

UNIVERSIDADE FEDERAL DOS VALES DO JEQUITINHONHA E MUCURI
FACULDADE DE CIÊNCIAS EXATAS
CURSO DE SISTEMAS DE INFORMAÇÃO

**APLICAÇÃO DE CIÊNCIA DE DADOS PARA AVALIAÇÃO DOS RESULTADOS DO
PISA**

BRUNO CARLOS CESAR FERNANDES

Trabalho de Conclusão de Curso

Diamantina/MG

2019

BRUNO CARLOS CESAR FERNANDES

**APLICAÇÃO DE CIÊNCIA DE DADOS PARA AVALIAÇÃO DOS RESULTADOS DO
PISA**

Trabalho apresentado ao curso de Sistemas de Informação da Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, como pré-requisito para obtenção do grau de bacharel, sob orientação do Prof. Dr. Alessandro Vivas Andrade.

Diamantina/MG

2019

BRUNO CARLOS CESAR FERNANDES

**APLICAÇÃO DE CIÊNCIA DE DADOS PARA AVALIAÇÃO DOS RESULTADOS DO
PISA**

Trabalho apresentado ao curso de Sistemas de Informação da Universidade Federal dos Vales do Jequitinhonha e Mucuri – UFVJM, como pré-requisito para obtenção do grau de bacharel, sob orientação do Prof. Dr. Alessandro Vivas Andrade.

COMISSÃO EXAMINADORA

DIAMANTINA, 10 DE JULHO DE 2019

Prof. Alessandro Vivas Andrade, Dr.
Universidade Federal dos Vales do Jequitinhonha e Mucuri
Orientador

Profa. Josiane Magalhães Teixeira, Dra.
Universidade Federal dos Vales do Jequitinhonha e Mucuri
Examinador

Profa. Luciana Pereira de Assis, Dra.
Universidade Federal dos Vales do Jequitinhonha e Mucuri
Examinadora

À todos familiares e amigos.

AGRADECIMENTOS

A Deus por ter me dado saúde e força para superar as dificuldades. Aos meus pais, pelo amor e incentivo, em especial a minha mãe Terezinha, pelo apoio nas horas mais difíceis de desânimo e cansaço. A toda minha família que sempre se mostrou disposta a me apoiar, em específico a Nilma que me motivou a ingressar na universidade, Geralda pelo suporte tão grande durante meus anos de universitário. Ao meu orientador Alessandro Vivas, pela oportunidade, confiança e muita paciência. A minha irmã de outra mãe Amanda Lopes, que sempre esteve presente. Aos meus amigos pela amizade, incentivo e contribuições. Agradeço também aos professores e técnicos do Departamento de Computação da UFVJM, que foram extremamente importantes e me guiaram durante essa minha jornada. E por final a todos que direta ou indiretamente fizeram parte da minha formação, o meu muito obrigado.

*“O importante não é aquilo que fazem de nós,
mas o que nós mesmos fazemos do que os outros fizeram de nós.”*

(JEAN-PAUL SARTRE)

FERNANDES, C Bruno Carlos Cesar. **Aplicação de Ciência de Dados para Avaliação dos Resultados do PISA**. Diamantina, 2019. 74. Trabalho de Conclusão de Curso Superior em Sistemas de Informação. Faculdade de Ciências Exatas, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, 2019.

RESUMO

O PISA (Programa Internacional de Avaliação de Estudantes) é um teste de nivelamento aplicado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE) para realizar um estudo sobre o desempenho de alunos do ensino básico em vários países. Este trabalho tem como objetivo aplicar conceitos de análise de Ciência de Dados para correlacionar as principais variáveis que determinam o sucesso ou fracasso dos estudantes neste teste.

Palavras-chave: PISA. Educação. Linguagem R. Ciência de dados.

FERNANDES, C Bruno Carlos Cesar. **Application of Data Science for Avaluation the Results of PISA** Diamantina, 2016. 74. Trabalho de Conclusão de Curso Superior em Sistemas de Informação. Faculdade de Ciências Exatas, Universidade Federal dos Vales do Jequitinhonha e Mucuri, Diamantina, 2019.

ABSTRACT

PISA (Programme for International Student Assessment) is a leveling test applied by the Organization for Economic Cooperation and Development (OECD) to conduct a study on the performance of primary school students in several countries. This work aims to apply Data Science analysis concepts to correlate the main variables that determine the success or failure of the students in this test.

Key-words: PISA. Education. Language R. Data Science.

LISTA DE ILUSTRAÇÕES

Figura 1	– Fluxograma - Os passos da Aplicação do PISA.	25
Figura 2	– Notas do Brasil nas edições do PISA	26
Figura 3	– Logo R	27
Figura 4	– Logo Python	28
Figura 5	– Interface MATLAB	29
Figura 6	– Interface Octave	30
Figura 7	– Interface Scilab	30
Figura 8	– Interface Julia	31
Figura 9	– R Consoler	33
Figura 10	– Interface RStudio.	33
Figura 11	– Importando arquivos pela interface RStudio - Passo 1 e 2	35
Figura 12	– Importando arquivos pela interface RStudio - Passo 3 e 4	35
Figura 13	– Importando arquivos pela interface RStudio - Passo 5 - Arquivo importado .	36
Figura 14	– Resposta do comando: <code>packageStatus()</code>	36
Figura 15	– Resposta do comando: <code>summary(packageStatus())</code>	37
Figura 16	– Instalando pacotes pela interface	37
Figura 17	– Removendo pacotes pela interface	38
Figura 18	– Atualizar pacotes pela interface	38
Figura 19	– Baixando arquivo .deb	39
Figura 20	– Interface RStudio	40
Figura 21	– Gráfico exemplo com uso - GGPlot2	42
Figura 22	– Gráfico com o tamanho médio das turmas em cada país.	50
Figura 23	– Países com as maiores médias	51
Figura 24	– Países com as menores médias	51
Figura 25	– Nota de Ciências versus Tamanho da Turma	52
Figura 26	– Nota de Leitura versus Tamanho da Turma	52
Figura 27	– Nota de Matemática versus Tamanho da Turma	52
Figura 28	– Gráfico de escolas com administração religiosa.	54
Figura 29	– Tamanho médio das turmas (maiores médias) + Adm. Religiosa	55
Figura 30	– Tamanho médio das turmas (menores médias) + Adm. Religiosa	55
Figura 31	– Países com maior N ^a Adm. Religiosa + Notas	56
Figura 32	– Países com menor N ^a Adm. Religiosa + Notas	56
Figura 33	– Gráfico de escolas administradas por organizações sem fins lucrativos. . . .	57
Figura 34	– Tamanho médio das turmas (maiores médias) + Adm. Sem Fins Lucrativos	58
Figura 35	– Tamanho médio das turmas (menores médias) + Adm. Sem Fins Lucrativos	59
Figura 36	– Países com maior N ^a Adm. Sem Fins Lucrativos + Notas	60

Figura 37 – Países com menor N ^a Adm. Sem Fins Lucrativos + Notas	60
Figura 38 – Gráfico de escolas administradas por organizações com fins lucrativos.	61
Figura 39 – Tamanho médio das turmas (maiores médias) + Adm. Com Fins Lucrativos	62
Figura 40 – Tamanho médio das turmas (menores médias) + Adm. Com Fins Lucrativos	62
Figura 41 – Países com maior N ^a Adm. Com Fins Lucrativos + Notas	63
Figura 42 – Países com menor N ^a Adm. Com Fins Lucrativos + Notas	63
Figura 43 – Gráfico professores com status de efetivo.	64
Figura 44 – Países com maior N ^a de professores efetivos + notas	65
Figura 45 – Países com menor N ^a de professores efetivos + notas	65
Figura 46 – Gráfico de professores com nível A5.	66
Figura 47 – Nível A5 + Ciência	67
Figura 48 – Nível A5 + Leitura	68
Figura 49 – Nível A5 + Matemática	69

SUMÁRIO

1 INTRODUÇÃO	13
1.1 JUSTIFICATIVA	14
1.2 OBJETIVOS	14
1.2.1 OBJETIVO GERAL	14
1.2.2 OBJETIVOS ESPECÍFICOS	15
1.3 ESTRUTURA DO TRABALHO	15
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 PISA	17
2.1.1 QUESTIONÁRIO - ALUNO	18
2.1.2 QUESTIONÁRIO - PROFESSOR	20
2.1.3 QUESTIONÁRIO - ESCOLA	22
2.1.4 APLICAÇÃO DO PISA	24
2.2 PISA NO BRASIL	25
3 FERRAMENTA UTILIZADA	27
3.1 LINGUAGEM R	27
3.1.1 LINGUAGENS CONCORRENTES AO - R	28
3.1.1.1 PYTHON	28
3.1.1.2 MATLAB	28
3.1.1.3 OCTAVE	29
3.1.1.4 SCILAB	30
3.1.1.5 JULIA	31
3.1.1.6 COMPARATIVO DAS LINGUAGENS	31
3.1.2 INTERFACES DE TRABALHO	32
3.1.3 IMPORTAÇÃO DE ARQUIVOS	34
3.1.4 GERENCIAMENTO DO R	36
3.1.5 PROCESSO DE INSTALAÇÃO DO - R:	38
4 METODOLOGIA	41
5 ANÁLISE DOS RESULTADOS	45
5.1 INTRODUÇÃO	45
5.2 QUESTIONÁRIO - ESCOLA	47
5.2.1 TAMANHO MÉDIO DAS TURMAS	47
5.2.2 TIPOS DE ADMINISTRAÇÃO	53
5.3 QUESTIONÁRIO - PROFESSOR	64
5.3.1 STATUS PROFESSOR	64

5.3.2 NÍVEL DE FORMAÇÃO	65
6 CONCLUSÃO	71
6.1 CONSIDERAÇÕES FINAIS	71
6.2 TRABALHOS FUTUROS	71
REFERÊNCIAS	73

1 INTRODUÇÃO

É inegável que a educação seja um dos temas de aspecto social mais debatido da contemporaneidade. Nesse aspecto o PISA - Programa Internacional de Avaliação de Estudante aponta que a educação foi e sempre será em qualquer país um importante assunto. Tal afirmativa se faz a partir de entender que os indivíduos por meio da educação são capazes de tomarem conhecimento, responsabilidade social, dentre muitas outras qualidades cidadã (PISA, 2001).

Assim, para que se tenha educação de qualidade é necessário que se entenda o sistema educacional como um todo, desde o professor até aos alunos e também a comunidade em que estão inseridos. Contudo, apesar de fatores separados, uma educação de qualidade só é construída a partir da inter-relação desses fatores, pois para se ter qualidade em um é necessária a qualidade em um sistema como todo, como aponta (GADOTTI, 2013).

O PISA além de aplicar provas para medir o nível educacional, também aplica questionários, com objetivo de coletar informações socioeconômicas e demográficas. Com dados disponibilizados pelo programa, o presente trabalho busca aplicar a Ciência de Dados, área de pesquisa que trabalha com análise de dados.

De acordo com (OLIVEIRA; GUERRA; MCDONNELL, 2018), p. 9) Ciência de Dados é:

“Trata-se de um termo cada vez mais utilizado para designar uma área de conhecimento voltada para o estudo e a análise de dados, onde busca-se extrair conhecimento e criar novas informações. É uma atividade interdisciplinar, que concilia principalmente duas grandes áreas: Ciência da Computação e Estatística. A Ciência de Dados vem sendo aplicada como apoio em diferentes outras áreas de conhecimento, tais como: Medicina, Biologia, Economia, Comunicação, Ciências Políticas etc. Apesar de não ser uma área nova, o tema vem se popularizando cada vez mais, graças à explosão na produção de dados e crescente dependência dos dados para a tomada de decisão.”

Assim este trabalho busca analisar os resultados e correlacioná-los com as respostas obtidas dos questionários aplicado pelo PISA do ano de 2015. Em Ciência de Dados, diversas linguagens de programação são utilizadas para manipular dados brutos e extrair informações. No trabalho com o uso da linguagem R fomos em busca de padrões que possam explicar os resultados alcançados pelos países.

1.1 JUSTIFICATIVA

Por ser assunto tão importante, diversos programas são utilizados para avaliar o ensino pelos países ao redor do mundo.

Como tentativa de quantificar e qualificar os sistemas de ensino, mecanismos cuja finalidade fora o de medir a qualidade educacional foram criados, a exemplo: TIMSS - Trends in International Mathematics and Science, e o PIRLS - Progress in International Reading Literacy Study. A principal função do TIMSS consiste em avaliar em ciclos de quatro anos o desenvolvimento e tendência matemática e conquistas científicas, já o PIRLS busca medir com ciclos de cinco anos a compreensão de leitura dos estudantes (TIMSSPIRLS, 2019).

No Brasil, programas de avaliações próprias são desenvolvidos em sincronia com os programas internacionais. A DAEB - Diretoria de Avaliação da Educação Básica junto com o INEP - Instituto Nacional de Estudos e pesquisas Educacionais Anísio Teixeira, são os responsáveis por coordenar os mais importantes sistemas de avaliação no território nacional. Dentre os vários mecanismos de avaliação de qualidade educacional no país, cabe aqui destacar alguns: Avaliação da Alfabetização Infantil (Provinha Brasil), Sistema de Avaliação da Educação Básica (SAEB), Avaliação Nacional da Educação Básica (Aneb), Avaliação Nacional do Rendimento Escolar (Anresc) e Avaliação Nacional de Alfabetização (ANA), são alguns importantes expoentes avaliativos (MEC, 2009).

Assim apresentado, pode-se aqui apontar que este se justifica devido à necessidade de se entender a importância desses mecanismos avaliativos. De tal maneira, os resultados aqui gerados poderão subsidiar outras pesquisas que tangenciam a essa área. Assim, os dados apoiam a construção de mecanismos que contribuem para um melhor entendimento dos resultados apresentados pelo PISA.

1.2 OBJETIVOS

Nesta seção são apresentados os objetivos do presente trabalho.

1.2.1 OBJETIVO GERAL

Este trabalho consiste em analisar os resultados e dados gerados a partir do Programa Internacional de Avaliação de Estudante no ano de 2015. Tem-se como objetivo produzir padrões que auxiliem na explicação dos resultados obtidos pelos países participantes das avaliações do PISA.

1.2.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos deste trabalho são:

- Estudar a teoria a respeito do que se entende por educação de qualidade, programas avaliadores de educação e educação no Brasil;
- Criar uma base de dados com os questionários aplicados a partir dos mecanismos de avaliação;
- Desenvolver código de linguagem R para manipular os dados, como as manipulações dos dados conseguir chegar a padrões;

1.3 ESTRUTURA DO TRABALHO

O presente trabalho é composto por quatro capítulos organizados da seguinte maneira:

- Capítulo 1: Introdução do trabalho e breve descrição.
- Capítulo 2: Apresentação do PISA - Programa Internacional de Avaliação de Alunos.
- Capítulo 3: Apresentação da ferramenta utilizada para desenvolvimento do trabalho.
- Capítulo 4: Apresentação das etapas do desenvolvimento, sua metodologia.
- Capítulo 5: Resultados obtidos.
- Capítulo 6: Conclusão do trabalho, considerações finais e propostas para projetos futuros.

2 FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta os conceitos necessários para o entendimento do trabalho.

2.1 PISA

Com o objetivo de melhorar os indicadores internacionais de desempenho educacional e apoiar os governos no desenvolvimento de novas políticas sociais e econômicas a OECD - Organização para a Cooperação e Desenvolvimento Econômico, formada por um aglomerado de 35 países membros e outros convidados, criou o PISA - Programa Internacional de Avaliação de Estudante o programa é uma avaliação internacional que busca avaliar estudantes da educação básica (PISA, 2001).

O programa avalia o conhecimento e habilidade dos jovens por meio de uma prova que se repete a cada três anos e se divide em três áreas de conhecimento: leitura, ciência e matemática. A cada edição uma área recebe uma ênfase maior dentro do teste, desta forma tomando dois terços do tempo de prova e exigindo uma maior atenção do aluno (PISA, 2001). Sua primeira edição ocorreu no ano de 2000 e contou com a participação de 32 países. No mesmo ano o foco esteve presente sobre a área de leitura, agora em sua aplicação mais recente em 2018 o foco novamente se repetiu em leitura.

O PISA é aplicado aos jovens que estejam ativamente matriculados a partir do sétimo ano do ensino fundamental, com faixa etária de quinze anos de idade, idade esta que se supõe que seja o fim da escolaridade básica (PISA, 2016). No ano de 2015, o programa foi aplicado para 70 países, 35 sendo membros da OECD e os outros 35 sendo convidados (PISA, 2016). Em cada país no qual ocorre a aplicação, uma organização é escolhida como coordenadora geral. No Brasil a organização responsável pela aplicação é o INEP - Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira, uma Instituição Federal associada ao MEC - Ministério da Educação (PISA, 2016).

Além de avaliar o conhecimento e habilidades dos estudantes, o PISA busca associar as habilidades alcançadas pelos alunos nas provas com fatores demográficos, sociais, econômicos e educacionais (PISA, 2016). Com isso os resultados do programa contribuem para uma discussão a respeito da qualidade do ensino e das políticas sociais presentes no país, e, para criar essas relações, são utilizados questionários que coletam informações a respeito dos alunos, professores e escolas. Os questionários tem um foco específico de acordo com quem vai respondê-lo e se divide em tópicos.

A seguir são apresentados os questionários aplicados no programa. Nos blocos divididos por assunto, encontram-se os códigos das questões e o que cada questão pretende extrair de quem esta respondendo.

2.1.1 QUESTIONÁRIO - ALUNO

O questionário respondido pelos alunos tem por objetivo coletar informações a respeito dos próprios e de sua família, além do ambiente familiar em que ele está inserido. Busca também coletar informações do seu ambiente escolar e como ele se organiza em suas atividades escolares (PISA, 2016). Assim os resultados podem ser utilizados para avaliar o que influencia o desempenho deles no programa.

Questionário - Estudante

Você, Sua Família e Sua Casa	
Código	Questão
ST001	Grade em que o aluno esta inserido.
ST002	Programa escolar que o aluno está inserido.
ST003	Data de nascimento.
ST004	Sexo do aluno.
ST005	Nível de escolaridade da mãe.
ST006	Qualificações acadêmicas da mãe.
ST007	Nível de escolaridade do pai.
ST008	Qualificações acadêmicas do pai.
ST011	Objetos que possui na residência do aluno: (TV, computador...)
ST012	Quantidade de objetos na residência.
ST013	Quantidade de livro em residência.
ST014	Em que a mãe trabalha.
ST015	Em que o pai trabalha.
ST123	Discorda ou concorda sobre suporte da família recebido.
ST019	País de origem. (aluno, pai, mãe)
ST021	Que idade o aluno chegou ao país.
ST022	Idioma falado em casa.
ST125	Idade de iniciou a educação primária.
ST126	Idade de iniciou a educação básica.
ST127	Aluno já repetiu algum ano letivo.

O que você pensa sobre sua vida**Código Questão**

- ST016 Satisfação com a vida atual.
- ST111 Nível de educação que o aluno pretende atingir.
- ST114 Tipo de trabalho quando tiver 30 anos.
- ST118 Discorda ou concorda sobre afirmações sobre si mesmo.
- ST119 Discorda ou concorda sobre afirmações sobre si mesmo.
- ST121 Discorda ou concorda sobre motivação de aprendizado.

Sua Escola**Código Questão**

- ST082 Discorda ou concorda sobre afirmações sobre si mesmo.
- ST034 Discorda ou concorda sobre sua escola.
- ST039 Experiências escolares nos últimos 12 meses.
- ST038 Experiências escolares nos últimos 12 meses.

Sua Programação Escolar e Tempo de Aprendizagem**Código Questão**

- ST059 Participação de atividades durante a semana.
- ST060 Períodos de aulas durante a semana.
- ST061 Quantos minutos dura uma aula.
- ST062 Coisas que aconteceram nas últimas duas semanas de aula.
- ST071 Tempo gasto para aprender matérias da escola. (estudo fora da escola)
- ST031 Quantos dias frequentou aulas de educação física.
- ST032 Atividades físicas praticadas fora da escola.

<Ciência Escolar> Aprendendo na Escola**Código Questão**

ST063	Cursos de ciências que participar no ano letivo.
ST064	Há opção de escolher o curso na escola.
ST097	Coisas que acontecem nas lições.
ST098	Frequência de atividades após aprender tópicos na escola.
ST065	Nome do curso de ciência na escola.
ST100	Coisas que acontecem nas lições.
ST103	Coisas que acontecem nas lições.
ST104	Coisas que acontecem nas lições.
ST107	Coisas que acontecem nas lições.

Sua Visão Sobre Ciência**Código Questão**

ST092	Como o aluno esta informado sobre questões ambientais.
ST093	Problemas ambientais que podem melhorar daqui 20 anos.
ST094	Discorda ou concorda sobre declarações sobre você mesmo.
ST095	Interesse em tópicos de ciência.
ST113	Discorda ou concorda sobre o que aprende na escola.
ST129	Facilidade em realizar tarefas sobre ciência.
ST131	Discorda ou concorda sobre ciência.
ST146	Frequências que o aluno faz certas ações.(atividades pessoais)
ST076	Atividades que o aluno realizou antes de ir para escola. (atividades pessoais)
ST078	Atividades que o aluno realizou depois da escola. (atividades pessoais)

2.1.2 QUESTIONÁRIO - PROFESSOR

O questionário submetido aos professores foi aplicado pela primeira vez no ano 2015 para poucos países. O questionário permite coletar informações sobre qualificação e desenvolvimento profissional, práticas de ensino, ambiente para aprendizagem, liderança e gerenciamento escolar (PISA, 2016).

Informação de fundo**Código Questão**

TC001	Sexo do profissional.
TC002	Idade do profissional.
TC004	Status de funcionário: (efetivo, temporário)
TC005	Status de funcionário: (tempo de trabalho)
TC006	Quantas escolas exerceu atividade.
TC007	Quantos anos de experiência.

Educação Inicial e Desenvolvimento Profissional**Código Questão**

TC012	Nível de formação.
TC013	Carreira de docente é objetivo.
TC014	Formação/treinamento de professor completo.
TC015	Qualificações de ensino.
TC018	Profissional ensina aos alunos o que aprendeu na formação.
TC029	Proporção de dedicação em qualificação.
TC020	Profissional participando de alguma atividade externa.
TC030	Dedicação nas atividades profissionais.
TC021	Profissional obrigado a participar de atividades desenvolvimento profissional.
TC045	O que foi estudado no programa de treinamento de professor.

Sua Escola**Código Questão**

TC028	Capacidade de evitar determinado problema.
TC039	Existe algum currículo formal para ciência escolar.
TC041	Quanta ênfase é dada ao currículo para ciência escolar.
TC043	Pais são informados sobre o conteúdo do currículo.
TC031	Profissional discorda ou concorda sobre cooperação no trabalho.
TC026	Como o profissional se sente em relação ao emprego.
TC060	Profissional discorda ou concorda equipe de trabalho
TC046	Frequência de realização de atividades

Práticas de Ensino de Ciência**Código** **Questão**

TC037	Ocorrência de atividades.
TC033	Liberdade na execução de atividades.
TC034	Liberdade na execução de atividades.

Práticas de Ensino**Código** **Questão**

TC048	Frequência que o professor atribui atividades a alunos.
TC051	Frequência que o professor avalia estudantes durante atividades colaborativas.
TC052	Tipos de colaboração entre os alunos que o professor utiliza.
TC053	Frequência de atividades em grupos.
TC054	Métodos usados para avaliar o aprendizado do aluno.
TC055	Atribuição de notas.

2.1.3 QUESTIONÁRIO - ESCOLA

O questionário aplicado às escolas busca coletar informações a respeito do ambiente escolar e práticas institucionais (PISA, 2016).

Informação de fundo Escolar**Código** **Questão**

SC001	Qual tamanho da comunidade em que sua escola está localizada.
SC002	Qual foi o total de matrículas escolares. (número de alunos)
SC003	Qual é o tamanho médio das turmas na escola.
SC004	Coletar informações sobre a proporção aluno/computador.
SC053	Quais atividades as escolas oferecem aos alunos. (atividades extras)
SC059	Departamento de ciências da escola. (equipamentos)
SC052	Escola oferece a ajuda no estudo. (sala de estudo)

Gestão Escolar**Código** **Questão**

- SC009 Frequência de atividades na sua escola durante o ano letivo.
- SC010 Indicar profissional responsável por tarefa.
- SC012 Fatores que são considerados na admissão de novos alunos.
- SC013 Se a escola é pública ou privada.
- SC014 Tipo de organização que administra a escola.
- SC016 De onde vem o financiamento da escola.
- SC017 Informações do que pode prejudicar a escola.

Corpo Docente**Código** **Questão**

- SC018 Tamanho do corpo docente.
- SC019 Tamanho do corpo docente em ciência.
- SC025 Professores que frequentou o programa de desenvolvimento profissional.
- SC027 Tipo de profissional a escola possui.

Avaliação**Código** **Questão**

- SC032 Método usado para avaliar a prática do professor.
- SC034 Qual frequência os alunos são avaliados.
- SC035 Testes padronizados ou desenvolvidos pelo profissional.
- SC036 Onde são usados os dados de desempenho.
- SC037 De onde vem os arranjos para garantir a qualidade na escola.
- SC040 Escola implementou alguma medida de melhoria após os resultados do programa.
- SC041 Mudanças que ocorrem na escola depois dos resultados do programa.

Grupo Orientado**Código** **Questão**

- SC042 Políticas presente na escola destinada aos alunos.
- SC048 Características dos estudantes da escola.

Clima Escolar**Código** **Questão**

SC061 O que dificulta o aprendizado dos alunos.

SC063 Envolvimento dos pais na escola.

2.1.4 APLICAÇÃO DO PISA

Por não ser possível aplicar o PISA em todas as escolas de todos os países, a coordenadora responsável de cada país seleciona uma amostra de instituições que são contactadas e orientadas sobre o que fazer para que ocorra a aplicação. A aplicação do PISA ocorre da seguinte forma:

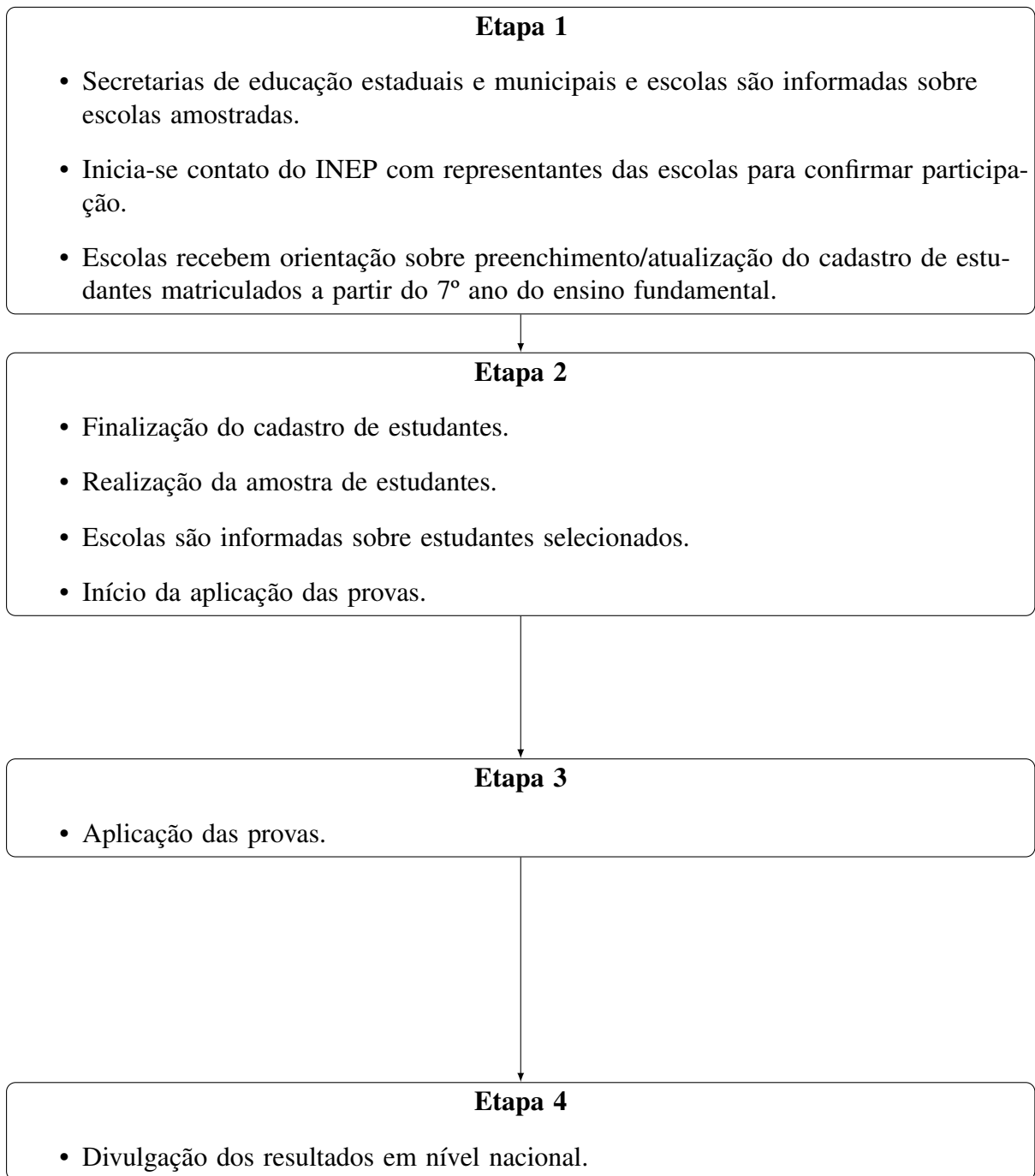


Figura 1 – Fluxograma - Os passos da Aplicação do PISA.

2.2 PISA NO BRASIL

No ano de 2000, o Brasil participou pela primeira vez do PISA como convidado e desde então sempre participa das edições do programa. A última ocorreu no ano de 2018, mas os resultados só irão estar disponíveis no segundo semestre do ano de 2019. Logo, temos à nossa disposição os resultados referentes ao ano de 2000 até o ano de 2015. Ao analisar as seis edições, podemos observar um dado interessante em relação ao número de estudantes que participaram do PISA, pois este número teve oscilações, mas teve um aumento considerável ao decorrer dos

anos.

Na tabela é possível ver a variação do número de alunos ao decorrer dos anos.

Ano	Alunos
2000	4.893
2003	4.452
2006	9.295
2009	20.127
2012	18.589
2015	23.141

Tabela 1 – Número de Alunos por Aplicação

Com os dados do PISA referente ao ano de 2015, encontramos o Brasil com notas abaixo da média nas três áreas: ciência(401), leitura(407) e matemática(377). Em relação a média geral do programa, que foi de: ciência(493), leitura(493) e matemática(490).

A seguir é possível observar a evolução dos resultados do Brasil nas seis edições do PISA.

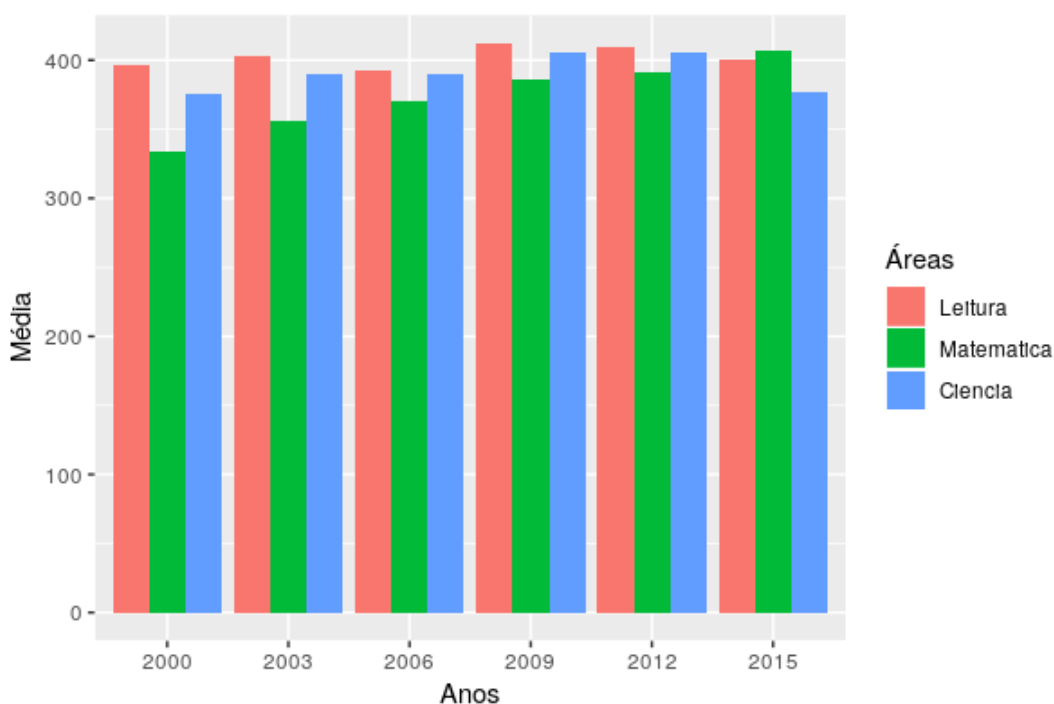


Figura 2 – Notas do Brasil nas edições do PISA

Ao observar o gráfico, a área de matemática é a única que apresenta um crescimento constante mesmo que pequeno, representada pela cor verde.

Um outro dado apresentado referente ao Brasil é o aumento de jovens com faixa etária de quinze anos se encontravam matriculados nas escolas. No ano de 2015 aproximadamente 71% dos jovens se encontravam matriculados, o que mostra uma evolução em relação ao ano de 2003, que tinha apenas 56%, mostrando assim um aumento na escolaridade no Brasil (OECD, 2016).

3 FERRAMENTA UTILIZADA

Este capítulo apresenta a linguagem R.

3.1 LINGUAGEM R

A linguagem de programação R foi desenvolvida por Ross Ihaka e Robert Gentleman por volta dos anos 90. O foco inicial foi na computação estatística, mas atualmente, a linguagem é muito usada na análise e manipulação de dados (DIONISIO, 2015).

A linguagem permite ao seu usuário realizar uma enorme variedade de modelagens com dados como: modelagem de regressão, linear e não linear, análise de séries temporais, análise fatorial e outras modelagem (AQUINO, 2014). Seu crescente uso se dá por trazer praticidade na manipulação e exibição gráfica de dados, fornecer uma certa simplicidade ao realizar inserção de fórmulas e símbolos matemáticos, além de gerar plotagens de alta qualidade (DIONISIO, 2015).

Com a linguagem é possível aplicar diversas funções estatística durante a manipulação de dados. Abaixo algumas das muitas funções que a linguagem permite realizar (OLIVEIRA; GUERRA; MCDONNELL, 2018) :

- `mean()`: calcula a média.
- `sum()`: o somatório dos elementos.
- `summary()`: mostra o resumo estatístico de um conjunto.
- `var()`: calcula a variância.
- `sd()`: calcula o desvio padrão.
- `median()`: calcula a mediana.
- `lm()`: retorna a estimativa dos parâmetros do modelo linear.



Figura 3 – Logo R

Fonte: The R Project for Statistical Computing

Por ser uma linguagem multiparadigma, R se tornou uma linguagem de sucesso por ser capaz de solucionar problemas de diferentes estilos, assim fazendo com que atenda um número

maior de usuários. Alguns paradigmas que a linguagem utiliza é o funcional e orientado a objeto. O paradigma funcional teve sua origem a partir de funções matemáticas e sua execução ocorrer por meio de expressões e funções (MANSSOUR, 2018). O paradigma orientado a objeto fornece apoio para tipos abstratos, herança e amarração dinâmica, a criação dos objetos, permitindo envio e recebimento de mensagens dentro do programa (MANSSOUR, 2018).

3.1.1 LINGUAGENS CONCORRENTES AO - R

Dentro da área de ciência de dados a linguagem R já é referência, mas junto com ela algumas linguagens auxiliam pesquisadores nesta área. Em outras palavras há linguagens que concorrem com a R como: Python, Matlab, Octave, Scilab, Julia e outras. Abaixo vamos falar um pouco de cada uma.

3.1.1.1 PYTHON

A linguagem Python foi desenvolvida na década de 90, mesma época que a linguagem R, por Guido Van Rossum baseada na linguagem ABC. Inicialmente o público alvo eram pessoas atuantes nas áreas de física e engenharia (BORGES, 2010). Por possuir uma sintaxe clara e objetiva, usuários encontram uma facilidade em aprender-lá. Vale ressaltar também que a linguagem é de software livre sendo possível utilizá-la em qualquer sistema operacional, além de ser sempre atualizada pela Python Foundation, organização sem fins lucrativos dedicados a linguagem (MENEZES, 2014).

Por sua simplicidade de aprendizagem Python cresce cada vez mais, assim é possível encontrar diversas aplicações desenvolvidas na linguagem como jogos, aplicações web, aplicações em banco de dados, inteligência artificial, animações 3D e muitas outras. Isso só é possível por ser uma linguagem completa que torna o trabalho do desenvolvedor mais produtivo e com menos erros durante o desenvolvimento (MENEZES, 2014).



Figura 4 – Logo Python

Fonte: Python.org

3.1.1.2 MATLAB

Matlab foi criada por volta dos anos de 1970 e seu criador é Cleve Moler. A linguagem é atualmente utilizada no ensino e pesquisa, trabalha exclusivamente com análise numérica,

cálculo matricial e possui ótimos recursos para criação de gráfico para representação de resultados (SIQUEIRA, 2007)

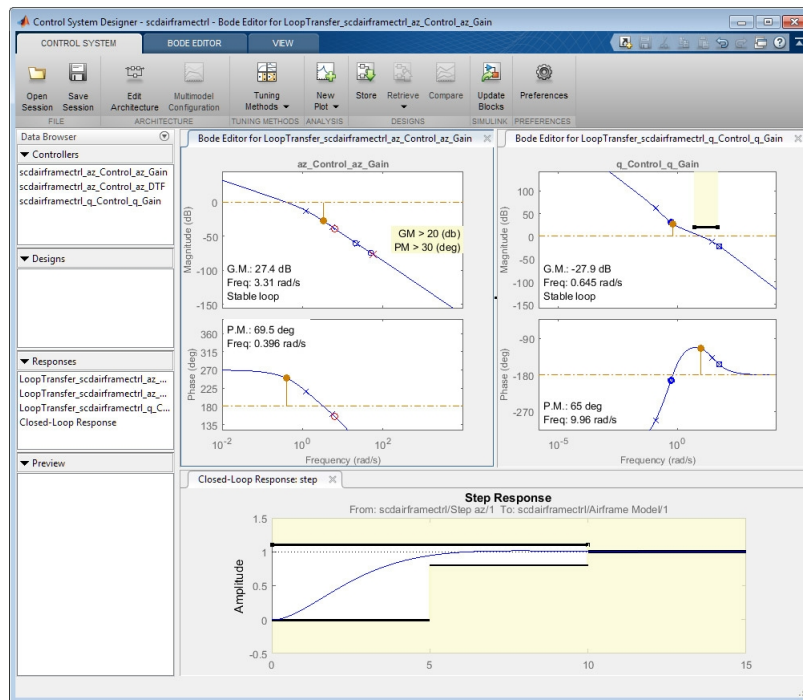


Figura 5 – Interface MATLAB

Fonte: Mathworks.com

3.1.1.3 OCTAVE

A linguagem teve sua origem a partir de um projeto de um reator químico através das mãos dos estudantes de graduação James B. Rawlings e John G. Ekerdt. Além de ter sido usada para ensinar o projeto, ela foi utilizado para ensinar equações diferenciais, álgebra linear pelos departamentos de química e matemática daquela época. Esta linguagem possui uma similaridade forte com Matlab e R citados anteriormente (SOUZA, 2003). Octave é uma linguagem com foco na resolução de problemas de computação matemática, por meio de cálculos aritméticos, escalares complexos, matrizes e diversos outros cálculos (MARQUES; MORGADO, 2010).

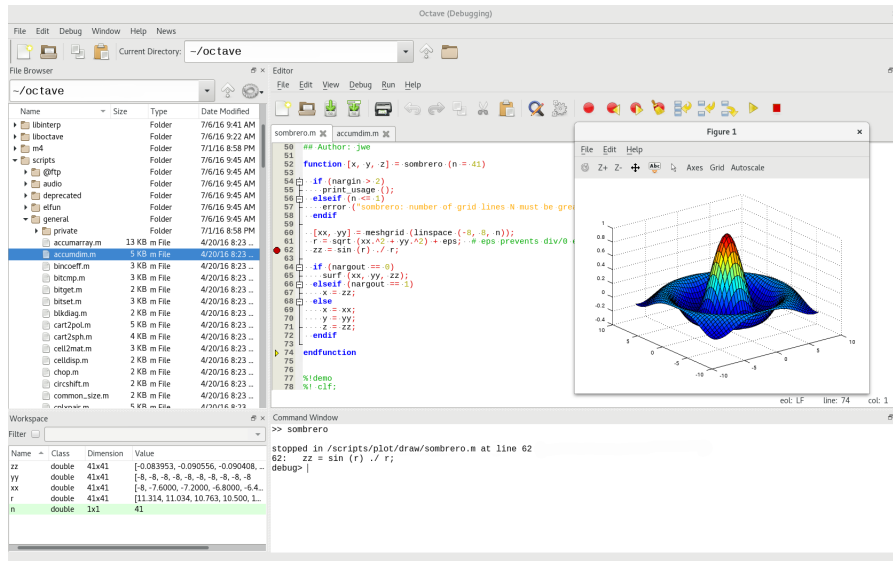


Figura 6 – Interface Octave

Fonte: Gnu.org

3.1.1.4 SCILAB

Uma linguagem de código aberto e de software livre semelhante às linguagens de programação R e python, mas sua sintaxe é similar ao Matlab sendo possível transitar de uma para a outra com facilidade, pois ambos trabalham para resolver problemas por meio de métodos numéricos (FONTANA, 2018). A Scilab permite ao usuário criar representações gráficas bidimensionais, tridimensionais e até animações, além de possibilita trabalhar com polinômio, sistemas lineares e muitas outras possibilidades (PIRES, Julho, 2004).

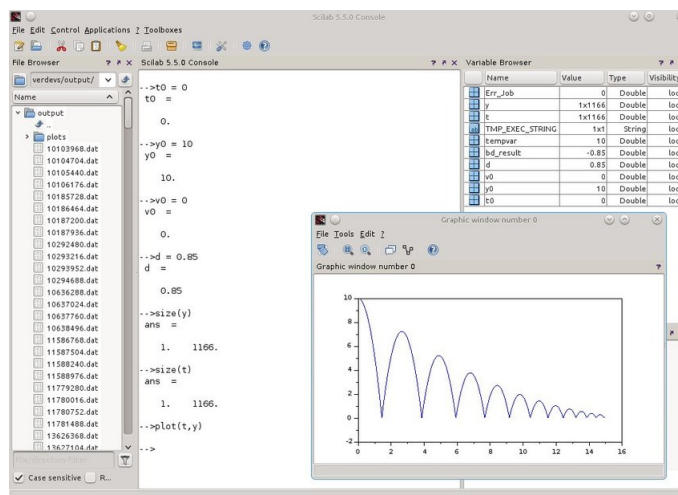


Figura 7 – Interface Scilab

Fonte: Researchgate.net

3.1.1.5 JULIA

A linguagem é relativamente nova se comparada com R e Python por exemplo. Julia foi criada por Stefan Karpinski, Jeff Bezanson, Alan Edelman e Viral shaa em 2009 com o foco em computação científica. Inicialmente a linguagem era de código fechado mas no ano de 2012 se tornou de código aberto (PEREIRA, 2017).

Julia possui uma linguagem intuitiva e similar ao Matlab, além de possuir um alto desempenho computacional, sendo próximo ao desempenho da linguagem C, além de outras características que faz com que a linguagem ganhe cada vez mais espaço (PEREIRA, 2017).

```
bruno@bruno-S460-L-BK26P1: ~
Arquivo Editar Ver Pesquisar Terminal Ajuda
bruno@bruno-S460-L-BK26P1:~$ ~/opt/julia-0.5.0/bin/julia
| A fresh approach to technical computing
| Documentation: http://docs.julialang.org
| Type "?help" for help.
|
| Version 0.5.0 (2016-09-19 18:14 UTC)
| Official http://julialang.org/ release
| x86_64-pc-linux-gnu
julia> Pkg.add("PyPlot")
INFO: Initializing package repository /home/bruno/.julia/v0.5
INFO: Cloning METADATA from https://github.com/JuliaLang/METADATA.jl
ERROR: GitError(Code:ERROR, Class:Net, SSL error: error:1407742E:SSL routines:SSL
L23_GET_SERVER_HELLO:tlsv1 alert protocol version)
in macro expansion at ./libgit2/error.jl:99 [inlined]
in clone(::String, ::String, ::Base.LibGit2.CloneOptions) at ./libgit2/reposito
ry.jl:191
in #clone#109(::String, ::Bool, ::Ptr{Void}, ::Nullable{Base.LibGit2.AbstractCr
edentials}, ::Function, ::String, ::String) at ./libgit2/libgit2.jl:327
in (::Base.LibGit2.#kw##clone)::Array{Any,1}, ::Base.LibGit2.#clone, ::String,
::String) at ./<missing>:0
in (::Base.Pkg.Dir.##4#6{String,String})() at ./pkg/dir.jl:49
in cd(::Base.Pkg.Dir.##4#6{String,String}, ::String) at ./file.jl:59
```

Figura 8 – Interface Julia

3.1.1.6 COMPARATIVO DAS LINGUAGENS

Um teste realizado por (KOUATCHOU, Dezembro, 2016) com objetivo de observar como as linguagens Python, Julia, R, Matlab e outras, lidavam com loops e vetorização. O teste foi realizado em um: Intel Xeon Haswell, cada nó tem 28 núcleos (2,6 GHz cada) e 128 Gb de memória disponível.

A seguir é apresentado apenas os resultados em milissegundos das linguagens Python, Julia, R e Matlab.

Problema: Considere uma matriz ($n \times n \times 3$) arbitrária. Queremos executar as seguintes operações em A:

- $A(i,j,1) = A(i,j,2)$
- $A(i,j,3) = A(i,j,1)$
- $A(i,j,2) = A(i,j,3)$

Linguagens	Versão	n=5000	n=7000	n=9000
Python	(2.7.1)	19.12	37.49	61.97
Julia	(0.6.0)	0.10	0.22	0.34
R	(3.2.2)	233.78	451.77	744,93
Matlab	(R2016a)	2.20	4.11	6.80

Tabela 2 – Tempo de processamento

Fonte: (KOUATCHOU, Dezembro, 2016)

Nesta comparação Julia, a linguagem mais nova, teve o melhor desempenho em relação às outras linguagens, mostrando assim seu potencial.

3.1.2 INTERFACES DE TRABALHO

Ao trabalhar com R é possível utilizar o R consoler, RStudio, Emacs e ESS ou outras interface que a linguagem disponibiliza. O R consoler é ambiente que demanda do usuário um conhecimento maior da linguagem, por ser apenas uma área para execução de comandos, o que torna o aprendizado mais exaustivo e pode tornar o trabalho improdutivo (OLIVEIRA; GUERRA; MCDONNELL, 2018).

Nesse trabalho a opção escolhida foi a interface gráfica RStudio que disponibiliza um ambiente organizado e repleto de funcionalidades, no qual torna o aprendizado da linguagem mais rápido e mais produtivo (COSTA, Dezembro, 2017).

```

Error in install.packages : 'match' requires vector arguments
> "ggplot2"
[1] "ggplot2"
> install.packages("haven")
Error in install.packages : Updating loaded packages
> install.packages("haven")
Installing package into '/home/bruno/R/x86_64-pc-linux-gnu-library/3.4'
(as 'lib' is unspecified)
trying URL 'https://cloud.r-project.org/src/contrib/haven_2.1.0.tar.gz'
Content type 'application/x-gzip' length 263276 bytes (257 KB)
=====
downloaded 257 KB

* installing *source* package 'haven' ...
** package 'haven' successfully unpacked and MD5 sums checked
** libs
gcc -std=gnu99 -I/usr/share/R/include -DNDEBUG -I"/home/bruno/R/x86_64-pc-linux-gnu-library/3.4/Rcpp/include" -Ireadstat -DHAVE_ZLIB -fpic -g -O2 -fdebug-prefix-map=/build/r-base-AitvI6/r-base-3.4.4=. -fstack-protector-strong -Wformat -Werror=format-security -Wdate-time -D_FORTIFY_SOURCE=2 -g -c tagged_na.c -o tagged_na.o
gcc -std=gnu99 -I/usr/share/R/include -DNDEBUG -I"/home/bruno/R/x86_64-pc-linux-gnu-library/3.4/

```

Figura 9 – R Consoler

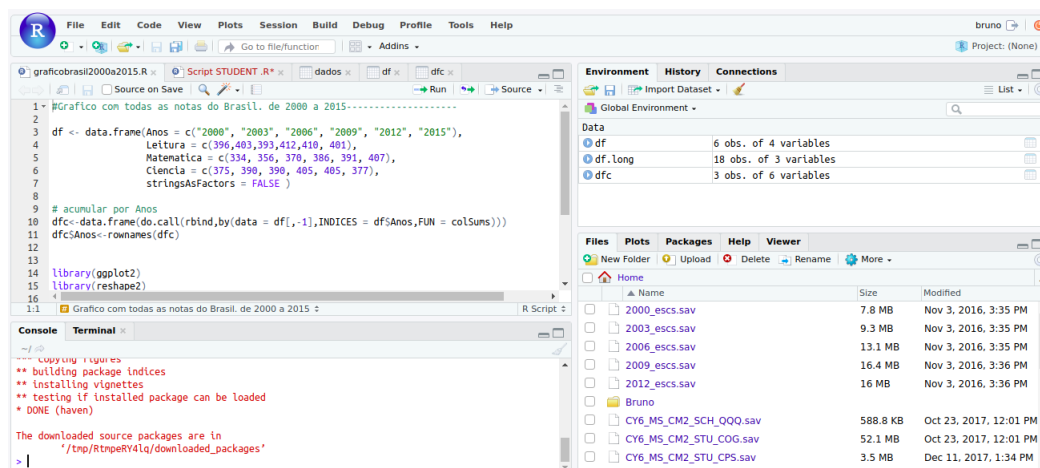


Figura 10 – Interface RStudio.

A interface é dividida em quatro partes e (OLIVEIRA; GUERRA; MCDONNELL, 2018) explica o que cada uma faz.

Editor de Código: Nesta parte é onde o usuário escreve e edita os scripts. Script nada mais é do que uma sequência de comandos que serão executados em sequência pelo R.

Console: É no console que o R mostrará a maioria dos resultados. Também é possível escrever os comandos diretamente no console, sem o uso do editor de código. É muito utilizado para testes e experimentos rápidos.

Environment e History: No Environment ficarão guardados todos os objetos que forem criados. Na aba History, o RStudio cria um histórico de comandos utilizados.

Files, Plots, Packages, Help e Viewer: Nesta janela, estão várias funcionalidades do RStudio.

- Files: Apresenta os arquivos constantes do diretório de trabalho.

- Plots: Aba onde ficam os gráficos gerados.
- Packages: Onde se encontram os pacotes instalados.
- Viewer: Utilizada para visualização de conteúdo da web.

3.1.3 IMPORTAÇÃO DE ARQUIVOS

A linguagem possibilita trabalhar com arquivos de diferentes formatos, de forma prática e precisa.

Arquivo CSV

Para importar e trabalhar com dados CSV é necessário ter o pacote (readr) instalado no R, assim é só chamar suas funcionalidades e realizar a leitura dos arquivos.

Exemplo de leitura de arquivo .CSV:

```
library(readr)
dados <- read_csv("Bruno/arquivo.csv")
View(dados)
```

Arquivo SPSS

Esses arquivos recebem extensão (.sav), para sua leitura é necessário utilizar o pacote (haven).

Exemplo de leitura de arquivo .CSV:

```
library(haven)
dados <- read_sav("arquivo.sav")
View(dados)
```

Arquivo SAS

No caso de arquivos com formato SAS também se usa o pacote (haven) para realizar a importação, além de informar ao software o caminho do arquivo.

Exemplo de leitura de arquivo .CSV:

```
library(haven)
dados <- read_sas("Bruno/Base_Dados/arquivo.sas7bdat", NULL)
View(dados)
```

Arquivo JSON

Para realizar a leitura de arquivos com esse formato é importante instalar o pacote (rjson) no R.

Exemplo de leitura de arquivo .JSON:

```
library(rjson)
dados <- fromJSON(file = "arquivo.json")
View(dados)
```

Novamente o uso da interface RStudio traz maior simplicidade no momento de importação de arquivos. Abaixo um exemplo de importação de arquivo no formato (SAS).

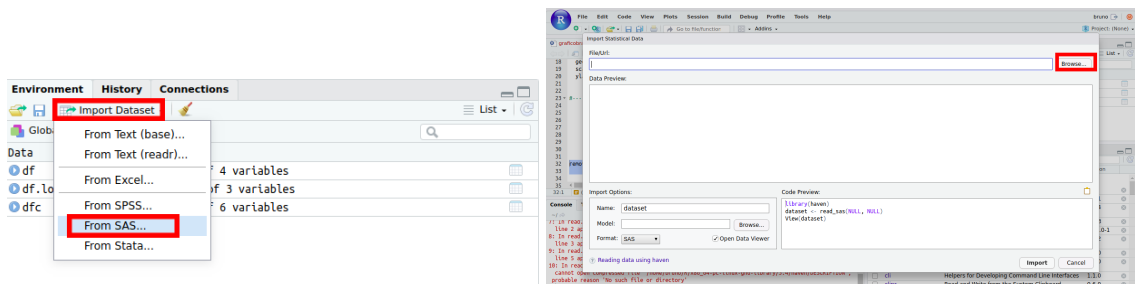


Figura 11 – Importando arquivos pela interface RStudio - Passo 1 e 2

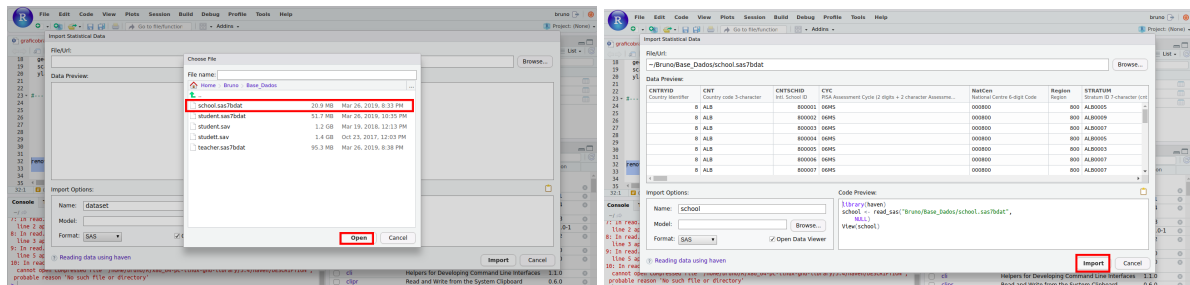


Figura 12 – Importando arquivos pela interface RStudio - Passo 3 e 4

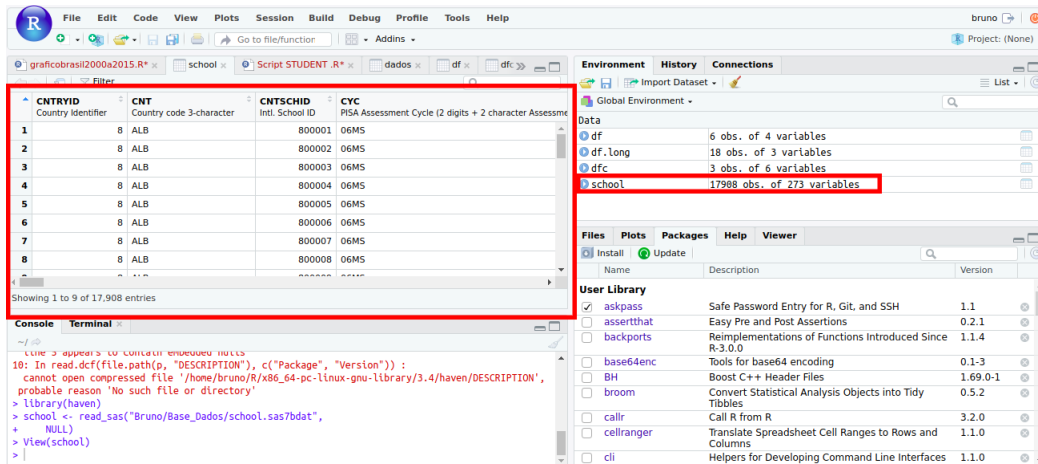


Figura 13 – Importando arquivos pela interface RStudio - Passo 5 - Arquivo importado

3.1.4 GERENCIAMENTO DO R

Como já mencionado o R permite uma extensa variedade de modelagem com os dados, mas isso só é possível com o auxílio de pacotes que se encontram disponíveis no CRAN, repositório oficial da linguagem. (MAYER, 2015) explica como o usuário pode realizar essas ações de gerenciamento.

Para instalar um pacote é necessário utilizar o comando abaixo.

```
install.packages("ggplot2")
```

Por meio do código abaixo é possível obter informações como: pacotes atualizados, pacotes que podem ser atualizados, pacotes indisponíveis, pacotes instalados e não instalados.

```
packageStatus()
```

```
> packageStatus()
Number of installed packages:

                ok upgrade unavailable
/home/bruno/R/x86_64-pc-linux-gnu-library/3.4 75      12         0
/usr/local/lib/R/site-library                0         0         0
/usr/lib/R/site-library                      0         0         1
/usr/lib/R/library                           14         0         0

Number of available packages (each package counted only once):

                installed not installed
https://cloud.r-project.org/src/contrib      87      13926
> |
```

Figura 14 – Resposta do comando: packageStatus()

O comando abaixo permite observar ver os nomes dos pacotes, até mesmo todo os repositórios do CRAN.

```
summary(packageStatus())
```

```
> summary(packageStatus())
```

```
Installed packages:
```

```
-----
```

```
*** Library /home/bruno/R/x86_64-pc-linux-gnu-library/3.4
```

```
$ok
```

```
[1] "askpass"      "assertthat"  "backports"   "base64enc"   "BH"          "broom"
[7] "callr"        "cellranger"  "cli"         "colorspace"  "crayon"      "curl"
[13] "DBI"          "evaluate"    "fansi"       "forcats"     "foreign"     "fs"
[19] "gdtools"     "generics"    "ggiraph"     "glue"        "gtable"      "haven"
[25] "highr"       "hms"         "htmltools"   "htmlwidgets" "httr"        "jsonlite"
[31] "knitr"       "labeling"    "lattice"     "lazyeval"    "lubridate"   "magrittr"
[37] "modelr"      "munsell"     "nlme"        "openssl"     "pkgconfig"   "plogr"
[43] "plyr"        "prettyunits" "processx"    "ps"          "purrr"       "R6"
[49] "RColorBrewer" "Rcpp"        "readr"       "readxl"      "rematch"     "reprex"
[55] "reshape2"    "rmarkdown"  "rstudioapi"  "rvest"       "scales"      "selectr"
[61] "stringi"     "stringr"    "sys"         "TidlyWikiR" "tidyr"       "tidyselect"
[67] "tidyverse"   "tinytex"    "utf8"        "viridisLite" "whisker"     "withr"
[73] "xfun"        "xml2"       "yaml"
```

```
Supgrade
```

```
[1] "clipr"      "dbplyr"     "digest"     "dplyr"       "ellipsis"   "ggplot2"   "markdown"   "mime"
[9] "pillar"    "progress"  "rlang"      "tibble"
```

Figura 15 – Resposta do comando: `summary(packageStatus())`

Também o R permite: Atualiza e Remover pacotes instalados.

```
update.packages("ggplot2")
```

```
remove.packages("ggplot2")
```

A interface RStudio permite realizar essas funções sem a necessidade de uma linha de código como podemos ver nas imagens a seguir.

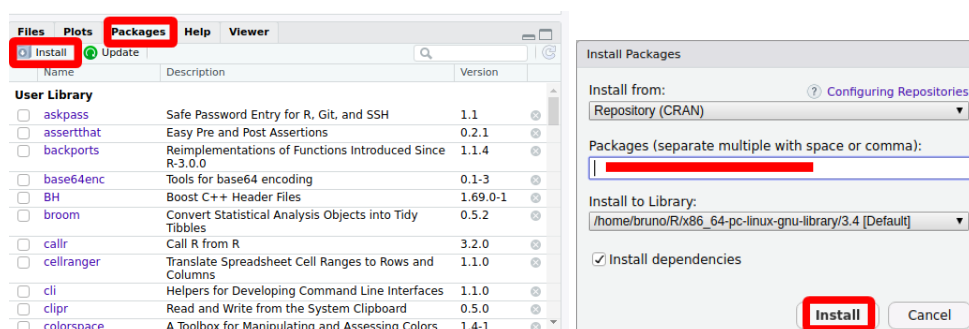


Figura 16 – Instalando pacotes pela interface

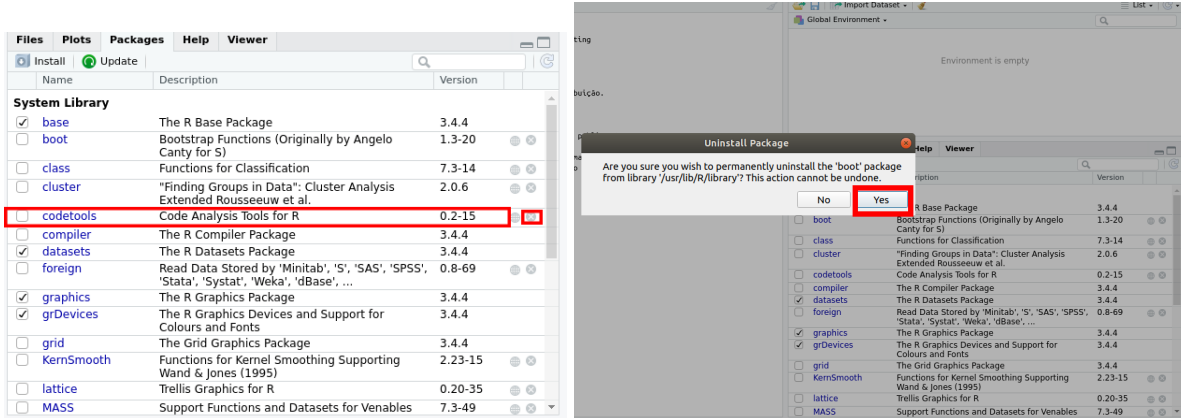


Figura 17 – Removendo pacotes pela interface

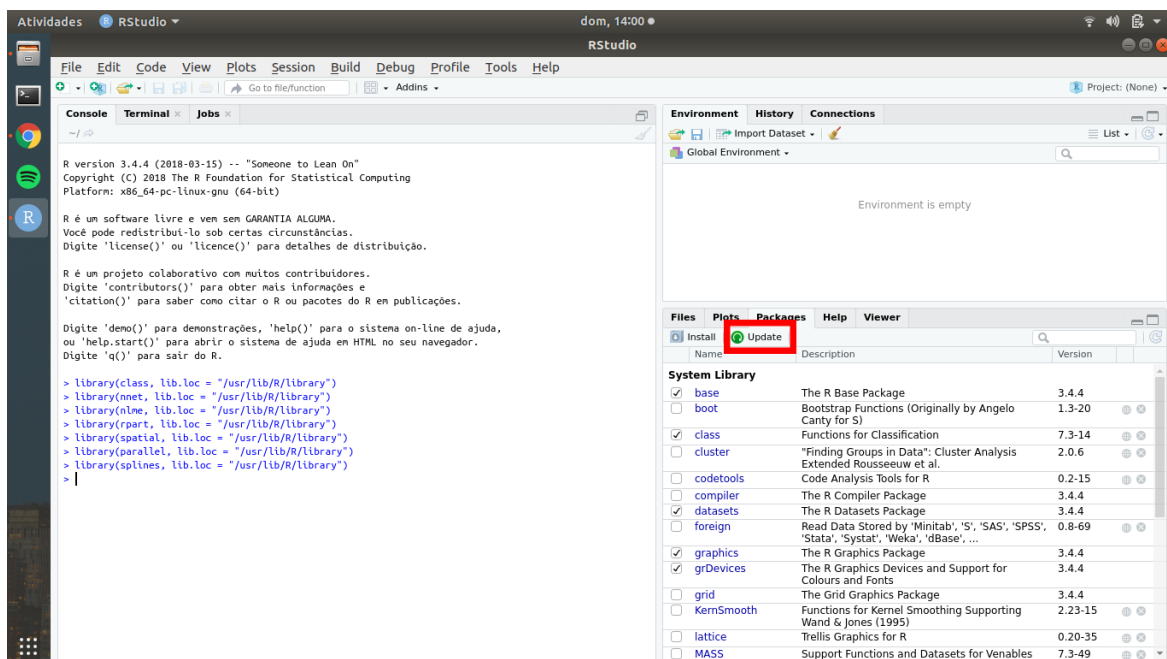


Figura 18 – Atualizar pacotes pela interface

3.1.5 PROCESSO DE INSTALAÇÃO DO - R:

Para instalação do R e do RStudio no sistema operacional Linux.

```

--- No terminal
--- Instalar o R
sudo apt-get install r-base r-base-dev

--- Adicionar o repositório
sudo echo deb http://cran.rstudio.com/bin/linux/ubuntu xenial/

--- Atualiza
sudo apt-get update}

```



```
--- Instalar
sudo apt-get install r-base r-base-dev}
```

Fonte: vivaolinux.com.br

Após a instalação do R no ubuntu é momento de instalar a interface RStudio.

```
--- Ir no site
https://www.rstudio.com/products/rstudio/download/
e baixar o arquivo para a distribuição pertinente de Linux.
```

Fonte: vivaolinux.com.br

The screenshot shows the RStudio website's download page. On the left, there are four license options: RStudio Desktop (Open Source License, FREE), RStudio Desktop (Commercial License, \$995 per year), RStudio Server (Open Source License, FREE), and RStudio Server Pro (Commercial License, \$4,975 per year). The 'DOWNLOAD' button for the Open Source Desktop version is highlighted with a red box. On the right, there is a table titled 'Installers for Supported Platforms' with columns for Installers, Size, Date, and MD5. The row for 'RStudio 1.2.1335 - Ubuntu 18 (64-bit)' is highlighted in red.

Installers	Size	Date	MD5
RStudio 1.2.1335 - Windows 7+ (64-bit)	126.9 MB	2019-04-08	d9e2470f1f8ef4cd35a669aa323a2136
RStudio 1.2.1335 - Mac OS X 10.12+ (64-bit)	121.1 MB	2019-04-08	6c570b0e21445837f7c48c284ce290eef
RStudio 1.2.1335 - Ubuntu 14 (Debian 8) (64-bit)	92.2 MB	2019-04-08	c1b0709511469ab7e582919b183eeeb3
RStudio 1.2.1335 - Ubuntu 16 (64-bit)	99.3 MB	2019-04-08	c142d69c216257fb1b018c045fff13c7
RStudio 1.2.1335 - Ubuntu 18 (64-bit)	100.4 MB	2019-04-08	71a8d1398c0d07939804b46cf0b0aa375
RStudio 1.2.1335 - Fedora 29 (Redhat 7+) (64-bit)	114.1 MB	2019-04-08	296b6e188969b912977ab6545f256a7a
RStudio 1.2.1335 - Debian 9+ (64-bit)	100.6 MB	2019-04-08	1e32d4d6f6e216f086a81ca82ef65a91
RStudio 1.2.1335 - OpenSUSE 15+ (64-bit)	101.6 MB	2019-04-08	2795a63c7ef08e2aa26d
RStudio 1.2.1335 - SLES/OpenSUSE 12+ (64-bit)	94.4 MB	2019-04-08	c65424b6e6f673727999

Figura 19 – Baixando arquivo .deb

Fonte: rstudio.com

```
--- No terminal
--- Instalar o gdebi:
sudo apt install gdebi

--- Abrir pelo terminal a pasta onde o arquivo foi baixado.
--- Instalando o Rstudio.
sudo gdebi -n rstudio-xenial-1.0.153-amd64.deb
```

Fonte: vivaolinux.com.br

Na figura representa a interface do RStudio.

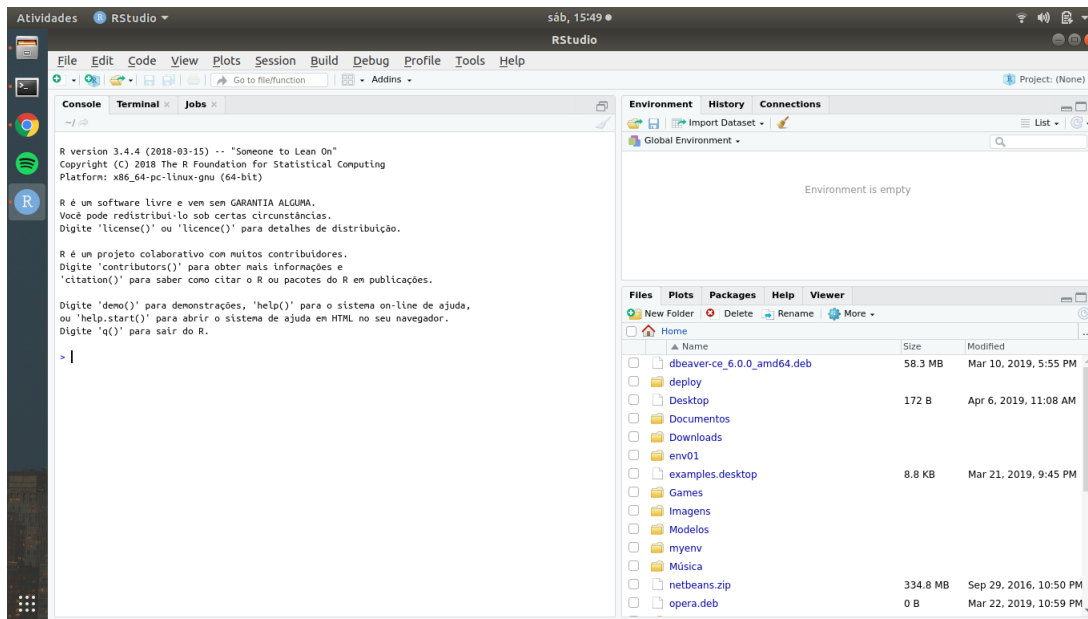


Figura 20 – Interface RStudio

No capítulo seguinte vamos explicar os métodos aplicados no trabalho, e, como será feita a manipulações dos dados do PISA.

4 METODOLOGIA

Os dados dos questionários que irão ser analisados se encontram na plataforma da OECD, todos no formato .SAS e contem informações a respeito de cada país como: sigla, região, linguagem, etc. Também se encontram nos arquivos as respostas dos professores, alunos e escolas. Esses arquivos serão manipulados com auxílio da linguagem R (SILVA; FARIA; FERNANDES, 2018). Para trabalhar de forma eficiente com os dados, serão usados pacotes que irão auxiliar no processo de manipulação dos dados, estes pacotes foram criados por Hadley Wickhamum, estatístico da Nova Zelândia e atual cientista chefe do RStudio.

O uso do HAVEN permite a leitura dos arquivos utilizados.

```
read_sas ("dados.sas7bdat ")
```

O pacote DPLYR é repleto de gramática para manipulação de dados, o que torna a criação do código mais simples e intuitivo, tornando a manipulação mais simples.

Algumas das funções do DPLYR são:

- filter() - filtra linhas.
- select() - seleciona colunas.
- mutate() - cria/modifica colunas.
- arrange() - ordena a base.
- summarise() - sumariza a base.

Abaixo um exemplo utilizado no trabalho para gerar a média do tamanho de turmas e ao mesmo tempo agrupa o resultado de acordo com as siglas dos países, representada pela (CNT).

```
tamanho_medio_classes = school%>%
group_by(CNT) %>%
summarise(SC003Q01TA = mean(SC003Q01TA, na.rm = TRUE))
```

Para representar os resultados em visualização gráfica, o uso GGLOT2 possibilita criar diferentes estilos de gráficos além de tornar possível adaptações no gráfico de acordo com o problema, não tornando a representação restrita a criação de gráficos de um conjunto já pré-selecionado.

Abaixo uma exemplo do uso do pacote.

```
ggplot(dados, aes(x = sigla, y = valor)) +  
geom_point(color = "black", fill = "yellow", shape = 24, size=4)+  
theme_light()+  
geom_hline(size = 0.3, color = "red", yintercept=4)+  
xlab(" ") + ylab(" ")
```

A figura mostra o resultado do código acima.

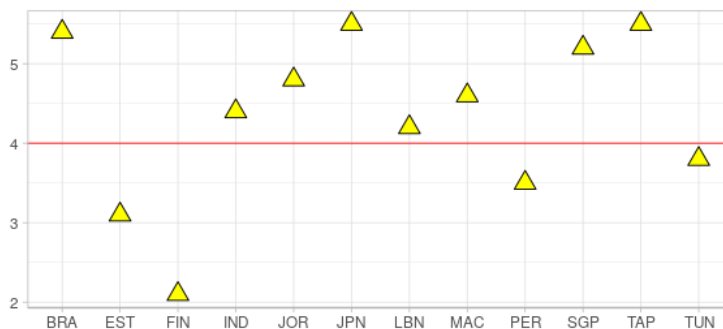


Figura 21 – Gráfico exemplo com uso - GGLOT2

SELEÇÃO DE DADOS

Foram selecionado os questionários respondidos pelas escolas e professores, por apresentar muitas questões foram escolhidas duas questões para serem analisadas.

Questionário Escola

Questão (cod: SC003Q01TA) - Qual é o tamanho médio das classes na sua escola?

Tamanho médio das turmas que realizaram a avaliação. Para responder essa questão a escola tinha 9 opções para o tamanho das turmas:

Número de Alunos	
1	15 ou menos
2	16 a 20
3	21 a 25
4	26 a 30;
5	31 a 35
6	36 a 40
7	41 a 45
8	46 a 40
9	mais de 50

Tabela 3 – Possíveis respostas

Com os resultados, foi feita uma soma de todas as respostas de cada país e no final uma média foi gerada. Assim é possível ter um resultado que represente o tamanho médio das turmas dos países.

Questão (cod: SC014Q01NA) - Que tipo de organização administra sua escola?

As escolas participantes do programa tinham como resposta as seguintes administrações.

Administração	
1	Organizações sem fins lucrativos
2	Organizações com fins lucrativos
3	Organizações religiosas

Tabela 4 – Tipos de Administração

Com esses dados, vamos representar graficamente os países com suas respectivas administrações. Além de comparar esses países com seus resultados, para descobrir se alguma administração é melhor.

Questionário Professor

Questão (cod: TC004Q01NA) Qual é o seu status de emprego como professor nessa escola?

Assim é possível analisar a quantidade de professores efetivos, temporários com contrato de mais de um ano e temporários com contrato menor que um ano.

Tempo	
1	Emprego permanente
2	Contrato por tempo determinado para um período superior a 1 ano
3	Contrato a termo certo por um período de um ano letivo ou menos

Tabela 5 – Status d Empregado

Para analisar optamos por analisar apenas os professores que responderam que possuem o status de (Emprego permanente).

Questão (cod: TC012Q01NA) Qual é o nível mais alto de educação formal que você completou?

Nível de formação dos professores, o nivelamento é feito por ISCED - International Standard Classification of Education, uma forma de classificar os níveis educacionais começando do nível 0 até o nível 9. Para essa pergunta os professores tinham como opção responder:

Nível	
Abaixo do nível 5 (<i>nível educacional pós secundário, mas não sendo considerado superior, podendo ser uma especialização.</i>)	
1	
Nível 5B (<i>um nível educacional terciário, de formação curta, mas ainda não sendo considerado uma graduação</i>)	
2	
Nível 5 A (<i>um nível educacional terciário, uma graduação</i>)	
3	
Nível 5 A (<i>um nível educacional terciário, um mestrado</i>)	
4	
Nível 6 (<i>um nível educacional terciário, um doutorado</i>)	
5	

Tabela 6 – Nível de formação

Dentre as opções de respostas que o professor possui, a análise será feita sobre a terceira opção (Nível 5 A). Que representa o número de professores com nível de graduação.

5 ANÁLISE DOS RESULTADOS

5.1 INTRODUÇÃO

Este capítulo tem por objetivo analisar os dados referentes aos resultados do PISA no ano de 2015. A Tabela a seguir apresenta a lista dos países avaliados com suas respectivas siglas e notas nas três áreas avaliadas.

PAÍS	SIGLA	Ciencia	Leitura	Matematica
Albania	ALB	427	405	413
Emirados Arabes Unidos	ARE	437	434	427
Australia	AUS	510	503	494
Austria	AUT	495	485	497
Belgica	BEL	502	499	507
Bulgaria	BGR	446	432	441
Brasil	BRA	401	407	377
Canada	CAN	528	527	516
Suica	CHE	506	492	521
Chile	CHL	447	459	423
Colombia	COL	416	425	390
Costa Rica	CRI	420	427	400
Republica Checa	CZE	493	487	492
Alemanha	DEU	509	509	506
Dinamarca	DNK	502	500	511
Republica Dominicana	DOM	332	358	328
Argelia	DZA	376	350	360
Espanha	ESP	493	496	486
Estonia	EST	534	519	520
Finlandia	FIN	531	526	511
Franca	FRA	495	499	493
Reino Unido	GBR	509	498	492
Georgia	GEO	411	401	404
Grecia	GRC	455	467	454
Hong Kong	HKG	523	527	548
Croacia	HRV	475	487	464
Hungria	HUN	477	470	477
Indonesia	IDN	403	397	386
Irlanda	IRL	503	521	504

Islandia	ISL	473	482	488
Israel	ISR	467	479	470
Italia	ITA	481	485	490
Jordania	JOR	409	408	380
Japao	JPN	538	516	532
Coreia	KOR	516	517	524
Libano	LBN	386	347	396
Lituania	LTU	475	472	478
Luxemburgo	LUX	483	481	486
Letonia	LVA	490	488	482
Macao - China	MAC	529	509	544
Moldavia	MDA	428	416	420
Mexico	MEX	416	423	408
Malta	MLT	465	447	479
Montenegro	MNE	411	427	418
Paises Baixos	NLD	509	503	512
Noruega	NOR	498	513	502
Nova Zelandia	NZL	513	509	495
Peru	PER	397	398	387
Polonia	PLO	501	506	504
Portugal	PRT	501	498	492
Romenia	ROU	435	434	444
Russia	RUS	487	495	494
Singapura	SGP	556	535	564
Republica Eslovaca	SVK	461	453	475
Eslovenia	SVN	513	505	510
Suecia	SWE	493	500	494
Tailandia	THA	421	409	415
Trinidad e Tobago	TTO	425	427	417
Tunisia	TUN	386	361	367
Turquia	TUR	425	428	420
Uruguai	URY	435	437	418
Estados Unidos	USA	496	497	470
Vietna	VNM	525	487	495
Qatar	QAT	418	402	390
B-S-J-G (China)	QCH	518	494	531
Kosovo	KSV	378	347	362
Cyprus	QES	433	443	437
Chinese Taipei	TAP	532	497	542

CABA (Argentina)	QAR	475	475	456
FYROM	MKD	384	352	371

Tabela 7 – Países avaliados pelo PISA

5.2 QUESTIONÁRIO - ESCOLA

5.2.1 TAMANHO MÉDIO DAS TURMAS

Questão (SC003Q01TA) Qual é o tamanho médio das classes na sua escola?

A Tabela 8 apresenta os dados da média de alunos por turma, o desvio, mínimo e máximo. Os valores são codificados como descritos na tabela.

País	Média	Desvio	Mínimo	Máximo
ALB	3.53	1.5	1	7
DZA	4.18	1.68	1	7
AUS	3.36	0.81	1	6
AUT	3.0	1.49	1	9
BEL	2.24	0.98	1	9
BRA	5.41	1.79	1	9
BGR	3.31	1.09	1	9
CAN	3.31	1.19	1	9
CHL	4.64	1.7	1	9
TAP	5.51	1.39	2	9
COL	5.49	2.4	1	9
CRI	3.9	1.37	1	9
HRV	3.32	0.7	1	5
CZE	3.1	1.04	1	5
DNK	2.69	0.83	1	8
DOM	5.47	2.21	1	9
EST	3.08	2.37	1	9
FIN	2.14	0.63	1	4
FRA	4.17	1.16	1	9
GEO	5.36	3.2	1	9
DEU	3.29	1.05	1	9
GRC	3.01	1.22	1	9
HKG	4.54	0.95	1	7
HUN	3.73	1.73	1	9
ISL	2.0	1.09	1	7
IDN	4.47	1.85	1	9

IRL	3.24	0.83	1	4
ISR	4.33	1.23	1	8
ITA	2.84	1.3	1	9
JPN	5.57	1.19	1	9
JOR	4.86	2.02	1	9
KOR	4.57	0.95	2	6
KSV	3.54	1.22	1	6
LBN	4.29	2.43	1	9
LVA	2.39	1.4	1	9
LTU	2.66	1.26	1	5
LUX	2.64	0.53	2	4
MAC	4.62	2.01	1	9
MLT	2.33	0.74	1	3
MEX	5.71	2.42	1	9
MDA	2.98	1.76	1	9
MNE	3.41	1.39	1	8
NLD	3.52	1.08	1	9
NZL	3.28	0.83	1	5
NOR	3.1	0.95	1	5
PER	3.51	1.85	1	9
POL	3.23	1.41	1	9
PRT	3.2	0.97	1	5
QAT	3.83	1.95	1	9
ROU	3.62	1.45	1	9
RUS	2.71	1.08	1	5
SGP	5.2	1.27	1	9
SVK	2.57	0.98	1	5
VNM	6.17	1.96	1	9
SVN	3.02	1.0	1	5
ESP	3.72	1.49	1	9
CHE	2.5	1.61	1	9
THA	5.05	1.96	1	9
TTO	4.12	1.49	1	9
ARE	3.99	2.03	1	9
TUN	3.89	1.26	1	9
TUR	7.42	2.45	1	9
MKD	3.4	1.77	1	9
GBR	3.24	0.82	1	9
USA	3.5	1.23	1	6

URY	3.63	1.53	1	9
QCH	6.74	1.88	1	9
QES	3.69	1.8	1	9
QAR	6.09	2.79	1	9

Tabela 8 – Dados de Tamanho de Turma por País

O gráfico abaixo apresenta os países e suas respectivas médias em relação a resposta a respeito do tamanho das turmas em ordem crescente.

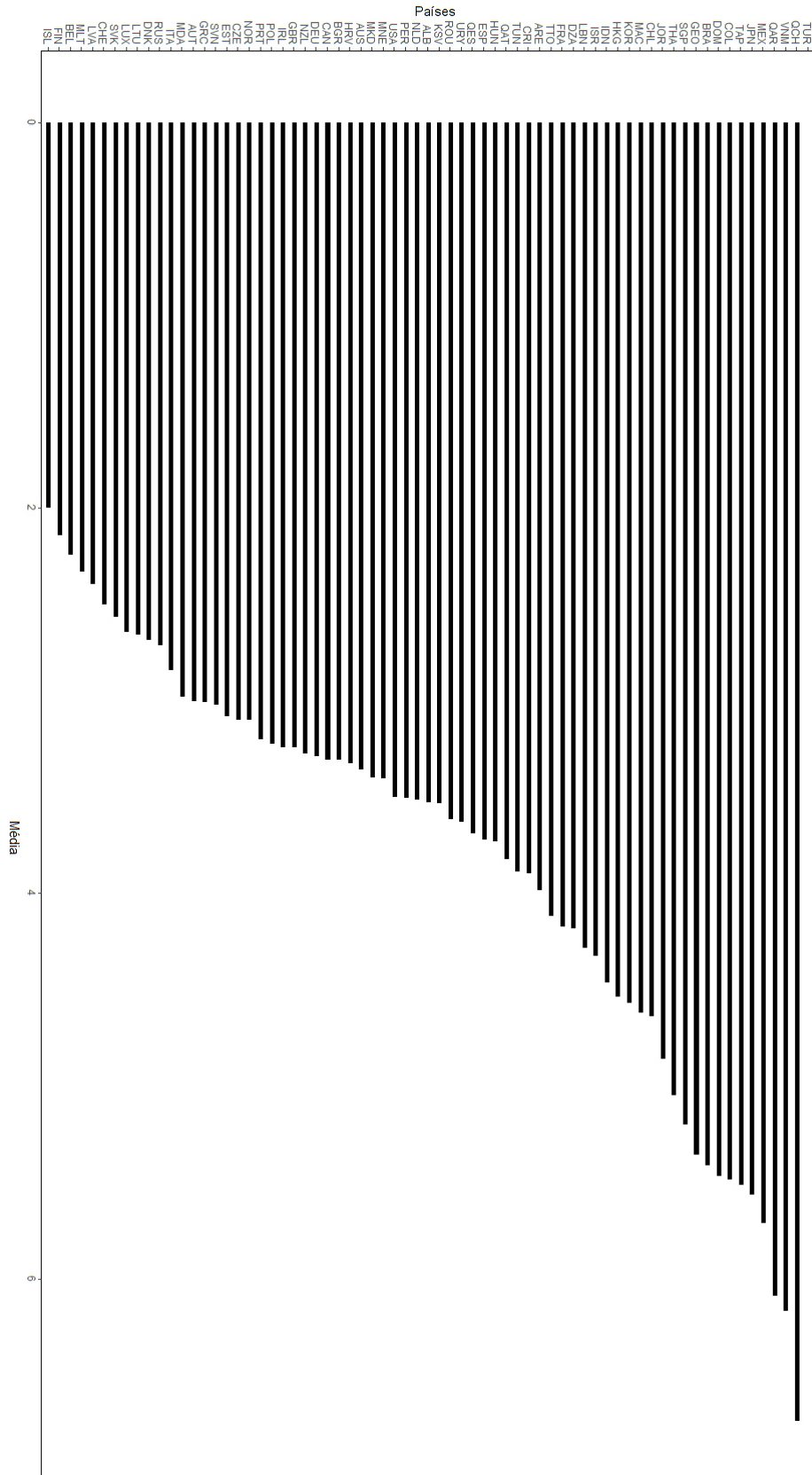


Figura 22 – Gráfico com o tamanho médio das turmas em cada país.

Por apresentar um grande número de países no gráfico, a visualização não é muito agradável. Por este motivo, foram selecionado os dez países que apresentaram as maiores e as menores médias para ter uma melhor visualização dos dados.

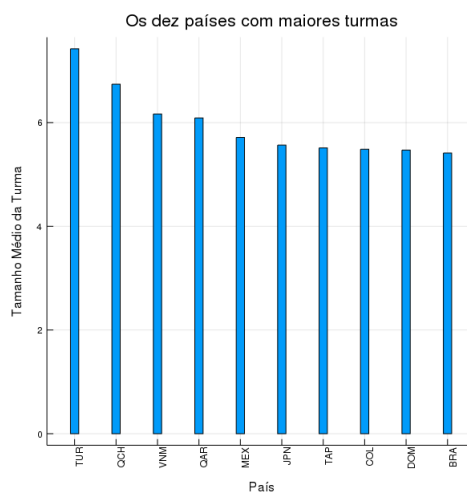


Figura 23 – Países com as maiores médias

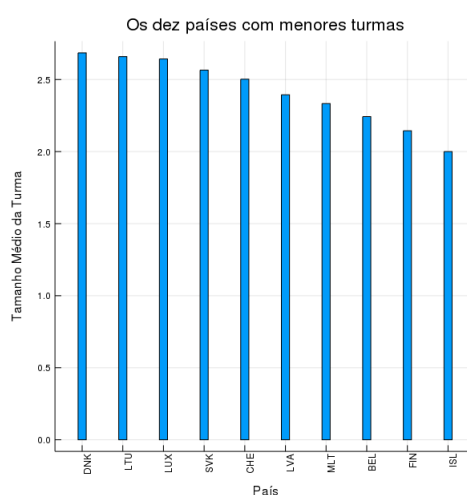


Figura 24 – Países com as menores médias

Comparando os gráficos, observamos a Turquia como o país com a maior média em relação ao tamanho de turma e a Islândia com a menor média. A tabela abaixo mostra ambos países e seus respectivos resultados no programa.

País	Ciência	Leitura	Matemático
Islândia	473	428	488
Turquia	425	428	410

Tabela 9 – Notas de ISL e TUR

A seguir os gráficos apresentam uma comparação entre o tamanho médio das turmas e as notas das três áreas: Ciência, Leitura e Matemática.

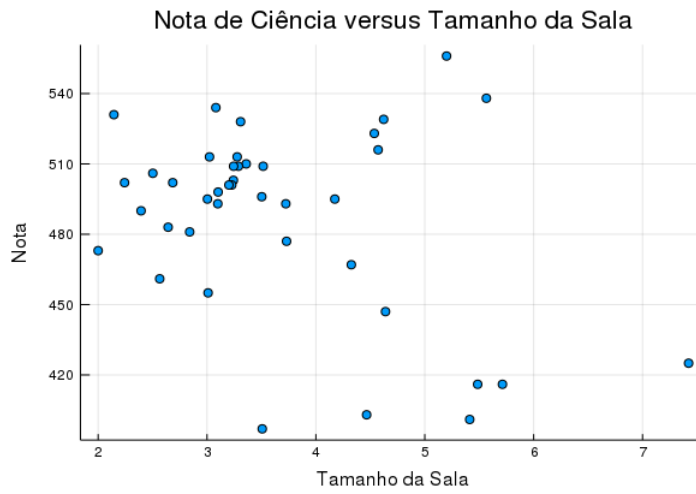


Figura 25 – Nota de Ciências versus Tamanho da Turma

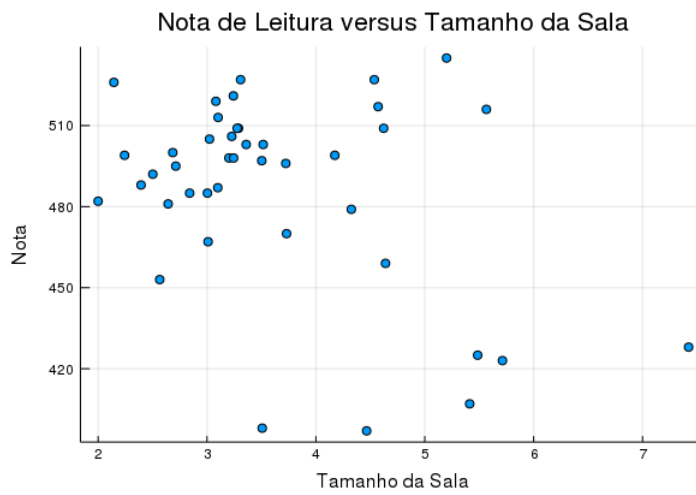


Figura 26 – Nota de Leitura versus Tamanho da Turma

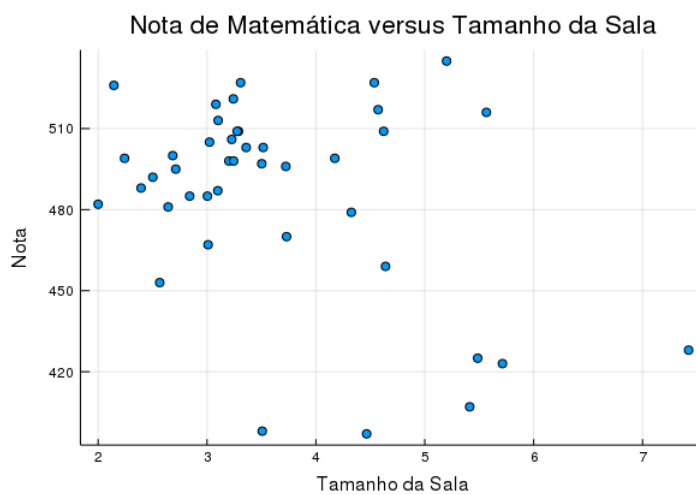


Figura 27 – Nota de Matemática versus Tamanho da Turma

Ao analisar os três gráficos é possível observar um padrão que se repete nas três áreas avaliadas, esse padrão ocorre entre o intervalo 2 e 4. Neste intervalo se encontra uma concentração de notas dos países, mostrando assim que a maioria dos países que realizaram o programa no ano de 2015 possuíam turmas formadas de 15 a 30 alunos.

5.2.2 TIPOS DE ADMINISTRAÇÃO

Questão (SC014Q01NA) - Que tipo de organização administra sua escola?

As escolas participantes do PISA possuem administrações distintas, podendo ser: administradas por organizações religiosas, organizações sem fins lucrativos e com fins lucrativos. A maneira como são administradas podem interferir diretamente no resultado do país no programa. Na questão escolhida para análise temos um número de 4.218 escolas que responderam o questionário, ficando 13.690 sem responder.

Administração Religiosa

O gráfico abaixo representa a quantidade de escolas que responderam que possuem uma administração religiosa. Apenas 1215 escolas responderam.

A seguir uma comparação entre os dez países que apresentaram as maiores e as menores média em relação ao tamanho das turmas e sua quantidade de escolas administradas por instituições religiosas.

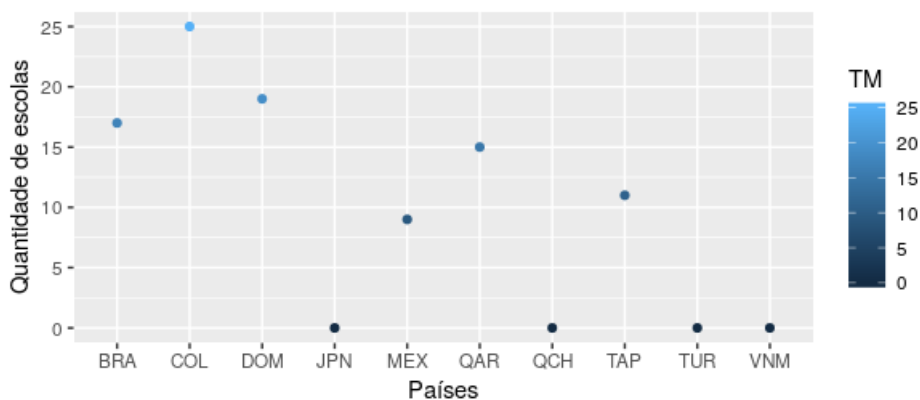


Figura 29 – Tamanho médio das turmas (maiores médias) + Adm. Religiosa

No gráfico, os três países que aparecem com o maior número de escolas sendo administradas por religiosa são: Brasil, Colômbia e Republica Dominicana.

País	Ciência	Leitura	Matemática
Colômbia	416	425	390
Republica Dominicana	332	358	328
Brasil	401	407	377

Tabela 10 – Notas de (COL, DOM, BRA)

Países com as menores médias e sua quantidade de administração religiosa.

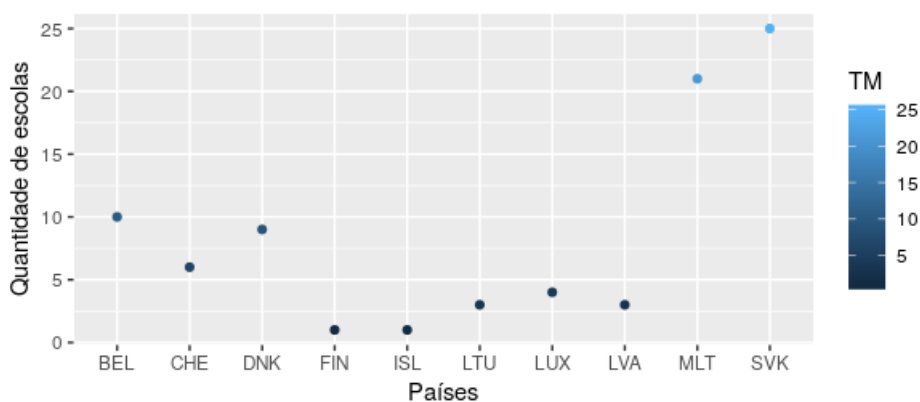


Figura 30 – Tamanho médio das turmas (menores médias) + Adm. Religiosa

No gráfico os três países com o menor número de escolas sendo administradas por religiosa são: Finlândia, Islândia e Lituânia.

País	Ciência	Leitura	Matemática
Finlândia	531	526	511
Islândia	473	482	488
Lituânia	475	472	478

Tabela 11 – Notas de (FIN, ISL, LTU)

Com auxílio das tabelas. Os países selecionados com a menor presença de administração religiosa, apresentaram notas superiores em todas as áreas em relação aos países com maior número de administração religiosa.

A baixo os gráficos apresentam os dez países que apresentaram o maior e o menor número de escolas com administração religiosa (*sem levar em consideração as escolas que responderam 0*) e seus resultados nas três áreas.

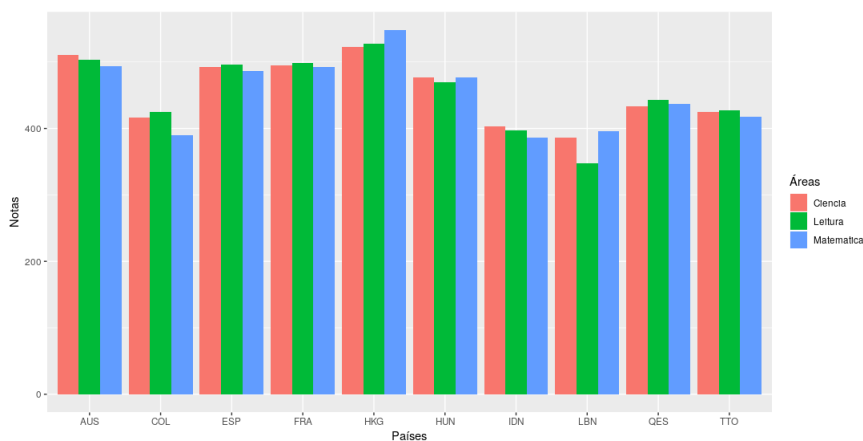


Figura 31 – Países com maior N^a Adm. Religiosa + Notas

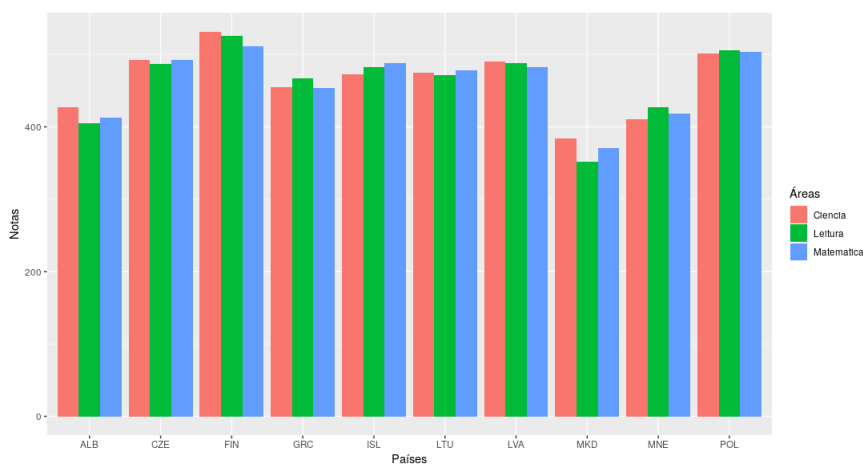


Figura 32 – Países com menor N^a Adm. Religiosa + Notas

Abaixo os dez países com as maiores média em relação ao tamanho das turmas e sua quantidade de escolas administradas por instituições sem fins lucrativos temos:

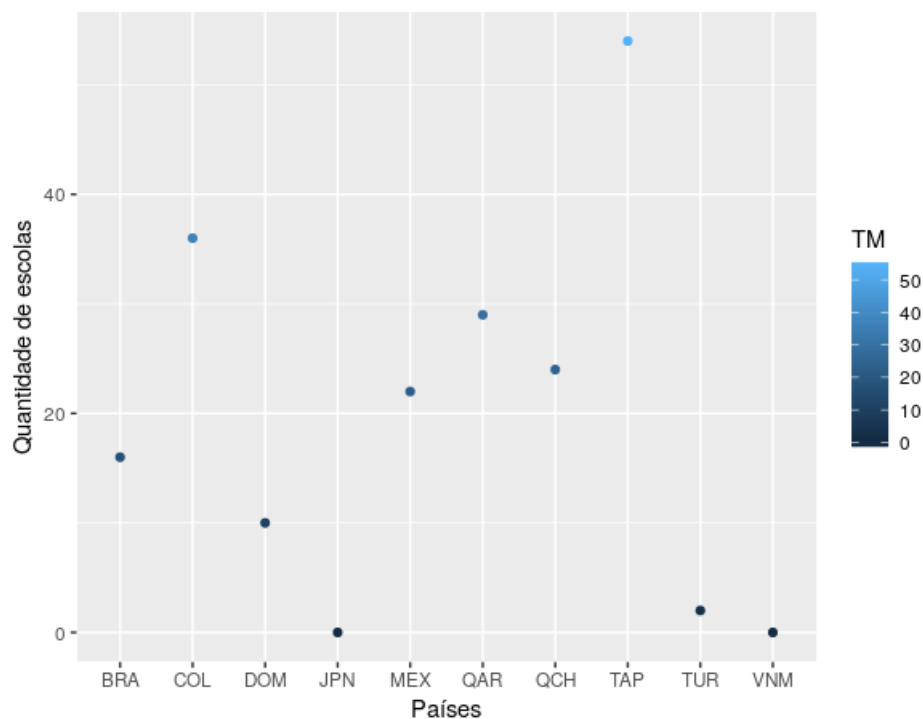


Figura 34 – Tamanho médio das turmas (maiores médias) + Adm. Sem Fins Lucrativos

Os três países que apresentaram o maior número foram: Colômbia, Chinese Taipei e CABA (Argentina).

País	Ciência	Leitura	Matemática
Colômbia	416	425	390
Chinese Taipei	532	497	542
CABA (Argentina)	475	475	456

Tabela 12 – Notas de (COL, TAP, QAR)

Países com as menores médias.

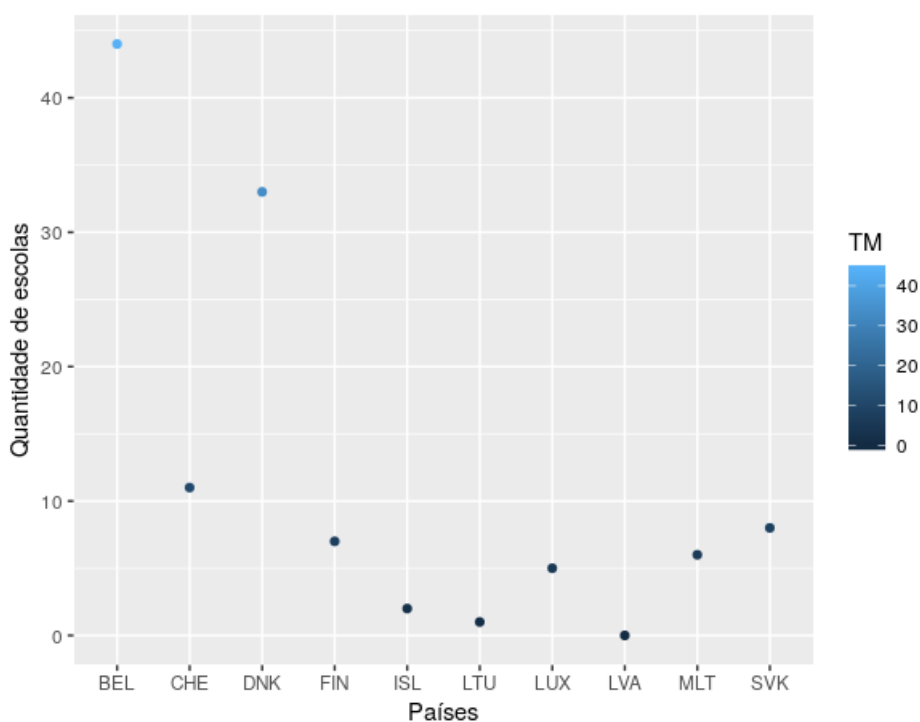


Figura 35 – Tamanho médio das turmas (menores médias) + Adm. Sem Fins Lucrativos

Os três países que apresentaram o menor número foram: Letônia, Islândia e Lituânia.

País	Ciência	Leitura	Matemática
Letônia	490	488	482
Islândia	473	482	488
Lituânia	475	472	478

Tabela 13 – Notas de (LVA, ISL, LTU)

Com auxílio das tabelas. Os países selecionados com a menor presença de administração sem fins lucrativo, apresentaram notas semelhantes aos países com maior número de administração religiosa.

A baixo os gráficos apresentam os dez países que apresentaram o maior e o menor número de escolas com administração sem fins lucrativos (*sem levar em consideração as escolas que responderam 0*) e seus resultados nas três áreas.

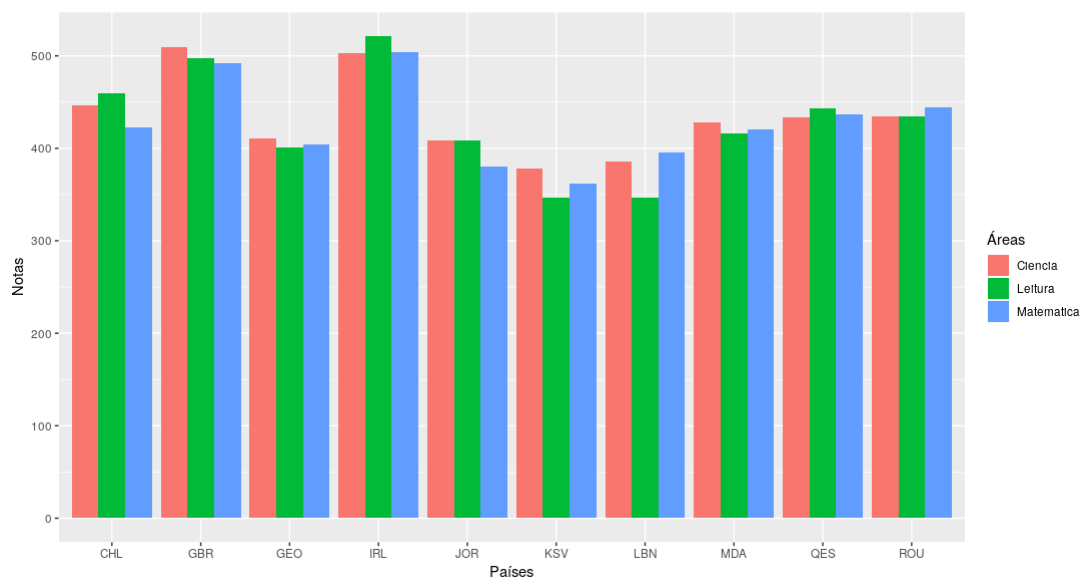


Figura 36 – Países com maior N^a Adm. Sem Fins Lucrativos + Notas

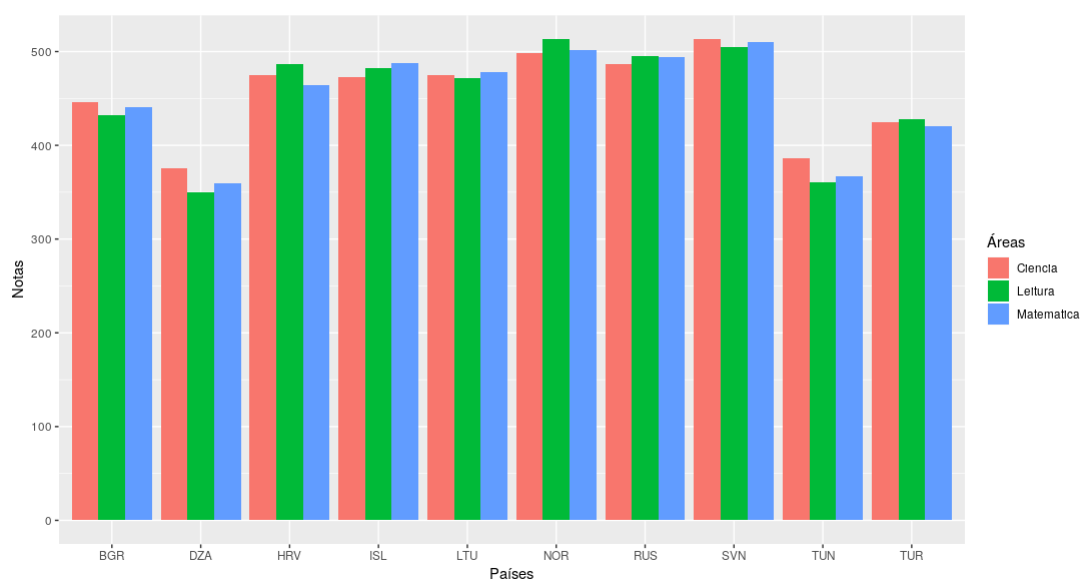


Figura 37 – Países com menor N^a Adm. Sem Fins Lucrativos + Notas

Administração Com Fins Lucrativos

O gráfico abaixo apresenta a quantidade de escolas com administração com fins lucrativos. Apenas 778 escolas responderam o questionário.

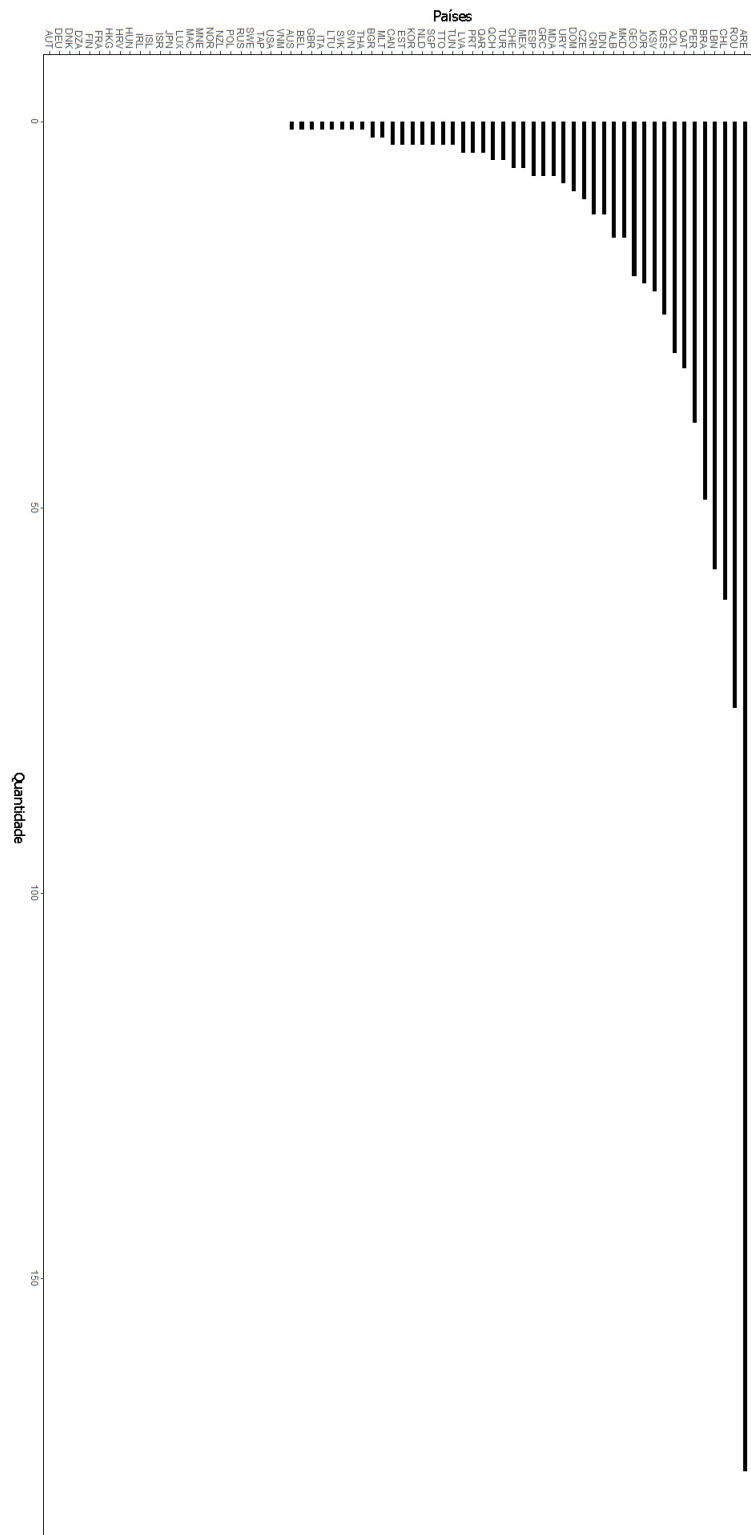


Figura 38 – Gráfico de escolas administradas por organizações com fins lucrativos.

O gráfico abaixo apresenta a comparação entres os dez países com as maiores média em relação ao tamanho das turmas e suas quantidade de escolas administradas por instituições com fins lucrativos temos:

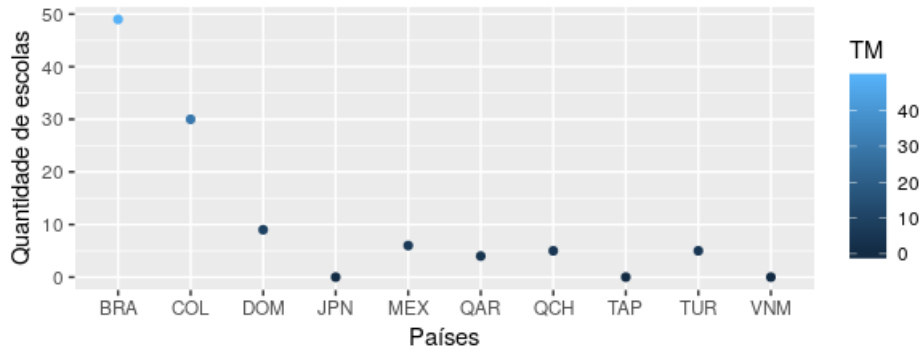


Figura 39 – Tamanho médio das turmas (maiores médias) + Adm. Com Fins Lucrativos

Os três países que apresentaram o menor número foram: Brasil, Colômbia e Republica Dominicana.

País	Ciência	Leitura	Matemática
Brasil	401	407	377
Colômbia	416	425	390
Republica Dominicana	332	358	328

Tabela 14 – Notas de (BRA, COL, DOM)

Países com as menores médias.

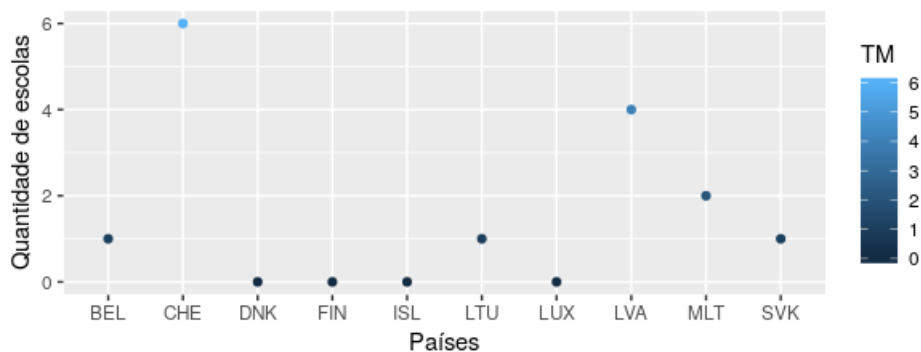


Figura 40 – Tamanho médio das turmas (menores médias) + Adm. Com Fins Lucrativos

Os três países que apresentaram o menor número foram: Finlândia, Islândia e Dinamarca.

País	Ciência	Leitura	Matemática
Finlândia	531	526	511
Islândia	473	482	488
Dinamarca	502	500	511

Tabela 15 – Notas de (FIN, ISL, DNK)

A seguir os gráficos apresentam os dez países que apresentaram o maior e o menor número de escolas com administração com fins lucrativos (*sem levar em consideração as escolas que responderam 0*) e seus resultados nas três áreas.

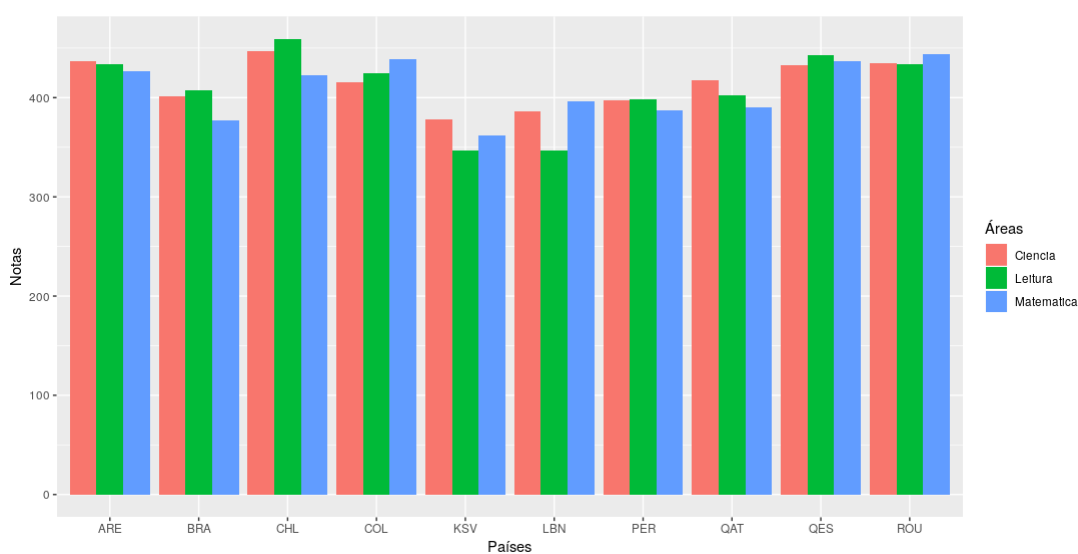


Figura 41 – Países com maior N^a Adm. Com Fins Lucrativos + Notas

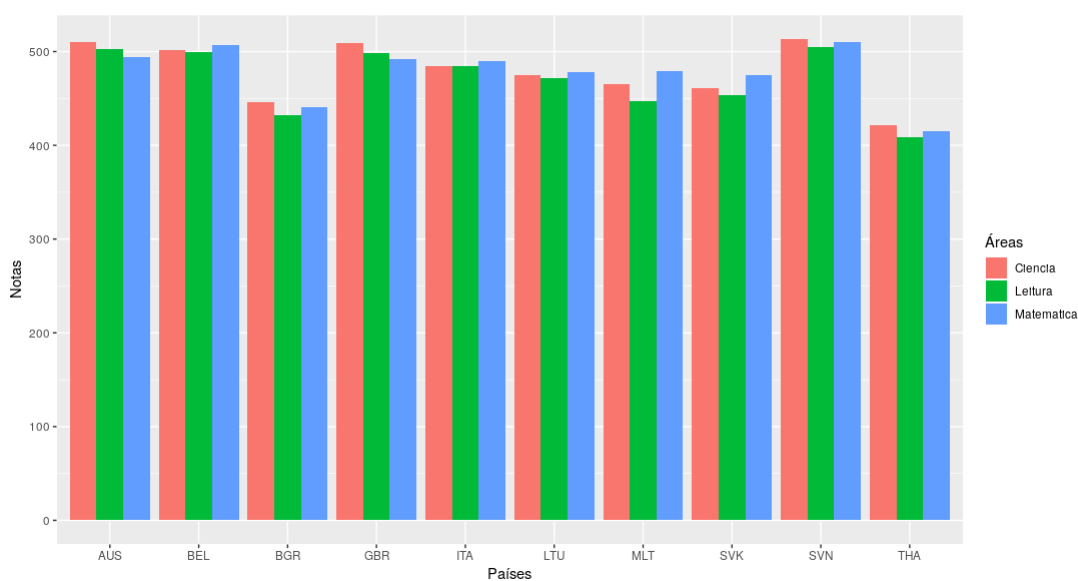


Figura 42 – Países com menor N^a Adm. Com Fins Lucrativos + Notas

5.3 QUESTIONÁRIO - PROFESSOR

5.3.1 STATUS PROFESSOR

Questão (TC004Q01NA) Qual é o seu status de emprego como professor nessa escola?

O questionário foi aplicado opcionalmente para alguns países no ano de 2015 e contou com 108.292 professores respondendo. A opção escolhida para análise contou com 62.977 respostas positivas. No gráfico podemos ver a quantidade de professores com status de efetivo por país.

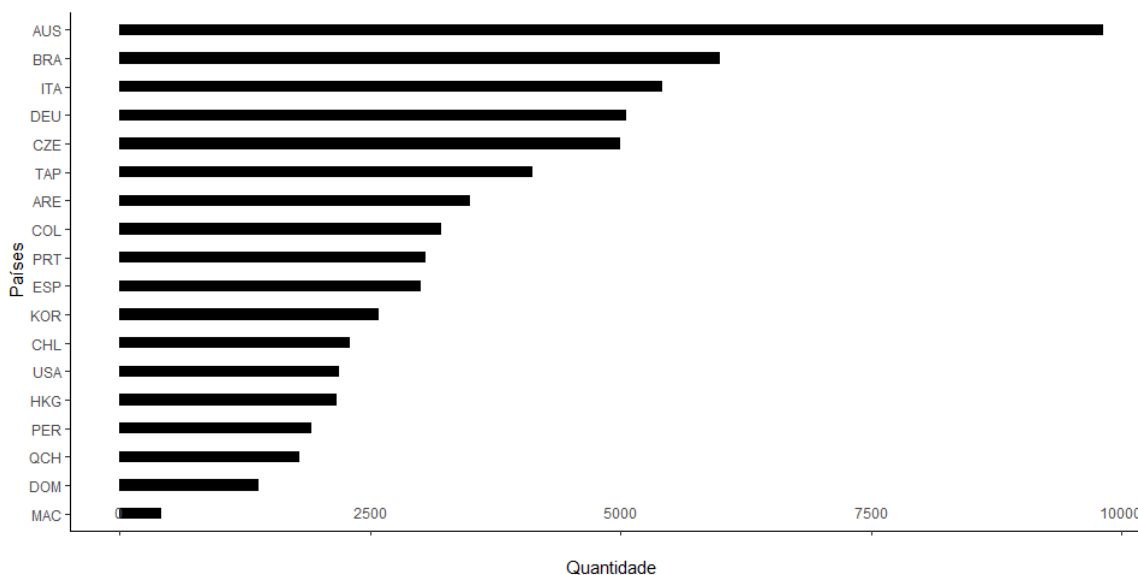


Figura 43 – Gráfico professores com status de efetivo.

Pelo gráfico, a Austrália aparece como o país que apresenta o maior número de professores com status de *Emprego Permanente*, também é um país que possui ótimos resultados, com notas acima da média: Ciência(510), Leitura(503) e Matemática(494).

Representando o menor número de professores com status de *Emprego Permanente*. Macao (China) uma região autônoma na costa sul da China, foi o país que apresentou a menor quantidade de professores permanentes. Seus resultados foram acima da média: Ciência(529), Leitura(509) e Matemática(544).

A seguir os sete países com o maior e o menor número de professores efetivos e suas respectivas notas nas três áreas de conhecimento.

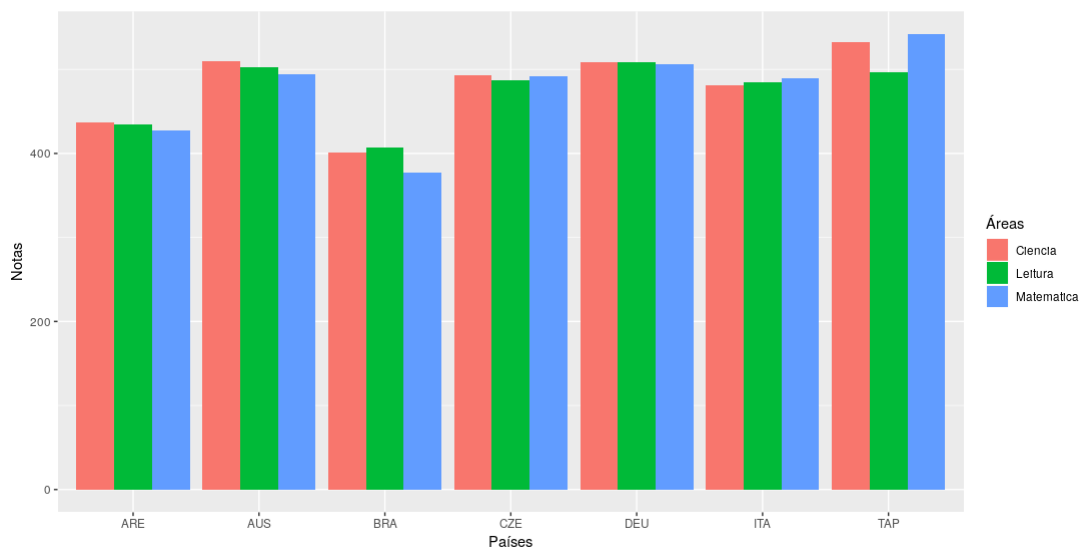


Figura 44 – Países com maior N^a de professores efetivos + notas

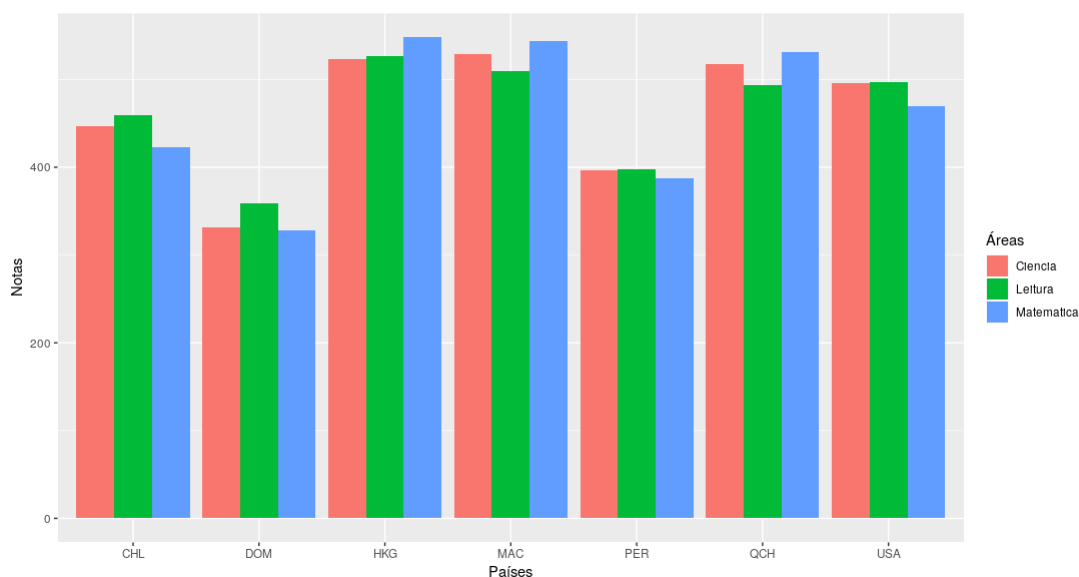


Figura 45 – Países com menor N^a de professores efetivos + notas

5.3.2 NÍVEL DE FORMAÇÃO

Questão (TC012Q01NA) Qual é o nível mais alto de educação formal que você completou?

Para esta questão, foi analisada as respostas de nível 5 A , o nível educacional terciário, ou seja, profissionais que tenham no mínimo graduação.

O gráfico mostra a quantidade de professores do nível A5 por país.

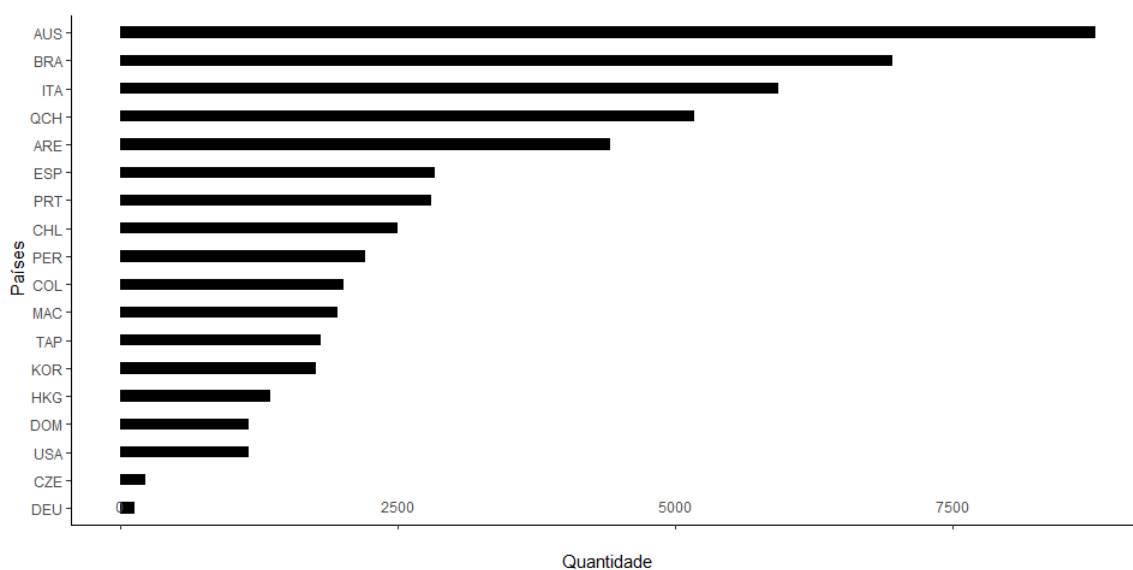


Figura 46 – Gráfico de professores com nível A5.

Vemos novamente a Austrália em destaque como o país com maior número de professores com o nível de formação 5A, e em comparação com seus resultados, o país se destacou com notas acima da média: Ciência(510), Leitura(503) e Matemática(494).

Por outro lado a República Checa, foi o país apresentou um número pequeno de professores com nível de formação A5, em relação aos seus resultados, o país obteve média apenas na área da ciência, nas outras suas notas foram abaixo da média: Ciência(493), Leitura(487) e Matemática(492).

Os gráficos representam os nove países com maior número de professores que responderam que são do nível A5. Os gráficos foram separados pelas áreas de conhecimento e para uma melhor análise colocamos uma linha vermelha que separa as notas acima de 500 em cada área.

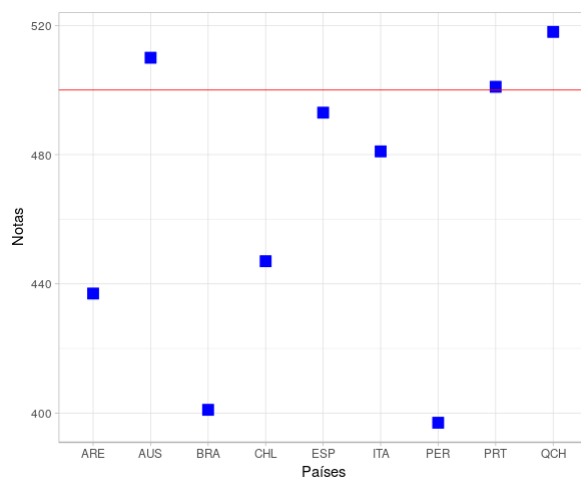


Gráfico com os nove países com maior número de professores nível A5

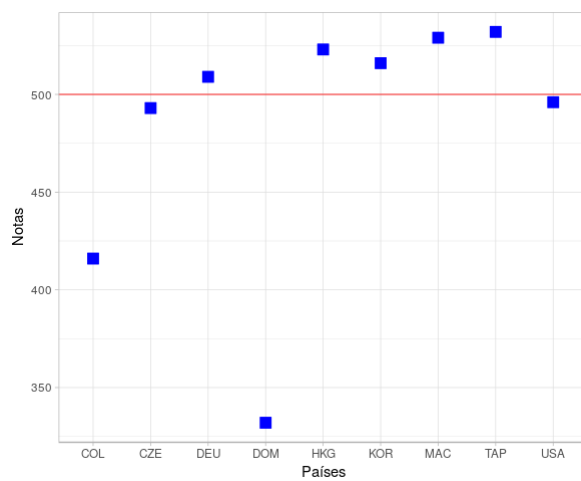


Gráfico com os nove países com menor número de professores nível A5

Figura 47 – Nível A5 + Ciência

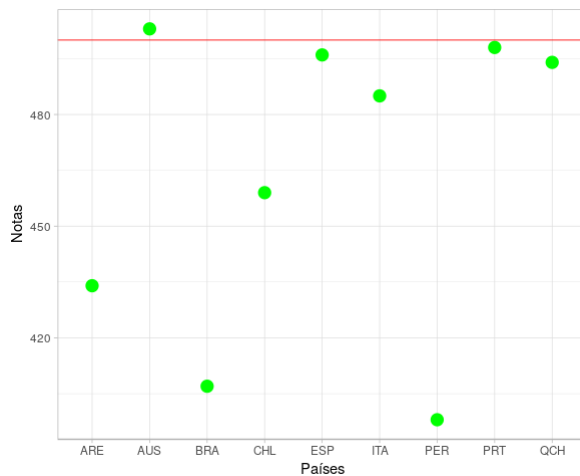


Gráfico com os nove países com maior número de professores nível A5

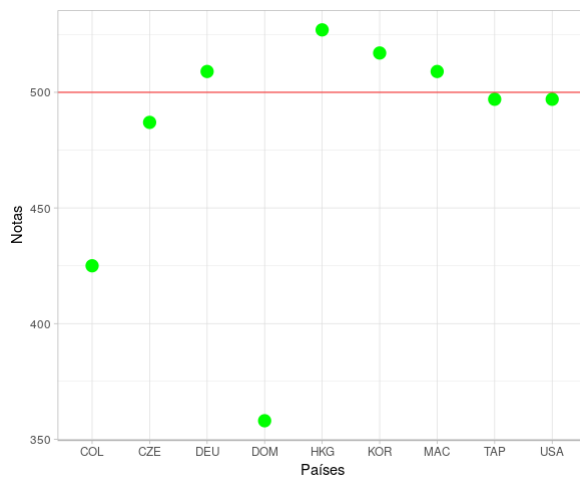


Gráfico com os nove países com menor número de professores nível A5

Figura 48 – Nível A5 + Leitura

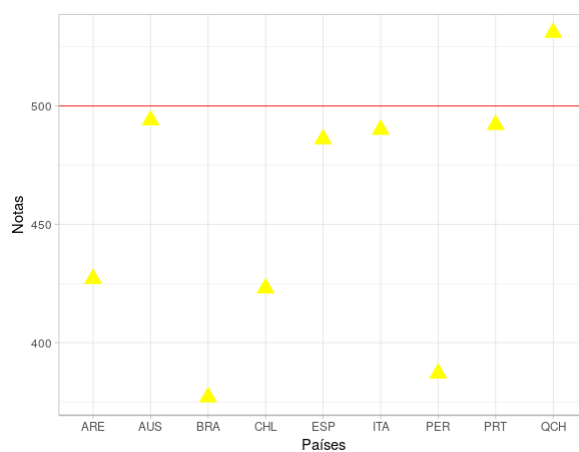


Gráfico com os nove países com maior número de professores nível A5

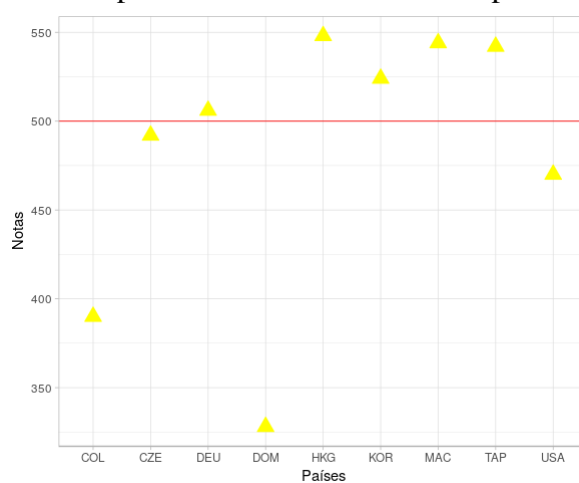


Gráfico com os nove países com menor número de professores nível A5

Figura 49 – Nível A5 + Matemática

Observando os gráficos é possível notar que os países com o menor número de professores que responderam que são do nível A5 tiveram um maior número de notas acima de 500. Esse padrão se repete nas três áreas.

6 CONCLUSÃO

Neste capítulo é apresentada uma síntese das conclusões obtidas no desenvolvimento do trabalho juntamente a sugestões para os trabalhos futuros.

6.1 CONSIDERAÇÕES FINAIS

Este trabalho abordou a importância dos meios que avaliam a educação em todo mundo, esses meios foram: programas internacionais e nacionais. Ao decorrer do trabalho observamos também que uma educação de qualidade não se aplica apenas ao ambiente escolar, mas sim em fatores que formam o sistema educacional: escola, comunidade, professores, alunos. Mas o proposto objetivo do trabalho foi realizado, e com suporte da linguagem R foi possível realizar a análise dos resultados dos questionários aplicados no PISA no ano de 2015.

Por meio das representações gráficas foi possível entender melhor os resultados, além de encontrar padrões que se repetiam em países com notas altas dentro do programa.

Foi possível mostrar a quantidade de dados brutos que o programa coleta, e como elas podem ser utilizadas de diversas formas para que se entenda a atual situação dos países. Assim podendo ser utilizada principalmente pelos governos dos países para criação de novas políticas educacionais e socioeconômicas. Como é o real objetivo da organização que criou o programa.

6.2 TRABALHOS FUTUROS

A partir desse trabalho, abre-se uma vasta quantidade de opções de análises a serem feitas como: comparações entre os países da América do sul, ou da Europa, ou simplesmente analisar outras questões disponibilizadas pelos questionários.

Também deixa em aberto a opção de atualizar os dados, pois resultados da última aplicação referente ao ano de 2018 ficarão disponíveis no segundo semestre do ano de 2019, para assim comparar quais as alterações que ocorreram após três anos.

REFERÊNCIAS

- AQUINO, J. A. de. *R para cientistas sociais*. Ilhéus, Bahia: EDITUS - EDITORA DA UESC, 2014. ISBN ISBN - 978-85-7455-369-6. Citado na página 27.
- BORGES, L. E. *Python para Desenvolvedores*. 2ª edição - edição do autor. ed. Rio de Janeiro, Brasil: [s.n.], 2010. ISBN ISBN - 978-85-909451-1-6. Citado na página 28.
- COSTA, H. V. V. da. *Introdução ao R*. [S.l.], Dezembro, 2017. Citado na página 32.
- DIONISIO, E. J. *Trabalhando com a Linguagem R*. 2015. Disponível em: <<https://www.devmedia.com.br/trabalhando-com-a-linguagem-r/33275>>. Citado na página 27.
- FONTANA Éliton. Breve introdução à programação em scilab 6.0. p. 1–16, Universidade Federal do Paraná - UFPR - Departamento de Engenharia Química 2018. Citado na página 30.
- GADOTTI, M. Qualidade na educação : Uma nova abordagem. In: *COEB 2013 - Congresso de Educação Básica : Qualidade na aprendizagem*. Florianópolis, Brasil: [s.n.], 2013. p. 1–18. Citado na página 13.
- KOUATCHOU, J. Basic Comparison of Python, Julia, R, Matlab and IDL. NASA - National Aeronautics and Space Administration Dezembro, 2016. Citado 2 vezes nas páginas 31 e 32.
- MANSSOUR, I. H. *Paradigmas de Linguagem*. 2018. Disponível em: <<https://www.inf.pucrs.br/~gustavo/disciplinas/pli/material/>>. Citado na página 28.
- MARQUES, N. C.; MORGADO, C. Octave: Guia de estudo. p. 1–46, Novembro 2010. Citado na página 29.
- MAYER, F. *Gerenciamento de pacotes no R*. 2015. Disponível em: <https://github.com/fernandomayer/R-pacotes-rautu/blob/master/R_gerenciamento_de_pacotes.md>. Citado na página 36.
- MEC. *Ministério da Educacao - Avaliações da aprendizagem*. 2009. Disponível em: <<http://portal.mec.gov.br/pdde/190-secretarias-112877938/setec-1749372213/18843-avaliacoes-da-aprendizagem>>. Citado na página 14.
- MENEZES, N. N. C. *Introdução à Programação com Python - Algoritmos e lógica de programação para iniciantes*. 2ª edição. ed. São Paulo, SP – Brasil: Novatec Editora Ltda, 2014. ISBN ISBN - 978-85-7522-408-3. Citado na página 28.
- OECD. *Programme for International Student Assessment - Results from PISA 2015 - Resumo de resultados nacionais PISA 2015 - Brasil*. [S.l.], 2016. Citado na página 26.
- OLIVEIRA, P. F. de; GUERRA, S.; MCDONNELL, R. *Ciência de Dados com R - Introdução*. Brasília – Brasil: IBPAD – Instituto Brasileiro de Pesquisa e Análise Dados, 2018. 240 p. ISBN ISBN - 978-85-54230-00-5. Citado 4 vezes nas páginas 13, 27, 32 e 33.
- PEREIRA, J. M. *Introdução à linguagem de programação Julia: A ambiciosa linguagem de programação que quer substituir Python, R e Matlab*. Piauí – Brasil: [s.n.], 2017. 315-335 p. ISBN ISBN - 978-85-7669-395-6. Citado na página 31.

PIRES, P. S. da M. Introdução ao scilab versão 3.0. p. 1–120, Departamento de Engenharia de Computação e Automação - Universidade Federal do Rio Grande do Norte - Natal Julho, 2004. Citado na página 30.

PISA. *PISA 2000 - Relatório Nacional*. [S.l.], 2001. Citado 2 vezes nas páginas 13 e 17.

PISA. *Brasil no PISA 2015 - Análise e reflexão sobre o desempenho dos estudantes brasileiros*. [S.l.], 2016. Citado 4 vezes nas páginas 17, 18, 20 e 22.

SILVA, E. B.; FARIA, B. L.; FERNANDES, B. C. C. A aplicação da ciência de dados como ferramenta de controle social: Análise da Publicização dos Gastos com as Universidades Públicas. In: *CONED 2018 - II Congresso Nacional em Educação*. Diamantina, Brasil: [s.n.], 2018. Citado na página 41.

SIQUEIRA, J. K. Guia Matlab - Para alunos de Engenharia Elétrica, Computação e Automação. p. 1–64, Junho 2007. Citado na página 29.

SOUZA, F. L. de. Introdução ao Octave. p. 1–31, Junho 2003. Citado na página 29.

TIMSSPIRLS. *TIMSS: Trends In International Mathematics And Science Study and PIRLS: Progress In International Reading Literacy Study*. 2019. Disponível em: <<https://timssandpirls.bc.edu/>>. Citado na página 14.