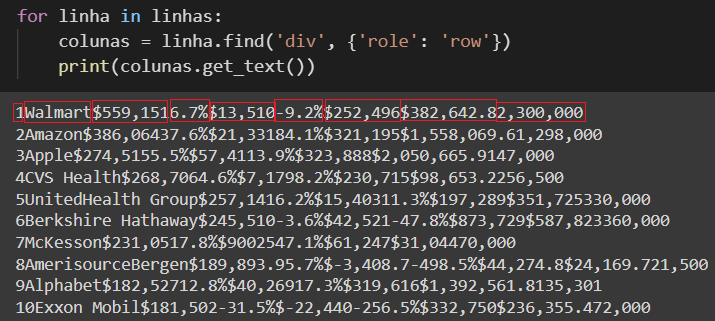
MÓDULO 12

Pergunta (relacionada ao exercício 1): como extrair os dados?

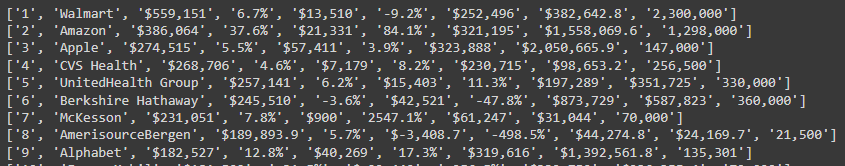
Resposta: como os dados estão sendo extraídos de um arquivo em formato html é importante entender essa estrutura organizada por tags para conseguir interpretar a fonte de dados. Você pode revisitar o tópico 2.1 do módulo 11 e estudar um pouco mais através desse material online: <https://www.w3schools.com/tags/ref_byfunc.asp>.

O exercício já disponibilizou o começo do código para ajudar na extração dos dados, se você imprimir o conteúdo da variável ‘colunas’, vai perceber que está no formato de tags. Então, podemos utilizar o método ‘get\_text()’ para verificar a informação textual da variável:



Veja que em cada linha, as informações solicitadas das colunas estão todas grudadas. Então a ideia é utilizar o ‘get\_text()’ para extrair a parte textual da variável ‘colunas’, usar um separador para delimitar cada informação (use como separador o ponto e vírgula) e depois quebrar os dados a partir do ponto e vírgula, utilizando o método ‘split()’.

A intenção é que ao final, você tenha listas com os dados separados, dessa forma:



Por fim, salve no arquivo CSV.