Comparing Healthcare Resources and Health Outcomes at Country and Regional Level

Authors: Hina Joshua (hinaj), Casey Dye (cadye), and Emin Ozkaya (eiozkaya)
February 2023

Motivation

Associated with the convergence of several global crises such as the COVID-19 pandemic, the Ukraine conflict, and the climate crisis, a reversal in several years of progress has been reported in the United Nation's 2022 Sustainable Development Goals Report.

We as data scientists at the UN's Statistics Division are assigned the task to study the performance of one of the 17 Sustainable Development Goals and the 2030 Agenda for Sustainable Development. Our team's focus is on the SDG, no. 3 - **Good Health and Well Being**, to study the relationship between the variability in health resources and the associated impact on health outcomes, at the country as well as regional level.

One would expect that countries with high healthcare resources will have better healthcare outcomes.

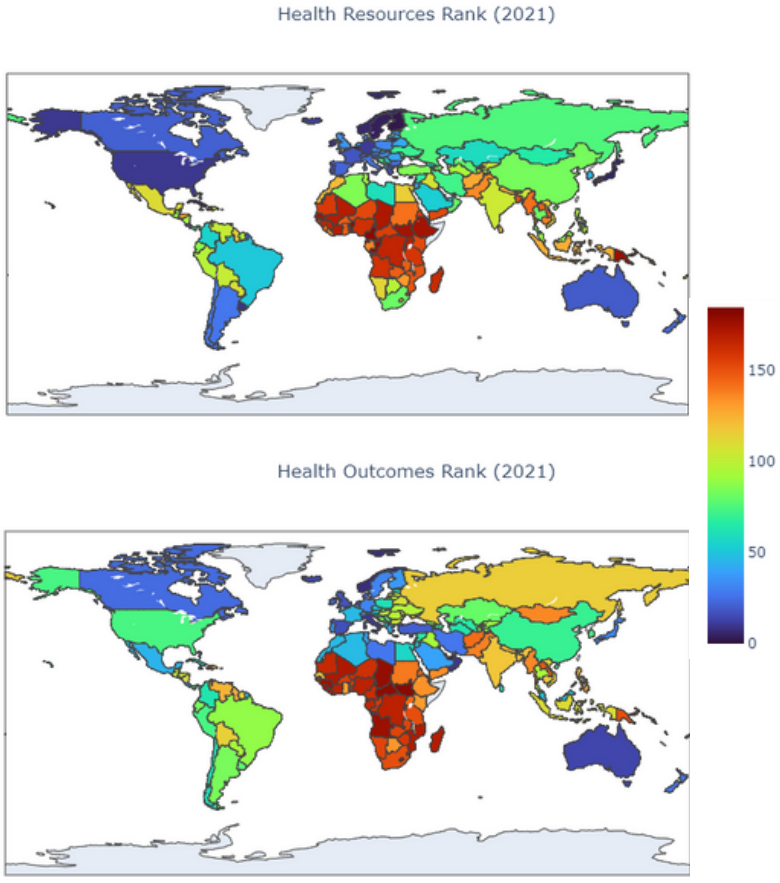
According to Filmer and Pritchett (1999), "While the link between expenditures and outcomes is never automatic in any country, it is generally positive when expenditures are managed and executed efficiently." Baldacci et Al's (2004) research suggests that it is important to couple good policies and governance with increased spending to improve outcomes.

By identifying efficient countries we can identify efficient models of healthcare management and delivery, to learn from as well as implement in countries struggling to meet the 2030 health goals.

Objectives

Create composite Health Resources Index and Health Outcomes index by grouping several health measures, and use these composite indices as a statistical tool to measure the relationship between the health resources and health outcomes in countries/regions.

Rank countries based on their Health Resources and Health Outcomes and use these rankings to identify efficient and inefficient countries/regions- that are better/worse at producing good healthcare outcomes within similar resources.



Executive Summary

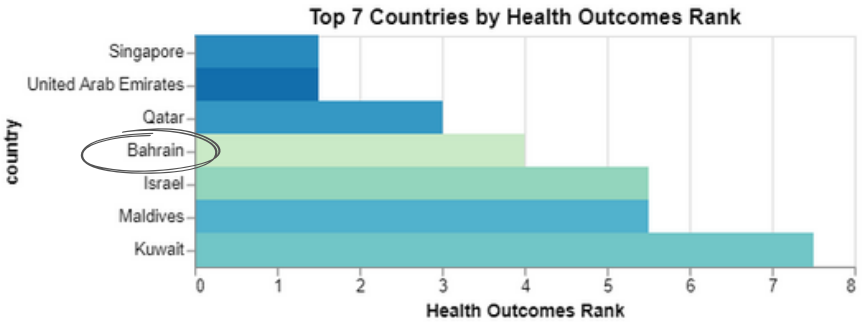
Health Resources Index and Health Outcomes Index are highly correlated among all countries and country groups.

Country Level Findings

- Singapore and the UAE are tied for rank 1 for good health outcomes. Followed by Qatar (no.2) and Bahrain (no.3)
- Among the top 7 countries by good health outcomes, **Bahrain** has the lowest health resources rank (no.124), suggesting that Bahrain might be the **most efficient** among these, at utilizing health resources
- Among top 7 countries ranked by health resources, Lithuania is ranked no.7, however, it's health outcomes rank (no. 127) is the lowest among these. This suggests that **Lithuania** might be the **least efficient** country, among this group, at utilizing health resources

Region Level Findings

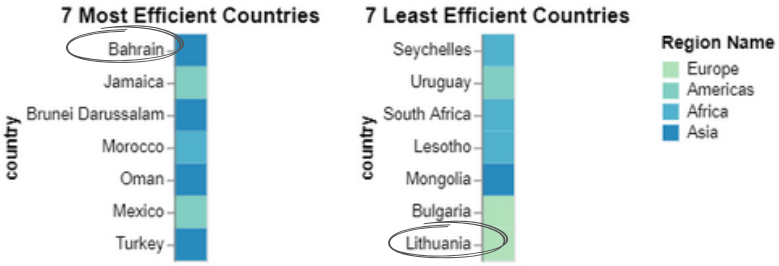
- **Asian** countries are lower in health resources but higher in good health outcomes. Only 2 Asian countries (**Japan** and **Israel**) are in the top 25 by health resources, however, all top 7 by health outcomes are Asian.
- **African** countries are low on both health resources and good health outcomes, and are neither represented in the top 25 by health resources or top 25 by good health outcomes (except **Seychelles** - no.16 in Health Resources)
- **European** countries are high in health resources but lower in good health outcomes. Of the top 25 Health Resources countries, 15 are in Europe, however, no European countries are in top 7 by Health Outcomes, (2 in top 10 are **Norway** and **Italy**)
- Of the 5 **Americas** countries in the top 25 by health resources, only **Canada** is also present in the top 25 by good health outcomes
- Oceania countries Australia and New Zealand are in top 25 by health resources; **Australia** is in top 25 by good health outcomes while New Zealand not too far behind at no.27



Efficiency Ranking

All countries we awarded an 'Efficiency' Rank

Efficiency rank is the difference in the health resources and health outcomes rank of the country. Countries with low resource rank and high outcomes rank are efficient and vice versa.



Next Steps

Explore additional measures or measure weighting to include in indices

To create indices even more reflective of the state of health resources and outcomes, we would like to explore other data sources that have measures related to these areas as well as weighting methodologies to see whether they would further improve the performance of our indices.

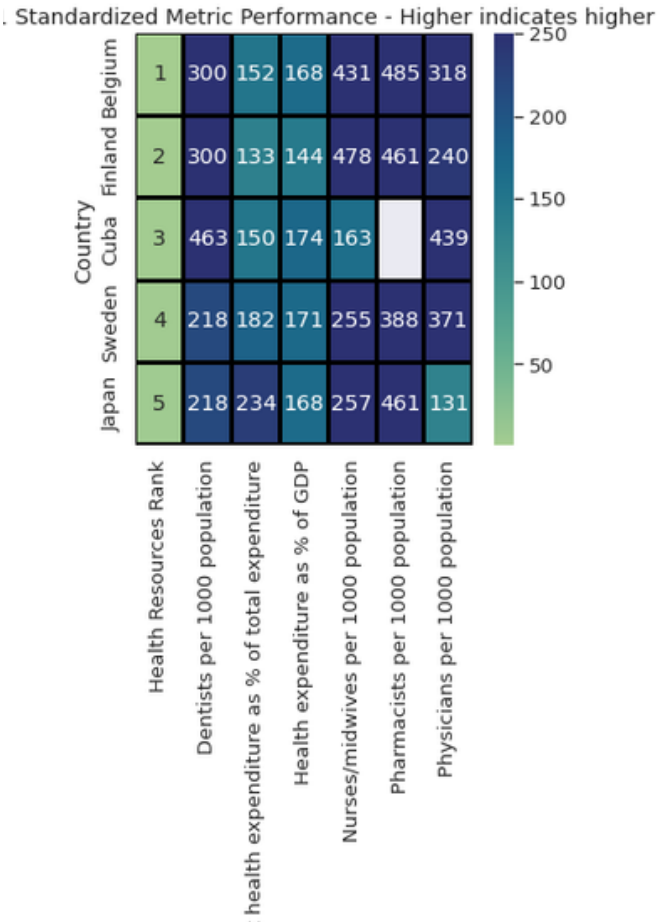
Explore country similarity and dissimilarity with health resources and outcomes

We would like to further explore the current landscape of health resources and outcomes by utilizing cosine similarity to get a score for how similar countries are to one another.

Composite Index Variables

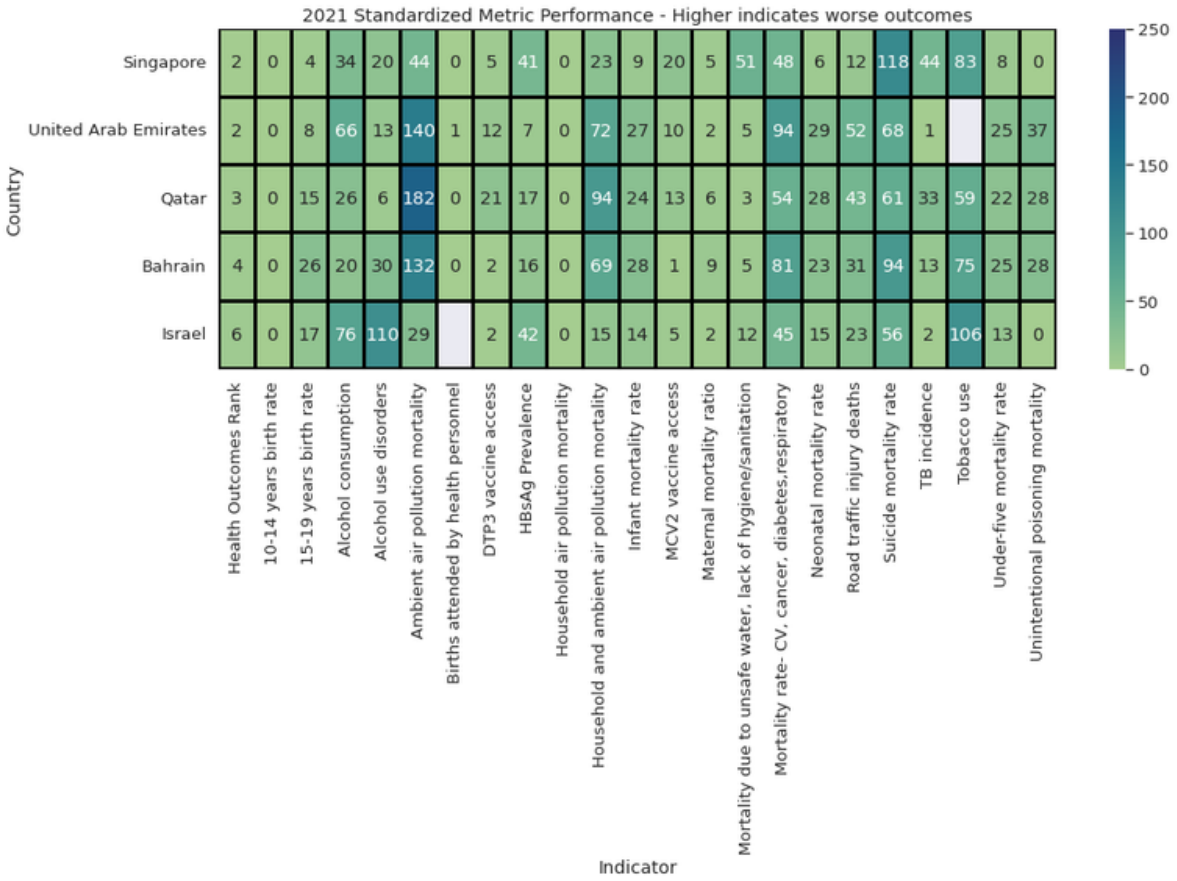
Health Resources Index Variables

There are 6 variables included in the composite Health Resource Index pertaining to both health expenditure (2 variables) and health personnel (4 variables) . The distribution of these variables among the top 5 countries by health resources rank can be visualized in the heatmap on the right.



Health Outcomes Index Variables

We selected 22 variables from the SDG goal 3 measures for the composite Health Outcomes Index. These measures pertain to cause-specific mortality rates, substance abuse-related diseases/mortality, and infectious and non-infectious disease rates. The SDG Goal does include other measures. We chose to exclude personnel, emergency preparedness, and expenditure-related measures to focus on health outcomes measures. Measures, where more than 30 countries were missing data, were also excluded from the analysis. The distribution of these variables among the top 5 countries by health outcomes rank can be visualized in the heatmap below. Visualization idea - (Lim et al. 2016)



Data Manipulation

Merging Datasets

Standardizing Dimensionality

Some measures were duplicated over multiple dimensions, disaggregated by sex, age, location type, et al. To ensure no duplication, we dropped any disaggregation that also had the measure with both groupings. For measures where this was not available, we appended the dimension to the measure name to create a unique measure.

Standardizing to Long Format

In order to merge our datasets, we utilized a long format table to concatenate the individual dataframes with yearly measures.

Standardizing ISO3 Codes and Country Names

We created a mapping dictionary from the various data sources with both country name and ISO3 code. We utilized this along with manual mapping or countries that did not have a match to add ISO3 code where necessary. We used ISO3 code to merge the region and HDI mapping to the measure data.

Creating Wide Format Table

We then utilized the pivot function to create a wide format data frame. We also filtered data to more recent years, including years 2000-2021.

Handling Missing Data

After an analysis of missing records by both measure and country, we used the wide format data frame to backfill and frontfill the missing years of data.

We looked at the number of missing records after doing so and determined an acceptable threshold for countries missing measures and measures missing countries. Any country or measure that was above this threshold was dropped. We did allow some missing data because we were combining metrics. This allows us to consider more countries, but could cause some inconsistency in comparison.

Creating Composite Indices

Standardizing Measures

In order to combine performance on multiple measures, we calculated distance to mean for each measure and each year (Freudenberg 2003). This involved dividing the country score by the total mean for that year for that measure and multiplying by 100. This allowed us to easily compare measures.

For the Health Resources Index, we averaged these for each measure. We utilized a measure grouping for the Health Outcomes Index, averaging the measures by group prior to averaging all groups. This allowed us to not overemphasize groupings measured in more ways. We created an index for each year as well as for all years in aggregate.

Standardizing Dimensions

User-defined fields

For the Health Outcomes Index, we utilized a dictionary to map the individual measures into measure groups, rename the measures, and denote whether a higher value indicated better performance. Most measures are lower better, so we inverted those that were higher better so that the resulting index could give some indication of performance.

Creating Ranks and Performance Categories

To further enhance our ability to compare countries, we created categorical variables by ranking yearly performance on the Health Outcomes Index and Health Resources Index. We also used the indices to create bins to group countries in terms of performance into Best, Good, Fair, and Bad categories.

Underlined words will bring you to relevant code (must be logged into UMICH account).

Data Manipulation

Evaluating the Health Resources Index

After creating the resources index, we explored the correlation between the index and its underlying measures as well as the distribution of the index.

Findings

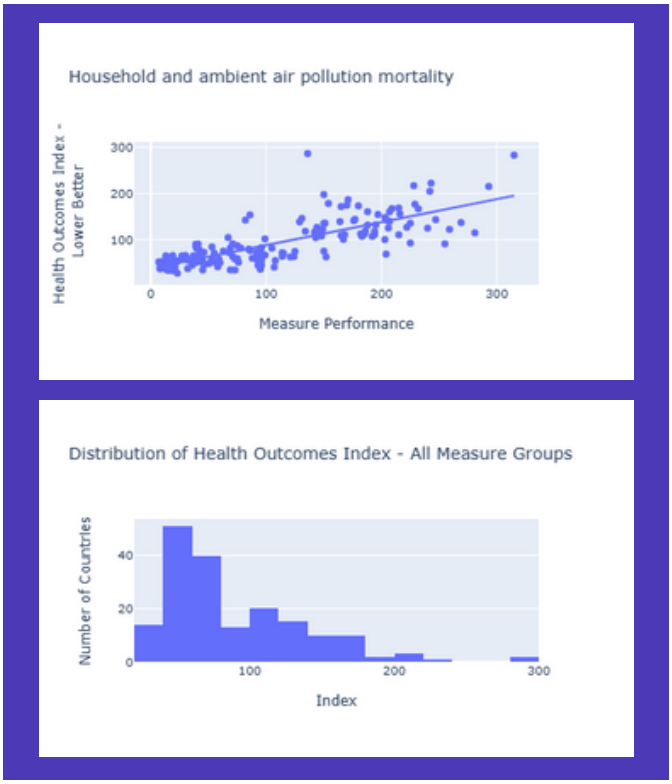
- The index was right skewed, indicating many countries with lower resources. This held true both when aggregated and disaggregated over time.
- Europe had the highest median resource index and Africa had the lowest median index when all regions were compared with a box plot.
- For all underlying measures, we found there to be a positive correlation between the index and the measure, indicating a higher index would mean a greater number of personnel per 1000 population or a greater percentage of expenditure.

Evaluating the Health Outcomes Index

We explored the correlation between the index and its underlying measures as well as the distribution of the index for the health outcomes index as well.

Findings

- The index was right skewed, indicating many countries with good healthcare outcomes. This held true both when aggregated and disaggregated over time.
- When visualized with a box plot, the outcomes by region were more similar than the resources. Africa had a higher median and 75th percentile along with more outliers, indicating worse performance.
- For most underlying measures, we found there to be a correlation between the index and the measure in the expected direction - this was positive if we expected a lower measure value to indicate better outcomes, as in the mortality rate measures and negative if we expected a higher measure value to indicate better outcomes, as in vaccination access. This did not hold for the alcohol consumption, alcohol use disorder, and tobacco use measures.

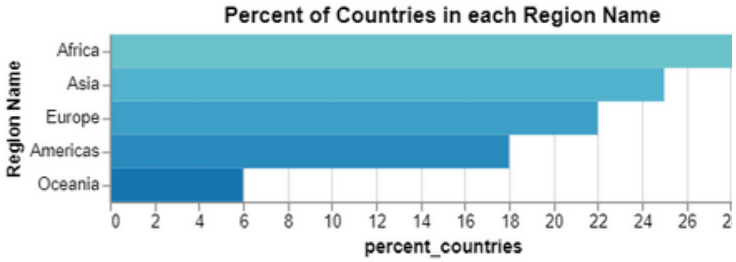


Underlined words will bring you to relevant code (must be logged into UMICH account).

Exploratory Data Analysis

Percent of Countries in Each Region

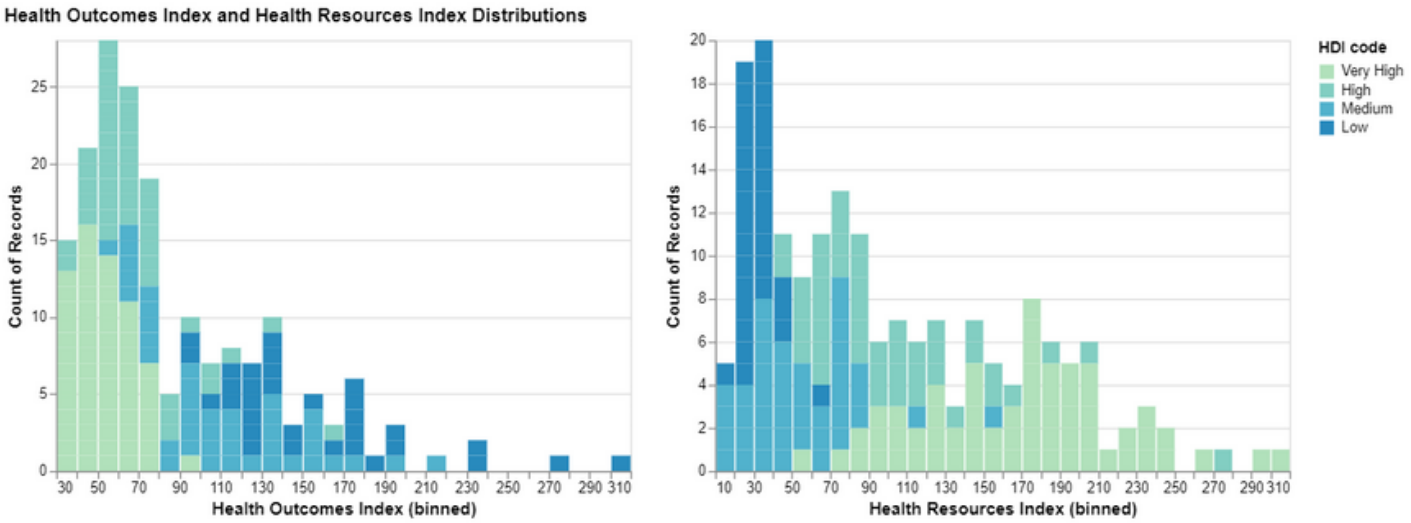
The dataset has a total of 181 countries. Africa has the highest percentage of countries (29%) and Oceania has the lowest (6%). Since the representation of countries in each region is not uniform, it will help interpret the representation of each region among the highest or lowest-tier country groups for variables of interest.



Distribution of Health Resources and Health Outcomes Index

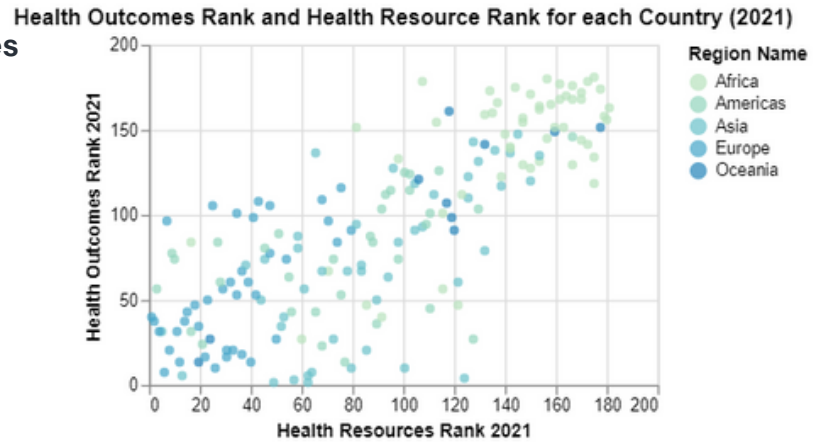
- Both Outcomes and Resources are right skewed, with a lot of countries with good outcomes and a lot of countries with low resources. Outcomes are more skewed than resources
- More Developed countries (very high -high HDI) have better (lower) healthcare outcomes
- More Undeveloped countries (low - medium HDI) a lot lower healthcare resources

.



Relationship Between Health Resources and Health Outcomes

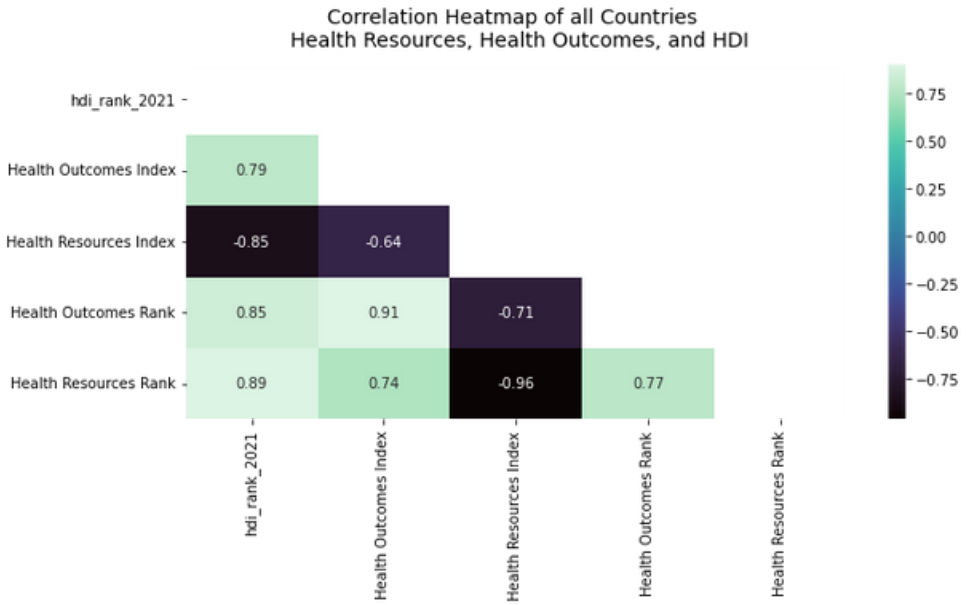
The scatterplot on the right indicates a linear relationship between the healthcare resources and healthcare outcomes of a country. This relationship is worth exploring, especially at the regional level. This finding became the basis for the direction of our analysis. We decided to explore which regions exhibit the strongest relationship between the two main variables

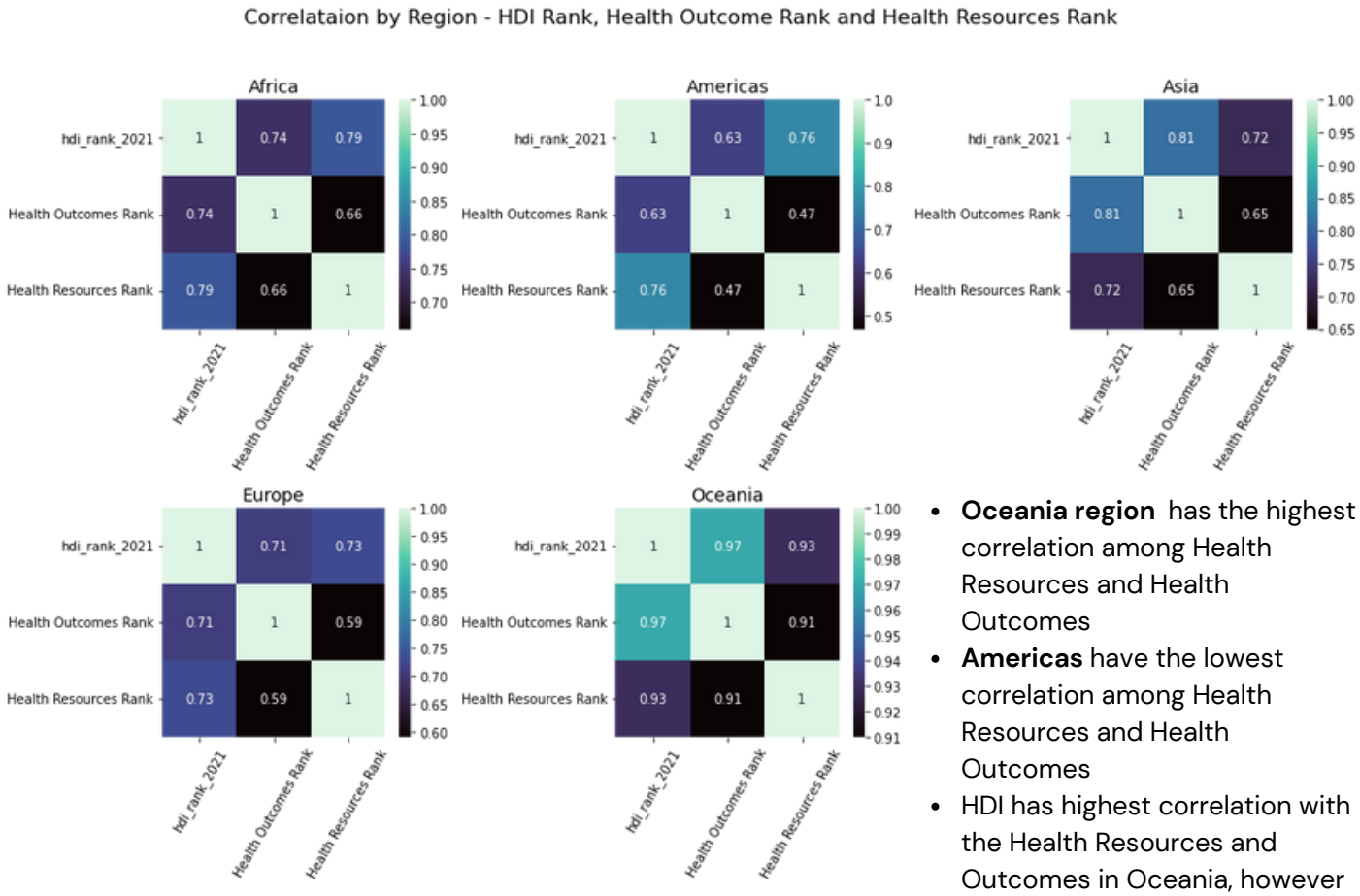


Analysis & Visualizations

Upon computing the Pearson Correlation (heatmap on right) between the two major variables (Health Resources and Health Outcomes composite indices) as well as the Human Development Index, we found that all three indices are strongly correlated with each other.

In order to answer the question if the human development, health resources and health outcomes have a stronger relationship in certain regions as opposed to others, the correlation was computed for individual regions.





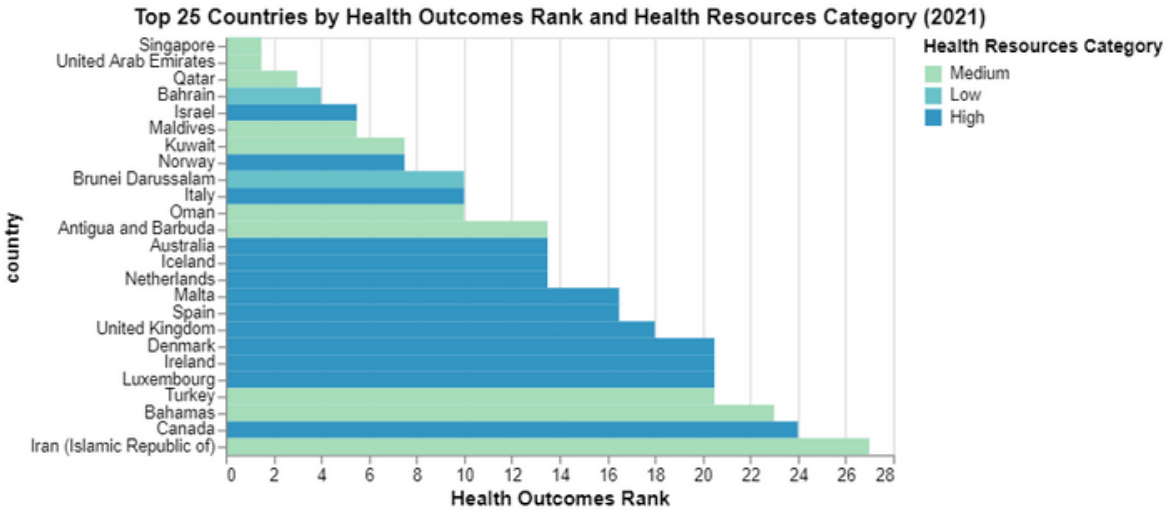
- **Oceania region** has the highest correlation among Health Resources and Health Outcomes
- **Americas** have the lowest correlation among Health Resources and Health Outcomes
- HDI has highest correlation with the Health Resources and Outcomes in Oceania, however this association is lowest in the Americas and Asia.

Analysis & Visualizations

The top countries by health resources and outcomes

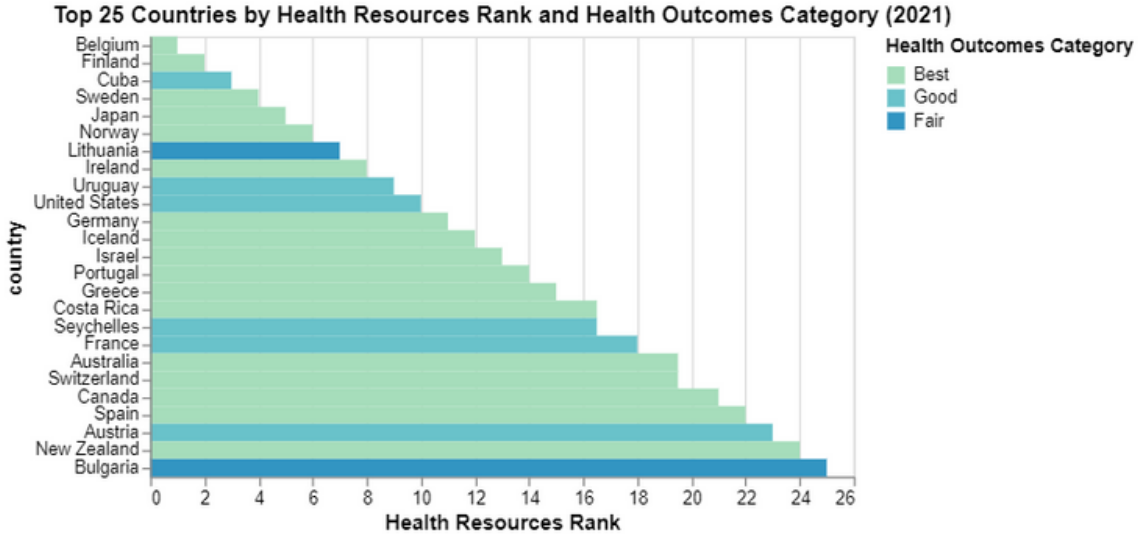
One of the objectives of this analysis was to identify countries with 'efficient' utilization of their healthcare resources. Countries that can produce exceptional health outcomes while utilizing less than high health resources were considered efficient.
We plotted, in the bar chart below, the 25 highest-ranked countries by healthcare outcomes and health resource category.

The top 4 countries, Singapore, UAE, Qatar and Bahrain do not utilize 'High' health resources but have produced the best health outcomes in 2021. Among these, **Bahrain** has the lowest health resources, indicating that it might be the most efficient country in this group of 25 best.



Similarly, in order to identify inefficient countries, we ranked countries by their Health Resources rank and visualized the top 25 along with their Health Outcomes rank. Countries utilizing the highest health resources but producing less than the best health outcomes, were considered inefficient.

In the bar chart below, we can see that most countries with the highest health resources also are in the 'Best' outcomes (above 75th percentile)category. However, there are a few (Cuba, Uruguay, USA, Seychelles, France and Austria) in the 'Good' outcomes category (between 50th and 75th percentile). The least efficient countries are **Lithuania and Bulgaria**, which are in the 'Fair' outcomes category (between 50th and 25th percentile).



Analysis – One-Way ANOVA

We wanted to know if the country groups grouped by "hdicode" or "Region Name" perform differently in the health resources and health outcomes indices. For this purpose, We planned to conduct the one-way ANOVA hypothesis test for the mean variances of countries within "hdicode" and "Region Name" groups if they pass the assumptions required before applying one-way ANOVA.

As One-Way ANOVA assumptions, We checked randomness and independence (trusted data resources), homogeneity of variance (using the ratio of the largest to the smallest sample standard deviation), and normality using Q-Q plots). We also set following hypothesis and decision rule as follows:

F test for differences in more than two means

H_0: $\mu_1 = \mu_2 = \mu_3 = \ldots = \mu_c$

H_1: Not all μ_i's are equal, where i = 1, 2, 3, ..., c.

Level of significance = α (will be used 0.05)

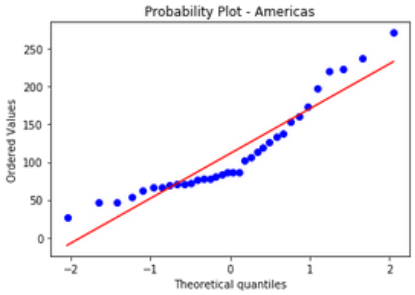
Decision Rule: Reject H_0 when F > F Critical Value OR p value < α

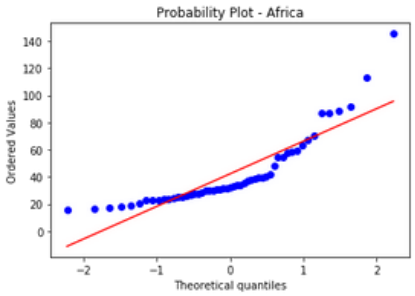
After testing all the assumptions, the only significant group and index to apply ANOVA was the "Region Name" group for health resources index. The other options unfortunately fail to pass homogeneity of variance tests.

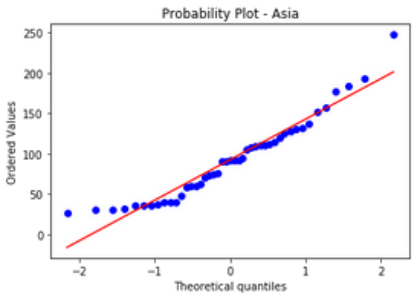
We calculated the F test statistic, critical value, and p-value to make inferences about the group and index mentioned above. The F statistic and p-value turn out to be equal to 53.850 and 1.105e-29 respectively. Since the F statistic is significantly bigger than the F critical value and the p-value is less than 0.05 hence we would reject the null hypothesis. This implies that we have sufficient proof to conclude that there exists a difference in the performance among five country groups (Africa, Americas, Asia, Europe, Oceania) in the context of health resources index.

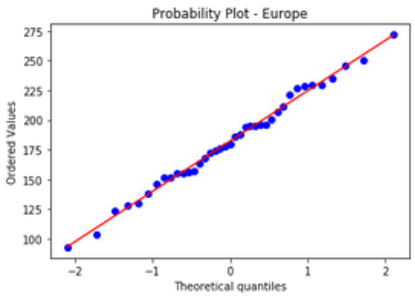
As a next step we may further analyze what groups are different from the others and why, using Tukey's Honestly-Significant Difference as a post-hoc test which will not be in our scope in this project.

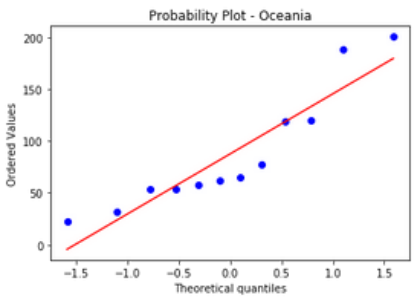
QQ-Plots
To test for Normality











The Q-Q plots show generally straight-line pattern if it is from a normal distribution. We may assume that the data for each group falls roughly on a straight line based on the Q-Q plots above.

Data Sources

SOURCE	DESCRIPTION AND FEATURES	ACCESS AND RESULTS
UN Health Personnel - CSV 1105 KB	Contains measures related to health personnel by country by year • Health personnel per 1000 population - numeric	• CSV from URL read into dataframe • 5,907 records • 2000-2020
UN Expenditure on Health - CSV 529 KB	Contains measures related to health expenditure by country by year • Health Expenditure as a percent of GDP - numeric • Health Expenditure as a % of Total Expenditure - numeric	• CSV from URL read into dataframe • 2,262 records • 2000-2019
UN Sustainable Development Goals - API Swagger page using the v1/sdg/Series/DataCSV endpoint 131200 KB	Contains measures related to progress on UN Sustainable Development Goal 3 relating to health by country by year with some additional dimensionality • Mortality - numeric • Births - numeric • Alcohol/tobacco/substance use - numeric • Hepatitis and TB incidence - numeric • Vaccination access - numeric	• CSV read into dataframe • CSV obtained manually using POST method from SDG API and filtering to Goal 3 • 224,468 records • 76,840 records after filtering • 1985-2021
UN Human Development Indices (HDI) - CSV 1740 KB	Contains measures included in Human Development Index calculation • Human Development Index Classification - ordinal • Human Development Index Rank - ordinal	• CSV accessed via URL read into dataframe • 206 records • 1990 -2021
UNSD Regions - CSV 21 KB	Contains region specification for countries • Region Classification - nominal	• CSV manually downloaded and read into dataframe • 249 records

Additional Sources Explored

Global Health Observatory | Global Health Expenditure Database

While we were interested in additional resource measures - percent of population pushed into poverty from healthcare expenses and inpatient and outpatient curative care as a percent of GDP - from these sources, we found that too many countries were not represented for the measures, so removed them from our analysis.

The Global Health Observatory (GHO) data was accessed via API and returned 4,002 results after filtering (60,086 intially). The GHO data ranged from 1985 to 2020. The Global Health Expenditure Database was loaded from an XLSX file manually downloaded with 4,225 rows. After filtering measures and melting to a long format, the data frame consisted of 8,448 records ranging from 2000 to 2021.

Data source download point can be accessed from link in Data Sources column. UN Health Personnel, Expenditure and HDI files download automatically. The code loading the data source is located in Access and Results column (must be logged into UMICH account).

Statement of Work

TEAM MEMBER	RESPONSIBLITIES
Casey Dye	Load UN Sustainable Development Goals, UNSD Regions, WHO Global Health Observatory, and WHO Global Expenditure Database | Merge datasets and create long/wide format tables | Create mapping dictionaries | Calculation and initial analysis of composite indices | Exploratory visualization
Hina Joshua	Load, merge and format UN Health Expenditure, Health Personnel, and Human Development Index Datasets. Data Cleaning and Manipulation- address missing values, create categorical variables from quantitative. Analysis of major variables including visualizations for univariate and multivaraite analysis
Emin Ozkaya	Data Manipulation: Identify missing values by using heat maps and sorting by features in percentages in ascending order. Visualization: Create pair-plots, histograms, and correlation matrices for visual exploration of data. Analysis: Apply one-way_ANOVA test for the mean differences of country groups for composite indices.
Methods and Tools Used for Collaboration	Method: Agile Methodology; Continues collaboration and improvement every stage. Zoom: Face to face meetings approximately twice a week. Slack: Continuous messaging in case of urgent matters. Jira: Task assignments and tracking. Google Colab: Code sharing. Canva: Project report design.

Limitations

Due to the incompleteness of data for all countries and all years, missing values were interpolated and those countries with missing values beyond our threshold were dropped. Dropping countries with incomplete/missing data might lead us to drop a certain group or category. Not representing this group could introduce selection bias in our analysis

References

- Baldacci, E., Benedict Clements, Sanjeev Gupta and Qiang Cui. 2004. Social Spending, Human Capital and Growth in Developing Countries: Implications for Achieving the MDGs. IMF Working Paper 04/217.
- Filmer, D. and Pritchett, L. 1999. The impact of public spending on health: Does money matter? Social Science & Medicine 49 (10):1309–23.
- Freudenberg, M. (2003), "Composite Indicators of Country Performance: A Critical Assessment", OECD Science, Technology and Industry Working Papers, No. 2003/16, OECD Publishing, Paris, https://doi.org/10.1787/405566708255.
- Lim, S. S., Allen, K., Bhutta, Z. A., Dandona, L., Forouzanfar, M. H., Fullman, N., Gething, P. W., Goldberg, E. M., Hay, S. I., Holmberg, M., Kinfu, Y., Kutz, M. J., Larson, H. J., Liang, X., Lopez, A. D., Lozano, R., McNellan, C. R., Mokdad, A. H., Mooney, M. D., … Murray, C. J. (2016). Measuring the health-related sustainable development goals in 188 countries: A baseline analysis from the global burden of disease study 2015. The Lancet, 388(10053), 1813–1850. https://doi.org/10.1016/s0140-6736(16)31467-2