

Impacts of Trade on National Productivity and Well-Being



TEAM EINHERJAR

TOM, MARIA, BEN, JEMI

Core Message & Hypothesis of the Project



If the GDP of a nation rises, then the nation will be able to import more and experience improvements in well-being (life expectancy & infant mortality).



However, the nation will export less as it becomes more productive, since it will be able to produce more of its own goods.

Questions & Motivations



We were most interested in first seeing the strength of the relationship between GDP per Capita and trade values (imports & exports), with the assumption that exports would decrease in correlation to GDP increase.

Then we were interested in seeing how the GDP per Capita would effect national well-being (life expectancy & infant mortality).

The main goal was to best identify how to improve the well-being of a nation.

Sources of data

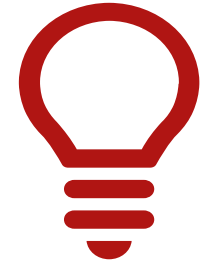


Originally, we started by downloading data for both GDP and World Development Indicators on Kaggle as well as the World Bank website.



We ran into the following issues:

We couldn't find columns to merge all the data sets.
The Kaggle data didn't have data for years past 1980.
There was a need to transpose the year rows in Kaggle data to get indicator value-years as columns.
This would exponentially increase the size of the merged data-frames, complicating merges and Git pushing.



For this reason, we decided to exclusively use the World Bank website.

Data cleanup process

```
In [3]: #Add Suffixes to Columns Except Country Code
#Code found on https://stackoverflow.com/questions/34049618/how-to-add-a-suffix-or-prefix-to-each-column-name
#Add a Suffix to Differentiate Columns For Merge
gdp_per_capita_df = gdp_per_capita_df.add_suffix("_GDP_Cap")
#Drop Suffix from 'CountryCode' column to Allow Merge
gdp_per_capita_df = gdp_per_capita_df.rename(columns = {"CountryCode_GDP_Cap":"CountryCode"})

#Add a Suffix to Differentiate Columns For Merge
life_expectancy_df = life_expectancy_df.add_suffix("_Life")
#Drop Suffix from 'CountryCode' column to Allow Merge
life_expectancy_df = life_expectancy_df.rename(columns = {"CountryCode_Life":"CountryCode"})

#Add a Suffix to Differentiate Columns For Merge
infant_mortality_df = infant_mortality_df.add_suffix("_Mortality")
#Drop Suffix from 'CountryCode' column to Allow Merge
infant_mortality_df = infant_mortality_df.rename(columns = {"CountryCode_Mortality":"CountryCode"})

#Add a Suffix to Differentiate Columns For Merge
exports_df = exports_df.add_suffix("_Exports")
#Drop Suffix from 'CountryCode' column to Allow Merge
exports_df = exports_df.rename(columns = {"CountryCode_Exports":"CountryCode"})

#Add a Suffix to Differentiate Columns For Merge
imports_df = imports_df.add_suffix("_Imports")
#Drop Suffix from 'CountryCode' column to Allow Merge
imports_df = imports_df.rename(columns = {"CountryCode_Imports":"CountryCode"})
```

Results after clean up

	CountryName_GDP_Cap	CountryCode	IndicatorName_GDP_Cap	IndicatorCode_GDP_Cap	1960_GDP_Cap	1961_GDP_Cap	1962_GDP_Cap	1963_GDP_Cap
0	Aruba	ABW	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
1	Afghanistan	AFG	GDP per capita (current US\$)	NY.GDP.PCAP.CD	59.773194	59.860874	58.458015	78.706
2	Angola	AGO	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
3	Albania	ALB	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
4	Andorra	AND	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
...
259	Kosovo	XKX	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
260	Yemen, Rep.	YEM	GDP per capita (current US\$)	NY.GDP.PCAP.CD	NaN	NaN	NaN	
261	South Africa	ZAF	GDP per capita (current US\$)	NY.GDP.PCAP.CD	443.009816	454.962013	473.011615	511.497
262	Zambia	ZMB	GDP per capita (current US\$)	NY.GDP.PCAP.CD	232.188867	220.042137	212.578123	213.896

► We merged the tables for GDP per Capita, trade values, and well-being indicators.

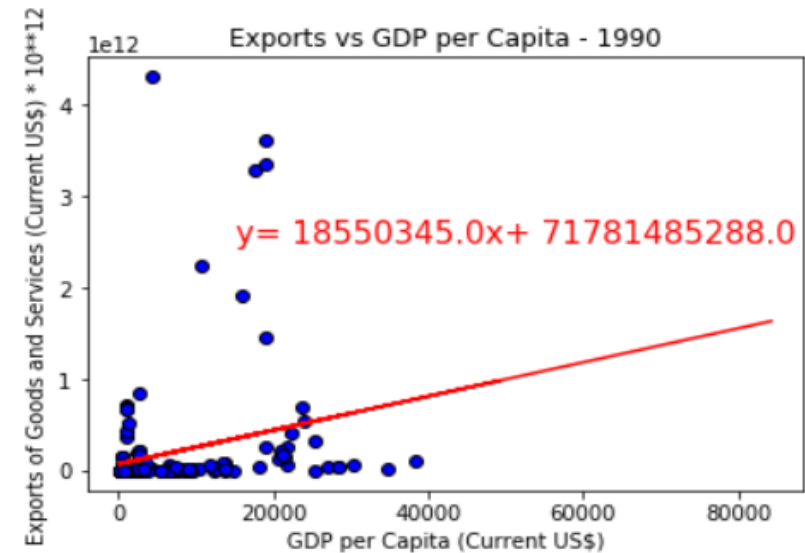
► In order to do that, we had to use the Add Suffix function to ensure that each set of columns could be properly identified.

► To enable merging by Country Code column, we then used the Rename function to drop the suffix for those.

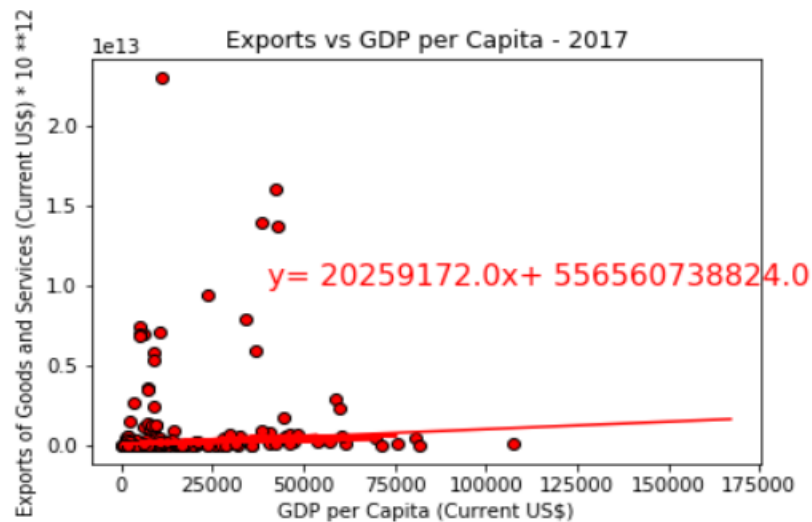
Data analysis process



The Rsquared value is 0.06.



The Rsquared value is 0.02.



Exports & Imports

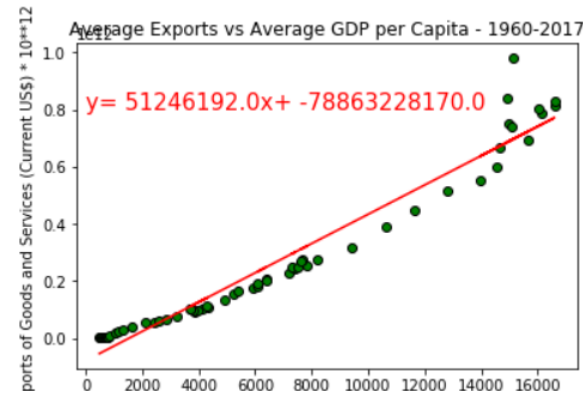
- ▶ We started by looking at scatter plots showing the relationship between the trade indicators and GDP per Capita for years 1990 & 2017.
- ▶ Linear regression couldn't initially be done, because there were NaN values in the data.
- ▶ The NumPy.isnan function was used to create a mask to filter out the NaN values to enable regression.
- ▶ The data showed that within any given year, the Exports didn't have a strong correlation with GDP per Capita, but this was because of outliers.
- ▶ The same held true for imports.
- ▶ Because the data was heavily skewed by outliers, a decision was made to further investigate this data.

Exports & imports (continued)

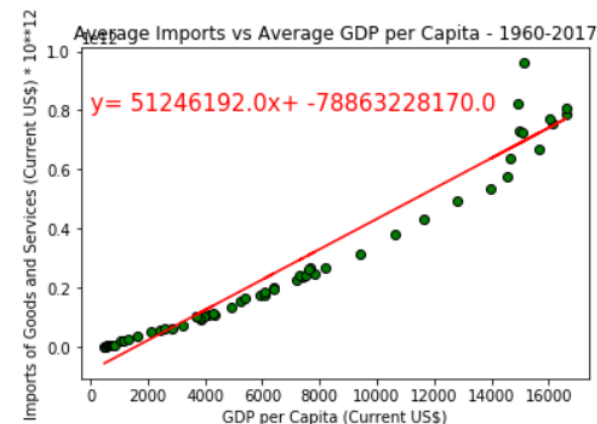
The Rsquared value is 1.0.



The Rsquared value is 0.95.



The Rsquared value is 0.95.



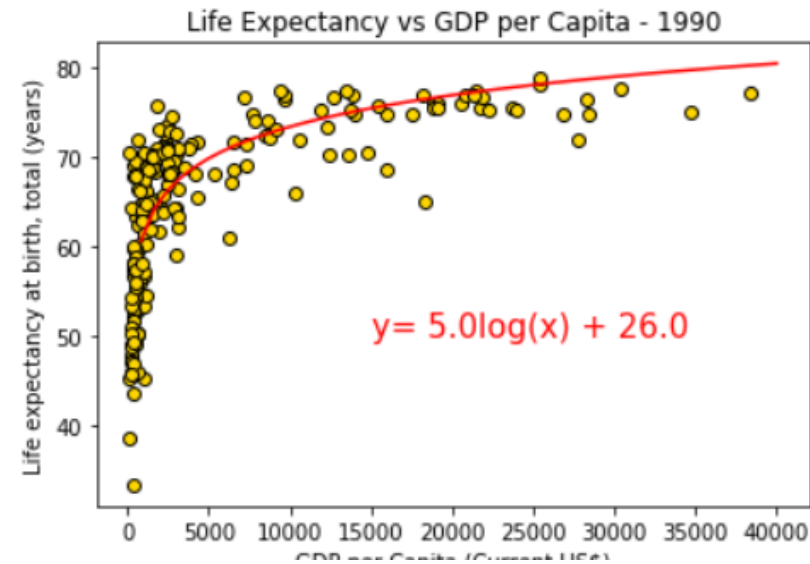
- ▶ To filter out the effects of the variation among countries, plots of average trade values vs average GDP per Capita were made.
- ▶ The results show that there is an extremely strong positive relationship between trade parameters and GDP per Capita
- ▶ As a means of validating the data from the World Bank, the average of Imports of all countries was compared to the average Exports of all countries for every available year.
- ▶ The data shows a virtually perfect linear correlation, indicating that the dollar value of each import was matched by the dollar amount of each export.
- ▶ This would tend to make the data more trustworthy.



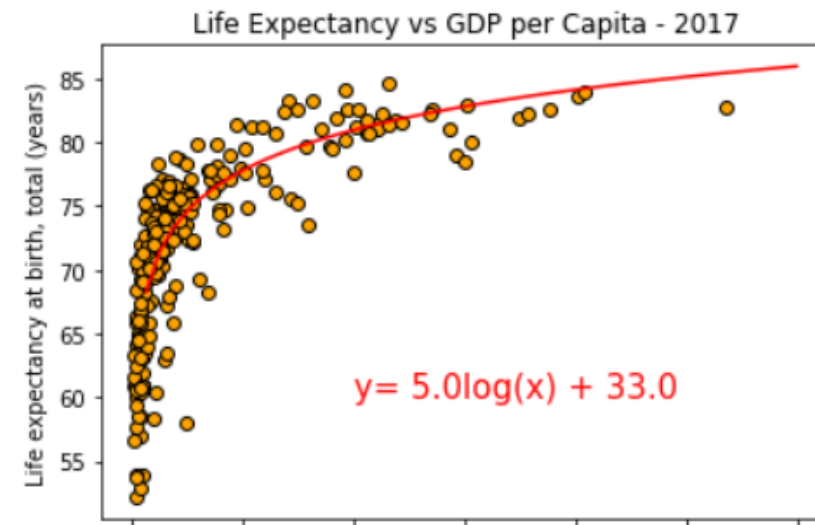
Data analysis
process – life
expectancy &
infant mortality

Life Expectancy VS GDP (1990 vs 2017)

- ▶ We started by looking at scatter plots showing the relationship between the life expectancy and GDP per Capita for years 1990 & 2017.
- ▶ The data showed that there wasn't a linear relationship between Life Expectancy and GDP per Capita.
- ▶ It has been shown that there is a strong logarithmic relationship between Life Expectancy and GDP per Capita.
- ▶ This means that Life Expectancy grows with GDP, but the rate of improvement slows as an economy advances.

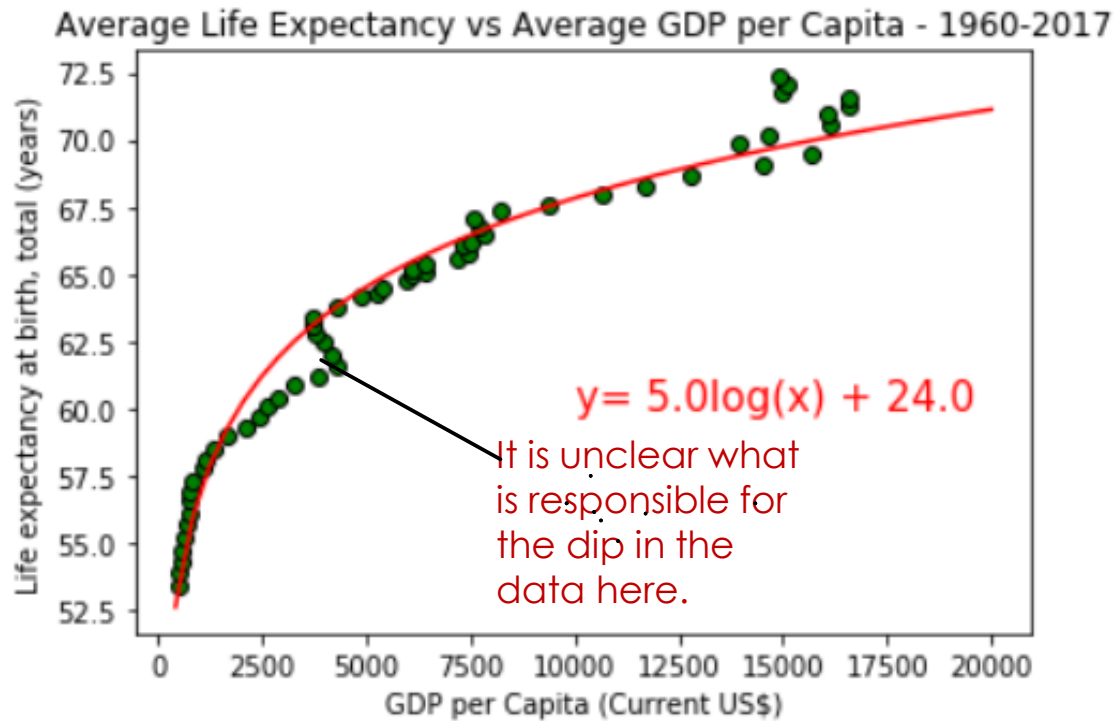


The Rsquared value is: 0.74.



Data analysis process – life expectancy (continued)

The Rsquared value is: 0.97.



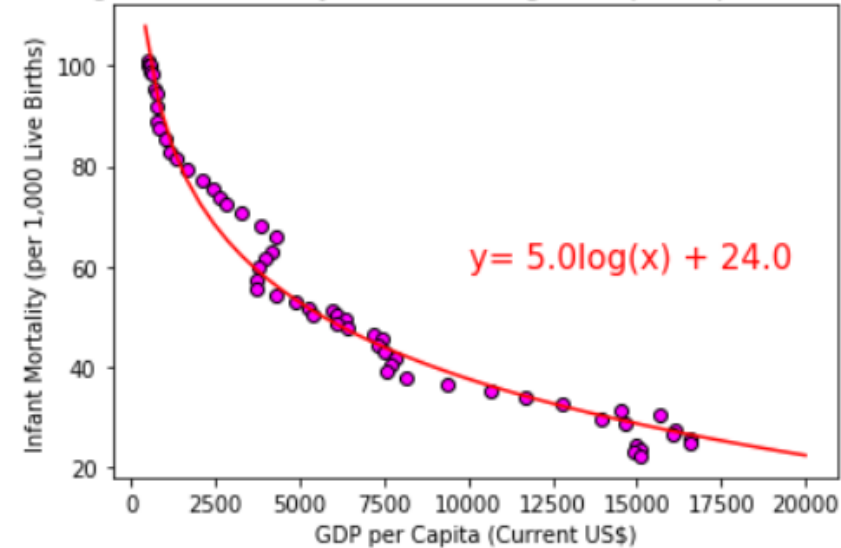
- ▶ To filter out the effects of the variation among countries, the plot of average Life Expectancy vs average GDP per Capita was made.
- ▶ The mean values of individual year columns were calculated and put in lists.
- ▶ The results show that there is an extremely strong positive relationship between life expectancy parameters and GDP per Capita.

Infant Mortality Rate (continued)

- ▶ We started by looking at scatter plots showing the relationship between infant mortality rate and GDP per Capita for years 1990 & 2017.
- ▶ The data showed that there wasn't a linear relationship between Infant Mortality Rate and GDP per Capita.
- ▶ It has been shown that there is a strong logarithmic relationship between Infant Mortality Rate and GDP per Capita.
- ▶ This means that Infant Mortality falls with GDP, but the rate of improvement slows as an economy advances.
- ▶ To filter out the effects of the variation among countries, the plot of average Infant Mortality Rate vs average GDP per Capita was made.
- ▶ The mean values of individual year columns were calculated and put in lists.
- ▶ The results show there is strong positive relationship between well-being and GDP per capita

The Rsquared value is: 0.98.

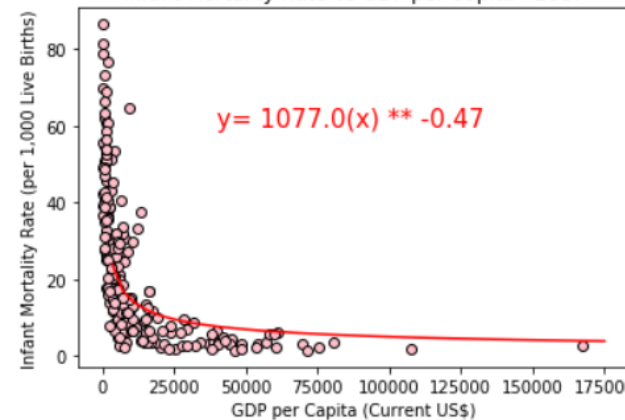
Average Infant Mortality Rate vs Average GDP per Capita - 1960-2017



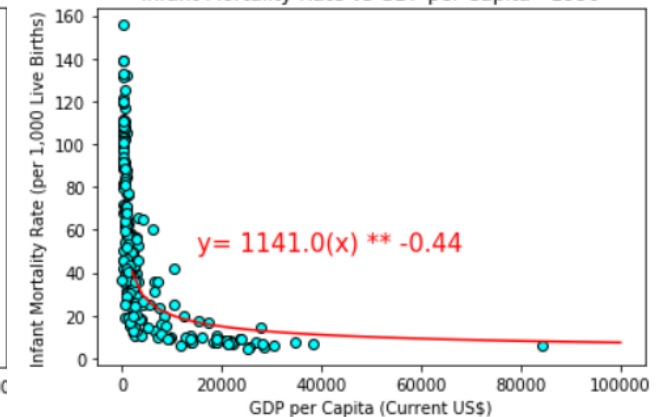
The Rsquared value is: 0.66.

The Rsquared value is: 0.73.

Infant Mortality Rate vs GDP per Capita - 2017



Infant Mortality Rate vs GDP per Capita - 1990

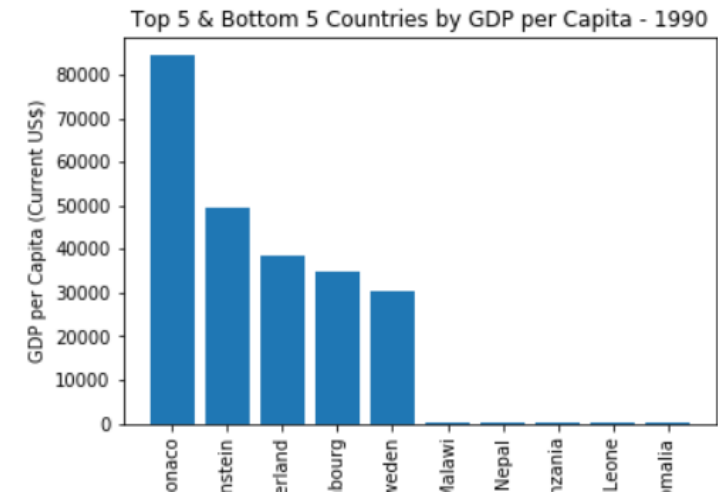
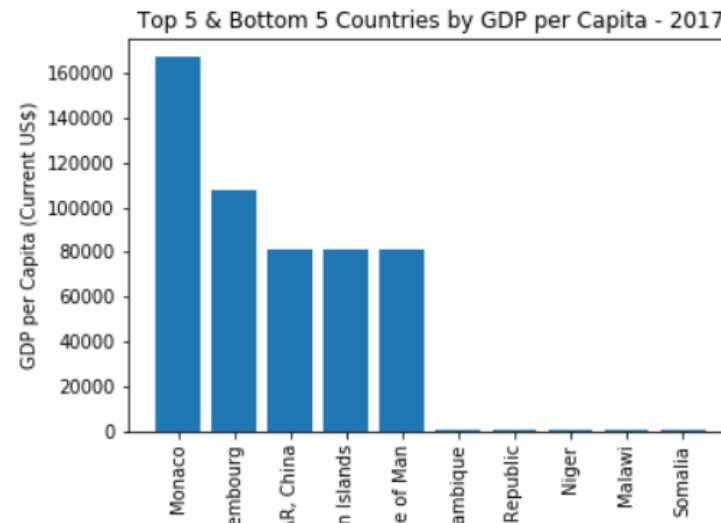




Predictions

GDP per Capita differences

- ▶ We wanted to see the level of differences between the top 5 & bottom 5 countries in a couple of years.
- ▶ It was also desired to see whether any of the bottom 5 countries in one year “escaped” the bottom 5 and whether it was because of sharp increases in trade and/or whether it led to improvements in well-being values.
- ▶ There is a several order of magnitude difference in GDP per Capita between the top 5 & bottom 5 countries.
- ▶ The countries of Nepal, Tanzania, and Sierra Leone “escaped” the bottom 5.

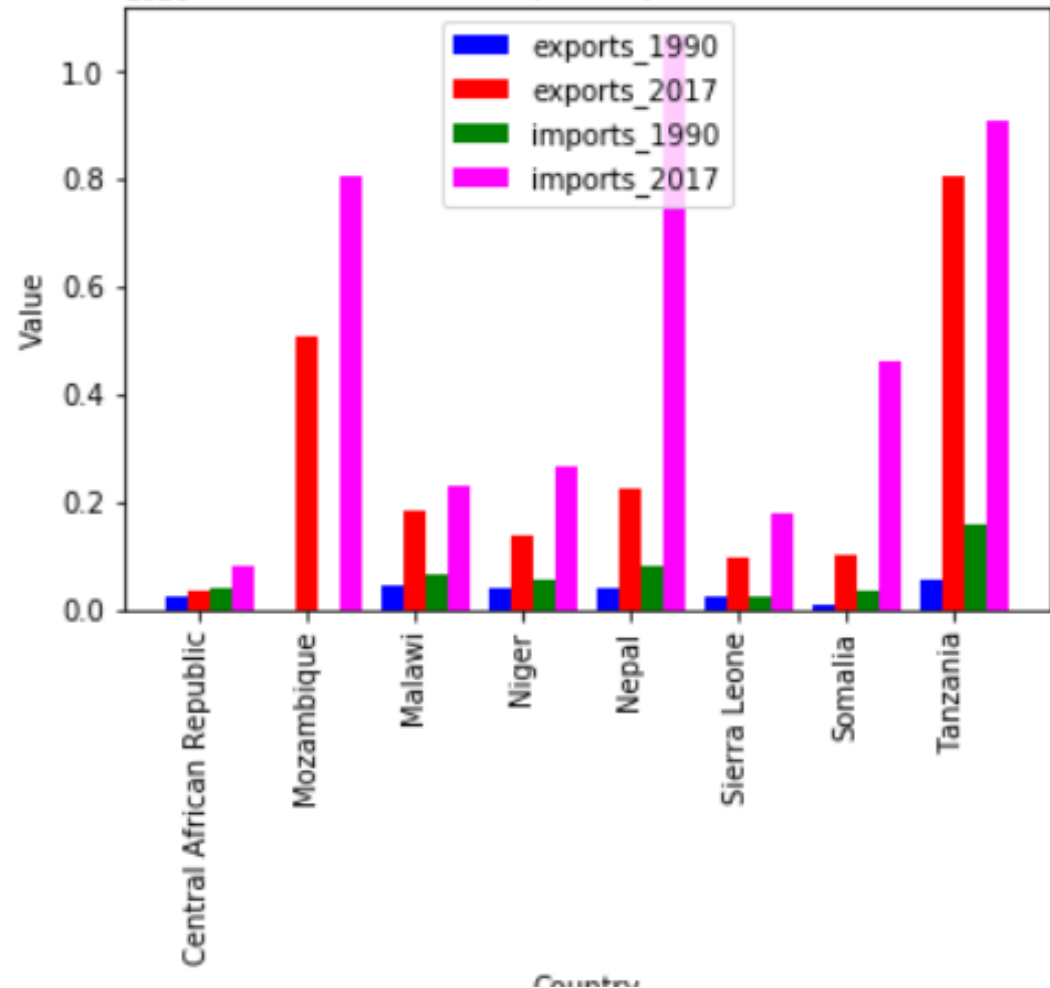


Trade Values of bottom GDP per Cap (1990-2017)

►Of the countries of Nepal, Tanzania, and Sierra Leone that “escaped” the bottom 5, Nepal & Tanzania had the largest increases in trade, but Sierra Leone didn’t.

►This would likely indicate that trade can improve the productivity of a country, but it is not a perfect predictor for individual countries between two separate years.

Trade Values of Bottom GDP per Capita Countries - 1990 & 2017

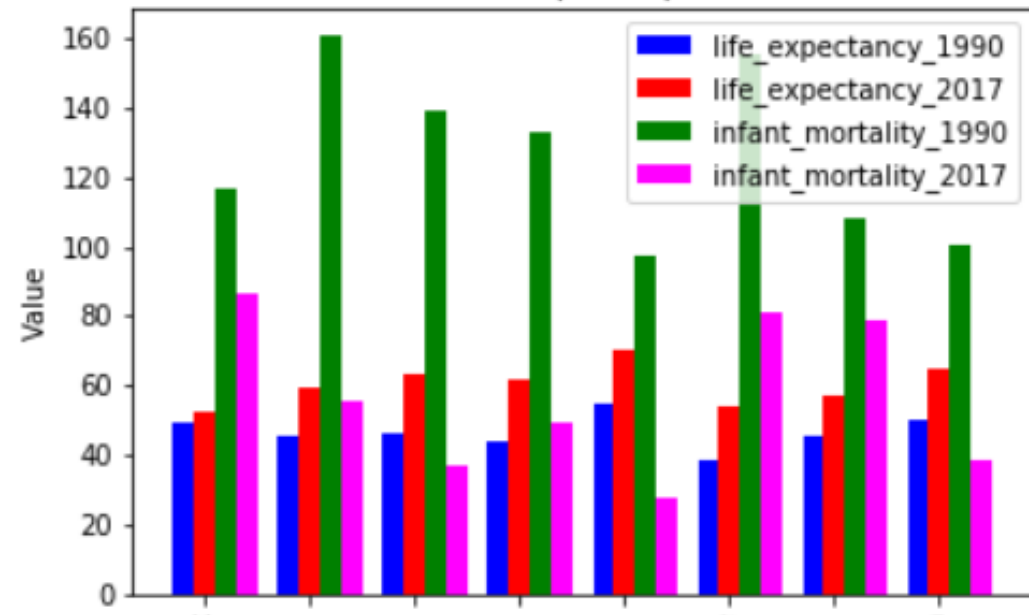



Wellness Values of bottom GDP per cap(1990-2017)

►Of the countries of Nepal, Tanzania, and Sierra Leone that “escaped” the bottom 5, all had relatively large improvements in life expectancy & infant mortality rates, but not always the largest.

►This would likely indicate that productivity of a country can improve well-being, but it is not a perfect predictor for individual countries between two separate years.

Wellness Values of Bottom GDP per Capita Countries - 1990 & 2017



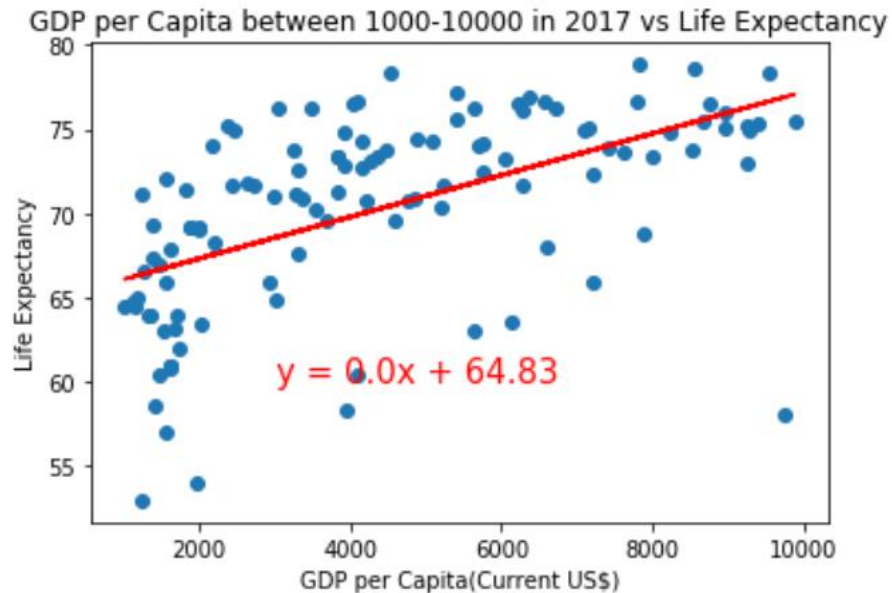
- 
1. At what point does GDP have the highest impact in life expectancy and infant mortality?
 2. At what point does GDP increase start to have less of an effect on wellbeing?
 3. If we had more time, we would want to look deeper into what other factors were in play regarding wellbeing, and test to see if there are nations that had an increase in GDP but not in wellbeing.

Questions after initial analysis.

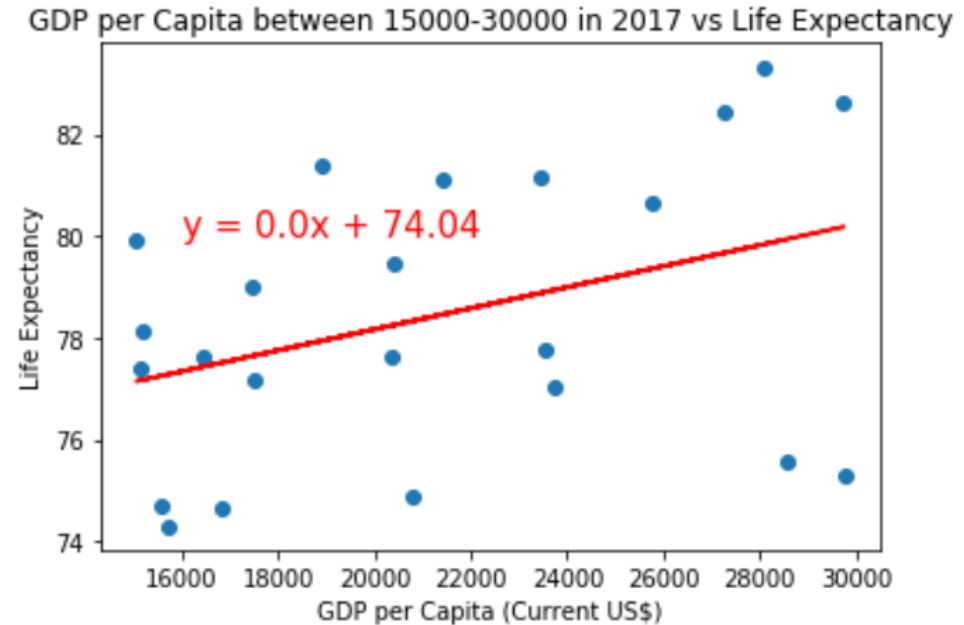
Life expectancy vs GDP (GDP between 2000 per cap – 10,000 per cap)

Life expectancy vs GDP (GDP between 16,000 per cap – 10,000 per cap)

The Rsquared value is 0.56.



The Rsquared value is 0.38.



While there is still a relationship with the rise of GDP and increase in life expectancy, after about 10,000 GDP per Cap, the relationship strength begins to decrease partly because there is only so high you can go with life expectancy, but the data shows it doesn't take a dramatic increase in productivity to improve well-being in poor countries .



EXPECTED

We initially expected exports to decrease as GDP increased, but the data indicates GDP and trade values move up or down together.



FOUND

Based on regression analysis the data comparing GDP and how it is impacted by trade values (import/export), we were able to conclude imports and exports have a direct relationship with GDP.



EXPECTED

We were correct in our initial hypothesis in saying, if GDP increases, then we can project life expectancy to increase and infant mortality rates to decrease based on analysis of regression.



FOUND

Keeping in mind that at a certain point GDP increase does not have the same effect on well being, we conclude, if a nation increases trade values then GDP will increase and subsequently wellbeing will likewise increase.

Conclusions