

# Estilos de Uso dos Espaços Virtuais de Aprendizagem: Diferentes Perfis e Formatos nas Gerações Digitais

Análise Exploratória para a Tese EUEV e Gerações

Priscilla Aparecida Sanatana Bittencourt

João Pedro Albino

29/07/2022

## 3 PROCEDIMENTOS METODOLÓGICOS

De maneira mais sucinta, a análise dos materiais realizou-se conforme os tópicos:

1º - Aplicação do RACI nos materiais levantados na revisão sistemática da literatura;

2º - Baseado nas respostas positivas e que foram obtidas do resultado da aplicação do roteiro RACI para com os materiais que foram levantados na revisão sistemática, mais especificamente na coluna 5 (a observação e descrição realizada dos aspectos observados nas investigações)

3º - Neste tópico serão demonstrados os procedimentos metodológicos para a coleta, análise e interpretação dos dados relativos à representação detalhada dos diferentes perfis e usos do digital nas gerações de acordo com os apontamentos identificados na revisão sistemática da literatura relativas ao questionário sobre os estilos de uso do espaço virtual.

## 4 RESULTADOS OBTIDOS E DISCUSSÃO DA INVESTIGAÇÃO

Neste Capítulo são apresentados os resultados desta investigação, bem como os resultados e a discussão das análises realizadas com dados secundários referentes aos estilos de uso do espaço virtual. Os dados primários foram coletados em trabalhos identificados na pesquisa bibliográfica exploratória e compartilhados mediante solicitação (Anexo 3) e autorizados para utilização neste trabalho.

Este Capítulo está subdividido em três partes:

- **primeira parte:** apresenta os resultados referentes a revisão sistemática da literatura;

- **segunda parte:** apresenta os procedimentos de coleta e análise exploratória de dados à luz do que foi identificado em relação a revisão sistemática da literatura concernentes aos estilos de uso do espaço virtual e relativas à representação detalhada dos diferentes perfis e usos do digital nas gerações;

- **terceira parte:** apresenta os resultados e a discussão das análises dos dados, em consonância com o objetivo principal da investigação.

### 4.1 Primeira parte: resultados da revisão sistemática da literatura

A apresentação dos dados coletados nos repositórios durante a leitura inicial foi realizada por meio de tabelas, com a sugestão do tema em questão de cada investigação e a quantidade de estudos com a palavra-chave da mesma temática.

Posteriormente, houve uma escolha mais criteriosa e objetiva dos materiais encontrados para cada tema proposto nesta investigação.

A exploração e análise dos materiais encontrados se deu por meio de leituras mais aprofundadas, nas quais se possibilitou a discussão e a confrontação do texto teórico descrito no Capítulo 2 - Fundamentação Teórica. Para a construção da base teórica do item 3.1, intitulado como: “A influência da tecnologia na vida das

peças e na sociedade”, utilizou-se a investigação exploratória e bibliográfica, considerando os seguintes conceitos-chave: “o impacto das tecnologias na vida das pessoas” e “importance of technologies in people’s lives”.

Limitou-se o período específico entre 2018 e 2021 e em diferentes idiomas, pois era importante encontrar dados mais recentes publicados sobre este tema, a fim de selecionar o material mais recente.

O levantamento bibliográfico foi realizado em bases de dados Brasileiras e Portuguesas:

- Google acadêmico de Portugal, acesso pelo link: <https://scholar.google.pt/>;
- Biblioteca do Conhecimento Online (b-on), acesso pelo link: <https://www.b-on.pt/>;
- Repositório aberto da UAb, acesso pelo link: <https://repositorioaberto.uab.pt/>;
- Repositórios Científicos de Acesso Aberto de Portugal (RCAAP), acesso pelo link: <https://www.rcaap.pt/>;
- Repositórios das UNESP, acesso pelo link: <https://repositorio.unesp.br/>;
- Google acadêmico do Brasil, acesso pelo link: <https://scholar.google.com.br/>;
- Scielo, acesso pelo link: <https://scielo.org/>

A partir da pesquisa bibliográfica, foram obtidos os resultados demonstrados no Quadro 3:

Quadro 3 : Resultado do levantamento Bibliográfico (2018-2021)

Base de dados	Palavra-chave 1 <sup>1</sup>	Palavra-chave 2 <sup>2</sup>
<a href="https://scholar.google.pt/">https://scholar.google.pt/</a>	32.800	16.600
<a href="https://www.b-on.pt/">https://www.b-on.pt/</a>	6.011	212.992
<a href="https://repositorioaberto.uab.pt/">https://repositorioaberto.uab.pt/</a>	2.312	1.371
<a href="https://www.rcaap.pt/">https://www.rcaap.pt/</a>	486	97
<a href="https://repositorio.unesp.br/">https://repositorio.unesp.br/</a>	42.993	16.270
<a href="https://scholar.google.com.br/">https://scholar.google.com.br/</a>	17.300	16.500
<a href="https://scielo.org/">https://scielo.org/</a>	4	0
Totais	101.906	263.830

Fonte: Elaborado pela pesquisadora (2022).

Como pode ser aferido pelo Quadro 3, a pesquisa exploratória realizada com as quatro palavras-chave selecionadas também resultou em um número considerável de artigos, dissertações, teses e livros, similar ao que pôde ser constatado nas buscas referentes ao Quadro 2.

No Anexo 3 pode-se observar o processo utilizado de busca para selecionar quais documentos incluir nos resultados da pesquisa.

Pode-se também observar no Quadro 3 que os resultados iniciais da busca foram muito amplos, mesmo restringindo-se o conteúdo das investigações por meio das palavras-chave e a limitação dos anos. Dentre os resultados obtidos no levantamento bibliográfico, os tipos de fontes que aparecem como resultados foram bastante relevantes e expressivos, constando-se de Revistas Acadêmicas, Resenhas, Relatórios, Publicações da Especialidade, entre outros.

Pelo conjunto de documentos obtidos na busca, inferiu-se que um número reduzido de conteúdo bibliográfico e documental alcançado seriam efetivamente pertinentes a esta investigação. Desta forma, de acordo com o roteiro RACI e por meio de leitura parcimoniosa dos resumos das publicações, selecionou-se de forma criteriosa os materiais que constam nas referências bibliográficas desta pesquisa.

Portanto, para fechar o foco no tema desta pesquisa, foi necessária uma seleção criteriosa do material levantado, utilizando-se para tanto os recursos do roteiro de análise de conteúdo para investigação - RACI (Tabela ???), além de leitura aprofundada do material bibliográfico. Desta forma, foram selecionados os seguintes autores como principais bases para o tema desta tese: Bates (2017), Martino (2014), Carr (2011), Kerckhove (2009), Castells (2004) e Carvalho (2006).

Para a construção da base teórica referentes a história da internet e das tecnologias de informação e comunicação, utilizou-se a investigação exploratória e bibliográfica, considerando os seguintes conceitos-chave: “história da internet”, “história das tecnologias”, “the history of the internet” e “history of technologies”.

<sup>1</sup>Palavra-chave 1: a importância das tecnologia na vida das pessoas

<sup>2</sup>Palavra-chave 2: importance of technologies in people’s lives

Limitou-se o período de buscas entre os anos de 2018 e 2021, em diferentes idiomas, buscando selecionar os conteúdos mais recentes publicados sobre o tema.

Esse levantamento bibliográfico também foi realizado em bases de dados Brasileiras e Portuguesas, tais como:

- Google acadêmico de Portugal, acesso pelo link: <https://scholar.google.pt/>;
- Biblioteca do Conhecimento Online (b-on), acesso pelo link: <https://www.b-on.pt/>;
- Repositório da Universidade Aberta (UAb) de Portugal, acesso pelo link: <https://repositorioaberto.uab.pt/>;
- Repositórios Científicos de Acesso Aberto de Portugal (RCAAP), acesso pelo link: <https://www.rcaap.pt/>;
- Repositórios das UNESP, acesso pelo link: <https://repositorio.unesp.br/>;
- Google acadêmico do Brasil, acesso pelo link: <https://scholar.google.com.br/>;
- Biblioteca Eletrônica Científica Online (SciELO), acesso pelo link: <https://scielo.org/>. A partir do levantamento bibliográfico, foram obtidos os resultados demonstrados no Quadro 4:

Quadro 4: Resultado do levantamento Bibliográfico (2018-2021)

Base de dados	Palavra-chave 1 <sup>3</sup>	Palavra-chave 2 <sup>4</sup>	Palavra-chave 3 <sup>5</sup>	Palavra-chave 4 <sup>6</sup>
<a href="https://scholar.google.pt/">https://scholar.google.pt/</a>	380	1.270	17.300	355
<a href="https://www.b-on.pt/">https://www.b-on.pt/</a>	0	0	0	0
<a href="https://repositorioaberto.uab.pt/">https://repositorioaberto.uab.pt/</a>	2.311	1.155	1.935	1.222
<a href="https://www.rcaap.pt/">https://www.rcaap.pt/</a>	476	405	621	524
<a href="https://repositorio.unesp.br/">https://repositorio.unesp.br/</a>	22.126	1.145	7.909	4.248
<a href="https://scholar.google.com.br/">https://scholar.google.com.br/</a>	17.000	1.280	16.700	356
<a href="https://scielo.org/">https://scielo.org/</a>	36	32	131	78
Totais	42.329	5.287	44.596	6.783

Fonte: Elaborado pela pesquisadora (2022).

Como pode ser aferido pelo Quadro 4, a pesquisa exploratória realizada com as quatro palavras-chave selecionadas também resultou em um número considerável de artigos, dissertações, teses e livros, similar ao que pôde ser constatado nas buscas referentes ao Quadro 3.

Dentre o total de resultados apresentados, destacam-se, no Quadro 4, os diversos tipos de fontes que aparecem como resultados, dentre eles: revistas acadêmicas, resenhas, relatórios, publicações da especialidade, entre outros.

Pelo conjunto de documentos obtidos na busca, inferiu-se que um número reduzido de conteúdo bibliográfico e documental alcançado seriam efetivamente pertinentes a esta investigação. Desta forma, de acordo com o roteiro RACI e por meio de leitura parcimoniosa dos resumos das publicações, selecionou-se de forma criteriosa os materiais que constam nas referências bibliográficas desta pesquisa.

Para a construção da fundamentação teórica do item 2.2, intitulado como: “os estilos de uso do espaço virtual”, utilizou-se o roteiro RACI para a investigação exploratória e bibliográfica, além de buscas em bases de dados disponíveis online utilizando-se das palavras-chave: “estilos de uso do espaço virtual”; “questionário estilos de uso do espaço virtual” e “uso do espaço virtual”. Restringiu-se o período específico de exploração nas bases de dados de publicações científicas entre os anos de 2008 e 2020 e em diferentes idiomas, pois este foi o período em que o questionário “Estilos de uso do espaço virtual” foi desenvolvido e aplicado por

<sup>3</sup>Palavra-chave 1 - A História da internet

<sup>4</sup>Palavra-chave 2 - the history of the internet

<sup>5</sup>Palavra-chave 3 - história das tecnologias

<sup>6</sup>Palavra-chave 4 - history of technologies

diferentes investigadores em diferentes países.

Parte desta etapa foi realizada em Portugal, durante o ano de 2018, período em que a pesquisadora realizou o estágio doutoral na Universidade Aberta (UAb) de Lisboa. Após este período imersivo da pesquisa, a pesquisa bibliográfica e documental continuou quando do retorno ao Brasil, estendendo-se até o ano de 2020. O resumo dos resultados desta busca exploratória e bibliográfica pode ser visualizado no Anexo 1. Foram selecionados trinta e nove resultados encontrados com base nesta busca específica, conforme demonstrado no Quadro 5. Parte deste resultado foi utilizado para compor as referências desta investigação, além de subsidiar o referencial teórico.

Quadro 5: Resultado do levantamento Bibliográfico (2008-2020)

Palavra-chave	Biblioteca do Conhecimento Online	Google acadêmico Portugal
estilos de uso del espacio virtual	2	10
estilos de uso do espaço virtual	6	15
questionário estilos de uso do espaço virtual	1	5
<b>Total</b>	<b>9</b>	<b>30</b>

Fonte: Elaborado pela pesquisadora (2022).

Após realizada a pesquisa exploratória com as palavras-chave: “estilos de uso del espacio virtual”, “estilos de uso do espaço virtual” e “questionário estilos de uso do espaço virtual” obteve-se um resultado relevante, mesmo limitando-se o período de pesquisa entre 2008 a 2020. Foram alcançados o total de 39 documentos, compreendendo-se artigos, dissertações, teses e livros. Os resultados da pesquisa nas bases de dados Biblioteca do Conhecimento Online (B-on) e Google acadêmico, ambos de Portugal, está demonstrado no Quadro 5.

A partir dessa busca exploratória e bibliográfica, foi realizada uma análise documental para a seleção do material com o auxílio do RACI, o roteiro de análise de conteúdo para investigação (Tabela ???), desenvolvido com objetivo de selecionar o material pertinente para esta tese. A escolha dos documentos baseou-se em critérios de análise nas dimensões de exploração, análise, e seleção para a identificação dos temas relacionados aos estilos de uso do espaço virtual, componentes do roteiro de análise de conteúdo. Além do RACI, a seleção também foi fundamentada em critérios baseados em estudos e pesquisas realizados anteriormente ao desenvolvimento deste trabalho pela pesquisadora.

Em síntese, o estudo bibliográfico e exploratório identificou o material utilizado para subsidiar inicialmente a investigação, que resultou nos documentos relacionados no Quadro 6:

Quadro 6: Estudo bibliográfico e exploratório no período de 2008 a 2020

Tipo de documento	Quantidade	Autores
Artigos	3	TERÇARIOL, A. A. L.; BARROS, D. M. V. (2017); MIRANDA, L.; MORAIS, C.; GOULÃO, F.; BARROS, D. M.V.; (2012); ARANDA, P. C. G.; POLANCO, P. O. P.; HERRERA, P. J. C. (2017).
Dissertações	3	FREITAS, J. M. A. (2013); MENDES, A. G. L. M. (2015); SILVA, A. C. (2011).

Tipo de documento	Quantidade	Autores
Teses	4	LESSA, V. E. (2018); GOMES, F. D. (2018); FERNANDES, F. (2018); BARRERA, A. V.; GAYTAN, B. G. R.; MANCILLA, M. A. A. (2016).

Fonte: Elaborado pela pesquisadora (2022).

Em linhas gerais, os artigos apresentaram estudos exploratórios, dissertando sobre o objetivo de identificar os estilos de uso do espaço virtual predominantemente em grupos de estudantes e professores, e os resultados desses estudos salientam a importância de se identificar tais estilos com o propósito de desenvolver estratégias para o uso colaborativo e criativo dos ambientes online neste grupo de usuários do espaço virtual. Os estudos exploratórios contidos nos artigos também revelaram uma breve reflexão com relação ao ensino presencial e à distância, assinalando que o uso das mídias sociais pode ampliar as oportunidades de aprendizagem individual e colaborativa. As dissertações expuseram, de forma geral, que os estilos de aprendizagem no virtual podem se modificar ou se adaptar de acordo com o contexto ao qual os indivíduos estão inseridos, influenciando no processo de ensino e aprendizagem dos estudantes, além de identificar o estilo de uso do espaço virtual predominante em cada investigação realizada.

As teses analisadas tratam de estudos realizados com estudantes, nas quais se pode perceber que, de maneira geral, todas observaram o processo de ensino e aprendizagem com o objetivo de entenderem os estudantes em seu contexto. Os resultados das teses analisadas mostraram maneiras de colaborar e incentivar o desenvolvimento de novos modelos de ensinar e aprender.

O levantamento bibliográfico e exploratório também foi importante no momento para identificar quais investigadores aplicaram e utilizaram a teoria dos “estilos de uso do espaço virtual” em suas investigações.

Posteriormente ao processo de consulta aos bancos de dados referente a temática dos estilos de uso do espaço virtual, houve a necessidade de atualização das obras consultadas durante a fase inicial da pesquisa, portanto, fez-se necessária nova pesquisa bibliográfica abrangendo o tema.

Realizou-se nova investigação exploratória e bibliográfica, consultando os seguintes conceitos-chave: “estilos de uso do espaço virtual”; “questionário estilos de uso do espaço virtual” e “uso do espaço virtual”. Ampliou-se o período de abrangência das buscas, compreendendo os anos entre 2018 e 2021 e expandiu-se também o número das bases de dados consultadas, pois pretendia-se visualizar as atualizações destes conceitos neste período.

A partir desta pesquisa atualizada, obteve-se os resultados, demonstrados no Quadro 7:

Quadro 7: Resultado do levantamento Bibliográfico (2018-2021)

Base de dados	Palavra-chave 1 <sup>7</sup>	Palavra-chave 2 <sup>8</sup>	Palavra-chave 3 <sup>9</sup>
https://scholar.google.pt/	77	18	3
https://www.b-on.pt/	0	0	0
https://repositorioaberto.uab.pt/	15	2	1
https://www.rcaap.pt/	2	1	0
https://repositorio.unesp.br/	41.141	42.024	0
https://scholar.google.com.br/	77	18	3
https://scielo.org/	0	0	0
Totais	41.582	5.287	42.063

<sup>7</sup>Palavra-chave 1 - estilos de uso do espaço virtual

Fonte: Elaborado pela pesquisadora (2022).

Como pode ser observado no Quadro 7, utilizando-se um período ampliado para as buscas nas bases de dados, a quantidade de documentos recuperados superou, em muito, os observados no Quadro 6, demonstrando que o tema “estilos de uso do espaço virtual” e correlatos encontra-se bastante frequente. Com tais resultados atualizados, observou-se também a necessidade de atualização e ampliação das buscas referentes aos conceitos-chave inerentes à questão de pesquisa desta tese, compreendendo então o conjunto de palavras-chave “espaço virtual” e “gerações digitais”. Para efetivar o levantamento bibliográfico e exploratório inerente ao tema deste trabalho, determinou-se que o período de buscas contivesse os anos entre 2018 (último ano da pesquisa bibliográfica anterior) e 2021 nas mesmas bases de dados já especificadas anteriormente.

A partir desta atualização e ampliação na pesquisa original realizada, uma ampla variedade de documentos foi recuperada, cujos resultados condensados estão demonstrados no Quadro 8. Uma visão mais detalhada dos resultados obtidos também pode ser visualizada no Anexo XXX.

Quadro 8: Resultado ampliado do levantamento Bibliográfico

Base de dados	Palavra-chave 1 <sup>10</sup>	Palavra-chave 2 <sup>11</sup>
<a href="https://scholar.google.pt/">https://scholar.google.pt/</a>	18.000	18.600
<a href="https://www.b-on.pt/">https://www.b-on.pt/</a>	0	0
<a href="https://repositorioaberto.uab.pt/">https://repositorioaberto.uab.pt/</a>	1.407	18.600
<a href="https://www.rcaap.pt/">https://www.rcaap.pt/</a>	11	3
<a href="https://repositorio.unesp.br/">https://repositorio.unesp.br/</a>	2.119	535
<a href="https://scholar.google.com.br/">https://scholar.google.com.br/</a>	12.600	129
<a href="https://scielo.org/">https://scielo.org/</a>	42	2
Totais	41.582	5.287

Fonte: Elaborado pela pesquisadora (2022).

A partir dessa extensa busca exploratória e bibliográfica e, novamente, com o auxílio do roteiro RACI (Tabela ???), efetuou-se uma análise documental para escolha de material relacionado à esta tese, considerando-se temas relacionados aos estilos de uso do espaço virtual e às gerações digitais. Além dos critérios de análise nas dimensões de análise de conteúdo do RACI, a escolha também foi abalizada em juízo crítico baseados em estudos e pesquisas realizados anteriormente pela pesquisadora, notadamente quando do desenvolvimento da dissertação (BITTENCOURT, 2016) e artigos BITTENCOURT e ALBINO, 2017).

Em linhas gerais, os artigos apresentaram... dissertando sobre o objetivo de .....

Com relação à questão do uso geracional da tecnologia ....

O levantamento bibliográfico e exploratório também foi importante para identificar quais documentos recuperados na busca nas bases de dados utilizaram o questionário de estilos de uso em seus trabalhos.

## 4.2 Segunda parte: procedimentos de coleta e análise exploratória de dados

Neste tópico serão demonstrados os procedimentos metodológicos para a coleta, análise e interpretação dos dados que foram compartilhados à esta tese por pesquisadores identificados, no processo de revisão sistemática da literatura, relativos ao uso do Questionário Estilo de Uso do Espaço Virtual – CEUEVE.

**4.2.1 Obtenção dos dados não submetidos a processamento** Os dados secundários utilizados nesta tese foram compartilhados por pesquisadores identificados no processo de revisão sistemática da literatura, relativos ao uso do Questionário Estilo de Uso do Espaço Virtual – CEUEVE. Tal instrumento de pesquisa, desenvolvido em Barros (2018) e especificado em Barros (2013), versa sobre identificar como as pessoas

<sup>8</sup>Palavra-chave 2 - estilos de uso del espacio virtual

<sup>9</sup>Palavra-chave 3 - questionário estilos de uso do espaço virtual

<sup>10</sup>Palavra-chave 1 - espaço virtual

<sup>11</sup>Palavra-chave 2 - gerações digitais

utilizam e aprendem no espaço virtual e quais estilos de aprendizagem podem ser considerados neste novo espaço.

Como citado anteriormente, Barros (2009) categorizou quatro tendências de uso dos estilos do espaço virtual, já discutidas na Fundamentação Teórica e que nesta etapa da tese serão utilizados com a terminologia:

- **Estilo de Uso A:** Uso Participativo no Espaço Virtual
- **Estilo de Uso B:** Busca e Pesquisa no Espaço Virtual
- **Estilo de Uso C:** Estruturação e Planejamento no Espaço Virtual
- **Estilo de Uso D:** Ação Concreta e Produção no Espaço Virtual

Em síntese, em Barros (2013), o estilo de uso do espaço virtual é definido como situações de utilização dos aplicativos e ferramentas determinam as características de um usuário na busca, no planejamento e na utilização de informação no digital.

Outro construto importante também tratado na Fundamentação Teórica em como as tecnologias digitais evoluem a cada dia e as pessoas, independentemente da nomenclatura utilizada para nomeá-las (nativos, imigrantes, habitantes, visitantes ou transeuntes), suas usuárias, estão adotando, cada grupo de indivíduos com suas características, mostrou-se interessante pesquisar as propensões de uso dos estilos do espaço virtual e as formas de utilização da tecnologia pelos conjunto de pessoas (gerações) e, assim, possivelmente compreender de que forma o uso dos espaços virtuais por pessoas em diferentes faixas etárias, impacta o seu cotidiano.

De forma concisa, a análise e interpretação dos dados dos questionários CEUEVE e as características geracionais relativas distinguiriam uma representação detalhada dos diferentes perfis e usos do digital nas gerações de acordo com os apontamentos identificados, convergindo diretamente à resposta da pergunta de pesquisa desta tese.

Portanto, para o andamento desta investigação tornou-se necessário coletar dados referentes ao questionário CEUEVE (Barros, 2013), junto a pesquisadores do tema identificados na revisão sistemática da literatura relativas a aplicação deste instrumento de pesquisa nos seus trabalhos. Após a identificação dos pesquisadores, foi necessário entrar em contato com os pesquisadores do tema. Uma carta, cujo modelo é mostrado no Anexo 2, foi enviada a cada um dos autores dos artigos, dissertações e teses selecionados, com o intuito de solicitar o compartilhamento dos dados, obtidos inteiramente dos questionários ou em formato de planilha de dados, referentes ao uso do CEUEVE, coletados em suas respectivas investigações. Porém, apesar da pluralidade de trabalhos recuperados na pesquisa aos bancos de dados e da tentativa de contato com os autores das pesquisas, houve morosidade para o retorno das mensagens enviadas eletronicamente e, após este tempo decorrido poucos pesquisadores responderam ao contato e compartilharam seus dados primários, resultando assim na primeira dificuldade encontrada para acesso aos dados obtidos diretamente dos questionários e não submetidos a processamento, definidos como dados brutos (HAI, KAMBER E PEI, 2011; WICKHAM E GROLEMUND, 2017). Desta forma, as fontes obtidas e que serão utilizadas nesta tese estão sumarizadas no Quadro nnn.

Quadro nnn: Fontes de dados utilizadas

Pesquisadores	País	Tipo de documento	Título	Ano de publicação
María Carmen López Berlanga et. al	Espanha	Artigo	El estilo de uso del espacio virtual con estudiantes de Educación Secundaria	2018
Alejandro Villegas Barrera	México	Monografia	Herramienta para la creación de un Entorno Personal de Aprendizaje	2017
Marcos Andrei Ota	Brasil	Tese	Adaptatividade em Ambientes Virtuais: uma proposta para personalizar a aprendizagem em cursos híbridos de ensino superior	2018

Pesquisadores	País	Tipo de documento	Título	Ano de publicação
Adriana Aparecida de Lima Terçariol et. al	Brasil	Artigo	Os Estilos de Uso dos Espaços Virtuais e as Redes Sociais na Pedagogia: Um Estudo Exploratório	2017
Daniela Melaré Vieira Barros	Brasil			2010 - 2019
Daniela Melaré Vieira Barros	Portugal			2018 - 2019
Cristina Sánchez Romero et. al	Espanha	Artigo	Estudio de Identificación de Los Estilos Del Uso Del Virtual de Los Seniors: perspectivas iniciales	2020

Fonte: Elaborado pela pesquisadora (2022).

**4.2.2 Processo de Mineração e análise dos dados** De acordo com Hai, Kamber e Pei (2011), mineração de dados (ou em inglês data mining) é o processo de extração e descoberta de padrões em conjuntos de dados (datasets) envolvendo métodos na interseção entre aprendizado de máquina, estatística e sistemas de banco de dados. A mineração, um subcampo interdisciplinar de ciência da computação e estatística, tem como objetivo geral extrair informações de um conjunto de dados e transformar as informações em uma estrutura compreensível para uso posterior. A mineração de dados é a etapa de análise do processo de “descoberta de conhecimento em bancos de dados”, ou **KDD** (Knowledge Discovery in Databases).

A descoberta de conhecimento em bases de dados tem como objetivo encontrar padrões intrínsecos aos dados nela contidos, apresentando-os de forma a facilitar sua assimilação como conhecimento, definem Silva, Peres e Boscarioli (2016). Ainda segundo esses autores, a mineração de dados é definida em termos do esforço para a descoberta de padrões e, a partir dos padrões descobertos, gerar conhecimento útil para um processo de tomada de decisão. Trata-se, portanto, da aplicação de técnicas, implementadas por meio de algoritmos computacionais, capazes de receber, como entrada, um conjunto de dados e devolver, como saída, um padrão de comportamento, o qual pode ser expresso, por exemplo, uma função de mapeamento, ou a modelagem de um perfil (SILVA, PERES e BOSCARIOLI, 2016).

Os processos fundamentais da mineração de dados a partir de fontes de dados (bancos de dados, relatórios, logs de acesso, transações etc.) consistem em: (a) limpeza (remoção de ruídos e redundâncias, remoção de duplicidades etc.); (b) preparação dos dados (preenchimento de informações, detecção de anomalias (outliers), dados faltantes etc.); (c) modelagem dos dados; e (d) implantação da solução. (HAI, KAMBE e PEI, 2011).

**4.2.3 Análise exploratória dos dados** Trabalhos recentes na área de ciência de dados, tais como Tyagi (2022), Kotu (2019) e Wickham e Grolemund (2017) consideram a prospecção de dados como pertencente ao campo da ciência de dados, no qual o processo de mineração abrange as técnicas, ferramentas e métodos relativos ao processo de obter insights valiosos a partir de dados estruturados e não estruturados. Muitas vezes, os termos mais genéricos e utilizados em larga escala, *analítica* - do termo em inglês *analytics* - e *análise de dados* - ou, quando se refere a métodos concretos, *inteligência artificial* e *aprendizado de máquina*, são mais apropriados, afirma Tyagi (2022). Por tanto, nesta tese, os termos mineração e análise de dados serão utilizados de forma alternada e com o mesmo objetivo: processo de obtenção de valiosos **insights** a partir da análise e interpretação dos dados. Um *insight* é um acontecimento cognitivo que pode ser associado a capacidade de ter uma compreensão clara, profunda e às vezes repentina de um problema ou situação complicada podendo ser sinônimo de compreensão, conhecimento, intuição<sup>12</sup>. *Insight* também pode

<sup>12</sup>Conforme definição em Significados. Porto: 7Graus, 2020. Disponível em: <https://www.significados.com.br/insight/#:~:text=O%20que%20significa%20Insight%3A&text=Um%20insight%20é%20um%20acontecimento,solução%20surge%20de%20forma%20repentina>. Acesso em: 26 jun. 2022.



ser definido como a perspicácia ou a capacidade de apreender alguma coisa e acontece quando uma solução surge de forma repentina<sup>13</sup>.

Para a análise e interpretação dos dados coletados pelos pesquisadores será utilizada neste tópico da tese o *framework* - termo inglês que, em tradução direta, significa estrutura ou arcabouço, em português - definido em Wickham e Grolemund (2017) e representado na Figura NNN.

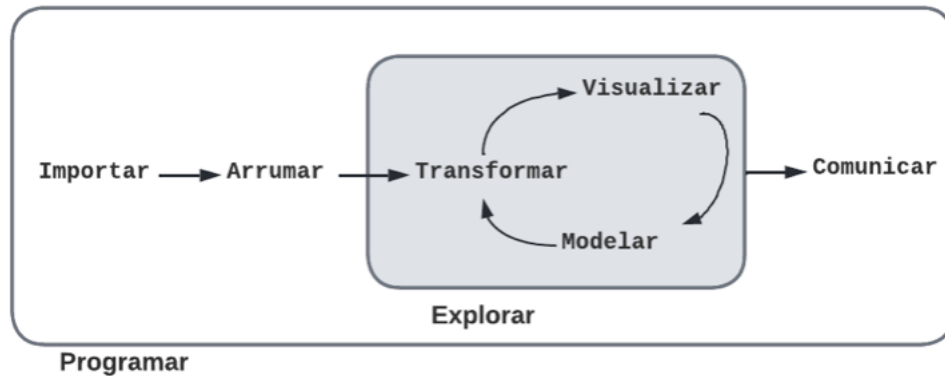


Figure 1: Figura NNN. Estrutura para análise e interpretação de dados.

Fonte: Baseado em Wickham e Grolemund, 2017.

De acordo com Kotu (2019), a grande maioria dos frameworks em ciência de dados apresentam como características um arcabouço genérico muito semelhante ao processo *CRISP-DM*, abreviação de *Cross Industry Standard Process for Data Mining*, que pode ser traduzido como Processo Padrão Inter-Indústrias para Mineração de Dados.

Ainda de acordo com o autor, o CRISP-DM é um framework de processo de mineração de dados que descreve abordagens comumente usadas por especialistas em mineração de dados para resolver um problema específico.

Como em qualquer estrutura de processo, um processo de ciência de dados recomenda a execução de um determinado conjunto de tarefas para obter a saída ideal.

O processo de extração de informações e conhecimento dos dados é iterativo, ou seja, é referente a um processo que é feito de novo, repetido ou reiterado, que é realizado inúmeras vezes<sup>14</sup>. As etapas dentro do processo de ciência de dados não são lineares e precisam passar por muitos loops<sup>15</sup>, ir e voltar entre as etapas e, às vezes, voltar à primeira etapa para redefinir a declaração do problema da ciência de dados.

O arcabouço de ciência de dados apresentada na Figura NNN é um conjunto genérico de etapas que é agnóstico de problema, algoritmo e ferramenta para análise de dados. O objetivo fundamental de qualquer processo que envolva a apreciação e a interpretação de dados é abordar a questão da análise, afirma Kotu (2019).

Em resumo, as etapas do processo da Figura NNN podem ser explicitadas como sendo compostas por:

1. **Importar** os dados de alguma fonte (Banco de dados, ou planilha) para o ambiente de ciência de dados;
2. **“Arrumar”** (no original, em inglês *tidy*) ou **organizar/pré-processar** os dados no formato de *dataframe* no ambiente de ciência de dados limpando os dados faltantes (*missing data*), campos digitados errados etc. Um *dataframe*, segundo Wickham e Grolemund (2017), são objetos de dados genéricos no ambiente de ciência de dados usados para armazenar dados tabulares. O *dataframe* pode ser visto como uma tabela em que cada linha corresponde a um registo ou uma observação e suas colunas têm nomes e podem conter diferentes tipos de dados. Em resumo, um *dataframe* é composto de três componentes principais: *dados, linhas e colunas*.

<sup>13</sup>Conforme definição em OxfordLanguages, 2022, Oxford University Press. Disponível em: <https://dictionary.cambridge.org/dictionary/english-portuguese/insight>. Acesso em: 01 jun. 2022.

<sup>14</sup>Conforme definido em LEXICO, Dicionário Online de Português. Porto: 7Graus, 2018. Disponível em: <https://www.lexico.pt/iterativo/>. Acesso em: 10/06/2022.

<sup>15</sup>Conjunto de instruções que um programa de computador percorre e repete um significativo número de vezes até que sejam alcançadas as condições desejadas. Conforme definido em OxfordLanguages, 2022, Oxford University Press. Disponível em: <https://languages.oup.com/google-dictionary-pt/>. Acesso em: 10 jun. 2022.

3. **Transformar** os dados, de forma que possam ser aplicadas as técnicas estatísticas. Em alguns casos, transformar os campos categóricos alfanuméricos (por exemplo sexo, ou faixa etária) em campos categóricos numéricos (exemplo: sexo feminino = 1, sexo masculino = 2, não respondeu = 3). Nesta fase também são desenvolvidos os códigos (programas) para a elaboração das análises.
4. **Visualizar**. Nesta etapa são gerados os gráficos e elaborado o relatório da análise.
5. **Modelar**. Elaboração de modelos com os dados após a análise exploratória dos dados. Um *modelo estatístico* é uma representação da realidade na qual definimos a relação entre variáveis para entender e prever o comportamento de um fenômeno.
6. **Comunicar**. Envio ou submissão do relatório final com as análises realizadas e as conclusões obtidas.

**4.2.3.1 Importação e preparação dos dados** Após o compartilhamento dos dados coletados pelos investigadores, houve necessidade de normalização e padronização dos mesmos para uso nas análises realizadas nesta tese. Tal processo é definido como limpeza, preparação e tratamento de dados, uma das etapas do processo de mineração de dados (Hai, Kamber e Pei, 2011), correspondendo a fase de pré-processamento (ou *arrumar*) da Figura NNN. Ainda na etapa inicial de pré-processamento dos dados brutos e com a finalidade de atingir os objetivos propostos e fundamentar, com a análise dos dados compartilhados, a resposta à pergunta de pesquisa desta tese, algumas colunas das planilhas originais foram selecionadas, logo, não foram utilizadas todas as variáveis (ou colunas), dos trabalhos apurados do Quadro nnn.

Assim, para esta investigação foram escolhidos os campos: nome do pesquisador; país onde foi aplicado o questionário; gênero e idade do respondente; e o estilo de uso do espaço virtual; variáveis estas que foram considerados relevantes para análise e interpretação.

De acordo com Wickham (2014), 80% do processo de análise de dados é consumido no processo de limpeza e preparação dos dados. Segundo o autor, a preparação de dados não é apenas uma primeira etapa, mas deve ser repetida muitas vezes ao longo da análise, à medida que novos problemas surgem ou novos dados são coletados e acrescentados. Apesar da quantidade de tempo que leva, ainda existem poucos trabalhos de pesquisa sobre como limpar bem os dados, processo também definido como *higienização* (da SILVA, 2021). Parte do desafio é a amplitude de atividades que a higienização de dados engloba atividades, tais como: verificação de valores discrepantes; análise das datas informadas; atribuição de valores ausentes; dentre outras define da Silva (2021). Para lidar com tal dificuldade, esta tese se concentrou, primeiramente em um aspecto importante da limpeza de dados, organizando os dados compartilhados e estruturando o conjunto de dados para facilitar a posterior análise.

Para tanto, a pesquisadora utilizou-se do apoio oferecido pelos recursos humanos e técnicos do Laboratório de Inteligência de Dados – LinDa, grupo multidisciplinar de pesquisa da UNESP-Bauru que tem como finalidade apoiar e desenvolver pesquisas relacionadas à “ciência de dados” e, em particular, estudar estruturas de dados e suas aplicações onde seja possível extrair “inteligência”.

Inicialmente, os dados recebidos foram compilados pela pesquisadora em um *workbook* utilizando o aplicativo de computador (software) Excel® da Microsoft®, constituído de diferentes tabelas formadas com todas as colunas e linhas das planilhas originais (*worksheets*) compartilhadas. Uma amostra do conteúdo das planilhas de trabalho em Excel® com os dados brutos, ou *dados não numericamente organizados*, é mostrada na Figura zzz.

[illegible]

Salvamento Automático										dados_brutos_transformados_Original																			
Página Inicial										Compartilhar										Comentários									
Calibri (Corpo)										Geral										Formato									
N I S										Formato										Formato									
F815										Estilo de Uso B - Busca e Pesquisa no Espaço Virtual.																			
A B C D E F G H I J K L M N O																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 50 a 60 anos.																													
16/11/2010 Brasil de 30 a 40 anos.																													
16/11/2010 Brasil de 20 a 30 anos.																													
16/11/2010 Brasil de 40 a 50 anos.																													
16/11/2010 Brasil de 30 a 40 anos.										</																			

na preparação dos dados recebidos, ainda utilizando os recursos do software Excel®. Segundo Wickham (2014), a preparação de dados é o processo de reunir, combinar, estruturar e organizar os dados brutos para que possam ser usados nos procedimentos de AED e também para os processos de visualização desses dados. Os componentes da preparação de dados incluem, de acordo com Silva (2021): o pré-processamento de dados; documentação dos procedimentos utilizados; criação dos dicionários de dados; limpeza; validação e transformação; e envolve também buscar nas fontes externas originais dos dados a complementação dos dados faltantes (missing data). Para finalizar esta etapa, em algumas das investigações recebidas a pesquisadora retornou o contato com os pesquisadores dos trabalhos selecionados para solicitar o envio dos dados faltantes bem como orientação e apoio no preenchimento desses campos, uma tarefa manual e criteriosa que demandou várias semanas. Após esta etapa inicial, os dados compartilhados foram compilados e geraram uma versão de trabalho do conjunto de dados em formato Excel®. Uma amostra do *dataset* pode ser visualizada na Figura aaa.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	pesquisador	data/ano	País	genero	idade	estilo												
2	pesquisador1	02/04/2018	Portugal	feminino		estilo de uso participativo												
3	pesquisador1	02/04/2018	Portugal	feminino		estilo de estruturação e planeamento												
4	pesquisador1	29/03/2018	Portugal	masculino		ESTILO DE USO BUSCA E PESQUISA												
5	pesquisador1	22/04/2018	Portugal	masculino		estilo de uso participativo												
6	pesquisador1	03/04/2018	Portugal	masculino		estilo de estruturação e planeamento												
7	pesquisador1	24/03/2018	Portugal	masculino		estilo de uso participativo												
8	pesquisador1	01/04/2018	Portugal	feminino		ESTILO DE USO BUSCA E PESQUISA												
9	pesquisador1	10/04/2018	Portugal	feminino		ESTILO DE USO BUSCA E PESQUISA												
10	pesquisador1	01/04/2018	Portugal	feminino		não identificado												
11	pesquisador2	11/11/2017	brasil	Masculino	35 a 44 anos													
12	pesquisador2	11/11/2017	brasil	Masculino	45 a 54 anos													
13	pesquisador2	11/11/2017	brasil	Masculino	35 a 44 anos													
14	pesquisador2	11/11/2017	brasil	Masculino	25 a 34 anos													
15	pesquisador2	11/11/2017	brasil	Masculino	45 a 54 anos													
16	pesquisador2	07/05/2018	brasil		37													
17	pesquisador2	07/05/2018	brasil		31													
18	pesquisador2	07/05/2018	brasil		53 anos													
19	pesquisador2	07/05/2018	brasil		30													
20	pesquisador2	07/05/2018	brasil		34													
21	pesquisador2	08/05/2018	brasil		61													
22	pesquisador2	08/05/2018	brasil		40													
23	pesquisador2	09/05/2018	brasil		61													
24	pesquisador2	11/04/2017	brasil	Feminino	de 17 a 20 anos	O Estilo de Uso A - Uso Participativo no Espaço Virtual.												
25	pesquisador2	12/04/2017	brasil	Feminino	de 20 a 30 anos	O Estilo de Uso A - Uso Participativo no Espaço Virtual.												
26	pesquisador2	13/04/2017	brasil	Feminino	de 20 a 30 anos	O Estilo de Uso B - Busca e Pesquisa no Espaço Virtual.												
27	pesquisador2	17/04/2017	brasil	Feminino	de 20 a 30 anos	O Estilo de Uso B - Busca e Pesquisa no Espaço Virtual.												
28	pesquisador2	18/04/2017	brasil	Feminino	de 20 a 30 anos	O Estilo de Uso A - Uso Participativo no Espaço Virtual.												
29	pesquisador2	19/04/2017	brasil	Feminino	de 30 a 40 anos	O Estilo de Uso A - Uso Participativo no Espaço Virtual.												
30	pesquisador3	25/11/2010	mexico	masculino	46													
31	pesquisador3	25/11/2010	mexico	masculino	46													
32	pesquisador3	25/11/2010	mexico	feminino	39													
33	pesquisador3	25/11/2010	mexico	feminino	56													
34	pesquisador3	25/11/2010	mexico	feminino														
35	pesquisador3	25/11/2010	mexico	feminino	45													
36	pesquisador4		espanha	masculino	14 anos													
37	pesquisador4		espanha	masculino	14 anos													
38	pesquisador4		espanha	feminino	14 anos													
39	pesquisador4		espanha	feminino	13 anos													
40	pesquisador4		espanha	feminino	14 anos													
41	pesquisador4		espanha	masculino	14 anos													

Figure 2: Figura aaa. Amostra do dataset inicial de trabalho.

Fonte: Elaborado pela pesquisadora (2022).

**4.2.3.2 Tratamento dos dados** Segundo da Silva (2021), um dos principais objetivos da preparação de dados é garantir que os dados brutos que estão sendo preparados para processamento e análise sejam precisos e consistentes para que os resultados dos aplicativos de análise sejam válidos.

Como visto na Figura zzz, os dados geralmente são criados com valores ausentes, imprecisões ou outros erros, e conjuntos de dados separados geralmente têm formatos diferentes que precisam ser reconciliados quando combinados. Corrigir erros de dados, validar a qualidade dos dados e consolidar conjuntos de dados são grandes partes dos projetos de preparação de dados, afirma Wickham (2014).

Após a fase inicial de preparação de dados, a próxima etapa envolve encontrar dados relevantes para garantir que os aplicativos de análise forneçam informações significativas e insights acionáveis para a tomada de decisões. Os dados nesta nova geralmente são enriquecidos e otimizados para torná-los mais informativos e úteis - por exemplo, combinando conjuntos de dados internos e externos -, criando novos campos de dados,

eliminando valores discrepantes e abordando conjuntos de dados desequilibrados que podem distorcer os resultados da análise. Esta fase, de acordo com Wickham (2014), da Silva (2021) e Askham (2022), pode ser definida como **tratamento dos dados**. Portanto, na fase de tratamento, os dados são estruturados e organizados para estar de acordo com os requisitos da **Análise Exploratória de Dados (AED)**. Além de serem estruturados, afirma da Silva (2021), os dados normalmente devem ser transformados em um formato padronizado e utilizável, como por exemplo, a criação de novas variáveis ou colunas que agreguem valores das existentes. O *enriquecimento de dados*, por outro lado, aprimora e otimiza os *datasets* conforme necessário, por meio de medidas como aumentar e adicionar dados, define Askham (2022).

Datasets são definidos por Wickham e Grolemond (2017) como grupos de dados relacionados compostos por uma ou mais tabelas, mas que podem ser manipuladas como uma unidade por um computador.

Na etapa de tratamento, rotinas automatizadas desenvolvidas em um ambiente de desenvolvimento integrado são executadas nos dados para transformar, organizar e estruturar e validar sua consistência, integridade e precisão (Wickham, 2014). Os dados assim preparados são então armazenados em um repositório e usados diretamente para a AED ou então disponibilizados para que outros usuários possam também utilizá-los. Para a realização das etapas de preparação, transformação, organização e estruturação serão utilizados nesta tese os recursos da Linguagem R, um software livre de programação voltada ao tratamento, análise e visualização de dados. A linguagem é largamente utilizada por acadêmicos, estatísticos e analistas de dados para realizar projetos de AED. Foi criada originalmente no departamento de Estatística da Universidade de Auckland, Nova Zelândia e sua manutenção é realizada por uma comunidade de colaboradores voluntários que contribuem para melhoria e otimização da linguagem e com a expansão das suas funcionalidades básicas por meio de bibliotecas, ou packages residentes no repositório CRAN-R (The Comprehensive R Archive Network) (LINGUAGEM R, 2022)<sup>16</sup>.

Com o objetivo de facilitar e agilizar o processo de desenvolvimento na linguagem R dos cálculos estatísticos e geração de gráficos apresentados nesta tese, utilizou-se também o Ambiente de Desenvolvimento Integrado (do inglês Integrated Development Environment – IDE) RStudio, um programa de computador de fonte aberta com características e ferramentas de apoio ao desenvolvimento acelerado de software (LINGUAGEM R, 2022).

Portanto, como tarefa inicial para a organização e estruturação dos dados, no Quadro yyy é possível visualizar o dicionário de dados com os nomes e as definições das variáveis utilizadas na AED desta tese.

Um glossário ou dicionário de dados, de acordo com Askham (2022), é utilizado para armazenar um conjunto limitado de metadados<sup>17</sup> concentrados nos nomes e definições relacionadas aos dados físicos e objetos relacionados.

Quadro yyy: Dicionário de dados

Nome da variável	Tipo de variável	Conteúdo	Descrição
pesquisador	texto		Nome do(s) pesquisador(es) responsável(is) pela pesquisa
data.ano	data	dd/mm/aaaa	Dia, mês e ano em que o questionário foi aplicado
pais	texto	Brasil, Espanha, México ou Portugal	Países onde o questionário CEUEV foi aplicado
genero	texto	Feminino, Masculino ou NR	Gênero do inquirido
idade	texto	de 11 a 14 anos, de 14 a 17 anos, de 17 a 20 anos, de 20 a 30 anos, de 30 a 40 anos, de 40 a 50 anos, de 50 a 60 anos, de 60 a 70 anos, acima de 70 anos.	Faixa etária dos respondentes

<sup>16</sup><https://cran.r-project.org/>.

<sup>17</sup>Metadados, ou Metainformação, são dados sobre outros dados. Um item de um metadado pode dizer do que se trata aquele dado, geralmente uma informação inteligível por um computador. Os metadados facilitam o entendimento dos relacionamentos e a utilidade das informações dos dados. METADADOS (2022).

Nome da variável	Tipo de variável	Conteúdo	Descrição
estilo	texto	<b>Estilo de Uso A:</b> Uso Participativo no Espaço Virtual; <b>Estilo de Uso B:</b> Busca e Pesquisa no Espaço Virtual; <b>Estilo de Uso C:</b> Estruturação e Planejamento no Espaço Virtual; <b>Estilo de Uso D:</b> Ação Concreta e Produção no Espaço Virtual; <b>Não Identificado</b>	Estilos de uso do espaço virtual definidos em BARROS (2013)

Fonte: Elaborado pela pesquisadora (2022).

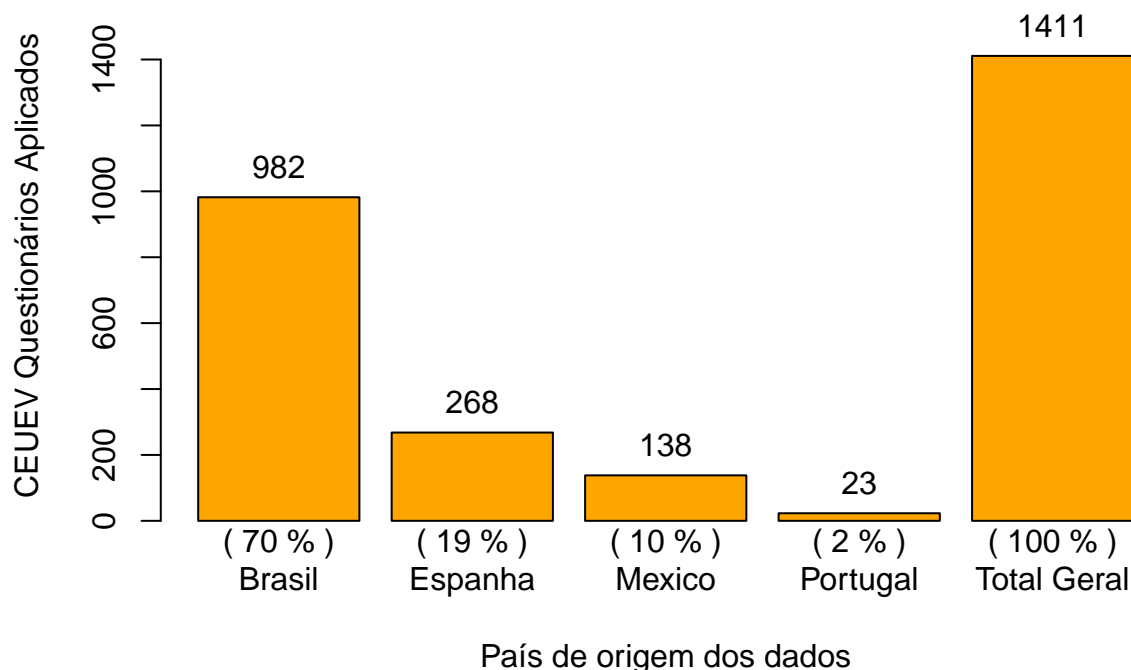
De acordo com Silva, Peres e Boscardioli (2016) a AED compreende, em conceitos estatísticos a análise descritiva, uma ferramenta capaz de descrever ou resumir dados, mostrando aspectos importantes do conjunto de dados como o tipo de distribuição associada aos valores mais representativos do conjunto, e permitindo criar visualizações referentes à tais aspectos.

De forma geral, a análise exploratória é uma abordagem à análise de conjuntos de dados de modo a resumir suas características principais, frequentemente com métodos visuais. Um modelo estatístico pode ou não ser usado, mas primariamente a AED tem como objetivo observar o que os dados podem nos dizer além da modelagem formal ou do processo de teste de hipóteses (WICKHAM e GROLEMUND, 2017).

```
#data("./dados/dados_minerados.csv")
euevgera <- read.table("./dados/dados_minerados.csv", header = TRUE, sep = ";", as.is=FALSE)
```

**4.2.3.3 Transformação dos dados** Os dados recebidos são referentes a quatro países, Brasil, Espanha, México e Portugal, totalizando **1411** observações, conforme pode-se visualizar no Gráfico 1.

**Gráfico 1: Quantidade casos por país**



Observa-se, no Gráfico 1, uma desigualdade referente a quantidade de dados recebidos de cada país, percebendo então que o Brasil foi o país com mais casos coletados 982 ou 70 %. Então considerou essa amostra de dados uma amostra com valores discrepantes pois as quantidades numéricas se diferem consideravelmente entre os países.

## Referências Bibliográficas

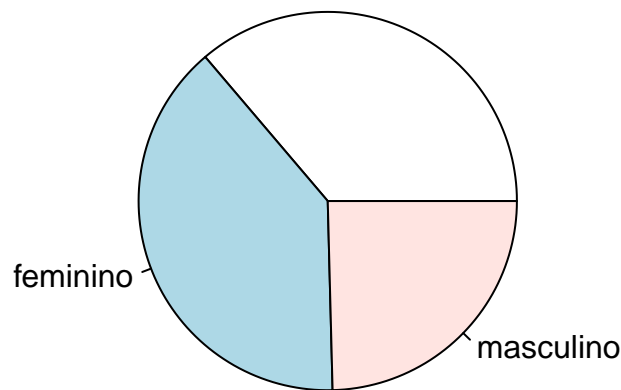
- López Berlanga, M. C., Vieira Barros, D.M., Sánchez Romero, C.(2019). El estilo de uso del espacio virtual de internet con estudiantes de Educación Secundaria. *Revista De Estilos De Aprendizaje*, 12(24), 77–88. <https://doi.org/10.55777/rea.v12i24.1389>
- Barrera, A. V. Herramienta para la creación de un Entorno Personal de Aprendizaje.2017. 186 p. Trabajo Terminal (Escuela Superior de Cómputo) - Instituto Politécnico Nacional, México, 2017.
- Ota, M. A. Adaptatividade em Ambientes Virtuais: uma proposta para personalizar a aprendizagem em cursos híbridos de ensino superior. 2018. 268 p. Tese (Doutorado em Ensino de Ciências e Matemática) - Universidade Cruzeiro do Sul, São Paulo, 2018.
- Terçariol, A. A.de L., Barros, D. M. V. (2017). Os estilos de uso dos espaços virtuais e as redes sociais na pedagogia: um estudo exploratório.*Revista De Estilos De Aprendizaje*, 10(20). <https://doi.org/10.55777/rea.v10i20.1067>
- Sánchez Romero, C.et al.(2020).Estudio de Identificación de Los Estilos Del Uso Del Virtual de Los Seniors: perspectivas iniciales, 14(2), 136–163.
- WICKHAM, H. Tidy Data. *Journal of Statistical Software*, [S. l.], v. 59, n. 10, p. 1–23, 2014. DOI: 10.18637/jss.v059.i10. Disponível em: <https://www.jstatsoft.org/index.php/jss/article/view/v059i10>. Acesso em: 26 jul. 2022.
- SILVA, D. Higienização de dados: o que é e 7 boas práticas para realizá-la. Blog da Zendesk, 11 ago 2021. Disponível em: <https://www.zendesk.com.br/blog/higienizacao-de-dados/>. Acesso em: 26 jun. 2022.
- LINGUAGEM R. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2022. Disponível em: [https://pt.wikipedia.org/w/index.php?title=R\\_\(linguagem\\_de\\_programa%C3%A7%C3%A3o\)&oldid=63783140](https://pt.wikipedia.org/w/index.php?title=R_(linguagem_de_programa%C3%A7%C3%A3o)&oldid=63783140). Acesso em: 13 jun. 2022.
- ASKHAM, N. Data Glossary or Data Dictionary? The Data Governance Coach. 14 nov. 2017. Disponível em: <https://www.nicolaaskham.com/blog/2017/11/8/what-is-a-data-glossary-and-how-is-it-different-from-a-data-dictionary>. Acesso em 08 mai. 2022.
- BARROS, D. M. V. Diseño y aplicación del Cuestionario Estilo de Uso Espacio Virtual (CEUEV). In CUE, J. L. G, *Estilos*



de aprendizagem y otras perspectivas pedagógicas del siglo XXI. México: Colegio de Postgraduados, 2013. P 149-164. HAN, J.; KAMBER, M.; PEI, J. (2011). Data Mining: Concepts and Techniques (3rd ed.). Morgan Kaufmann. ISBN 978-0-12-381479-1. KOTU, V. Data Science: Concepts and Practice. Cambridge: Morgan Kaufmann, 2019. METADADOS. In: WIKIPÉDIA, a enciclopédia livre. Flórida: Wikimedia Foundation, 2022. Disponível em: <https://pt.wikipedia.org/w/index.php?title=Metadados&oldid=63544067>. Acesso em: 08 mai. 2022. SILVA, Leandro Augusto da, PERES, Sarajane Marques, BOSCARIOLI, Clodis. Introdução à Mineração de Dados com Aplicações em R. Rio de Janeiro: Elsevier, 2016. TYAGI, Amit Kumar. Data Science and Data Analytics: Opportunities and Challenges. Flórida: CRC Press, 2022. WICKHAM, Hadley, GROLEMUND, Garret. R for Data Science: Imort, Tidy, Transform, Visualize and Model Data. Sebastopol: O'Rely Media, 2017.

## Including Plots

You can also embed plots, for example:



Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.