

PONTIFÍCIA UNIVERSIDADE CATÓLICA
DO RIO DE JANEIRO



DEPARTAMENTO DE ECONOMIA
MONOGRAFIA DE FINAL DE CURSO

**O USO DE PLATAFORMAS DE ENSINO ONLINE E O IMPACTO SOBRE O
RESULTADO ACADÊMICO DE ALUNOS DO ENSINO BÁSICO**

Luiza Puga Columa

Nº de matrícula: 1312740

Orientador: Claudio Ferraz

Rio de Janeiro,

Julho 2017

PONTIFÍCIA UNIVERSIDADE DO RIO DE JANEIRO
DEPARTAMENTO DE ECONOMIA
MONOGRAFIA DE FINAL DE CURSO

**O USO DE PLATAFORMAS DE ENSINO ONLINE E O IMPACTO SOBRE O
RESULTADO ACADÊMICO DE ALUNOS DO ENSINO BÁSICO**

Luiza Puga Columa

Nº de matrícula: 1312740

Orientador: Claudio Ferraz

Declaro que o presente trabalho é de minha autoria e que não recorri para realizá-lo, a nenhuma forma de ajuda externa, exceto quando autorizado pelo professor tutor.

Luiza Puga Columa

Rio de Janeiro,

Julho 2017

As opiniões expressas neste trabalho são de responsabilidade única e exclusiva do autor.

Agradecimentos

Agradeço aos meus pais, Cesar Puga e Luciana Andrade, por todo o apoio com a minha educação e com o meu preparo emocional para a vida. À Marina Puga, minha irmã, e ao Ricardo Marcondes, meu cunhado, por todo o apoio e companheirismo desde sempre e por terem lido minha monografia e dado insights importantes para seu aprimoramento. Aos meus tios, Renata Moura e Luis Henrique Moura por todo o incentivo e por terem contribuído para a minha educação e orientação profissional.

Agradeço aos meus amigos Julio Barros e Roni Band por estarem sempre dispostos a me auxiliar quando eu tinha dúvidas na utilização do software R e por todo o apoio emocional que me deram durante a confecção desse trabalho. Aos meus demais colegas de curso, pelo convívio durante esta fase e por terem me premiado com grandes amizades que carregarei para sempre.

Agradeço a todos meus professores pelo conhecimento e ensinamentos passados durante os últimos anos e em especial meu orientador, Claudio Ferraz, por toda disponibilidade e aprendizado durante a elaboração deste trabalho.

Sumário

1.	Introdução -----	6
2.	Revisão da Literatura -----	7
3.	Motivação -----	10
4.	Dados -----	11
	I. Fonte: A Eleva Educação-----	11
	II. Dados utilizados -----	11
5.	Modelos -----	15
6.	Resultados -----	28
7.	Conclusão -----	30
8.	Referências Bibliográficas -----	31
9.	Gráficos e Tabelas -----	32

Lista de tabelas e Gráficos

Gráfico 1: Resultados PM 1ª Série EM Pensi -----	13
Tabela 6: Resultados regressão <i>naive</i> t=6 -----	16
Gráfico 2: Utilização de Canal Eleva -----	17
Gráfico 3: Utilização de Monitoria Online -----	18
Gráfico 4: Utilização de Atualidades -----	18
Gráfico 5: Utilização de Eleva ao vivo -----	19
Gráfico 6: Nova utilização de Canal Eleva -----	20
Gráfico 7: Nova utilização de Monitoria Online -----	20
Gráfico 8: Nova utilização de Atualidades -----	21
Gráfico 9: Nova utilização de Eleva ao Vivo -----	21
Gráfico 10: Nova nova utilização de Monitoria online -----	22
Tabela 12: Resultados Regressão <i>naive</i> após retirada dos <i>outliers</i> em t=6 -----	23
Tabela 13: Resultados regressão com controle de notas defasadas -----	25
Tabela 14: Resultado da regressão em painel sobre a nota geral -----	28

As demais tabelas e gráficos se encontram ao final deste documento, a partir da página 32, em um capítulo destinado apenas a expô-las.

1. Introdução

A educação é o principal meio de ascensão social na vida de milhões de pessoas no Brasil e no mundo. Diversos economistas que estudaram desigualdade social e pobreza concluíram que a renda individual está direta e positivamente relacionada ao número de anos dedicados à educação. Ou seja, quanto mais uma criança ou jovem consegue se dedicar a seus estudos e se manter no ambiente escolar, mais chances ela terá de ascender socialmente no longo prazo. A explicação para isso está ligada à produtividade marginal do trabalho, que é definida por 3 principais fatores: capacidade nata; escolha da atividade laboral; e oportunidades.

Se considerarmos que os dois primeiros fatores não são facilmente influenciáveis, pois dependem das habilidades e vontades individuais, nos resta trabalhar em cima das oportunidades para garantir que a produtividade marginal do trabalho aumente e possa levar crianças e jovens à ascensão social.

Quando falamos de oportunidades podemos nos referir a uma gama enorme de variáveis que permeiam as vidas dos indivíduos por acaso ou devido a escolhas feitas por eles ao longo dos anos. As instituições escolares representam um grande motor precursor dessas oportunidades. Com o conjunto de ferramentas, experiências e metodologias que elas disponibilizam para cada aluno, conseguem moldar o desenvolvimento de crianças e jovens impactando seus futuros.

Contudo, nem todas as escolas são iguais e o resultado que elas imprimem no desenvolvimento de seus alunos pode variar bastante, evidenciando assim a diferença de oportunidades. Uma possível explicação para essas diferenças é a utilização de métodos de ensino diferentes ou a disponibilização de ferramentas de ensino diferenciadas.

Com o estudo das melhores práticas das escolas no país, é possível encontrar uma das chaves para desenvolver a qualidade do ensino de muitas escolas pelo Brasil. Se uma ferramenta utilizada em uma grande rede de escolas privadas gera impacto positivo sobre o resultado acadêmico dos seus alunos e é escalável, pode-se comercializar a um custo baixo, será possível dar melhores oportunidades de desenvolvimento a milhões de crianças e jovens gerando um impacto positivo no desenvolvimento socioeconômico do país como um todo.

2. Revisão da Literatura

Como a educação mudou nos últimos cinquenta anos? Como as salas de aula mudaram nesse período?

Ao longo das últimas décadas a tecnologia avançou em todas as esferas da vida humana. O campo da saúde evoluiu muito. Doenças que há pouco tempo traziam risco de vida para milhares de pessoas hoje possuem tratamentos modernos com uso de novos equipamentos que reduziram muito a taxa de mortalidade. O campo da economia mudou radicalmente com a introdução da econometria e de softwares que nos ajudam a analisar dados e precificar produtos com uma rapidez inimaginável há cinquenta anos atrás. Até mesmo o lazer mudou com a tecnologia. Hoje, o que crianças e adolescentes consideram como atividades prazerosas para seus tempos livres, como vídeo games, redes sociais, utilização de smartphones, há pouco tempo atrás nem existiam. Como no âmbito da educação, essa onda de modernização atingiu escolas e universidades?

Com o avanço das tecnologias é possível ver a diversificação do ensino. Ou seja, o conteúdo que antigamente seria absorvido apenas por alunos que estivessem sentados em sala de aula, pode ser acessado online por meio de vídeo aulas, jogos baseados no conteúdo, maior volume de materiais disponíveis e acessíveis a qualquer momento. Neste contexto, é preciso avaliar se tais mudanças estão sendo positivas ou não para o aprendizado.

No universo da graduação, existe uma tendência de se trocar aulas presenciais por aulas online. Um terço dos estudantes escolhe cursar no mínimo uma de suas aulas à distância, ou seja, online, durante a sua formação e esse número cresceu três vezes na última década, segundo estudos de Allen and Seaman¹ de 2013. Com isso as universidades ganham escala, pois um bom professor que antes ensinava para quarenta alunos em uma sala de aula, consegue ensinar mais de duzentos por meio da interação online.

Entretanto, uma pesquisa feita pelo centro de políticas educacionais da universidade de Stanford², utilizando dados da DeVry University, mostram que alunos que optam por aulas online obtêm resultados substancialmente piores do que aqueles que optam por assistir aulas de forma presencial. Ou seja, o efeito da substituição do conteúdo lecionado presencialmente pelo conteúdo lecionado online é negativo sobre o aprendizado dos jovens. Isso pode ser reflexo da qualidade das aulas em si ou da falta de engajamento dos

alunos decorrente da pouca interação que existe quando o conteúdo é transmitido por meio da tela de um computador.

Entretanto, é importante ressaltar, como é feito por Eric Bettinger² no artigo “Changing Distributions: How online college classes alter student and professor performance”, que tais resultados não são estáticos, ou seja, é possível que a qualidade do ensino online melhore com o tempo e que inclusive venha a superar as aulas presenciais. Uma pesquisa feita em 2003 por uma professora da University of California, San Francisco (UCSF)² com líderes acadêmicos, e posteriormente repetida em 2012, mostra que a percepção geral dos professores em relação a qualidade do ensino online de suas instituições tem melhorado nos últimos anos.

Outro ponto a ser destacado é o possível ganho de bem-estar que a criação de cursos online gera no agregado da sociedade. Por mais que alunos que optam por cursos online tenham uma qualidade de aprendizado pior, se não houvesse a opção online talvez muitos alunos nem tivessem a oportunidade de estudar. Com um custo marginal muito baixo, pelo fato de que ao disponibilizar uma vídeo-aula o custo de produção será o mesmo independentemente do número de pessoas que assistirem, quando uma universidade decide disponibilizar um curso online ela pode reduzir o preço cobrado por essa formação, gerando um ganho muito grande para aqueles alunos que antes, pelo preço cheio, não teriam acesso ao conteúdo.

Até aqui foi analisada a possibilidade de os estudos online substituírem as salas de aula. Adicionalmente, tais conteúdos podem ser utilizados como apoio para os estudos de sala de aula e não apenas como substitutos. Neste caso, será que a influência sobre os resultados acadêmicos dos alunos mudaria? Estudos nesse campo ainda são poucos e creio que isso se deve ao fato de que para se calcular o impacto da utilização de ferramentas online no aprendizado de crianças e jovens é necessário ter uma base de informações muito completa com notas de alunos em provas padronizadas e utilização de ferramentas online. Hoje existem muitas iniciativas online, algumas até gratuitas como a Kahn Academy, que buscam expandir a educação de qualidade para todos. Mas essas iniciativas não estão diretamente ligadas a alguma avaliação de desempenho realizada por alunos.

Para ir a fundo nesse tema é necessário fazer um experimento ou então encontrar uma grande rede de ensino básico que utilize alguma iniciativa online de estudos e

disponibilize seus dados, como é o caso da Eleva Educação, que serviu me cedeu sua base de dados para realizar a análise presente neste trabalho.

3. Motivação

Quem nunca ouviu que tal escola é melhor que outra? Ou que aquela instituição de ensino é muito mais séria? Comentários assim vêm de diferentes razões como qualidade do corpo docente, nível de dificuldade das avaliações, capacidade de gestão dos diretores escolares e ferramentas de ensino utilizadas pela instituição escolar.

Esta última variável pode ser muito importante se acreditarmos que uma ferramenta específica utilizada em uma escola pode ser escalável e distribuída para diversas escolas, dando melhores oportunidades de estudo para milhões de alunos e gerando impacto positivo no ensino. Por isso, esse trabalho tem como objetivo encontrar uma relação de causalidade entre a utilização de uma ferramenta, que chamaremos aqui de plataforma de ensino online, e os resultados acadêmicos dos alunos.

A Eleva educação, uma grande rede de ensino básico do Brasil, cedeu dados privados de utilização de sua plataforma de ensino online, assim como do resultado de seus alunos em provas padronizadas, para que este estudo pudesse ser realizado. Descobrir o impacto de tal plataforma pode ajudar a empresa em sua missão de transformar vidas por meio da educação. A ferramenta online já foi criada há dois anos e tem seu conteúdo constantemente atualizado e aprimorado. Além disso, por ser online, o custo fixo médio por aluno apenas diminui à medida em que mais e mais alunos utilizam essa ferramenta.

A grande motivação deste estudo é encontrar um resultado que mostre efeitos positivos da plataforma de ensino online no resultado acadêmico dos alunos e, assim, influenciar a disseminação dessa ferramenta por todo o Brasil afim de melhorar o ensino e dar mais oportunidade de ascensão social a crianças e adolescentes.

4. Dados

I. Fonte: A Eleva Educação

Com o objetivo de analisar a influência da plataforma de ensino online sobre o resultado acadêmico dos alunos, a análise foi baseada em dados do segundo maior grupo educacional do Brasil, a Eleva Educação.

A Eleva é uma empresa que teve início em 2013 quando o fundo de investimentos Gera Venture, que tem como principal investidor o megaempresário Jorge Paulo Lemann, comprou duas redes de escolas importantes no Rio de Janeiro, o Pensi e o Elite. Essas escolas privadas são voltadas para as classes B e C, tendo como foco garantir o sucesso profissional na vida de seus estudantes.

Nos anos seguintes mais duas redes de escolas foram adicionadas ao grupo, o Colegium de Minas Gerais e o Alfa no Paraná. Ao mesmo tempo, uma plataforma de ensino online foi criada com o objetivo de complementar o ensino em sala de aula e auxiliar os alunos na hora do estudo.

Hoje, a Eleva é a segunda maior holding de educação básica do Brasil em número de alunos, totalizando aproximadamente 30 mil alunos em escolas próprias e 65 mil alunos estudando por meio de seu material didático*, que é comercializado para escolas parceiras, colocando a empresa entre as seis maiores do mundo no ramo.

*Dados de 2016

II. Dados utilizados:

A Eleva educação, em 2016, possuía 54 unidades escolares próprias espalhadas em 3 estados do Brasil: Rio de Janeiro, Minas Gerais e Paraná. A rede possuía 31.273 alunos dos quais 6.288 cursavam ensino fundamental II e 5.458 cursavam o ensino médio. Os demais alunos estavam alocados no ensino infantil, fundamental I, pré-vestibular e cursos livres.

O objetivo desse trabalho é analisar os resultados acadêmicos dos 11.746 estudantes, entre o 6º ano a 2ª série do ensino médio, ao longo do ano de 2016 e comparar com o número de vezes que os mesmos acessaram a plataforma de ensino online. Tentando com isso encontrar uma influência positiva sobre tais resultados.

A plataforma de ensino online criada pela Eleva Educação é dividida em 4 principais ferramentas. São elas: Atualidades, Canal Eleva, Eleva ao Vivo e Monitoria Online. Abaixo segue uma breve descrição de cada uma delas.

(i) Em Atualidades os alunos encontram, a cada 15 minutos, um tema atual novo disponibilizado em formato de texto motivador. Com ele o estudante pode treinar redação, realizar debates em sala de aula ou ter ideias para projetos de disciplinas específicas.

(ii) O Canal Eleva funciona como um depósito de vídeo-aulas dos mais variados assuntos que os alunos podem acessar a qualquer momento para estudar. Os vídeos podem ser filtrados por matéria, conteúdo ou módulo do livro didático, facilitando a integração do estudo dentro e fora de sala de aula.

(iii) O Eleva ao vivo funciona como um programa de televisão. De segunda a quinta feira, das 19h às 22h, aulas ao vivo são exibidas sobre temas específicos como preparação para a olimpíada de matemática, história da arte, integração física e química, entre outros. Nessa ferramenta os alunos também podem interagir enviando perguntas que serão respondidas ao vivo pelos professores e além disso, nas vésperas de prova as aulas são sempre sobre o conteúdo que será cobrado.

(iv) A Monitoria Online funciona para aqueles alunos que não conseguiram sanar todas as suas dúvidas em sala de aula. O estudante pode entrar na ferramenta e solicitar a resolução de uma questão específica do livro e esperando apenas alguns minutos ele recebe um gabarito comentado do exercício para auxiliar o aprendizado.

Para fins dessa pesquisa foram utilizados os dados de número de visualizações que cada aluno teve em cada uma das 4 ferramentas citadas acima em 6 períodos diferentes do ano como variáveis explicativas. A variável dependente será o resultado de cada aluno nas 6 provas multidisciplinares realizadas ao longo do ano.

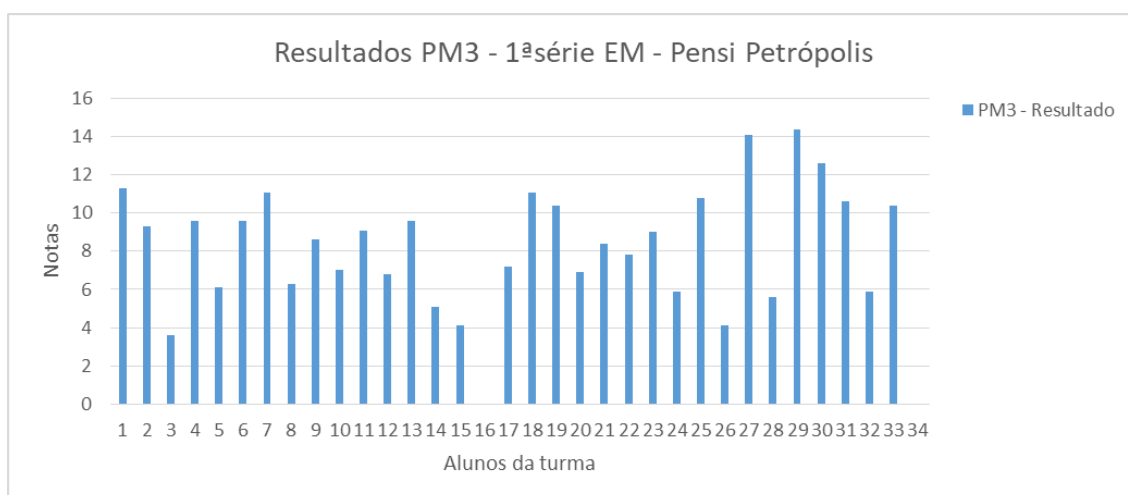
Com o objetivo de garantir um resultado não viesado, mediremos o resultado acadêmico dos alunos de 6º ano do ensino fundamental 2 à 2ª série do ensino médio, por meio de provas padronizadas submetidas a uma correção centralizada. Esse modelo de avaliação garante que diferenças no nível de cobrança dos professores, ao criar ou corrigir uma prova, ou até mesmo diferenciações que professores podem fazer entre alunos, não causem viés no estudo. O grupo de professores que cria e corrige as provas não tem nenhum contato com alunos próprios da rede Eleva.

O modelo das provas que foram utilizadas nas análises é conhecido como Prova Multidisciplinar, ou PM. Essa avaliação abrange as quatro áreas de conhecimento cobradas no ENEM - ciências da natureza, ciências humanas, linguagens e matemática - e é aplicada nos mesmos dias e horários para todos os alunos da rede. Ao longo do ano, esse tipo de avaliação é aplicada 6 vezes e as notas podem ir de zero a 20, sendo 5 pontos referentes a cada área do conhecimento.

Para ilustrar, o gráfico 1 abaixo mostra os resultados das turmas de primeira série do ensino médio da unidade de Petrópolis da rede Pensi na PM3 aplicada no dia 11 de junho de 2016.

O eixo X apresenta os alunos da primeira série do ensino médio da unidade citada, identificados por números. Enquanto o eixo Y apresenta a nota final dos alunos naquela PM.

Gráfico1:



Ao analisarmos os dados é possível ver que existe uma grande diferença entre as notas dos alunos de uma mesma unidade e se compararmos com unidades diferentes teremos mais variação ainda. A grande diferença encontrada pode ser resultado de inúmeras variáveis. Com isso, o desafio aqui é isolar apenas o efeito da utilização das ferramentas de estudo online.

Uma grande vantagem de se utilizar uma rede grande de escolas é que alguns efeitos que teriam muita influência sobre o resultado dos alunos podem ser considerados homogêneos no grupo, não causando assim qualquer tipo de viés. Por exemplo, a diferença de qualidade dos professores pode ser considerada mínima nesse estudo. Isso

ocorre por que a distribuição de professores entre as unidades escolares que a empresa possui se dá de forma parcialmente aleatória. Os docentes informam os horários que eles possuem disponíveis para dar aulas em um sistema, chamado Urânia, que gera uma combinação entre o horário dos diversos professores para construir a grade escolar de todas as unidades. Ao rodar esse sistema não existem preferências entre professores e unidades escolares. Considera-se apenas tempo de deslocamento de uma unidade para outra e a disponibilidade fornecida. Além disso, considerando que os critérios de seleção dos professores que fazem parte da rede é o mesmo, feito com base no currículo, em provas presenciais, e em uma aula expositiva ministrada pelo candidato a professor, e que esses critérios são avaliados pelo mesmo grupo de profissionais, podemos considerar que o padrão de qualidade docente de todas as unidades é aproximadamente igual e que isso não causaria nenhum efeito viesado por unidade.*

Além disso, para reforçar essa hipótese, vale destacar que os professores que fazem parte da rede Eleva têm pouca liberdade ao conduzir suas aulas. Ou seja, por precisarem seguir o mesmo cronograma de avaliações que é aplicado para todo o Brasil e utilizarem o material didático padronizado do grupo, as aulas acabam um pouco engessadas e sem espaço para os docentes fugirem do conteúdo e metodologia pré-definidos.

*Essa hipótese é forte, porém fundada e necessária para o andamento do estudo.

5. Modelos

Para montar o melhor modelo para análise desses dados optei por começar com regressões mais simples e ir aprimorando o modelo a cada etapa de análise. Este capítulo descreve todas as etapas percorridas até o modelo final utilizado.

1ª Etapa: Regressão *Naive*

Na primeira etapa utilizei os dados de forma rústica e gerei uma regressão simples, comumente chamada de regressão *naive*, pelo método dos mínimos quadrados ordinários onde tento explicar as notas dos alunos nas provas apenas por meio do número de acessos à plataforma por período, como mostra a fórmula abaixo.

Equação 1:

$$pm_t = \beta_1 canal_pm_t + \beta_2 monitoria_pm_t + \beta_3 atualidades_pm_t + \beta_4 aovivo_pm_t$$

Em minha primeira equação, o t representa o período em que a prova ocorreu e com isso eu rodei seis regressões considerando o t de 1 a 6. Os coeficientes β_i representam a influência da utilização das ferramentas de estudo online, indicadas por canal_pmt, monitoria_pmt, atualidades_pmt e aovivo_pmt, têm sobre a nota dos alunos (pmt).

A regressão descrita acima foi rodada para todos os seis períodos do ano, obtendo assim seis tabelas descritivas diferentes (tabelas 1 a 6 presentes ao final deste documento).

Essa regressão simples é chamada de *naive* pois é apenas com muita ingenuidade que poderíamos acreditar que os resultados obtidos com ela são fidedignos. Isso se dá devido ao fato de que os dados não passaram por nenhum tipo de tratamento e a regressão em si não se preocupa em expurgar nenhum tipo de viés que possa existir.

Dito isso, ao aplicar esse primeiro modelo a todos os 6 períodos encontrei, em geral, efeitos positivos das ferramentas sobre as notas. Apesar dos efeitos serem aparentemente pequenos e variarem bastante entre um período e o outro, podemos observar um alto nível de significância, como podemos ver na tabela 6 de estatística descritiva abaixo. A tabela mostra, como exemplo, o resultado obtido para o período 6.

Tabela 6:

Resultados Regressão naíve em $t=6$

Dependent variable:	

PM6	

Canal Eleva	0.001 (0.003)
Monitoria Online	0.003*** (0.0005)
Atualidades	0.206*** (0.015)
Eleva ao Vivo	-0.446*** (0.073)
Constante	10.153*** (0.036)

Observations	9,906
R2	0.029
Adjusted R2	0.029
Residual Std. Error	3.165 (df = 9901)
F Statistic	74.352*** (df = 4; 9901)
=====	
Note:	* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Na tabela podemos ver que a ferramenta Atualidades parece influenciar mais a nota dos alunos, dentre as quatro ferramentas apresentadas, e que a ferramenta Eleva ao Vivo parece ter um efeito negativo sobre a nota dos alunos. Além disso, vale ressaltar que, apesar de termos calculado um coeficiente para a ferramenta Canal Eleva, esse resultado tem nível de significância muito baixo, indicando que provavelmente o coeficiente é igual a zero nesse caso. Outro ponto importante de se observar é o R^2 da regressão, ou seja, a medida que indica o quanto o modelo estimado consegue explicar os valores observados da amostra. O R^2 desse modelo é muito baixo, indica que 2,91% da variável dependente consegue ser explicada pelos regressores do modelo.

Esse resultado encontrado, apesar de ainda estar distante do modelo mais confiável, é satisfatório pois mostra que provavelmente existe uma relação entre a utilização das ferramentas e as notas que os alunos obtêm em suas provas.

Para refinar os resultados analisei os dados de utilização da plataforma de ensino online a fim de excluir eventuais *outliers*, ou seja, padrões de utilização

muito fora da curva poderiam estar viesando o resultado e por essa razão preferi expurgar essas observações de minha análise.

2ª Etapa: Retirando *outliers*

Para retirar as observações de utilização da ferramenta que estão muito fora da curva optei por plotar o gráfico de número de acessos agregados no ano para cada ferramenta individualmente e por alunos, de forma que cada bolinha nos gráficos representa a soma da utilização da ferramenta em questão durante todo o ano por um aluno específico. Os quatro gráficos obtidos (gráficos 2 à 5) estão ilustrados nas figuras abaixo:

Gráfico 2:

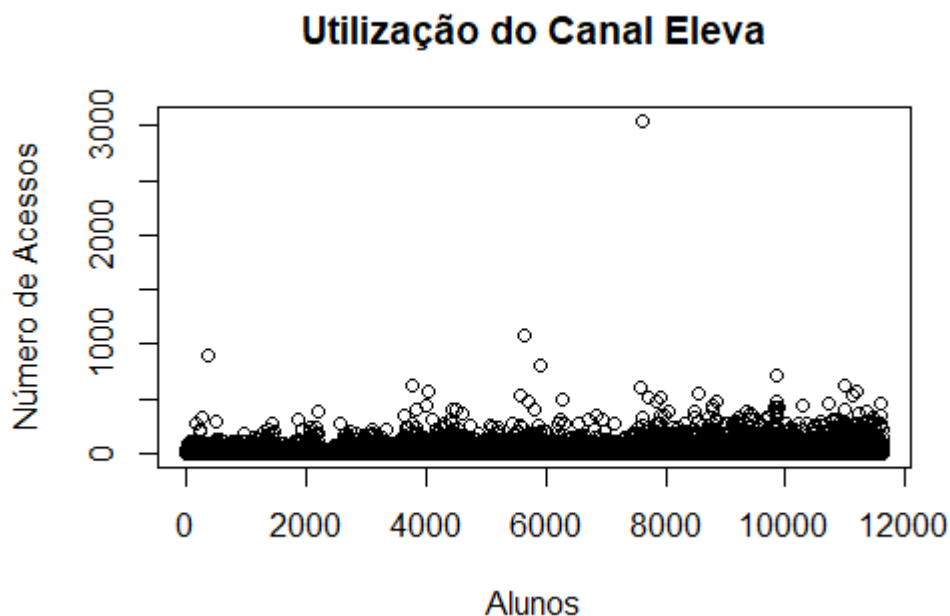


Gráfico 3:

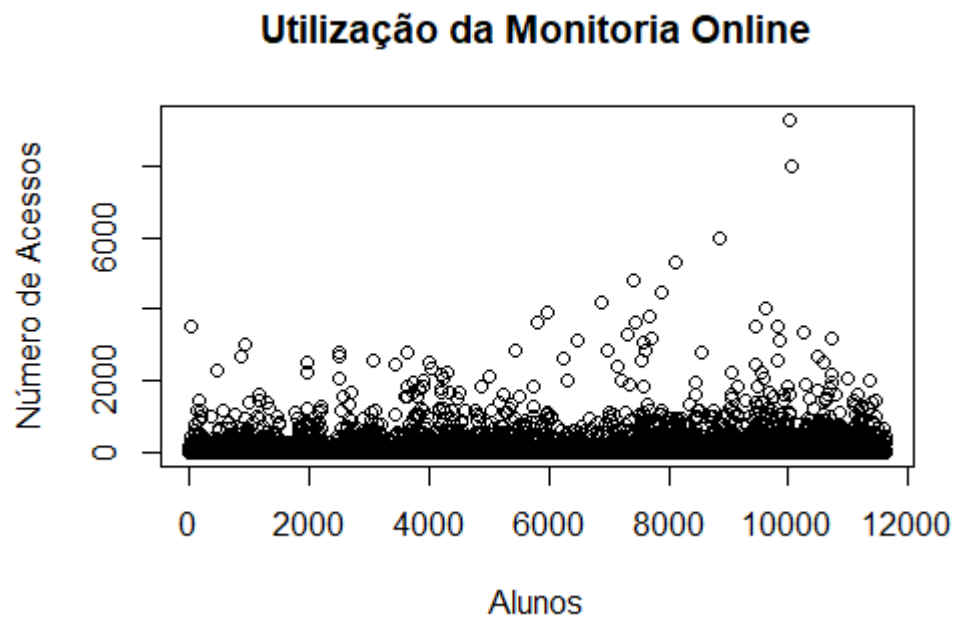


Gráfico 4:

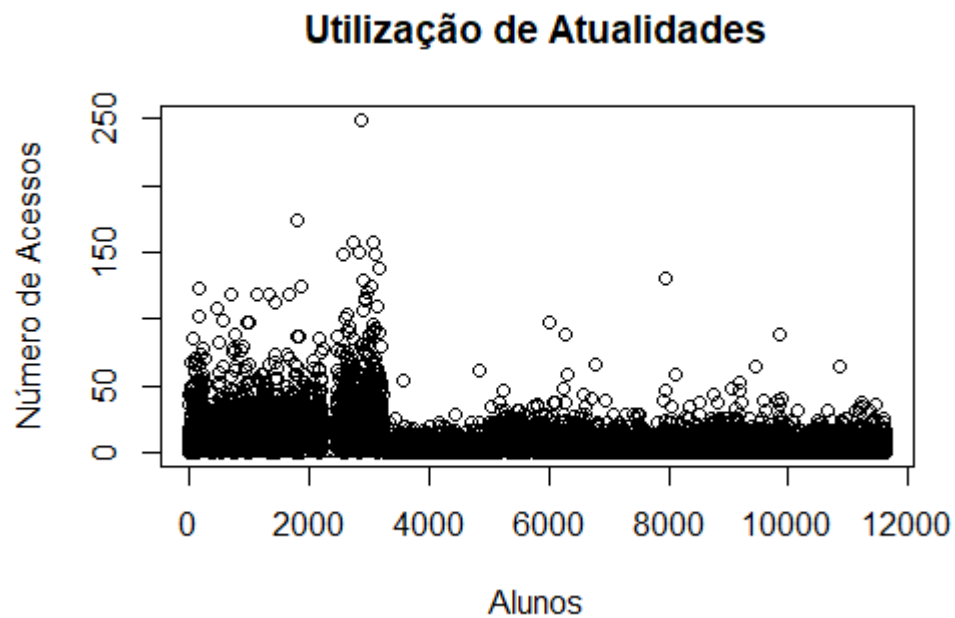
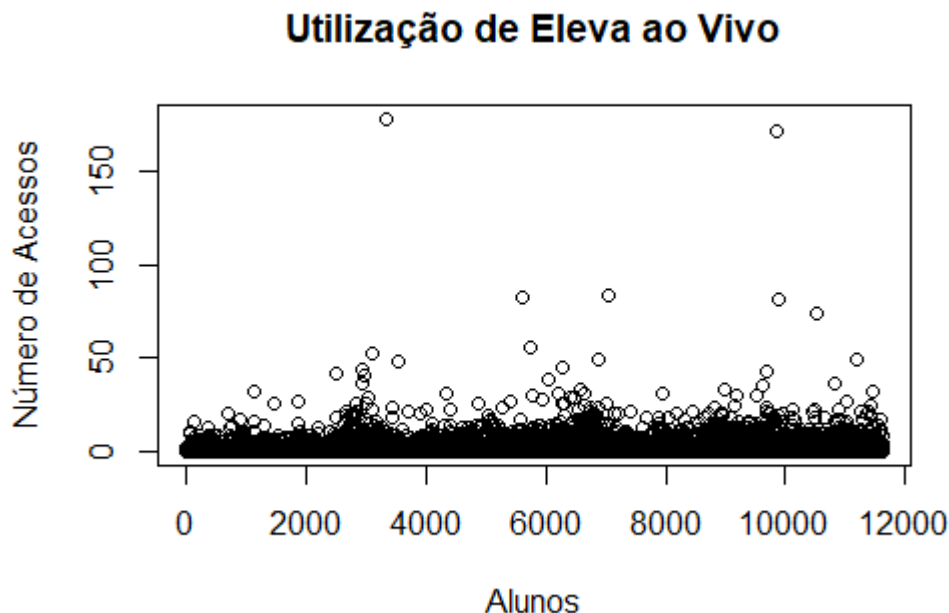


Gráfico 5:



Observando as imagens podemos ver que, para todas as ferramentas, existe uma enorme concentração de dados e alguns pontos se destacam como fora da curva. Essas observações, por terem um padrão muito distinto das demais, podem ser resultado de alguma falha no computador de algum aluno específico ou alguma testagem de equipamento em que a plataforma online foi acessada sucessivamente. Para que isso não interfira nos resultados da análise, optei por retirar da amostra todos os alunos que possuem padrões de utilização da ferramenta fora do comum.

Para tal feito, calculei o percentil 99% da utilização de cada ferramenta e retirei todos os alunos que estavam com número de acessos acima desse percentil, ou seja, retirei 1% dos alunos que mais utilizaram cada ferramenta.

Para verificar se a escolha pelo percentil 99% foi boa, plotei novamente o gráfico de número de acessos, por tipo de ferramenta e por aluno, e comparei com as imagens anteriores. Os gráficos abaixo (gráficos 6 à 9) ilustram a distribuição das observações de utilização das ferramentas após a retirada dos *outliers*.

Gráfico 6:

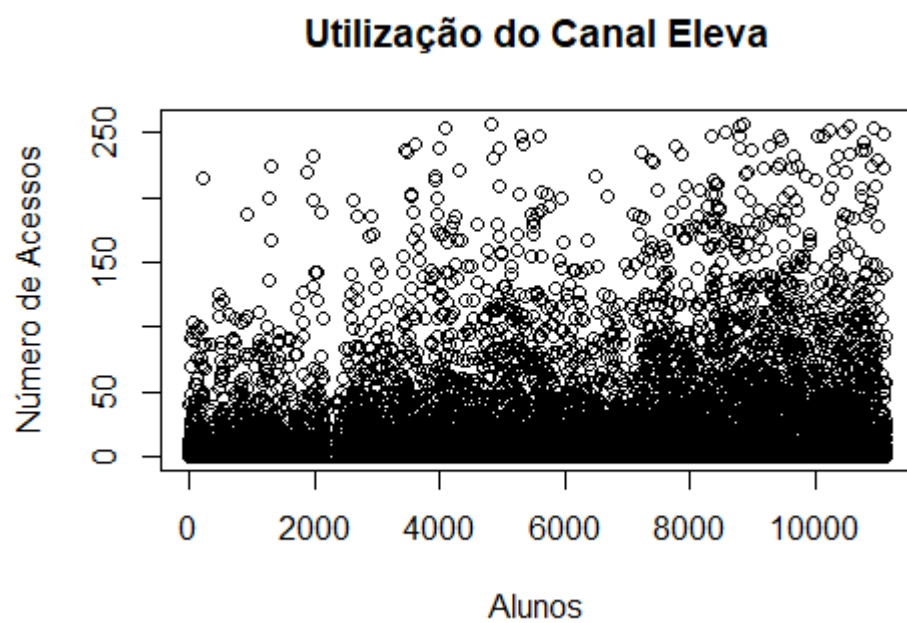


Gráfico 7:

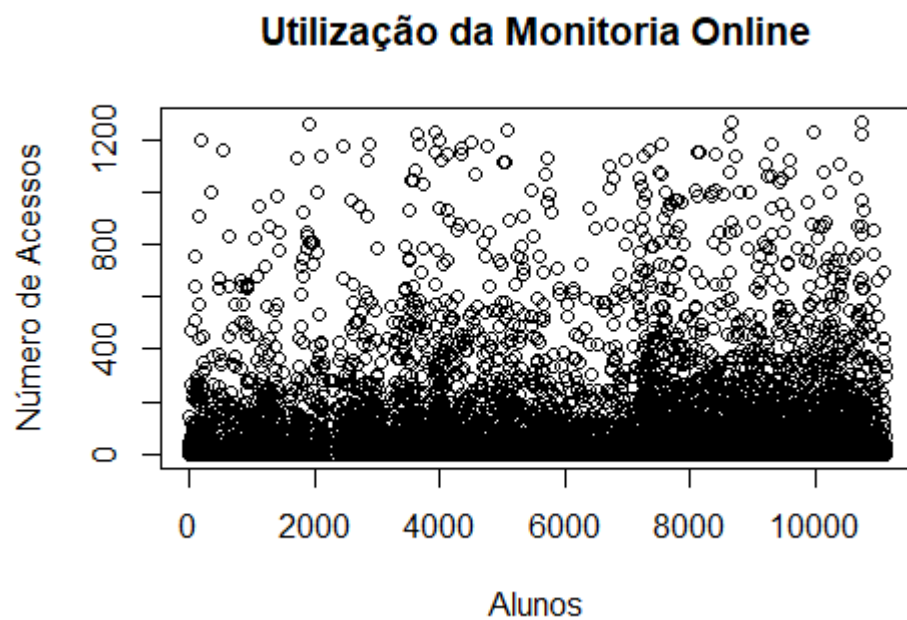


Gráfico 8:

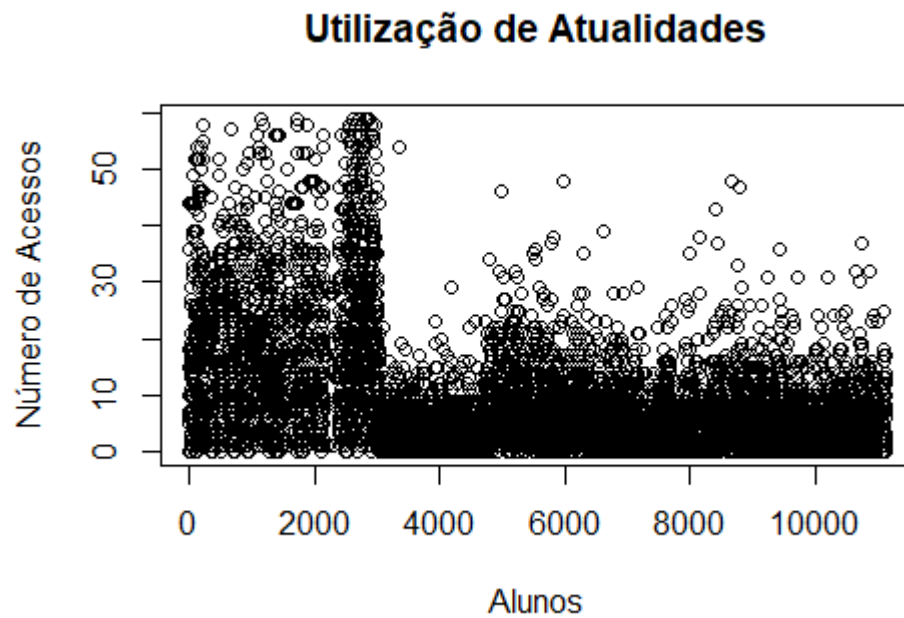
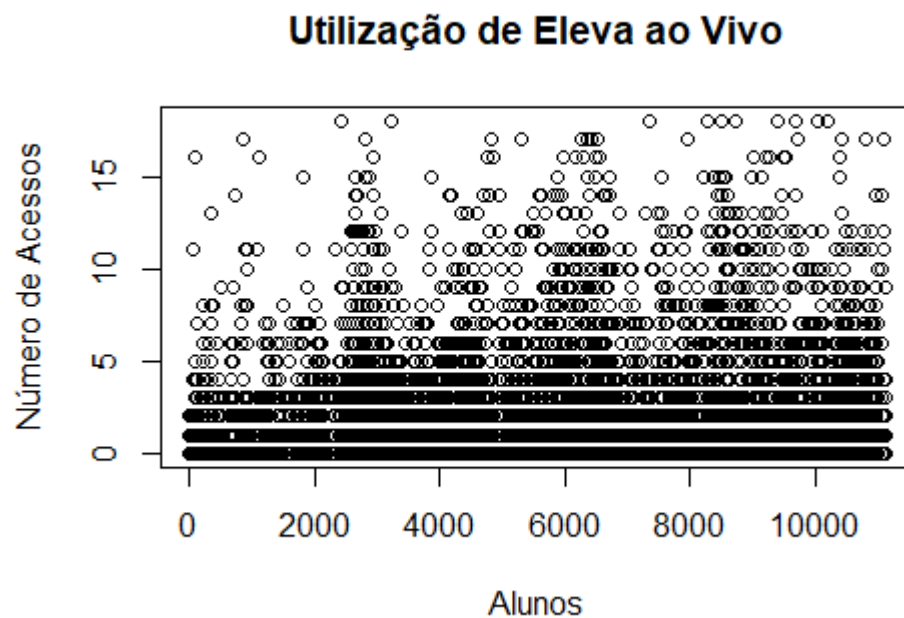


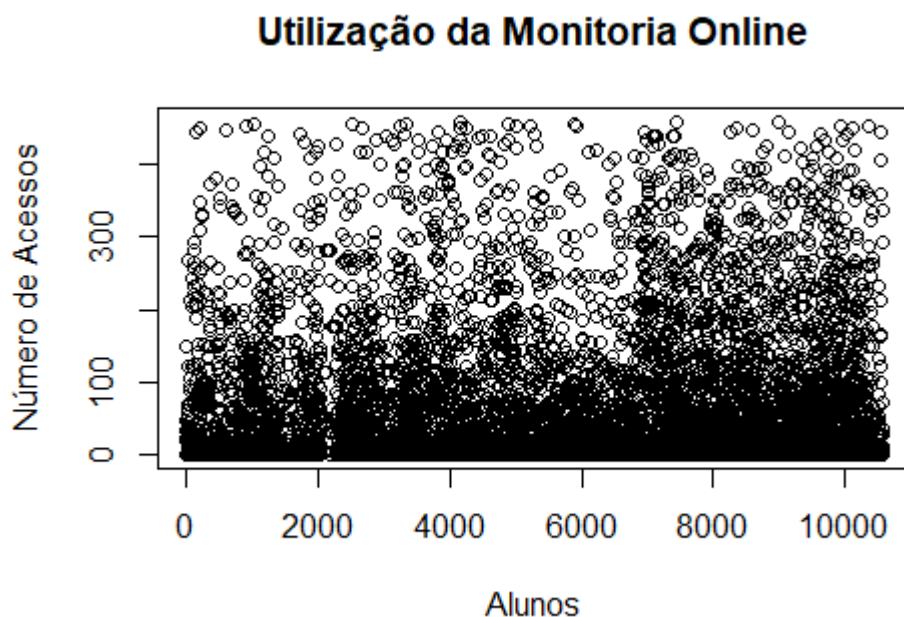
Gráfico 9:



Em primeira mão, as novas distribuições parecem mais harmônicas. Porém, analisando a média e mediana dos dados de utilização da ferramenta Monitoria Online, notei que ainda havia uma dispersão muito grande dos dados e que alguns alunos ainda

apresentavam padrões de acesso muito fora do que se espera. Sendo assim, foi realizado um novo ajuste retirando da amostra todos os alunos cuja utilização da plataforma de estudos online na ferramenta de Monitoria estava acima do percentil 95% da distribuição e com isso distribuição final deixou de apresentar distorções anormais, como podemos ver no novo gráfico gerado.

Gráfico 10:



É interessante observar, no gráfico 8, que existe uma concentração de alunos que utilizam mais a ferramenta Atualidades do que outros. Isso se deve ao fato de que uma das redes de escolas da Educação Básica, chamada Colegium, reforça mais a utilização dessa ferramenta junto aos alunos e tem a cultura mais forte de incentivar os debates com base nos temas que a ferramenta propõe para fins de treino de redação. O efeito disso parece muito claro quando observamos os resultados dessa rede nas redações do ENEM nos últimos 3 anos. Essa rede está constantemente entre os primeiros lugares de Belo Horizonte, onde a maior parte de suas unidades se encontra.

3ª Etapa: Nova regressão *naive*

Após a retirada dos *outliers* rodei uma nova leva de regressões *naive* utilizando o método dos mínimos quadrados ordinários novamente para verificar se os resultados se alteraram e, em geral, encontramos resultados ainda melhores. O nível de significância se manteve alto e a influência que cada tipo de ferramenta

tem sobre a nota do aluno em um período específico, ou seja, os betas de nossa equação 1, aumentaram um pouco. As tabelas das estatísticas descritivas obtidas nas regressões dos 6 períodos podem ser encontradas no final desse documento (tabelas 7 à 12).

Podemos ver o resultado mencionado acima observando a tabela de estatística descritiva que obtivemos na regressão *naive* com as informações referentes ao sexto período (tabela 12) e comparar com a tabela 6, utilizada como exemplo acima e que apresenta os mesmos resultados, porém antes da retirada dos *outliers*.

Tabela 12:

Resultados Regressão *naive* após retirada dos *outliers* em $t=6$

Dependent variable:	

PM6	

Canal Eleva	0.0001 (0.004)
Monitoria Online	0.006*** (0.001)
Atualidades	0.280*** (0.021)
Eleva ao Vivo	-0.455*** (0.085)
Constante	10.069*** (0.039)

Observations	9,020
R2	0.025
Adjusted R2	0.025
Residual Std. Error	3.174 (df = 9015)
F Statistic	58.776*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Como podemos ver, a utilização do Canal Eleva se manteve com um nível de significância que nos leva a pensar que essa ferramenta não influencia a nota dos alunos. As ferramentas Atualidades e Monitoria Online aumentaram bastante sua influência sobre

as notas enquanto o Eleva ao Vivo parece também aumentar sua influência, porém de forma negativa.

A influência negativa do Eleva ao Vivo me levou a suspeitar de um possível problema, a causalidade reversa. Isso significa que, não apenas a utilização das ferramentas pode estar influenciando as notas dos alunos, mas também a nota dos alunos pode influenciar seus padrões de utilização da ferramenta. Esse efeito talvez seja mais visível na ferramenta Eleva ao Vivo pois, por passar sempre aulas ao vivo com conteúdo das próximas provas, ela é a ferramenta que provavelmente os alunos desesperados mais buscam na véspera de avaliações. Ou seja, se um aluno foi mal em uma prova no período anterior tende a utilizar mais o Eleva ao Vivo nos períodos futuros afim de melhorar suas notas e, com isso, podemos ter um problema de causalidade reversa na regressão.

Para tentar isolar esse efeito adicionei nas regressões a nota do aluno nos períodos anteriores como controle.

4ª Etapa: Regressão linear com controle de notas defasadas

Ao adicionar em minhas regressões as notas dos alunos defasadas no tempo, pretendo controlar os resultados pelo fato do aluno ter ido bem anteriormente e poder se despreocupar e estudar menos para uma futura prova ou ter ido pior e precisar se dedicar mais aos estudos para melhorar suas notas. As notas defasadas representam os resultados passados dos alunos e refletem a necessidade que eles têm de focar mais em períodos futuros ou não.

Equação2:

$$pm_6 = \beta_1 canal_pm_t + \beta_2 monitoria_pm_t + \beta_3 atualidades_pm_t + \beta_4 aovivo_pm_t + \gamma_1 pm_5 + \gamma_2 pm_4 + \gamma_3 pm_3 + \gamma_4 pm_2 + \gamma_5 pm_1$$

A única diferença entre a nova equação e a equação 1, apresentada anteriormente, é a adição das notas das cinco provas passadas como variáveis explicativas. Aqui apliquei a fórmula apenas para o sexto período pois apenas nesse faz sentido colocarmos todas as provas sendo representadas na equação, já que estamos falando de notas defasadas como controle na equação

Ao incluir as notas defasadas na regressão e analisar os resultados para o período 6, identifiquei que as variáveis ligadas à utilização das ferramentas de estudo online deixaram de ser significativas a nível de significância abaixo de 0,1% e apenas, *monitoria_pm6* se manteve significativa a um nível aceitável. Os efeitos da utilização do Eleva ao Vivo mantêm-se negativos. Quando falamos da Monitoria Online e do Canal

Eleva a influência de ambos sobre o resultado acadêmico dos alunos altera um pouco em relação aos resultados anteriores obtidos na regressão da etapa 3, mas não chegam a alterar o sentido do efeito. A tabela 13 a seguir mostra os resultados obtidos.

Tabela13:

Resultados regressão com controle de notas defasadas

=====	
Dependent variable:	

	PM6

Canal Eleva	0.002 (0.003)
Monitoria Online	0.003*** (0.001)
Atualidades	0.014 (0.017)
Eleva ao Vivo	-0.085 (0.065)
PM1	0.140*** (0.011)
PM2	-0.061*** (0.011)
PM3	0.150*** (0.012)
PM4	0.316*** (0.010)
PM5	0.210*** (0.010)
Constante	2.148*** (0.124)

Observations	9,020
R2	0.439
Adjusted R2	0.438
Residual Std. Error	2.409 (df = 9010)

F Statistic 783.295*** (df = 9; 9010)
 =====
 Note: *p<0.1; **p<0.05; ***p<0.01

Algo que me chama atenção nesse novo modelo é que o R^2 obtido é bem mais alto, ou seja, utilizando esse novo modelo conseguimos explicar 43,9% dos valores observados em nossa amostra. Com isso, podemos concluir que o resultado do aluno em provas passadas consegue explicar boa parte dos resultados futuros.

Até agora obtivemos resultados muito variados, tanto quando comparamos os efeitos da utilização das ferramentas entre si, como quando comparamos o efeito a utilização de uma mesma ferramenta sobre o resultado do aluno em diferentes provas. Com isso, podemos supor aqui que existem outras variáveis, não observáveis, que podem estar afetando o nosso resultado, como por exemplo o fato de cada aluno ter naturalmente uma maior pré-disposição para estudar do que outros, viver em um ambiente diferente, ter uma criação diferente, entre outros fatores. Esse problema gera um viés de variável omitida.

Adicionalmente, podemos dizer que existe um problema de viés de seleção, ou seja, provavelmente os alunos que tiram maiores notas já são aqueles que naturalmente se dedicam mais a estudar e devem procurar mais essas ferramentas. Para resolver esses problemas, decidi reorganizar minha base de dados e colocá-la em formato de dados em painel. Com isso consigo comparar o aluno com ele mesmo ao longo do tempo e, além disso, tentei adicionar efeitos fixos para alunos, isolando assim os efeitos de variáveis que variam entre alunos e não ao longo do tempo como, por exemplo, a situação sócio econômica, criação, habilidades natas, entre outros.

5ª Etapa: Dados em painel e regressão com efeitos fixos

Como dito na etapa anterior, provavelmente existem outras variáveis que também afetam o resultado dos alunos e que não estão previstas nas regressões feitas até agora. Variáveis omitidas, como por exemplo: a influência da criação do aluno, a sua situação sócio econômica, a distância que ele mora da escola, entre outras, pode afetar seus resultados acadêmicos. Para isolarmos os efeitos dessas variáveis e conseguirmos analisar apenas o impacto da utilização da plataforma online de ensino sobre o resultado acadêmico dos alunos, podemos usar dados em painel e realizar uma regressão com efeitos fixos. Ou seja, nossos dados possuem duas dimensões de variação: alunos (i) e tempo (t) e ao montar os dados em formato de painel o aluno aparece t vezes, podendo

ser comparado. Com isso, controlei o resultado por efeitos fixos no tempo, porém que variam entre alunos.

Além disso, para controlarmos por possíveis efeitos relacionados à prova em si, adicionarei dummies de provas. Serão 6 dummies onde cada uma delas assumirá o valor de 1 se estivermos em determinada prova e zero caso estejamos em qualquer outra das 5 provas. Com isso, nossa regressão será:

Equação 3:

$$Nota_{it} = \alpha + \beta_1 Canal_{PM} + \beta_2 Monitoria_{PM} + \beta_3 Atualidade_{PM} + \beta_4 AoVivo_{PM} + \gamma_2 Dummy_2 + \gamma_3 Dummy_3 + \gamma_4 Dummy_4 + \gamma_5 Dummy_5 + \gamma_6 Dummy_6 + u_i$$

Onde i representa a unidade de alunos; t o período do tempo, que varia de 1 a 6 ao longo do ano; u_i é o efeito fixo por aluno e $Dummy_t$ representa a dummy para cada período t , totalizando cinco dummies distintas (retirei a dummy do primeiro período para não termos problemas de multicolinearidade perfeita).

Para a variável nota eu usei tanto a nota geral dos alunos nas provas multidisciplinares, como a nota para cada área do conhecimento, sendo elas: matemática, ciências da natureza, ciências humanas e linguagens. Isso me levou às cinco tabelas de resultados que analisarei no próximo capítulo e que se encontram no final deste documento numeradas de 14 a 18.

6. Resultados

Ao realizarmos a regressão mais simples, considerando apenas o agregado de acessos às ferramentas em um período t e as notas dos alunos no mesmo período já encontramos um resultado muito interessante de efeito positivo e altamente significativo. Entretanto, não paramos nesse ponto pois, essa regressão simples pode apresentar inúmeros problemas como viés de variável omitida e viés de seleção.

O viés de variável omitida pode existir por exemplo pelo fato de cada aluno ter uma família diferente, uma situação socioeconômica diferente, e isso influencia os resultados desse aluno. Já o viés de seleção está ligado ao fato de que melhores alunos tendem a estudar mais e usar mais a plataforma de ensino naturalmente.

Afim de acabar com esses dois tipos de viés, como explicado no capítulo anterior, rodei uma regressão de dados em painel com efeitos fixos. Ao realizar essa regressão considerando a nota geral dos alunos nas PMs encontrei o seguinte resultado:

Tabela 14:
Resultados da regressão em painel sobre a nota geral

Dependent variable:	

Notas nas PMS	

Canal Eleva	0.004*** (0.001)
Monitoria Online	0.003*** (0.001)
Atualidades	0.063*** (0.007)
Eleva ao vivo	-0.138*** (0.017)
Dummy PM1	-10.964*** (3.716)
DummyPM2	-4.642* (2.628)
DummyPM4	-10.604*** (3.716)

DummyPM5	-2.589 (2.628)

Observations	54,120
R2	0.005
Adjusted R2	-0.195
F Statistic	25.523*** (df = 8; 45092)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Essa tabela nos dá um resultado muito satisfatório. Podemos ver que, com esse modelo, todas as ferramentas de ensino online influenciam de forma significativa a nota geral dos alunos. Apesar da magnitude do efeito ser pequena, eles têm um bom nível de significância evidenciando esse resultado é diferente de zero de forma significativa.

Ao rodar a mesma regressão, porém considerando separadamente as notas obtidas em cada área do conhecimento pude observar que, provavelmente o conteúdo ou a metodologia funciona melhor para matemática e em seguida para ciências da natureza. Podemos afirmar isso ao analisar o nível de significância e os valores dos coeficientes encontrados nas tabelas 15 e 16.

Outro resultado interessante obtido é em relação à ferramenta Eleva ao Vivo. Essa ferramenta obteve, em quase todos os casos, um resultado negativo sobre as notas dos alunos. Como disse na hipótese levantada anteriormente, essa ferramenta é possivelmente mais utilizada por alunos em véspera de provas e, funcionando como um programa de televisão a noite, pode causar um efeito de maior cansaço nos alunos que a utilizam, assim como um aumento de nervosismo caso o aluno perceba que não está entendendo bem a aula e possui pouco tempo para estudar.

Infelizmente, o R^2 desse modelo indica que existe pouco ajustamento entre o modelo e os valores observados. Entretanto, isso não invalida os resultados. Sabemos que a influência das ferramentas online de estudo sobre a nota dos alunos é, em magnitude, muito baixa tendo assim baixa influência sobre o resultado final da nota dos alunos, porém não deixam de influenciar de forma positiva tornando o resultado relevante para aprendermos sobre como as diferentes metodologias de ensino impactam o resultado acadêmico de alunos.

7. Conclusão

Esse estudo mostrou que ferramentas de ensino online utilizadas por uma rede privada no Brasil podem ter influência positiva sobre a nota dos alunos em ensino fundamental e médio que nela estudam. Entretanto o impacto sobre as notas é bem pequeno.

O resultado encontrado é animador pois, apesar da influência ser baixa, mostra que podemos utilizar essas ferramentas para melhorar a qualidade do ensino em diferentes lugares. Nesse estudo o impacto do uso das plataformas de ensino online foi analisado considerando a utilização por parte de alunos que já tem acesso a um ensino de qualidade. Não foi calculada a influência da utilização dessas ferramentas sobre o aprendizado de alunos que não têm acesso a salas de aula decentes ou professores de qualidade durante o turno escolar. Com isso quero dizer que, provavelmente, o aprendizado online, por ser escalável e possuir um baixo custo marginal, pode gerar um impacto muito maior do que o calculado nesse trabalho se replicado para comunidades de baixíssima renda.

8. Referência Bibliográfica

1 - ALLEN, I. Elaine, and SEAMAN, Jeff. **Changing Course: Ten Years of tracking Online Education in the United States.** Newburyport, MA: Sloan Consortium, 2013

2 - BETTINGER, Eric, L Fox, S Loeb, E Taylor. **Changing distributions: how online college classes alter student and professor performance,** 2014

BETTINGER, Eric P., Susanna Loeb, and Eric S. Taylor. **Remote but influential: Peer Effects and Reflection on Online Higher Education Classrooms,** 2015

FIGLIO, David, RUSH, Mark, and Lu Yin. **Is It Live or Is It Internet? Experimental Estimates of the Effects of Online Instructions on Student Learning.** Journal of Labor Economics no.31(4):763-784, 2013

CARD, David. **Using geographic variation in college proximity to estimate the return to schooling,** National Bureau of Economic research, 1993

ROUSE, Cecilia; ASHENFELTER, Orley. **Income, schooling, and ability: evidence from a new sample of identical twins,** National Bureau of Economic research, 1997

MURALIDHAREN, Karthik. **Field Experiments in Education in Developing Countries.**

EVANS, David K.; POPOVA, Anna. **What Really Works to Improve Learning in Developing countries? An Analysis of Divergent Findings in Systematic Reviews.**

FERNANDES, Mauricio M.; FERRAZ, Claudio. **Conhecimento ou Práticas pedagógicas? Medindo os Efeitos da Qualidade dos Professores no Desempenho dos Alunos,** 2014

9. Gráficos e Tabelas:

Gráfico 1: Exemplo de distribuição das notas em uma turma de ensino médio

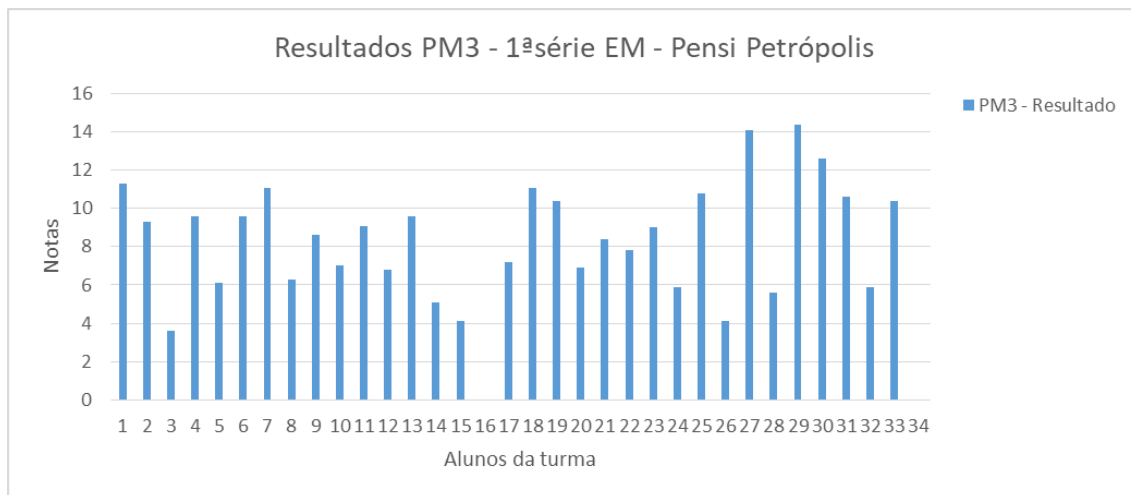


Gráfico 2: Distribuição das observações de número de acessos por aluno à ferramenta canal eleva antes da retirada de *outliers*.

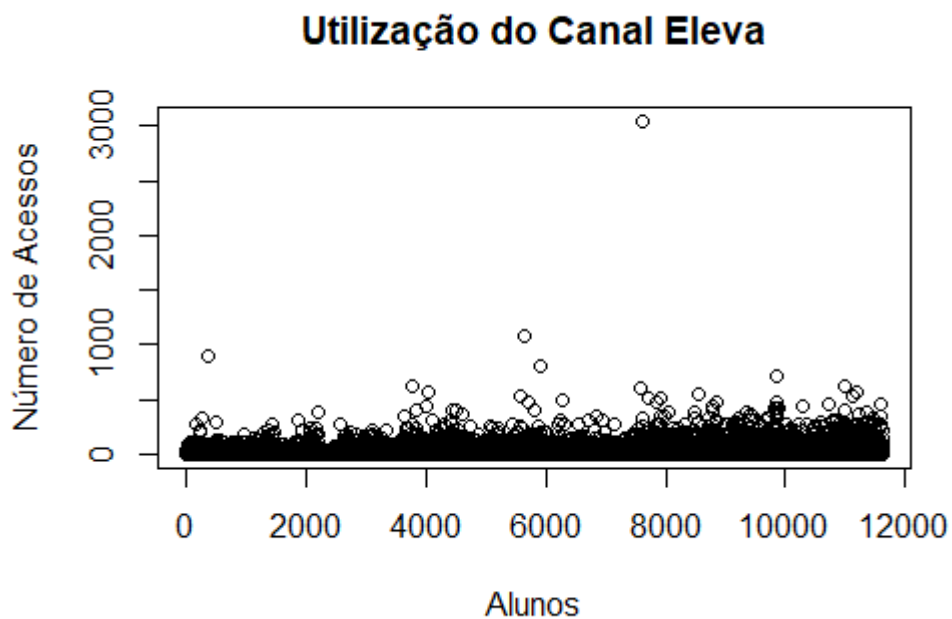


Gráfico 3: Distribuição das observações de número de acessos por aluno à ferramenta monitoria online antes da retirada de *outliers*.

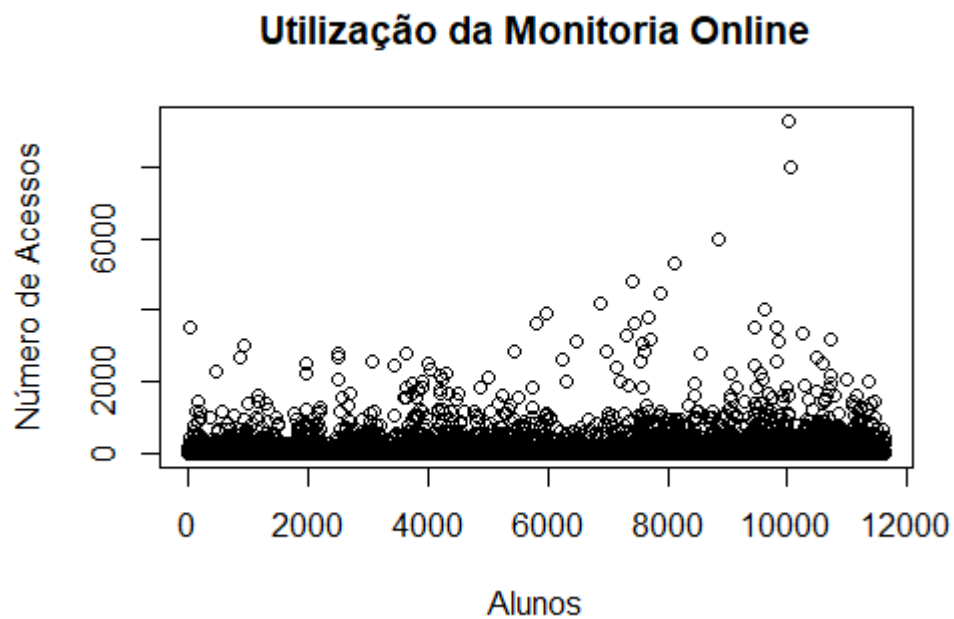


Gráfico 4: Distribuição das observações de número de acessos por aluno à ferramenta atualidades antes da retirada de *outliers*.

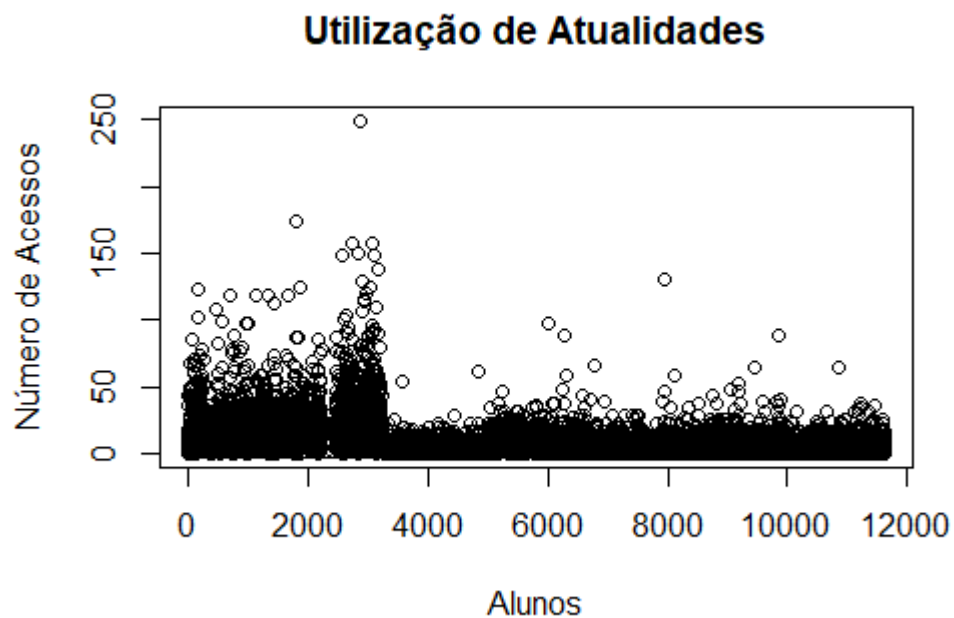


Gráfico 5: Distribuição das observações de número de acessos por aluno à ferramenta eleva ao vivo antes da retirada de *outliers*.

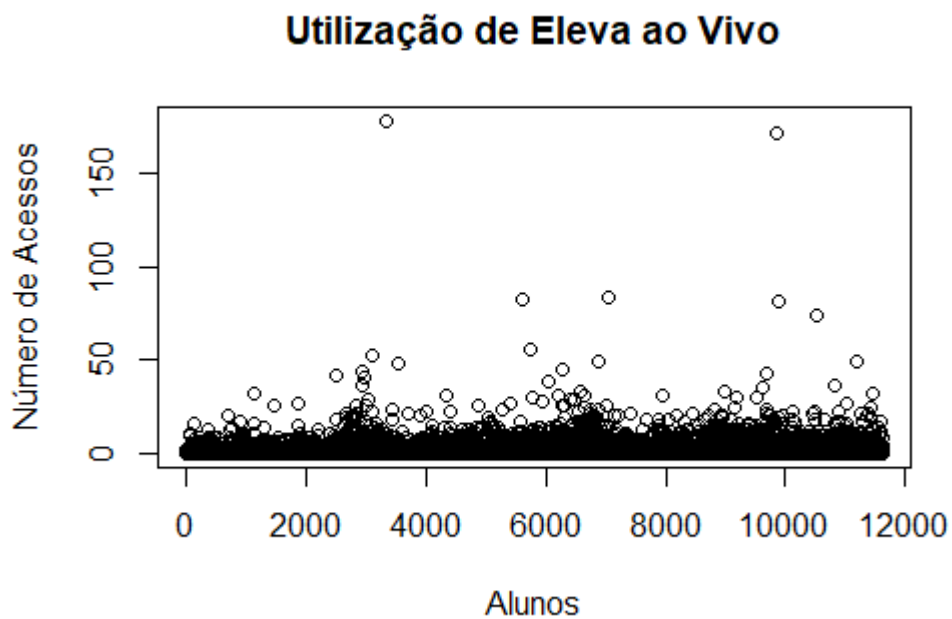


Gráfico 6: Distribuição das observações de número de acessos por aluno à ferramenta canal eleva após a retirada de *outliers*.

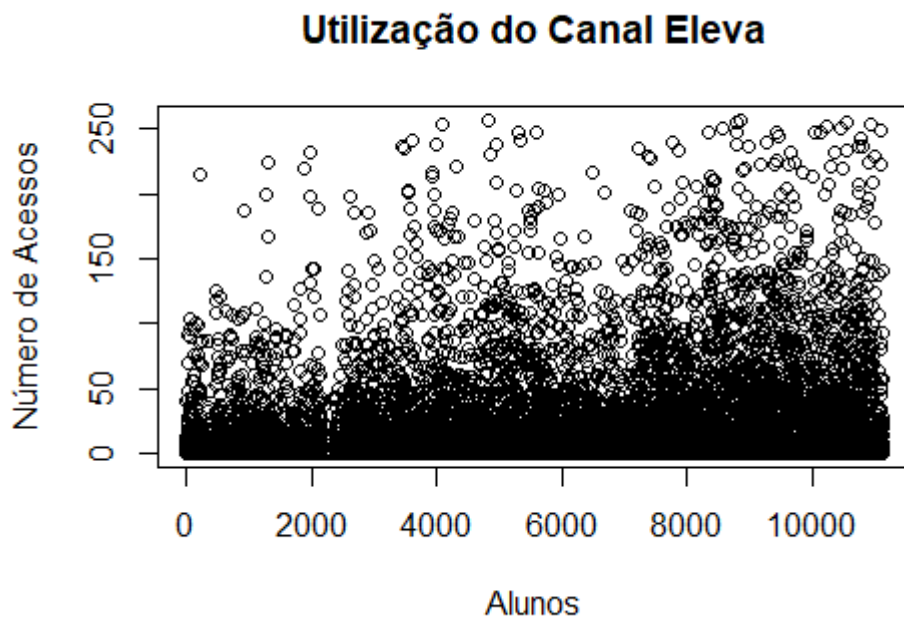


Gráfico 7: Distribuição das observações de número de acessos por aluno à ferramenta monitoria online após a retirada de *outliers*.

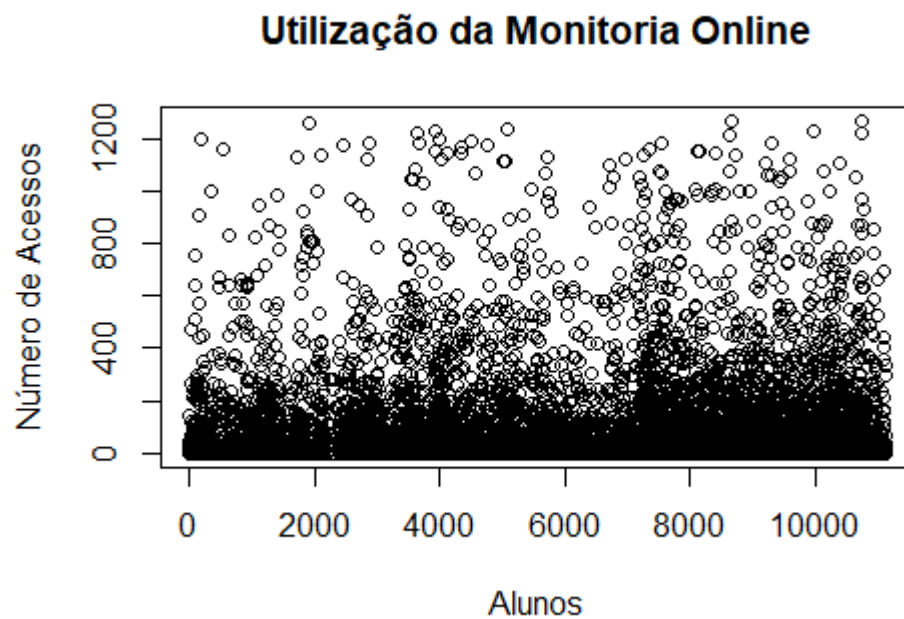


Gráfico 8: Distribuição das observações de número de acessos por aluno à ferramenta atualidades após a retirada de *outliers*.

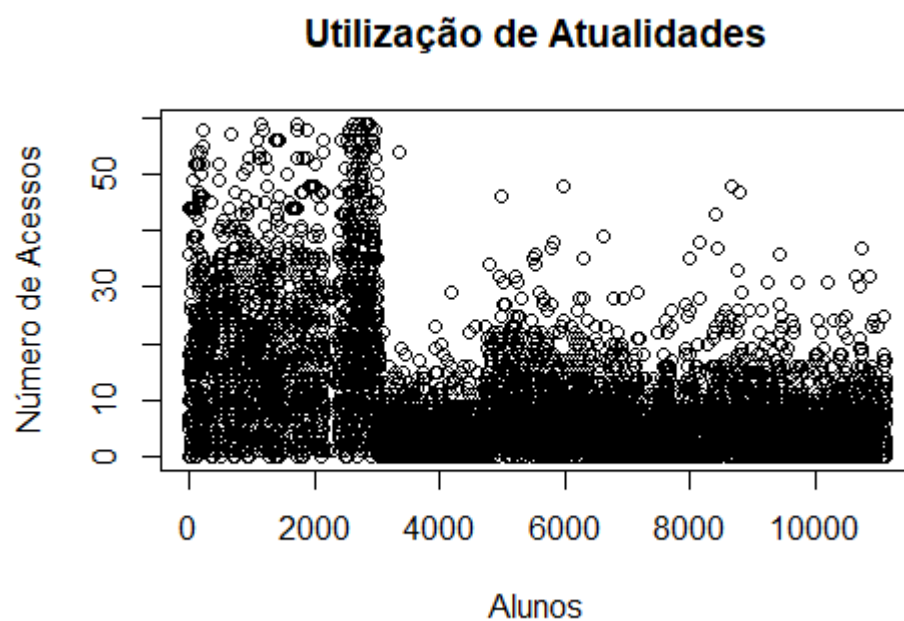


Gráfico 9: Distribuição das observações de número de acessos por aluno à ferramenta Eleva ao vivo após a retirada de *outliers*.

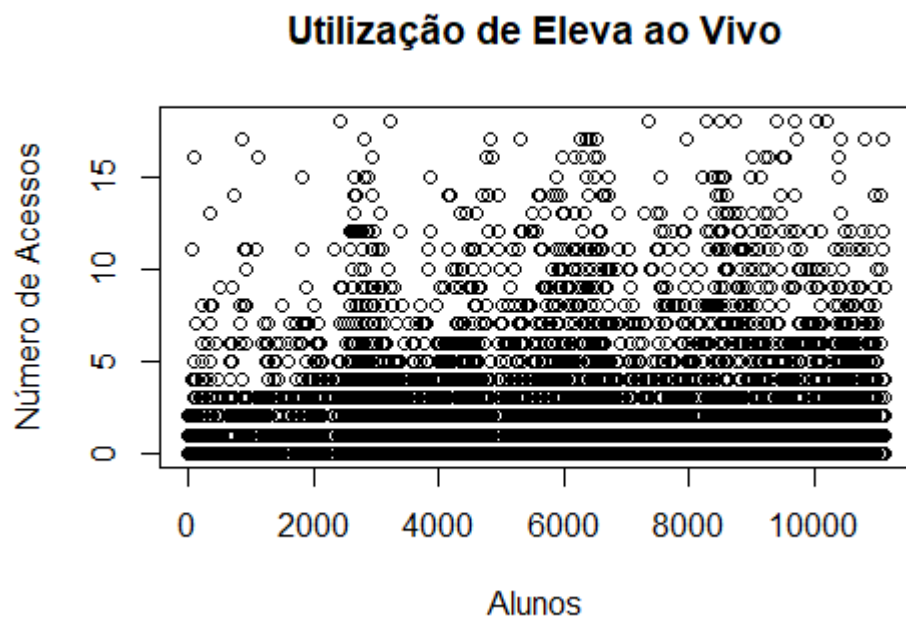


Gráfico 10: Distribuição das observações de número de acessos por aluno à ferramenta monitoria online após a segunda retirada de *outliers*.

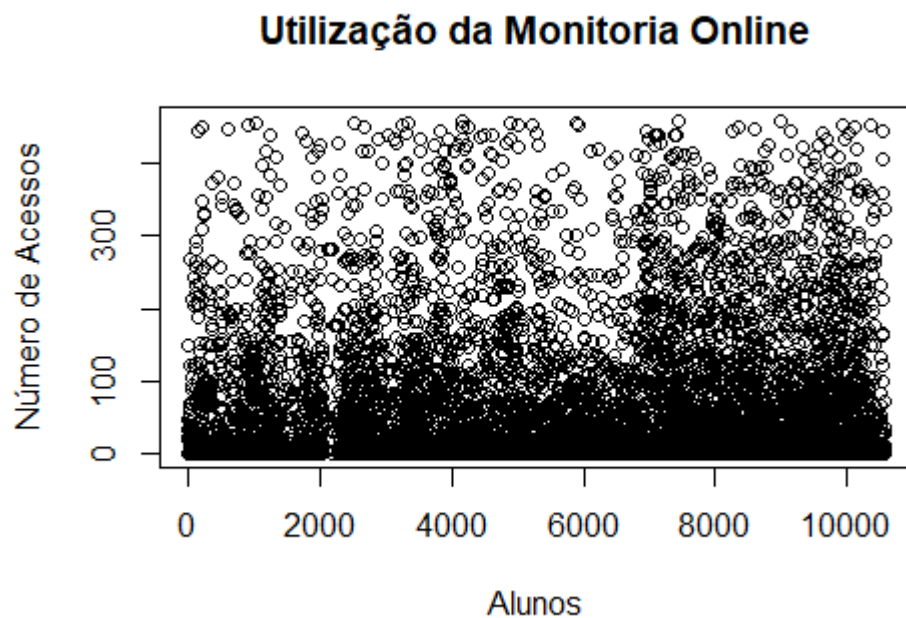


Tabela 1: Regressão *naïve* considerando o primeiro período de provas ($t=1$)Resultados Regressão *naïve* em $t=1$

Dependent variable:	

PM1	

Canal Eleva	0.012* (0.006)
Monitoria Online	0.002*** (0.001)
Atualidades	0.287*** (0.016)
Eleva ao Vivo	-0.374*** (0.028)
Constante	11.239*** (0.037)

Observations	9,906
R2	0.044
Adjusted R2	0.043
Residual Std. Error	3.274 (df = 9901)
F Statistic	112.918*** (df = 4; 9901)
=====	
Note:	* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Tabela 2: Regressão *naïve* considerando o segundo período de provas (t =2)Resultados Regressão *naïve* em t=2

Dependent variable:	

PM2	

Canal Eleva	-0.001 (0.001)
Monitoria Online	0.0001 (0.0003)
Atualidades	0.050*** (0.008)
Eleva ao Vivo	0.085*** (0.015)
Constante	13.032*** (0.037)

Observations	9,906
R2	0.009
Adjusted R2	0.008
Residual Std. Error	3.060 (df = 9901)
F Statistic	21.762*** (df = 4; 9901)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 3: Regressão *naïve* considerando o terceiro período de provas (t =3)Resultados Regressão *naïve* em t=3

Dependent variable:	

PM3	

Canal Eleva	0.003* (0.002)
Monitoria Online	0.003*** (0.0005)
Atualidades	0.135*** (0.009)
Eleva ao Vivo	-0.143*** (0.013)
Constante	10.166*** (0.038)

Observations	9,906
R2	0.037
Adjusted R2	0.036
Residual Std. Error	3.169 (df = 9901)
F Statistic	94.124*** (df = 4; 9901)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 4: Regressão *naïve* considerando o quarto período de provas ($t=4$)Resultados Regressão *naïve* em $t=4$

Dependent variable:	

PM4	

Canal Eleva	-0.004 (0.002)
Monitoria Online	0.003*** (0.001)
Atualidades	0.262*** (0.014)
Eleva ao Vivo	-0.449*** (0.032)
Constante	10.485*** (0.041)

Observations	9,906
R2	0.053
Adjusted R2	0.052
Residual Std. Error	3.466 (df = 9901)
F Statistic	137.624*** (df = 4; 9901)
=====	
Note:	* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Tabela 5: Regressão *naïve* considerando o quinto período de provas ($t=5$)Resultados Regressão *naïve* em $t=5$

Dependent variable:	

PM5	

Canal Eleva	-0.0002 (0.002)
Monitoria Online	0.001*** (0.0004)
Atualidades	0.063*** (0.009)
Eleva ao Vivo	0.191*** (0.046)
Constante	11.361*** (0.039)

Observations	9,906
R2	0.008
Adjusted R2	0.008
Residual Std. Error	3.297 (df = 9901)
F Statistic	21.059*** (df = 4; 9901)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 6: Regressão *naïve* considerando o sexto período de provas (t =6)Resultados Regressão *naïve* em t=6

Dependent variable:	

PM6	

Canal Eleva	0.001 (0.003)
Monitoria Online	0.003*** (0.0005)
Atualidades	0.206*** (0.015)
Eleva ao Vivo	-0.446*** (0.073)
Constante	10.153*** (0.036)

Observations	9,906
R2	0.029
Adjusted R2	0.029
Residual Std. Error	3.165 (df = 9901)
F Statistic	74.352*** (df = 4; 9901)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 7: Regressão *naive* após retirada dos *outliers* considerando o primeiro período de provas (t =1)

Resultados Regressão *naive* após retirada dos outliers em t=1

Dependent variable:	

	PM1

Canal Eleva	0.017** (0.009)
Monitoria Online	0.012*** (0.003)
Atualidades	0.338*** (0.019)
Eleva ao vivo	-0.528*** (0.040)
Constante	11.150*** (0.040)

Observations	9,020
R2	0.048
Adjusted R2	0.048
Residual Std. Error	3.278 (df = 9015)
F Statistic	113.647*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 8: Regressão *naive* após retirada dos *outliers* considerando o segundo período de provas (t =2)

Resultados Regressão *naive* após retirada dos outliers em t=2

Dependent variable:	

	PM2

Canal Eleva	0.0002 (0.003)
Monitoria Online	-0.001 (0.001)
Atualidades	0.063*** (0.010)
Eleva ao Vivo	0.160*** (0.027)
Constante	12.963*** (0.041)

Observations	9,020
R2	0.010
Adjusted R2	0.009
Residual Std. Error	3.066 (df = 9015)
F Statistic	22.370*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 9: Regressão *naive* após retirada dos *outliers* considerando o terceiro período de provas (t =3)

Resultados Regressão *naive* após retirada dos outliers em t=3

Dependent variable:	

	PM3

Canal Eleva	0.008*** (0.003)
Monitoria Online	0.005*** (0.001)
Atualidades	0.202*** (0.012)
Eleva ao vivo	-0.540*** (0.033)
Constante	10.138*** (0.042)

Observations	9,020
R2	0.052
Adjusted R2	0.051
Residual Std. Error	3.153 (df = 9015)
F Statistic	122.513*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 10: Regressão *naive* após retirada dos *outliers* considerando o quarto período de provas (t =4)

Resultados Regressão *naive* após retirada dos outliers em t=4

Dependent variable:	

	PM4

Canal Eleva	-0.009*** (0.003)
Monitoria Online	0.007*** (0.001)
Atualidades	0.351*** (0.021)
Eleva ao vivo	-0.637*** (0.047)
Constante	10.405*** (0.044)

Observations	9,020
R2	0.050
Adjusted R2	0.049
Residual Std. Error	3.470 (df = 9015)
F Statistic	117.831*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 11: Regressão *naive* após retirada dos *outliers* considerando o quinto período de provas (t =5)

Resultados Regressão *naive* após retirada dos outliers em t=5

Dependent variable:	

	PM5

Canal Eleva	-0.001 (0.003)
Monitoria Online	0.002* (0.001)
Atualidades	0.094*** (0.013)
Eleva ao vivo	0.204*** (0.058)
Constante	11.297*** (0.042)

Observations	9,020
R2	0.008
Adjusted R2	0.008
Residual Std. Error	3.310 (df = 9015)
F Statistic	18.633*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 12: Regressão *naive* após retirada dos *outliers* considerando o sexto período de provas (t =6)

Resultados Regressão *naive* após retirada dos outliers em t=6

Dependent variable:	

	PM6

Canal Eleva	0.0001 (0.004)
Monitoria Online	0.006*** (0.001)
Atualidades	0.280*** (0.021)
Eleva ao vivo	-0.455*** (0.085)
Constante	10.069*** (0.039)

Observations	9,020
R2	0.025
Adjusted R2	0.025
Residual Std. Error	3.174 (df = 9015)
F Statistic	58.776*** (df = 4; 9015)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela13: Resultados da regressão simples com adição de notas defasadas como controle para causalidade reversa.

Resultados regressão com controle de notas defasadas

Dependent variable:	

PM6	

Canal Eleva	0.002 (0.003)
Monitoria Online	0.003*** (0.001)
Atualidades	0.014 (0.017)
Eleva ao Vivo	-0.085 (0.065)
PM1	0.140*** (0.011)
PM2	-0.061*** (0.011)
PM3	0.150*** (0.012)
PM4	0.316*** (0.010)
PM5	0.210*** (0.010)
Constante	2.148*** (0.124)

Observations	9,020
R2	0.439
Adjusted R2	0.438
Residual Std. Error	2.409 (df = 9010)
F Statistic	783.295*** (df = 9; 9010)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 14: Resultados da regressão em painel com efeitos fixos para alunos sobre a nota geral

Resultados da regressão em painel sobre a nota geral

Dependent variable:	
Notas nas PMS	
Canal Eleva	0.004*** (0.001)
Monitoria Online	0.003*** (0.001)
Atualidades	0.063*** (0.007)
Eleva ao vivo	-0.138*** (0.017)
DummyPM1	-10.964*** (3.716)
DummyPM2	-4.642* (2.628)
DummyPM4	-10.604*** (3.716)
DummyPM5	-2.589 (2.628)
Observations	54,120
R2	0.005
Adjusted R2	-0.195
F Statistic	25.523*** (df = 8; 45092)
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 15: Resultados da regressão em painel com efeitos fixos para alunos sobre a nota de matemática.

Resultados da regressão em painel sobre a nota de Matemática

Dependent variable:	
Notas em Matemática	
Canal Eleva	0.002*** (0.001)
Monitoria Online	0.001*** (0.0002)
Atualidades	0.032*** (0.003)
Eleva ao vivo	-0.083*** (0.006)
Dummy PM1	-4.105*** (1.342)
DummyPM2	-1.419 (0.949)
DummyPM4	-5.196*** (1.342)
DummyPM5	-2.001** (0.949)
Observations	54,120
R2	0.009
Adjusted R2	-0.189
F Statistic	51.274*** (df = 8; 45092)
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 16: Resultados da regressão em painel com efeitos fixos para alunos sobre a nota de ciências da natureza.

Resultados sobre a nota em Ciências da Natureza

Dependent variable:	
Notas em Ciências da Natureza	
Canal Eleva	0.001** (0.001)
Monitoria Online	0.001** (0.0002)
Atualidades	0.017*** (0.003)
Eleva ao vivo	-0.018*** (0.006)
Dummy PM1	-2.412* (1.386)
DummyPM2	-0.732 (0.980)
DummyPM4	-0.761 (1.386)
DummyPM5	0.486 (0.980)
Observations	54,120
R2	0.002
Adjusted R2	-0.198
F Statistic	9.229*** (df = 8; 45092)
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 17: Resultados da regressão em painel com efeitos fixos para alunos sobre a nota de ciências humanas.

Resultados sobre a nota em Ciências Humanas	
=====	
Dependent variable:	

Notas em Ciências Humanas	

Canal Eleva	0.001 (0.0005)
Monitoria Online	0.0001 (0.0002)
Atualidades	-0.004* (0.002)
Eleva ao vivo	0.013** (0.006)
Dummy PM1	-0.830 (1.227)
DummyPM2	-0.810 (0.868)
DummyPM4	-0.942 (1.227)
DummyPM5	-0.069 (0.868)

Observations	54,120
R2	0.0003
Adjusted R2	-0.200
F Statistic	1.457 (df = 8; 45092)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01

Tabela 18: Resultados da regressão em painel com efeitos fixos para alunos sobre a nota de linguagens.

Resultados sobre a nota em Linguagens	
=====	
Dependent variable:	

Notas em Linguagens	

Canal Eleva	0.0003 (0.001)
Monitoria Online	0.001*** (0.0002)
Atualidades	0.017*** (0.002)
Eleva ao vivo	-0.050*** (0.006)
Dummy PM1	-3.618*** (1.266)
DummyPM2	-1.681* (0.895)
DummyPM4	-3.704*** (1.266)
DummyPM5	-1.005 (0.895)

Observations	54,120
R2	0.004
Adjusted R2	-0.196
F Statistic	21.734*** (df = 8; 45092)
=====	
Note:	*p<0.1; **p<0.05; ***p<0.01