**EXERCÍCIO DE TREINO\***

Use a base de dados disponibilizada, “*adult\_data.csv*”, e realize as seguintes etapas:

* Importe as bibliotecas necessárias;
* Importe os tipos de dados necessários;
* Importe as funções SQL;
* Crie uma sessão PySpark;
* Defina um esquema de tratamento (*Não esqueça de analisar visualmente o dataset antes*);
* Importe o dataset;

Após essas etapas, faça os seguintes tratamentos:

1. Imprima o Schema de seu DF;
2. Imprima os primeiros 5 linhas de seu DF;
3. Converta o campo idade do tipo inteiro para o tipo Float (*ou vice versa, dependendo do seu Schema*);
4. Exiba somente 5 itens com os campos ‘age’ e ‘*education*’;
5. Agrupe a quantidade de itens em ‘*education*’ ordenados de maneira ascendente;
6. Exiba um describe da tabela ‘capital\_gain’;

\*Este exercício não possui correção.