

Carbon Emissions Contributors

Kevin Xia, Jeffrey Mu

Project Description

At the beginning of the semester our team was interested in determining the extent to which the establishment of power plants affected CO2 emissions of that country. We wanted to learn if there would we statistically significant effects on the increase of CO2 when a power plant was established in a specific country. We eventually decided to look at power plants and vehicles as factors that affect CO2 emissions.

Methods and Datasets

- Datasets:
- Power Plants Established By Country (World Resources Institute)
 - Carbon Emissions by Country (World Bank)
 - Number of Vehicles by Country (World Bank)
 - Population By Country Each Year (World Bank)

- Created:
- Fossil Fuel Power Plants
 - For every country, all years where a power plant was established
 - Carbon Emissions Total
 - Multiplied carbon country and year values (per capita) with population year values to get total carbon emissions
 - Number Vehicles Country
 - Got total number of vehicles per year for each country

- Methods:
- Constraint Satisfaction, Linear Regression. Hypothesis Testing

Approach

Constraint Satisfaction

Our goal in employing constraint satisfaction as a technique was mainly to determine if our intuition about the effects of power plants and vehicles was correct. We assumed that if either factor had a significant effect on CO2 emissions for that country, we should expect to see that when we analyze the CO2 emissions before a power plant was established, and after a power plant was established. We did the same process for the vehicle data but less strict. We wanted to just ensure that a positive change in vehicles corresponded with a positive change in CO2 emissions. This property existed in most of the years, but there were a few countries that did not experience this property.

Linear Regression/Hypothesis Testing

To bolster our results from the constraint satisfaction analysis, we wanted to make sure that our model fit the data we had. We also wanted to do hypothesis testing, and actually figure out how significant our results are. We performed a linear regression analysis on both the number of power plants vs. CO2 emissions and number of vehicles per year vs CO2 emissions. In this portion our null hypothesis was that there is no relationship between the two variables. Our alternative hypothesis is that there is a linear relationship between the two variables. Using Scipy.stats.linreg we calculated the slope, intercept, R-value, R-squared, and P-value of the two variable pairs. With an alpha value of 0.05, any P-value that was under 0.05 or close to 0.05 meant that we can accept our alternative hypothesis and reject the null. Any P-value greater than 0.05 meant that we cannot reject the null. More analysis of these results in the results section.

Results

Country	Slope	Intercept	R-Value	R-Squared	P-Value
Austria	-6.9488611947947705	102729078.23176225	-0.26102888680827485	0.06813607974836716	0.532348256084377
Azerbaijan	-0.23245055199930809	32130079.313893676	-0.08677850055973631	0.007530508159396154	0.8896490126003144
Belgium	-10.065586883352301	165079611.4425215	-0.5125087461644895	0.26266521489509714	0.19405905449854105
Bangladesh	83.43307803488574	18524233.99854636	0.8932220881985942	0.7978456988458571	0.0067507391214175925
Bahrain	42.54556597438675	3274510.5284418054	0.935685307355121	0.8755069944002473	0.001945805786414783
Brazil	5.6931155673999445	153602353.69495332	0.9759974621553578	0.952571046133699	0.0008572685415455821
Canada	-2.1528303453638378	591872433.1497535	-0.17945138763910012	0.032202800525598575	0.7727476970430496

Figure 1: Screenshot from web application showing the slope, intercept, r-value, r-squared, and p-value for countries

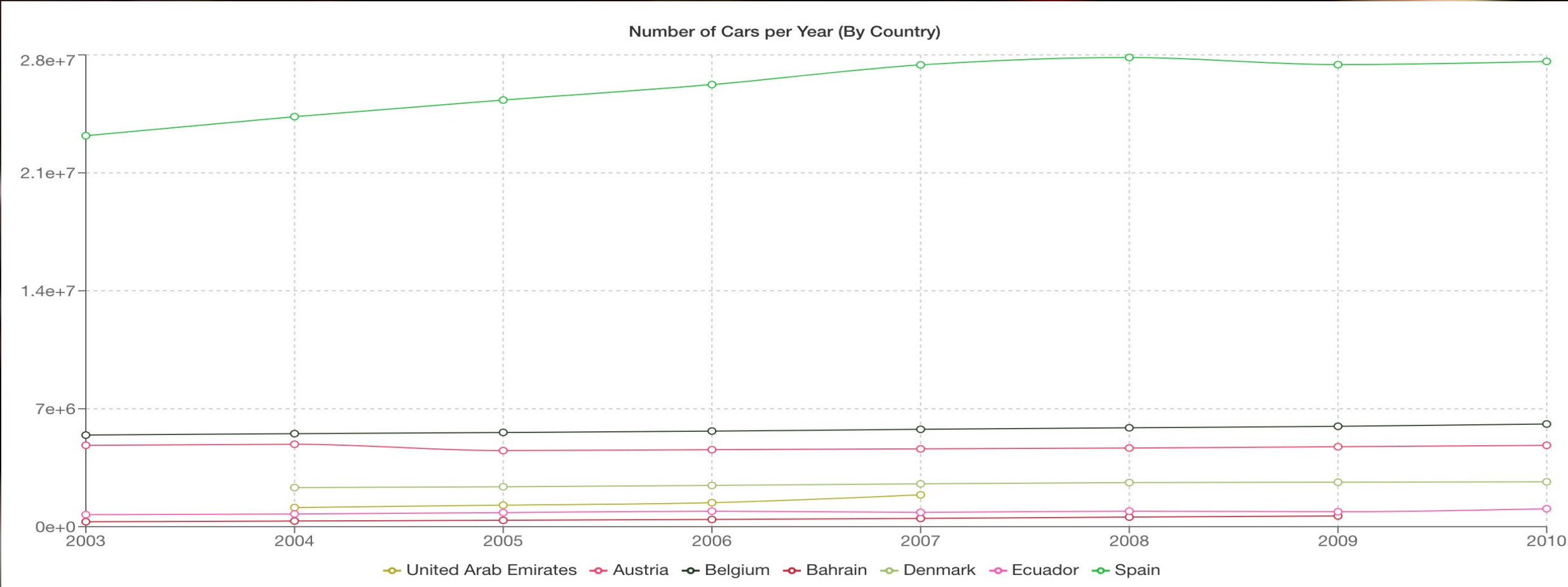


Figure 2: Number of cars per year for different countries

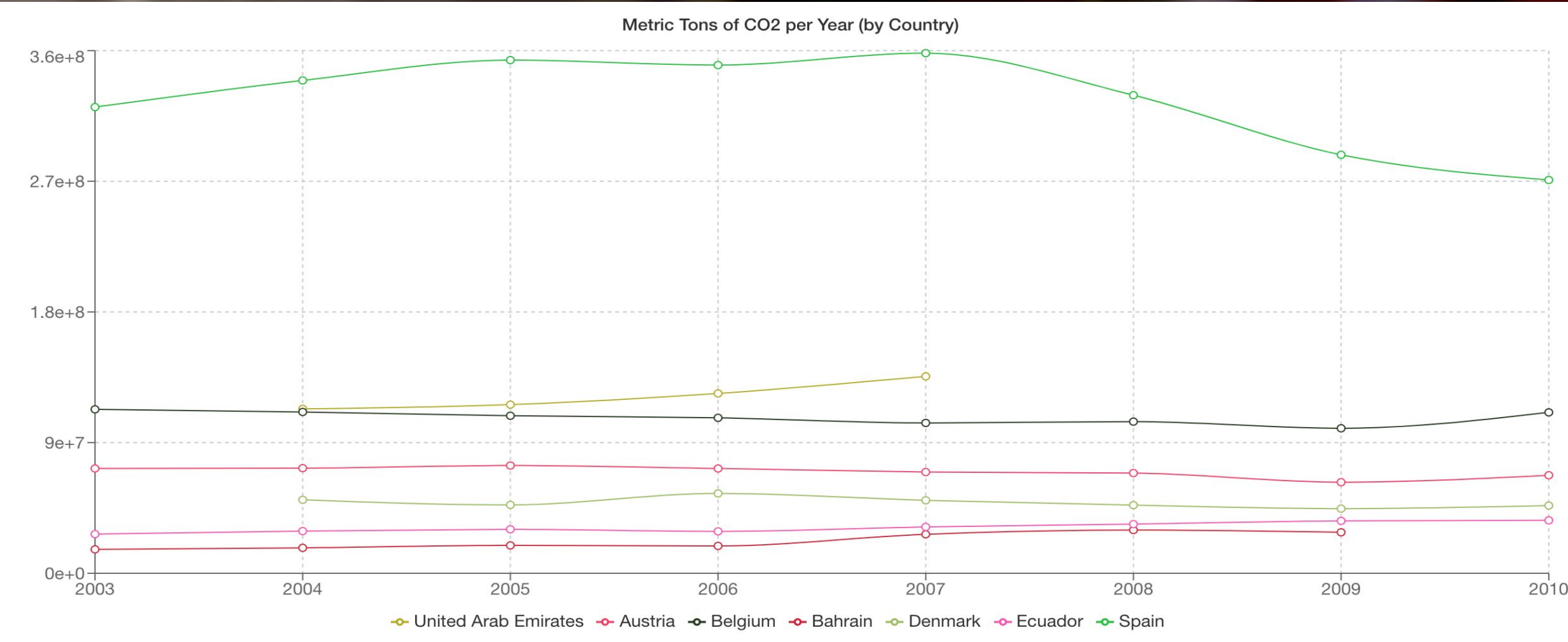


Figure 3: CO2 Emissions (Metric Tons) per year for different countries