Analysing the Global Air Traffic Network at Different Spatial Scales

Agnes Annilo, Karl Hendrik Tamkivi

Network Science (LTAT.02.011)

***Abstract*—This study project investigates the spatial structure and disruptive event resilience of the global air traffic network across various spatial scales using data about airports and flight routes while also incorporating additional metrics about economic wealth and travel restrictions. A general analysis of the network was conducted to give an overview of the central nodes and community structure to support further analysis. Air traffic risk scores in the case of potential local catastrophes were calculated for different countries and regions based on outgoing routes per capita, airport density and visa restrictions. The findings reveal significant variability in risk scores while also identifying key airports and countries that play crucial roles in interregional connectivity. The analysis also highlights a strong correlation between air traffic risk scores and the economic wealth of countries.**

The study provides valuable insights into safe areas with high catastrophe resilience and contrary into regions with potential vulnerabilities. Future research could further refine the methodology by incorporating additional parameters such as flight frequency, and environmental variables, enhancing the understanding of the air traffic network's robustness. These enhancements could lead to more targeted strategies to mitigate the impacts of potential disruptive events, ultimately contributing to a more resilient global air travel system as a result.

I. Introduction

Airline travel has become an integral part of modern society, playing a vital role in economic development, tourism, and global connectivity [1], with its growth experiencing an exponential increase in recent decades, reaching a total of 4.4 billion passengers in 2018 and the number being projected to almost double by 2040 [2]. Analysing human travel and tourism through a network perspective is a widely adopted and effective approach in research [3]. Such studies on airports and their passenger volume are common, mostly with analysis on socio-economic indicators such as population or GDP for additional insights [2]. The structure of airline route networks is typically modelled as a directed network, with airports serving as nodes and flight routes as edges. These edges can be characterised by multiple attributes such as the number of airlines per route, flight frequencies and distances, providing a comprehensive overview of the air travel network. [4], [5], [6]

In recent years studies focusing on airline travel analysis through the network analysis perspective have had a particular focus on identifying key airports and routes [7], [8], analysing their topological properties [5], [6], and exploring the dynamics of the network over time [4], [8]. Many of the studies have focused specifically on various network measures such as centrality, clustering and community detection to identify important airports and flight routes in the global air travel network [7], [8]. Metrics and methods often occurring in those papers are degree centrality, betweenness centrality and PageRank algorithm, which aim to provide insights into the connectivity and centrality of different airports and regions in the global air travel network [7], [9], [2].

There have also been a few studies that have attempted to use different centrality measures to assess the air travel network's robustness and vulnerability in the case of events, which may cause the need for mass evacuation, such as natural disasters and targeted attacks [6], [10]. Unfortunately, only a limited number of papers have analysed the relationships between different network measures and the air travel demand of nodes within the air travel network at various spatial scales, such as the metropolitan area, country, or regional level [2]. A better understanding of the spatial aspect of the characteristics of air travel network nodes would allow to detect problematic areas in the current network structure and better alleviate the possible negative effects caused by potential disruptive events in the future. Therefore, this project aims to provide deeper insights into the spatial structure of the global air traffic network at various spatial scales, to identify potential risk areas and address them more effectively in the future.

II. Dataset

The data used in the following analysis originates from five main sources, which include the OpenFlights (OF) datasets [11] containing foundational information about airports and routes; the World Bank (WB) dataset [12] containing information about demographic and economic measures of countries that belong to the United Nations alliance; the Passport Index Dataset [13] containing data about the visa requirements and lastly the geo-countries GitHub repository [14] which holds useful information for the geographic visualisations.

Airports and routes. For this analysis, we focus exclusively on commercial flights and their respective airports, limiting our study to flight routes that operated in 2014. Although we are aware that this data might not represent the latest conditions of the global air network, we expect that the findings of our study will still be relevant and the structure of the network has not radically changed within the last decade. As of June 2014, the airports dataset contains essential details for 7,698 airports, including identifiers such as names, IATA, and ICAO codes. This dataset also includes geographical information like latitude, longitude, and country name (but not country ISO codes).

The routes dataset, on the other hand, documents 67,663 routes connecting 3,321 airports across 548 airlines, with each entry providing details about the airline, the originating and destination airports, and the specific airplane model used. This dataset also notes instances of codesharing, a common practice in the aviation industry where two or more airlines operate the same flight [2].

Population, region labels and income group. The WB dataset we're using is derived from the United Nations Population Division and encompasses data for 266 geographical areas, which include entities that may not strictly qualify as independent countries. For our analysis, we are particularly focused on the population figures from 2014, aligning with the timeframe of our flights data, and also from 2022, which provides the most recent population metrics available.

In addition to population data, we are interested in the income group classification. This classification labels geographical areas on an ordinal scale with four discrete levels - ranging from low to high income - based on the GDP per capita, as determined by the WB. This income classification helps in understanding the economic backdrop of different regions, which is crucial for assessing potential air travel demand, especially in response to unforeseen events.

Moreover, the dataset offers a valuable variable regarding the geographic region of each area, which is instrumental in analyzing air travel dynamics at different spatial levels. This comprehensive set of variables enhances our ability to explore and interpret the intricacies of global air travel patterns.

Visa requirements. The data repository that details visa requirements between countries worldwide has been updated annually from 2019 to 2024. For our analysis, we are using the data as of January 2024. This dataset includes pairs of three-letter country ISO codes for 199 different countries, along with a visa requirement label for each pair. For clarity and ease of analysis, we have consolidated the numerous visa requirement labels into four distinct categories: no visa, e-visa, visa required, and no entry. This simplification will facilitate the calculation of visa restriction metrics in our further analysis.

We chose the January 2024 data because the visa requirement labels from 2019 to 2023 include travel bans related to COVID-19. Including these labels would complicate our analysis due to their temporary nature. To simplify, we have excluded this complexity for now. However, studying the impact of these travel restrictions on network operations at various spatial levels during and outside of pandemic conditions could be a valuable area for future research.

TABLE I
SUMMARY OF DATASETS

Data source	File names	Rows	Cols
OpenFlights	airports.txt	7698	14
	routes.txt	67663	9
World Bank	population.csv	266	4
	areas.csv	265	6
Visa requirements	visa_20240105_iso3.csv	39601	3
Geodata	countries.geojson	255	4

Data cleaning and merging. The further analysis in this project required some data preprocessing and table merging which due to different data inconsistencies and quality standards was always not straightforward and thus some more potentially confusing steps will be described in more detail here.

The first significant challenge encountered when integrating the OF datasets with the WB data tables stemmed from a discrepancy in country identifiers. The OF datasets use country names, while the WB datasets adhere to United Nations naming standards, leading to many mismatches in country names between the two datasets. Attempting to match these names based on string similarity proved inadequate; setting a reasonable similarity threshold failed to capture all necessary matches, and lowering this threshold resulted in numerous inaccurate pairings. To achieve a high-quality merge, it became necessary to manually remap the country names in the WB dataset to align with the OF data. This process resulted in assigning the appropriate country ISO codes to the airports listed in the OF datasets, which was crucial for subsequent data merging activities as well.

However, this manual intervention came with the drawback of excluding airports not listed in the WB datasets, most of which fortunately probably have minimal impact on global air traffic. It is important to highlight however that this remapping process led to the exclusion of Taiwan from the further analysis due to it not belonging to United Nations. As of 2014, Taiwan had approximately 23 million inhabitants [15] and omitting this region from the further analysis could significantly influence the overall findings of the study. This example underscores the complexities and potential limitations involved in data integration, especially when dealing with geopolitical sensitivities and varying international standards.

The second major challenge in data integration involved enhancing the OF routes dataset with meaningful visa requirement data. This task started by simplifying the existing visa requirement categories into a format that could be more readily analyzed. Given that the OF dataset does not contain any routes where 'no entry' is applicable (as such routes are inherently not possible), we were left with three primary categories: no visa, e-visa, and visa. These categories were then converted into numerical values to facilitate analysis. Specifically, 'no visa' was assigned a value of 0, indicating no visa restrictions, while both 'e-visa' and 'visa' were assigned a value of 1, despite the fact that these two categories can represent significantly different levels of travel restriction in practice. Following this categorization, the visa restriction values were incorporated into the OF routes data table. This merging was based on pairs of source and destination country ISO codes, which were sourced from the WB data set previously.

III. METHODOLOGY

Our research project aims to assess the robustness of the airline network across various spatial scales, with a specific focus on country and world region levels. To achieve this, we have devised a risk score that aims to quantify the network's vulnerability. The detailed description of our risk score calculation methodology is provided in a separate section below. In addition to calculating the risk score, we will conduct a comprehensive analysis of the network structure on the level of airports and countries. This analysis will enhance the understanding of critical nodes within the network. The entire workflow is visually depicted in Figure 1.

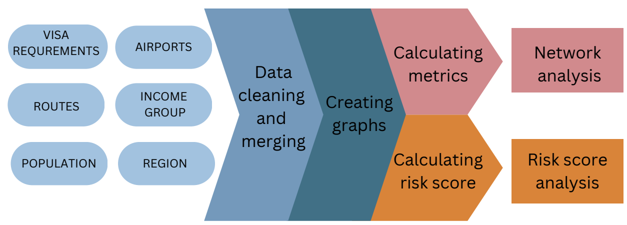

Fig. 1. Flowchart of workflow.

Our analysis is based on a directed graph $G = (V, E)$ where each vertex V represents an airport. The set V includes all the airports' starting points or destinations in the route dataset. Each edge, defined as a tuple (u, v), where u and v are vertices in V, represents a flight route between two airports. The set E includes all the direct flight connections from one airport to another as defined in the routes dataset. Since multiple airlines can serve the same node, we will sum all the possible routes between two airports, keeping the number of duplicate routes as a weight for the edges.

The final graph is created by selecting the largest strongly connected component. The weights used in calculations will differ depending on the context of the calculations, but will be based either on the visa requirements or the number of routes that connect different nodes. All the metrics described below will be calculated using the Python package *networkx*.

Weighted degree centrality, edges weighted by visa restrictions. Degree centrality is a straightforward metric that describes the number of edges a node has, in this case with the added weights of the visa restrictions. This metric will give an overview of how well-connected the different airports are while also reflecting on the potential ease or difficulty of travelling from the respective airports.

In order to calculate the weighted degree centrality, we will weight the edges by the visa restrictions and to calculate the weights we will create the following mapping for the visa requirement labels, where a higher value for the weights signifies a better connection:

$$\text{Visa Liberty} = \begin{cases} 1 & \text{if no visa} \\ 0 & \text{if e-visa} \\ 0 & \text{if visa.} \end{cases}$$

To calculate the weighted degree centrality we will calculate the weighted degree of each node, which is the sum of the edge weights attached to the node, then normalise this based on the maximum degree value within the network.

PageRank. PageRank was chosen to calculate the importance of a node in G. PageRank, a variant of eigenvector centrality, measures the relative importance of an airport, where an airport gains importance based on the amount of routes which it is a destination for. PageRank also incorporates random transitions, making it robust against cyclic structures.

Community detection. We will use either modularity maximisation and Infomap algorithms to find groups of airports which are more densely connected to each other. Based on the communities found, we will conduct an analysis of whether the communities align with world region borders or if they are based on some other factors.

Analysis on country or world region level. To analyse the network dynamics and characteristics at different spatial scales, we will create graphs $G = (V, E)$, where the vertices are either countries or regions, the edges are directed and the weights for the edges are based on either the visa liberty mappings or the number of routes that connect different nodes. A similar analysis to the airport-based analysis can then be conducted as described previously, but now on a wider spatial scale.

Risk score calculations. The final assessment on the resilience of the air traffic network against potential sudden disruptive events like natural catastrophes or epidemics at different spatial scales will be based on the risk score calculation for each country and world region. The comparison of countries will be based on the risk score (RS), which is defined as follows:

$$RS = \frac{(1 - ORPC_N) + (1 - A_N) + VR}{3} \cdot 100,$$

where

- $ORPC_N$ represents the Outgoing Routes Per Capita normalised,
- A_N represents the Airports Per Country normalised,
- VR represents the Visa Restriction score (1 - Visa Liberty).

The domains for each variable are given by:

- $RS \in (0, 100]$,
- $ORPC_N \in (0, 1]$,
- $A_N \in (0, 1]$,
- $VR \in [0, 1]$.

The previously described formula uses quantile-based normalisation technique [16] from Python package *sklearn* to make respective metrics follow a uniform distribution. This normalisation method was chosen due to the presence of small island countries, that had extremely high metric values even after applying frequently used min-max normalisation technique. In this analysis a quantile-based transformer with a uniform distribution and 100 quantiles was used. Removing these outlying countries was not an option since we did not want to exclude any countries from the global analysis. The quantile-based transformer was a suitable choice since it maps the data into a more balanced uniform distribution, ensuring that the final risk score estimate is meaningful and comparable across all countries.

IV. Results

Overview of Graph Analysis

Centrality Measures: The general analysis was conducted to give an overview of the network structure and support the findings of the risk score calculations. First, degree centrality and betweenness centrality were calculated without weights, with Frankfurt Main Airport and Charles de Gaulle International Airport being the most central airports respectively. It was found that the top ten most important airports in terms of the total number of flight routes are involved in 9.84% of all the possible routes. The average degree for the nodes in the unweighted network was 23.4.

Further analysis was conducted on airports with the additional route weight being the visa restriction mapping score. The weighted degree centrality was calculated for which the results can be seen in Table IV. The most important node in terms of the weighted degree centrality amongst all airports turned out to be the Denver International Airport, with the United States and Spain both having three airports within the top ten most influential nodes. Since this analysis is conducted on a directed graph, both incoming and outgoing nodes are included in the calculation, highlighting the most connected airports. Although many U.S. airports primarily serve domestic routes (e.g., 89% of flights in and out of Denver are domestic), this information remains valuable. It demonstrates the overall connectivity of an airport, which is crucial for evacuation scenarios that may not involve international travel (especially for big countries) but still require robust domestic connectivity. Some of the least influential airports in this regard available in the dataset are located in Italy, Czech Republic, Turkmenistan, North and South Korea.

PageRank scores for the nodes were calculated without including any weight parameters. The maximum centrality value was 0.005, indicating that there are no overwhelmingly dominant hubs. The damping parameter was adjusted to 0.95 to better analyze flights as connecting flights between multiple airports, linking major airports together The top three most influential nodes based on the PageRank calculation were the Hartsfield Jackson Atlanta International Airport, the Istanbul Atatürk International Airport and the Chicago O'Hare International Airport. These results are consistent with the characteristics of these airports. Since 1998, the Hartsfield Jackson Airport has been the world's busiest airport by passenger traffic [17]. The Atatürk International Airport was the hub for Turkish Airlines until 2019 [18] and the Chicago O'Hare airport is considered the world's most connected airport [19].

Community Detection: Community detection was performed using both greedy modularity maximisation (Fig. 2) as well as Infomap algorithm, out of which the first one produced more stable and realistic results. With both detection algorithms, communities often aligned with world region borders, indicating that on the airports spatial scale there are more shorter routes inside of a region than longer interregional flights, for example in the United States most of the flights are domestic. Interestingly it could aslo be seen that African airports do not form their own community and depend on other communities, while many of the Russian airports together with some from former Soviet Union countries form a separate air traffic community of their own.

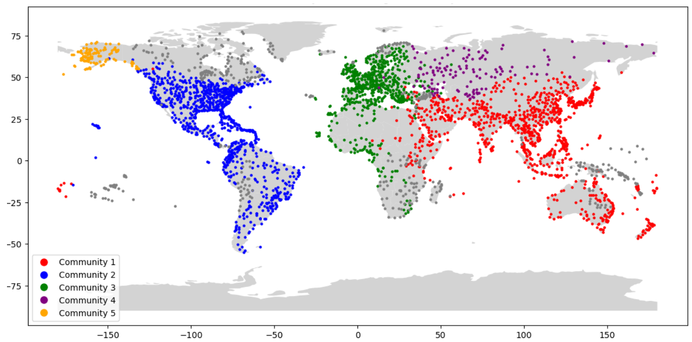

Fig. 2. Greedy modularity maximisation with five largest communities colored. Airports as nodes.

A graph with countries as nodes was created to get a better overview of the connectivity between countries rather than just between airports, which might often have the majority of the connections within a country and thus a similar analysis was conducted on a country scale. The least restricted country based on visa restriction weighted degree centrality was Germany, while the most restricted being Afghanistan, Turkmenistan and North Korea.

Based on PageRank score calculations, the three most connected countries within the air traffic network were France, Turkey and The United Kingdom, additional information in this regard can be seen in Table V.

As with previous analysis methods, also the community detection was performed on a country scale using the country graph and the same two algorithms, out of which once again the greedy modularity maximisation algorithm performed best (Fig. 3). The network communities form differently when airports are grouped together per country, since domestic flights are not included into the analysis any more. Results reveal that most European countries remain in one community while interestingly larger Western European countries like Great Britain and France are grouped together with overseas countries, highlighting their role as gateways in interregional travel. Interestingly The United States for example is not grouped together with European countries, but with countries from Asia, Oceania and parts of Northern Africa.

Air Traffic Risk Score

Countries: The potential risk resilience of the air traffic network was assessed for each country using the risk scores described before. The risk score was an equally-weighted average of three components: outgoing routes per capita, airports (with routes out of the country) per square kilometre, and the visa restriction score. A total of 1102 airports and

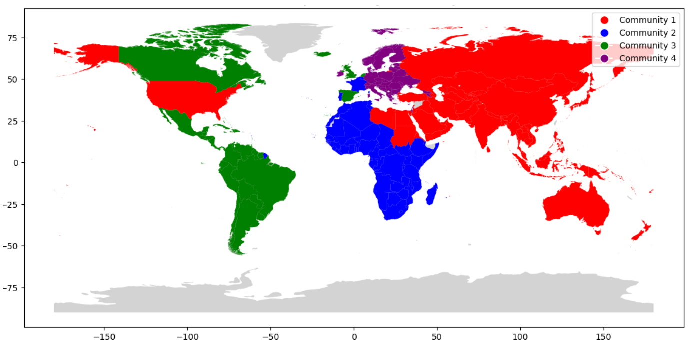

Fig. 3. Modularity maximisation. Countries as nodes.

TABLE II
DISTRIBUTION OF AIR TRAFFIC RISK SCORES BY INCOME GROUP

Income group	Countries	Average risk score	Low risk (<30)	Medium risk (30-60)	High risk (>60)
High	58	22.335	45	13	0
Upper-middle	52	38.905	18	25	9
Lower-middle	54	51.763	6	27	21
Low	25	61.887	0	11	14

18413 flight routes were included in the risk score calculations, as these were deemed valid and part of the set of international flights. The results reveal significant variability in risk scores across different countries worldwide (Fig. 4).

In each of the three risk score components, certain somewhat unexpected high performers stood out. For instance, some Arabian Peninsula countries excelled in outgoing routes per capita, some Caribbean countries had a high number of airports per square kilometre, and some Southern African countries scored positive low results on visa restrictions. However, when these components were averaged, the results showed a more predictable distribution. Countries with lower risk scores were mostly Western European nations or former British Empire colonies. However, due to the nature of the risk score formula calculation, the countries with the very lowest risk scores were still predominantly small island nations in the Oceania region. In contrast, countries with higher risk scores were often located in Northern and Central Africa, as well as in Central Asia. Detailed results for each country are provided in the appendix of the report.

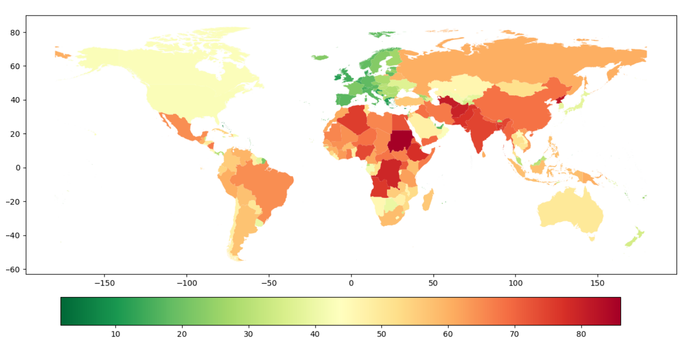

Fig. 4. Risk scores of the world countries

The air traffic risk score showed a significant positive correlation with the UN income group labels of the countries. Countries with higher average GDP per capita tend to have substantially lower air traffic risk scores compared to those with lower economic security (Table II).

World Regions: The potential risk resilience of the air traffic network at the regional level was assessed similarly to the process used for countries. However, for regions, all risk score component metrics were recalculated by treating each region as a single entity, akin to a fictional country. The polygons, areas, and populations of the world regions were determined by merging the corresponding countries based on their region labels. For the risk score assessment, 577 airports and 5788 flight routes were deemed valid, as they represented interregional flights.

Although the region-level risk score analysis indicated that the assessment logic is less effective for regions than for individual countries, it still provides valuable insights. Overall, the risk scores are lowest in the northern hemisphere, where population densities are lower and air traffic options are more diverse due to better transport infrastructure on average. In contrast, regions in the southern hemisphere tend to have higher population densities and a greater proportion of low-income countries, which is strongly associated with higher air traffic risk scores. The risk score is particularly high for Sub-Saharan Africa (91.033) and the East Asia and Pacific region (80.341) (Fig. 5 and Table III).

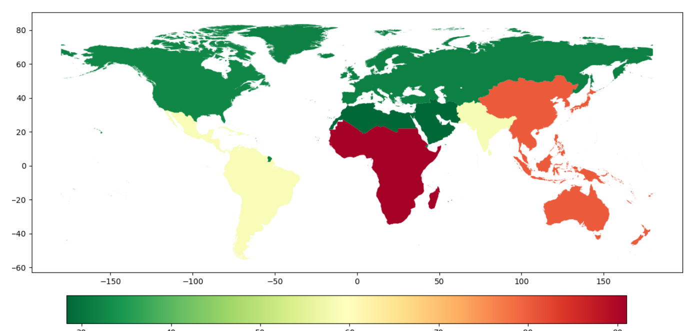

Fig. 5. Risk scores of the world regions

V. CONCLUSION

In conclusion, the analysis provided valuable insights into the spatial structure and resilience of the global air traffic network across various spatial scales. The findings revealed significant variability in air traffic risk scores among different countries and regions, highlighting safe areas of relatively high catastrophe resilience while also revealing areas with potential vulnerabilities that should be addressed. Notably, the analysis identified key airports and countries that play crucial roles in interregional connectivity, serving as gateways for international travel. The analysis also showed that the air

TABLE III
AIR TRAFFIC RISK SCORES OF WORLD REGIONS

Region	Average risk score	Population density	Low income countries %
East Asia & Pacific	80.341	92.090	3.12
Europe & Central Asia	31.531	29.951	0
Latin America & Caribbean	58.634	30.266	0
Middle East & North Africa	28.288	39.374	5.26
North America	32.274	18.232	0
South Asia	58.205	346.826	12.5
Sub-Saharan Africa	91.033	41.0135	45.83

traffic risk scores also tend to strongly correlate with economic wealth of the countries, but the causes and connections behind the risk score distributions are probably much more complex.

However, while the results are interesting and promising, they are not without limitations and flaws. The current methodology of network risk score assessment could be further refined by incorporating additional relevant parameters such as airplane capacity, flight frequency and the number of mainland neighbors to account for potential land evacuation routes for example. Moreover, the assessment metrics could be made more specific to different types of potential catastrophes by considering environmental variables, political stability, and many other potential risk factors. The risk score calculation should also be more balanced in the future so that it would be more resilient against outliers, which currently often happen to be small oceanic island countries. It also is important to note that due to high variability within world regions, risk score estimates at the regional level can give a biased end results and should be interpreted carefully in conjunction with risk estimates at finer scales.

Future research could benefit from these enhancements, offering a more comprehensive understanding of the air traffic network's robustness at various spatial scales. This would allow for more targeted strategies and informative future predictions to mitigate the negative impacts of potential disruptive events, ultimately contributing to a more resilient global air travel system.

REFERENCES

[1] J. Bowen, "Airline hubs in Southeast Asia: national economic development and nodal accessibility," *Journal of Transport Geography*, vol. 8, no. 1, pp. 25–41, Jan. 2000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0966692399000307

[2] H. N. Huynh, K. L. Ng, R. Toh, and L. Feng, "Understanding the impact of network structure on air travel pattern at different scales," *PLOS ONE*, vol. 19, no. 3, p. e0299897, Mar. 2024. [Online]. Available: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10923468/

[3] N. Scott, R. Baggio, and C. Cooper, "Network Analysis and Tourism: From Theory to Practice," Jan. 2009.

[4] J. Zhang, X.-B. Cao, W.-B. Du, and K.-Q. Cai, "Evolution of Chinese airport network," *Physica A*, vol. 389, no. 18, pp. 3922–3931, Sep. 2010.

[5] J. Lin and Y. Ban, "Complex Network Topology of Transportation Systems," *Transport Reviews*, vol. 33, no. 6, pp. 658–685, Nov. 2013, publisher: Routledge _eprint: https://doi.org/10.1080/01441647.2013.848955. [Online]. Available: https://doi.org/10.1080/01441647.2013.848955

[6] Y. Zhou, J. Wang, and G. Q. Huang, "Efficiency and robustness of weighted air transport networks," *Transportation Research Part E: Logistics and Transportation Review*, vol. 122, pp. 14–26, Feb. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1366554518311165

[7] R. Guimerà, S. Mossa, A. Turtschi, and L. A. N. Amaral, "The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles," *Proceedings of the National Academy of Sciences*, vol. 102, no. 22, pp. 7794–7799, May 2005, publisher: Proceedings of the National Academy of Sciences. [Online]. Available: https://www.pnas.org/doi/full/10.1073/pnas.0407994102

[8] T. K. Y. Cheung, C. W. H. Wong, and A. Zhang, "The evolution of aviation network: Global airport connectivity index 2006–2016," *Transportation Research Part E: Logistics and Transportation Review*, vol. 133, p. 101826, Jan. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1366554519301243

[9] P. Mazzarisi, S. Zaoli, F. Lillo, L. Delgado, and G. Gurtner, "New centrality and causality metrics assessing air traffic network interactions," *Journal of Air Transport Management*, vol. 85, p. 101801, Jun. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0969699719305307

[10] L. Siozos-Rousoulis, D. Robert, and W. Verbeke, "A study of the U.S. domestic air transportation network: temporal evolution of network topology and robustness from 2001 to 2016," *Journal of Transportation Security*, vol. 14, no. 1, pp. 55–78, Jun. 2021. [Online]. Available: https://doi.org/10.1007/s12198-020-00227-x

[11] "OpenFlights: Airport and airline data." [Online]. Available: https://openflights.org/data

[12] "World Bank Open Data." [Online]. Available: https://data.worldbank.org

[13] I. Ilyankou, "ilyankou/passport-index-dataset," Apr. 2024, original-date: 2017-06-01T19:23:47Z. [Online]. Available: https://github.com/ilyankou/passport-index-dataset

[14] "datasets/geo-countries," Apr. 2024, original-date: 2016-02-15T11:02:19Z. [Online]. Available: https://github.com/datasets/geo-countries

[15] "Taiwan Population 1950-2024." [Online]. Available: https://www.macrotrends.net/global-metrics/countries/TWN/taiwan/population

[16] "QuantileTransformer." [Online]. Available: https://scikit-learn/stable/modules/generated/sklearn.preprocessing.QuantileTransformer.html

[17] "List of busiest airports by passenger traffic (2010–2015)." [Online]. Available: https://en.wikipedia.org/wiki/List_of_busiest_airports_by_passenger_traffic_(2010%E2%80%932015)

[18] "Atatürk Airport." [Online]. Available: https://en.wikipedia.org/wiki/Atat%C3%BCrk_Airport

[19] "Chicago O'Hare Airport named 'World's Best-Connected Airport' on 2022 Megahub Connectivity Index," *ABC7 Chicago*. [Online]. Available: https://abc7chicago.com/chicago-ohare-international-airport-worlds-best-connected-flights/12247801/

SOURCE CODE AND DATA

https://github.com/khtamkivi/project_OpenFlights/tree/main

Appendix

TABLE IV
TOP 10 AIRPORTS BY VISA WEIGHTED DEGREE CENTRALITY

Degree Centrality	Name	City	IATA
1.000000	Denver International Airport	Denver	DEN
0.993827	Munich Airport	Munich	MUC
0.901235	Barcelona International Airport	Barcelona	BCN
0.876543	Atatürk International Airport	Istanbul	ISL
0.814815	Adolfo Suárez Madrid–Barajas Airport	Madrid	MAD
0.790123	Leonardo da Vinci–Fiumicino Airport	Rome	FCO
0.787037	McCarran International Airport	Las Vegas	LAS
0.783951	Minneapolis-St Paul International/Wold-Chamberlain Airport	Minneapolis	MSP
0.762346	Palma De Mallorca Airport	Palma de Mallorca	PMI
0.731481	Brussels Airport	Brussels	BRU

TABLE V
PAGERANK OF COUNTRIES IN NETWORK

Country	PageRank
France	0.023253
Turkey	0.021358
United Kingdom	0.021100
Germany	0.020178
United Arab Emirates	0.020088
United States	0.018910
Italy	0.017483
Netherlands	0.017207
Spain	0.016275
Russia	0.015626

TABLE VI: AIR TRAFFIC RISK SCORES OF WORLD COUNTRIES

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Afghanistan	AFG	South Asia	Low	32716210	0.17	0.265	0.825	79.662	high
Angola	AGO	Sub-Saharan Africa	Lower middle	27128337	0.221	0.12	0.625	76.128	high
Albania	ALB	Europe & Central Asia	Upper middle	2889104	0.569	0.58	0.056	30.193	medium
United Arab Emirates	ARE	Middle East & North Africa	High	8835951	0.888	0.734	0.141	17.295	low
Argentina	ARG	Latin America & Caribbean	Upper middle	42669500	0.262	0.151	0.122	56.988	medium
Armenia	ARM	Europe & Central Asia	Upper middle	2889930	0.695	0.727	0.235	27.136	low
Antigua and Barbuda	ATG	Latin America & Caribbean	High	89236	0.99	0.938	0.227	9.96	low

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Australia	AUS	East Asia & Pacific	High	23475686	0.553	0.102	0.153	49.942	medium
Austria	AUT	Europe & Central Asia	High	8546356	0.779	0.737	0.049	17.768	low
Azerbaijan	AZE	Europe & Central Asia	Upper middle	9535079	0.515	0.668	0.318	37.822	medium
Burundi	BDI	Sub-Saharan Africa	Low	10494913	0.101	0.597	0.0	43.375	medium
Belgium	BEL	Europe & Central Asia	High	11209057	0.789	0.816	0.044	14.65	low
Benin	BEN	Sub-Saharan Africa	Lower middle	10614844	0.355	0.352	0.194	49.596	medium
Burkina Faso	BFA	Sub-Saharan Africa	Low	18169842	0.209	0.312	0.267	58.211	medium
Bangladesh	BGD	South Asia	Lower middle	155961299	0.094	0.434	0.819	76.377	high
Bulgaria	BGR	Europe & Central Asia	Upper middle	7223938	0.615	0.585	0.053	28.395	low
Bahrain	BHR	Middle East & North Africa	High	1311134	0.842	0.916	0.362	20.16	low
Bahamas	BHS	Latin America & Caribbean	High	389131	0.972	0.886	0.041	6.08	low
Bosnia and Herzegovina	BIH	Europe & Central Asia	Upper middle	3571068	0.492	0.703	0.0	26.835	low
Belarus	BLR	Europe & Central Asia	Upper middle	9448515	0.498	0.24	0.524	59.573	medium
Belize	BLZ	Latin America & Caribbean	Upper middle	352335	0.869	0.65	0.538	33.981	medium
Bolivia	BOL	Latin America & Caribbean	Lower middle	10916987	0.283	0.134	0.143	57.522	medium
Brazil	BRA	Latin America & Caribbean	Upper middle	203459650	0.161	0.131	0.235	64.785	high
Barbados	BRB	Latin America & Caribbean	High	277493	0.921	0.94	0.147	9.55	low
Brunei	BRN	East Asia & Pacific	High	416656	0.849	0.849	0.038	11.362	low
Bhutan	BTN	South Asia	Lower middle	736357	0.59	0.535	0.0	29.162	low
Botswana	BWA	Sub-Saharan Africa	Upper middle	2260376	0.524	0.298	0.0	39.283	medium
Central African Republic	CAF	Sub-Saharan Africa	Low	4798734	0.177	0.121	0.333	67.841	high
Canada	CAN	North America	High	35437435	0.645	0.142	0.077	42.986	medium
Switzerland	CHE	Europe & Central Asia	High	8188649	0.839	0.779	0.077	15.269	low
Chile	CHL	Latin America & Caribbean	High	17687108	0.369	0.333	0.136	47.812	medium
China	CHN	East Asia & Pacific	Upper middle	1371860000	0.117	0.272	0.44	68.376	high
Cote d'Ivoire	CIV	Sub-Saharan Africa	Lower middle	22995555	0.237	0.176	0.31	63.211	high
Cameroon	CMR	Sub-Saharan Africa	Lower middle	22299585	0.256	0.274	0.375	61.524	high
Congo (Kinshasa)	COD	Sub-Saharan Africa	Low	76035588	0.08	0.091	0.545	79.137	high
Congo (Brazzaville)	COG	Sub-Saharan Africa	Lower middle	4944861	0.474	0.256	0.289	51.962	medium
Colombia	COL	Latin America & Caribbean	Upper middle	46677947	0.302	0.361	0.328	55.494	medium
Comoros	COM	Sub-Saharan Africa	Lower middle	714612	0.667	0.881	0.071	17.479	low
Cape Verde	CPV	Sub-Saharan Africa	Lower middle	546076	0.902	0.9	0.846	34.821	medium
Costa Rica	CRI	Latin America & Caribbean	Upper middle	4844288	0.62	0.615	0.615	46.012	medium
Cuba	CUB	Latin America & Caribbean	Upper middle	11332026	0.511	0.741	0.875	54.081	medium
Cyprus	CYP	Europe & Central Asia	High	1176995	0.934	0.878	0.047	7.799	low

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Czech Republic	CZE	Europe & Central Asia	High	10525347	0.658	0.716	0.094	24.028	low
Germany	DEU	Europe & Central Asia	High	80982500	0.74	0.764	0.046	18.083	low
Djibouti	DJI	Middle East & North Africa	Lower middle	989087	0.697	0.658	0.5	38.145	medium
Dominica	DMA	Latin America & Caribbean	Upper middle	69371	0.97	0.907	0.0	4.094	low
Denmark	DNK	Europe & Central Asia	High	5643475	0.836	0.804	0.045	13.479	low
Dominican Republic	DOM	Latin America & Caribbean	Upper middle	10282115	0.672	0.837	0.767	41.902	medium
Algeria	DZA	Middle East & North Africa	Lower middle	38760168	0.404	0.218	0.89	75.606	high
Ecuador	ECU	Latin America & Caribbean	Upper middle	15957994	0.343	0.414	0.385	54.252	medium
Egypt	EGY	Middle East & North Africa	Lower middle	95592324	0.328	0.367	0.873	72.605	high
Eritrea	ERI	Sub-Saharan Africa	Low	3323425	0.397	0.334	0.857	70.858	high
Spain	ESP	Europe & Central Asia	High	46480882	0.808	0.692	0.025	17.52	low
Estonia	EST	Europe & Central Asia	High	1314545	0.791	0.644	0.036	20.05	low
Ethiopia	ETH	Sub-Saharan Africa	Low	99746766	0.157	0.131	0.614	77.527	high
Finland	FIN	Europe & Central Asia	High	5461512	0.752	0.526	0.08	26.729	low
Fiji	FJI	East Asia & Pacific	Upper middle	915560	0.818	0.797	0.364	24.929	low
France	FRA	Europe & Central Asia	High	66312067	0.708	0.717	0.087	22.09	low
Micronesia	FSM	East Asia & Pacific	Lower middle	109024	0.895	0.971	0.0	4.485	low
Gabon	GAB	Sub-Saharan Africa	Upper middle	1966855	0.632	0.322	0.206	41.721	medium
United Kingdom	GBR	Europe & Central Asia	High	64602298	0.777	0.825	0.044	14.753	low
Georgia	GEO	Europe & Central Asia	Upper middle	3719414	0.641	0.641	0.088	26.888	low
Ghana	GHA	Sub-Saharan Africa	Lower middle	28196358	0.243	0.205	0.462	67.135	high
Guinea	GIN	Sub-Saharan Africa	Lower middle	11333365	0.195	0.197	0.188	59.839	medium
Gambia	GMB	Sub-Saharan Africa	Low	2189019	0.457	0.777	0.429	39.817	medium
Guinea-Bissau	GNB	Sub-Saharan Africa	Low	1743309	0.442	0.556	0.2	40.075	medium
Equatorial Guinea	GNQ	Sub-Saharan Africa	Upper middle	1295183	0.566	0.748	0.25	31.177	medium
Greece	GRC	Europe & Central Asia	High	10892413	0.858	0.829	0.028	11.364	low
Grenada	GRD	Latin America & Caribbean	Upper middle	117972	0.931	0.957	0.375	16.256	low
Guatemala	GTM	Latin America & Caribbean	Upper middle	15306316	0.278	0.485	0.579	60.547	high
Guyana	GUY	Latin America & Caribbean	High	751115	0.686	0.377	0.5	47.916	medium
Hong Kong	HKG	East Asia & Pacific	High	7229500	0.757	0.898	0.462	26.923	low
Honduras	HND	Latin America & Caribbean	Lower middle	9127846	0.446	0.587	0.519	49.505	medium
Croatia	HRV	Europe & Central Asia	High	4238389	0.849	0.812	0.008	11.565	low
Haiti	HTI	Latin America & Caribbean	Lower middle	10412740	0.263	0.747	0.667	55.231	medium
Hungary	HUN	Europe & Central Asia	High	9866468	0.618	0.509	0.032	30.138	medium
Indonesia	IDN	East Asia & Pacific	Upper middle	256229761	0.108	0.378	0.223	57.892	medium

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
India	IND	South Asia	Lower middle	1307246509	0.084	0.311	0.634	74.636	high
Ireland	IRL	Europe & Central Asia	High	4657740	0.901	0.789	0.032	11.41	low
Iran	IRN	Middle East & North Africa	Lower middle	79961672	0.267	0.413	0.697	67.229	high
Iraq	IRQ	Middle East & North Africa	Upper middle	36746488	0.334	0.405	0.798	68.644	high
Iceland	ISL	Europe & Central Asia	High	327386	0.961	0.496	0.136	22.641	low
Israel	ISR	Middle East & North Africa	High	8215700	0.65	0.656	0.077	25.683	low
Italy	ITA	Europe & Central Asia	High	60789140	0.717	0.798	0.034	17.281	low
Jamaica	JAM	Latin America & Caribbean	Upper middle	2784543	0.737	0.849	0.698	37.066	medium
Jordan	JOR	Middle East & North Africa	Lower middle	8658026	0.576	0.517	0.736	54.775	medium
Japan	JPN	East Asia & Pacific	High	127276000	0.394	0.757	0.3	38.294	medium
Kazakhstan	KAZ	Europe & Central Asia	Upper middle	17288285	0.504	0.227	0.122	46.389	medium
Kenya	KEN	Sub-Saharan Africa	Lower middle	45831863	0.321	0.288	0.341	57.717	medium
Kyrgyzstan	KGZ	Europe & Central Asia	Lower middle	5835500	0.556	0.385	0.091	38.324	medium
Cambodia	KHM	East Asia & Pacific	Lower middle	15210817	0.395	0.398	0.297	50.129	medium
Kiribati	KIR	East Asia & Pacific	Lower middle	114985	0.888	0.931	0.4	19.366	low
Saint Kitts and Nevis	KNA	Latin America & Caribbean	High	47789	0.983	0.973	0.111	5.185	low
South Korea	KOR	East Asia & Pacific	High	50746659	0.477	0.759	0.394	38.618	medium
Kuwait	KWT	Middle East & North Africa	High	3761584	0.727	0.699	0.357	31.059	medium
Laos	LAO	East Asia & Pacific	Lower middle	6691454	0.434	0.466	0.25	44.994	medium
Lebanon	LBN	Middle East & North Africa	Lower middle	6274342	0.601	0.787	0.656	42.268	medium
Liberia	LBR	Sub-Saharan Africa	Low	4519398	0.36	0.505	0.25	46.175	medium
Libya	LBY	Middle East & North Africa	Upper middle	6097764	0.526	0.172	0.667	65.637	high
Saint Lucia	LCA	Latin America & Caribbean	Upper middle	174804	0.956	0.961	0.353	14.556	low
Sri Lanka	LKA	South Asia	Lower middle	21239457	0.381	0.553	0.744	60.319	high
Lesotho	LSO	Sub-Saharan Africa	Lower middle	2095242	0.132	0.564	0.0	43.452	medium
Lithuania	LTU	Europe & Central Asia	High	2932367	0.808	0.667	0.029	18.47	low
Luxembourg	LUX	Europe & Central Asia	High	556319	0.955	0.865	0.0	6.011	low
Latvia	LVA	Europe & Central Asia	High	1993782	0.862	0.446	0.051	24.781	low
Macau	MAC	East Asia & Pacific	High	604167	0.924	0.986	0.026	3.879	low
Morocco	MAR	Middle East & North Africa	Lower middle	34248603	0.559	0.456	0.852	61.225	high
Moldova	MDA	Europe & Central Asia	Upper middle	2857815	0.634	0.553	0.12	31.125	medium
Madagascar	MDG	Sub-Saharan Africa	Low	24215976	0.195	0.386	0.265	56.126	medium
Maldives	MDV	South Asia	Upper middle	416738	0.919	0.99	0.1	6.37	low
Mexico	MEX	Latin America & Caribbean	Upper middle	118755887	0.371	0.478	0.786	64.562	high
Marshall Islands	MHL	East Asia & Pacific	Upper middle	50419	0.96	0.98	0.0	1.993	low

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Macedonia	MKD	Europe & Central Asia	Upper middle	2067471	0.685	0.619	0.091	26.213	low
Mali	MLI	Sub-Saharan Africa	Low	17551814	0.236	0.081	0.25	64.433	high
Malta	MLT	Middle East & North Africa	High	434558	0.981	0.96	0.059	3.956	low
Burma	MMR	East Asia & Pacific	Lower middle	51072436	0.126	0.233	0.458	69.964	high
Montenegro	MNE	Europe & Central Asia	Upper middle	621810	0.832	0.838	0.111	14.704	low
Mongolia	MNG	East Asia & Pacific	Lower middle	2902823	0.454	0.105	0.111	51.763	medium
Mozambique	MOZ	Sub-Saharan Africa	Low	26038704	0.171	0.364	0.125	52.988	medium
Mauritania	MRT	Sub-Saharan Africa	Lower middle	3843174	0.415	0.168	0.556	65.737	high
Mauritius	MUS	Sub-Saharan Africa	Upper middle	1261208	0.764	0.875	0.083	14.816	low
Malawi	MWI	Sub-Saharan Africa	Low	16477966	0.141	0.454	0.0	46.807	medium
Malaysia	MYS	East Asia & Pacific	Upper middle	30606459	0.514	0.658	0.096	30.793	medium
Namibia	NAM	Sub-Saharan Africa	Upper middle	2243001	0.543	0.16	0.083	46.028	medium
Niger	NER	Sub-Saharan Africa	Low	19372014	0.113	0.089	0.25	68.257	high
Nigeria	NGA	Sub-Saharan Africa	Lower middle	179379016	0.099	0.225	0.542	73.907	high
Nicaragua	NIC	Latin America & Caribbean	Lower middle	6208676	0.285	0.324	0.625	67.201	high
Netherlands	NLD	Europe & Central Asia	High	16865008	0.773	0.821	0.052	15.302	low
Norway	NOR	Europe & Central Asia	High	5137232	0.898	0.568	0.017	18.323	low
Nepal	NPL	South Asia	Lower middle	27462106	0.21	0.284	0.587	69.778	high
Nauru	NRU	East Asia & Pacific	High	10940	0.993	0.995	0.167	5.958	low
New Zealand	NZL	East Asia & Pacific	High	4516500	0.607	0.439	0.088	34.73	medium
Oman	OMN	Middle East & North Africa	High	4009267	0.713	0.273	0.277	43.027	medium
Pakistan	PAK	South Asia	Lower middle	208251628	0.141	0.42	0.871	77.017	high
Panama	PAN	Latin America & Caribbean	High	3888793	0.764	0.628	0.284	29.714	low
Peru	PER	Latin America & Caribbean	Upper middle	30353951	0.298	0.174	0.28	60.274	high
Philippines	PHL	East Asia & Pacific	Lower middle	101325201	0.19	0.525	0.434	57.314	medium
Palau	PLW	East Asia & Pacific	Upper middle	17796	1.0	0.928	0.25	10.739	low
Papua New Guinea	PNG	East Asia & Pacific	Lower middle	8464153	0.307	0.215	0.333	60.386	high
Poland	POL	Europe & Central Asia	High	38011735	0.606	0.576	0.014	27.758	low
North Korea	PRK	East Asia & Pacific	Low	25126131	0.082	0.335	0.875	81.943	high
Portugal	PRT	Europe & Central Asia	High	10401062	0.814	0.69	0.03	17.536	low
Paraguay	PRY	Latin America & Caribbean	Upper middle	6090721	0.345	0.243	0.1	50.381	medium
Qatar	QAT	Middle East & North Africa	High	2214465	0.912	0.771	0.422	24.625	low
Romania	ROU	Europe & Central Asia	High	19908979	0.536	0.606	0.005	28.789	low
Russia	RUS	Europe & Central Asia	Upper middle	143819667	0.483	0.159	0.451	60.284	high
Rwanda	RWA	Sub-Saharan Africa	Low	11368451	0.293	0.622	0.333	47.309	medium

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Saudi Arabia	SAU	Middle East & North Africa	High	32125564	0.594	0.255	0.283	47.802	medium
Sudan	SDN	Sub-Saharan Africa	Low	37003245	0.15	0.099	0.825	85.886	high
Senegal	SEN	Sub-Saharan Africa	Lower middle	13970308	0.406	0.245	0.5	61.65	high
Singapore	SGP	East Asia & Pacific	High	5469724	0.803	0.925	0.093	12.157	low
Solomon Islands	SLB	East Asia & Pacific	Lower middle	597375	0.586	0.596	0.25	35.625	medium
Sierra Leone	SLE	Sub-Saharan Africa	Low	7140688	0.332	0.43	0.364	53.393	medium
El Salvador	SLV	Latin America & Caribbean	Upper middle	6209526	0.535	0.678	0.5	42.912	medium
Somalia	SOM	Sub-Saharan Africa	Low	13309235	0.229	0.348	0.583	66.867	high
Serbia	SRB	Europe & Central Asia	Upper middle	7130576	0.654	0.538	0.059	28.916	low
South Sudan	SSD	Sub-Saharan Africa	Low	11213284	0.176	0.119	0.286	66.353	high
Sao Tome and Principe	STP	Sub-Saharan Africa	Lower middle	197497	0.732	0.898	0.5	28.989	low
Suriname	SUR	Latin America & Caribbean	Upper middle	569682	0.699	0.295	0.143	38.289	medium
Slovakia	SVK	Europe & Central Asia	High	5418649	0.47	0.633	0.0	29.896	low
Slovenia	SVN	Europe & Central Asia	High	2061980	0.687	0.679	0.023	21.902	low
Sweden	SWE	Europe & Central Asia	High	9696110	0.786	0.566	0.03	22.61	low
Swaziland	SWZ	Sub-Saharan Africa	Lower middle	1125865	0.222	0.707	0.0	35.686	medium
Seychelles	SYC	Sub-Saharan Africa	High	91359	0.944	0.941	0.062	5.913	low
Chad	TCD	Sub-Saharan Africa	Low	13697126	0.187	0.081	0.222	65.159	high
Togo	TGO	Sub-Saharan Africa	Low	7288383	0.378	0.469	0.286	47.944	medium
Thailand	THA	East Asia & Pacific	Upper middle	69960943	0.429	0.465	0.443	51.644	medium
Tajikistan	TJK	Europe & Central Asia	Lower middle	8326348	0.576	0.506	0.123	34.7	medium
Turkmenistan	TKM	Europe & Central Asia	Upper middle	5663152	0.384	0.149	0.955	80.73	high
East Timor	TLS	East Asia & Pacific	Lower middle	1184830	0.423	0.723	0.167	34.039	medium
Tonga	TON	East Asia & Pacific	Upper middle	106626	0.873	0.914	0.375	19.596	low
Trinidad and Tobago	TTO	Latin America & Caribbean	High	1450661	0.743	0.866	0.348	24.615	low
Tunisia	TUN	Middle East & North Africa	Lower middle	11428948	0.667	0.608	0.821	51.566	medium
Turkey	TUR	Europe & Central Asia	Upper middle	77181884	0.546	0.494	0.72	55.999	medium
Tuvalu	TUV	East Asia & Pacific	Upper middle	10899	0.947	1.0	0.0	1.768	low
Tanzania	TZA	Sub-Saharan Africa	Lower middle	50814552	0.151	0.21	0.232	62.362	high
Uganda	UGA	Sub-Saharan Africa	Low	36336539	0.125	0.2	0.471	71.521	high
Ukraine	UKR	Europe & Central Asia	Lower middle	45272155	0.444	0.485	0.03	36.711	medium
Uruguay	URY	Latin America & Caribbean	High	3391662	0.488	0.402	0.071	39.392	medium
United States	USA	North America	High	318386329	0.462	0.293	0.04	42.839	medium
Uzbekistan	UZB	Europe & Central Asia	Lower middle	30757700	0.424	0.533	0.196	41.288	medium
Saint Vincent and the Grenadines	VCT	Latin America & Caribbean	Upper middle	106912	0.872	0.953	0.0	5.84	low

TABLE VI: (continued)

Country	ISO	Region	IncGroup	Pop2014	ORPCN	AN	VR	RS	RiskClass
Venezuela	VEN	Latin America & Caribbean		30193258	0.316	0.359	0.456	59.366	medium
Vietnam	VNM	East Asia & Pacific	Lower middle	91235504	0.222	0.444	0.537	62.366	high
Vanuatu	VUT	East Asia & Pacific	Lower middle	269927	0.823	0.844	0.357	22.998	low
Samoa	WSM	East Asia & Pacific	Lower middle	201757	0.726	0.859	0.667	36.059	medium
Yemen	YEM	Middle East & North Africa	Low	27753304	0.349	0.424	0.859	69.507	high
South Africa	ZAF	Sub-Saharan Africa	Upper middle	54729551	0.313	0.183	0.241	58.154	medium
Zambia	ZMB	Sub-Saharan Africa	Lower middle	15737793	0.273	0.192	0.132	55.566	medium
Zimbabwe	ZWE	Sub-Saharan Africa	Lower middle	13855753	0.244	0.323	0.0	47.779	medium