

STAT 350: Dataset Description

The dataset provided for student use has been modified and sourced primarily from an oil price prediction contest hosted on [Kaggle](#). Kaggle is a popular platform where users can share data, code, and engage in discussions related to machine learning and data science.

The objective of this dataset is to analyze the relationship between historical prices and various factors related to global oil and natural gas production, consumption, and prices. The dataset has been preprocessed to make it more accessible. To address the inherent variability in oil prices, we have performed a natural logarithm transformation on several variables. This transformation helps in stabilizing the data and allows for more meaningful analysis and predictions.

Students can utilize this dataset to explore the dynamics between historical oil and natural gas data, including production, consumption, and prices. They can apply various analytical techniques, such as data visualization, and statistical analysis, to gain insights into the factors influencing oil prices and make future price predictions.

Please note that the dataset has been modified from its original form for educational purposes. It provides a valuable opportunity for students to apply their knowledge of data science and machine learning to real-world energy market scenarios.

The following table describes the columns/variables in this dataset:

Variable Name	Dataset Column	Description
CtrName	1	Country name.
CtrCode	2	Country ISO code.
ContName	3	Continent name.
Year	4	The year the data was collected.
IPop	5	Log population of the country. Note that the first character is a small 'el.'
IGasPrice1	6	Log import natural gas price per MCM (thousand cubic meters) converted to 2000 using a predicted constant inflation factor in USD (US Dollars). Note that the first character is a small 'el.'
IGasPrice2	7	Log import natural gas price per MCM converted to 2000 using the actual inflation factor in USD. Note that the first character is a small 'el.'
GasProd01	8	Whether or not the country produces natural gas.
GasProd	9	Yearly natural gas production amount in MCM
GasCons	10	Yearly natural gas consumption amount in MCM.
GasVal	11	Yearly nature gas consumption value.
IOilPrice	12	Log oil price per barrel converted to 2000 using the actual inflation factor in USD. Note that the first character is a small 'el.'
OilProd01	13	Whether or not the country produces oil.
OilProd	14	Yearly oil production amount in barrels.
OilCons	15	Yearly oil consumption amount in barrels.
OilVal	16	Yearly oil consumption value.

Disclaimer: the original datasets were modified and changed to suit the purposes of STAT 350. No serious conclusions should be drawn from the analysis.