, examinaremos as implicações dessa mudança na pesquisa acadêmica de marketing. Terceiro, abordaremos as transformações que provavelmente continuarão evoluindo no futuro. Por fim, incluiremos reflexões sobre os desenvolvimentos mais recentes que impactam a pesquisa de marketing, como a pandemia de Covid-19 e as consequentes interrupções na cadeia de suprimentos. Nossa estrutura para esses comentários é um artigo que publicamos antes desses eventos que não antecipava esses desenvolvimentos recentes (Hair, Harrison & Risher, 2018).

2 | BASE TEÓRICA

2.1 | O Papel emergente do marketing e sua pesquisa

Inúmeros desenvolvimentos estão influenciando o marketing e, deste modo, a pesquisa de marketing. Entre os desenvolvimentos mais importantes estão os seguintes:

2.1.1 | *Mercados modernos*

Talvez uma das tendências mais importantes que impactam a pesquisa de marketing sejam os mercados baseados na Internet. Modelos de negócios como a economia de compartilhamento ou empresas de correspondência digital surgiram de plataformas de tecnologia avançada e se tornaram mercados altamente bem-sucedidos e conectados digitalmente (Harrison & Hair 2017). Exemplos incluem empresas baseadas na Internet, como Uber, Amazon, Airbnb, DoorDash, Netflix, bem como serviços de armazenamento e computação em nuvem disponíveis em Salesforce.com, Microsoft Azure, IBM e Amazon Web Services, que criaram modelos de negócios bem-sucedidos, finalmente revolucionando as indústrias. Embora esses desenvolvimentos tenham se materializado gradualmente nos últimos 20 anos, seu impacto total está se tornando evidente.

Os clientes e as empresas tornaram-se coletivamente menos avessos aos riscos frequentemente associados às transações online, e isso impulsionou o crescimento desses modelos de negócios emergentes. Além disso, o comportamento dinâmico do cliente e as mudanças do mercado contribuíram para a necessidade de maior conveniência e experiências mais personalizadas do que inicialmente oferecidas pelos mercados online. Por exemplo, tendências que se movem mais rapidamente em modas de vestuário é um dos muitos desenvolvimentos de mercado que podem ser atribuídos às capacidades de produtos e serviços baseados na Internet. Mas há muitos outros exemplos. Em essência, as empresas inovadoras desenvolveram e dominaram propostas de valor que respondem de forma mais eficaz às oportunidades de mercado e atendem mais de perto aos desejos e necessidades dos clientes.

Um componente crítico que surgiu junto com essas novas propostas de valor são os dados digitais, que permitiram que empresas e consumidores não apenas estivessem mais informados do que nunca, mas também vendessem e buscassem quase todos os produtos e serviços totalmente online. Por exemplo, se você tiver uma conta My Google, ela está rastreando suas atividades 24 horas por dia, 7 dias por semana. Um indivíduo verificou o número de itens rastreados durante um período e descobriu que, em média, mais de 200 atividades foram adicionadas à sua conta My Google por dia, ou seja, mais de 73.000 por ano. Quem teria pensado há alguns anos que os indivíduos não apenas compartilhariam seus dados pessoais, mas bicicletas e carros em vez de comprá-los, sem mencionar que estariam dispostos a dar a organizações como Google ou Facebook acesso a grandes quantidades de seus dados pessoais, ou usar o celular para pesquisar e comprar mantimentos, ou quase qualquer outro produto ou serviço que deseja que seja entregue à sua porta.

2.1.2 | *Surgimento de Big Data*

Empresas e clientes estão produzindo grandes quantidades de dados de vários

pontos de contato. Do lado do cliente, os dados são capturados por meio de experiências omnicanal altamente promovidas. Como exemplos, considere a compra on-line da Target e a retirada imediata na loja, a associação principal da Amazon e as relações de vendas e remessa associadas, ou o uso da metodologia *blockchain* da FedEx para gerenciar dados de disputas de clientes e como um método eficiente para rastrear pacotes, todos estão produzindo montanhas de informações úteis para o lado comercial e do consumidor das transações. Talvez o coletor de dados mais difundido seja o Google, que tem rastreadores em quase 80% de todos os sites (exceto na China, de onde o Google se retirou porque se recusou a permitir a censura de informações). Da mesma forma, no lado *B-to-B*, General Electric e Boeing estão instalando sensores em motores a jato, robôs estão sendo coordenados para fabricar muitos produtos, como automóveis e aviões, sensores estão habilitando marcas de veículos como Volvo, Tesla, Mercedes, BMW, e outros para comercializar carros autônomos e muitas outras inovações de distribuição e fabricação, todas também coletando dados em todo o ecossistema de fabricação e distribuição para gerar insights úteis e valiosos. Como resultado, as empresas podem gerenciar melhor a produção, entender e responder aos requisitos dos clientes e agregar valor. Com esse recurso emergente de dados, eles podem gerenciar seus negócios de maneira mais inteligente, melhorar os tempos de resposta, promover a inovação, reduzir custos e aumentar as receitas e os lucros.

Os dados alteram as relações de poder dentro da organização. Hoje, apenas cerca de dez por cento da economia dos EUA é a economia de dados - pequena agora, mas a porcentagem está se expandindo exponencialmente. No passado, os dados eram de responsabilidade do *Chief Information Officer* (CIO), um cargo da equipe. No ano passado (2021) nos EUA, a tendência de uma proporção maior de orçamentos de tecnologia da informação sendo gastos por gerentes de linha do que CIOs continuou. Para os CIOs, a segurança dos sistemas de dados continua sendo uma alta prioridade. Mas os CMOs e CFOs se concentram mais na execução dos negócios. A violação de dados da Equifax há vários anos foi provavelmente resultado dessa transição na responsabilidade pela segurança dos dados. Ao mesmo tempo, para indivíduos treinados na faculdade de administração, é provável que seja surpreendente que mais de 15% dos CEOs sejam agora cientistas de dados - eles não possuem MBA!

Dados e análise de dados têm o potencial de mudar a ordem econômica do mundo de maneiras que serão perturbadoras para muitas empresas e pessoas. Além disso, os dados baseados em conhecimento provavelmente se tornarão uma fonte de poder difícil de desafiar para países e empresas que não possuem esse conhecimento e essas habilidades. Em suma, dados e análises tornaram-se para as economias do século 21 o que o petróleo foi para o século 20 - um ativo valioso essencial para uma vida econômica melhor. Isso foi evidenciado com a adição da Salesforce.com (CRM) ao Dow Jones Industrial Index em 2020, não apenas substituindo a Exxon-Mobile, mas marcando a primeira vez na história da Bolsa de Valores de Nova York que um CRM/Marketing A empresa de pesquisa foi considerada relevante por representar uma parcela substancial de toda a atividade industrial. Parece que estamos à beira de um novo mundo de dados, que definitivamente mudará tanto o marketing quanto sua pesquisa.

No lado da análise de dados, tecnologia avançada e custos mais baixos agora são capazes de armazenar e processar grandes quantidades de dados dessas várias interações. Os custos de armazenamento de um megabyte de informação, por exemplo,

caíram de US\$ 15,00 em 1992 para menos de 10 centavos de dólar por um megabyte em 2021. Mas, à medida que mais dados são produzidos e armazenados, quase todos os dados localizados no armazenamento em nuvem, há uma necessidade maior de aplicar métodos avançados de análise de marketing. Esses métodos emergentes permitem que os pesquisadores de marketing obtenham uma compreensão mais sofisticada dos dados, não apenas relacionados ao que está vendendo e através de quais canais da cadeia de suprimentos, mas também para produzir medidas aprimoradas de retorno sobre investimentos em marketing, valor da vida útil do cliente, impulsionadores da fidelidade do cliente, e assim por diante. Claramente, sem essas ferramentas de análise emergentes, as interrupções na cadeia de suprimentos associadas à pandemia teriam sido muito piores para consumidores e empresas do que as experimentadas.

2.1.3 | *Marketing Transformador*

À medida que a quantidade, o tipo (estruturado e não estruturado) e a complexidade dos dados mudam, também mudam os métodos de pesquisa necessários para apoiar esses desenvolvimentos. A variedade, velocidade e volume de dados surgem de ambientes de marketing em constante mudança, tecnologia emergente e o menor custo de coleta e armazenamento de dados (Erevelles, Fukawa & Swayne 2015; Kumar 2018; Varadarajan 2018; Davenport 2018). Exemplos de impulsionadores da criação de dados incluem a evolução do comportamento de compra do cliente, a internet das coisas, *wearables* digitais, o aumento da inteligência artificial, tecnologia aprimorada da cadeia de suprimentos e a capacidade de incorporar análises em vendas, distribuição e sistemas de produção, para citar alguns (Kumar 2018). Só o Facebook tem mais de 200 bilhões de fotos armazenadas na nuvem e mais de 350 milhões de novas fotos são carregadas todos os dias. Além disso, os usuários do Facebook carregam mais de 4 petabytes de dados por dia, o que representa mais de um milhão de gigabytes. Em resposta a esses fatores, as estratégias de marketing, as decisões organizacionais e o desempenho estão sendo revolucionados por ativos de informação, recursos de análise e maior conhecimento do cliente.

2.1.4 | *Mudando o papel da pesquisa de marketing*

Há valor em ambas as abordagens de pesquisa tradicionais que examinam, por exemplo, hipóteses derivadas da prática e da teoria, bem como análises preditivas que comumente enfatizam a resolução de problemas de negócios com dados (Babin, Griffin & Hair 2015; Delen & Zolbanin 2018). De fato, ferramentas e técnicas que informam diferentes níveis de inteligência sofisticada estão cada vez mais disponíveis (Davenport & Harris 2017). Essas ferramentas e técnicas variam de análises simples de planilhas usadas pelas empresas para entender quantos, com que frequência ou por que algo ocorreu, ferramentas preditivas que ajudam a explicar e prever o que provavelmente acontecerá, ferramentas automatizadas de aprendizado de máquina aplicadas para obter insights de dados e mais recentemente, um livro-texto universitário sobre *Essentials of Marketing Analytics*, McGraw-Hill Education (Hair, Harrison & Ajjan, 2021). Em suma, todos os níveis de sofisticação oferecem potencial para vantagens competitivas (Davenport & Harris, 2017), mas cada vez mais as empresas devem adotar e contar com as ferramentas mais sofisticadas disponíveis. O que muitas vezes é esquecido, no entanto, é que o *machine learning* cada vez mais aliado à inteligência artificial, quase sempre requer intervenção humana para executar e concluir as análises. Assim, o papel dos pesquisadores de marketing e a complexidade se expandirá e aumentará, exigindo treinamento muito mais avançado, criando assim oportunidades de trabalho adicionais para

graduados que estudam áreas relevantes de análise. Os parágrafos a seguir fornecem uma visão geral de como três abordagens de pesquisa de marketing - análise descritiva, preditiva e prescritiva - estão evoluindo e mudando.

2.1.5 | *Análise Descritiva Tradicional*

A pesquisa tradicional geralmente usa métodos descritivos e diagnósticos. Esse tipo de pesquisa é de natureza mais explicativa e confirmatória (Sivarajah, Kamal, Irani & Weerakkody 2017). Por exemplo, os principais indicadores de desempenho, *key performance indicators* (KPIs) fornecem visualizações de painel em tempo real para entender o quão eficaz e eficiente a empresa é em alcançar objetivos de negócios, como desempenho de vendas, gerenciamento de controle de estoque ou capacidade de resposta da cadeia de suprimentos. Embora valioso como métricas históricas, esse feedback confirmatório oferece apenas insights preditivos limitados sobre se os resultados futuros serão semelhantes. Mas continuará sendo muito útil para os profissionais de marketing como meio de monitorar as operações do dia a dia, e os pesquisadores de marketing precisarão fornecer suporte para melhor compreensão e explicação da evolução do mercado, bem como sugerir formas inovadoras de aplicar essas informações.

Os dados usados na análise descritiva podem ser encontrados dentro e fora da organização. As empresas geralmente coletam dados regularmente sobre clientes e concorrentes, mas também usam dados de fontes externas, como data.com e salesforce.com. Embora a quantidade de dados internos seja muito mais fácil de coletar neste momento, é provável que o futuro traga recursos de tecnologia avançada, permitindo que as organizações coletem dados de fontes externas com mais eficiência.

2.1.6 | *Análise preditiva*

A tecnologia avançada combinada com algoritmos complexos facilita as abordagens de análise preditiva. Essas técnicas emergentes examinam montanhas de dados para promover narrativas para avaliar oportunidades de forma mais eficaz, bem como prever desenvolvimentos futuros. Em suma, a análise preditiva aplica ferramentas de modelagem aos dados para prever tendências e desenvolvimentos futuros do mercado (Davenport & Harris 2017). Os pesquisadores categorizaram a análise preditiva em duas categorias: técnicas de regressão e técnicas de *machine learning* (Gandomi & Haider 2015), ambas cada vez mais elementos centrais dos programas de pesquisa de marketing.

Ao passo que a análise preditiva normalmente exija algum nível de intervenção humana, geralmente há flexibilidade no nível de envolvimento (por exemplo, selecionar variáveis-alvo, desenvolver modelos teoricamente derivados, permitir que os dados relatem uma história). *Big data* obtido de muitas fontes, potencialmente contendo milhares e até milhões de variáveis de *driver*, é analisado usando centenas de modelos analíticos. Mas esse tipo de modelagem apresenta desafios substanciais em comparação com os métodos tradicionais (Davenport & Harris 2017) e requer analistas de dados mais experientes do que no passado. Por exemplo, com muitas fontes e tipos de dados (estruturados e não estruturados) integrados em um único banco de dados, é praticamente impossível formular hipóteses de todos os relacionamentos possíveis. De acordo com Peter Norvig (2009), diretor de pesquisa do Google, a grande quantidade de dados e a complexidade resultante dos relacionamentos geralmente tornam vantajoso

para os computadores derivar modelos dos dados, em vez de os humanos inicialmente gastarem tempo usando abordagens científicas para identificar possíveis modelos. Os computadores, portanto, exploram os dados para localizar relacionamentos ocultos que, de outra forma, muitas vezes não seriam identificados e, em seguida, desenvolvem modelos preditivos que suportam conclusões provisórias e facilitam a tomada de decisões estratégicas.

As aplicações de *machine learning* (ML), *natural language processing* (NLP) e rede neural (RN) aumentaram substancialmente tanto na pesquisa quanto na prática acadêmica de marketing. Os algoritmos de *machine learning* são projetados para coletar conhecimento de bancos de dados disponíveis e combiná-lo com processos automatizados de comportamento de aprendizado, com o objetivo de melhorar nossa compreensão de relacionamentos complexos e melhorar a previsão dos resultados desejados. A maioria desses métodos era inerente à análise de dados multivariada tradicional e agora está sendo estendida e adaptada para resolver problemas adicionais com maior precisão. Exemplos incluem combinar esses métodos com inteligência artificial (IA) e as fontes de dados digitais cada vez mais difundidas (Davenport et al. 2020) para desenvolver algoritmos e software de computador que desenvolvam soluções inteligentes (Shankar 2018). As soluções melhoram a capacidade dos pesquisadores de mercado de interpretar dados internos e externos de forma mais eficaz, desenvolver soluções usando os dados e facilitar adaptações mais flexíveis (Kaplan & Haenlein, 2019). Exemplos de métodos relacionados e extensões analíticas incluem metaheurísticas como algoritmos genéticos e métodos probabilísticos como filtros de Kalman.

A análise preditiva também está sendo cada vez mais aplicada no setor corporativo. Não apenas pelas maiores empresas, mas agora algumas de médio porte estão usando a análise preditiva. Muitos algoritmos de *machine learning* estão disponíveis publicamente por meio de programas de código aberto, como R e Python. Por exemplo, R e Python oferecem a capacidade de realizar uma análise de cesta de mercado aplicando regras de associação simples usando o algoritmo FP-Growth. Para dados não estruturados, algoritmos de processamento de linguagem natural podem ser empregados para determinar a classificação, o segmento e a contagem do número de vezes que as palavras aparecem em sites de feedback corporativo, postagens de mídia social e blogs relacionados ao consumidor. Quando combinadas com dados de transações de clientes, técnicas mais probabilísticas, como análise discriminante linear ou baías ingênuas, podem ser úteis para prever qual Google AdWords fornecerá o melhor retorno sobre o investimento em marketing. Uma desvantagem de usar R e/ou Python é que eles não são fáceis de usar. Isso levou ao surgimento de empresas de nicho, como Creatio e SugarCRM, que atendem a empresas de médio porte e promovem plataformas de CRM de baixo código/sem código acessíveis, impulsionadas por IA. Para análise preditiva baseada em regressão, o software SmartPLS é uma interface gráfica relativamente barata que é fácil de usar e funciona bem mesmo com dados limitados.

Muitas empresas maiores e pesquisadores avançados empregam aplicações automáticas de *machine learning*. Os algoritmos de aprendizado de *machine learning* (Auto ML) obtêm informações de dados anteriores e usam entradas de dados subsequentes para reduzir erros e melhorar a capacidade preditiva (Hair, Harrison e Ajjan, 2021). Softwares de dados avançados, como Databot, Microsoft Azure ML e Rapidminer, oferecem aos usuários a

capacidade de empregar (interfaces de programação de aplicativos) *Application Programming Interface* (API) para importar dados automaticamente para repositórios virtuais dentro do software ou por meio de serviços de hospedagem de repositório, como o Github, que podem ser imediatamente acessado pelo software. Isso permite atualizações virtuais de dados de clientes e negócios a partir de dados transacionais gerados por pontos de contato do cliente, cookies de navegador da web, dados de terceiros, *web-scraping*, dados econométricos e informações fornecidas por meio de aplicativos móveis que podem fornecer informações pessoais, como atividade de mídia social, dados geoespaciais e até dados biométricos (Hair & Sarstedt 2021). Esses aplicativos Auto ML, juntamente com a atualização contínua de dados, permitem atualizações em tempo real para modelos e resultados preditivos, o que pode melhorar os relacionamentos comerciais por meio de melhores previsões e relacionamentos com o consumidor por meio da hiperpersonalização (Ma & Sun 2020). Uma aplicação simples de hiperpersonalização em tempo real seria o *geofencing*, no qual uma empresa pode enviar anúncios personalizados para clientes identificados como andando ou dirigindo dentro de um raio definido de seu local de varejo. Um exemplo muito mais complexo pode ser visto na plataforma PROS que é empregada por muitas companhias aéreas e agências de viagens on-line para ajustar automaticamente as tarifas aéreas usando uma série de entradas, como histórico de pesquisa do consumidor, dados de cookies e comportamento de compras anteriores, dados de mercado externo, sazonalidade e disponibilidade para prever o preço que melhor combina a capacidade com a demanda.

O avanço da pesquisa de marketing, tanto na academia quanto na prática, exige que os pesquisadores estendam o objetivo do *machine learning* das saídas preditivas para a inteligência artificial explicável, que ocorre quando as saídas preditivas podem ser compreendidas e explicadas usando a teoria, a razão e a lógica existentes (Cabelo e Sarstedt 2021). Um aumento na IA explicável não apenas melhoraria a adoção do *machine learning* na prática, mas também poderia ser utilizado para estender a teoria existente e criar hipóteses testáveis que poderiam ser validadas por meio de métodos de pesquisa tradicionais (Ma & Sun 2020). Para fazer isso, no entanto, ainda envolve considerável intervenção humana de analistas de marketing e dados. Indivíduos competentes devem ser capazes de selecionar o método analítico apropriado e, quando as soluções são identificadas, outros analistas qualificados devem interpretá-las em termos de lógica, relevância para problemas de negócios e, em última análise, valor estratégico. Em última análise, os pesquisadores humanos continuarão sendo necessários para facilitar o “porquê” em resposta às descobertas “o quê” e “quando” geradas pelos resultados do aprendizado de máquina. O problema é que não há cientistas de dados qualificados suficientes para tomar essas decisões. A McKinsey estima que apenas os EUA enfrentam uma escassez de 140.000 a 190.000 pessoas com experiência analítica e 1,5 milhão de gerentes e analistas com habilidades para entender e tomar decisões com base na análise de *big data*. Em resposta a essa necessidade, as universidades dos EUA adicionaram mais de 100 programas de bacharelado em análise de dados nos últimos cinco anos, com mais sendo adicionados todos os anos.

2.1.7 | Análise Prescritiva

A análise prescritiva é um nível avançado de análise que examina o que deve ser feito ou o que pode ser feito para que algo aconteça. A análise prescritiva permite que os pesquisadores de marketing determinem comportamentos ideais e avaliem o eventual impacto nos negócios (Davenport & Harris 2017; Sivarajah et al. 2017). Para conseguir isso, a análise prescritiva experimenta e otimiza vários cenários em um ritmo acelerado e corre menos risco de interpretação humana subjetiva (Davenport & Harris 2017). Isso pode ser bom se a análise funcionar bem, mas pode não ser tão bom de outra forma. Assim, a análise prescritiva, embora não completamente autônoma neste momento, pode produzir implicações e gerar recomendações. Como exemplo, a FleetPride, fornecedora de peças para caminhões e reboques para clientes em vários setores, obteve sucesso na implementação de

análises prescritivas. Usando as soluções oferecidas pela IBM, os resultados forneceram aos gerentes da cadeia de suprimentos *insights* para operar sem problemas; permitindo que a empresa otimize as redes de distribuição, aumente a velocidade do fluxo de estoque e reduza os erros de embalagem do armazém. Embora o número de organizações que usam a análise prescritiva esteja crescendo, é difícil saber o nível de adoção. Infelizmente, não é incomum onde a adoção ocorreu em silos departamentais ou de unidades, que outros em toda a organização muitas vezes não saibam que a análise prescritiva foi implementada (Rossi 2015). É claro que a chave para a análise prescritiva bem-sucedida depende, assim como da análise preditiva, das habilidades dos cientistas de dados que a executam, e há uma escassez global de indivíduos treinados nessa área.

2.1.7.1 | Para onde vai a Pesquisa de Marketing a partir daqui?

Ferramentas e técnicas sofisticadas estão sendo prontamente adotadas na prática de pesquisa de mercado, com orçamentos cada vez mais dedicados à compra de soluções de análise avançada (Mela & Moorman 2018). Com as entradas de dados em constante evolução (por exemplo, tecnologias), tipos de dados (por exemplo, dados não estruturados) e mudanças no comportamento do cliente, é importante que a pesquisa de marketing permaneça relevante e oportuna para a gestão. Várias questões relevantes são: até que ponto os pesquisadores de marketing e os cientistas de dados devem permitir que essas novas ferramentas analíticas desenvolvam modelos e resultados sem primeiro desenvolver hipóteses derivadas teoricamente? Qual é o lugar na pesquisa de marketing para a integração de ferramentas preditivas e prescritivas? Como esse método poderia desenvolver novas teorias e agregar valor ao campo? Como a pesquisa de marketing pode aplicar softwares como SAS/STAT para modelagem preditiva, LIWK para análise de texto, ferramentas de análise de redes de mídia social como NodeXL ou ferramentas de *machine learning* como DataRobot? Como as soluções de inteligência artificial (IA) são aplicadas, como o Salesforce Einstein, que não apenas oferece previsões, mas também faz recomendações sobre como responder aos clientes, impactando os recursos da força de vendas? A pesquisa de marketing deve permanecer ciente da evolução em *big data* e *analytics*, entender como e quando usar esses métodos na pesquisa e acompanhar as prioridades gerenciais na prática.

3 | Mudanças na Metodologia e Desenho de Pesquisa

Paralelamente à evolução das técnicas e ferramentas avançadas, também estamos testemunhando mudanças nos projetos e metodologias de pesquisa.

3.1 | O uso de dados de pesquisa

Ainda na década de 1990, mais de 70% dos trabalhos acadêmicos publicados eram baseados em pesquisas ou entrevistas. Mas em 2013, o número de artigos publicados baseados em pesquisas nos três principais periódicos de marketing [Journal of the Academy of Marketing Science (JAMS), Journal of Marketing (JM) e Journal of Marketing Research (JMR)] era apenas cerca de 35. % (Hulland, Baumgartner & Smith 2018). Para coletar dados, os pesquisadores confiam cada vez mais em métodos de pesquisa on-line, como Qualtrics, Google Survey, Mechanical Turk e Survey Monkey (Hulland & Miller 2018). Empresas e pesquisadores estão capitalizando a transformação digital para alcançar os

participantes da pesquisa por meio de plataformas online. As plataformas online forneceram acesso conveniente a amostras de negócios e consumidores que eram difíceis, caras e muitas vezes impossíveis de alcançar. Infelizmente, as populações estão sendo mais pesquisadas, resultando em recusas em responder, e a proporção de artigos baseados em pesquisas começou a diminuir (Hulland, Baumgartner & Smith 2018). Outra questão observada por várias disciplinas de ciências sociais, e particularmente por grupos de gestão e cadeia de suprimentos, são os designs de pesquisa de respondente único. Nesse tipo de desenho, o mesmo entrevistado responde a perguntas relacionadas tanto às variáveis independentes quanto às dependentes, produzindo o que se chama de viés de métodos comuns (Krause, Luzzini & Lawson 2018; Kull, Kotlar & Spring 2018; Flynn, Pagell & Fugate 2018; Roh, Whippe & Boyer 2013). Outras disciplinas, incluindo marketing, levantaram preocupações semelhantes e estão sugerindo projetos que combinem dados secundários com dados primários como solução. Infelizmente, até o momento, nenhum método para medir com precisão a extensão do viés de métodos comuns em projetos de pesquisa foi desenvolvido.

3.2 | Mudança para dados objetivos

Com o surgimento de dados digitais, e muitos deles, os pesquisadores passaram a coletar dados objetivos e secundários. Informações de arquivo ou proprietárias, como dados históricos, estão prontamente disponíveis para pesquisadores (Verma, Agarwal, Kachroo & Krishen 2017). Usando dados históricos, estudos de eventos (Sorescu, Warren & Ertekin 2017) podem examinar o impacto de longo prazo de uma decisão específica na empresa. Empresas e pesquisadores estão focados em estratégias preditivas e prescritivas, aprimoradas por inteligência artificial (Huang & Rust 2018) e *machine learning* (Antons & Breidbach 2018). Até certo ponto, essas técnicas são projetadas para serem capazes de produzir análises autocorretivas, enquanto outras, como a otimização matemática, “determinam a melhor solução para problemas definidos matematicamente” (Snyman 2005). Nos mercados de consumo, por exemplo, a digitalização facilitou a disponibilidade de dados fisiológicos. De fato, em geral, os métodos de pesquisa mais técnicos incluem dados biométricos e neurológicos (Chan, Boksem & Smidts 2018; Harris, Ciorciara & Gountas 2017). Essas técnicas, inicialmente emergindo em economia, finanças, ciência da computação e pesquisa médica, agora estão se tornando mais prevalentes e infundidas na pesquisa de marketing e nas ciências sociais em geral. À medida que olhamos para o futuro, esses métodos de pesquisa em evolução se tornarão um recurso crítico nos mercados dinâmicos e em constante mudança entre empresas e consumidores.

3.3 | Software de visualização para comunicar resultados

É fundamental relatar claramente os resultados da pesquisa de uma maneira que comunique rápida e facilmente resultados significativos. A visualização de informações tem sido cada vez mais adotada como a principal abordagem para facilitar a comunicação do resultado de análises estatísticas complexas com foco na exploração de dados, descoberta e exibição de resultados de pesquisas de marketing. Ao usar métodos de visualização, os pesquisadores podem criar relatórios não técnicos para os tomadores de decisão entenderem, independentemente de sua formação. Existem várias ferramentas populares de visualização de dados, como Tableau e Microsoft BI, que incluem gráficos avançados de computador, gráficos e mapeamento. Essas ferramentas permitem que os pesquisadores identifiquem rapidamente padrões visuais, apresentem informações em configurações em tempo real, criem painéis interativos para revelar informações

filtradas, melhorando a tomada de decisões de marketing.

3.4 | A ascensão dos dados não estruturados

O marketing digital por meio de áreas como mídias sociais produziu conteúdo substancial não estruturado (Balducci & Marinova 2018) que pode ser submetido à análise de texto (Humphreys & Wang 2017), e cada vez mais codificado para modelagem quantitativa. Na verdade, a conversão de dados não estruturados para análise quantitativa está se expandindo rapidamente, já que 90% de todos os dados não são estruturados. Dados não estruturados incluem tweets, fotos, registros de chamadas de serviço, blogs, reclamações de clientes e dados de comentários, comportamento de compra do cliente e assim por diante, todos os quais devem ser analisados e codificados antes de serem enviados ao software de análise de dados quantitativos. Programas como NodeXL e LIWC são usados para aprimorar a modelagem quantitativa (Harrison, Ajjan & Coughlan 2018). Por meio da coleta e análise de dados não estruturados, como texto digital ou comentários, empresas e pesquisadores ampliam seus conhecimentos em segmentação, nível de envolvimento e experiência dos clientes, redes sociais e análise de sentimentos, o que também cria oportunidades para pesquisadores de mercado.

3.5 | Maior foco na iniciativa, rigor e qualidade

À medida que os dados e os métodos analíticos se tornam mais complexos, a pesquisa de marketing também está passando por uma mudança em direção a técnicas mais rigorosas. Os pesquisadores devem usar todas as abordagens razoáveis para refutar suas hipóteses em vez de sentir que é necessário apoiar estatisticamente as hipóteses - para que eles contribuam (Babin, Griffin & Hair 2015). Por exemplo, endogeneidade, poder estatístico (tamanho do efeito), validade externa e relevância preditiva desempenham um papel importante na avaliação precisa das relações entre variáveis exógenas (independentes) e endógenas (dependentes).

Por muitos anos, pesquisadores de marketing examinaram a heterogeneidade observada. A heterogeneidade observada baseia-se no conhecimento de que subgrupos de populações provavelmente apresentam diferenças, e essas diferenças são conhecidas quando a pesquisa é projetada e conduzida. Por exemplo, diferentes faixas etárias de consumidores geralmente exibem diferentes padrões de pesquisa e compras e, em última análise, comportamento de compra. Para os consumidores, a heterogeneidade observada normalmente se baseia em dados demográficos. Para as organizações, a heterogeneidade observada foi baseada no tamanho da empresa, ofertas de produtos ou serviços, conveniência e assim por diante. Mais recentemente, a tecnologia permitiu que os analistas de dados examinassem a possibilidade de heterogeneidade não observada, que é a noção de que as populações consistem em subgrupos da população que não são facilmente observáveis. O conhecimento sobre a presença de heterogeneidade não observada é fundamental nos estudos empíricos de hoje e, se não for explorado, potencialmente ameaça a validade dos modelos de medição e estruturais examinados ao aplicar análises descritivas e preditivas (Becker, Rai, Ringle & Volckner 2013; Sarstedt & Ringle 2010; Hair, Matthews, Matthews & Sarstedt 2017; Hair et al., 2022a). Tanto os pesquisadores da indústria quanto os acadêmicos estão avançando rapidamente para desenvolver análises estatísticas avançadas que abordam efetivamente possíveis

problemas de confiabilidade e validade.

Embora essas questões tenham sido uma preocupação na pesquisa de marketing, métodos estatísticos aprimorados permitem que os pesquisadores as abordem com mais rigor (Hult, Hair et al. 2018; Worm, Bharadwaj, Ulaga & Reinartz 2017). Análises baseadas em regressão funcionam na suposição de multicolinearidade mínima. Ao mesmo tempo, porém, a moderação ocorre quando dois ou mais fatores covariam o suficiente para alterar o impacto no resultado. Se a moderação for conhecida ou suspeita, ela pode ser testada com variáveis contínuas ou categóricas. Isso é particularmente útil para identificar variáveis de higiene ou condições de contorno.

Acadêmicos e profissionais estão usando cada vez mais softwares sofisticados para executar análises preditivas. Um exemplo, o software SmartPLS (www.SmartPLS.de) é de ponta em termos de análise e uma opção preferida para muitos estudiosos e profissionais porque é muito fácil de usar. O software executa a análise de vários grupos no SmartPLS e pode ser usado para examinar a heterogeneidade observada, na qual grupos conhecidos são propostos para se comportarem de forma diferente (por exemplo, comparações femininas/masculinas; grandes/pequenas empresas), bem como heterogeneidade não observada (Sarstedt et al., 2017; Hair et al., 2016; Matthews et al. 2016, Hair et al., 2022a). As análises de vários grupos (*Multi-group analyses* -MGA) ajudam os pesquisadores a descobrir se existem diferenças nas estimativas de parâmetros (por exemplo, pesos externos, cargas externas e estimativas de parâmetros) de grupos de dados predefinidos (Matthews, Hair & Matthews 2018). No PLS-SEM o MGA é benéfico ao examinar com eficiência a moderação em vários relacionamentos (Hair, Sarstedt, Ringle & Mena 2012). Outro software cada vez mais popular é o R, amplamente utilizado entre estatísticos e cientistas de dados para executar análises de dados. O software é de código aberto e gratuito, e cada vez mais aplicado (Hair et al., 2022b).

Se os moderadores são desconhecidos antecipadamente, esse tipo de análise se torna mais difícil. Frequentemente, aparecem clusters com valores semelhantes para determinadas variáveis. Esses tipos de grupos não podem ser definidos com precisão por nenhuma variável de dados demográfica ou descritiva específica. Esse tipo de moderação pode ser difícil de prever ou mesmo identificar. Nesses casos, técnicas baseadas em previsão devem ser usadas para identificar segmentos que podem estar criando uma heterogeneidade inexplicável. Além disso, vários pacotes de software estatísticos oferecem aos pesquisadores a oportunidade de verificar ainda mais a validade externa e a relevância preditiva (www.smartpls.de, www.r-project.org; Schmueli, Ray et al. 2016).

Com foco no PLS-SEM os pesquisadores propuseram a aplicação metodológica da segmentação orientada a predição (*Partial Least Squares - prediction-oriented segmentation* PLS-POS) e segmentação por mínimos quadrados parciais de mistura finita (*Finite mixture partial least squares* - FIMIX-PLS) como métodos para superar limitações anteriores de identificar heterogeneidade não observada. Outras técnicas heurísticas devem ser utilizadas quando a heterogeneidade encontrada nesses efeitos for sequencial. Isso vai muito além da análise longitudinal, e muitas vezes esse tipo de variância resulta de efeitos aninhados, devido à ordem dos efeitos, sazonalidade, agrupamento etc. Tradicionalmente, o HLM tem sido usado para identificá-los na econometria. No *machine learning*, o teorema de Bayes é usado para superar a natureza transversal dessas situações. O teorema de Bayes assume que os termos de erro variam dependendo da situação ou ordem dos efeitos. A modelagem bayesiana usa o que é conhecido para prever

o desconhecido. Além disso, a modelagem de Markov Chain Monte Carlo (MCMC), como as técnicas Bayesianas, consiste em algoritmos focados em distribuições de probabilidade para grandes modelos multidimensionais, hierárquicos e parâmetros desconhecidos. O método é vantajoso quando existe um alto nível de incerteza e a informação histórica é escassa. Mas também pode ser usado para investigar informações em várias fontes de dados (Rossi & Allenby 2003). As aplicações dos métodos bayesianos aumentaram em paralelo com a transformação digital impulsionada pelo uso da internet.

A pesquisa de marketing, particularmente acadêmica, acolheu a adoção de técnicas Bayesianas e aplicou o método para abordar muitos aspectos do marketing (Rossi & Allenby 2003), como publicidade (Agarwal, Hosanager & Smith 2011), escolha do cliente (Chung, Rust & Wedel 2009) e relacionamento com clientes (Netzer, Lattin & Srinivasan 2008). Embora os dados sejam abundantes, a digitalização ampliou o desconhecido em ambientes dinâmicos e em constante evolução. A modelagem de equações estruturais de mínimos quadrados parciais (PLS-SEM) é um elemento essencial da caixa de ferramentas metodológica dos pesquisadores de marketing. O campo PLS-SEM passou por amplo desenvolvimento nos últimos anos, mas os pesquisadores de marketing precisam continuar adotando as extensões mais recentes (Sarstedt, Hair, Pick, Liengaard, Radomir & Ringle, 2022). Os desenvolvimentos incluem PLSpredict (Shmueli, Sarstedt, Hair, Cheah, Ting, Vaithilingam e Ringle, 2019; Manley, Hair, Williams e McDowell, 2021; Hair & Sarstedt, 2021b); análise composta confirmatória (CCA) (Hair, Howard & Nitzl, 2020); mediação (Sarstedt, Hair, Nitzl, Ringle & Howard, 2020); e métricas de comparação de modelos (Hair 2021).

Uma tendência paralela aos inúmeros avanços nas técnicas analíticas é a ênfase na qualidade (Hair, Moore & Harrison, 2022). Os métodos analíticos emergentes que estão sendo aplicados na pesquisa de marketing são úteis para a tomada de decisões apenas quando os dados usados com as técnicas são válidos. Assim, a qualidade dos dados é uma preocupação fundamental na sua aplicação. A ausência de qualidade de dados se manifesta de várias maneiras, incluindo dados ausentes, incompletos, inconsistentes, imprecisos, desatualizados e duplicados (Gudivada et al. 2017). Por exemplo, valores discrepantes em conjuntos de dados podem resultar em estimativas de parâmetros tendenciosos, dados ausentes podem facilmente levar a erros de previsão substanciais e conjuntos de dados pequenos normalmente reduzem o poder de um modelo (Bosu & McDonell 2019; Gudivada et al. 2017).

Outro critério essencial para garantir a qualidade dos resultados das pesquisas é a capacitação dos analistas de pesquisa de mercado. Considere, por exemplo, um sistema de recomendação de um varejista on-line em que um cliente normalmente compra para si mesmo, mas ocasionalmente também para terceiros. Os analistas de dados devem ter conhecimento e experiência suficientes para determinar se os algoritmos estatísticos automatizados que determinam o que é normal e o que é uma anomalia são projetados para identificar padrões válidos. Se os dados analisados não tiverem um número suficiente de respostas para representar cada tipo de situação de compra, as medidas das frequências relativas dos diferentes tipos de compras não serão válidas, e os resultados não fornecerão uma compreensão precisa do que é “normal.” e o que não é. Além disso, os conjuntos de dados geralmente incluem zeros em excesso como parte do processo de coleta de dados, chamados de modelos de dados de contagem inflados de zero (*zero-inflated*), e isso pode produzir resultados tendenciosos quando não analisados corretamente (Spriensma, Hajos, de Boer, Heymans & Twisk, 2013 ; Favero, Cabelo,

Souza, Albergaria, & Brugni, 2021).

A necessidade de monitorar com precisão a qualidade dos dados se torna mais desafiadora ao combinar várias fontes de dados díspares, geralmente uma característica de dados secundários, que precisam ser harmonizados para uma única análise. Diferenças no tratamento de dados, por exemplo, em termos de identificação e tratamento de valores discrepantes, queixa de valores ausentes ou conversão de dados não estruturados em estruturados, podem alterar os resultados analíticos ao prever variáveis de resultado. Identificar tais inconsistências é muito desafiador quando os dados são obtidos de várias fontes que utilizam procedimentos diferentes para garantir a qualidade dos dados. Além disso, o método de amostragem é de igual importância na era da *big data* porque as abordagens de amostragem com *big data* raramente são baseadas em critérios científicos. De fato, não é incomum que analistas menos experientes nem mesmo apliquem amostragem aos seus dados. Em vez disso, a amostragem é frequentemente baseada em fatores sociais, políticos, econômicos e técnicos que determinam quais dados acabam no conjunto de dados final (Leonelli, 2014).

Embora o princípio “*garbage in - garbage out*” se aplique a todas as análises de dados, as preocupações com a qualidade dos dados parecem ser muito menos debatidas entre os pesquisadores de marketing que usam *machine learning* e inteligência artificial. Em contraste, a pesquisa sobre qualidade de dados tem uma longa história em ciência da computação e sistemas de informação gerencial. Pesquisadores de qualidade de dados nesses campos formaram sua própria comunidade (Madnick et al. 2009) com inúmeras contribuições intelectuais na descoberta de padrões de viés, tratamento de valores ausentes e desenvolvimento de métricas de precisão, muitas das quais estão documentadas no *ACM Journal of Data and Information Quality*. Os insights dessa comunidade desempenharão um papel cada vez mais importante na pesquisa de marketing para garantir que as recomendações dos profissionais de marketing não sejam baseadas em dados e análises caracterizadas por problemas de qualidade (Hair & Sarstedt, 2021).

Moore, Harrison e Hair (2022) propuseram um processo de seis etapas para garantir a qualidade dos dados na pesquisa de marketing. As etapas incluem: (1) começar com pesquisa qualitativa para garantir um projeto de pesquisa rigoroso, (2) projetar questionário e coletar dados, (3) filtrar e limpar dados, (4) aplicar o método analítico correto, (5) interpretar os resultados para representar todas as perspectivas e (6) comunicar o processo realizado para garantir a qualidade dos dados em todas as etapas do processo de pesquisa. Seguir um processo abrangente de garantia de qualidade de dados é particularmente importante, pois em muitas organizações a qualidade dos dados é muitas vezes questionável, a presença de problemas de qualidade de dados raramente é conhecida e a qualidade dos dados frequentemente não pode ser resolvida após a coleta dos dados. Como a baixa qualidade dos dados leva a resultados analíticos incorretos, interpretações distorcidas ou deturpadas das descobertas e previsões imprecisas do provável sucesso ou fracasso das estratégias e táticas de marketing, os pesquisadores de marketing devem se concentrar em minimizar sua presença em seus dados.

4 | Observações Finais

Como deve ser evidente, muita coisa mudou em marketing e sua pesquisa nos últimos anos. Muitas outras mudanças estão previstas para os próximos anos. O que se segue resume os nossos pensamentos sobre os desenvolvimentos futuros. As previsões do futuro são

claramente arriscadas. Muitos indivíduos inteligentes tentaram prever o futuro - e muitas vezes se mostraram errados. Nossos pensamentos são baseados no que acreditamos ser um raciocínio sólido, mas, como outros no passado, é improvável que nosso pensamento seja inteiramente como nos parece hoje. Acreditamos que estamos corretos em sua maioria e pedimos que você veja nossos pensamentos como pelo menos direcionais - pois estamos confiantes sobre a direção, mas não necessariamente os detalhes.

4.1 | Os dados continuarão aumentando exponencialmente

As primeiras discussões sobre a rápida expansão dos dados começaram há pelo menos 70 anos, bem antes do atual interesse generalizado em big data e seu impacto. Foi então que começaram as primeiras tentativas de quantificar a taxa de crescimento dos dados e, pelos próximos 50 anos, essa tendência foi chamada de “explosão da informação”. Mais ou menos na mesma época (1941), o *Oxford English Dictionary* pela primeira vez incluiu o termo explosão de informações em sua lista, e muitos outros termos de dados foram adicionados desde então.

Entre os primeiros autores de negócios a escrever sobre esse fenômeno foi Tom Peters em seu livro de 1991 intitulado *Thriving on Chaos*. Embora seu livro previsse muitas mudanças econômicas amplas ocorrendo naquela época, a frase mais relevante para a *big data* era: “Estamos nos afogando em informações e famintos por conhecimento.” Ele continuou observando que a diferença entre informação e conhecimento é que a informação consiste apenas em palavras e números que podem ser interessantes, mas não são úteis para a tomada de decisões. Além disso, o pensamento naquela época era que talvez apenas cinco por cento de todas as informações fossem conhecimento, e o desafio era converter mais informações disponíveis em conhecimento. Peters e outros acreditavam que isso seria possível se conscientizassem sobre esse problema, mas não previam o impacto da internet e, principalmente, o surgimento de dados digitais.

No final da década de 1990, o livro intitulado *Quanta informação existe no mundo?* por Michael Lesk (1997), concluiu “Pode haver alguns milhares de petabytes de informação ao todo; e . . . em apenas alguns anos. . . seremos capazes de salvar tudo - nenhuma informação terá que ser descartada, e . . . a informação típica nunca será vista por um ser humano.” A essa altura, estudiosos e profissionais estavam claramente cientes não apenas da enorme quantidade de dados produzidos, mas também da importância da análise de dados e, particularmente, do papel do aprendizado de máquina na análise dos dados disponíveis. É claro que os cientistas de dados também estavam desenvolvendo novos termos em um esforço para esclarecer como alguém poderia medir as grandes quantidades de dados e começaram a usar palavras como terabyte, petabyte, exabyte, zettabyte e, finalmente, yottabyte - com cada termo sucessivo referindo-se a uma quantidade 1.000 vezes maior. De acordo com um relatório da Forbes, a quantidade total de informações em 2013 foi de 4,4 zettabytes, e em 2020 a quantidade de dados chegará a 44 zettabytes (Kanellos, 2016). Para o leigo, 44 zettabytes tem pouco significado, mas para o cientista de dados indica claramente uma enorme quantidade de dados. É claro que esta e outras estimativas são apenas previsões informadas, mas a direção e a quantidade de aumento no acúmulo e armazenamento de dados são inquestionavelmente corretas.

Conforme observado, acreditamos que os dados continuarão a aumentar exponencialmente, e isso afetará tanto os profissionais de pesquisa de mercado quanto os acadêmicos. Os profissionais devem ajudar as organizações a entender e responder efetivamente a esse excesso de dados. Eles devem aconselhar os profissionais de marketing sobre quais dados são relevantes e quais podem ser descartados. Além disso, os pesquisadores de marketing precisarão ser mais eficazes em informar os gerentes de marketing sobre como os dados podem ser usados e quais dados serão mais úteis. A maior parte dos dados disponíveis agora e no futuro será secundária e digital. Isso contrasta fortemente com o passado, no qual os pesquisadores de marketing coletavam dados primários e controlavam que tipo de dados seriam coletados na esperança de que fossem direta e imediatamente aplicáveis aos problemas enfrentados por seus clientes. Com o aumento dos dados secundários, os pesquisadores de marketing devem ajudar as organizações a vasculhar as montanhas de dados e determinar o que é relevante e o que tem pouco ou nenhum valor. Além disso, futuros dados longitudinais estarão disponíveis para realizar análises retrospectivas, bem como preparar previsões mais eficazes.

Os acadêmicos de marketing devem seguir um caminho semelhante - determinar e obter acesso a dados secundários que sejam relevantes para suas pesquisas e permitirão que eles publiquem em periódicos de qualidade revisados por pares. Até agora, os acadêmicos acessaram e aplicaram dados secundários e digitais em apenas alguns casos. Mas há oportunidades cada vez mais substanciais para acessar esse tipo de dados, principalmente em marketing, e à medida que os dados aumentam exponencialmente, também aumentam as oportunidades para os acadêmicos buscarem essas oportunidades. Ao mesmo tempo, os textos de pesquisa de marketing também devem passar por grandes revisões, pois a maioria inclui apenas material limitado sobre análise de dados, ou mesmo sobre a explosão de dados secundários e digitais.

4.2 | A qualidade dos dados melhorará

Todos os aspectos da pesquisa acadêmica e profissional estão sendo impactados pela disponibilidade de dados e pelos avanços na análise de dados. Uma área afetada é o exame de dados e, à medida que o exame de dados é implementado nas organizações, ele melhorará a qualidade dos dados e, em última análise, o sucesso ou fracasso do esforço de pesquisa. Enquanto muitos pesquisadores continuam a operar com dados primários, tanto qualitativos quanto quantitativos, muitos outros estão lidando cada vez mais com fontes de dados amplamente díspares (por exemplo, dados em nível de cliente em várias divisões de empresas, mídia social e outros dados digitais, dados de localização etc.) que possuem diferentes estruturas e formatos de dados. O exame de dados expandido, geralmente chamado de limpeza de dados (*data cleaning*), é necessário e melhorará a qualidade dos dados e, portanto, os resultados.

Outra área em evolução, e provavelmente a mais desafiadora da era da big data, é o gerenciamento de dados - organizar, limpar e codificar os dados de várias fontes e em vários formatos para preparar a análise. Muitos pesquisadores pensam que a principal questão é qual técnica usar e não percebem que o tipo de dados geralmente dita a seleção da tecnologia e da técnica de análise. Na maioria dos casos, o que precede e dita a técnica de análise apropriada é o gerenciamento de dados de várias fontes, o que pode facilmente levar 50% ou mais do tempo total para concluir um projeto de análise de dados.

Combinar e gerenciar dados de várias fontes torna-se complexo muito rapidamente. O que pode parecer uma simples fusão de dados pode ser muito difícil quando o analista precisa extrair dados de vários bancos de dados, combinar formatos e prazos e muitos outros

ajustes necessários. E como 90% dos dados geralmente não são estruturados, essa tarefa é claramente assustadora. Portanto, pesquisadores acadêmicos e profissionais de todas as áreas terão que se tornar “gerentes de dados” além de analistas de dados.

4.3 | A análise de dados melhorará

A principal força por trás da ampla aplicação da análise de dados é a potencial melhoria na tomada de decisões. A análise de dados aprimorada beneficiará profissionais e acadêmicos, bem como governos e organizações com ou sem fins lucrativos. Mas, para obter esses benefícios, os analistas precisarão estar mais envolvidos no fornecimento de informações críticas sobre quais dados coletar, em que formato, como gerenciar e combinar os dados e, finalmente, selecionar a melhor técnica analítica.

Até recentemente, os acadêmicos demoravam a adotar e aplicar big data e análise de dados em seu ensino e pesquisa. Algumas disciplinas das ciências (biologia, biomedicina e neurociência) estão se movendo mais rapidamente na adoção da análise de dados, mas as ciências sociais estão começando a reconhecer o potencial de métodos analíticos mais complexos. Esses métodos de análise aprimorados possibilitam que tipos adicionais de questões de pesquisa sejam examinados e novos desafios sejam superados. À medida que acadêmicos e profissionais de pesquisa de marketing se tornam mais conscientes dessas oportunidades, tanto em dados quanto em técnicas, sua utilização aumentará, esperançosamente, mais cedo ou mais tarde.

4.4 | As decisões serão mais conscientes

Tanto a gestão quanto as decisões do cliente serão cada vez mais baseadas no conhecimento. A era da big data forneceu novas e variadas fontes de dados e colocou requisitos adicionais nas técnicas analíticas que devem lidar com essas fontes de dados. Existem, portanto, desafios únicos enfrentados pelo analista de hoje com os muitos problemas de big data e análise. Acreditamos que com um foco expandido em desafios de dados e análise de dados, os cientistas de dados serão capazes de criar conscientização e motivar a gestão a investir nesses desenvolvimentos promissores. Se isso acontecer, a pesquisa de marketing e as decisões de negócios em geral serão baseadas em melhores dados e técnicas analíticas mais eficazes e, portanto, mais baseadas em conhecimento. Claramente, as organizações e indivíduos que se movem mais rápido nessas áreas emergentes irão adquirir e estabelecer verdadeiras vantagens competitivas nos mercados de amanhã.

4.5 | Problemas e desafios de privacidade

Em seu livro *The Assault on Privacy*, Arthur Miller (1971) observou que “Muitos manipuladores de informações parecem medir [um indivíduo pelo] número de bits da capacidade de armazenamento que [seu] dossiê ocupará”. Hoje, nas economias avançadas, os dados existem sobre todos e cresceram exponencialmente devido aos registros digitais. Cada vez mais, no entanto, as mudanças nas políticas implementadas pelo governo estão afetando a maneira como coletamos, acessamos, armazenamos e usamos dados. Logo após os problemas de segurança da informação da Cambridge Analítica envolvendo o Facebook, duas leis de privacidade do consumidor foram ou estão em processo de adoção - o Regulamento Geral de Proteção de Dados (*General Data Protection Regulation-GDPR*) e a Lei de Privacidade do Consumidor da Califórnia, A.B. 375 foram desenvolvidos para proporcionar maior controle aos clientes dos dados que foram coletados sobre eles. Como resultado, as empresas devem agora aumentar os controles e a transparência de como os dados são coletados, quais dados são coletados, por que os dados estão sendo coletados, quem tem acesso aos dados e como as empresas estão usando os dados coletados. Os

pesquisadores de marketing também são obrigados a cumprir esses regulamentos, pois enfrentam as mesmas regras na coleta, proteção e distribuição de dados identificáveis.

4.6 | As capacidades de análise de dados fornecem vantagens competitivas substanciais

Os dados são tão bons quanto a inteligência que podemos obter deles, e isso envolve análises de dados eficazes e muito poder de computação para lidar com o aumento exponencial no volume e no tipo de dados. Quase todas as organizações (grandes, médias e até pequenas) podem aplicar a análise de dados para melhorar as atividades de fabricação, cadeia de suprimentos, gerenciamento e marketing e para serem mais eficientes e eficazes no alcance das metas organizacionais. Embora a aplicação da análise de dados esteja produzindo descobertas úteis nas áreas de agricultura, saúde, design urbano, redução do crime e energia, além de negócios em geral, ainda há muito a ser feito. O campo da análise está evoluindo rapidamente e mais oportunidades estão surgindo para aplicar as novas fontes de dados e os métodos de análise adicionais e mais sofisticados. A velocidade de adoção e aplicação à tomada de decisão influenciará claramente a aquisição de vantagens competitivas, mas os líderes nesse campo acumularão as vantagens, e suas vantagens podem se tornar intransponíveis para os retardatários.

4.7 | A demanda por cientistas de dados treinados excederá a oferta

Mais universidades precisam se envolver no treinamento de novos cientistas de dados. O treinamento e a conscientização sobre big data e análise de dados incentivarão os analistas a definir não apenas suas questões de pesquisa de forma mais ampla, mas também o escopo de sua responsabilidade em liderar o caminho para usar dados de forma mais eficaz para resolver problemas de negócios. A utilização bem-sucedida da análise de dados será influenciada substancialmente pelos analistas e pelas decisões que eles tomam, não apenas ao informar os gerentes seniores sobre o potencial da análise de dados, mas também na seleção e aplicação das técnicas a serem usadas no futuro.

De acordo com o site Study Portals, nos EUA existem pouco mais de 100 programas de bacharelado com concentração em análise de dados. Esse número precisa ser cinco e talvez dez vezes maior - e mudanças semelhantes devem ocorrer globalmente. Mas existe uma necessidade ainda maior de mestrado em análise de dados (apenas cerca de 40 são oferecidos nos EUA), já que ser um cientista de dados requer treinamento substancial. As universidades em todo o mundo precisam se mover rapidamente para treinar esses cientistas de dados para atender não apenas à demanda atual, mas para continuar produzindo o que será uma necessidade substancial e crescente de cientistas de dados no futuro, principalmente no campo da pesquisa de marketing. Instituições nos EUA que oferecem mestrado em análise de dados, <https://www.hotcoursesabroad.com/study/training-degrees/us-usa/masters/data-analysis-courses/loc/211/slevel/57-3-3/cgory/cb.451-4/sin/ct/programs.html>, acessado Agosto 2018.

5 | REFERÊNCIAS

- Abubakar, A.M., Namin, B.H., Harazneh, I., Arasli, H., & Tunç, T. (2017). Does gender moderate the relationship between favoritism/nepotism, supervisor incivility, cynicism, and workplace withdrawal: A neural network and SEM approach. *Tourism Management Perspectives*, 23, 129-139.
- Antons, D., & Breidbach, C. F. (2018). Big data, big insights? Advancing service innovation and design with machine learning. *Journal of Service Research*, 21(1), 17-39.
- Agarwal, A., Hosanagar, K., & Smith, M. D. (2011). Location, location, location: An analysis of profitability of position in online advertising markets. *Journal of marketing research*, 48(6), 1057-1073.

- Bachelor's Degrees in Data Science & Big Data in the United States, <https://www.bachelorsportal.com/study-options/269779226/data-science-big-data-united-states.html>, accessed August 2018.
- Balducci, B., & Marinova, D. (2018). Unstructured data in marketing. *Journal of the Academy of Marketing Science*, 46(4), 1-34.
- Becker, J. M., Rai, A., Ringle, C. M., & Völckner, F. (2013). Discovering unobserved heterogeneity in structural equation models to avert validity threats. *MIS Quarterly*, 37(3), 665-694.
- Bosu, M. F., & MacDonell, S. G. (2019). Experience: Quality Benchmarking of Datasets Used in Software Effort Estimation. *ACM Journal of Data and Information Quality*, 11(4), Article No. 19.
- Chan, H. Y., Boksem, M., & Smidts, A. (2018). Neural profiling of brands: Mapping brand image in consumers' brains with visual templates. *Journal of Marketing Research*. 55(4), 600-615.
- Chung, T. S., Rust, R. T., & Wedel, M. (2009). My mobile music: An adaptive personalization system for digital audio players. *Marketing Science*, 28(1), 52-68.
- Davenport (2018) Will data scientist continue to be the sexiest job? Downloaded June 2018 at <https://fianalytics.com/research/will-data-scientist-continue-to-be-the-sexiest-job>.
- Davenport, T., Guha, A., Grewal, D., & Bressgott, T. (2020). How artificial intelligence will change the future of marketing. *Journal of the Academy of Marketing Science*, 48, 24-42.
- Davenport, T., & Harris, J. (2017). *Competing on Analytics: Updated, with a New Introduction: The New Science of Winning*. Harvard Business Press.
- Delen, D., & Zolbanin, H. M. (2018). The analytics paradigm in business research. *Journal of Business Research*, 90, 186-195.
- Erevelles, S., Fukawa, N., & Swayne, L. (2016). Big Data consumer analytics and the transformation of marketing. *Journal of Business Research*, 69(2), 897-904.
- Favero, Luiz Hair, J. F., Jr., Souza, Rafael de Freitas, Albergaria, Matheus, and Brugni, T. V. (2021). Zero-Inflated Generalized Linear Mixed Models: A Better Way to Understand Data Relationships, *Mathematics*, 9, 1100, <https://doi.org/10.3390/math9101100>.
- Flynn, B., Pagell, M., & Fugate, B. (2018). Survey Research Design in Supply Chain Management: The Need for Evolution in Our Expectations. *Journal of Supply Chain Management*, 54(1), 1-15.
- Gandomi, A., & Haider, M. (2015). Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2), 137-144.
- Gudivada, V., Apon, A., & Jing, J. (2017). Data Quality Considerations for Big Data and Machine Learning: Going Beyond Data Cleaning and Transformations". In: *International Journal on Advances in Software* 10.1 (2017), pp. 1 - 20.
- Hair, J. F. (2021). Next Generation Prediction Metrics for Composite-based PLS-SEM. *Industrial Management & Data Systems*. 121(1), 5-11. <https://doi.org/10.1108/IMDS-08-2020-0505>.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate Data Analysis* (8th ed.). Cengage Learning, U.K.
- Hair, J. F., Harrison, D., & Ajjan, H. (2022). *Essentials of Marketing Analytics*, McGraw-Hill, N.Y., New York.
- Hair, J. F., Howard, M., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101-110. <https://doi.org/10.1016/j.jbusres.2019.11.069>.
- Hair, J.F., Hult, G.T.M., Ringle, C.M. and Sarstedt, M. (2022a), *A Primer on Partial Least Squares Structural Equation Modeling (PLS-SEM)*, 3rd ed., Sage, Thousand Oaks, CA.

- Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., and Ray, S. (2022b). *Partial Least Squares Structural Equation Modeling (PLS-SEM) Using R*. Springer, New York.
- Hair, J.H., Matthews, L., Matthews, R., & Sarstedt, M. (2017). PLS-SEM or CB-SEM: Updated Guidelines on Which Method To Use,” (2017), *International Journal of Multivariate Data Analysis*, Vol. 1 (2), pp 107-123.
- Hair, J.H., Matthews, L., Ringle, C., & Sarstedt, M. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part I – Method, *European Business Review*, 28(1), 63-76.
- Hair, J. F., Moore, Z., & Harrison, D. (2022). Data Quality Assurance Begins Before Data Collection and Never Ends: What Marketing Researchers Absolutely Need to Remember (2022). *International Journal of Market Research*, <https://doi.org/10.1177/14707853211052183>.
- Hair, J., Risher, J., Sarstedt, M., & Ringle, C. (2019). When to use and how to report the results of PLS-SEM. *European Business Review*, 31(1), 2–24.
- Hair, J. F., & Sarstedt, M. (2021a). Data, Measurement, and Causal Inferences in Machine Learning: Opportunities and Challenges for Marketing, *Journal of Marketing Theory and Practice*, 29(1), pp 65-77, doi.org/10.1080/10696679.2020.1860683.
- Hair, J. F., & Sarstedt, M. (2021b). Explanation plus prediction – The logical focus of project management research. *Project Management Journal*. 52(4), 319-322. <https://doi.org/10.1177%2F8756972821999945>
- Hair, J. F., Sarstedt, M., Ringle, C. M., & Mena, J. A. (2012). An Assessment of the use of Partial Least Squares Structural Equation Modeling in Marketing Research. *Journal of the Academy of Marketing Science*, 40(3), 414-433.
- Harris, J. M., Ciorciari, J., & Gountas, J. (2018). Consumer neuroscience for marketing researchers. *Journal of Consumer Behaviour*, 17(3), 239-252.
- Harrison, D.E., Ajjan, H. and Coughlan, A. (2018). Working paper: Understanding social media sentiment, positioning and engagement: the impact on direct selling sales performance. *Informs Society of Marketing Science Proceedings*.
- Harrison, D. E., & Hair, J. F. (2017). The Use of Technology in Direct-Selling Marketing Channels: Digital Avenues for Dynamic Growth. *Journal of Marketing Channels*, 24(1-2), 39-50.
- Henke, N., Levine, J., McInerney, P. (2018). You don't have to be a data scientist to fill this must have analytics role. *Harvard Business Review*. February 2005.
- Huang, M. H., & Rust, R. T. (2018). Artificial intelligence in service. *Journal of Service Research*, 21(2), 155-172.
- Hulland, J., Baumgartner, H., & Smith, K. M. (2018). Marketing survey research best practices: evidence and recommendations from a review of JAMS articles. *Journal of the Academy of Marketing Science*, 46(1), 92-108.
- Hulland, J., & Miller, J. (2018). “Keep on Turkin”? *Journal of the Academy of Marketing Science*, 46(5), 789-794.
- Hult, T., Hair, J. F., Sarstedt, M., Ringle, C., Proksch, D., and Pinkwart, A. (2018). Addressing Endogeneity in International Marketing Applications of Partial Least Squares Structural Equation Modeling, *Journal of International Marketing*, forthcoming.
- Humphreys, A., & Wang, R. J. H. (2017). Automated text analysis for consumer research. *Journal of Consumer Research*, 44(6), 1274-1306.
- Kanellos, M. (2016). 152,000 smart devices every minute in 2025: IDC outlines the future of smart things. March 2016, retrieved from <https://www.forbes.com/sites/michaelkanellos/2016/03/03/152000-smart-devices-every-minute-in-2025-idc-outlines-the-future-of-smart-things/>.
- Kaplan, A., & Haenlein, M. (2019). Siri, Siri, in my hand: Who's the fairest in the land? On the

- interpretations, illustrations, and implications of artificial intelligence. *Business Horizons*, 62(1), 15-25.
- Krause, D., Luzzini, D., & Lawson, B. (2018). Building the Case for A Single Key Informant in Supply Chain Management Survey Research. *Journal of Supply Chain Management*, 54(1), 42-50.
- Kull, T. J., Kotlar, J., & Spring, M. (2018). Small and Medium Enterprise Research in Supply Chain Management: The Case for Single-Respondent Research Designs. *Journal of Supply Chain Management*, 54(1), 23-34.
- Kumar, V. (2018). Transformative Marketing: The Next 20 Years. *Journal of Marketing*. 82(4), 1-12.
- Leonelli, S. (2014) What difference does quantity make? On the epistemology of Big Data in biology. *Big Data Soc* 1: 1–11.
- Madnick, S. E., Wang, R. Y, Lee, Y. W., & Zhu, H. (2009) Overview and framework for data and information quality research. *ACM Journal of Data and Information Quality*, 1(1), Article 2.
- Manley, S. C., Hair, J. F., Williams, R. I., and McDowell, W. C. (2021). Essential new PLS-SEM analysis methods for your entrepreneurship analytical toolbox. *International Entrepreneurship and Management Journal*. 17(1), 1-21. DOI: 10.1007/s11365-020-00687-6
- Matthews, L., Hair, J., Matthews, R. (2018). PLS-SEM: The holy grail for advanced analysis. *Marketing Management Journal*, 28(1), 1-13.
- Matthews, L., Hair, J.H., Sarstedt, M., and Ringle, C. (2016). Identifying and Treating Unobserved Heterogeneity with FIMIX-PLS: Part II – Case Study, *European Business Review*, 28(2).
- Ma, L., & Sun, B. (2020). Machine learning and AI in marketing—Connecting computing power to human insights. *International Journal of Research in Marketing*, 37(3), 481-504.
- Mela, C.F. & Moorman, C. (2018). Why Marketing Analytics hasn't lived up to its promise. May 30, retrieved from <https://hbr.org/2018/05/why-marketing-analytics-hasnt-lived-up-to-its-promise>.
- Moore, Z., Harrison, D., and Hair, J. F. (2021). Data Quality Assurance Begins Before Data Collection and Never Ends: What Marketing Researchers Absolutely Need to Remember (2022). *International Journal of Market Research*, 63 (6), <https://doi.org/10.1177%2F14707853211052183>.
- Netzer, O., Lattin, J. M., & Srinivasan, V. (2008). A hidden Markov model of customer relationship dynamics. *Marketing science*, 27(2), 185-204.
- Roh, J. A., Whipple, J. M., & Boyer, K. K. (2013). The effect of single rater bias in multi-stakeholder research: A methodological evaluation of buyer-supplier relationships. *Production and Operations Management*, 22(3), 711-725.
- Rossi, B. (2015), From insight to action: why prescriptive analytics is the next big step for data. February 2015, retrieved from <https://www.information-age.com/insight-action-why-prescriptive-analytics-next-big-step-big-data-123458977/>.
- Rossi, P. E., & Allenby, G. M. (2003). Bayesian statistics and marketing. *Marketing Science*, 22(3), 304-328.
- Sarstedt, M., Hair, J. F., Nitzl, C., Ringle, C. M. and Howard, M. C. (2020). Beyond a tandem analysis of SEM and PROCESS: Use PLS-SEM for mediation analyses! *International Journal of Market Research*, 62(3), 288-299. <https://doi.org/10.1177/1470785320915686>.
- Sarstedt, M., Hair, J. F., & Ringle, C. M. (2022). “PLS-SEM: Indeed a silver bullet” – Retrospective observations and recent advances, *Journal of Marketing Theory & Practice*, <https://doi.org/10.1080/10696679.2022.2056488> .

- Sarstedt, M., Hair, J. F., Pick, M., Liengard, B., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade, *Psychology and Marketing*, 1-30.
- Sarstedt, M., & Ringle, C. M. (2010). Treating unobserved heterogeneity in PLS path modeling: a comparison of FIMIX-PLS with different data analysis strategies. *Journal of Applied Statistics*, 37(8), 1299-1318.
- Sarstedt, M., Ringle, C., & Hair, J. (2017). Identifying and Treating Unobserved Heterogeneity in PLS: A State of the Art Review,” chapter in *Recent Developments on Partial Least Squares Structural Equation Modeling: Basic Concepts, Methodological Issues and Applications*, H. Latan, and Richard Noonan, Editors, Springer International Publishing AG, Switzerland.
- Shankar, V. (2018). How artificial intelligence (AI) is reshaping retailing. *Journal of retailing*, 94(4), vi-xi.
- Shmueli, G., Ray, S., Velasquez Estrada, J. M., & Chatla, S. B. (2016). The elephant in the room: Predictive performance of PLS models, *Journal of business Research*, 69, 4552-4564.
- Shmueli, G., Sarstedt, M., Hair, J. F., Cheah, J., Ting, H., Vaithilingam, S., & Ringle, C. M. (2019). Predictive model assessment in PLS-SEM: Guidelines for using PLSpredict. *European Journal of Marketing*, 53(11), 2322-2347. <https://doi.org/10.1108/EJM-02-2019-0189>.
- Sivarajah, U., Kamal, M. M., Irani, Z., & Weerakkody, V. (2017). Critical analysis of Big Data challenges and analytical methods. *Journal of Business Research*, 70, 263-286.
- Sorescu, A., Warren, N. L., & Ertekin, L. (2017). Event study methodology in the marketing literature: an overview. *Journal of the Academy of Marketing Science*, 45(2), 186-207.
- Snyman, J. A. and Wilke, D. N. (2005). *Practical mathematical optimization: basic optimization theory and gradient-based algorithms*. Springer.
- Varadarajan, R. (2018). A Commentary on “Transformative Marketing: The Next 20 Years”. *Journal of Marketing*. 82(4), 15-18.
- Verma, P., Agarwal, S., Kachroo, P., & Krishen, A. (2017). Declining transportation funding and need for analytical solutions: dynamics and control of VMT tax. *Journal of Marketing Analytics*, 5(3-4), 131-140.
- Wedel, M., & Kannan, P. K. (2016). Marketing analytics for data-rich environments. *Journal of Marketing*, 80(6), 97-121.
- White, S., (2017). What is a data scientist? A key data analytics role and a lucrative career. <https://www.cio.com/article/3217026/data-science/what-is-a-data-scientist-a-key-data-analytics-role-and-a-lucrative-career.html>, accessed August 2018.
- Worm, S., Bharadwaj, S. G., Ulaga, W., & Reinartz, W. J. (2017). When and why do customer solutions pay off in business markets?. *Journal of the Academy of Marketing Science*, 45(4), 490-512.
-