

Desafio Forecast

- Esse desafio utiliza como base de dados as ocorrências de acidentes de trânsito na cidade de Recife¹;
- O objetivo deste desafio é prever a quantidade de ocorrências de acidentes de trânsito do próximo dia. Desta forma deseja-se utilizar a quantidade de ocorrências do dia $t-1$, $t-2$, $t-n$ para prever o dia t , no qual n é o número máximo de lags (janela temporal).
- Base de dados:
 - Ocorrências de acidentes de trânsito na cidade de Recife, contendo ocorrências entre os anos de 2015 e 2019 (em anexo);
 - Será necessário transformar os dados disponíveis em uma série temporal diária YYYY-MM-DD (iniciando de 2015-06-01 até 31-12-2019), contendo a quantidade de acidentes de trânsito do dia;
 - Dica: para criar a série temporal desejada, basta utilizar a coluna “data” para identificar a quantidade de ocorrências do dia;
- Entregáveis:
 1. Análise descritiva da série temporal, dicas:
 - 1.1. Existem lags relevantes ?
 - 1.2. A série é estacionária ? baseado em qual análise ?
 - 1.3. Quais outras estatísticas podem ajudar no entendimento da distribuição?
 2. Modelagem, sugerir um modelo para prever a série temporal:
 - 2.1. Avaliar diferentes modelos;
 - 2.2. Prever 1 passo à frente, utilizando $t-1$, $t-2$, $t-n$ para prever o dia t ;
 - 2.3. Exibir resultados no conjunto de teste dos modelos selecionados. O conjunto de teste compreende o todo ano de 2019 (os últimos 365 pontos).
 - 2.4. Sugira e utilize métricas para mensurar o resultado do teste, ie: MSE, MAPE ...;
 - 2.5. Extra: É possível adicionar outras variáveis para ajudar no processo de previsão? Quais? melhoraram os resultados?
 - 2.6. Qual abordagem você usaria nesse contexto?
 3. Report de resultados:
 - 3.1. Criar dois jupyter notebooks com os entregáveis 1 e 2 (um notebook para cada);
 - 3.2. Pode utilizar o próprio notebook para responder questionamentos e discutir sobre os resultados;
 - 3.3. Importante que os notebooks estejam organizados e que possuam uma linha lógica (Data Storytelling);
 - 3.4. Diferencial: apresentação dos resultados em ppt;
 - 3.5. Subir os entregáveis no git e enviar o link por email;

Contato para envio e dúvidas: mauren.micalichen@gavb.com.br.

¹ dados retirados de: <https://www.emprel.gov.br/>