



Trabalho de Introdução à Data Science

Previsão da demanda de passagens aéreas para Porto Seguro

Frederico Cardozo Bianchessi



Objetivo do trabalho

- Desenvolver um modelo de regressão capaz de prever a demanda de passagens aéreas para vôos fretados por uma operadora de turismo com destino à Porto Seguro, Bahia, saindo de Porto Alegre.



Contextualização

- Relevância da situação: aeronaves são fretadas pela operadora de turismo a partir de análises qualitativas feitas por pessoas que trabalham há muitos anos no mercado.
- A partir de análise exploratória prévia, foi possível verificar que o principal destino da operadora é Porto Seguro.
- Simplificação do trabalho para trabalhar com a previsão de demanda apenas com vôos fretados para Porto Seguro



Dados



Coleta dos dados

Dados foram coletados a partir do sistema gerencial da operadora

Transformação dos dados

Base de dados foi transformada em uma série temporal com frequência diária para adequar aos padrões do Pycaret

Base de dados final

Série temporal de frequência diária, que se inicia em 01/01/2015 e vai até 31/12/2019. Comporta informação de passageiros que compraram a passagem com a operadora de turismo confirmados nos voos fretados pela operadora, com origem em Porto Alegre com destino à Porto Seguro, Bahia,

Dicionário de dados

Dicionário de dados		
Coluna	Tipo	Descrição
Data	datetime64	Dia do embarque
Confirmados	int64	Número de passageiros que compraram a passagem com a operadora de turismo confirmados no voo

Metodologia

- Python
- Bibliotecas utilizadas: Pandas e Pycaret
- Análise exploratória da série temporal
 - Sazonalidade detectada de 7 dias
 - Tendência de manutenção.

Construção do modelo

- Tipo: Univariada
- Horizonte de previsão: 90 dias
- Cross-validation: Expanding window (3 folds)
- Avaliação dos modelos com base na métricas de erro MAE e RMSE

Resultados obtidos

- Melhor modelo para os dados: Random Forest w/ Cond. Deseasonalize & Detrending
- Mean Absolute Error (MAE): 12.1638
- Root Mean Squared Error (RMSE): 34.1248

Próximos desafios

- Buscar mais dados para fazer uma análise univariada com uma série temporal de frequência mensal.
- Utilizar dados específicos da região de Porto Seguro para realizar análise multivariada



OBRIGADO

Frederico Cardozo Bianchessi

