# A base de dados escolhida para o projeto A3 é a **“Amazon Products Dataset 2023”** referente ao mês de setembro, estando disponível no Kaggle: (<https://www.kaggle.com/datasets/asaniczka/amazon-products-dataset-2023-1-4m-products?select=amazon_products.csv>).

Utilizaremos esta base de dados para a UC de “Análise de Dados e Big Data” e “Inteligência Artificial”.

As colunas da base de dados estão divididas da seguinte forma:

**Arquivo “amazon\_products”:**

* “asin”: Coluna referente aos ID dos produtos (Tipo String)
* “title”: Coluna referente aos nomes dos produtos (Tipo String)
* “imgUrl”: Coluna referente ao Url da imagem dos produtos (Tipo String)
* “productUrl”: Coluna referente ao Url dos produtos (Tipo String)
* “stars”: Coluna referente a avaliação dos clientes sobre os produtos, variando de 0 a 5 (Tipo Float)
* “reviews”: Coluna referente ao número de comentários sobre os produtos (Tipo Int)
* “price”: Coluna referente ao preço dos produtos (Tipo float)
* “listPrice”: Coluna referente ao preço inicial dos produtos, caso o produto não tenha sofrido alteração de preço, o valor será 0 (Tipo float)
* “category\_id”: Coluna referente a categoria dos produtos, se baseando no arquivo “amazon\_categories”(Tipo Int)
* “isBestSeller”: Coluna referente ao produto estar ou não na categoria de “mais vendido” (Tipo Boolean)
* “boughtInLastMonth”: Coluna referente ao número de vendas no último mês (Tipo Int)

**Arquivo “amazon\_categories”**

* “id”: Coluna referente ao ID das categorias dos produtos (Tipo Int)
* “category\_name”: Coluna referente ao nome das categorias dos produtos (Tipo String)