



Lisandra Godoi &lt;lysfernandagodoi@gmail.com&gt;

## #7DaysOfCode - Python Pandas 4/7: Análise exploratória de dados e Variáveis categóricas

1 mensagem

**Francisco Foz** <contato@7daysofcode.io>

29 de agosto de 2025 às 08:00

Responder a: contato@alura.com.br

Para: lysfernandagodoi@gmail.com



Python Pandas



Olá, Lisandra Fernanda de. Animado(a) para mais um dia de #7DaysOfCode?

Hoje você vai continuar a questionar ainda mais os seus dados para extrair ainda mais informações deles.

O objetivo será entender a quantidade de empréstimos a partir das variáveis categóricas do seu conjunto de dados.

Variável categórica? Mas o que é isso, Francisco?

Bom, um conjunto de dados é descrito pelas suas **características** que variam de acordo com cada um, por exemplo: peso, idade, quantidade de páginas, notas, nome da biblioteca, número de empréstimos, número de filhos, etc.

Elas podem ser classificadas em dois grandes grupos que se subdividem em mais dois outros grupos cada.

**Quantitativas:** são características numéricas que podemos

enumerar:

1. Discreta: possuem apenas valores finitos enumeráveis.  
Exemplo: número de empréstimos, número de filhos, número de páginas.
2. Contínua: possuem valores dentro de um intervalo de números.  
Exemplo: altura, peso, notas.

**Categóricas (ou qualitativas):** são características não-numéricas para as quais podemos definir categorias.

1. Nominal: possuem categorias não ordenadas naturalmente.  
Exemplo: nomes, cores, tema do livro.
2. Ordinal: são categorias onde está implícito um tipo de ordenamento ou hierarquia. Exemplo: classe social, nível de escolaridade.

Vamos explorar algumas das variáveis categóricas das quais precisaremos extrair mais informações. Elas são:

- Tipo de vínculo
- Coleção
- Biblioteca
- Classificação geral da CDU

Ao se analisar uma nova variável, é bem interessante verificar cada uma listando **quais são os valores únicos** dentro delas.

Para explorar os dados, alguns questionamentos serão pertinentes para a diretoria das bibliotecas, como:

- “*Como se distribuem os empréstimos de exemplares pelos **tipos de vínculo dos usuários**?*”

Desta forma, a diretoria poderá entender qual é o público que está utilizando a biblioteca e assim tomar decisões em continuar com a estratégia de negócio atual ou modificá-la.

- *Quais **coleções** são mais emprestadas?*

Da mesma forma, as coleções. Ranquear as coleções mais emprestadas pelo público, será bastante importante para a estratégia atual.

- *Quais são as **bibliotecas** com mais ou menos quantidade de empréstimos?*

Assim, a diretoria conseguirá entender onde ela deverá melhorar e focar suas iniciativas.

Você visualizará os dados dos empréstimos que aconteceram, mas e os dados que não estão aqui na análise?

Por quais outras perspectivas você pode entender essa tabela de como ela se formou? Comparar bibliotecas de tamanhos diferentes, com públicos diferentes, acervos diferentes... isso faz sentido para a diretoria do sistema de bibliotecas?

Aponte algumas outras métricas que poderiam entrar aqui para enriquecer essa análise.

- *De quais **temas da CDU** são os exemplares emprestados?*

Entender quais os temas mais procurados pelos usuários é fundamental para o desenvolvimento de novos planos de marketing do acervo. Para que possam não apenas fortalecer o que está sendo utilizado, mas também promover o que não está.

**Gere uma tabela de frequência e com o percentual para cada variável.**

Como é um trabalho repetitivo, **crie uma função** que gere a tabela com os valores.

Após geradas as tabelas, traga algumas das suas percepções para as análises com o que você poderá contribuir para a diretoria da biblioteca.

---

## DICA

Para arredondar os números do percentual, você pode utilizar a função built-in do Python Round().

---

## EXTRA

Confira a documentação da função Round() do Python.

Dúvidas sobre como criar funções com Python? Veja esse capítulo da apostila de “Python e orientação a objetos” da Alura.

---

## RESPOSTA DO EXERCÍCIO DO DIA 3

– Solução do Dia 3

---

Você terminou o quarto dia do #7DaysOfCode de Python Pandas e está cada vez mais se aprofundando nas análises.

Após finalizado, não deixe de postar seus resultados nas redes sociais e me marcar.

Ficarei muito feliz de ver o seu trabalho!

Boa jornada!

**Francisco Foz**

Bibliotecário e Analista de Dados

---

Alura, Vila Mariana, São Paulo, SP

[Não quero mais receber os Desafios](#)