Data Visualisation

Assignment 2

A dive into Ukraine Forced Migration in 2010-2023

Monday 8:00

Semester 1, 2024

Word count: 1907

Nguyen Duy Anh- 104212294

Lai Gia Khanh- 104222015

https://mercury.swin.edu.au/cos30045/s104222015/project/project.html

Table of Contents

Contents

1. Introduction 3
   1.1. Background and Motivation 3
   1.2. Visualization Purpose 3
   1.3. Project Schedule 3
2. Data 4
   2.1. Data Source 4
   2.2. Data Processing 6
3. Requirements 9
   3.1. Must-Have Features 9
   3.2. Optional Features 10
4. Visualizations Design 10
   4.1. Scatter Plot 10
   4.2. Line plot 12
   4.3. Bar chart 14
5. Conclusion 15

References 16

1. Introduction

1.1. Background and Motivation

Human migration has been a significant phenomenon in Ukraine from 2010 to 2023, marked by dynamic shifts influenced by various socioeconomic, political, and environmental factors. Understanding the complexities and patterns of migration in Ukraine during this period requires comprehensive analysis and visualization. This report aims to provide an overview of forced migration issues in Ukraine from 2010 to 2023, where they went, country and region and the trend of forced migration in the period.

1.2. Visualization Purpose

With our provided visualizations, users will be able to answer various questions about human migration issues in the world such as:

- Forced migration trend accross regions of the world from 2010 to 2023
- Forced migration of Ukraine in the last decade
- Pattern of movement of Ukraine

1.3. Project Schedule

Week 1-2: Start the project:

- Form the group.
- Create the GitHub repository.
- Create a communication channel (Discord server).

Week 3-5: Prepare for the project:

- Looking for information about migration, doing a literature review.
- Doing research and analyzing datasets about migration.
- Decide the visualizations that are suitable for the found dataset.

Week 6-7: Start coding the website and visualizations:

- Using HTML to build the structure of the website.
- Using JS to draw the graphs based on the found datasets.

Week 8: Finish the website:

- Using CSS to improve the interface of the website.
- Add some transitions or effects to the visualizations.

Week 9-10: Writing the Project Process Book:

- Complete the content of the Process Book based on the prepared information.
- Fix the grammar and spelling mistakes if they exist.

Week 11: Make improvements:

- Receive feedback from the instructor to fix all problems with the Process Book and the website. Week 12: Presentation:
- Submit the project.
- Present and interview with the instructor.

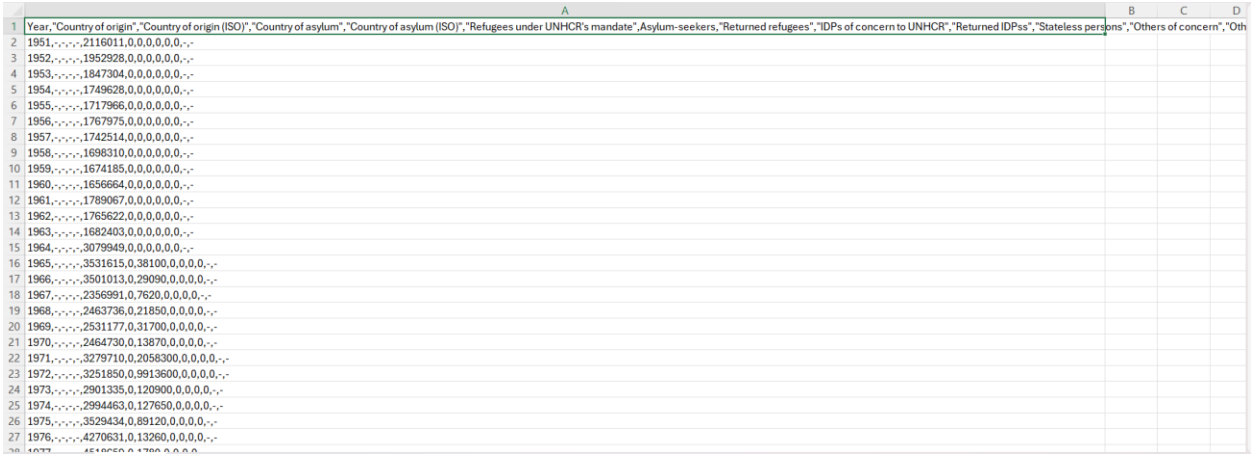
2. Data

2.1. Data Source

All the data for this project come from the UN, particularly, the UNHCR and UN Department of Economic and Social Affairs. There are three datasets used for this project which can be found here:

https://www.unhcr.org/refugee-statistics/insights/explainers/forcibly-displaced-flow-data.html

https://www.unhcr.org/refugee-statistics/download/?url=bf8OWP



	A	B	C	D
1	Year,"Country of origin","Country of origin (ISO)","Country of asylum","Country of asylum (ISO)","Refugees under UNHCR's mandate",Asylum-seekers,"Returned refugees","IDPs of concern to UNHCR","Returned IDPss","Stateless persons","Others of concern","Oth			
2	1951,-,-,-,-,2116011,0,0,0,0,0,0,-,-			
3	1952,-,-,-,-,1952928,0,0,0,0,0,0,-,-			
4	1953,-,-,-,-,1847304,0,0,0,0,0,0,-,-			
5	1954,-,-,-,-,1749628,0,0,0,0,0,0,-,-			
6	1955,-,-,-,-,1717966,0,0,0,0,0,0,-,-			
7	1956,-,-,-,-,1767975,0,0,0,0,0,0,-,-			
8	1957,-,-,-,-,1742514,0,0,0,0,0,0,-,-			
9	1958,-,-,-,-,1698310,0,0,0,0,0,0,-,-			
10	1959,-,-,-,-,1674185,0,0,0,0,0,0,-,-			
11	1960,-,-,-,-,1656664,0,0,0,0,0,0,-,-			
12	1961,-,-,-,-,1789067,0,0,0,0,0,0,-,-			
13	1962,-,-,-,-,1765622,0,0,0,0,0,0,-,-			
14	1963,-,-,-,-,1682403,0,0,0,0,0,0,-,-			
15	1964,-,-,-,-,3079949,0,0,0,0,0,0,-,-			
16	1965,-,-,-,-,3531615,0,38100,0,0,0,0,-,-			
17	1966,-,-,-,-,3501013,0,29090,0,0,0,0,-,-			
18	1967,-,-,-,-,2356991,0,7620,0,0,0,0,-,-			
19	1968,-,-,-,-,2463736,0,21850,0,0,0,0,-,-			
20	1969,-,-,-,-,2531177,0,31700,0,0,0,0,-,-			
21	1970,-,-,-,-,2464730,0,13870,0,0,0,0,-,-			
22	1971,-,-,-,-,3279710,0,2058300,0,0,0,0,-,-			
23	1972,-,-,-,-,3251850,0,9913600,0,0,0,0,-,-			
24	1973,-,-,-,-,2901335,0,120900,0,0,0,0,-,-			
25	1974,-,-,-,-,2994463,0,127650,0,0,0,0,-,-			
26	1975,-,-,-,-,3529434,0,89120,0,0,0,0,-,-			
27	1976,-,-,-,-,4270631,0,13260,0,0,0,0,-,-			

Fig1. Dataset 1 original

As you can see from figure 1, the dataset tracks refugee-related metrics from 1951 to 2023, including countries of origin and asylum, numbers of refugees, asylum-seekers, internally displaced persons (IDPs), stateless persons, and others of concern to the UNHCR. It provides insights into global refugee movements, returns, and international protection needs over time. For the purpose of put project, the dataset will be clean up to only contain data of forcibly displaced population which contain: “Refugees”, “Asylum-seekers IDPs of concern to UNHCR” and “Other people in need of international protection”, the “Country of Asylum will be filter to Region and the timespan of the data will be limited from 2010 to 2023

For the second dataset, it contains 3 tables: Summary of migrant by years, Summary of Asylum and the last one contain the total data of migrant across the world from the origin country to destination country and amount of migrant from 1962 to 2023. For the second visualization, we will use the last table as our dataset. This dataset comprises information on asylum seekers originating from different countries seeking refuge in various destination countries. It includes categorical attributes such as the country of origin, its ISO code, and name, along with corresponding details for the destination country. Additionally, asylum region, type of protection sought, year of asylum application, and the count of asylum seekers are provided. These data, spanning from 1962 onwards, offer insights into the flow of asylum seekers across different regions globally, aiding in the analysis of migration patterns and refugee movements over time.

	A	B	C	D	E	F	G	H	I	J
1	origin	OriginISO	OriginName	asylum	AsylumISO	AsylumName	AsylumRegion	PT	Year	Count
2	ALG	DZA	Algeria	MTA	MLT	Malta	Europe	REF	1962	5
3	ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1962	20.000
4	ANG	AGO	Angola	NAM	NAM	Namibia	Southern Africa	REF	1962	277
5	BDI	BDI	Burundi	NAM	NAM	Namibia	Southern Africa	REF	1962	13
6	CHI	CHN	China	NEP	NPL	Nepal	Asia and the Pacific	REF	1962	5
7	COB	COG	Congo, Republic of	NAM	NAM	Namibia	Southern Africa	REF	1962	8
8	COD	COD	Dem. Rep. of the Congo	NAM	NAM	Namibia	Southern Africa	REF	1962	13
9	IRQ	IRQ	Iraq	MTA	MLT	Malta	Europe	REF	1962	17
10	SOM	SOM	Somalia	NAM	NAM	Namibia	Southern Africa	REF	1962	5
11	SUD	SDN	Sudan	MTA	MLT	Malta	Europe	REF	1962	10
12	TIB	TIB		NEP	NPL	Nepal	Asia and the Pacific	REF	1962	2.236
13	ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	5.000
14	ARE	EGY	Egypt	LEB	LBN	Lebanon	Middle East and North Africa	REF	1963	160
15	ROM	ROU	Romania	LEB	LBN	Lebanon	Middle East and North Africa	REF	1963	70
16	RWA	RWA	Rwanda	BDI	BDI	Burundi	East and Horn of Africa, and Great Lakes	REF	1963	4.000
17	RWA	RWA	Rwanda	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	1.000
18	RWA	RWA	Rwanda	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1963	2.000
19	SUD	SDN	Sudan	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	8.000
20	SUD	SDN	Sudan	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1963	5.000
21	UKN	UKN	Unknown/other	UKN	UKN	Unknown/other		REF	1963	40
22	ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1964	25.000
23	ARM	ARM	Armenia	ARE	EGY	Egypt	Middle East and North Africa	REF	1964	90
24	COB	COG	Congo, Republic of	BDI	BDI	Burundi	East and Horn of Africa, and Great Lakes	REF	1964	20.000
25	COB	COG	Congo, Republic of	CAR	CAF	Central African Rep.	West and Central Africa	REF	1964	3.000
26	COB	COG	Congo, Republic of	TAN	TZA	United Rep. of Tanzania	East and Horn of Africa, and Great Lakes	REF	1964	1.700
27	COB	COG	Congo, Republic of	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1964	34.000
28	GUI	GIN	Guinea	SEN	SEN	Senegal	West and Central Africa	REF	1964	24.000
29	MOZ	MOZ	Mozambique	TAN	TZA	United Rep. of Tanzania	East and Horn of Africa, and Great Lakes	REF	1964	10.000

Fig2. Dataset 2 original

2.2. Data Processing

The data cleanup process is an essential part of the project. We aim to remove redundant data and make sure it fits the overall aim of the visualization. Data are extracted with accuracy and an important aspect is to simplify the final dataset in order to have an easy coding process.

In the first dataset, we have query tool from the UN website to make the process easier. In the query, we only choose the demographic to be Forcibly displaced population, limited the time from 2010 to 2023. From the dataset, we further process it by grouping country into region. An id value are also added for each datapoint for the coding process. You can see the final data in figure 3.

	A
1	Id,Year,Region,Count,Distance
2	1,2010,Asia and the Pacific,52,5500
3	2,2011,Asia and the Pacific,46,5500
4	3,2012,Asia and the Pacific,42,5500
5	4,2013,Asia and the Pacific,34,5500
6	5,2014,Asia and the Pacific,138,5500
7	6,2015,Asia and the Pacific,190,5500
8	7,2016,Asia and the Pacific,152,5500
9	8,2017,Asia and the Pacific,172,5500
10	9,2018,Asia and the Pacific,214,5500
11	10,2019,Asia and the Pacific,252,5500
12	11,2020,Asia and the Pacific,191,5500
13	12,2021,Asia and the Pacific,178,5500
14	13,2022,Asia and the Pacific,4772,5500
15	14,2023,Asia and the Pacific,4671,5500
16	15,2010,Europe,22515,580
17	16,2011,Europe,22702,580
18	17,2012,Europe,22765,580
19	18,2013,Europe,3017,580
20	19,2014,Europe,247496,580
21	20,2015,Europe,336311,580
22	21,2016,Europe,257339,580
23	22,2017,Europe,154192,580
24	23,2018,Europe,105165,580
25	24,2019,Europe,70189,580
26	25,2020,Europe,41914,580
27	26,2021,Europe,30473,580

Fig3. Dataset 1 after processing.

The second dataset are much more complicated. As mentioned before, it contains comprises information on asylum seekers originating from different countries seeking refuge in various destination countries on the span of 61 years.

origin	OriginISO	OriginName	asylum	AsylumISO	AsylumName	AsylumRegion	PT	Year	Count
ALG	DZA	Algeria	MTA	MLT	Malta	Europe	REF	1962	5
ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1962	20.000
ANG	AGO	Angola	NAM	NAM	Namibia	Southern Africa	REF	1962	277
BDI	BDI	Burundi	NAM	NAM	Namibia	Southern Africa	REF	1962	13
CHI	CHN	China	NEP	NPL	Nepal	Asia and the Pacific	REF	1962	5
COB	COG	Congo, Republic of	NAM	NAM	Namibia	Southern Africa	REF	1962	8
COD	COD	Dem. Rep. of the Congo	NAM	NAM	Namibia	Southern Africa	REF	1962	13
IRQ	IRQ	Iraq	MTA	MLT	Malta	Europe	REF	1962	17
SOM	SOM	Somalia	NAM	NAM	Namibia	Southern Africa	REF	1962	5
SUD	SDN	Sudan	MTA	MLT	Malta	Europe	REF	1962	10
TIB	TIB		NEP	NPL	Nepal	Asia and the Pacific	REF	1962	2.236
ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	5.000
ARE	EGY	Egypt	LEB	LBN	Lebanon	Middle East and North Africa	REF	1963	160
ROM	ROU	Romania	LEB	LBN	Lebanon	Middle East and North Africa	REF	1963	70
RWA	RWA	Rwanda	BDI	BDI	Burundi	East and Horn of Africa, and Great Lakes	REF	1963	4.000
RWA	RWA	Rwanda	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	1.000
RWA	RWA	Rwanda	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1963	2.000
SUD	SDN	Sudan	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1963	8.000
SUD	SDN	Sudan	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1963	5.000
UKN	UKN	Unknown/other	UKN	UKN	Unknown/other		REF	1963	40
ANG	AGO	Angola	COD	COD	Dem. Rep. of the Congo	Southern Africa	REF	1964	25.000
ARM	ARM	Armenia	ARE	EGY	Egypt	Middle East and North Africa	REF	1964	90
COB	COG	Congo, Republic of	BDI	BDI	Burundi	East and Horn of Africa, and Great Lakes	REF	1964	20.000
COB	COG	Congo, Republic of	CAR	CAF	Central African Rep.	West and Central Africa	REF	1964	3.000
COB	COG	Congo, Republic of	TAN	TZA	United Rep. of Tanzania	East and Horn of Africa, and Great Lakes	REF	1964	1.700
COB	COG	Congo, Republic of	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1964	34.000
GUI	GIN	Guinea	SEN	SEN	Senegal	West and Central Africa	REF	1964	24.000
MOZ	MOZ	Mozambique	TAN	TZA	United Rep. of Tanzania	East and Horn of Africa, and Great Lakes	REF	1964	10.000
RWA	RWA	Rwanda	BDI	BDI	Burundi	East and Horn of Africa, and Great Lakes	REF	1964	24.000
RWA	RWA	Rwanda	TAN	TZA	United Rep. of Tanzania	East and Horn of Africa, and Great Lakes	REF	1964	1.300
RWA	RWA	Rwanda	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1964	13.000
SUD	SDN	Sudan	CAR	CAF	Central African Rep.	West and Central Africa	REF	1964	300
SUD	SDN	Sudan	UGA	UGA	Uganda	East and Horn of Africa, and Great Lakes	REF	1964	7.000
UKN	UKN	Unknown/other	LEB	LBN	Lebanon	Middle East and North Africa	REF	1964	370
UKN	UKN	Unknown/other	UKN	UKN	Unknown/other		REF	1964	30

Fig3. Pre-cleanup dataset 2

To create a focused dataset analyzing migration patterns in Ukraine from 2010 to 2023, redundant columns such as country codes are excluded. The original dataset is filtered to include only records where Ukraine is the origin country, and the time range is narrowed down accordingly. Subsequently, redundant columns such as 'origin' and 'OriginISO', are removed to streamline the dataset. The processed data will then be split for 2 visualizations, one will serve the purpose of showing which country Ukraine migrate to the most and the other will show the fluctuation of Ukrainian refugee in the specific country. Finally, the cleaned dataset is exported as a CSV file and ready for further analysis to elucidate migration trends

```
1  Countries;Counts
2  Russian Federation;1789711
3  Czechia;492006
4  United Kingdom;214958
5  Germany;1111841
6  Poland;1081130
7  Other;1912217
```

Fig4. Post-cleanup Ukraine dataset for line bar chart

	Year	Count
2	2010	23
3	2011	11
4	2012	13
5	2013	19
6	2014	271237
7	2015	149142
8	2016	23534
9	2017	12044
10	2018	5877
11	2019	6056
12	2020	4726
13	2021	3612
14	2022	1310748
15	2023	6027

Fig5. Post-cleanup Ukraine refugee to Russia dataset overtime

3. Requirements

3.1. Must-Have Features

For our project, the requirement is to have a one-page website with HTML, CSS and d3.js for visualization. The webpage will include an introduction and visualizations and text:

- Scatter Plot: The scatter plot visualizes data pertaining to Ukrainian refugees displaced to various regions from 2010 to 2023. Each data point represents an observation, with the x-axis denoting the year of displacement, the y-axis indicating the distance traveled, the size of each point representing the count of refugees, and the color indicating the destination region.

- Bar chart: Use to show the most popular destination country for migration of Ukrainian citizen from the time frame of 2010-2023. When a bar chart represent one country if clicked, the user will be directed to a line chart that show number of Ukrainian refugees to that specific country overtime.
- Line Chart of selected countries: This visualization main purpose is to show the recent history of forced migration of chosen country.

In addition to charts, explanations will also be provided on why we chose this data, what happened in the period in the dataset and what we found out after analyzing.

3.2. Optional Features

- Connection Map: The Connection Map is a powerful tool designed to enhance the presentation and understanding of immigration patterns across the globe. This feature provides users with a visual representation of the interconnectedness of different regions through immigration flows. By leveraging a world map interface, users can explore and analyze the movement of people across borders, gaining valuable insights into migration trends and patterns.

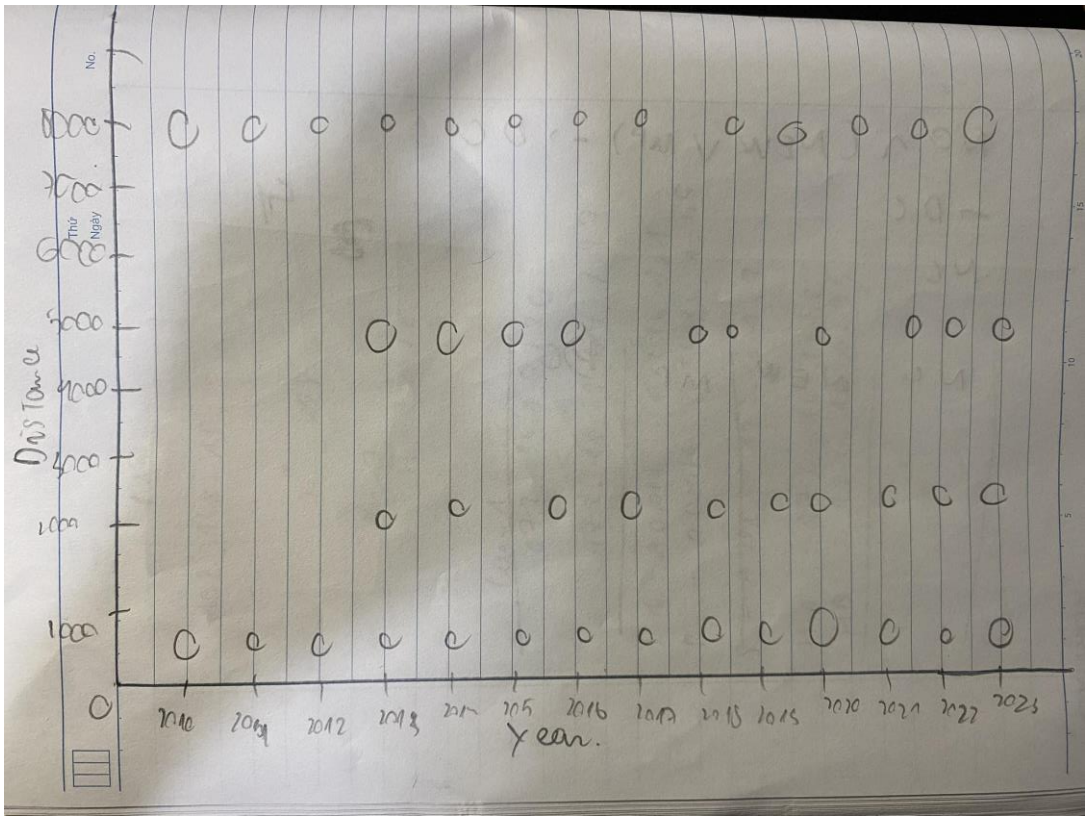
4. Visualizations Design

4.1. Scatter Plot

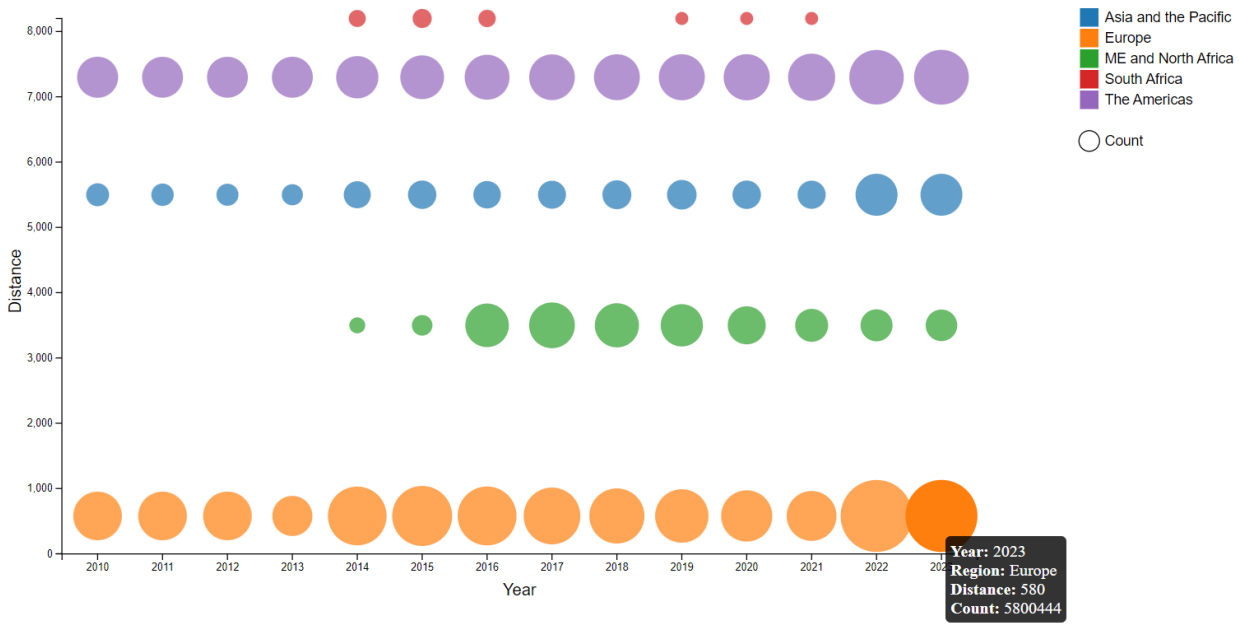
The idea of the Scatter plot is to show the forced migration of Ukraine to different region of the world from 2010 to 2023. Distance of travel between Ukraine and said continent will also be included.

- X axis will represent years from 2010 to 2023
- Y axis will represent Distant (km)
- Each bubble will represent a datapoint, the size of the dot will show how big the count(number of people migrating) is.

Sketch:



Final Design:



Justification:

In the final design, we changed the line plot with multiple groups to scatter plot. The scatter plot serves the same idea as the line plot but adding 1 more axis which is distance of travel.

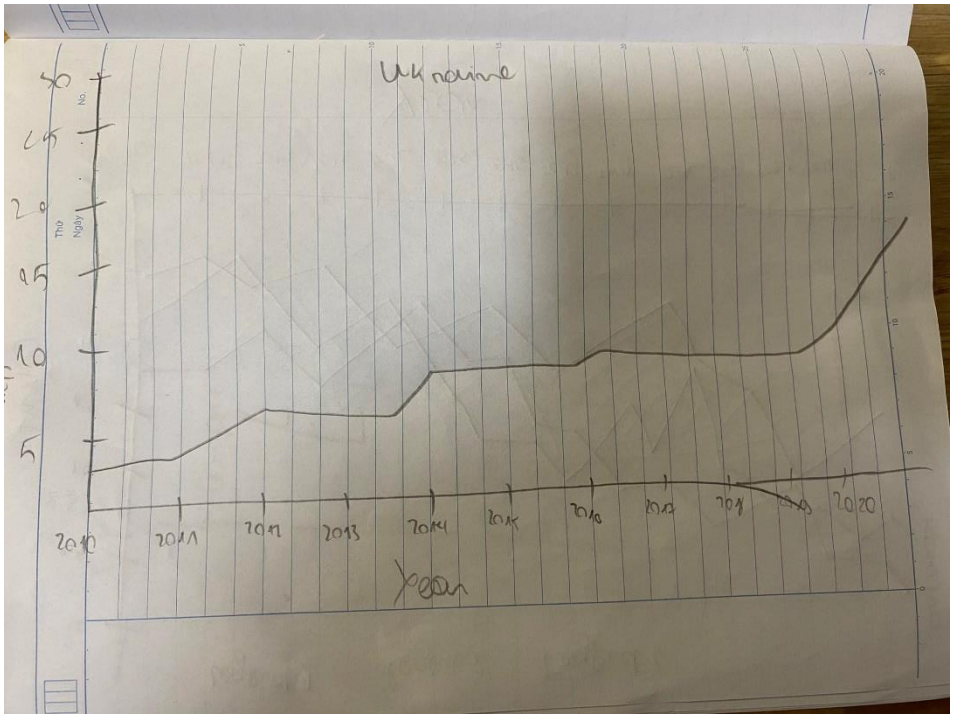
- The X axis shows years while the Y axis shows distance of travel.
- The size of the bubble shows the counts.
- By using a logarithmic scale, our chart can emphasize the difference in number of people, making it easier to spot the change in counts.
- Using different colors for different regions will help the user understand more easily.
- Legend will be implemented for clarity.
- Tooltips added when hovering over each bubble for more details.

4.2. Line plot

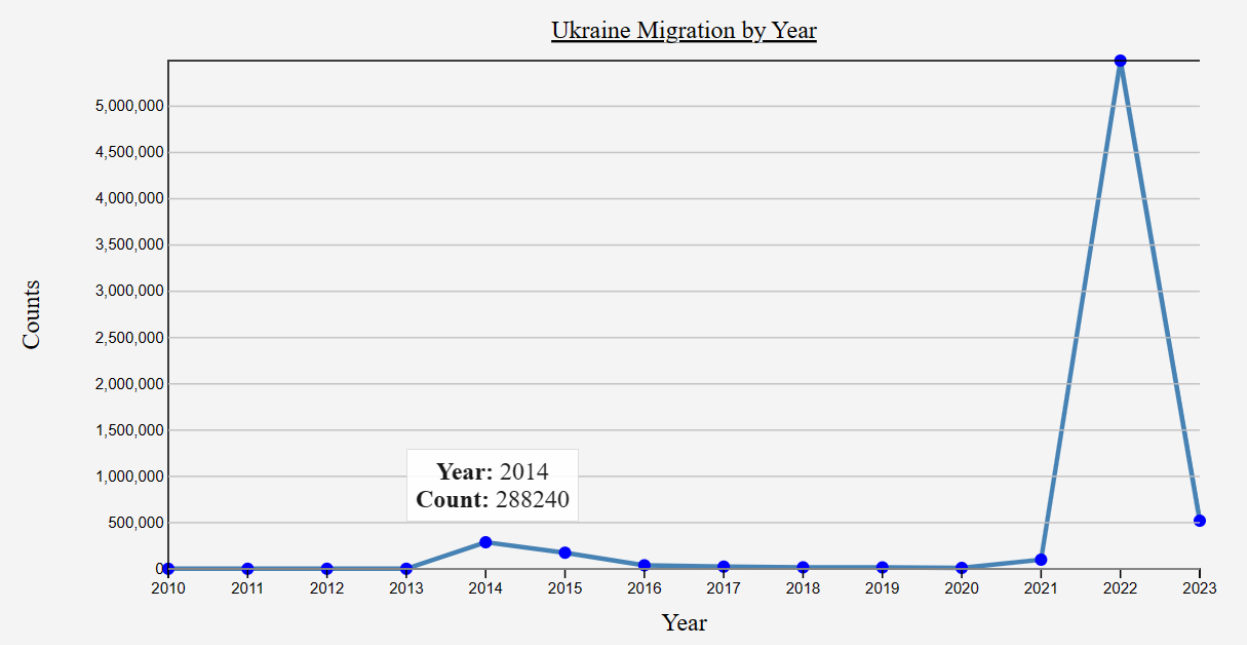
This graph is used to depict the change of forced migration of a country from 2010 to 2023.

- X axis will represent years from 2010 to 2023
- Y axis will represent Count(people)
- Dots will be used to indicate the migration number for each year when hover mouse over, allowing user to know the data better.

Sketch:



Final Design:



Justification:

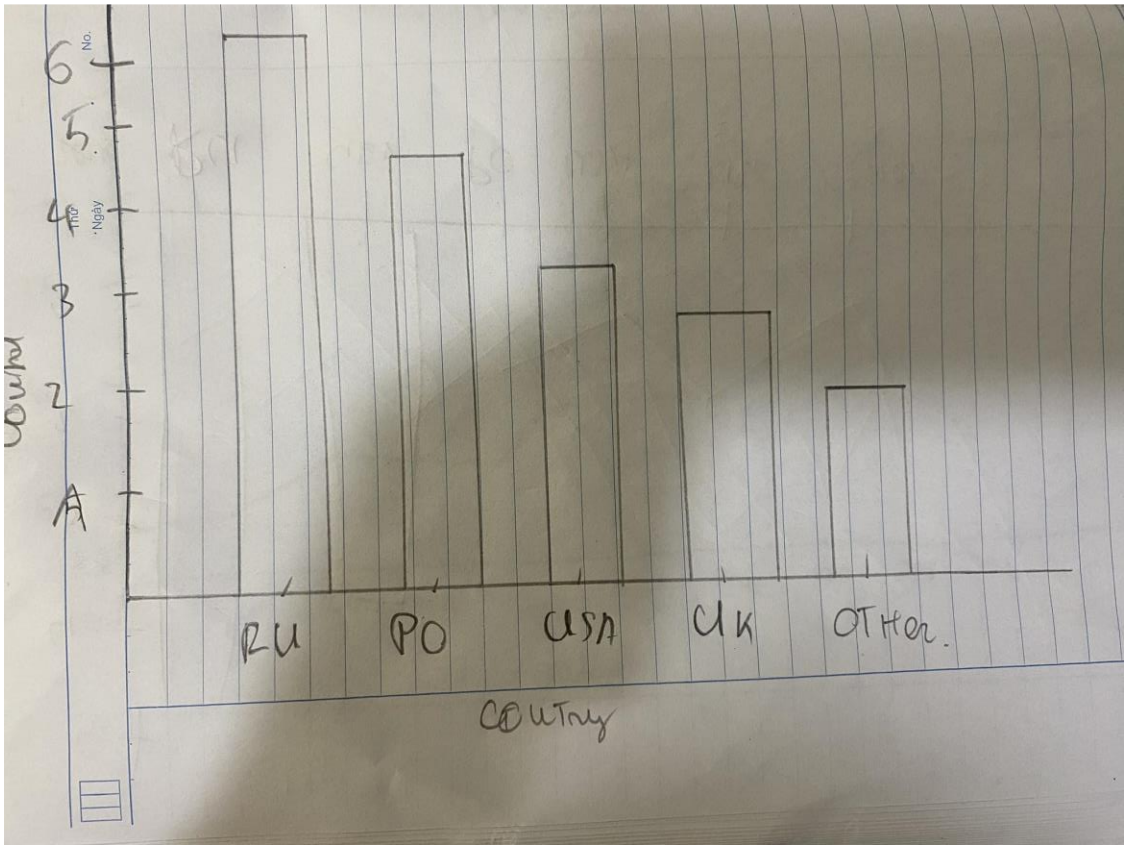
- A line graph is an effective option to identify the changes and trends over time.
- A tooltip is provided to display accurate data.

4.3. Bar chart

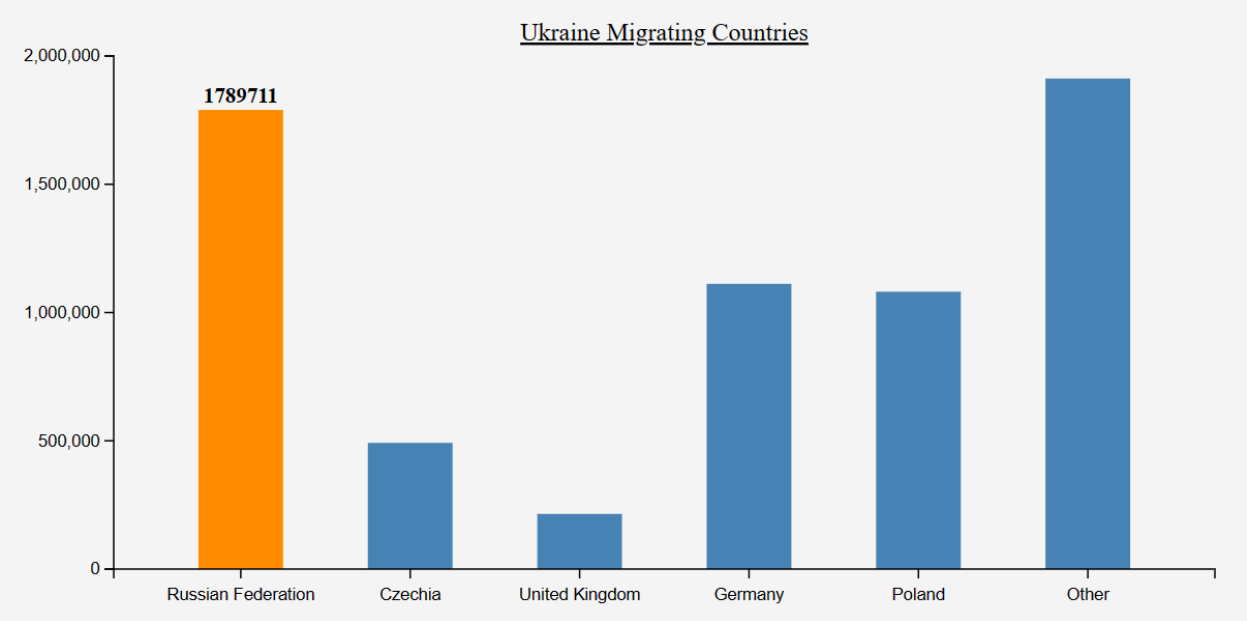
Bar chart will be used to show which country receive the most migrant number from depicted country.

- X axis will show the Asylum destination
- Y axis show counts(people)
- Color gradient for easy comparison.

Sketch:



Final Design:



Justification:

- Bar chart is an effective tool for easy comparison of data between different categories, in this case, asylum destinations.
- Tooltips for showing exact numbers help with accuracy.
- When the user press on a bar, they will be directed to a line chart of the selected country, showing Ukraine refugee flow into that country.

5. Conclusion

In conclusion, this project has been a valuable learning experience. Our team aim as to create a website that hightlighted the trend an pattern of forced migration of Ukraine during the evetful time from 2010 to 2023 by meticulous data analysis and visualization.

Using statistics from trustworthy sources such as the UNHCR and the UN Department of Economic and Social Affairs, we were able to gain significant insights regarding the movement of refugees, asylum seekers, and internally displaced persons (IDPs) across borders. Our visualizations portrayed Ukrainian migrants' journeys, including their destinations, distances traveled, and oscillations in migration numbers over time.

The use of numerous visualizations, such as scatter plots, line charts, and bar graphs, enabled us to present the data in a clear and understandable manner. Each graphic had a specific purpose, allowing people to understand various aspects of the forced migration situation in Ukraine. Our visualizations presented a thorough perspective of the complex environment of human displacement, ranging from identifying popular asylum locations to assessing migratory flow trajectory.

However, the project has not run smoothly. We had hurdles throughout the process, the most notable of which was the link and story behind our project. We originally struggled with selecting how we wanted to control the flow of our website, how to tie each graph together to provide consumers with a pleasant and instructive experience.

Overall, through this project, we have gained valuable experiences of designing data visualization, and how much impact it can have on people. Data preparation is also a important aspect as we learned the better we prepare the data, the easier the coding process.

References

https://www.unhcr.org/external/component/header. (n.d.). UNHCR. Retrieved March 31, 2024, from https://www.unhcr.org/refugee-statistics/download/?url=bf8OWP

Forced displacement flow dataset. (n.d.). UNHCR Refugee Statistics. https://www.unhcr.org/refugee-statistics/insights/explainers/forcibly-displaced-flow-data.html