

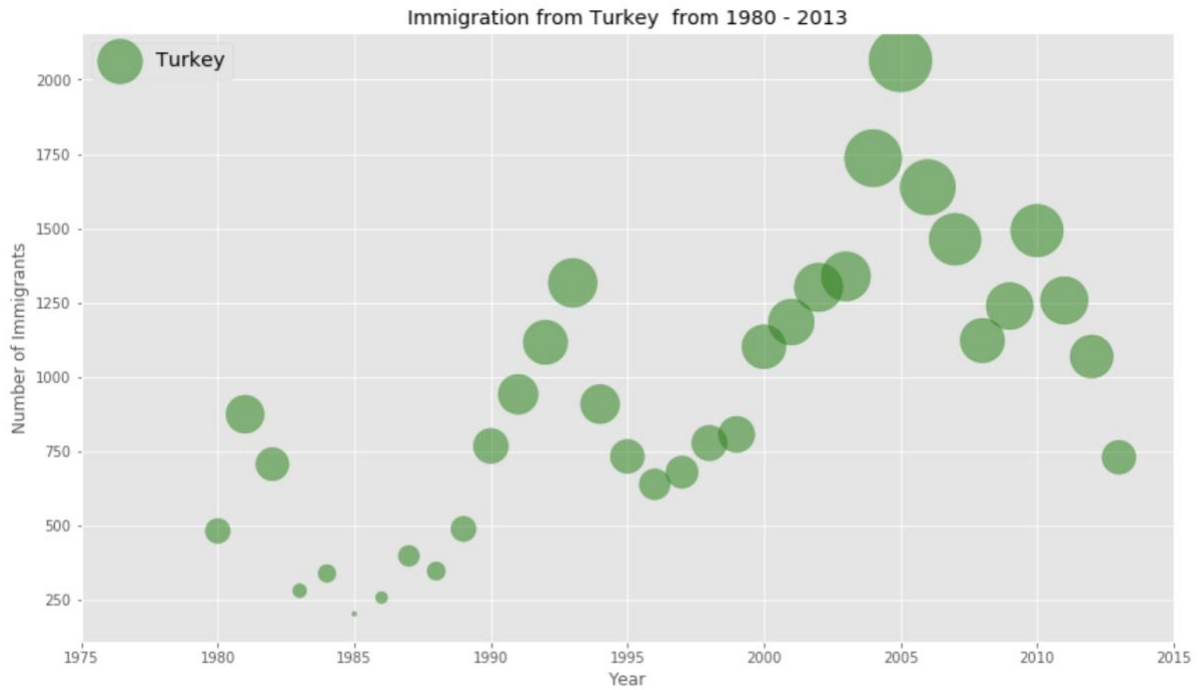
Introduction

Background and problem

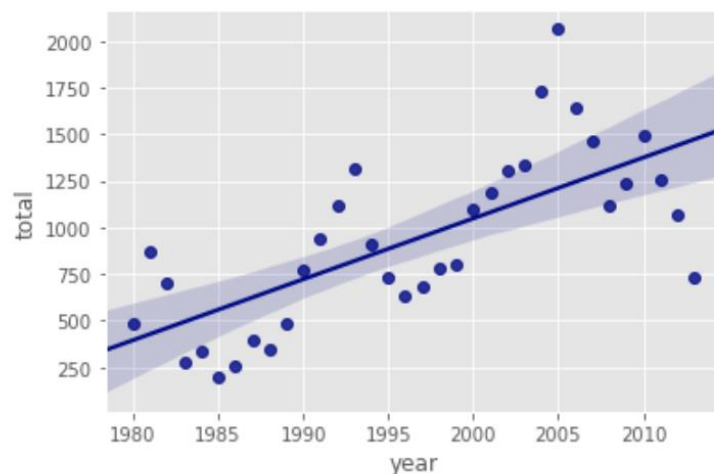
People have moved from their countries for centuries, for all sorts of reasons such as escaping conflict, violence, poverty, or past or future persecution based on race, religion, nationality, and/or membership in a particular social group or political opinion. Nowadays, people also migrate to developed countries to seek superior healthcare, education, jobs and business opportunities. In 2013, the percentage of international migrants worldwide increased by 33% with 59% of migrants targeting developed regions (https://en.wikipedia.org/wiki/Human_migration). Many people go through difficult period in their adaptation which may be related to having lived under very different conditions and circumstances in their countries of origin. I firmly believe that similarities between the neighborhoods of people's home and host countries might help them better adapt to the host country. Therefore, for this project I have picked one of the top migration corridors: Turkey – Canada and clustered neighborhoods by venues in their economic centers: Istanbul and Toronto. Clustering of neighborhoods might help migrants to find similar neighborhoods in Toronto and accelerate the adaptation process in their host country.

Migration facts

According to Canadian Census data, in the period of 1980 – 2013 more than 30.000 Turkish nationals migrated to Canada for different reasons. The majority of Turkish Canadians live in Ontario, mostly Toronto. The graph below shows the trend of immigration from Turkey to Canada in the period between 1980 and 2013:



It is seen from the graph that immigration from Turkey to Canada peaked in 2005 with more than 2,000 migrants. The potential reason of decrease in immigration to Canada after 2005 might be that Turkey's EU membership negotiations officially launched after intense bargaining. The process of softening EU's immigration policies for Turkish citizens might cause them to prefer EU countries over Canada for migration. Nevertheless, the overall trend of migration from Turkey to Canada is upward:



Data

Data of this project come from different sources. Migration data was retrieved from Statistics Canada (<https://www12.statcan.gc.ca/census-recensement/index-eng.cfm?MM=1>). After cleaning data and extracting yearly data for Turkey, I used Python “matplotlib” and “seaborn” libraries to obtain visualizations above.

The next task was to obtain data of Toronto’s neighborhoods. To do so, I used several libraries to scrape Wikipedia page (https://en.wikipedia.org/wiki/List_of_postal_codes_of_Canada:_M) which contained the list of postcodes in Toronto (*i.e. BeautifulSoup4, lxml, html5lib*). Then I downloaded coordinates of each postcodes from *colc.us* and constructed a final data set which looked like this:

	Postcode	Borough	Neighbourhood	Postal Code	Latitude	Longitude
0	M3A	North York	Parkwoods	M3A	43.753259	-79.329656
1	M4A	North York	Victoria Village	M4A	43.725882	-79.315572
2	M5A	Downtown Toronto	Harbourfront	M5A	43.654260	-79.360636
3	M5A	Downtown Toronto	Regent Park	M5A	43.654260	-79.360636
4	M6A	North York	Lawrence Heights	M6A	43.718518	-79.464763
5	M6A	North York	Lawrence Manor	M6A	43.718518	-79.464763
6	M7A	Queen's Park	Queen's Park	M7A	43.662301	-79.389494
7	M9A	Etobicoke	Islington Avenue	M9A	43.667856	-79.532242
8	M1B	Scarborough	Rouge	M1B	43.806686	-79.194353
9	M1B	Scarborough	Malvern	M1B	43.806686	-79.194353
10	M3B	North York	Don Mills North	M3B	43.745906	-79.352188
11	M4B	East York	Woodbine Gardens	M4B	43.706397	-79.309937
12	M4B	East York	Parkview Hill	M4B	43.706397	-79.309937

Next, I needed to obtain the list of Istanbul's neighborhoods with latitudes and longitudes. I extracted Istanbul's 39 neighborhoods with coordinates from a web page (<https://worldpostalcode.com/turkey/istanbul/>). I couldn't automate this process using Python's scrapping libraries, as every neighborhood was a hyperlink and it was needed to click on it to see its coordinates. Therefore, I manually collected coordinates of Istanbul's 39 neighborhoods:

	Neighbourhood	Latitude	Longitude
0	Adalar	40.876259	29.091027
1	Bagcilar	41.045061	28.833649
2	Bayrampasa	41.051248	28.898469
3	Beyoglu	41.028387	28.974045
4	Esenler	41.062002	28.869538
5	Gaziosmanpasa	41.073421	28.901556
6	Kartal	40.915036	29.203717
7	Sancaktepe	40.999438	29.225250
8	Sisli	41.061672	28.984261
9	Umraniye	41.027219	29.127459

Lastly, I extracted data on venues in both cities from Foursquare. I applied two criteria for venue data: venues had to be within 500m radius of the neighborhood and the maximum number of

venues around the neighborhood was limited to 100. The final data contained names and coordinates of neighborhoods, and names, coordinates and categories of venues in both cities.

	Neighborhood	Neighborhood Latitude	Neighborhood Longitude	Venue	Venue Latitude	Venue Longitude	Venue Category
0	Adalar	40.876259	29.091027	Merit Halki Palace Hotel	40.878802	29.090974	Hotel
1	Adalar	40.876259	29.091027	İnönü Evi Müzesi	40.878251	29.093647	History Museum
2	Adalar	40.876259	29.091027	L'isola Guesthouse	40.877038	29.096136	Bed & Breakfast
3	Adalar	40.876259	29.091027	Aqua Green Beach	40.880498	29.090354	Beach
4	Adalar	40.876259	29.091027	Huseyin Rahmi Gulpinar Muzesi	40.877224	29.092228	Museum
5	Adalar	40.876259	29.091027	Asaf Beach Club	40.879211	29.088317	Surf Spot
6	Adalar	40.876259	29.091027	heybeliada tenis kortu	40.878337	29.091221	Tennis Court
7	Adalar	40.876259	29.091027	heybeliada cicekli orman	40.878291	29.090232	Mountain
8	Adalar	40.876259	29.091027	Heybeliada Bisiklet Turu	40.878426	29.091641	Bike Rental / Bike Share
9	Adalar	40.876259	29.091027	Heybeliada Bayrakepe	40.874992	29.094010	Scenic Lookout