Web Scraping - Cadeiras Gamer da Kabum

Este projeto tem como objetivo realizar o web scraping do site da Kabum para obter informações sobre cadeiras gamer. Os dados extraídos são tratados e armazenados para posterior análise.

Objetivo do Projeto:

O objetivo deste projeto é obter dados sobre cadeiras gamer disponíveis no site da Kabum, incluindo informações de marca e preço. Esses dados podem ser utilizados para análise de mercado, comparação de preços e outras investigações relacionadas ao mercado de cadeiras gamer.

Tratamento de Dados:

Após a extração dos dados, é realizado um tratamento para garantir a consistência e qualidade das informações. Neste projeto, o tratamento de dados consiste nas seguintes etapas:

1. Limpeza de caracteres especiais: os valores de marca e preço são submetidos a um processo de remoção de caracteres especiais utilizando a biblioteca unidecode.

2. Renomeação de colunas: os nomes das colunas são renomeados para 'Marca' e 'Valor' para uma melhor compreensão dos dados.

3. Formato de saída: os dados tratados são armazenados em um arquivo CSV com o separador ';' e a codificação UTF-8.

Análise de Dados:

Após o tratamento dos dados, é possível realizar análises e investigações adicionais. Alguns exemplos de análises que podem ser feitas com os dados obtidos neste projeto são:

1. Estatísticas descritivas: calcular estatísticas básicas, como média, mediana e desvio padrão dos preços das cadeiras gamer.

2. Comparação de marcas: identificar as marcas mais populares ou mais caras com base nos dados coletados.

3. Tendências de preços: observar variações de preço ao longo das páginas ou ao longo do tempo para identificar tendências de mercado.

Tecnologia Utilizada:

- Python: linguagem de programação utilizada para desenvolver o script de web scraping e manipulação de dados.

- Biblioteca requests: utilizada para fazer solicitações HTTP ao site da Kabum e obter o conteúdo das páginas.

- Biblioteca BeautifulSoup: usada para analisar o HTML e extrair as informações desejadas.

- Biblioteca pandas: utilizada para manipulação e análise dos dados, incluindo a criação do DataFrame e a escrita em arquivo CSV.

- Biblioteca math: utilizada para cálculos matemáticos, como o arredondamento do número de páginas.

- Biblioteca unidecode: utilizada para remover caracteres especiais dos dados.

Conclusão:

Este projeto demonstra como realizar o web scraping de um site específico para obter dados relevantes. Através do tratamento adequado dos dados extraídos, é possível realizar análises e obter insights valiosos sobre o mercado de cadeiras gamer. Porém, é importante ressaltar que o uso de web scraping deve ser feito em conformidade com os termos de uso do site alvo e com as leis e regulamentações aplicáveis.