



I EXPOTECH

Ênfase em Análise de Dados em *Big Data*

07-09 Novembro 23

EXPLORANDO A RELAÇÃO ENTRE ATIVOS FINANCEIROS E GUERRAS

Angelo Souza de Oliveira



I Exposição de Projetos de Extensão de Tecnologia da Informação do Campus UNESA - Niterói

Introdução

Este trabalho de extensão, apresentado na disciplina extensionista de Tópicos de Big Data, tem como objetivo realizar uma análise de Regressão Simples, acompanhada de uma análise descritiva, abordando a complexa relação entre a incidência dos últimos 5 anos da palavra 'Guerra' no Google Trends e os preços semanais entre outubro de 2018 a outubro de 2023 das ações da PETROBRAS (Petróleo Brasileiro SA), do IBOVESPA (Índice Bovespa) e do índice do *Dollar DXY* (Índice do Dólar Norte-Americano). Ao considerar os impactos das guerras na geopolítica e as instabilidades nas regiões produtoras de petróleo, como o Oriente Médio e a Rússia, essa análise busca corroborar com o padrão dos impactos das guerras na volatilidade dos ativos financeiros citados acima.

Big Data e Regressão Linear

Big Data é uma área que se dedica a explorar, analisar, processar e armazenar conjuntos extensos de informações, frequentemente originados de fontes diversas. Dentro desse contexto, a análise de dados se configura como uma colaboração interdisciplinar que abrange várias áreas do conhecimento e oferece a base fundamental para a coleta de dados, a detecção de padrões e o tratamento de informações, fundamentando o processo de tomada de decisão.

Além disso, a utilização da regressão linear é um método estatístico que visa avaliar as interações entre variáveis independentes e dependentes, proporcionando *insights*, sejam eles de natureza preditiva ou não.

Metodologia

A metodologia aplicada no presente trabalho envolve o uso do *Python 3.8* em um ambiente de programação com a interface interativa do *Jupyter Notebook*. Utilizou-se a biblioteca 'Pandas' para a análise e manipulação de dados de arquivos CSV por meio de suas funções. O *NumPy* foi empregado para facilitar o empilhamento de *array*, bem como na transformação de *dataframes* em *arrays NumPy* por meio da função *to_numpy()*. Por fim, utilizou-se a função *polyfit()* para o cálculo dos coeficientes angulares e interceptos. Adicionalmente, as bibliotecas *Matplotlib* e *Seaborn* foram empregadas para a geração de gráficos estatísticos. Os dados brutos dos ativos financeiros foram obtidos do Yahoo Finance, e os registros no mundo da palavra "Guerra" foi retirado do Google Trends.

Resultados da Regressão Linear com Python

O código a seguir lê um arquivo CSV, realiza a limpeza dos dados, renomeia uma coluna e calcula a média, desvio-padrão, mediana e moda. Esse processo é repetido para os dados da PETR4, Índice do Dólar e para a ocorrência da palavra "guerra":

```
df_ibov = pd.read_csv('./data/ibov.csv').dropna().iloc[: -1]
df_ibov = df_ibov.rename(columns={'Close': 'preco_fechamento'})
ibov_mean = df_ibov['preco_fechamento'].mean()
ibov_desvio = df_ibov['preco_fechamento'].std()
ibov_mediana = df_ibov['preco_fechamento'].median()
ibov_moda = df_ibov['preco_fechamento'].mode()
```

O cálculo da regressão linear foi realizado com três séries temporais em *X_series*, que são transpostas. *Y_word_war* é a série temporal da ocorrência da palavra "guerra". Um loop calcula a regressão linear simples entre cada série (*X*) e a série de ocorrência da palavra "guerra" (*Y*), obtendo os coeficientes angular e o intercepto da reta:

```
X_series = [X_ibov, X_petr4, X_dolar_dx]
X_matrix = np.vstack(X_series).T
Y_word_war = df_word_war['data_word'].to_numpy()
coeficientes, interceptos, y_pred = [], [], []
for X in X_series:
    coef_angular, intercepto = np.polyfit(X, Y_word_war, 1)
    coeficientes.append(coef_angular)
    interceptos.append(intercepto)
    y_pred.append(coef_angular * X + intercepto)
```

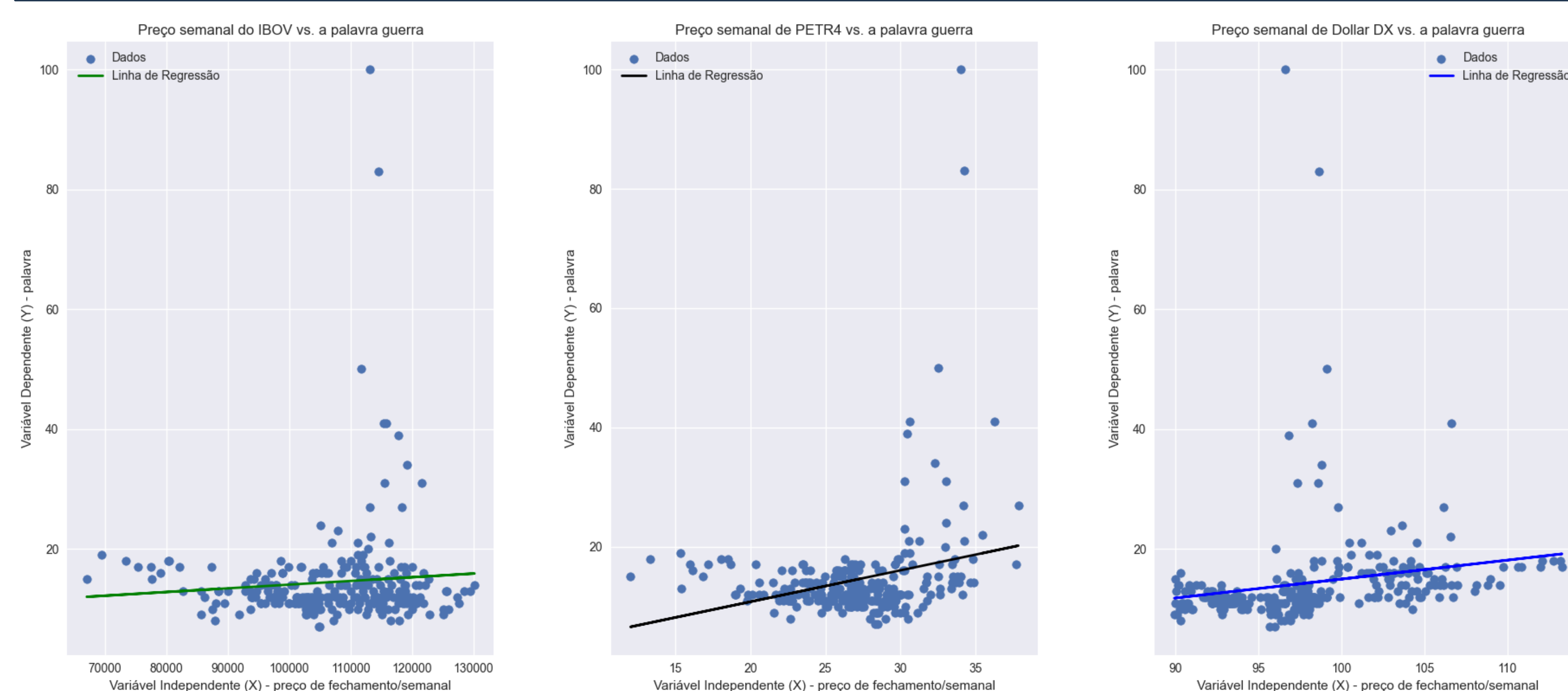


Figura 1. Gráfico de Regressão Linear resultante; IBOVESPA, PETR4 e DOLLAR INDEX.

Conclui-se que existe uma correlação positiva entre o preço das ações da PETR4, o *Dollar Index* e a incidência da palavra "guerra". No entanto, a correlação é mais fraca com o Ibovespa devido à representação de apenas 7,061% das ações da PETR4 no índice IBOVESPA, resultando em uma correlação positiva, porém modesta.

Referências

YAHOO FINANCE. Yahoo Finance. Disponível em: <<https://finance.yahoo.com/>>. Acesso em: 28/10/23.

Erl, Thomas; Khattak, Wajid; Buhler, Paul. Capítulo 1. In: **Fundamentos de Big Data: conceitos, motivadores e técnicas**. São Francisco: O'Reilly Media, 2016.

Disponível em: <<https://learning.oreilly.com/library/view/big-data-fundamentals/9780134291185/ch01.xhtml#ch01>>. Acesso em 28/10/2023.