

Capítulo IX – Análise de dados em transformação digital: índice de transformação digital da Agência Brasileira de Desenvolvimento Industrial

Matheus Freire Pessoa ¹

Sofia Auer ²

Tiago Galhoto ³

Tuany Kasiorowski Neves ⁴

Suellen Terroso de Mendonça Ferreira Jacoboski ⁵

Jorge Ferreira De Sa Junior ⁶

João Carlos Silva ⁷

Gustavo Dambiski Gomes de Carvalho ⁸

Ana Cristina Fermino Deschamps ⁹

Eli Carlos Dal Pupo ¹⁰

Joair Rossetto Schelela Junior ¹¹

Thiago Schaedler Uhlmann ¹²

Denice Lusa ¹³

Cassiana Fagundes da Silva ¹⁴

RESUMO

O objetivo deste artigo científico é analisar dados acerca do índice de transformação digital de países elaborado pela Agência Brasileira de Desenvolvimento Industrial. Para a organização dos dados, foi utilizada uma base de dados no formato Excel, a qual possui recursos para manipulação de dados e integração com software, como SPSS. Na análise de dados, foi desenvolvida uma análise estatística de correlação com o software IBM SPSS acerca do sucesso na promoção de TIC, índice de segurança global e contratação de tecnologia avançada. Os resultados indicam correlações positivas e significativas entre as variáveis estudadas.

Palavras-chave: Transformação Digital. Índice de Transformação Digital. Análise de Dados.

¹ UniSenaiPR - Campus São José dos Pinhais

² UniSenaiPR - Campus São José dos Pinhais

³ UniSenaiPR - Campus São José dos Pinhais

⁴ Docente UniSenaiPR - Campus São José dos Pinhais, tuany.neves@sistemafiep.org.br

⁵ Docente UniSenaiPR - Campus CIC, suellen.jacoboski@sistemafiep.org.br

⁶ Docente UniSenaiPR - Campus CIC, jorge.junior@sistemafiep.org.br

⁷ Docente UniSenaiPR - Campus CIC, joao.carlsilva@sistemafiep.org.br

⁸ Docente UniSenaiPR - Campus CIC, gustavo.carvalho@sistemafiep.org.br

⁹ Docente UniSenaiPR - Campus da Indústria, ana.deschamps@sistemafiep.org.br

¹⁰ Docente UniSenaiPR - Campus São José dos Pinhais, eli.pupo@sistemafiep.org.br

¹¹ Docente UniSenaiPR - Campus São José dos Pinhais, joair.junior@sistemafiep.org.br

¹² Docente UniSenaiPR - Campus São José dos Pinhais, thiago.uhlmann@sistemafiep.org.br

¹³ Docente UniSenaiPR - Campus São José dos Pinhais, denice.lusa@sistemafiep.org.br

¹⁴ Docente UniSenaiPR - Campus São José dos Pinhais, cassiana.silva@sistemafiep.org.br

1. INTRODUÇÃO

A transformação digital é um tema emergente tanto para a teoria quanto para a prática devido à evolução da tecnologia e seus potenciais impactos, especialmente as tecnologias habilitadoras que caracterizam a quarta revolução industrial, também conhecida como Indústria 4.0, como cunhado originalmente na Alemanha (ZELLER; HOCKEN; STICH, 2018). As principais tecnologias que caracterizam essa revolução incluem Inteligência Artificial, Impressão 3D, Big Data, Analytics, Computação em Nuvem, Internet das Coisas (IoT), entre outras (CARVALHO et al., 2022; MACHADO; SCHROEDER; CARVALHO, 2022).

O conceito de transformação digital não se limita apenas a essas tecnologias e vai além. O Centro Internacional de Tecnologia de Software (CITS, 2023, local. 1) define Transformação Digital como:

“É o processo contínuo das organizações (de comércio, indústria ou serviços) para integrar tecnologias habilitadoras em todas as áreas, por meio de mudanças fundamentais em estratégia, liderança, estrutura, pessoas, processos, cultura e negócios digitais, com o objetivo de suportar as tomadas de decisões a partir de dados, reduzir desperdícios e agregar valor aos seus clientes e à sociedade.” (CITS, 2023, local. 1)

Além disso, devido ao seu potencial de mudança, a transformação digital tem sido estudada no âmbito das pessoas, das organizações e até mesmo de países e regiões (SANTOS et al., 2023). Nesse contexto, a Agência Brasileira de Desenvolvimento Industrial (ABDI) desenvolveu o Índice de Transformação Digital, que compara o Brasil com outros países em relação a essa transformação. Em especial, o índice busca acompanhar “o desenvolvimento da migração das relações entre governo, mercado (empresas) e consumidores (cidadãos) para o universo digital” (ABDI, 2024, local.1).

Alguns indicadores interessantes incluem segurança cibernética, sucesso de governos na promoção das tecnologias de informação e comunicação (TICs) e contratação de tecnologias avançadas por governos. Apesar da disponibilidade aberta dos dados dessa pesquisa, até o momento, ainda não há estudos ou pesquisas que explorem os dados do índice de transformação digital da ABDI.

Assim, essa pesquisa tem como objetivo analisar dados acerca do índice de transformação digital de países elaborado pela Agência Brasileira de Desenvolvimento Industrial. Em especial, são analisadas as relações entre indicadores de transformação digital como segurança cibernética, sucesso na promoção das TICs e contratação de tecnologias.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. Análise de dados

Análise de dados é o processo de aplicação de técnicas de estatística e lógica para o estudo de informações obtidas, para que resultados relevantes sejam obtidos e decisões possam ser tomadas (SILVA, 2022). Estes dados estão presentes em vários ramos da sociedade, como saúde, tecnologia e educação, além de ser possível aplicar a análise de dados em órgãos públicos e empresas privadas.

A análise de dados possui alguns tipos, a serem aplicados de acordo com as necessidades de cada indivíduo ou organização. As ramificações desta área do conhecimento são: análise preditiva, análise prescritiva, análise descritiva e análise diagnóstica (TAYRANE, 2022). Uma análise preditiva é realizada para realizar projeções futuras com base nos resultados da análise obtida. Ela evidencia possíveis resultados que podem acontecer a partir de determinadas condições (SILVA, 2022).

Além disso, uma análise prescritiva diz respeito às consequências de determinadas ações ou acontecimentos. Esta vertente diz respeito ao melhor caminho para o atingimento de uma determinada meta (TAYRANE, 2022). Um outro ramo de análise de dados, a análise descritiva, é a descrição de eventos, objetos ou hipóteses avaliadas, sendo usada para descrever comportamentos e tendências (TAYRANE, 2022).

O outro tipo de análise de dados é a análise diagnóstica. Ela é utilizada para estudar relações de causa e efeito nos itens analisados (TAYRANE, 2022). Com ela, é possível desenvolver ações com base nas tendências e cenários obtidos.

Um projeto de análise de dados faz o uso de uma intensa aplicação de estatística, com os respectivos tópicos sendo usados conforme os requisitos e necessidades de resultados. Constituem-se ramificações estatísticas aplicadas à análise de dados correlações, média, variância e desvio padrão. Uma média diz respeito aos valores que podem representar uma tendência geral do estudo.

Enquanto isso, a variância é um conceito que avalia o quão distante um valor ou conjunto de valores estão da média. Quanto maior é a variância, maior é a distância de itens em relação à média. Desvio padrão representa os quão distantes determinados itens estão da média (FIELD, 2005, p.33).

Uma representação gráfica utilizada na estatística é o histograma, o qual pode ser utilizado para resumir uma base de dados, de forma a facilitar a representação das informações e tirar conclusões (SIQUEIRA, 2023). Outra representação é o gráfico de

barras, o qual é usado para representar duas variáveis, bem como os comportamentos entre elas.

Outra análise de estatística é o teste de Kruskal-Wallis. Esse teste permite comparar três ou mais grupos de amostras, com base na média desses grupos (OLIVEIRA, 2020). Correlação é um conceito de estatística que permite avaliar se há alguma relação entre as variáveis estudadas e mudanças entre elas. Uma correlação positiva significa que se uma variável aumenta, a outra também aumenta. Uma correlação negativa significa que se uma variável diminui, a outra também diminui. Uma correlação nula significa que não há uma correlação entre variáveis (FIELD, 2005, p.125). A análise fatorial serve para encontrar e agrupar variáveis, de acordo com características em comum (A análise fatorial é empregada para compreender melhor a estrutura de um grupo de variáveis, e assim melhor gerenciar a base de dados, reduzindo o tamanho da base de dados e mantendo a informação original (CARVALHO, 2021; FIELD, 2005, p.660).

A área de análise de dados também se relaciona com a área de inteligência artificial, como a área de aprendizado de máquina. Esta área diz respeito ao aprendizado com base em dados, hipóteses e tendências, com o uso de uma base de dados, estatística e programação. Os conceitos de análise de dados também podem ser empregados para fazer com que os dispositivos aprendam a reconhecer tendências, padrões e cenários, para auxiliar na tomada de decisões.

3. METODOLOGIA

Neste trabalho, a base de dados utilizada foi o Índice de Transformação Digital da Agência Brasileira de Desenvolvimento Industrial (ABDI, 2023). A base de dados diz respeito ao índice de transformação digital do Brasil comparado a outros países. As dimensões presentes descritas na página de internet do índice são: estrutura produtiva, ambiente empresarial, segurança e participação e cultura digital. Por outro lado, na base de dados disponibilizada, as dimensões presentes são: E-cidadania, Emprego, Inclusão Digital; Infraestrutura; Inovação; e Negócios. Apesar da busca de informações, não foi identificado um documento sobre a metodologia da relação entre essas dimensões da página da internet e da base de dados.

Neste artigo, as ramificações presentes de análise de dados são a análise descritiva, uma vez que há a descrição das variáveis envolvidas, com base nos conceitos de estatística. Os dados foram inicialmente baixados em .CSV, conforme disponibilizado na

plataforma da ABDI. Contudo, os dados não estavam pré-organizados para uma análise direta por países. Para solucionar a questão, especificou-se o uso de uma base de dados no formato Excel. Desta forma, aplicou-se diferentes filtros relacionados ao índice “indicador”, para estudar o comportamento deste item em relação aos outros filtros. Os resultados filtrados foram então colocados em diferentes abas da mesma planilha, para efeito de comparação e organização por países.

Nesta base de dados, existe uma grande quantidade de países. O estudo abrange quase todos os continentes do mundo, exceto a Antártida, além do Continente Americano estar dividido entre América do Norte e América do Sul. Além disso, os países permanecem agrupados de acordo com o continente a qual pertencem.

Há a presença de uma coluna chamada de “Dimensão”. Ademais, há as colunas “Indicador” e “Descrição_Indicador”. Assim, observa-se a composição das dimensões por indicadores. “Fonte” é a variável que detém as organizações responsáveis pelos dados. Essas organizações, ao analisar a base de dados, são OECD (Organização Para Cooperação e Desenvolvimento Econômico), The Fletcher School, BID (Banco Interamericano de Desenvolvimento), WEF (Fórum Econômico Mundial, na sigla em Português), Tufts University e ONU (Organização das Nações Unidas). Ademais, existe a variável “Ano”, que se refere ao período no qual a informação foi retirada, sendo esse período compreendido entre 2012 e 2018.

Neste artigo, foram discutidos os indicadores “Índice Global de Segurança Cibernética”, que está relacionado com a dimensão “Infraestrutura”. Enquanto isso, os indicadores “Sucesso do Governo na Promoção das TIC” e “Contratação de Tecnologia Avançada Pelo Governo” estão relacionados com a dimensão “E-Cidadania”.

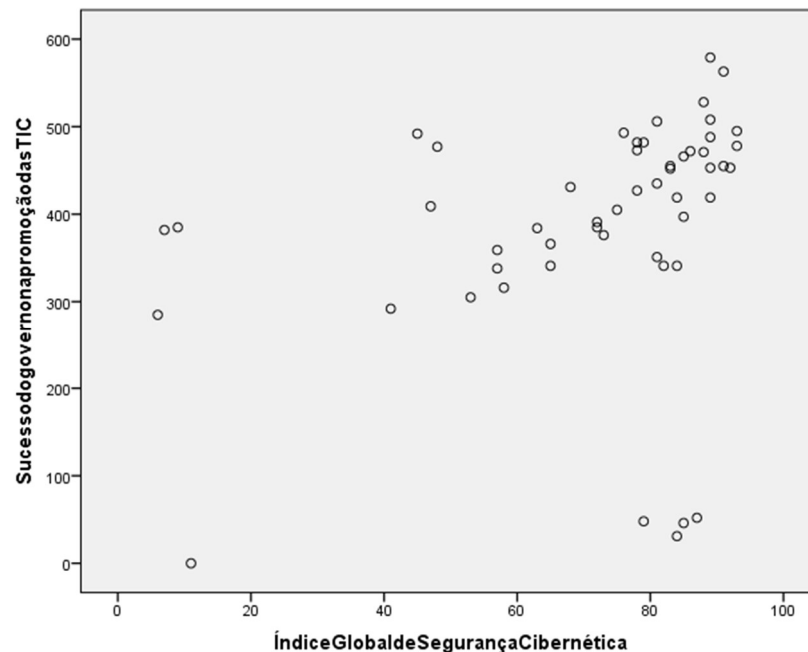
Para a análise de dados, envolveu-se o uso da linguagem do software SPSS, a qual é muito usado em áreas como inteligência artificial, aprendizado de máquina e análise de dados. Ele foi usado para melhor entender o comportamento das variáveis, em função de seus valores. Ele foi empregado em virtude da rápida possibilidade em desenvolver representações estatísticas e gráficos, para a tomada de decisões.

Com o uso do SPSS, foram geradas algumas tabelas, as quais trazem algumas informações sobre as variáveis, como estatísticas descritivas e correlações. A análise de correlação corresponde à principal análise de relação de variáveis neste artigo e, por isso, gráficos de dispersão de variáveis também foram elaborados.

4. APRESENTAÇÃO E DISCUSSÃO DOS RESULTADOS

Na Figura 1, existe a comparação entre o sucesso do governo na promoção de TIC e índice global de segurança cibernética. Este gráfico analisa se pode ser também papel do Estado criar ações no que diz respeito a evitar ataques cibernéticos em empresas e organizações. Neste gráfico, é possível notar a proximidade dos pontos no canto superior direito, o que mostra uma menor dispersão de dados. Essa menor dispersão de dados mostra uma maior correlação de variáveis ($r=0,291$; $p\text{-valor}<0,05$), ou seja, se o valor de uma aumentar, o valor correspondente à outra também cresce.

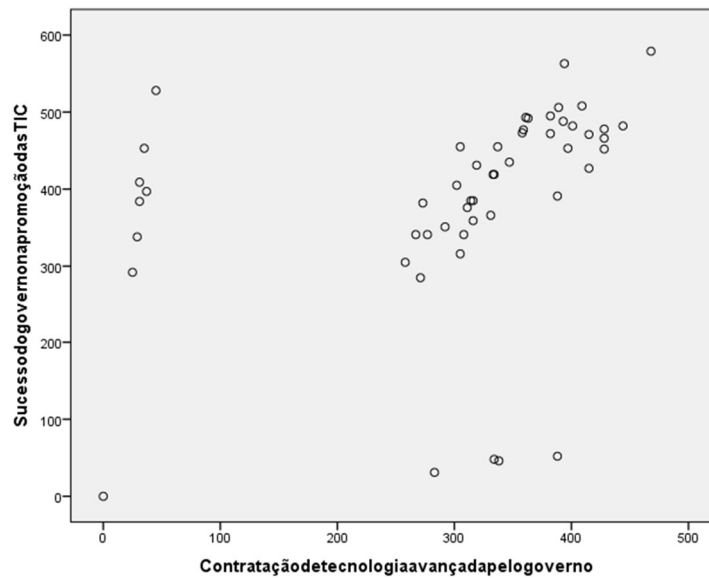
Figura 1 – Sucesso Governamental x Segurança



Fonte: Dos autores.

A Figura 2 mostra uma comparação entre o sucesso do governo em suas políticas de TIC e o uso de tecnologia avançada por parte do governo. Neste gráfico, existe uma avaliação se o uso de novas tecnologias por parte do governo pode ajudar o mesmo a ter êxito em suas políticas científicas e tecnológicas. O agrupamento dos pontos evidencia uma maior relação entre essas variáveis ($r=0,295$; $p\text{-valor} <0,05$), ou seja, se houver um maior uso de novas tecnologias por parte do governo, as políticas governamentais relacionadas a ciência e tecnologia têm mais chances de êxito.

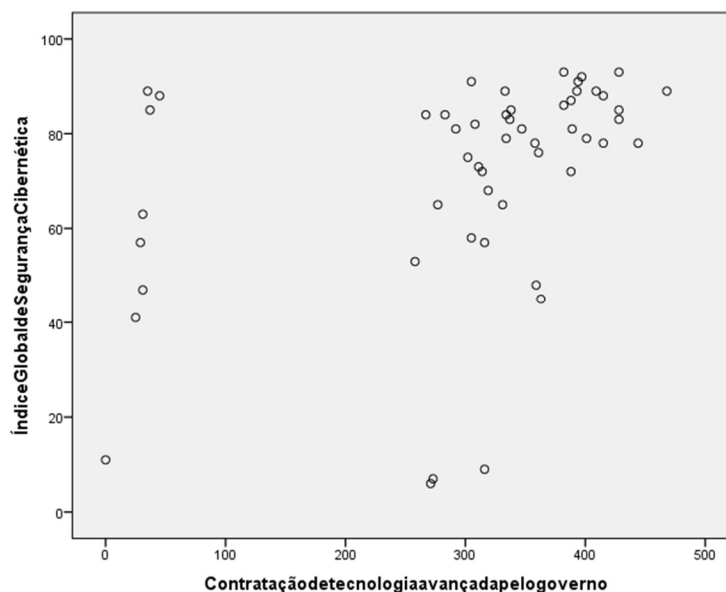
Figura 2 – Uso de tecnologia avançada x Sucesso Governamental



Fonte: Dos autores.

A Figura 3 mostra uma comparação entre o índice global de segurança cibernética e a contratação de tecnologia avançada pelo governo. Nesse gráfico, existe uma avaliação se o uso de novas tecnologias torna possível uma maior chance de evitar ataques virtuais. Pelo gráfico, nota-se que existe um maior agrupamento de pontos no canto superior direito, ainda que haja alguns pontos mais dispersos. Isso significa que há uma correlação significativa ($r=0,375$; $p\text{-valor} < 0,01$) entre essas variáveis, ou seja, usar novas tecnologias pode ajudar a reduzir ataques cibernéticos.

Figura 3 – Contratação de tecnologia avançada x índice global de segurança cibernética



Fonte: Dos autores.

A tabela 01 mostra uma comparação de correlações entre as variáveis analisadas. Ao analisar a tabela, evidencia-se que não há valores negativos, ao passo que há valores destacados com o símbolo * (asterisco). Este símbolo mostra que os valores destacados possuem correlações significativas, isto é, nível de significância (p-valor) inferior a 0,05.

A correlação indica que, uma vez que o valor de uma variável aumenta, o valor da outra variável também aumenta. Assim, ao pegar a variável “Índice Global de Segurança Cibernética” e “Sucesso do Governo na Promoção das TIC”, nota-se que quando o índice de segurança cibernética aumenta, o sucesso do governo em promover TIC também cresce.

Tabela 1 – Correlações entre variáveis

		Índice Global de Segurança Cibernética	Sucesso do governo na promoção das TIC	Contratação de tecnologia avançada pelo governo
Índice Global de Segurança Cibernética	Correlação de Pearson	1	,291*	,375**
	Sig. (2 extremidades)		,036	,006
	N	52	52	52
Sucesso do governo na promoção das TIC	Correlação de Pearson	,291*	1	,295*
	Sig. (2 extremidades)	,036		,034
	N	52	52	52
Contratação de tecnologia avançada pelo governo	Correlação de Pearson	,375**	,295*	1
	Sig. (2 extremidades)	,006	,034	
	N	52	52	52

*. A correlação é significativa no nível 0,05 (2 extremidades).

**. A correlação é significativa no nível 0,01 (2 extremidades).

Além disso, as variáveis “Sucesso do Governo na Promoção das TIC” e “Contratação de Tecnologia Avançada Pelo Governo” mostram a correlação significativa, quando analisada a tabela 01. Isso significa que, quando o valor de uma variável sube, o valor da outra variável também cresce. Esse conceito também é válido ao comparar as variáveis “Índice Global de Segurança Cibernética” e “Contratação de Tecnologia Avançada Pelo Governo”. Essas correlações indicam que a participação governamental é relevante para o índice de transformação digital nos países.

5. CONSIDERAÇÕES FINAIS

Em geral, demonstra-se que há uma relação entre as variáveis. Os resultados também demonstram que há variáveis mais correlacionadas entre si do que outras. As

representações gráficas existentes demonstram os resultados obtidos e podem auxiliar em tomadas de decisões.

O trabalho desenvolvido trouxe uma melhor compreensão das etapas de um projeto de análise de dados, ao propor a avaliação de tendências e comparação entre as variáveis envolvidas. Permitiu-se também o melhor entendimento de como os conceitos necessários podem ser combinados para a definição de resultados, ao empregar a estatística com a programação lógica de etapas, como algoritmos e regras de negócios.

Todavia, ao concluir o artigo científico, percebe-se a necessidade de um estudo mais contínuo e prolongado dos fatores estatísticos até então estudados, por serem conceitos novos e mais sofisticados do que os assuntos de níveis educacionais antecedentes. Em virtude do proposto, sugere-se a continuação de aplicação de uma linguagem de programação, como a linguagem Python, uma vez que isso torna possível uma análise mais direcionada e aprofundada dos dados. Sugere-se também o estudo da relação entre a análise de dados com outros ramos da tecnologia da informação e da sociedade em geral, ao ser possível a obtenção de informações por diferentes meios, como um sistema web, além de ser possível trazer os resultados para diferentes sistemas já existentes.

REFERÊNCIAS

ABDI – AGÊNCIA BRASILEIRA DE DESENVOLVIMENTO INDUSTRIAL. **Índice de Transformação Digital**. Disponível em: <https://itd.abdi.com.br/>. Acesso em: out. 2023.

CARVALHO, G. D. G. **Proposta de um modelo para a análise das relações entre a capacidade de gestão e a capacidade de inovação de micro e pequenas empresas do sul do Brasil**. 2021. Tese (Doutorado em Engenharia de Produção). Universidade Tecnológica Federal do Paraná. Ponta Grossa, 2021. Disponível em: <https://repositorio.utfpr.edu.br/jspui/handle/1/25063> . Acesso em: dez. 2023.

CARVALHO, H. G.; DIOGO, T. M.; PIMENTEL; R. CARVALHO, G. D. G. **Gestão da Inovação em Cooperativas: um caminho para Inovar**. Curitiba: ISAE, 2022.

CITS – CENTRO INTERNACIONAL DE TECNOLOGIA DE SOFTWARE. **Pesquisa e Desenvolvimento**. Disponível em: cits.br. Acesso em: 02 dez. 2023

DA SILVA, Douglas. **Conheça dos quatro tipos de análise de dados para criar estratégias certeiras**, 2022. Disponível em: < <https://www.zendesk.com.br/blog/tipos-analise-de-dados/>>

Acesso em: 23 de set. 2023

FIELD, Andy. **Descobrimo a Estatística usando o SPSS**. Sussex: Sage Publications of London, Thousand Oaks and New Delhi, 2005.

LOPES, Vanessa. **Como calcular a porcentagem acumulada de elementos**, 2017. Disponível em: < <https://educacao.umcomo.com.br/artigo/como-calculer-a-percentagem-acumulada-de-elementos-1970.html> >. Acesso em: 15 de out. 2023

MACHADO, A. C. M.; SCHROEDER, R. T.; CARVALHO, G. D. G. Transformação digital 4.0 em Cooperativas. In: CARVALHO, H. G.; DIOGO, T. M.; PIMENTEL; R. CARVALHO, G. D. G. (orgs.) **Gestão da Inovação em Cooperativas: um caminho para Inovar**. Curitiba: ISAE, 2022.

OLIVEIRA, Bruno. **Teste de Kruskal-Wallis e Nemenyi**, 2020. Disponível em: <<https://statplace.com.br/blog/teste-de-kruskal-wallis-e-o-teste-de-nemenyi/>>. Acesso em: 15 de out. 2023

SANTOS, Juliana Moreira et al. **Empreendedorismo digital por mulheres: uma revisão integrativa da literatura**. Revista Gestão e Desenvolvimento, v. 20, n. 2, p. 150-175, 2023.

SIQUEIRA, Daniel. **Histograma: O que é, Exemplos, Gráficos e Tipos**, 2023. Disponível em: <<https://www.alura.com.br/artigos/o-que-e-um-histograma#:~:text=O%20objetivo%20de%20um%20histograma,em%20torno%20deste%20valor%20central.>>. Acesso em: 01 de out. 2023

TAYRANE. **Metodologia de análise de dados: um guia completo sobre o tema**, 2022. Disponível em: < <https://blog.ploomes.com/analise-de-dados/> >. Acesso em: 23 de set. 2023

ZELLER, Violet; HOCKEN, Christian; STICH, Volker. Acatech Industrie 4.0 maturity index – a multidimensional maturity model. In: **Advances in Production Management Systems**. Smart Manufacturing for Industry 4.0: IFIP WG 5.7 International Conference, APMS 2018, Seoul, Korea, August 26-30, 2018, Proceedings, Part II. Springer International Publishing, 2018. p. 105-113.