**01 - Nomes da próxima turma**

Sempre quem uma nova turma de nossa escola se inicia, nós passamos os nomes dos alunos para a portaria do prédio, facilitando a identificação das salas. Esse arquivo geralmente é uma planilha de Excel, que precisaremos gerar.

Separaremos essa etapa do projeto em duas etapas. Na primeira verificaremos quais são os alunos de um determinado curso. Começaremos criando uma nova seção "Nomes dos alunos e alunas da próxima turma".

Antes de buscarmos esses nomes, vamos relembrar os dataframes que compõem nosso projeto atualmente, começando pelo matriculas\_por\_curso, que armazena a quantidade de alunos em cada turma.

matriculas\_por\_curso.head()COPIAR CÓDIGO

| **id\_curso** | **nome\_do\_curso** | **quantidade\_de\_alunos** |
| --- | --- | --- |
| 1 | Lógica de programação | 62 |
| 2 | Java para Web | 52 |
| 3 | C# para Web | 25 |
| 4 | Ruby on Rails | 33 |
| 5 | Cursos de Python | 59 |

Dentre os cursos da lista completa, escolheremos o "Estatística básica", cujo ID é 16 e que, por coincidência, possui 16 alunos. Se preferir, você pode escolher outro curso da lista.

Temos também o dataframe matriculas, que nos mostra os IDs dos alunos matriculados em cada curso.

matriculas.head()COPIAR CÓDIGO

|  | **id\_aluno** | **id\_curso** |
| --- | --- | --- |
| 0 | 235 | 15 |
| 1 | 235 | 4 |
| 2 | 43 | 6 |
| 3 | 43 | 4 |
| 4 | 43 | 10 |

Isso nos permite, por exemplo, filtrar todos os IDs dos alunos matriculados em determinado curso. Pensando nisso, criaremos a variável id\_curso que receberá o ID que escolhemos inicialmente, 16. Em seguida, criaremos uma variável proxima\_turma que receberá uma *query* buscando todos os alunos matriculados no curso escolhido.

Para gerarmos essa *query*, usaremos a função query() recebendo como parâmetro id\_curso == {}".format(id\_curso).

id\_curso = 16

proxima\_turma = matriculas.query("id\_curso == {}".format(id\_curso))COPIAR CÓDIGO

Ao visualizarmos, receberemos o ID de todos os alunos matriculados no curso 16.

proxima\_turma.head()COPIAR CÓDIGO

|  | **id\_aluno** | **id\_curso** |
| --- | --- | --- |
| 95 | 259 | 16 |
| 155 | 151 | 16 |
| 176 | 179 | 16 |
| 178 | 179 | 16 |
| 334 | 368 | 16 |

Agora queremos saber quais os nomes desses alunos, de forma a passarmos as informações corretas para a portaria da escola. Para isso, uniremos dus tabelas: proxima\_turma e nomes.

nomes.sample(3)COPIAR CÓDIGO

|  | **nome** | **id\_aluno** | **dominio** | **email** | **matriculas** |
| --- | --- | --- | --- | --- | --- |
| 182 | POLIANA | 369 | @servicodoemail.com | poliana@servicodoemail.com | 2 |
| 12 | AMANDA | 217 | @servicodoemail.com | amanda@servicodoemail.com | 1 |
| 185 | SHEILA | 116 | @dominiodoemail.com | sheila@dominiodoemail.com | 1 |

Para isso, setaremos o índice de proxima\_turma como a coluna id\_aluno e chamaremos, a partir dela, a função join(). A esta função passaremos como parâmetro o dataframe nomes, também setando seu índice como id\_aluno.

proxima\_turma.set\_index('id\_aluno').join(nomes.set\_index('id\_aluno'))COPIAR CÓDIGO

O resultado dessa operação será um dataframe que consiste em todos os alunos matriculados no curso 16, "Estatística básica".

| **id\_aluno** | **id\_curso** | **nome** | **dominio** | **email** | **matriculas** |
| --- | --- | --- | --- | --- | --- |
| 147 | 16 | VINICIOS | @servicodoemail.com | vinicios@servicodoemail.com | 2 |
| 151 | 16 | TEREZA | servicodoemail.com | tereza@servicodoemail.com | 1 |
| 179 | 16 | TAIS | @dominiodoemail.com | tais@dominiodoemail.com | 1 |

Porém, existem informações nesse conjunto que não fazem tanto sentido no contexto em que estamos trabalhando - a portaria da escola -, como o domínio do e-mail e a quantidade de matrículas. Na realidade, poderíamos passar somente o nome dos alunos. Para isso, após a junção dos nossos dataframes, passaremos entre colchetes a coluna nome, especificando que queremos exibir somente ela.

proxima\_turma.set\_index('id\_aluno').join(nomes.set\_index('id\_aluno'))['nome']COPIAR CÓDIGO

id\_aluno 147 VINICIOS 151 TEREZA 179 TAIS 179 TAIS 184 JUNIOR 202 ESTER 206 ITALO 214 JAIME 259 RAIMUNDA

Note que nos foi devolvida apenas uma lista, o que não traz o visual elegante de um dataframe. Corrigiremos isso chamando, ao final do código, a função to\_frame(). Atribuiremos esse dataframe a uma variável proxima\_turma.

proxima\_turma = proxima\_turma.set\_index('id\_aluno').join(nomes.set\_index('id\_aluno'))['nome'].to\_frame()COPIAR CÓDIGO

| **id\_aluno** | **nome** |
| --- | --- |
| 147 | VINICIOS |
| 151 | TEREZA |
| 179 | TAIS |
| 179 | TAIS |
| 184 | JUNIOR |
| 202 | ESTER |
| 206 | ITALO |
| 214 | JAIME |
| 259 | RAIMUNDA |

Assim temos um resultado mais interessante, mas ainda existem pontos a solucionar. Por exemplo, não estamos passando o curso que esses alunos estão matriculados, o que pode resultar em confusões caso apareçam alunos com o mesmo nome. A ideia é incluirmos o nome do curso no nosso dataframe.

Criaremos então uma variável nome\_curso que receberá o comando cursos.loc[id\_curso], com o qual conseguiremos recuperar o nome do curso cujo ID está armazenado na variável id\_curso.

nome\_curso = cursos.loc[id\_curso]

nome\_cursoCOPIAR CÓDIGO

nome\_do\_curso Estatística básica Name: 16, dtype: object

Repare que a variável nome\_curso contém informações extras que não precisamos. Corrigiremos isso atribuindo a ela o valor nome\_curso.nome\_do\_curso, ou seja, apenas o conteúdo do campo nome\_do\_curso desse objeto.

nome\_curso = nome\_curso.nome\_do\_curso

nome\_cursoCOPIAR CÓDIGO

'Estatística básica'

Por fim, renomearemos a coluna nome do dataframe proxima\_turma para "Alunos do curso de" seguido do nome armazenado na variável nome\_curso. Faremos isso com o auxílio da função rename().

proxima\_turma.rename(columns = {'nome':'Alunos do curso de {}'.format(nome\_curso)})COPIAR CÓDIGO

| **id\_aluno** | **Alunos do curso de Estatística básica** |
| --- | --- |
| 147 | VINICIOS |
| 151 | TEREZA |
| 179 | TAIS |
| 179 | TAIS |
| 184 | JUNIOR |
| 202 | ESTER |
| 206 | ITALO |
| 214 | JAIME |
| 259 | RAIMUNDA |

Se você escolheu outro curso, será esse o nome a ser exibido. Atribuiremos esse resultado novamente à variável proxima\_turma e pediremos 5 amostras.

proxima\_turma = proxima\_turma.rename(columns = {'nome':'Alunos do curso de {}'.format(nome\_curso)})COPIAR CÓDIGO

| **id\_aluno** | **Alunos do curso de Estatística básica** |
| --- | --- |
| 214 | JAIME |
| 179 | TAIS |
| 184 | JUNIOR |
| 319 | INES |
| 398 | CRISTIAN |

Assim obtivemos o ID desses alunos, o curso que eles irão fazer e os seus nomes. Só falta exportarmos esse dataframe para uma planilha do Excel que será enviada à portaria.