Canadian Immigration

Analyzing immigration data and global indicators.

Group members

Felipe Scarpelli

Heidi Zarse

Karan Singh

Overview

Why this topic

Not our first idea...

We originally wanted to look at how much the sharing economy has been taking from traditional markets.

We tried to find datasets to compare Uber vs. taxis, or Airbnb vs. hotels.

We learned that commercial datasets are hard to find and ultimately couldn't find what we needed to answer our initial question.

Plan B

Plan B involved looking for government data that was openly available.

We picked immigration as that's something that impacts all Canadians in one way or another, and the fact that we are all Canadian immigrants.

With that in mind, we set out to analyze Canadian immigration in an attempt to identify patterns, similarities and curiosities.

Datasets

Open Government

open.canada.ca

Annual number of immigrant arrivals to Canada by source country from 1980 to 2017.

TheGlobalEconomy

theglobaleconomy.com

National socioeconomic indicators and indexes from 1960 to 2018.

OECD

stats.oecd.org

International migration database with the number of emigrants leaving their countries from 2000 to 2016.

Data prep and cleaning

- In Excel, consolidated datasets to get a single CSV file from each source

  e.g. Open Canada didn't have a single report with immigration data covering from 1980 to 2017, so multiple reports had to be downloaded with different years and combined using VLOOKUP.

- In Excel, compared countries to standardize country names and codes

  e.g. Two datasets used country codes, so we compared them to ensure the codes used were the same. We then added those codes to the third dataset using VLOOKUP.

- In MySQL, removed unnecessary data and consolidated values as needed

  e.g. The emigration dataset entries consisted of origin country, destination country and year. We didn't need the destination country for our study, so we consolidated those to get the total number [SUM(amount)] of emigrations from each country per year.

Challenges

- Country names and country codes were different in different datasets

  e.g. China vs. People's Republic Of China
  e.g. CRC vs. CRI as country codes for Costa Rica

- Countries have been unified or divided

  e.g. East Germany was unified with West Germany
  e.g. USSR, Yugoslavia and Czechoslovakia were divided

- Data was on the Y axis on a couple reports, but on the X axis on another

  e.g. The "Year" field for one specific dataset was on the X axis, with each column being a year. The other two datasets had the year as rows in the Y axis, so we couldn't link them upon importing them into Tableau. Instead, we had to go back to Excel, transpose the data, and bring it back to Tableau.

- Too much data

  e.g. We ended up with 20 socioeconomic indicators for all countries from 1960, Canadian immigration data from 1980, and global emigration data from 2000. This huge amount of data made it challenging to filter out what was unnecessary and focus on key points that really mattered.

Data sample

	A	B	C	D	E	F	G	H	I	J
1	Country of Citizenship	Country Cc	Popular Name	1980	1981	1982	1983	1984	1985	1986
2	Afghanistan	AFG	Afghanistan	16	39	39	47	71	340	496
3	Africa NES		Africa (NES)	0	0	1	0	4	0	0
4	Albania	ALB	Albania	1	0	0	0	0	0	1
5	Algeria	DZA	Algeria	80	67	71	69	63	44	69
6	Andorra	AND	Andorra	0	0	0	0	0	0	2
7	Angola	AGO	Angola	1	3	6	6	4	3	5
8	Antigua and Barbuda	ATG	Antigua and Ba	0	0	0	0	42	52	51
9	Argentina	ARG	Argentina	368	426	626	241	237	196	213
10	Armenia	ARM	Armenia	0	0	0	0	0	0	0
11	Asia NES		Asia (NES)	0	0	2	0	2	0	0
12	Australia	AUS	Australia	702	639	484	317	317	319	356
13	Austria	AUT	Austria	234	238	201	117	127	165	196
14	Azerbaijan	AZE	Azerbaijan	0	0	0	0	0	0	0
15	Bahama Islands, The	BHS	Bahamas	26	23	38	12	21	28	23
16	Bahrain	BHR	Bahrain	0	2	1	1	1	3	0
17	Bangladesh	BGD	Bangladesh	83	84	86	81	98	92	486
18	Barbados	BRB	Barbados	372	376	299	244	265	285	251
19	Belarus	BLR	Belarus	0	0	0	0	0	0	0
20	Belgium	BEL	Belgium	511	540	519	297	183	181	197
21	Belize	BLZ	Belize	16	27	13	21	37	26	25
22	Benin, Republic of	BEN	Benin	2	5	4	3	4	3	6
23	Bhutan	BTN	Bhutan	0	0	0	0	1	0	0
24	Bolivia	BOL	Bolivia	44	52	42	49	38	44	79
25	Bosnia-Herzegovina	BIH	Bosnia and Her	0	0	0	0	0	0	0
26	Botswana, Republic of	BWA	Botswana	10	1	3	3	7	4	2
27	Brazil	BRA	Brazil	211	220	192	139	145	130	205
28	Brunei	BRN	Brunei	79	6	8	2	2	4	12

	A	B	C	D
1	country_code	year	value	
2	AFG	2000	2744	
3	AFG	2001	4112	
4	AFG	2002	2726	
5	AFG	2003	2228	
6	AFG	2004	2352	
7	AFG	2005	2175	
8	AFG	2006	1994	
9	AFG	2007	1580	
10	AFG	2008	2173	
11	AFG	2009	2525	
12	AFG	2010	2609	
13	AFG	2011	2545	
14	AFG	2012	4152	
15	AFG	2013	4096	
16	AFG	2014	4031	
17	AFG	2015	7643	
18	AFG	2016	23411	
19	AGO	2000	593	
20	AGO	2001	485	
21	AGO	2002	578	
22	AGO	2003	651	
23	AGO	2004	627	
24	AGO	2005	666	
25	AGO	2006	638	
26	AGO	2007	733	
27	AGO	2008	832	
28	AGO	2009	852	
29	AGO	2010	757	
30	AGO	2011	722	
31	AGO	2012	777	
32	AGO	2013	972	
33	AGO	2014	860	

Tableau

Analysis

But First

A cool visualization depicting immigration flows to Canada


Immigration to Canada
(1980 - 2017)
Inflows by Country and Year
1980
Canada from:
1 United Kingdom 22,045
2 United States 9378
3 India 8880
4 Philippines 6051
5 China 5123
6 Portugal 4473
7 Jamaica 3198
8 Guyana 2334
9 Netherlands 1889
10 Italy 1820

Analysis #1

Immigrant arrivals by country of origin, 1980-2017

Who is coming to Canada?

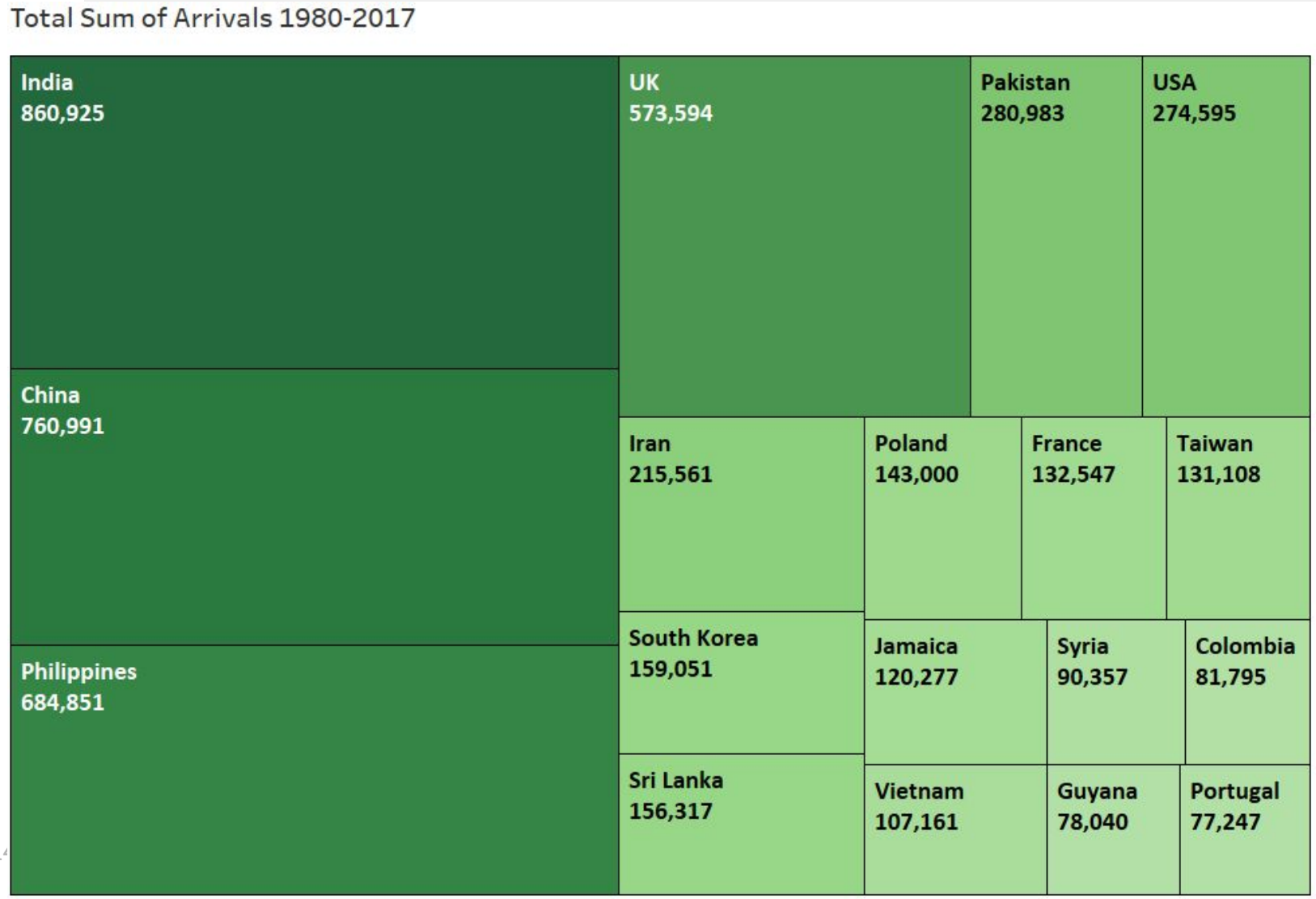
The next few maps showcase the source country of immigrants arriving in Canada, broken down by decade.

The darker shades of green indicate more immigrants coming from those countries.

Canada

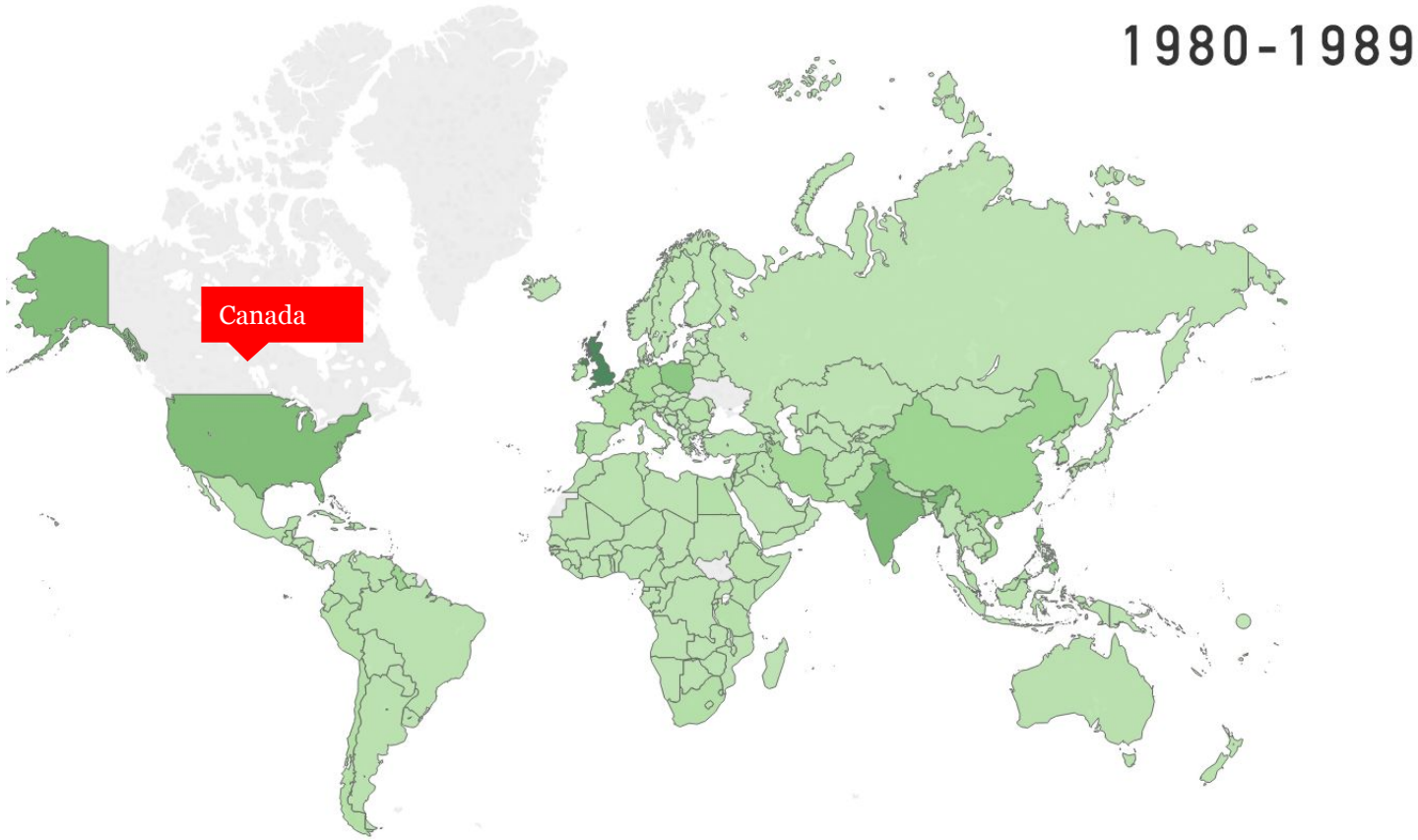
Analysis #1

Immigrant arrivals by country of origin, 1980-2017



Analysis #1

Immigrant arrivals by country of origin, 1980-2017

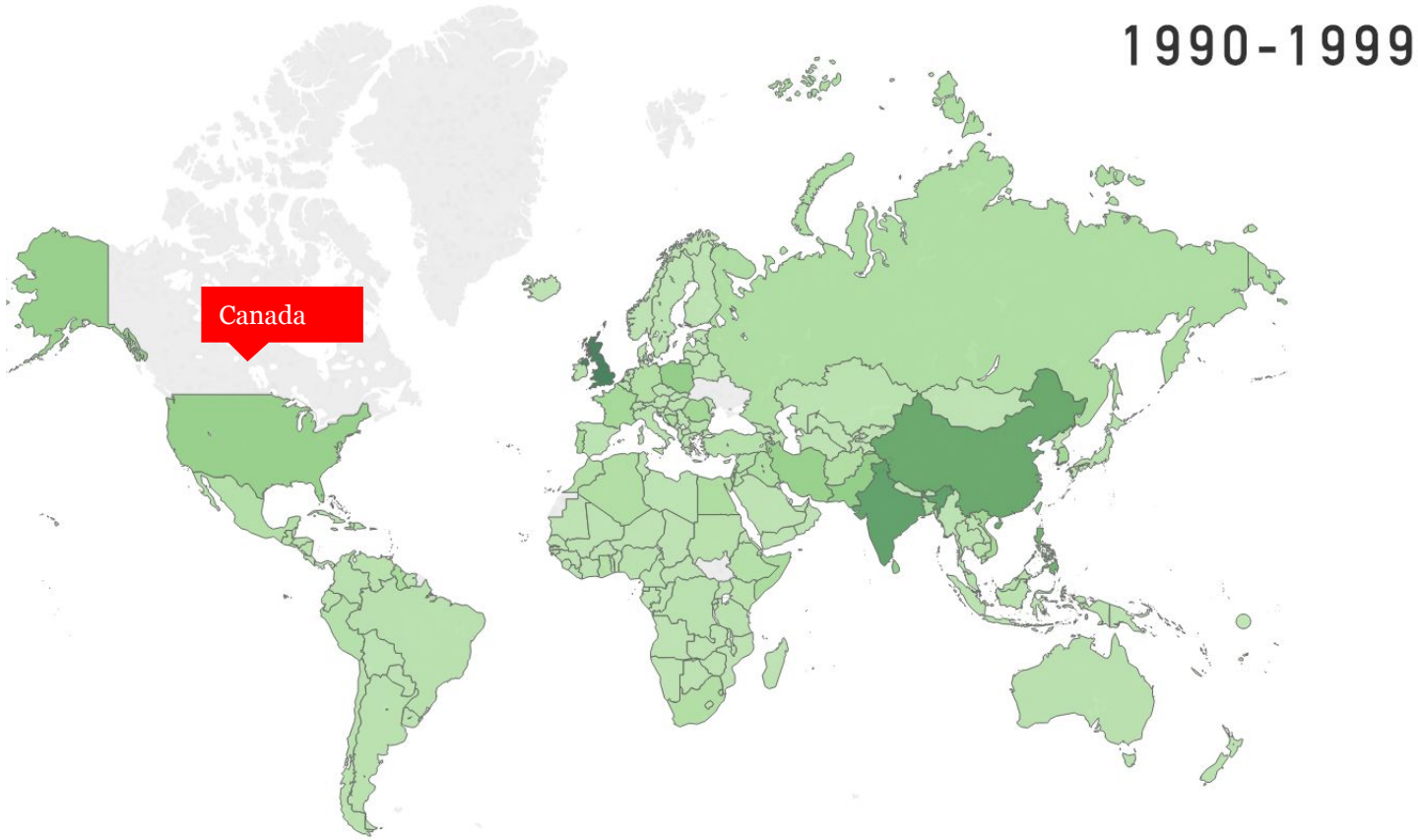


Top 10 Countries

United Kingdom	India	USA	Philippines	Poland	Portugal	Jamaica	Guyana	China	Vietnam
179,171	82,156	76,824	60,764	57,602	40,432	34,328	32,096	32,003	30,639

Analysis #1

Immigrant arrivals by country of origin, 1980-2017

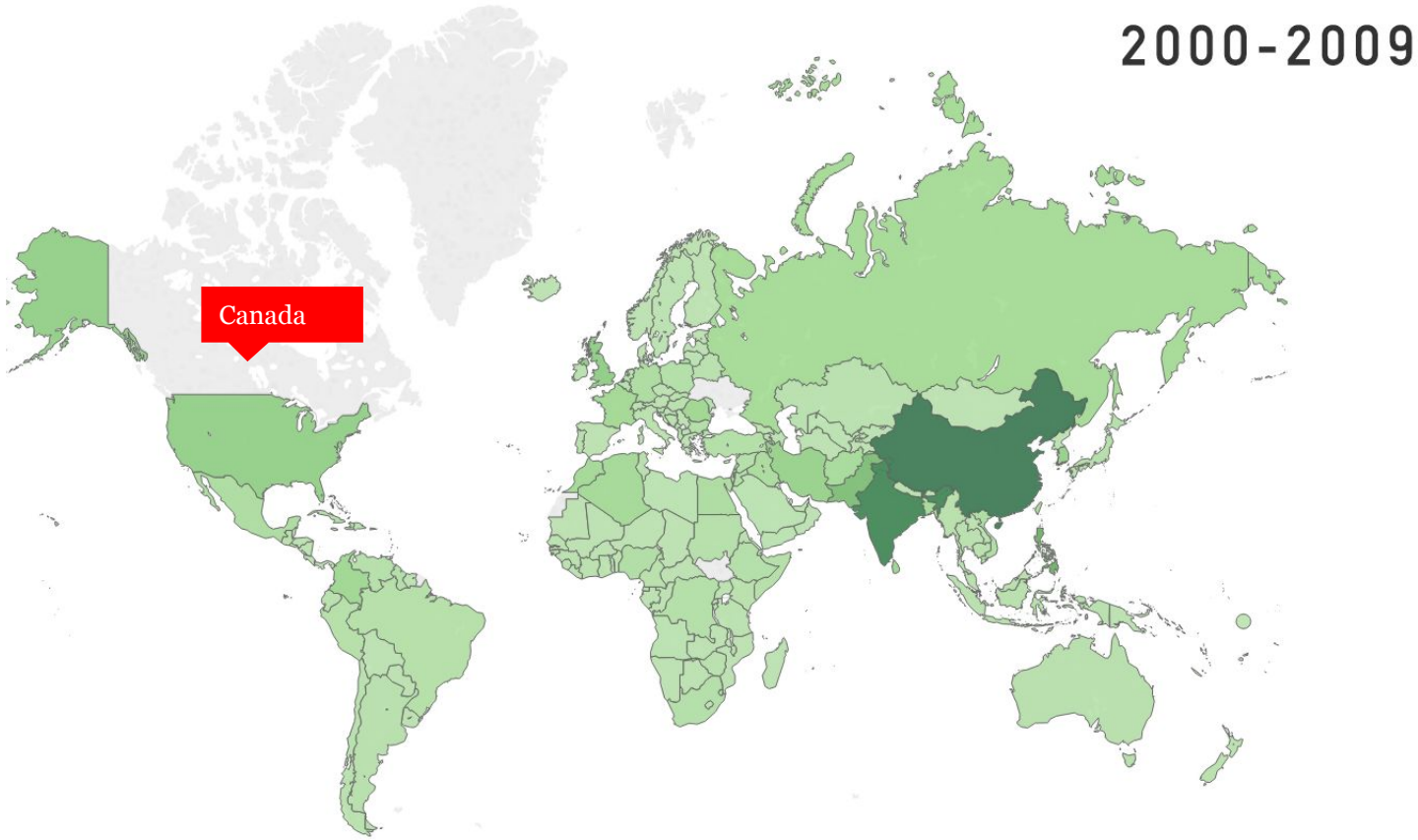


Top 10 Countries

United Kingdom	India	China	Philippines	Taiwan	Sri Lanka	Pakistan	Poland	USA	Iran
261,954	180,393	161,528	138,482	81,032	70,421	65,302	64,864	56,915	54,871

Analysis #1

Immigrant arrivals by country of origin, 1980-2017

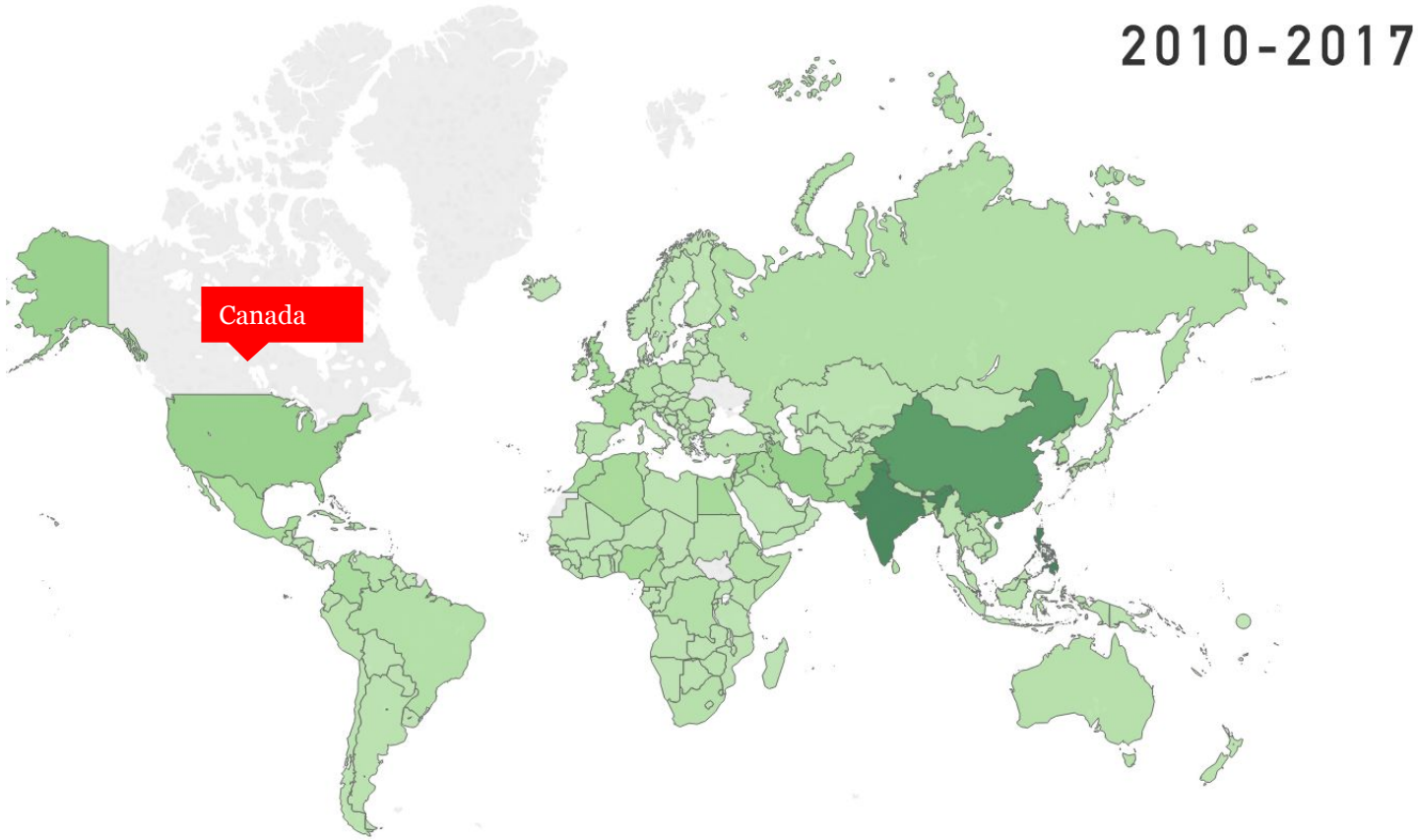


Top 10 Countries

China	India	Philippines	Pakistan	United Kingdom	USA	South Korea	Iran	Sri Lanka	Colombia
340,161	303,359	172,857	127,559	83,323	75,163	68,121	65,781	49,544	45,813

Analysis #1

Immigrant arrivals by country of origin, 1980-2017



Top 10 Countries

Philippines	India	China	Pakistan	Iran	USA	Syria	United Kingdom	France	South Korea
312,748	295,017	227,299	77,531	73,432	65,693	62,578	49,146	44,093	36,482

Analysis #1

Highlights/Conclusion

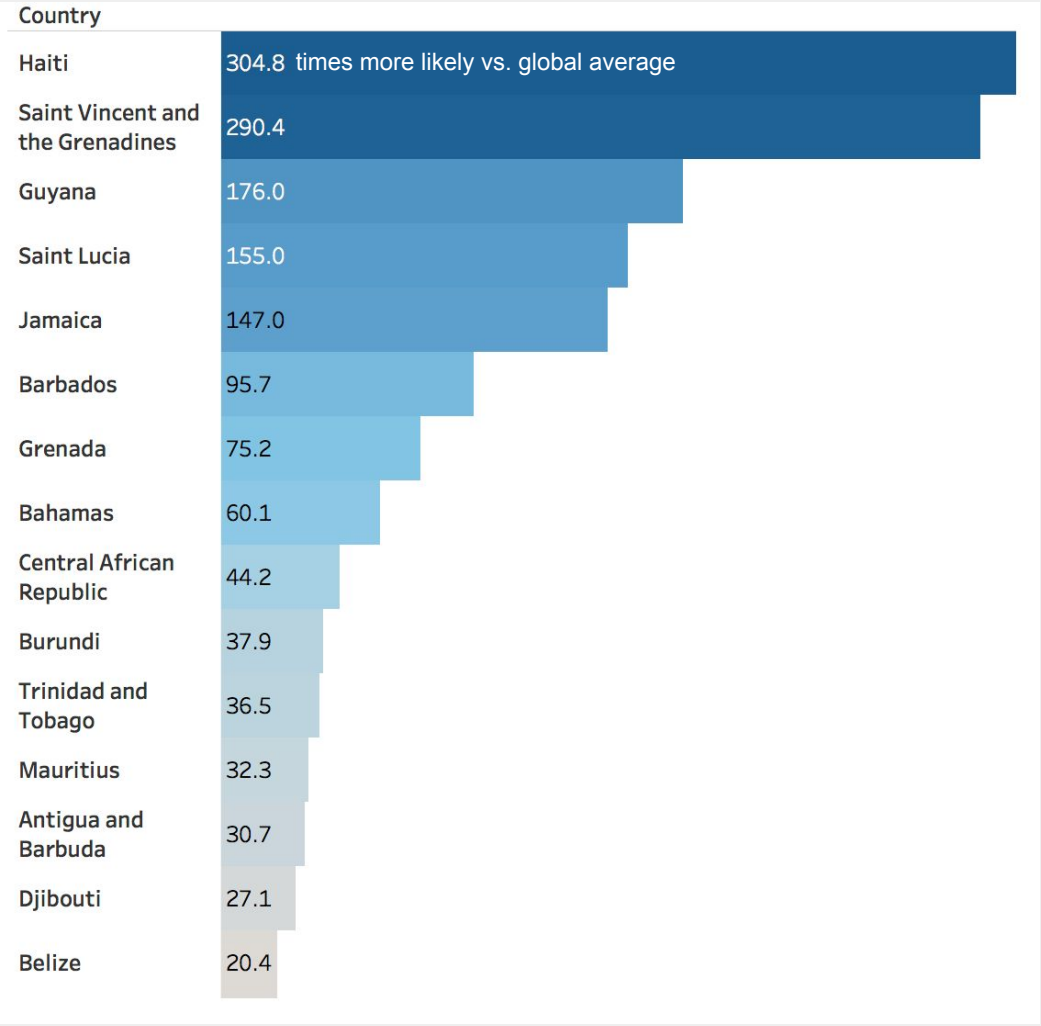
• UK was the #1 country of origin in the 80s and 90s, dropping back in the 2000s and 2010s. Still, it's always been in the top 10.

• The number of immigrants from the US is surprisingly consistent during the four decades, with a minimum of 56,915 in the 90s and 76,824 in the 80s. It's been in the top 10 during the four decades.

• Clear increase in immigration from India, China and the Philippines over the past 4 decades.

Analysis #2

Canadian arrivals vs. global emigration (deviation) 2016, top 15 countries



How likely is someone to choose Canada?

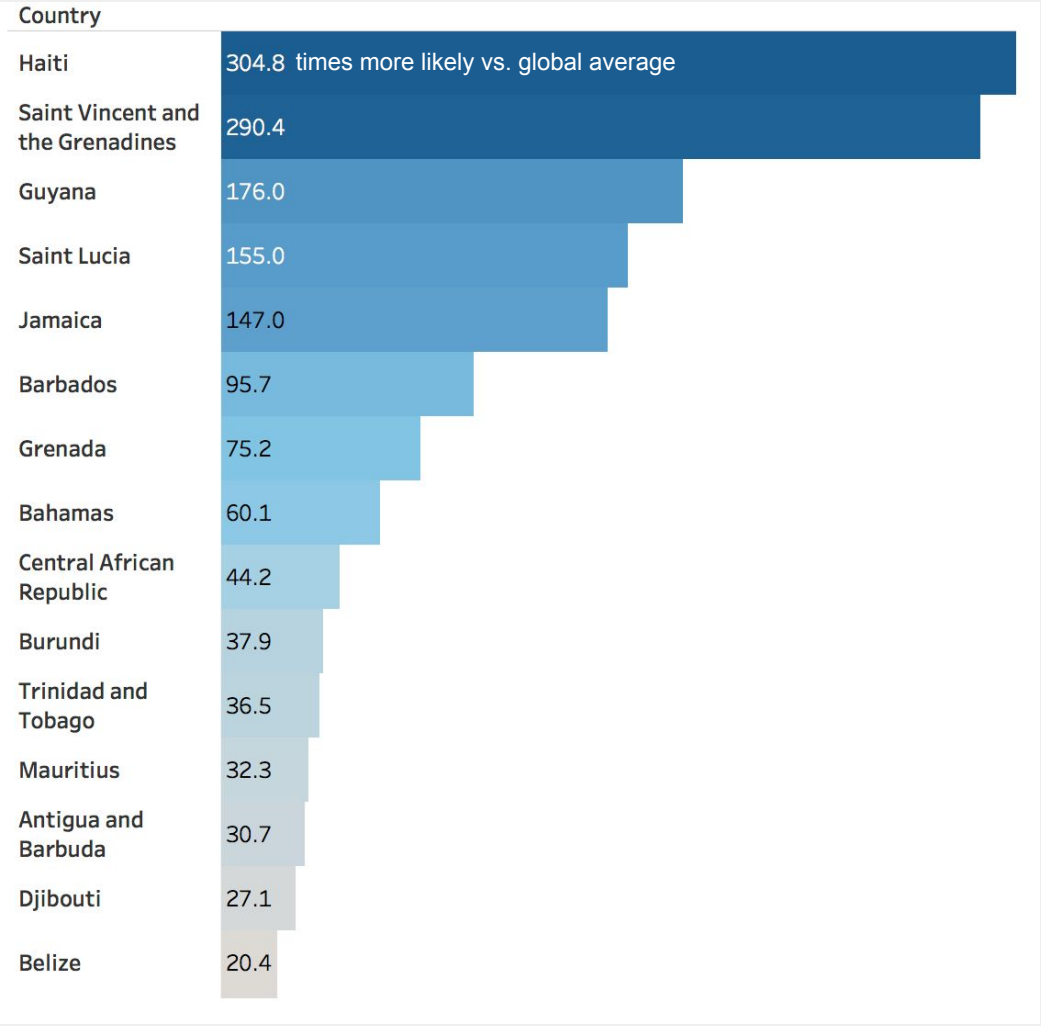
This chart compares the percentage breakdown of arrivals to Canada with the percentage breakdown of all emigrants leaving their countries in 2016.

The higher the value, the more likely it will be that a person leaving that country will choose Canada as their destination.

This means someone from Haiti is 304.8 times more likely to choose Canada as destination versus the global average.

Analysis #2

Canadian arrivals vs. global emigration (deviation) 2016, top 15 countries

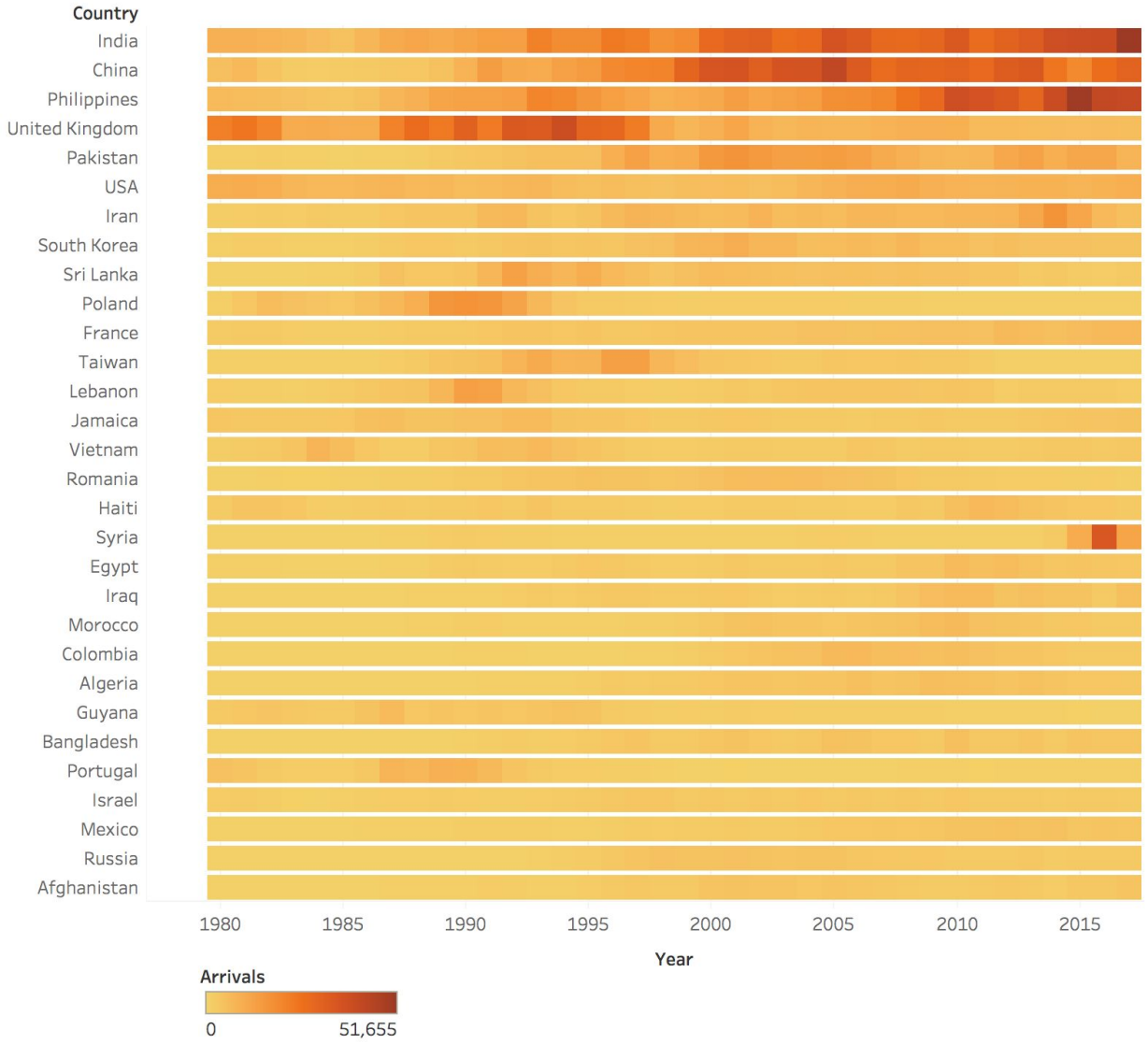


Highlights / Conclusion

Even though Canada doesn't receive a large influx of Caribbean immigrants compared to immigrants from Europe and Asia immigrants, a person leaving a country in the Caribbean is much more likely to choose Canada as their destination compared to the global average.

Analysis #3

Canadian arrivals by origin country by year 1980-2017, top 30 countries



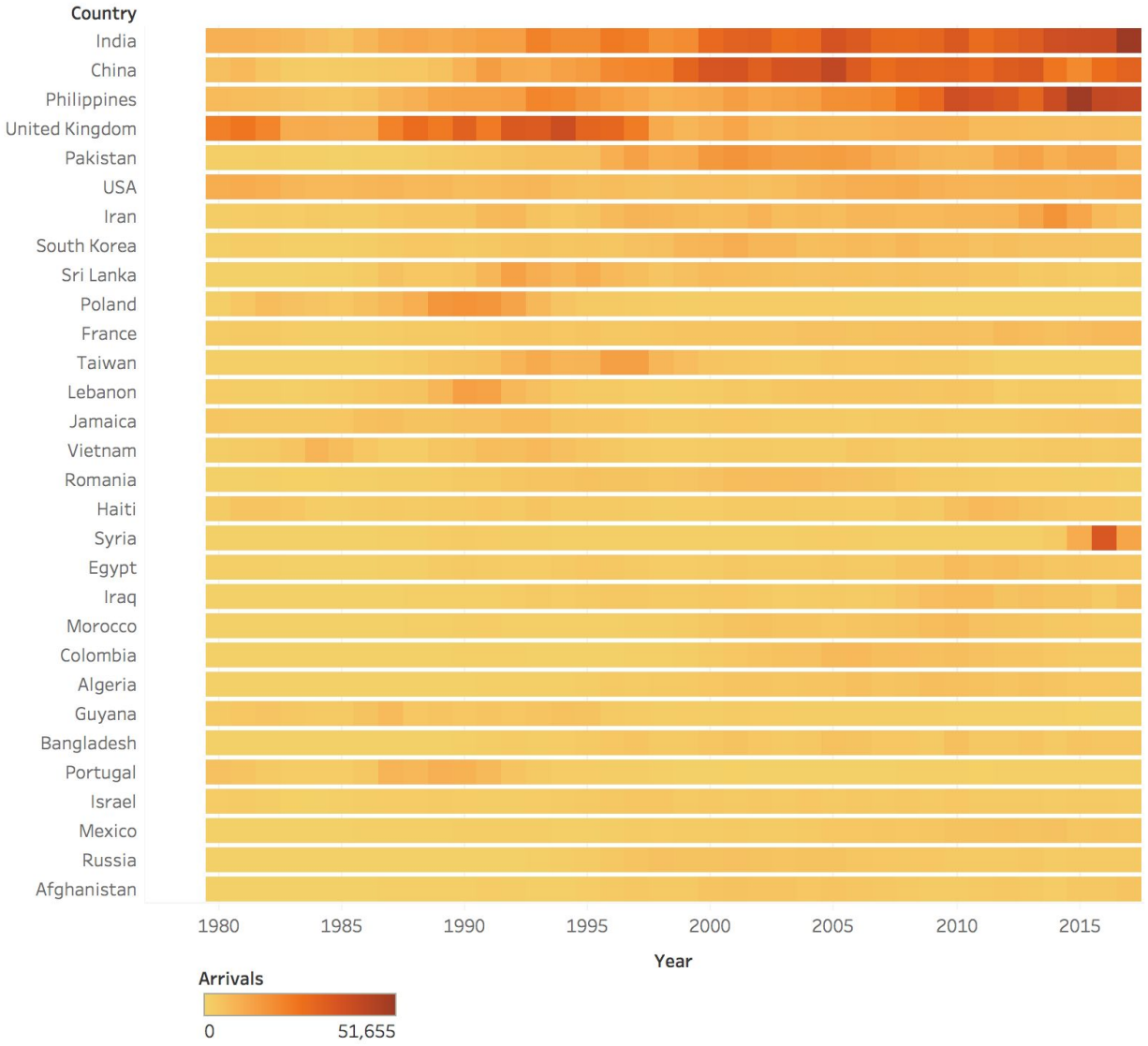
How much has immigration to Canada varied on an yearly basis?

This is a heatmap of the influx of immigrants per country since 1980, broken down by year.

The darker the shade of red, the higher the number of immigrants from that country that year.

Analysis #3

Canadian arrivals by origin country by year 1980-2017, top 30 countries

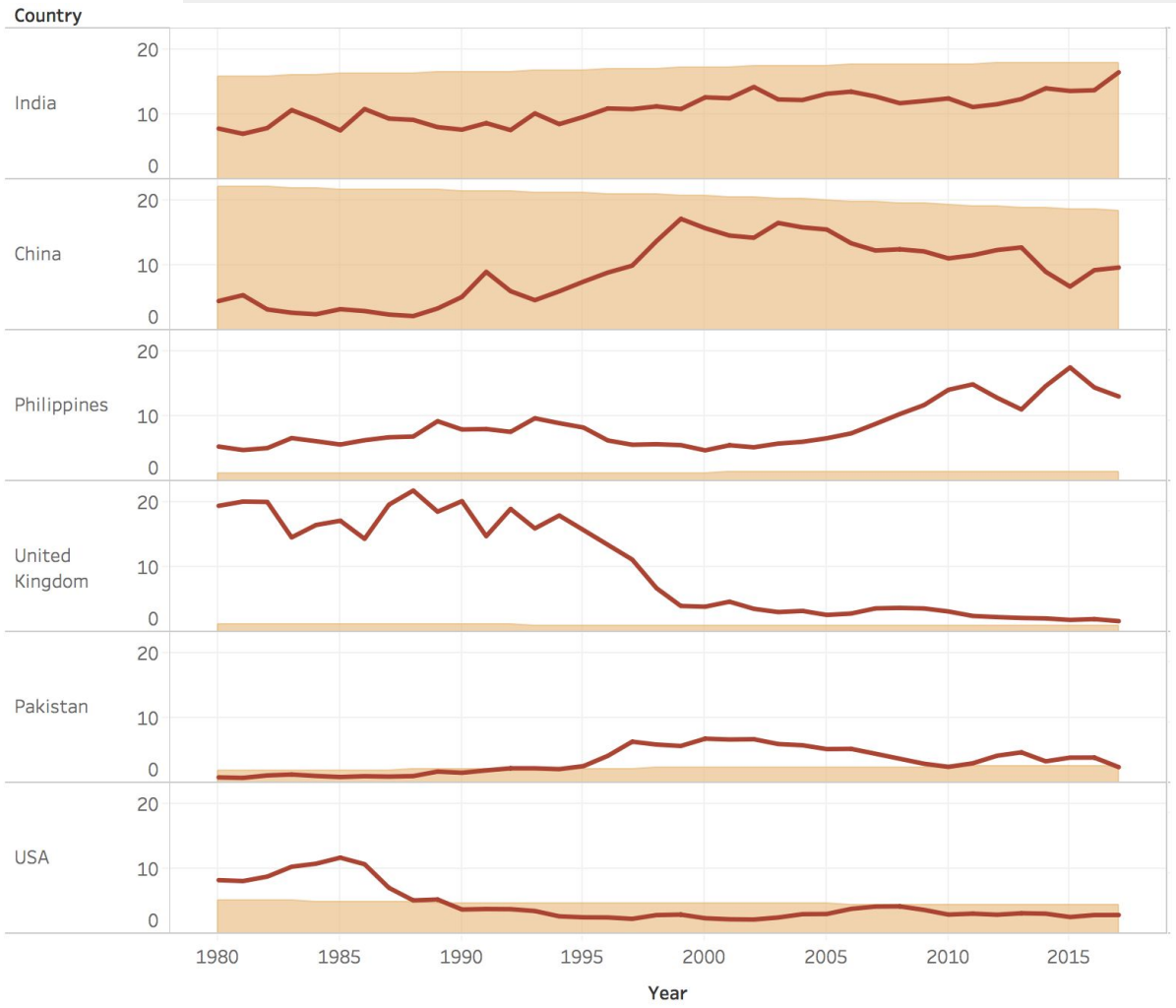


Highlights / Conclusion

- Visible increase in immigration from India, China and the Philippines over the years.
- Visible decrease in immigration from the UK over the years.
- Subtle immigration spikes from Poland, Portugal, Lebanon, Taiwan, Sri Lanka and Iran that lasted a few years.
- Strong immigration spike from Syria from 2015 to 2017.
- Relatively stable influx from Afghanistan, Russia, Mexico, Israel over the last 4 decades.

Analysis #4

Canadian immigration vs. global population 1980-2017, top 6 countries

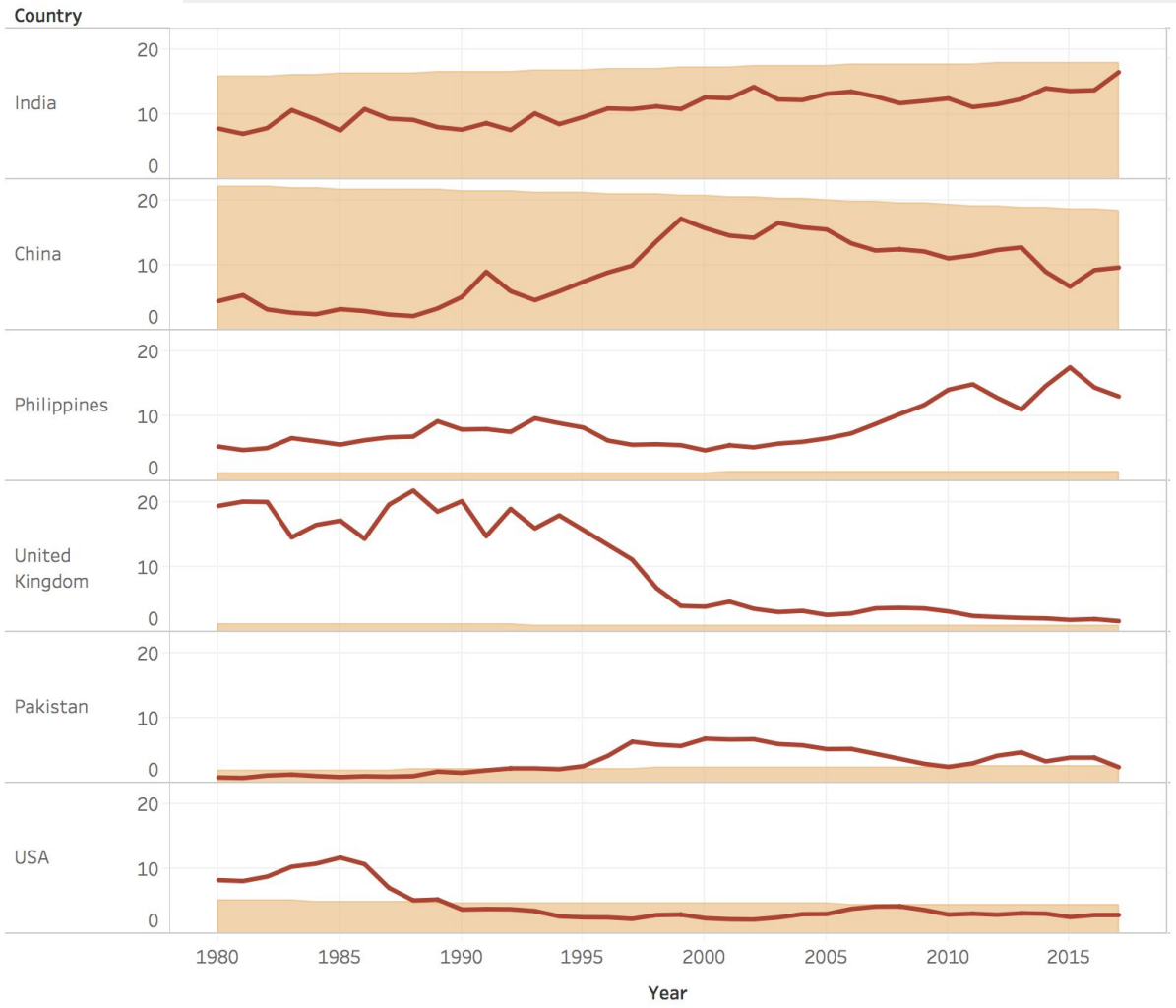


Does the percentage breakdown of Canadian immigration match the global population distribution?

This chart compares the percentage breakdown by country of all immigrant arrivals to Canada with the percentage breakdown of the global population.

Analysis #4

Canadian immigration vs. global population 1980-2017, top 6 countries



Highlights / Conclusion

- India and China account for a large part of the global population, as well as immigration to Canada, especially since the last 90s.
- Immigration from Pakistan and the USA are somewhat in line with their share of the global population, with some variation.
- British were coming to Canada at a disproportionate rate until the late 90s compared to their share of the global population.
- Starting in the early 2000s, Filipinos are immigrating to Canada at a much higher volume compared to their share of the global population.

Analysis #5

Canadian immigration vs. socioeconomic indicators

Can socioeconomic indicators explain immigration patterns?

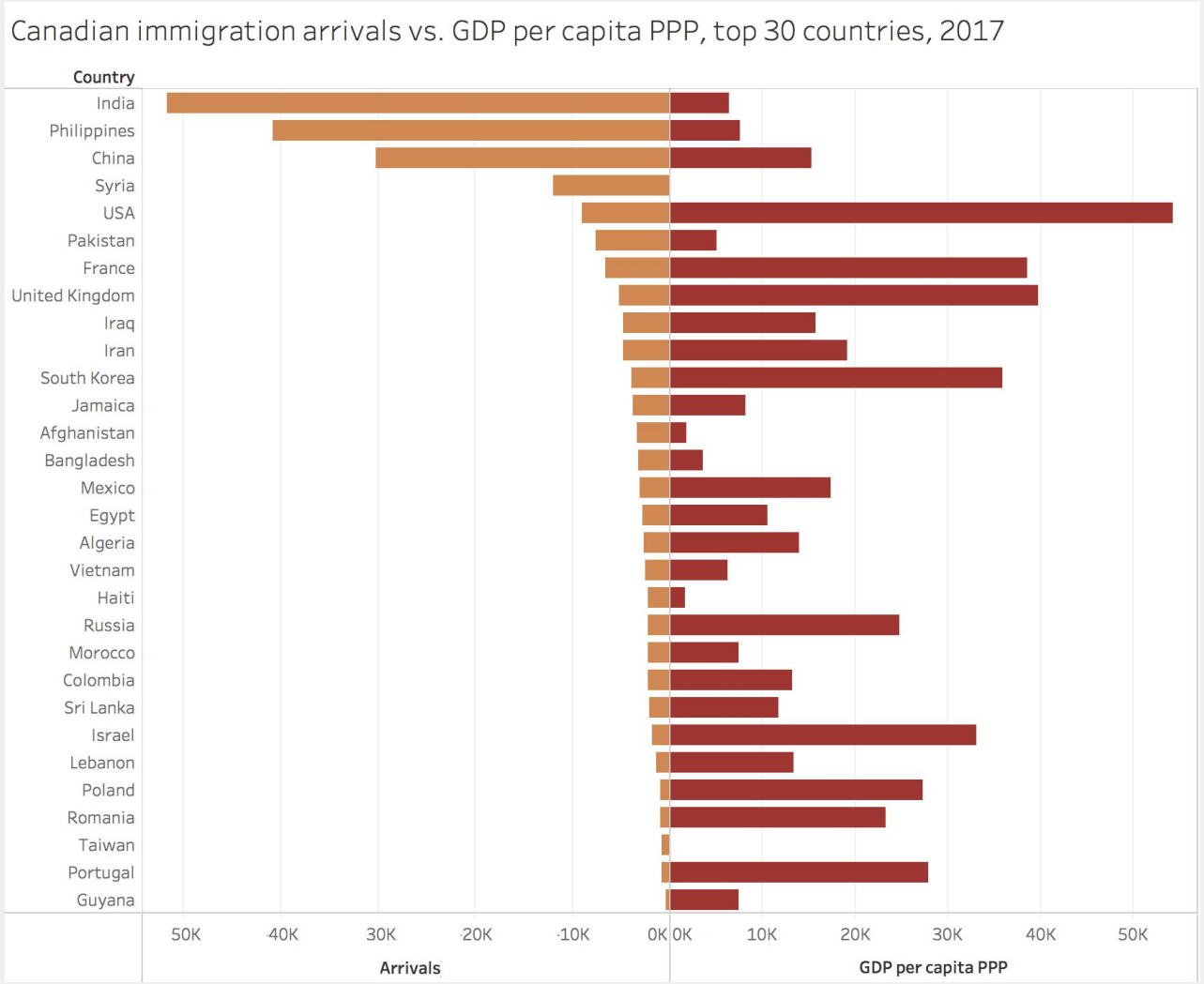
The next few graphs compare the number of immigrants arriving in Canada from each country and their respective socioeconomic score or indicator.

This is an attempt to find quick correlations between immigration and socioeconomic factors on a broad scale.

We created similar graphs to analyze 20 different indicators, but only included a few in the presentation as examples.

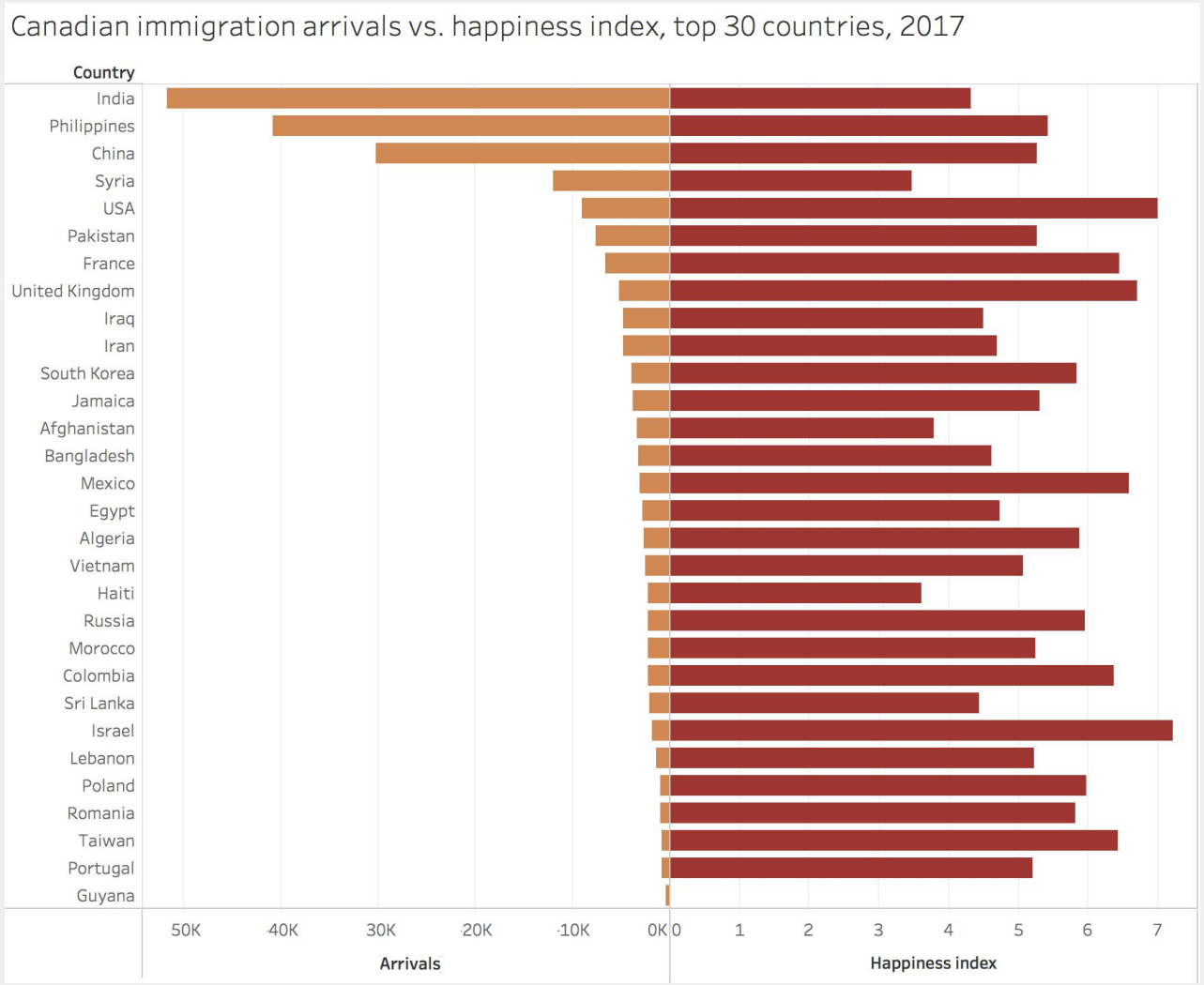
Analysis #5

Canadian immigration vs. socioeconomic indicators



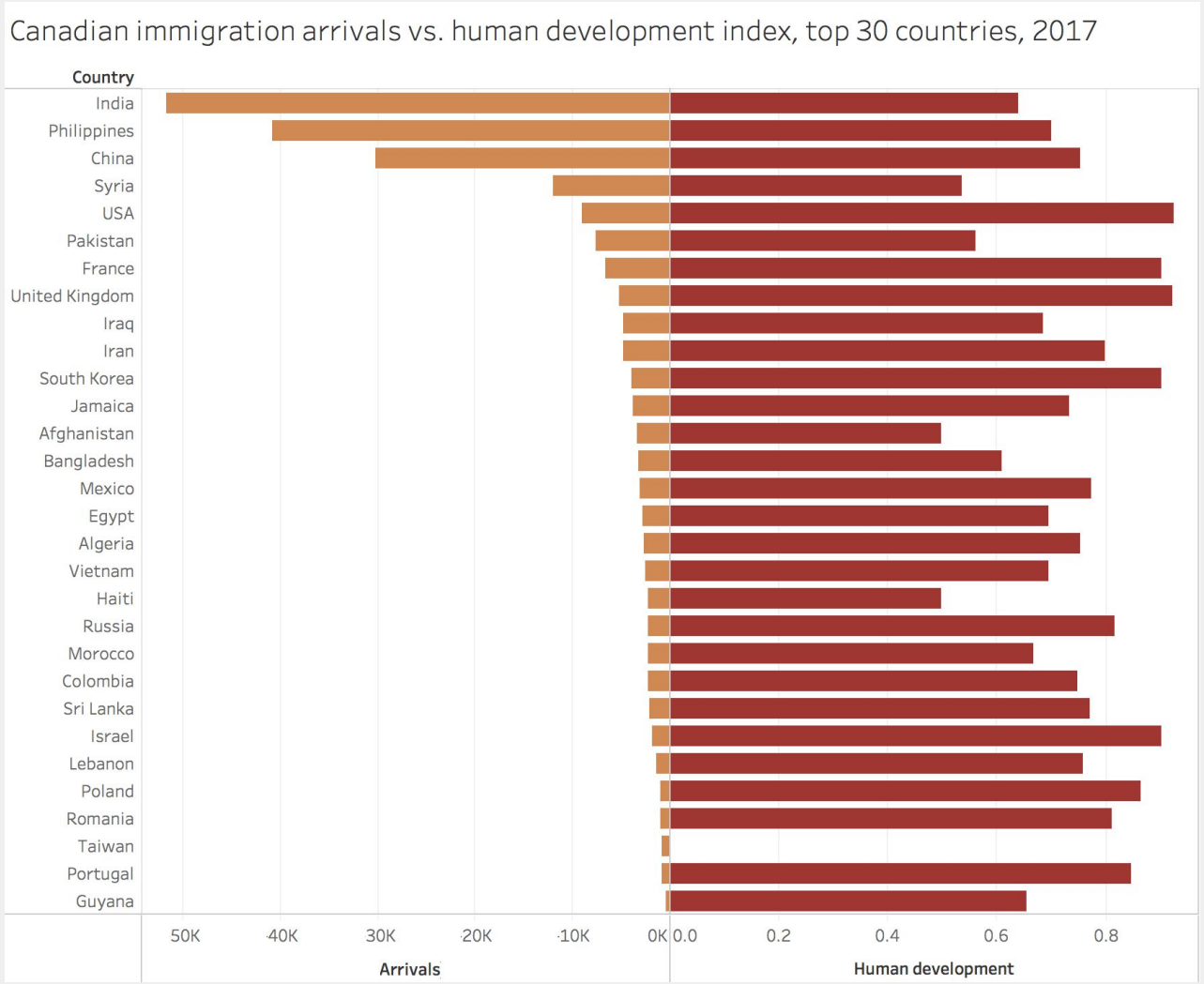
Analysis #5

Canadian immigration vs. socioeconomic indicators



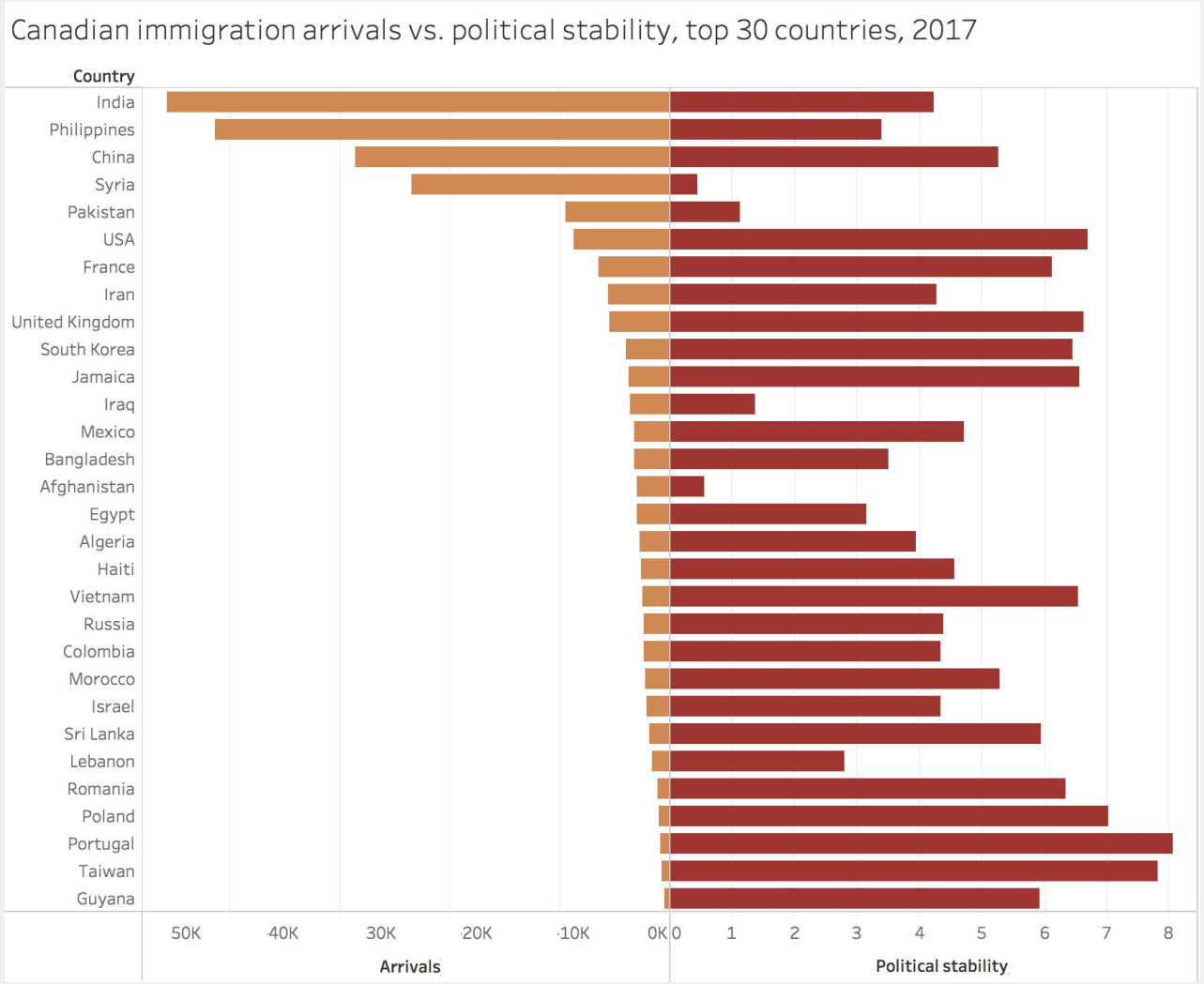
Analysis #5

Canadian immigration vs. socioeconomic indicators



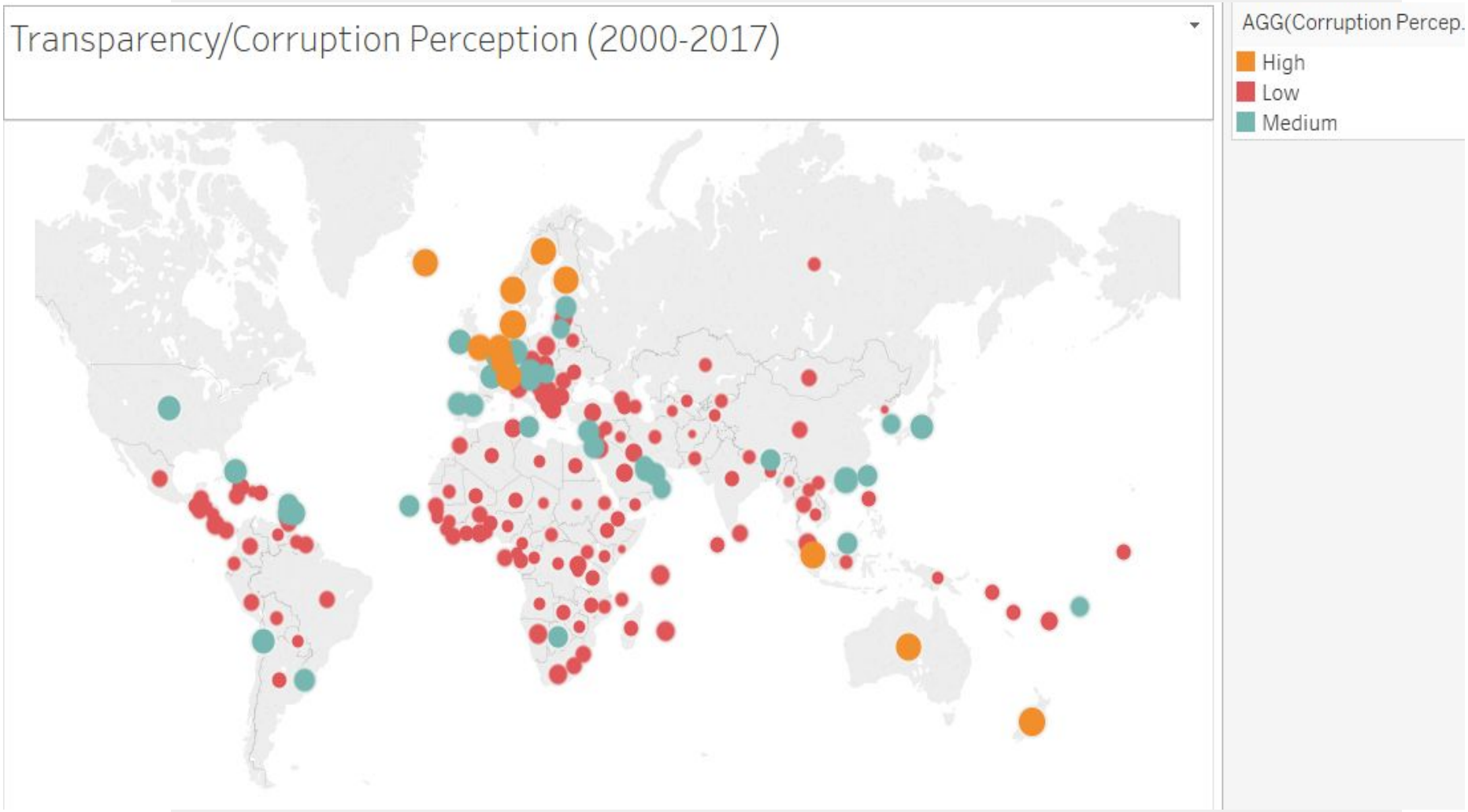
Analysis #5

Canadian immigration vs. socioeconomic indicators



Analysis #5

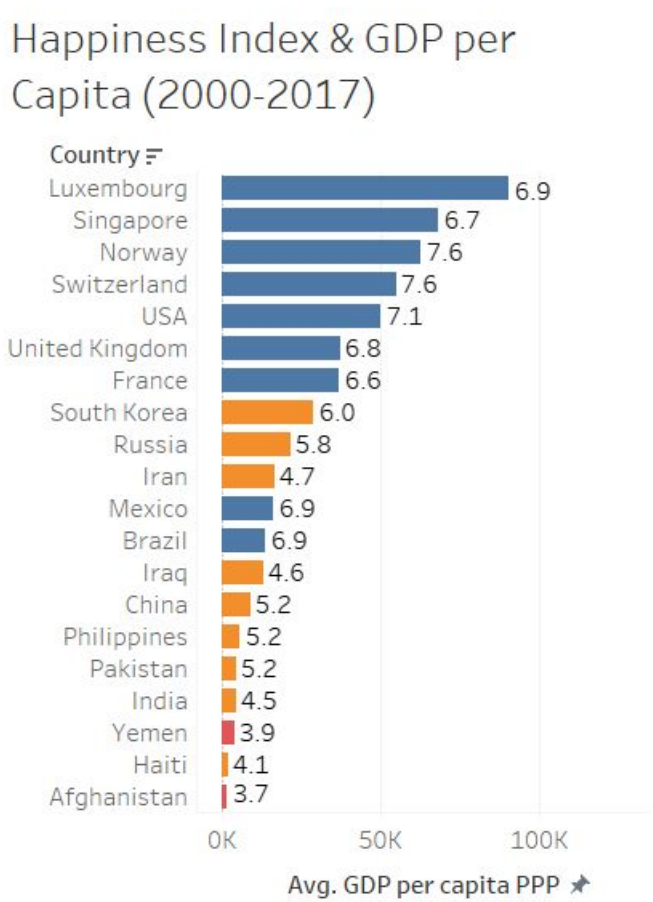
Canadian immigration vs. socioeconomic indicators

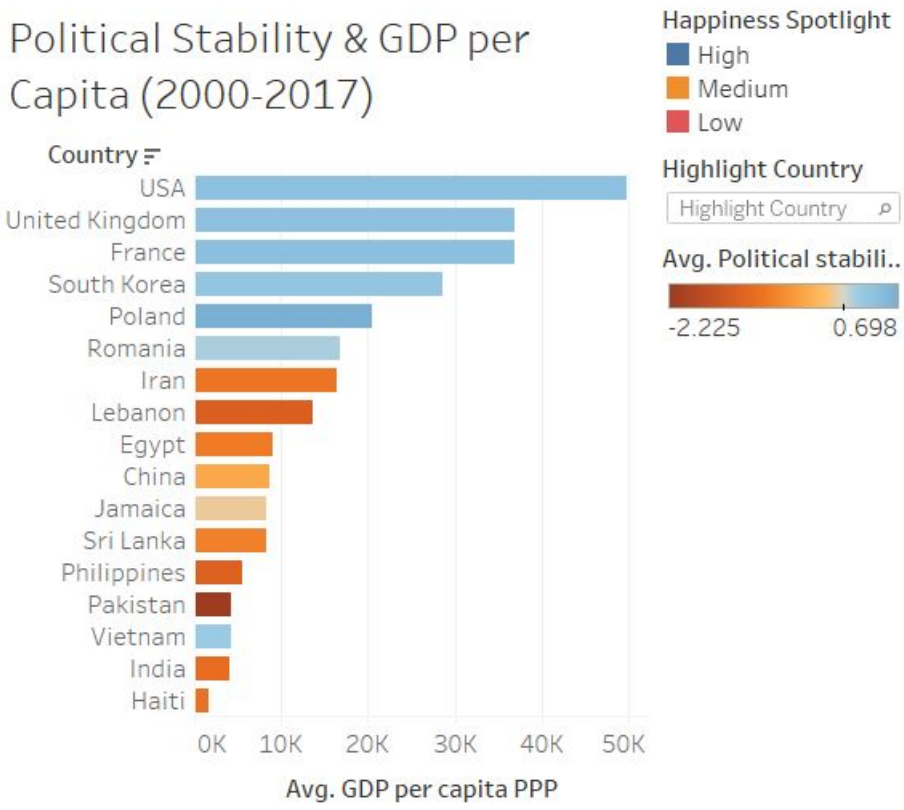


Analysis #5

Canadian immigration vs. socioeconomic indicators

Some additional Insights





Highlights/Conclusion

• No clear or direct correlation between 20 single socioeconomic indicators analyzed and Canadian immigration.

• This shows that there is no one size fits all when it comes to immigration, and the reasons for one group of people to immigrate are not necessarily the same as other groups.

• Some direct correlation between economic and socio-political factors across countries (subject to further statistical analysis)

Conclusion

Takeaways

What we learned....

Immigration patterns are extremely complex. As such, although it's possible to notice trends over time, it's not entirely possible to understand the motives or explain the reasons for immigration without looking at hundreds of other variables unique to each group of people.

Had we had more time...

We would have looked at the immigration policies of the Canadian government, as that might explain some of the variation and patterns in this analysis.

Thank you

Any questions?