Happiness and representation: How well do UN General Debate speeches reflect the state of a country?

Bas van den Brink[11322195], Florens Douwes[11254483], Laura Hilhorst[11048999], David Jung[13646168], and Kaj Meijer[10509534]

University of Amsterdam

Abstract. This paper investigates the General Debate speeches of the United Nations in order to gain insights whether the sentiment within a speech correlates with the demographic developments within the respective country. This is done by conducting a in-depth data exploration with a sentiment analysis of the speeches followed by a linear regression which tries to discover any relationships between the World Happiness Report and the analyzed sentiment. However, no significant relationship could be identified between sentiment and demographic developments.

Keywords: Exploratory data analysis · Sentiment analysis · Text analysis · Linear Regression

1 Introduction

Every year at the beginning of the United Nations General Assembly, member states hold speeches during the General Debate. These speeches are not restricted to any topic; member states are free to talk about whatever they deem important to bring up. This gives the General Debate an interesting character which is different from the rest of the General Assembly. For one, it allows states to give a clear overview of their priorities, regardless of whether or not those topics are on the agenda for the General Assembly. Secondly, and more covertly, it presents the states with an opportunity to strategically profile themselves in front of the world, influencing other members' perception of them (Baturo, Dasandi & Mikhaylov [1]).

Keeping in mind this unique character of General Debate speeches, this exploratory data analysis aims to gain understanding into the representativeness of these speeches of the nation itself. Do the speeches accurately reflect the state of a country? In other words, do countries discuss their own losses and successes, and does the sentiment of the speech match the happiness of its citizens? This exploration is interesting because it may provide insight into the motivation behind the speeches. For example, a relationship between positive speeches and happy citizens might mean that nations use their speeches to celebrate prosperous times. Similarly, speeches with a negative sentiment that correlate with low scores of citizen happiness might mean they use the opportunity to mourn their losses, perhaps to convince other states to aid them.

The basis for this experiment is a combined dataset taken from the UN General Debate Corpus and the World Happiness report. All countries that appear in the set from 2005-2020 are included to see if there is any relationship between the sentiment of a country's speech and several metrics relating to the happiness of its citizens. This is done by first exploring the dataset thoroughly, after which linear regression is used to try and discover relations between the World Happiness Report dataset and the analyzed sentiment of General Debate speeches.

2 Methodology

2.1 Data sources

The data for this work was obtained from the Happiness Report 2021 [2], the UN General Debate Corpus (UNGDC) [3] and from the UN Statistics Division [4]. The World happiness report is

a yearly effort of the United Nations to try to estimate how “happy” the average resident of a country is. The happiness report of 2021 was focused on the effects of COVID-19, which had a big impact in all our lives. The report also includes all results from previous years. The happiness report includes attributes such as the GDP of a country, the average life span, and relative social support and freedom. Residents were questioned these with a Cantril ladder. Residents were asked to think of a ladder, with the best possible life being a 10, and the worst possible life being a 0.
The UN General Debate Corpus contains records all UN speeches given by country representatives at the General Debate from 1970 onward. These speeches are about all kinds of subjects, but mostly focused on major issues in the world politics. All country speeches from 2005 on to 2021 was used in this paper. The UN Statistical Division dataset is used to combine the different usages of naming the countries. For example, the happiness report uses “Afghanistan” as the descriptor, but the Debate Corpus uses “AFG”.

2.2 Preprocessing & merging

In order to gain the full names of the countries, a natural join of the two datasets with the speeches and the detailed information about the countries was performed on the variable 'ISO-alpha3 Code'. Upon these speeches a sentiment analysis was conducted with the help of the NLTK SentimentIntensityAnalyzer. For each speech, a value of neutral (neu), positive (pos) and negative (neg) sentiment was calculated and added to a pandas dataframe. This will be discussed in further detail in Section 2.3. Lastly, a natural join of this sentiment dataset and the happiness dataset was performed using the two variables 'Country' and 'Year'.

2.3 Sentiment analysis

We applied sentiment analysis to the speeches of the UN by utilizing the SentimenIntensity-Analyzer from the nltk package. This analyzer works with English text, and analyzes words such as *don't*, *aren't*, *doesn't* (negative connotation), *absolutely*, *enormously*, *incredibly* (positive connotation), and even idioms such as *the bomb* and *cut the mustard*[5]. While effective for such a simple analyzer, there are some constraints. The analyzer can only analyze words (and very specific idioms), but cannot perform sentiment analysis on the structure of the whole text. It has some other constraints, such as the incapability of handling quoted text or negated text. However, its simplicity and speed make it appropriate for the scope of this experiment.

The analyzer bundles the different connotations or (*moods* in a single score. This score is a a dictionary with the positive (‘pos’), neutral (‘neu’), negative (‘neg’) and compound (combined, ‘compound’) sensitivity scores. The UN Debate corpus was calculated for all preprocessed documents and was stored with the same index as the UNGDC dataset ('Year', 'ISO-alpha3 Code').

2.4 Overview and distribution of data

It is clearly visible that not all of the variables follow a normal distribution, which can be seen in Figure 1. A more detailed visualization of the data used in prediction can also be found in Figures 7-12 in the Appendix.

- **Normally distributed:** Neu, Pos and Readability
- **Left-Skewed:** Social Support, Freedom to make life choices and Perceptions of corruption
- **Right-Skewed:** Neg, Generosity and Negative affect.

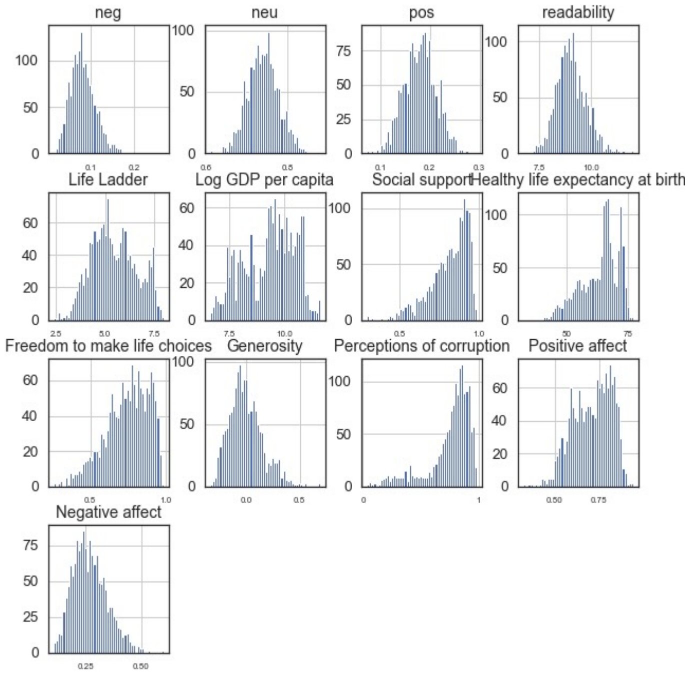

Fig. 1. Distribution of the different data sources

2.5 Correlation in the data

Sentiment analysis of the speeches is combined with the UN Happiness Report. This data is combined with the sentiment analysis that was performed in Section 2.3. The data has been normalized by re-scaling it to fit the range $[-1, 1]$.

The correlation matrix shown in Figure 2 shows the relation of the negative, neutral and positive sentiment of the speeches given by the different countries, related to their happiness attributes such as the "Life Ladder" and the "Generosity" of a country.

2.6 Linear regression

Linear regression was first applied to both datasets. This method is a linear approach to predict the value of a numeric variable based on the value of another numeric variable. In our case, we try to predict the different sentiments given in the corpus of the UN Speeches by looking at the different happiness statistics given by the questionnaire of the happiness report. We do this by applying the sklearn LinearRegression model. We first divide the dataset into a training set and a test set. This is an important step: testing on the training set will always give good results. We use the `train_test_split` method of sklearn for this. Then – first fitting with `.fit` – and then predicting – with `.predict` – we analyze the results in a few graphics. See Figure 4. While the analyzer applied some coefficient to the regression, we can see that the data does not fit the linear approach. The data is too much in a "cloud", which linear regression is not necessarily good at.

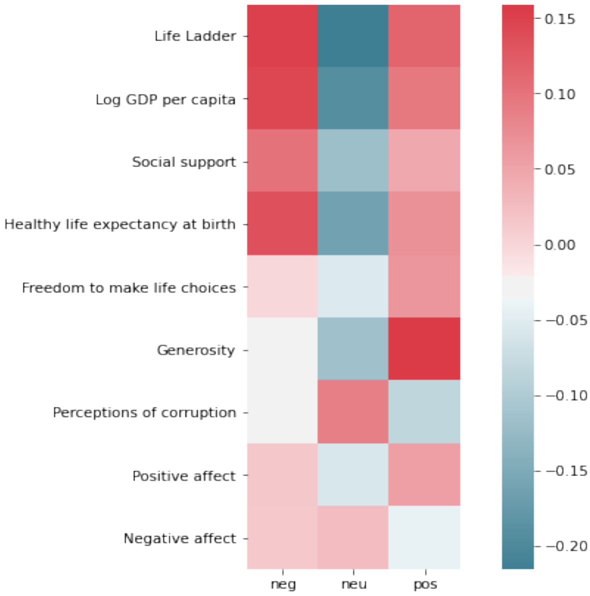

Fig. 2. Correlation matrix between the happiness and speech sentiment data

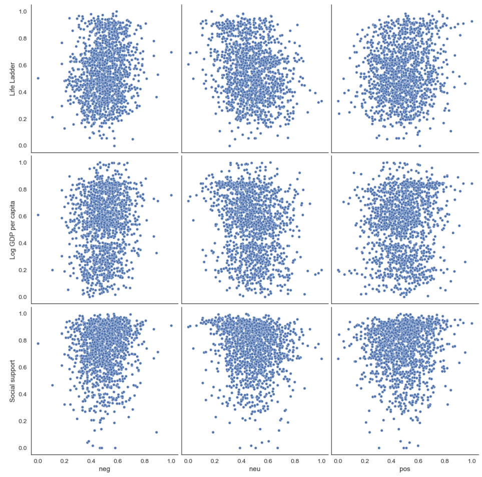

Fig. 3. Scatter plots of the Life Ladder, Log GDP and Social Support together with the negative, neutral and positive sentiments of the speeches

3 Results and Discussion

3.1 Results

Correlation between features A thorough data exploration was performed, but the features showed little correlation. Some combinations did show a slight correlation, namely ('Life Ladder', 'neg'), ('Life Ladder', 'neu') and ('Generosity', 'pos') which can be seen in Figure 2. The lack of strong correlations can be explained by using Figure 3. The data points of the scatter plots, between the sentiment analysis data and the happiness data, are forming one big blob. These blobs of data points do not have a perfect center or mean, so that results in a slight positive or negative correlation between the sentiment and happiness data.

Linear Regression Another method used to look for relationships in the data is linear regression. One can see in Figure 3 that there is hardly any linear relationship between 'Social Support', 'Log GDP per capita' and 'Life Ladder' against the different sentiment scores. These observations can be supported by the linear regression model, seen in Figure 4. An average R^2 of 0.03 was obtained, which indicates that the linear regression has little to no predictive power. This R^2 is to be expected if you look at Figure 4.

3.2 Discussion

The experiment showed no visible relations between the sentiment of the UN General Debate speeches and the metrics in the World Happiness report. Neither the correlation tests nor linear regression showed significant results.

There are multiple possible explanations for the lack of result in this experiment. Firstly, the outcome of the experiment might suggest that UN General Debate speeches are not really geared towards internal issues, but focused more on international issues, as this is generally more the case in the UN General Assembly.

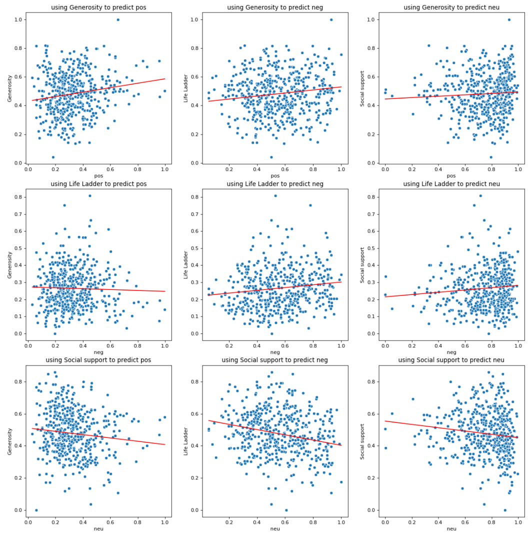

Fig. 4. Scatter plots and linear regression of the Generosity, Life Ladder and Social Support together with the negative, neutral and positive sentiments of the speeches

Another possible explanation is that the sentiment analyzer used in the experiment is not suitable for this type of document. The NLTK Sentiment Analyzer is a pre-trained method with a default dictionary of negative and positive words. A premise for future research could be to repeat the experiment using a custom sentiment analyzer geared toward diplomatic language.

4 Conclusion

In this experiment, an attempt was made to find relationships between the sentiment of UN General Debate speeches and the World Happiness Reports. The goal of this experiment was to see whether speeches made by country representatives during the UN General Debate accurately reflect its internal state of happiness. Using correlation and linear regression, the conclusion can be made that no relationships were found. Multiple explanations for this are possible, but further research is needed to confirm these speculations.

References

1. Baturo, A., Dasandi, N., & Mikhaylov, S. J. (2017). Understanding state preferences with text as data: Introducing the UN General Debate corpus. Research & Politics, 4(2), 2053168017712821.
2. Helliwell, John F., Richard Layard, Jeffrey Sachs, and Jan-Emmanuel De Neve, eds. 2021. World Happiness Report 2021. New York: Sustainable Development Solutions Network
3. Harvard Dataverse, `https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/0TJX8Y`. Last accessed 30 Sep 2021
4. UN Statistics Devision, `https://unstats.un.org/unsd/methodology/m49/overview/`. Last accessed 30 Sep 2021
5. NLTK Vader Sentiment Analyzer, `https://www.nltk.org/_modules/nltk/sentiment/vader.html`. Last accessed 1 Oktober 2021.
6. @articlekingma2014adam, title=Adam: A method for stochastic optimization, author=Kingma, Diederik P and Ba, Jimmy, journal=arXiv preprint arXiv:1412.6980, year=2014

5 Appendix

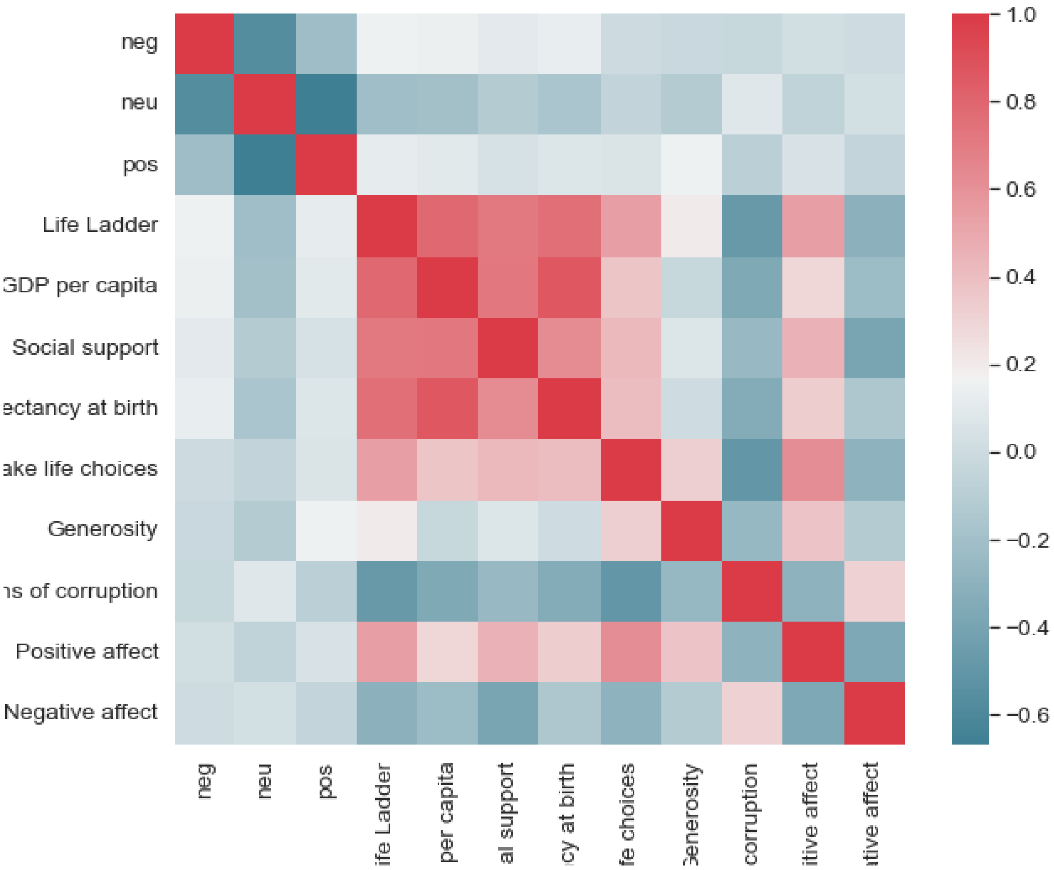

Fig. 5. Correlation between all columns

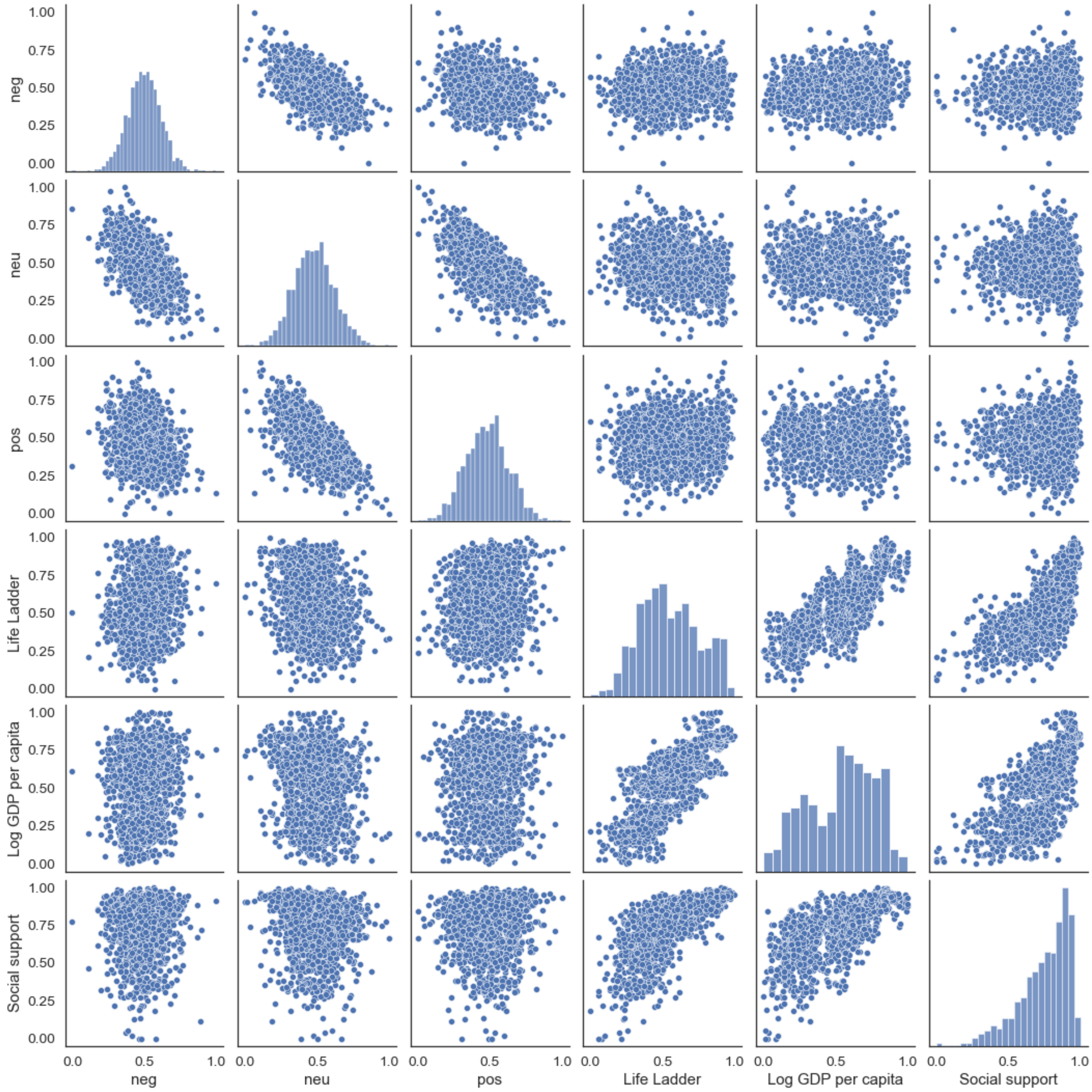

Fig. 6. Relationships between all columns

	neg	neu	pos	Liflad	LogGDP	Soc.Sup.	H.L expact	Life joices	Gener	PercepCorr	P.affect	N.affect
count	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00	1507.00
mean	0.50	0.48	0.48	0.55	0.54	0.75	0.69	0.66	0.33	0.75	0.63	0.33
std	0.11	0.14	0.15	0.20	0.23	0.18	0.17	0.20	0.16	0.20	0.18	0.16
min	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
25%	0.43	0.39	0.38	0.39	0.35	0.65	0.58	0.53	0.22	0.70	0.49	0.21
50%	0.50	0.48	0.48	0.53	0.56	0.79	0.73	0.69	0.31	0.81	0.65	0.31
75%	0.58	0.57	0.57	0.70	0.72	0.89	0.82	0.82	0.41	0.89	0.77	0.43
max	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 1. Numerical summary of the data

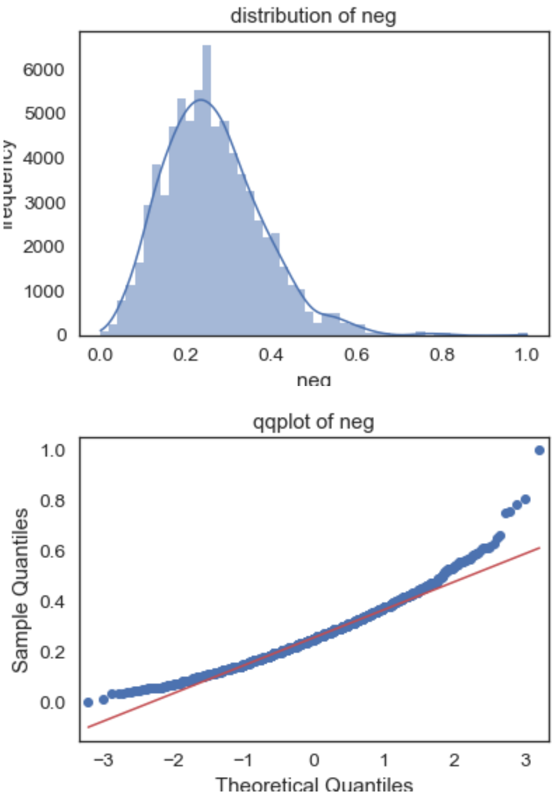

Fig. 7. Distibution & QQ-plot of neg

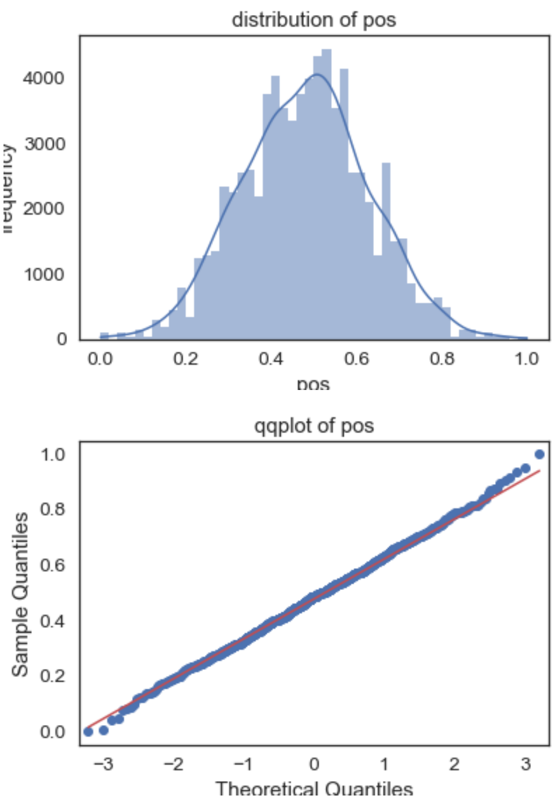

Fig. 8. Distibution & QQ-plot of pos

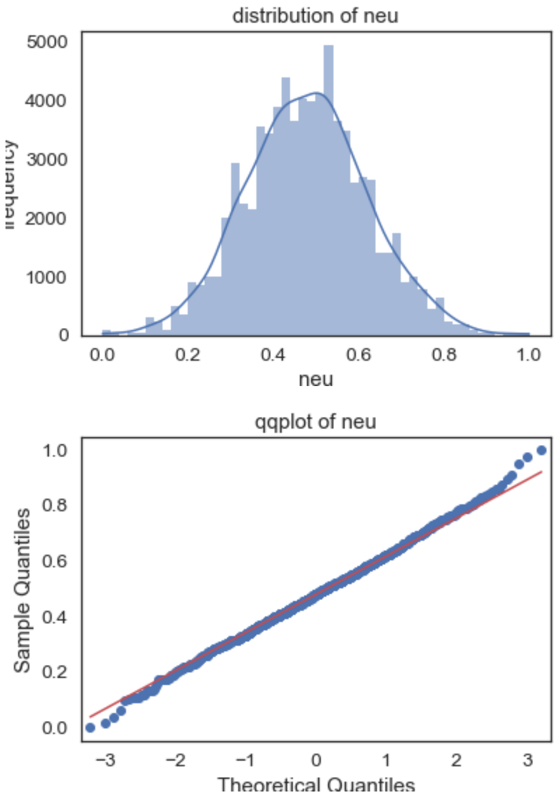

Fig. 9. Distibution & QQ-plot of neu

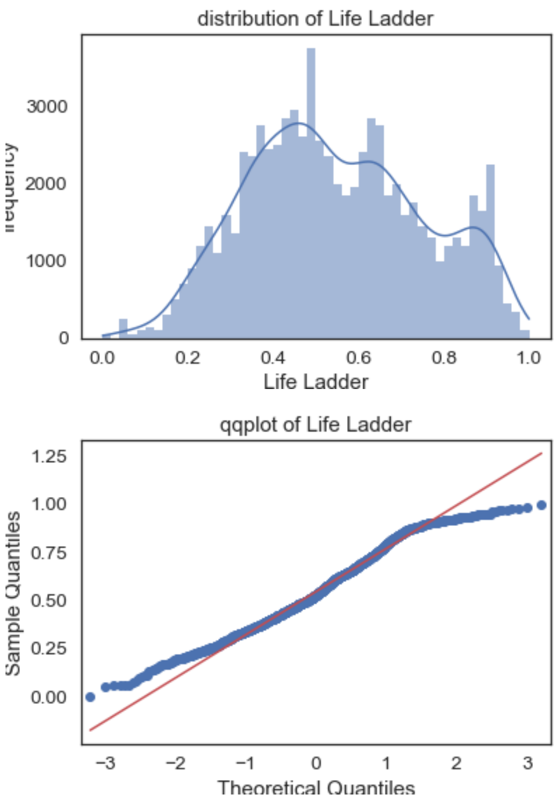

Fig. 10. Distibution & QQ-plot of Life Ladder

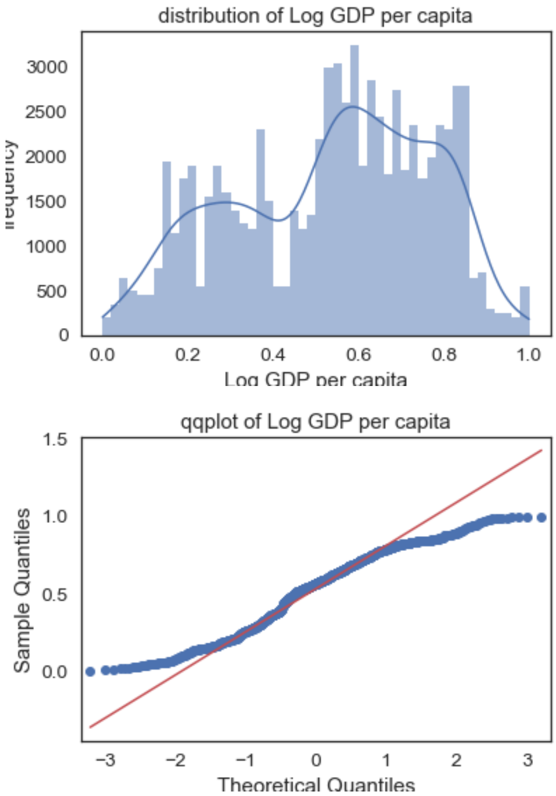

Fig. 11. Distibution & QQ-plot of Log GDP per capita

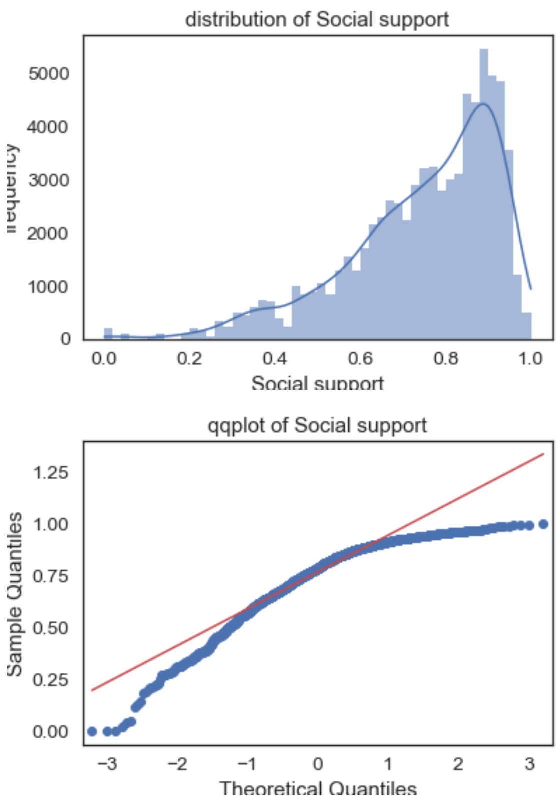

Fig. 12. Distibution & QQ-plot of Social support

	Total missing	Percent missing
Perceptions of corruption	99	0.059
Generosity	55	0.033
Freedom to make life choices	25	0.015
Positive affect	19	0.011
Negative affect	14	0.008
Log GDP per capita	12	0.007
Social support	10	0.006
Life Ladder	0	0.000
Healthy life expectancy at birth	0	0.000
iso3	0	0.000
ISO-alpha3 Code	0	0.000
neu	0	0.000
neg	0	0.000
Speech	0	0.000
Session	0	0.000
Developed / Developing Countries	0	0.000
Sub-region Name	0	0.000
Region Name	0	0.000
pos	0	0.000

Table 2. Missing data