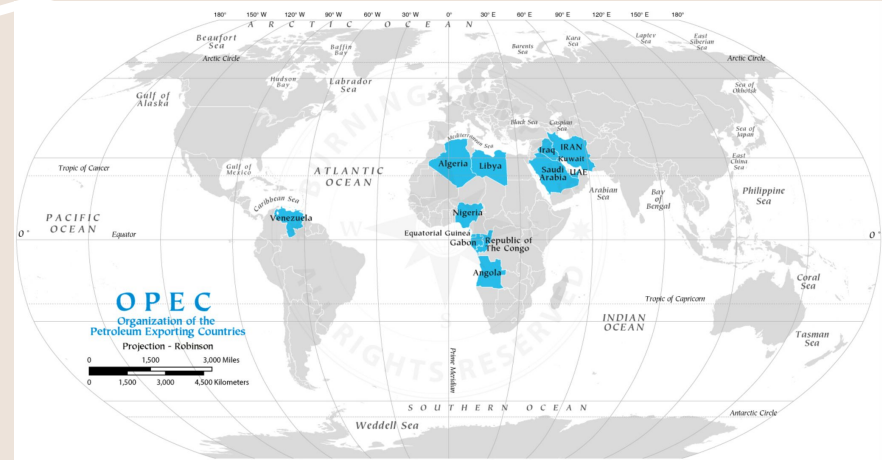


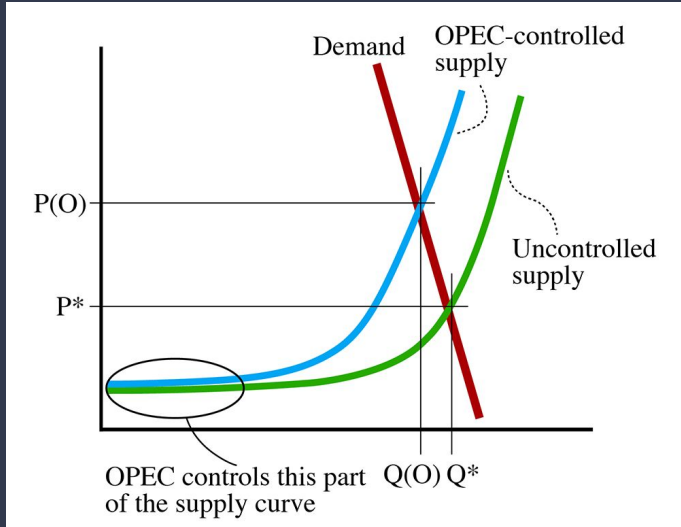
# Analyzing OPEC Member Crude Oil Production Quotas

Will Rodman

[willcrodman.github.io](https://willcrodman.github.io)



# Problem

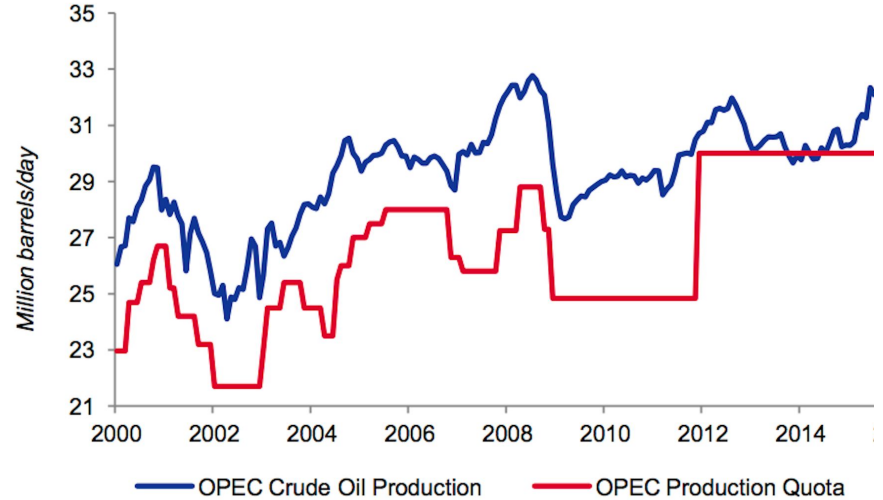


OPEC is an international pact of 13 petroleum exporting nations.

The organization's primary object is to control the crude oil supply of its member nations.

Behaves like **cartel supplier** in microeconomics.

# Question



OPEC members consistently deviate from crude oil supply quotas set by the cartel, causing overproduction (or underproduction).

Can the magnitude of overproduction for individual members be modeled?

# Data Wrangling

Datasets include six crude oil economic indicators by nation and year from 1960 to 2022; including OPEC and non-OPEC nations.

Sourced directly from OPEC's archive at [asb.opec.org](https://asb.opec.org). Original format are Microsoft Excel files.

Loaded into Juniper Notebooks as six separate Pandas DataFrames.

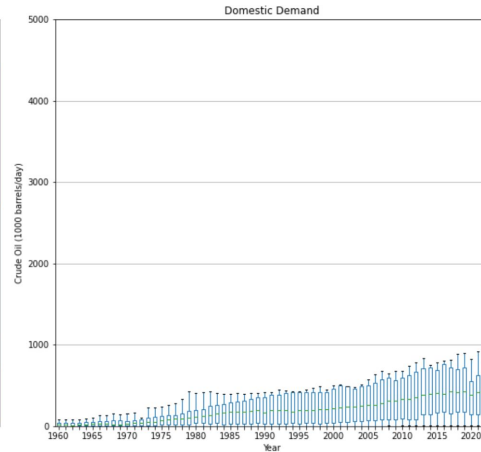
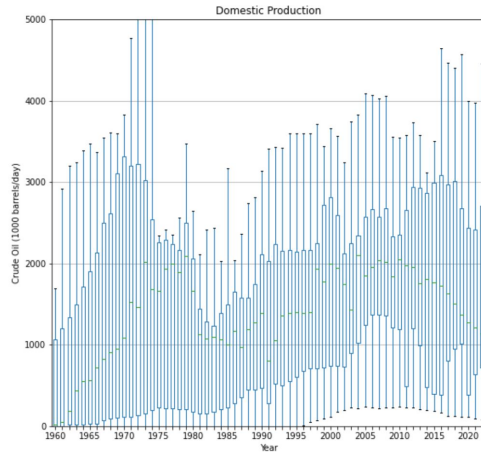
<b>Economic Indicator</b>	<b>Date Range</b>	<b>Nation Count</b>
Crude Oil Demand	1960 - 2022	59
Crude Oil Production	1960 - 2022	64
Crude Oil Refinery Capacity	1980 - 2022	54
Refinery Crude Oil Throughput	1980 - 2022	69
Crude Oil Spot Price	1983 - 2022	21
Crude Oil Production Quota*	1982 - 2022	13

\* Only applicable for OPEC member nations.

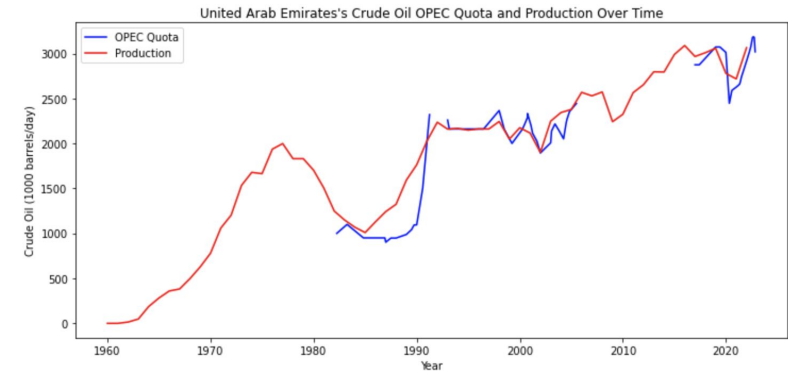
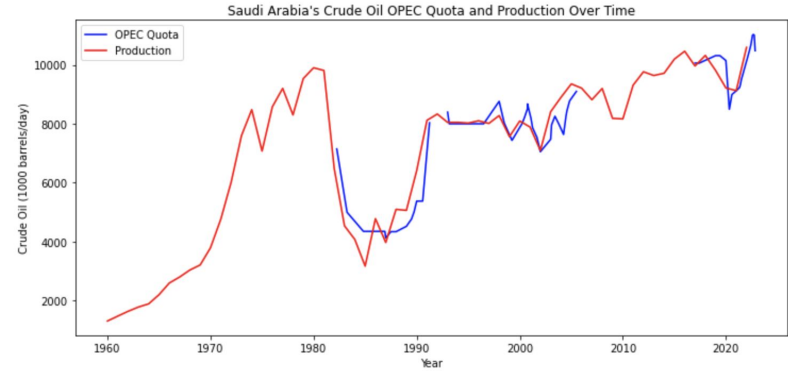
# Data Analysis

Many OPEC nation data points missing across datasets.

Data visualization of crude oil production, demand, and production quota datasets:



OPEC Member Countries Crude Oil Metrics



# Model: Random Forest Regression

Predicts individual nations overproduction for a single year.

Trained and tested on 54 observations from Saudi Arabia and the United Arab of Emirates.

Feature Space:

1. **Domestic Demand:** OPEC nations domestic demand for crude oil.
2. **Refinery Throughput:** The amount of crude oil refined into final products domestically by an OPEC nation.
3. **Spot Price Premium:** The percentage a nation's crude oil spot price is above (or below) the OPEC Oil Reference Basket.

Letting overproduction be a percentage:

$$\text{Overproduction}_{\text{year}} = \frac{\text{Target}_{\text{year}} - \text{Production}_{\text{year}}}{\text{Production}_{\text{year}}}$$

The regression model is defined as:

$$\hat{\text{Overproduction}}_{\text{year}} = \frac{1}{100} \sum_{i=1}^{100} h_i(x)$$

$100 := \text{Number of Decision Trees}$

$h_i(x) := \text{Decision Tree Prediction}$

# Conclusion

Feature Importance:

1. Demand: **34%**
2. Throughput: **34%**
3. Premium: **32%**

**Possible Improvements:**

Increase number of nation observations or model only aggregate OPEC data.

