

Predicting CO₂ Emissions

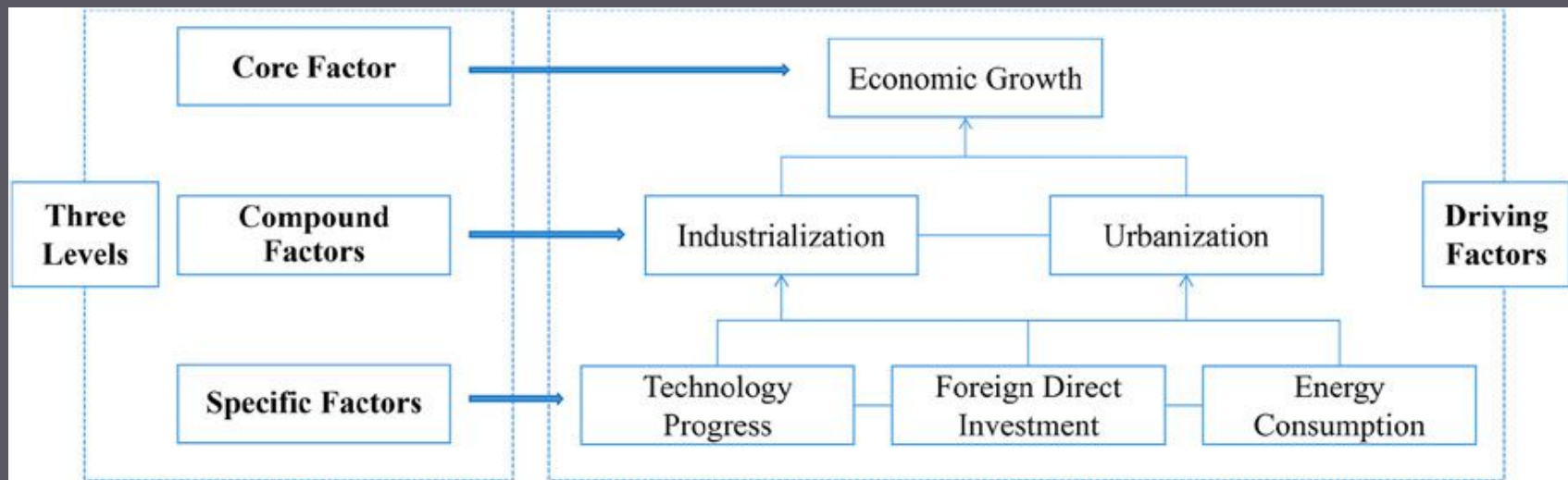
Jack Hlavka and Gabrielle Walrath
STATS/DATASCI 112

Research Question

**How well can we
predict a country's
CO₂ emissions?**

**What about if we don't use
any factors relating to
what type of energy it
uses?**

Data Collection



Our World in Data

Population Division World Urbanization Prospects 2018



THE WORLD BANK
IBRD • IDA

FORTUNE

kaggle

1 Walmart

2 Amazon

3 State Grid

4 China National Petroleum

5 Sinopec Group

6 Saudi Aramco

7 Apple

8 Volkswagen

9 China State Construction Engineering

10 CVS Health

Annual Total Population at Mid-Year (thousands)		Note		1950	1951	1952	1953	1954
ISO 3166-1 numeric	Location							
900	World			2 526 275	2 583 817	2 630 584	2 677 230	2 724 262
901	More developed region			814 865	824 219	834 071	844 264	854 832
902	Less developed region			1 721 410	1 759 604	1 796 513	1 832 967	1 869 431
941	Least developed countries			185 289	189 052	202 905	206 885	211 045
954	Least developed region			1 526 151	1 569 552	1 593 608	1 626 082	1 658 386
948	Least developed region, excluding China			1 187 187	1 176 838	1 203 955	1 229 485	1 256 404
1003	High-income countries			472 286	480 430	485 859	487 445	488 262
1004	Upper-middle-income countries			1 734 481	1 772 394	1 808 889	1 844 730	1 880 558
1005	Lower-middle-income countries			778 277	792 199	807 013	822 677	839 156
1006	Upper-middle-income countries			958 204	980 195	1 001 876	1 022 054	1 041 402
1007	Lower-middle-income countries			129 087	129 689	135 143	136 594	137 754
947	Sub-Saharan Africa			197 621	193 039	188 628	190 378	194 285
903	Africa			228 670	233 277	238 113	243 178	248 471
910	Eastern Africa			86 758	88 206	89 712	91 272	92 896
108	Burundi			2 309	2 359	2 404	2 445	2 487
174	Cameroon			159	163	167	170	173
202	Dominican Republic			62	63	65	66	68
232	Eritrea			1 142	1 160	1 180	1 201	1 224
231	Ethiopia			18 128	18 487	18 820	19 184	19 560
404	Kenya			6 077	6 240	6 412	6 593	6 782
420	Madagascar			4 084	4 188	4 287	4 389	4 494
454	Malawi			2 954	3 008	3 065	3 125	3 187
480	Mauritius			493	506	521	537	554
178	Mayotte			15	16	16	17	18
908	Mozambique			6 152	6 249	6 353	6 463	6 580
628	Niger			248	259	269	277	284
646	Rwanda			2 186	2 251	2 313	2 378	2 449
690	Seychelles			36	37	37	38	38
706	Somalia			2 264	2 308	2 352	2 397	2 444
728	South Sudan			2 853	2 902	2 928	2 953	2 972
800	Uganda			5 158	5 309	5 456	5 601	5 748
834	United Republic of Tanzania			2 7 650	7 847	8 056	8 275	8 503
884	Zambia			2 310	2 367	2 428	2 494	2 562
716	Zimbabwe			2 747	2 830	2 917	3 009	3 104
911	Madagascar			26 484	26 965	27 490	28 052	28 653
34	Angola			4 548	4 629	4 775	4 880	5 005
120	Cameroon			4 307	4 383	4 460	4 539	4 621
140	Central African Republic			1 327	1 340	1 353	1 368	1 383

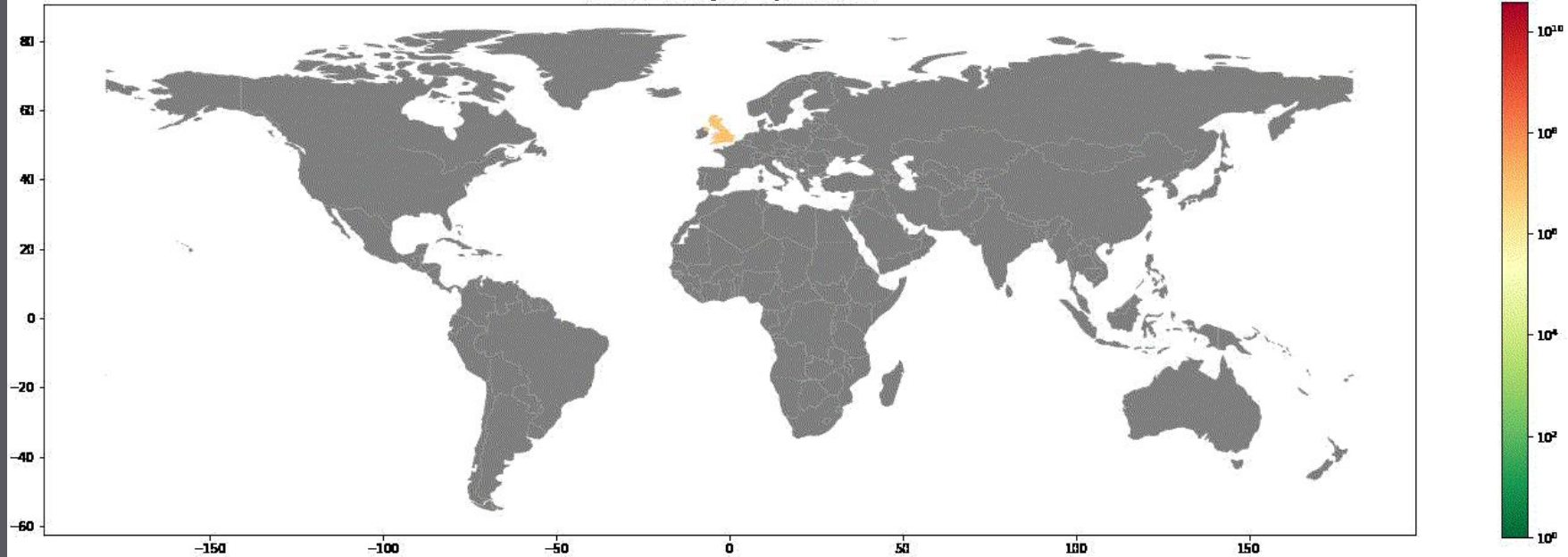
Entity	Code	Year	Annual CO ₂ emissions
Algeria	AFG	1969	14658
Algeria	AFG	1970	84272
Algeria	AFG	1971	91600
Algeria	AFG	1972	91600
Algeria	AFG	1973	100236
Algeria	AFG	1974	100236
Algeria	AFG	1975	103888
Algeria	AFG	1976	103888
Algeria	AFG	1977	292120
Algeria	AFG	1978	329700
Algeria	AFG	1979	384971
Algeria	AFG	1980	413885
Algeria	AFG	1981	480726
Algeria	AFG	1982	688564
Algeria	AFG	1983	700736
Algeria	AFG	1984	838351
Algeria	AFG	1985	1008977
Algeria	AFG	1986	1031188
Algeria	AFG	1987	1281885
Algeria	AFG	1988	1223391
Algeria	AFG	1989	347232
Algeria	AFG	1990	1870307
Algeria	AFG	1991	1880554
Algeria	AFG	1992	1530347
Algeria	AFG	1993	1020556
Algeria	AFG	1994	1913152
Algeria	AFG	1995	2121383
Algeria	AFG	1996	1980859
Algeria	AFG	1997	2381175
Algeria	AFG	1998	2153300
Algeria	AFG	1999	2232754
Algeria	AFG	2000	1790302
Algeria	AFG	2001	1978483
Algeria	AFG	2002	2046818
Algeria	AFG	2003	2518954
Algeria	AFG	2004	2821540
Algeria	AFG	2005	3501423
Algeria	AFG	2006	3313645
Algeria	AFG	2007	3113628
Algeria	AFG	2008	2808096
Algeria	AFG	2009	2746555
Algeria	AFG	2010	2024226

Data Analysis

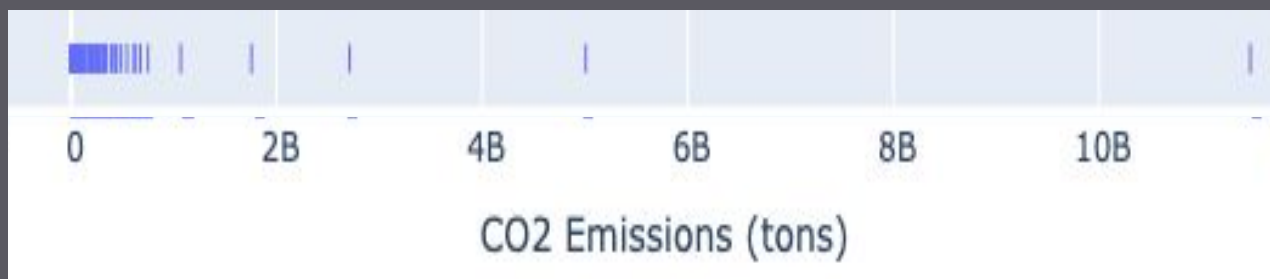
	Country	Code	Year	Population	Urban Percentage	GDP Per Capita (US\$)	Electricity Generation Per Capita (kWh)	FDI Inward (US\$)	Biofuel Share (%)	Coal Share (%)	Fossil Fuel Share (%)	Gas Share (%)	Low Carbon Share (%)	Oil Share (%)
0	Australia	AUS	1990	17041000	85.4	18249	100	8457776859	0	42.032	95.922	15.889	4.078	38.021
1	Australia	AUS	1991	17272000	85.4	18860	100	2612066526	0	43.151	95.669	15.31	4.331	37.207
2	Australia	AUS	1992	17486000	85.3	18624	100	4941906671	0	42.972	95.648	15.68	4.352	36.996
3	Australia	AUS	1993	17687000	85.2	17700	100	5312435141	0	42.05	95.654	15.806	4.346	37.798
4	Australia	AUS	1994	17883000	85	18129	100	4458484243	0	41.263	95.871	16.881	4.129	37.727
5	Australia	AUS	1995	18077000	84.9	20446	100	13268875155	0	41.562	96.069	16.603	3.931	37.904
6	Australia	AUS	1996	18272000	84.8	22020	100	4563952446	0	42.196	96.058	16.313	3.942	37.549
7	Australia	AUS	1997	18468000	84.6	23645	100	8088068982	0	43.239	96.146	15.623	3.854	37.284
8	Australia	AUS	1998	18665000	84.5	21478	100	7597610928	0	44.139	96.246	15.718	3.754	36.389
9	Australia	AUS	1999	18864000	84.4	20698	100	2210917991	0	44.911	96.283	15.425	3.717	35.948
10	Australia	AUS	2000	19066000	84.2	21853	100	14892978180	0	45.039	96.291	15.649	3.709	35.604
11	Australia	AUS	2001	19269000	84.1	19681	100	10717133150	0	44.177	96.406	16.724	3.594	35.505
12	Australia	AUS	2002	19475000	84.2	20291	100	14656321800	0	44.511	96.341	16.959	3.659	34.87
13	Australia	AUS	2003	19697000	84.3	23705	100	8985246029	0	45.106	96.275	16.839	3.725	34.33
14	Australia	AUS	2004	19948000	84.5	30819	100	42907672820	0	45.491	96.284	16.623	3.716	34.169
15	Australia	AUS	2006	20574000	84.7	36570	100	30551100656	0.063	43.976	96.146	17.407	3.854	34.763
16	Australia	AUS	2007	20947000	84.8	41023	100	4444090037	0.239	43.046	96.19	19.226	3.81	33.918
17	Australia	AUS	2008	21342000	84.9	49679	100	4510024270	0.564	44.14	96.148	18.598	3.852	33.41
18	Australia	AUS	2009	21739000	85.1	42810	100	28683266147	0.835	43.041	95.689	19.152	4.311	33.496
19	Australia	AUS	2010	22120000	85.2	52134	100	35210733743	1.255	39.77	94.963	22.123	5.037	33.071
20	Australia	AUS	2011	22480000	85.3	62596	100	65548930649	1.278	37.415	93.707	22.299	6.293	33.992
21	Australia	AUS	2012	22822000	85.4	68044	100	57590426822	1.152	35.471	93.859	22.643	6.141	35.745
22	Australia	AUS	2013	23151000	85.5	68158	100	54465480054	1.084	33.414	93.14	23.606	6.86	36.119
23	Australia	AUS	2014	23475000	85.6	62513	100	632022728079	1.025	32.667	93.756	25.103	6.244	35.866
24	Australia	AUS	2015	23800000	85.7	56710	100	46892808567	0.917	33.326	93.496	25.97	6.504	34.2
25	Australia	AUS	2016	24126000	85.8	49875	100	42969505135	0.776	33.045	92.793	25.535	7.207	34.213
26	Australia	AUS	2017	24451000	85.9	53936	100	48198614234	0.878	31.977	93.115	25.289	6.885	35.849
27	Australia	AUS	2018	24772000	86	57207	100	60685908478	0.952	30.764	91.647	24.873	8.353	36.01
28	Australia	AUS	2019	25089000	86.1	54541	100	38954597515	0.87	27.819	91.433	30.184	8.567	33.43
42	Iceland	ISL	1992	260000	91.1	27124	100	8648079	0	2.904	40.448	0	59.552	37.545
43	Iceland	ISL	1994	265000	91.5	24018	100	2394198	0	2.821	40.389	0	59.611	37.568
44	Iceland	ISL	1996	270000	91.8	27614	100	83381484	0	3.006	41.624	0	58.376	38.619
45	Iceland	ISL	1997	273000	92	27919	100	148012830	0	2.976	39.155	0	60.845	36.179
46	Iceland	ISL	1998	275000	92.1	31030	100	154082591	0	2.47	36.741	0	63.259	34.271
47	Iceland	ISL	1999	278000	92.3	32381	100	66639312	0	2.637	33.872	0	66.128	31.234
48	Iceland	ISL	2000	280000	92.4	32096	100	155160734	0	3.435	33.419	0	66.581	29.984
49	Iceland	ISL	2001	283000	92.5	28897	100	163445772	0	3.688	31.539	0	68.461	27.851
50	Iceland	ISL	2002	285000	92.7	32409	100	91905493	0	3.537	31.43	0	68.57	27.893
51	Iceland	ISL	2003	288000	92.8	39476	100	335355205	0	3.456	31.51	0	68.49	28.054
52	Iceland	ISL	2004	291000	92.9	47334	100	755317000	0	3.548	32.307	0	67.693	28.759
53	Iceland	ISL	2005	295000	93	56794	100	3093680067	0	3.373	32.543	0	67.457	29.17
54	Iceland	ISL	2006	300000	93.2	57492	100	3876554983	0	3.146	30.161	0	69.839	27.015
55	Iceland	ISL	2007	305000	93.3	69495	100	6872169567	0	2.946	26.672	0	73.328	23.727
56	Iceland	ISL	2008	311000	93.4	56943	100	1203686964	0	2.054	19.666	0	80.334	17.612
57	Iceland	ISL	2009	316000	93.5	41301	100	63588140	0	1.903	18.254	0	81.746	16.351
58	Iceland	ISL	2010	320000	93.6	43237	100	256982201	0	2.006	17.666	0	82.334	15.66
59	Iceland	ISL	2011	323000	93.6	47714	100	1106958185	0	2.005	17.452	0	82.545	15.447
60	Iceland	ISL	2012	326000	93.6	45995	100	1024164328	0	2.031	17.162	0	82.835	15.13
61	Iceland	ISL	2013	327000	93.6	49804	100	472607545	0	2.265	17.588	0	82.367	15.324
62	Iceland	ISL	2014	328000	93.7	54576	100	766775639	0	1.858	17.905	0	82.011	16.047
63	Iceland	ISL	2015	330000	93.7	52551	100	1145767444	0	1.969	18.744	0	81.019	16.775
127	Singapore	SGP	1990	3013000	100	11861	100	557473855	0	0.091	99.914	0	0.086	99.822
128	Singapore	SGP	1991	3097000	100	14502	100	4887094440	0	0.053	99.755	0	0.245	99.702
129	Singapore	SGP	1992	3189000	100	16135	100	2204342221	0	0.072	99.772	3.501	0.228	96.199
130	Singapore	SGP	1993	3288000	100	18290	100	4686312017	0	0.073	99.791	4.37	0.209	95.349
131	Singapore	SGP	1994	3385000	100	21553	100	8550165226	0	0.076	99.818	3.821	0.182	95.52
132	Singapore	SGP	1995	3479000	100	24914	100	11942852103	0	0.034	99.825	3.665	0.175	96.125
133	Singapore	SGP	1996	3566000	100	26233	100	11432363956	0	0.001	99.827	3.632	0.173	96.154

CO₂ Emissions

CO2 Emissions by Country 1750 (tons)



count	1.710000e+02
mean	2.102999e+08
std	9.869335e+08
min	0.000000e+00
25%	5.548892e+06
50%	2.032710e+07
75%	8.857515e+07
max	1.147237e+10



Predicting Variables

1

Population

4

**Foreign Direct
Investment**

2

GDP Per Capita

5

**Fortune Global
500 Companies**

3

Electrification

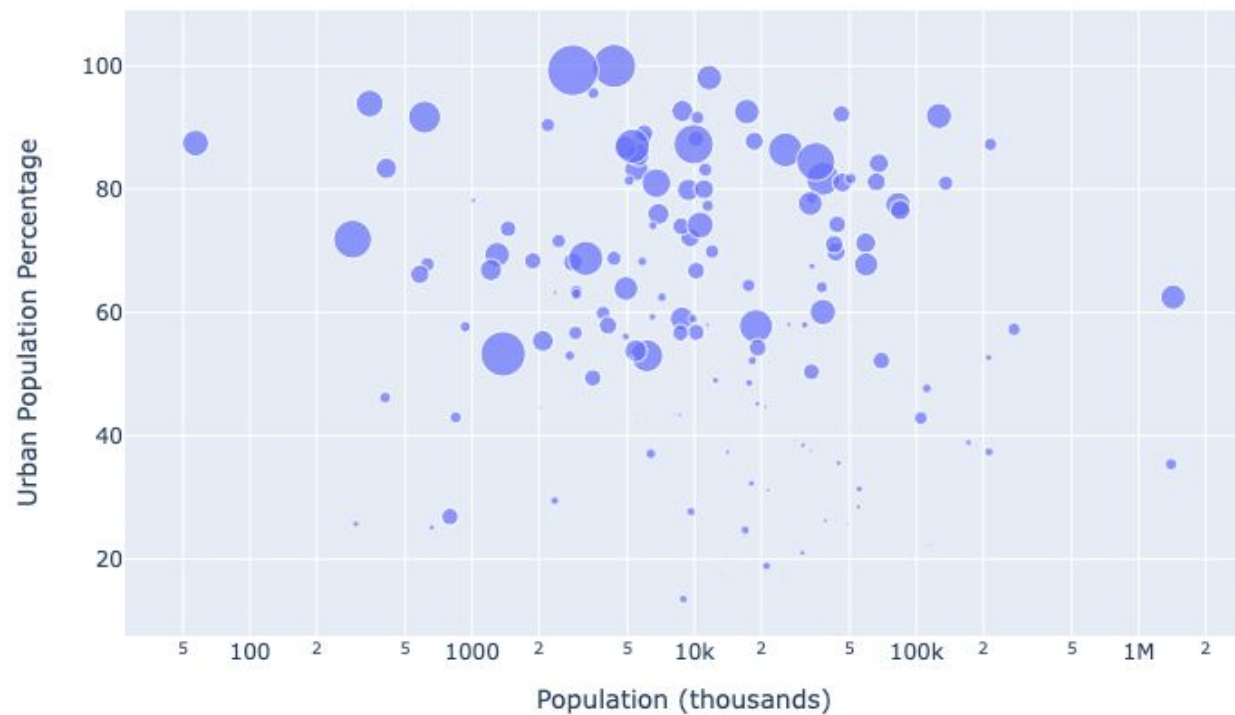
6

Energy Sectors

01

Population

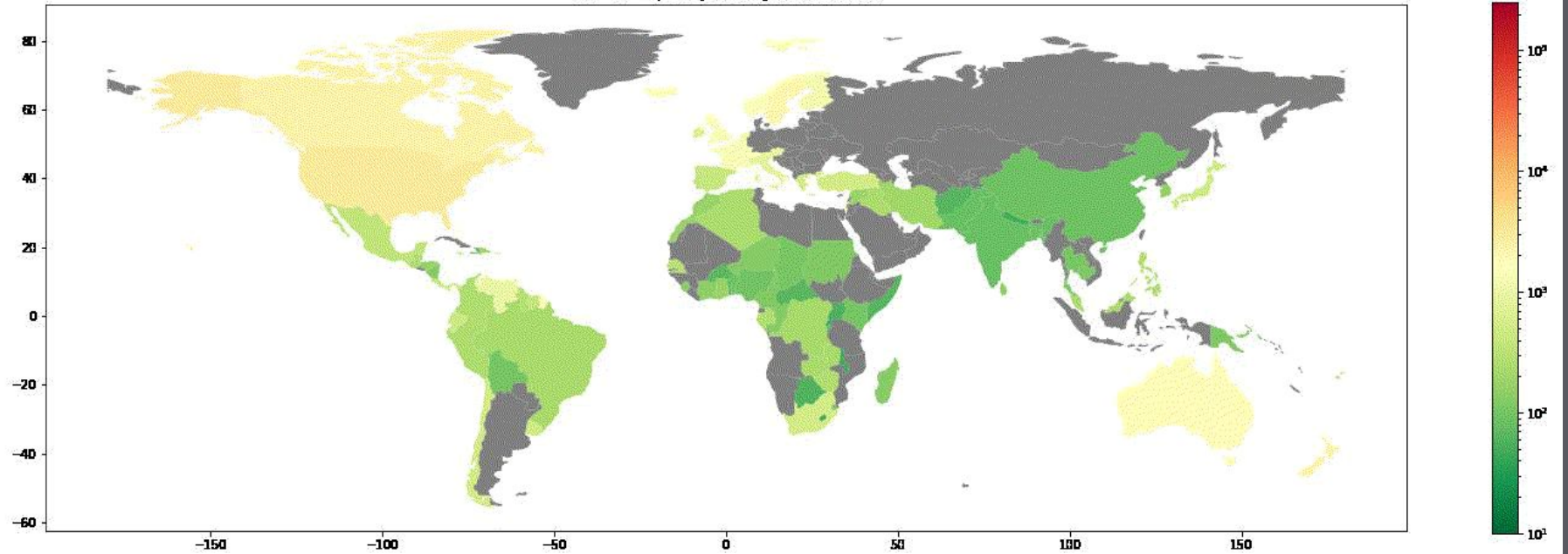
Per Capita Emissions vs Population and Urban Population Percentage



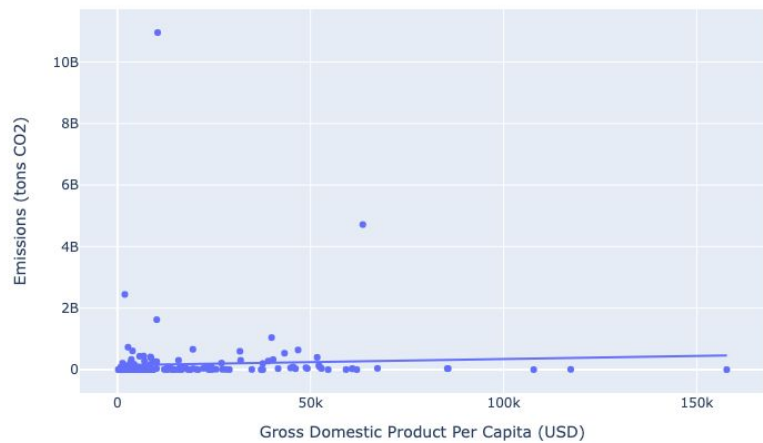
02

GDP Per Capita

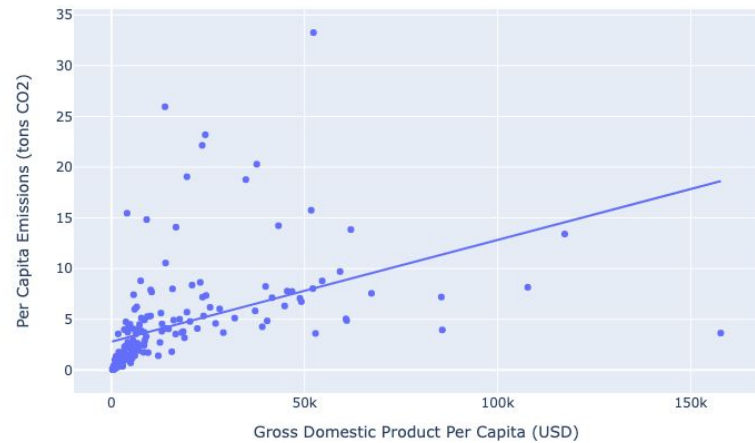
GDP Per Capita by Country in 1960 (USD)



Emissions vs GDP Per Capita



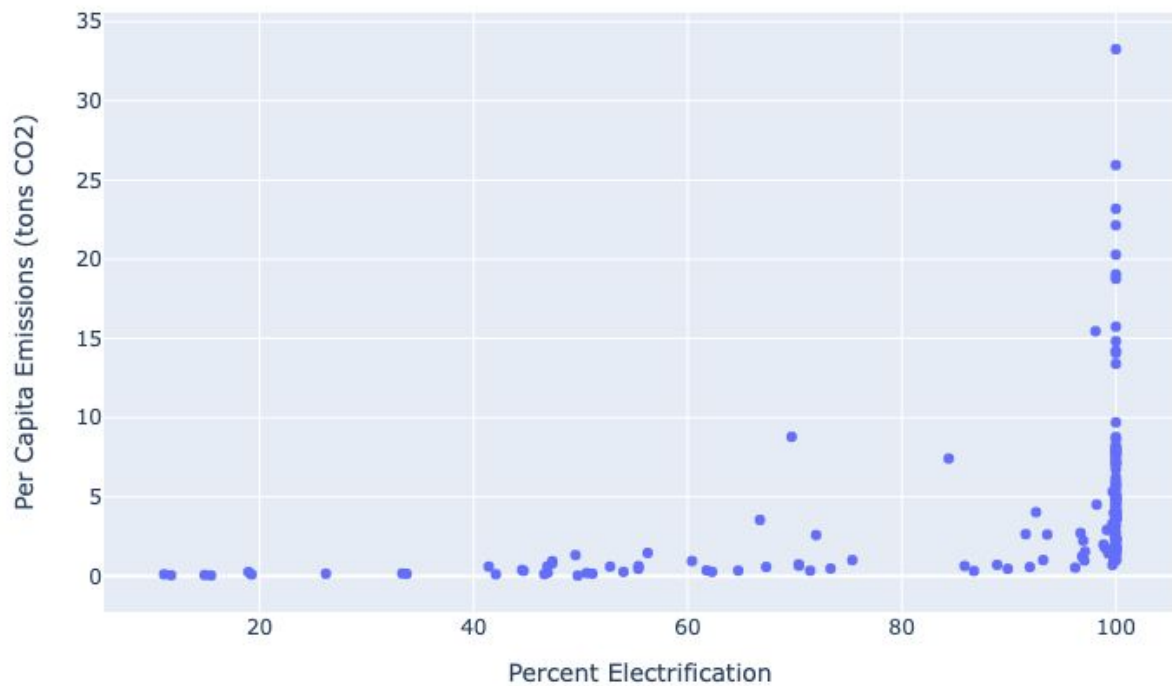
Per Capita Emissions vs GDP Per Capita



03

Electrification

Electrification vs Per Capita Emissions



04

Foreign Direct Investment

FDI vs Emissions



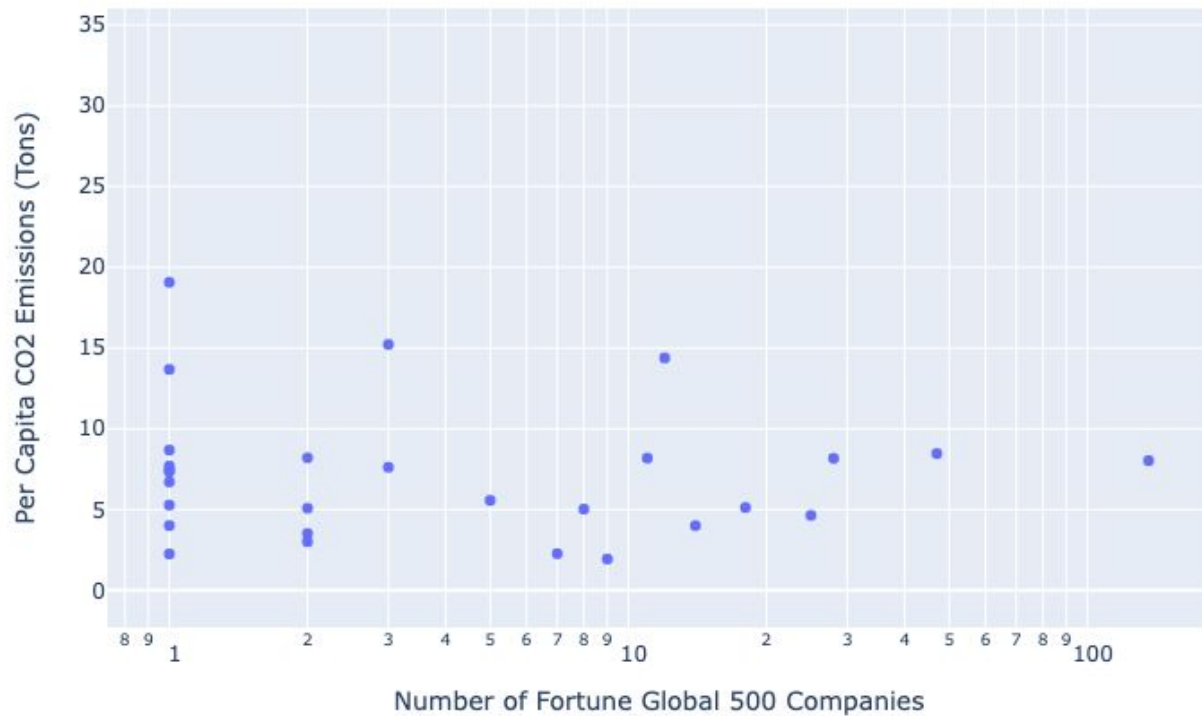
05

Fortune Global 500 Companies

Fortune Global 500 Countries by City



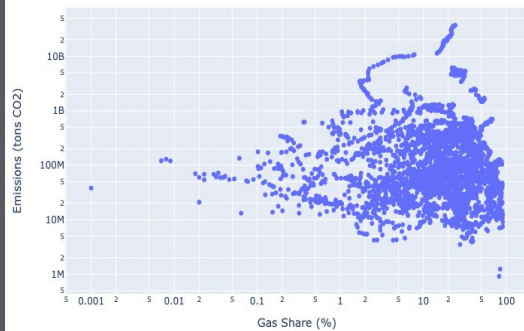
Fortune 500 Companies vs Per Capita Emissions



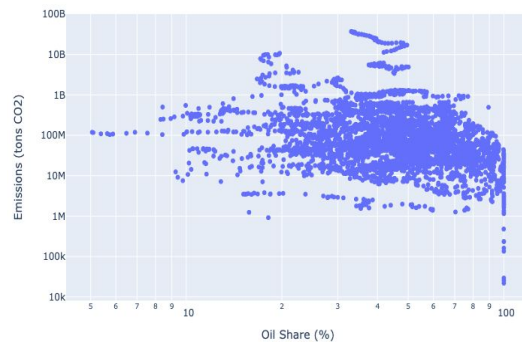
06

Energy Sectors

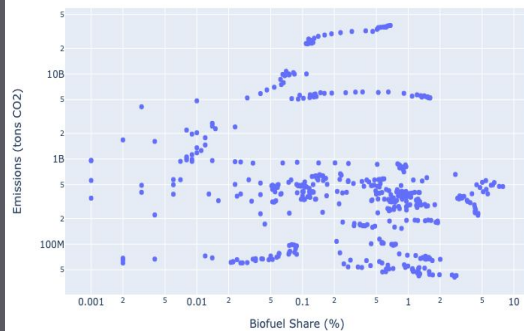
Gas Share (%) vs Emissions



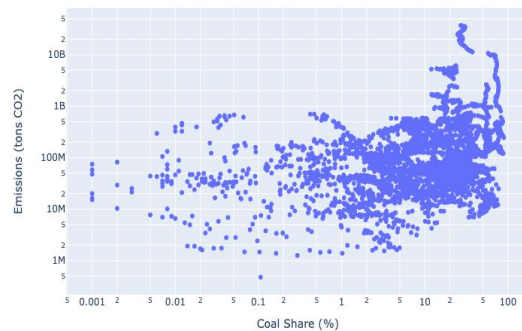
Oil Share (%) vs Emissions



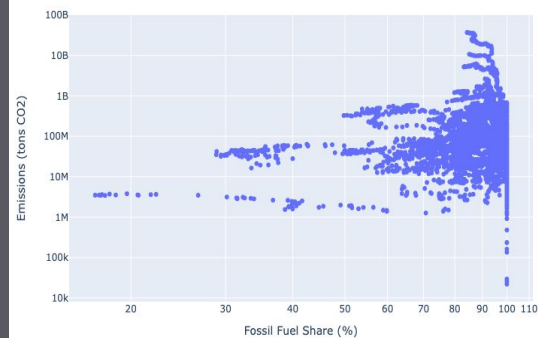
Biofuel Share (%) vs Emissions



Coal Share (%) vs Emissions

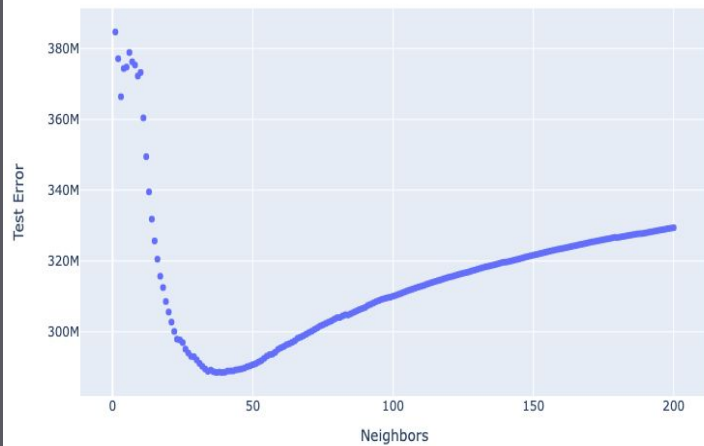


Fossil Fuel Share (%) vs Emissions

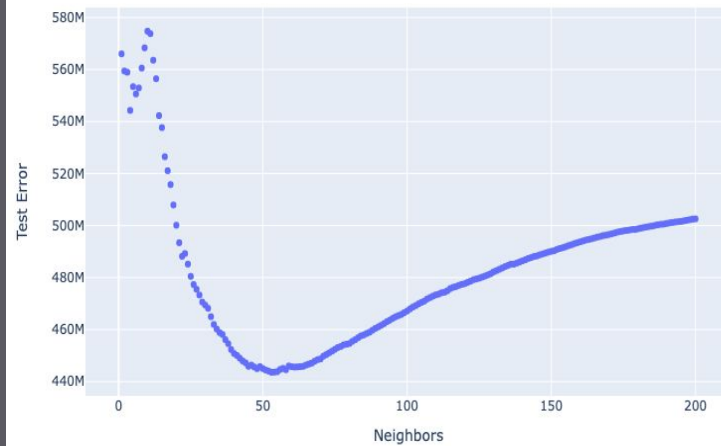


Machine Learning

Test Error for K-Nearest Neighbors Regressor without Energy

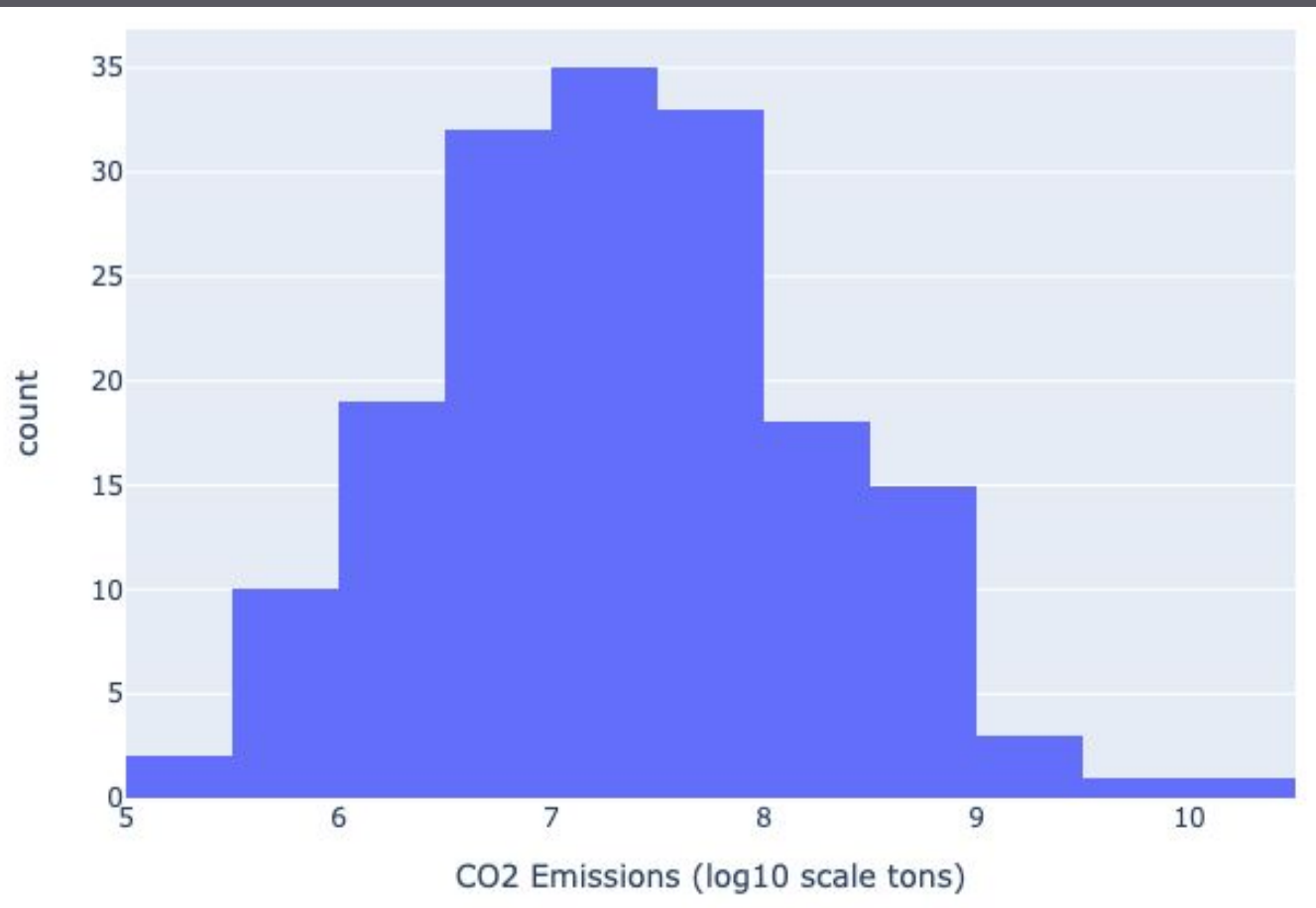


Test Error for K-Nearest Neighbors Regressor with Energy



	Without Energy Data	With Energy Data
K-Nearest Neighbors	2.9×10^8 tons CO ₂	4.4×10^8 tons CO ₂
Linear Regression	4.4×10^8 tons CO ₂	7.9×10^8 tons CO ₂
Voting Ensemble	3.3×10^8 tons CO ₂	5.6×10^8 tons CO ₂
XGBoost	3.2×10^8 tons CO ₂	6.4×10^8 tons CO ₂
XGBoost with Hyperparameter Tuning	2.7×10^8 tons CO ₂	5.1×10^8 tons CO ₂

Log Normal Distribution



	Without Energy Data	With Energy Data
K-Nearest Neighbors	0.711	0.411
Linear Regression	0.846	0.423
Voting Ensemble	0.739	0.374
Stacking Ensemble	0.707	0.385
XGBoost	0.325	0.268
XGBoost with Hyperparameter Tuning	0.301	0.258

Accuracy

Normal RMSE

RMSE > Mean

RMSE \approx 30% std

ML is about 3x as accurate as
guessing the average value

Log Scale RMSE

RMSE = 0.258

$$y_pred * 10^{0.258} = y_true$$
$$y_pred * 1.8 = y_true$$

ML is within a factor of 2
on average

Questions?