

GLOBAL MIGRATION DYNAMICS: VISUALIZING PATTERNS AND CONTEXTS

mercury.swin.edu.au/cos30045/s104221423/WorldMigrationProject/

Team member: Hoang Minh Khue, Le Tran Bao Kien

Tutor Date: Monday morning, 9 AM – 12 PM

Spring 2024

Word Count:

Hoang Minh Khue – Le Tran Bao Kien
S104221423 – s104223584

Table of Contents

1. Introduction.....	2
1.1. Background and Motivation	2
1.2. Visualization Purpose	2
1.3. Project Schedule	3
2. Data.....	4
2.1. Data Source	4
2.2. Data Processing	5
Hand Processing:	5
World Bank(www.worldbank.org):.....	6
United Nations (www.un.org):	9
Code Processing:.....	12
Additional files:	14
3. Requirements	14
3.1. Must-have Features:	14
3.2. Optional Features:	14
4. Visualization Design:	15
4.1. Conceptualize (Design Iteration) stage:	15
4.2. Visualize Stage:	16
4.3. Materialize Stage:	19
4.4. Implementation Stage:.....	20
5. Final Design	23
6. Conclusion	24
References.....	31

1. Introduction

1.1. Background and Motivation

One of the key characteristics of human history has been migration, which has shaped civilizations, economies, and cultures on many continents and during many eras (Siddiqui, T., 2012). People have moved about throughout history as a necessary component of human existence, from the ancient nomadic tribes to the contemporary diasporas fueled by globalization. The topic of global migration has become increasingly important and well-known in the modern era due to a variety of causes, including political unrest, economic inequality, environmental damage, and climate change (Hammar, T., Brochmann, G., Tamas, K., & Faist, T., 2021).

In a time of growing globalization and interconnection, the dynamics of human mobility are more intricate and nuanced than in the past. There are many different reasons why people move, such as to pursue better job prospects, further their education, be with family, or seek safety from persecution or violence (Castles, 2010). Furthermore, population movements are driven by environmental causes including natural catastrophes, environmental degradation, and climate change, which exacerbate already-existing vulnerabilities and inequities.

It is impossible to exaggerate the importance of the global migration issue in the modern world. As the number of individuals living in forced displacement reaches previously unheard-of heights, migration has become a significant humanitarian, social, political, and economic concern. It has significant ramifications for host communities, transit nations, and international geopolitics and has a significant impact on the rights and well-being of migrants (Faist, 2013). Furthermore, a wide range of other urgent challenges, including urbanization, sustainable development, labor markets, social cohesion, and human rights, are intertwined with migration (Castles, 2010).

1.2. Visualization Purpose

It is becoming more and more necessary to visualize migration patterns and contexts due to the complexity and difficulties involved in global migration (Zambotti, G., Guan, W., & Gest, J. D. 2015) . A potent tool for deciphering the complex dynamics of migratory flows, exposing temporal and spatial trends, finding patterns and correlations, and locating underlying causes and effects is data visualization (Bertin, 2011). Visualization helps stakeholders, policymakers, scholars, and the public better understand the scope and complexity of migration phenomena by converting complex statistics into understandable and visually appealing representations (van Biljon, J., & Osei-Bryson, K. M. 2020).

Furthermore, migration data visualization acts as a catalyst for comprehending and resolving the underlying reasons and effects of migration (Sander, N., Abel, G. J., Bauer, R., & Schmidt, J. 2014). It offers insightful information about the geographic distribution of migrants, their travel routes and corridors, the demographics of migrant communities, and the social, economic, and environmental settings in which migration takes place. Equipped with such discernments, policymakers can devise empirically grounded tactics and measures to tackle the obstacles presented by migration, alleviate its

adverse effects, and capitalize on its prospective advantages for both migrant populations and receiving communities.

To summarize, the visualization of migration patterns and contexts serves two purposes: first, it improves comprehension of the intricate dynamics of global migration; second, it provides guidance and information for policy and decision-making processes that aim to address the obstacles and maximize the benefits associated with human mobility. Through the effective use of visualization, we can shed light on the paths and histories of migration, give voice to the opinions and experiences of migrants, and help create more just inclusive, and sustainable societies in a global community that is becoming more interconnected by the day.

1.3. Project Schedule

Week 2-4: Project Kickoff and Data Collection

- Define project objectives and scope.
- Assign roles and responsibilities within the team.
- Identify and gather relevant data sources on global migration.
- Begin initial exploration and assessment of the collected data.

Week 5-6: Data Cleaning and Processing

- Conduct thorough data cleaning to address missing values, inconsistencies, and outliers.
- Standardize data formats and ensure data compatibility.
- Perform any necessary data transformations or aggregations to prepare the data for visualization.

Week 7-8: Design and Prototyping

- Brainstorm visualization ideas and design concepts.
- Develop initial prototypes of visualization layouts and elements.
- Gather feedback from team members and stakeholders to refine the design.

Week 9-10: Implementation

- Translate finalized design concepts into code using chosen visualization tools or libraries.
- Integrate cleaned and processed data into the visualization.
- Conduct iterative testing and debugging to ensure functionality and usability.

Week 11: Refinement and Evaluation

- Fine-tune visual elements, layouts, and interactions based on user feedback.
- Conduct thorough testing of the completed visualization across different devices and browsers.
- Evaluate the effectiveness and usability of visualization in conveying migration patterns and contexts.

Week 11-12: Documentation and Presentation

- Compile comprehensive documentation of the project process, including data sources, cleaning and processing methods, design decisions, and implementation details.
- Prepare a final presentation summarizing key findings, insights, and implications of the visualization.
- Submit the completed project process book and present the visualization to the instructor and class.

2. Data

2.1. Data Source

The primary sources for raw data in this project are the World Bank ([worldbank.org](http://www.worldbank.org)) and the United Nations (un.org). These reputable organizations provide comprehensive datasets related to global migration, offering insights into migration trends, demographic characteristics, and socio-economic contexts across countries and regions worldwide.

- **The World Bank** (www.worldbank.org):
 - The World Bank offers a wealth of demographic and socio-economic data, including information on international migration patterns, immigrant populations, and migration-related variables such as country of birth, citizenship status, and migration flows. Data from the World Bank's Migration and Remittances dataset and other sources provide valuable insights into migration trends within countries and globally.
 - **Data Set 1:** Life Expectancy data from the World Bank's World Development Indicators dataset.
 1. **Data Type:** The data set includes the following variables:
 - i. Country Name (Categorical)
 - ii. Life Expectancy (Quantitative)
 - **Data Set 2:** GDP per Capita data from the World Bank's World Development Indicators dataset.
 1. **Data Type:** The data set includes the following variables:
 - i. Country Name (Categorical)
 - ii. GDP per Capita (Quantitative)
 - **Data Set 3:** Population data from the World Bank's World Development Indicators dataset.

1. **Data Type:** The data set includes the following variables:
 - i. Country Name (Categorical)
 - ii. Population (Quantitative)
- **File Type:** Raw data obtained in .csv format.
- **United Nations (www.un.org):**
 - The United Nations is a leading source of data and research on global migration, providing comprehensive statistics and analyses on international migration trends, refugee movements, asylum seekers, internally displaced persons (IDPs), and other migration-related topics. Datasets from UN agencies such as the International Organization for Migration (IOM), the United Nations High Commissioner for Refugees (UNHCR), and the United Nations Department of Economic and Social Affairs (UN DESA) offer valuable insights into the scale, dynamics, and impacts of migration worldwide.
 - **Data Set:** Global Migration Statistics from the United Nations
 - **Data Type:** The data set includes the following variables:
 1. Country Name (Categorical)
 2. Number of Immigration People (Quantitative)
 - **File Type:** Raw data obtained in .xlsx format, converted to .csv format using www.cloudconvert.com/xlsx-to-csv website.

These two main data sources serve as foundational pillars for this project, providing reliable and authoritative data that underpins the visualization of global migration patterns and contexts. Additional supplementary data sources may be explored to enrich the analysis and address specific research questions or objectives identified during the project's development.

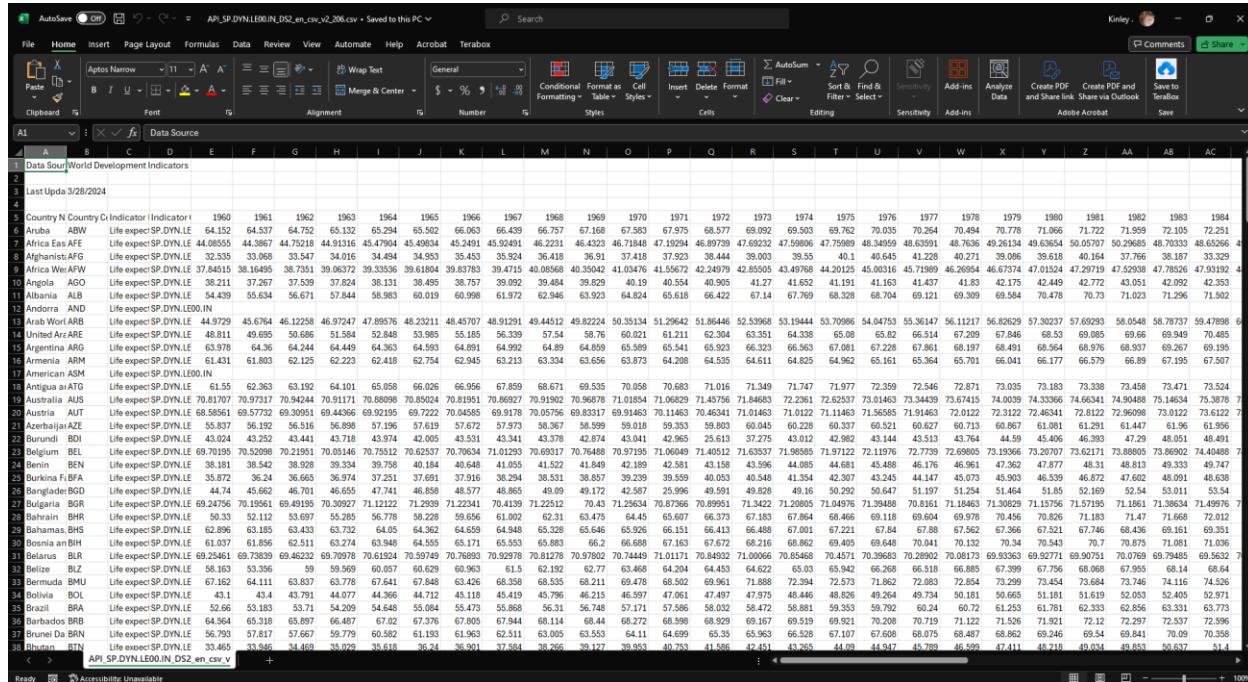
2.2. Data Processing

Hand Processing:

Although the datasets obtained from the organizations guarantee high integrity and timeliness, data cleaning and reorganization are essential to achieve the desired level of "conformity" necessary for visualization in D3.js. The data processing procedure involves distinct steps for each of the retrieved datasets, as each one necessitates pruning, pivoting, and aggregating to a certain degree. Once the design selection is finalized, the dataset will be further filtered to align with the specifications of the data visualization design, marking the milestone of "conformity" and indicating that the data is prepared for visualization.

World Bank(www.worldbank.org):

The original data is structured as a CSV (Comma-Separated Values) file, with the first dataset for **Life Expectancy** containing rows representing different countries or regions and columns representing attributes such as the name of the country, country code, indicator (in this case, life expectancy at birth), and numerical values indicating life expectancy for each year. The second dataset was for **GDP per capita (Current US\$)**, in which each row represents a country, and the columns include the country name, country code, indicator name (GDP per capita), indicator code, and GDP per capita values for different years. The years range from at least 1960 to 2022, with missing data represented by empty fields. The final dataset from the World Bank is for the **Total Population** data. Each row represents a country or region, with columns indicating the name of the country or region, the country code, the type of data (in this case, total population), and population values for different years. The years from 1960 to a more recent year, likely reflect the most up-to-date available data. The population values are in numeric format. Additionally, there are rows representing aggregates such as "World" and "Upper middle income.". The following three figures depict the three original datasets downloaded from the World Bank.



The screenshot shows the Microsoft Excel interface with the 'Data Source' sheet of the 'World Development Indicators' CSV file. The table has 24 columns and 2,000+ rows. The columns are labeled A through Z, with some columns (e.g., C, I, O, Q, R, S, T, U, V, W, X, Y, Z, AA, AB, AC) being merged into a single header 'Data Source'. The data includes columns for Country, Country Code, Indicator, Indicator Name, and various years from 1960 to 2022. The 'Indicator' column contains values like 'Life expectancy', 'GDP per capita (Current US\$)', and 'Total population'. The 'Indicator Name' column provides more detail, such as 'Life expectancy at birth', 'GDP per capita (constant 2010 US\$)', and 'Population total'. The 'Indicator Code' column contains codes like 'SP.DYN.LE00.IN.DS2_en.csv_v2_206' and 'SP.DYN.LE00.IN.DS2_en.csv_v'. The 'Value' column contains the actual data points for each country and year.

FIGURE 1: LIFE EXPECTANCY AT BIRTH FROM WORLD BANK.ORG

Country	Country	Indicator	Indicator	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984				
Aruba	ABW	GDP per capita, Nominal	GDP per capita, Nominal	141,386	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Africa, East	AZE	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Afghanistan	AFG	GDP per capita, Nominal	GDP per capita, Nominal	62,4437	60,9506	82,0217	85,51107	105,2432	143,1032	167,1657	134,0128	134,2504	162,6422	166,2248	141,3653	149,7441	181,5988	194,669	205,6741	232,7767	255,0456	284,7555	291,6498	311,8536	322,7767	342,7767	362,7767	382,7767	402,7767	422,7767		
Africa, West	AFW	GDP per capita, Nominal	GDP per capita, Nominal	107,0537	112,1284	117,8147	122,3701	130,7003	137,3018	143,0126	127,3233	128,3668	142,3275	193,5763	167,637	198,4034	243,8755	300,0787	374,0155	439,8321	450,0741	477,194	577,5416	709,8489	129,911	112,5904	805,0335	649,0258	637,9036			
Angola	AGO	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Albania	ALB	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Andorra	AND	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Arab World	ARB	GDP per capita, Nominal	GDP per capita, Nominal	303,3184	321,2694	353,0987	396,7319	450,8064	566,8916	1046,2064	120,2739	136,546	150,1662	205,1279	219,0849	257,1211	277,2208	299,259	321,2694	351,0987	396,7319	450,8064	566,8916	1046,2064	120,2739	136,546	150,1662	205,1279	219,0849			
United Arab Emirates	ARE	GDP per capita, Nominal	GDP per capita, Nominal	303,3184	321,2694	353,0987	396,7319	450,8064	566,8916	1046,2064	120,2739	136,546	150,1662	205,1279	219,0849	257,1211	277,2208	299,259	321,2694	351,0987	396,7319	450,8064	566,8916	1046,2064	120,2739	136,546	150,1662	205,1279	219,0849			
Argentina	ARG	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Armenia	ARM	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
American Samoa	ASM	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Antigua and Barbuda	ATG	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Australia	AUS	GDP per capita, Nominal	GDP per capita, Nominal	181,597	187,776	185,642	196,109	213,138	228,012	234,832	250,111	272,131	291,386	303,848	345,142	394,373	4770,619	6482,831	703,744	748,596	777,497	822,650	924,359	1026,65	1185,51	1277,89	1515,62	1241,47	1197,93			
Austria	AUT	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Azerbaijan	AZE	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Belarus	BLR	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Belgium	BEL	GDP per capita, Nominal	GDP per capita, Nominal	127,692	130,859	148,523	155,024	170,846	185,595	195,606	208,636	222,262	245,082	262,791	280,539	303,833	363,493	490,952	572,798	671,377	745,047	842,947	1029,779	1181,62	1484,102	1626,273	1844,237	2057,299	2267,360			
Benin	BEN	GDP per capita, Nominal	GDP per capita, Nominal	90,0383	92,3748	96,2566	100,4681	105,9669	108,602	107,682	112,476	111,6953	110,3469	108,4987	124,9996	156,2803	168,0123	200,8123	211,4929	265,3324	317,1078	326,5205	327,565	328,6932	362,2777	444,8444	444,8444	444,8444	444,8444	444,8444	444,8444	
Burkina Faso	BFA	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Bangladesh	BGD	GDP per capita, Nominal	GDP per capita, Nominal	84,82534	92,65511	95,04781	96,30221	94,86756	100,9679	108,8557	116,7983	116,9405	128,6079	133,1429	127,9955	104,6787	114,2185	124,8115	132,1288	155,7555	176,2975	224,3286	259,0376	276,1955	295,2455	329,619	322,487	338,1855	348,3558	368,3558	388,3558	408,3558
Bulgaria	BGR	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Bahrain	BHR	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Bahamas	BHS	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Bosnia and Herzegovina	BH	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Bosnia and Herzegovina	BH	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Barbados	BRB	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Brunei Darussalam	BRN	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Rhodesia	RTN	GDP per capita, Nominal	GDP per capita, Nominal	142,692	144,342	148,774	157,047	166,849	177,769	186,630	199,598	210,078	229,711	243,158	258,705	270,761	345,573	412,831	426,309	414,468	447,927	485,666	560,175	720,771	729,912	668,777	692,732	625,466				
Rhodesia	RTN	GDP per capita, Nominal	GDP per capita, Nom	1059	117.095	1181.943	1309.979	1259.337	1342.995	1242.472	1882.204	2905.414	695.187	7317.772	621.124	1015.01	1020.68	1549.887	2628.208	2494.887	1059.173	1254.228	1456.177	1653.223	1852.269	2052.269	2252.269	2452.269	2652.269	2852.269	3052.269	

FIGURE 2: GDP PER CAPITA (CURRENT US\$) FROM WORLD BANK.ORG

Country	Country	Cr.	Indicator	Indicator	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984						
6	Aruba	ABW	Population	SP.TOT.POP.TC	54608	55745	56187	56592	56922	59330	59106	58816	58855	59365	60028	60715	61193	61465	61739	62090	62267	62614	63116	63683	64174										
7	Africa, Eas EAF	Population	SP.TOT.POP.TC	1.31E+08	1.34E+08	1.38E+08	1.42E+08	1.46E+08	1.51E+08	1.54E+08	1.58E+08	1.61E+08	1.64E+08	1.72E+08	1.76E+08	1.81E+08	1.88E+08	1.94E+08	1.99E+08	2.05E+08	2.11E+08	2.17E+08	2.24E+08	2.31E+08	2.38E+08	2.45E+08	2.53E+08	2.61E+08							
8	Afghanistan AFG	Population	SP.TOT.POP.TC	86,228	87,903	89,574	91,145	93,555	95,561	97,800	100,229	104,850	109,297	110,185	112,835	115,730	118,987	121,735	126,070	128,760	131,200	134,680	137,150	140,550	143,950	147,350	150,750	154,150	157,550						
9	Angola ANG	Population	SP.TOT.POP.TC	53,571,985	54,413,933	55,240,400	55,988,927	57,738,582	57,604,444	58,620,033	58,620,856	59,620,070	57,670,000	57,645,271	57,645,271	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000	57,670,000					
10	Albania ALB	Population	SP.TOT.POP.TC	16,088,000	17,111,919	18,141,833	18,671,941	19,145,733	19,655,988	20,222,016	20,815,479	21,415,216	22,067,625	22,710,485	23,412,166	24,112,847	24,812,527	25,512,208	26,211,889	26,911,569	27,611,240	28,310,911	29,010,572	29,710,233	30,410,894	31,110,555	31,810,216	32,510,877	33,210,538	33,910,199	34,610,860	35,310,521			
11	Andorra AND	Population	SP.TOT.POP.TC	9,443	10,216	11,014	11,839	12,599	13,353	14,546	15,175	17,079	18,499	19,866	21,232	22,582	23,433	25,260	27,240	29,294	30,394	31,374	32,344	33,315	34,287	35,259	36,231	37,203	38,175	39,147					
12	Arab World AWD	Population	SP.TOT.POP.TC	93,706,489	95,630,484	98,203,608	101,008	104,08	108,08	112,08	116,08	120,08	124,08	128,08	132,08	136,08	140,08	144,08	148,08	152,08	156,08	160,08	164,08	168,08	172,08	176,08	180,08	184,08	188,08	192,08					
13	United Arz ARE	Population	SP.TOT.POP.TC	13,342,466	14,099,4	14,887,7	15,700,6	16,530,5	17,379,7	18,250,9	19,140,4	21,352,6	23,252,1	29,980,4	34,451,3	39,236,9	44,154,6	49,195,5	54,339,6	61,417,7	70,696,1	80,521,1	90,845,2	101,404,1	110,193,8	115,785,2	123,757,2	130,833,1							
14	Argentina ARG	Population	SP.TOT.POP.TC	20,477,444	20,860,653	20,120,593	20,510,209	20,910,217	21,707,487	20,536,961	24,031,111	22,757,014	23,121,971	24,720,472	23,480,263	24,127,949	24,612,179	25,205,958	25,449,754	25,755,558	26,292,057	26,717,170	27,141,212	27,508,418	28,040,803	28,471,285	28,937,219	29,377,262	29,737,157	30,191,230					
15	Armenia ARM	Population	SP.TOT.POP.TC	19,044,181	19,715,030	20,934,6	21,061,42	21,710,82	22,405,1	23,559,7	24,164,61	24,763,0	25,437,7	25,909,23	26,480,4	27,068,52	27,658,2	28,230,79	28,851,1	29,455,22	30,022,1	30,592,3	31,152,5	31,678,7	32,201,1	32,773,2	33,272,10								
16	American ASM	Population	SP.TOT.POP.TC	20,085	20,626	21,217	21,949	22,656	23,391	24,122	24,848	25,608	26,396	27,075	27,793	28,063	28,529	29,012	29,573	30,194	30,816	31,428	32,045	32,684	33,409	33,367	36,735	36,316							
17	Antigua atG	Population	SP.TOT.POP.TC	5,532,42	5,625,45	5,709,4	5,777,8	5,864,4	5,964	6,015,1	6,167	6,265,8	6,374,2	6,451,7	6,479	6,469,6	6,446,3	6,421,8	6,403,5	6,399,0	6,411,7	6,432,9	6,460,9	6,488,9	6,507,0	6,521,1	6,540,8	6,562,6							
18	Australia AUS	Population	SP.TOT.POP.TC	10,420,000	10,430,000	10,450,000	10,500,000	11,670,000	11,810,000	11,950,000	12,090,000	12,230,000	12,370,000	12,510,000	12,650,000	12,790,000	12,930,000	13,070,000	13,210,000	13,350,000	13,490,000	13,630,000	13,770,000	13,910,000	14,050,000	14,190,000	14,330,000	14,470,000	14,610,000	14,750,000					
19	Azerbaijan AZE	Population	SP.TOT.POP.TC	7,703,459	7,809,459	7,905,459	7,991,459	8,078,459	8,164,459	8,250,459	8,442,459	8,630,459	8,817,459	9,004,459	9,191,459	9,378,459	9,565,459	9,752,459	9,940,459	10,127,459	10,314,459	10,501,459	10,688,459	10,875,459	11,062,459	11,250,459	11,437,459	11,624,459	11,811,459	12,000,459					
20	Bosnia and Herzegovina BHR	Population	SP.TOT.POP.TC	38,640,000	40,457,000	41,850,000	42,650,000	44,852,000	47,450,000	49,450,000	51,450,000	53,450,000	55,450,000	57,450,000	59,450,000	61,450,000	63,450,000	65,450,000	67,450,000	69,450,000	71,450,000	73,450,000	75,450,000	77,450,000	79,450,000	81,450,000	83,450,000	85,450,000	87,450,000	89,450,000					
21	Burundi BDI	Population	SP.TOT.POP.TC	27,462,88	28,162,7	28,762,78	29,491,33	30,322,1	31,161,94	31,903,78	32,744,93	33,544,97	34,344,97	35,147,95	35,973,75	36,792,47	37,602,20	38,402,92	39,204,67	40,002,41	40,797,00	41,593,75	42,390,50	43,187,25	43,984,00	44,780,75	45,577,50	46,374,25	47,170,00	47,966,75	48,763,50				
22	Belgium BEL	Population	SP.TOT.POP.TC	9,519,848	9,813,848	10,208,72	10,587,72	10,967,72	11,346,72	11,725,72	12,104,72	12,483,72	12,862,72	13,241,72	13,619,72	14,017,72	14,415,72	14,813,72	15,211,72	15,609,72	16,007,72	16,405,72	16,803,72	17,201,72	17,600,72	18,000,72	18,400,72	18,800,72	19,200,72	19,600,72					
23	Benin BEN	Population	SP.TOT.POP.TC	25,122,84	25,212,6	25,932,0	27,432,0	29,132,0	30,832,0	32,532,0	34,232,0	35,932,0	37,632,0	39,332,0	41,032,0	42,732,0	44,432,0	46,132,0	47,832,0	49,532,0	51,232,0	52,932,0	54,632,0	56,332,0	58,032,0	59,732,0	61,432,0	63,132,0	64,832,0	66,532,0	68,232,0				
24	Burkina Faso BFA	Population	SP.TOT.POP.TC	47,832,99	48,523,33	49,424,997	49,4967	50,761,11	51,579,23	52,427,533	53,315,74	54,211,96	55,114,21	56,016,46	56,917,71	57,819,96	58,721,21	59,622,46	60,523,71	61,425,96	62,327,21	63,228,46	64,129,71	65,029,96	65,929,21	66,829,46	67,729,71	68,629,96	69,529,21	70,429,46	71,329,71	72,229,96			
25	Bangladesh BGD	Population	SP.TOT.POP.TC	50,934,229	51,882,958	53,616,951	55,904,115	58,595,000	60,625,952	62,104,881	64,595,628	65,618,958	67,837,694	69,346,704	71,114,881	72,947,087	74,707,045	76,467,003	78,226,958	80,085,921	81,845,989	83,605,957	85,325,925	87,045,993	88,765,961	90,485,929	92,205,897	93,925,865	95,645,833	97,365,791	99,085,759	100,805,727			
26	Bulgaria BGR	Population	SP.TOT.POP.TC	7,687,734	7,943,118	10,204,876	11,843,400	13,421,840	15,000,240	16,578,640	18,144,240	19,711,640	21,278,640	22,845,640	24,412,640	26,019,640	27,626,640	29,233,640	30,840,640	32,447,640	34,054,640	35,661,640	37,268,640	38,875,640	40,482,640	42,091,640	43,699,640	45,297,640	46,895,640	48,493,640	50,091,640	51,691,640			
27	Bahrain BHR	Population	SP.TOT.POP.TC	10,691	11,697	13,735	17,891	18,657	19,302	19,919	20,481	21,019	21,669	22,255	23,843	24,565	26,207	26,516	27,104	27,703	28,302	28,901	29,500	30,100	30,700	31,300	31,900	32,500	33,100	33,700	34,300	34,900			
28	Bahamas BHS	Population	SP.TOT.POP.TC	11,450,000	12,061,250	12,625,000	13,200,250	13,783,000	14,364,000	14,943,000	15,521,000	16,100,000	16,678,000	17,253,000	17,828,000	18,403,000	18,978,000	19,553,000	20,130,000	20,705,000	21,280,000	21,855,000	22,430,000	23,005,000	23,576,000	24,146,000	24,716,000	25,286,000	25,856,000	26,426,000	27,096,000				
29	Bosnia and Herzegovina BHN	Population	SP.TOT.POP.TC	32,625,93	33,253,93	33,875,12	34,485,93	35,086,93	35,685,93	36,280,93	36,873,93	37,707,93	38,541,93	39,373,93	40,204,93	41,035,93	41,866,93	42,697,93	43,528,93	44,360,93	45,191,93	46,012,93	46,833,93	47,654,93	48,475,93	49,296,93	50,117,93	50,938,93	51,759,93	52,580,93					
30	Bolivia BOL	Population	SP.TOT.POP.TC	37,747,44	38,440,11	39,447,95	40,492,74	41,155,11	42,045,95	42,925,65	43,804,95	44,686,95	45,665,95	46,642,95	47,730,94	48,922,48	50,112,94	51,302,94	52,491,94	53,681,94	54,871,94	56,061,94	57,251,94	58,441,94	59,631,94	60,821,94	62,011,94	63,201,94	64,391,94	65,581,94	66,771,94	67,960,94			
31	Brazil BRA	Population	SP.TOT.POP.TC	70,325,000	75,330,000	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555	77,991,555			
32	Barbados BAR	Population	SP.TOT.POP.TC	22,550,000	23,236,000	23,848,000	24,363,000	24,876,000	25,390,000	25,895,000	26,399,000	26,893,000	27,397,000	27,891,000	28,395,000	28,899,000	29,393,000	29,897,000	30,391,000	30,895,000	31,399,000	31,893,000	32,397,000	32,891,000	33,395,000	33,899,000	34,393,000	34,897,000	35,391,000	35,895,000	36,399,000	36,893,000	37,397,000	37,891,000	
33	Brunei DA BRN	Population	SP.TOT.POP.TC	8,534	8,956	9,383	9,827	10,294	10,766	11,259	11,762	12,264	13,201	13,334	13,386	13,438	13,490	13,543	13,595	13,650	13,706	13,761	13,816	13,871	13,926	14,081	14,136	14,181	14,231	14,281	14,331	14,381	14,431	14,481	14,531
34	Bhutan BTN	Population	SP.TOT.POP.TC	21,226,000	23,317,000	25,474,000	25,550,000	25,629,000	27,139	28,010	28,926	29,884	30,857	31,192	31,528	31,864	32,367	32,787	33,202																

FIGURE 3: POPULATION TOTAL FROM WORLD BANK ORG

The first step of cleaning is to remove unnecessary rows and columns that do not contain important information needed for the visualization. This process was done by hand since all three files only contain

a single table, and by cleaning and renaming the file to *WB_LifeExpect.csv*, *GDP.csv*, and *Population.csv* respectively, we can better handle the file here on out. Figures 3, 4, and 5 depict the aforementioned.

FIGURE 4: LIFE EXPECTANCY DATA AFTER CLEANING BY HAND

FIGURE 5: GDP PER CAPITA DATA AFTER CLEANING BY HAND

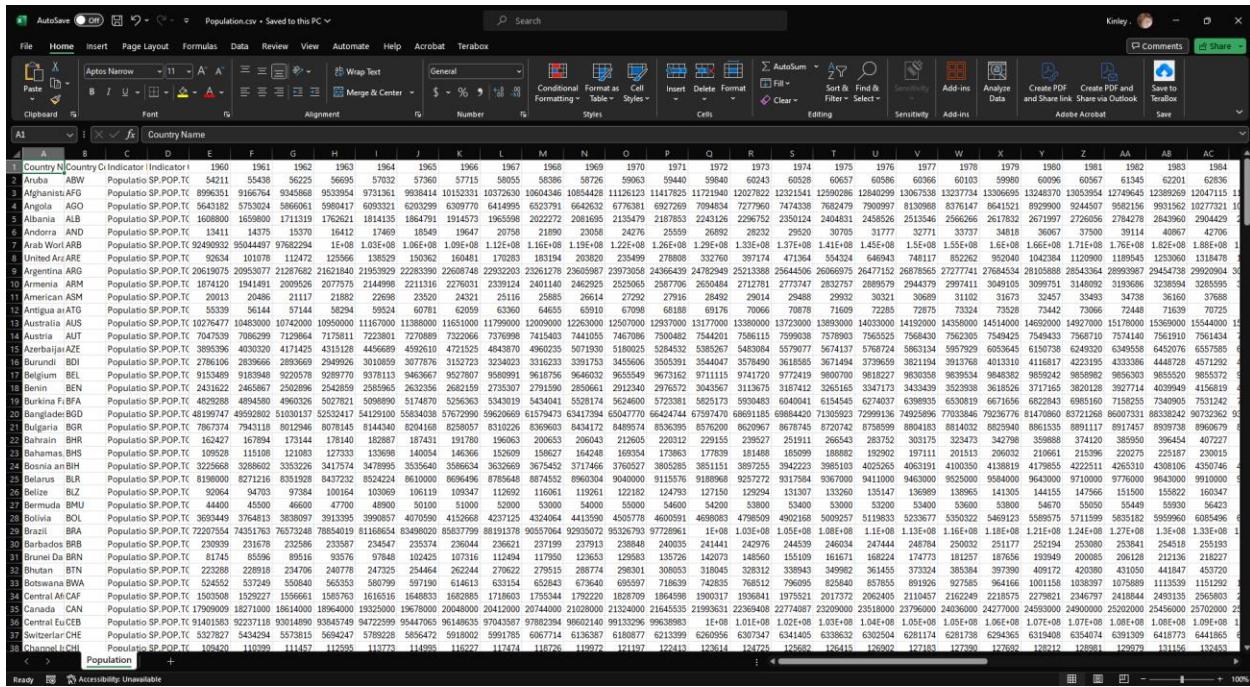


FIGURE 6: TOTAL POPULATION DATA AFTER CLEANING BY HAND

United Nations (www.un.org):

Different from the dataset gathered from the World Bank, the original data from the United Nations is structured as an XLSX (Microsoft Excel Open XML Format Spreadsheet) file, with multiple tables each representing a different dataset. The following figure depicts the original data downloaded from the United Nations.

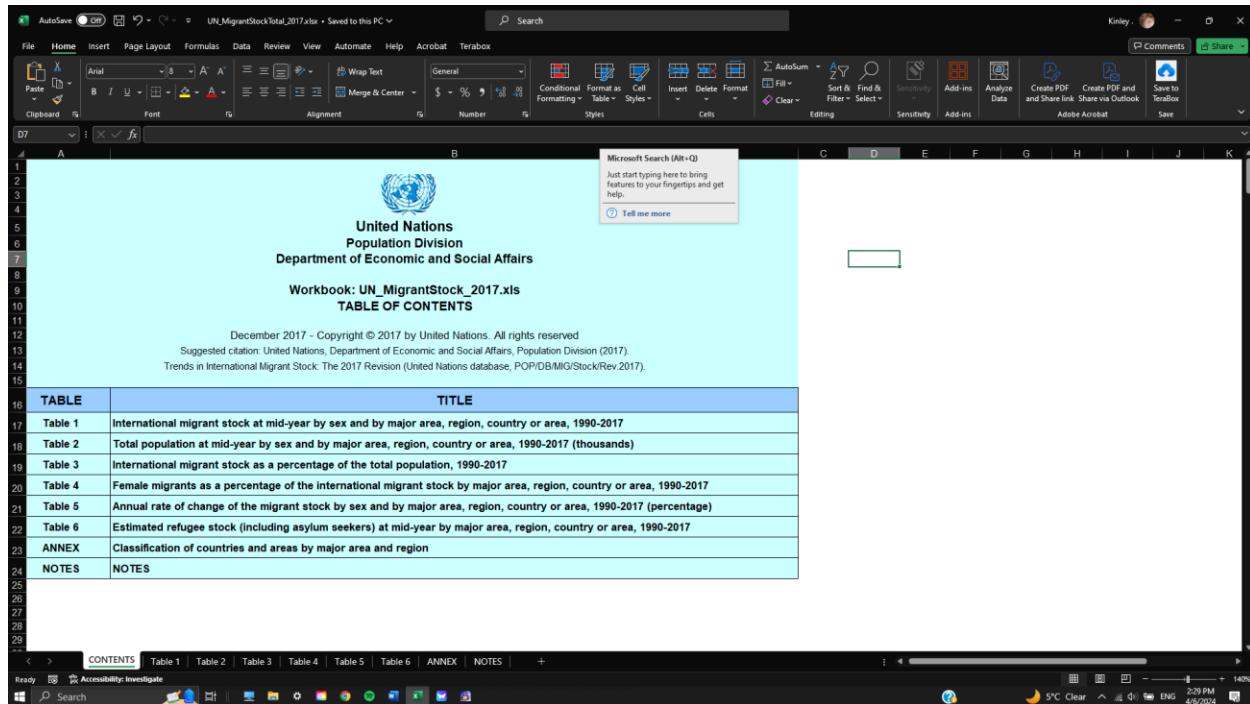


FIGURE 7: RAW DATA FILE DOWNLOADED FROM UN

With the first step of making the dataset usable for our project, the file needs to be converted to a CSV format. By using a web-based application called “cloudconvert” (www.cloudconvert.com/csv-to-xlsx), we were able to obtain the required file. Based on the above figure, the file contains multiple tables, each representing an aspect of immigrant and emigrant when being considered for comparison between countries. This was an obstacle for us when trying to process the data. Instead of processing the whole file, we decided to pick out the tables that held information that was crucial to the aim of the visualization. The table we have chosen is Table 1 (International migrant stock at mid-year by sex and by major area, region, country, or area, 1990 – 2017). With the process of handling the dataset by hand, we have split the said table into different files, with each name containing the format “*Migrate_XXXX.xlsx*” where “XXXX” corresponds to each respective year. The original Table 1 is as follows:

Table 1 - International migrant stock at mid-year by sex and by major area, region, country or area, 1990-2017

December 2017 - Copyright © 2017 by United Nations. All rights reserved

United Nations
Population Division
Department of Economic and Social Affairs

Workbook: UN_MigrantStock_2017.xlsb • Saved to this PC

File Home Insert Page Layout Formulas Data Review View Automate Help Acrobat Terabox

Font Alignment Number Styles Cells Editing

Comments Share

Sort & Filter Select Sensitivity Add-ins Analyze Data Create PDF and Share link. Share via Outlook Adobe Acrobat Save

AA2

Major area, region, country or area of destination Notes Code Type of data (id) 1990 1991 1992 1993 1994 1995 1996 1997 1998 1999 2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 1990 1995 2000 2005 2010 2015 2016 2017 International migrant stock at mid-year (male) International migrant stock at mid-year (female)

Sort and filter

1 WORLD

2 Major developed regions

3 Less developed regions

4 Least developed countries

5 Least developed regions, excluding least developed countries

6 High-income countries

7 Middle-income countries

8 Lower-middle-income countries

9 Lower-middle-income countries

10 Low-income countries

11 Sub-Saharan Africa

12 Africa

13 Eastern Africa

14 Central Africa

15 Comoros

16 Djibouti

17 Eritrea

18 Ethiopia

19 Madagascar

20 Malawi

21 Maldives

22 Mauritius

23 Mayotte

24 Mozambique

25 Namibia

26 Rwanda

27 Seychelles

28 South Sudan

29 Uganda

30 United Republic of Tanzania

31 Zambia

32 Zimbabwe

33 Mongolia

34 Montenegro

35 Angola

36 Azerbaijan

37 Central African Republic

38 Chad

FIGURE 8: THE ORIGINAL TABLE 1 FROM THE UN DATASET

After the process of hand-cleaning data for table 1, we have the following files:

Ready CONTENTS Table 1 Table 2 Table 3 Table 4 Table 5 Table 6 ANNEX NOTES +

4°C Clear 4/9/2024 ENG 4/9/2024

GDP.csv update 2 days ago

ISO_codes.csv update 6 minutes ago

LatLong.csv update 2 days ago

Migrate_1995.xlsx update 2 days ago

Migrate_2000.xlsx update 2 days ago

Migrate_2005.xlsx update 2 days ago

Migrate_2010.xlsx update 2 days ago

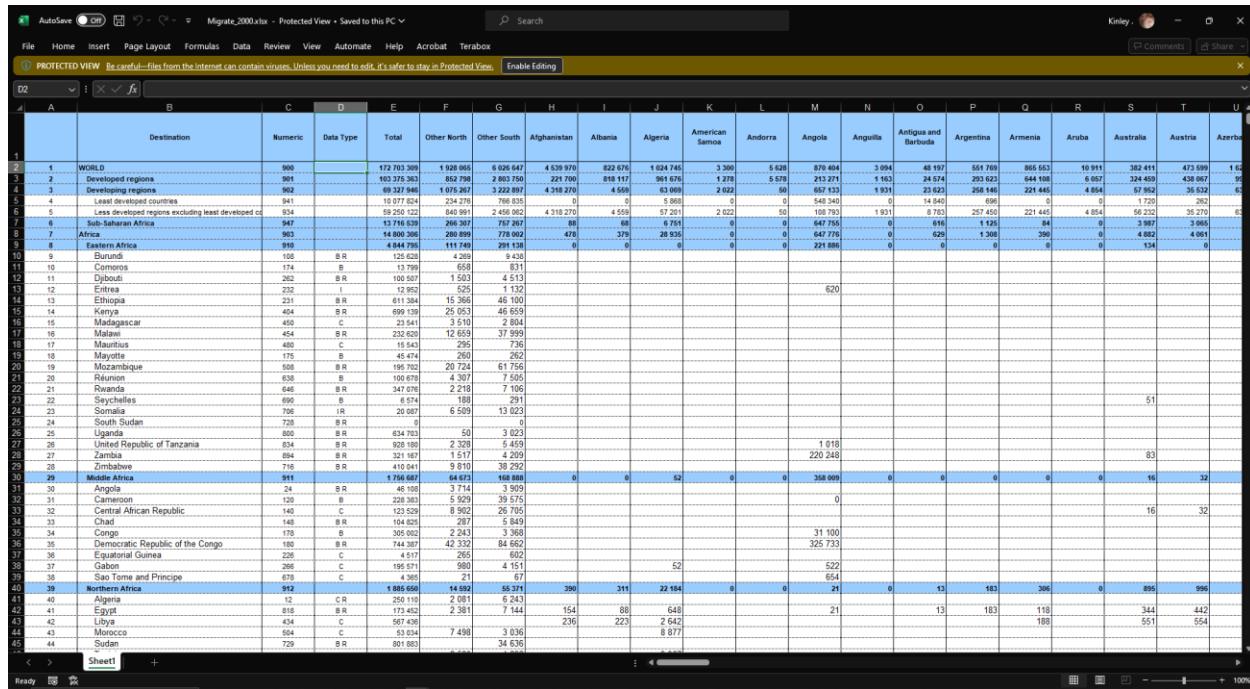
Migrate_2015.xlsx update 2 days ago

Population.csv update 2 days ago

WB_LifeExpect.csv update 2 days ago

FIGURE 9: SPLIT THE TABLE MENTIONED ABOVE INTO DIFFERENT FILES

Each of these files has the same format, with the only difference between each one being the year they represent. For example, the following figure is a representation of the format, more specifically the file “*Migrate_2000.csv*”.



		Destination	Numeric	Data Type	Total	Other North	Other South	Afghanistan	Albania	Algeria	American Samoa	Andorra	Angola	Anguilla	Antigua and Barbuda	Argentina	Armenia	Aruba	Australia	Austria	Azerbaijan	
2	1	WORLD	900		172 703 309	1 928 065	6 026 647	4 539 970	922 676	1 024 745	3 300	5 628	970 404	3 094	48 197	561 769	865 553	10 911	382 411	473 599	1 640	
3	2	Developed regions	901		163 375 363	852 798	2 860 750	221 700	918 117	961 676	1 278	5 578	213 071	1 163	24 574	293 623	644 108	6 057	324 450	438 067	95	
4	3	Developing regions	902		69 328 946	1 079 274	2 009 897	4 318 270	4 559	5 889	2 923	58	58	1 931	23 623	258 146	221 445	4 854	67 952	35 532	63	
5	4	Least developed countries	904		19 077 770	2 027 279	768 100	0	5 989	0	0	0	548 549	0	1 447	1 060	0	179	260	260	260	
6	5	Less developed regions excluding least developed countries	924		59 250 122	840 991	2 450 092	4 318 270	4 559	57 201	2 022	58	108 793	1 931	0	8 753	257 450	221 445	4 854	56 232	35 270	63
7	6	Sub-Saharan Africa	947		13 716 539	266 307	757 267	88	68	6 755	0	0	647 765	0	0	616	1 125	84	0	3 987	3 065	0
8	7	Africa	963		14 800 306	269 899	778 062	478	379	28 355	0	0	647 776	0	0	629	1 308	390	0	0	4 882	4 051
9	8	Eastern Africa	916		4 800 798	119 149	0	0	0	0	0	0	221 886	0	0	0	0	0	0	0	134	0
10	9	Burundi	160	B	124 000	4 200	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
11	10	Comoros	174	B	13 799	658	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
12	11	Djibouti	262	B	100 507	1 503	4 513	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
13	12	Entebbe	232	I	12 952	525	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
14	13	Ethiopia	231	B	111 546	18 201	46 559	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
15	14	Kenya	404	B	899 139	26 053	46 559	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
16	15	Madagascar	450	C	23 541	3 510	2 804	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
17	16	Malawi	454	B	232 620	12 659	37 599	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
18	17	Morocco	486	C	15 543	295	736	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
19	18	Mayotte	115	B	45 474	2 304	263	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
20	19	Mozambique	508	B	195 702	20 724	61 756	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
21	20	Réunion	638	B	100 678	4 307	7 505	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
22	21	Rwanda	646	B	347 076	2 218	7 106	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
23	22	Seychelles	696	B	6 574	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
24	23	Senegal	766	B	20 270	6 569	13 023	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
25	24	South Sudan	728	B	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
26	25	Uganda	800	B	634 703	50	3 023	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
27	26	United Republic of Tanzania	834	B	928 180	2 328	5 459	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
28	27	Zambia	894	B	22 101	1 574	2 297	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
29	28	Zimbabwe	894	B	410 041	9 810	38 292	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
30	29	Middle Africa	911		1 756 087	64 073	108 888	0	0	0	0	0	0	358 009	0	0	0	0	0	16	32	
31	30	Angola	24	B	46 108	3 714	3 909	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
32	31	Cameroon	126	B	228 383	9 529	39 575	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
33	32	Central African Republic	142	C	133 459	8 952	28 261	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
34	33	Chad	145	B	104 821	287	1 849	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
35	34	Congo	178	B	305 002	2 243	3 368	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
36	35	Democratic Republic of the Congo	180	B	744 387	42 332	84 662	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
37	36	Equatorial Guinea	228	C	4 517	265	602	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
38	37	Gabon	266	C	195 200	9 982	44 111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
39	38	Sao Tome and Principe	676	C	4 365	21	67	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
40	39	Northern Africa	912		1 885 658	14 592	55 371	390	311	22 184	0	0	21	0	0	13	183	206	0	895	996	
41	40	Algeria	12	C	250 110	2 081	6 243	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
42	41	Egypt	184	B	172 425	2 381	7 144	154	88	648	0	0	21	0	0	13	183	118	0	344	442	
43	42	Lybia	434	C	567 439	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
44	43	Morocco	504	C	53 034	7 498	3 036	0	0	0	0	0	0	0	0	0	0	0	0	0	0	
45	44	Sudan	729	B	801 853	34 636	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	

FIGURE 10: MIGRATE_2000.CSV EXAMPLE

All the “Migrate_XXXX.xlsx” was formatted the same for ease of use and cleaner data representation.

Code Processing:

With each of the previously hand-processed data files, we put them into the same directory called “`raw_data`” for ease of navigation and differentiation. All the previous files were not suitable for generating a visualization, so we had to come up with a Python-based program to help us process the data better and faster. The code can be explained as the following:

1. Data Processing: The program reads and processes various demographic and economic data from different sources, including data related to countries, population, life expectancy, GDP, and migration.
2. Statistical Calculations: The program calculates various statistics related to migration, such as positive and negative stock, net migration, and immigration fraction.
3. Data Output: After processing and analyzing the data, the program outputs the results into structured formats, including JSON and CSV files, which can be used for further analysis, visualization, or integration with other systems.

```

1 # Define constants
2 labels = ['Country Name', 'Country Code']
3 years = [1990, 1995, 2000, 2005, 2010, 2015]
4 used = labels + years
5 not_used = ['destination', 'Numeric', 'Data Type', 'Total', 'Other North', 'Other South']
6
7 # Read the manually prepared data
8 countries = pd.read_csv('./raw_data/countrieslist.txt', delimiter=' ', index_col='Country Code', engine='python')
9 latlong = pd.read_csv('./raw_data/latlong.csv', index_col='Alpha 3', engine='python', usecols=['Alpha 3', 'lat', 'long'])
10 iso_codes = pd.read_csv('./raw_data/ISO_codes.csv', index_col='Country Code')
11 pop = pd.read_csv('./raw_data/population.csv', index_col='Country Code', usecols=used)
12 lifeexp = pd.read_csv('./raw_data/lifeexpect.csv', index_col='Country Code', usecols=used)
13 gdp = pd.read_csv('./raw_data/gdp.csv', index_col='Country Code', usecols=used)
14 migrate_data = pd.read_excel('./raw_data/migrate_1990.xlsx', header=0)
15
16 # Initialize variables to store read data and results
17 countries_dict = countries.to_dict()
18 countries_list = countries.index.tolist()
19 latlong_list = latlong.index.tolist()
20
21 shared_index = countries_list[not_used] & set(latlong_list)].index
22 shared_index = pop.index.intersection(shared_index)
23
24 pop = pop.reindex(shared_index)
25
26 immigrants = {}
27 emigrants = {}
28 maxs = {}
29
30 in_lines = {}
31 em_lines = {}
32
33 positive_stock = pd.DataFrame(0.0, index=pop.index, columns=years)
34 negative_stock = pd.DataFrame(0.0, index=pop.index, columns=years)
35 net_stock = pd.DataFrame(0.0, index=pop.index, columns=years)
36 total_stock = pd.DataFrame(0.0, index=pop.index, columns=years)
37 in_frac = pd.DataFrame(0.0, index=pop.index, columns=years)
38
39 # Reusable functions
40
41 # Scale values for easier reading
42 def scale(value, max_val):
43     return ((value * 0.2) / max_val) + 0.2
44
45 # Calculate migration data for each defined year
46 def calc_migration(year):
47     migration = pd.read_excel('./raw_data/Migrate_{}_year.xlsx', header=0)
48
49     migration = migration.query('Numeric < 000 and Numeric > 999')
50     migration['Country Code'] = iso_codes.loc[migration['Numeric'], 'alpha-3'].values
51     migration.set_index('Country Code', inplace=True)
52     migration_shard_index = migration.index.intersection(pop[year].dropna().index)
53     migration = migration.reindex(migration_shard_index)
54     pop[year] = pop[year].reindex(migration_shard_index)
55
56     good_countries = set(migration['Destination'])
57
58     for code in migration.columns[G]:
59         if code not in good_countries:
60             migration.drop(code, axis=1, inplace=True)
61         else:
62             migration.rename(columns={code: migration.index[migration['Destination'] == code][0]}, inplace=True)
63
64     new_not_used = []
65     for not_used_col in not_used:
66         if not_used_col in migration:
67             new_not_used.append(not_used_col)
68
69     return migration.drop(new_not_used, axis=1)
70
71 # Calculate the origin and end of a migration
72 def mig_line(origin, destination, status, value, scaled_value):
73     d = dict()
74     d['origin'] = ('latitude': latlong.loc[origin]['lat'],
75                   'longitude': latlong.loc[origin]['long'])
76     d['destination'] = ('latitude': latlong.loc[destination]['lat'],
77                         'longitude': latlong.loc[destination]['long'])
78     d['value'] = value
79     d['scaled_value'] = scaled_value
80     if status == 'in':
81         d['id'] = origin
82         d['name'] = countries_dict['Country Name'][origin]
83         d['label'] = origin
84         d['id'] = destination
85         d['name'] = countries_dict['Country Name'][destination]
86         return d
87
88 # Process each row in the migration data file
89 def mig_row(row_name, row, status, max_val):
90     if status == 'in':
91         target = destination + migration
92         target += ('latitude': latlong.loc[destination]['lat'],
93                    'longitude': latlong.loc[destination]['long'])
94         target += ('latitude': latlong.loc[origin]['lat'],
95                    'longitude': latlong.loc[origin]['long'])
96     else:
97         target = origin + migration
98         target += ('latitude': latlong.loc[origin]['lat'],
99                    'longitude': latlong.loc[origin]['long'])
100
101     target['value'] = value
102     target['scaled_value'] = scaled_value
103
104     if target['value'] > max_val:
105         target['value'] = max_val
106     if target['scaled_value'] > max_val:
107         target['scaled_value'] = max_val
108
109     return target
110
111 # Process the migration data
112 def mig_lines(migration, status, threshold, maximum):
113     assert status in ['in', 'em']
114     d = {}
115     if status == 'in':
116         for code, row in migration.iterrows():
117             d[mig_line(code, row['Destination'], 'in', row['value'])] = row['value']
118     else:
119         for code, row in migration.iterrows():
120             d[mig_line(row['Origin'], code, 'em', row['value'])] = row['value']
121
122     return d
123
124 # Process data for each defined year
125 for year in years:
126     immigrants[year] = calc_migration(year)
127     emigrants[year] = immigrants[year].transpose()
128     emigrants[year] = emigrants[year].transpose()
129     in_lines[year] = mig_lines(emigrants[year], 'in', 1000, maxs[year])
130     em_lines[year] = mig_lines(emigrants[year], 'em', 1000, maxs[year])
131
132     positive_stock[year] = immigrants[year].fillna(0.0).sum(axis=0)
133     negative_stock[year] = emigrants[year].fillna(0.0).sum(axis=0)
134
135     total_stock[year] = positive_stock[year] + negative_stock[year]
136     net_stock[year] = positive_stock[year] - negative_stock[year]
137
138     in_frac[year] = net_stock[year] / total_stock[year]
139
140     lifeexp = lifeexp.reindex(lifeexp.index.intersection(shared_index))
141     lifeexp.drop('SNP', inplace=True)
142
143     pop = pop.reindex(pop.index.intersection(countries.index))
144     gdp = gdp.reindex(gdp.index.intersection(pop.index))
145
146     gdp_per_cap = gdp.copy()
147     for year in years:
148         gdp_per_cap[year] = gdp[year] / pop[year]
149
150     gdp_per_cap.rename(columns={'Country Name': 'name'}, inplace=True)
151
152     # Write to processed files for use in analysis
153     with open('./processed_data/immigrant.json', 'w') as out: json.dump(in_lines, out, indent=2)
154     with open('./processed_data/emigrant.json', 'w') as out: json.dump(em_lines, out, indent=2)
155     with open('./processed_data/countries.json', 'w') as out: json.dump(countries_dict['Country Name'], out, indent=2)
156     gdp_per_cap.to_csv('./processed_data/gdp_per_cap.csv', index_label='id')
157     lifeexp.to_csv('./processed_data/lifeTotalMigrants.csv', index_label='id')
158     total_stock.to_csv('./processed_data/NetTotalMigrants.csv', index_label='id')
159     in_frac.to_csv('./processed_data/NetTotalRatio.csv', index_label='id')
160
161

```

FIGURE 11: PYTHON-BASED PROGRAM FOR DATA CLEANING

Additional files:

Apart from all the above dataset files, two more were considered to be crucial during the process of creating the visualizations: *ISO_codes.csv* and *LatLong.csv*.

3. Requirements

3.1. Must-have Features:

Interactive Visualization:

- a. Hover Effect:
 - Implemented the world map to display detailed data and explanations when hovering over specific countries or regions.
 - Added to the map to highlight migration pathways and showcase detailed data.
- b. Drill-Down:
 - Different description box tabs are shown when clicking on specific countries, providing a How-to-read section and a summary/background section. Additionally, a detailed data table is displayed.
- c. Selectors:
 - Implemented the world map, allowing users to select a country and view migration patterns to and from that country.

Additional Information:

- Providing ample context is crucial for visualization understanding, so we provide additional labeling.

Clear Data Labeling:

- Ensures users grasp the represented data.

Color Contrast:

- Different colors represent distinct migration flows, aiding readers' visual clarity.

3.2. Optional Features:

Tooltip Integration:

Tooltips offer instant context, eliminating the need to memorize information. While this functionality couldn't be implemented, an alternative has been provided to serve a similar purpose.

4. Visualization Design:

We structure our visualization design process into four stages, each applicable to both initial development and refinement. These stages are as follows:

1. Conceptualize:

- Brainstorming & Sketching: Generate rough ideas based on project goals and data. This includes sketching layouts, exploring visual elements (maps, charts, graphs), and storytelling approaches.
- Research & Inspiration: Analyze existing migration visualizations to learn from best practices.
- Target Audience: Consider user needs (policymakers, researchers, public) for effective communication.
- Technical Feasibility: Evaluate visualization techniques considering data complexity, performance, and resources.

2. Visualize:

- Refine Sketches: Polish initial sketches based on feedback, using design software or prototyping tools.
- Prototype Layouts: Develop layouts balancing visual appeal and functionality. Experiment with data element arrangement, scales, and annotations.
- Color Palette & Style: Explore color schemes, typography, and graphic styles for a cohesive and engaging presentation. Apply color theory for readability and meaning.

3. Materialize:

- Finalize Design: Make final decisions on styling, colors, and additional visual elements. Create a style guide for consistency.
- Iterate on Feedback: Refine based on usability testing and user research, ensuring accessibility for diverse users.
- Interactive Elements: Consider tooltips, filters, and animations to enhance engagement and exploration. Design intuitive navigation to guide users.

4. Implementation:

- Coding the Design: Translate final design concepts into code using appropriate libraries (d3.js).
- Data Integration: Integrate cleaned and processed data, ensuring accuracy and relevance. Use data-driven visual encoding techniques (size, color, position).
- Testing & Debugging: Conduct thorough testing to identify and resolve technical issues or compatibility concerns. Test across devices, screen sizes, and browsers.
- Iterative Changes: Propose changes based on usability testing to optimize engagement and comprehension. Document the implementation process for future updates and maintenance.

4.1. Conceptualize (Design Iteration) stage:

We want to offer a straightforward, interactive layout in the first early stages of the visualization design process that will work for the public, who is our target audience. We adopt distinct approaches to various visualizations while working with classified sorted data, all the while maintaining consistency to guarantee a user-friendly experience.

Without thoroughly reviewing the data we have discovered, we created a rough sketch for the first visualization, Migration Pattern, to convey our ideas to the reader. We concentrated on "World Map" style visualizations and carefully considered our options because our goal is to visualize the flow of immigrants and emigrants from different countries to one another.

The figure below is our attempt to implement a "World Map" visualization, which in our opinion, was the go-to way to help users grasp the concept of Immigrant and Emigrant flows relative to countries all over the world.



FIGURE 12: AN ATTEMPT AT THE WORLD MAP SKETCH

Throughout our research, we have found other visualization alternatives but mostly aren't suitable for the scope our project is going for. A bar chart is a great implementation of a visualization to describe the flow of immigrations and emigrations over time. Still, sadly that's only applicable for a single country for each country, since implementing for all countries over the world can make a bar chart seem inadequate when the user wants to compare between countries. Instead, we have chosen the Bar Chart as a supporting visualization for the main one, i.e. the World Map above. Another type of supporting visualization we decided to implement was a Pie Chart, for which the goal was to visualize the men and women percentage in immigrating or emigrating between countries.

4.2. Visualize Stage:

We now produce a set of higher fidelity designs after looking into our design decisions, which aids in contextualizing the idea behind these visualizations. We would like to see a justified, centered look at the Migration Pattern visualization, with different sections representing different rates of immigration and emigration.



FIGURE 13: FIRST ITERATION OF THE WORLD MAP

The figure depicts our first iteration of a World Map. The key feature is the clear delineation of countries by black borders. There's no indication of country labels or other details on this specific screenshot, and the scale isn't evident, making it difficult to gauge the relative size of different regions. However, this map serves as a simple yet informative way to visualize the Earth's landmasses divided into individual nations. There is no additional information about the World Migration Project on the screenshot. It is unclear what the purpose of the website is, or what kind of information it provides about migration.

With these ideas in mind, a mockup for the website was made by me in an Excel sheet to gather feedback from the lecturer, Dr Hoang Xuan Tung. We made suggestions on how to implement the GDP per Capita data into the visualization to help users better understand the correlation between the immigration/emigration rate and how the GDP of a country can affect those numbers. We have chosen a side-to-side Pie Chart depicted in the figure below, to help users compare developing and developed countries. It would contain the percentage of men and women when immigrating or emigrating, and the user can specify which country to visualize.

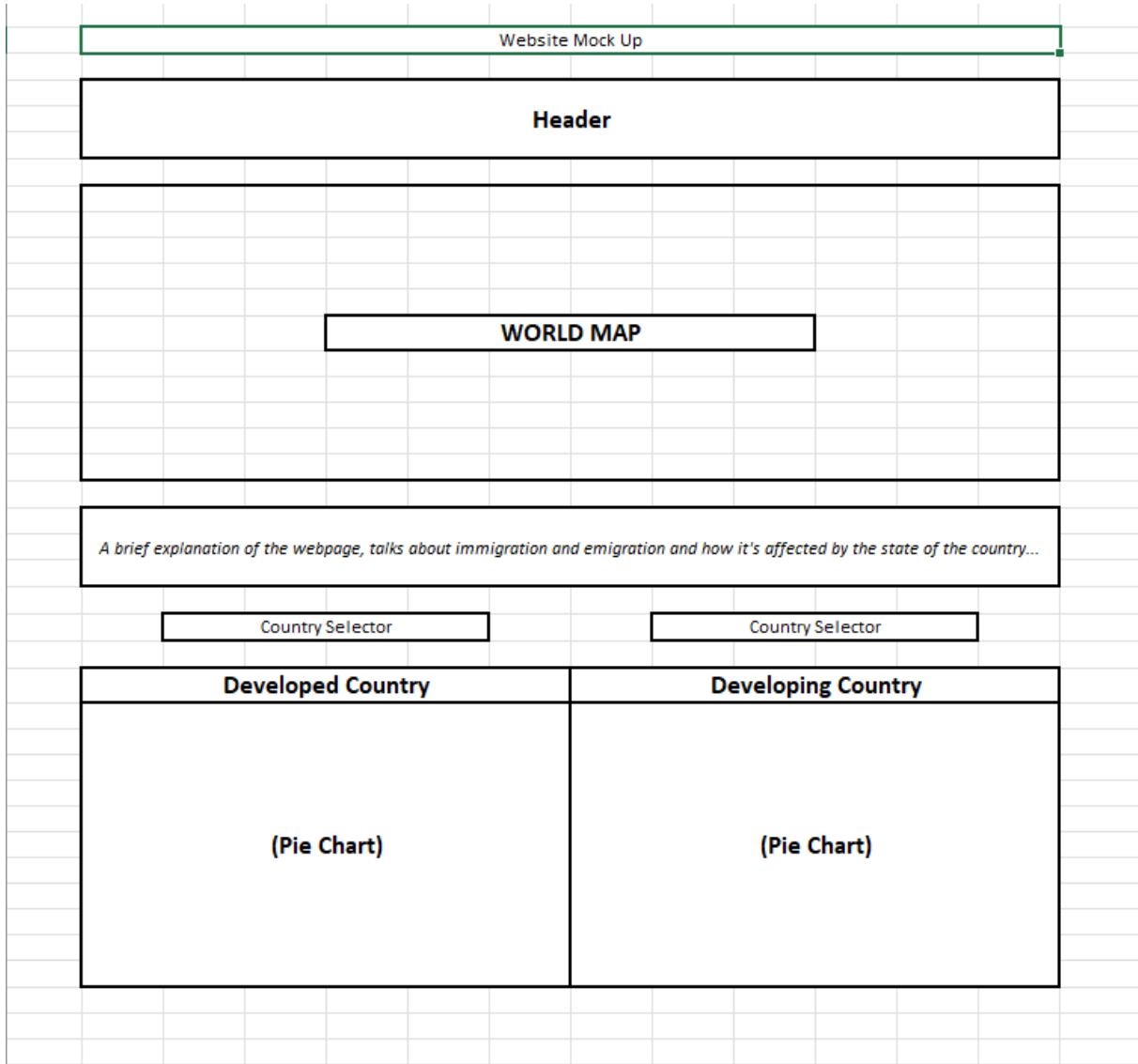


FIGURE 14: MOCKUP DESIGN OF THE WEBSITE

After receiving feedback from Dr. Hoang Xuan Tung, we have decided to implement a color system within the World Map to display the GDP per capita, making the two Pie charts below the main visualization redundant. This will be further elaborated in the next section of this report.

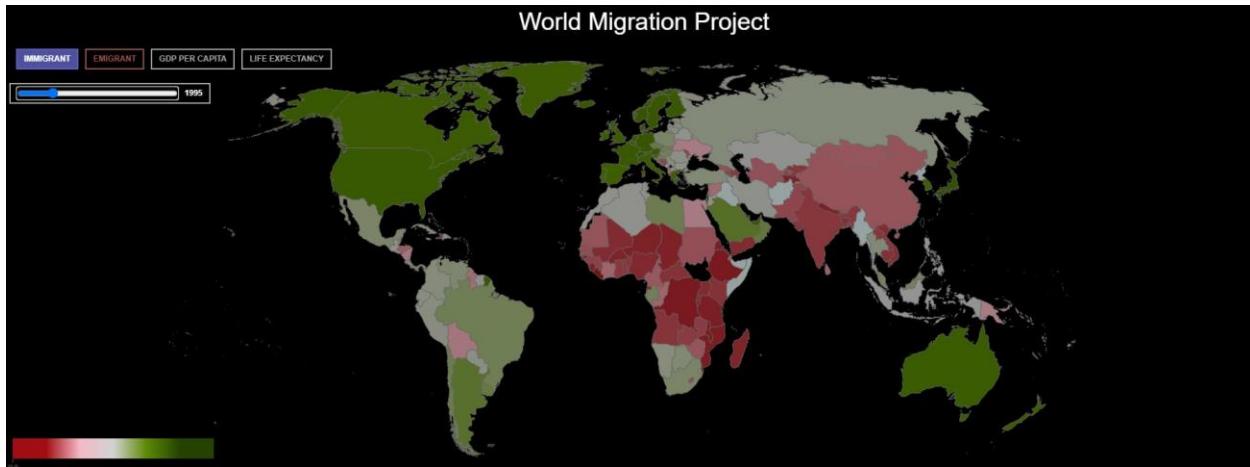


FIGURE 14: SECOND ITERATION OF THE WORLD MAP

With suggestions from Dr. Hoang Xuan Tung, the idea was to use a single World Map to display the flow of immigration and emigration using colored lines, flowing in or out of a selected country. Options are displayed at the top-left of the World Map to prompt the users between viewing the immigrant flow, which would be represented by a curved line flowing in from other countries, with the emigrant being represented in the same way. At this stage, we were still working on a simple way of showing the difference between the in-flow of immigrants and the out-flow of immigrants. A simple color scale between red and green was used to represent the differences in **GDP per capita**, and **Life expectancy**. However, there was a problem as our team felt the lack of supporting charts, as the website lacked basic information about the said World Map, so we were prompted to raise concerns during the next Stand-Up meeting.

4.3. Materialize Stage:

Between the two Stand-Up meetings 2 and 3, our team was able to add additional user controls to help navigate the world map, while changing the color scale to better represent the changes in the data (i.e. GDP per capita and Life expectancy). The result was the following figure.



FIGURE 15: THIRD ITERATION OF THE WORLD MAP

With Stand-Up Meeting 3 coming to an end, we decided to add another supporting visualization, a simple line graph that shows the net immigrant/emigrant rate of a country. With the idea of helping the users grasp the concept of immigrant and emigrant better, we feel that the said supporting visualization would provide ample support, displaying numbers that help users to conceive the flow of immigrants and emigrants better from the World Map alone.

4.4. Implementation Stage:



FIGURE 16: FOURTH ITERATION OF THE WORLD MAP

The final design for our Stand-Up Meeting 4 is displayed above, with almost every feature we aim to create, the map contains everything needed for a user to navigate and use the World Map effectively. The

newly implemented user interface showcases a Migration Data showcase of selected countries, which would display every country immigrating or emigrating relative to the selected country.



FIGURE 17: EXAMPLE OF THE NEW USER INTERFACE (U.S.A. IN 1995)

With the previously mentioned supporting visualization, we have found a way to implement it. By selecting a country on the World Map, a second visualization will appear beneath the World Map. This visualization aims to help users understand more about the numbers of immigrants compared to emigrants of a country.

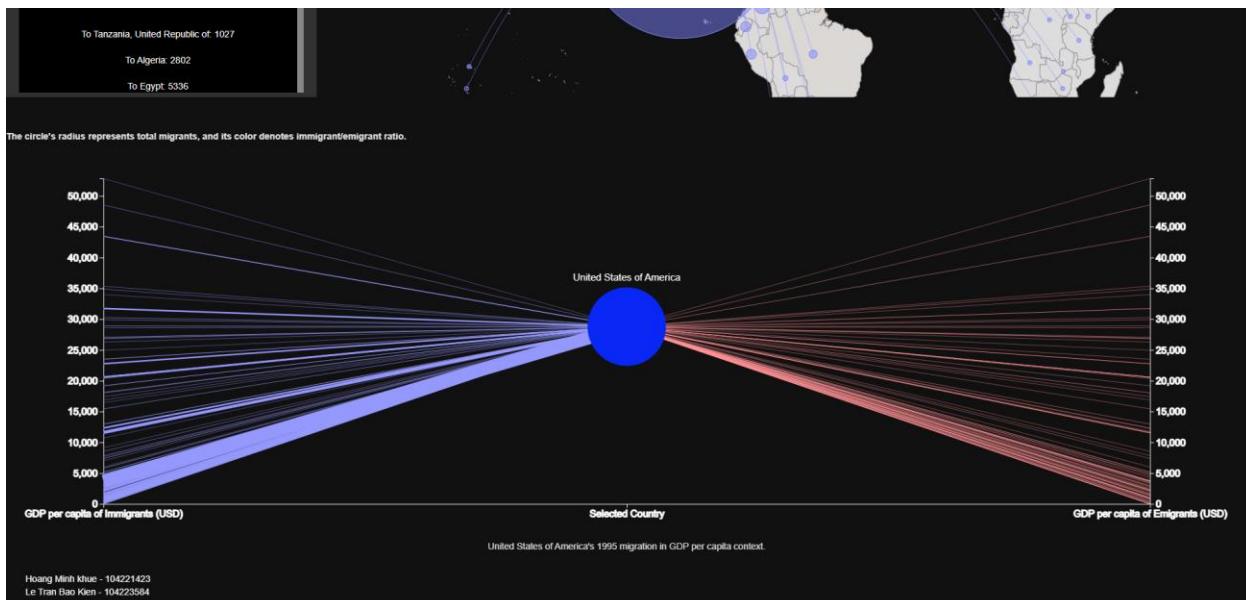


FIGURE 18: EXAMPLE OF THE SUPPORTING VISUALIZATION (U.S.A. IN 1995)

With the final product completed, we felt the need to help users find countries easier, since for some, locating countries through geographical is not an easy task, so we opted to implement a search bar function, to help users find specific countries.

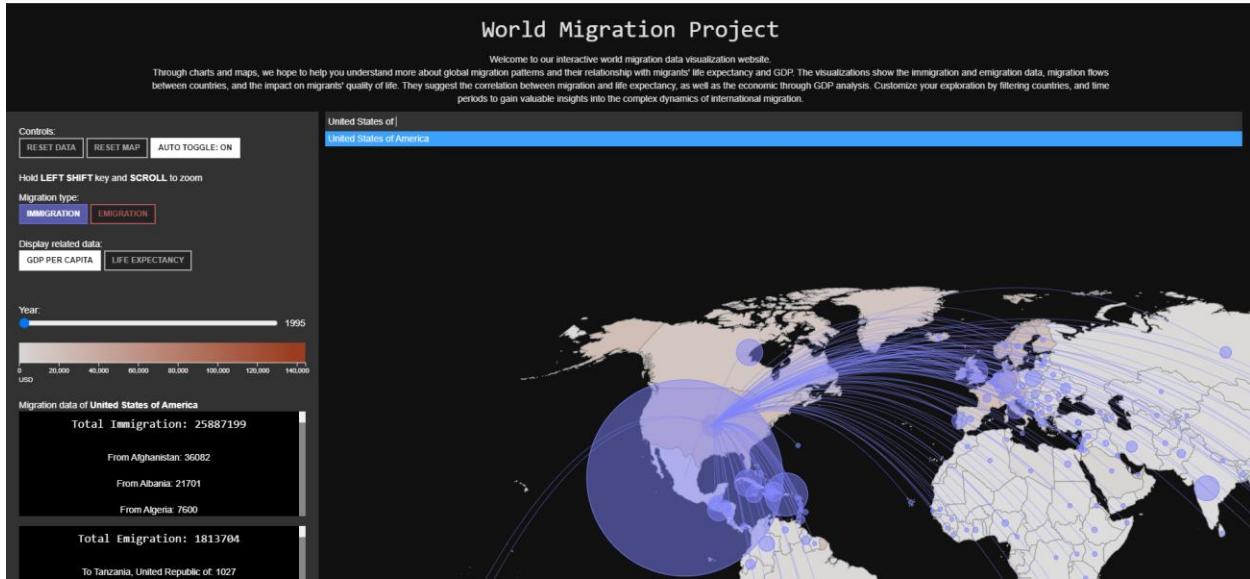


FIGURE 19: IMPLEMENTATION OF THE SEARCH FUNCTION, FINAL ITERATION OF THE WORLD MAP

5. Final Design

The final version of our data visualization set is shown below; you can also view it by clicking the provided Mercury link (mercury.swin.edu.au/cos30045/s104221423/WorldMigrationProject/).

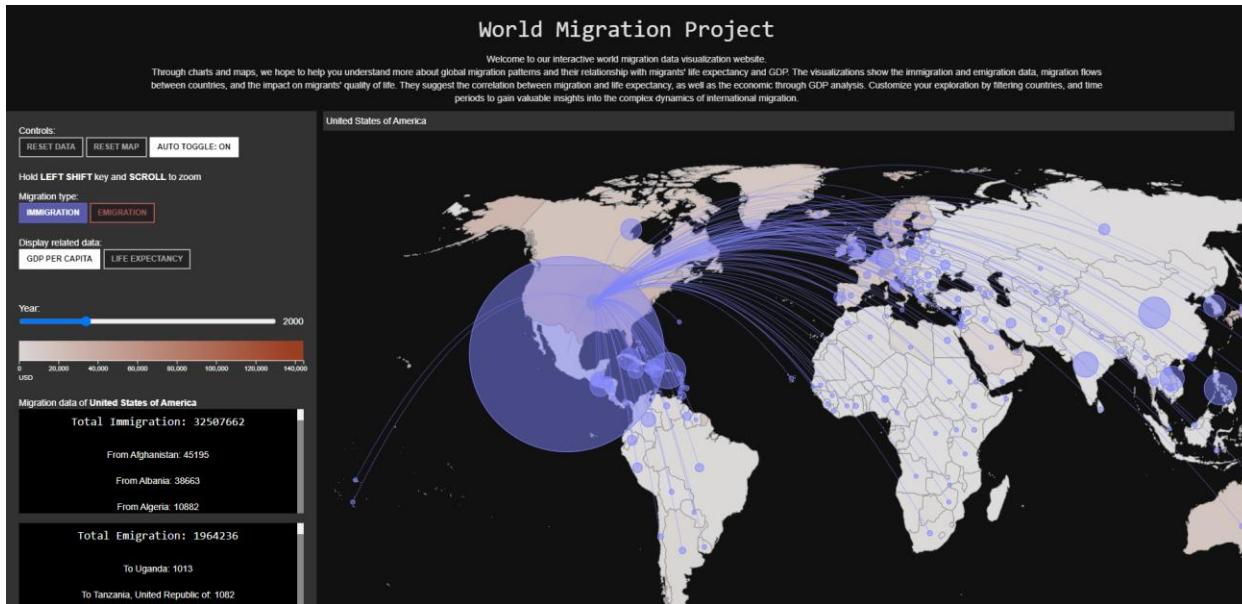


FIGURE 20: FINAL DESIGN OF THE WEBSITE 1

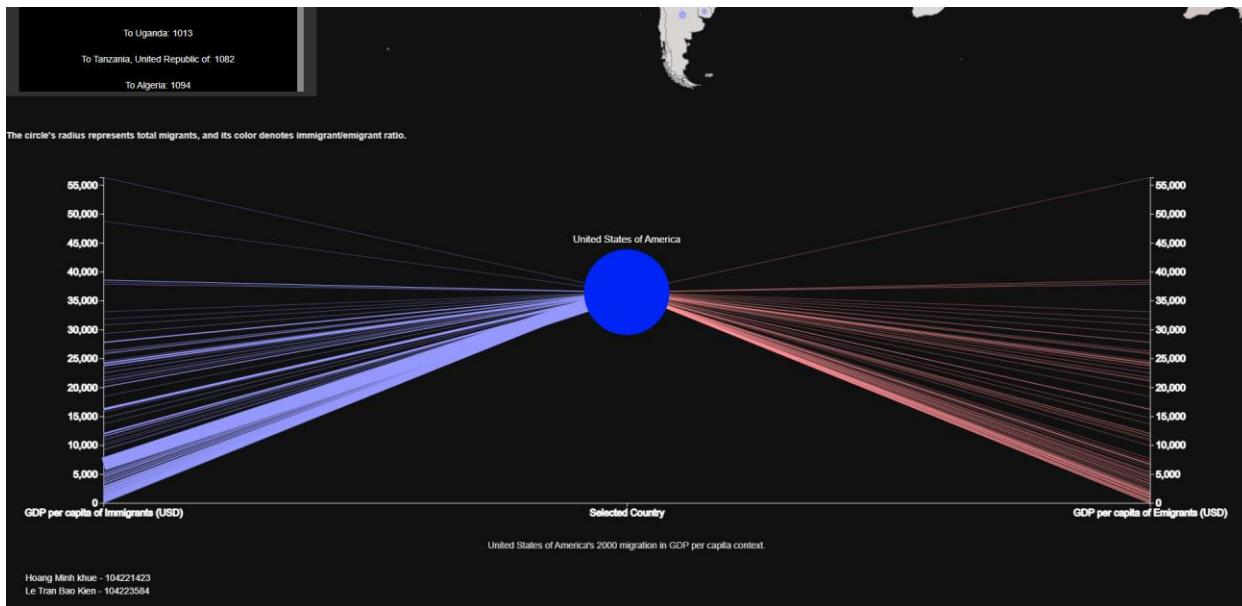


FIGURE 21: FINAL DESIGN OF THE WEBSITE 2

6. Validation

In addition, the user experience validation phase was conducted to gather valuable insights and feedback from participants regarding the usability and effectiveness of the visualization on global migration, life expectancy, and GDP per capita. A total of 24 participants took part in the user test, representing a diverse range of backgrounds and expertise. This section presents the findings and analysis of the user test, focusing on the participants' experiences, observations, and recommendations. The validation process aims to ensure that the visualization meets the needs of its intended user and provides an engaging and informative user experience. The insights gained from this phase can lead to further improvements of the overall usability and effectiveness of our project.

The following is a list of the questions that I asked in the form:

- The user's experience with UI/UX and Data Visualization:
 - o What is your experience with UI/UX and design patterns? (scale 0 - 10)
 - o What is your experience with Data Visualization? (scale 0 - 10)
- The user's experience with our project:
 - o How intuitive was it to interact with the visualization? (scale 0 - 10)
 - o Which features or functionalities of the visualization did you find most useful or engaging? (select boxes)
 - o Which features or functionalities of the visualization did you find most useless? (select boxes)
 - o How engaging is the visualization? (scale 0 – 10)
 - o Were there any features or functionalities that you felt were missing or could be improved? (text entry)
 - o To what extent does the visualization answer your question regarding the world migration topic? (scale 0 – 10)
 - o Did the visualization effectively convey the global migration trends and their relationship with life expectancy and GDP per capita? (Yes – No)
 - o On a scale of 1 to 5, how would you rate the overall usability of the visualization? (scale 1 - 5)
 - o On a scale of 1 to 5, how visually appealing did you find the visualization? (scale 1 – 5)
 - o Would you recommend this visualization to others? (Yes – No)

6.1. Participants' experience with UI/UX and Data Visualization

What is your experience with UI/UX and design patterns?

24 câu trả lời

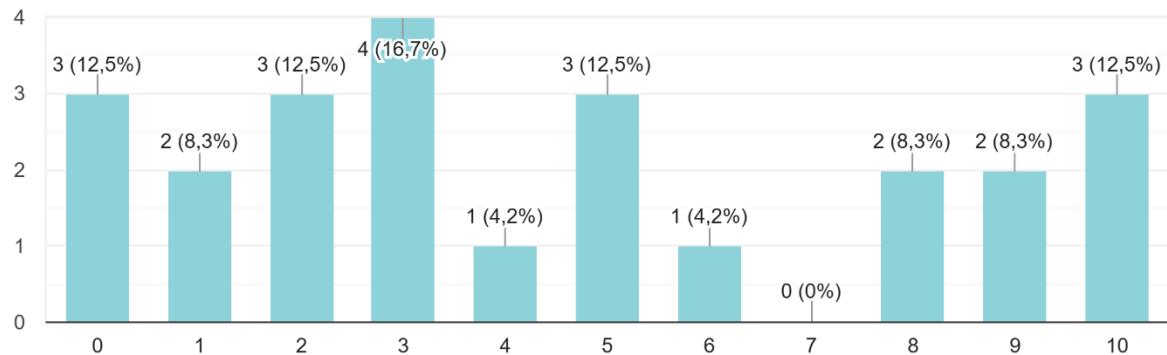


FIGURE 22: USERS' EXPERIENCE WITH UI/UX RESPONSE

What is your experience with Data Visualization?

24 câu trả lời

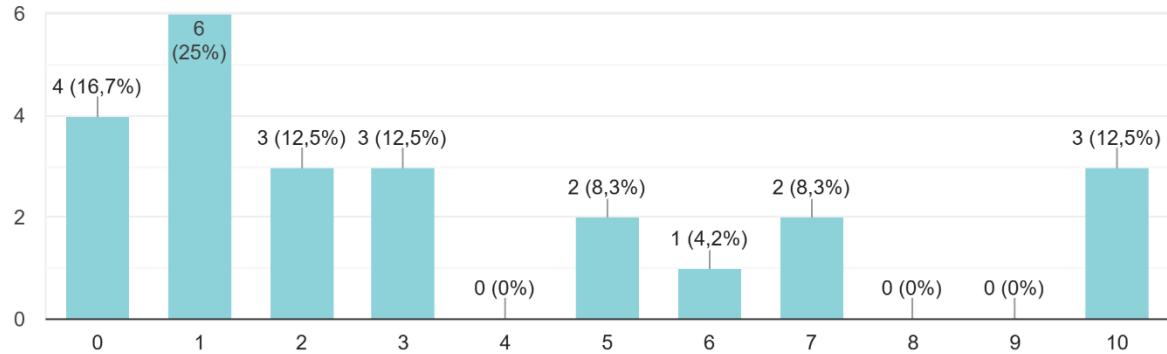


FIGURE 23: USERS' EXPERIENCE WITH DATA VISUALIZATION RESPONSE

It can be observed that a significant portion of the participants indicated a relatively low level of familiarity with both UI/UX and data visualization. For UI/UX and design patterns, 12.5% of the participants reported a score of 0, indicating no prior experience in this domain. Similarly, for data visualization, 16.7% of the participants reported a score of 0, indicating no previous exposure to data visualization concepts. These findings suggest that the majority of participants are relatively unfamiliar with UI/UX principles and data visualization techniques. It is important to consider this limited familiarity when interpreting their feedback and recommendations. The insights provided by these participants can offer an unbiased perspective and highlight areas where the visualization can be improved to cater to users with diverse levels of expertise.

These varying levels of experience with UI/UX, design patterns, and data visualization will provide valuable context when analyzing the participants' feedback on the usability and effectiveness of the visualization. The insights gained from individuals with different backgrounds and expertise will

contribute to a comprehensive understanding of the user experience and guide improvements in the visualization's design and functionality.

6.2. Participants' experience with our project

How intuitive was it to interact with the visualization?

24 câu trả lời

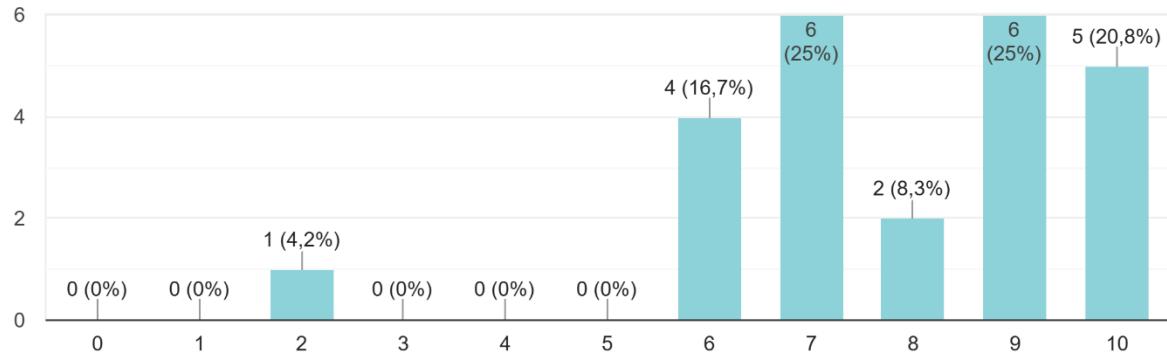


FIGURE 24: USERS' RESPONSE TO INTUITIVENESS OF THE PROJECT

The responses received for the question "How intuitive was it to interact with the visualization?" indicate that a majority of participants found the interaction with the visualization to be relatively intuitive, with ratings ranging from 6 to 10. However, it is worth noting that a small percentage of participants did not find the interaction to be intuitive, as they didn't provide ratings in the higher range. This suggests that there may be room for improvement in terms of enhancing the overall intuitiveness of the visualization's interaction design.

Which features or functionalities of the visualization did you find most useful or engaging?

23 câu trả lời

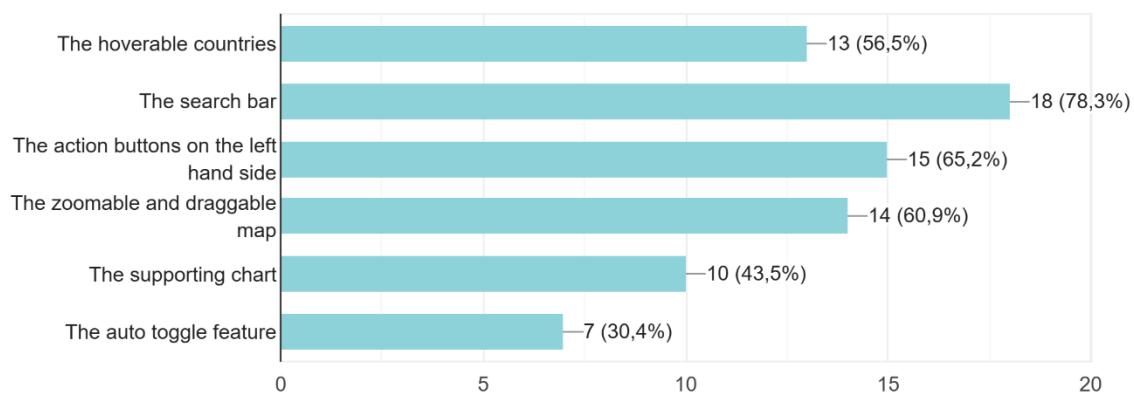


FIGURE 25: USERS' RESPONSE TO USEFUL FEATURES

Which features or functionalities of the visualization did you find most useless?

12 câu trả lời

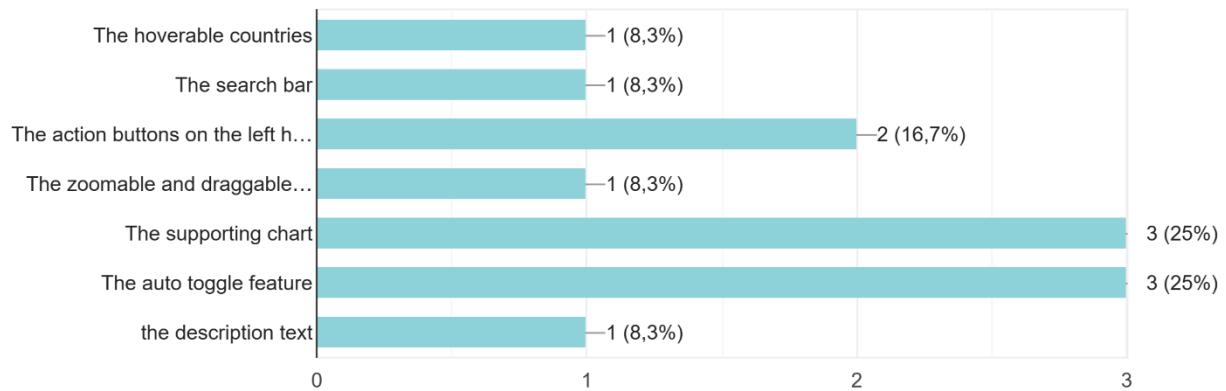


FIGURE 26: USERS' RESPONSE TO USELESS FEATURES

23 responses received for the question "Which features or functionalities of the visualization did you find most useful or engaging?" highlight the features that resonated most with the participants. The majority found the "search bar" feature engaging, indicating its effectiveness in capturing users' attention. The "hoverable countries" and "action buttons" were considered useful by a significant percentage, suggesting their contribution to the overall functionality of the visualization. The "zoomable and draggable map" feature attracted the interest of a majority of participants, while the "supporting chart" was found useful by a considerable proportion. However, the "auto toggle feature" did not generate as much engagement as the other features.

Only 12 responses for the opposite question which suggest that a small percentage of participants did not find certain features of the visualization to be useful. The "supporting chart" and "auto toggle feature" received a higher percentage of respondents considering them useless, while other features had relatively lower percentages. It's important to note that these opinions represent the views of a subset of participants and may not reflect the overall user sentiment.

How engaging is the visualization?

24 câu trả lời

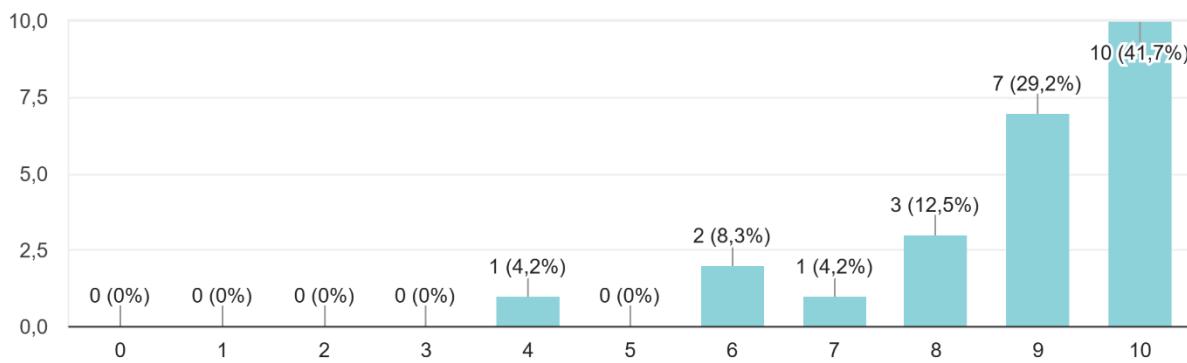


FIGURE 27: USERS' RESPONSE TO ENGAGEMENT

Based on the responses received for the question "How engaging is the visualization?", it is clear that the majority of participants found the visualization to be highly engaging, with a significant

percentage giving it a rating of 9 or 10. Ratings from 6 to 8 were less common, while lower ratings (4 and below) were provided by a smaller percentage of participants.

For the question “Were there any features or functionalities that you felt were missing or could be improved?”, there are 2 valid suggestions. One states that the coloring of our visualization should be improved and the other says that we should include more numeric detail for the migration data.

To what extent does the visualization answer your question regarding the world migration topic?

24 câu trả lời

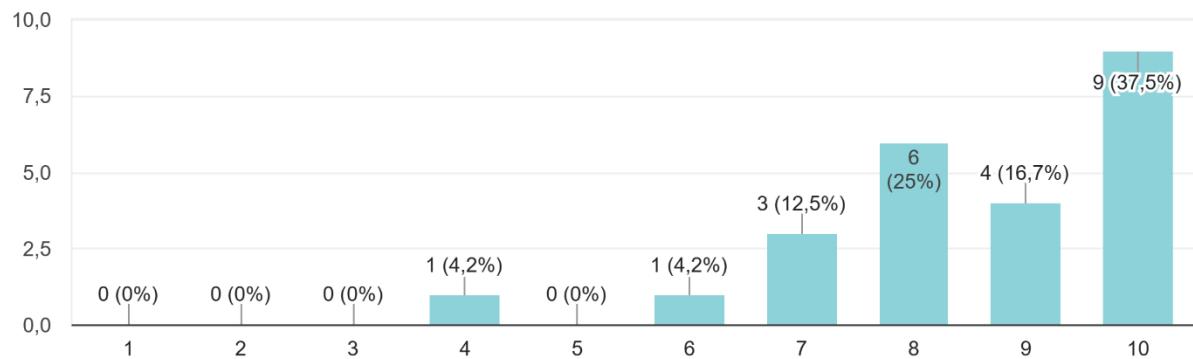


FIGURE 27: USERS’ RESPONSE TO INFORMATIVENESS

The responses to the question “To what extent does the visualization answer your question regarding the world migration topic?” indicate that the majority of participants found that the visualization effectively answered their questions regarding the world migration topic. A significant percentage rated it as 9 or 10, indicating a high level of satisfaction with the visualization’s ability to address their queries. Ratings of 7 and 8 were also relatively common, while lower ratings (4 and 6) were provided by a smaller percentage of participants.

Did the visualization effectively convey the global migration trends and their relationship with life expectancy and GDP per capita?

24 câu trả lời

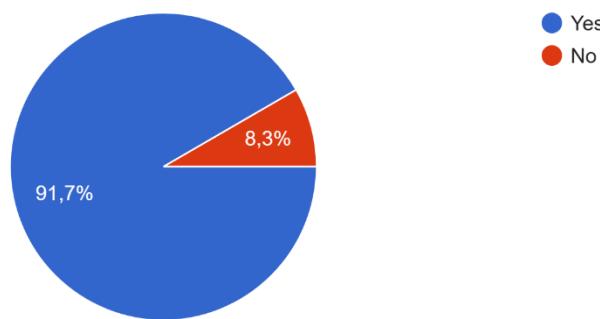


FIGURE 28: USERS’ RESPONSE TO RELATIONS BETWEEN DATA

The feedback indicates that the visualization was successful in effectively conveying the global migration trends and their relationship with life expectancy and GDP per capita for the majority of participants. The high percentage of participants who answered "YES" suggests that the visualization provided clear and understandable information, allowing users to grasp the connections between

migration, life expectancy, and GDP per capita. However, it's important to note that a small percentage of participants (8.3%) responded with "NO," indicating that they did not feel the visualization effectively conveyed these relationships.

On a scale of 1 to 5, how would you rate the overall usability of the visualization?

24 câu trả lời

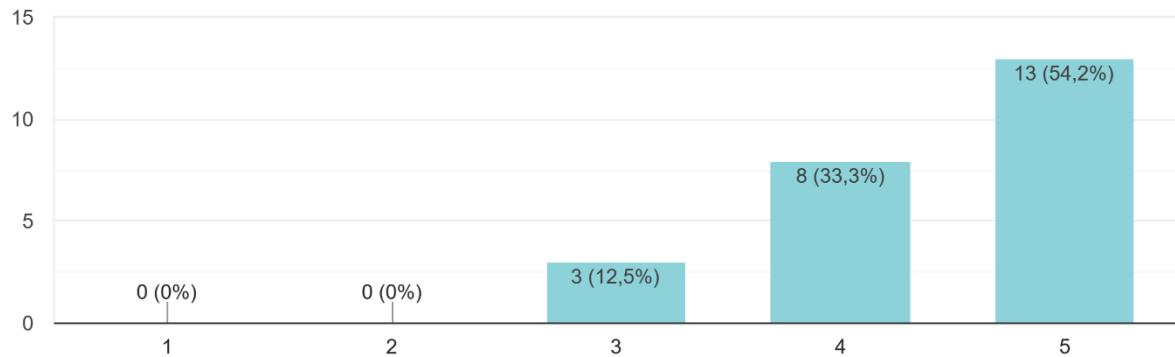


FIGURE 29: USERS' RESPONSE USABILITY

The responses indicate that the majority of participants found the overall usability of the visualization to be highly satisfactory. The highest percentage of participants, 54.2%, rated it as 5, indicating a high level of usability. Additionally, a significant portion of participants, 33.3%, rated it as 4, suggesting that they still found it to be quite usable.

On a scale of 1 to 5, how visually appealing did you find the visualization?

24 câu trả lời

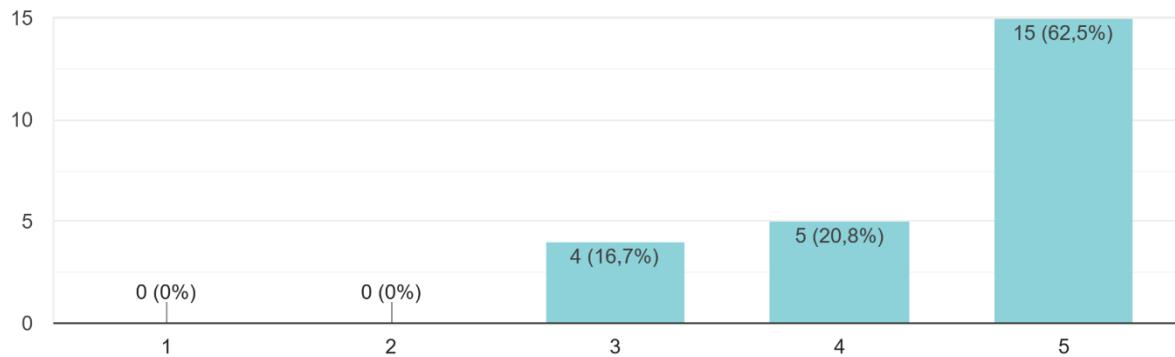


FIGURE 30: USERS' RESPONSE TO APPEARANCE

The majority of participants found the visualization to be visually appealing. The highest percentage of participants, 62.5%, rated it as 5, indicating a high level of visual appeal. Additionally, a significant portion of participants, 20.8%, rated it as 4, suggesting that they still found it visually appealing.

Would you recommend this visualization to others?

24 câu trả lời

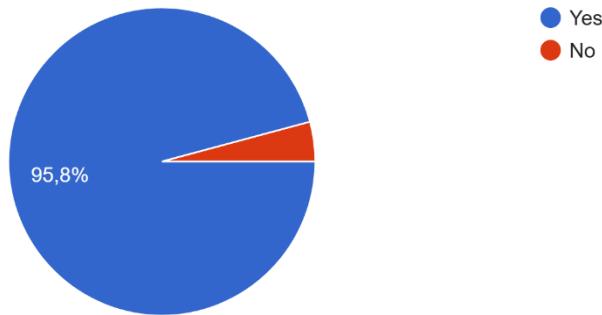


FIGURE 31: USERS' RESPONSE TO RECOMMENDATION TO OTHERS

Most of the participants would likely to recommend our visualization to others

7. Conclusion

In today's globalized world, migration—especially cross-country migration—has become extremely common. There are pull and push factors at play in this trend. We extracted, processed, analyzed, and visualized the global migration to highlight the aforementioned phenomena. We selected two distinct kinds of diagrams. A world map is utilized to illustrate and contrast the migration flow, GDP per capita, and life expectancy; a line chart is employed to illustrate the variations in the migration numbers. The JavaScript D3.js library was used to plot the data. The essential features for both visualizations are interactive visualization, extra background information, a distinct data label, and contrasting colors for the various data points. Lastly, usability testing was done to make sure the visualization is easy to use, effectively conveys information, and meets the needs of the intended audience.

References

- Siddiqui, T. (2012). Impact of migration on poverty and development. Brighton: University of Sussex.
- Hammar, T., Brochmann, G., Tamas, K., & Faist, T. (Eds.). (2021). International migration, immobility and development: Multidisciplinary perspectives. Routledge.
- Castles, S. (2010). Understanding Global Migration: A Social Transformation Perspective. *Journal of Ethnic and Migration Studies*, 36(10), 1565-1586
- Faist, T. (2013). The Mobility Turn: A New Paradigm for the Social Sciences? *Ethnic and Racial Studies*, 36(11), 1637-1646.
- Zambotti, G., Guan, W., & Gest, J. D. (2015). Visualizing human migration through space and time. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*
- Bertin, J. (2011). *Semiology of graphics: Diagrams, networks, maps*. Esri Press.
- van Biljon, J., & Osei-Bryson, K. M. (2020). The communicative power of knowledge visualizations in mobilizing information and communication technology research. *Information Technology for Development*, 26(4), 637-652.
- Sander, N., Abel, G. J., Bauer, R., & Schmidt, J. (2014). Visualising migration flow data with circular plots (No. 2/2014). Vienna Institute of Demography Working Papers.