



UNIVERSIDADE PAULISTA

UNIP – Universidade Paulista

Ciência da Computação (CC)

Web-Scraping

Disciplina: APS

Prof. Paulo Dias

Integrantes:

- F0359E1 - Lucas Campanuchi Corrêa
- N485HJ6 - Vinicius Alves Panobianco
- N429160 - Wellington F. De Oliveira
- N3855J3 - Wesley Marcos M. Dos Santos

Ribeirão Preto - Maio de 2021

Índice

OBJETIVO E MOTIVAÇÃO DO TRABALHO	3
INTRODUÇÃO	4
FUNDAMENTOS DA COMUNICAÇÃO DE DADOS EM REDE	6
PLANO DE DESENVOLVIMENTO DA APLICAÇÃO	10
PROJETO	19
RELATÓRIO COM AS LINHAS DE CÓDIGO DO PROGRAMA	23
APRESENTAÇÃO DO PROGRAMA EM FUNCIONAMENTO	32
BIBLIOGRAFIA	33
FICHA DE ATIVIDADES	34

Objetivo e motivação do trabalho

Com o desenvolvimento tecnologico cada vez mais acelerado nos dias atuais, e com a alta competitividade no mercado consumidor entre empresas decorrido desses avanços tecnologicos, deu origem a novos meios de comércio, os quais trouxeram muito mais benecifios a aquele que se adaptaram a esses novos sistemas.

muitas empresas que que antigamente só faziam comercio de produtor unicamente por meio de lojas fisicas, tiveram que se integrar ao comercio digital por meio de Websites, os quais tornaram muito mais pratico e rapido se realizar compras e vendas, visto que poderia expandir muito mais o alcance de suas lojas.

Hoje possuir um web site para sua empresa ou negócio é extremamente vital para a qual possa competir e se manter, com relação às demais concorrentes. Mas com essa grande quantidade de site de vendas criados devido a essa concorrência, criou-se também um excesso de informações e ofertas sendo jogadas ao consumidor, o que as vezes até o atrapalha ou ocasiona compras erradas.

Objetivo deste trabalho é desenvolver com uso da linguagem Python, um pagina Web para coletar informações de sites de equipamentos eletrônicos focado em peças para computadores, atravez do precesso de raspagem(scraping) junto com os módulos Requests e BeautifulSoup.

Nós da LVW² Soluções, apartir dessa ideia desenvolvemos o projeto Hardware Finder, que tem a finalidade de possibilitar que qualquer pessoa possa fazer procurar as melhores ofertas hardware e acessorios, pois foi construido visando ser intuitivo, de fácil manuseio, em que o usuário possa buscar informações de produtos a venda e comparar entre as lojas.

Introdução

Para o desenvolvimento dessa página Web, utilizamos a linguagem de programação Python.

“O Python é uma linguagem de programação de alto nível e muito versátil. Ela suporta tanto a programação orientada a objetos quanto a programação estruturada. Com Python, você pode acessar bibliotecas nativas que oferecem funcionalidades para desenvolvimento de projetos e implementação de aplicações complexas. A tecnologia está presente nos códigos do Instagram, Netflix, Spotify, Reddit, Facebook, Google e muitos outros.”

Python dispõe de recursos como tipagem dinâmica, orientação a objetos, multiparadigmas, além de recursos excelentes em sua biblioteca padrão e via módulos e frameworks desenvolvidos pela sua comunidade.

“Desenvolvida pelo matemático holandês Guido van Rossum, atualmente a linguagem faz parte de um modelo de desenvolvimento comunitário, gerenciado pela organização sem fins lucrativos Python Software Foundation.

Lançada no começo dos anos 1990, a linguagem tem ganhado crescente notoriedade nos últimos anos, tornando-se uma das mais populares entre programadores, principalmente pela sua funcionalidade com dados, big data e inteligência artificial.”

Guido van Rossum se propôs a criar uma linguagem de script simples e incrementada com as melhores propriedades da linguagem ABC, que por sua vez havia muitas limitações. conceitos como lista, dicionários, bibliotecas e obrigatoriedade de indentação, tornavam a linguagem Python mais aceita pelos os programadores.

Uma de suas principais características em relação às outras linguagens de programação são a facilidade e praticidade de se criar códigos, sendo utilizado para pequenos ou grandes projetos.

Python possui também uma vasta biblioteca que fornece ao usuário um leque maior de possibilidades de aprimorar e aperfeiçoar seu projeto. Ele também possuindo um sistema que o permite ser multiplataforma, sendo capaz de rodar em várias

plataformas diferentes em relação a que foi originalmente criada como: Windows, Linux, Android, IOS, MAC, etc.

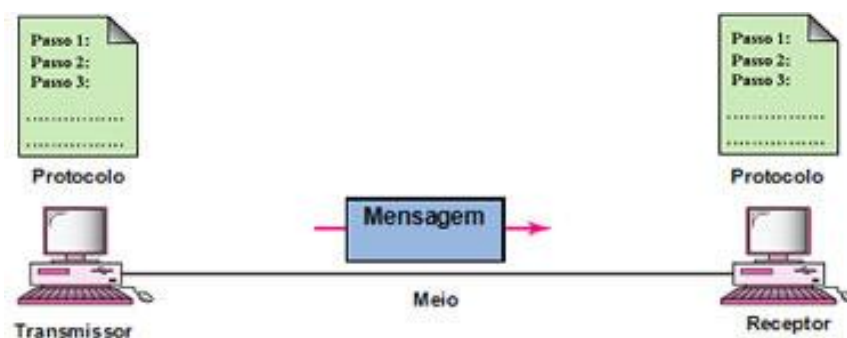
“Em vários aplicativos Web, é comum conectar-se a vários serviços de terceiros usando APIs. Ao usar essas APIs, você tem acesso a dados como informações meteorológicas, placares esportivos, listas de filmes, tweets, resultados de mecanismo de busca e fotos. Também é possível usar as APIs para adicionar funcionalidades ao seu app. Exemplos dessas funcionalidades são pagamentos, agendamentos, e-mails, traduções, mapas e transferência de arquivos. Se você tivesse que criar qualquer uma dessas funcionalidades por conta própria, levaria muito tempo, mas com as APIs pode levar somente alguns minutos para se conectar a uma funcionalidade e acessar seus recursos e dados.”

A biblioteca Request que permite o uso de uma API em Python implica em enviar solicitações HTTP e receber respostas. As solicitações HTTP representam o sistema que a Web funciona. quando se navega em uma página Web, seu navegador realiza várias solicitações para o servidor da página da Web, o qual manda as informações solicitadas e seu navegador renderiza a página.

“Beautiful Soup é uma biblioteca Python de extração de dados de arquivos HTML e XML. Ela funciona com o seu interpretador (parser) favorito a fim de prover maneiras mais intuitivas de navegar, buscar e modificar uma árvore de análise (parse tree). Ela geralmente economiza horas ou dias de trabalho de programadores ao redor do mundo.”

Fundamentos da comunicação de dados em rede

“Conforme Forouzan (2006), comunicação de dados é a troca de informação entre dois dispositivos através de algum meio de comunicação como, por exemplo, um par de fios.”



A Comunicação Dados trata a transmissão de informação entre sistemas computacionais e dispositivos diferentes através de um meio de transmissão. A transmissão de informação acontece pela passagem de sinais através dos meios físicos de comunicação que fazem parte das redes.

Um sistema de comunicação de dados possui cinco Elementos básicos que são: mensagem, transmissor, meio, receptor e protocolo.

A mensagem é o objeto de comunicação que pode ser transmitido tanto visual como auditivamente.

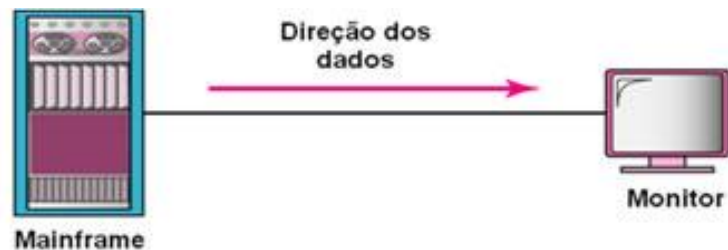
O transmissor é o dispositivo responsável por enviar a mensagem de dados como um computador, celular, televisão, câmera de vídeo, etc. Até a um receptor.

Meio é o caminho físico por onde viaja uma mensagem direcionada do transmissor até ao receptor.

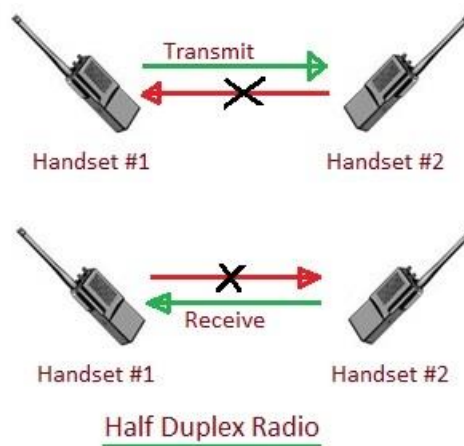
Receptor: é o dispositivo que recebe a mensagem. Pode ser um computador, celular, uma câmera, etc.

A comunicação entre dois dispositivos pode ocorrer de três maneiras: simplex, half-duplex ou full-duplex.

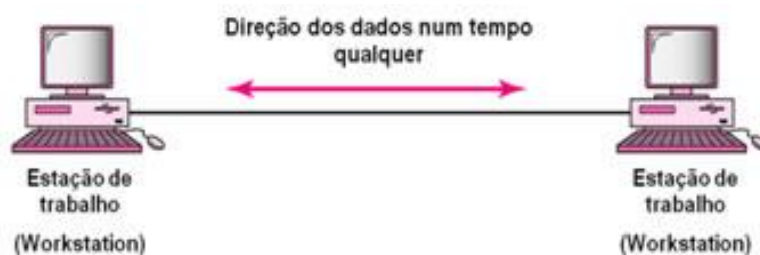
Simplex: a comunicação é unidirecional, somente um dispositivo no link é capaz de transmitir; logo o outro só será capaz de receber.



Half-duplex: transmissão bidirecional. quando um dispositivo está transmitindo o outro está recebendo e vice-versa, mas não podem transmitir e receber ao mesmo tempo.



Full-duplex: transmissão bidirecional. ambas estações podem transmitir e receber simultaneamente.



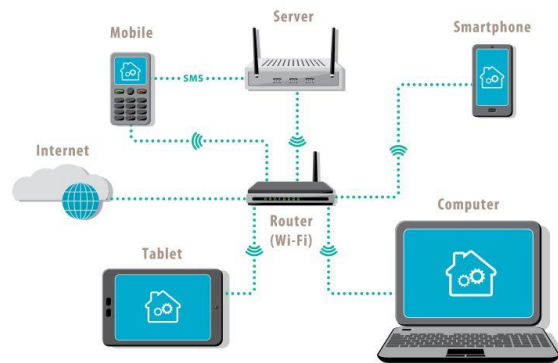
“rede de computadores é um conjunto de equipamentos interligados de maneira a trocarem informações e compartilhem recursos, como arquivos de dados gravados, impressoras, modems, softwares e outros equipamentos”.

Uma das formas mais utilizadas para a classificação das redes é em relação à abrangência geográfica:

LAN (Local Area Network): é a chamada rede doméstica. A distância entre os módulos é no máximo de alguns quilômetros.

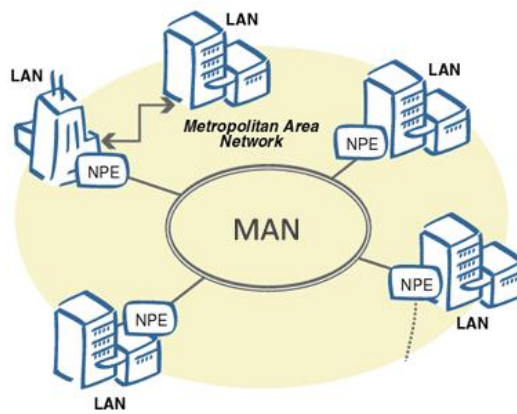
- Alta velocidade e baixa latência

“LAN é a tecnologia que apresenta uma boa resposta para interligação de dispositivos com distâncias relativamente pequenas e com uma largura de banda considerável.”



MAN (Metropolitan Area Network) atingem distâncias maiores e podem conectar cidades ou até um Estado inteiro provendo a interligação das redes locais em uma área metropolitana.

- Alta velocidade e baixa latência



WAN (Wide Area Network): podem conectar todo um país ou até diferentes pontos pelo mundo, sendo necessário quando distâncias envolvidas na interligação dos computadores são superiores a uma região metropolitana.

- baixa velocidade e alta latência



A topologia de rede é o modo no qual o meio de rede está conectado aos computadores e outros componentes de rede, podendo ser descritos como estrutura física ou lógica.

“Podemos dizer que a topologia física de uma rede local compreende os enlaces físicos de ligação dos elementos computacionais da rede, enquanto a topologia lógica da rede se refere à forma através da qual o sinal é efetivamente transmitido entre um computador e outro.”

Plano de desenvolvimento da aplicação

Ao pensar como iniciar, desenvolver, finalizar e entregar de forma íntegra e discriminada um projeto, ideias e subjeções veem à tona, trazendo inúmeras possibilidades, um tanto quanto inusitadas, difíceis, e as mais prováveis, as alternativas simples, porém conclusivas. É de modo, macro, limitante escolher uma que encaixe nos trilhos da incongruência argumentativa. Contudo, todavia, quando já se tem fatores, intermediariamente, requisitados, propostos e obrigatórios, a tarefa de alcançar uma solução viável fica mais fácil, compondo em sua natureza elementos que agem de maneira manuteníveis e refutáveis.

Sendo elementos que em sua integridade possuem propriedades básicas de uma boa resposta para o propósito de desenvolvimento, pode-se dizer que, contém, ao menos, as estruturas bases para a inicialização do projeto, seja ele qual for e seu afim. Desta forma, tomado como molde, toda essa visão de planejamento, o curso de desdobramento desta aplicação foi iniciado dentro da perspectiva de um propósito, um ideal (problema solúvel) e encapsulado de ferramentas (obrigatórias e auxiliares), além de obter, é claro, um toque de autenticidade e tomada própria de decisão.

No desenvolvimento da aplicação foi utilizado e absorvido ao longo do período inteiro de construção projeto de web-scraping diversas tecnologias e ferramentas que possibilitaram a manipulação dos elementos e demonstração visual do jeito correto que foi proposto e idealizado para a aplicação. Conforme foi dito, ao decorrer do curso de inicialização à finalização do projeto, foi inserido lentamente elementos, ferramentas, bibliotecas, tecnologias que ajudassem a moldar o resultado final esperado. Sendo assim, foi ajustado a implementação de módulos bases que facilitassem a continuação da aplicação, alguns deles são discriminados a seguir.

Como parte fundamental do projeto, a escolha por algumas tecnologias foram sucintas à utilidade, como exemplo, a linguagem python proporcionando em sua natureza, volatilidade e liberdade para construir e mudar constantemente o ecossistema do ambiente de dev., além de oferecer módulos que trabalhem e

manipulem os dados recebidos e enviados com eficiência. Módulos como `beautifulsoup`; `requests`; `pandas`; `json`, entre outros, são bibliotecas de suma importância para o favorecimento de uma boa solução. De maneira similar ao ambiente python e seus complementares (módulos), outra linguagem de programação, mas agora voltada ao lado servidor, foi relevante para o desenvolvimento, o php trazendo sua utilidade para a web, proporcionou um atributo de separação e desenvolvimento paralelo, oportuno na questão de poder trabalhar com o back-end e front-end ao mesmo tempo, fato esse, que só houve, graças ao xampp, unindo esses arcos de códigos-fontes limpos e coesos. Finalizando, por cima da arte visual e da persuasão, o bootstrap acoplado com css e javascript, fidelizando uma interface bem distribuídas e tentando ao máximo, utilizando técnicas de comportamento visual, atingir ao máximo as heurísticas de Nielsen.

Abordando um pouco sobre as tecnologias utilizadas:

Python

Python é uma linguagem de programação de alto nível — ou High Level Language —, dinâmica, interpretada, modular, multiplataforma e orientada a objetos — uma forma específica de organizar softwares onde, a grosso modo, os procedimentos estão submetidos às classes, o que possibilita maior controle e estabilidade de códigos para projetos de grandes proporções.

Por ser uma linguagem de sintaxe relativamente simples e de fácil compreensão, ganhou popularidade entre profissionais da indústria tecnológica que não são especificamente programadores, como engenheiros, matemáticos, cientistas de dados, pesquisadores e outros.

Um de seus maiores atrativos é possuir um grande número de bibliotecas, nativas e de terceiros, tornando-a muito difundida e útil em uma grande variedade de setores dentro de desenvolvimento web, e também em áreas como análise de dados, machine learning e IA.

Beautiful Soup

Beautiful Soup é um pacote Python para análise de documentos HTML e XML (incluindo marcação malformada, isto é, tags não fechadas, assim chamadas após tag soup). Ele cria uma árvore de análise para páginas analisadas que pode ser usada para extrair dados de HTML que é útil para web scraping.

O Beautiful Soup foi iniciado por Leonard Richardson, que continua a contribuir com o projeto, e é adicionalmente suportado pelo Tidelift, uma assinatura paga para manutenção de código aberto.

Requests

Requests é uma biblioteca Python HTTP, lançada sob a licença Apache 2.0. O objetivo do projeto é tornar as solicitações HTTP mais simples e mais amigáveis. A versão atual é 2.25.0

Requests é uma das bibliotecas python mais populares que não está incluída com o python. Foi proposto que as solicitações sejam distribuídas com o python por padrão.

Requests é uma biblioteca HTTP licenciada Apache2, escrita em Python, para seres humanos.

O módulo urllib2 padrão do Python fornece a maioria dos recursos HTTP de que você precisa, mas a API está totalmente corrompida. Foi construído para uma época diferente - e uma rede diferente. Requer uma enorme quantidade de trabalho (até mesmo substituições de métodos) para realizar as tarefas mais simples.

JSON

JSON é basicamente um formato leve de troca de informações/dados entre sistemas. Mas JSON significa JavaScript Object Notation, ou seja, só posso usar com JavaScript correto? Na verdade, não e alguns ainda caem nesta armadilha.

O JSON além de ser um formato leve para troca de dados é também muito simples de ler. Mas quando dizemos que algo é simples, é interessante compará-lo com algo mais complexo para entendermos tal simplicidade, não é? Neste caso podemos comparar o JSON com o formato XML.

```

1 <note>
2 <to>Tove</to>
3 <from>Jani</from>
4 <heading>Reminder</heading>
5 <body>Don't forget me this weekend!</body>
6 </note>

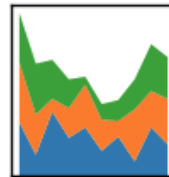
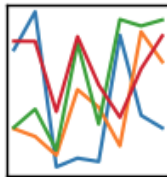
```

```

1 {
2   "id":1,
3   "nome":"Alexandre Gama",
4   "endereco":"R. Qualquer"
5 }

```

pandas

$$y_{it} = \beta' x_{it} + \mu_i + \epsilon_{it}$$


Pandas é uma biblioteca para manipulação e análise de dados, escrita em Python. Essa é a biblioteca perfeita para iniciar suas análises exploratórias de dados, pois ela nos permite ler, manipular, agregar e plotar os dados em poucos passos. Em uma simples definição, DataFrame é como se fosse uma planilha de Excel ou uma tabela de banco de dados. É composto por colunas, linhas e índice. Quando nós lemos algum arquivo de dados, ele se torna um DataFrame para o Pandas.

Pandas não anda sozinho, segue algumas bibliotecas que nos ajudam a refinar nossas análises:

seaborn



Seaborn, Estatística e visualizações.



NumPy, Bibliotecas para funções matemáticas.



Matplotlib, Visualizações de dados.

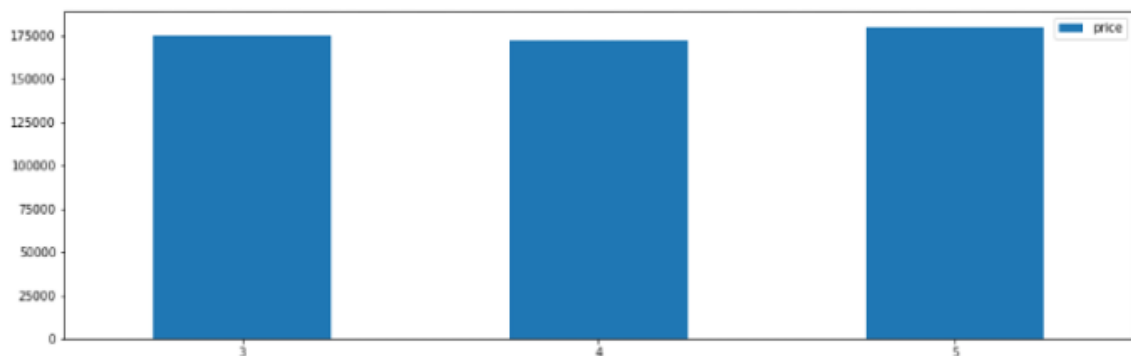
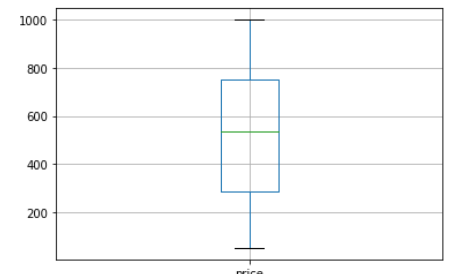
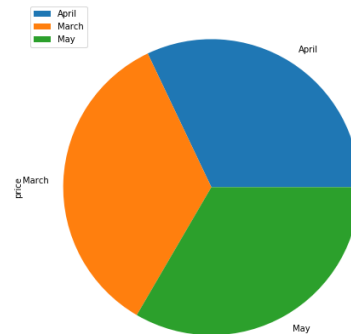
Scikit-Learn, Classificação, clusterização e regressão.

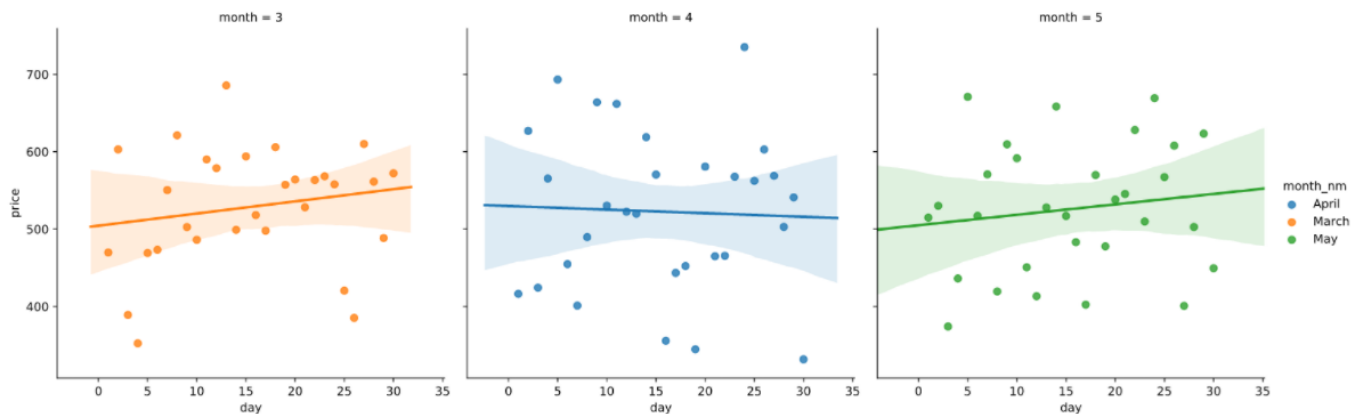
Alguns exemplos de demonstração visual de dados:

Antes

order_id	date	price	month_m	day
0	1 2020-03-01	501.17	March	1
1	2 2020-03-01	731.26	March	1
2	3 2020-03-01	596.96	March	1
3	4 2020-03-01	185.23	March	1
4	5 2020-03-01	722.01	March	1

Depois				
month_m	April	March	May	
day				
1	416.360000	469.784615	514.861250	
2	626.921429	602.972727	530.154286	
3	424.375714	389.098750	374.160000	
4	565.341333	352.375714	436.177778	
5	693.210000	469.860000	670.887778	
6	454.779167	473.246667	517.853636	
7	401.183333	550.424000	570.726429	
8	489.657500	621.204286	419.441338	
9	663.880667	502.648750	689.596000	
10	530.310769	406.150000	591.529891	





Média do preço, Dias do mês x Mês Regressão Linear em linha

Flask

Flask é um pequeno framework web escrito em Python. É classificado como um microframework porque não requer ferramentas ou bibliotecas particulares, mantendo um núcleo simples, porém, extensível. Não possui camada de abstração de banco de dados, validação de formulário ou quaisquer outros componentes onde bibliotecas de terceiros pré-existentes fornecem funções comuns. No entanto, o Flask oferece suporte a extensões que podem adicionar recursos do aplicativo como se fossem implementados no próprio Flask. Existem extensões para mapeadores objeto-relacional, validação de formulário, manipulação de upload, várias tecnologias de autenticação aberta e várias ferramentas comuns relacionadas ao framework.

```
from flask import Flask
app = Flask(__name__)

@app.route("/")
def index():
    return 'Bem-vindo a TreinaWeb!'

if __name__ == "__main__":
    app.run()
```

XAMPP

XAMPP é um pacote com os principais servidores de código aberto do mercado, incluindo FTP, banco de dados MySQL e Apache com suporte as linguagens PHP e Perl. De plataforma, software livre, que consiste principalmente na base de dados MySQL, o qual foi substituído pelo MariaDB(embora ainda seja utilizado MySql em algumas versões), o servidor web Apache e os interpretadores para linguagens de script: PHP e Perl,além de um cliente FTP. O nome provém da abreviação de X (para qualquer dos diferentes sistemas operativos), Apache, MariaDB, PHP, Perl. É um método que torna extremamente fácil para os desenvolvedores a criar um servidor web local para fins de teste.

PHP

O PHP (um acrônimo recursivo para PHP: Hypertext Preprocessor) é uma linguagem de script open source de uso geral, muito utilizada, e especialmente adequada para o desenvolvimento web e que pode ser embutida dentro do HTML. Em vez de muitos comandos para mostrar HTML (como acontece com C ou Perl), as páginas PHP contêm HTML em código mesclado que faz "alguma coisa" (neste caso, mostra "Olá, eu sou um script PHP!"). O código PHP é delimitado pelas instruções de processamento (tags) de início e fim `<?php` e `?>` que permitem que você entre e saia do "modo PHP".

O que distingue o PHP de algo como o JavaScript no lado do cliente é que o código é executado no servidor, gerando o HTML que é então enviado para o navegador. O navegador recebe os resultados da execução desse script, mas não sabe qual era o código fonte. Você pode inclusive configurar seu servidor web para processar todos os seus arquivos HTML com o PHP, e então não há como os usuários dizerem o que você tem na sua manga.

A melhor coisa em usar o PHP é que ele é extremamente simples para um iniciante, mas oferece muitos recursos avançados para um programador profissional. Não tenha medo de ler a longa lista de recursos do PHP. Pode entrar com tudo, o mais rápido que puder, e começar a escrever scripts simples em poucas horas.

PostgreSQL

O PostgreSQL é uma ferramenta que atua como sistema de gerenciamento de bancos de dados relacionados. Seu foco é permitir implementação da linguagem SQL em estruturas, garantindo um trabalho com os padrões desse tipo de ordenação dos dados.

Nos últimos anos, o uso desse sistema tem crescido consideravelmente, muito por conta de sua praticidade e pela sua alta compatibilidade com diferentes padrões de linguagem. Seu funcionamento é desenvolvido para ser, na prática de grande suporte para que qualquer trabalho seja feito sem maiores dificuldades.

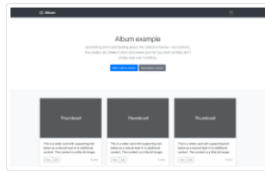
Um de seus pontos principais é sua adequação em padrões de conformidade, ajudando a construir bancos de dados otimizados. Nesse trabalho, com suas qualidades principais, o PostgreSQL ajuda a armazenar informações de forma segura e, se necessário, restaurá-las sempre que houver solicitação de outras aplicações integradas.

O PostgreSQL é um sistema que lida bem com altos volumes de solicitações e com cargas de trabalho grandes, ou seja, funciona muito bem para sites com intensidade de acesso. E-commerces famosos, por exemplo, é um ótimo exemplo de estrutura que precisa desse sistema para ter um desempenho otimizado, devido ao alto número de acessos simultâneos recebidos.

Bootstrap

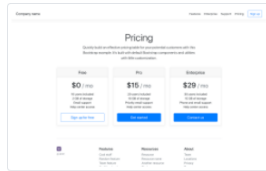
O Bootstrap nada mais é que um framework Front End utilizados por milhares de desenvolvedores web pelo mundo. A partir dele, muitas etapas dos desenvolvimentos web se tornam mais rápidas e dinâmicas, pois já trazem consigo diversos elementos prontos e estilizados.

Ele é uma ferramenta gratuita para desenvolvimento HTML, CSS e JS. Portanto, é uma ferramenta de código fonte aberto para todos. Contudo, hoje em dia, é o framework Front End mais conhecido e mais utilizado pelo mundo.



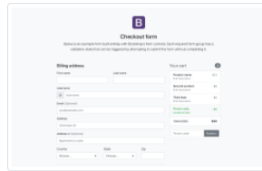
Album

Template de página simples para galerias de fotos, portfólios, etc.



Pricing

Exemplo de página de preços feita com card e personalizados cabeçalhos e rodapé.



Checkout

Formulário de checkout customizado para mostrar nossos componentes de formulário e seus recursos de validação.



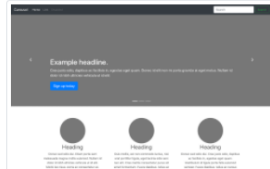
Product

Página de marketing focada em produto e bem trabalhada com o grid e imagens.



Cover

Um template para construir simples e lindas home pages.



Carousel

Customize a barra de navegação o e carousel, então, adicione alguns novos componentes.



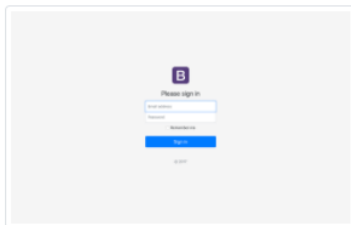
Blog

Blog no estilo de revista com cabeçalho, navegação e conteúdo destacado.



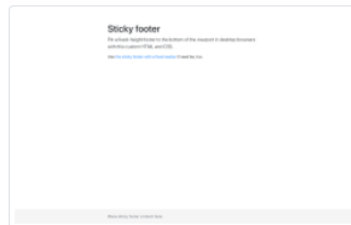
Dashboard

Painel de controle administrativo básico com sidebar fixo e barra de navegação.



Sign-in

Layout e design de formulário customizado para um simples formulário de login.



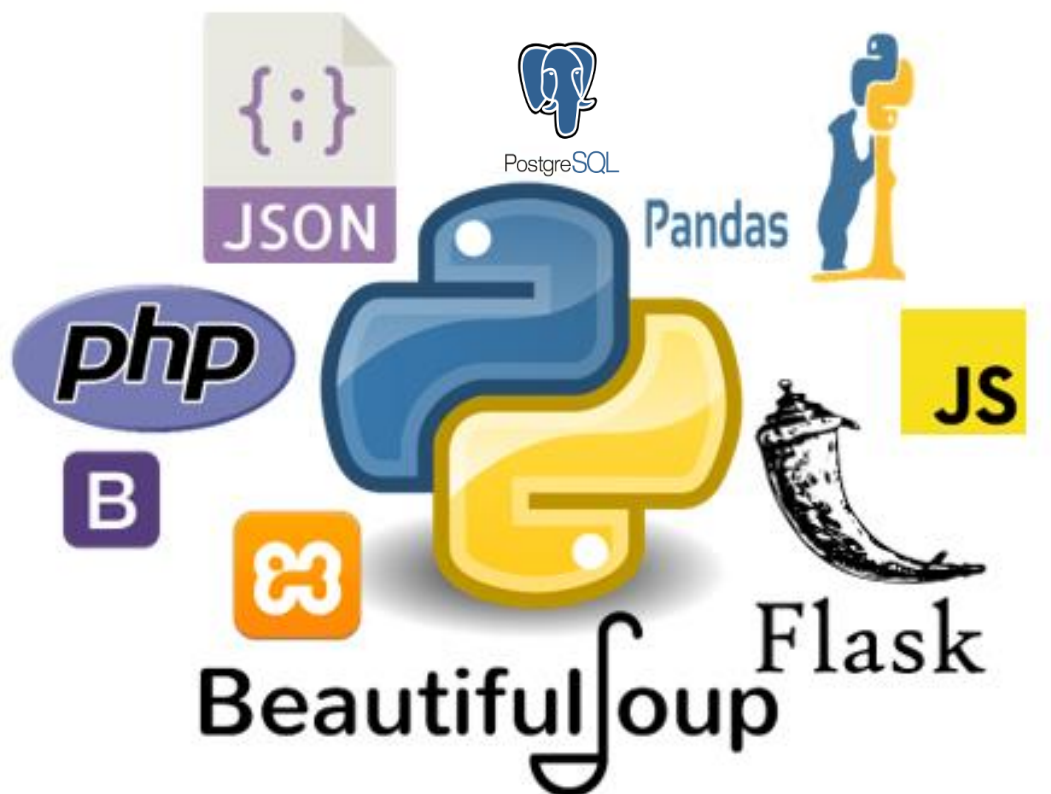
Sticky footer

Grude o rodapé a parte inferior da viewport, quando o conteúdo da página é curto.



Sticky footer navbar

Grude o rodapé a parte inferior da viewport e uma barra de navegação no topo.

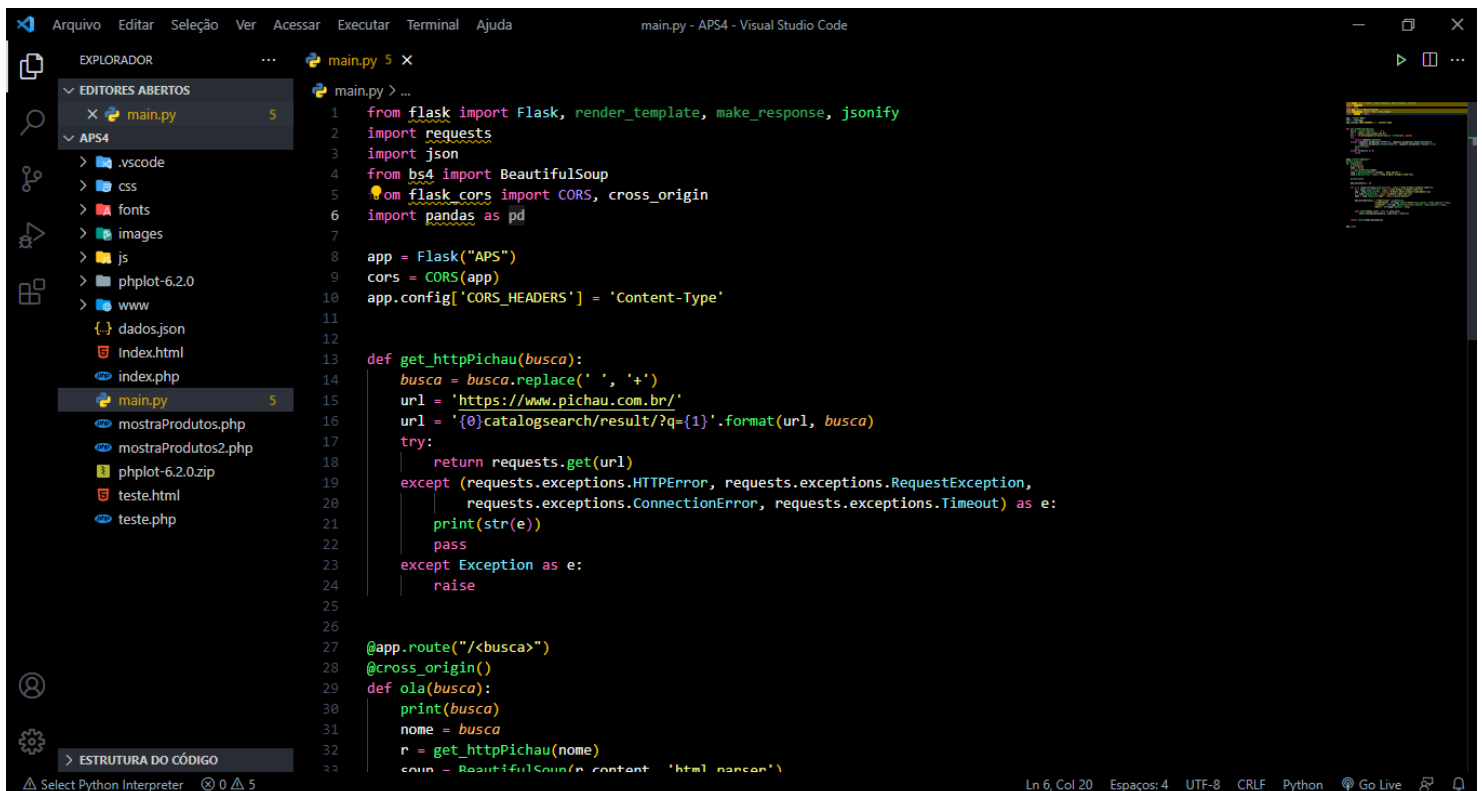


Projeto

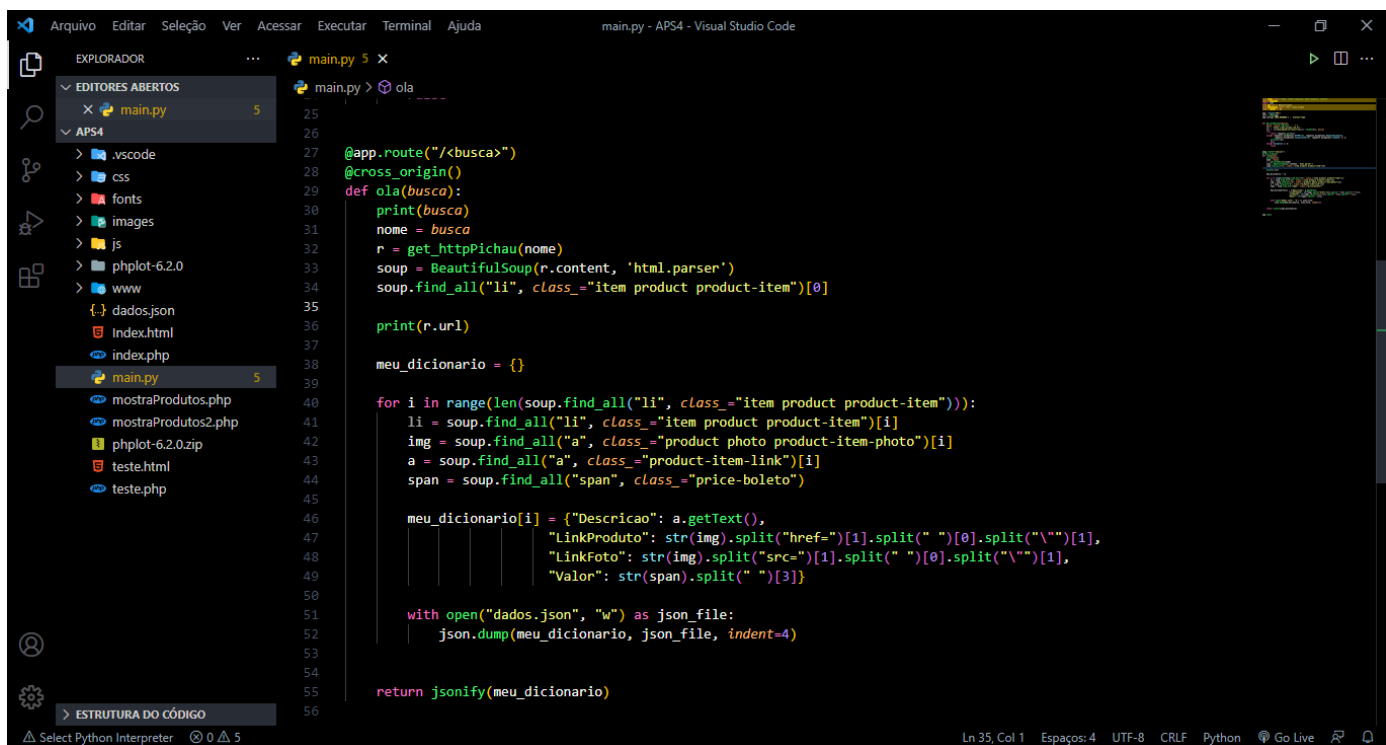
No desenvolvimento do projeto de web-scraping, foram acoplados módulos de forma separados, porém, íntegros, isto é, as tecnologias e ferramentas utilizadas são relacionáveis uma com a outra de maneira que não interferem na programação, adição, exclusão, etc. de cada uma.

Quatro pilares fundamentais foram implementados: python para trabalhar com o back-end e regras de negócios; banco de dados e json para guardar as informações e dados do projeto; e para finalizar o php para manipular as informações e dados do lado do servidor. Entretanto, essas são as tecnologias mais relevantes do projeto, pois outras um pouco menos extensas são usadas também no código-fonte. Objetos como módulos de python, bootstrap para visualização de conteúdo html e javascript, xampp para unificar php com banco de dados, entre outros são exemplos de ideias desenvolvidas no web-scraping.

A estrutura foi planejada em pastas para distinguir as diferentes tecnologias, como php, python, flask, json, etc. Segue a exploração da estrutura.



```
1 from flask import Flask, render_template, make_response, jsonify
2 import requests
3 import json
4 from bs4 import BeautifulSoup
5 from flask_cors import CORS, cross_origin
6 import pandas as pd
7
8 app = Flask("APS4")
9 cors = CORS(app)
10 app.config['CORS_HEADERS'] = 'Content-Type'
11
12 def get_httpPichau(busca):
13     busca = busca.replace(' ', '+')
14     url = 'https://www.pichau.com.br/'
15     url = '{0}catalogsearch/result/?q={1}'.format(url, busca)
16     try:
17         return requests.get(url)
18     except (requests.exceptions.HTTPError, requests.exceptions.RequestException,
19           requests.exceptions.ConnectionError, requests.exceptions.Timeout) as e:
20         print(str(e))
21         pass
22     except Exception as e:
23         raise
24
25 @app.route("/<busca>")
26 @cross_origin()
27 def ala(busca):
28     print(busca)
29     nome = busca
30     r = get_httpPichau(nome)
31     soup = BeautifulSoup(r.content, 'html.parser')
```



```
25
26
27 @app.route("/<busca>")
28 @cross_origin()
29 def ola(busca):
30     print(busca)
31     nome = busca
32     r = get_httpPichau(nome)
33     soup = BeautifulSoup(r.content, 'html.parser')
34     soup.find_all("li", class_="item product product-item")[0]
35
36     print(r.url)
37
38     meu_dicionario = {}
39
40     for i in range(len(soup.find_all("li", class_="item product product-item"))):
41         li = soup.find_all("li", class_="item product product-item")[i]
42         img = soup.find_all("a", class_="product photo product-item-photo")[i]
43         a = soup.find_all("a", class_="product-item-link")[i]
44         span = soup.find_all("span", class_="price-boleto")
45
46         meu_dicionario[i] = {"Descricao": a.getText(),
47                             "LinkProduto": str(img).split("href=")[1].split(" ")[0].split("\")[1],
48                             "LinkFoto": str(img).split("src=")[1].split(" ")[0].split("\")[1],
49                             "Valor": str(span).split(" ")[3]}
50
51     with open("dados.json", "w") as json_file:
52         json.dump(meu_dicionario, json_file, indent=4)
53
54     return jsonify(meu_dicionario)
55
56
```

O arquivo main.py desenvolvido em python foi estruturado para trabalhar e manipular com a parte de back-end e regras de negócios do web-scraping, como interagir com a página requisitada, converter dos dados html para dados mais visualmente entendíveis, como o json. Nele estão os módulos que manipulam a web, varrendo os dados e jogando para as outras tecnologias, como o framework flask utilizado para prover modulo simples para o desenvolvimento web, o request que possibilitar o acesso as paginas web e seus dados, o beautifulsoup que permiti ajustar e modelar as informações de forma mais bonita e bem apresentável, o pandas que também ajuda no processo de tratamento e visualização dos dados, trazendo diversas funções e ferramentas de análise estatística e o modulo json para guardar, transferir e representar os dados recebidos.

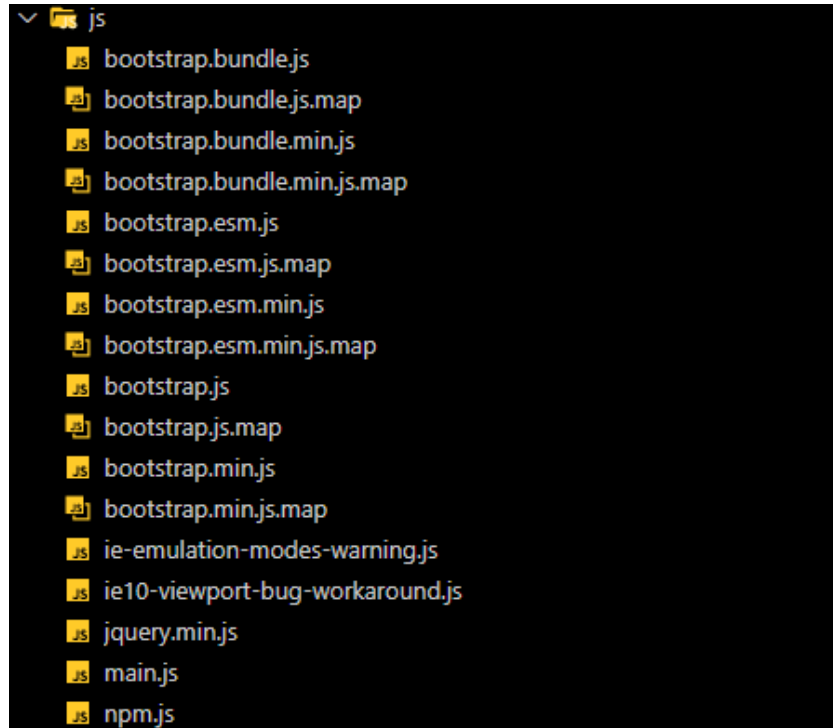
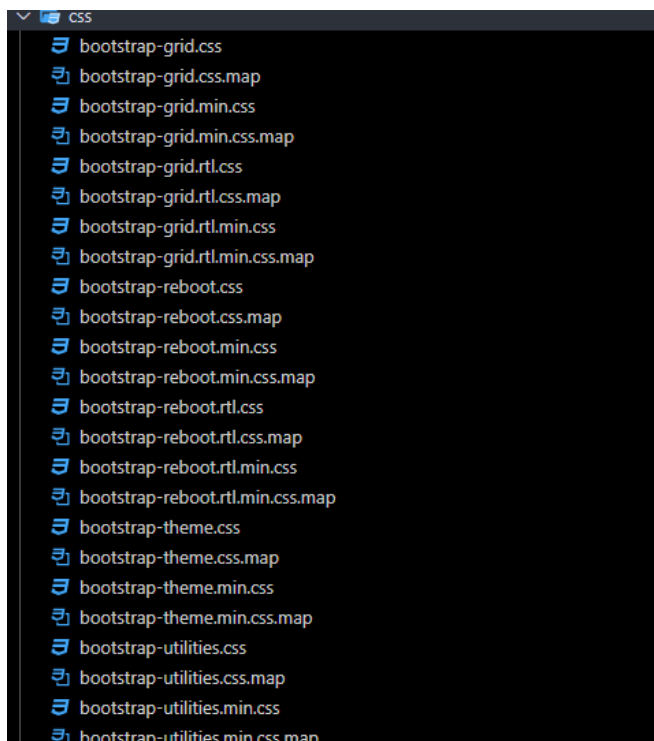
```
1 <!DOCTYPE html>
2 <html Lang="en">
3 <head>
4 <meta charset="utf-8">
5 <meta http-equiv="X-UA-Compatible" content="IE=edge">
6 <meta name="viewport" content="width=device-width, initial-scale=1">
7 <!-- The above 3 meta tags *must* come first in the head, any other head content must come *after* these tags -->
8
9 <meta name="description" content="">
10 <meta name="author" content="">
11 <link rel="icon" href="images/icon.png">
12
13 <title>Hardware Finder</title>
14
15 <!-- Bootstrap core CSS -->
16 <link href="css/bootstrap.css" rel="stylesheet">
17
18 <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
19 <link href="css/ie10-viewport-bug-workaround.css" rel="stylesheet">
20
21 <!-- Custom styles for this template -->
22 <link href="css/jumbotron.css" rel="stylesheet">
23
24 <!-- Just for debugging purposes. Don't actually copy these 2 lines! -->
25 <!--[if lt IE 9]><script src="../../assets/js/ie8-responsivse-file-warning.js"></script><![endif]>-->
26 <script src="js/ie-emulation-modes-warning.js"></script>
27
28 <!-- HTML5 shim and Respond.js for IE8 support of HTML5 elements and media queries -->
29 <!--[if lt IE 9]>
30 <script src="https://oss.maxcdn.com/html5shiv/3.7.3/html5shiv.min.js"></script>
31 <script src="https://oss.maxcdn.com/respond/1.4.2/respond.min.js"></script>
32 <![endif]>-->
33 </head>
```

```
102 <?php
103
104
105 $a = 0;
106
107 for ($i = 1; $i <= 3; $i++) {
108     if($i == 1){
109         $conteudo = file_get_contents("http://127.0.0.1:5000/".str_replace(" ", "%20", $buscar));
110
111         $dados = json_decode($conteudo, False);
112
113         foreach($dados as $chave => $valor)
114         {
115             <div class="col-sm-4" border="1">
116                 <p style="font-weight:500;">Descrição:<?php echo $valor->{'Descricao'};></p>
117                 <p>preço:<?php echo $valor->{'Valor'};></p>
118                 <img src=<?php echo $valor->{'LinkFoto'};> alt="HTML5 Icon" style="width:128px;height:128px;
119                 ">
120                 <a href=<?php echo $valor->{'LinkProduto'};>><button class="btn_aceso">Conferir!</button></a>
121                 <!--<p><a class="btn btn-default" href="#" role="button">View details &raquo;</a></p>-->
122             </div>
123             <?php
124
125             $array[$a] = [
126                 "Descricao" => $valor->{'Descricao'},
127                 "Valor" => $valor->{'Valor'},
128                 "LinkFoto" => $valor->{'LinkFoto'},
129                 "LinkProduto" => $valor->{'LinkProduto'},
130             ];
131
132             $a++;
133         }
134     }
135 }
```

Com os arquivos php é possível realizar conexão com o banco de dados, inserindo, deletando e alterando os dados salvos nas tabelas. Dados esses que são recebidos do python e resolvidos no lado do servidor, através do php.

```
{
  "0": {
    "Descricao": "\nSSD WD Blue 250gb M.2 2280, WDS250G2B0B ",
    "LinkProduto": "https://www.pichau.com.br/hardware/ssd/ssd-wd-blue-250gb-m-2-2280-wds250g2b0b",
    "LinkFoto": "https://media.pichau.com.br/media/catalog/product/cache/cd2e75e2636e632c04e681018268259c/w/d/wds250g2b0b.jpg",
    "Valor": "R$419,01</span>"
  },
  "1": {
    "Descricao": "\nSSD Gigabyte 1TB 2.5\" SATA 6GB/s, GP-GSTFS31100TNTD ",
    "LinkProduto": "https://www.pichau.com.br/hardware/ssd/ssd-gigabyte-1tb-2-5-sata-6gb-s-gp-gstfs31100tntd",
    "LinkFoto": "https://media.pichau.com.br/media/catalog/product/cache/cd2e75e2636e632c04e681018268259c/g/p/gp-gstfs31100tntd3212.jpg",
    "Valor": "R$419,01</span>"
  },
  "2": {
    "Descricao": "\nSSD Pichau Gaming 256GB 2.5\" Sata 6GB/s, PG256X ",
    "LinkProduto": "https://www.pichau.com.br/hardware/ssd/ssd-pichau-gaming-256gb-2-5-sata-6gb-s-pg256x",
    "LinkFoto": "https://media.pichau.com.br/media/catalog/product/cache/cd2e75e2636e632c04e681018268259c/p/g/pg256x1.jpg",
    "Valor": "R$419,01</span>"
  },
  "3": {
    "Descricao": "\nSSD WD Black SN750 1TB M.2 2280 NVMe, WDS100T3X0C ",
    "LinkProduto": "https://www.pichau.com.br/hardware/ssd/ssd-wd-black-1tb-m-2-2280-sn750-nvme-wds100t3x0c",
    "LinkFoto": "https://media.pichau.com.br/media/catalog/product/cache/cd2e75e2636e632c04e681018268259c/w/d/wds100t3x0c1.jpg",
    "Valor": "R$419,01</span>"
  },
  "4": {
    "Descricao": "\nSSD Team Group GX2 128GB 2.5\" Sata 6GB/s, T253X2128G0C101 ",
    "LinkProduto": "https://www.pichau.com.br/hardware/ssd/ced-tasm-anna-av3-128gb-2-5-sata-6gb-ct253v2128n0r101"
  }
}
```

Esse é um exemplo do arquivo json e sua forma de representação dos dados, usando combinações de chave-valor.



Segue as pastas com os arquivos de css e javascript em bootstrap.

Relatório das linhas de código

```
main.py 6 X
main.py > ...
1  from flask import Flask, render_template, make_response, jsonify
2  import requests
3  import json
4  from bs4 import BeautifulSoup
5  from flask_cors import CORS, cross_origin
6  import pandas as pd
7  from pytz import exceptions
8
9  app = Flask(__name__)
10 cors = CORS(app)
11 app.config['CORS_HEADERS'] = 'Content-Type'
12
13
14 def get_httpPichau(busca):
15     busca = busca.replace(' ', '+')
16     url = 'https://www.pichau.com.br/'
17     url = '{0}catalogsearch/result/?q={1}'.format(url, busca)
18     try:
19         return requests.get(url)
20     except (requests.exceptions.HTTPError, requests.exceptions.RequestException,
21             requests.exceptions.ConnectionError, requests.exceptions.Timeout) as e:
22         print(str(e))
23         pass
24     except Exception as e:
25         raise
26
27 def get_httpKabum(busca):
28     print(busca)
29     busca = busca.replace(' ', '+')
30     url = 'https://www.kabum.com.br/'
31     url = '{0}cgi-local/site/listagem/listagem.cgi?string={1}'.format(url, busca)
32     print(url)
```

```
main.py 6 X
main.py > get_httpPichau
33     try:
34         return requests.get(url)
35     except (requests.exceptions.HTTPError, requests.exceptions.RequestException,
36             requests.exceptions.ConnectionError, requests.exceptions.Timeout) as e:
37         print(str(e))
38         pass
39     except Exception as e:
40         raise
41
42 @app.route("/<busca>")
43 @cross_origin()
44 def ola(busca):
45     print(busca)
46     nome = busca
47     r = get_httpPichau(nome)
48     soup = BeautifulSoup(r.content, 'html.parser')
49
50
51     print(r.url)
52
53     meu_dicionario = {}
54
55     for i in range(len(soup.find_all("li", class_="item product product-item"))):
56         li = soup.find_all("li", class_="item product product-item")[i]
57         img = soup.find_all("a", class_="product photo product-item-photo")[i]
58         a = soup.find_all("a", class_="product-item-link")[i]
59         span = li.find_all("span", class_="price-boleto")
60
61         valor = ""
62         try:
63             valor = str(span).split(" ")[3]
64         except:
```

```

main.py 6 X
main.py > get_httpPichau
64         except:
65             valor = str("PRODUTO INDISPONÍVEL")
66
67             meu_dicionario[i] = {"Descricao": a.getText(),
68                                 "LinkProduto": str(img).split("href=")[1].split(" ")[0].split("\")[1],
69                                 "LinkFoto": str(img).split("src=")[1].split(" ")[0].split("\")[1],
70                                 "Valor": valor}
71
72
73
74         return jsonify(meu_dicionario)
75
76 @app.route("/kabum/<busca>")
77 @cross_origin()
78 def olaa(busca):
79     print(busca)
80     nome = busca
81
82     meu_dicionario = {}
83
84     r = get_httpKabum(nome)
85     soup = BeautifulSoup(r.content, 'html.parser')
86
87     try:
88         return str(soup).split("const listagemDados =")[1].split("const listagemErro =")[0]
89     except:
90         return "500"
91
92 @app.route("/magalu/<busca>")
93 @cross_origin()
94 def olaaa(busca):
95     print(busca)

```

```

main.py 6 X
main.py > get_httpPichau
96     nome = busca
97
98     url = "https://www.magazineluiza.com.br/busca/"+busca.replace(" ", "%20")
99     r = get_httpPichau(nome)
100
101     url = requests.get(url)
102     soup = BeautifulSoup(url.content, 'html.parser')
103
104     h3 = soup.find_all("h3", class_="productTitle")
105     img = soup.find_all("img", class_="product-image")
106     span = soup.find_all("span", class_="price")
107     a = soup.find_all("a", class_="product-li")
108
109     meu_dicionario = {}
110
111     for i in range(len(h3)):
112         try:
113             valor = str(span[i].getText()).replace(" ", "").split("R$")[1].split("\n")[0]
114         except:
115             valor = "PRODUTO INDISPONÍVEL"
116
117         meu_dicionario[i] = {"Descricao": h3[i].getText(),
118                             "LinkProduto": str(a).split("href=\"")[i + 1].split("\")[0],
119                             "LinkFoto": str(img[i]).split("data-original=\"")[1].split("\")[0],
120                             "Valor": valor}
121
122     return jsonify(meu_dicionario)
123
124
125 app.run()

```



```
mostraProdutos.php X Configurações
mostraProdutos.php > PHP Inteliphense > html > head > title
1 <!DOCTYPE html>
2 <html Lang="pt-br">
3 <head>
4 <meta charset="utf-8">
5 <meta http-equiv="X-UA-Compatible" content="IE=edge">
6 <meta name="viewport" content="width=device-width, initial-scale=1">
7
8 <!-- Página destinada para apresentados resultados das buscas feitas na API. -->
9 <meta name="APS" content="">
10 <link rel="icon" href="images/icon.png">
11
12 <title>Hardware Finder</title>
13
14 <!-- Bootstrap core CSS -->
15 <link href="css/bootstrap.css" rel="stylesheet">
16
17 <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
18 <link href="css/ie10-viewport-bug-workaround.css" rel="stylesheet">
19
20 <!-- Custom styles for this template -->
21 <link href="css/jumbotron.css" rel="stylesheet">
22
23 <!-- Just for debugging purposes. Don't actually copy these 2 lines! -->
24 <!--[if lt IE 9]><script src="../../assets/js/ie8-responsive-file-warning.js"><![endif]-->
25 <script src="js/ie-emulation-modes-warning.js"></script>
26
27 <!-- HTML5 shim and Respond.js for IE8 support of HTML5 elements and media queries -->
28 <!--[if lt IE 9]>
29 <script src="https://oss.maxcdn.com/html5shiv/3.7.3/html5shiv.min.js"></script>
30 <script src="https://oss.maxcdn.com/respond/1.4.2/respond.min.js"></script>
31 <![endif]-->
32 </head>
33 <?php
34 if(!@($conexao=pg_connect ("host=localhost dbname=aps port=5432 user=postgres password=adm")) {
35     print "Não foi possível estabelecer uma conexão com o banco de dados.";
36 }
37 $buscar = $_GET['buscar'];
38
```

```
mostraProdutos.php X
mostraProdutos.php > PHP Inteliphense > html > head > title
39
40 for ($i = 1; $i <= 2; $i++) {
41     if($i == 1){
42         $conteudo = file_get_contents("http://127.0.0.1:5000/".str_replace(" ", "%20", $buscar));
43
44         $dados = json_decode($conteudo, False);
45
46         $a = 0;
47
48         foreach($dados as $chave => $valor)
49         {
50
51             $array[$a] = [
52                 "Descricao" => $valor->{'Descricao'},
53                 "Valor" => floatval(str_replace("R$", "", str_replace(",",".", str_replace(".", "", $valor->{'Valor'})))),
54                 "LinkFoto" => $valor->{'LinkFoto'},
55                 "LinkProduto" => $valor->{'LinkProduto'},
56             ];
57             $a++;
58         }
59     }
60     elseif($i==2){
61         $conteudo = file_get_contents("http://127.0.0.1:5000/kabum/".str_replace(" ", "%20", $buscar));
62         $i = 3;
63
64         while ($i <= sizeof(explode("nome", $conteudo))) {
65             $array[$a] = [
66                 "Descricao" => str_replace("", "", explode("\n", explode("nome", $conteudo)[$i])[2]),
67                 "Valor" => floatval(explode(",", explode("preco_desconto":", $conteudo)[$i/3])[0]),
68                 "LinkFoto" => explode("\n", explode("img", $conteudo)[$i])[2],
69                 "LinkProduto" => "https://www.kabum.com.br/".explode("\n", explode("link_descricao", $conteudo)[$i/3])[2],
70             ];
71             $a++;
72             $i = $i + 3;
73         }
74     }
75 }
76
```

```

mostraProdutos.php X
mostraProdutos.php > PHP Intelephense > html > array_sort
77     function array_sort($array, $on, $order=SORT_ASC)
78     {
79         $new_array = array();
80         $sortable_array = array();
81
82         if (count($array) > 0) {
83             foreach ($array as $k => $v) {
84                 if (is_array($v)) {
85                     foreach ($v as $k2 => $v2) {
86                         if ($k2 == $on) {
87                             $sortable_array[$k] = $v2;
88                         }
89                     }
90                 } else {
91                     $sortable_array[$k] = $v;
92                 }
93             }
94
95             switch ($order) {
96                 case SORT_ASC:
97                     asort($sortable_array);
98                     break;
99                 case SORT_DESC:
100                     arsort($sortable_array);
101                     break;
102             }
103
104             foreach ($sortable_array as $k => $v) {
105                 $new_array[$k] = $array[$k];
106             }
107         }
108
109         return $new_array;
110     }
111 }
112
113 <body>
114 <nav class="navbar navbar-dark bg-dark" style="margin-top: -50px; background: linear-gradient(0deg, rgba(0,0,50,0.0) 35%, rgba(0,0,80,0.5) 50%, rgba(0,0,150,0.001) 100%); backdrop-filter: blur(10px);">

```

```

mostraProdutos.php X
mostraProdutos.php > PHP Intelephense > html > body > nav.navbar.navbar-dark.bg-dark
114 <div class="container">
115 <a href="index.php"></a>
116 <form class="form-group d-flex align-items-center">
117 <input class="col-xs-3 form-control m-1 container-md" type="text" placeholder="Buscar" name="buscar" aria-label="Search">
118 <input hidden type="text" id="inputBusca" class="form-control align-self-md-center" placeholder="relevante" name="ordena" value="relevante">
119 <input type="submit" class="btn btn-primary m-2"></input>
120 </form>
121 </div>
122 </div>
123 </nav>
124
125 <div class="container-busca">
126 <div class="pt-3"><h3>Resultados da Busca:</h3></div>
127 <div class="row justify-content-left">
128 <div class="btn-group" style="width: 250px;">
129 <select name="select" onchange="javascript:mostraAlerta(this);" type="button" class="btn btn-primary dropdown-toggle" data-toggle="dropdown" aria-haspopup="true" aria-expanded="false">
130 <option id="bt-ordenar" >Selecione...</option>
131 <option class="dropdown-item" value="valor1">Preço Crescente</option>
132 <option class="dropdown-item" value="valor2">Preço Decrescente</option>
133 <option class="dropdown-item" value="valor3">Mais Relevantes</option>
134 <option class="dropdown-item" value="valor4">A-Z</option>
135 <div class="dropdown-divider"></div>
136 <option class="dropdown-item" value="valor5">Z-A</option>
137 </select>
138 </div>
139 </div>
140 </div>
141
142 <div class="container">
143 <div class="row justify-content-center">
144 <?php
145
146     try {
147         $ordena = $_GET['ordena'];
148     } catch (Throwable $e) {
149         echo $e;
150     }

```



```

mostraProdutos.php X
mostraProdutos.php > PHP Intelephense > html > body > div.container
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259

```

```

    }
  }
</script>

</div> <!--/row-->
<footer class="panel-footer">
  <p>&copy; 2021 APS, SA.</p>
</footer>
</div> <!-- /container -->
<!-- Footer -->
<footer class="bg-dark text-center text-white">

  <div class="container p-4">
    <div class="text-center p-3">
      © 2021 Copyright:
      <a class="text-white" href="#">APS LVW</a>
    </div>
  </div>

</footer>
<!-- Footer -->

<!-- Bootstrap core JavaScript
===== -->
<!-- Placed at the end of the document so the pages load faster -->
<script src="https://ajax.googleapis.com/ajax/libs/jquery/1.12.4/jquery.min.js"></script>
<script>window.jQuery || document.write('<script src="js/jquery.min.js"></script>')</script>
<script src="js/bootstrap.min.js"></script>
<!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
<script src="js/ie10-viewport-bug-workaround.js"></script>
</body>
</html>

```

```

index.php X
index.php > html > head
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

```

```

<!DOCTYPE html>
<html lang="pt-br">
  <head>
    <meta charset="utf-8">
    <meta http-equiv="X-UA-Compatible" content="IE=edge">
    <meta name="viewport" content="width=device-width, initial-scale=1">

    <!-- Página destinada para apresentados resultados das buscas feitas na API. -->
    <meta name="APS" content="">
    <link rel="icon" href="images/icon.png">

    <title>Hardware Finder</title>

    <!-- Bootstrap core CSS -->
    <link href="css/bootstrap.css" rel="stylesheet">

    <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
    <link href="css/ie10-viewport-bug-workaround.css" rel="stylesheet">

    <!-- Estilo personalizado para esta página. -->
    <link href="css/jumbotron.css" rel="stylesheet">

    <script src="js/ie-emulation-modes-warning.js"></script>

  </head>
  <body style="width: 100%; height: 100%; padding-bottom:0px; padding-top:0px; background-image: url('images/pc.jpg'); background-size: 100% 100%; background-repeat: no-repeat;">
    <div style="margin-top: 0%; " class="frontPage align-middle">
      <form style="padding-top:15%; padding-bottom:30%; border: none; background: linear-gradient(0deg, rgba(30,30,30,0.5) 35%, rgba(0,0,0,0.1) 70%, rgba(150,150,150,0.0) 100% ); backdrop-filter: blur(10px);" class="
        "form-control form-control-lg frontPage" action="mostraProdutos.php?ordena=relevante">
        <div class="form-label-group justify-content-center align-self-center align-middle " >

```

```

index.php X
index.php > html > head
30 <div class="text-center mb-4">
31 <a class="justify-content-center align-self-center" href="index.php"></a>
32 </div>
33 <input type="text" id="inputBusca" class="form-control align-self-md-center" placeholder="Busca de
   Hardware" name="buscar" required="" autofocus="">
34 <input hidden type="text" id="inputBusca" class="form-control align-self-md-center" placeholder="relevante"
   name="ordena" text="relevante" value="relevante">
35 <div class="text-center mb-4" style="margin-top:20px;">
36 <button style="width: 150px;" class="btn-lg btn-primary m-2" type="submit">Buscar</button>
37 </div>
38 </div>
39 </form>
40 </div>
41
42 <!-- Bootstrap core JavaScript
43 ===== -->
44 <!-- Placed at the end of the document so the pages load faster -->
45 <script src="https://ajax.googleapis.com/ajax/libs/jquery/1.12.4/jquery.min.js"></script>
46 <script>window.jQuery || document.write('<script src="js/jquery.min.js"></script>')</script>
47 <script src="js/bootstrap.min.js"></script>
48 <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
49 <script src="js/ie10-viewport-bug-workaround.js"></script>
50 </body>
51 </html>

```

```

historicoProduto.php X
historicoProduto.php > html > head > link
1 <!DOCTYPE html>
2 <html lang="pt-br">
3 <head>
4 <meta charset="utf-8">
5 <meta http-equiv="X-UA-Compatible" content="IE=edge">
6 <meta name="viewport" content="width=device-width, initial-scale=1">
7
8 <!-- Página destinada para apresentados resultados das buscas feitas na API. -->
9 <meta name="APS" content="">
10 <link rel="icon" href="images/icon.png">
11
12 <title>Hardware Finder</title>
13
14 <!-- Bootstrap core CSS -->
15 <link href="css/bootstrap.css" rel="stylesheet">
16
17 <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
18 <link href="css/ie10-viewport-bug-workaround.css" rel="stylesheet">
19
20 <!-- Custom styles for this template -->
21 <link href="css/jumbotron.css" rel="stylesheet">
22
23 <!-- Just for debugging purposes. Don't actually copy these 2 lines! -->
24 <!--[if lt IE 9]><script src="../../assets/js/ie8-responsive-file-warning.js"></script><![endif]-->
25 <script src="js/ie-emulation-modes-warning.js"></script>
26
27 <!-- HTML5 shim and Respond.js for IE8 support of HTML5 elements and media queries -->
28 <!--[if lt IE 9]>
29 <script src="https://oss.maxcdn.com/html5shiv/3.7.3/html5shiv.min.js"></script>
30 <script src="https://oss.maxcdn.com/respond/1.4.2/respond.min.js"></script>
31 <![endif]-->
32 </head>

```

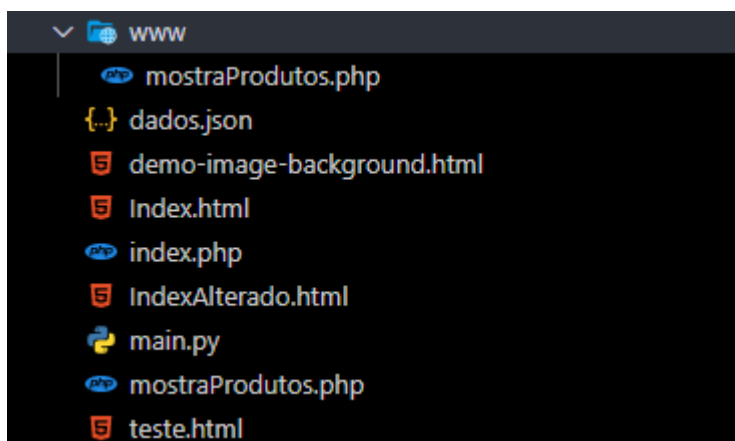
```
historicoProduto.php X
historicoProduto.php > html
33
34 <?php
35 if(!@($conexao=pg_connect ("host=localhost dbname=aps port=5432 user=postgres password=1234"))) {
36     print "Não foi possível estabelecer uma conexão com o banco de dados.";
37 }
38
39 $recebe = $_GET['recebe'];
40 $link = $_GET['link'];
41
42
43 $result = pg_query($conexao, "SELECT * FROM preco_produtos WHERE nomeproduto = '".rtrim($recebe)."' ORDER BY
data_preco");
44 ?>
45
46 <script type="text/javascript" src="https://www.gstatic.com/charts/loader.js"></script>
47 <script type="text/javascript">
48     google.charts.load('current', {'packages':['corechart']});
49     google.charts.setOnLoadCallback(drawChart);
50
51     function drawChart() {
52         var data = google.visualization.arrayToDataTable([
53             ['Data', 'Valores'],
54
55             <?php
56             while ($row = pg_fetch_assoc($result)) {
57                 $date = new DateTime($row['data_preco']);
58                 $a = strval($date->format('d/m/Y'));
59
60             ?>
61             [<?php echo '\'.$a.\'', <?php print_r(strval($row['preco'])); ?>],
62
63             <?php
```

```
historicoProduto.php X
historicoProduto.php > html > script > drawChart > options
63 <?php
64 }
65 ?>
66
67 ]);
68
69 var options = {
70     title: 'Histórico de Preços ',
71     curveType: 'function',
72     legend: { position: 'bottom' }
73 };
74
75 var chart = new google.visualization.LineChart(document.getElementById('curve_chart'));
76
77 chart.draw(data, options);
78 }
79 setTimeout(function() {
80     window.location.reload(1);
81 }, 6000);
82 function voltar() {
83     window.history.back();
84 }
85 </script>
86 <body>
87 <nav class="navbar navbar-dark bg-dark" style="margin-top: -50px; background: linear-gradient(0deg, rgba(0,0,50,
0.0) 35%, rgba(0,0,80,0.5) 50%, rgba(0,0,150,0.001) 100% ); backdrop-filter: blur(10px);">
88     <div class="container-busca">
89         <a href="index.php"></a>
90         <form class="form-group d-flex align-items-center">
91
92         </form>
93     </div>
94 </div>
```

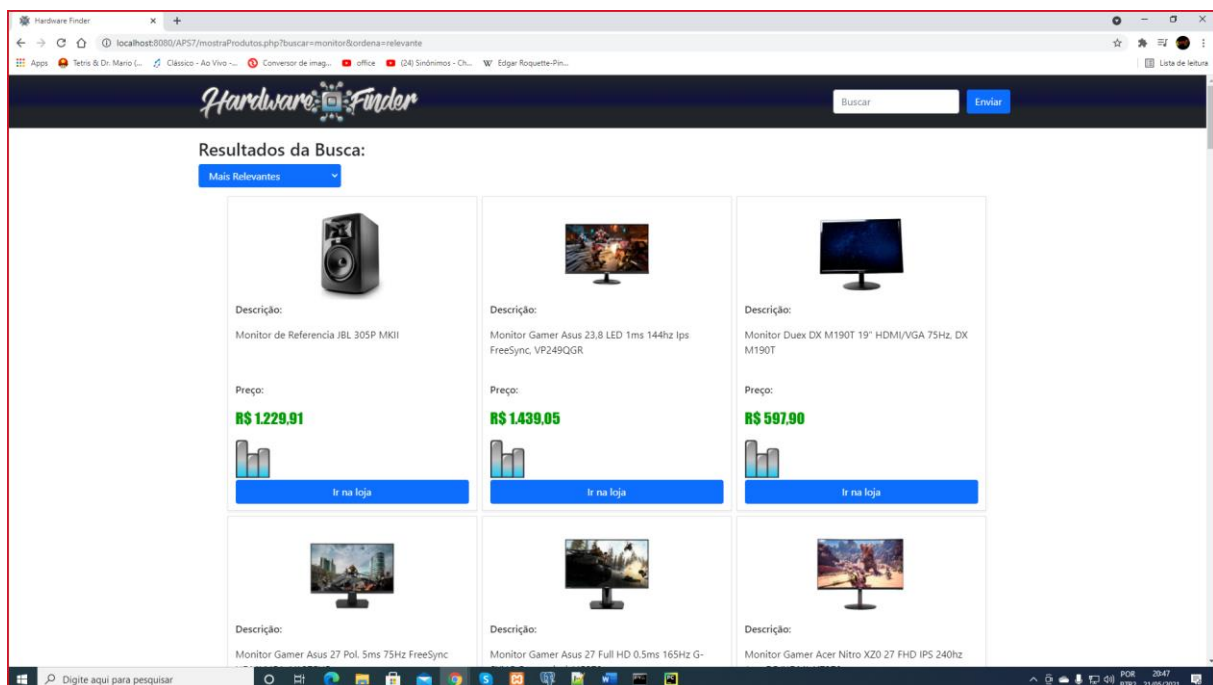
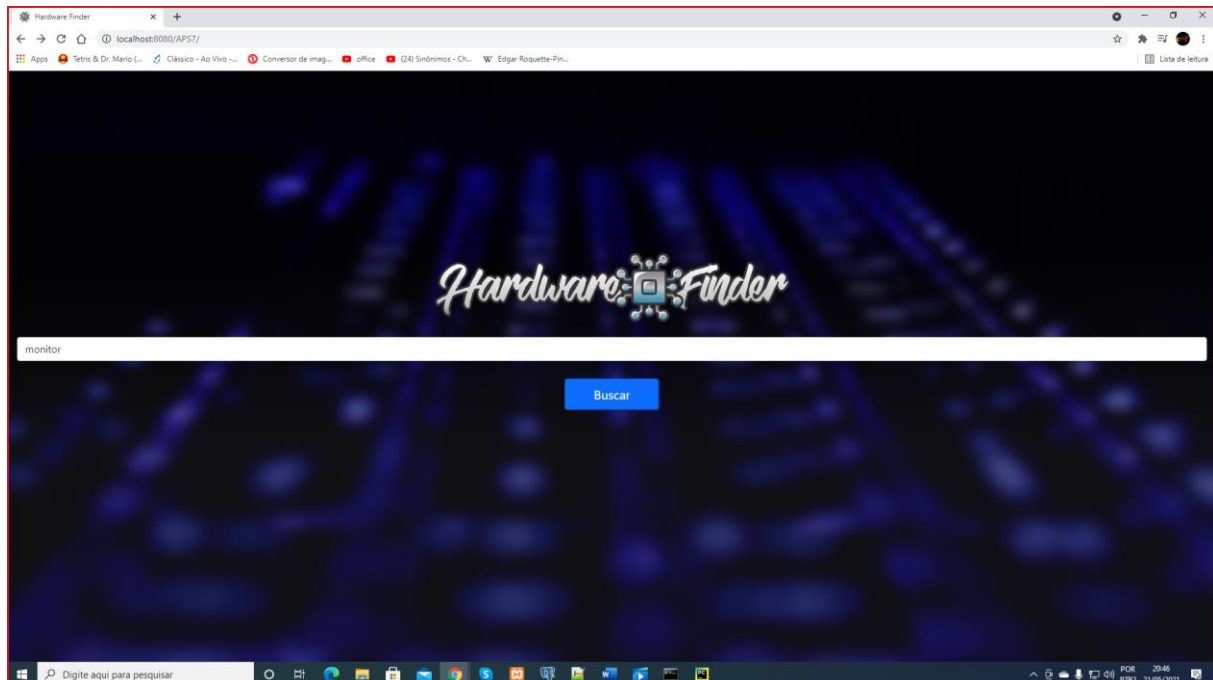
```

historicoProduto.php X
historicoProduto.php > html > body
94     </div>
95 </nav>
96
97     <div id="curve_chart" style="width: 100%; height: 80%"></div>
98     <div style="margin-left: 10%;">
99         <a href=?php echo $link; ?><button class="btn btn-primary" style="margin-bottom:10px; position: relative;
width:250px;">Ir na loja</button></a>
100         <button onclick="voltar()" class="btn btn-primary" style="margin-bottom:10px; position: relative;
width:250px;">Voltar</button>
101     </div>
102
103 <!-- Footer -->
104 <!-- Bootstrap core JavaScript
105 ===== -->
106 <!-- Placed at the end of the document so the pages load faster -->
107 <script src="https://ajax.googleapis.com/ajax/libs/jquery/1.12.4/jquery.min.js"></script>
108 <script>window.jQuery || document.write('<script src="js/jquery.min.js"></script>')</script>
109 <script src="js/bootstrap.min.js"></script>
110 <!-- IE10 viewport hack for Surface/desktop Windows 8 bug -->
111 <script src="js/ie10-viewport-bug-workaround.js"></script>
112 </body>
113 </html>

```



Apresentação do programa em funcionamento









Hardware Finder

localhost:8080/AP57/mostraProdutos.php?busca=s&ordenar=crescente

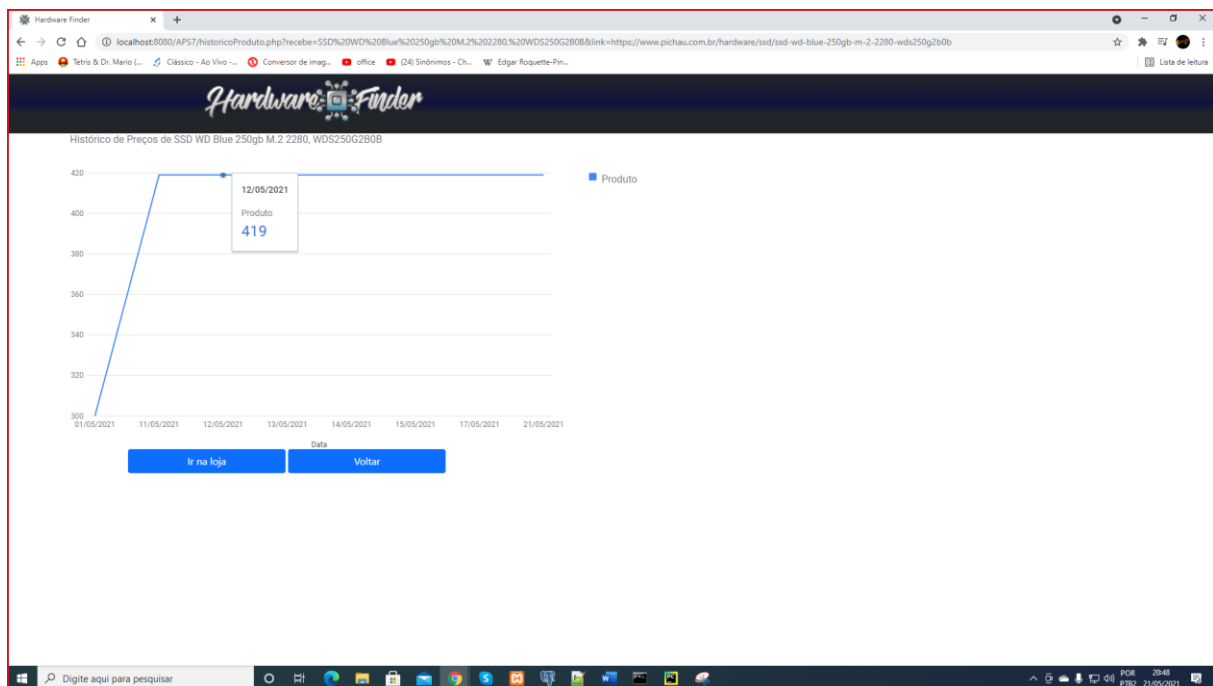
Buscar Enviar

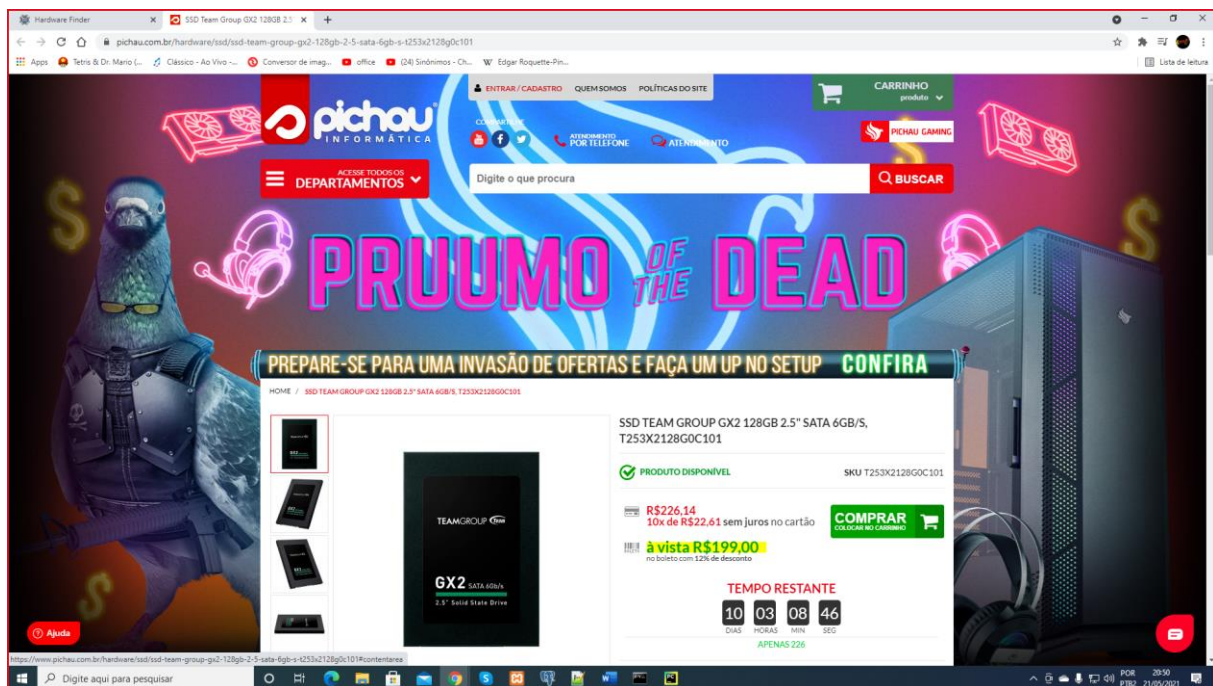
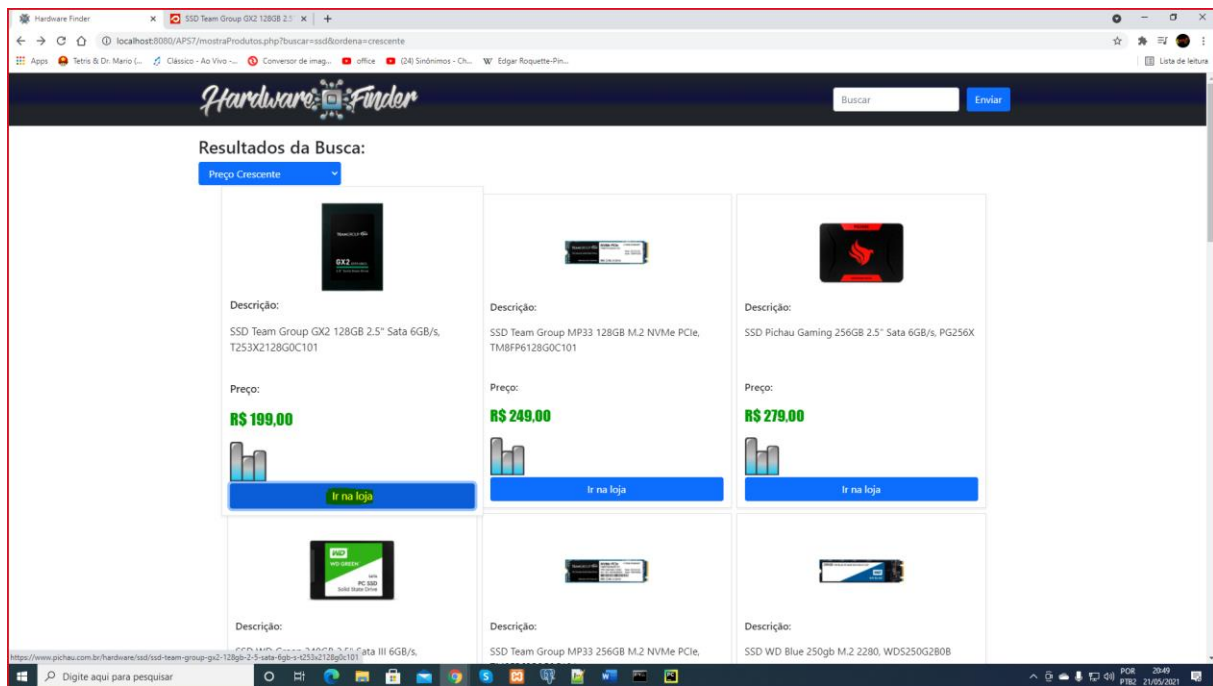
Resultados da Busca:

Preço Crescente

 <p>Descrição: SSD Team Group GX2 128GB 2.5" SATA 6GB/s, T253K2128GOC101</p> <p>Preço: R\$ 199,00</p> <p>Ir na loja</p>	 <p>Descrição: SSD Team Group MP33 128GB M.2 NVMe PCIe, TM8FP6128GOC101</p> <p>Preço: R\$ 249,00</p> <p>Ir na loja</p>	 <p>Descrição: SSD Pichau Gaming 256GB 2.5" SATA 6GB/s, PG256X</p> <p>Preço: R\$ 279,00</p> <p>Ir na loja</p>
 <p>Descrição: SSD WD Green 240GB 2.5" SATA III 6GB/s</p>	 <p>Descrição: SSD Team Group MP33 256GB M.2 NVMe PCIe</p>	 <p>Descrição: SSD WD Blue 250gb M.2 2280, WDS250G2B0B</p>

20:47 21/05/2021





Bibliografia

<https://kenzie.com.br/blog/o-que-e-python/>

[https://en.wikipedia.org/wiki/Beautiful_Soup_\(HTML_parser\)](https://en.wikipedia.org/wiki/Beautiful_Soup_(HTML_parser))

[https://en.wikipedia.org/wiki/Requests_\(software\)](https://en.wikipedia.org/wiki/Requests_(software))

<https://www.devmedia.com.br/o-que-e-json/23166>

<https://medium.com/tech-grupozap/introdu%C3%A7%C3%A3o-a-biblioteca-pandas-89fa8ed4fa38>

<https://medium.com/data-hackers/uma-introdu%C3%A7%C3%A3o-simples-ao-pandas-1e15eea37fa1>

<https://www.treinaweb.com.br/blog/o-que-e-flask/>

[https://pt.wikipedia.org/wiki/Flask_\(framework_web\)](https://pt.wikipedia.org/wiki/Flask_(framework_web))

<https://pt.wikipedia.org/wiki/XAMPP>

https://www.php.net/manual/pt_BR/intro-what-is.php

<https://www.homehost.com.br/blog/tutoriais/o-que-e-bootstrap/>

<https://getbootstrap.com.br/docs/4.1/examples/>

<https://rockcontent.com/br/blog/postgresql/>




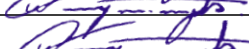





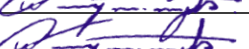
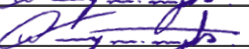

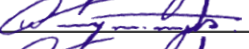
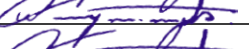
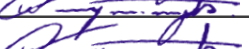

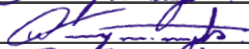
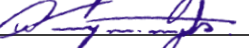

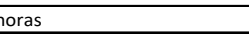

<https://computerworld.com.br/carreira/python-10-motivos-para-aprender-a-linguagem-em-2019/>

[\(fonte: https://brasilecola.uol.com.br/informatica/comunicacao-dados.htm\)](https://brasilecola.uol.com.br/informatica/comunicacao-dados.htm)

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: WESLEY MARCOS MORAIS DOS SANTOS**RA:** N3855J3**TURMA:** CC5P18**CURSO:** CIÊNCIA DA COMPUTAÇÃO**CAMPUS:** RIBEIRÃO PRETO**SEMESTRE:** 5º SEMESTRE**TURNO:** NOTURNO

ASSINATURA				
DATA	ATIVIDADE	TOTAL DE HORAS	ALUNO	PROFESSOR
12/03/2021	Discussão e escolha do tema	3 horas		
10/04/2021	Pesquisa sobre tema e ferramentas	4 horas		
17/04/2021	Distribuição de função de membros	3 horas		
18/04/2021	Desenvolvimento inicial do projeto	4 horas		
23/04/2021	Pesquisa sobre o desenvolvimento e utilização do Beautifull Soup	4 horas		
24/04/2021	Desenvolvimento da API	5 horas		
25/04/2021	Pesquisa sobre tecnologias do front-end	6 horas		
25/04/2021	Desenvolvimento de front-end	5 horas		
26/04/2021	Implementações na API	4 horas		
27/04/2021	Implementações no front-end	5 horas		
01/05/2021	Revisões e Testes	3 horas		
02/05/2021	Ajustes de front-end e API	4 horas		
08/05/2021	Implementação de banco de dados	5 horas		
09/05/2021	Ajustes de páginas de recebimento de gráficos	6 horas		
10/05/2021	Testes e revisões	3 horas		
15/05/2021	Início de elaboração de documentação	5 horas		
16/05/2021	Complementação de documentação	5 horas		
17/05/2021	Revisão de documentação	3 horas		
21/05/2021	Revisão de ajustes de todo projeto	4 horas		
22/05/2021	Revisão e ajustes de documentação	5 horas		
23/05/2021	Finalização e testes finais	4 horas		

TOTAL DE HORAS: 90 horas

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: WELLINGTON FAGUNDES DE OLIVEIRA**RA:** N429160**TURMA:** CC5P18**CURSO:** CIÊNCIA DA COMPUTAÇÃO**CAMPUS:** RIBEIRÃO PRETO**SEMESTRE:** 5º SEMESTRE**TURNIO:** NOTURNO**ASSINATURA**

DATA	ATIVIDADE	TOTAL DE HORAS	ALUNO	PROFESSOR
12/03/2021	Discussão e escolha do tema	3 horas	Wellington F	
10/04/2021	Pesquisa sobre tema e ferramentas	4 horas	Wellington F	
17/04/2021	Distribuição de função de membros	3 horas	Wellington F	
18/04/2021	Desenvolvimento inicial do projeto	4 horas	Wellington F	
23/04/2021	Pesquisa sobre o desenvolvimento e utilização do Beautifull Soup	4 horas	Wellington F	
24/04/2021	Desenvolvimento da API	5 horas	Wellington F	
25/04/2021	Pesquisa sobre tecnologias do front-end	6 horas	Wellington F	
25/04/2021	Desenvolvimento de front-end	5 horas	Wellington F	
26/04/2021	Implementações na API	4 horas	Wellington F	
27/04/2021	Implementações no front-end	5 horas	Wellington F	
01/05/2021	Revisões e Testes	3 horas	Wellington F	
02/05/2021	Ajustes de front-end e API	4 horas	Wellington F	
08/05/2021	Implementação de banco de dados	5 horas	Wellington F	
09/05/2021	Ajustes de páginas de recebimento de gráficos	6 horas	Wellington F	
10/05/2021	Testes e revisões	3 horas	Wellington F	
15/05/2021	Início de elaboração de documentação	5 horas	Wellington F	
16/05/2021	Complementação de documentação	5 horas	Wellington F	
17/05/2021	Revisão de documentação	3 horas	Wellington F	
21/05/2021	Revisão de ajustes de todo projeto	4 horas	Wellington F	
22/05/2021	Revisão e ajustes de documentação	5 horas	Wellington F	
23/05/2021	Finalização e testes finais	4 horas	Wellington F	

TOTAL DE HORAS: 90 horas

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: LUCAS CAMPANUCHI CORRÊA**RA:** F0359E1**TURMA:** CC5Q18**CURSO:** CIÊNCIA DA COMPUTAÇÃO**CAMPUS:** RIBEIRÃO PRETO**SEMESTRE:** 5º SEMESTRE**TURNO:** NOTURNO**ASSINATURA**

DATA	ATIVIDADE	TOTAL DE HORAS	ALUNO	PROFESSOR
12/03/2021	Discussão e escolha do tema	3 horas	lucas	
10/04/2021	Pesquisa sobre tema e ferramentas	4 horas	lucas	
17/04/2021	Distribuição de função de membros	3 horas	lucas	
18/04/2021	Desenvolvimento inicial do projeto	4 horas	lucas	
23/04/2021	Pesquisa sobre o desenvolvimento e utilização do Beautifull Soup	4 horas	lucas	
24/04/2021	Desenvolvimento da API	5 horas	lucas	
25/04/2021	Pesquisa sobre tecnologias do front-end	6 horas	lucas	
25/04/2021	Desenvolvimento de front-end	5 horas	lucas	
26/04/2021	Implementações na API	4 horas	lucas	
27/04/2021	Implementações no front-end	5 horas	lucas	
01/05/2021	Revisões e Testes	3 horas	lucas	
02/05/2021	Ajustes de front-end e API	4 horas	lucas	
08/05/2021	Implementação de banco de dados	5 horas	lucas	
09/05/2021	Ajustes de páginas de recebimento de gráficos	6 horas	lucas	
10/05/2021	Testes e revisões	3 horas	lucas	
15/05/2021	Início de elaboração de documentação	5 horas	lucas	
16/05/2021	Complementação de documentação	5 horas	lucas	
17/05/2021	Revisão de documentação	3 horas	lucas	
21/05/2021	Revisão de ajustes de todo projeto	4 horas	lucas	
22/05/2021	Revisão e ajustes de documentação	5 horas	lucas	
23/05/2021	Finalização e testes finais	4 horas	lucas	

TOTAL DE HORAS: 90 horas

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: VINICIUS ALVES PANOBIANCO**RA:** N485HJ6**TURMA:** CC5Q18**CURSO:** CIÊNCIA DA COMPUTAÇÃO**CAMPUS:** RIBEIRÃO PRETO**SEMESTRE:** 5º SEMESTRE**TURNO:** NOTURNO**ASSINATURA**

DATA	ATIVIDADE	TOTAL DE HORAS	ALUNO	PROFESSOR
12/03/2021	Discussão e escolha do tema	3 horas	Vinicius A. Panobianco	
10/04/2021	Pesquisa sobre tema e ferramentas	4 horas	Vinicius A. Panobianco	
17/04/2021	Distribuição de função de membros	3 horas	Vinicius A. Panobianco	
18/04/2021	Desenvolvimento inicial do projeto	4 horas	Vinicius A. Panobianco	
23/04/2021	Pesquisa sobre o desenvolvimento e utilização do Beautifull Soup	4 horas	Vinicius A. Panobianco	
24/04/2021	Desenvolvimento da API	5 horas	Vinicius A. Panobianco	
25/04/2021	Pesquisa sobre tecnologias do front-end	6 horas	Vinicius A. Panobianco	
25/04/2021	Desenvolvimento de front-end	5 horas	Vinicius A. Panobianco	
26/04/2021	Implementações na API	4 horas	Vinicius A. Panobianco	
27/04/2021	Implementações no front-end	5 horas	Vinicius A. Panobianco	
01/05/2021	Revisões e Testes	3 horas	Vinicius A. Panobianco	
02/05/2021	Ajustes de front-end e API	4 horas	Vinicius A. Panobianco	
08/05/2021	Implementação de banco de dados	5 horas	Vinicius A. Panobianco	
09/05/2021	Ajustes de páginas de recebimento de gráficos	6 horas	Vinicius A. Panobianco	
10/05/2021	Testes e revisões	3 horas	Vinicius A. Panobianco	
15/05/2021	Início de elaboração de documentação	5 horas	Vinicius A. Panobianco	
16/05/2021	Complementação de documentação	5 horas	Vinicius A. Panobianco	
17/05/2021	Revisão de documentação	3 horas	Vinicius A. Panobianco	
21/05/2021	Revisão de ajustes de todo projeto	4 horas	Vinicius A. Panobianco	
22/05/2021	Revisão e ajustes de documentação	5 horas	Vinicius A. Panobianco	
23/05/2021	Finalização e testes finais	4 horas	Vinicius A. Panobianco	

TOTAL DE HORAS: 90 horas