

LA INFRAESTRUCTURA DE LAS ESCUELAS Y EL DESEMPEÑO DE SUS ALUMNOS EN ENEM: UN ANÁLISIS

Giuliano Nunes

giulianonunes0@gmail.com

Universidade Luterana do Brasil

Renato P. dos Santos

renatopsantos@ulbra.edu.br

Universidade Luterana do Brasil

Recibido: 10.01.2019 **Aceptado:** 30.04.2019

RESUMEN

La educación juega un papel fundamental en el empoderamiento de los ciudadanos, que contribuye a su bienestar y también al desarrollo económico del país y, cuando es de alta calidad, de suma importancia para una nación. En Brasil tenemos indicadores educativos que buscan medir la calidad de la enseñanza en las escuelas, incluidos el IDEB y el ENEM. Hay algunos factores que determinan el rendimiento cognitivo de los estudiantes, como la estructura de las escuelas. Este trabajo busca encontrar relaciones entre la estructura escolar obtenida en el Censo Escolar 2015 y su desempeño, medido por medio de la calificación promedio en el ENEM de ese mismo año, de cada entidad docente, si existe y qué elementos tienen mayor influencia. Después de la clasificación de las variables, fue posible enviar estos datos para las pruebas de correlación a través del lenguaje del análisis estadístico R. Los resultados muestran una mayor correlación entre las calificaciones en el ENEM con la existencia en la escuela de acceso a una red pública de suministro de agua, así como una correlación positiva cuando las aguas residuales son a través del servicio público y la red negativa cuando se hace a través de un tanque séptico. La literatura explica esta correlación por una mayor incidencia de enfermedades infecciosas, gastrointestinales y neuro cognitivas en niños en edad escolar expuestos a instalaciones inadecuadas de agua y saneamiento que podrían impactar negativamente la asistencia escolar de los estudiantes. Esto indica que se podría lograr una mejora significativa en el rendimiento escolar simplemente al extender el saneamiento básico a las escuelas.

Palabras clave: Data Science, infraestructura escolar, ENEM, IDEB, desigualdades escolares.

A INFRAESTRUTURA DAS ESCOLAS E O DESEMPENHO DE SEUS ALUNOS NO ENEM: UMA ANÁLISE

RESUMO

A educação tem papel fundamental para a capacitação do cidadão, o que contribui para seu bem-estar e também para o desenvolvimento econômico do país e, em sendo de qualidade, de extrema importância para uma nação. No Brasil contamos com indicadores educacionais que buscam aferir a qualidade do ensino das escolas, entre eles o IDEB e o ENEM. Existem alguns fatores que determinam o desempenho cognitivo dos alunos, como por exemplo a estrutura das escolas. Este trabalho busca encontrar relações entre a estrutura das escolas, obtidas no Censo

Escolar 2015, e seu desempenho, medido através da nota média no ENEM desse mesmo ano, de cada entidade de ensino, se elas existem e quais itens com maior influência na média da escola. Após a classificação das variáveis, foi possível submeter estes dados para testes de correlação através da linguagem de análise estatística R. Os resultados apresentam uma maior correlação entre as notas do ENEM com a existência na escola de acesso a uma rede pública de abastecimento de água, bem como uma correlação positiva quando o esgoto sanitário é por meio de rede pública e negativa quando o é por meio de fossa. A literatura explica essa correlação por uma maior incidência de doenças infecciosas, gastrointestinais e neurocognitivas em crianças em idade escolar expostas a instalações inadequadas de água e saneamento que impactaria negativamente na frequência escolar dos estudantes. Isso indica que uma melhoria significativa do desempenho escolar poderia ser obtida apenas ao se estender saneamento básico às escolas.

Palavras chave: Ciência de Dados, infraestrutura escolar, ENEM, IDEB, desigualdades escolares.

SCHOOL INFRASTRUCTURE AND THE PERFORMANCE OF ITS STUDENTS IN THE ENEM: AN ANALYSIS

ABSTRACT

Education plays a fundamental role in empowering citizens, which contributes to their well-being and also to the economic development of the country and, when of high quality, of extreme importance for a nation. In Brazil we have educational indicators that seek to measure the quality of teaching in schools, including IDEB and ENEM. There are some factors that determine students' cognitive performance, such as the structure of schools. This work seeks to find relationships between the school structure obtained in the 2015 School Census and its performance, measured by means of the average grade in the ENEM of that same year, of each teaching entity, if they exist and what items with greater influence. After the classification of the variables, it was possible to submit these data for correlation tests through the language of statistical analysis R. The results show a greater correlation between the grades in the ENEM with the existence in the school of access to a public network of water supply as well as a positive correlation when sanitary sewage is through public service and negative network when it is through septic tank. The literature explains this correlation by a higher incidence of infectious, gastrointestinal and neurocognitive diseases in school-aged children exposed to inadequate water and sanitation facilities that would negatively impact students' school attendance. This indicates that a significant improvement in school performance could be achieved simply by extending basic sanitation to schools.

Keywords: Data Science, school infrastructure, ENEM, IDEB, school inequalities.

INTRODUÇÃO

A educação vem se transformando ao longo dos anos e é um tema sempre colocado em debate. Para Mayer e Rodrigues (2013), educação pode ser vista como a composição dos processos de ensinar e aprender, sendo um fenômeno existente em toda sociedade responsável por passar às próximas gerações os modos culturais de ser, estar e agir necessários para uma melhor convivência de uma pessoa em seu grupo. Não podemos negar a importância da

educação para o desenvolvimento econômico de um país, pois segundo Rocha (2004), a educação proporciona ganhos diretos para as pessoas com melhor formação e favorece a mudança social através da melhoria da qualidade de vida, contribuindo assim para o progresso econômico.

No entanto, para Carvalho (2017), a escola não deve ser tratada exclusivamente como um local ao qual se atribuem certas funções sociais e finalidades econômicas, mas também como um potencial lugar de experiências. Um local para encontros intergeracionais mediados pelo diálogo, com um conjunto de objetos e práticas culturais, contribuindo para a formação do sujeito.

No Brasil, é atribuição da escola a função de preparar e capacitar as novas gerações através da aquisição de conhecimentos, tornando os indivíduos participantes ativos e críticos na vida social do país. Cabe à escola a aprendizagem dos conteúdos cognitivos. Segundo Soares (2004), três fatores determinam o desempenho cognitivo dos alunos: os associados à estrutura escolar, os associados à família e àqueles relacionados ao próprio aluno.

A fim de aferir o desempenho cognitivo dos alunos e, indiretamente, a qualidade de ensino propiciadas por suas escolas, são utilizados indicadores educacionais. Tais indicadores apresentam uma nota ao desempenho dos estudantes em determinadas etapas do processo de ensino. No Brasil podemos citar o Índice de Desenvolvimento da Educação Básica (IDEB), que avalia o ensino fundamental em seus anos iniciais e finais, e o Exame Nacional de Ensino Médio (ENEM), que avalia o ensino médio. Para a obtenção destes indicadores também são avaliadas as questões sociais e econômicas dos estudantes. Desta maneira, é possível utilizar os dados obtidos através dos indicadores para avaliar os sistemas educacionais e o desempenho cognitivo de seus alunos e repensar as políticas públicas referentes ao ensino.

Desde sua criação em 1998, o Exame Nacional de Ensino Médio (ENEM) vem sendo utilizado para avaliar os estudantes no final da educação básica em todo o país. Primeiramente utilizado na seleção de candidatos a uma bolsa no Programa Universidade Para Todos (PROUNI), hoje é um dos critérios para admissão de estudantes em diversas instituições federais e privadas (MALUSÁ; ORDONES; RIBEIRO, 2014). Muitas escolas utilizam a média dos seus alunos no ENEM como forma de atestar sua qualidade de ensino.

Em 2016 a Diretoria de Análise de Políticas Públicas da Fundação Getúlio Vargas (FGV/DAPP) apresentou um estudo chamado Universo Escolar: o estado da educação no

Brasil¹, com o intuito de discutir o impacto da infraestrutura disponível no ensino médio no país, no desempenho escolar. Para tal, foram cruzadas as notas de matemática do ENEM por escola, obtidas em 2014, com as informações do Censo Escolar do mesmo ano. No referido estudo, foram selecionados os dados de 10% das instituições que obtiveram alto desempenho no ENEM e compararam com 10% das que apresentaram baixo desempenho. Segundo os dados apresentados pela FGV, entre as escolas de baixo desempenho, 239 declararam não ter acesso à rede pública de água. Este número representa 15% das escolas classificadas com baixo desempenho. Enquanto para as escolas de alto desempenho, menos de 4% não tem acesso à rede pública de água. Referente ao esgoto, 93,35% das escolas de alto desempenho possuem acesso à rede pública e apenas 37,6% das escolas de baixo desempenho dispõem do mesmo.

Jasper et al. (2012) identificaram através de uma revisão literária, os efeitos do saneamento e do fornecimento da água nas escolas sobre o desempenho dos alunos. A revisão apresenta estudos que documentam uma maior incidência de doenças infecciosas, gastrointestinais e neurocognitivas em crianças em idade escolar, que foram expostas a instalações inadequadas de água e saneamento, impactando diretamente na frequência escolar dos estudantes.

Nas palavras de Soares (2004), “o efeito de uma escola no aprendizado de seus alunos é em grande parte determinado pelo professor, por seus conhecimentos, seu envolvimento e sua maneira de conduzir as atividades da sala de aula”. O ato de ensinar é um trabalho conjunto entre professor e aluno. Cabe aqui salientar que características de personalidade do professor, envolvimento com a turma, expectativas em relação ao futuro do aluno e a capacidade de promover um ambiente favorável para o aprendizado, são variáveis que não podemos medir, mas que contribuem para o bom desempenho dos alunos. Visto isso, alguns estudos mostram que turmas pequenas, favorecem o aprendizado dos alunos, principalmente os de famílias com menor capital cultural e/ou dos níveis iniciais. Nessa situação o professor tem um contato mais próximo e frequente com os alunos, supervisionando o trabalho de cada um. Uma boa infraestrutura escolar pode auxiliar o professor a desempenhar suas atividades com melhor aproveitamento.

¹<http://dapp.fgv.br/universo-escolar/>

Como citado anteriormente, um dos fatores determinantes no desempenho cognitivo dos alunos é relacionado à estrutura escolar. Mas qual a influência da estrutura escolar no desempenho dos alunos nos indicadores educacionais? Quais itens da estrutura podem colaborar com melhores resultados?

Desta forma, o presente trabalho busca encontrar relações entre indicadores referentes à estrutura das escolas e o desempenho de seus no ENEM, medido através da nota média de cada entidade de ensino, se elas existem e quais itens com maior influência na média da escola.

METODOLOGIA

Para a análise proposta, foram utilizadas técnicas de Ciência de Dados (*Data Science*). Segundo Wickham e Grolemund (2016, p. 1, tradução nossa) “a Ciência de Dados é uma disciplina emocionante que permite transformar os dados brutos em compreensão, percepção e conhecimento”, afirmando assim a capacidade desta ciência para transformar dados em informação e conhecimento.

A Ciência de Dados deriva do processo de Descoberta de Conhecimento em Bancos de Dados (*Knowledge Discovery in Databases – KDD*) proposto por Fayyad, Piatetsky-Shapiro e Smyth (1996), que visa à identificação de padrões novos, válidos, potencialmente úteis e compreensíveis, que estivessem embutidos nos dados. As cinco etapas desse processo são: a seleção de bases adequadas, a limpeza das inconsistências que podem afligir esses dados, a transformação dos dados em formatos mais adequados, escolher de técnicas e algoritmos para a mineração e a avaliação e interpretação dos padrões extraídos, em forma do novo conhecimento. Hoje, as três primeiras dessas etapas são agrupadas numa fase denominada *análise exploratória de dados* (PENG, 2016) e fazem parte da chamada ‘mineração de dados’ (*data mining*).

Mais especificamente, foram utilizados os recursos da linguagem R (R CORE TEAM, 2017), versão 3.4.0, disponibilizada em 21 de abril de 2017. Apesar de R ser muito versátil e contar com muitos recursos, todo o trabalho foi realizado em poucos minutos num notebook convencional, com processador Dual-Core de 1,0 GHz 32 bits, memória RAM de 4 GB DDR3 e com sistema operacional Windows 10 x32 (*build 2011*). Por isso, a linguagem R é uma das ferramentas mais utilizadas na área de Ciência de Dados.

Desta forma, para a consecução do objetivo acima expresso, foram escolhidas as seguintes bases de dados:

- Os microdados do Censo da Educação Básica 2015, para as informações sobre a estrutura das escolas, e
- Os dados referentes às médias por escola do desempenho no ENEM 2015 dos alunos matriculados no 3º ano do Ensino Médio regular.

As informações sobre a estrutura das escolas são fornecidas pela própria unidade de ensino, por meio do Censo Escolar da Educação Básica, por meio do cadastro da escola no sistema *Educacenso*², e disponibilizados por meio do arquivo 'Microdados do Censo da Educação Básica 2015', obtido do site do Inep³.

Os dados referentes ao desempenho dos alunos no ENEM 2015 provêm do conjunto 'Microdados do Enem 2015', também obtido do site do Inep⁴.

O censo escolar traz informações das instituições de ensino em 166 variáveis, onde avaliamos 96, referentes à estrutura das escolas. Para melhor compreensão, essas 96 variáveis foram organizadas em 4 grupos: Prédio Escolar, Água Luz Saneamento, Dependências da Escola e Equipamentos Disponíveis.

No grupo Prédio Escolar estão contidas as variáveis sobre o local de funcionamento da escola, podendo ser em um prédio próprio e exclusivo para fins da instituição de ensino ou salas adaptadas em empresas, unidades prisionais, igrejas, dentre outras (Tabela 1).

Tabela 1: Nome das variáveis do grupo Prédio Escolar

Item	Variável	Descrição
1	CO_ENTIDADE	Código da Entidade
2	LOCAL_FUNC_PREDIO_ESCOLAR	Local de funcionamento da escola – Prédio Escolar
3	TP_OCUPACAO_PREDIO_ESCOLAR	Forma de ocupação do prédio escolar
4	LOCAL_SALAS_EMPRESA	Local de funcionamento da escola – Sala de empresa
5	LOCAL_SOCIOEDUCATIVO	Local de funcionamento da escola –

² <http://censobasico.inep.gov.br>

³ <http://portal.inep.gov.br/microdados>

⁴ http://download.inep.gov.br/microdados/microdados_enem2015.zip

		Unidade de Atendimento Socioeducativo
6	LOCAL_UNID_PRISIONAL	Local de funcionamento da escola – Unidade Prisional
7	LOCAL_PRISIONAL_SOCIO	Local de funcionamento da escola – Unidade Prisional ou Unidade de Atendimento Socioeducativo
8	LOCAL_TEMPLO_IGREJA	Local de funcionamento da escola – Templo/Igreja
9	LOCAL_CASA_PROFESSOR	Local de funcionamento da escola – Casa do Professor
10	LOCAL_GALPAO	Local de funcionamento da escola – Galpão/Rancho/Paiol/Barracão
11	TP_OCUPACAO_GALPAO	Forma de ocupação do Galpão/Rancho/Paiol/Barracão
12	LOCAL_SALAS_OUTRA_ESC	Local de funcionamento da escola – Salas em outra escola
13	LOCAL_OUTROS	Local de funcionamento da escola – Outros
14	PREDIO_COMPARTILHADO	Prédio compartilhado com outra escola

Fonte: Inep

As variáveis relacionadas ao fornecimento e tratamento de água e esgoto, assim como fornecimento de energia elétrica, estão contidas no grupo Água Luz Saneamento (Tabela 2).

Tabela 2: Nome das variáveis do grupo Água Luz Saneamento

Item	Variável	Descrição
1	CO_ENTIDADE	Código da Entidade
2	AGUA_FILTRADA	Água consumida pelos alunos na escola passa por um processo de filtragem
3	AGUA_RD_PUB	Abastecimento de água – Rede pública
4	AGUA_POCO_ART	Abastecimento de água – Poço artesiano
5	AGUA_CACIMBA	Abastecimento de água – Cacimba/Cisterna/Poço
6	AGUA_FTE_RIO	Abastecimento de água – Fonte/Rio/Igarapé/Riacho/Córrego
7	AGUA_INEX	Abastecimento de água - Inexistente
8	ENERGIA_RD_PUB	Abastecimento de energia elétrica – Rede pública
9	ENERGIA_GER	Abastecimento de energia elétrica - Gerador
10	ENERGIA_OUTROS	Abastecimento de energia elétrica – Outros (Energia alternativa)
11	ENERGIA_INEX	Abastecimento de energia elétrica -

		Inexistente
12	ESGOTO_RD_PUB	Esgoto sanitário – Rede pública
13	ESGOTO_FOSSA	Esgoto sanitário – Fossa
14	ESGOTO_INEX	Esgoto sanitário – Inexistente
15	LIXO_COL_PERI	Destinação do lixo – Coleta periódica
16	LIXO_QUEIMA	Destinação do lixo – Queima
17	LIXO_JG_OUTRA_AREA	Destinação do lixo – Joga em outra área
18	LIXO_RECICLA	Destinação do lixo – Recicla
19	LIXO_ENTERRA	Destinação do lixo – Enterra
20	LIXO_OUTROS	Destinação do lixo – Outros

Fonte: Inep

O grupo Dependências da Escola contém as variáveis que caracterizam a estrutura como: quantidade de salas existentes, salas utilizadas, a existência ou não de quadras esportivas e banheiros, dentre outras (Tabela 3).

Tabela 3: Nome das variáveis do grupo Dependências da Escola

Item	Variável	Descrição
1	CO_ENTIDADE	Código da Entidade
2	SALA_DIRETORIA	Sala de diretoria
3	SALA_PROF	Sala de professores
4	LAB_INFO	Laboratório de informática
5	LAB_CIENCIAS	Laboratório de ciências
6	SALA_ATEND_ESP	Sala de recursos multifuncionais para Atendimento Educacional Especializado (AEE)
7	QDRA_ESP_COBERTA	Quadra de esportes coberta
8	QDRA_ESP_DESCOBERTA	Quadra de esportes descoberta
9	QDRA_ESP	Quadra de esportes coberta ou descoberta
10	COZINHA	Cozinha
11	BIBLIOTECA	Biblioteca
12	SALA_LEITURA	Sala de leitura
13	BIBLIO_SALA_LEITURA	Biblioteca e/ou Sala de leitura
14	PARQUE_INFANTIL	Parque infantil
15	BERCARIO	Berçário
16	BANHEIRO_FORA_PRED	Banheiro fora do prédio
17	BANHEIRO_DENTRO_PRED	Banheiro dentro do prédio
18	BANHEIRO_EI	Banheiro adequado à Educação Infantil
19	BANHEIRO_PNE	Banheiro adequado ao uso dos alunos com deficiência ou mobilidade reduzida
20	DEPENDENCIAS_PNE	Dependências e vias adequadas a alunos com deficiência ou mobilidade reduzida
21	SECRETARIA	Sala de secretaria

22	BANHEIRO_CHUVEIRO	Banheiro com chuveiro
23	REFEITORIO	Refeitório
24	DESPENSA	Despensa
25	ALMOXARIFADO	Almoxarifado
26	AUDITORIO	Auditório
27	PATIO_COBERTO	Pátio coberto
28	PATIO_DESCOBERTO	Pátio descoberto
29	ALOJAM_ALUNO	Alojamento de aluno
30	ALOJAM_PROF	Alojamento de professor
31	AREA_VERDE	Área verde
32	LAVANDERIA	Lavanderia
33	DEPENDENCIAS_OUTRAS	Nenhuma das dependências relacionadas
34	SALAS_EXISTENTES	Número de salas de aula existentes na escola
35	NU_SALAS_UTILIZADAS	Número de salas utilizadas como salas de aula (dentro e fora do prédio)

Fonte: Inep

Em Equipamentos Disponíveis, estão listados os itens relacionados aos equipamentos encontrados nas escolas. Algumas variáveis deste grupo são Aparelho de TV, projetor multimídia, computador, quantidade de computadores para alunos e para o setor administrativo, dentre outras (Tabela 4).

Tabela 4: Nome das variáveis do grupo Equipamentos Disponíveis

Item	Variável	Descrição
1	CO_ENTIDADE	Código da Entidade
2	EQ_TV	Aparelho de televisão
3	EQ_VIDEOCASSETE	Aparelho de videocassete
4	EQ_DVD	Aparelho de DVD
5	EQ_PARABOLICA	Antena parabólica
6	EQ_COPIADORA	Copiadora
7	EQ_RETROPROJETOR	Retroprojektor
8	EQ_IMPRESSORA	Impressora
9	EQ_IMPRES MULT	Impressora multifuncional
10	EQ_SOM	Aparelho de som
11	EQ_MULTIMIDIA	Projetor multimídia (Datashow)
12	EQ_FAX	Aparelho de Fax
13	EQ_FOTO	Máquina fotográfica/filmadora
14	COMPUTADOR	Computador
15	NU_EQ_TV	Quantidade de aparelhos de televisão
16	NU_EQ_VIDEOCASSETE	Quantidade de aparelhos de videocassete
17	NU_EQ_DVD	Quantidade de aparelhos de DVD
18	NU_EQ_PARABOLICA	Quantidade de antenas parabólicas

19	NU_EQ_COPIADORA	Quantidade de copiadoras
20	NU_EQ_RETROPROJETOR	Quantidade de retroprojetores
21	NU_EQ_IMPRESSORA	Quantidade de impressoras
22	NU_EQ_IMPRES MULT	Quantidade de impressoras multifuncionais
23	NU_EQ_SOM	Quantidade de aparelhos de som
24	NU_EQ_MULTIMIDIA	Quantidade de projetores multimídia (Datashow)
25	NU_EQ_FAX	Quantidade de aparelhos de Fax
26	NU_EQ_FOTO	Quantidade de máquinas fotográficas/filmadoras
27	NU_COMPUTADOR	Quantidade de computadores na escola
28	NU_COMP_ADMIN	Quantidade de computadores para uso administrativo
29	NU_COMP_ALUNO	Quantidade de computadores para uso dos alunos
30	INTERNET	Acesso à Internet
31	BANDA_LARGA	Internet banda larga

Fonte: Inep

Por sua vez, as médias de cada escola no ENEM estão disponibilizadas nas 4 áreas de conhecimento, que são: Linguagens, Códigos e suas Tecnologias; Ciências Humanas e suas Tecnologias; Ciências da Natureza e suas Tecnologias e; Matemática e suas Tecnologias. Também consta a média das notas da Redação, outro item avaliado dentro da área de conhecimento Linguagens, Códigos e suas Tecnologias.

RESULTADOS E DISCUSSÃO

Inicialmente, os dois arquivos, contendo as informações sobre a estrutura das escolas e as referentes ao desempenho dos alunos no ENEM, foram unidos e interseccionados em um novo arquivo de dados pela variável que contém o Código de Identificação da Escola, formado por oito dígitos, criado pelo Inep no momento do cadastro da escola no sistema Educacenso, e é comum a ambos os arquivos, originando a base de dados a ser efetivamente analisada. Há um cuidado em identificar correta e unicamente cada estudante, inicialmente por meio do seu CPF, quando informado, ou outras formas de relacionamento para efetivar essa integração. Além disso, é aplicado pelo Inep um processo de consistência, por meio de *script* de verificação fonética, entre os dados informados pelo participante na inscrição do Enem 2015 e os dados cadastrados no Censo Escolar (BRASIL. INEP, 2016).

Na sequência, separamos os dados quanto à dependência administrativa das instituições, como Federais, Estaduais, Municipais e Privadas, com o propósito de mensurar a influência de cada variável ligada à estrutura da escola dos setores público e privado. Em seguida, separamos então as médias obtidas pelas escolas em cada área de conhecimento e redação, conforme o tipo de dependência administrativa (Federal, Estadual, Municipal ou Privada), resultando no arquivo final de dados (Tabela 5).

Tabela 5: classificação das áreas de conhecimento x dependência administrativa

Item	Área de conhecimento	Dependência Administrativa
ENEM_CH_FED ENEM_CN_FED ENEM_LC_FED ENEM_MT_FED ENEM_RED_FED	Ciências Humanas Ciências da Natureza Linguagem, Códigos Matemática Redação	Federal
ENEM_CH_EST ENEM_CN_EST ENEM_LC_EST ENEM_MT_EST ENEM_RED_EST	Ciências Humanas Ciências da Natureza Linguagem, Códigos Matemática Redação	Estadual
ENEM_CH_MUN ENEM_CN_MUN ENEM_LC_MUN ENEM_MT_MUN ENEM_RED_MUN	Ciências Humanas Ciências da Natureza Linguagem, Códigos Matemática Redação	Municipal
ENEM_CH_PRI ENEM_CN_PRI ENEM_LC_PRI ENEM_MT_PRI ENEM_RED_PRI	Ciências Humanas Ciências da Natureza Linguagem, Códigos Matemática Redação	Privada

Fonte: Inep

Após esta classificação, obtivemos 328 observações para cada área de conhecimento para instituições federais, 8836 para instituições estaduais, 109 para instituições municipais e 6324 para instituições privadas.

A partir de então foi possível submeter estes dados para testes de correlação através da linguagem de análise estatística R, utilizando o coeficiente de correlação r , para, por exemplo, as relações da nota média obtida no ENEM por uma instituição federal, na área de conhecimento Linguagens, Códigos e suas Tecnologias, com os indicadores do grupo Água, Luz e Saneamento para essa mesma instituição.

A partir destes, então, foi possível realizar testes de correlação através da linguagem de análise estatística R, comparando, por exemplo, ENEM_LC_FED (nota média obtida no ENEM por uma instituição federal, na área de conhecimento Linguagens, Códigos e suas Tecnologias) × os indicadores do grupo Água, Luz e Saneamento para essa mesma instituição.

Os resultados obtidos ao cruzar os itens que compõem as demais variáveis com as notas médias no ENEM das escolas, em cada área de conhecimento e de acordo com as respectivas dependências administrativas, são apresentados na Tabela 6, onde r é o coeficiente de correlação entre as variáveis e p é a significância dessa correlação. Por economia de espaço, as tabelas apresentam somente os casos em que $r > 0.3$, com significância $p < 0.01$ e que ocorreram mais de 6 vezes.

Não foram encontrados resultados significativos na correlação entre as notas médias no ENEM, em nenhuma das quatro áreas do conhecimento, com as variáveis do grupo Prédio Escolar (Tabela 1). Aparentemente, o fato da escola funcionar em prédio próprio e exclusivo para fins da instituição de ensino ou em salas adaptadas em empresas, unidades prisionais, igrejas, dentre outras não tem grande impacto sobre o desempenho de seus alunos, conforme medido pelas médias no ENEM de seus alunos. Sendo assim, não serão apresentados os resultados referentes a este grupo de variáveis.

Tabela 6: Principais coeficientes de correlação (r) entre médias no ENEM × indicadores de estrutura das escolas para Instituições Federais

Grupo/Variável	Área de conhecimento				
	CH	CN	LC	MT	RED
AGUA LUZ SANEAMENTO					
Esgoto sanitário – Rede pública	0.46	0.43	0.48	0.44	0.38
Esgoto sanitário - Fossa	-0.45	-0.45	-0.47	-0.44	-0.40
Abastecimento de água – Rede pública	0.42	0.35	0.44	0.37	0.35
DEPENDÊNCIAS da ESCOLA					
Alojamento de aluno	-0.34	-	-0.37	-	-0.32

Fonte: Esta pesquisa

Nota: CH – Ciências humanas e suas tecnologias; CN – Ciências da natureza e suas tecnologias; LC – Linguagens, códigos e suas tecnologias; MT – Matemática e suas tecnologias; RED - Redação

Tabela 7: Principais coeficientes de correlação (r) entre médias no ENEM × indicadores de estrutura das escolas para Instituições Estaduais

Grupo/Variável	Área de conhecimento
----------------	----------------------

	CH	CN	LC	MT	RED
AGUA LUZ SANEAMENTO					
Esgoto sanitário – Rede pública	-	-	0.34	-	-
Esgoto sanitário - Fossa	-	-	-0.35	-	-
DEPENDÊNCIAS da ESCOLA					
Número de salas de aula existentes na escola	-	-	0.31	-	-
EQUIPAMENTOS DISPONÍVEIS					
Quantidade de computadores na escola	0.42	0.44	0.41	0.43	0.34
Quantidade de computadores para uso dos alunos	0.38	0.40	0.36	0.40	0.31
Quantidade de computadores para uso administrativo	0.38	0.40	0.39	0.39	0.30
Quantidade de impressoras	-	0.32	0.30	-	-

Fonte: Esta pesquisa

Nota: Vide tabela 6

Tabela 8: Principais coeficientes de correlação (r) entre médias no ENEM × indicadores de estrutura das escolas para Instituições Municipais

Grupo/Variável	Área de conhecimento				
	CH	CN	LC	MT	RED
AGUA LUZ SANEAMENTO					
Esgoto sanitário – Rede pública	0.43	0.41	0.44	0.35	0.33
Esgoto sanitário – Fossa	-0.40	-0.39	-0.41	-0.34	-0.31
DEPENDÊNCIAS da ESCOLA					
Número de salas de aula existentes na escola	-	-	0.30	-	-

Fonte: Esta pesquisa

Nota: Vide tabela 6

Tabela 9: Principais coeficientes de correlação (r) entre médias no ENEM × indicadores de estrutura das escolas para Instituições Privadas

Grupo/Variável	Área de conhecimento				
	CH	CN	LC	MT	RED
EQUIPAMENTOS DISPONÍVEIS					
Quantidade de computadores para uso administrativo	-	-	-	0.30	-

Fonte: Esta pesquisa

Nota: Vide tabela 6

Embora as correlações apresentadas nas Tabelas 6, 7, 8 e 9 não sejam fortes, já se podem observar correlações indicativas entre algumas das variáveis:

- Nota-se que as variáveis do grupo Água Luz Saneamento possuem as maiores correlações com as notas do ENEM. Quando existe uma rede pública de abastecimento de água, encontramos uma correlação de 35% a 44% com as notas médias das escolas federais no ENEM, dependendo da área de conhecimento e do tipo de dependência administrativa.
- Também no o grupo Água Luz Saneamento, observa-se uma correlação negativa de 31% a 47%, quando o esgoto sanitário é por meio de fossa, resultado apresentado nos dados das instituições federais, estaduais e municipais. Quando há uma rede pública de esgoto sanitário, encontramos uma correlação positiva de 33% a 48% nas mesmas dependências administrativas.
- No grupo de Dependências das Escolas, observamos uma correlação negativa para as notas das escolas federais e municipais com a variável Alojamento dos Alunos. Para a variável Laboratório de Ciências, observa-se uma esperada correlação positiva, com maiores valores para as escolas municipais. Ainda neste grupo, encontramos uma correlação positiva, em instituições estaduais e municipais, para a variável que indica o número de salas de aula existentes na escola;
- Finalmente, no grupo Equipamentos Disponíveis as maiores correlações estão registradas, como seria de se esperar, nas variáveis referentes à quantidade de computadores nas instituições de ensino, onde os resultados mais expressivos estão nas de administração estadual.

Na busca pelo entendimento destes resultados, comparamos com os indicados na revisão de literatura pertinentes aos grupos de variáveis nos quais foram identificadas as maiores correlações.

Assim, pode-se verificar que os resultados apresentados nas correlações do grupo Água Luz Saneamento, mais precisamente quando estas se referem ao abastecimento de água e coleta do esgoto sanitário, convergem para os resultados obtidos pelo estudo Universo Escolar da FGV (2016). O citado estudo observou um menor acesso à rede pública de água e de esgoto pelas escolas classificadas com baixo desempenho do que as escolas de alto desempenho. Por outro lado, a revisão literária de Jasper et al. (2012) explica essa relação por uma maior incidência de doenças infecciosas, gastrointestinais e neurocognitivas em crianças em idade

escolar expostas a instalações inadequadas de água e saneamento que impacta negativamente na frequência escolar dos estudantes.

No grupo Equipamentos Disponíveis, as variáveis que indicam a quantidade de computadores tanto para uso administrativo quanto para uso dos alunos, nas instituições estaduais, apresentaram correlações positivas entre 30% e 44%. Tomando os computadores como recurso pedagógico, Mayer (2009), pesquisador na área de cognição, tecnologia e ensino, destaca a importância de desenvolver conteúdos multimídia para auxiliar na aprendizagem. Souza (2014) apresenta a influência da gestão escolar sobre o desempenho dos alunos. Considerando o computador como uma ferramenta para auxílio na tarefa de gerir uma escola, pode-se explicar, em parte, o fato da correlação positiva para a presença de computadores para uso do setor administrativo.

Quanto ao resultado da correlação positiva para a variável que contabiliza o número de salas de aula existentes na escola, encontra-se uma possível relação com a possibilidade de turmas com menor quantidade de alunos, proporcionando uma maior interação entre professor e aluno. Vale a pena destacar que, segundo Soares (2004), o professor possui uma influência muito grande sobre o desempenho dos seus educandos.

Não foram encontrados estudos que pudessem auxiliar na compreensão da correlação negativa para a existência de alojamentos para alunos.

CONSIDERAÇÕES FINAIS

Naturalmente, este estudo não é conclusivo há a necessidade de uma melhor análise sobre os dados obtidos. No entanto, os resultados encontrados até agora, já permitem observar algumas correlações sugestivas.

Observou-se que as variáveis do grupo Água Luz Saneamento possuem as maiores correlações com as notas do ENEM, enquanto que, no grupo Equipamentos Disponíveis, as maiores correlações estão registradas nas variáveis referentes à quantidade de computadores e número de salas de aula existentes nas instituições de ensino.

Mais estudos serão necessários para comprovar a explicação aventada do menor desempenho escolar a partir de redução na frequência escolar dos estudantes como consequência de maior incidência de doenças derivadas de pior acesso às redes públicas de água e esgoto.

O mesmo vale para a compreensão da correlação positiva do desempenho com as variáveis quantidade de computadores e número de salas de aula existentes. Quanto à utilização dos computadores, podemos citar como hipótese, seu uso para auxiliar na gestão da escola e como recurso didático em sala de aula, tornando os momentos de aprendizagem mais atrativos para a geração atual.

Ainda que não seja possível explicar todas as correlações de maneira plena, há uma indicação não desprezível do impacto de cada variável da infraestrutura escolar sobre o desempenho dos alunos, algo não visto nos estudos encontrados durante a revisão bibliográfica.

O fato de a correlação mais importante ter sido com o acesso à rede pública de água e de esgoto indica que uma melhoria significativa do desempenho escolar poderia ser obtida apenas ao se estender saneamento básico às escolas.

Finalmente, é de se ressaltar o potencial da utilização da Ciência de Dados como metodologia de análise.

AGRADECIMENTOS

Os autores agradecem à FAPERGS pela bolsa PROBIC que viabilizou esta pesquisa.

REFERÊNCIAS

- BRASIL. INEP. *Nota Explicativa ENEM 2015 por escola*. Brasília: INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, 2016. Retrieved from http://download.inep.gov.br/educacao_basica/enem/nota_tecnica/2015/nota_explicativa_e_nem2015_por_escola.pdf
- CARVALHO, José Sérgio Fonseca de. Os ideais da formação humanista e o sentido da experiência escolar. *Educação e Pesquisa*, São Paulo, v. 43, n. 4, p. 1023-1034, out./dez., 2017.
- FAYYAD, Usama; PIATETSKY-SHAPIRO, Gregory; SMYTH, Padhraic. From Data Mining to Knowledge Discovery in Databases. *AI Magazine*, v. 17, n. 3, p. 87, 1996.
- FGV. *Universo escolar: o estado da educação no Brasil*. Coordenação Marco Aurélio Ruediger. Rio de Janeiro: FGV, DAAP, 2016. 26 p.: il. Disponível em: <http://dapp.fgv.br/universo-escolar/>. Acesso em: 30 de agosto de 2018.
- JASPER, Chistian; LE, Thanh-Tam; BARTRAM, Jamie. Water and sanitation in schools: a systematic review of the health and educational outcomes. *International Journal of Environmental Research and Public Health*, v. 9, p. 2772– 2787. 2012
- MAYER, Fernanda Gimenes; RODRIGUES, Waldemar. A influência do capital humano sobre o desenvolvimento econômico: um olhar sobre a educação. *Revista de Administração do UNISAL*, v. 3, n.3, p.1-16, abr. 2013.

- MALUSÁ, Silvana; ORDONES, Luma Lemos de Medeiros; RIBEIRO, Elisabete. Enem: pontos positivos para a educação brasileira. *Revista Educação e Políticas em Debate*, v.3, n.2, p. 358-382, ago./dez 2014.
- PENG, Roger D. *Exploratory Data Analysis with R*. Victoria, CA-BC: Leanpub, 2016.
- ROCHA, Fernando José Meira. Educação e economia: uma abordagem sobre as consequências e condicionantes econômicos do desenvolvimento humano, com ênfase em educação. *Cadastro de Finanças Públicas, Brasília*, n. 5, p. 51-171, 2004.
- SOARES, José Francisco. O efeito da escola no desempenho cognitivo de seus alunos. *REICE: Revista Iberoamericana sobre Calidad, Eficacia y Cambio en Educación*, vol. 2, n. 2, p. 83-104, 2004.
- SOUZA, Caroline Beatriz Rodrigues de. *Disfuncionalidade escolar: uma análise teórica de identificação dos fatores que afetam o desempenho das escolas no Brasil*. UFRGS. Porto Alegre, 2014.
- WICKHAM, Hadley; GROLEMUND, Garrett. *R for Data Science: Import, Tidy, Transform, Visualize, and Model Data*. Sebastopol: O'Reilly, 2016. 522 p.

Autores

Giuliano Nunes

Licenciado em Física e Bolsista de Iniciação Científica
Universidade Luterana do Brasil.
giulianonunes0@gmail.com.

Renato P. dos Santos

Doutor em Física
Docente e pesquisador no
Programa de Pós-Graduação em Ensino de Ciências e Matemática (PPGECIM)
da Universidade Luterana do Brasil.
renatopsantos@ulbra.edu.br