

Séries Temporelles

14/02/2022



Cronograma

- Manipulação de Séries Temporais com Pandas
- Regressão linear em séries temporais
- Time-step e lag features
- Tendência (Trend)
- Sazonalidade
- Aplicação



Consideração importante

A análise, o pré-processamento, a limpeza e modelagem dos dados, na grande maioria dos casos é MUITO mais importante do que o algoritmo de predição que está sendo utilizado.



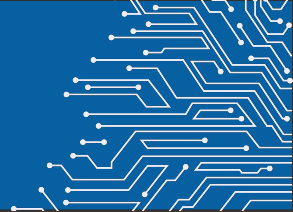
O que são séries temporais?

Uma série temporal pode ser definida como um **conjunto de observações ordenadas no tempo**, como, por exemplo:

- Os valores mensais de temperatura registrados na cidade de São Paulo;
- O salário mensal de uma pessoa durante um ano;
- O preço de ações em um determinado período de tempo.



Manipulação de Séries Temporais com pandas



Antes de avançarmos nos conceitos de Séries Temporais, vamos ver como manipulá-las usando a biblioteca **pandas**!



Regressão linear com Séries Temporais

Uma regressão linear aprende o seguinte:

$$\text{target} = \text{weight_1} * \text{feature_1} + \text{weight_2} * \text{feature_2} + \text{bias}$$

Durante o treinamento, o algoritmo de regressão aprende valores para os parâmetros **weight_1**, **weight_2**, e **bias** que melhor dão o fit com o target. O fit irá minimizar a soma das distâncias ao quadrado entre os dados reais e as previsões.



Regressão linear com Séries Temporais

Uma regressão linear aprende o seguinte:

$$\text{target} = \text{weight}_1 * \text{feature}_1 + \text{weight}_2 * \text{feature}_2 + \text{bias}$$

Durante o treinamento, o algoritmo de regressão aprende valores para os parâmetros **weight_1**, **weight_2**, e **bias** que melhor dão o fit com o target. O fit irá minimizar a soma das distâncias ao quadrado entre os dados reais e as previsões.

Vamos ver agora algumas das features que usaremos:

- time-step features
- lag features



Time-step features

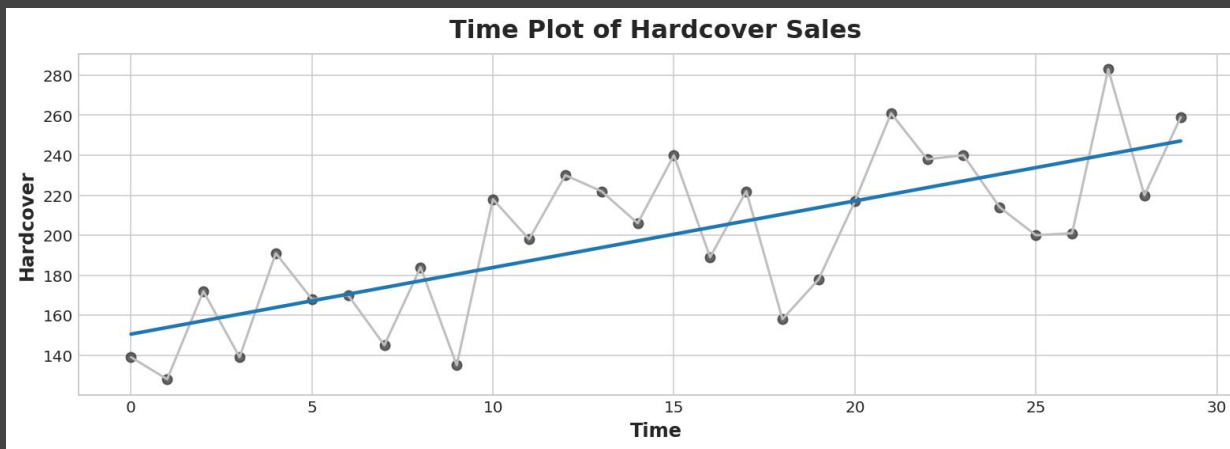
Time-step features podem ser derivadas diretamente do time index. A feature time-step mais básica é a **time dummy**, que conta o número de passos em uma série do início ao fim:

	Hardcover	Time
Date		
2000-04-01	139	0
2000-04-02	128	1
2000-04-03	172	2
2000-04-04	139	3
2000-04-05	191	4



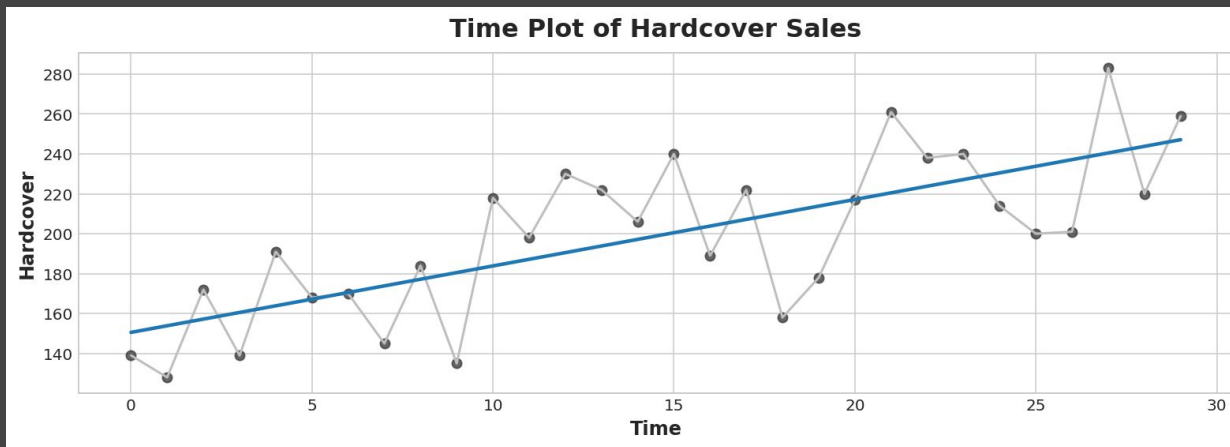
Time-step features

Em seguida, podemos simplesmente colocar o **time-step** no eixo X, as vendas feitas no eixo Y, e plotar a regressão linear:



Time-step features

Time-step features permite uma modelagem de dependência temporal. Uma série segue esse padrão se os seus valores podem ser preditos de acordo com o momento temporal em que aconteceram.



Nesse exemplo, podemos ver que as vendas preditas no final do mês são maiores do que as que ocorrem mais cedo no mês



Lag features

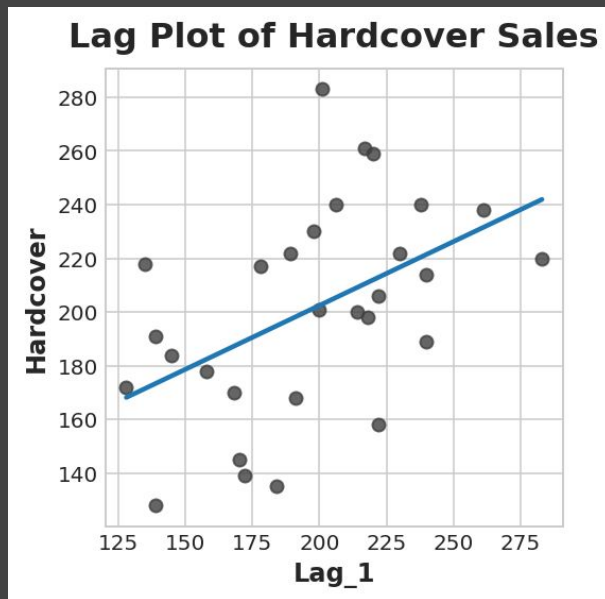
Para fazer uma lag feature, nós deslocamos as observações para elas parecerem como se tivessem ocorrido mais tarde no tempo:

	Hardcover	Lag_1
Date		
2000-04-01	139	NaN
2000-04-02	128	139.0
2000-04-03	172	128.0
2000-04-04	139	172.0
2000-04-05	191	139.0



Lag features

Para fitar uma regressão linear aqui, cada observação será plotada contra a observação anterior:

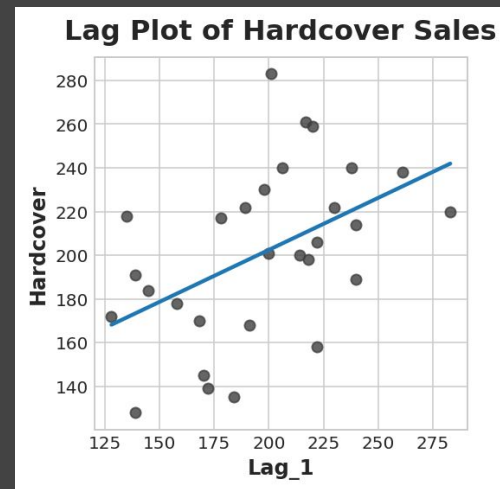


Lag features

Podemos observar uma correlação entre as vendas do dia anterior (Lag_1).

De forma geral, lag features permite modelar dependência serial. Uma série temporal possui esse padrão quando uma observação pode ser predita a partir da observação anterior.

Podemos prever a partir do gráfico que vendas altas em um dia geralmente significam vendas altas no dia seguinte também.



Trends

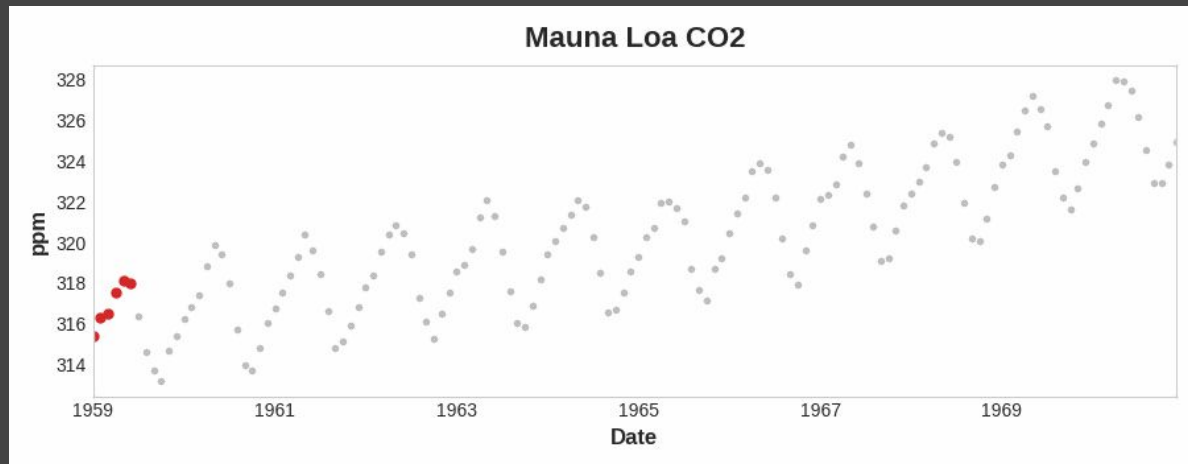
Uma **trend** (tendência) representa uma mudança de longo prazo na série temporal.

Em uma série temporal de vendas de produtos, uma trend crescente pode ser o efeito de uma expansão do mercado conforme mais pessoas consomem o produto a cada ano.



Média móvel

- Média dos valores em uma janela de tamanho n



A ideia é suavizar quaisquer flutuações de curto prazo na série para que apenas as mudanças de longo prazo permaneçam.



Sazonalidade

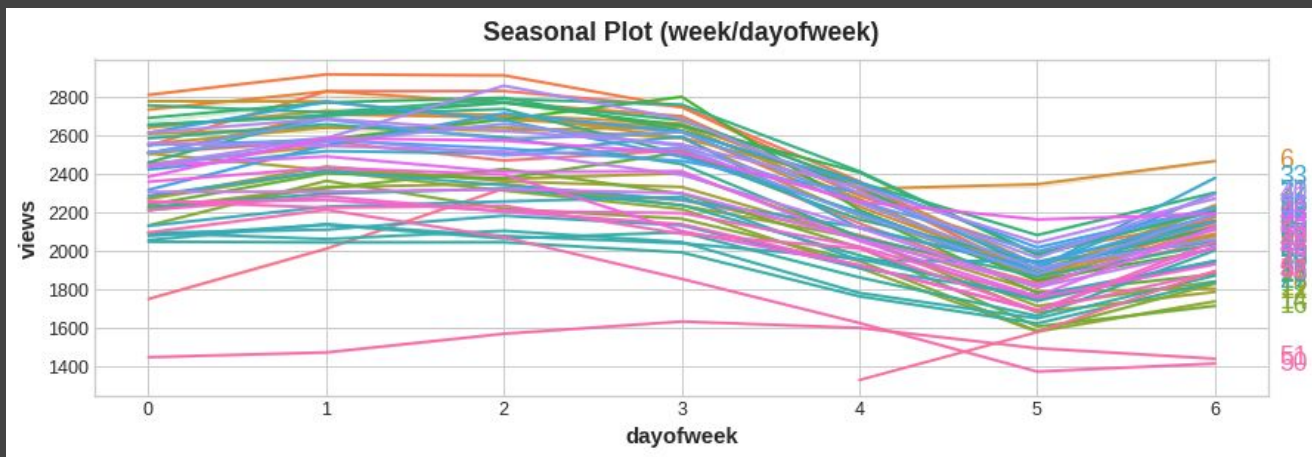
Dizemos que uma série temporal possui **sazonalidade** quando há uma mudança periódica regular na média da série.

Mudanças sazonais geralmente seguem as nossas medidas de tempo, como repetições ao longo de uma semana, ou no ano, por exemplo.



Plots sazonais

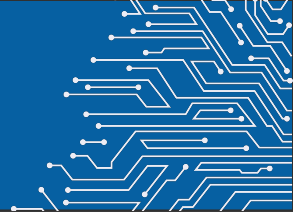
Esse tipo de plot mostra simplesmente uma série VS um período de tempo:



Aqui temos as visualizações de uma página da wikipedia sobre trigonometria ao longo dos dias da semana.



Indicadores sazonais

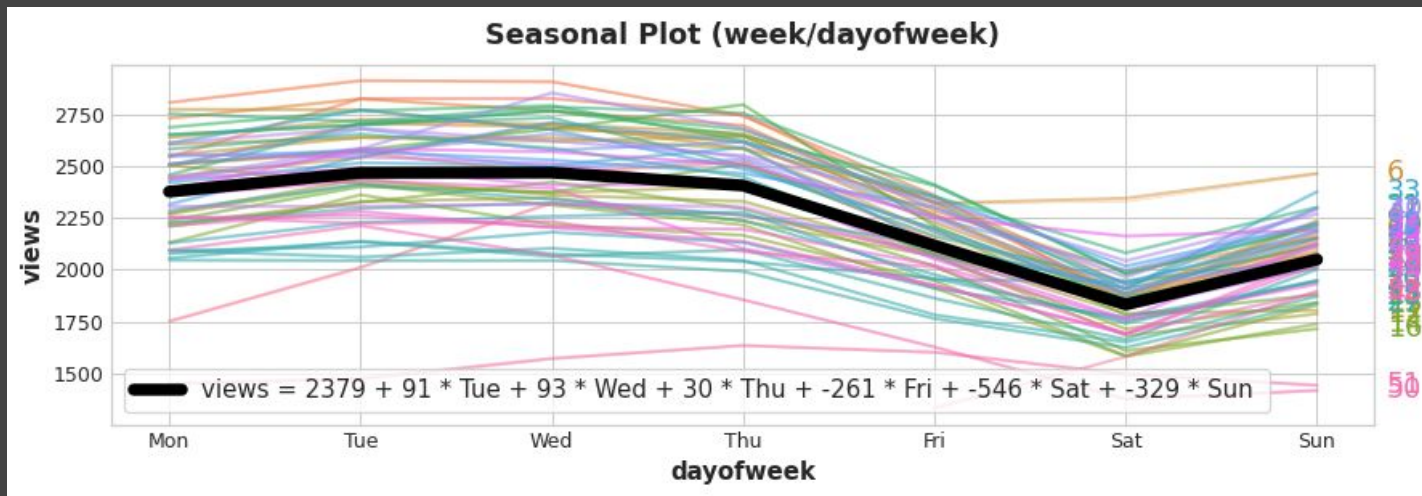


Date	Tuesday	Wednesday	Thursday	Friday	Saturday	Sunday
2016-01-04	0.0	0.0	0.0	0.0	0.0	0.0
2016-01-05	1.0	0.0	0.0	0.0	0.0	0.0
2016-01-06	0.0	1.0	0.0	0.0	0.0	0.0
2016-01-07	0.0	0.0	1.0	0.0	0.0	0.0
2016-01-08	0.0	0.0	0.0	1.0	0.0	0.0
2016-01-09	0.0	0.0	0.0	0.0	1.0	0.0
2016-01-10	0.0	0.0	0.0	0.0	0.0	1.0
2016-01-11	0.0	0.0	0.0	0.0	0.0	0.0
...



Indicadores sazonais

Adicionando os indicadores no treinamento do modelo, ficamos com o seguinte resultado:



Cronograma seguinte

- Processo de criação uma regressão em time series
- Modelos de regressão geral em Machine Learning
- Modelos clássicos de regressão de time series
- Facebook Prophet
- Métricas de Avaliação de modelos de regressão
- Previsão em time series de maneira recorrente

