English Proficiency Correlation on Gross National Income Per Capita

Introduction

For decades, English has dominated the global arena as the language of education, business, pop culture, and the Internet, which is why considering English replaceable or unimportant couldn't be farther from the truth. The study, "A Ranking of 112 Countries and Regions by English Skills," conducted by *Education First*, shows that countries with higher English proficiency indices (EPI) enjoy more social equality, better job opportunities, and increased involvement in the environment (Education First). Imagine a day when the world can communicate in the same language and still enjoy traditions through their mother tongue. Then, there would be better communication among countries, resulting in fewer wars, more educated citizens, more advances in medicine, the end of world hunger, and a better quality of life.

For this project, I will use data visualization techniques to provide supporting evidence that high proficiency in the English language plays a role in a (non-native English speaking) country's wealth, specifically in their gross national income per capita (GNI PPP), which economists use as an indicator of country's social, economic, and environmental health. The findings presented in this paper can be used by governments, social scientists, and departments of education worldwide when planning a country's well-being in the long term through accessible, quality programs of English as a second language.

Video Presentation Link: https://youtu.be/ZuoRlo6tems

Dataset and Data Abstraction (The What)

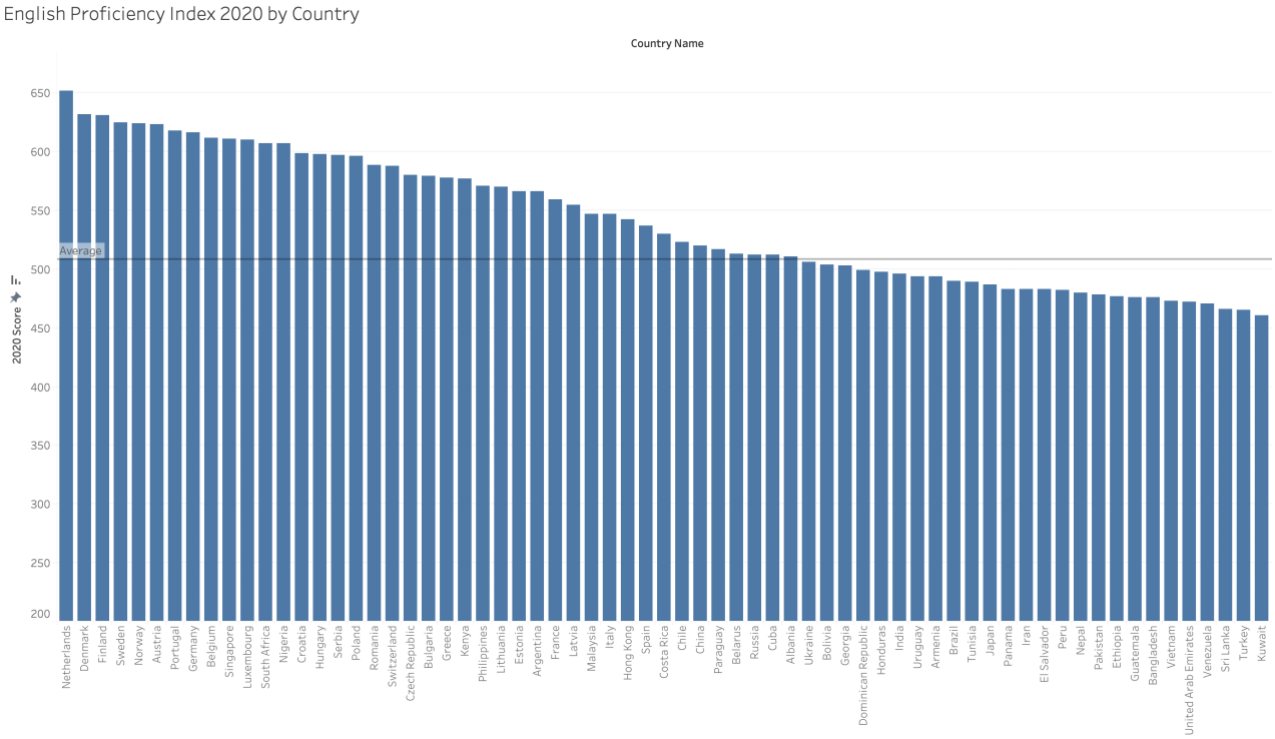
- The first dataset is, "GNI per capita, Atlas method (current US$)" published by *The World Bank* on 10/28/2021 (GNI PPP) (The World Bank)
- The original format is an Excel spreadsheet and contains tables with rows and columns
- Data type: Items (rows) and attributes (columns)
- Data and dataset type: Tables with items and attributes (country name or year)
- Attribute type: Ordered (dollar amounts highest to lowest, and country names A-Z)
- The second dataset is, "A Ranking of 112 Countries and Regions by English Skills" published by *Education First* in 2021 (EPI) (Education First)
- The original format is in PDF, which I converted into an Excel spreadsheet to compare with the first dataset
- Data type: Items (rows) and attributes (columns)
- Data and dataset type: Tables with items and attributes (country name or EPI score)
- Attribute type: Ordered (EPI scores highest to lowest, and country names A-Z)
- Both datasets have been uploaded to Tableau to produce bar graphs
- Because the GNI PPP dataset was more extensive than EPI, I used the information visualization rule of thumb of keeping data as simple as possible without losing its essence, and only considered the years 2011-2021 to come up with a 10-year average GNI PPP
- Due to EPI showing information only 112 countries and GNI PPP on 267 countries, I used Excel's advanced filter to reduce the number of countries to 112
- Understandably, the number of countries ended up being around 90 due to differences in the naming system of both datasets by the companies, but small enough to be negligible

Domain Tasks and Task Abstraction (The Why)

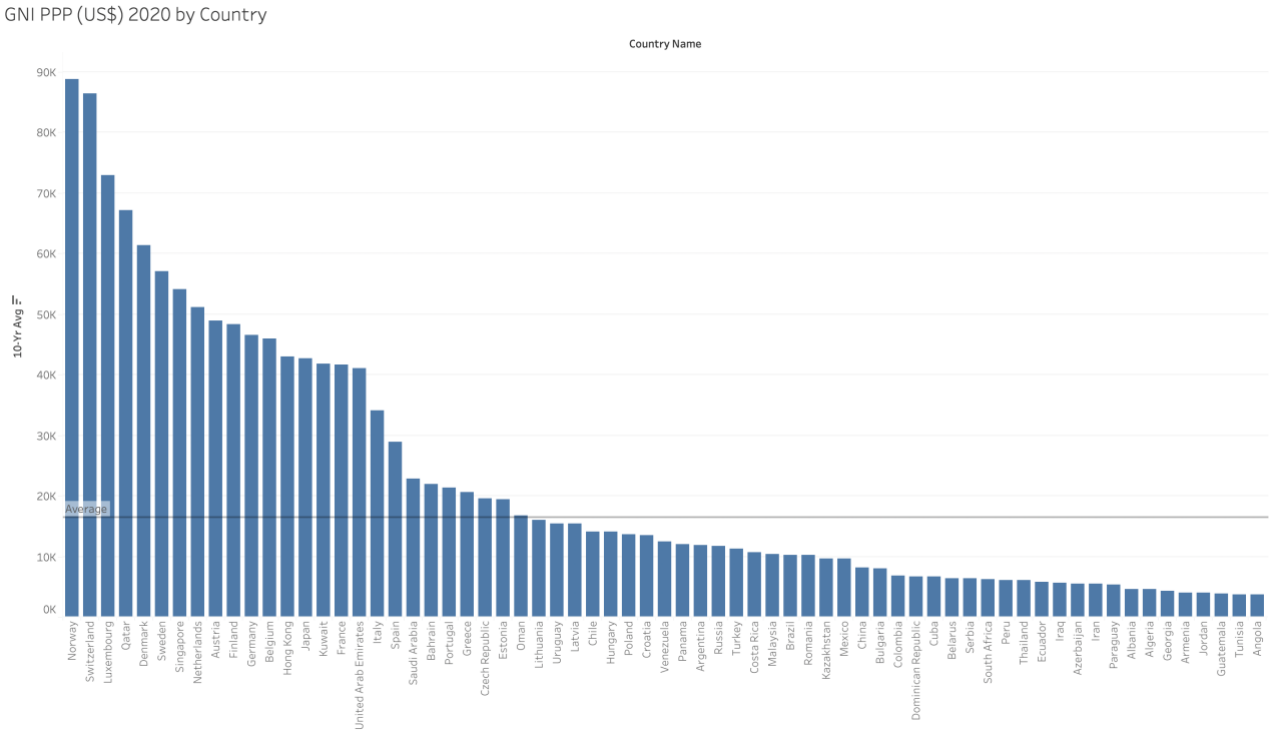
- Analyze: Consume: Discover and present
- Analyze: Produce: Record and derive
- Designed for governments, social scientists, and departments of education
- Provides statistical analysis and comparison between GNI PPP 2020 and EPI 2020
- Query: Identify, compare, and summarize datasets
- All data: Trends
- Attributes: One: Distribution
- Attributes: Many: Correlation and similarity
- Users will be able to see an average line drawn on both bar graphs (EPI by Countries, and GNI PPP by countries) that separates countries in above and below average
- We are interested in the countries that lie above the average curve
- The datasets have been organized to show only countries that both lists have in common
- Countries not in common, whose names were spelled differently, or whose data was unavailable were removed from the datasets to prevent errors
- Users will be able to compare two distinct datasets and quickly see how they relate
- Curated datasets enrich statistical analysis in communication stages

Visual Representations and Interactions (The How)

- Encode: Arrange: Align values in descending order
- Encode: Map: Color: Hue solid blue against white background
- Manipulate: Select only greater-than-average values by tracing an average line
- Reduce: Filter only the country names in common between both datasets
- Reduce: Aggregate the above-average EPI and GNI PPP values into a master dataset
- From the master dataset, take consider only the countries that share both a greater-than-average GNI PPP and EPI
- Whether a higher GNI PPP results in a higher EPI, or vice versa, is a matter of debate
- I would like to hypothesize that countries with a high GNI PPP have disposable income that can be invested in English-as-a-second language programs, but the results of education could take years to see
- A reason to believe that a high GNI PPP precedes a high is that countries with scarce resources may decide to tackle other social problems like health, infrastructure, and basic education, such as native language, math, and science, rather than invest them in English education, if countries can even invest in education at all
- An additional reason to believe that GNI PPP precedes EPI is schools provide certificates of completion, while self-learning does not always



EF English Proficiency Index 2020



The World Bank 2020

The first bar graph above shows the English Proficiency Index provided by Education First of countries where the English language exam was conducted. According to EF, for countries to appear on the dataset, a minimum number of participants was required. Similarly, the second bar graph shows a 10-year average of the gross national income (in US dollars) as updated by The World Bank on 10/28/2021. Both lists have been filtered to contain the names of the countries that both have in common.

The following charts have been generated using only the above-average EPI and GNI PPP values from both bar graphs above. On the table to the left, the average EPI score is 508.07, whereas, on the table to the right, the average GNI PPP value is $16,483. Moreover, countries that do not belong to both lists have been intentionally greyed-out, and 18 out of the 20 countries on the chart to the right have above-average GNI PPP as well as an above-average EPI.

Country Name	2020 Score	2020 Proficiency Band
Netherlands	652	Very High Proficiency
Denmark	632	Very High Proficiency
Finland	631	Very High Proficiency
Sweden	625	Very High Proficiency
Norway	624	Very High Proficiency
Austria	623	Very High Proficiency
Portugal	618	Very High Proficiency
Germany	616	Very High Proficiency
Belgium	612	Very High Proficiency
Singapore	611	Very High Proficiency
Luxembourg	610	Very High Proficiency
Nigeria	607	Very High Proficiency
South Africa	607	Very High Proficiency
Serbia	597	High Proficiency
Romania	589	High Proficiency
Switzerland	588	High Proficiency
Greece	578	High Proficiency
Kenya	577	High Proficiency
Philippines	571	High Proficiency
Lithuania	570	High Proficiency
Estonia	566	High Proficiency
France	559	High Proficiency
Latvia	555	High Proficiency
Italy	547	Moderate Proficiency
Malaysia	547	Moderate Proficiency
Hong Kong	542	Moderate Proficiency
Spain	537	Moderate Proficiency

EF English Proficiency Index 2020

Country Name	10-YR AVG
Norway	$ 88,805.00
Switzerland	$ 86,389.00
Luxembourg	$ 72,937.00
Qatar	$ 67,124.00
Denmark	$ 61,350.00
Sweden	$ 57,138.00
Singapore	$ 54,178.00
Netherlands	$ 51,167.00
Austria	$ 48,884.00
Finland	$ 48,321.00
Germany	$ 46,480.00
Belgium	$ 45,974.00
Hong Kong	$ 43,055.00
France	$ 41,600.00
Italy	$ 34,165.00
Spain	$ 28,881.00
Portugal	$ 21,402.00
Greece	$ 20,554.00
Czech Republic	$ 19,630.00
Estonia	$ 19,394.00

The World Bank 2020

Results

It is fascinating to see first-hand how the English language has taken over world communication—from education to movies, from business to social media. It is no wonder why developing countries see English proficiency as an opportunity to increase their wealth and well-being. Intrigued by this observation, I searched for evidence on the relationship between high gross national income per capita (GNI PPP) and high English proficiency indices (EPI). I was pleasantly surprised to find two datasets backed by renowned organizations, such as The World Bank for GNI PPP, and Education First for EPI. Afterward, I used Excel's advanced filter to leave only the countries in common in both datasets and created a 10-year average out of GNI PPP (leaving out English speaking countries), imported into Tableau where I organized values from high to low and trace an average line, took only the values above the average line, and created comparison lists.

The findings were more than I expected. First, 18 out of 20 countries with an above-average EPI also enjoyed an above-average GNI PPP, except for Qatar (a relatively new country) and The Czech Republic. Second, around half of the countries that share high GNI PPP and EPI also speak a Germanic language, such as Dutch, German, and Swedish. Third, more people seem to be learning English to find better education and job opportunities, but reaching proficiency takes time. Fourth, Nigeria and South Africa ranked high on the EPI list, but not on the GNI PPP list, even though both countries consider English a native language. Fifth, the country with the highest GNI PPP (the Netherlands) was not the one with the highest GNI PPP (Norway), which means that this relationship is not perfectly linear, and this could be due to productivity, fiscal policies, education investment, among other factors. In conclusion, these findings support that higher proficiency of English is a good indicator of a better quality of life.

Works Cited

Education First. “A Ranking of 112 Countries and Regions by English” EF English Proficiency Index, https://www.ef.com/wwen/epi/. Accessed November 22, 2021.

Munzner, Tamara. *Visualization Analysis and Design (AK Peters Visualization Series).* 1st. Boca Raton, London, New York: A K Peters/CRC Press, 2014.

The World Bank. “GNI per capita, Atlas method (current US$)” World Development Indicators, The World Bank Group, 10-28-2021, https://data.worldbank.org/indicator/NY.GNP.PCAP.CD. Accessed November 22, 2021.