

Relatório Desafio Seazone

Comecei o programa pegando o URL da OLX e verificando o que modificava em cada lista de anúncios. Depois que identifiquei que apenas alterava o ultimo numero do endereço (referente a qual pagina estava), coloquei como parâmetro in range o numero de páginas que sera desejado fazer a aquisição.

Após isso utilizei uma função com o BeautifulSoup para fazer a requisição e encontrar a lista de todos os anúncios da pagina referente e analisar os dados de cada anuncio por partes. Identificando a lista, fiz um for que retirada todas as informações mais facéis de capturar, dali saíram o tamanho do terreno, local e horário em que foi postado cada anuncio.

Posteriormente percebi que alguns dados so poderiam ser adquiridos se meu scrapper entrasse em cada URL dos anúncios de forma individual, repeti os mesmos passos do inicio, identificando quais alterações o endereço recebia para acessar os anúncios, criei outra requisição em uma nova função para esses anuncios individuais. Acessando isoladamente cada um, consegui outros dados importantes para o desafio, sendo eles: Descrição do anuncio, título do anuncio e preço do terreno.

Chegamos no impasse para a conclusão do meu programa: Conseguir o nome do vendedor. Fiquei bastante tempo tentando entender porque eu não conseguia fazer a requisição do nome sendo que ele estava sendo exibido na pagina dos anuncios, mas, após certo tempo, descobri que essa informação estava sendo carregada por um post em uma API separada. Achando essa informação ficou fácil conseguir a informação faltante para a conclusão do desafio.

Criei uma última função onde eu fazia um request para essa API separa e pedia para pegar um item especifico no JSON que correspondia ao nome do Vendedor, depois juntei todas as informações adquiridas em um dicionário para organizar por tópicos e finalmente adicionei todos em uma listaJson e converti para CSV.

Por fim para conseguir as medias, máximos e mínimos, fiz alguns comandos simples no python e coloquei em uma nova lista que no final gera um arquivo xlsx para uma visualização facilitada dos dados.

Análise sobre a programação:

Acredito que poderia ter feito de uma forma mais sucinta o código de um modo geral, penso que seria possível efetuar todos os request em uma so função e seria mais eficiente. Concluo também que existem algumas redundâncias durante o código que pela falta de experiencia em web scrapping foram as soluções encontradas por mim para o desafio.

Feedback sobre o desafio:

Para mim foi um desafio incrível, pois não tinha nenhum conhecimento sobre Scrapper's no inicio da proposta e aprender tudo de uma forma pratica foi inspirador e muito motivador para mim, me dando mais motivos para querer trabalhar com isso na empresa. Mostrou, para mim, a minha capacidade de explorar o mundo da programação e o quanto é gratificante concluir um projeto desse com sucesso. Agradeço a oportunidade de participar desse desafio.