

**APS – Ciência da Computação 2º Semestre**

**Relatório - “Análise das escolas na mesorregião de Ribeirão Preto”**

**Professora:** Camila Prado

**Pandas**



**Aluno:**

João Lucas Roldão Rocha – RA: N9614H4 – Turma: CC2P18

Email: [joaolucasrr0210@gmail.com](mailto:joaolucasrr0210@gmail.com)

# Índice

1 – Introdução

2 – Tema escolhido

2.1 – O que é o Censo Escolar

2.2 – Qual é a finalidade do Censo Escolar

2.3 – Sobre o tema

2.4 – Definição de “Mesorregião”

3 – Cinco tópicos principais a serem analisados

3.1 – Quantidade total de escolas que realizaram a pesquisa

3.2 – Escolas que não realizam o tratamento do lixo

3.3 - Escolas que possuem um laboratório de informática;

3.4 - Escolas que possuem um refeitório

3.5 - Escolas que não possuem nenhuma forma de acessibilidade a pessoas com necessidades especiais

3.6 - Escolas que possuem banheiros acessíveis a pessoas com necessidades especiais

4 – Script

4.1 – Script: Estrutura do código

4.1.1 – Script: Estrutura do código – Biblioteca *Pandas*

4.1.2 – Script: Estrutura do código – Biblioteca *Matplotlib*

4.1.3 – Script: Estrutura do código – Biblioteca *Openpyxl*

4.2 – Script: Relatório com as linhas do código

4.3 – Script em funcionamento

5 – Bibliografia

6 – Ficha de atividades práticas supervisionadas

## **1 – Introdução**

Nosso trabalho possui como objetivo realizar um script utilizando a biblioteca “NumPy” e/ou “Pandas” na linguagem de programação “Python”, para que seja feita uma análise de uma das planilhas disponíveis no site proposto pelos professores (<https://dados.gov.br/group/dados-em-destaque>), e através desta, fundamentarmos a escolha, discutindo de forma clara qual a importância e relevância dos dados.

## **2 – Tema escolhido**

Através da análise das planilhas propostas pelos professores, decidi usar como tema principal do meu trabalho: “Análise das escolas na mesorregião de Ribeirão Preto”.

Antes de explicar o tema, precisamos definir exatamente o que é o “censo” realizado pelo INEP (*Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira*).

### **2.1 – O que é o Censo Escolar**

O Censo Escolar é o principal instrumento de coleta de informações da educação básica e a mais importante pesquisa estatística educacional brasileira. É coordenado pelo Inep e realizado em regime de colaboração entre as secretarias estaduais e municipais de educação e com a participação de todas as escolas públicas e privadas do país. A pesquisa estatística abrange as diferentes etapas e modalidades da educação básica e profissional:

- Ensino regular (educação infantil, ensino fundamental e médio);
- Educação especial – escolas e classes especiais;
- Educação de Jovens e Adultos (EJA);

- Educação profissional (cursos técnicos e cursos de formação inicial continuada ou qualificação profissional).

A coleta de dados das escolas tem caráter declaratório e é dividida em duas etapas. A primeira etapa consiste no preenchimento da Matrícula Inicial, quando ocorre a coleta de informações sobre os estabelecimentos de ensino, gestores, turmas, alunos e profissionais escolares em sala de aula. A segunda etapa ocorre com o preenchimento de informações sobre a Situação do Aluno, e considera os dados sobre o movimento e rendimento escolar dos alunos, ao final do ano letivo.

O Censo Escolar é regulamentado por instrumentos normativos que instituem a obrigatoriedade, os prazos, os responsáveis e suas responsabilidades, bem como os procedimentos para realização de todo o processo de coleta de dados.

## **2.2 – Qual é a finalidade do Censo Escolar**

O Censo Escolar é uma ferramenta fundamental para que os atores educacionais possam compreender a situação educacional do país, das unidades federativas, dos municípios e do Distrito Federal, bem como das escolas e, com isso, acompanhar a efetividade das políticas públicas.

A compreensão da situação educacional ocorre por intermédio de um conjunto amplo de indicadores que possibilitam monitorar o desenvolvimento da educação brasileira, como o Índice de Desenvolvimento da Educação Básica (*Ideb*), as taxas de rendimento e de fluxo escolar, a distorção idade-série, entre outros, que servem de referência para as metas do Plano Nacional da Educação (*PNE*), que podem ser acompanhadas no Observatório do *PNE*. Todos esses indicadores são calculados com base nos dados do Censo Escolar.

Além disso, as matrículas e os dados escolares coletados servem de base para o repasse de recursos do governo federal e para o planejamento e divulgação de dados das avaliações realizadas pelo Inep. Todos esses Indicadores Educacionais são calculados com base nos dados do Censo Escolar.

## **2.3 – Sobre o tema**

Em um dos conjuntos de dados disponibilizados no site, a sexta opção é chamada de “Inep: Microdados do Censo Escolar”, e ao acessarmos a página em questão, somos apresentados aos dados entre os anos de 1995 e 2015. Porém, ao solicitar o download das informações no site, somos redirecionados a uma outra página(<https://www.gov.br/inep/pt-br/acesso-a-informacao/dados-abertos/microdados/censo-escolar>), dessa vez em um outro site do governo, onde podemos acessar todos os dados dos censos escolares realizados pelo INEP. Dessa vez, as opções contam com dados entre os anos de 1996 e 2021. Por se tratar do ano com dados disponibilizados mais recente, decidi que seria uma boa opção escolher o ano de 2021 para realizar a análise. Ao clicar para acessar os dados, o site nos disponibiliza o download de um arquivo compactado (.zip), dentro dele temos acesso à uma pasta “2021”, dentro dela possuímos mais três pastas, sendo elas: “leia-me”, que possui um arquivo .pdf com o mesmo nome dentro, com a apresentação sobre a pesquisa e o crédito pelos organizadores; “Dados”, que possui um arquivo .csv chamado “microdados\_ed\_basica\_2021”, com os dados do censo de 2021 realizado em todas as escolas brasileiras (dados esses que serão usados na construção do nosso script); e a última chamada “Anexos”, com outras duas pastas dentro, sendo elas “ANEXO I - Dicionário de Dados”, que possui um arquivo .xlsx chamado “dicionário\_dados\_educação\_básica”, que possui um dicionário em tabela para ajudar o nosso entendimento na tabela referente ao censo; e na outra pasta, chamada “ANEXO II - Questionários do Censo da Educação Básica”, temos acesso à vários arquivos em .pdf, cada um referente a um dos questionário do censo. Temos um arquivo em .pdf chamado “Aluno”, referente ao questionário disponibilizado aos alunos no dia em que o censo é realizado, outro arquivo chamado “Escola”, que coleta dados sobre a escola em que o estudo está sendo feito, outro chamado “Gestor\_Escolar”, que é disponibilizado à equipe de gestão da escola, outro chamado “Profissional\_Escolar”, disponibilizado aos profissionais (professores, funcionários) da escola em questão, e um último arquivo chamado “Turma”, que coleta dados da turma escolar onde o estudo é realizado.

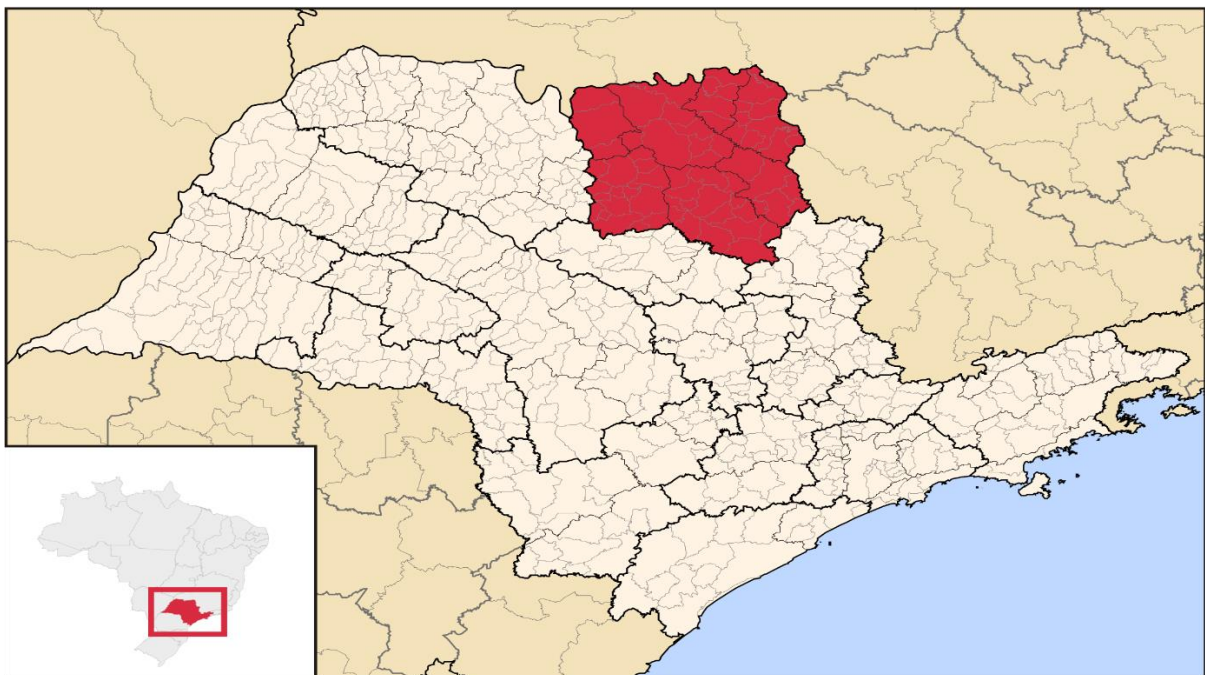
Ao tentar abrir o arquivo “microdados\_ed\_basica\_2021” no *python*, acabei presenciando um problema em relação ao tamanho da planilha. Por se tratar de uma tabela que representa todas as escolas do Brasil inteiro, temos aqui um total de 221141 linhas, distribuídas entre 344 colunas, o que nos totaliza uma planilha com mais de 76 milhões de células, fazendo com que não só o arquivo se torne

extremamente grande (se tratando de um arquivo .csv), como também nos impossibilitando de abrir ele através do *python*, pois o computador simplesmente não carrega o arquivo. Por conta disso, acabei sentindo a necessidade de formatar a planilha, de modo a torná-la menor, facilitando o trabalho sobre ela. Para isso, decidi selecionar somente a mesorregião de Ribeirão Preto para realizar a minha pesquisa, fazendo com que a planilha fique com cerca de 2000 colunas, o que ainda é um número grande se tratando de uma tabela, mas bem menor do que antes, se tornando um trabalho bem mais fácil.

## **2.4 – Definição de “Mesorregião”**

Mesorregiões são regiões compostas por diversos municípios de uma determinada área geográfica de um estado brasileiro. Nestas divisões é levada em consideração similaridades destes municípios em áreas como sociais e econômicas.

O estado de São Paulo foi dividido geograficamente pelo IBGE em 15 mesorregiões, sendo uma delas a mesorregião de Ribeirão Preto, contando com 66 municípios, dividida entre sete microrregiões, sendo elas: Barretos, São Joaquim da Barra, Ituverava, Franca, Jaboticabal, Ribeirão Preto e Batatais. A mesorregião de Ribeirão Preto possui cerca de mais de duas mil escolas, divididas entre os seus municípios.



*Localização da Mesorregião de Ribeirão Preto, no estado de São Paulo.*

Existem cerca de 50 faculdades em atividade em Ribeirão Preto. Desse total, apenas duas são públicas. A grande quantidade de universidades presentes na cidade acaba sendo um grande atrativo aos estudantes da região, devido à escassez de faculdades nas cidades presentes na microrregião de Ribeirão Preto. Por isso, a população da cidade tem uma grande quantidade de estudantes. Muitos dos que estudam na cidade as vezes nem moram nela, usam do transporte intermunicipal para estudar. Eu sou um desses alunos, e achei que seria interessante trabalhar em um tema sobre a região em que moro.

### **3 – Cinco tópicos principais a serem analisados**

Como citado na apresentação do trabalho, a partir dos dados que escolhermos trabalhar, precisamos escolher e ressaltar cinco tópicos principais, à escolha do aluno. Os cinco tópicos escolhidos por mim são:

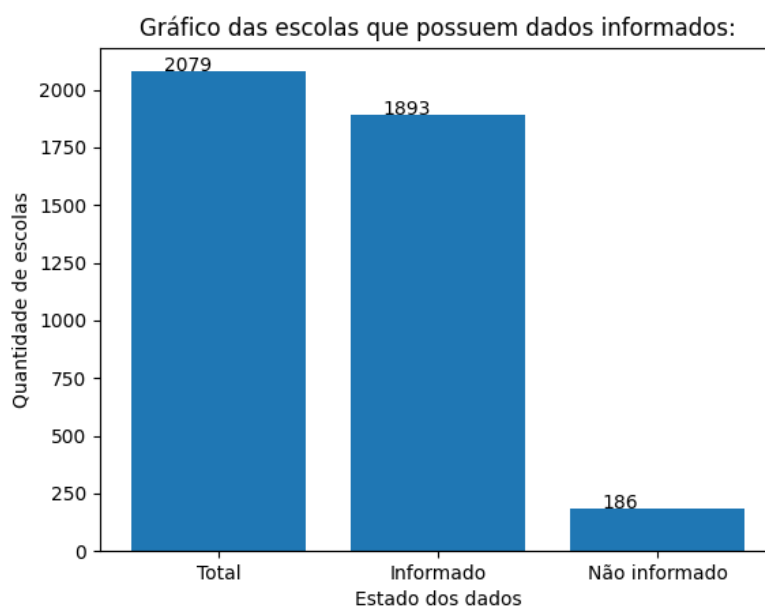
- Escolas que não realizam o tratamento do lixo;
- Escolas que possuem um laboratório de informática;
- Escolas que possuem um refeitório;
- Escolas que possuem banheiros acessíveis a pessoas com necessidades especiais;
- Escolas que não possuem nenhuma forma de acessibilidade a pessoas com necessidades especiais.

A seguir, irei falar sobre a Importância e a relevância de cada um desses tópicos, assim como a conclusão que tive com as pesquisas.

#### **3.1 – Quantidade total de escolas que realizaram a pesquisa**

A planilha original do Censo não informa a quantidade total de escolas que realizaram e das que não realizaram a pesquisa, portanto, senti a necessidade de realizar o cálculo, para que tenhamos mais informações a respeito da pesquisa. Através de meu código, consegui chegar à seguinte conclusão: de um total de 2079

escolas na nossa região, 1893 realizaram o Censo, enquanto 186 não possuem informações.



*Gráfico que representa a quantidade total de escolas que possuem dados informados. O gráfico foi gerado pelo meu script, através da biblioteca “Matplotlib” do Python, baseado nos dados coletados pelo Censo escolar em 2021.*

### **3.2 - Escolas que não realizam o tratamento do lixo**

O ambiente escolar é muito propício à coleta seletiva. Isso porque, além de destinar o lixo de forma adequada, é um ato de ensinar aos alunos sobre responsabilidade ambiental e ensinar a importância da reciclagem para a construção de um mundo melhor para todos.

Vale ressaltar que a formação dos alunos deve ser muito mais ampla do que os assuntos abordados nos livros e envolver a formação da cidadania como forma de aumentar a autoestima e valorizar o meio ambiente. A consciência dos recursos naturais cada vez menores do planeta prepara os alunos com uma atitude mais positiva em relação à sustentabilidade e responsabilidade social.

Todos os dias, uma escola pode gerar muitos resíduos. A maior parte desse lixo se acumula durante o recreio, mas também há lixo espalhado em salas de aula, banheiros e outras áreas de tráfego intenso.



Ninguém está isento de gerar lixo, mas sempre é possível trabalhar com coleta seletiva refletindo sobre o desperdício e aprendendo mais sobre a forma correta de descartar cada item.

Crianças e adolescentes tendem a copiar comportamentos e transmiti-los, principalmente quando estão em grupos sociais. É por este comportamento que é importante expor os jovens a ambientes que apostam em bons valores como a empatia, a importância da educação, a generosidade e a consciência das circunstâncias em que vivemos.

As escolas devem tomar a iniciativa dessas ações, convidando entidades que possam falar diretamente com os alunos sobre o lixo e como os materiais gerados no ambiente escolar são reciclados. Em seguida, é necessária uma análise do lixo local e da melhor forma de descartá-lo.

Através do estudo da pesquisa realizada pelo Censo, cheguei a infeliz conclusão de que ainda estamos muito longe do ideal. Para se ter noção, na nossa região, cerca de 53% das escolas não realizam nenhuma forma de tratamento de lixo, ou seja, simplesmente descartam o lixo de qualquer maneira. O desenvolvimento de uma educação ambiental envolve um conjunto de fatores no universo educativo, diversos sistemas de conhecimento e o envolvimento de toda a sociedade. Ocorre a necessidade de reflexão e mudanças de hábitos, de pensar e agir em torno da questão ambiental.



*Gráfico que representa a quantidade total de escolas que não realizam o tratamento do lixo. O gráfico foi gerado pelo meu script, através da biblioteca "Matplotlib" do Python, baseado nos dados coletados pelo Censo escolar em 2021.*

### **3.3 - Escolas que possuem um laboratório de informática;**

Não é segredo o quanto a informática é importante para a formação do aluno. Os computadores trouxeram consigo uma gama de softwares e aplicativos, trazendo uma interatividade que não é possível em salas de aula comuns.

Essa interatividade é fundamental para ensinar novos aspectos e motivar o que aprendem. Em uma idade mais jovem, a interatividade pode ser o fator determinante no interesse de um aluno pelo que está aprendendo. Este aspecto perpassa toda a vida escolar. Você precisa estar sempre atualizado para manter a atenção de seus alunos, principalmente em uma era digital como esta.

Outro ponto essencial da introdução da informática nas escolas é preparar os alunos para um mundo mais tecnológico, fornecendo-lhes uma base fácil para desenvolver suas competências futuras.

A introdução da tecnologia da informação nas escolas não é mais uma opinião, mas um consenso. A importância dos laboratórios de informática nas escolas quase pode ser exagerada. Não há razão para não haver uma sala como esta em um ambiente escolar.

No laboratório, os alunos trabalharão com computadores. Essa conexão pode ser feita por meio de aulas básicas de informática ou mesmo por meio de aulas complementares de disciplinas escolares comuns, incluindo jogos educativos de matemática, química, português, etc. Ao fornecer aos alunos essas oportunidades diferenciadas de aprendizado, seu interesse em aprender pode ser mantido.

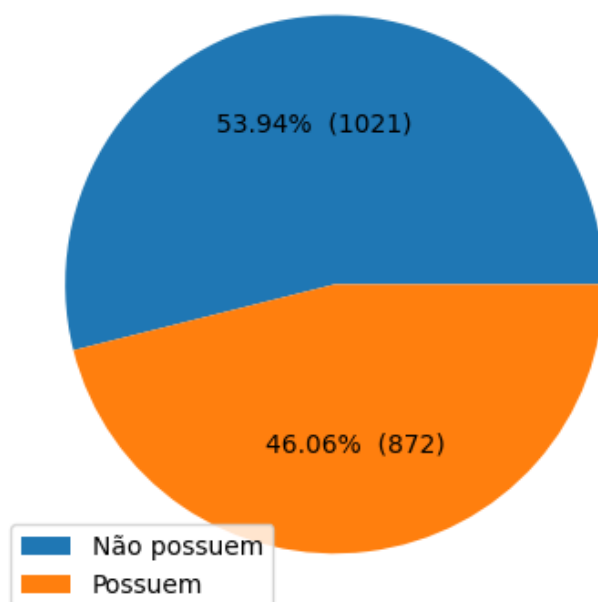
Outra característica importante dos laboratórios de informática é que muitas vezes é lá que as crianças e jovens são expostos pela primeira vez aos computadores, confirmando a importância social desses ambientes, pois é por meio deles que há a inclusão digital de pessoas que não poderiam ter acesso a tecnologia de outra forma.

Esses laboratórios não só auxiliam no aprendizado, mas tem um importante papel para que crianças com condições menos favoráveis possam ter chances iguais no futuro.

Através do estudo da pesquisa realizada pelo Censo, cheguei à conclusão de que muita coisa ainda precisa mudar. Para se ter noção, cerca de 54% das escolas da nossa região ainda não possuem um laboratório de informática disponível aos alunos.

É importante salientar que isso se dá ao fato de que muitas dessas escolas infelizmente não possuem verba suficiente para realizarem a construção de um laboratório adequado aos alunos. Devemos conscientizar os nossos representantes políticos acerca desses dados, e cobrarmos uma posição adequada aos mesmos. Não podemos simplesmente “passar pano” a qualquer tipo de atitude maliciosa como desvios de verba, principalmente se tratando da área da educação.

Gráfico das escolas que possuem um laboratório de informática:



*Gráfico que representa a quantidade total de escolas que possuem um laboratório de informática. O gráfico foi gerado pelo meu script, através da biblioteca “Matplotlib” do Python, baseado nos dados coletados pelo Censo escolar em 2021.*

### **3.4 - Escolas que possuem um refeitório**

Ninguém faz uma refeição apenas para matar a fome. Ao longo da história, as refeições durante as rotinas sempre serviram a uma função social. A hora do almoço, jantar ou café também é utilizada pelas pessoas para compartilhar diversão e estilo de vida. Sem contar o aspecto emocional evocado quando os aromas de feijão torrado, leite fervido e café aguçam os sentidos. Na escola, a hora do almoço das crianças não foge à regra: elas precisam comer em um ambiente agradável e acolhedor, o que mostra que a escola as respeita.

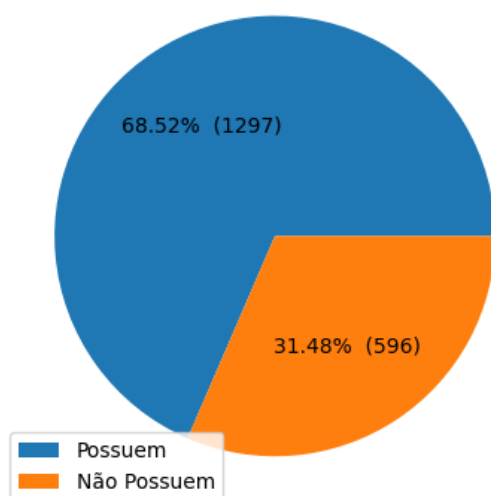
Apenas uma em cada quatro escolas públicas do Brasil possui refeitório. Nos outros três, os alunos comiam em pé, em salas de aula, em bancos de espera ou sentados no chão. Nas redes privadas, a situação é bem parecida: apenas 32% possuem um ambiente específico para refeições. Essa instabilidade reflete a saída do prédio da escola do programa de educação abrangente em evolução.

O mais dramático é o patamar em que não há qualquer local adaptado e as crianças sentam-se no chão. O mais comum, porém, é a instalação de bancos e mesões no pátio, no mesmo local em que se brinca, e a falta de ações para que a refeição seja também um momento prazeroso.

Através da minha pesquisa, descobri que cerca de 31% das escolas na nossa região ainda não possuem um refeitório ou um local adequado para a alimentação dos alunos. Assim como no último tópico, ressalto que devemos conscientizar os nossos representantes políticos acerca desses dados, e cobrarmos uma posição adequada aos mesmos. É inadmissível uma escola não possuir um

ambiente adequado para a alimentação dos alunos.

Gráfico das escolas que possuem um refeitório:



*Gráfico que representa a quantidade total de escolas que possuem um refeitório. O gráfico foi gerado pelo meu script, através da biblioteca “Matplotlib” do Python, baseado nos dados coletados pelo Censo escolar em 2021.*

### **3.5 - Escolas que não possuem nenhuma forma de acessibilidade a pessoas com necessidades especiais**

Muito mais do que uma obrigação, a acessibilidade nas escolas é interesse de todos.

Em vigor desde 2008, a Política Nacional de Educação Especial na Perspectiva da Educação Inclusiva está presente tanto em escolas da rede pública, quanto privada. Todas as pessoas com deficiência física, intelectual, auditiva ou visual possuem o direito à igualdade de oportunidades e cabe às instituições o dever de manter assegurado o sistema educacional inclusivo em todos os níveis de ensino para promover acessibilidade nas escolas.

Sabemos que pessoas deficientes têm o direito de igualdade em qualquer instituição de ensino, mas muito além das obrigações implementadas por lei, não podemos deixar a empatia e humanidade de lado e termos total consciência de que antes de tudo, somos todos seres humanos e lutamos pelos mesmos direitos.

Uma instituição regular de ensino deve conceder educação especial para todo e qualquer tipo de aluno com deficiência. É um processo rigoroso, que exige mudanças, desde o modo de condução e adaptação das aulas, até a maneira de como o indivíduo vai se relacionar com o restante da turma e com as famílias. O preparo deve partir dos educadores e de todos os envolvidos na gestão.

A educação inclusiva é um tema de extrema importância, mas muitas instituições ainda não contam com as adaptações adequadas para atender acessibilidade nas escolas. As melhorias devem ocorrer, para chegarem num patamar em que determinados paradoxos sejam rompidos e essa realidade seja de extrema naturalidade, minimizando os empecilhos causados por diferentes condições.

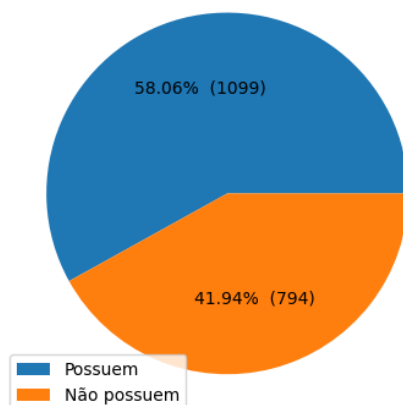
Existem diversas questões em relação ao sistema inclusivo no ambiente escolar, e aqui podemos esclarecer sobre as principais formas de acessibilidade que devem estar presentes nas instituições:

- **Arquitetônica:** a acessibilidade arquitetônica corresponde ao espaço e aos equipamentos que devem ser implementados para atender os alunos com necessidades especiais, garantindo sua segurança durante a locomoção no ambiente.
- **Comportamental:** consiste na cultura de valores inclusivos dentro da instituição, conscientizando à todos sobre a importância da inclusão e prezando sempre pelo respeito às diferenças.
- **Comunicação:** envolve diversos elementos favoráveis às pessoas com deficiência, desde a acústica das salas de aula e demais ambientes, até as sinalizações e figuras que possam auxiliar os mesmos na comunicação e interação com os demais.

- Técnica: a acessibilidade técnica trata-se da inclusão de equipamentos e produtos que possam auxiliar a rotina de uma pessoa com necessidades especiais, como texturas nos pisos, barras de apoio no banheiro (vaso sanitário), corrimãos etc.
- Pedagógica: são itens primordiais para a inclusão, que permitem aos alunos deficientes uma experiência e vivência saudável dentro do ambiente escolar. Isso inclui elementos como programas (softwares) para pessoas com deficiência visual, bem como adaptadores para lápis, caneta, tesoura, entre outros.

Através das minhas pesquisas, cheguei à conclusão de que a situação é extremamente grave quando se trata de acessibilidade nas escolas da nossa região. Para se ter noção, cerca de quase 42% das escolas ainda não possuem nenhuma forma de acessibilidade a pessoas com necessidades especiais. Isso quer dizer que quase metade das nossas escolas não oferecem nenhuma forma de possibilitar um melhor ensino a esses alunos. Se tratando de escolas, principalmente em meios públicos, o meu ponto de vista continua o mesmo: devemos exigir de nossos representantes uma transposição acerca dos fatos, e cobrá-los uma melhor condição a esses alunos.

Gráfico das escolas que não possuem nenhuma forma de acessibilidade à pessoas com necessidades especiais:



*Gráfico que representa a quantidade total de escolas que não possuem nenhuma forma de acessibilidade à pessoas com necessidades especiais. O gráfico foi gerado pelo meu script, através da biblioteca "Matplotlib" do Python, baseado nos dados coletados pelo Censo escolar em 2021.*

### **3.6 - Escolas que possuem banheiros acessíveis a pessoas com necessidades especiais**

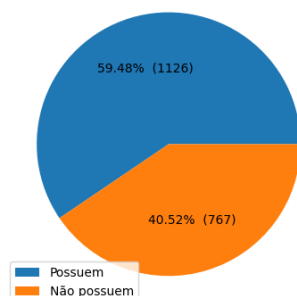
Proporcionar acessibilidade e infraestrutura adequada para pessoas com deficiência é fundamental para garantir qualidade de vida e autonomia a elas. Cada vez mais, os profissionais de arquitetura desenvolvem soluções que facilitam a usabilidade de todos os ambientes. Entre as transformações necessárias, destacamos a importância dos banheiros adaptados.

Devemos sempre buscar que as pessoas com deficiência tenham o máximo de autonomia possível. Elas devem manter sua funcionalidade e poder realizar as tarefas cotidianas com segurança. Assim, a adaptação da casa em conjunto com o fornecimento de equipamentos adequados é uma medida de qualidade de vida.

Como discutido no tópico anterior, uma instituição regular de ensino deve conceder educação especial para todo e qualquer tipo de aluno com deficiência, principalmente em relação à sua deslocação dentro da instituição e aos banheiros, pois é de extrema importância que esses alunos não possuam nenhum empecilho na hora de realizar as suas necessidades.

Através de minhas pesquisas, é possível observar que cerca de 40% das escolas ainda não possuem um banheiro adaptado às necessidades especiais desses alunos.

Gráfico das escolas que possuem banheiros acessíveis a pessoas com necessidades especiais:



*Gráfico que representa a quantidade total de escolas que possuem banheiros acessíveis à pessoas com necessidades especiais. O gráfico foi gerado pelo meu script, através da biblioteca "Matplotlib" do Python, baseado nos dados coletados pelo Censo escolar em 2021.*



## **4 – Script**

Através do próximo tópico, irei falar sobre o script feito para a análise dos dados, assim como as bibliotecas que utilizei, a forma como o código é estruturado, a forma que foi feito e um breve relatório sobre como foi trabalhar nesse código.

### **4.1 – Script: Estrutura do código**

Para a realização do meu código, utilizei as seguintes bibliotecas presentes no *Python*: “*Pandas*” (responsável pela interação entre o *Python* e as planilhas), “*Matplotlib*” (responsável pela impressão dos gráficos) e “*Openpyxl*” (responsável por criar planilhas). Através do *Python*, eu queria criar um código que não só calculasse e me fornecesse a diferença entre as informações da planilha, como também me imprimisse um gráfico baseado nas informações e gerasse uma nova planilha baseada nos dados anteriormente analisados, tudo isso de forma automática, buscando automatizar a análise de dados e planilhas.

#### **4.1.1 – Script: Estrutura do código – Biblioteca**

##### **Pandas**

Em programação de computadores, *pandas* é uma biblioteca de software criada para a linguagem *Python* para manipulação e análise de dados. Em particular, oferece estruturas e operações para manipular tabelas numéricas e séries temporais. É software livre sob a licença BSD. O nome é derivado do termo inglês “panel data”(dados em painel), um termo usado em estatística e econometria para conjunto de dados que incluem várias unidades amostrais (indivíduos, empresas, etc) acompanhadas ao longo do tempo.

## **4.1.2 – Script: Estrutura do código – Biblioteca**

### **Matplotlib**

O *Matplotlib* é uma biblioteca para a visualização de dados em *Python*. Ele apresenta uma API orientada a objetos que permite a criação de gráficos em 2D de uma forma simples e com poucos comandos. A ferramenta disponibiliza diversos tipos de gráficos, como em barra, em linha, em pizza, histogramas entre outras opções. A API foi projetada para ser compatível com o MATLAB que, apesar de ser referência na área de processamento numérico, não é opensource.

## **4.1.3 – Script: Estrutura do código – Biblioteca**

### **Openpyxl**

*Openpyxl* é uma biblioteca *Python* para leitura e gravação de arquivos Excel (com extensão *xlsx* / *xlsm* / *xlsx* / *xltm*). O módulo *openpyxl* permite que o programa *Python* leia e modifique arquivos Excel.

Por exemplo, os usuários podem ter que percorrer milhares de linhas e selecionar algumas informações para fazer pequenas alterações com base em alguns critérios. Usando o módulo *Openpyxl*, essas tarefas podem ser realizadas de forma muito eficiente e fácil.

## **4.2 – Script: Relatório com as linhas do código**

Segue a seguir as linhas de código presentes do script:

```
# Aqui, estou importando os módulos necessários para o funcionamento do
# código, sendo eles: "matplotlib", responsável por gerar os gráficos em
# imagem; "pandas", responsável pela
# interação entre o python e as planilhas; e "openpyxl", responsável por
# gerar as novas planilhas baseadas nos dados analisados. Caso não possua os
# módulos, eles podem ser
# instalados com o comando "pip3 install".
import matplotlib.pyplot as plt
import pandas as pd
import openpyxl
```

```
# Aqui, estou definindo qual é a planilha para o módulo "pandas" poder ler,
e logo em seguida definindo uma variável à aquela planilha, no formato de
DataFrame.
```

```
planilha = 'Planilhas Censo 2021/Censo 2021 - Mesorregião de Ribeirão
Preto.xlsx'
```

```
df_planilha = pd.read_excel(planilha)
```

```
# Como nem todas as escolas possuem dados informados, fiz com que o python
me informasse a quantidade total de escolas que foram analisadas, a
quantidade de escolas que não
```

```
# possuem dados informados (fiz com que o módulo "pandas" me informasse as
células da planilha com dados vazios), e logo em seguida calculei a
quantidade de escolas que possuem dados informados.
```

```
qntescolas = df_planilha[df_planilha.columns[0]].count()
```

```
qntescolasna = df_planilha.IN_TRATAMENTO_LIXO_INEXISTENTE.isnull().sum()
```

```
qntescolasenso = (qntescolas) - (qntescolasna)
```

```
# Aqui, estou imprimindo ao usuário os dados calculados acima.
```

```
print("Quantidade total de escolas na mesorregião de Ribeirão Preto: ",
qntescolas)
```

```
print("Quantidade de escolas na mesorregião de Ribeirão Preto que possuem
dados informados: ", qntescolasenso)
```

```
print("Quantidade de escolas na mesorregião de Ribeirão Preto que não
possuem dados informados: ", qntescolasna)
```

```
# Como os dados calculados acima não são informados na planilha principal,
fiz com que o módulo "openpyxl" criasse uma nova planilha baseada nos dados
analisados.
```

```
df_escolasenso = pd.DataFrame([[qntescolas, qntescolasenso,
qntescolasna]], index = ['Quantidade'], columns = ['Total de escolas',
'Dados informados', 'Dados não informados'])
```

```
df_escolasenso.to_excel('Novas planilhas
geradas/PlanilhaTotaisCenso.xlsx', sheet_name = 'TotaisCenso')
```

```
print("Uma nova planilha com a quantidade total das escolas com dados
informados foi gerada!")
```

```
# Agora que calculamos a quantidade total de escolas, escolas que possuem e
que não possuem dados informados, montei uma lista baseada nessas
variáveis, e fiz com que o módulo
```

```
# matplotlib analise e imprima um gráfico de barras para o usuário, baseado
nos dados já analisados. O gráfico é salvo em um arquivo .png logo em
seguida. Como o módulo matplotlib
```

```
# não possui nenhuma forma nativa de exibir a quantidade dos dados em cima
de cada barra, usei uma estrutura de repetição "for" para mostrar os dados
em cima das barras.
```

```
lista_escolas = [qntescolas, qntescolasenso, qntescolasna]
```

```
xlista = [x for x in range(len(lista_escolas))]
```

```
for i, v in enumerate(lista_escolas):
```

```
    plt.text(xlista[i] - 0.25, v + 0.01, str(v))
```

```
plt.bar(xlista, lista_escolas)
```

```
plt.title("Gráfico das escolas que possuem dados informados:")
```

```
plt.xlabel("Estado dos dados")
```

```
plt.ylabel("Quantidade de escolas")
```

```
plt.xticks([0, 1, 2], ['Total', 'Informado', 'Não informado'])
```

```
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que possuem
dados informados.png')
```

```
plt.show()
```

```
# A partir daqui, começamos a trabalhar com as cinco colunas
correspondentes aos dados que quero analisar. Criei duas variáveis, em uma
fiz com que o python calculasse a quantidade
```

```

# de dados que possuem em cada uma das possíveis representações na planilha
("Realizam " e "Não realizam"), e na outra pedi que ele calculasse a
porcentagem entre esses dados. Em
# seguida, imprimir esses dados ao usuário, tive que usar o parâmetro
".to_string()" para evitar que o script mostrasse informações da variável
(caso não esteja em string, ele exibe
# a tipagem da variável junto com os dados, informação desnecessária
durante a análise), e pedi para que o módulo "matplotlib" imprimisse um
gráfico em formato
# de pizza baseado nessas informações.
qntescolaslixo =
df_planilha['IN_TRATAMENTO_LIXO_INEXISTENTE'].value_counts()
pctescolaslixo =
df_planilha['IN_TRATAMENTO_LIXO_INEXISTENTE'].value_counts(normalize =
True).mul(100).apply(lambda x: "{:, .2f}".format(x))+ '%'
print ("Quantidade de escolas que não realizam o tratamento do lixo: \n",
qntescolaslixo.to_string())
print ("Porcentagem de escolas que não realizam o tratamento do lixo: \n",
pctescolaslixo.to_string())
plt.pie(qntescolaslixo, autopct = lambda p : '{:.2f}% ({:.0f})'.format(p,
p * sum(qntescolaslixo)/100))
plt.title("Gráfico das escolas que não \nrealizam o tratamento do lixo:")
plt.legend(["Não realizam", "Realizam"], loc = 'lower left')
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que não
realizam o tratamento do lixo.png')
plt.show()

qntescolasref = df_planilha['IN_REFEITORIO'].value_counts()
pctescolasref = df_planilha['IN_REFEITORIO'].value_counts(normalize =
True).mul(100).apply(lambda x: "{:, .2f}".format(x))+ '%'
print ("Quantidade de escolas que possuem um refeitório: \n",
qntescolasref.to_string())
print ("Porcentagem de escolas que possuem um refeitório: \n",
pctescolasref.to_string())
plt.pie(qntescolasref, autopct = lambda p : '{:.2f}% ({:.0f})'.format(p, p
* sum(qntescolasref)/100))
plt.title("Gráfico das escolas que possuem um refeitório:")
plt.legend(["Possuem", "Não Possuem"], loc = 'lower left')
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que possuem um
refeitório.png')
plt.show()

qntescolaslab = df_planilha['IN_LABORATORIO_INFORMATICA'].value_counts()
pctescolaslab =
df_planilha['IN_LABORATORIO_INFORMATICA'].value_counts(normalize =
True).mul(100).apply(lambda x: "{:, .2f}".format(x))+ '%'
print ("Quantidade de escolas que possuem um laboratório de informática:
\n", qntescolaslab.to_string())
print ("Porcentagem de escolas que possuem um laboratório de informática:
\n", pctescolaslab.to_string())
plt.pie(qntescolaslab, autopct = lambda p : '{:.2f}% ({:.0f})'.format(p, p
* sum(qntescolaslab)/100))
plt.title("Gráfico das escolas que possuem\n um laboratório de
informática:")
plt.legend(["Não possuem", "Possuem"], loc = 'lower left')
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que possuem um
laboratório de informática.png')
plt.show()

qntescolasbanpne = df_planilha['IN_BANHEIRO_PNE'].value_counts()

```

```

pctescolasbanpne = df_planilha['IN_BANHEIRO_PNE'].value_counts(normalize =
True).mul(100).apply(lambda x: "{:, .2f}".format(x))+ '%'
print ("Quantidade de escolas que possuem banheiros acessíveis à pessoas
com necessidades especiais: \n", qntescolasbanpne.to_string())
print ("Porcentagem de escolas que possuem banheiros acessíveis à pessoas
com necessidades especiais: \n", pctescolasbanpne.to_string())
plt.pie(qntescolasbanpne, autopct = lambda p : '{:.2f}%
({:.0f})'.format(p, p * sum(qntescolasbanpne)/100))
plt.title("Gráfico das escolas que possuem banheiros\n acessíveis à pessoas
com necessidades especiais:")
plt.legend(["Possuem", "Não possuem"], loc = 'lower left')
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que possuem
banheiros acessíveis à pessoas com necessidades especiais.png')
plt.show()

```

```

qntescolaslpne = df_planilha['IN_ACESSIBILIDADE_INEXISTENTE'].value_counts()
pctescolaslpne =
df_planilha['IN_ACESSIBILIDADE_INEXISTENTE'].value_counts(normalize =
True).mul(100).apply(lambda x: "{:, .2f}".format(x))+ '%'
print ("Quantidade de escolas que não possuem nenhuma forma de
acessibilidade à pessoas com necessidades especiais: \n",
qntescolaslpne.to_string())
print ("Porcentagem de escolas que não possuem nenhuma forma de
acessibilidade à pessoas com necessidades especiais: \n",
pctescolaslpne.to_string())
plt.title("Gráfico das escolas que não possuem nenhuma\n forma de
acessibilidade à pessoas com necessidades especiais:")
plt.pie(qntescolaslpne, autopct = lambda p : '{:.2f}% ({:.0f})'.format(p, p
* sum(qntescolaslpne)/100))
plt.legend(["Possuem", "Não possuem"], loc = 'lower left')
plt.savefig('Imagens geradas dos gráficos/Gráfico de escolas que não
possuem nenhuma forma de acessibilidade à pessoas com necessidades
especiais.png')
plt.show()

```

```

print("Foi gerado um arquivo .png para cada gráfico analisado!")

```

```

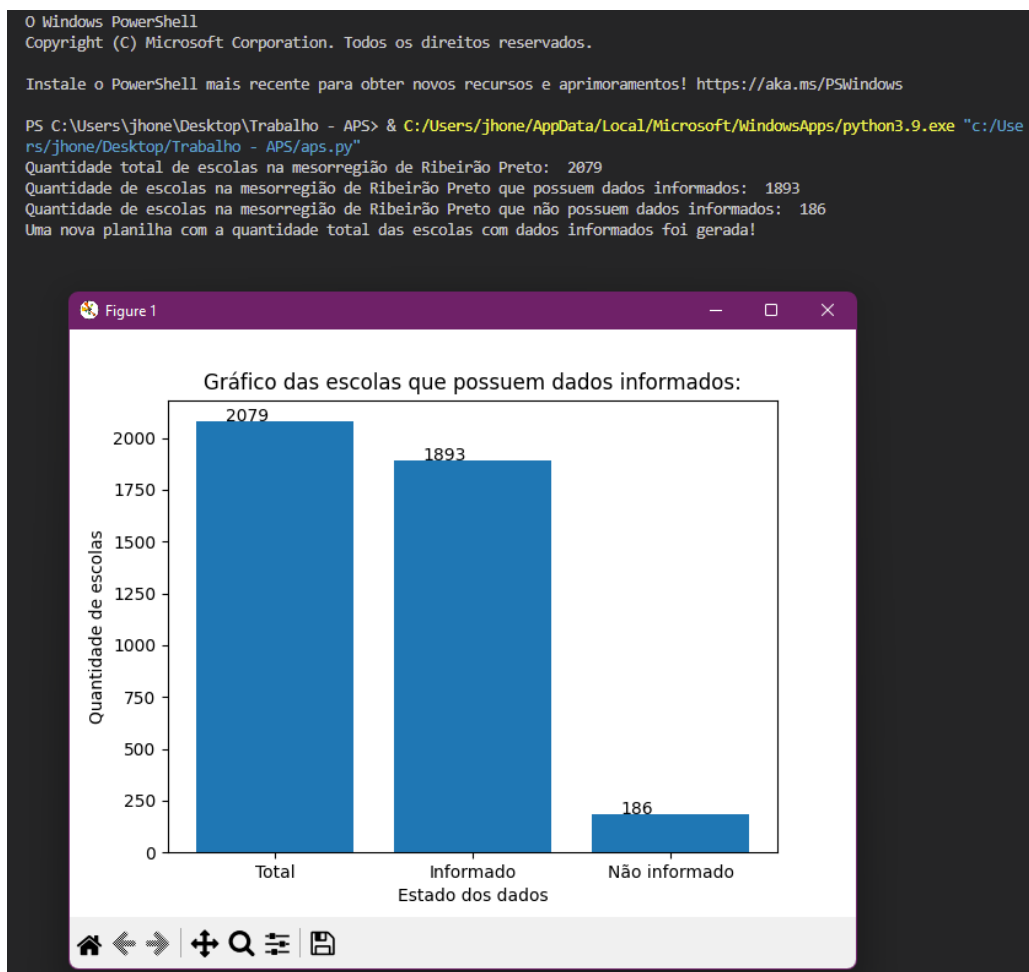
# Para finalizar o script, criei um novo DataFrame baseado nos cinco
tópicos analisados, e pedi para que o módulo "openpyxl" criasse uma nova
planilha baseada nesses dados.
nova_planilha = pd.DataFrame([[qntescolaslixo.to_string(),
pctescolaslixo.to_string()], [qntescolasref.to_string(),
pctescolasref.to_string()],
[qntescolaslab.to_string(), pctescolaslab.to_string()],
[qntescolasbanpne.to_string(), pctescolasbanpne.to_string()],
[qntescolaslpne.to_string(),
pctescolaslpne.to_string()]], index = ['Escolas que não realizam o
tratamento do lixo', 'Escolas que possuem refeitório', 'Escolas que possuem
um laboratório de informática',
'Escolas que possuem banheiros acessíveis à pessoas com necessidades
especiais', 'Escolas que não possuem nenhuma forma de acessibilidade à
pessoas com necessidades especiais'],
columns = ['quantidade', 'porcentagem'])
nova_planilha.to_excel('Novas planilhas geradas/NovaPlanilha.xlsx',
sheet_name = 'NovaPlanilha')
print("Uma nova planilha com os dados analisados foi gerada!")

```

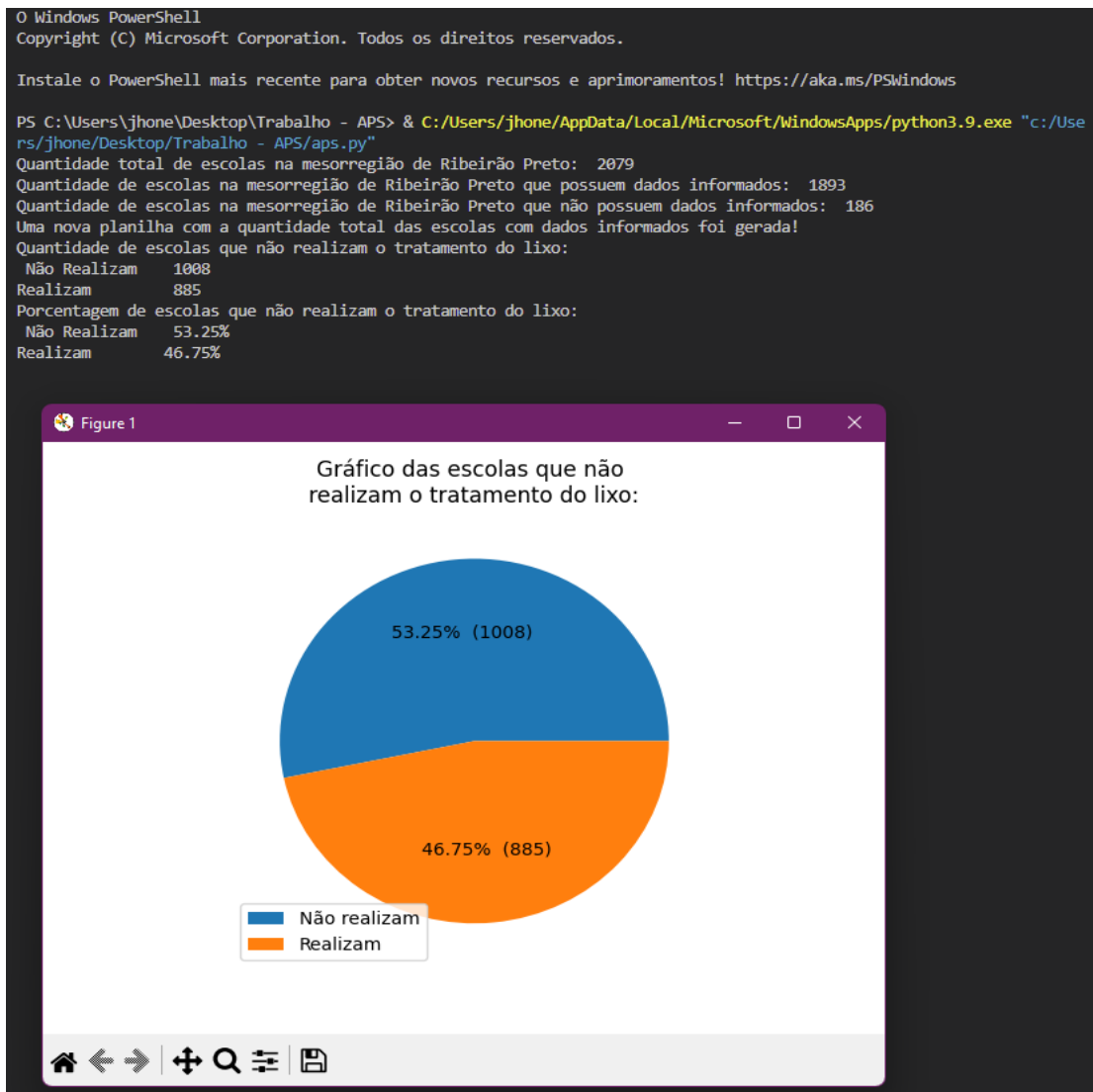
## 4.3 – Script em funcionamento

Ao solicitar para o VSCode executar o código, sua execução ocorre da seguinte maneira:

**1º:** O código nos fornece a quantidade total das escolas que realizaram a pesquisa, gerando uma nova planilha baseada nesses dados e nos imprimindo um gráfico no formato de colunas, que facilita a análise dos dados. O gráfico logo em seguida é salvo como imagem em uma pasta.



**2º:** Ao fecharmos o gráfico, o código parte para o próximo passo, que é imprimir os dados baseados no tópico escolhido, tanto em quantidade quanto em porcentagem. O gráfico também é salvo como imagem, indo assim sucessivamente para os próximos cinco tópicos.



**3º:** Após termos verificado cada tópico, os seus dados (em quantidade e em porcentagem) e seus gráficos, o código gera uma nova planilha baseada nos tópicos que acabamos de analisar. O fim do programa é exibido da seguinte forma:

```

O Windows PowerShell
Copyright (C) Microsoft Corporation. Todos os direitos reservados.

Instale o PowerShell mais recente para obter novos recursos e aprimoramentos! https://aka.ms/PSWindows



PS C:\Users\jhone\Desktop\Trabalho - APS> & C:/Users/jhone/AppData/Local/Microsoft/WindowsApps/python3.9.exe "c:/Users/jhone/Desktop/Trabalho - APS/aps.py"
Quantidade total de escolas na mesorregião de Ribeirão Preto: 2079
Quantidade de escolas na mesorregião de Ribeirão Preto que possuem dados informados: 1893
Quantidade de escolas na mesorregião de Ribeirão Preto que não possuem dados informados: 186
Uma nova planilha com a quantidade total das escolas com dados informados foi gerada!
Quantidade de escolas que não realizam o tratamento do lixo:
    Não Realizam    1008
    Realizam        885
Porcentagem de escolas que não realizam o tratamento do lixo:
    Não Realizam    53.25%
    Realizam        46.75%
Quantidade de escolas que possuem um refeitório:
    Possuem        1297
    Não Possuem     596
Porcentagem de escolas que possuem um refeitório:
    Possuem        68.52%
    Não Possuem     31.48%
Quantidade de escolas que possuem um laboratório de informática:
    Não Possuem    1021
    Possuem        872
Porcentagem de escolas que possuem um laboratório de informática:
    Não Possuem    53.94%
    Possuem        46.06%
Quantidade de escolas que possuem banheiros acessíveis à pessoas com necessidades especiais:
    Possuem        1126
    Não Possuem     767
Porcentagem de escolas que possuem banheiros acessíveis à pessoas com necessidades especiais:
    Possuem        59.48%
    Não Possuem     40.52%
Quantidade de escolas que não possuem nenhuma forma de acessibilidade à pessoas com necessidades especiais:
    Possuem        1099
    Não Possuem     794
Porcentagem de escolas que não possuem nenhuma forma de acessibilidade à pessoas com necessidades especiais:
    Possuem        58.06%
    Não Possuem     41.94%
Foi gerado um arquivo .png para cada gráfico analisado!
Uma nova planilha com os dados analisados foi gerada!
PS C:\Users\jhone\Desktop\Trabalho - APS>

```

Como podemos ver, as novas planilhas e os gráficos que representam os dados analisados são salvos em pastas, dessa forma, facilitando e automatizando a análise desses dados. As imagens podem ser usadas posteriormente em, por exemplo, a montagem de uma apresentação sobre os dados, assim como as planilhas podem ser usadas para posterior análise:



Este Computador > Área de Trabalho > Trabalho - APS > Novas planilhas geradas

Nome	Data de modificação	Tipo	Tamanho
 NovaPlanilha.xlsx	17/11/2022 23:32	Planilha do Micro...	6 KB
 PlanilhaTotaisCenso.xlsx	17/11/2022 23:26	Planilha do Micro...	5 KB

Este Computador > Área de Trabalho > Trabalho - APS > Imagens geradas dos gráficos

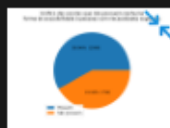


Gráfico de escolas que não possuem nenhuma form...

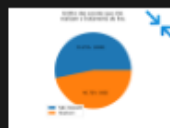


Gráfico de escolas que não realizam o tratamento do...

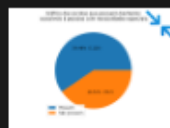


Gráfico de escolas que possuem banheiros a...

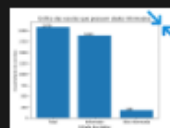


Gráfico de escolas que possuem dados informados.png

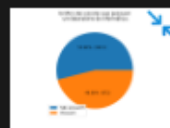


Gráfico de escolas que possuem um laboratório de...

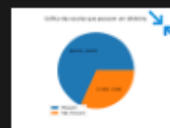


Gráfico de escolas que possuem um refeitório.png

## **5 – Bibliografia**

<https://dados.gov.br/dataset/microdados-do-censo-escolar>

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

<http://projetoportunadinho.ufmg.br/escola/eu-quero-saber/mesorregiao-microrregiao#:~:text=Mesorregi%C3%A3o%3A%20s%C3%A3o%20regi%C3%B5es%20compostas%20por,%C3%A1reas%20como%20sociais%20e%20econ%C3%B4micas.>

[https://pt.wikipedia.org/wiki/Lista\\_de\\_mesorregi%C3%B5es\\_e\\_microrregi%C3%B5es\\_de\\_S%C3%A3o\\_Paulo](https://pt.wikipedia.org/wiki/Lista_de_mesorregi%C3%B5es_e_microrregi%C3%B5es_de_S%C3%A3o_Paulo)

<https://www.pensamentoverde.com.br/reciclagem/lixo-na-escola-qual-importancia-e-como-podemos-reciclar-os-materiais-gerados-neste-ambiente/#:~:text=A%20escola%20deve%20tomar%20a,melhor%20forma%20de%20descart%C3%A1%2Dlo.>

<https://www.revistaea.org/artigo.php?idartigo=3735>

<https://canaldoensino.com.br/blog/laboratorio-de-informatica-a-importancia-nas-escolas>

<https://www.cartacapital.com.br/educacao/a-falta-dos-refeitorios-nas-escolas/>

<https://gestaoescolar.org.br/conteudo/740/o-espaco-da-refeicao>

<https://www.proesc.com/blog/acessibilidade-nas-escolas-conheca-importancia/>

<https://blog.freedom.ind.br/a-importancia-de-banheiros-adaptados-para-pessoas-com-deficiencia/>

[https://pt.wikipedia.org/wiki/Pandas\\_\(software\)](https://pt.wikipedia.org/wiki/Pandas_(software))

<https://king.host/blog/2018/03/visualizacao-de-dados-matplotlib/>

<https://acervolima.com/lendo-um-arquivo-excel-usando-o-modulo-python-openpyxl/>

**FICHA DE ATIVIDADES PRÁTICAS SUPERVISIONADAS - APS**

(Atividades Práticas supervisionadas (laboratórios, atividades em biblioteca, iniciação Científica, trabalhos individuais e em grupo, práticas de ensino e outras)

**NOME:** João Lucas Roldão Rocha - CC2P18

**RA:** N9416H4 **CURSO:** Ciência da Computação

**CAMPUS:** Ribeirão Preto - Vargas **SEMESTRE:** 2º **TURNO:** Noturno

DATA	ATIVIDADE	TOTAL DE HORAS	ASSIN	
			ALUNO	
14/11	Estudar e analisar um meio de realizar a atividade proposta	5	João Lucas R Rocha	
15/11	Realizar o código para a atividade proposta	10	João Lucas R Rocha	
16/11	Realizar a pesquisa para cada tópico escolhido para análise	5	João Lucas R Rocha	
17/11	Realizar o relatório sobre a atividade	5	João Lucas R Rocha	

**TOTAL DE HORAS:** 25