

Análise da influência dos países em relação ao Aquecimento Global

Equipe

Nome: Os Remanescentes

Repositório GitHub da equipe:

<https://github.com/thiagorbernardo/UTFPR-Data-Science-CSB55>

Membros:

- | | | |
|--|----------|------------------|
| • Pedro Sodré dos Santos, | 2126745, | pedro_sodre, |
| psantos.2000@alunos.utfpr.edu.br , | EC, | UTFPR |
| • Thiago Ramos Bernardo, | 2126753, | thiagorbernardo, |
| thiagobernardo@alunos.utfpr.edu.br , | EC, | UTFPR |

Introdução

Sabe-se que estamos com diversos problemas ambientais no planeta, como o Efeito Estufa, medidas ambientais são necessárias para intervir nos problemas que já estamos sentindo, como o **Aquecimento Global**. Também, nessa linha existem diversos locais no planeta com uma qualidade do ar horrível, enquanto outros com uma qualidade perfeita. Para isso foi realizada essa pesquisa, que consiste em entender como os países se comportam em relação aos índices econômicos, como IDH e PIB, para entender se é possível um país continuar crescendo economicamente e emitir poucos gases. Também nessa linha foi procurado entender como é a relação entre os gases de efeito estufa e a qualidade do ar, um país com muitas emissões tem uma qualidade do ar ruim?

Em resumo, procuramos responder 3 perguntas de pesquisa:

- Um país com muita emissão de gases estufa, possui uma qualidade de ar significativamente pior em relação a um país com pouca emissão?
- É possível um país dissociar aumento econômico com a quantidade de emissões estufa?

Para a primeira pergunta, a conclusão foi que não há relação entre emissões de gases de efeito estufa e o índice de qualidade do ar nos países, isso principalmente devido à problemas com os dados e a distribuição deles discutidos nos resultados e limitações. Já para a segunda questão, foi concluído que é sim possível, principalmente se seu investimento na economia for no terceiro setor.

Processamento de Dados

Primeiramente, o foco seria a análise da relação entre as emissões de gases de efeito estufa e a qualidade do ar de cada país/continente. Primeiramente foi preciso utilizar os dados do IQAr do site <https://aqicn.org/data-platform/covid19/>, visto na Figura 1, são dados de cada estação de captura de dados do ar e dados meteorológicos, entre eles os principais analisados (monóxido de carbono, partículas com diâmetro inferior a 10 e 2,5 micrômetros, ozônio, dióxido de enxofre e dióxido de nitrogênio).

Date	Country	City	Specie	count	min	max	median	variance
2015-01-06	KR	Jeonju	co	124	0.1	12.3	4.5	55.74
2015-01-22	KR	Jeonju	co	116	4.5	10.0	6.7	16.09
2015-03-30	KR	Jeonju	co	118	1.2	11.2	5.6	35.98
2015-05-27	KR	Jeonju	co	93	2.3	5.6	3.4	6.54
2015-02-03	KR	Jeonju	co	133	4.5	13.4	7.8	39.24
2015-02-11	KR	Jeonju	co	111	3.4	11.2	6.7	28.73
2015-02-18	KR	Jeonju	co	108	3.4	8.9	6.7	22.24
2015-03-05	KR	Jeonju	co	129	2.3	12.3	5.6	24.22
2015-03-19	KR	Jeonju	co	100	2.3	7.8	4.5	15.28
2015-04-04	KR	Jeonju	co	87	1.2	10.0	4.5	31.48
2015-05-28	KR	Jeonju	co	109	2.3	5.6	3.4	6.38
2015-05-05	KR	Jeonju	co	79	1.2	6.7	3.4	8.20
2015-05-09	KR	Jeonju	co	72	1.2	4.5	3.4	7.26
2015-05-14	KR	Jeonju	co	63	2.3	5.6	4.5	11.23
2015-06-03	KR	Jeonju	co	85	1.2	5.6	3.4	7.64
2015-06-05	KR	Jeonju	co	90	1.2	8.9	3.4	28.40
2015-06-06	KR	Jeonju	co	61	2.3	6.7	3.4	8.42

Figura 1. Dataset IQAr.

Após filtrar apenas os poluentes interessantes para a análise, foi agrupado todas as estação por país e ano, foram criados 6 datasets, um para cada poluente, esse pré-processamento foi repetido para os continentes. Essa análise inicial deu uma ideia de como os continentes e países se comportam em cada poluente, visto na Figura 2.

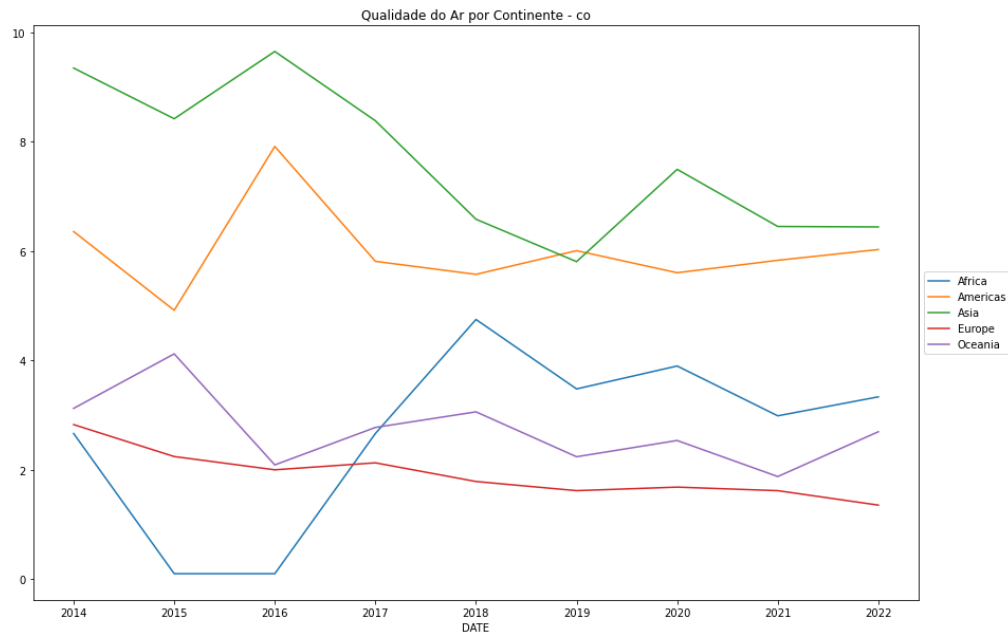


Figura 2. Qualidade do Ar por Continente - Monóxido de Carbono.

Após ter os dados separados por poluente, país e continente, foi necessário os classificar e normalizar, para isso foi utilizado uma tabela customizada para dar um nível de significância indo do BOM ao PÉSSIMO, como visto na Figura 3.

	INDEX	QUALITY	pm25	pm10	so2	no2	co	o3
0	0	BOA	25	50	20	200	9.00	100
1	1	REGULAR	60	120	125	260	9.05	140
2	2	INADEQUADA	124	249	799	1129	14.90	199
3	3	RUIM	209	419	1599	2259	29.90	399
4	4	PESSIMO	249	499	2099	2999	39.90	599

Figura 3. Tabela de Classificação.

Com os dados classificados e normalizados foi necessário agregar com os dados de emissões de cada continente, como visto na Figura 4. Para o dataset de países foi preciso um trabalho maior para remover países com poucas cidades com estação (foi mantido países com mais de 10 cidades), discutido melhor nas limitações do trabalho, resultando no dataset da Figura 5.

	DATE	REGION	MEAN_BY_CONTINENT	CLASSIFICATION	EMISSIONS
0	2014	Africa	8.243333	0.00	0.109344
1	2014	Americas	18.772202	0.20	0.394471
2	2014	Asia	37.822036	1.00	0.918972
3	2014	Europe	21.363126	0.40	0.190249
4	2014	Oceania	18.392635	0.20	0.000000
5	2015	Africa	8.298631	0.00	0.110449

Figura 4. Dataset com Índice de Qualidade do Ar e Emissões por Continente.

	DATE	CODE	CODE-2	MEAN	CLASSIFICATION	COUNTRY	EMISSIONS
8	2014	CHN	CN	52.435908	1.000000	China	0.929342
16	2014	FRA	FR	20.851889	0.166667	France	0.001246
23	2014	JPN	JP	23.278667	0.166667	Japan	0.077248
39	2014	USA	US	11.911514	0.000000	United States	0.463753
53	2015	CHN	CN	41.766540	0.666667	China	0.919121
62	2015	FRA	FR	17.610029	0.000000	France	0.001711

Figura 5. Dataset com Índice de Qualidade do Ar e Emissões por país.

Para analisar a outra hipótese, foram utilizados os mesmo dados de emissões da análise anterior, onde possui a emissão de cada país em cada ano (desde 1990), e também dados de PIB de cada país, ambos retirados do Our World in Data (<https://ourworldindata.org/grapher/gross-domestic-product>, <https://ourworldindata.org/co2-and-other-greenhouse-gas-emissions>). Também foi utilizado o dataset de países e regiões, para verificar o continente de cada país.

Primeiramente foi feito um filtro no PIB dos países, que haviam dados desde 1960, porém não tínhamos dados de emissões nessa época, então o dataset foi filtrado para anos superiores a 1990. Após isso, todos os datasets foram juntados com base no ano e no país, assim o dataset ficaria com todos os dados desejados.

O modelo, posteriormente, também foi simplificado para conseguir visualizar os resíduos de uma forma melhor, fazendo a média de todos os anos de PIB e de emissões.

Resultados

Para responder a primeira pergunta com base nos continentes foi feito uma regressão linear a partir da variável dependente (*classification*, classificação com base na tabela do índice de qualidade do ar) e da variável independente (*emissions*, emissões de gases de efeito estufa), com isso foi obtido o seguinte resultado demonstrado na Figura 6.

OLS Regression Results			
Dep. Variable:	CLASSIFICATION	R-squared:	0.437
Model:	OLS	Adj. R-squared:	0.417
Method:	Least Squares	F-statistic:	21.72
Date:	Fri, 02 Dec 2022	Prob (F-statistic):	7.01e-05
Time:	23:20:09	Log-Likelihood:	12.431
No. Observations:	30	AIC:	-20.86
Df Residuals:	28	BIC:	-18.06
Df Model:	1		
Covariance Type:	nonrobust		

Figura 6. Métricas da regressão linear dos continentes de índice x emissões.

Com esse resultado podemos ver uma forte relação entre as emissões e o índice de qualidade do ar, porém por se tratar de uma análise de continentes ficaria difícil realmente analisar os atores, e propor ideias e soluções. A partir disso foi feita uma regressão linear com o dataset dos países com o mesmo modelo, e com isso foi obtido o resultado demonstrado na Figura 7.

OLS Regression Results			
Dep. Variable:	CLASSIFICATION	R-squared:	0.112
Model:	OLS	Adj. R-squared:	0.104
Method:	Least Squares	F-statistic:	14.56
Date:	Fri, 16 Dec 2022	Prob (F-statistic):	0.000219
Time:	15:37:38	Log-Likelihood:	20.529
No. Observations:	118	AIC:	-37.06
Df Residuals:	116	BIC:	-31.52
Df Model:	1		
Covariance Type:	nonrobust		

Figura 7. Métricas da regressão linear dos países de índice x emissões.

Considerando os dados disponíveis e as limitações discutidas no próximo tópico, os dados apontam para uma relação fraca entre as emissões e o índice de qualidade do ar, logo, a resposta para a pergunta seria: “Não, um país com muitas emissões de gases de efeito estufa não necessariamente tem uma qualidade do ar pior em relação a um país com poucas emissões”. Porém, outros dados e informações indicam alguns problemas nessa análise, como o comentário do Dr. Jan Mateu Armengol que o índice de qualidade do ar é muito afetado por poluições de países vizinhos, e que principalmente na Europa isso é um problema, uma vez que os países e consequentemente as estações de captura estão muito próximas.

Para responder a segunda pergunta, foi feito uma regressão linear com análise de resíduos, para verificar, baseado na linha de base da regressão, quais países estão abaixo da linha (possuem desenvolvimento de PIB porém sem grandes emissões) e os países que estão acima da linha (possuem emissões acima do desenvolvimento do PIB).

Realizando a regressão linear de emissões como variável dependente e PIB como variável independente, verificamos, como esperado, que apresenta uma regressão linear crescente (com coeficiente positivo), ou seja, quanto maior o crescimento do PIB, maior a emissão de gases estufa, e um R^2 de 0.627, explicando mais de 60% da variância. Porém, o que queríamos verificar é se realmente um país consegue dissociar desenvolvimento econômico com baixa emissão de gases estufa. Então foi feita uma regressão, agora colocando o continente como variável independente para verificar o coeficiente de cada continente (quanto maior o coeficiente, maior a emissão de poluentes por PIB). Dessa forma conseguimos o modelo abaixo:

OLS Regression Results						
Dep. Variable:	Total_Emissions		R-squared:		0.644	
Model:	OLS		Adj. R-squared:		0.643	
Method:	Least Squares		F-statistic:		1650.	
Date:	Tue, 13 Dec 2022		Prob (F-statistic):		0.00	
Time:	15:39:15		Log-Likelihood:		-97971.	
No. Observations:	4571		AIC:		1.960e+05	
Df Residuals:	4565		BIC:		1.960e+05	
Df Model:	5					
Covariance Type:	nonrobust					
	coef	std err	t	P> t	[0.025	0.975]
Intercept	4.612e+07	1.37e+07	3.366	0.001	1.93e+07	7.3e+07
region[T.Americas]	1.433e+07	2.18e+07	0.657	0.511	-2.84e+07	5.71e+07
region[T.Asia]	1.923e+08	2.05e+07	9.362	0.000	1.52e+08	2.33e+08
region[T.Europe]	-1.166e+08	2.08e+07	-5.603	0.000	-1.57e+08	-7.58e+07
region[T.Oceania]	-2.565e+07	2.97e+07	-0.863	0.388	-8.39e+07	3.26e+07

Figura 8. Métricas da regressão linear das emissões ~ PIB + continente.

Com esse modelo, é possível destacar que o continente asiático de modo geral possui um coeficiente positivo, não conseguindo dissociar seu crescimento econômico com emissões estufa. Porém, quando verificamos os continentes da Europa e da Oceania, observa-se que o coeficiente apresentado é negativo, ou seja, quanto maior o desenvolvimento de um país, menos ele polui (de acordo com a regressão gerada), nesse caso, o continente da Europa e da Oceania conseguiram dissociar crescimento econômico com suas emissões, possuindo um desenvolvimento sustentável.

Quando realizamos essa análise por país, e verificando os resíduos no nosso modelo, verificamos que a China possui um resíduo muito grande, ou seja, é o país que polui mais se comparado ao seu desenvolvimento (mesmo com um crescimento de PIB grande comparado ao resto do mundo, também houve aumento de poluição muito maior em consequência). Por outro lado, existem países que ficaram com resíduos negativos, ou seja, abaixo dos valores esperados, como Japão, Alemanha e Estados Unidos. Em destaque a esse último que, apesar de ter suas emissões aumentadas nos últimos anos, as suas emissões estão muito abaixo do crescimento do PIB. Como podemos ver nos gráficos de resíduos abaixo:

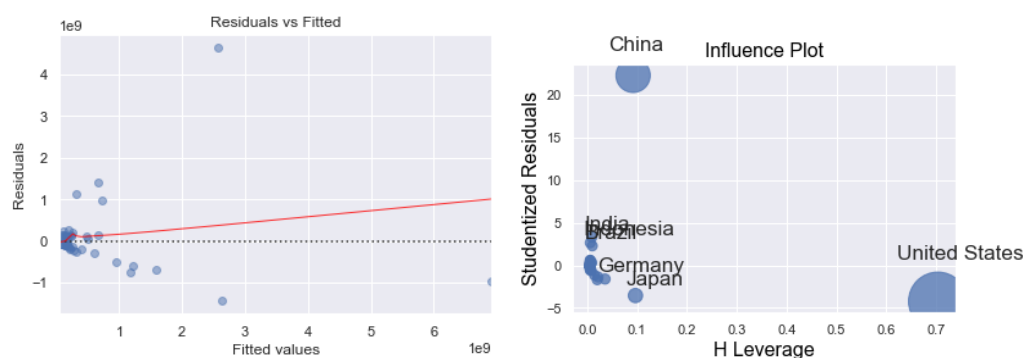


Figura 8. Resíduos do Modelo.

Dessa forma, vemos que alguns países conseguiram crescer e mesmo assim emitiram poucos poluentes, dissociando seu desenvolvimento econômico com suas emissões. Porém é importante ressaltar que todos esses países importam inúmeros produtos da China, que é a grande responsável por fabricar grandes quantidades de produtos para todo o mundo, com seu setor industrial responsável por mais de 60% do PIB. E se verificarmos esses outros países que tiveram emissões abaixo do esperado, tem sua economia baseada no terceiro setor da economia, com comércio e prestação de serviços, onde todos esses possuem esse setor responsável por mais de 60% do PIB. Respondendo a pergunta, se um país consegue dissociar crescimento econômico com o emissões de poluentes, é sim possível, principalmente se o investimento da sua economia for no terceiro setor, porém algum outro país do mundo vai ser responsável por fabricar os seus insumos e vai acabar terceirizando sua emissão para outro país.

Limitações e Trabalhos Futuros

O maior problema para responder a primeira pergunta foi a falta de dados para muitos países, enquanto alguns países tinham dezenas de estações, outros tinham uma ou nenhuma, como visto na Figura 8. O Brasil, por exemplo, país que era pra ter sido considerado na análise, mas foi cortado por ter apenas 3 cidades com estações (região sudeste), Argentina possui apenas uma. Os países que entraram na análise foram Estados Unidos (57 estações), China (52 estações), França (27 estações)...

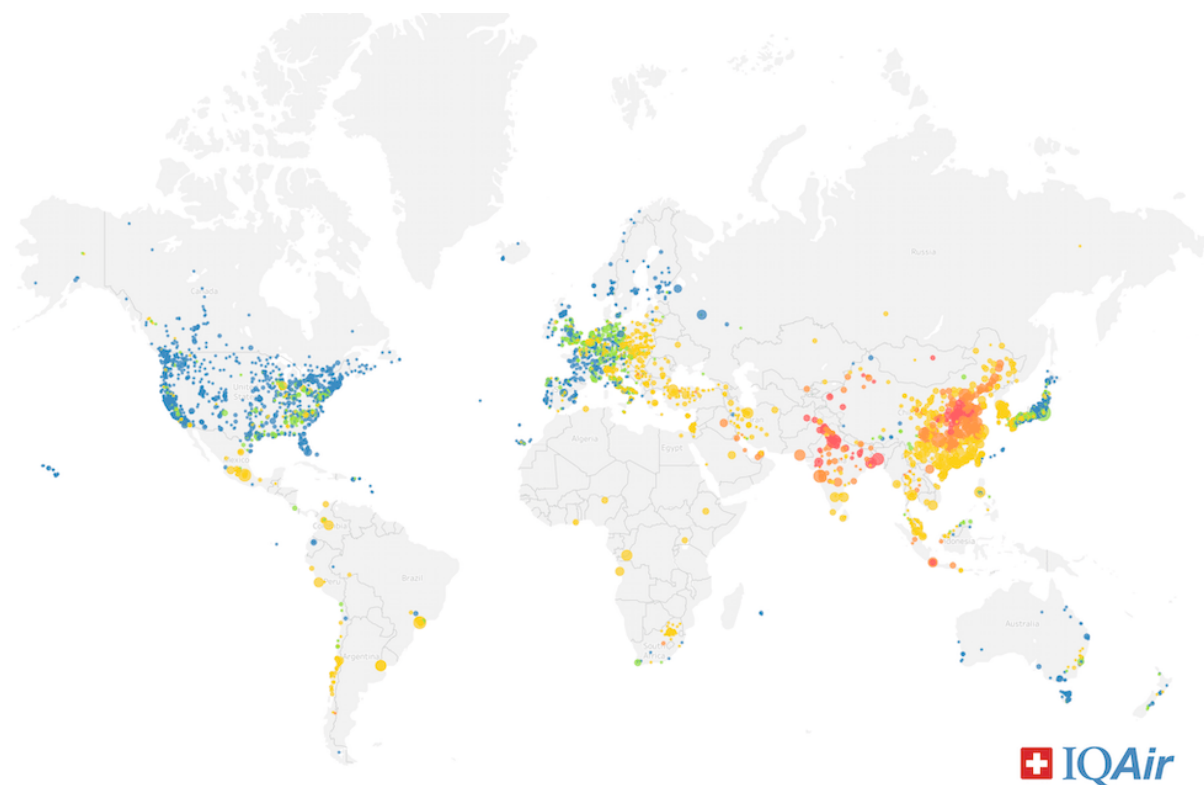


Figura 9. Mapa Mundial em de estações de qualidade do ar.

Como os dados do índice eram por estações, países com muitas estações aproximavam mais de um índice generalizado pro país, isso porque os dados das emissões eram por país. Um país com apenas poucas estações era muito prejudicado, exemplo disso é o Brasil, onde 44% das emissões provém do desmatamento de vegetação nativa e a região Norte concentra 60% das emissões desse setor, como visto na Figura 9 para o monóxido de carbono, entretanto não existe nenhuma estação de qualidade do ar nessa região, dados segundo o Sistema de Estimativas de Emissões e Remoções de Gases do Efeito Estufa (SEEG). O Brasil é o claro contraste dessa análise, onde os lugares que mais emitem não tem estações de qualidade do ar (Figura 8 x Figura 9). É válido acreditar que esse padrão aconteça para vários outros países.



Figura 10. Mapa do Brasil em Emissões Totais de Monóxido de Carbono.

Uma forma de melhorar essa análise seria reduzir para uma menor quantidade de países, assim podendo usar emissões por região / estado, isso melhora a precisão do modelo. Considerando esses pontos discutidos é visível que os dados foram um problema, pois são contraditórios em alguns países.

Já na segunda questão, uma limitação encontrada foi a falta de alguns dados de emissões de alguns países, que não havia em todos os anos da análise (1990 - 2019), podendo levar a conclusões inadequadas.

Referências

<https://www.capitalreset.com/afinal-de-onde-vem-as-emissoes-de-gases-de-efeito-estufa-no-brasil/>
<https://ipam.org.br/regiao-norte-emite-60-dos-gases-de-efeito-estufa-no-brasil/>
<https://china2brazil.com.br/crescimento-da-economia-e-pib-da-china-nos-ultimos-anos/>
<https://noticias.r7.com/internacional/china-e-responsavel-por-31-das-emissoes-de-co-em-2021-29062022>