

FERRAMENTAS APLICADAS À ANÁLISE DE DADOS: UM ESTUDO BIBLIOMÉTRICO FEITO A PARTIR DO ANAIS DO ENCONTRO NACIONAL DE ENGENHARIA DE PRODUÇÃO NOS ANOS DE 2019 E 2020.

Gabrielly Amaral Costa Martins¹
Jonas da Conceição Ricardo²

Resumo: Com a revolução industrial que presenciamos atualmente e a evolução da tecnologia, temos um aumento no número de dados gerados a cada dia. Estes dados são armazenados todos os dias em volumes cada vez maiores, porém dados desconexos, não trazem nenhum tipo de benefício a empresa. Diante deste fato há a necessidade de tratar essas grandes bases de dados, a fim de transformar estes em informações que possam trazer conhecimento para tomada de decisões. Esses trabalhos de estratificações e análises, juntamente com conhecimentos e melhoria contínua, acarreta processos e serviços cada vez mais robustos para as empresas, gerando competitividade em um mercado onde cada vez mais o diferencial traz a liderança. Este trabalho apresenta uma visão micro desse universo, onde foi verificado diversos artigos dos anos de 2019 e 2020 publicados pela ENEGEP (Encontro Nacional de Engenharia de Produção), sendo feito as análises desses artigos de forma mista, usando tanto a análise qualitativa quanto a quantitativa. Diante dessa análise foi possível a identificação das principais ferramentas utilizadas para a realização destes processos, assim como as regiões do Brasil onde mais ocorrem, o número de artigos sobre o tema e o gênero que mais escreve a respeito.

Palavras-chave: Tecnologia. Dados. ENEGEP. Ferramentas. Análise.

Abstract: With the industrial revolution we are currently witnessing and the evolution of technology, we have an increase in the number of data generated every day. The data is stored every day in increasingly larger volumes, but disconnected data does not bring any kind of benefit to the company. Given this fact, there is a need to deal with these large databases, in order to transform these into information that can bring knowledge for decision making. These stratification and analysis work, together with knowledge and continuous improvement, lead to increasingly robust processes and services for companies, generating competitiveness in a market where the differential increasingly brings leadership. In order to connect these data, analysis and stratification tools are developed and improved to facilitate the visualization and understanding of the material. During this work, we verified several articles from the years 2019 and 2020 published by ENEGEP, and these articles were analyzed in a mixed way, using both qualitative and quantitative analysis. It was possible to identify, through this analysis, the main tools used to carry out these processes, as well as the regions of Brazil where they occur the most, the number of articles on the topic and the gender that writes the most about it

Keywords: Technology. Data. ENEGEP. Tools. Analysis.

¹ Acadêmica do Curso de Bacharelado em Engenharia de Produção da Universidade Estácio de Sá - UNESA, e-mail: gaby__martins@live.com.

² Professor e Pesquisador da Universidade Estácio de Sá, Doutorando em Ciências Tecnologia e Educação (CEFET/RJ), e-mail: jnsricardo@gmail.com.





1 INTRODUÇÃO

Atualmente, estamos vivenciando a 4ª Revolução Industrial, ou também conhecida como Indústria 4.0 (TROYJO, MARCOS); trata-se de uma revolução onde há integração de tecnologia em itens essenciais ou cotidianos que anteriormente, poderiam ser desprovidos de “inteligência”, ou acesso à *internet*. Chamamos esta transformação de Internet das Coisas *Internet of Things* (IOT).

Segundo Daniel Laper, a IoT é um passo essencial para uma série de outras tecnologias como a Inteligência Artificial e *Big Data* e por definição, a Inteligência Artificial se alimenta e beneficia de grandes volumes de dados, em relação direta em que quanto mais informação disponível, maior a capacidade de identificar padrões, aprender e gerar insights e valor. Nesse contexto, o grande valor de soluções IOT está na maximização da coleta e acesso aos dados.

Por sua vez, aumenta-se a cada dia a produção de dados de fato, mais de 2,7 *zettabytes* de dados existem no universo digital de hoje, e isso deverá crescer para 180 *zettabytes* em 2025 (CETAX, 2021). Todo esse volume de dados gerados, necessita de análises para a disseminação de informação e ações para melhoria do processo ou produto estudado.

O uso de ferramentas para executar estas análises acaba por ser essencial, já que com a quantidade gigantesca de dados gerados, torna-se impossível fazer isto de forma “Braçal”.

Neste estudo será apresentado um panorama sobre o histórico e origem dos dados e suas análises, o crescimento do volume de dados e o surgimento do termo *Big Data*, seguido pela contextualização das características e tipos de dados existentes e assim chegando ao ápice do estudo: as ferramentas mais utilizadas e suas minúcias.

2 REFERENCIAL TEÓRICO

Segundo estudos realizados, a história sugere que os primeiros humanos recorreram a paus e pedras para ajudar a prever as tendências prováveis nas vendas de itens que eram considerados ‘novidades’. Em 1800, quando surgiu a era industrial, introduziu-se a necessidade de gestão de negócios como disciplina científica: Henry Ford, por exemplo, aplicou isso em sua nova linha

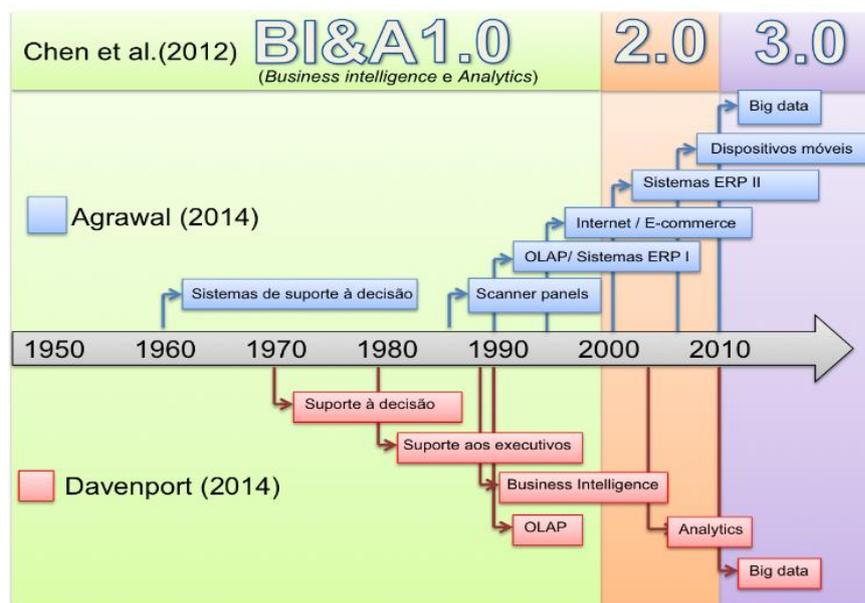


de produção para o Ford Modelo T, com o objetivo de medir o tempo de montagem dos componentes de seus veículos (BORINI, 2017).

À medida que a tecnologia evoluiu no decorrer do século XX, computadores ganharam papel fundamental na implementação e adoção da análise de negócios por meio da introdução da tecnologia. A ferramenta usada foi chamada de “Sistemas de Apoio para a Decisão” (*Decision Support Systems*) no início dos anos 70 (BORINI, 2017).

Sabendo que o chamado *analytics* baseia-se na análise de informações (dados), surgiu a necessidade de aliar não só o sustento do armazenamento desses dados, mas também sua organização e estruturação de forma sistêmica, facilitando a análise dessas informações por motores analíticos já mencionados. É por isso que, durante os últimos anos, os *Data Warehouses*, considerados a primeira geração desse tipo de tecnologia, ganharam popularidade no meio corporativo, ajudando a organizar grandes quantidades de dados obtidos (BORINI, 2017).

Figura 3: Linha do tempo da evolução dos sistemas de informação nas empresas



. Fonte: Adaptado de Chen et al.(2012), Agrawal (2014) e Davenport (2014).

Até pouco tempo atrás, a maioria dos resultados analíticos eram destinados a gestores de primeira linha dentro das empresas e corporações. No entanto, nessas últimas décadas tudo



mudou com o advento das redes de computadores, internet, *Big Data* e serviços em nuvem (BORINI, 2017).

2.1- Tipos de Dados

2.1.1 -Dados Estruturados

Os dados estruturados são aqueles que possuem uma estrutura determinada, com categorias, clusters e definições, como localização, vendas e informações sobre o perfil de clientes, contatos entre outros (NASCIMENTO, 2017).

Dados estruturados são encontrados nos bancos de dados que, para armazenar qualquer dado, precisam ter muito bem definidos onde cada informação estará. *Softwares* de empresas como ERP, CRM, sistemas financeiros, sistemas de RH entre outros possuem dados estruturados. (NASCIMENTO, 2017).

2.1.2 Dados não estruturados

Os dados não estruturados são os mais complexos de se trabalhar, pois não existe neles uma estruturação sequer, sendo necessária intervenção humana para sua preparação. Estamos falando dos dados de mídias sociais, como *YouTube*, *Facebook*, *Instagram*, portais de notícias etc. Todas essas mídias que citamos lidam com dados em vídeos, imagens, textos e até mesmo áudio, por isso a complexidade para o tratamento desses dados é maior (NASCIMENTO, 2017).

Hoje, por exemplo, conseguimos monitorar as mídias sociais extraindo comentários e menções públicas sobre determinada palavra-chave. Ou seja, podemos monitorar o que as pessoas estão dizendo sobre sua empresa ou até mesmo o mercado em que você atua de forma geral. Porém, forma mais confiável de estruturar esses dados ainda é por meio da intervenção humana. Pois deve ser realizada uma análise prévia do que está sendo comentado e o contexto daquela menção. Temos diversos tipos de comentários que podem parecer positivos, mas que na realidade são palavras com sarcasmo, ironias e, na maioria das vezes os robôs não conseguem captar (NASCIMENTO, 2017).

Além disso, é necessário criar *tags*, como se fossem categorias para o que estão dizendo em um determinado contexto. Esse tipo de trabalho (por enquanto) deve ser realizado por um ser



humano pois envolve diversas particularidades de acordo com cada projeto (NASCIMENTO, 2017).

Uma dúvida muito comum é “Se eu encontrar 20 mil comentários, um ser humano precisa categorizar, taggear e estruturar esses dados?”. Exatamente. Por esse motivo consideramos a complexidade dos dados não estruturados maior, mais trabalhosa e morosa (NASCIMENTO, 2017).

2.2 Análise de Dados

Em termos bem gerais, a prática consiste no processo de transformação de dados em informações úteis aos mais diversos propósitos de negócios. Um dos objetivos mais comuns para a realização dessa tarefa é a melhora da performance das empresas no mercado em tempos de transformação digital (CORTEX..., 2021).

2.3 Tipos de Análise de dados

Após coletar dados, é necessário decidir um meio para realizar uma análise. Estas podem ser divididas em 4 tipos: Análise Preditiva, Análise Descritiva, Análise Prescritiva, Análise Diagnóstica. Na *tabela 1*, podemos entender a aplicação para cada tipo de análise citada acima e os métodos utilizados para sua realização.

Tabela 1: Tipos de análise e suas peculiaridades.

Tipo	Definição	Métodos Utilizados
Análise Preditiva	Este tipo de análise visa criar padrões para explicar certos fenômenos e prever futuros acontecimentos.	Isso é feito, por exemplo, por métodos como a regressão, no qual um certo fato é condicionado a um outro, podendo ou não se confirmar essa relação (FIVE ACTS, 2021).
Análise Descritiva	Este tipo de análise busca explicar a realidade atual ou próxima, através de dados reais.	Tem como método principal procedimentos relacionados a técnica de estatística descritiva (FIVE ACTS, 2021).
Análise Prescritiva	Esse tipo de análise busca prever comportamentos através de simulações auxiliando na tomada de decisões quanto o futuro.	São utilizadas ferramentas estatísticas que também podem ser aplicados em análises descritivas e preditivas.
Análise Diagnóstica	Visa detectar causas de um certo fenômeno ou comportamento (FIVE ACTS, 2021).	São usadas ferramentas mais específicas, com destaque para os testes A/B (FIVE ACTS, 2021).

Fonte: Elaborada pelos autores



No banco de dados da ENEGEP, foram pesquisados artigos dos anos de 2019 e 2020 que tratavam de análises de dados, e avaliados quais as ferramentas utilizadas para a realização desta. Foram catalogados 63 artigos de diversas regiões do Brasil, que será apresentado na seção seguinte.

3 - CORPUS DA PESQUISA e PROCEDIMENTO METODOLÓGICO

O método utilizado para este estudo foi a revisão bibliográfica descritiva, pois para proporcionar o avanço em um campo do conhecimento é preciso primeiro conhecer o que já foi realizado por outros pesquisadores e quais são as fronteiras do conhecimento naquela (VIANNA, 2001). Já a análise dos artigos utilizou o que é caracterizado como método misto, pois trata-se tanto de uma análise qualitativa e quantitativa (CRESWELL, 2007).

O corpus da pesquisa é constituído por 63 artigos originados do banco de dados da ENEGEP, que está alocado no site da ABEPRO (Associação Brasileira de Engenharia de Produção) como mostra a **Figura 1**.

Figura 1: Página de pesquisa no portal ENEGEP.



The image shows a screenshot of the search interface on the ABEPRO website. At the top left is the ABEPRO logo (Associação Brasileira de Engenharia de Produção). The main heading is "Anais do Encontro Nacional de Engenharia de Produção - Enegep". Below this, there are several tabs: "Evento", "Pesquisa de Trabalhos" (which is highlighted in blue), "Áreas", "Sub-áreas", and "Autores". Under the "Pesquisa de Trabalhos" tab, there are four search criteria: "Evento: (Event)" with a dropdown menu showing "2020 - ENEGEP"; "Área: (Area)" with a dropdown menu showing "Selecione a área (Select the area)"; "Pesquisa: (Search)" with a text input field containing "ferramenta de analise de dados"; and "Autor: (Author)" with an empty text input field. A blue "Pesquisar" button is located at the bottom right of the search criteria. Below the search area, the text "Resultado da Pesquisa" is displayed, followed by "Página (Page) : [1] | 2 | 3".

Fonte: portal.abepro.org.br (2021)



Para realizar a escolha de artigos, foram pesquisadas as palavras chaves “Ferramentas de análise de dados” nos anos de 2019 e 2020 respectivamente. Todos os artigos retornados pela pesquisa foram retornados.

Os artigos foram catalogados utilizando a ferramenta *Google Sheets*, e foram levados em consideração os pontos abaixo:

-  Ano
-  Região
-  Estado
-  Gênero dos autores
-  Ferramentas utilizadas
-  Área de pesquisa
-  Palavras-chave
-  Título do artigo

O estudo seguiu o fluxo representado abaixo na **Figura 2**

Figura 2: Fluxo de processos para realização do estudo



Fonte: Elaborada pelos autores



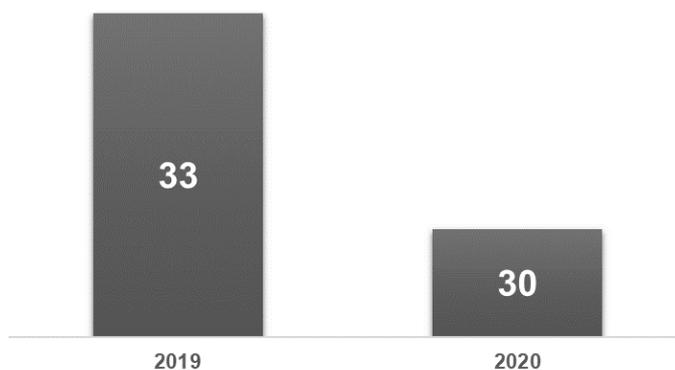
Os artigos selecionados A partir desse fluxo são apresentados e discutidos os dados extraídos dos trabalhos encontrados.

4- RESULTADOS E DISCUSSÃO

Ao analisarmos os 63 trabalhos que compõe o corpus do estudo, pudemos observar o volume de artigos que utilizaram ferramentas para análise de dados nos anos de 2019 e 2020 demonstrado no Gráfico 1.

Podemos dizer que o número segue uma certa linearidade entre os dois anos com uma variação de apenas 9% que pode ser justificado devido a pandemia que comprometeu e modificou diversos aspectos da vida humana.

Gráfico 1: Quantitativo de artigos analisados.



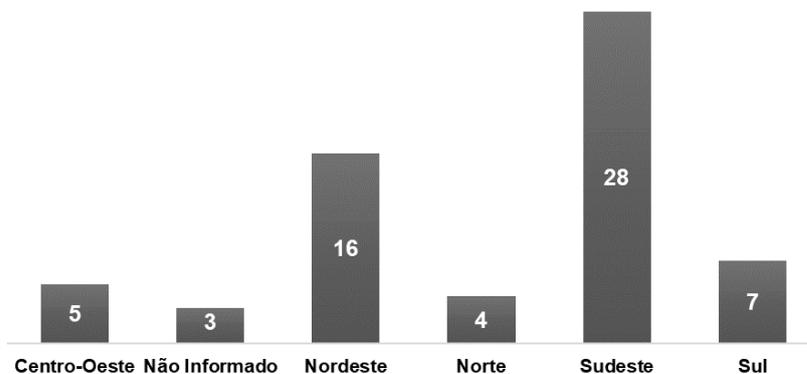
Fonte: elaborado pela autora.

Falando a respeito da distribuição geográfica dos artigos, podemos ver pelo Gráfico 2 que o Sudeste lidera as publicações. A geografia da produção e colaboração científica no país é marcada por intensa heterogeneidade espacial, com concentração sistemática da produção e dos fluxos de conhecimento na região Sudeste. (SIDONE, 2016).





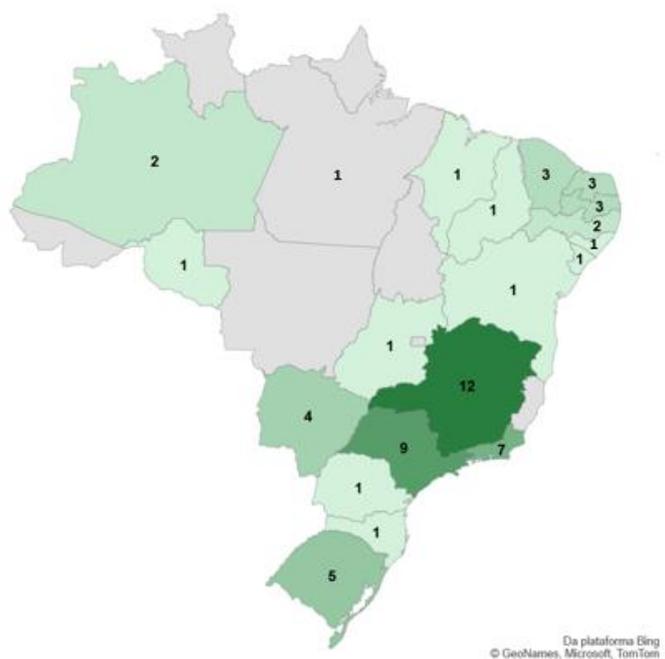
Gráfico 2: Representatividade de regiões brasileiras nos artigos.



Fonte: elaborado pelos autores.

Verificando mais a fundo a distribuição geográfica dos artigos no gráfico de mapa (Gráfico 3), podemos verificar que o estado de Minas Gerais lidera as publicações no Sudeste, que como visto anteriormente, representa o maior publicador das pesquisas analisadas. As únicas regiões que não detêm publicações em todos os seus estados são as Norte e Centro-Oeste. O que representa um aumento na representatividade regional científica por parte de outras regiões do Brasil que anteriormente não tinham participação no ramo de pesquisas científicas.

Gráfico 3: Mapa representando estados detentores de artigos publicados.

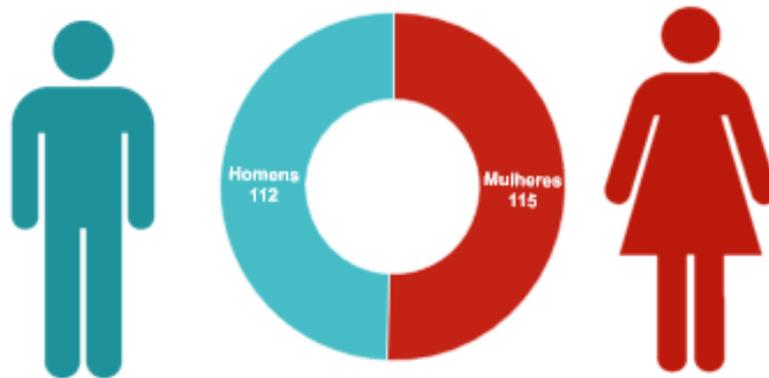


Fonte: elaborado pelos autores.



Quanto ao gênero dos autores que redigiram os artigos analisados, temos o **gráfico 4**, todos os artigos analisados contavam com mais de um autor. Catalogando a todos, temos um número superior de mulheres em relação ao número de homens. O que mostra uma maior participação feminina na área de engenharia voltada a pesquisas científicas.

Gráfico 4: Representatividade de gênero de autores.



Fonte: elaborado pelos autores.

Todos os artigos analisados foram categorizados em nichos de áreas de aplicação e estudo. No **gráfico 5**, podemos verificar que a maioria de ferramentas para análises de dados foram aplicadas em pesquisas voltadas para Engenharia operacional.

Gráfico 5: Artigos categoriados por temas



Fonte: elaborado pelos autores.



No Gráfico 6 podemos encontrar as ferramentas utilizadas para realizar a análise de dados para cada artigo analisado. Foram encontradas 80 diferentes ferramentas, mas em unanimidade o *Excel* é a mais utilizada, sendo muitas vezes combinado com outras ferramentas para obter resultados eficazes. Dos artigos em debate, 49% utilizaram o *Excel* como ferramenta.

O que é esperado, já que o *Excel* é utilizado por mais de 90% das corporações atuais, incluindo a própria NASA (Administração Nacional da Aeronáutica e Espaço) e agências do governo Norte-Americano (G1, 2020). O programa apresenta soluções para as mais variadas aplicações para quem conhece suas funcionalidades, possibilitando a criação de *Dashboards*, gráficos, relatórios, controles etc.

Gráfico 6: Gráfico Mapa contendo ferramentas para análises utilizadas nos artigos.



Fonte: elaborado pelos autores.

Dos artigos que utilizaram a ferramenta *Excel* (39 artigos), podemos visualizar que 65% puderam concluir sua análise apenas com conclusões obtidas através da ferramenta e suas aplicações, já 35% usaram o *Excel* para auxiliar junto de outro *software* ou ferramenta.

5- CONSIDERAÇÕES FINAIS

Nesta análise preliminar, foram encontradas uma ampla gama de materiais com aplicações na área de dados, mostrando a existência de diversos tipos de dados que podem ser



trabalhados em situações diferentes para obter as soluções desejadas. Podendo destacar que com o aumento crescente da tecnologia, podemos esperar um aumento significativo de estudos e aplicações nesta área.

Durante o estudo dos artigos pode-se observar diversos fatores sociais, bem como geográficos e de aplicabilidade. Vimos que majoritariamente a região sudeste domina as áreas de pesquisa verificadas. O que é esperado devido ao privilégio econômico em que ela se encontra. Quanto ao sexo dos autores, houve um equilíbrio grande, com uma pequena margem positiva para o sexo feminino. Os temas voltados a processos de engenharia dominaram as pesquisas indicando que a busca por melhoria através de análises nas atividades operacionais é grande gerando assim a evolução e melhoria contínua dos processos que cada vez mais se busca para alcançar competitividade no mercado. Quanto as ferramentas utilizadas, o *Excel* domina todos as pesquisas, o que se pode atribuir ao fato de ser um *software* oferecido dentro de um computador *Windows* comum ou encontrado facilmente online. O software que já é consolidado também sofre atualizações constantes aumentando ainda mais sua vasta aplicabilidade e mantendo-o no topo da preferência de uso para pesquisadores.

Conclui-se que para ampliar o conhecimento da área, é necessário o aumento de pesquisas futuras sobre o tema utilizando uma amostra maior de dados para realizar a comparação quantitativa assim como realizar revisões mais detalhadas sobre o tema que podem vir a auxiliar em futuras decisões de ferramentas mais adequadas para análises mais robustas e eficientes.

REFERÊNCIAS.

ABEPRO. **Anais do Encontro Nacional de Engenharia de Produção – Enegep**. Disponível em: <<http://www.abepro.org.br/publicacoes/index.asp?pesq=ok&ano=&area=&pchave=ferramenta+de+analise+de+dados&autor>>. Acesso em: 10 de julho de 2021

BLOG CETAX, **Data Analytics, Big Data, Data Science**. Disponível em: <<https://www.cetax.com.br/blog/data-science-vs-big-data-vs-data-analytics/>>. – Acesso em: 08 de agosto de 2021 às 21:19.

BORINI , **Análise de dados nos negócios: um passeio pela história**. Disponível em: <<https://itforum.com.br/noticias/analise-de-dados-nos-negocios-um-passeio-pela-historia/>>. – Acesso em: 28 de agosto de 2021 às 15:52.

CIDONE, (2016). **A ciência nas regiões brasileiras: evolução da produção e das redes de colaboração científica**. São Paulo.

CORTEX, **O que é análise de dados? Saiba tudo sobre o tema!**. Disponível em: <<https://cortex-intelligence.com/blog/inteligencia-de-mercado/o-que-e-analise-de-dados/>>. – Acesso em: 02 de outubro de 2021 às 15:12.



CRESWELL, (2007). Projeto de Pesquisa. Métodos Qualitativo, Quantitativo e Misto, 117 (1).

G1. Excel: domínio desta ferramenta aumenta as chances de crescimento profissional. Disponível em: <<https://g1.globo.com/rj/sul-do-rio-costa-verde/especial-publicitario/ubm/conhecimento-transforma/noticia/2020/02/15/excel-dominio-desta-ferramenta-aumenta-as-chances-de-crescimento-profissional.ghtml>>. Acesso em: 17 de outubro de 2021 às 14:38.

MARKETING POR DADOS, Afinal, o que é Big Data?. Disponível em: <<https://marketingpordados.com/analise-de-dados/o-que-e-big-data/>>. – Acesso em: 28 de julho de 2021 às 21:42.

NASCIMENTO. O que é Big Data? Disponível em: <https://blog.maxieduca.com.br/big-data-investimento/> – Acesso em: 28 de julho de 2021 às 20:58.

SILVEIRA, MARCOLIN E FREITAS. O BIG DATA E SEU USO CORPORATIVO: UMA REVISÃO DE LITERATURA. Rio Grande do Sul, 2015.

TROYJO , Tipos de análise de dados: Conheça os 4 principais!. Disponível em: <<https://www.fiveacts.com.br/tipos-de-analise-de-dados/>>. – Acesso em: 02 de outubro de 2021 às 15:45.

VIANNA, (2001). Metodologia do Trabalho Científico. Um Enfoque Didático da Produção Científica, 68 (2).

