The raw data is named ‘MPI\_national.csv’ which is obtained in Kaggle and here is the reference: <https://www.kaggle.com/jamesmuniu/mpi-nationalcsv>. The data was first published in 2010 in the United Nations Development Program’s Human Development Report and has been abstracted and uploaded to the Kaggle. Kaggle is the world largest community of data scientists and machine learners. As such a famous and reliable community, the data is believed to be true and reliable. This data is downloaded for free and can be read and used with the appropriate reference.

The size of data is 103(rows) \* 8(columns). In other words, there are 102 unique values for every column except for the first row. The data is about multidimensional poverty index of most of country in the world in brief. Following is the introduction for every column of the data:

Column1: ISO, International Organization for Standardization, is used as an abbreviation for the country.

Column2: Country, which is the full name of the country.

Column3: MPI Urban, which is the multidimensional poverty index. It is an international measure of acute poverty covering over 100 developing countries. It complements traditional income-based poverty measures by capturing the severe deprivation of education, health and living standards faced by everyone at the same time.

Column4: Headcount Ratio Urban (%), which is the rate of the people who lives below the poverty line in urban area of the country.

Colunn5: Intensity of Deprivation Urban (%), which is the rate of people’s dismissal in the condition of poor MPI in urban area of the country.

Column6: MPI Rural, which is the MPI in rural area of the country.

Column7: Headcount Ratio Rural, which is the rate of the people who lives below the poverty line in rural area of the country.

Column8: Intensity of Deprivation Rural, which is the rate of people’s dismissal in the condition of poor MPI in rural area of the country.

The dataset descripts every country’s average MPI and MPI which is separate in urban area and rural area in detail. However, it didn’t give the data which may be related to MPI. For example, if a country’s MPI is low, which features of the country will be low or high? On other hand, if one feature of the country is low and this country’s MPI is low, will other country also show the same result? As a result, this dataset is combined to another dataset to work out which feature is related to the MPI.