

UN MigrantStockTotal_2015_ Visualisation_WriteUP

This document is a United Nations publication on Trends in International Migrant Stock from 1990-2015 with 6 tables. Table1 represents the International Migrant Stock at mid-year by sex and by major area, region, country or area from 1990 to 2015. Table2 shows that Total Population at mid-year by sex and by major area, region, country or area from 1990 to 2015. Table 3 represents International migrant stock as a percentage of the total population between 1990 and 2015. Tables 4 shows Female migrants stock as a percentage of the international migrant stock by major area, region, country or area from 1990 to 2015. Table5 represents Annual rate of change of the migrant stock by sex and by major area, region, country or area among these years. The last table represents that Estimated refugee stock at mid-year by major area, region, country or area between 1990 and 2015. In the ANNEX screen you will find the names of all countries and their codes, as well as whether they are developed regions or not.

As we have learned some Tufts's visualization principles, one of this is clear, detailed and thorough labeling should be used to defeat graphical distortion and ambiguity. Write out explanations of the data on the graph itself. Label important events in the data. In this project, the main study is the relationship between immigrant population, male and female ratio, over country regions during 1990-2015. Because of the need to visually analyze each variable in the table, labeling is very important. It will also be shown in the code in the jupyter notebook.

TABLE1

For the first table, I want to get the conclusion about comparison of male and female migration during this period and between the various regions. The first step is cut out the unnecessary parts of the table, and setting the first column as headers. Then, dropping the Notes column and type of data(a) column, which is 'a' that we used in the following parts of code. Setting a new index on the cleaned data frame, creating a copied dataset of the cleaned dataset, labeling each column, and replacing the null value

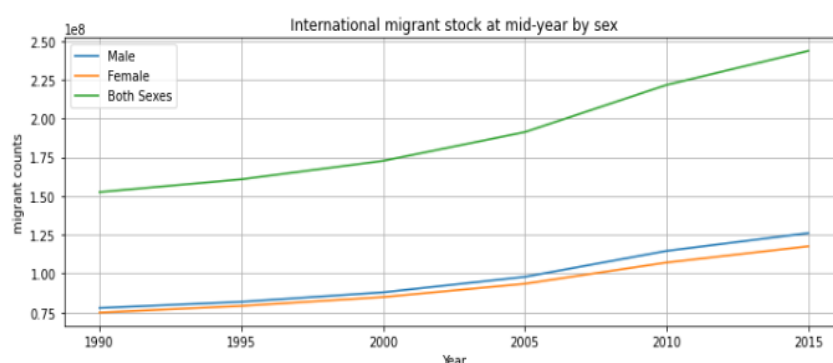
with destination. These steps will also be used in the analysis of the following tables.

	Destination	Both sexes	Both sexes	Both sexes	Both sexes	Both sexes	Both sexes	Male	Male	Male
0	Destination	1990	1995	2000	2005	2010.0	2015.0	1990	1995	2000
1	WORLD	152563212	160801752	172703309	191269100	221714243.0	243700236.0	77747510	81737477	87884839
2	Developed regions	82378628	92306854	103375363	117181109	132560325.0	140481955.0	40263397	45092799	50536796
3	Developing regions	70184584	68494898	69327946	74087991	89153918.0	103218281.0	37484113	36644678	37348043
4	Least developed countries	11075966	11711703	10077824	9809634	10018128.0	11951316.0	5843107	6142712	5361902
...
261	Samoa	3357	4694	5998	5746	5122.0	4929.0	1771	2451	3101
262	Tokelau	270	266	262	258	429.0	487.0	150	147	144
263	Tonga	2911	3274	3684	4301	5022.0	5731.0	1488	1718	1981
264	Tuvalu	318	263	217	183	154.0	141.0	180	148	121
265	Wallis and Futuna	1402	1680	2015	2365	2776.0	2849.0	726	859	1018

	Male	Male	Male	Male	Male	Female	Female	Female	Female	Female
	1995	2000	2005	2010.0	2015.0	1990	1995	2000	2005	2010.0
81737477	87884839	97866674	114613714.0	126115435.0	74815702	79064275	84818470	93402426	107100529.0	117584801.0
45092799	50536796	57217777	64081077.0	67618619.0	42115231	47214055	52838567	59963332	68479248.0	72863336.0
36644678	37348043	40648897	50532637.0	58496816.0	32700471	31850220	31979903	33439094	38621281.0	44721465.0
6142712	5361902	5383009	5462714.0	6463217.0	5236216	5573685	4721920	4432371	4560536.0	5493028.0
...
2451	3101	2940	2594.0	2469.0	1586	2243	2897	2806	2528.0	2460.0
147	144	133	206.0	233.0	120	119	118	125	223.0	254.0
1718	1981	2328	2727.0	3127.0	1423	1556	1703	1973	2295.0	2604.0
148	121	101	85.0	78.0	138	115	96	82	69.0	63.0
859	1018	1194	1401.0	1438.0	676	821	997	1171	1375.0	1411.0

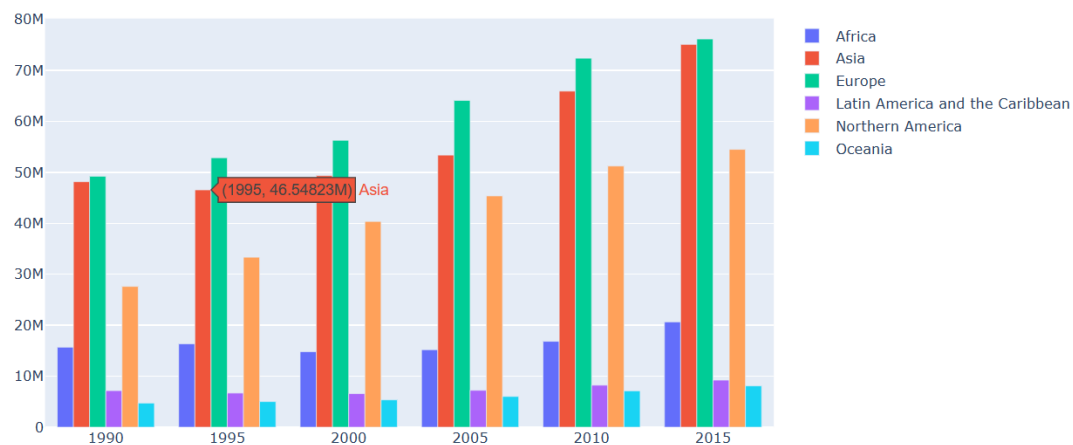
(This table shows the number of people in these years according to destination as well as male, female and both together as a variable of gender)

In the later visualizations, the data are also tabulated using such tables, some of which are differentiated by a duration of five years each, 1990, 1995, 2000, 2005, 2010 and 2015.



This linear graph shows the International migrant stock at mid-year by gender. What is clear is that the number of male and female migrants has increased at a similar rate over the years, with an overall trend of increasing and steady growth.

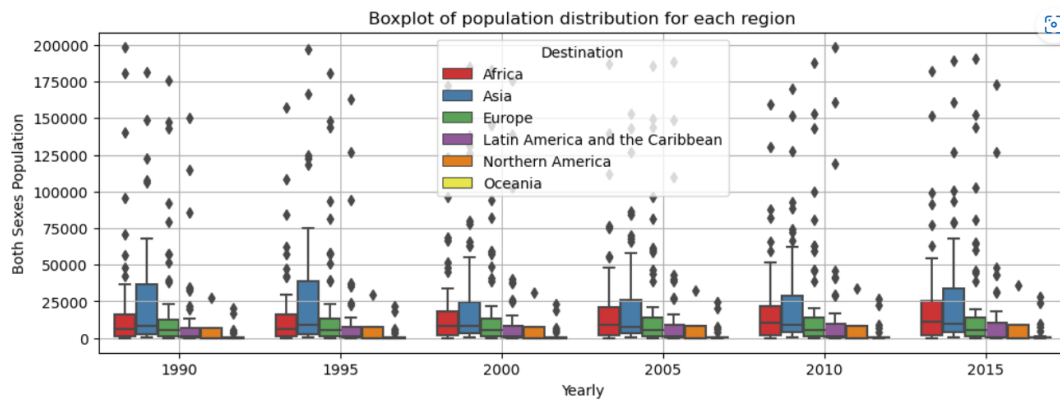
From this line chart, we can clearly see what the X-axis and Y-axis represent and what the chart indicates, as well as the variables represented by the individual color lines. This also conforms to the Tufts's visualization principles.



From this boxplot, it can be seen that Oceania has the lowest number of international migrants, the blue part of the graph, which has not changed much in the intervening years. This is followed by Latin America, which has also not changed much as well. These two areas have remained almost constant and small in number compared to other areas. Both Europe and North America have seen regular and steady upward growth. The stock of migrants in Asia has been increasing at an increasing rate in each duration, except for the period 1995, when it was declining.

TABLE2

Before making the graphs, I will repeat some of the same steps as in TABLE1, making a table of the male, female and total population of the world, the developing regions, the developed regions and the least developed countries. For TABLE2 I am going to make a box plot of the population distribution for each region, extracting the data for each state for analysis.



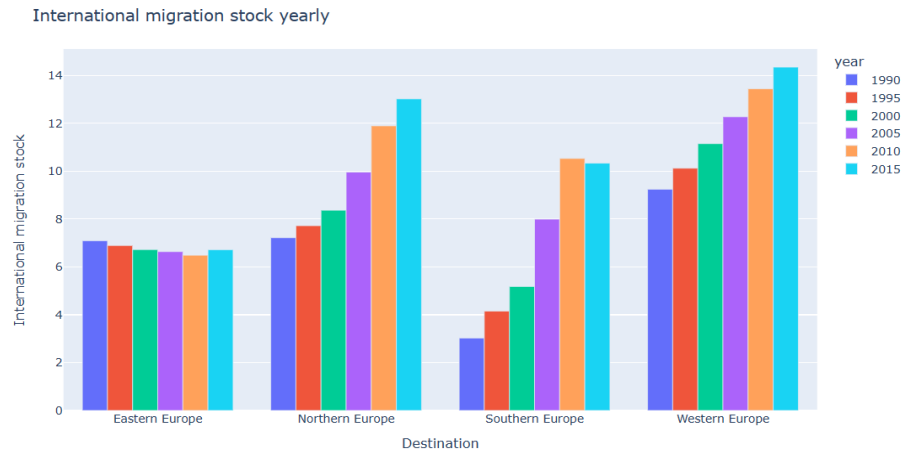
It is easily to find that Asia ranked #1 among these years. In 1995 Asia had the highest population, but in comparison with the previous table, Asia had the lowest stock of immigrants in that year.

TABLE3

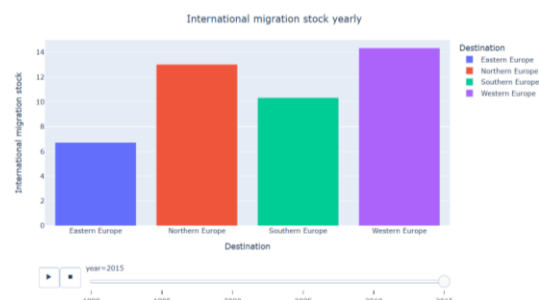
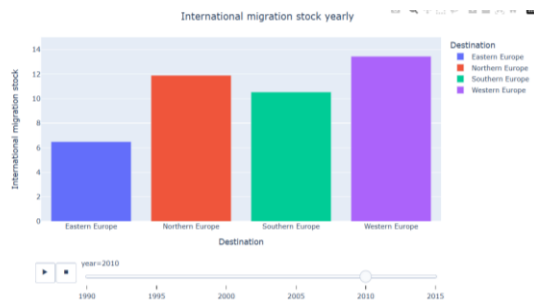
In the above analysis, we know that Europe has the largest stock of international migrants, and here I explore the stock of international migrants in Europe like a percentage of the total population. In the table, Europe is divided by region into Eastern Europe, Northern Europe, Southern Europe, and Western Europe. To get a better read on the numbers, I first extracted the data into the following form.

	1990	1995	2000	2005	2010	2015
Destination						
Eastern Europe	7.094647	6.894779	6.724573	6.638181	6.489308	6.719523
Northern Europe	7.220322	7.723637	8.369902	9.963591	11.896200	13.024358
Southern Europe	3.026736	4.152493	5.1821	7.996996	10.533461	10.336661
Western Europe	9.246246	10.125542	11.151449	12.274909	13.445415	14.351878

From this table, we can easily find the percentage of the stock of international migrants in each area in Europe of the total population.



To better show the results, I have made the data into a bar chart and a dynamic chart, so that you can pull the scrolls directly in jupyter to visualise the change in the percentage of international migrants in each region of Europe.



Using the above bar chart and the dynamic distribution, it is easy to see that all three regions, with the exception of Eastern Europe, are on an upward trend during this period. However, the eastern region has not changed much and has changed erratically over this 25-year period.

	Destination	1990	1995	2000	2005	2010	2015
0	Eastern Europe	6.942883	6.828278	6.747861	6.67634	6.520844	6.749442
1	Northern Europe	7.093783	7.510984	8.065235	9.693087	11.649945	12.655967
2	Southern Europe	2.940433	4.059356	5.095241	7.956131	10.456902	9.928114
3	Western Europe	9.949418	10.576792	11.39549	12.292275	13.158167	14.115656

【International migrant stock as a percentage of the total population (male)】

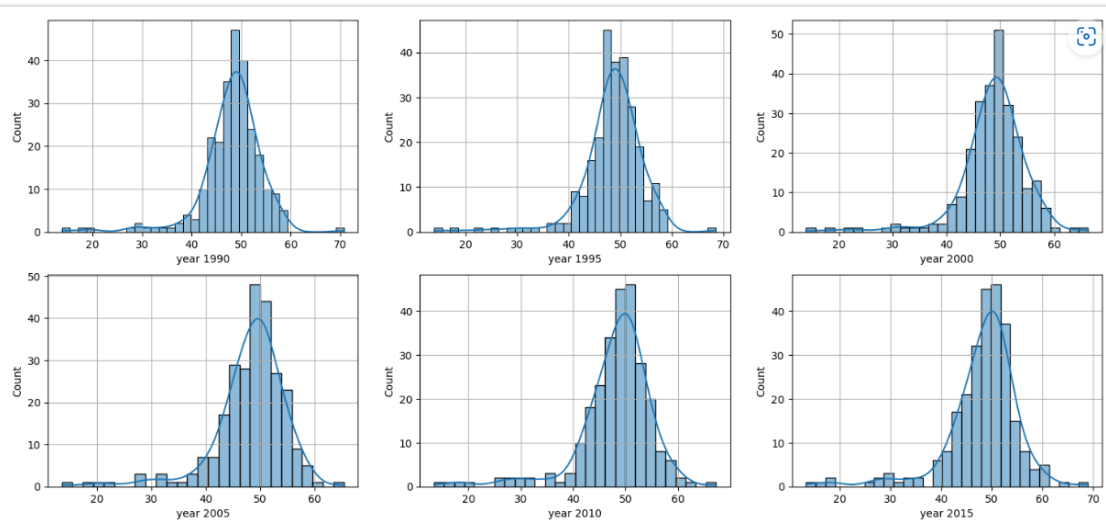
	Destination	1990	1995	2000	2005	2010	2015
0	Eastern Europe	7.23137	6.954738	6.703653	6.604148	6.461315	6.692922
1	Northern Europe	7.340208	7.925139	8.659656	10.222758	12.133701	13.382147
2	Southern Europe	3.109389	4.241346	5.26507	8.036188	10.606817	10.725404
3	Western Europe	8.579175	9.694106	10.917612	12.258264	13.720862	14.579278

【International migrant stock as a percentage of the total population (female)】

Looking at the data graphs as well as the line graphs, we see that in both sexes, except in the eastern part of Europe, there has been an increase in most years. Therefore, the percentage of immigrants in the total population in the European region is not related to gender.

TABLE4

Tables 4 shows Female migrants stock as a percentage of the international migrant stock during these years, which can display a large amount of data and the frequency of the data values.



It can be seen that in each year the percentage of female migrants in the total number of international migrants is around 50%, with no significant differences and all of them largely conforming to a normal distribution. Although there are some outliers present, the impact is not significant in general.

TABLE5

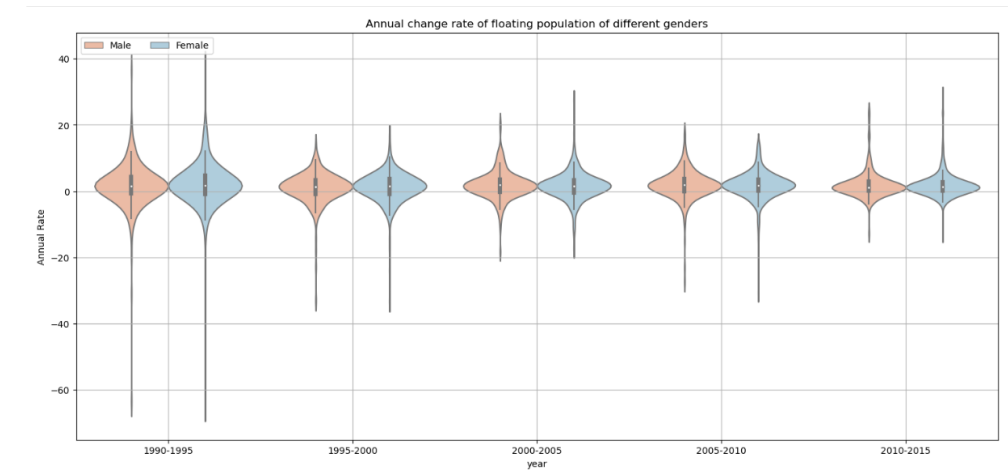
Table5 shows the annual rate of change of the migrant stock, so I plan to explore on differences in the annual rate of change of the migrant population in terms of gender.

Destination	Both sexes	Both sexes	Both sexes	Both sexes	Both sexes	Male	Male	Male	Male	Male	Female	Female	Female	Female
NaN	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	1990-1995	1995-2000	2000-2005	2005-2010
WORLD	1.051865	1.428058	2.042124	2.95416	1.890991	1.000922	1.450294	2.151575	3.159228	1.912603	1.104667	1.405044	1.92808	2.737012
Developed regions	2.275847	2.264965	2.50708	2.466343	1.160824	2.265595	2.279583	2.483259	2.265689	1.074685	2.285643	2.250995	2.529838	2.655595
Developing regions	-0.487389	0.241777	1.328107	3.702217	2.929634	-0.45298	0.380246	1.693824	4.352954	2.927058	-0.526904	0.081268	0.89236	2.881555
Least developed countries	1.118175	-3.001139	-0.539636	0.419137	3.526927	1.000073	-2.718952	0.078575	0.293964	3.363629	1.249146	-3.316818	-1.265617	0.57011

This table represents the annual rate of all gender, male and female in world, developed regions, developing regions, and least developed countries.

	Destination	Male	Male	Male	Male	Male	Female	Female	Female	Female	Female
0	NaN	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015	1990-1995	1995-2000	2000-2005	2005-2010	2010-2015
1	WORLD	1.000922	1.450294	2.151575	3.159228	1.912603	1.104667	1.405044	1.92808	2.737012	1.867837
2	Developed regions	2.265595	2.279583	2.483259	2.265689	1.074685	2.285643	2.250995	2.529838	2.655595	1.241097
3	Developing regions	-0.45298	0.380246	1.693824	4.352954	2.927058	-0.526904	0.081268	0.89236	2.881555	2.933003
4	Least developed countries	1.000073	-2.718952	0.078575	0.293964	3.363629	1.249146	-3.316818	-1.265617	0.57011	3.72079
...
261	Samoa	6.499035	4.704571	-1.066301	-2.50417	-0.987758	6.931983	5.117229	-0.638315	-2.08863	-0.545343
262	Tokelau	-0.404054	-0.412386	-1.589283	8.750541	2.463246	-0.167365	-0.168777	1.152582	11.577161	2.60325
263	Tonga	2.874558	2.848819	3.228155	3.163851	2.737439	1.787022	1.80546	2.943277	3.023552	2.526318
264	Tuvalu	-3.914892	-4.028435	-3.613401	-3.449385	-1.718849	-3.646431	-3.611679	-3.152579	-3.452255	-1.819436
265	Wallis and Futuna Islands	3.364378	3.396526	3.189382	3.197545	0.52134	3.886601	3.884553	3.217252	3.211913	0.516899

And then, I made a table of the annual rate of change by region and by gender, but I found that it was not easy to find a pattern. We don't quite find too regular a marked change from these two tables above. So I decided to draw grouped violin diagrams.



From this graph, there is no significant differences can be seen between female and male in annual rate of population.

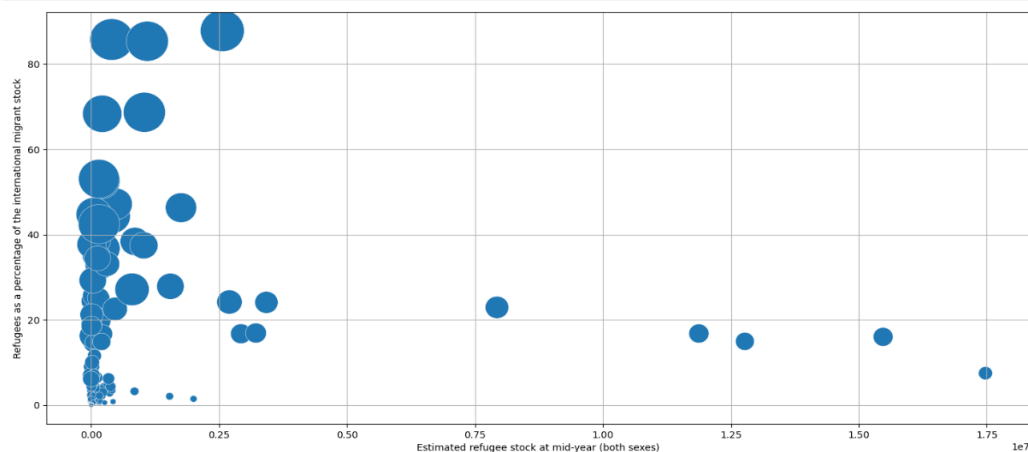
TABLE6

Table 6 represents the estimated refugee stock at mid-year by major area from 1995 to 2010. Since the table contains a lot of "0" information, we need to extract it.

Major area, region, country or area of destination	Estimated refugee stock at mid-year (both sexes)						Refugees as a percentage of the international migrant stock						Annual rate of change of the refugee stock				
	NaN	1990	1995	2000	2005	2010.0	2015.0	1990	1995	2000	2005	2010.000000	2015.000000				
WORLD	18836571	17853840	15827803	13276733	15370755.0	19577474.0	12.346732	11.103013	9.164736	6.941389	6.932687	8.033424	-2.123497	-3.837069	-5.557223	-0.025089	2.947267
Developed regions	2014564	3609670	2997256	2361229	2046917.0	1954224.0	2.445494	3.910511	2.899391	2.015025	1.544140	1.391085	9.388424	-5.983348	-7.277379	-5.323293	-2.087656
Developing regions	16822007	14244170	12830547	10915504	13323838.0	17623250.0	23.968236	20.795958	18.507035	14.733162	14.944759	17.073768	-2.839417	-2.332154	-4.561	0.285195	2.663652
Least developed countries	5048391	5160131	3047488	2363782	1957884.0	3443582.0	45.56588	44.041961	30.221557	24.08243	19.533425	28.801534	-0.680327	-7.531747	-4.541459	-4.187109	7.766031

And then, I gather estimated refugee stock, bubbles for all countries and regions, distribution between the annual rate of change of the refugee stock and Refugees as a percentage of the international migrant stock overall, and average Annual rate of change of the refugee stock by each mid-year, and draw a diagram (below).

	Estimated refugee stock at mid-year (both sexes)	Refugees as a percentage of the international migrant stock	Annual rate of change of the refugee stock
1	17474114.5	7.483056	8.033424
2	2000570.5	1.467613	1.391085
3	15473544.0	16.009263	17.073768
4	2700733.0	24.167480	28.801534
5	12772811.0	14.950419	15.537313



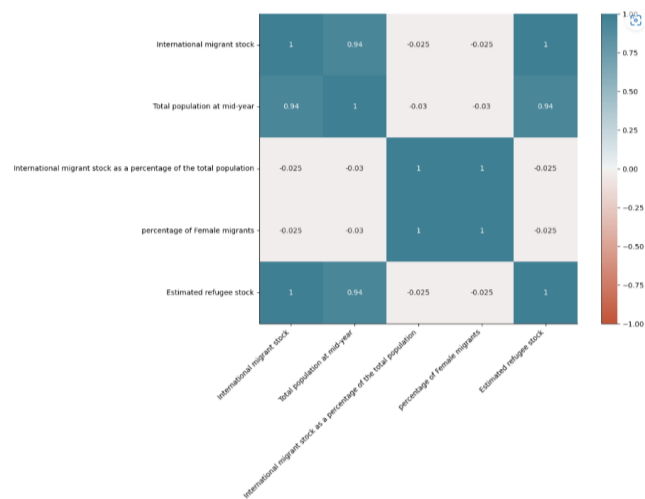
It can be seen that the more Refugees as a percentage of the international migrant stock, the greater Annual rate of change of the refugee stock, and the smaller the annual rate of change of the refugee stock when both Refugees as a percentage of the international migrant stock and Refugees as a percentage of the international migrant stock are in the smaller range.

In order to visualize the last step with the heat map, we extracted some information and made a table.

13	Country or area	Sort order	Major area	Code	Sort order	Region	Code	Sort order	Developed region	Least developed country	Sub-Saharan Africa
14	Afghanistan	99	Asia	935	71	Southern Asia	5501	98	No	Yes	No
15	Albania	154	Europe	908	127	Southern Europe	925	153	Yes	No	No
16	Algeria	40	Africa	903	7	Northern Africa	912	39	No	No	No
17	American Samoa	257	Oceania	909	238	Polynesia	957	256	No	No	No
18	Andorra	155	Europe	908	127	Southern Europe	925	153	Yes	No	No

	International migrant stock	Total population at mid-year	International migrant stock as a percentage of the total population	percentage of Female migrants	Estimated refugee stock
0	2.012500e+03	2.005000e+03	2012.500000	2012.500000	2.012500e+03
1	1.551382e+08	6.532116e+06	3.257677	3.257677	8.737061e+06
2	9.101409e+07	1.210444e+06	10.987094	10.987094	1.000286e+06
3	6.412407e+07	5.321672e+06	1.628865	1.628865	7.736780e+06
4	7.324823e+06	7.607586e+05	1.217487	1.217487	1.350379e+06
...
261	3.350333e+03	1.807914e+02	2.652103	2.652103	0.000000e+00
262	3.053333e+02	1.333400e+00	38.378678	38.378678	0.000000e+00
263	3.584333e+03	1.009524e+02	5.114627	5.114627	0.000000e+00
264	9.833333e+01	9.616600e+00	1.494528	1.494528	0.000000e+00
265	1.875000e+03	1.392040e+01	21.064091	21.064091	0.000000e+00

Combine all the above tabel for heat map visualization and visualize the correlation heat map between each table.



It can be seen that the correlation coefficient between Total population at mid-year and International migrant stock is the highest, reaching 94% correlation, and the correlation coefficient between Estimated refugee stock and International migrant stock is also high reaching 80% correlation.