Alex Ram
STAT287 – Final Project
12/07/18

Migration Relates to Happiness

Abstract

How does a shift in population within a country relate to the happiness there? People will be left behind in some countries and locals will be met with new neighbors in others. We show that countries with high emigration rates are likely to have a relatively unhappy population and that countries with high immigration rates have happier populations. That is, there's a strong and positive correlation between a countries happiness score and its migration rate. To help show this relationship, we compare foreign-born and local-born happiness scores of countries to their migration rates and show that locals are happier in countries with higher migrant populations. To show a pattern between migration in general and happiness we also use local migration. This is shown by comparing urbanization rates of countries to their happiness scores, which similarly has a positive correlation. We also investigate these topics within regions of the world. We show that each region has a bimodal distribution of happiness scores and migration rates, each with a mode above and below the world average. This suggests that migration may be related to proximity, while still following the pattern shown in international and local migration.

Introduction

Most first world countries share the common conservative view that immigration is a problem[1]. They express how unhappy it makes them that foreign populations are growing in their country. Here we show that this view is either naïve or personal because the country is happier with more migrants. It is possible that migrants are moving to happier countries, but these countries would be even happier without them. Although it's true that people want to live in places with better social fabrics, GDP's and other factors, it's not true that migrants are reducing the happiness of locals. We show that countries with higher immigration tend to have happier local-born residents than foreign-born. This means immigrants aren't piggybacking on success of countries but creating a better wellbeing for locals than themselves.

1. "Does Migration Increase Happiness? It Depends."

The largest form of migration is rural to urban within a country, or urbanization[2]. This is a conflicting issue because it produces economic growth but may make people's homes unrecognizable. It turns out that urbanization follows the trend of international migration, meaning there is a positive correlation to happiness. This may not have always been the case, but recently countries with higher rural-urban migration rates tend to be the happier countries in the world. Using these trends, we've shown that both major forms of migration are positively related to happiness.

It makes sense that migrants wouldn't want to travel too far from home. This suggests that migration would happen regionally. To test this, we compared all countries happiness and migration rates within each region. The distribution tended to be spread bimodally with a mode above and bellow the world averages. Each distribution compliments the other, suggesting people are leaving their country to live nearby in search for a better life. Western Europe and the Middle East/North Africa are great examples of this. However, the distribution isn't a perfect compliment for every region, suggesting that some regions are more desirable than others. Still, the regions in these cases are close. For example, South Asia has almost all countries with positive migration rates whereas all countries from East Asia have negative rates. When these regions are combined, the bimodal distribution appears.

Data and Materials

There were three datasets used, The World Happiness Report (WHR), international migration and urbanization. In the World Happiness Report, happiness has been calculated for 156 countries and they've been doing this for the past ten years. They use seven factors in their calculation: dystopia, GDP, life expectancy, social support, freedom, generosity and corruption. These factors are weighted based on their importance in the culture of the country. This project uses the most recent WHR dataset (2018) and splits it into four .csv files: change_hap.csv contains the change in happiness of each country from 2008-2010 and now. foreign-local-hap.csv contains the foreign-born and local-born happiness of residents of each country. hap_factors.csv shows the score that each of the seven factors received for each country. happiness.csv shows the happiness of each country and how the seven factors were weighted.

2. The World Happiness Report, page 4

The international migration dataset (UNdata_migration.csv) came from UNdata and contains the migration rates of countries over the years 2005-2010 and 2015-2020. Mostly the recent migration rates were used but the earlier data was used to calculate the change in migration which was needed for one statistic. The urbanization dataset (urban-and-rural-population.csv) is from World in Data and contains the rural population and urban population of each country for the last fifty years. Both data sets were compared with the WHR data set and encountered similar problems. When manipulating the data, it was common to have different names for the same country, for example, “USA” and “United States”. These changes were done manually to these .csv files. Other times these migration datasets were missing countries entirely. Other online resources were used to find migration rates for missing countries making it possible to use every country on the WHR.

Results

Using pandas, each dataset was loaded as its own data frame (happiness_migration.py, line 27). To see that there was overlap (name of countries) in the datasets a simple visualization was used (happiness_migration.py, line 40). This also helped see if there could be a correlation between happiness and migration rates. This visualization printed the ten happiest countries in ascending order and ten least happy countries in descending order along with their migration rates:

Happiest	Least Happy
Finland -> 2.531	Botswana -> 1.297
Norway -> 5.258	Malawi -> -0.634
Denmark -> 2.647	Haiti -> -3.17
Iceland -> 1.128	Liberia -> -1.041
Switzerland -> 5.886	Syria -> -13.171
Netherlands -> 0.938	Rwanda -> -0.728
Canada -> 5.982	Yemen -> -1.05
New Zealand -> 3.15	Tanzania -> -0.686
Sweden -> 4.023	South Sudan -> 2.354
Australia -> 6.911	Central African Republic -> -6.338

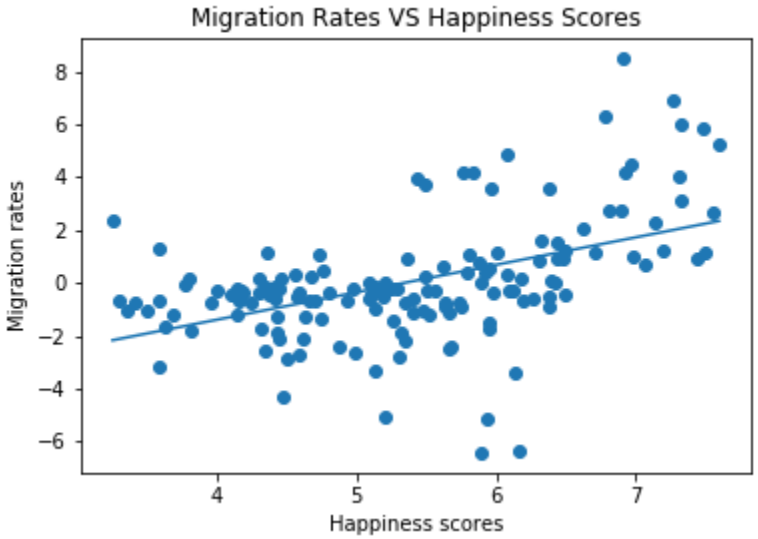
The chart shows that the happiest countries have higher migration rates than the least happy, backing up the hypothesis. The next steps were to shape the datasets, check for missing values and preprocess them. Each test that was preformed had a NumPy array created, taking what was needed from the data frames. This allowed for working with smaller datasets containing data only relevant to the current problem. Sometimes this required calculations between multiple columns of a data frame in order to fill one column of a NumPy array. Each array was a len(countries) by 2 shape. Below is a walkthrough of how each array was created.

- Happiness scores and migration rates (2018) (happiness_migration.py, line 60) – From the data frame created by happiness.csv select all countries happiness scores and enter it in column one. From the international migration data frame select the recent migration rate associated with each country and enter it into column two.
- Change in happiness scores and change in migration rates (2008/2010 – 2015/2017) (happiness_migration.py, line 85) – From the change in happiness data frame select all delta scores and enter it into column one. From the international migration dataset take the old and recent migration rates, find the difference and make it positive if migration rate increased on negative if it decreased. Enter it into column two.
- The difference in local vs foreign born happiness levels and migration rates (2018) (happiness_migration.py, line 110) – The difference in local-born and foreign-born happiness scores per country in column one (from the data frame produced by foreign-local-hap.csv) and the countries migration rate in column two.
- Migrations relationship to other factors used to calculate happiness score (2018) (happiness_migration.py, line 128) – Produces seven arrays, one for each factor used to calculate happiness. This factor has a score (taken from the data frame produced by hap_factors.csv), put it in column one and that countries migration rate in column two.
- Happiness scores and Urban-rural migration rates (2018) (happiness_migration.py, line 178) – Happiness scores taken from the happiness.csv data frame put in column one and the respective urban-rural migration rates taken from urban-and-rural-population.csv data frame put in column two

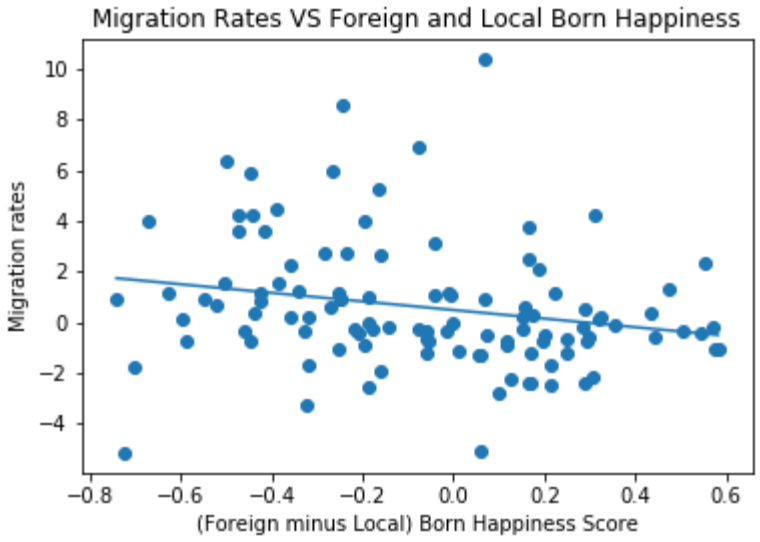
- Migration rates and happiness of regions (2018) (happiness_migration.py, line 214) – Regions can be found in the data frame produced by hap_factors.csv and a list can be created for each region containing the index of countries happiness/migration.

By creating arrays of these shapes, it was easy to pass them into functions for preprocessing, graphing and statistical analysis. Each array was preprocessed separately, based on how they were going to be used. For correlation plots, outliers were removed (line 258). This showed to greatly increase the correlation coefficient even though only a couple points were removed. For histograms the data was normalized (line 223) in order to fit both migration rates and happiness scores on the same plot. An average is shown by a dashed line for each variable to show where countries in the region stand compared to the world.

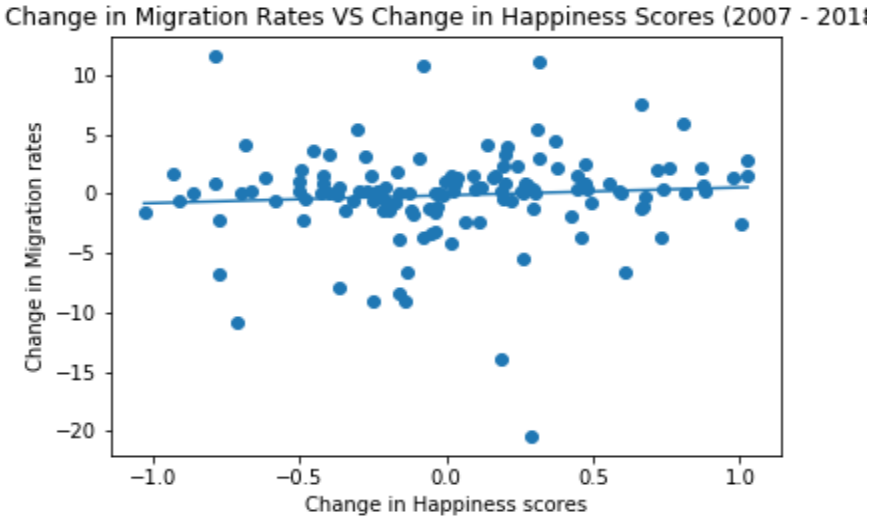
Bellow is a description for each linear correlation plot (happiness_migration.py, line 297). Pearson's correlation coefficient and p-value were used to determine how strong of a relationship each plot showed and if the data was significant.



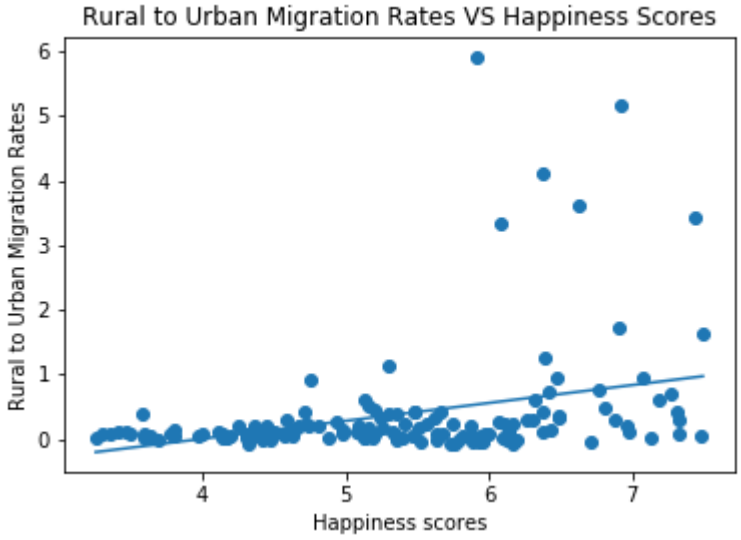
For migration rates vs happiness scores, we receive a Pearson's correlation coefficient of 0.48 and a p-value of 6.5e-10. This shows that there's a strong positive correlation between migration rates and happiness scores and that this correlation is very significant. This plot supports the hypothesis that happier countries have higher immigration and less happy countries have higher emigration.



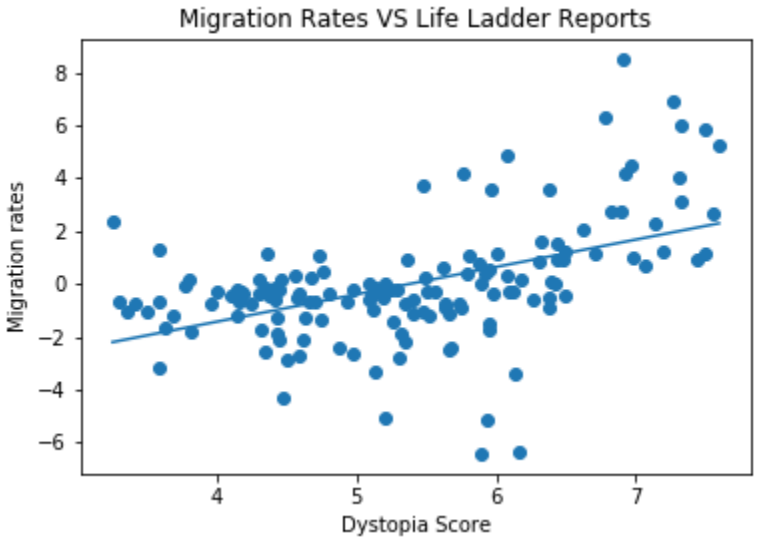
For migration rates vs foreign-born and local-born happiness, we receive a Pearson's correlation coefficient of -0.221 and a p-value of 0.02. This shows that the data is mediumly negatively correlated and the correlation is significant. On the x-axis, zero represents foreign-born being equally as happy as locally-born. When the x-value is negative we have that local-born are happier than foreign-born. This plot supports the hypothesis that locals are happier than migrants in countries with higher migration rates.



For change in migration rates vs change in happiness scores (2007 – 2018), we receive a Pearson's correlation coefficient of 0.078 and a p-value of 0.374. This shows there is no evidence that there is weak positive correlation. With more data and data manipulation this kind of plot may be able to show that a change in migration to countries will change the happiness score, or vice versa.



For rural to urban migration rates and happiness scores, we receive a Pearson's correlation coefficient of 0.335 and a p-value of 3.83e-05. This shows that the data is mediumly positively correlated and the correlation is very significant. This supports the hypothesis that happier countries are urbanizing. This leads to the idea that major forms of migration all can be positively correlated to happiness scores.

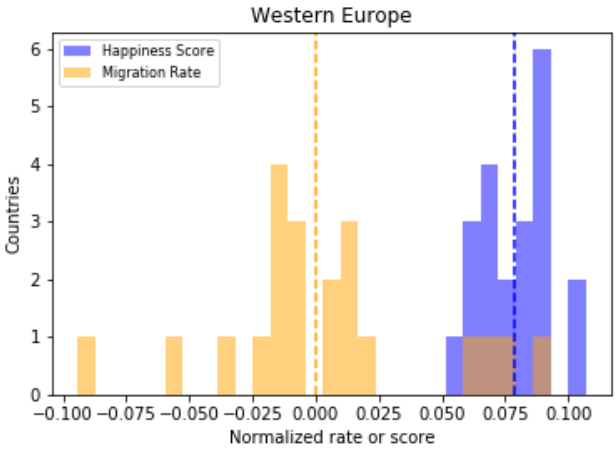


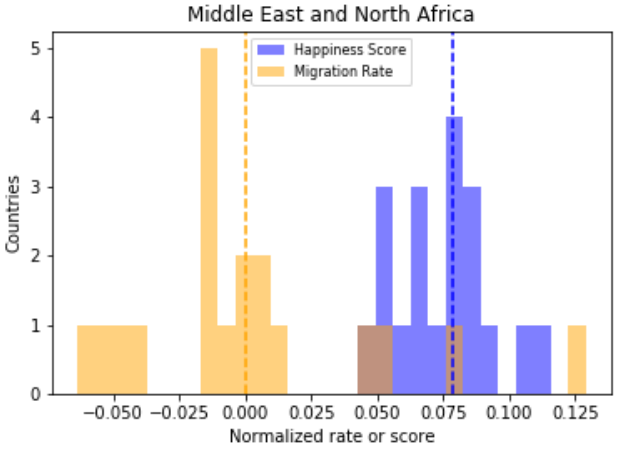
This is one of the graphs for the seven factors compared to migration rates. The others can be seen by running the happiness_migration.py script or checking the stats folder for the project. More information on how these categories were calculated can be found in the World Happiness Report[3]. A quick summary of the seven plots:

3. The World Happiness Report, page 18

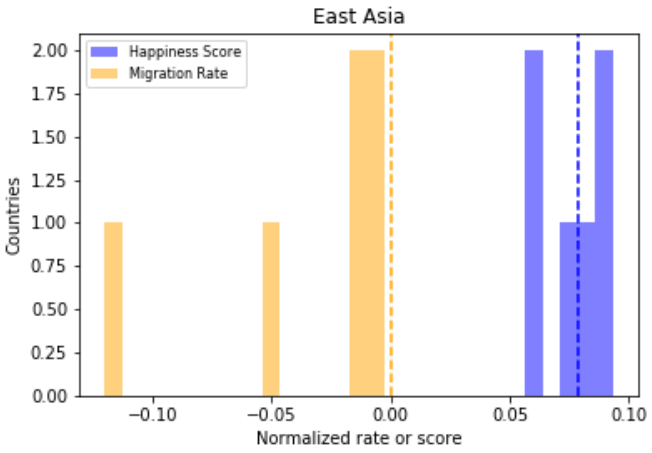
- Migration Rates VS Life Ladder Reports – The life ladder reports are surveys given to residents of a country where they rank their own happiness, averaged per country. Correlation coefficient: 0.488 p-value: 4.92e-10 meaning strongly positively correlated with very high significance.
- Migration Rates VS GDP - Correlation coefficient: 0.636 p-value: 3.79e-17 meaning very strongly positively correlated with very high significance.
- Migration Rates VS Life Expectancy - Correlation coefficient: 0.366 p-value: 8.24e-06 meaning very mediumly positively correlated with very high significance.
- Migration Rates VS Social Support - Correlation coefficient: 0.341p-value: 3.83e-05 meaning mediumly positively correlated with very high significance.
- Migration Rates VS Freedom - Correlation coefficient: 0.282 p-value: 0.0007 meaning mediumly positively correlated with high significance.
- Migration Rates VS Generosity - Correlation coefficient: 0.457 p-value: 1.42e-08 meaning strongly positively correlated with very high significance.
- Migration Rates VS Corruption - Correlation coefficient: -0.243 p-value: 0.0061 meaning mediumly negatively correlated with high significance.

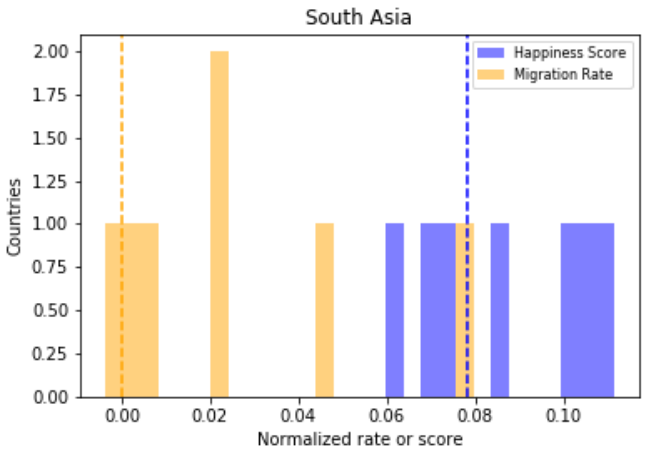
For the histograms we show both the migration rates and happiness scores (happiness_migration.py, line 241). These plots are only to suggest that migration happens regionally. Much more data would be needed to prove this hypothesis. Bellow are a few of the plots produced that were discussed earlier. The rest can be found by running the happiness_migration.py script or in the stats folder of the project.





These plots show that regions are split by countries that have negative and positive migration rates, and happiness scores above and below the average world happiness (represented by the dashed line). These splits form bimodal distributions with one mode on either side of the average. This suggests that migration may be happening regionally.





Some histograms show regions that are undesirable or very desirable, shown by migration rates. When combines a similar bimodal structure is formed, like the first histograms shown. These can be combined because the regions are next to each other. Migration happening regionally is just a hypothesis that is reasonably backed up by these figures, there may be other explanations and with further work this can be determined.

Discussion

We've learned that happiness in countries relates positively to migration rates around the world. This shows that migrants will move to countries with the ability for a better wellbeing and that countries with high emigration likely have low ability for wellbeing. It also shows that migrants contribute to the rise of wellbeing for locals in a country with more immigration. This is true because countries with higher migration rates tend to have happier local-born residents than foreign-born resident. Other forms of migration follow this pattern, such as urbanization. Lastly, we predict that migration may happen regionally due to bimodal patterns in migration rates above and below zero.

Future work is promising. We can show that changes in migration leads to change in happiness within a country, further proving that migrants positively contribute to the happiness of a country. This can be used to advise immigration policies around the world to help improve the happiness of people everywhere. We can also show movement of migrants internationally by collecting data on the origin ratios of migrants per country and use this to show that migration tends to happen regionally (or not). This can be used to predict if evenly spread ratios of migrants (more diversity) relates to happier countries, or maybe countries with migrants from closer countries are happier.

Bibliography

1. Helliwell, John F., et al. "World Happiness Report." *Worldhappiness.report*, 2018, s3.amazonaws.com/happiness-report/2018/WHR_web.pdf.
2. Polgreen, Linnea A., and Nicole B. Simpson. "Happiness and International Migration." *SpringerLink*, Springer, 22 Oct. 2010, link.springer.com/article/10.1007/s10902-010-9229-3.
3. Hsiao-Chien, Tsui. "What Affects Happiness: Absolute Income, Relative Income or Expected Income?" *NeuroImage*, Academic Press, 18 Nov. 2014, www.sciencedirect.com/science/article/pii/S0161893814000945.
4. Yarom, Barak. "The Immune System and Happiness." *NeuroImage*, Academic Press, 21 Mar. 2006, www.sciencedirect.com/science/article/pii/S1568997206000279.
5. Haybron, Daniel M. "Philosophy and the Science of Subjective Well-Being." *Googlegroups*, 2006, 0b0687ee-a-62cb3a1a-s-sites.googlegroups.com/site/danhaybron/happiness-and-well-being/PhilosophyandthescienceofSWB-finalrevisedsingle.pdf
6. Brüntrup, Michael. "Rising Food Prices – Causes, Implications, and Challenges for Development Policy ." *Die-Gdi*, Apr. 2004, www.die-gdi.de/uploads/media/BP_4.2008.pdf.
7. Hendriks, Martijn. "Does Migration Increase Happiness? It Depends." *Migrationpolicy.org*, 18 July 2018, www.migrationpolicy.org/article/does-migration-increase-happiness-it-depends.

Links to datasets

1. 2015 – 2017 WHR - https://www.kaggle.com/unsdsn/world-happiness
2. 2018 WHR - http://worldhappiness.report/ed/2018/
3. UN data migration rates - http://data.un.org/Data.aspx?d=PopDiv&f=variableID%3A85
4. Our World in Data Urbanization report - https://ourworldindata.org/urbanization