

Lista de Exercícios 4a

1. Considere as tabelas de vendas diárias de três lojas:

```
Mercearia = pd.DataFrame({"loja":["A", "A", "A"],  
                           "dia": ["sex", "sab", "dom"],  
                           "valor": [7500, 9500, 8200]})
```

```
Farmácia = pd.DataFrame({"loja":["B", "B", "B"],  
                           "dia": ["sex", "sab", "dom"],  
                           "valor": [5100, 8250, 9900]})
```

```
Restaurante = pd.DataFrame({"loja":["C", "C"],  
                             "dia": ["sab", "dom"],  
                             "valor": [7500, 11800]})
```

Crie dois novos dataframes por meio da junção dos três acima, agrupando, para isso, os dataframes no sentido de linha para criar um, e no sentido de coluna para criar o segundo. Explique as diferenças entre as duas saídas, principalmente a ocorrência de valores NaN.

2. Com base nos dataframes a seguir, que guardam informações de matrículas de alunos em dois cursos, um de SQL e outro de Python:

```
SQL = pd.DataFrame({"email":["rakesh@xyz.com",  
                             "ecg@acmecorpus.com"]})
```

```
PYTHON = pd.DataFrame({"email":["ana@xyz.com",  
                                "jonas@acmecorpus.com",  
                                "rakesh@xyz.com"]})
```

responda:

- (a) quais são os alunos distintos, considerando ambas as listas?
- (b) quais são os alunos que fizeram ambos os cursos?

- (c) quais os alunos que realizaram o curso de SQL, mas não o de PYTHON (e vice-versa)?
3. Considere os dois dataframes a seguir, referentes aos dados sobre os departamentos e os funcionários de uma empresa.

```
dic_emp = {"id": ["D1", "D2", "D3", "D4"],
           "nomDeppto": ["Compras", "RH", "TI", "Vendas"],
           "local": ["SP", "RJ", "RJ", "SP"]}

dic_func = {"num": [3199, 3269, 3555, 3788, 3844],
            "nome": ["Ana", "David", "José", "Marina", "Luís"],
            "salario": [1600, 2975, 1500, 5000, 3000],
            "idDeppto": ["D2", "D3", None, "D2", "D4"]}

empresa = pd.DataFrame(dic_emp)
funcionarios = pd.DataFrame(dic_func)
```

Combine os dados dos funcionários com os dados dos departamentos para criar um novo dataframe completo sobre a empresa, usando `merge`.