

## SITE DE RESULTADOS - CIÊNCIA DE DADOS APLICADA AO PISA 2018

### *Produto Técnico-Tecnológico 4 - Software/Aplicativo*



#### APRESENTAÇÃO

A educação está presente e ligada a várias áreas de interesse coletivo. Sabendo-se dessa importância, Avaliações em Larga Escala ou *Large-Scale Assessments* - LSAs são aplicadas em diversos países, objetivando-se identificar tais cenários educacionais, bem como monitorar as habilidades cognitivas relacionadas à literacia e à numeracia (HOYOS, ESTRADA, VARGAS, 2021).

Para o Brasil, assim como outros países, é de suma importância compreender a real situação de seus sistemas educacionais para promover mudanças que

Rochelly Fernandes Andrade

Prof. Dr. Alessandro Vivas Andrade

Mestrado Profissional em Educação  
em Ciências, Matemática e Tecnologia





**UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI**

**Reitor** Janir Alves Soares

**Vice-Reitor** Marcus Henrique Canuto

## **APOIO**

Grupo de Pesquisa em Educação em Ciências, Matemática e Tecnologia.



Rochelly Fernandes Andrade  
Prof. Dr. Alessandro Vivas Andrade

## **PRODUTO EDUCACIONAL:**

### **4 -Software/Aplicativo**

# **SITE DE RESULTADOS CIÊNCIA DE DADOS APLICADA AO PISA 2018**

Produto Educacional apresentado como requisito à obtenção do grau de Mestre em Educação em Ciências, Matemática e Tecnologia pelo Programa de Mestrado Profissional em Educação em Ciências Matemática e Tecnologia da Universidade Federal dos Vales do Jequitinhonha e Mucuri, campus Diamantina. Aprovado em banca de defesa de mestrado no dia 11/08/2022 pelos seguintes membros:

***Prof. Dr. ALESSANDRO VIVAS ANDRADE/UFVJM.***

***Prof. Dr. CRISTIANO GRIJÓ PITANGUI/UFSJ.***

***Prof. Dr. DIEGO NUNES BRANDÃO/CEFET-RJ.***

***Prof. Dra. HELEN ROSE DE CASTRO SILVA ANDRADE/UFVJM.***

**1ª Edição**

**UFVJM  
Diamantina, MG  
2022**

O conteúdo desta publicação é de inteira responsabilidade dos autores.  
Permitida a reprodução total ou parcial, desde que citada a fonte.

**Editoração eletrônica e projeto gráfico/capa:**

Programa de Pós-Graduação em Educação em Ciências, Matemática e Tecnologia

Elaborado com os dados fornecidos pelo (a) autor(a).

A554s	<p>Andrade, Rochelly Fernandes Site de resultados: ciência de dados aplicada ao PISA 2018 [recurso eletrônico] / Rochelly Fernandes Andrade.– Diamantina: UFVJM, 2022. 18 p. :il.</p> <p>Produto Educacional da pesquisa Ciência de Dados Aplicada ao PISA 2018, desenvolvida no Programa de Mestrado Profissional em Educação em Ciências Matemática e Tecnologia, da Universidade Federal dos Vales do Jequitinhonha e Mucuri, Orientador: Prof. Dr. Alessandro Vivas Andrade/UFVJM.</p> <p>1. Avaliações em larga escala. 2. Aprendizado de máquina. 3. Ciência de dados. 4. Educação. 5. PISA. I. Andrade, Rochelly Fernandes . II. Título. III. Universidade Federal dos Vales do Jequitinhonha e Mucuri.</p> <p style="text-align: right;"><b>CDD 370</b></p>
-------	---

Ficha Catalográfica – Serviço de Bibliotecas/UFVJM  
Bibliotecária Viviane Pedrosa– CRB-6/2641

# SUMÁRIO

<b>APRESENTAÇÃO DO PRODUTO EDUCACIONAL.....</b>	<b>1</b>
<b>CONTEXTUALIZAÇÃO.....</b>	<b>2</b>
O PISA.....	2
Ciência de Dados.....	3
Disseminação da Informação Científica.....	5
<b>METODOLOGIA DE DESENVOLVIMENTO.....</b>	<b>8</b>
<b>CONSIDERAÇÕES FINAIS .....</b>	<b>16</b>
<b>REFERÊNCIA BIBLIOGRÁFICA.....</b>	<b>17</b>

# APRESENTAÇÃO DO PRODUTO EDUCACIONAL

Este material apresentado como Produto Educacional, é parte integrante da nossa pesquisa intitulada Ciência de Dados Aplicada ao PISA 2018, desenvolvida no Programa de Mestrado Profissional em Educação em Ciências Matemática e Tecnologia, da Universidade Federal dos Vales do Jequitinhonha e Mucuri, sob orientação do Professor Doutor Alessandro Vivas Andrade.

Este Produto Educacional – PE, consiste em um *site* que apresenta todas as descobertas da pesquisa Ciência de Dados Aplicada ao PISA 2018. O *site* tem por objetivo disseminar a informação de maneira simples e acessível, promovendo interações e discussões por meio de comentários.

O site foi desenvolvido e composto com partes das informações referentes aos processos/etapas do projeto de Ciência de Dados sobre as bases de dados do PISA 2018.

Este PE, tem aderência na linha de pesquisa Ensino e Aprendizagem em Ciências e Matemática, sendo do tipo Protótipo e subtipo PTT4 - Software/Aplicativo, elaborado no programa, sendo aplicável, porém não aplicado. Tem-se por objetivo, responder a seguinte pergunta: Como apresentar as descobertas de uma pesquisa de uma maneira interativa e acessível?

O PE aqui apresentado, tem como público-alvo, não somente professores e formadores de políticas públicas da área da educação, mas também, todos os interessados em melhorias na educação Brasileira, apresentando alto teor inovador, visto que foi desenvolvido baseado em conhecimento inéditos.

O *site* que constitui o presente PE, pode ser acessado através do link <https://cienciadedadospisa.wordpress.com/> em diferentes plataformas, como *desktops* e *mobile*, e será validado em 2ª instância.

# CONTEXTUALIZAÇÃO

O presente PE é derivado da pesquisa intitulada Ciência de Dados Aplicada ao PISA 2018, que teve como objetivo principal a extração de conhecimento dos dados oriundos do PISA 2018, com o intuito de identificar características do questionário contextual aplicado aos professores, que têm relevância sobre os resultados dos alunos nas avaliações de matemática e ciências.

A pesquisa foi provocada, devido ao fato de os resultados do Brasil serem preocupantes, pois, nas três áreas de conhecimento avaliadas - leitura, matemática e ciências -, apresentou-se uma baixa proficiência, quando comparadas aos 78 países participantes do PISA INEP (2019). Este fato demonstra o grave cenário educacional que o Brasil enfrenta nas três áreas de conhecimento, levando-se em consideração a importância que a educação tem em várias áreas. Pesquisas para a melhoria desses índices de proficiência são indispensáveis.

Para este Produto Educacional, foi projetado um *web site* para realizar-se a disseminação dos conhecimentos e descobertas obtidos com a Pesquisa Ciência de Dados Aplicada ao PISA 2018.

## **O PISA**

O Programa Internacional de Avaliação de Estudantes, ou *Programme for International Student Assessment* - PISA, é uma Avaliação em Larga Escala, ou *Large-scale Assessment* - LSA, na qual uma LSA pode ser considerada, como testes e/ou questionários, aplicados à um grande número de estudantes simultaneamente (THURLOW, 2010).

O PISA é coordenado pela Organização para Cooperação e Desenvolvimento Econômico, ou *Organisation for Economic Co-operation and Development* – OECD, em parceria com órgãos coordenadores dos países participantes. No caso do Brasil, a coordenação é realizada pelo Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira - Inep (INEP, 2015).

O PISA é aplicado a cada três anos e tem por objetivo principal avaliar três áreas do conhecimento: leitura, matemática e ciências. Essa LSA é aplicada aos alunos do

ensino fundamental, com faixa etária de 15 anos de idade, que estão finalizando a educação obrigatória (INEP, 2020).

Além das três áreas de conhecimento avaliadas, o PISA busca coletar informações referentes ao aluno, escola, professores, assim como variáveis demográficas, socioeconômicas e educacionais, através de questionários que são aplicados na ocasião do teste.

A partir dos dados coletados pelos questionários, diversas técnicas podem ser aplicadas com a finalidade de extrair conhecimentos relevantes para a melhoria da educação como um todo. Dentre estas técnicas passíveis de serem aplicadas a tais dados, temos a Ciência de Dados.

A pesquisa que foi base para a elaboração desse PE, utilizou os dados coletados pelo questionário dos professores e resultados dos alunos nas avaliações de ciências e matemática da 7ª edição.

### ***Ciência de Dados***

Ciência de Dados ou *Data Science*, é uma área multidisciplinar que envolve, principalmente, estatística e aprendizado de máquina e vem crescendo no mundo nos últimos anos, devido às necessidades de manipulação de grandes quantidades de dados, tanto em volume, quanto em variedade (CURTY; CERVANTES, 2016). De maneira simplificada, a Ciência de Dados pode ser entendida com um processo de extração de informações em base de dados (BUGNION; MANIVANNAN; NICOLAS, 2017).

O processo de Ciência de Dados pode seguir algumas etapas, iniciando-se com o entendimento do problema, que é considerado uma das etapas mais importantes, visto que delimita o escopo do problema, bem como define os objetivos a serem alcançados, além disso, a má compreensão deste pode levar a resultados sem valor ou errados (PACHECO et al., 2014).

Com a delimitação e compreensão do problema feitos, é realizada a coleta, o pré-processamento e compreensão dos dados, abordando a limpeza de registros irrelevantes, ou que possam atrapalhar no processo, bem como o estudo de sua estruturação e notações (WITTEN et al., 2016).



Para um entendimento maior sobre os dados adquiridos, bem como uma descoberta prévia de conhecimento, são aplicadas técnicas de análise exploratória de dados, que de maneira geral, aplicam métodos gráficos estatísticos (MEDRI, 2011).

Outra etapa bem utilizada no processo de ciência de dados é a *Feature Engineering*, na qual é realizada uma análise de quais informações podem ser retiradas ou agrupadas, com o objetivo de simplificar a base de dados. Para Nargesian et al. (2017), a *Feature Engineering* é uma etapa que aprimora o desempenho da etapa de modelagem de dados, envolvendo aplicação de funções de transformação, as quais geram novos recursos, que por sua vez levam ao aprimoramento do modelo.

A etapa central de todo o processo é a chamada Modelagem de Dados e aplica modelos estatísticos e de aprendizado de máquina para validar uma hipótese (GONÇALVES, 2018), na qual Aprendizado de Máquina ou *Machine Learning* - ML é um campo de estudo presente na Inteligência Artificial (IA), dedicado à criação de sistemas de aprendizado computacional, que neste contexto, são programas de computador que aprendem e tomam decisões, com base em experiências de problemas anteriormente solucionados (SANTOS, 2005).

As técnicas de ML são utilizadas para diversas finalidades e aplicações, como por exemplo, o uso em problemas de classificação, reconhecimento de padrões e jogos (BIANCHI, 2004). Além do objetivo de adquirir conhecimento de maneira automática, muitos sistemas de aprendizado empenham-se em simular outros tipos de aprendizado, como por exemplo, o aprendizado humano (ARAÚJO, 2004).

Segundo Tech (2022), um modelo pode ser entendido como “a equação final gerada através de um algoritmo de *machine learning*, que será utilizada para definir os valores de saída a partir de novos dados apresentados” e para Microsoft (2022) “é um arquivo que foi treinado para reconhecer determinados tipos de padrões”.

De maneira mais simples, pode-se dizer que um modelo de ML é um *software* que foi preparado e alimentado com exemplos já classificados ou rotulados com valores, em que se aprende com esses exemplos e é capaz de realizar previsões de classes, ou valores em exemplos não vistos anteriormente.

Através dos modelos de ML, é possível realizar previsões do status de uma variável dependente a partir de um conjunto de variáveis independentes.

Para o presente PE, a variável independente corresponde a um conjunto de características coletadas de um questionário aplicado a professores e a variável dependente ao status de desempenho vinculado a um possível aluno do professor.

Para que um modelo de *machine learning* tenha credibilidade, ele deve ser avaliado. Em problemas de classificação, algumas métricas de avaliação podem ser calculadas a partir da contagem de previsões corretas e incorretas, sob a perspectiva de uma determinada classe.

Dentre as métricas de avaliação de modelos, podemos citar: *Precision*, que mede a proporção sob a perspectiva de uma determinada classe da predição que está correta; *Recall*, que indica a relação entre os exemplos de uma determinada classe corretamente identificadas pelo classificador e todas as previsões que realmente são dessa classe em questão; *F-measure*, calculado pela média harmônica ponderada dos resultados de *Recall*; e *Precision, Accuracy*, que tem o objetivo de medir a proporção dos exemplos preditos (classificados) de maneira correta, independentemente da classe.

Após todas as etapas mencionadas anteriormente, as descobertas devem ser comunicadas e disseminadas de maneira clara e compreensível, levando em conta que, um dos determinantes do sucesso da ciência de dados é justamente a compreensão das descobertas por parte do público-alvo, para que esse conhecimento adquirido seja posteriormente utilizado, finalizando assim, o projeto de ciência de dados.

### ***Disseminação da Informação Científica***

Levando em consideração que, para que o sucesso do processo de ciência de dados seja alcançado deva ser realizada a comunicação adequada ao seu público-alvo, faz-se necessário refletir acerca de como o conhecimento científico é disseminado.

Kuramoto (2006) cita a dificuldade em relação ao acesso à informação científica, principalmente nos países em fase de desenvolvimento como o Brasil, visto que, tal tipo de informação, é base para desenvolvimento científico e tecnológico do país.

O conhecimento científico não deve estar acessível apenas a uma parte da população. Ele deve ser democratizado e compartilhado, visto que é um produto social em prol da prosperidade, que melhora a qualidade de vida dos indivíduos, promovendo a inclusão social e o desenvolvimento do senso crítico (PADILHA; PRESSER; ZARIAS, 2016).

Para disseminar a informação científica, devemos considerar o conceito de fonte de informação. Araújo e Fachin (2015, p. 84) citam que:

As fontes de informações são registros utilizados ao longo da vida do ser humano, possibilitando ampliar a visão do mundo em que vive e sobre as coisas que estão a sua volta. No campo científico são aquelas que nos permitem criar, recriar e ter acesso ao conhecimento sobre um assunto ou área de nosso interesse ou pesquisas. De modo que, as fontes de informações são referências sobre o que está registrado e disponível ao ser humano, possibilitando reinventar ou compreender melhor seu objeto de estudo.

Com as novas tecnologias de informação, como fontes de informação, foi possível disseminar o conhecimento científico de maneira rápida e eficiente (ARAUJO; FACHIN, 2015), sendo que palavras como intangibilidade, conectividade, velocidade e inovação, ficaram cada vez mais relacionadas ao contexto da informação (SUAIDEN, 2007).

Segundo Costa (1997, p. 3):

as tecnologias de informação geram diversidades e mudanças na sociedade, provocando diferentes impactos e exigindo uma nova postura por parte dos indivíduos diante dos novos cenários então vivenciados.

Baseado nessa concepção, adaptar-se a essa nova realidade é uma necessidade, em vários contextos, inclusive no contexto da disseminação de conhecimento científico.

Dentre várias maneiras de disseminar conhecimento por meios tecnológicos, existem os chamados *web sites*. Os *web sites*, ou simplesmente *sites*, são um método muito comum de disseminação de informações de todos os tipos. De maneira simplificada, um *site* pode ser definido como um conjunto organizado de páginas relacionadas e interligadas entre si, através dos chamados *hiperlinks* ou *links* (UNIT, 2021).

Os *sites* podem ser utilizados em diversos contextos, para diversos propósitos. Atualmente, existem ferramentas que facilitam muito a criação de sites e, dentre essas ferramentas, existem os chamados *Content Management System* – CMS, ou Sistema de Gerenciamento de Conteúdo.

Um CMS é um *software* que provê um certo nível de automação para o gerenciamento de conteúdo em *site*, o qual, nesse contexto, é a informação produzida através de um processo editorial para o consumo via publicação. Os CMSs, usualmente, são baseados em servidores, permitindo aos usuários manipularem conteúdos e os disponibilizarem para outras pessoas (BARKER, 2016).

SYDLE (2021) cita algumas características dos CMS como: simplicidade para a criação de páginas *web*, atualização simples de conteúdo, funcionalidades extensíveis

através de extensões e plugins, escalabilidade, integração com outros *softwares*, dentre outros.

Para o presente PE, utilizou-se o CMS chamado *Wordpress*. O *Wordpress* foi criado em 2003 por Mike Little e Matt Mullenweg e hoje está presente em cerca de 43% de todos os *sites* na *Internet*, apresentando-se como uma plataforma elegante e bem arquitetada, que dá oportunidades para que qualquer pessoa crie e compartilhe vários tipos de conteúdos (WORDPRESS, 2022a).

## METODOLOGIA DE DESENVOLVIMENTO

O *site* foi desenvolvido utilizando o plano gratuito do *Wordpress*, que disponibiliza um endereço com domínio *wordpress.com*, diversos temas, número ilimitado de páginas e *posts*, painel de administração e 1 GB de armazenamento de mídia (WORDPRESS, 2022b).

Dentro do *site*, criou-se um menu de navegação, onde o visitante tem uma visão simplificada das principais informações que foram trazidas do projeto “Ciência de Dados Aplicada ao PISA 2018”. O menu principal pode ser observado na Figura 1.

**Figura 1.** Menu Principal



Fonte: Próprio Autor.

Através do Menu Principal apresentado na Figura 1, o visitante tem acesso a todas as páginas do site.

A página *Home*, é a página chamada de “Apresentação”, que traz um resumo sobre o trabalho, bem como um panorama das informações contidas no *site*.

A intenção da página “Apresentação”, é ambientar o visitante/leitor ao trabalho que foi desenvolvido, para que ele possa compreender melhor as demais informações apresentadas por outras páginas dos *sites*, trazendo um apanhado geral sobre os objetivos do trabalho, metodologias, resultados e algumas conclusões. A página “Apresentação” pode ser observada na Figura 2.

**Figura 2.** Página *Home* - Apresentação



## APRESENTAÇÃO

A

educação está presente e ligada a várias áreas de interesse coletivo. Sabendo-se dessa importância, Avaliações em Larga Escala ou *Large-Scale Assessments* – LSAs são aplicadas em diversos países, objetivando-se identificar tais cenários educacionais, bem como monitorar as habilidades cognitivas relacionadas à literacia e à numeracia (HOYOS; ESTRADA; VARGAS, 2021).

Fonte: Próprio Autor.

A página “Etapas do Projeto”, busca fazer com que o visitante tenha uma noção de como o projeto foi conduzido. As etapas apresentadas nessa página são:

- + Entendimento do problema;
- + Aquisição dos dados na base do PISA;
- + Compreensão dos dados;
- + Análise exploratória de dados;
- + Tratamento de dados faltantes;
- + *Feature engineering*, modelagem de dados;;
- + Avaliação dos modelos
- + Extração e ranqueamento das 10 características que mais influenciaram no desempenho dos alunos.

Nesta página, cada fase do projeto foi separada em blocos, para melhor compreensão do leitor. Uma parte da página “Etapas dos Projetos” pode ser observada na Figura 3.

**Figura 3.** Página - Etapas do Projeto



## ETAPAS DO PROJETO

As tarefas desenvolvidas neste trabalho foram baseadas nas etapas de um projeto de *Data Science* incorporadas à metodologia da pesquisa com abordagem quantitativa descritiva tendo em vista as características de mensuração em números e uso de técnicas estatísticas (DALFOVO; LANA; SILVEIRA, 2008) sobre os dados obtidos do PISA 2018.

Fonte: Próprio Autor.

O menu de “Resultados” foi dividido em submenus, sendo eles: “Análise Exploratória de Dados”, “Avaliação de Modelos” e “Características Influenciadoras de Desempenho”.

Dentro do submenu “Análise Exploratória de Dados”, foi criado ainda outro submenu, contendo as análises sobre “Resultados dos Alunos no PISA 2018”, “Capacitação de Professores”, “Idade de Professores” e “Uso de Tecnologias Como Instrumento de Ensino”. Os submenus criados podem ser observados na Figura 4.

**Figura 4.** Submenus



Fonte: Próprio Autor.

A página “Resultados dos Alunos no PISA 2018”, traz a análise referente às notas dos alunos de todos os países participantes do PISA 2018, das avaliações de matemática

e ciências. Nessa página, são analisadas questões referentes à frequência de notas médias, concentração em faixas de notas e resultados dos alunos pelo mundo. Para ambos os testes, a página pode ser observada na Figura 5.

**Figura 5.** Página - Resultados dos Alunos no PISA 2018



## RESULTADOS DOS ALUNOS NO PISA 2018

**A**s técnicas de Análise Exploratória de Dados foram aplicadas sobre a perspectiva da estatística descritiva. Primeiramente, foram explorados dados relacionados aos resultados dos alunos no PISA 2018, para ambos os testes de ciências e matemática.

A Figura 1 apresenta a distribuição de frequências das notas médias de cada um

Fonte: Próprio Autor.

A Página “Capacitação de Professores”, apresenta as análises oriundas de questões relacionadas à capacitações de professores, que foram coletadas pelo questionário do PISA 2018.

Essa página apresenta análises relacionadas ao tempo de licenciatura ou programa de capacitação dos professores, tempo que os professores estudaram em algum país estrangeiro, realização de atividades de desenvolvimento profissional e a percepção do professor sobre a capacidade da escola em fornecer uma boa educação em relação a qualificação deles. Uma parte da página pode ser observada na Figura 6.



**Figura 6. Página - Capacitação de Professores**



## CAPACITAÇÃO DE PROFESSORES

Analisando as características relacionadas à capacitação dos professores, notou-se que a maioria dos países apresentam mais capacitações com tempos maiores que um ano conforme a Figura 5.



Fonte: Próprio Autor.

A página “Idade de Professores”, apresenta informações sobre a distribuição das idades dos professores e a perspectiva de idades em relação ao gênero dos professores que responderam ao questionário. Uma parte dessa página pode ser observada na Figura 7.

**Figura 7. Página - Idade de Professores**



## IDADES DE PROFESSORES

Além dos desempenhos dos países, foram analisadas as características dos professores coletados pelo questionário do PISA. A distribuição das idades de todos os professores pode ser observada na Figura 9 e na Figura 10.

Esta informação é segmentada por país, na qual se nota que o Azerbaijão possui mais professores com idades maiores e o Macau com idades menores, que é perceptível por apresentar uma distância interquartil menor, indicado maior

Fonte: Próprio Autor.

Na página “Uso de Tecnologias Como Instrumento de Ensino”, é possível encontrar informações referentes à abordagem tecnológica para ensino, relacionada a jogos, multimídia, recursos computacionais, redes sociais e política de uso de dispositivos eletrônicos. Um trecho da página pode ser observado na Figura 8.

**Figura 8.** Página - Uso de Tecnologias Como Instrumento de Ensino



## USO DE TECNOLOGIAS COMO INSTRUMENTO DE ENSINO

As Figuras 12, 13, 14, 15, 16 e 17 apresentam o uso de tecnologias como instrumento de ensino em sala de aula.

Fonte: Próprio Autor.

Os resultados das avaliações dos modelos de *Machine Learning*, foram apresentados na página “Avaliação de Modelos”. Nessa página, são apresentadas as métricas dos modelos de *Decision Tree* e *Random Forest* avaliados com *Precision*, *Recall*, *F1-score*, *Accuracy* e *Matthews Correlation Coefficient*, que foram obtidos com a técnica *Holdout*, para ambos os testes de matemática e ciências. Uma parte dessa página pode ser observada na Figura 9.

**Figura 9.** Página - Avaliação de Modelos



## AValiação DE MODELOS

Conforme as Etapas do Projeto adotadas, foram aplicados os modelos de *Decision Tree* e *Random Forest* avaliados com as métricas *Precision Recall*, *F1score*, *Accuracy* e *Matthews Correlation Coefficient*. Os resultados das métricas foram obtidos com a técnica *Holdout*, para ambos os testes de matemática e ciências.

Fonte: Próprio Autor.

As características encontradas que tiveram mais influência sobre o desempenho dos alunos, foram apresentadas na página “Características Influenciadoras de Desempenho”, que pode ser observada na Figura 10.

**Figura 10.** Página - Características Influenciadoras de Desempenho



Fonte: Próprio Autor.

A página “Conclusões”, como o próprio nome diz, apresentou conclusões e discussões acerca de todas as etapas do trabalho, analisando os resultados das avaliações dos modelos, bem como as características influenciadoras de desempenho encontradas. Uma parte da página pode ser observada na Figura 11.

**Figura 11.** Página - Conclusões

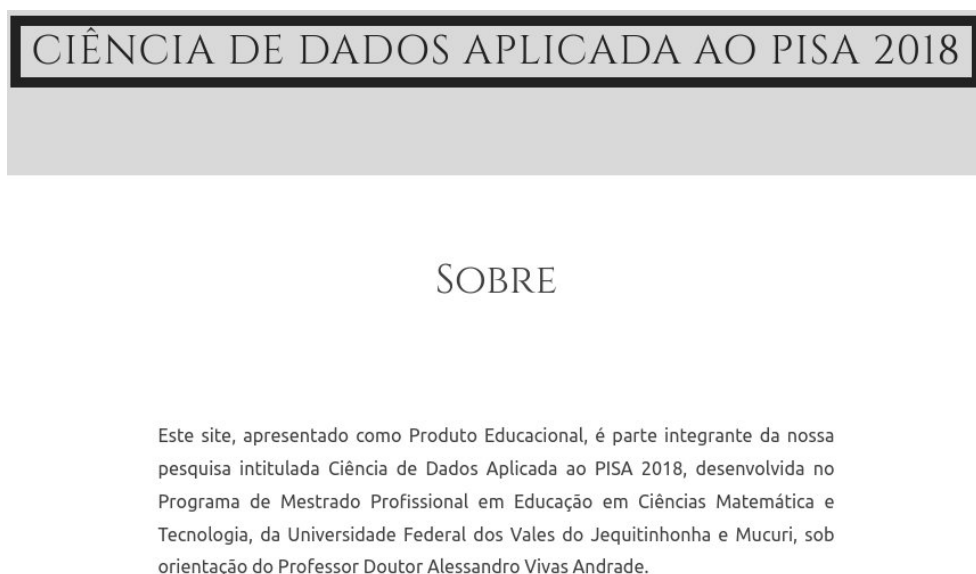


O presente estudo aplicou técnicas da Ciência de Dados sobre dados contextuais dos professores coletados pela Avaliação em Larga Escala PISA edição 2018. Observou-se a existência de várias possibilidades de caminhos para a condução do processo de Ciência de Dados.

Fonte: Próprio Autor.

Foi exibido através da página “Sobre”, uma explicação sobre os vínculos do *site* apresentado como PE, bem como sua caracterização. Ainda nessa página, é possível acessar o tutorial para a execução dos códigos utilizados na pesquisa. A página pode ser observada na Figura 12.

**Figura 12.** Página - Sobre



Fonte: Próprio Autor.

Em todas as páginas, foi ainda adicionado um campo para que o usuário pudesse comentar acerca do conteúdo apresentado, conforme demonstra a Figura 13.

**Figura 13.** Campo de Comentário



Fonte: Próprio Autor.

## CONSIDERAÇÕES FINAIS

O presente PE teve o intuito de apresentar os resultados de uma produção científica, realizados de maneira alternativa, através de um *site* em que várias pessoas poderão acessá-lo em diversos locais e através de variados tipos de dispositivos, maximizando os efeitos para a disseminação da informação científica.

Os comentários poderão ser utilizados como um método de *feedback* entre os leitores/visitantes e os pesquisadores que, além de provocar discussões, poderão ser insumo para *insights* de novas pesquisas.

O *Wordpress* se mostrou como uma ótima ferramenta para a criação de *sites*, provendo uma plataforma completa para a produção, bem como hospedagem e um nome de domínio grátis, fator facilitador para a construção desse PE.

Através das informações apresentadas de maneira resumida e simplificada, o leitor/visitante do *site*, mesmo que não ambientado ao contexto de trabalhos científicos, poderá despertar o interesse, posteriormente, na leitura completa do trabalho Ciência de Dados Aplicada ao PISA 2018.

Como demonstrado através do *site* criado, foi possível apresentar as descobertas da pesquisa de uma forma interativa e acessível, fato esse que responde à questão deste Produto Educacional, fazendo com que ele cumpra seu propósito.

## REFERÊNCIA BIBLIOGRÁFICA

ARAÚJO, N. C.; FACHIN, J. Evolução das fontes de informação. **Revista do Instituto de Ciências Humanas e da Informação**, v. 29, n. 1. 2015.

ARAÚJO, R. M. D. **Aprendizado de máquina em sistemas complexos multiagentes: estudo de caso em um ambiente sob racionalidade limitada**. Dissertação (Mestrado) - Universidade Federal Do Rio Grande Do Sul, Porto Alegre, dez. 2004.

BARKER, D. **Web content management: Systems, features, and best practices**. [S.l.]: O'Reilly Media, Inc., 2016.

BIANCHI, R. A. D. C. **Uso de heurísticas para a aceleração do aprendizado por reforço**. Tese (Doutorado) — Universidade de São Paulo, São Paulo, 2004.

BUGNION, P.; MANIVANNAN, A.; NICOLAS, P. R. **Scala: Guide for Data Science Professionals**. [S.l.]: Packt Publishing Ltd, 2017.

COSTA, S. M. de S. Impactos sociais das tecnologias de informação. **Revista de Biblioteconomia de Brasília**, Brasília, v. 19, n. 1, p. 3-22, jan./jun. 1995.

CURTY, R. G.; CERVANTES, B. M. N. Data science: Ciência orientada a dados. **Informação & Informação**, v. 21, n. 2, p. 1-4, 2016.

GONÇALVES, P. **Afinal, como se desenvolve um projeto de Data Science?** 2018. Disponível em: <<https://medium.com/techbloghotmart/afinal-como-se-desenvolve-um-projeto-de-data-science-233472996c34>>. Acesso em: 16 mai. 2019.

INEP. **Pisa 2018 revela baixo desempenho escolar em leitura, matemática e ciências no Brasil**. 2019. Disponível em: <[http://portal.inep.gov.br/artigo/-/asset\\_publisher/B4AQV9zFY7Bv/content/pisa-2018-revela-baixo-desempenho-escolar-em-leitura-matematica-e-ciencias-no-brasil/21206](http://portal.inep.gov.br/artigo/-/asset_publisher/B4AQV9zFY7Bv/content/pisa-2018-revela-baixo-desempenho-escolar-em-leitura-matematica-e-ciencias-no-brasil/21206)>. Acesso em: 16 jun. 2021.

INEP. **Relatório Brasil no PISA 2018**. Brasília-DF: Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2020. v. 53. 185 p. 2018.

INEP, M. d. E. **Programa Internacional de Avaliação de Estudantes (Pisa)**. 2015. Disponível em: <<http://inep.gov.br/pisa>>. Acesso em: 25 abr. 2019.

KURAMOTO, H. Informação científica: proposta de um novo modelo para o brasil. **Ciência da informação**, SciELO Brasil, v. 35, p. 91-102, 2006.

MEDRI, W. Análise exploratória de dados. **Centro De Ciências Exatas – CCE Departamento De Estatística - Universidade Estadual de Londrina**, v. 15, p. 05-13, 2011.

MICROSOFT. **O que é um modelo de machine learning?** 2022. Disponível em: <<https://docs.microsoft.com/pt-br/windows/ai/windows-ml/what-is-a-machine-learning-model>>. Acesso em: 03 jul. 2022.

NARGESIAN, F.; SAMULOWITZ, H.; KHURANA, U.; KHALIL, E. B.; TURAGA, D. Learning Feature Engineering for Classification. In: **Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence**. California: International Joint Conferences on Artificial Intelligence Organization, v. 0, n. August, p. 2529–2535. 2017.

PACHECO, F.; RANGEL, C.; AGUILAR, J.; CERRADA, M.; ALTAMIRANDA, J. Methodological framework for data processing based on the data science paradigm. In: IEEE. **2014 XL Latin American Computing Conference (CLEI)**. [S.l.], p. 1–12. 2014.

PADILHA, S. C.; PRESSER, N. H.; ZARIAS, A. **Divulgação científica: uso social do produto dos estudos científicos na fundação joaquim nabuco**. Em **Questão**, p. 161–187, 2016.

SANTOS, C. N. dos. **Aprendizado de máquina na identificação de sintagmas nominais: o caso do português brasileiro**. Tese (Doutorado) — Instituto Militar de Engenharia, 2005.

SUAIDEN, E. J. Informação científica e tecnológica—a web e a teia da vida. **Ponto de Acesso**, v. 1, n. 1, 2007.

SYDLE. **O que é CMS e quais seus benefícios?** 2021. Disponível em: <<https://www.sydle.com/br/blog/cms-6139106984679b13bf02ae61/>>. Acesso em: 03 jul. 2022.

TECH, D. **O que é um Modelo de Machine Learning**. 2022. Disponível em: <<https://didatica.tech/o-que-e-um-modelo-de-machine-learning/>>. Acesso em: 03 jul. 2022.

THURLOW, M. Large Scale Assessment and Accountability and Students with Special Needs. In: **International Encyclopedia of Education**. Elsevier, 2010. p. 752–758. Disponível em: <<https://linkinghub.elsevier.com/retrieve/pii/B9780080448947011581>>. Acesso em: 03 jul. 2022.

UNIT, D. **What is a site?** 2021. Disponível em: <<https://www.digitalunite.com/technology-guides/using-internet/searching-browsing/what-site>>. Acesso em: 03 jul. 2022.

WITTEN, I. H.; FRANK, E.; HALL, M. A.; PAL, C. J. Data mining: Practical machine learning tools and techniques. In: **Data Mining: Practical machine learning tools and techniques**. 4. ed. Cambridge-MA, United States: Morgan Kaufmann, p. 56–65. 2016.

WORDPRESS. **Nossa história**. 2022. Disponível em: <<https://br.wordpress.org/about/>>. Acesso em: 03 jul. 2022.

WORDPRESS. **Planos do WordPress.com**. 2022. Disponível em: <<https://wordpress.com/pt-br/support/recursos-do-plano/>>. Acesso em: 03 jul. 2022.

22. Disponível em: <<https://wordpress.com/pt-br/support/recursos-do-plano/>>. Acesso em: 03 jul. 2022.