

02

Para saber mais

O MultiLabelBinarizer é utilizado para casos onde temos várias labels dentro da mesma “ocorrência”. Em caso mais simples onde temos uma coluna com uma categoria por linha, uma forma bem simples de fazer esse tratamento é com o método “get_dummies” do Pandas.

```
pd.get_dummies(df['sx'])
```

	sx
0	male
1	male
2	male
3	female
4	male

female	male
0	1
0	1
0	1
1	0
0	1

No exemplo acima, para cada valor único da coluna “sx”, será criado uma coluna que é marcada com “1” quando há ocorrência do valor da coluna, e “0” quando não há.