The effect of oil prices on the economies of major oil exporters

December 2023

Table of Contents

Introduction

In our project, we aim to explore the relationship between crude oil prices, GDP, the inflation of G20 and OPEC countries, and export values. As a proxy variable for crude oil prices, we are using West Texas Intermediate (WTI), a specific type of crude oil primarily produced in the United States, particularly in West Texas. WTI is a widely recognized benchmark for crude oil pricing worldwide. Our objective is to delve into the dynamics of crude oil pricing and understand how it interacts with various factors, including exports and economic conditions. At the end of the project, we will present visualizations and correlations between key economic indicators.

Dataset

The source of national economic and environmental footprint data is the World Bank, Inc., including:

- [Electricity production from coal sources](#) (% of total)
- [Energy use (kg of oil equivalent per capita)](#)
- [CO2 emissions from liquid fuel consumption](#) (kt)
- [CO2 emissions from manufacturing industries and construction](#) (% of total fuel combustion)
- [Exports of goods and services](#) (% of GDP)
- [Inflation, GDP deflator](#) (annual %)
- [GDP per capita, PPP](#) (current international $)

API

We have enriched the economic and environmental indicator data with real asset prices such as commodities using the Yahoo Finance API. West Texas Intermediate (WTI) serves as a proxy variable for oil prices as it is the world's largest producer.

Navigation

- [Github Repo](#)
- The CSV files are available [here](#).
- SQL dumps for [importing](#) and [transforming](#) data using pipeline.
- ER diagram that is indicating the [relationship between economic indicators](#) and [tables after transformation](#)
- Knime [pipeline](#).

Setup

1. Upload the data in MySQL using [Data_Dump_2.sql](#).
2. Run the code from [Data_Pipeline.sql](#)
3. Download [Knime pipeline](#)
4. Setup your SQL connection using your credentials
5. Run the pipeline
6. Check the resulting correlation matrices.
7. Check the visualizations

Technical choices

The data is provided on an annual granularity as the World Bank provides data only on an annual basis, making it infeasible to assess monthly performance on a country level.

We have included two additional columns, one indicating G20 membership and another for OPEC country status, allowing us to conduct a correlation analysis for each group.

As a result, we have created one visualization and one correlation matrix for each scenario, depending on whether the country belongs to the G20 or OPEC. For the correlation analysis, we have primarily relied on data concerning inflation, exports, and WTI, as many other indicators contained a significant number of missing values, making it challenging to obtain dependable results.
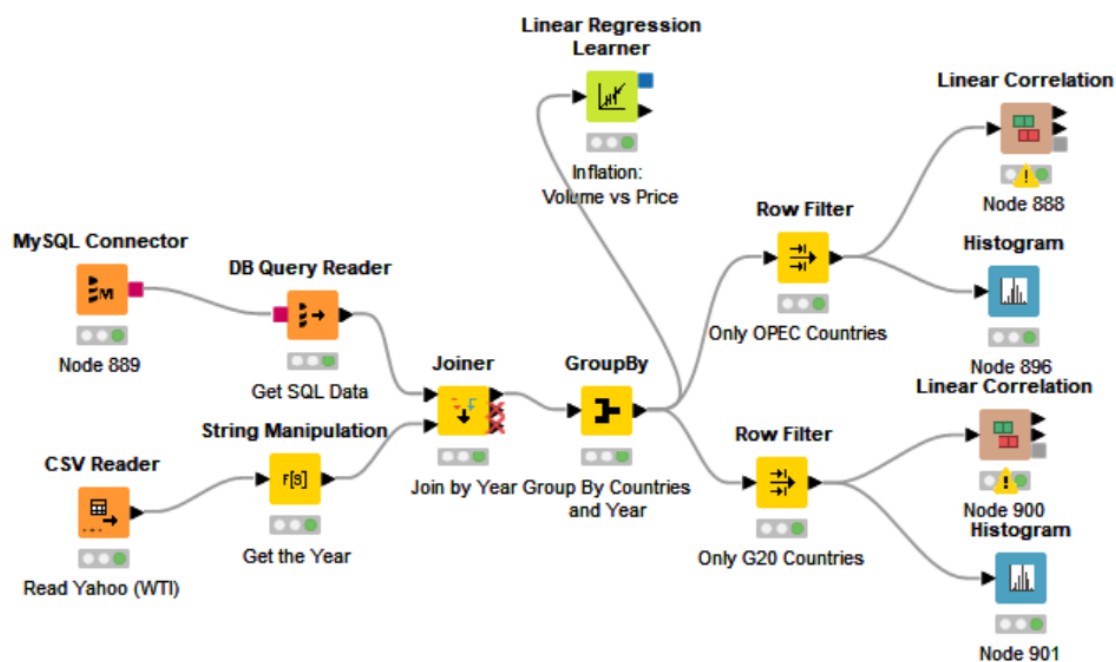
EER diagram of base data
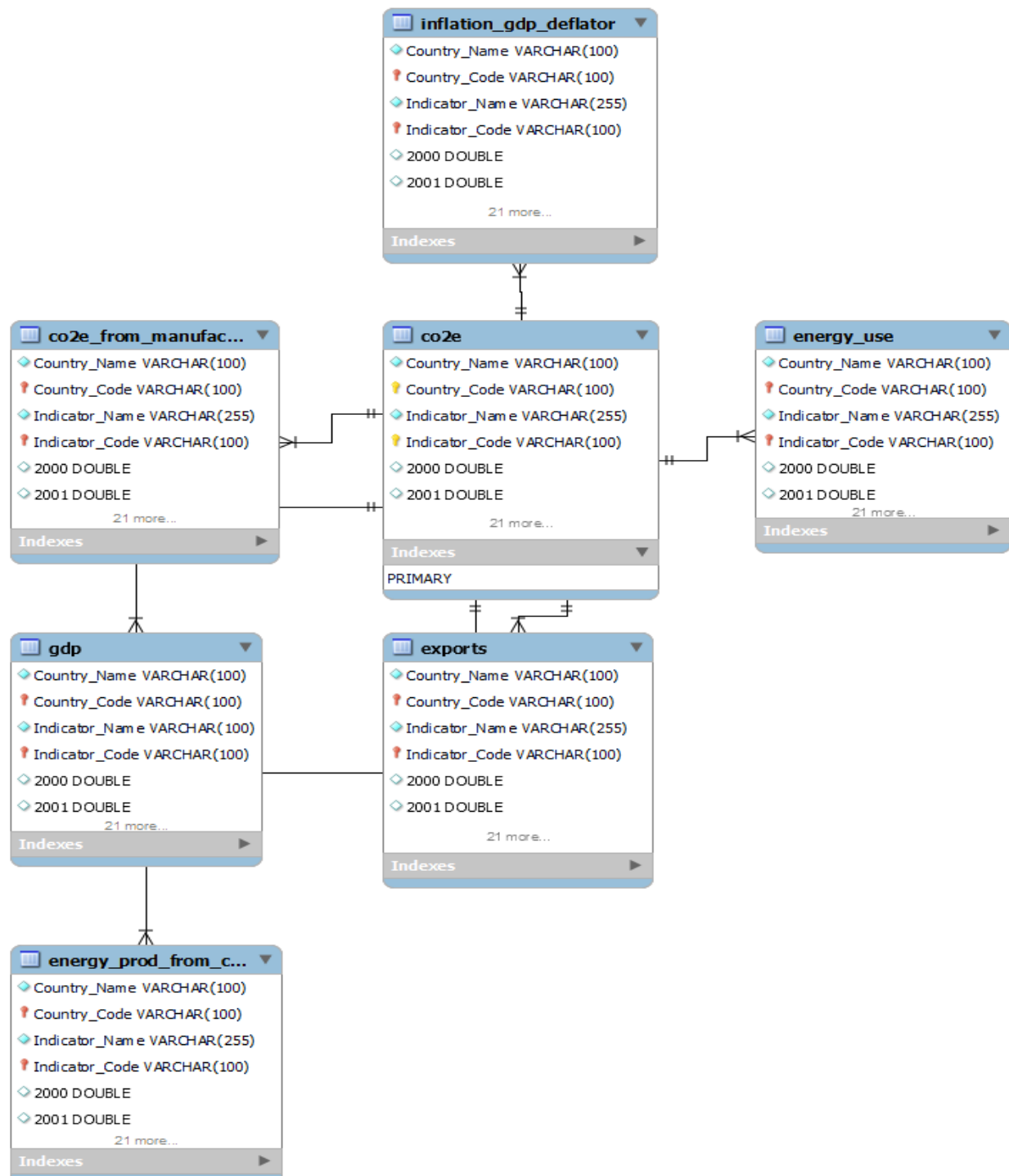
# EER diagram with API data
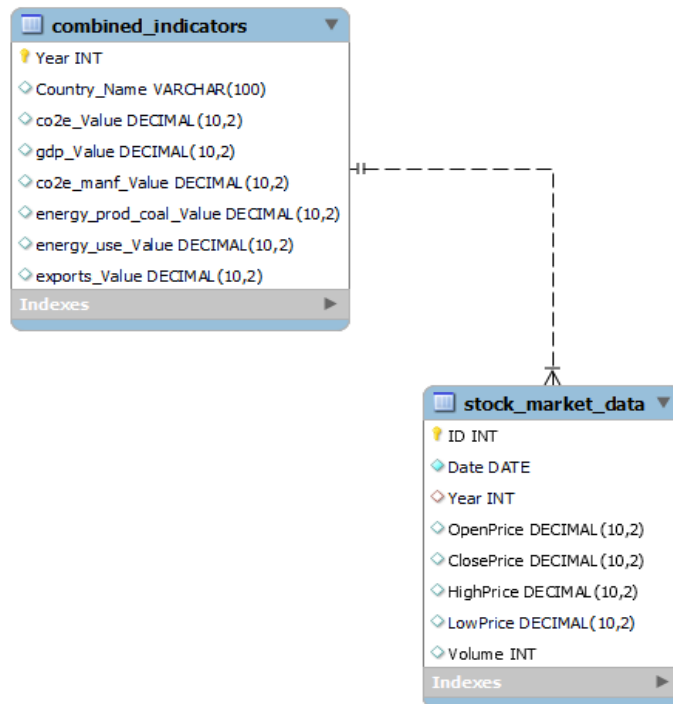
Knime Pipeline

Pipeline Explained

We have used a MySQL connector to connect Knime with our database. The pipeline includes the following steps:

1. Get the transformed analytical table using DB Query Reader.
2. Obtain WTI data using the Yahoo Finance API.
3. Add the year as a column using a string manipulation node.
4. Join these two datasets using the year column from both tables.
5. Group the joined dataset by year and country to get one observation per country and year. Note that we used average values to aggregate WTI.
6. Using a row filter based on columns is_G20 and is_OPEC, we created two scenarios.
7. For each scenario, we built a correlation matrix and histogram visualization.

Results

OPEC Countries

Based on the correlation matrix for OPEC countries, we observe that the correlation between exports and the average WTI, as well as the correlation between inflation value and the average WTI, is positive and moderate. Also, there is a strong positive correlation between GDP and exports.
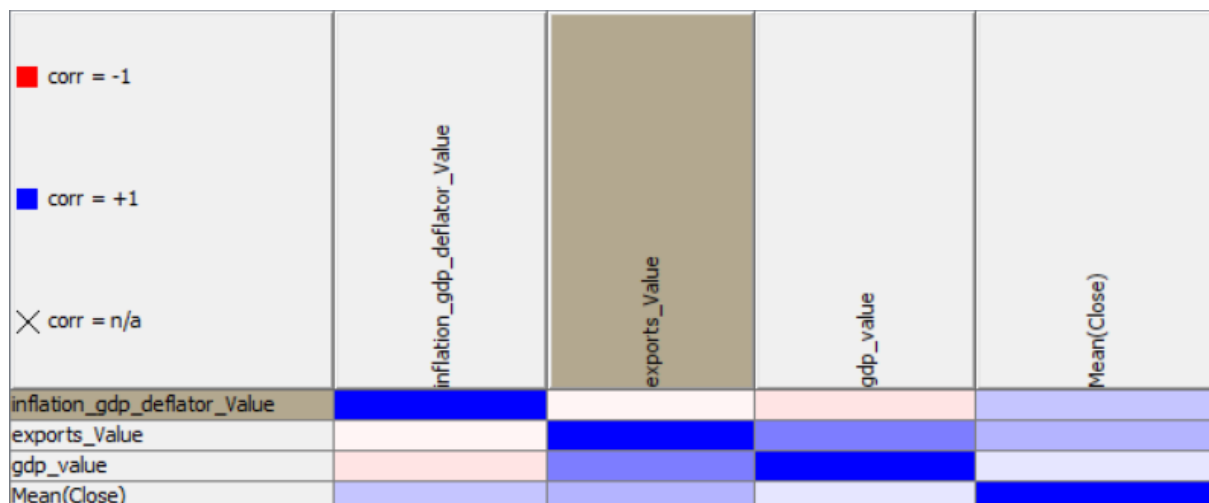
Upon closer examination, we note that the correlations are not statistically significant, as the p-value exceeds 0.1 for all detected correlations.

Ian
David Szabados
Jana Hochel
Artyom

Homework 2
Data Engineering 2
CEU
12/02/2023

Row ID	S First column name	S Second column name	D Correlation value	D p value	I Degree...
Row0	inflation_gdp_deflator_Value	exports_Value	-0.038885526385014...	0.579879308579...	203
Row1	inflation_gdp_deflator_Value	gdp_value	-0.10454177168175259	0.125589357673...	214
Row2	inflation_gdp_deflator_Value	Mean(Close)	0.22847617337136744	7.162606351724...	214
Row3	exports_Value	gdp_value	0.5106709764405124	5.107025913275...	203
Row4	exports_Value	Mean(Close)	0.29236760918458987	1.315260213141...	213
Row5	gdp_value	Mean(Close)	0.09918143909587804	0.146282110624...	214

G20 Countries

Row ID	S First column name	S Second column name	D Correlation value	D p value	I Degree...
Row0	inflation_gdp_deflator_Value	exports_Value	-0.14758772156592...	0.015602723470084...	266
Row1	inflation_gdp_deflator_Value	gdp_value	-0.250313762471257	3.176695973012212E-5	268
Row2	inflation_gdp_deflator_Value	Mean(Close)	0.021703721757721...	0.7225777707262717	268
Row3	exports_Value	gdp_value	0.2917036736131235	1.178824759806573...	266
Row4	exports_Value	Mean(Close)	0.07082404208514552	0.2478973509269058	266
Row5	gdp_value	Mean(Close)	-0.18958295440837...	0.001753193740707...	268

Turning our attention to the correlations for G20 countries, we uncover different outcomes. A negative moderate correlation of -0.14 exists between exports and inflation, and it proves to be statistically significant at the 5% level. This suggests that as one indicator increases, the other decreases. The other correlation that was detected was between GDP and average WTI. The correlation coefficient is -0.19 indicating a negative relationship between GDP and average WTI.

The Effect of Volume and Price on Inflation

Statistics on Linear Regression

On the global level, the volume of oil stocks sold seems to decrease inflation whilst the higher price of the stock seems to have a positive effect on inflation. This may be linked to the volume of oil traded on the market. This is consistent with the law of supply and demand although it explains only a small percentage of variation of inflation.

Variable	Coeff.	Std. Err.	t-value	P>|t|
Year=2006	-5.48E13	2.26E13	-2.4248	0.0154
Year=2007	-4.14E12	5.94E12	-0.6972	0.4857
Year=2008	-9.85E12	2.35E13	-0.4184	0.6757
Year=2009	1.31E14	4.48E13	2.9326	0.0034
Year=2010	1.20E14	4.11E13	2.9258	0.0035
Year=2011	5.77E13	2.11E13	2.735	0.0063
Year=2012	6.36E13	2.17E13	2.926	0.0035
Year=2013	7.53E13	2.58E13	2.9213	0.0035
Year=2014	1.17E14	4.01E13	2.9216	0.0035
Year=2015	2.17E14	7.45E13	2.9108	0.0036
Year=2016	2.16E14	7.35E13	2.933	0.0034
Year=2017	2.35E14	8.07E13	2.9146	0.0036
Year=2018	3.00E14	1.10E14	2.7328	0.0063
Year=2019	2.60E14	9.16E13	2.8332	0.0046
Year=2020	4.22E14	1.59E14	2.6527	0.008
Year=2021	2.99E14	1.07E14	2.796	0.0052
Year=2022	3.58E14	1.35E14	2.66	0.0078
Mean(Close)	6.03E12	2.29E12	2.636	0.0084
Mean(Volume)	-61,733,296.2492	28,694,209.7269	-2.1514	0.0315
Intercept	-1.34E14	5.48E13	-2.4422	0.0146

R-Squared: 0.0332