

Desafio Ciência dos Dados

*The value of
an idea lies in
the using of it.*

*- Thomas A.
Edison,
American
Inventor*

*The goal is to
turn data into
information,
and
information
into insight.*
*- Carly Fiorina,
Former CEO of
HP*

*If you torture
the data long
enough, it will
confess.*

*- Ronald
Coase,
Economist*

Dados não servem para nada se não podem ser transformados em informação. Informações não servem para nada se não nos levarem a tomar alguma ação.

Em anexo você encontrará o dump de um banco de dados MySQL, contendo diversas informações de produtos retirados da Internet, como por exemplo nome, url, categoria, conteúdo HTML, etc.

Seu objetivo

Objetivo 1: Extrair informações dos dados

Desenvolva um algoritmo que extraia a marca de cada produto. A saída do seu programa deve ser um arquivo `marcas.txt`, que deve seguir o exemplo ao lado.

```
#Exemplo de marcas.txt
#ID_PRODUTO=MARCA_EM_LOW
ERCASE 9=passe bem
106=samsung 95=aoc 79=
62=philco 104=lg
```

Shell ▾

Objetivo 2 (plus): Extrair ações das informações

Desenvolva um algoritmo que extraia a

marca de cada produto. A saída do seu programa deve ser um arquivo `marketshare.txt`, que deve seguir o exemplo ao lado.

```
#Exemplo de marketshare.txt
#MARCA_EM_LOWERCASE,MARKET=MARKETSHARE passe bem,58=0.32%
#significa que 0.32% dos produtos do Walmart (58) são da marca
'passe bem' samsung,62=12% #significa que 12% dos produtos das
Casas Bahia (62) são da marca 'samsung'
```

Shell ▾

Anexos

https://mega.nz/#!rnY1zICJ!WalE-CX94Rmb8NMgh0ra2QeHmv98F6Q2jn8Zs_rAlnM

Banco de Dados MySQL com os dados a serem processados

Regras

Não existem regras, desde que o objetivo seja atingido!

- Você pode usar qualquer linguagem de programação, dicionários, frameworks, bibliotecas de processamento de texto, bancos de dados de marcas da internet, qualquer outra informação pública da internet, etc
- Faça o que for necessário para atingir o seu objetivo, com a maior precisão possível.
- Use a imaginação, pense fora da caixa.
- Para este desafio, não existe resposta certa ou errada: Além do resultado, você será avaliado pela sua criatividade e inteligência ao abordar o problema