

Bootcamp IGTI

Desafio

Módulo 2	Python para a Análise de Dados
-----------------	---------------------------------------

Objetivos

Exercitar os seguintes conceitos trabalhados no Módulo:

- ✓ Realizar a análise de uma base de dados real em Python.

Enunciado

O uso de bicicletas como meio de transporte ganhou muita força nos últimos anos, seja por questões ambientais, de saúde ou até mesmo de infraestrutura de trânsito. Para incentivar seu uso, cidades em todo o mundo têm implementado programas de compartilhamento de bicicleta. Nesses sistemas, elas são retiradas e devolvidas em quiosques automatizados que ficam espalhados por diversos pontos da cidade.

As plataformas de compartilhamento de bicicletas costumam coletar diversos tipos de dados, como a duração da viagem, as localizações iniciais e finais dos percursos, entre outros. Esses dados, em conjunto com informações sobre o clima, o trânsito e o relevo, por exemplo, possibilitam uma análise mais robusta do compartilhamento de bicicletas.

Segue um descritivo dos dados coletados:

- **rec_id**: índice do registro de locação;
- **datetime** : data;
- **season** : estação do ano (1: inverno, 2: primavera, 3: verão, 4: outono). Relativo ao hemisfério norte;
- **year** : ano (0: 2011, 1:2012);

- **month** : mês (1 a 12);
- **hour** : hora do dia (0 a 23);
- **is_holiday** : booleano indicando feriado;
- **weekday** : dia da semana (0: domingo, 1: segunda-feira, ..., 6: sábado);
- **is_workingday** : booleano indicando dia útil;
- **weather_condition**: (1: limpo, 2: nublado, 3: chuva leve, 4: chuva forte);
- **temp** : Temperatura escalada entre 0 e 1. Valor original em graus Celsius: -8 a 39;
- **atemp**: Sensação térmica escalada entre 0 e 1. Valor original em graus Celsius: -16 a 50;
- **humidity**: Humidade relativa (0 a 1);
- **windspeed**: Velocidade do vento escalada entre 0 e 1 (máximo original: 67);
- **casual**: número de locações para usuários casuais;
- **registered**: número de locações para usuários registrados;
- **total_count**: contador total de alugueis (casual+registered).

Esta atividade tem como objetivo analisar os dados de compartilhamento de bicicletas em uma cidade, coletados pela Universidade do Porto.

Atividades

Os alunos deverão desempenhar as seguintes atividades:

1. Acessar o ambiente [Google Colaboratory](#) (recomendado) ou qualquer ambiente de desenvolvimento em Python.
2. Importar as bibliotecas do numpy e pandas conforme:

```
# importando as bibliotecas
import numpy as np
import pandas as pd
```

3. Carregar o dataset <https://pycourse.s3.amazonaws.com/bike-sharing.csv> para análise utilizando o pandas. Exemplo:

```
df = pd.read_csv("https://pycourse.s3.amazonaws.com/bike-sharing.csv")
```

4. Responder as seguintes perguntas:

- a. Qual o tamanho desse dataset?
- b. Qual a média da coluna **windspeed**?
- c. Qual a média da coluna **temp**?
- d. Quantos registros existem para o ano de 2011?
- e. Quantos registros existem para o ano de 2012?
- f. Quantas locações de bicicletas foram efetuadas em 2011?
- g. Quantas locações de bicicletas foram efetuadas em 2012?
- h. Qual estação do ano contém a maior média de locações de bicicletas?
- i. Qual estação do ano contém a menor média de locações de bicicletas?
- j. Qual horário do dia contém a maior média de locações de bicicletas?
- k. Qual horário do dia contém a menor média de locações de bicicletas?
- l. Que dia da semana contém a maior média de locações de bicicletas?
- m. Que dia da semana contém a menor média de locações de bicicletas?
- n. Às quartas-feiras (`weekday = 3`), qual o horário do dia contém a maior média de locações de bicicletas?
- o. Aos sábados (`weekday = 6`), qual o horário do dia contém a maior média de locações de bicicletas?