

**Final Report:**  
**Analysis of Geopolitical Events on Oil Prices**  
**Sraavika Vaddadi**

**Introduction:**

Crude oil is one of the world's most important commodities. It plays a role in shaping everything from global trade and national security decisions, to impacting our everyday individual expenses. This makes crude oil sensitive to changes in the geopolitical landscape such as wars or international threats. Through this project, I aim to understand just how sensitive WTI and Brent crude oil prices are to international political conflicts and to observe those trends over time. I also hope to gain deeper understanding of the most significant geopolitical events and the repercussions from them.

**Data:**

For this project, I utilized two datasets:

1. EIA Petroleum Spot Prices:

The EIA, or Energy Information Administration, has an open-source API that I scraped to get data on crude oil prices for my project. In order to do this, I had to sign up to receive my API key and had to write a request code to the API to get the spot pricing data. To ensure my analysis was globally-oriented, I used the two main global benchmarks for oil pricing – WTI or West Texas Intermediate priced in Cushing, Oklahoma or the Brent Crude priced on North Sea production.

This dataset starts from 1986 and is updated daily. During the time I scraped the data, there were 918 rows excluding the header and 11 columns, but only three are relevant to my project – Period, Product and Price. The scraping file, 00\_EIA-API\_scraping\_raw.ipynb, can be found in the source data folder.

There was not much data preprocessing and cleaning required for this dataset. My only changes were converting the Period column to a date-time format and creating some calculated fields like the logarithmic price and normalizing the value. The cleaned dataset was saved as a CSV called price\_data.csv and can be found in the data final folder.

2. Geopolitical Risk Index:

The Geopolitical Risk Index was developed by two Federal Reserve Board economists - Dario Caldara and Matteo Iacoviello. Their methodology is to measure the proportion of newspaper articles covering geopolitical tensions to the monthly total articles of several major newspapers.

At the time I downloaded the GPR CSV file from Matteo Iacoviello's website, there were 115 columns and 1499 rows including the header row. This was a very extensive dataset and required several steps of data cleaning like renaming columns, filtering to include data only from 1986 to match the EIA dataset, converting datatypes, and creating logarithmic and normalized calculated fields.

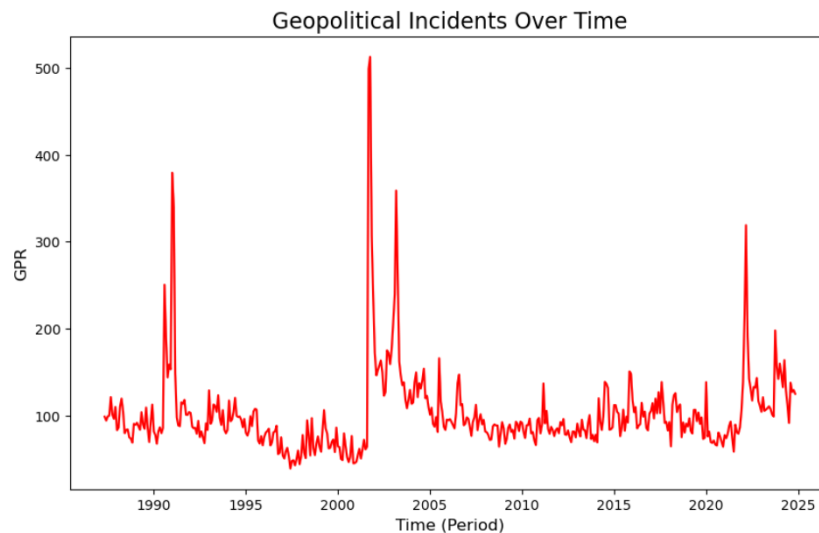
The two datasets were inner merged using the period column to ensure that only the rows with a GPR score and oil price were kept. Following the merge, the dataset comprised of 902 rows and 20 columns and was saved as a new CSV file named gpr\_price.csv and can be found in the final data folder.

## Data Dictionary

Field Name	Data Type	Source	Description
Period	Date	Both	The month and year for pricing, GPR, and crisis data
Product	String	EIA	The type of petroleum – WTI or Brent
Price	Numeric	EIA	Price of oil
GPR	Numeric	GPR	The measure of geopolitical risk
GPR_Threats	Numeric	GPR	The measure of geopolitical threats
GPR_Acts	Numeric	GPR	The measure of geopolitical threats that escalated into acts
Cat_1	Numeric	GPR	Frequency of articles on War Threats
Cat_2	Numeric	GPR	Frequency of articles on Peace Threats
Cat_3	Numeric	GPR	Frequency of articles on Military Buildups
Cat_4	Numeric	GPR	Frequency of articles on Nuclear Threats
Cat_5	Numeric	GPR	Frequency of articles on Terror Threats
Cat_6	Numeric	GPR	Frequency of articles on Beginning of War
Cat_7	Numeric	GPR	Frequency of articles on Escalation of War
Cat_8	Numeric	GPR	Frequency of articles on Terror Acts
GPR_Articles	Numeric	GPR	Proportion of total articles categorized in the GPR index
Total_Articles	Numeric	GPR	Total number of articles published by the 10 major newspapers

## Analysis:

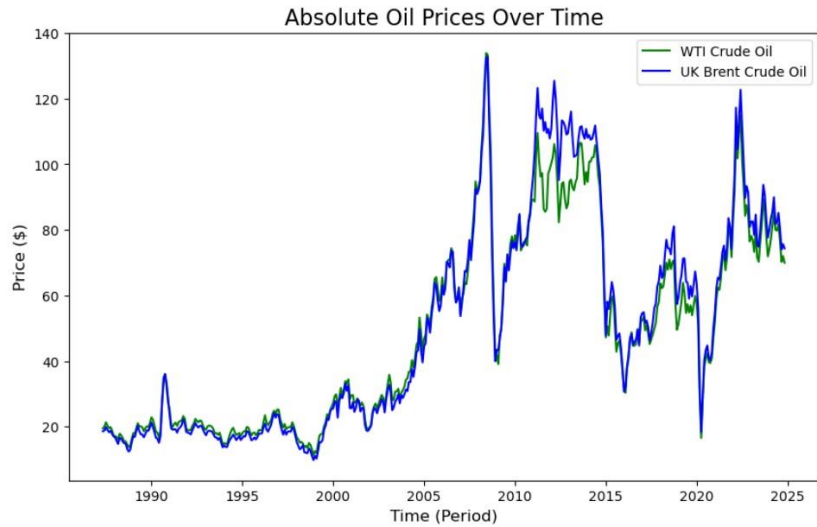
### *Geopolitical Incidents Over Time*



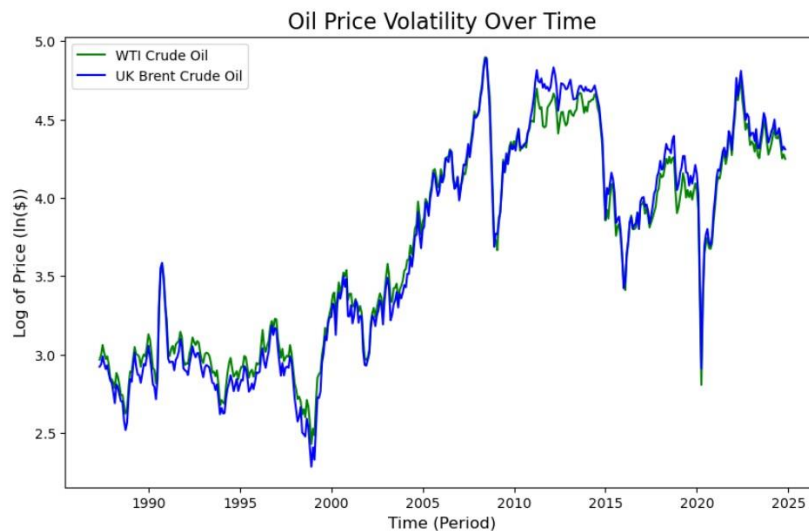
*Figure 1 GPR Index visualized over time*

The first thing I hoped to understand was what were the major geopolitical incidents within the time frame of my dataset. Through this line chart, it is easy to identify that there are four significant events that resulted in a spike in geopolitical risk and can be attributed to the Gulf War in 1990, the September 11 attacks and following outbreak of the Afghanistan war in the early 2000s, and lastly the Russian Invasion in Ukraine and the Israel and Palestine war in 2022 and 2023 respectively.

## *Relationship Between WTI Crude and Brent Crude Prices*



*Figure 2 Absolute Oil Prices*



*Figure 3 Logarithmic Oil Prices*

Both WTI and Brent crude are representative of two pricing benchmarks crucial for the global market. However, as they are based in different producing regions, I analyzed the relationship between these two prices to later conclude if geopolitical risk affects the prices in different ways.

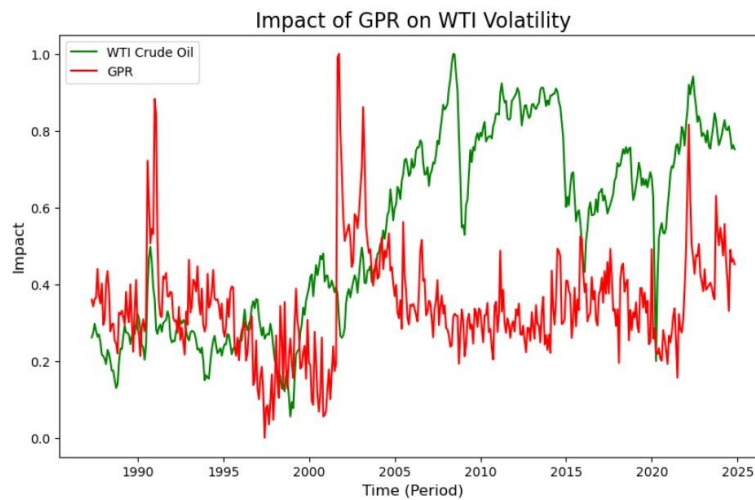
Originally, I plotted both prices on an absolute scale, but later realized that using a logarithmic scale is more representative of volatility. Following the shale revolution in 2009, oil prices were trending from \$60-120 range unlike pre-shale discovery which trended between \$20-40 as seen in Figure 2. The absolute pricing scale is misleading to viewers as it looks like there were drastic drops in the more recent years and remained relatively flat before the 2000s.

After visualizing both crude grades together, the analysis surprised me. It looks like there is minimal variance between the two crude grades and the only noticeable deviation was in the 2010 to 2014 range which again can be attributed to the aftereffects of the US shale revolution. As there was an excess of

supply in the US, WTI prices were trading below Brent. To verify this relationship beyond the visual, I also calculated the Pearson correlation and found a 99% correlation implying a strong linear correlation.

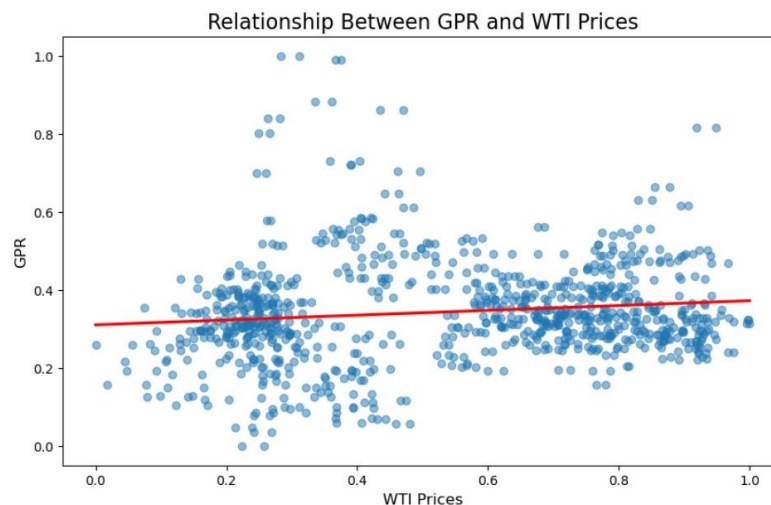
### ***Analyzing Geopolitical Effects on Crude Prices***

Following my understanding of both datasets individually, I next analyzed the relationship between the geopolitical risk index and crude prices. After concluding that there is a strong linear relationship between WTI and Brent Crude, for simplicity, my further analysis just uses WTI. First, I visualized the normalized logarithmic scales of both features over time to see if there is a noticeable pattern over time. However, I discovered that the relationship looks more complex than I originally thought. As seen in Figure 4, it is not easily deducible if there is a linear relationship between GPR and crude prices.



*Figure 4 GPR and WTI Over Time*

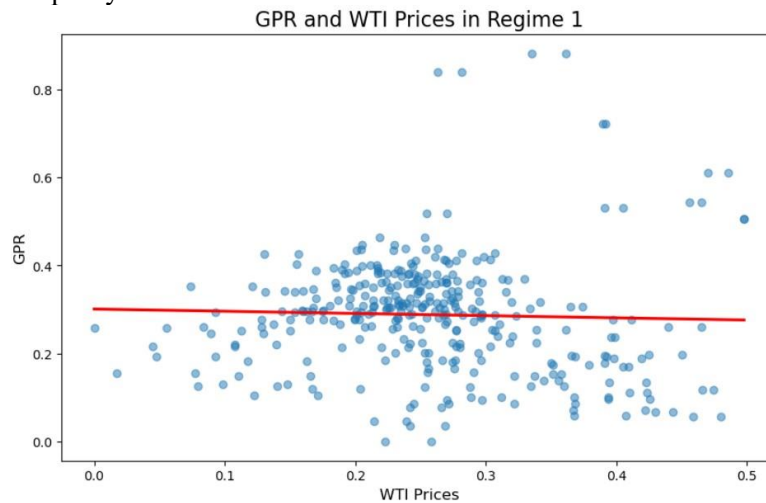
Instead, I created a scatter plot to further decode this relationship. I observed that the data was naturally being grouped. From here I decided to individually analyze the GPR WTI relationship based on the groups that I categorized as regime 1 and 2.



*Figure 5 Linear Relationship between GPR and WTI*

### ***Regime 1 Correlation***

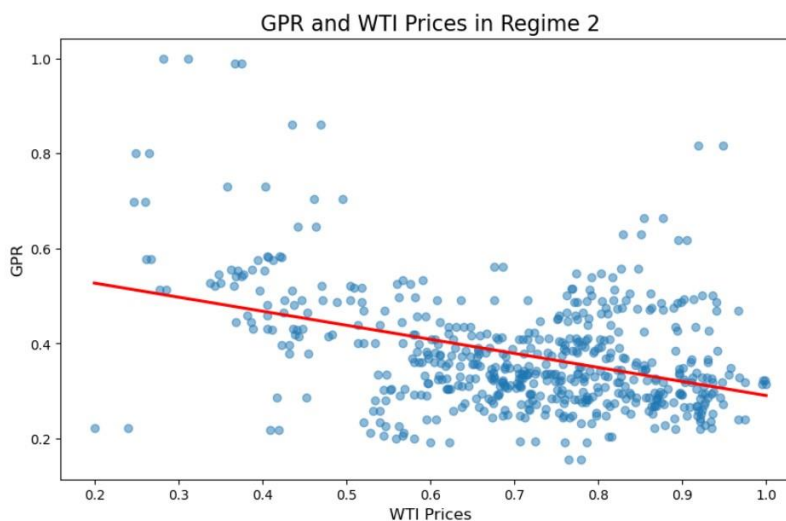
Upon further analysis, I realized that the natural groupings were occurring during the time of the September 11 terrorist attacks and accordingly split the data. Regime 1 consists of data prior to these attacks and after performing a Pearson correlation, there is a negative correlation of 16.8%. Additionally, after calculating the p-value, 0.4675, it can be concluded that there is no statistically significant relationship, and the correlation is weak and not meaningful. This can also be seen in Figure 6 as the regression line remains pretty flat.



*Figure 6 Linear Relationship in Regime 1*

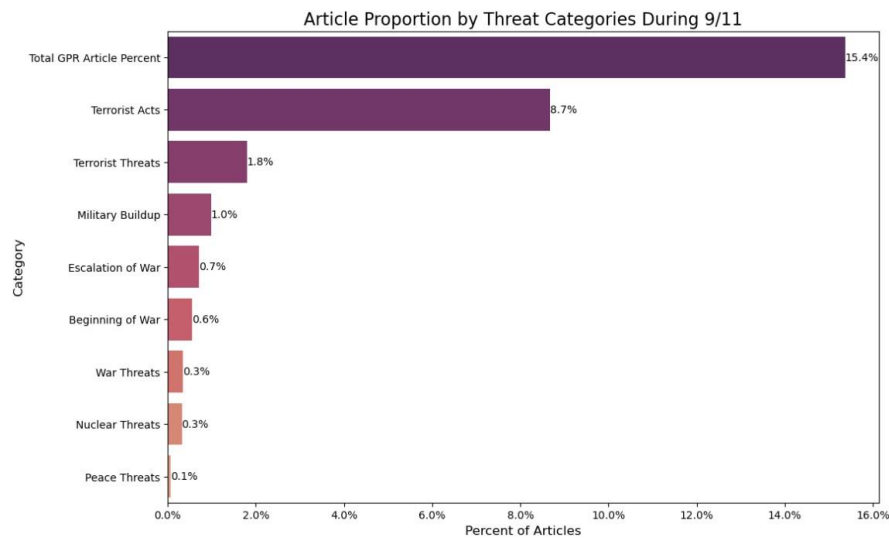
### ***Regime 2 Correlation***

Next, I examined the relationship in regime 2 or post September 11 attacks. It can be seen that the relationship is a little more observable in Figure 7. There is a Pearson correlation of negative 48.16% and the p-value is extremely low. This implies that there is a moderate negative relationship between both features and that the correlation is statistically significant.



*Figure 6 Linear Relationship in Regime 2*

### ***Further Analysis on the Biggest Geopolitical Event***



*Figure 7 News Coverage Categories in Oct. 2001*

After all the above analysis, I was curious to see why, despite the 9/11 terrorist attacks to this day having the highest geopolitical risk score, oil prices had a less drastic drop going from \$26 in September to reaching its lowest point in November at \$19.64. While this is nearly a 25% change, I expected there to be a greater impact seeing as the attacks were in New York, a global financial hub and where oil futures are traded on the markets.

Looking at the news coverage during the attacks and what categories these articles fell under, it looks like there was some speculation at the time that there would be military buildup and potentially even a war. Figure 7 shows the breakdown of articles by category in October which had the highest news coverage regarding the attacks. However, it is possible that prices were not as drastic as the allegations and responsible terrorist groups were based in Afghanistan which is not in OPEC or the Organization of Petroleum Exporting Countries.

Whereas looking at the case of the Russian invasion of Ukraine, there was a much more significant price impact as Russia is a major exporter of oil and natural gas to European markets and is within the top 5 producing oil countries in the world. There was a much bigger impact and concern of global oil supply leading to more significant price fluctuations.

### **Conclusion**

In conclusion, there is still room for further and more statistical analysis on the relationship between geopolitical events and oil prices, but my project is the first step in exploratory analysis on this topic.

My biggest findings were that there were 4 major geopolitical events since 1986 but have varying effects on oil prices. For example, the 9/11 attack was the biggest geopolitical event and had the greatest amount of news coverage at nearly 16% of all articles published that month, but oil prices were not as drastically impacted relative to other incidents such as the Russia Ukraine war.

Additionally, I found that there is a very significant correlation between Brent and WTI oil prices despite representing two separate producing regions.

There is also a natural grouping of oil and GPR pre and post 9/11 attacks. This can be due to a shift in media and news coverage following the attacks, entering a new period of pricing and the increase in geopolitical events in significant oil producing regions leading to a higher correlation in recent times compared to pre 9/11 times.

Some limitations for this analysis include a shift in media and news coverage. The GPR index is based on how many news articles contain certain words and phrases relating to geopolitics that can be attributed to events and incidents occurring across the world. However, with a changing digital age and how and where we get our information, there is some room for speculation on whether the GPR index is still an accurate measure. Future work for this project is a more statistical in-depth analysis of the relationship between oil and geopolitical events.

**Works Cited:**

Caldara, Dario and Matteo Iacoviello (2022). “Measuring Geopolitical Risk,” International Finance Discussion Papers 1222r1. Washington: Board of Governors of the Federal Reserve System, <https://doi.org/10.17016/IFDP.2022.1222r1>.