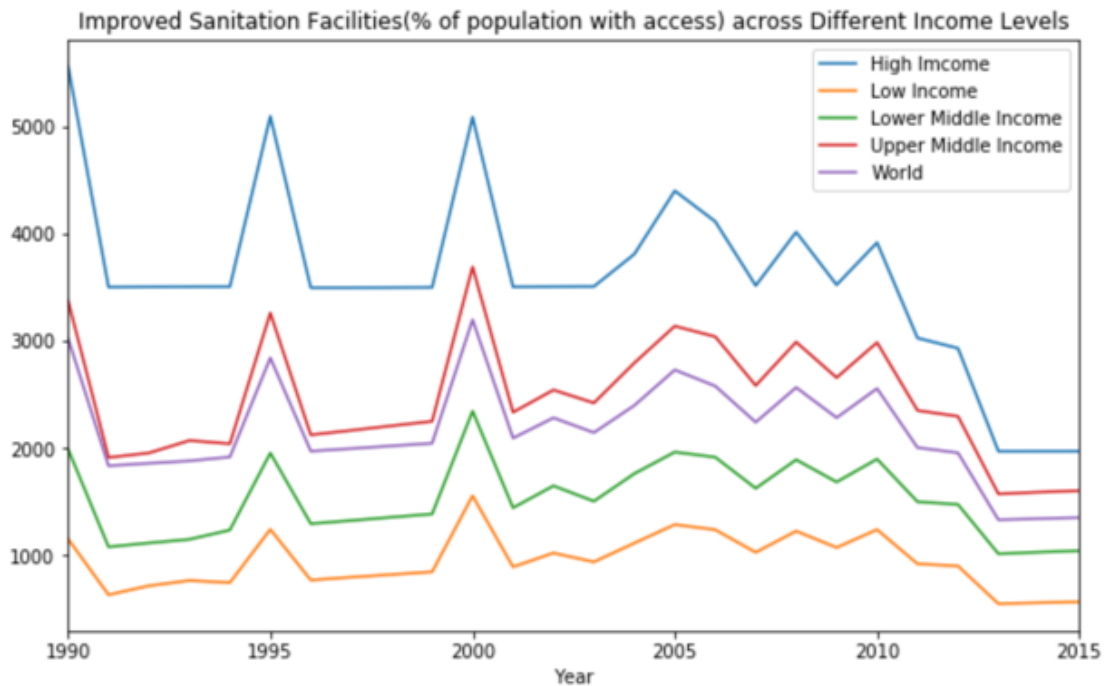


Part I Summary



I' ve applied a request module to retrieve data from WDI API. A caching layer is implemented to optimize the performance and accommodate any disruption in service. The data extracted from API is stored in a variable called `raw_data`. Data formatting converts the retrieved data into a dictionary where every country or group name is the key and the list of improved sanitation facilities index from 1960 to 2017 is the value. It' s stored in `country_index` variable. As countries have been aggregated into different income levels in the data set, I put together the values of world, high income, upper middle income, lower middle income and low income together in a pandas data frame with year as the index. The data from 1960 to 1989 is largely missing so I

performed data cleaning before next step. The cleaned data frame is stored in a variable called `income_df`.

The line chart above is the visualization of improved sanitation facilities index across different income levels. The index varies similarly across different income levels and world level. From 1991 to 2001, the index was stable with bumps in 1995 and 2000. The index trended upward from 2001 to 2010 with fluctuations. After 2010, the index has trended downward with a plunge in 2010 and 2012. In 2015, the index is set at 1969.70 for high income level, 1601.57 for upper middle income level, 1041.50 for lower middle income level, 564.75 for low income level and 1350.52 for the world level.