

Questão 16 – LINQ – Nível Aprendiz

Você precisa analisar dois arquivos de dados para um trabalho de ciência de dados. A análise será sobre os dados de aluguel de bicicletas de um serviço (real) nos Estados Unidos. O primeiro dataset chamado *dayinfo.csv* possui 5 colunas:

1. instant: Dia desde o começo da análise.
2. season: Estação do ano.
3. workingday: 1 se o dia é um dia de trabalho e 0 caso contrário, ou seja, fim de semana ou feriado.
4. weathersit: Indica as condições do tempo, quanto menor, melhor as condições naquele dia.
5. temp: indica temperatura normalizada.

No segundo arquivo chamado *bikeSharing* você possui os seguintes dados:

1. instant: Dia desde o começo da análise.
2. casual: Quantidade de usuários casuais que alugaram uma bicicleta, ou seja, não tem conta e vínculo com a empresa mas pagaram para alugá-las.
3. registered: Quantidade de usuários com vínculo que alugou uma bicicleta.

Seu trabalho é ler os dados e cruzar as informações de tempo com o aluguel para responder as seguintes perguntas:

1. Qual a média de alugueis de bicicletas em todo período? Sempre considere os aluguéis casuais junto aos registrados.
2. A empresa parece ter crescido, ou seja, aumentado os alugueis de bicicletas ao longo do tempo? Utilize as médias a cada 30 dias para analisar isso. Dica: Você pode resolver isso com um GroupBy com uma divisão.
3. Como a estação, condições de tempo e temperatura impactam nos resultados? Responda para os três casos separadamente.
4. Qual a média de aluguel de bicicletas nos dias de trabalho? E nos dias que não se trabalha?
5. Quais são os picos, tanto de alta quanto de baixa para o aluguel de bicicletas e quais eram as condições (dia de trabalho, condições do tempo, etc) nesses dias.