Migration and Happiness Scores

Andres, Alex and Imogen

September 17, 2020

Introduction

Question: correlations between migration and happiness?

- Project Goals
  - check correlations between migration and happiness.
  - To obtain data from different sources.
  - We merged several csv files from various sources to look for trends, check for correlations and conduct analyses.
  - To check correlations and to make regression analyses.
  - To visualise our results (Tableau).

IRON HACK

Questions and Hypotheses

Our guiding question was: Which factors cause migration? This question has two parts:

- Which factors cause people to emigrate?
- Which factors influence people's choice of destination country?

Starting out, we had two main hypotheses:

- Countries with high emigration rates have a low happiness rating.
- Countries with high immigration rates have a high happiness rating.

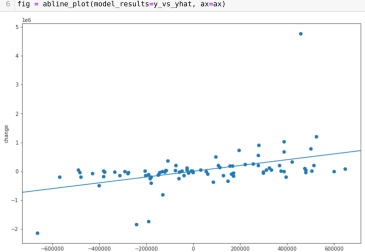
Workflow

We used linear regression (OLS) to check which features from our collected data were more likely to affect the migration (net and rate). First, we took a general approach, using all the information collected getting a very good approx (80%). To avoid noise (multicolinearity) in our model, we broke down our analysis using only a few factors for iteration, getting an accuracy only of 19%. So we concluded that among other factors, the happiness score of a country, while having a statistically significant correlation with migration, is not enough to explain migration patterns.

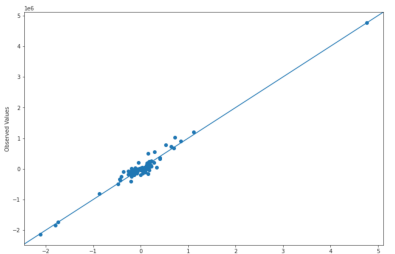
OLS Models

```
x=migration_income['Net Migration 2017']
```

```
from statsmodels.graphics.api import abline_plot
fig = plt.figure(figsize=(12,8))
ax = fig.add_subplot(111, ylabel='change', xlabel='Fitted Values')
ax.scatter(Net_2017_hap, x)
y_vs_yhat = sm.OLS(Net_2017_hap, sm.add_constant(Net_2017_hap, prepend=True)).fit()
fig = abline_plot(model_results=y_vs_yhat, ax=ax)
```

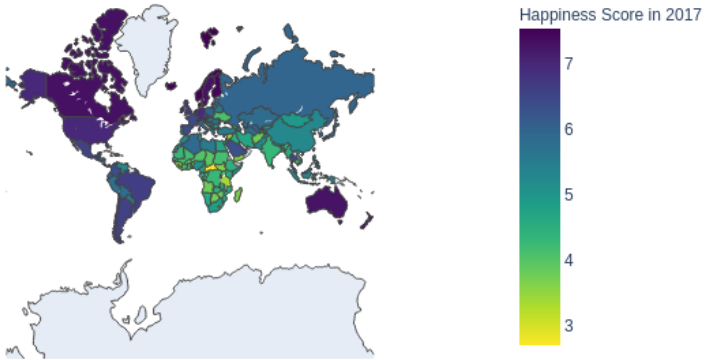



```
from statsmodels.graphics.api import abline_plot
fig = plt.figure(figsize=(12,8))
ax = fig.add_subplot(111, ylabel='Observed Values', xlabel='Fitted Values')
ax.scatter(Net_2017_hat, x)
y_vs_yhat = sm.OLS(Net_2017_hat, sm.add_constant(Net_2017_hat, prepend=True)).fit()
fig = abline_plot(model_results=y_vs_yhat, ax=ax)
```



link to Tableau

Happiness Score



Dataset

We utilised 3 separate databanks for our analysis.

- World Happiness Reports for years 2015 - 2019 Source: Gallup World Poll (https://www.kaggle.com/unsdsn/world-happiness)
- Global net migration for every 5 years between 1962 and 2017 Source: United Nations Population (https://data.worldbank.org/indicator/SM.POP.NETM?end=1981start=1981view=mapyear=2017)
- Global population for years 1960 - 2017 Source: United Nations Population Division (https://data.worldbank.org/indicator/SP.POP.TOTL)

Further questions

For a future project, it would be interesting to locate additional datasets containing more information on different migration categories, in particular irregular immigration, to reflect immigration flows more accurately.

VIELEN DANK!!!
MUCHAS GRACIAS!!!
THANKS!!!

This presentation was made using LaTeX