
Fatores extraescolares e a qualidade da educação na rede estadual paulista

Carla Costa de Freitas Soares¹, Luiz Antonio Lima Rodrigues², Seiji Isotani³

Resumo

Os setores que amparam as ações pedagógicas são múltiplos, no entanto, não são periodicamente vistos ou analisados como indicadores que viabilizam o aprimoramento do sistema e a qualidade da educação. Nesse sentido, o presente estudo objetivou identificar a relação entre fatores extraescolares e os resultados educacionais da rede estadual paulista, realizando uma pesquisa quantitativa-descritiva, com a aplicação de ferramentas de Mineração de Dados Educacionais e aspectos extraescolares das 91 diretorias regionais de ensino. Além disso, com o uso do algoritmo Apriori, buscou por regras de associação relacionadas ao indicador de desempenho. Tais resultados propiciaram novos conhecimentos acerca da influência, em especial, das regiões administrativas e do quantitativo de escolas no desempenho dos órgãos regionais.

Abstract

The sectors that support the pedagogical actions are multiple, however, they are not periodically seen or analyzed as indicators that make possible the improvement of the system and the quality of education. In this sense, the objective was to identify the relationship between extra-school factors and the educational results of the São Paulo state network, carrying out quantitative-descriptive research, with the application of educational data mining tools and extra-school aspects of 91 regional teaching. And, using the Apriori algorithm, find association rules related to the performance indicator - IDESP. These results provided new knowledge about the influence, in particular, of the administrative regions and the number of schools in the performance of regional bodies.

¹ Pós-Graduando(a) em Computação Aplicada à Educação, USP, <carlacosta@usp.br >.

² Co-orientador, USP, lalrodrigues@usp.br.

³ Orientador, USP, damascenodiego@alumni.usp.br.

Introdução

Dentro do âmbito educacional muitos são os estudos que envolvem os resultados educacionais e, conseqüentemente, a qualidade do ensino. Tais pesquisas problematizam acerca das diversidades de fatores, dentro e fora da escola, que atuam sobre o processo ensino e aprendizagem, os quais se realizam por meio de processos e práticas escolares, bem como por aspectos sociodemográficos, econômicos, organizacionais e institucionais.

Nesse viés, a oferta de uma educação de qualidade se apresenta como um desafio para os agentes executores, ou seja, as burocracias públicas. Azevedo (2014), ressalta que para entendermos a qualidade do ensino, é necessário sabermos que a mesma está relacionada a vários significados, e esses, por sua vez, estão atrelados a diferentes filosofias de ação que orientam a ação pública. Oliveira e Abrucio (2018), complementa ao citar que a burocracia de médio escalão, setor composto por cargos e posições que não estão localizados nem no topo nem na base, busca promover o diálogo entre as necessidades do público que atende com as prioridades estabelecidas pelo alto escalão.

Com base nas citações e por meio das informações coletadas na plataforma, Dados Abertos da Educação do Estado de São Paulo (2020), este estudo busca identificar relações entre os indicadores educacionais derivados do Índice de Desenvolvimento da Educação de São Paulo (IDESP) e as dimensões regionais, municipais e internas da burocracia de nível médio, neste caso, as diretorias regionais de ensino (DRE).

A partir desse objetivo, e por meio das ferramentas de Mineração de Dados Educacionais (MDE) - do inglês, *Educational Data Mining* – buscou-se as relações entre os indicadores e, por fim, a construção de novos conhecimentos, que transformados em evidências possam nortear decisões pedagógicas mais focadas e assertivas.

Dentro do contexto da MDE, trabalhos na literatura discorrem acerca das ferramentas de Mineração de Dados Educacionais para geração de conhecimento, como por exemplo, Ferreira (2015), ao utilizar os dados disponíveis no Censo Escolar da Educação Básica do Instituto Nacional de Estudos e Pesquisas (INEP) de 2014, para identificar as características pessoais e sociais que impactam na evasão escolar.

Seguindo a direção das citações acima, este estudo se valeu do método quantitativo-descritivo para identificar e analisar padrões de ocorrência através de combinações de indicadores. E, com base em regras de associação, entre os dados e seus valores (intervalos), buscou estabelecer características comuns e evidências de relações.

As seções para este trabalho foram organizadas da seguinte forma: A seção 2 discorre acerca da fundamentação teórica e estrutural do estudo; A seção 3 traz um amostral dos trabalhos relacionados; A seção 4 aponta a metodologia utilizada na tarefa de mineração; A seção 5 desenvolve a descrição dos resultados obtidos; A seção 6 analisa os dados e respectivos resultados e, por fim; A seção 7 expõe as conclusões obtidas através do estudo, além de propor temáticas para trabalhos futuros.

2. Fundamentação Teórica

Após análise da literatura, identificou-se que estudos referentes aos aspectos extraescolares da burocracia de nível médio foram pouco explorados, mesmo possuindo fundamental importância na qualidade da educação do sistema de ensino. Oliveira e Abrucio (2011) referem que boa parte das pesquisas desenvolvidas no Brasil centram-se na burocracia de alto escalão (os órgãos mais elevados) e como esta atua no processo de produção de políticas públicas e em relação à burocracia de nível de rua (unidades escolares).

Entretanto, a descoberta de padrões relevantes acerca da burocracia de nível médio, transformados em conhecimentos, podem subsidiar o desenvolvimento de práticas pedagógicas e tomadas de decisão que anteriormente não foram tratadas como objeto de estudo ou mecanismos de aprendizagem. Para tal empreitada, buscou-se as referências de Fayyad et al. (1996), onde a descoberta de conhecimentos em múltiplas bases - do inglês, *Knowledge Discovery in Database* - se mostra como um mecanismo propulsor para a exploração de informações úteis que geralmente não estão visíveis. Sua utilização proporciona o desenvolvimento de etapas ligadas à seleção, pré-processamento dos dados, transformação, mineração dos dados e avaliação (pós-processamento dos resultados obtidos na mineração dos dados), figura 2.1.

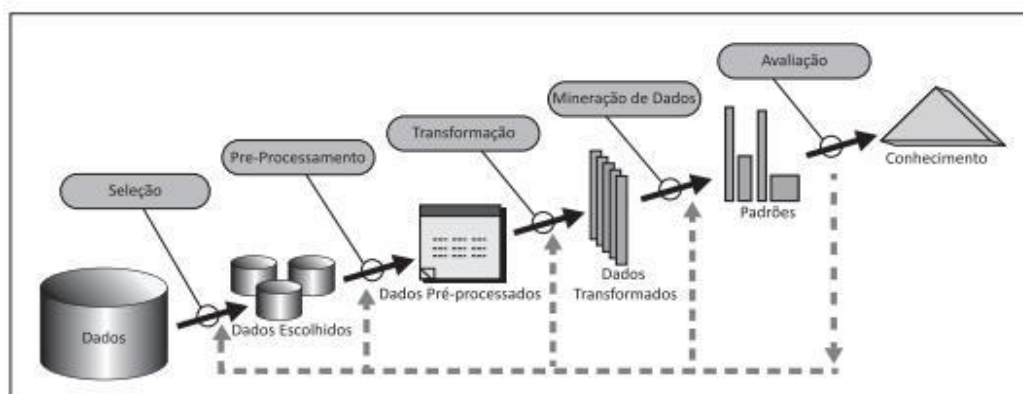


Figura 2.1. Etapas operacionais do processo KDD

Com base na Figura 2.1, observa-se o seguinte cenário: Os dados selecionados (valores referentes à medição) são integrados à fase de pré-processamento: captação, organização e tratamento dos dados. Em seguida, ocorre a transformação por meio da higienização e padronização e, subsequentemente, a aplicação do algoritmo na etapa de mineração. Já na fase de avaliação ou pós-processamento se realiza a análise dos resultados de forma quantitativa ou qualitativa, dependendo da tarefa escolhida.

Outro fator extremamente importante se caracteriza pela escolha da ferramenta, que deve levar em conta as especificidades da área educacional. Nesse contexto, o *software* de aprendizado de máquina Weka se destaca como uma boa escolha, pois pode ser acessado, segundo Hall et al. (2009), através de um código aberto em uma interface gráfica do usuário. É amplamente utilizado por profissionais com pouca ou nenhuma

habilidade na área computacional, permitindo que os valores possam ser aplicados à base de dados sem a necessidade de criação de códigos.

Os tipos de técnicas a serem utilizadas na mineração de dados são determinados pela escolha de uma tarefa e de um algoritmo, subordinados pelos atributos descritos na base de dados. A ferramenta Weka dispõe de algoritmos de *Data Mining*, tais como: regras de classificação, regras de agrupamento e regras de associação e, respectivamente, os algoritmos que fazem o processo de mineração dos dados.

Já os algoritmos são divididos pelas funções que desenvolvem, levando em consideração o objetivo da implementação. Entre as muitas possibilidades destaca-se positivamente a tarefa descritiva (padrões que possibilitam a descrição dos dados) em relação à procura de regras de associação com a aplicação do algoritmo de mineração Apriori.

2.1. Regras de associação

As regras de associação se caracterizam como uma parte da mineração de dados que se predispõem a análise de atributos, os quais são chamados de itens. Tem como pressupostos encontrar itens que implicam na presença de outros itens, ou seja, descobrir relações ou padrões frequentes entre o conjunto de dados.

São muitas as possibilidades entre os algoritmos de regras de associação, contudo, o algoritmo Apriori, proposto por Agrawal et al. (1993), é amplamente utilizado, pois constrói um padrão de comportamento com base nos cálculos de todas as combinações de itens e, por fim, a extração de subconjuntos de itens frequentes.

O processo de mineração dos dados com o uso do algoritmo Apriori ocorre por meio de dois passos: geração e poda. O primeiro passo se destina a varredura do arquivo, com o objetivo de criar um conjunto de itens. O segundo passo busca desconsiderar certas combinações que não atendam uma frequência mínima pré-fixada.

Assim, para aplicação da regra é necessário, antes de iniciar o processo de busca por associações, estabelecer dois parâmetros:

- Suporte mínimo - Quantidade mínima de transações analisadas que apresente a mesma regra, por exemplo: itens A e B.
- Confiança mínima - Quantidade mínima de transações que conste o item A e também o item B.

Outra tarefa extremamente importante se caracteriza através do processo de discretização, o qual ocorre por meio de uma técnica que consiste em transformar valores/itens numéricos em valores nominais ou discretos. Além disso, o processo também cria a aplicação de intervalos, ou seja, agrupamento de itens em categorias. Nessa etapa é fundamental que os agentes envolvidos tenham conhecimento técnico acerca dos itens, a fim de que os intervalos de cada item estejam bem distribuídos e representem a realidade ao qual pertencem.

3. Trabalhos Relacionados

Após pesquisas de artigos relacionados nas mais diversas fontes de informação, este estudo não identificou trabalhos que abordam, diretamente, o problema desta pesquisa. Contudo, cinco textos relacionados foram identificados e analisados, compondo o escopo desta seção.

O estudo de Dourado, Oliveira e Santos (2007) problematiza o conceito de qualidade da educação, considerando suas múltiplas significações e dimensões. Apresenta os fatores intra e extraescolares como fundamentais para a construção de uma escola de qualidade para todos e desenvolve uma análise articulada dos diferentes aspectos que interferem no processo de construção da aprendizagem.

Peixoto et al. (2017) abordam os fatores que interferem diretamente no fracasso escolar, bem como desenvolvem uma análise histórica de sua perpetuação. Compreendem que os fatores externos e internos que envolvem as escolas estão atrelados à relação do indivíduo com o mundo a sua volta e, fatos históricos, a própria escola e a reestruturação do sistema de ensino corroboram para o sucesso ou insucesso escolar. Além disso, a psicologia surge como mecanismo interventivo nessa situação, considerando as dinâmicas das unidades escolares, o contexto social e as especificidades da região como fatores que influenciam, de forma positiva e negativa, o processo ensino e aprendizagem.

Na dissertação de Miranda (2018) as demandas educacionais, o processo de informatização do ambiente escolar e o baixo desempenho dos alunos foram estudados. Para isso, utilizou-se a MDE como ferramenta de pesquisas e análise para a construção de perfis de alunos e identificação de fatores escolares e extraescolares com vistas a distinguir metodologias para a otimização do processo ensino e aprendizagem.

Barros (2018) em “O elo institucional regional da educação básica: um estudo das diretorias regionais de ensino paulista” debruça-se sobre a questão do alinhamento de processos necessários, entre os níveis institucionais implementadores, para o provimento de suporte adequado às escolas e o alcance de metas. Entre os eixos que abarcam o documento destaca-se o papel da burocracia de nível médio na execução do trabalho pedagógico e no alcance das metas projetadas pelo sistema de ensino.

Vasconcelos e Carvalho (2004) em “Aplicação de Regras de Associação para Mineração de Dados na *Web*” buscam por meio da mineração de dados e de regras de associação as correlações, padrões e tendências entre as informações da *Web*. Ademais, suscitam a coleta, o armazenamento dos dados e a capacidade de entendê-los e transformá-los como elemento preponderante para a construção de conhecimentos.

4. Metodologia

Buscando responder o problema desta pesquisa ao “Compreender como os fatores extraescolares, em grupos de diferentes gestões regionais, correlacionam-se no intuito de subsidiar a melhoria da qualidade da educação” este estudo se debruçou acerca de aspectos demográficos e estruturais, os quais influenciam os atores e as estruturas administrativas, pedagógicas e de desempenho das escolas.

Nesse contexto, a perspectiva de Fayyad et al. (1996) se tornou um elemento basilar, pois ao abordar a Mineração de Dados Educacionais como um processo que apresenta como entrada uma base de dados educacional e como saída um conhecimento, viabilizou o encontro de possíveis relações entre algumas variáveis extraescolares relacionadas às diretorias regionais de ensino e, conseqüentemente, construção de saberes e possíveis mudanças nas políticas públicas.

4.1. Seleção dos indicadores e criação da base de dados

A base de dados para este estudo se constituiu por meio da seleção de indicadores provenientes de levantamentos estatísticos e avaliativos da Educação Básica, com vistas nas 91 diretorias regionais de ensino pertencentes à rede. E, estão disponíveis para visualização através dos endereços eletrônicos dispostos nas referências.

Para a escolha dos indicadores estruturais e organizacionais levou-se em conta a estabilidade dos dados (não sofrem variações frequentes em relação à mudança de ano letivo ou diferentes gestões) e, em muitas situações, podem ser revistos e reestruturados. Já a escolha dos indicadores de rendimento segue as diretrizes da rede estadual paulista, visando o diagnóstico da qualidade da educação.

4.1.1. IDESP do Ensino Médio das DRE

O Índice de Desenvolvimento da Educação do Estado de São Paulo, Dados Abertos da Educação (2020), é o indicador que avalia a qualidade das escolas e diretorias regionais de ensino. É composto por dois critérios: desempenho dos alunos nos exames de proficiência do Sistema de Avaliação do Rendimento Escolar do Estado de São Paulo (SARESP), i.e., o quanto aprenderam; e fluxo escolar, i.e., em quanto tempo aprenderam.

Dentre os segmentos de ensino da rede estadual paulista, escolheu-se o Ensino Médio de forma estratégica, pois a trajetória do ciclo se apresenta como o maior desafio da Educação Básica. Além disso, buscando a verificação de indicadores que apresentem a fotografia real das DRE, construiu-se a média dos resultados educacionais do IDESP dos últimos 2 anos (2018 e 2019).

4.1.2. Divisões regionais das DRE

As diretorias regionais de ensino foram divididas segundo o mapa das regiões da Secretaria da Educação do Estado de São Paulo e seus respectivos polos administrativos: Capital, Grande São Paulo e Interior. No entanto, a fim de uma análise dos dados mais

assertiva e ajustada ao objetivo, desmembrou-se o Interior em dois novos blocos: Interior Norte e Interior Sul, tal como apresentado na Tabela 4.1.

Tabela 4.1. Polos Regionais

Junção dos polos regionais da Secretaria da Educação do Estado de São Paulo	Quantidade de diretorias regionais de ensino
Capital (Polos 1, 2 e 3)	13 DRE
Grande São Paulo (Polos 4,5 e 6)	15 DRE
Interior Norte (Polos 7,8, 9, 14 e 15)	34 DRE
Interior Sul (Polos 10, 11, 12 e 13)	29 DRE

4.1.3. Quantidade de municípios, escolas e servidores para cada DRE

Os municípios pertencentes ao Estado de São Paulo foram agrupados em áreas de jurisdições. Cada jurisdição apresenta um número específico de municípios e/ou sub-regiões, variando entre uma sub-região ou município de até vinte e cinco municípios. As escolas estão organizadas de acordo com as jurisdições a que pertencem seus municípios. A quantidade de unidades para cada DRE varia entre 17 e 118 escolas. As DRE apresentam um quadro diversificado em relação ao número total de servidores. A soma dos funcionários de todos os centros e núcleos para cada órgão regional varia entre 48 e 128.

4.2. Pré-Processamento

Por meio da definição acerca do problema que se deseja responder e do domínio de aplicação, foi necessário selecionar e agrupar os indicadores em uma única base de dados, processos fundamentais para a mineração dos dados. Os indicadores e seus respectivos valores foram retirados de diferentes bases estruturadas e alocados em um editor de texto, *Notepad⁺⁺*, no formato em que foram extraídos.

Para que os dados fossem lidos pela ferramenta Weka foi criado um arquivo de texto no formato ARFF (*Attribute Relation File Format*). Este formato de arquivo é representado basicamente em marcações de cabeçalho: *@relation* (nome do arquivo) e *@attribute* (nome do atributo/item e a descrição dos valores que ele pode representar); e *dos dados: @data* (instância - registros dos valores a serem minerados).

O cabeçalho assumiu uma forma ordenada de instruções e, foi representado em cada linha pelos nomes dos atributos: região, municípios, escolas, servidores e IDESP e, os tipos de dados: *nominal, numeric, numeric, numeric, numeric*. As linhas das instâncias representam os valores dos atributos separados pelos espaçamentos, os quais foram colocados na mesma ordem em que foram declarados na seção do cabeçalho. Por exemplo, se o atributo escola foi o terceiro a ser declarado, o Weka entende que todos os valores da terceira coluna, em todas as linhas, pertencem ao atributo.

4.3. Transformação dos dados

O processo de higienização e padronização ocorreu por meio da transformação dos dados, para um formato que permitiu a leitura do algoritmo. Segundo Tang et al. (2005) o principal objetivo da transformação dos dados está em sua modificação para diferentes formatos em termos de tipo de dados e valores. Portanto, o estabelecimento de grupos dentro de um mesmo item se fez necessário.

Os valores dos atributos IDESP, quantidade de municípios, quantidade de escolas e quantidade de servidores foram transformados/discretizados para valores nominais e agrupados em categorias com nomenclaturas que buscaram representar o grupo do conjunto. Além disso, as divisões dos outros itens foram estabelecidas por meio do método não supervisionado, ou seja, a própria ferramenta determinou os valores dos intervalos entre os itens, como apresentado na Tabela 4.2.

Tabela 4.2. Transformação dos dados

Item	Valores	Intervalos	Representação dos Intervalos
região	-	nominal	Interior Norte (IN) Interior Sul (IS) Grande São Paulo (G) Capital (C)
municípios	Entre 1 e 25	nominal	$\leq 9 / 10 - 17 / \geq 18$
escolas	Entre 17 e 118	nominal	$\leq 50 / 51 - 84 / \geq 85$
servidores	Entre 48 e 128	nominal	$\leq 74 / 75 - 101 / \geq 102$
IDESP	Entre 2,26 e 3,20	nominal	Baixo: $\leq 2,26$ Medio: 2,27-2,77 Alto: $\geq 2,78$

4.4. Tarefa de Mineração

Procurando compreender as relações entre os itens, bem como associações mais específicas com base em um dado determinante, nesse caso o IDESP, o estudo construiu com base na ferramenta Weka, nas regras de associação e no algoritmo Apriori a identificação de relações entre o desempenho das diretorias regionais de ensino e suas especificidades (fatores extraescolares).

As regras de associação se estabeleceram no relacionamento entre os conjuntos de itens, os quais passaram por diversas leituras na base de dados: dados organizacionais, estruturais e de rendimento. Portanto, viabilizando a obtenção de regras do tipo “se IDESP alto ocorre, então uma quantidade menor ou igual à 50 escolas por DRE ocorre”, ou seja, os órgãos regionais com IDESP alto e menos de 50 escolas aparecem conjuntamente com certa frequência na base de dados.

Além disso, determinou-se o suporte e a confiança mínima para aplicação da regra. O suporte “0.1” eliminou os itens que não atenderam a quantidade mínima de registros estabelecida, portanto, a quantidade mínima razoável de coincidências para justificar a associação. A confiança “0.8” estabeleceu a frequência mínima de associações

entre os itens, dessa forma, somente as regras com probabilidades maiores que o grau de confiança mínimo foram consideradas.

5. Avaliação

A fase de avaliação envolveu a visualização e interpretação dos dados, os quais foram submetidos à tarefa de mineração. Nesse sentido, analisou-se as regras disponibilizadas através do algoritmo Apriori, bem como os gráficos fornecidos pela própria ferramenta, após a inserção do documento no formato ARFF.

5.1. Geração de regras com o Apriori

Na análise dos resultados selecionou-se as regras de associação que atenderam o foco do problema desta pesquisa – Fatores extraescolares que interferem na qualidade da educação. Portanto, tem-se o indicador de desempenho “IDESP” como a variável foco para a seleção das regras que se tornaram o escopo da análise. A Tabela 5.1 apresenta os resultados obtidos, conforme disponibilizado na ferramenta Weka.

Tabela 5.1. Regras de associação geradas pelo algoritmo Apriori

Índice	Regra de Associação	Confiança	Suporte
4	região G & IDESP médio (10) ⇒ municípios ≤ 9 (10)	1	1.44
6	região IN & servidores ≤ 74 & IDESP alto (10) ⇒ escolas ≤ 50	1	2.76
10	servidores ≤ 74 & IDESP alto (14) ⇒ escolas ≤ 50 (13)	0.93	2.56
12	região IS & IDESP alto (12) ⇒ municípios ≤ 9 (11)	0.92	1.32
17	região IN & escolas ≤ 50 & IDESP médio (10) ⇒ servidores ≤ 74 (9)	0.9	2.41
21	região IS & IDESP médio (16) ⇒ municípios ≤ 9 (14)	0.88	2.26
23	região IN & IDESP alto (14) ⇒ escolas ≤ 50 (12)	0.86	2.36
25	escolas ≥ 85 & IDESP médio (14) ⇒ municípios ≤ 9 (12)	0.86	1.24

No processo de seleção identificou-se 8 regras, do total de 25, que suscitam o indicador de desempenho. Neste montante, 4 estão relacionadas com IDESP alto ($\geq 2,78$) e 4 com IDESP médio (2,26 - 2,77). Contudo, o algoritmo não gerou nenhuma regra relacionada com IDESP baixo ($\leq 2,26$). Para melhor interpretação dos dados, realizou-se a análise descritiva para cada regra gerada, em que o item IDESP se fez presente:

- Regra 4 – 66% das DRE que estão na Grande São Paulo tem IDESP médio e poucos municípios (intervalo mínimo);

- Regra 6 - 29% das DRE que estão no Interior Norte tem IDESP alto, poucos servidores (intervalo mínimo) e poucas escolas (intervalo mínimo) dentro de sua jurisdição;
- Regra 10 – 41% das DRE com poucos servidores (intervalo mínimo) tem IDESP alto e poucas escolas (intervalo mínimo);
- Regra 12 - 41% das DRE que estão no Interior Sul tem IDESP alto e poucos municípios (intervalo mínimo);
- Regra 17 - 29% das DRE que estão no Interior Norte tem IDESP médio, poucas escolas em sua jurisdição (intervalo mínimo) e também poucos servidores (intervalo mínimo);
- Regra 21 - 55% das DRE do Interior Sul tem IDESP médio e poucos municípios dentro de suas jurisdições (intervalo mínimo);
- Regra 23 - 41% das DRE que estão no Interior Norte tem IDESP alto e poucas escolas (intervalo mínimo);
- Regra 25 - 60% das DRE com a maior quantidade de escolas em suas jurisdições (intervalo máximo) tem IDESP médio e poucos municípios (intervalo mínimo).

5.2. Relação IDESP e indicadores extraescolares

Após a geração das regras com o algoritmo Apriori, a ferramenta disponibilizou o ícone “*visualize*”, no qual foi possível fazer uma análise mais minuciosa dos dados, a partir da relação entre apenas dois itens.

5.2.1. IDESP x região

A representação gráfica dos dados no *software* Weka, Gráfico 5.1, evidencia que as diretorias regionais de ensino com IDESP alto estão concentradas no Interior Norte e Sul. Basicamente com percentual igual entre as duas regiões. Também apresentam IDESP médio e baixo, mas vale salientar: com percentual mínimo em relação ao IDESP baixo. Já a região da Grande São Paulo apresenta IDESP médio (com dois terços do total) e baixo (com um terço do total). O maior percentual de IDESP baixo está localizado na Capital, que também não apresenta percentual de IDESP alto.

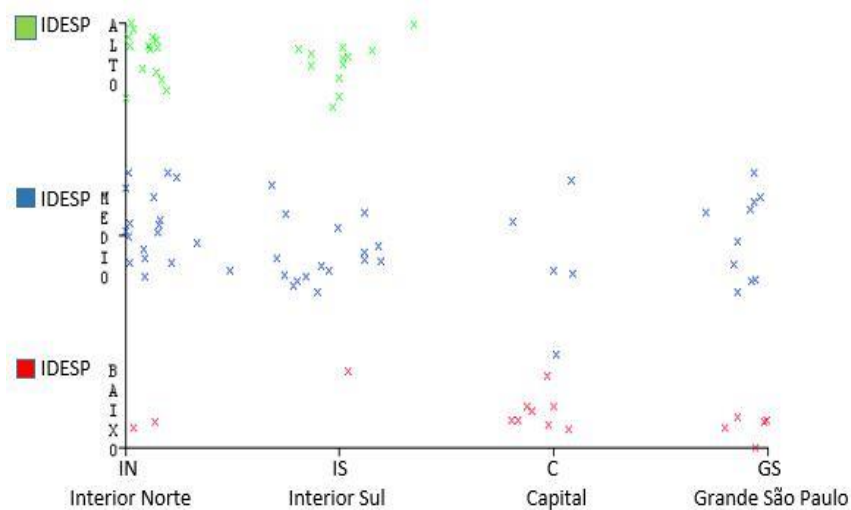


Gráfico 5.1. Relação - IDESP x região

5.2.2. IDESP x escolas

A fim de evidenciar a relação entre os indicadores IDESP e escolas, fez-se necessário uma análise descritiva das informações contidas no Gráfico 5.2. Deste modo, as diretorias regionais de ensino apresentaram o seguinte cenário em relação à quantidade de escolas e os resultados educacionais que medem a qualidade da educação:

- Intervalo mínimo de escolas – 6% de IDESP baixo / 45,5% de IDESP médio / 48,5% de IDESP alto;
- Intervalo médio de escolas – 20% de IDESP baixo / 57,1% de IDESP médio / 22,8% de IDESP alto.
- Intervalo máximo de escolas – 30% de IDESP baixo / 60,8% de IDESP médio / 8,6% de IDESP alto.

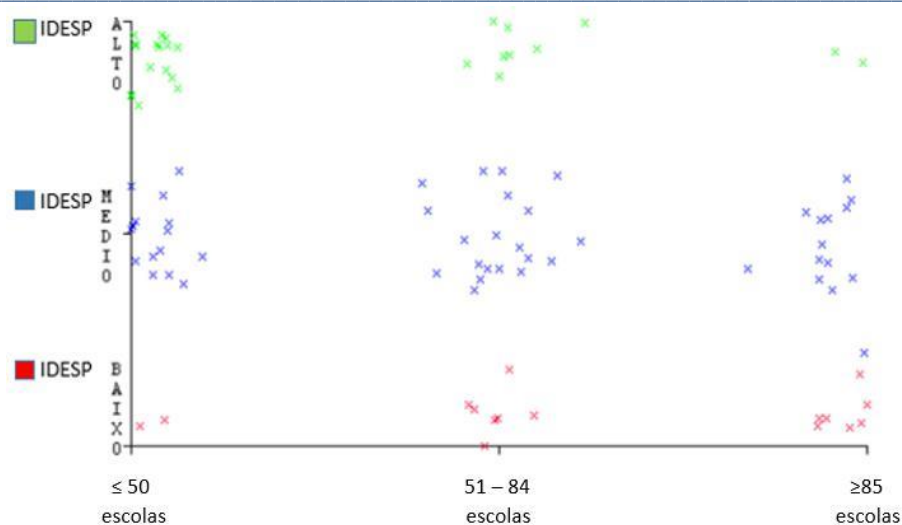


Gráfico 5.2. Relação - IDESP x escolas

5.2.3. IDESP x municípios

A fim de evidenciar a relação entre os indicadores IDESP e municípios, fez-se necessário uma análise descritiva das informações contidas no Gráfico 5.3. Deste modo, as diretorias regionais de ensino apresentaram o seguinte cenário em relação à quantidade de municípios e os resultados educacionais que medem a qualidade da educação:

- Intervalo mínimo de municípios – 22,2% de IDESP baixo / 57,1% de IDESP médio / 20,6% de IDESP alto.
- Intervalo médio de municípios – 12,5% de IDESP baixo / 54,1% de IDESP médio / 33,3% de IDESP alto.
- Intervalo máximo de municípios – 0% de IDESP baixo / 25% de IDESP médio / 75% de IDESP alto.

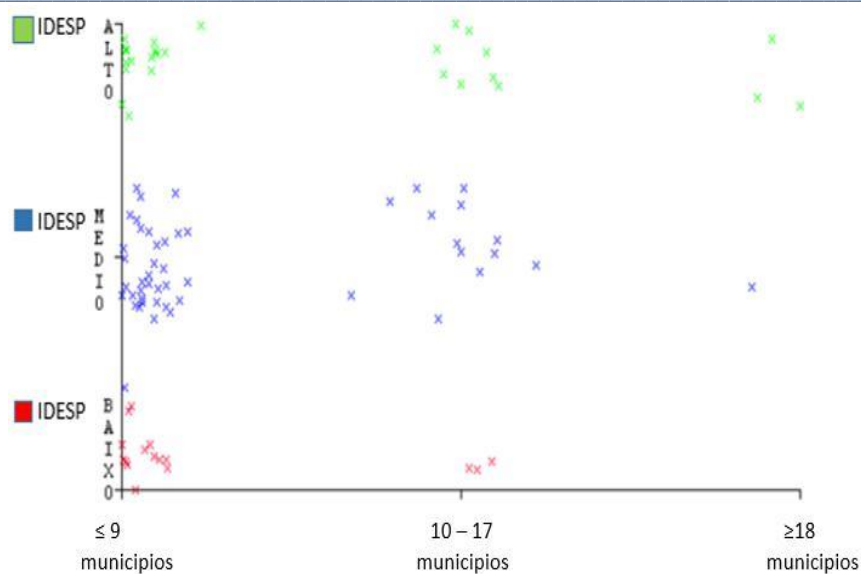


Gráfico 5.3. Relação - IDESP x municípios

5.2.3. IDESP x servidores

A fim de evidenciar a relação entre os indicadores IDESP e servidores, fez-se necessário uma análise descritiva das informações contidas no Gráfico 5.4. Deste modo, as diretorias regionais de ensino apresentaram o seguinte cenário em relação à quantidade de servidores e os resultados educacionais que medem a qualidade da educação:

- Intervalo mínimo de servidores – 8,8% de IDESP baixo / 50% de IDESP médio / 41,1% de IDESP alto.
- Intervalo médio de servidores – 23,6% de IDESP baixo / 47,3% de IDESP médio / 28,9% de IDESP alto.
- Intervalo máximo de servidores – 21% de IDESP baixo / 73,6% de IDESP médio / 5,2% de IDESP alto.

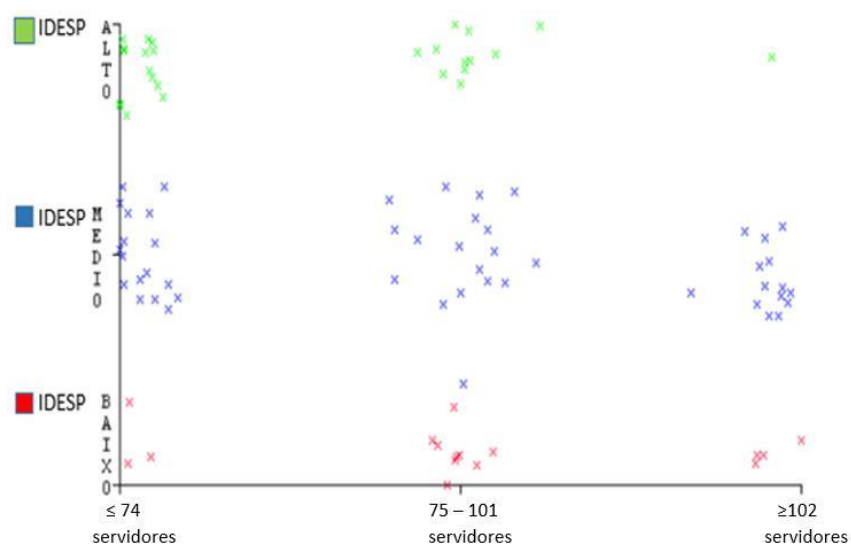


Gráfico 5.4. Relação - IDESP x servidores

6. Discussão

A fim de analisar todas as variáveis atreladas às diretorias regionais de ensino e a qualidade da educação, esta etapa iluminará cada um dos itens selecionados, associando-os ao indicador IDESP. Procurará estabelecer alguns caminhos de investigação e verificará a correlação entre as variáveis extraescolares e os resultados educacionais.

Após análise do indicador “municípios”, percebeu-se, com base nas duas variáveis de análise, que as diretorias regionais de ensino que apresentaram os indicativos de desempenhos mais altos não estão relacionadas a nenhum tipo específico de intervalo. Além disso, não se observou caminhos diretos e significativos de análise para a relação entre a quantidade de municípios por DRE e os resultados educacionais.

O indicador “servidores” não apresentou evidências robustas de sua influência em relação ao indicador de desempenho, na verdade, o item como um todo se relaciona e está distribuído de forma compatível com o indicador escola, ou seja, quanto maior a quantidade de escolas, maior a quantidade de servidores internos nas diretorias regionais de ensino.

As regras geradas com vistas no item “região” aparecem com maior recorrência em relação a todas as outras combinações de itens - no total de 8 regras analisadas, o item aparece por 6 vezes. Nesse panorama, um fator extremamente interessante se estabelece a partir dos itens região Interior Norte e escolas ≤ 50 (intervalo mínimo) pois, se o item região Interior Norte aparece, o item intervalo mínimo de escolas também aparece, ambos ocorrendo conjuntamente com bastante frequência.

As regras geradas a partir do item “escolas” também apresentam variáveis interessantes. No total de 8 regras o item aparece por 5 vezes e, em 4 situações, é visualizado com o intervalo ≤ 50 . Além disso, a relação deste item com o indicador IDESP demonstra que as diretorias regionais de ensino com poucas escolas em suas jurisdições apresentam os melhores resultados educacionais.

Atentando para os resultados obtidos é justificável apresentar os indicadores (região e escolas) que se destacam em relação aos demais, vão ao encontro do objetivo da pesquisa e podem servir de mote para outros estudos. Entretanto, vale salientar que os próximos parágrafos não devem ser vistos como as únicas possibilidades de análise ou imutáveis em relação a outras variáveis de interpretações.

Primeiramente, os itens escolas (intervalos na vertical) e região (intervalos na horizontal), Gráfico 6.1, apresentam algumas correlações: as diretorias regionais de ensino que estão no Interior Norte e Sul (IN e IS) apresentam a menor quantidade de escolas e os melhores resultados educacionais. E, subsequentemente, as diretorias regionais de ensino que estão nas regiões Capital e Grande São Paulo demonstraram resultados menos promissores e uma quantidade elevada de escolas por órgão regional.

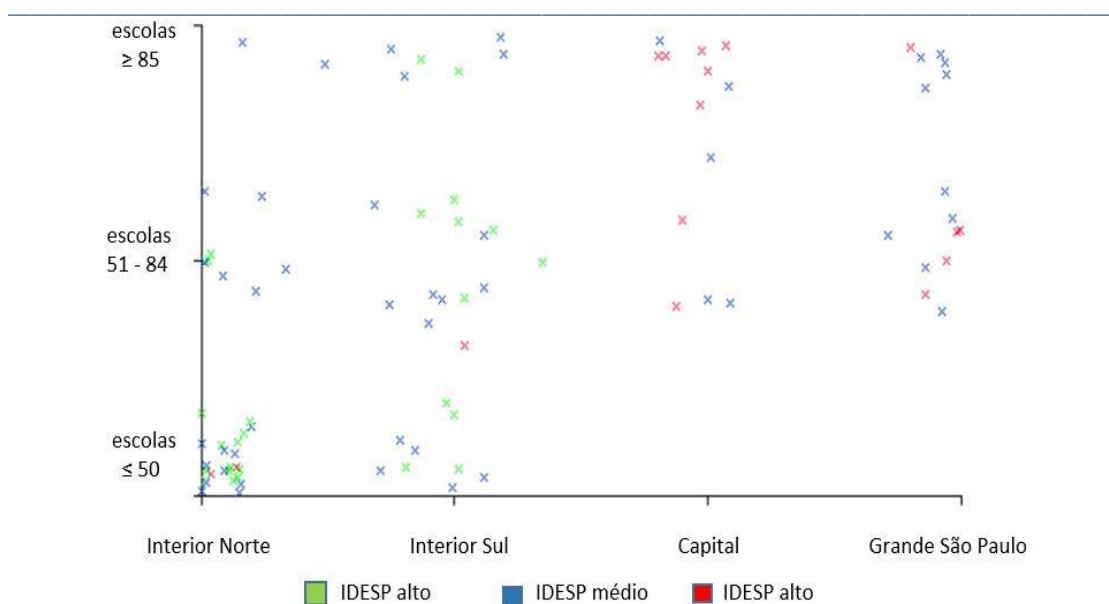


Gráfico 6.1. Relação do IDESP com os indicadores regiões e escolas

Em consequência, evidenciam a extração de subconjuntos de itens frequentes, sendo difícil determinar qual item corrobora de forma mais significativa em relação aos resultados educacionais, consequentemente, abrindo caminhos para interpretações:

- As divisões regionais e suas especificidades sociais, econômicas, regionais, culturais, familiares e acadêmicas propiciam particularidades relevantes para a qualidade de ensino que está sendo ofertado e, pela localização espacial em que se encontram, determinam a quantidade de escolas para cada jurisdição?
- Uma quantidade específica de escolas para suporte educacional e administrativo pode determinar que as ações e demandas pedagógicas e operacionais das diretorias regionais de ensino sejam mais assertivas e ajustadas com as necessidades dos grupos escolares e, reverberar em qualidade de educação?

Da perspectiva dos limites, este estudo não apresentou indicadores necessários para a identificação das características específicas de cada órgão regional. Para isso, necessitaria de outros indicadores, intra e extraescolares, e mecanismos de análise para a validação das hipóteses suscitadas. Contudo, é possível observar um valor significativo acerca da sua validade e importância, uma vez que cada item proposto abre diferentes possibilidades para novos estudos, problemas e validação de hipóteses. Além disso, cumpre seu objetivo, pois traz luz para alguns fatores extraescolares diretamente relacionados ao indicador que mede a qualidade do ensino na rede estadual paulista.

7. Conclusão

O acesso à educação gratuita e de qualidade tem representado um desafio para as gestões educacionais. Sobretudo, em relação ao alinhamento dos processos pedagógicos e operacionais entre os níveis institucionais. Ademais, observa-se que o diagnóstico das fragilidades e potencialidades dos setores que dão suporte às ações e demandas pedagógicas das escolas pouco tem sido explorado pela literatura, principalmente em relação ao órgão que estabelece o elo entre a burocracia de alto escalão e a burocracia de nível de rua.

Nesse ínterim, as diretorias regionais de ensino (burocracia de nível médio) encontram dificuldades em alcançar, em muitas situações, o fortalecimento de suas práticas, pois, além de desenvolverem muitas funções: monitorar, assistir, dimensionar, propor, gerenciar, acompanhar, supervisionar e implementar as políticas públicas formuladas pelo nível central, não estão estruturadas e organizadas de forma a obter o máximo de eficiência; tendo como exemplo, as metas projetadas e os resultados obtidos no IDEB de 2019 para os Anos Finais do Ensino Fundamental e Ensino Médio.

Considerando os indicadores selecionados e a ferramenta de mineração de dados educacionais Weka – instrumento essencial para a análise dos dados e para a construção de conhecimentos - diversas técnicas foram testadas e combinadas, a fim de que associações e padrões fossem estabelecidos.

Os resultados obtidos tiveram como escopo os aspectos estruturais e organizacionais e de desempenho das 91 DRE que compõem o sistema estadual de ensino paulista do Ensino Médio de 2018 e 2019, tratando assim, das necessidades que estão em voga. E, apesar de sinalizarem indicadores que já foram iluminados pela análise dos índices educacionais - caso das regiões com os melhores coeficientes no IDESP. Ressalta-se, aqui, o estabelecimento de novas associações entre os indicadores extraescolares e a validação dos resultados.

Do ponto de vista das seções resultados e discussão, este estudo traz luz há fatores extraescolares (quantidade de escolas por diretoria regional de ensino e divisões regionais) que sinalizam e fomentam a necessidade de novos estudos quantitativos e qualitativos. Pois, para que mudanças estruturais e organizacionais sejam efetivas e impactem positivamente nos resultados das escolas é necessário a compreensão do todo.

A partir do exposto, várias são as alternativas para novos trabalhos, tendo como possibilidade de temáticas: A análise das relações entre as diretorias regionais de ensino e unidades escolares (especificidades pedagógicas e administrativas), a partir de diferentes quantitativos de escolas por órgão regional; A análise das características e especificidades do Interior paulista em relação ao desenvolvimento do processo ensino e aprendizagem; A análise do grau de influência dos fatores regionais da rede estadual paulista na oferta de um ensino de qualidade; E, um estudo acerca da relação (causa ou correlação) entre os indicadores extraescolares “região” e “quantitativo de escolas”.

Para finalizar, é importante salientar que não basta que os órgãos regionais mobilizem seu corpo burocrático em direção às demandas do seu público-alvo, adotem parâmetros de qualidade, recursos humanos dotados de habilidades e conhecimentos para formação do professorado, instrumentos de controle e monitoramento da implementação

de projetos e da aplicação do Currículo. É imprescindível, que também tenham uma estrutura organizacional que corrobore positivamente no trabalho que precisa ser realizado.

Referências

Abrucio, F.L.; Oliveira, V. E. (2018) “Burocracia e políticas públicas no Brasil : interseções analíticas”. Brasília : Ipea : Enap. 413 p.

Abrucio, F. L. (2005) “Reforma do Estado no Federalismo Brasileiro: A Situação das Administrações Públicas Estaduais”. RAP. Revista Brasileira de Administração Pública, Rio de Janeiro/RJ, v. 39, n. 2, p. 401-420.

Abrucio, F. L. (2010) “Gestão Escolar e Qualidade da Educação: Um Estudo sobre Dez Escolas Paulistas”. Estudos e Pesquisas Educacionais, v. 1, p. 241-274.

Agrawal, R., Imieliński, T. e Swami, A. (1993). “Regras de associação de mineração entre conjuntos de itens em grandes bancos de dados”. In: Proc. Conf. on Management of Data , 207-216. Nova York: ACM Press.

Azevedo, J. M. L. (20214) “Plano Nacional de Educação e planejamento - A questão da qualidade da educação básica”. Revista Retratos da Escola, Brasília, v. 8, n. 15, p. 265-280.

Baker, Ryan; Isotany, Seiji; Carvalho, Adriana. (2011) “Mineração de Dados Educacionais: Oportunidades para o Brasil”. Revista Brasileira de Informática na Educação, [S.l.], v. 19, n. 02, p. 03, ago. 2011. ISSN 2317-6121.

Barros, Maria Camila Mourão Mendonça (2018). “O elo institucional regional da educação básica: um estudo das diretorias regionais de ensino paulista”. Tese (Doutorado em Administração Pública e Governo) – Escola de Administração de Empresas de São Paulo da Fundação Getúlio Vargas, São Paulo. 204 f.

Cabena, P., Hadjinian, P., Stadler, R., Verhees, J., and Zanasi, A. (1998). “Discovering data mining: from concept to implementation”. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

Carrano, Davi et al. (2019) “Combinando Técnicas de Mineração de Dados para Melhorar a Detecção de Indicadores de Evasão Universitária”. Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação - SBIE), [S.l.], p. 1321, nov. 2019. ISSN 2316-6533.

Carvalho, J. V. de; Mongiovi, M. C. S. G (2000). “Utilização de técnicas de data mining para o reconhecimento de caracteres manuscritos”. Universidade Federal da Paraíba. Campina Grande, PB.

Costa, Evandro et al. (2013) “Mineração de Dados Educacionais: Conceitos, Técnicas, Ferramentas e Aplicações”. Jornada de Atualização em Informática na Educação, [S.l.], p. 1-29, fev. 2013. ISSN 23167734.

Dados Abertos da Educação. (2020) “Índice de Desenvolvimento da Educação do Estado de São Paulo (IDESP) por Diretoria de Ensino”. Disponível em: <https://dados.educacao.sp.gov.br/dataset/%C3%ADndice-de-desenvolvimento-da-educa%C3%A7%C3%A3o-do-estado-de-s%C3%A3o-paulo-idesp-por-diretoria-de-ensino>. Acesso em 30 de jun. de 2020.

Dourado, L.F. Oliveira, J.F.; Santos, C.A. (2007) “A qualidade da educação: conceitos e definições. Série Documental: Texto para discussão”. Brasília, DF, v. 24, n. 22p. 5-34.

Dourado, L. F. Oliveira J.F. (2009). "A qualidade da educação: perspectivas e desafios". Cad. CEDES vol.29 no.78 Campinas May/Aug. 2009.

Fayyad, U. M.; Piatetsky-Shapiro, G.; Smity, P. (1996) "From data mining to knowledge discovery in databases". AI Magazine, American Association for Artificial Intelligence, Califórnia, USA, v. 17.

Ferreira, G. S. (2015) "Investigação acerca dos fatores determinantes para a conclusão do ensino fundamental utilizando mineração de dados educacionais no censo escolar da educação básica do Inep". Anais dos Workshops do IV Congresso Brasileiro de Informática na Educação (CBIE), p. 1034–1043.

Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I. H. (2009) "The Weka data mining software: an update". SIGKDD Explor. Newsl., 11(1):10–18.

INEP (2019) "Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira". Disponível em: <<http://ideb.inep.gov.br/resultado>>. Acesso em: 13 de dezembro de 2020.

Jaques, Patrícia Augustin; Pimentel, Mariano; Siqueira, Sean; Bittencourt, Ig. (2020) "Metodologia de Pesquisa Científica em Informática na Educação: Concepção de Pesquisa". Porto Alegre: SBC, 2020. (Série Metodologia de Pesquisa em Informática na Educação, v. 1). Disponível em: <<https://metodologia.ceie-br.org/livro-1/>>. Acesso em: 30 de jun. de 2020.

Júnior, José G. de Oliveira; Noronha, Robinson Vida; Kaestnerk, Celso Antônio Alves. (2015) "Análise da Correlação da Evasão de Cursos de Graduação com o Empréstimo de Livros em Biblioteca". Anais dos Workshops do Congresso Brasileiro de Informática na Educação, [S.l.], p. 601, jan. 2015. ISSN 2316-8889. Disponível em: <<https://www.br-ie.org/pub/index.php/wcbie/article/view/3284>>. Acesso em: 05 out. 2020.

Minussi, M. M. et al. (2008) "Metodologia de mineração de dados para detecção de desvio de comportamento do uso de energia em concessionária de energia elétrica". Dissertação de Mestrado, Pontifícia Universidade Católica do Rio Grande do Sul, Porto Alegre-RS.

Miranda, L. P. (2018). "Mineração de dados como suporte educacional. 53f. Dissertação (Mestrado em Matemática) - Programa de Mestrado Profissional em Matemática em Rede Nacional, Universidade Federal do Triângulo Mineiro, Uberaba.

Oliveira, V. E. & Abrucio, F. L. (2011) "Entre a política e a burocracia: a importância dos burocratas de nível médio para a produção de políticas públicas em saúde e educação". In *Anais do 35o Encontro Anual da Anpocs*. Caxambu, MG.

Peixoto, S.P. et al. (2017) "O impacto dos fatores intra e extraescolares para o fracasso escolar: desmistificando as visões psicologizantes". Ciências Humanas e Sociais. Alagoas, v. 4, n 2, p.235-248.

São Paulo. (2019) Decreto nº 64.187, de 17 de abril de 2019. Reorganiza a Secretaria da Educação e dá providências correlatas, São Paulo: Diário Oficial do Estado de São Paulo, Executivo, 18 de abril de 2019, p. 3.

Secretaria de Educação do Estado de São Paulo. (2020) "Visão da Diretoria", Secretaria Escolar Digital. Disponível em: <<https://sed.educacao.sp.gov.br/SedBI/VisaoDiretoria>>. Acesso em: 30 de jun. de 2020.

Secretaria da Educação do Estado de São Paulo. (2020) "Localize uma Diretoria de Ensino". Disponível em: <http://www.educacao.sp.gov.br/central-de-atendimento/Index_Mapas_Dir.asp>. Acesso em: 30 de jun. de 2020.

Silva, Leandro A. Silva, Luciano. (2015) "Fundamentos de Mineração de Dados Educacionais". Anais dos Workshops do Congresso Brasileiro de Informática na Educação, [S.l.], p. 568, jan. 2015. ISSN 2316-8889.

Singh, S. and Kumar, V. (2012) "Classification of Student's data Using Data Mining Techniques for Training & Placement Department in Technical Education", (IJCSN) International Journal of Computer Science and Network, Vol. 1(4), pages.121-126, ISSN: 2277-5420.

Tang, Zhaohui; Maclennan, Jaime. Data Mining with SQL Server (2005). Indianapolis: Wiley.

Vasconcelos, M. R. Carvalho C. L. (2004) "Aplicação de Regras de Associação para Mineração". Technical Report - RT-INF004-04 - Relatório Técnico, novembro - 2004.