

USO DE *BUSINESS INTELLIGENCE* PARA AVALIAÇÃO DE INDICADORES DE DESEMPENHO NA EDUCAÇÃO BÁSICA NO ESTADO DO ACRE

USING BUSINESS INTELLIGENCE TO EVALUATE PERFORMANCE INDICATORS IN PRIMARY EDUCATION IN THE STATE OF ACRE

Pedro Farias Wanderley¹

Cláudio de Souza Baptista²

Ronald Polanco Ribeiro³

Anselmo Cardoso de Paiva⁴

Resumo: o referido trabalho apresenta uma solução de *Business Intelligence (BI)* voltada à construção de painéis de Indicadores-Chave de Desempenho (*Key Performance Indicators - KPI*) da Educação Básica que, de forma interativa, fornecem informações do desempenho escolar dos Municípios do Estado do Acre. A metodologia consiste na escolha dos indicadores educacionais, na implantação do processo de extração-transformação-carga dos dados educacionais de diversas fontes, e implementação dos indicadores em painéis de *Business Intelligence*. Por fim, realiza-se uma avaliação de usabilidade da ferramenta de BI. Conclui-se que a ferramenta proposta, através da simplificação das narrativas, facilita os processos decisórios de gestores e o entendimento do cidadão das políticas educacionais.

Palavras-chave: *Business Intelligence, ETL, Indicadores de Desempenho, KPI, Educação.*

Abstract: the work presents a Business Intelligence (BI) solution aimed at building Key Performance Indicator (KPI) dashboards for Primary Education, which interactively provide information on school performance

1 Aluno concluinte do Curso de Bacharelado em Ciência da Computação da UFCG. É pesquisador de Iniciação Científica no Laboratório de Sistemas de Informação da UFCG. pedro.wanderley@ccc.ufcg.edu.br.

2 Ph.D. em Ciência da Computação pela University of Kent at Canterbury, Inglaterra. Mestre em Informática pela Universidade Federal da Paraíba. Bacharel em Ciência da Computação e Bacharel em Direito pela Universidade Estadual da Paraíba. É Professor Titular do Departamento de Sistemas e Computação da Universidade Federal de Campina Grande, onde atua no Doutorado, Mestrado e Graduação e coordena o Laboratório de Sistemas de Informação. É Pesquisador 2 do CNPq. baptista@computacao.ufcg.edu.br.

3 Mestre em Desenvolvimento Regional pela Universidade Federal do Acre. Especialista. MBA Executivo em Gestão Empresarial pela FGV. Bacharel em Economia pela Universidade Federal do Acre. Tecnólogo em Tecnologia da Construção Civil – Modalidade Estradas e Topografia pela Universidade Federal do Acre. Conselheiro do Tribunal de Contas do Estado do Acre, onde exerce a Presidência no biênio 2021-2023. Deputado Estadual do Estado do Acre por três mandatos pelo Partido dos Trabalhadores (1995-2011). polanco@tce.ac.gov.br.

4 Doutor em Informática e Mestre em Engenharia Civil pela PUC-Rio, Engenheiro Civil pela Universidade Estadual do Maranhão. É professor Titular do Departamento de Informática da UFMA onde atua no Doutorado, Mestrado e Graduação e coordena o Núcleo de Computação Aplicada. É também Pesquisador 2 do CNPq. paiva@nca.ufma.br.

in the Municipalities of the State of Acre. The methodology consists of choosing educational indicators, implementing the extraction-transformation-loading process of academic data from different sources, and implementing the indicators in Business Intelligence dashboards. Finally, a usability assessment of the BI tool is performed. It concludes that the proposed tool, through the simplification of the narratives, facilitates the decision-making processes of managers and the citizen's understanding of educational policies.

Keywords: Business Intelligence, ETL, Performance Indicators, KPI, Education

1 INTRODUÇÃO

Com o advento da nova gestão pública brasileira, um dos maiores desafios enfrentados pelos gestores públicos passou a ser a organização dos dados acumulados nas diferentes esferas do governo, trazendo com isso limitações no acompanhamento das políticas públicas, através de índices e indicadores de desempenho de setores como o educacional, que têm papel destacado na redução das desigualdades regionais. No caso específico dos municípios acreanos, observa-se o quanto é fundamental para a governança pública regional a utilização de ferramentas que sintetizam informações e facilitam a gestão de políticas públicas não somente do setor educacional, mas também, daqueles como: a saúde, meio ambiente, o social, o econômico, a mobilidade urbana, a segurança, entre outros.

No setor da educação, as esferas administrativas necessitam de um meio que viabilize um entendimento concreto e ágil da real situação educacional de populações indígenas, ribeirinhas e mesmo daquelas residentes nas periferias dos municípios acreanos. Na maioria dos casos, as informações vinculadas à educação estão dispostas através de dados brutos de difícil entendimento e leitura, em forma de microdados, resultantes de pesquisas educacionais realizadas por entidades públicas, ou em séries históricas, muitas vezes com lapsos temporais. Em seu formato original, esses dados não têm tanto poder informacional para gestores públicos e a cidadania, pois o formato utilizado torna seu entendimento e leitura complexos e a informação é apresentada com características abstratas.

O entendimento de dados obtidos através de avaliações e pesquisas censitárias, como o censo escolar⁵, é crucial para que se possa ter uma noção do que precisa ser melhorado, o que deve ser mantido, quais as perspectivas para o futuro e quais ações devem ser tomadas a curto, médio e longo prazo. Além desse entendimento, um estudo detalhado sobre quais indicadores educacionais são relevantes para um determinado contexto é fundamental, a fim de que haja um *feedback construtivo e focal, atendendo à diversidade das populações residentes nessas localidades, muitas delas com acesso tão somente por via fluvial, possibilitando com isso ações voltadas às necessidades específicas de cada segmento social.*

Sistemas de Business Intelligence (BI) combinam arquiteturas, bancos de dados, ferramentas analíticas, aplicativos e metodologias visando dar suporte à

⁵ Disponível em: <https://www.gov.br/inep/pt-br/areas-de-atuacao/pesquisas-estatisticas-e-indicadores/censo-escolar>.

tomada de decisão. O principal objetivo do BI é possibilitar o acesso interativo a informações e dados, para assim permitir que os gestores dos negócios e analistas públicos possam conduzir análises apropriadas. Através do processo de análise dessas informações, históricas ou atuais, os indivíduos responsáveis pela tomada de decisão obtêm insights valiosos que os permitem tomar decisões melhores e mais embasadas (SHARDA; DELEN; TURBAN, 2018, p. 16). Sistemas de BI apresentam-se como soluções ideais para tarefas analíticas, tal qual a análise de índices educacionais (DIEPEN; BREDEWEG, 2016; VITAL et al., 2019).

A integração de soluções de BI, no âmbito da gestão educacional, tem uma capacidade imensa de compilar informações de diversas fontes, provendo ferramentas de fácil acesso e entendimento que transformam dados brutos em informação e conhecimento (ZHU, 2018). Esta visão sistêmica habilita a aferição das métricas de eficiência, eficácia e efetividade nas políticas públicas voltadas para a educação.

Este artigo propõe uma arquitetura de BI para indicadores de desempenho em educação, com integração de dados de diversas fontes, possibilitando ao gestor de educação um avanço no processo de tomada de decisão. Os indicadores aqui propostos advêm do Projeto Integrar⁶, projeto este que é fruto de uma parceria do Tribunal de Contas da União⁷ (TCU) com a Organização para Cooperação e Desenvolvimento Econômico⁸ (OCDE) e vários Tribunais de Contas dos estados brasileiros. Foi implementado um estudo de caso para o Acre, pela facilidade no acesso às informações desse Estado, com apoio do Tribunal de Contas do Estado do Acre. No entanto, o sistema proposto pode ser facilmente adaptado para ser utilizado em qualquer estado do Brasil, bastando que se tenham os dados necessários para representar os indicadores de desempenho propostos.

O presente trabalho teve como principal objetivo a construção de painéis interativos (do inglês, *dashboards*), que trazem informações atuais e séries históricas, sobre indicadores educacionais dos municípios do Estado do Acre. O intuito deste projeto é levar informação para gestores públicos e cidadãos comuns sobre a situação dos índices educacionais das suas localidades. Além disso, os painéis interativos funcionam como uma ferramenta de suporte à tomada de decisão de gestores educacionais, sempre visando a melhoria dos índices vinculados à Educação Básica.

O restante deste artigo está assim organizado: a seção 2 contempla a fundamentação teórica, na qual são discutidos os trabalhos relacionados ao tema. Na seção 3 é apresentada a metodologia utilizada para desenvolver a solução de BI. Na seção 4, os principais resultados obtidos são apresentados. Na seção 5 são expostos os resultados do processo de avaliação qualitativa da solução proposta e, por fim, na seção 6, são abordadas as considerações finais e perspectivas para trabalhos futuros.

6 Disponível em: <https://sites.tcu.gov.br/integrar>.

7 Disponível em: <https://portal.tcu.gov.br>.

8 Disponível em: <https://www.oecd.org>.

2 FUNDAMENTAÇÃO TEÓRICA

Sistemas de apoio à decisão, utilizando-se de *business intelligence* e KPI (*Key Performance Indicators*), vêm sendo amplamente explorados no domínio da gestão educacional em todo o mundo.

Autores como Diepen e Bredeweg (2016) usaram algoritmos genéticos para determinar indicadores de aprendizagem para alunos da educação secundária. Weeda (et al., 2020, p. 1) investigaram indicadores que melhor caracterizam a qualidade de código em programas elaborados por alunos da educação secundária. Já Rezgui, Gómez e Maaouia (2017) usaram KPI para aprimorar processos de apoio à tomada de decisão baseados em sistemas de gestão de qualidade aplicados ao ensino superior. Petiot (2018) apresentou um experimento para o cálculo de indicadores educacionais usando técnicas de data mining e sistemas de apoio à decisão. Vital (et al., 2019, p. 743) utilizaram técnicas de aprendizagem não supervisionada, através de algoritmos de *clustering* (hierárquico e k-means), para analisar o desempenho de estudantes. Os teóricos Panic e Lozanov-Crvenkovic (2018) usaram modelos financeiros e regressão linear para formular modelos de predição do comportamento de indicadores relacionados ao desempenho de alunos no ensino superior. Moscoso-Zea (et al., 2019, p. 38778) propuseram uma infraestrutura de *Business Intelligence & Analytics* (BI&A) e gestão do conhecimento para um *Datawarehouse* educacional com repositório para digitalização do conhecimento e análise de componentes organizacionais como pessoas, processos e tecnologias. Salaki e Ratnam (2018) focaram na importância de adotar um processo ágil de *analytics* no desenvolvimento de *Business Intelligence* e *Datawarehouse* em instituições de ensino superior. Por fim, Zhu (2018) apresentou um sistema de apoio à decisão dirigido a dados como uma plataforma de auxílio do processo de tomada de decisão educacional.

3 METODOLOGIA

Nesta seção, são apresentados os procedimentos metodológicos utilizados no desenvolvimento desta pesquisa. Quanto ao objetivo, a abordagem caracteriza-se como uma pesquisa exploratória, visando contribuir para o esclarecimento das questões abordadas inicialmente e proporcionando maiores informações acerca do tema. Em relação ao problema, a pesquisa apresenta abordagem qualitativa e quantitativa, apoiando-se em conceitos apresentados em pesquisa bibliográfica e baseada, também, na opinião de especialistas membros da equipe do projeto. Foi realizado um estudo de caso, no qual dados primários de Educação foram coletados e analisados. Para a análise qualitativa, foi aplicado um questionário que visou averiguar o nível de usabilidade do sistema. O questionário foi baseado no método *System Usability Scale* (SUS). (LEWIS; SAURO, 2009; BANGOR; KORTUM; MILLER, 2008).

Inicialmente, na primeira etapa metodológica, foi realizado um estudo detalhado sobre os indicadores descritos no documento do Projeto Integrar (TCU, 2020). O intuito dessa primeira etapa foi entender e avaliar quais desses indicadores seriam essenciais para compor os painéis, de acordo com o contexto do Estado do Acre. A coleta de dados deu-se a partir de uma amostra de dados educacionais do referido Estado, no período de 2013 a 2019. Esta coleta de dados ocorreu a partir do banco de dados das diversas fontes como: Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) e o Sistema de Informações sobre Orçamentos Públicos em Educação (SIOPE). Foi então modelado o banco de dados e executado o processo de Extração, Transformação e Carga (em inglês, *Extract, Transform and Load* - ETL) dos dados. Em seguida foi elaborada a arquitetura da solução de BI aqui proposta e, por fim, foram desenvolvidos os painéis de BI para os indicadores escolhidos. A seguir, descrevemos cada uma das etapas metodológicas utilizadas.

3.1 COLETA DE DADOS E O PROCESSO DE ETL

Com as fontes definidas e a base de dados modelada, foi realizado um dos processos mais importantes na construção de um sistema de BI, que é a etapa de *Extract, Transform and Load* (ETL). (KIMBALL; CASERTA, 2004, p. 3).

ETL é um processo de extração, transformação e carregamento de dados de diversas fontes para uma base de dados. A etapa de coleta de dados pode ser originada de variadas fontes de dados heterogêneas como planilhas, arquivos *JavaScript Object Notation* (JSON), bancos de dados, páginas web, dentre outros. A etapa de transformação consiste na limpeza e tratamento dos dados, que são advindos de diversas fontes, além de operações de modificação e cruzamento de informações. A etapa de carregamento consiste na inserção dos dados no banco de dados, que abriga grandes quantidades de informações das diversas fontes utilizadas.

O primeiro passo do ETL foi a coleta dos dados das fontes oficiais definidas na etapa de estudo de indicadores e fontes. Um dos grandes facilitadores foram os formatos de arquivos que os dados estavam disponibilizados, geralmente arquivos *Excel* (XLSX) e *Comma-separated Values* (CsV). Com os dados extraídos do respectivo formato de arquivos, foi realizada uma etapa de pré-processamento, com o intuito de eliminar ruídos das informações, como, por exemplo, a ocorrência de valores nulos, vazios e inconsistentes, além de um processo de filtragem, pois para os painéis, o foco foi as informações do Estado do Acre.

Utilizando a linguagem Python, e bibliotecas como *pandas*⁹ e *pyodbc*¹⁰, o pré-processamento dos dados, além das operações de *Data Manipulation Language* (DML) voltadas para o banco de dados, foram implementadas, removendo o trabalho

9 Disponível em: <https://pandas.pydata.org/docs>.

10 Disponível em: <https://github.com/mkleehammer/pyodbc/wiki>.

manual de processar informações advindas dos arquivos originários. Os scripts Python desenvolvidos nesta etapa foram essenciais para a otimização de tempo e robustez no processo de ETL, principalmente no processo de leitura de dados advindos dos arquivos carregados das diversas fontes, que por vezes, continham um grande volume de informações. Sabe-se, pois, que a manipulação destas informações de forma manual dificultaria o tratamento coerente de todo conteúdo dos arquivos.

3.2 FONTES DE DADOS

Os dados primários educacionais estão disponíveis em diversas fontes. As principais fontes de dados utilizadas foram:

- INEP¹¹: que possui uma seção de dados abertos com séries históricas de diversos indicadores, formada principalmente por metadados, a exemplo dos dados do Censo Escolar.
- IBGE¹²: utilizada para informações geográficas e de contexto. As informações foram buscadas diretamente do Instituto Brasileiro de Geografia e Estatística (IBGE), que fornece informações sobre os municípios e suas características.
- SIOPE¹³: contém dados sobre orçamentos e contas públicas vinculadas à educação, disponibilizados através do portal do Fundo Nacional de Desenvolvimento da Educação (FNDE), via Sistema de Informações sobre Orçamentos Públicos em Educação (SIOPE).

3.2.1 Scripts Python

De posse dos arquivos das fontes primárias, foram desenvolvidos scripts em Python, com três funções primordiais: ler os dados dos arquivos originários, manipular e transformar alguns desses dados, caso fosse necessário e, por fim, conectar com o banco de dados, fazendo manipulações nas tabelas, principalmente scripts DML, mais voltados para inserção de dados.

O processo de leitura e transformação dos dados foi feito com o auxílio da Biblioteca Pandas que fornece ferramentas de análise de dados e estrutura de dados de alta performance, onde é possível ler arquivos extensos e manipular informações neles contidas. Já para conectar com o banco de dados, foi utilizada a Biblioteca Pyodbc que facilita a conexão e manipulação direta com bancos de dados relacionais, simplificando o processo de acesso ao banco de dados. Os scripts desenvolvidos estão disponíveis no site repositório do projeto¹⁴.

11 Disponível em: <https://www.gov.br/inep/pt-br/aceso-a-informacao/dados-abertos>.

12 Disponível em: <https://www.ibge.gov.br/>.

13 Disponível em: https://www.fnde.gov.br/fnde_sistemas/siope.

14 Disponível em: https://github.com/pedrofwalkerley/scripts_tcc_indicadores_educacionais_Acre/tree/main/scripts.

3.3 MODELAGEM DO BANCO DE DADOS DE INDICADORES EDUCACIONAIS

Após o estudo do material fornecido pelo Projeto Integrar, que serviu de base para o desenvolvimento deste trabalho, foi iniciado o processo de modelagem do banco de dados. O banco de dados representa o ponto central no processo analítico posteriormente provido pelos painéis informativos. Nesta etapa, foram executados processos de SQL *Data Definition Language* (DDL), com criação de tabelas e colunas, com os respectivos índices e *constraints*.

Depois de feito um estudo detalhado sobre as fontes de dados, e o que seria aproveitado das mesmas, foi projetado o esquema do banco de dados a ser utilizado na construção dos painéis. Devido aos dados serem bem-estruturados e o volume deles não ser *big data* (o tamanho do banco de dados é de aproximadamente 145 MB), optou-se por utilizar um banco de dados relacional, que é uma solução simples e eficiente para o armazenamento de informações com essas características. Foi escolhido o Microsoft SQL Server que é um SGBD relacional bastante utilizado no mercado. Esse sistema situa-se no primeiro quadrante do quadrante mágico do Gartner¹⁵ sendo, portanto, uma das ferramentas mais atrativas e utilizadas no mercado, por ter uma interface amigável, bom desempenho e prover facilidades de integração com o Power BI.

No primeiro momento, foram selecionados dez indicadores e três entidades voltadas para os filtros da interface. Os componentes (tabelas) do esquema do banco de dados estão descritos a seguir:

- ADEQUACAO_FORMACAO_DOCENTE: porcentagem de docentes em cada um dos níveis de adequação de formação.
- ATENDIMENTO_EDUCACAO_INFANTIL: quantidade de matrículas nas creches e pré-escolas.
- AVALIACAO_NAC_ALFABETIZACAO: desempenho nas avaliações de escrita, leitura e matemática.
- DISTORCAO_IDADE_SERIE: taxas de distorção idade-série nas diferentes etapas de ensino.
- ESFORCO_DOCENTE: porcentagem de docentes em cada nível da variável categórica que indica o nível de esforço do docente.
- LOCALIDADE_ENSINO: localidade do ensino (rural ou urbana).
- MEDIA_ALUNOS: média de alunos por turma em cada etapa de ensino.
- MUNICIPIO: localidade, podendo ser município, microrregiões e mesorregiões.
- PROVA_BRASIL: resultados nas provas de Matemática e Língua Portuguesa da Prova Brasil.

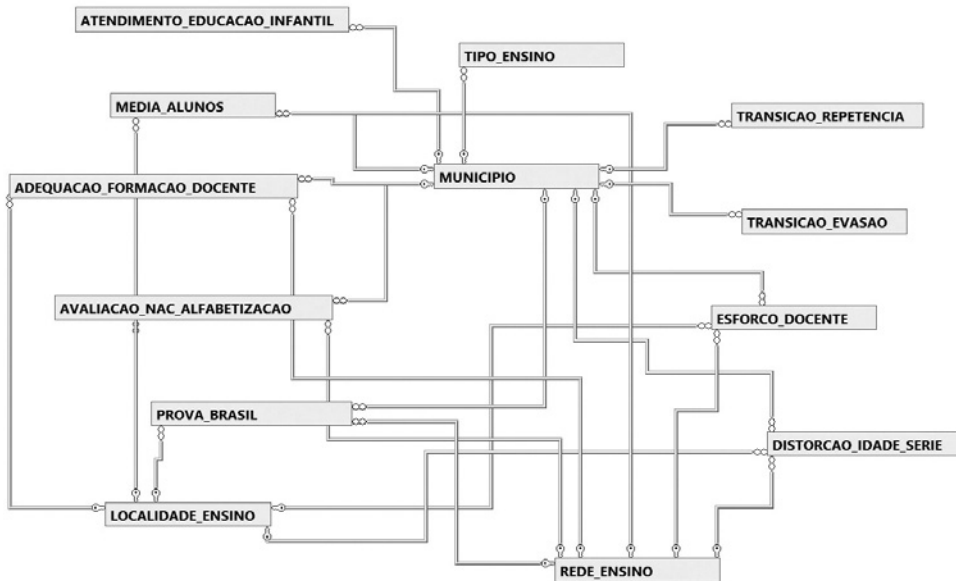
¹⁵ Disponível em: <https://azure.microsoft.com/en-us/resources/gartner-2020-magic-quadrant-for-cloud-database-management-sys>.

- REDE_ENSINO: Rede de ensino, podendo ser: Municipal, Estadual, Federal e Privada.
- TIPO_ENSINO: tipo de ensino escolar (parcial, integral).
- TRANSICAO_EVASAO: taxas de evasão nas diferentes etapas de ensino.
- TRANSICAO_REPETENCIA: taxas de repetência nas diferentes etapas de ensino.

O relacionamento entre as entidades é formado por três indicadores centrais: município, rede e localidade. O indicador município pode ser estendido para os subindicadores de microrregião e mesorregião. Essas subcategorias de localidade, vinculadas aos municípios em sistemas de BI, podem ser exploradas por operações conhecidas como *drill down* e *drill up*, que aumentam ou diminuem a granularidade das informações dispostas no painel, fornecendo uma visão mais ou menos detalhada e vinculada a um dado filtro. O indicador rede refere-se à Rede de ensino pública ou privada, e tem como subgrupo de filtros as sub-redes: municipal, estadual e federal. Por fim, o indicador localidade diz respeito à localização do ambiente escolar, apontando se a escola encontra-se no meio rural ou urbano.

Os três componentes mencionados anteriormente são essenciais no processo comparativo, possibilitando a introdução de relacionamentos cruzados entre os indicadores disponíveis, além de filtrações dinâmicas. A figura 1, a seguir, apresenta o diagrama relacional do banco de dados modelado.

Figura 1 - Diagrama Relacional do Banco de Dados de Indicadores da Educação Básica

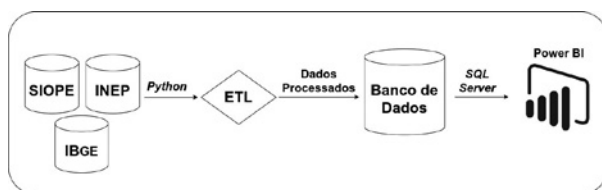


Fonte: Elaborado pelo autor (2021).

3.4 ARQUITETURA DO BI DE EDUCAÇÃO

A arquitetura da solução de BI de Educação, mostrada na figura 2, consiste basicamente em três camadas: o processo de ETL (*Extract, Transform and Load/* Extração, Transformação e Carga) dos dados originais; a implementação de um banco de dados relacional e a criação do *front end* de BI com os indicadores educacionais, organizados em painéis interativos.

Figura 2 - Arquitetura da solução



Fonte: Elaborado pelo autor (2021).

O processo de ETL foi realizado através de scripts escritos em Python, que acessaram os dados originais em planilhas ou bancos de dados relacionais, realizando transformações nos dados e fazendo a carga destes no banco de dados.

O banco de dados foi implementado utilizando o sistema de gerenciamento de banco de dados (SGBD) Microsoft SQL Server 2019¹⁶. Esse SGBD possui uma ferramenta facilitadora, o SQL Server Management Studio¹⁷, um ambiente integrado com interface gráfica para gerenciar infraestruturas sql, como realizações de consultas e criação de diagramas de bancos de dados.

Para o *front end*, foi utilizado o Power BI, que é um framework para desenvolvimento de painéis e relatórios voltados para a área de BI. Esse framework provê uma grande quantidade de funcionalidades para exibição dos indicadores. Além disso, o Power BI fornece suporte à linguagem *Data Analysis Expressions* (DAX), uma linguagem de expressão de fórmulas que permite a criação de consultas dinâmicas, cálculos, dentre outras funcionalidades, integradas diretamente com os *dashboards* desenvolvidos na ferramenta.

3.5 PROJETO DOS DASHBOARDS

Com os indicadores definidos e os dados carregados no banco de dados foi necessário um trabalho de *User Experience* (UX) e *User Interface* (UI), pensando no *front end* da aplicação. Por se tratar de um número grande e diversificado de indicadores foi definida, nesta etapa, a estrutura da interface e organização dos seus componentes. O agrupamento de indicadores seguiu o que foi proposto originalmente pelo Projeto Integrar, que recomendou a separação dos indicadores em dois grupos: indicadores de acesso

¹⁶ Disponível em: <https://www.microsoft.com/pt-br/sql-server/sql-server-downloads>.

¹⁷ Disponível em: <https://docs.microsoft.com/pt-br/sql/ssms/sql-server-management-studio-ssms>.

à educação e indicadores de proficiência. Os indicadores de acesso dizem respeito à viabilidade, facilidade e condições em geral que possibilitam o ingresso, a permanência e o progresso de um estudante nas inúmeras etapas de ensino que compõem a matriz curricular da Educação Básica. Por sua vez, os indicadores de proficiência estão atrelados aos resultados obtidos por esses estudantes em avaliações censitárias, como a Prova Brasil e a Avaliação Nacional da Alfabetização (descontinuada em 2016), que permitem a dedução de *insights* sobre os índices de aprendizagem do corpo discente. Com essa definição, os usuários finais podem seguir uma sequência lógica no momento de visualização dos painéis, facilitando o entendimento das informações dispostas.

Como o foco do trabalho são os indicadores educacionais, foram excluídos do planejamento dos painéis os indicadores de contexto e indicadores de financiamento, ambos propostos pelo Projeto Integrar.

Definido o layout da interface, deu-se início ao desenvolvimento dos painéis. Para a homologação de cada versão finalizada, utilizou-se uma metodologia baseada em Provas de Conceito (POC), devido ao enorme número de possibilidades de funcionalidades fornecidas pelo Power BI. À medida que as POCs eram desenvolvidas, as mesmas eram submetidas à análise de um usuário independente, que tinha a função de opinar sobre as visualizações escolhidas, ponto fundamental para garantir uma melhor experiência do usuário.

4 RESULTADOS

Os dashboards desenvolvidos contemplaram os dois grupos definidos no Projeto Integrar: os indicadores de acesso e os indicadores de proficiência e foram implementados os seguintes painéis:

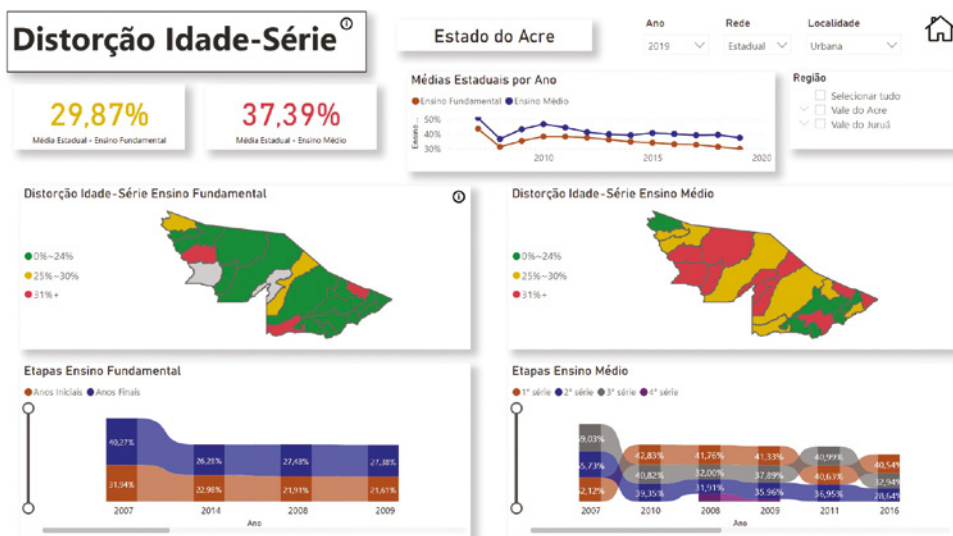
- Painel Distorção Idade-Série: informações sobre as taxas de distorção idade-série na Educação Básica;
- Painel Esforço Docente: informações sobre as taxas de cada um dos níveis de esforço docente;
- Painel Adequação da Formação Docente: informações sobre as taxas de formação docente adequada para cada um dos grupos;
- Painel Prova Brasil: resultados e comparativos dos resultados nas avaliações de Língua Portuguesa e Matemática da Prova Brasil;
- Painel Avaliação Nacional da Alfabetização: resultados das avaliações aplicadas nas modalidades de leitura, escrita e matemática;
- Painel Evasão Escolar: taxas de evasão escolar nas diferentes etapas de ensino;
- Painel Repetência Escolar: taxas de repetência escolar nas diferentes etapas de ensino;
- Painel Atendimento à Educação Infantil: informações sobre a quantidade de alunos matriculados nas Creches e Pré-escolas do Estado;

- Painel Média de Alunos por Turma: informações sobre a média de alunos por turma nas etapas de ensino da Educação Básica;
- Painel Ensino Integral: informações sobre a oferta de ensino integral para as etapas de ensino da Rede pública.

4.1 PAINEL DISTRORÇÃO IDADE-SÉRIE

O indicador de distorção idade-série acompanha a taxa de alunos que têm idade acima da esperada para a etapa de ensino em que estão matriculados. O painel desenvolvido para esse indicador é composto por um conjunto de filtros, município, Rede de ensino, localidade de ensino e ano. Além disso, como existem dados de todo o Estado do Acre, é possível realizar operações de *drill up* e visualizar dados das mesorregiões e microrregiões do Estado, conforme apresentado na figura 3.

Figura 3 - Painel do indicador de Distorção Idade-Série



Fonte: Elaborado pelo autor (2021).

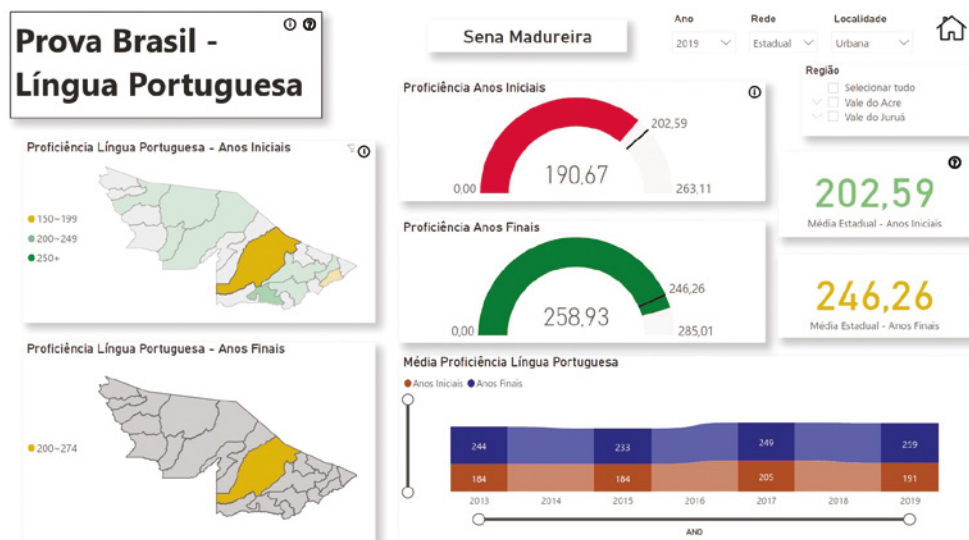
Na figura 3, os dois indicadores abaixo do título trazem informações sobre a média estadual da taxa de distorção idade-série, para o Ensino Fundamental (superior esquerdo) e Ensino Médio (superior direito), onde sua coloração muda de acordo com o valor da taxa. O gráfico de linhas acima dos mapas faz um comparativo entre as médias estaduais das taxas de distorção idade-série ao longo dos anos, para o Ensino Fundamental e Médio, podendo assim ter uma visão geral sobre as taxas para cada etapa de ensino ao longo dos anos. Os dois mapas trazem as taxas de distorção idade-série vinculadas ao Ensino Fundamental (mapa à esquerda) e ao Ensino Médio (mapa à direita) e, de acordo com a cor, é possível se ter uma noção se o nível do indicador está crítico (vermelho),

regular (amarelo) ou ideal (verde) para cada um dos municípios. Por fim, os gráficos de faixas no final do painel mostram um comparativo entre as taxas de distorção nos anos iniciais e finais do Ensino Fundamental (inferior esquerdo) e ao lado um comparativo entre as três ou quatro séries do Ensino Médio (inferior direito), mostrando uma série histórica com dados vinculados a um município, caso seja ativado o filtro do mesmo.

4.2 PAINEL PROVA BRASIL (LÍNGUA PORTUGUESA)

A Prova Brasil é uma avaliação censitária, aplicada bianualmente, que contempla provas de Língua Portuguesa e Matemática e tem como principais objetivos medir a qualidade do ensino público, bem como observar como está o nível de aprendizagem dos alunos. A figura 4 exibe este painel com mapas e gráficos referentes ao desempenho na Prova Brasil no ano de 2019, mais especificamente na avaliação de Língua Portuguesa.

Figura 4 - Painel do indicador de Prova Brasil para a avaliação de Língua Portuguesa/2019



Fonte: Elaborado pelo autor (2021).

Os mapas trazem os valores das médias municipais. O mapa superior é referente aos anos iniciais do Ensino Fundamental, já o mapa inferior é referente aos anos finais. A cor em cada município segue uma escala definida por uma equipe científica do movimento Todos pela Educação¹⁸. Mais à direita, na figura 4, os indicadores trazem informações sobre as médias estaduais na prova, também diferindo pela etapa de ensino. A coloração do valor da média segue a mesma escala dos mapas temáticos.

Os gauges, no centro do painel, indicam se o município alcançou a média estadual para sua etapa de ensino, no caso, os anos iniciais e finais. Caso o valor da média municipal ultrapasse ou seja igual à média estadual, a cor do gauge será

18 Disponível em: <https://todospelaeducacao.org.br/>.

verde, caso o valor fique abaixo da média, o preenchimento fica vermelho. Ainda na figura 4, para o município acreano de Sena Madureira, o resultado dos anos iniciais do Ensino Fundamental ficou abaixo da média estadual; já nos anos finais, o valor ficou maior que a média do estado.

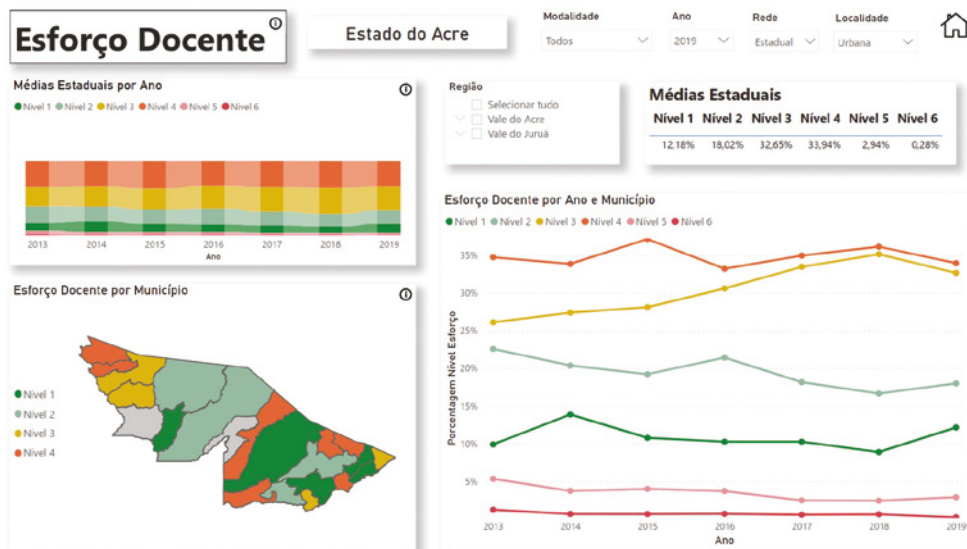
O gráfico de faixas (inferior direito) é um comparativo das médias de resultados municipais entre os anos iniciais e finais, e tem um caráter temporal, sendo possível uma avaliação dos resultados de anos anteriores.

4.3 PAINEL ESFORÇO DOCENTE

Este indicador é calculado utilizando dados disponíveis no Censo Escolar da Educação Básica e leva em consideração quatro pontos, segundo nota técnica¹⁹ do INEP: número de escolas que o (a) docente atua, número de turnos de trabalho, número de alunos atendidos e o número de etapas de ensino que o (a) mesmo (a) leciona.

Existem seis níveis de esforço docente que, em ordem crescente, representam a carga de trabalho de um (a) docente, sendo o nível 1, o de menos esforço e o nível 6, o mais crítico. A figura 5 mostra o painel deste indicador. O primeiro gráfico (superior esquerdo) traz os valores das médias estaduais ao longo dos anos, e representa a porcentagem de docentes em cada um dos níveis de esforço docente. O mapa mostra qual dos níveis teve uma maior porcentagem de docentes em cada um dos municípios.

Figura 5 - Painel do indicador de Esforço Docente



Fonte: Elaborado pelo autor (2021).

19 Disponível em: https://download.inep.gov.br/informacoes_estatisticas/indicadores_educacionais/2014/docente_esforço/nota_tecnica_indicador_docente_esforço.pdf.

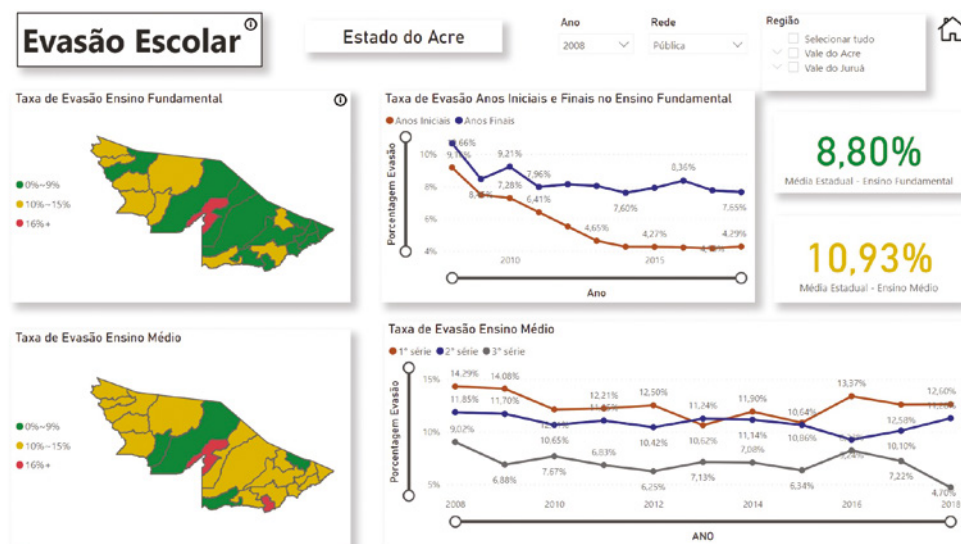
No lado superior direito da figura 5, a tabela traz os valores das médias estaduais para cada um dos níveis de esforço docente, vinculadas ao ano selecionado no filtro da tela. O gráfico abaixo da tabela mostra uma série histórica das porcentagens de cada nível de esforço docente para um determinado município, respeitando os filtros de Rede e localidade de ensino.

4.4 PAINEL EVASÃO ESCOLAR

A evasão escolar faz referência ao aluno que foi reprovado em um dado ano, ou abandonou a escola e, no ano seguinte, não fez a matrícula novamente para dar continuidade aos estudos.

O painel para este indicador, mostrado na figura 6, traz dois mapas temáticos, o superior está vinculado às porcentagens de evasão no Ensino Fundamental e o inferior está vinculado às porcentagens de evasão no Ensino Médio, onde a cor que preenche o município muda de acordo com o valor daquela localidade específica, indo de desejável (verde) a crítico (vermelho). Os gráficos de linhas trazem séries históricas sobre os valores deste indicador, também voltados para o Ensino Fundamental e Médio. No gráfico superior, é mostrada uma comparação entre os anos iniciais e finais do Ensino Fundamental, já no gráfico inferior um comparativo entre os valores de evasão escolar para a 1ª, 2ª e 3ª séries do Ensino Médio.

Figura 6 - Painel do indicador de Evasão Escolar



Fonte: Elaborado pelo autor (2021).

Os dois indicadores mais à esquerda da figura 6 representam os valores das médias estaduais dos índices de evasão para o Ensino Fundamental e Médio, mudando sua coloração de acordo com a escala de valores presente na legenda do mapa. Por fim, as funcionalidades do painel podem ser filtradas por município, microrregião, mesorregião, Rede de ensino e ano.

Para o filtro por município, podem ser executadas operações de *drill down* e *drill up*, permitindo uma filtragem por microrregiões e mesorregiões do Estado do Acre.

5 AVALIAÇÃO DE USABILIDADE

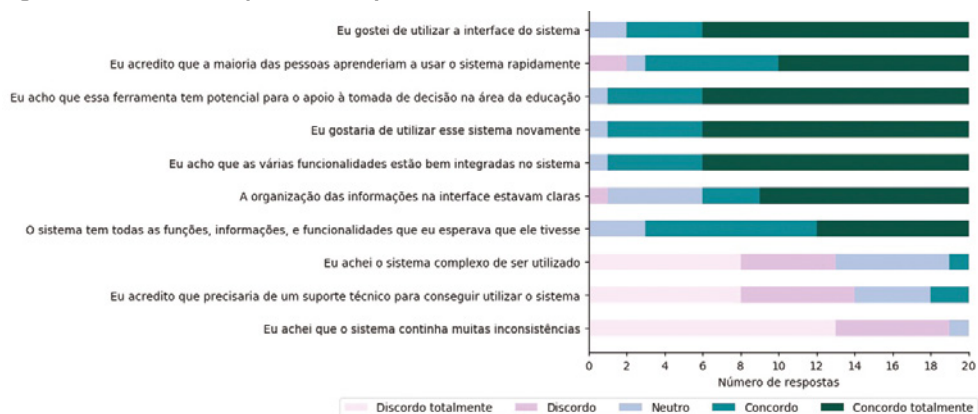
O processo de avaliação de usabilidade da ferramenta de BI proposta consistiu na aplicação de um questionário que visou averiguar o nível de usabilidade do sistema. O questionário foi baseado no método *System Usability Scale* (SUS). (LEWIS; SAURO, 2009; BANGOR; KORTUM; MILLER, 2008). O SUS é um questionário que se propõe avaliar a usabilidade de sistemas. O questionário é composto por 10 afirmativas e 5 opções de respostas que indicam o nível de concordância do usuário com o que está sendo afirmado.

Para a avaliação do painel foi utilizada uma escala *Likert* (JOSHI et al., 2015, p. 396), com o intuito de observar o grau de concordância dos usuários em relação as afirmativas dispostas no questionário. Para cada afirmativa o usuário poderia selecionar um número de 1 a 5, onde 1 significa discordância total com o que está sendo afirmado e 5, concordância total. O valor 3, portanto, seria um valor neutro.

O processo de avaliação contou com 20 indivíduos que, anonimamente, responderam ao questionário. Dentre eles, alguns eram gestores educacionais. Para a avaliação, foi disponibilizada uma versão do sistema, acompanhada de um vídeo instrucional demonstrando como se utiliza a ferramenta, sobretudo as consultas dinâmicas que o Power BI possibilita realizar.

A figura 7 traz informações sobre as perguntas do formulário, bem como as respostas para cada um dos itens. Pode-se concluir que o resultado foi excelente, considerando o alto grau de concordância dos usuários com o que diziam as afirmações do questionário, com exceção dos últimos três itens mostrados na figura 7, que tratavam sobre a complexidade e inconsistências no sistema. Nessas afirmativas, o grau de discordância foi alto, o que representa um ótimo indício.

Figura 7 - Resultados do questionário aplicado



Fonte: Elaborado pelo autor (2021).

Os resultados para duas afirmativas presentes no questionário representam bem o potencial da solução para atender os dois principais públicos, ao qual a solução foi desenvolvida: usuários cidadãos e gestores. A primeira afirmativa diz respeito ao item 3 (“*Eu acho que essa ferramenta tem potencial para o apoio à tomada de decisão na área da educação*”) que afirma que o sistema tem potencial para auxiliar processos de tomada de decisão no âmbito educacional, esse item alcançou uma média de respostas de 4,65 indicando a validade para utilização da ferramenta por gestores educacionais. A segunda afirmativa representada pelo item 8, faz referência a complexidade do sistema (“*Eu achei o sistema complexo de ser utilizado*”) e obteve como média de resposta o valor 2, indicando uma alta discordância com o que está sendo dito, demonstrando, portanto, ser um sistema simples de ser utilizado para ambos os públicos.

Diante das respostas dos usuários conclui-se que a ferramenta de BI Educacional proposta atende aos dois públicos-alvo para os quais a ferramenta foi proposta: usuários cidadãos e gestores em educação.

6 CONSIDERAÇÕES

Este artigo abordou a construção de um sistema de apoio à decisão, baseado em *Business Intelligence*, que provê uma ferramenta analítica para indicadores educacionais, tendo como estudo de caso o Estado do Acre.

O principal objetivo desta pesquisa foi desenvolver painéis com narrativas de fácil entendimento para cidadãos e gestores, assim como trazer informações cruciais para a melhor gestão da Educação Básica nos municípios acreanos.

Uma das limitações da ferramenta proposta está na dificuldade de obtenção de dados, muitos dos quais contemplam informações imprecisas e séries históricas de

dados muito curtas, que impedem uma análise mais profunda da situação, bem como a realização de prognósticos, utilizando-se de técnicas de aprendizado de máquina.

A avaliação da ferramenta por usuários trouxe um feedback muito positivo, que comprovou a eficácia e importância da solução proposta. Os painéis conseguiram levar com clareza informações para os usuários finais, que foram aptos a compreender a situação educacional no Estado do Acre. Desse modo, o BI Educacional mostra-se como um importante instrumento para auxílio de gestores educacionais do referido Estado, podendo ser uma base para futuras pesquisas quantitativas experimentais.

Recomenda-se para trabalhos futuros, a incorporação de um modelo ou algoritmo, para direcionamento de gestores, sobre quais ações serão mais eficazes para o alavanque da Educação Básica em cada localidade específica. Através da análise de indicadores, como os expostos neste documento, e cruzamento de informações, utilizando técnicas de *Machine Learning*, o sistema, utilizando um modelo de aprendizagem, seria capaz de recomendar o direcionamento de recursos para um determinado setor da educação, otimizando o processo de tomada de decisão dos gestores, principalmente em função da grande quantidade de variáveis que se deve levar em consideração para esses processos decisórios.

Recomenda-se, também, a criação de um Modelo Educacional Nacional, também utilizando como base municípios que obtiveram resultados satisfatórios, vinculados à escolarização e aprendizado. Para tal modelo, técnicas de *Machine Learning* seriam ideais para análise de dados educacionais, com objetivo de encontrar uma solução ideal para impulsionar índices educacionais em todo território nacional.

REFERÊNCIAS

BANGOR, A.; KORTUM, P. T.; MILLER, J. T. An empirical evaluation of the system usability scale. **International Journal of Human-Computer Interaction**, Taylor Francis, v. 24, n. 6, 2008, p. 574-594. Disponível em: <https://doi.org/10.1080/10447310802205776>. Acesso em: 17 abr. 2021.

DIEPEN, P. V.; BREDEWEG, B. Performance indicators for online secondary education: a case study. *In*: SPRINGER. **Benelux Conference on Artificial Intelligence**. 2016, p. 169-177.

JOSHI, A. [*et al.*]. Likert scale: Explored and explained. **British Journal of Applied Science Technology**, v. 7, 2015, p. 396-403.

KIMBALL, R.; CASERTA, J. **The Data Warehouse ETL Toolkit: Practical Techniques for Extracting, Cleaning, Conforming and Delivering Data**. Hoboken, NJ, USA: John Wiley Sons, Inc., 2004.

LEWIS, J. R.; SAURO, J. The factor structure of the system usability scale. *In*: KUROSU, M. (Ed.). **Human Centered Design**. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009. p. 94-103.

MOSCOSO-ZEA, O. [et al.] A hybrid infrastructure of enterprise architecture and business intelligence & analytics for knowledge management in education.

IEEE Access, v. 7, 2019, p. 38778. Disponível em: <https://doi.org/10.1109/ACCESS.2019.2906343>. Acesso em: 9 abr. 2021.

PANIC, A. A.; LOZANOV-CRVENKOVIC, Z. Panel regression analysis of the financial and performance indicators in higher education in central east europe. *In*: 16th IEEE International Symposium on Intelligent Systems and Informatics, **SISY 2018**. Subotica, Serbia, September 13-15, 2018. IEEE, 2018, p. 317-320. Disponível em: <https://doi.org/10.1109/SISY.2018.8524766>. Acesso em: 9 abr. 2021.

PETIOT, G. Merging information using uncertain gates: An application to educational indicators. *In*: MEDINA, J. et al. (Ed.). **Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations - 17th International Conference, IPMU 2018, Cádiz, Spain, June 11-15, 2018, Proceedings, Part I**. Springer, 2018. (Communications in Computer and Information Science, v. 853), p. 183–194. Disponível em: https://doi.org/10.1007/978-3-319-91473-2_16. Acesso em: 09 abr. 2021.

REZGUI, A.; GÓMEZ, J. M.; MAAOUIA, R. B. Kpi-based decision evaluation system to enhance qmss for higher educational institutes. **Int. J. Decis. Support Syst. Technol.**, v. 9, n. 2, 2017, p. 39-55. Disponível em: <https://doi.org/10.4018/IJDSST.2017040103>. Acesso em: 9 abr. 2021.

SALAKI, R. J.; RATNAM, K. A. Agile analytics: Applying in the development of data warehouse for business intelligence system in higher education. *In*: ROCHA, Á. et al. (Ed.). **Trends and Advances in Information Systems and Technologies**. v. 1, WorldCIST'18, Naples, Italy, March 27-29, 2018. Springer, 2018. (Advances in Intelligent Systems and Computing, v. 745), p. 1038–1048. Disponível em: https://doi.org/10.1007/978-3-319-77703-0_101. Acesso em: 9 abr. 2021.

SHARDA, R.; DELEN, D.; TURBAN, E. **Business Intelligence, Analytics, and Data Science: a Managerial Perspective**. Pearson, 2018.

TRIBUNAL DE CONTAS DA UNIÃO. **Projeto Integrar**: propostas para o fortalecimento do controle externo de políticas públicas descentralizadas. Tribunal de Contas da União, Organização para Cooperação e Desenvolvimento Econômico. Brasília, 2020.

VITAL, T. P. [et al.]. Student performance analysis with using statistical and cluster studies. *In*: **Soft Computing in Data Analytics**. Springer, 2019. p. 743-757.

WEEDA, R. [et al.]. Towards an assessment rubric for eipe tasks in secondary education: Identifying quality indicators and descriptors. *In*: FALKNER, N.; SEPPÄLÄ, O. (Ed.). **Koli Calling '20: 20th Koli Calling International Conference on Computing Education Research**. Koli, Finland, November 19-22, 2020. ACM, 2020. Artigo 30, p. 1-10. Disponível em: <https://doi.org/10.1145/3428029.3428031>. Acesso em: 9 abr. 2021.

ZHU, Y. A data driven educational decision support system. **International Journal of Emerging Technologies in Learning (IJET)**, v. 13, n. 11, 2018 p. 4-16.