COUNTRY RECOMMENDATIONS TOOLS

André Rocha
Sapienza
`magalhaesrocha.1991375`

Hanna Solstad
Sapienza
`solstad.1991023`

Tito Griné
Sapienza
`trindadegrine.1992095`

ABSTRACT

A connected world provides a wide dimension of data that can be analyzed. Comparison between countries is a common exercise, whether it is to discuss where to emigrate, or it is to discuss what could be improved in a given country. While analyzing a single dataset is an exercise simply done with the tools today available, it's not simple to combine these datasets. Our tool provides a way of analyzing data from similar categories combined and compare it with data relating to categories, allowing to take conclusions on combined areas such as education and democracy, or safety and happiness.

1. INTRODUCTION

As the world grows more connected over time, the amount of data available about countries has increased significantly. From data related to education, to climate, and even gender rights, the huge amount of data can definitely be overwhelming. It is natural for it to be nearly impossible for any citizen to be able to compare these, specially over a wide number of countries.

With this tool, we intend to provide a means for data to become more easily comparable between countries, whether the user's goal is to decide which country he would consider living in, or even if they're just making an analysis on countries over some data points of some categories.

It is recognized the importance of defining some guidelines regarding the goals of the visual tool:

- the number of of datasets on countries is overwhelmingly big, a way of combining them must be found, even if precision of data will suffer from it;

- the range of possible users is wide, so the tool must be generic enough that it can tailor to most user's relevant needs;

- the tool must be able to provide insights without much latency delay, since users are compelled to constantly change the data they are considering;

- the tool must provide different means of visualization that help them understand the data, query it and take conclusions from it.

2. DATA

Given the field of country recommendations, it is recognized the importance of combining several different datasets on countries. This leads to the problem of handling how to aggregate these various datasets.

One of the biggest sources of data of countries is *Our World In Data* [1], containing, at the time of writing, over 3000 different datasets on countries, organized into almost 300 different topics. Considering the wealth of information presented in *Our World In Data*, it was undeniably one of the best sources of data for this sort of problem. While there's heterogeneity as to how the data in *Our World In Data* is presented, there are some common fields observable in all datasets:

- **Entity** - the name of the country was always present in the dataset. All datasets in *Our World In Data* are presented with the exact same name, something of importance since there are multiple countries with different names (e.g. Czech Republic and Czechia). This field is usually a *string*;

- **Code** - Countries have several different unique identifiers associated, defined by ISO 3166 [2]. *Our World In Data* makes use of their alphabetical code of 3 letters, usually denominated of Alpha-3 Code (e.g. Italy's code is *ITA*). This field is usually a *string*

- **"Fields"** - All datasets will then possess the fields that are of importance to them. Those do not have generalized names, like *Entity* or *Code*, but are related to their field in particular. The way these fields are represented is also not generic. Fields may just be a *string* field, usually representing a *number* or may be a list of objects, once again depending on their particular use case.

Listing 1. *Our World In Data* JSON Example. The current dataset represents a reference to the number of Adolescent women birth rates

```
[
    ( . . . )
        {
                "Entity": "Australia",
                "Code": "AUS",
                "1st (Incomes across
                    the Distribution
                    Database (2016))":
                "9283.941406"
```

```
            }
    ( . . . )
]
```

While the number of datasets from *Our World In Data* can cover a wide range of categories of importance, there are still some topics of importance that were not covered by it:

- **Language** - it is important, specially when deciding to travel to a given country, to understand how would it possible to communicate to the population of said country. A dataset on the percentage of English speakers of each country has been made available through Kaggle, which uses data from Wikipedia [3]. No data for other was found.

- **LGBTQI+ Rights** - considering how LGBTQI+ rights can vary substantially between countries, it was of importance to find some data on this matter. Serwatka [4] has compiled information regarding sexual orientation and gender identity laws for all countries.

Listing 2. *Kaggle* data JSON Example. The percentage of english speakers ends up being the field of most importance

```
[
        {
                "Country": "United␣
                    States",
                "Eligible␣population":
                    "296603003",
                "Total␣English␣
                    speakers": "
                    283160411",
                "Total␣English␣
                    speakers␣(%)": "
                    95.46",
                "As␣first␣language": "
                    234171556",
                "As␣first␣language␣(%)
                    ": "79",
                "As␣an␣additional␣
                    language[note␣1]":
                    "48988855",
                "As␣an␣additional␣
                    language␣(%)": "
                    16.5",
                "Comments": "(...)",
                "Q3_iso_alpha": "USA"
        }
]
```

Listing 3. *Serwatka's* data JSON Example. Several fields on laws are presented

```
[
{
```

```
        "country": "Zimbabwe",
        "HAPPINESS": "3.692",
        "Per␣capita␣GDP": "$2
            ,147.00",
        "%␣with␣tertiary␣educ.
            ": "0.38",
        "GRI": "2.8",
        "Religiosity": "88%",
        "SHI": "2.1",
        "legality␣of␣sam−sex␣
            sexual−activity": "
            0.25",
        "marriage/␣civil␣
            unions": "0",
        "same−sex␣
            coupleadoptions": "
            0",
        "serving␣in␣the␣
            military": "0",
        "antidiscrimination␣
            laws": "0",
        "gender␣ID␣markers": "
            0",
        "Democracy␣Status": "A
            ",
        "electoral␣process␣and
            ␣pluralism": "0.5",
        "political␣culture": "
            2.5",
        "civil␣liberties": "
            3.24",
        "Satisfaction␣with␣GDP
            ␣per␣capita": "
            1.151",
        "Social␣support": "
            1.479",
        "Healthy␣life␣
            expectancy": "0.599
            ",
        "Freedom␣to␣make␣life␣
            choices": "0.399",
        "Generosity": "0.065",
        "Perceptions␣of␣
            corruption": "0.025
            "
        }
]
```

Taking into account this information regarding all datasets, we can infer the dimensions of our data:

- 195 countries in the world;

- 92 datasets are being used;

- Each of these datasets have, at the very least, one field that is being taken into account

- Minimum Dimensionality $= 92 \times 195 = 17940$

All of these datasets can be grouped into several different categories for better browsing:

- **Air Pollution** - Detailing information on several important factors related to indoor and outdoor pollution

- **Natural Disaster** - Describing numbers on several natural disasters, such as earthquakes and volcanic eruptions

- **Climate Change** - Exhibiting countries' contributions to climate change and against it

- **Food** - Detailing the quality of countries' nutrition as well as the cost of this in populations' expenses

- **Income** - Explaining the economical status of countries as well as their capability of purchase

- **Education** - Reporting the quality of a country's education in the several different levels of education

- **Political** - Agnostically describing the quality of country's democracies and their political participation

- **Gender Rights** - Explaining how fair are countries when it comes to gender rights, from an economical perspective but also from an opportunities' perspective.

- **Language** - Describing the percentage of English speakers for each country

- **Safety** - Offering data on crime, bully, terrorism and war rates

- **Happiness** - Comprehending the happiness of a population is a non-trivial matter, tho rates of self satisfaction, suicide and depression are taken into account

- **LGTBQI+ Rights** - Detailing countries' laws supporting different sexual orientations and gender identities

- **Health** - Describing the quality of health services, as well as their cost to citizens

3. ANALYTICS

Reducing the number of dimensions of the data is key to ensuring a proper visualization is given to the user. Some of these datasets might also be dependent on each other (e.g. the share of a country's electricity produced by fossil fuels is related to their share of electricity produced by renewable energies). Understanding how these dimensions can be ignored, was an important step. Furthermore, our ideal visualization would provide users with a single normalized scored for each category. How could such a score be produced was important to detail.

The technique used in this visualization for solving these problems was based on dimensionality reductions techniques. All algorithms that make use of Singular-Value-Decomposition (SVD) [5] will detail a matrix of the variance of each field. This variance matrix can be used to understand which fields are of most importance.

For this particular case, the Principal Component Analysis algorithm (PCA) [6] was used, since it was able to provide variance matrix of the data, making it possible to understand the importance of each field. Given this, it is possible to make use of the variance matrix provided to compute a single value out of all fields. This value will optimize the "importance" of each field, while allowing them to be portrayed into a single dimension.

With a single dimension, visualization for the user of the quality of the data will be much simpler, even if some depth of the data may be lost.

4. VISUAL ENVIRONMENT

The environment is composed of 3 main sections in a total of 5 distinct panels. These sections are 4.1 Data Selection, 4.2 Data Visualization, and 4.3 World Map View.

Figure 1. Country Recommender application

4.1 Data Selection

The Data Selection section encompasses both the Category and Parameter selection panel and, as the name suggests, is where the user can select which metrics are considered in scoring the countries.
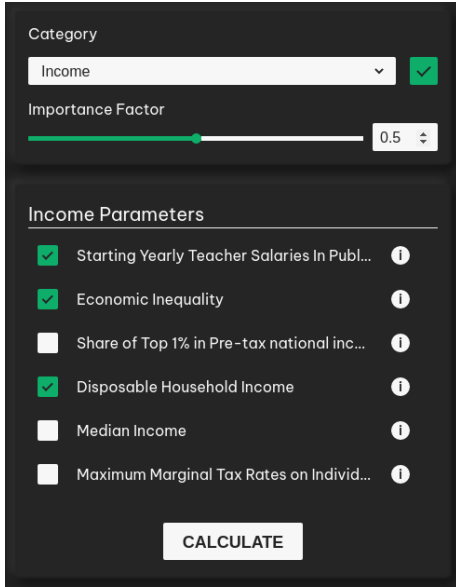
Figure 2. Data Selection section

The usual flow begins by selecting a category from a dropdown, which consequently updates the Parameter Panel to display the parameters that exist for that category. After a category is selected, the user must explicitly click the checkbox to the right of the dropdown in order to begin interacting with the rest of the selection. Once checked, the user can change the importance factor attributed to that category, either by using the slider or the number range input below the dropdown, to a value ranging from 0 to 1.

This importance factor changes the weight of the score of that category in the calculation of the overall country score. The overall score of a given country is obtained by the following equation:

$$overallScore = \sum_{c}^{C} s_c \times \frac{f_c}{F} \qquad (1)$$

Where C is all the categories selected by the user, s_c is the country's score for category c, f_c if the importance factor the user gave to category c and F is the sum of all importance factors.

As an example, if only one category has been selected, its weight in the overall score is all 1, regardless of the importance factor attributed to it. If two categories are selected, one with an importance factor of 0.25 and the other 0.5, their resulting weight in the overall score will be 33.(3)% and 66.(6)% respectively, or in other words, the score of the second category is two times more significant than the score of the first category.

As mentioned, each category will have its own list of parameters for the user to choose from. Each parameter roughly correlates with data taken from a dataset. The user can select any combination of the existing parameters by clicking their respective checkmark. Each parameter also has an info icon to the right of its name which the user can hover in order to toggle a tooltip where further information

regarding that parameter is displayed, like its full name and the fields/attributes that are apart of it.

For example, if the user selects the category **Climate Change**, the following parameters will be displayed: *"Share Electricity From Renewables"*, *"Fossil Fuel Production"*, *"Fossil Fuel Primary Energy"*, *"Share Electricity From Fossil Fuels"*. When hovering the information icon for the **Fossil Fuel Production** parameter, the tooltip shows that it is comprised of the fields *"Coal Production"*, *"Oil Production"* and *"Gas Production"* all given in TWh.

To complete the flow, as long as one of the parameters from the selected category are selected, the user can click the **Calculate** button in order to compute the PCA matrix of the scores of each country, for that category, using the selected parameters.

If the selected category is already being used for calculating scores, then clicking the Calculate button will replace the previously used PCA matrix for the new one. Alternatively, if the user untoggles the checkbox for that category, the **Remove** button can be used to remove the currently used PCA matrix for that category from the score calculation, effectively removing the category from being considered for scoring.

4.2 Data Visualization

Once at least one PCA matrix of a category has been calculated, the two panels that make up the Data Visualization section will update to show a Parallel Coordinates and a Histogram graph representing the data.
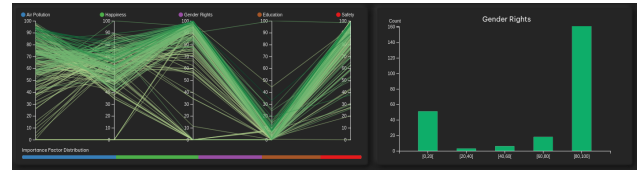

Figure 3. Data Visualization section

Since the user can use multiple categories for calculating the score, we decided that the most efficient and effective way to visualize the overall data would be through a Parallel Coordinates graph. This type of graph allows for a consistent visualization experience while categories are added or removed, and still be easy to identify the distribution of the around 240 countries among the categories. Each country is represented by a line, where the color varies depending on the percentile on which that country is included. That is, if the country is in the top 1 percentile of all countries in its overall score, its color will be a light green, more specifically *#c2e699*. If on the other hand it is only within the top 50 percentile, its color will be a darker green, namely *#31a354*. This aids the user in quickly getting a field for the overall score distribution. Additionally, each axis can be used to select a range of scores for that axis' category. This means that only the countries with a score within that range will be considered. To make this

more apparent in the Parallel Coordinates graph, the lines corresponding to countries that do not fall into this score range will have its opacity significantly reduced. This results in the countries that are select to stand out while still preserving the overall distribution of the scores. Finally, clicking on the legend of a category's axis updates the category considered for the histogram panel.

The Histogram Panel displays a histogram of the score distribution of a particular category within ranges of 20. That is, a histogram where the bar height signifies the number (count) of countries whose score for the selected category falls within that range, where the ranges are [0, 20[, [20, 40[, [40, 60[, [60, 80[, [80, 100]. The user can click one or more of these histogram bars in order to select only the countries belonging to those ranges, which will override any range selection of that category that may have existed in the Parallel Coordinates graph.

In this section, below the Parallel Coordinates graph, one can also find a bar representing the importance factor distribution of all the currently selected categories. As explained in the section 4.1, the user selects for each category an importance factor meant to convey the importance of that category in calculating the overall score. However, the correlation between the value of a category's importance factor and its corresponding weight is not immediately apparent. The fact that a category has an importance factor of 0.5, doesn't mean its score contributes 50% to the overall score. As such, in an effort to make this correlation more easily understood, an importance factor distribution bar was added. Each category has a given color, which can be viewed as the color of the circle next to the legend of each category's axis in the Parallel Coordinates graph. The distribution bar is colored with each category, in the same order as the axis on the graph, where the proportion of the width matches the weight of that category in the overall score. As an example, if a category has an importance factor that makes it so its score contributes to half of the overall score, the distribution bar will have 50% of it colored with that category's color. Since the width (length) is being used to perceive the differences between the factors of each category, ideally they would be aligned to one side for an easy perception of the differences. Since space is limited, we decided to not use this approach, but instead order the category's in the bar in ascending order given their importance factor. Although this still makes it difficult to perceive actual minute differences, it makes it immediate to compare two categories in terms of which has a larger proportion. Besides this, the user can also access the actual weight of each category, by hovering a category's color on the distribution bar where a tooltip will show its exact percentage.

It is important to mention that when the user is changing the importance factor of a category on the Selection Section, even if the category has not yet been added, the distribution bar is updated at the same time to reflect this change and thus allow the user to get an immediate feel for how the current importance factor will reflect the category's weight in the overall score.

4.3 World Map View

The World Map View Section is made up of a panel with a world map and a few legends that contextualize it.
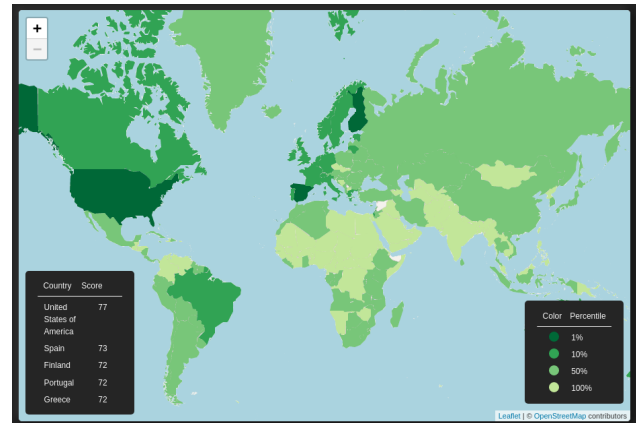

Figure 4. World Map view

The map starts of all countries having the same blue color, signifying that there is no scoring. Once at least one PCA matrix has been calculated and introduced to calculate scores, each country will have a color related to the percentile it is included in, exactly the same logic as per the line colors in the Parallel Coordinates graph. In order for the user to understand the meaning behind these colors, a legend is present at the bottom right corner of the map. On the opposite side, at the bottom left corner, a small table with the top 5 countries and the corresponding scores are shown. If the user hovers on a country, this table is updated to show the breakdown of that country's scores instead, depending on the currently used categories.

Although the user can't select a range of countries using a map, if there is some filter either from the Parallel Coordinates or the Histogram graphs, the selected countries will stand out on the map using the same approach as in the Parallel Coordinates, that is, the selected countries will keep their color with full opacity, while the rest have their opacity significantly reduced, but keeping the color matching its percentile. Additionally, the table showing the countries with the top scores will only include the countries that are part of the current selection.

5. INSIGHTS

In order to prove the utility of the tool for certain scenarios, this section will provide detail into some insights that can be concluded from this tool. The main goal of this section, given how ambitious the analytical techniques are, will be to display insights that should be commonly known to the reader, either through intuition or known statistics.

- A high quality democracy does not necessarily result in a high quality education system - while this insight can be intuitively agreed on, it is important to find backing on data for it. Several countries who

don't have a democratic system, or what is classified as a flawed democracy (where the elected party has been the same one for a long number of years), that don't provide any sorts of freedom of press or protest, can happen to have some of the best education systems in the world. The data that is taken into account to consider a country's education good is the funding (per capita) that is applied to making that country, the number of teachers (per capita) and their PISA scores.
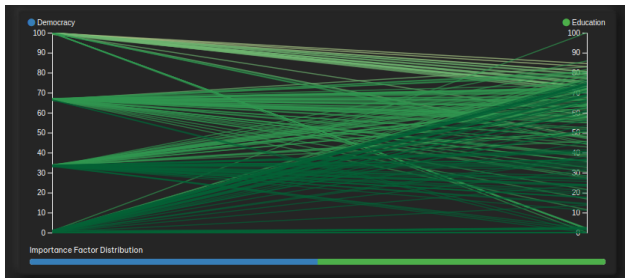

Figure 5. Parallel Coordinates for the comparison of Democracy with Education.

- Safer countries are usually the happiest, but higher income countries are not necessarily so - many articles have already been published on how some of the best countries economically speaking end up not having the happiest population [7], while intuitively it is simple to comprehend how safety can allow their population to feel more happy. It is important to mention how the measurements of safety are taken into account information such as bullying rates.
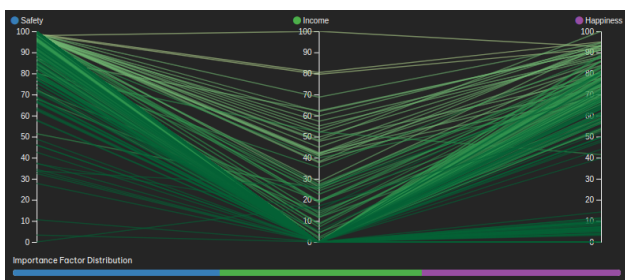

Figure 6. Parallel Coordinates for the comparison of Safety, Income and Happiness.

- Western Europe provides a much higher quality of life, generally speaking - if taking into account all of the datasets available through the application, it can be observed how most of Western Europe (with the exception of Portugal) happen to be in the top 10% of the countries in the world. This supports the claims commonly heard throughout the European Union of the sovereignty of western countries over eastern countries of the European Union. It is also important to state that the two countries in the

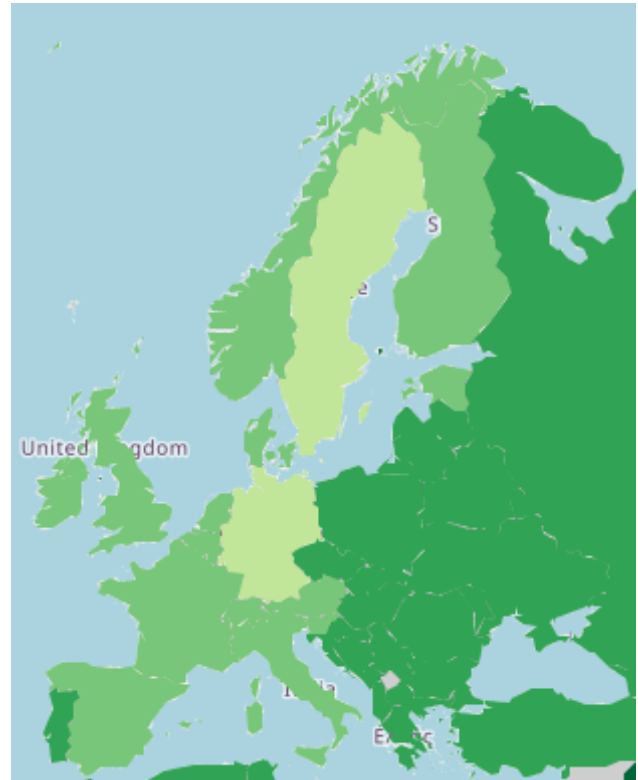top 1% are part of the European Union: Germany and Sweden.

Figure 7. Map of the European Union when considering all datasets in all categories available in the app.

6. RELATED WORKS

In order to understand the quality of this work, it is important to compare it to other works in the field of comparing data between countries.

One of the biggest sources of data for this work also provides its own visualization, *Our World In Data* [1]. Our World in Data provides their information in various visualizations: world maps, line charts and tables are some of the most common ones. Their goal is to allow an overview of a given dataset, whether the user's intentions is to have a general overview of every countries or to search for the specific value of a collection of countries. The difference between their charts to this work is defined by their lack of dimensionality: all of their charts are always over one or two variables, while this work provides a way of combining various datasets into a logical interpretable value.

In a similar fashion, the European Union also exposes several datasets through its DataEuropa platform [8]. The Eurostat [9] is a European Union organization who provides several different visualizations for these datasets. While *Our World In Data* tends to provide the visualizations for all datasets, the Eurostat usually provides visualizations tailored to the data they're portraying, though, once again, they don't attempt to aggregate multiple datasets' dimensions like our work. It is also important

to point out how the Eurostat doesn't always provide interactive visualizations but may only create reports with data displayed.

A completely different approach to visualizing multiple datasets, without losing the dimensionality, is the usage of Linked Open Data (LOD). LOD is a different kind of dataset from the ones used for this work. While the data used for this work presents itself as sparse data that is later aggregated by countries, LOD has as a core principle linking data in a machine interpretable way. In principle, LOD does not care about actually the storage of this data, only caring about properly linking it, creating a huge data space distributed across the Web [10]. Klímek [11] explored this structure to show data on maps, similar to this work. LOD facilitates working the data since it is already linked. Klímek makes use of the of the Czech Linked Open Data (CzLOD) to explore his visualization. Even if this visualization is specific to a single country, the principle is the same, as it portrays data by regions. Klímek doesn't try to aggregate data and simplify to show it to the users, rather focusing on visualizations for high dimensionality data: a sunburst visualizer, a treemap visualizer and columns on a map.

7. CONCLUSIONS AND FUTURE WORK

This project attempts to bring an innovative way of looking into real world data and making comparisons between countries simpler to users. While there are clear imperfections of reducing the dimensionality of the data so much, it still offers an interestingly different approach towards comparing several datasets.

While the current state of the project is usable, there are still several improvements that could be explored in the future. Regarding the analytics, using only the PCA method to reduce the dimensionality of the data is a valid approach, but perhaps some other dimensionality reduction algorithms, such as t-SNE or MDS.

Although more datasets would likely make the tool more interesting and expand its use, a particulary interesting addition would be to included passaport strength as a filter for viable countries. Essentially, the passaports a person holds dictate significantly how easily it is to travel to a particular destination. As such, it would be a useful addition to add the possibility for the user to set which passports they hold and then only consider the countries that accept them.

Regarding the user interaction and visual representations some more features could be added. Simpler features like a dropdown for selecting the histogram's category would improve the UX. More complex additions could augment the tool even further, like allowing the user to select two categories and show a scatter plot graph with each category as an axis. Through this plot graph, more insights related to the relationship between two categories could be perceived, besides allowing for another way to filtered the data by brushing through the graph area.

Furthermore, as Klímek uses LOD [11] as a means to obtain the data, this visualization could also attempt to make use of such data linking abilities to access different kinds of data, and even displaying those relationships to the users.

8. REFERENCES

[1] "Our world in data," Accessed in February 2022. [Online]. Available: https://ourworldindata.org/

[2] "List of country codes by alpha-2, alpha-3 code (iso 3166)," Accessed in February 2022. [Online]. Available: https://www.iban.com/country-codes

[3] "Wikipedia english speaking population | kaggle," Accessed in February 2022. [Online]. Available: https://www.kaggle.com/dmi3kno/wikipedia-english-speaking-population/version/2

[4] T. Serwatka, "Dataset for "sexual orientation and gender-identity (sogi) laws that support and/or limit international development"," *Research Datasets*, 1 2020. [Online]. Available: https://digitalcommons.unf.edu/datasets/1

[5] G. W. Stewart, "On the early history of the singular value decomposition," *http://dx.doi.org/10.1137/1035134*, vol. 35, pp. 551–566, 7 2006.

[6] H. Abdi and L. J. Williams, "Principal component analysis," *Wiley Interdisciplinary Reviews: Computational Statistics*, vol. 2, pp. 433–459, 7 2010. [Online]. Available: https://onlinelibrary.wiley.com/doi/full/10.1002/wics.101https://onlinelibrary.wiley.com/doi/abs/10.1002/wics.101https://wires.onlinelibrary.wiley.com/doi/10.1002/wics.101

[7] "14% young europeans at risk of depression, sweden tops the list: Study." [Online]. Available: https://www.euronews.com/2019/07/05/14-young-europeans-at-risk-of-depression-sweden-tops-the-list-stu

[8] "data.europa.eu," Accessed in February 2022. [Online]. Available: https://data.europa.eu/en

[9] "Eurostat," Accessed in February 2022. [Online]. Available: https://ec.europa.eu/eurostat/web/main/home

[10] I. Ermilov, M. Martin, J. Lehmann, and S. Auer, "Linked open data statistics: Collection and exploitation," *Communications in Computer and Information Science*, vol. 394, pp. 242–249, 2013. [Online]. Available: https://link.springer.com/chapter/10.1007/978-3-642-41360-5_19

[11] J. Klímek, J. Helmich, and M. Necaskỳ, "Application of the linked data visualization model on real world data from the czech lod cloud," in *LDOW*, 2014.

9. APPENDIX

9.1 All datasets used

9.1.1 Air pollution

- Exposure to air pollution: https://ourworldindata.org/grapher/pm25-air-pollution

- Ambient particulate air pollution: https://ourworldindata.org/grapher/death-rates-from-ambient-particulate-air-pollution

- Indoor and outdoor air pollution: https://ourworldindata.org/grapher/death-rates-total-air-pollution

- Outdoor air pollution death rate: https://ourworldindata.org/grapher/outdoor-pollution-death-rate

- Attributed to total air pollution – outdoor and indoor – as a risk factor: https://ourworldindata.org/grapher/share-deaths-air-pollution

- Share of deaths outdoor air pollution: https://ourworldindata.org/grapher/share-deaths-outdoor-pollution

9.1.2 Natural Disaster

- Number of known significant earthquakes: https://ourworldindata.org/grapher/significant-earthquakes

- Direct economic loss attributed to disasters: https://ourworldindata.org/grapher/direct-economic-loss-attributed-to-disasters

- Number of significant volcanic eruptions: https://ourworldindata.org/grapher/significant-volcanic-eruptions

- Internally displaced persons from natural disasters: https://ourworldindata.org/grapher/internally-displaced-persons-from-disasters

9.1.3 Climate Change

- Share of electricity production from renewables: https://ourworldindata.org/grapher/share-electricity-renewables

- Fossil fuel production: https://ourworldindata.org/grapher/fossil-fuel-production

- Fossil fuel consumption: https://ourworldindata.org/grapher/fossil-fuel-primary-energy

- Share of electricity production from fossil fuels: https://ourworldindata.org/grapher/share-electricity-fossil-fuels

9.1.4 Food

- How does diet compare to the EAT-Lancet diet? (The EAT-Lancet diet is a diet recommended to balance the goals of healthy nutrition and environmental sustainability for a global population). The dataset shows how the diet is composed (vegetables, meat, etc..): https://ourworldindata.org/grapher/eat-lancet-diet-comparison?country=USA~GBR~EAT-Lancet~IND~KEN

- Annual food expenditure per person vs. GDP per capita (food at home), 2016: https://ourworldindata.org/grapher/annual-food-expenditure-per-person-vs-gdp-per-capita

- Cost of a healthy diet: https://ourworldindata.org/grapher/cost-healthy-diet?country=USA~GBR~IND~CHN~AUT~MWI~MYS~

- Share of consumer expenditure spent on food (2016), (unfortunately lacks most of Africa): https://ourworldindata.org/grapher/share-of-consumer-expenditure-spent-on-food

- Daily protein supply of animal origin vs. plant origin (2013): https://ourworldindata.org/grapher/daily-protein-supply-of-animal-origin-vs-plant-origin

- Antibiotic use in livestock (2010): https://ourworldindata.org/grapher/antibiotic-use-in-livestock

- Daily meat consumption per person: https://ourworldindata.org/grapher/daily-meat-consumption-per-person

- Daily per capita fat supply: https://ourworldindata.org/grapher/daily-per-capita-fat-supply

- Daily per capita supply of calories: https://ourworldindata.org/grapher/daily-per-capita-supply-of-calories

- Daily per capita protein supply: https://ourworldindata.org/grapher/daily-per-capita-protein-supply?country=MDV~SUR

- Per capita sugar food supply vs GDP: https://ourworldindata.org/grapher/per-capita-sugar-food-supply-vs-gdp-per-capita

9.1.5 Income

- Starting salaries for teachers in public primary education: https://ourworldindata.org/grapher/starting-yearly-teacher-salaries-in-public-primary-education

- GDP per capita vs. economic inequality: https://ourworldindata.org/grapher/gdp-per-capita-vs-economic-inequality

- Share of Total Income going to the Top 1%: https://ourworldindata.org/grapher/share-of-top-1-in-pre-tax-national-income

- Disposable household income of the lowest decile in the income distribution: https://ourworldindata.org/grapher/1st-d

- Median household disposable income: https://ourworldindata.org/grapher/median-income-eeg

- Top marginal income tax rates: https://ourworldindata.org/grapher/maximum-marginal-tax-rates-on-individual-income?time=1979..2002

9.1.6 Education

- Pisa scores (not all countries covered): https://ourworldindata.org/grapher/oecd-pisa-test-scores-for-selected-countries-2000-2012

- Gross enrollment ratio in secondary education (2016): https://ourworldindata.org/grapher/gross-enrollment-ratio-in-secondary-education

- Share of population with completed tertiary education: https://ourworldindata.org/grapher/share-of-the-population-with-completed-tertiary-education

- Number of teachers across education levels: https://ourworldindata.org/grapher/number-of-teachers-across-education-levels?tab=table&country=~OWID_WRL

- Completion rate of lower secondary education (2016), not all countries included: https://ourworldindata.org/grapher/completion-rate-of-lower-secondary-education

- Duration of compulsory education: https://ourworldindata.org/grapher/duration-of-compulsory-education

- Total government expenditure on education, 1970 to 2019: https://ourworldindata.org/grapher/total-government-expenditure-on-education-gdp?country=KOR~CMR~~BGD~USA~GBR

- National average learning outcomes vs. government expenditure per primary student: https://ourworldindata.org/grapher/national-average-learning-outcomes-vs-government-expenditure-per-primary-student

- Share of youth not in education, employment or training (2020): https://ourworldindata.org/grapher/youth-not-in-education-employment-training

9.1.7 Political/ democracy

- Political regime type vs Human Rights Score: https://ourworldindata.org/grapher/political-regime-type-vs-human-rights-score

- Political regime, full autocracy to full democracy: https://ourworldindata.org/grapher/political-regime-updated2016

- Political regime, 2020: https://ourworldindata.org/grapher/political-regimes

- Corruption perception index: https://ourworldindata.org/grapher/share-of-people-paying-bribes-vs-corruption-perception

- Human rights score (2017): https://ourworldindata.org/grapher/human-rights-scores?country=~CHN~HUN~PRK~KOR~NOR~ALB

- Human rights violations: https://ourworldindata.org/grapher/human-rights-violations

- Freedom of the press status, 2017: https://ourworldindata.org/grapher/world-map-of-the-freedom-of-the-press-status

- Refugee population by country or territory of asylum, 2020: https://ourworldindata.org/grapher/refugee-population-by-country-or-territory-of-asylum

- Trust in national government (2018): https://ourworldindata.org/grapher/share-who-trust-government

- Political Participation: https://ourworldindata.org/grapher/political-participation

- GDP per capita vs State fragility: https://ourworldindata.org/grapher/gdp-per-capita-vs-state-fragility

- World Press Freedom: https://ourworldindata.org/grapher/world-press-freedom

9.1.8 *Gender rights*

- Share of women between 15 and 19 years old with no education (2010): https://ourworldindata.org/grapher/share-of-women-between-15-and-19-years-old-with-no-education

- Net enrollment rate in primary education by sex: https://ourworldindata.org/grapher/total-net-enrollment-rate-in-primary-education-by-sex

- Gender Inequality Index from the Human Development Report: https://ourworldindata.org/grapher/gender-inequality-index-from-the-human-development-report

- Proportion of women in senior and middle management positions: https://ourworldindata.org/grapher/proportion-of-women-in-senior-and-middle-management-positions?tab=chart

- Share of firms with female top managers: https://ourworldindata.org/grapher/firms-with-female-top-manager-of-firms-bars

- Women's Economic Rights: https://ourworldindata.org/grapher/womens-economic-rights

- Women's Economic Opportunity Index: https://ourworldindata.org/grapher/womens-economic-opportunity-2012-index

- Unadjusted gender gap in median earnings (lots of countries missing): https://ourworldindata.org/grapher/gender-wage-gap-oecd?tab=map

- Does the law mandate equal remuneration for females and males for work of equal value?: https://ourworldindata.org/grapher/law-mandate-equal-pay

- Gross enrolment ratio, primary: https://ourworldindata.org/grapher/gpi-primary-education

- Gross enrolment ratio, secondary education: https://ourworldindata.org/grapher/gross-enrolment-gpi-secondary

- Gross enrolment ratio, tertiary: https://ourworldindata.org/grapher/gpi-tertiary-education

- Share of male vs female employment in agriculture: https://ourworldindata.org/grapher/share-of-male-vs-female-employment-in-agri

- Share of male vs female employment in industry: https://ourworldindata.org/grapher/share-of-male-vs-female-employment-in-industry

- Share of male vs female employment in services: https://ourworldindata.org/grapher/share-of-male-vs-female-employment-in-services

- Unemployment rate, women: https://ourworldindata.org/grapher/unemployment-rate-women

- Maternal Mortality Ratio: https://ourworldindata.org/grapher/maternal-mortality-slope-chart?tab=map

- Adolescent birth rate in women aged 10-19 years: https://ourworldindata.org/grapher/adolescent-fertility-ihme

- Proportion of women in ministerial positions: https://ourworldindata.org/grapher/proportion-of-women-in-ministerial-positions

-

9.1.9 *Language*

- Population that is able to speak english in a country: https://www.kaggle.com/dmi3kno/wikipedia-english-speaking-population/version/2

9.1.10 Peacefulness/Safety

- Homicide rate: https://ourworldindata.org/grapher/homicide-rate-ihme-vs-world-bank

- Suicide rate vs. homicide rate: https://ourworldindata.org/grapher/suicide-vs-homicide-rate

- Number of deaths by age group: https://ourworldindata.org/grapher/deaths-by-age-group-stacked?country=~OWID_WRL

- Total number of nuclear warheads in the inventory of the nuclear powers: https://ourworldindata.org/grapher/number-of-nuclear-warheads-in-the-inventory-of-the-nuclear-powers

- Deaths from conflict and terrorism: https://ourworldindata.org/grapher/deaths-from-conflict-and-terrorism

- Number of terrorist attacks: https://ourworldindata.org/grapher/terrorist-incidents

- Share of children who report being bullied: https://ourworldindata.org/grapher/share-of-children-ages-13-15-who-report-being-bullied

9.1.11 Happiness

- Suicide rate vs. homicide rate: (also in Peacefullness/Safety) https://ourworldindata.org/grapher/suicide-vs-homicide-rate

- Self-reported Life Satisfaction: https://ourworldindata.org/grapher/happiness-cantril-ladder

- Share of population with depression: https://ourworldindata.org/grapher/depression-ihme-vs-who

- Suicide rates vs. prevalence of depression: https://ourworldindata.org/grapher/suicide-rates-vs-prevalence-of-depression

9.1.12 LGBTQ Rights

- Sexual Orientation and Gender-Identity (SOGI) Laws That Support and/or Limit International Development: https://digitalcommons.unf.edu/datasets/1/

9.1.13 Health

- Hospital beds per 1,000 people: https://ourworldindata.org/grapher/hospital-beds-per-1000-people

- Healthcare Access and Quality Index: https://ourworldindata.org/grapher/healthcare-access-and-quality-index

- Universal Health Coverage Index: https://ourworldindata.org/grapher/universal-health-coverage-index?country=LBR~HND~LVA~KHM

- Healthcare access and quality by level of healthcare spending: https://ourworldindata.org/grapher/haq-by-level-of-healthcare-spending-endpoints

- Health expenditure per person: https://ourworldindata.org/grapher/annual-healthcare-expenditure-per-capita

- Share of Out-of-Pocket Expenditure on Healthcare: https://ourworldindata.org/grapher/share-of-out-of-pocket-expenditure-on-healthcare

- Share of the population at risk of catastrophic expenditure when surgical care is required: https://ourworldindata.org/grapher/share-of-the-population-at-risk-of-catastrophic-expenditure-when-surgical-care-is-required

- Percentage of population covered by health insurance: https://ourworldindata.org/grapher/health-protection-coverage

- Private health expenditure per person: https://ourworldindata.org/grapher/private-health-expenditure-per-person