

Exploring the Factors Affecting Happiness Level Using Regression Analysis

Research Aim: In my project, I want to examine the economic factors influencing happiness levels across different countries using regression analysis of World Values Survey (WVS) data. By analyzing survey responses on socio-demographic characteristics and economic factors, I wish to identify significant predictors of happiness.

Method:

For data collection, I will gather data from World Values Survey (WVS) from the official WVS website or data repositories such as the GESIS Data Archive. I thought about using The WVS dataset since it contains a wide range of variables related to socio-demographic characteristics and economic indicators collected through surveys conducted in many countries. Since in WVS, data consists of answers given to questions regarding the specific topic, for the happiness level, answers to the question related to “happiness” will be used. Same logic will be applied while collecting data for economic indicators which will be **income level and employment status**. After the collection of relevant data, I will clean the data and handle the missing values for clear analysis. Then, before doing the regression analysis, I will visualize the data via scatter plots to detect outliers and see whether there is an association between happiness and economic factors or not. Also, I will use histograms, and correlation matrices to see the type of relationship between happiness and economic indicators to be selected, as mentioned above. After this step, I will develop my regression model to make a prediction about happiness level based on income level and employment status. I will use a linear regression model. Also, I will use logistic regression too by taking the happiness variable as a binary outcome. After running the regression models, I will evaluate the models to make a comparison between them and see which one is more efficient in explaining the relationship between happiness and economic factors. For that part, I will look at R squared values to see how much of the change in outcome variables can be attributed to changes in economic factors. Also, I will look at mean absolute error, and root mean squared error to see the significance of coefficients for each factor examined, and make an interpretation on their impact on happiness levels. After those steps, I will interpret the results of regression analysis models to identify whether income and employment status are significant predictors of happiness levels. Since I will repeat the process for several countries (will be selected later on), I will also make a comparison between how the relationship between happiness and economic indicators are affected based on the country. Then, in my written report, I will connect my model to my interpretation which will be connected to possible policy implications to promote happiness and well-being across countries in relation to economic factors. Jupyter Notebook will be used throughout the entire process.

