

Multiple Regression, Clustering, Correlation and Visualization Analysis for Human Freedom Index

Yuzhe Yang, Siyuan Liu, Yuqian Dai, Laiyuan Zhang and Lei Guo
{yy1a19, sl18n19, yd6u19, lz4y19 and lg1y19} @soton.ac.uk

1. Introduction

Our team name is “Deep Dark Fantasy” and we choose to analyze the Human Freedom Index (HFI). The Human Freedom Index (HFI) presents the state of human freedom in the world based on a broad measure that encompasses personal, civil, and economic freedom. The utilized datasets in this research include multi factors, which contains Rule of Law, Security and Safety, Movement, Religion, Association, Assembly, and Civil Society, and etc. Our motivation is to figure out the principal components in HFI, analyze the correlation among different factors, dig out the impacts of different factors and visualize the data from different levels to give an intuitive observation. Data can be obtained from: <https://www.cato.org/human-freedom-index-new>.

2. Expected Approaches

In this research, to understand the data thoroughly, we decide to decompose the overall mission into seven sections and we figure out some potential problems in different section. In addition, we also propose potential solutions or approaches related to different mission, which is shown as below:

- A. Data cleaning (Missing data, Outlier Value, etc.)
- B. Visualization (Raw data level; Feature level; Output level)
- C. Machine learning based approaches comparison.
(i.e. Logistic Regression, Random Forest, SVR, MLP, Bayes Regression, AdaBoost and XGBoost)
- D. Comparison with deep learning based approaches.
- E. Correlation analysis (i.e. PLCC, SRCC, KLCC)
- F. Clustering analysis (i.e. K-Means, Spectral Clustering)
- G. Academic report

3. Task Division

Based on the aforementioned mission sections and personal research interests, we allocate the tasks to our group members separately. The allocation is shown as below:

- a) Yuzhe Yang is responsible for section C, D, E and G.
- b) Siyuan Liu is responsible for section A and B.
- c) Yuqian Dai is responsible for section C and G.
- d) Laiyuan Zhang is responsible for section D, E and F.
- e) Lei Guo is responsible for section A and G.

(Note: Each member needs to submit a sub-report related to the allocated tasks, which will be re-organized and adapted into the final group report. Therefore, their contributions to the final group report are equal.)