# Machine Learning Analysis of Resource Exports and HDI Trends

Richard Cai and Zhao (Kevin) Wenkai

Cornell, rjc432@cornell.edu and wz459@cornell.edu

# Abstract

In the greater realm of macroeconomics, the precise factors dictating growth is a continuing field of research and exploration filled with uncertainty and unquantifiable variables even hundreds of years after the publication of *The Wealth of Nations (1776)*. However, it is widely acknowledged that it is far easier to quantify detracting variables that take away from the productive energies of a country, than to quantify variables that contribute. Corruption actively syphons resources away from productive measures and leads to negative impacts ranging from the development of crony capitalism (*rent seeking*) to the rise of state capture. One relevant form of corruption to this study is that incurred by the *resource* curse in *rentier states*—where the rise of corruption is largely ignored due to a large dearth of natural resource production. Growth driven by primary sector natural resource production is quicker to achieve than long-run growth generated by secondary sector manufacturing or tertiary sector service industries, which typically require greater investment, infrastructure, and stability and are more sensitive to global economic trends (though events such as the Oil Crisis of 1973 illustrate that resource-based growth can also be vulnerable). We explore potential hypothesis, including the *resource curse* but also competing *resource-driven-growth* hypotheses, for correlations between HDI and resource exports across countries and continents

“While high HDI might not necessarily translate to high resource exports, high exports should translate into high HDI, ceteris paribus. Why such wealth does not is the question.”

# Introduction

Corruption has been endemic to humanity since the dawn of time. Even Plato himself advised that philosophers "shelter behind a wall" when it came to corruption due to its ironically incorrigible nature. However, it has been proven time and time again by reformers that its entrenched presence can be ridden from societies. What concerns this study is whether exports of certain types of natural resources are or are not correlated with the expected increase in human HDI in a naïve assumption. This naïve assumption is that the large extraneous amounts of sudden wealth accumulated through natural resource exports often raise a country’s growth above a natural growth rate and should lead to large improvements in quality of life (income), lifespan (healthcare), and education (average years in school), the three primary benchmarks used in calculating HDI, corresponding to increasing personal and government wealth. This assumption is characterized as naïve because many real-world instances of extraneous wealth gained from resources exports often fall into the wrong hands or are used for efforts other than improving the country to which it belongs. Of course, it is possible to increase HDI without large resource exports— we just suppose that the large revenue produced *if* such exports were to exist should increase a country’ HDI. This study aims to firstly explore whether the *resource curse* hypothesis holds; that is, if more exports immediately correlates to increased corruption and doesn’t stimulated expected growth; and secondly explores whether resource exports are absolutely correlated to corruption and whether a metric utilizing HDI / resource export ratios can be used as a metric to measure corruption. We specifically focus on the export of cereal, oil, lumber, rare earth metals, and ores from countries worldwide during the period 2010-2020.

# Methods

Analysis was performed on the same, uniform dataset of exports-per-capita / HDI data across 2010-2020 for all countries. Manual adjustments of hyperparameters was used for all methods, as well as filtering.

## Data

All data was sourced from the [UN Comtrade database](https://comtradeplus.un.org/) for broad export types: all mineral spirits (crude & refined oil), all nonorganic mined resources (rare earth metals), ores (particularly steel of note), wood (forestry), and wheat. Population data used for per-capita comparison was sourced from the [World Bank database](https://databank.worldbank.org/source/population-estimates-and-projections) and HDI data was sourced from the [UNDP](https://hdr.undp.org/data-center/human-development-index#/indicies/HDI), which produces the index.

## Method A: Regression Trends

Regression trends were derived from curves with manually derived hyperparameters that excluded specific clustering to improve performance and identify differing trends.

### Oil

The sheer magnitude of oil exports make it notable— in the range of $10000, when cereals are in the range of $500.

Oil perhaps was the most interesting out of all of our findings with this method, with two instead of one distinct curves upward and some of the highest r2 values in nonlinear regression fitting. Instead of the expected singular pattern expected, two distinct lines emerged with different intercepts. Historically, it has been shown countries could have high HDI without exports, but a low reasonable low HDI would be correlated with small material wealth (exports) and a country with high exports of oil would be expected to have high HDI from development. This does not appear to be the case, with Gulf states forming the separate line with a lower intercept and other oil exporters (i.e. Norway) forming the upper line. This seems to suggest two different theories of resource-driven exports: that where it may have been used to fuel economical development in otherwise unfavorable environments (i.e. Norway) or where it was used in less productive measures (i.e. Gulf States), such as regime-building, defense, or embezzled.

### Ores

Exports of ores (such as raw iron) saw the most stereotypical fit with a pattern, suggesting a lack of the *resource-curse* phenomenon— instead a *resource-driven* hypothesis seems more likely for these industrializing countries.

### Cereals

Cereal exports fit into a linear trend with extremely low r2 values even when compared to other exports, suggesting that exports of such foodstuffs had nothing to do with development.

### Rare Earth Metals

Rare earth metals were the exception to any trend, with a resounding *negative* correlation between HDI and export levels, suggesting that exporting rare earth metals was something more developed countries quickly abandoned (environmental damage, for example), or something forced upon lesser developed countries by necessity (*pollution haven hypothesis (PHH)*).

## Method B: Clustering

Clustering for *n* variable clusters for each resource was performed on plots of all data from 2010-2020 simultaneously after regression in order to identify specific parts of trends— the countries and years involved, namely. The amount of clusters for each individual resource was manually adjusted as a hyperparameter to achieve relevant clusterings supporting our hypothesis. Of particular note are clusterings in the lower half or left of graphs indicating low HDI with medium to high resource exports, which would be one possible indicative sign of a resource curse. Also interesting are patterns shown by distributions of extremely high resource exports (>$5000 per capita).

### Oil

It would be expected that with such massively profitable export (with outliers going to $50000 per capita) should immediately improve the quality of life, but it appears not so, especially with Cluster 0 and Cluster 1 (Fig. 1a). After further analysis, it was further revealed that Cluster 0 represented a cluster of oil-rich Gulf countries and Cluster 1 represented a series of highly corrupt or unstable countries (Fig 2). A particularly emblematic country representative of Cluster 1 was Angola, which can be singled out for its rife corruption. Despite the state-owned oil Sonangol making massive revenues in excess of $10 billion on the monopoly it enjoys on petroleum and natural gas exports, equating to over $1000 per capita, much of the country outside of Luanda is still in poverty and practices subsistence farming.

### Wood

Clustering of wood export data indicate a relatively normal trend, except for outliers in Cluster 1 and 3 (Fig. 1b). Cluster 3 shows that extremely high lumber exports tend to correlate to a *lower* HDI index, perhaps indicative of how anti-deforestation environmental awareness takes place at higher HDIs. Cluster 1 also displays a few higher export countries stuck at a relatively low HDI, but it is crucial to remember the scale of the exports. Wood exports per capita never exceeded $1400 per capita in any case, which would render its significant in development rather limited given the small amount of derived revenue.

### Rare Earth Metals

Apart from its notable *negative* trend overall, specific clusters are notable for skewing the trend further towards negative or outliers. Cluster 1 shows a series of rather undeveloped countries (HDI ~0.7) exporting large amounts of ore, suggesting specific countries impacted by the *pollution haven hypothesis (PHH)*. Similarly, extremely high exports of rare earth metals in Cluster 3 similar to the Gulf States of Oil Cluster 0 have extremely middling HDIs (~0.8). However, unlike even Oil, larger exports of rare earth metals resolutely lead to lower HDI. Possible hypotheses include the PHH, but also that the fact that rare earth metal mining is an extremely polluting and dangerous activity often cost-prohibitive in more developed countries.

Figure 1a, b, c, d, e: clusters and various fitted nonlinear/linear regressions overlaid upon resource export per capita in dollars against HDI.

|  |  |
| --- | --- |
| A graph of different colored dots  AI-generated content may be incorrect. | |
| A graph with colored dots and lines  AI-generated content may be incorrect. | A graph of a graph with colored lines  AI-generated content may be incorrect. |
| A graph of different colored lines  AI-generated content may be incorrect. | A graph of different colored dots  AI-generated content may be incorrect. |

Figure 2: table of selected cluster data from oil exports.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cluster 0 | | | | | Cluster 1 | | | | |
| Year | Country | HDI | Exports per C. ($) | Year | | Country | HDI | Exports per C. ($) |
| 2013 | Brunei | 0.835 | 26870.93 | 2013 | | Congo | 0.6 | 1629.37 |
| 2013 | Bahrain | 0.839 | 8121.1 | 2013 | | Angola | 0.555 | 2541.71 |
| 2013 | Kuwait | 0.813 | 30788.14 | 2013 | | Gabon | 0.678 | 4483.35 |
| 2013 | Oman | 0.814 | 10820.48 | 2013 | | Iraq | 0.648 | 2538.39 |
| 2013 | Saudi Arabia | 0.831 | 11656.53 | 2014 | | Angola | 0.565 | 2110.99 |
| 2013 | UAE | 0.847 | 17028.4 | 2014 | | Gabon | 0.687 | 4033.12 |
| 2014 | Bahrain | 0.839 | 11394.55 | 2014 | | Iraq | 0.651 | 2309.43 |
| 2014 | Brunei | 0.834 | 23335 | 2015 | | Angola | 0.591 | 1132.72 |
| 2014 | Kuwait | 0.816 | 25785.54 | 2015 | | Gabon | 0.692 | 2189.95 |
| 2014 | Oman | 0.818 | 9632.86 | 2015 | | Iraq | 0.656 | 1312.48 |
| 2014 | Saudi Arabia | 0.836 | 10067.37 | 2016 | | Gabon | 0.696 | 1397.37 |
| 2014 | UAE | 0.853 | 12780.04 | 2016 | | Iraq | 0.661 | 1137.73 |
| 2015 | Brunei | 0.832 | 13989.83 | 2017 | | Angola | 0.597 | 1093.17 |
| 2015 | Kuwait | 0.829 | 12643.81 | 2017 | | Gabon | 0.699 | 1666.76 |
| 2015 | Qatar | 0.852 | 27507.31 | 2018 | | Angola | 0.598 | 1243.08 |
| 2016 | Brunei | 0.83 | 10020.23 | 2018 | | Congo | 0.603 | 1672.66 |
| 2016 | Kuwait | 0.832 | 10352.23 | 2018 | | Gabon | 0.699 | 2074.18 |
| 2016 | Qatar | 0.853 | 18868.44 | 2019 | | Angola | 0.597 | 1031.26 |
| 2017 | Brunei | 0.829 | 11528.97 | 2019 | | Gabon | 0.702 | 1866.24 |
| 2017 | Kuwait | 0.835 | 11932.87 | 2020 | | Gabon | 0.704 | 1673.52 |
| 2017 | Qatar | 0.862 | 22320.92 |  | |  |  |  |
| 2018 | Brunei | 0.826 | 13663.28 |  | |  |  |  |
| 2018 | Kuwait | 0.836 | 15124.58 |  | |  |  |  |
| 2018 | Qatar | 0.866 | 28336.14 |  | |  |  |  |
| 2019 | Brunei | 0.827 | 14484.66 |  | |  |  |  |
| 2019 | Kuwait | 0.838 | 13211.52 |  | |  |  |  |
| 2019 | Qatar | 0.869 | 23724.86 |  | |  |  |  |
| 2020 | Brunei | 0.827 | 12039.49 |  | |  |  |  |
| 2020 | Kuwait | 0.826 | 8146.67 |  | |  |  |  |
| 2020 | Qatar | 0.863 | 15080.78 |  | |  |  |  |

## Method C: Ratios

Among many other tools included in this project, ratios and various indices of HDI-to-Exports were developed to create a final metric by which to judge presence of *resource curse*. However, it was found that direct ratios proved to be a poor measure as countries that achieved high development independent of resource exports were flagged as *resource cursed* due to their lower HDI-to-Export ratio. One of the measures developed to mitigate this shortcoming was an upper- and lower-limit to the function, such that countries that didn’t export “enough” relative to their HDI to be considered relevant were excluded. However, this proved to be far too arbitrary even for an exploratory study, as adjustments tended to confine results to the singular expected line and removed actual outliers. Additionally, even with filters, countries that simply exported a lot with high development were repeatedly flagged (i.e. Norway). In short, we found using a HDI-to-Export ratio or adjusted index to be unhelpful for general analysis. Within clusters known to be representative of certain conditions however, this ratio index (when properly filtered) was useful for quantifying severity of the suspected *resource curse* phenomenon (Fig. 3).

Figure 3: exports and ratio data for selected 2013 data of Custer 0 of Oil Exports in Dollars per Capita

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Year | Country | HDI | Exports per C. ($) | HDI-Export Ratio |
| 2013 | Brunei | 0.835 | 26870.93 | 32180.75 |
| 2013 | Bahrain | 0.839 | 8121.1 | 9679.499 |
| 2013 | Kuwait | 0.813 | 30788.14 | 37869.79 |
| 2013 | Oman | 0.814 | 10820.48 | 13292.97 |
| 2013 | Saudi Arabia | 0.831 | 11656.53 | 14027.11 |
| 2013 | UAE | 0.847 | 17028.4 | 20104.37 |

## Conclusions

We find the certain resource exports of are correlated with HDI , while others are not. When resource exports are correlated with HDI, deviations or secondary trends are important for determining any hypothesis of *resource curse* or differentiating *resource-driven growth*. Clustering can be further used to pick out groups of outliers or notable countries over time, which can translate into further case analysis over a generalized group.

Oil had the most major evidence of the *resource curse*, with an entirely separate curve for high exporters and several lower-HDI outliers. Closer inspection yields instances of corruption or malfeasance in most countries outlined in those patterns. Rare earth metals followed a new hypothesis, as high exports of such did not appear to drive growth but rather decreased paradoxically as exports increased. Ores, wheat, and wood together mostly followed the same pattern of increasing HDI if exports increased, as would be expected of such more generalized and less valuable exports.

Further efforts on this exploratory topic should mostly focus less on additional clustering or more accurate regression, but rather on:

1. The development of a more accurate metric or ratio to quantify the prevalence of possible “*resource curse*.” The current cropped HDI-exports/capita ratio is not very useful nor specific and requires extensive manual adjustment to the point where personal bias may entirely dictate results.
2. Improvement of heatmap processes. Although not mentioned in-article, a HDI/exports heatmap and cropping tool was included amongst dozens of other of undocumented tools. Heatmaps are the next step in clustering visualization.
3. Automatic fitting. All hyperparameters were manually adjusted according to calculations and intuition, leading to a very human-fraught analysis best characterized as exploratory. Automatic cropping detection such as percentage cropping has already been implemented, as well as various scaling algorithms such as min-max. Much works need to be done *objectively* to properly achieve a meaningful high r2 value.

Overall, the numerous tools (especially clustering and regression) developed in this project proved to be extremely enlightening in this exploratory study. Other tools included such as numerical correlation functions, heatmaps, histograms, and 2D/3D plots (not mentioned) are incredibly valuable for visualization to understand why some countries and regions remain afflicted by poverty despite such abundant wealth.