

ANÁLISE ESTATÍSTICA E MODELAGEM PREDITIVA DE SÉRIES TEMPORAIS EM PYTHON (DSA)

Aula 3 – Tipos de Séries Temporais

Nem todas as séries temporais são iguais ou do mesmo tipo. Aliás, raramente serão, pois estamos falando de diferentes fenômenos observados ao longo do tempo.

Se você comparar uma série temporal mostrando a venda de geladeiras em 1982, ela será diferente de uma outra série dessa mesma variável (venda de geladeiras) de 2020. Os fenômenos mudam de comportamento ao longo do tempo, pois são afetados por diversos fatores.

Vejamos quais são os Tipos de Séries Temporais e na imagem abaixo você encontra um resumo completo.

Estacionária x Não Estacionária

Para modelos estatísticos de séries temporais, os dados não estacionários são aqueles em que a média e variância não são constantes e a covariância é dependente do tempo. São imprevisíveis e não podem ser modelados ou previstos.

Dados não estacionários devem ser transformados em dados estacionários, para que sejam consistentes e confiáveis.

Essa transformação é assegurada pela operação de “Diferenciação”, em que os valores de dados são substituídos pela diferença entre seus valores e os valores anteriores (subtraindo $Y(t-1)$ de $Y(t)$, assumindo a diferença:

$$Y(t) - Y(t-1)$$

Às vezes, a operação de diferenciação deve ser repetida um certo número de vezes até que a função resultante fique estacionária.

A estacionariedade é uma característica importante das séries temporais. Uma série temporal é considerada estacionária se suas propriedades estatísticas não mudarem ao longo do tempo. Em outras palavras, ela tem média e variância constantes, e a covariância é independente do tempo.

Frequentemente, os preços das ações não são um processo estacionário, pois podemos ver uma tendência crescente ou sua volatilidade pode aumentar ao longo do tempo (o que significa que a variação está mudando).

Ou será que alguém previu a queda abrupta no preço das ações que ocorreu esta semana? Aliás, será que essa queda se deve apenas ao Corona Vírus, ou estão aproveitando a oportunidade para desfazer a bolha financeira que tinha se formado? Alguns bons livros (como o *The Black Swan: The Impact of the Highly Improbable* - O Cisne

Negro: O Impacto do Altamente Improvável, de Nassim Nicholas Taleb, um best-seller imperdível) relatam que de tempos em tempos algum evento promove uma grande transformação e que o mercado financeiro é afetado a cada ciclo de 10 a 15 anos. A última crise foi em 2008/2009 (quebradeira dos bancos por pura ganância e irresponsabilidade), a anterior tinha sido em 2011 (atentado em Nova York) e agora em 2020 (Corona Vírus). 🤔

Podemos filosofar sobre isso em outra hora. Voltemos aos Tipos de Séries Temporais.

Linear x Não Linear

Para uma série temporal linear, os pontos de dados $X(t)$ são representados como uma combinação linear de valores passados:

$$X(t) = \alpha_1 X(t-1) + \dots + \alpha_n X(t-n) + \varepsilon(t)$$

onde α que representa um desvio e $\varepsilon(t)$ um ruído branco.

Devido à sua simplicidade, as séries temporais lineares às vezes não conseguem capturar recursos essenciais.

Uma série temporal não linear é muito mais complexa, exibe recursos que não podem ser modelados por função linear: variação de mudança de tempo, ciclos assimétricos, estruturas de momentos mais altos, limites e quebras.

Sazonal x Não Sazonal

Sazonalidade é a presença de variações que ocorrem em intervalos regulares específicos há menos de um ano, representa padrões periódicos, repetitivos e geralmente regulares e previsíveis de período fixo e conhecido (dia, semana, mês, ...).

Se a sazonalidade é estritamente periódica pode eliminar-se da série com um componente determinista. Um truque poderia ser estudar separadamente as séries correspondentes em períodos equivalentes.

Sazonalidade refere-se a flutuações periódicas. Por exemplo, o consumo de eletricidade é alto durante o dia e baixo à noite, ou as vendas de sorvete são maiores no verão e menores no inverno.

Univariada x Multivariada

As séries temporais univariadas dependem de uma única variável, enquanto o conjunto de dados de séries temporais multivariadas contém diversas variáveis dependentes, que são levadas em consideração na previsão de valores futuros.

Neste caso, $X(t)$ inclui várias séries temporais que podem contribuir para prever $Y(t+1)$.

Por exemplo, para séries temporais univariadas, usamos apenas dados históricos de temperatura para prever valores futuros. Para séries temporais multivariadas, combinamos 5 séries temporais (temperatura, nebulosidade, ponto de orvalho, umidade e vento) para fazer uma previsão mais precisa, já que também capturamos padrões das outras variáveis além da variável de temperatura que estamos estudando.

Outro exemplo de uma série temporal multivariada é a previsão do saldo da conta usando como entrada todos os tipos de despesas.

Cada um dos tipos acima tem uma técnica diferente de análise e cada técnica poderia ser um curso por si só, pois há muito conhecimento envolvido em todo processo de análise de séries temporais.

Como é mesmo que aquele instrutor fala nas aulas??.....Ahh sim: “Por isso que o Cientista de Dados deve ser muito bem remunerado”. 😂

Ao final do curso passarei todas as referências, mas para aqueles que estão ansiosos por uma boa fórmula matemática, recomendamos esse excelente material (em português) sobre análise de séries temporais:

<http://www.dme.ufri.br/dani/pdf/slidespartefrequentista.pdf>

A imagem abaixo resume o que estudamos na Aula 3.

Na Aula 4 colocamos Python para trabalhar.

