

11

Uma rápida classificação

Vimos que os histogramas dividem em faixas as idades antes de fazer o gráfico. Mas isso não muda as nossas bases. Para ajudar com as análises dos clientes da AluraPlay, é interessante que façamos uma classificação da variável `idade`, para separar os clientes em “jovens”, “adultos”, etc.

Na parte 1 do curso fizemos uma classificação da data de lançamento dos jogos na variável `flag_lancamento`, determinando uma data específica que separa os jogos em duas classes. Com a idade precisamos fazer algo diferente, pois queremos que cada classe de idade tenha, aproximadamente, a mesma quantidade de clientes. Também, pela análise do histograma feito na aula anterior, concluímos que cinco é um bom número de classes em que podemos usar para dividir a idade.

Para fazer essa divisão existe um procedimento de “ranqueamento”, o `PROC RANK`, que recebe os seguintes parâmetros:

- Começamos dizendo a base de entrada após o parâmetro `data=`, que é a base `CADASTRO_CLIENTE_V3`, que está na biblioteca `alura`.
- Para não substituir a base de entrada, determinamos uma base de saída diferente através do parâmetro `out=`. Vamos chamar ela de `base_ranks`.
- A quantidade de categorias a serem criadas é expressa no parâmetro `groups=`.
- A variável a ser categorizada (`Idade`) é colocada após o parâmetro `var`.
- É importante também escolher o nome da variável que conterá essas classes que iremos criar. Para isso serve o parâmetro `ranks`, onde escolhemos o nome `faixa_idade` para esta nova variável.

O código resultante fica assim:

```
PROC RANK  
  data = alura.cadastro_cliente_v3  
  out = base_ranks  
  groups = 5;  
  var Idade;  
  ranks Faixa_Idade;  
RUN;
```

Isso cria a variável de categorização `faixa_idade`, mas não nos fornece nenhuma informação sobre as características das classes em si.

Construa uma base que contenha, para cada faixa de idade dessa nossa variável categorizada, as informações de quantidade de clientes, assim como o valor mínimo e o valor máximo da variável `Idade`. Chame esta base de `FAIXAS_IDADE` e salve no diretório da AluraPlay.