Hunter Breen

Patrick Conway

Info 150

Discrete Mathematics Project Write-Up

For the Informatics 150 Discrete Mathematics project, we chose to do a study of the economic standing of all of the countries in the world. The reason why we chose to do such a topic is because we both have an interest in data science, world dynamics, and global economics. Pat studies economics, and we both have a strong interest in political and socio-economic standings of the various countries of the world. We felt as if this was the most suitable and interesting topic to do for our project as it was an opportunity to apply the skills and methodology of discrete mathematics to a field that we had genuine interest and knowledge in. At the heart of our project we specifically wanted to compare the standard of living, GDP, GDP per Capita, GNI, inflation, and population of the top and bottom 10% of countries in the world in order to draw implications on what makes a country flourish or not.

To start our project, we first had to find a data set to bring us the raw data. From there, we scoured the internet to find the most suitable data set. After sifting through a few sheets, we ended up using the official World Development Indicators data set from the World Bank[1]. This Data Sheet is made public by the World Bank, and it is updated quarterly with the last update being on the 15th of September, 2017. It is a Time Series data set, meaning it shows the data of every variable starting from 1960 to 2017 with the most accuracy available for each variable. The data, according to the World Bank, is “compiled from officially-recognized international sources”[2].

When we downloaded and imported the data set into Microsoft Excel, we found the sheer size of the data set to be incredibly massive and dense, as it had 6 sheets, all with a large amount of rows and columns. As our first real data science project, we found the approximately 410,000 rows within the main data sheet to be pretty intimidating. While the both of us have some experience with Excel, neither of us knew how to use the software to its maximum potential. Our solution to this was to use key searches and organizational skills learned in other areas of mathematics and computing to represent the data in such a way that it would be comprehensive to us, and would allow us to draw the best conclusions from the massive data set. To start, we went through each sheet and judged its relevance to our project. Most of the sheets just had code words to abbreviate certain variables. After whittling the sheets down to just two sheets, we were then stuck with the two largest sheets. In order to bring some cohesion to the vast sea of data, we had sorted the variables column to be in lexicographical order, so we could then view all of the countries and regions included in the data set. An alternate way in which we had sorted the data was by lexicographically sorting the countries, so we could then view how many variables were available per country. Using the built in search functions, we had searched through the variables to see if we could find key terms that were related to the project as the staggering number of variables muddied up finding the data by just looking at it. For instance, we wanted to find GDP per Capita for every country, however variables such as “Merchandise imports from economies in the Arab World” were deemed far to specific for our project and watered down the information we were looking at. We used specific search queries within Excel to find and optimized the data we were looking for.

Once we had found a search and ordering method to optimize our data visualization and representation, we then looked for the variables that we felt represented a country to its most representational variables. These variables were GDP, GNI, GDP per capita, unemployment, inflation, net exports, life expectancy, CO2 emissions, land size, rural & urban populations, and urban population growth. Once we had decided on these variables, we had imported the relevant data sheets into Google Sheets, so we could both systematically manipulate the data to increase productivity. To get the most accurate data we chose the year 2016 for all the variables, as it had the most columns filled out. After searching and sifting through the data, we had taken each of our selected variables for every country and region in the world, and copied it over to a seperate sheet for organizational purposes. Once we had compiled all of our relevant pieces of data, we decided it would be more accurate to remove all of the country entries that depicted entire regions and continents, as we wanted to do this on a country to country basis. To draw up some formal conclusions and implications, we thought that we should compare certain variables to one another. We had chose to compare the following:

- Unemployment(% of total labor force), Inflation of consumer prices(annual %)
- GDP, GNI, GDP Per Capita (All in the current US $)
- GDP, Net Exports (al in US $)
- GDP per capita(US $), Unemployment(% of total labor force), Life Expectancy (years), CO2 Emissions(metric tons per capita)
- Land Area(Sq. KM), Urban Population, Rural Population, Total Population
- Urban population growth (annual %), Inflation of consumer prices (annual %)

Upon drawing up these comparisons, the density of approximately 200 countries was somewhat overwhelming, so we had narrowed it down to the top 10% of countries in the world, and the bottom 10% based on GDP alone. There seemed to be a lot of data missing for the bottom 10% of countries based on GDP, so we decided to use a negative statistic to compare against the top 10% GDP. That statistic was top 10% unemployment. We then decided to use the comparisons listed above to compare the top 10% GDP to the top 10% Unemployment as our main study. We organized the comparisons, and then pitted the GDP based countries to the unemployment based countries against each other. Making mock-pivot tables is how we went about arranging this data.

In economics, there is an inverse relationship between the rate of change of unemployment and the rate of change of inflation. It is represented by the phillips curve. This relationship occurs because when the economy begins to speed up more workers are required. As the economy speeds up more and more wages must rise to be competitive in hiring a new employee. As the nominal wages rise, companies increase the price of the product to account for the lost profit from the wage increase. This sequence of events is the reason that unemployment and inflation have an inverse relationship. In our data set, the top 20 unemployment countries had a significantly higher inflation rate than that of the top 20 GDP countries. Some of the countries on our list experienced a negative level of inflation. In economics, this is called deflation. It occurs when the price of good increases faster than wages. An easy way to think about it is an increase in the value of a currency. We experienced a slight positive relationship between unemployment in 2016 and inflation in 2016, which we initially thought was odd. After some research, we realized that this was not odd at all. The phillips curve only presents itself

when you are taking the rates of change of unemployment and inflation into account, so when we are comparing only data from 2016 the phillips curve will not appear because we are not taking the rates of change into account, only the unemployment level and the inflation level.

Country	Urban population growth (annual %)	Inflation, consumer prices (annual %)
United States	0.902097108	1.261583206
China	2.612874387	2
Japan	0.343564293	-0.116666667
Germany	1.471035055	0.483355422
United Kingdom	1.07132712	0.641613199
France	0.696308865	0.183334861
India	2.329109667	4.941447235
Italy	0.005885377	-0.123335389
Brazil	1.104235786	8.739478523
Canada	1.431183937	1.428759547
Korea, Rep.	0.594291337	0.97
Russian Federation	0.295827456	7.049766291
Spain	0.271783712	-0.202671741
Australia	1.556450202	1.276990945
Mexico	1.64312093	2.821707752
Indonesia	2.475676863	3.525805157
Turkey	2.238395514	7.775134153
Netherlands	1.052785082	0.316666667
Switzerland	1.182683765	-0.434632563
Saudi Arabia	2.492920855	3.523510972
Country	**Urban population growth (annual %)**	**Inflation, consumer prices (annual %)**
Solomon Islands	4.006945688	
Gambia, The	4.021492946	7.220092431
Lesotho	3.247536073	6.609116058

Macedonia, FYR	0.266201315	-0.239290788
South Africa	2.376645257	6.3262638
Bosnia and Herzegovina	-0.108867286	-1.251444391
Swaziland	1.842346623	7.84641182
Namibia	4.255007914	6.711963782
West Bank and Gaza	3.188556818	-0.219106608
Mozambique	3.789714313	9.966216216
Greece	-0.275605998	-0.825657581
Comoros	2.720188525	
Libya	1.184934295	
St. Vincent and the Grenadines	0.858405264	-0.149641648
St. Lucia	0.650517539	-3.09324257
Botswana	2.303663536	3.769304598
Spain	0.271783712	-0.202671741
Gabon	2.77849466	2.106707128
Oman	5.79180675	1.101052116

We decided to compare the gross domestic product, the gross national income, and the gross domestic product per capita due to their perceived similarity. Gross domestic product (GDP) is equal to government spending plus investment spending plus consumer spending plus net exports. Gross national income is equal to the gross domestic product plus net income from overseas business. Gross domestic product per capita is equal to the gross domestic product divided by the population. As expected, GDP and GNI had a strong relationship. The relationship was so strong that the Top 20 GDP and Top 20 GNI match up exactly. We also expected the GDP for the top 20 countries to have a relationship with the GDP per capita, but this was not the case. There was no relationship whatsoever between the top 20 GDP countries,

but the top 20 GDP countries GDP and GDP per capita were much higher than the top 20 Unemployment countries. For example, Switzerland has the highest GDP per capita in the world, but they have the 19th highest GDP. This difference in GDP per capita is based upon the population difference of the countries. Switzerland has a relatively low GDP and population compared to the other countries, resulting in a high GDP per capita. Comparing different indicators to GDP is very useful because GDP is one of the most important indicators in the entirety of the economy.

Country	GDP (current US$)	GNI (current US$)	GDP per capita (current US$)
United States	1.86E+13	1.87E+13	57466.78711
China	1.12E+13	1.12E+13	8123.180873
Japan	4.94E+12	5.11E+12	38894.46773
Germany	3.47E+12	3.52E+12	41936.05858
United Kingdom	2.62E+12	2.59E+12	39899.38839
France	2.47E+12	2.52E+12	36854.96828
India	2.26E+12	2.24E+12	1709.387921
Italy	1.85E+12	1.85E+12	30527.2682
Brazil	1.80E+12	1.76E+12	8649.948492
Canada	1.53E+12	1.51E+12	42157.92799
Korea, Rep.	1.41E+12	1.41E+12	27538.80613
Russian Federation	1.28E+12	1.25E+12	8748.364504
Spain	1.23E+12	1.23E+12	26528.49179
Australia	1.20E+12	1.18E+12	49927.81951
Mexico	1.05E+12	1.04E+12	8201.306253
Indonesia	9.32E+11	9.01E+11	3570.294888
Turkey	8.58E+11	8.49E+11	10787.60934
Netherlands	7.71E+11	7.62E+11	45294.78
Switzerland	6.60E+11	6.68E+11	78812.65069
Saudi Arabia	6.46E+11	6.62E+11	20028.64821

Country	GDP (current US$)	GNI (current US$)	GDP per capita (current US$)
Solomon Islands	1202125000	1153375000	2005.483643
Gambia, The	964599177.5	936815006.9	473.1904363
Lesotho	2199709489	2462946647	998.1343716
Macedonia, FYR	10899583155	10437003632	5237.14767
South Africa	2.95E+11	2.87E+11	5273.59388
Bosnia and Herzegovina	16559695719	16650427012	4708.718261
Swaziland	3727303664	3563097682	2775.153908
Namibia	10267157280	10235043777	4140.461932
West Bank and Gaza	13397100000	14975800000	2943.404534
Mozambique	11014858592	10760007923	382.0693304
Greece	1.95E+11	1.95E+11	18103.96932
Comoros	616654490.5	617610069.8	775.0800847
Libya			
St. Vincent and the Grenadines	770796555.6	765861555.6	7030.057145
St. Lucia	1378627407	1358227074	7744.445173
Botswana	15274861068	14954625955	6788.042745
Spain	1.23E+12	1.23E+12	26528.49179
Gabon	14213558130	13125284553	7179.340661
Oman	66293368010		14982.35792

Comparing gross domestic product and net exports to see if there was a relationship between the two was a waste of time. Net exports are equal to the exports of a country minus the imports of a country. If a net export value is negative, that means that the country imports more than it exports. We were expecting the GDP of a country to have a positive relation with the net exports, but the data we were able to compile for this comparison was very slim. Missing data

was one of the biggest problems we faced while trying to break down the large data set into smaller, more understandable packages.

Country	GDP (current US$)	Exports of goods and services (current US$)	Imports of goods and services (current US$)	Net Exports Current US Dollar
United States	1.86E+13		620304980	
China	1.12E+13	231491150.2		
Japan	4.94E+12		2048855657	
Germany	3.47E+12	5367444081	6533387362	-1165943281
United Kingdom	2.62E+12	89401819220	1.00E+11	-1.07E+10
France	2.47E+12		8625867056	
India	2.26E+12		1973838378	
Italy	1.85E+12		7916798686	
Brazil	1.80E+12	2432400000	7501704675	-5069304675
Canada	1.53E+12	3828497482	10598848530	-6770351048
Korea, Rep.	1.41E+12	58666448637	58805896735	-139448098
Russian Federation	1.28E+12	106681226.9		106681226.9
Spain	1.23E+12			
Australia	1.20E+12	198069222.2		
Mexico	1.05E+12	658518518.5		
Indonesia	9.32E+11		7712275862	
Turkey	8.58E+11	4.07E+11	3.67E+11	4.08E+10
Netherlands	7.71E+11	5837511878		
Switzerland	6.60E+11		36665926391	
Saudi Arabia	6.46E+11			
Country	**GDP (current US$)**	**Exports of goods and services (current US$)**	**Imports of goods and services (current US$)**	**Net Exports Current US Dollar**
Solomon Islands	1202125000		620304980	

Gambia, The	964599177.5	231491150.2		
Lesotho	2199709489		2048855657	
Macedonia, FYR	10899583155	5367444081	6533387362	-1165943281
South Africa	2.95E+11	89401819220	1.00E+11	-1.07E+10
Bosnia and Herzegovina	16559695719		8625867056	
Swaziland	3727303664		1973838378	
Namibia	10267157280		7916798686	
West Bank and Gaza	13397100000	2432400000	7501704675	-5069304675
Mozambique	11014858592	3828497482	10598848530	-6770351048
Greece	1.95E+11	58666448637	58805896735	-139448098
Comoros	616654490.5	106681226.9		
Libya				
St. Vincent and the Grenadines	770796555.6	198069222.2		
St. Lucia	1378627407	658518518.5		
Botswana	15274861068		7712275862	
Spain	1.23E+12	4.07E+11	3.67E+11	4.08E+10
Gabon	14213558130	5837511878		
Oman	66293368010		36665926391	

In order to determine a standard of living, we created a table that compared the GDP per capita of a country, the unemployment of a country, the average life expectancy of a country, and the CO2 emissions per capita. As expected, there was a positive relationship between GDP per capita and life expectancy. This was expected because countries with higher GDP per capita typically have better hospitals and therefore a longer life expectancy. With this comes the expectation that the top 20 unemployment countries have a much lower life expectancy than the top 20 GDP countries. The top 20 Unemployment countries have a significantly lower volume of CO2 emissions than the top 20 GDP countries.One of the most interesting parts of our project was the fact that there was no relation between CO2 emissions per capita and life expectancy in 2016.

Country	GDP per capita (current US$)	Unemployment, total (% of total labor force) (modeled ILO estimate)	Life expectancy at birth, total (years)	CO2 emissions (metric tons per capita)
United States	57466.78711	4.897999763	78.74146341	16.49366613
China	8123.180873	4.639999866	76.11709756	7.543907641
Japan	38894.46773	3.000999928	83.84365854	9.538706103
Germany	41936.05858	4.157000065	81.0902439	8.889370395
United Kingdom	39899.38839	5.012000084	81.60487805	6.49744049
France	36854.96828	9.840999603	82.67073171	4.572088353
India	1709.387921	3.427999973	68.33260976	1.730000432
Italy	30527.2682	11.40799999	83.4902439	5.270866786
Brazil	8649.948492	12.35499954	75.19912195	2.594388285
Canada	42157.92799	7.127999783	82.12902439	15.1132392
Korea, Rep.	27538.80613	3.614000082	82.15585366	11.57034541
Russian	8748.364504	5.798999786	70.90853659	11.85752777

Federation				
Spain	26528.49179	18.32600021	83.3804878	5.033824487
Australia	49927.81951	5.460000038	82.45121951	15.39859985
Mexico	8201.306253	4.050000191	76.88136585	3.866241113
Indonesia	3570.294888	5.769000053	69.03863415	1.819363319
Turkey	10787.60934	10.83699989	75.40709756	4.491479026
Netherlands	45294.78	5.605999947	81.70731707	9.920138075
Switzerland	78812.65069	4.581999779	83.19756098	4.311562994
Saudi Arabia	20028.64821	5.539000034	74.5744878	19.5292718
Country	**GDP per capita (current US$)**	**Unemployment, total (% of total labor force) (modeled ILO estimate)**	**Life expectancy at birth, total (years)**	**CO2 emissions (metric tons per capita)**
Solomon Islands	2005.483643	31.43600082	70.47668293	0.350449345
Gambia, The	473.1904363	29.70599937	60.95743902	0.267684889
Lesotho	998.1343716	27.50499916	53.57009756	1.150111032
Macedonia, FYR	5237.14767	27.25399971	75.52431707	3.614938183
South Africa	5273.59388	25.99799919	61.93409756	9.045270345
Bosnia and Herzegovina	4708.718261	25.4489994	76.64029268	6.234719162
Swaziland	2775.153908	25.11199951	56.91236585	0.928714992
Namibia	4140.461932	24.9090004	63.63643902	1.583728667
West Bank and Gaza	2943.404534	24.1989994	73.30168293	
Mozambique	382.0693304	24.1439991	57.60926829	0.309666607
Greece	18103.96932	23.02799988	81.58780488	6.180337268
Comoros	775.0800847	20.01399994	63.46068293	0.202814119
Libya		19.21999931	71.83039024	9.186845393

St. Vincent and the Grenadines	7030.057145	19.13699913	73.06495122	1.91134541
St. Lucia	7744.445173	18.95999908	75.2835122	2.307191321
Botswana	6788.042745	18.56500053	65.75109756	3.243287637
Spain	26528.49179	18.32600021	83.3804878	5.033824487
Gabon	7179.340661	18.14299965	65.68441463	2.768265721
Oman	14982.35792	17.82099915	77.12182927	15.44316719

In any given country, the urban population and the rural population have an inverse relationship. This is because as the people leave the country (or the city) they go to the city (or the country). Something to note is that the countries in the top 20 GDP have a higher population than the countries in the top 20 unemployed. When comparing different countries, the higher the population the higher the urban and rural population of a country. It will be interesting to see if the urban population has anything to do with economic growth.

Country	Land area (sq. km)	Urban population	Population, total	Rural Population
United States	83600	264279530	323127513	58847983
China	23618361	782778414	1378665000	595886586
Japan	10120	119283404	126994511	7711107
Germany	1259200	62422369	82667685	20245316
United Kingdom	579290	54370607	65637239	11266632
France	176520	53349647	66896109	13546462
India	101000	438777420	1324171354	885393934
Italy	3660	41884704	60600590	18715886
Brazil	23154350	178442336	207652865	29210529
Canada	64200	29757046	36286425	6529379
Korea, Rep.	410450	42324854	51245707	8920853
Russian Federation	397300	106959159	144342396	37383237

Spain	20140	37063208	46443959	9380751
Australia	4238213	21606836	24127159	2520323
Mexico	62180	101416318	127540423	26124105
Indonesia	42390	142219144	261115456	118896312
Turkey	5130	58749346	79512426	20763080
Netherlands	180	15492197	17018408	1526211
Switzerland	17200	6194515	8372098	2177583
Saudi Arabia	8870	26895653	32275687	5380034
Country	**Land area (sq. km)**	**Population, total**	**Urban population**	**Rural Population**
Solomon Islands	87460	599419	136542	462877
Gambia, The	240	2038501	1227667	810834
Lesotho	2973190	2203821	613544	1590277
Macedonia, FYR	88780	2081206	1190533	890673
South Africa	72180	55908865	36505693	19403172
Bosnia and Herzegovina	94659576	3516816	1404616	2112200
Swaziland	2376000	1343098	286281	1056817
Namibia	300	2479713	1180963	1298750
West Bank and Gaza	350	4551566	3435568	1115998
Mozambique	94280	28829476	9371886	19457590
Greece	9388211	10746740	8417814	2328926
Comoros	129733172.7	795601	226046	569555
Libya	1628760	6293253	4956188	1337065
St. Vincent and the Grenadines	500210	109643	55807	53836
St. Lucia	1213090	178015	33004	145011
Botswana	13382221	2250260	1298625	951635
Spain	20140	46443959	37063208	9380751
Gabon	9093510	1979786	1729660	250126
Oman	653080	4424762	3455208	969554

We paired urban population growth and inflation together because we assumed there would be a positive relationship between the two, since inflation is a sign of economic growth and much of the economic growth takes place in urban areas. We were correct in assuming that there was a positive relationship between urban population growth and inflation. This can be shown by the fact that there is higher urban populations growth than most of the top 20 unemployed countries.

Country	Urban population growth (annual %)	Inflation, consumer prices (annual %)
United States	0.902097108	1.261583206
China	2.612874387	2
Japan	0.343564293	-0.116666667
Germany	1.471035055	0.483355422
United Kingdom	1.07132712	0.641613199
France	0.696308865	0.183334861
India	2.329109667	4.941447235
Italy	0.005885377	-0.123335389
Brazil	1.104235786	8.739478523
Canada	1.431183937	1.428759547
Korea, Rep.	0.594291337	0.97
Russian Federation	0.295827456	7.049766291
Spain	0.271783712	-0.202671741
Australia	1.556450202	1.276990945
Mexico	1.64312093	2.821707752
Indonesia	2.475676863	3.525805157
Turkey	2.238395514	7.775134153
Netherlands	1.052785082	0.316666667
Switzerland	1.182683765	-0.434632563
Saudi Arabia	2.492920855	3.523510972

Country	Urban population growth (annual %)	Inflation, consumer prices (annual %)
Solomon Islands	4.006945688	
Gambia, The	4.021492946	7.220092431
Lesotho	3.247536073	6.609116058
Macedonia, FYR	0.266201315	-0.239290788
South Africa	2.376645257	6.3262638
Bosnia and Herzegovina	-0.108867286	-1.251444391
Swaziland	1.842346623	7.84641182
Namibia	4.255007914	6.711963782
West Bank and Gaza	3.188556818	-0.219106608
Mozambique	3.789714313	9.966216216
Greece	-0.275605998	-0.825657581
Comoros	2.720188525	
Libya	1.184934295	
St. Vincent and the Grenadines	0.858405264	-0.149641648
St. Lucia	0.650517539	-3.09324257
Botswana	2.303663536	3.769304598
Spain	0.271783712	-0.202671741
Gabon	2.77849466	2.106707128
Oman	5.79180675	1.101052116

To reiterate, we found a positive relationship between unemployment and inflation, which went against what we were expecting. The GDP and the GNI were directly related, but the GDP per capita is not related to GDP or GNI. We expected net exports to increase as the GDP went up, but there was no definitive proof of this in our data set. In any given country, the urban and rural populations are inverse. Urban population growth and inflation had a positive relationship.

As our first true data science project, we found this to be a very fun way to throw ourselves into the world of big data. It also helped that we had the freedom to use something that we mutually liked as a platform to demonstrate our skills that we learned in discrete mathematics and computing. Viewing global socio-economic data was a fun exercise for us. It allowed us to bring together our fields of study in economics, discrete mathematics, and computer science to an area of interest, which was a huge motivating factor for us. Along the way, we learned how to manage massive amounts of quantitative data in Excel, using searches and organizational methods. Also, we were able to draw many implications and comparisons from the information that we found, giving us some further insight as to how dynamic socio economics can be in the world, employing skills learned in computer science and discrete mathematics.

Works Cited

[1] The World Bank, "World Development Indicators" datasheet, 1960 [Revised Sept. 15, 2017]

[2] The World Bank, "World Development Indicators" *World Bank,* [Online]. Available: https://data.worldbank.org/data-catalog/world-development-indicators [Accessed: Nov. 28, 2017].