Statistical Analysis on Factors influencing Life Expectancy

Priyank Koul, Arif Sidiq Wani, Shevgoor Adithya Kamath, and Ashutosh Rout

PES University Bengaluru,priyankkoul14, wanixarif, kingdomofadithya, ashu24159@gmail.com

Abstract - **Although there have been lot of studies undertaken in the past on factors affecting life expectancy considering demographic variables, income composition and mortality rates. It was found that affect of immunization and human development index was not taken into account in the past. Also, some of the past research was done considering multiple linear regression based on data set of one year for all the countries. Hence, this gives motivation to resolve both the factors stated previously by formulating a regression model based on mixed effects model and multiple linear regression while considering data from a period of 2000 to 2015 for all the countries. Important immunization like Hepatitis B, Polio and Diphtheria will also be considered. In a nutshell, this study will focus on immunization factors, mortality factors, economic factors, social factors and other health related factors as well. Since the observations this dataset are based on different countries, it will be easier for a country to determine the predicting factor which is contributing to lower value of life expectancy. This will help in suggesting a country which area should be given importance in order to efficiently improve the life expectancy of its population.**

INTRODUCTION

The World Health Statistics series is WHO's annual compilation of health-related data for its 194 Member States, and includes a summary of the progress made towards achieving the health-related Millennium Development Goals (MDGs) and associated targets. The series is produced by the WHO Department of Health Statistics and Information Systems of the Health Systems and Innovation Cluster. As in previous years, World Health Statistics 2015 has been compiled using publications and databases produced and maintained by WHO technical programmes and regional offices. A number of demographic and socioeconomic statistics have also been derived from data produced and maintained by a range of national and international organizations. The latter include the United Nations Children's Fund (UNICEF), the United Nations Department of Economic and Social Affairs (UNDESA) and its Population Division, the United Nations Educational, Scientific and Cultural Organization (UNESCO), the United Nations International Telecommunication Union (ITU) and the World Bank. The indicators used in this report have been included on the basis of their relevance to global public health, on data availability and quality, and on the reliability and comparability of the resulting estimates. Taken together, these indicators provide a comprehensive summary of the current status of national health and health systems in the following nine areas:

■ life expectancy and mortality
■ cause-specific mortality and morbidity
■ selected infectious diseases
■ health service coverage
■ risk factors
■ health systems
■ health expenditure
■ health inequities
■ demographic and socioeconomic statistics.

For WHO, 2015 represents the target year for the Millennium Development Goals (MDGs).

By assessing the progress made based upon data available up to 2014, this report provides a clear indication of whether the regional and global targets set for 2015 are likely to be met. Since our dataset is published by WHO as well, we can look at what they have done in order to determine some statistics, and we can look forward for similar inferences, and some more of them. They have summarised the statistics for different countries, and have done a comparison amongst veraious aspects.

The report represents some best estimates of WHO for a broad range of key public health indicators based on evidence available at the beginning of 2015.

WHAT DOES IT RELATE TO? (PROBLEM, DATA OR TECHNIQUE PLANNED TO USE)

For WHO, 2015 represents the target year for the Millennium Development Goals (MDGs). It mostly relates to factors such as Life Expectancy, Mortality Rates. Most of the aspects and factors analysed in the dataset are similar to those of ours, considering the dataset was released by WHO too. So, whatever tables which are mentioned in the summary, we are planning to go with the same technique of analytics and infer things of some countries and then later comparing them.

The MDG on drinking-water called for the proportion of the population without sustainable access to safe drinking-water to be halved between 1990 and 2015, and was met globally in 2010.

Progress in child survival worldwide has been described as one of the greatest success stories of international development, with child deaths being almost halved over the last two decades compared to the 1990 MDG baseline.

Between 1990 and 2013 under-five mortality rates declined by 49%, falling from an estimated 90 deaths per 1000 live births to 46 deaths per 1000 live births.

GLOBAL, NATIONAL AND REGIONAL MORTALITY TRENDS (1990-2017)

From 1990 to 2016, the mortality of children younger than 5 years decreased by more than half, and there are plentiful data regarding mortality in this age group through which we can track global progress in reducing the under-5 mortality rate. By contrast, little is known on how the mortality risk among older children (5–9 years) and young adolescents (10–14 years) has changed in this time. We aimed to estimate levels and trends in mortality of children aged 5–14 years in 195 countries from 1990 to 2016. The reducing neonatal mortality is an essential part of the third Sustainable Development Goal (SDG), to end preventable child deaths. To achieve this goal the paper believes that it requires an understanding of the levels of and trends in neonatal mortality. Therefore it estimates the levels of and trends in neonatal mortality by use of a statistical model that can be used to assess progress in the SDG era. With these estimates of neonatal mortality between 1990 and 2017, the paper then aimed to assess how different targets for neonatal mortality could affect the burden of neonatal mortality from 2018 to 2030.

ASSUMPTIONS
- The ARR in each country from 2000 to 2017 would continue from 2018 to 2030. If the ARR from 2000 to 2017 was negative, the country's NMR is kept constant.
- Scenario 1 assumes a continuation of 2017 NMRs.
- Scenario 2 assumes a continuation of the 2000–17 annual rate of reduction in NMRs.
- Scenario 3 assumes an annual rate of reduction in NMR in each country that is equal to that of the country in the region that had the best annual rate of reduction in 2000–17.
- Scenario 4 assumes the necessary annual rate of reduction to achieve the Sustainable Development Goal target of 12 deaths per 1000 livebirths or fewer by 2030, across all countries.
- Scenario 5 assumes the necessary annual rate of reduction to achieve the 2017 average annual NMR

across high-income countries—three deaths per 1000 livebirths—across all countries.
(NMR=neonatal mortality rate)
- In most low-income and middle-income countries, vital statistics systems that record births and deaths on a continuous basis are not sufficiently complete to generate accurate mortality rates.
- In the absence of vital statistics data, estimates from the World Population Prospects (WPP) report and WHO are mainly derived from model life tables.
- Estimates of deaths by age in the WPP are not provided with uncertainty intervals and are available by 5-year periods. They calculated the annual numbers of deaths by use of Beers' six-term modified formula and aggregated them by GBD region.

MAIN CLAIMS

- The paper claims that the top five global causes of death in children aged 5–14 years in 2015 were lower respiratory tract infections, diarrhoeal diseases, drowning, meningitis, and road injuries.This finding indicates that substantial progress could still be achieved with public health interventions covering this age group.
- The paper also claims that this is the first study to provide estimates of mortality in children aged 5–14 years in all countries in the world that were made by use of a comprehensive database of measurements from nationally representative surveys, censuses, and vital registration.
- Further, since 2000, the pace of progress in reducing mortality among older children and young adolescents has been slower than that for children younger than 5 years, which calls for increased attention towards this age group.
- The risk of mortality of children aged 5–14 years has substantially decreased since 1990, despite this age range not being specifically targeted by health interventions.
- In 129 of the 195 countries analysed, the mortality in children aged 5–14 years was reduced by more than half (based on point estimates) between 1990 and 2016. 32 countries, including Bangladesh, Cambodia, Nepal, Vietnam, and Ethiopia, reduced the mortality rate of children aged 5–14 years by two-thirds between 1990 and 2016.

- Overall, 98% (98–99) of all deaths of children aged 5–14 years occurred in low-income and middle-income countries, which comprise 89·5% of the global population of children of this age group.

The paper claims although substantial progress has been made in reducing neonatal mortality since 1990, increased efforts to improve progress are still needed to achieve the SDG target by 2030. Accelerated improvements are most needed in the regions and countries with high NMR, particularly in sub-Saharan Africa and south Asia.

STATE OF INEQUALITY: REPRODUCTIVE, MATERNAL, NEWBORN AND CHILD HEALTH

This report, *State of inequality: reproductive, maternal, newborn and child health*, was developed to demonstrate best practices in reporting the results of health inequality monitoring, and to introduce innovative, interactive ways for audiences to explore inequality data. The report draws on data about reproductive, maternal, newborn and child health (RMNCH) in low- and middle-income countries – an important topic in global health – but the approach and underlying concepts can be applied to any health topic.

Until recently, development goals and agendas have lacked a systematic focus on the reduction of within-country inequality. Emphasis has tended to be placed on improving the overall national situation (that is, the national average), with too little attention devoted to narrowing the gaps that exist between subgroups of the population. The Millennium Development Goals (MDGs), adopted in the year 2000, called for improvements in national averages, but did not address how efforts to achieve the goals might affect within-country inequality. While a country may be on track to achieve national MDG targets, the situation with respect to some subgroups of the population may have stagnated or even worsened over time. In the absence of inequality monitoring, it remains unknown whether countries have narrowed or widened the gap between the advantaged and the disadvantaged. Acknowledging this deficit, several subsequent initiatives have emerged that promote the practice of monitoring the state of inequality alongside actions to attain the targets specified in the MDGs.

An understanding of the state of inequality reveals gaps in population health and lends insight into how policies, programmes and practices can be aligned to promote the ideal of health for all. This report presents selected examples of the state of inequality in low- and middle-income countries, highlighting important and relevant stories in RMNCH.

MAIN CLAIMS

- Health inequalities are observable differences in health between subgroups of a population. Subgroups can be defined by demographic, geographic or socioeconomic factors such as age, economic status, education, place of residence and sex.

- The paper claims that inequalities exist wherever there are differences in health indicators between subgroups. When health data are disaggregated – broken down by subgroups – they reveal differences between social groups that might have otherwise remained hidden behind the overall average.

MAIN TAKEAWAY

A brief understanding of how the analysis is performed.

Also, we would try our analytics to infer if certain goals and targets which were planned were quite achieved or not, by different visualisation techniques. This report is a breakthrough for going deeper into what actually our dataset is about.

## EDA AND VISUALIZATION

Correlation plot between variables
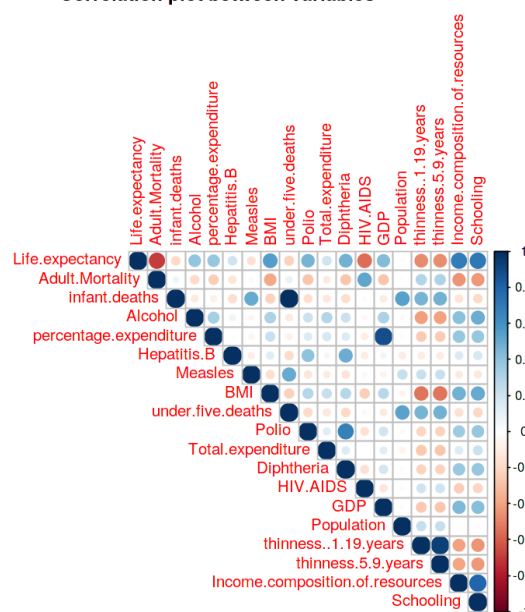
Here, after comparing the correlation plot between variables, we can see that **Infant Deaths are highly positively correlated with Under Five Deaths.** The causes for this correlation might be the fact that many instances of Infant Deaths come under Under Five Deaths. Same is the case with **GDP and Percentage Expenditure.** They are highly **positively correlated.** On the contrary, **Adult Mortality**

seems to be having a negative correlation with Life Expectancy. Adult.Mortality, thinness 1-19 years, thinness 5-9 years , HIV.AIDS are negatively correlated with Life expectancy.

Also, there seems to be multicollinearity present because there is correlation between independent variables.
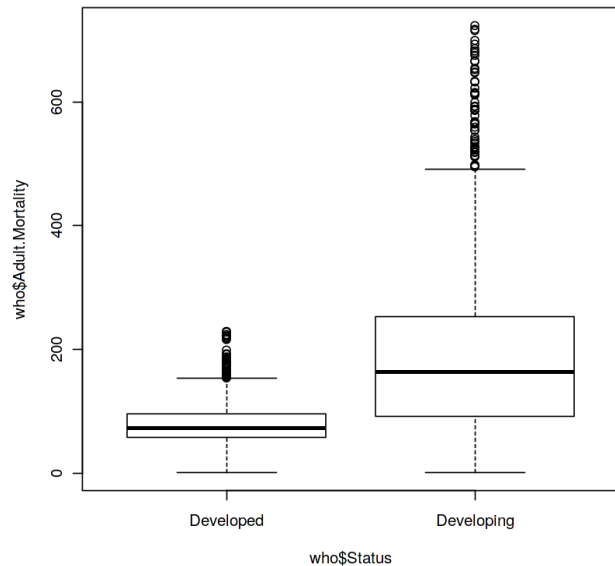
Comparing the box plot of Adult Mortality in Developing and Developed Countries, we can observe that the number is pretty high in Developing Countries. The reason might be because of lack of facilities and healthcare.
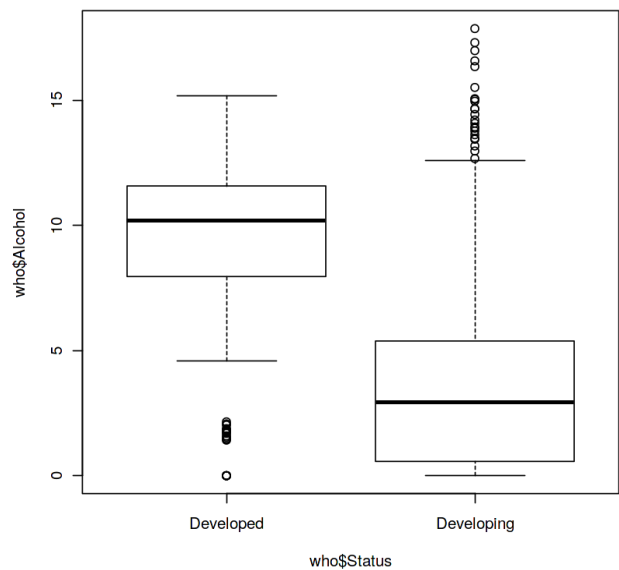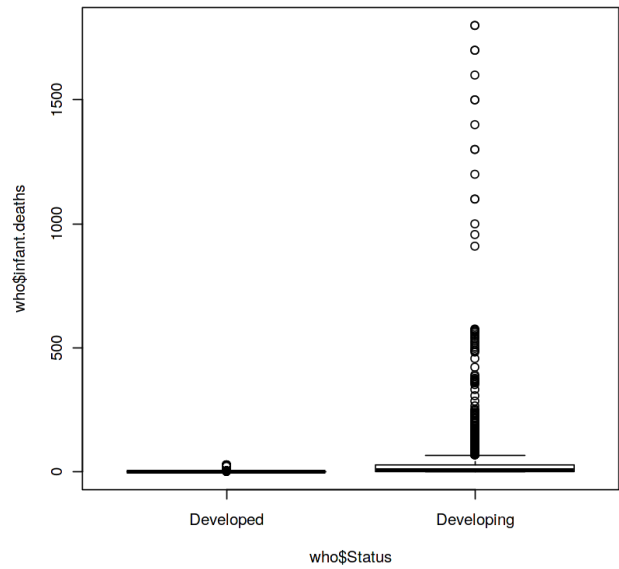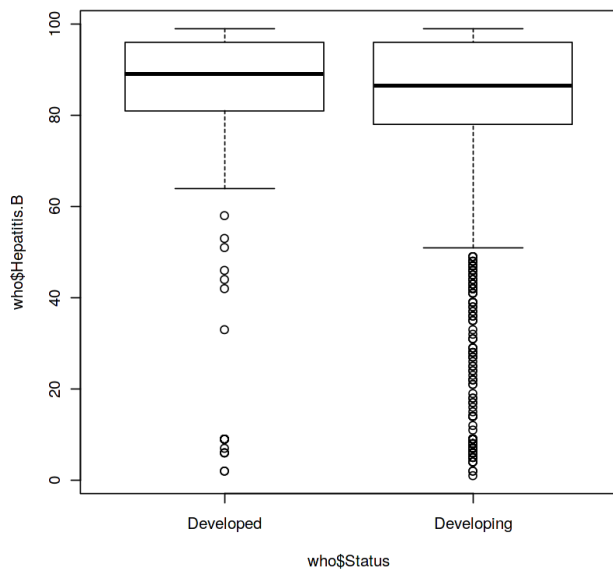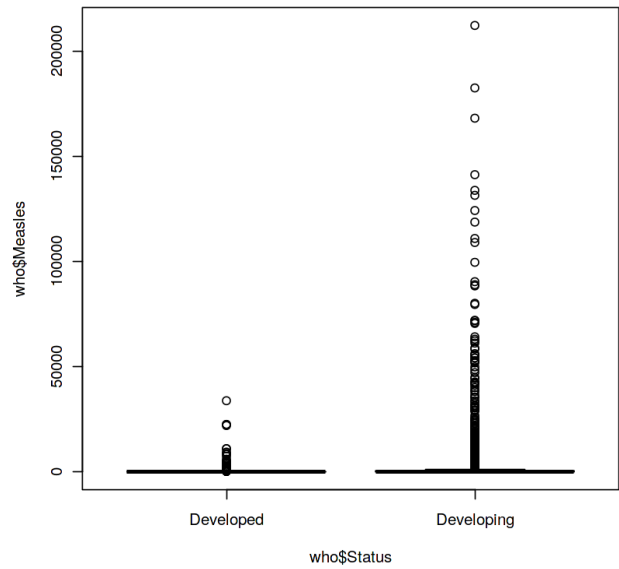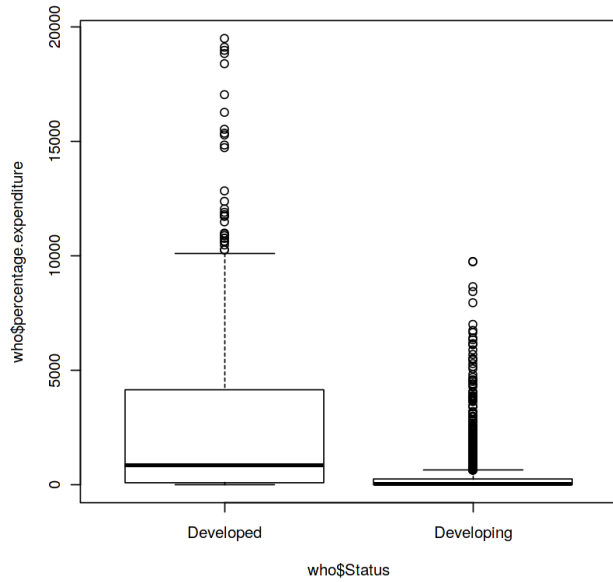
INFERENCES:

- The distributions of Total.expenditure, Income.composition.of.resources and Schooling are normally distributed.
- Alcohol ,infant.deaths, under.five.deaths, percentage.expenditure, Measles,Adult.Mortality, thinness 1-19 years, thinness 5-9 years, and HIV.AIDS are distributed with skewness at right.
- Hepatitis.B, Polio, Diphtheria are distributed with skewness at left.
- BMI has bimodal distribution.
- The distributions of Total.expenditure, Income.composition.of.resources and Schooling are normally distributed.
- Alcohol ,infant.deaths, under.five.deaths, percentage.expenditure, Measles,Adult.Mortality, thinness 1-19 years, thinness 5-9 years, and HIV.AIDS are distributed with skewness at right.
- Hepatitis.B, Polio, Diphtheria are distributed with skewness at left.
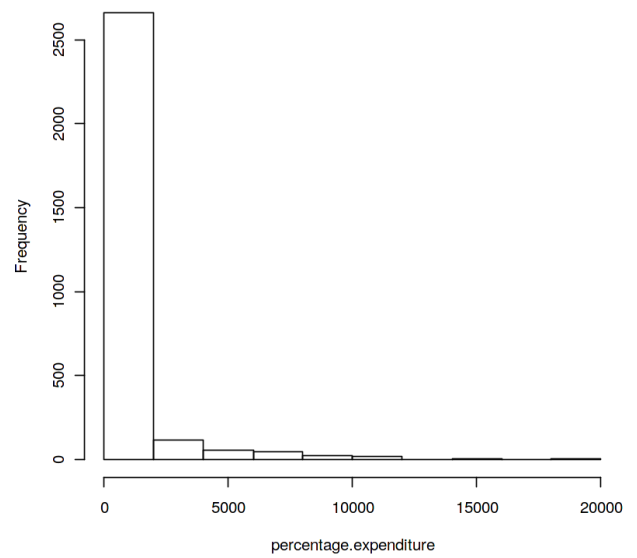- BMI has bimodal distribution.
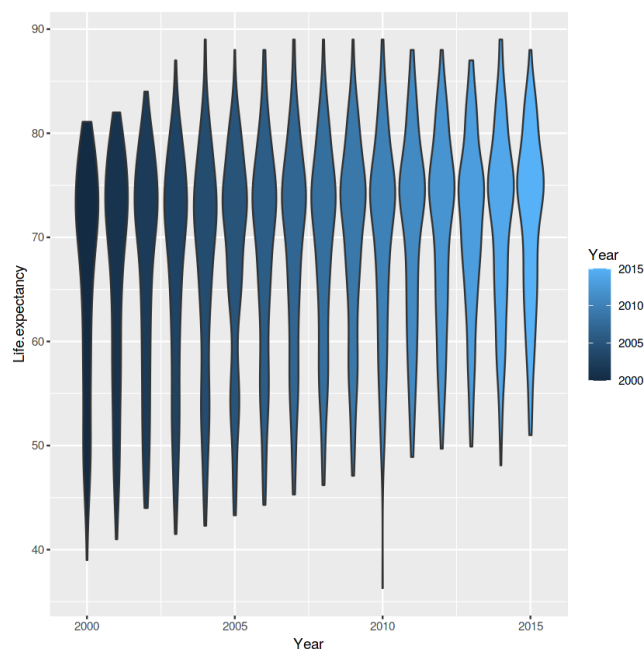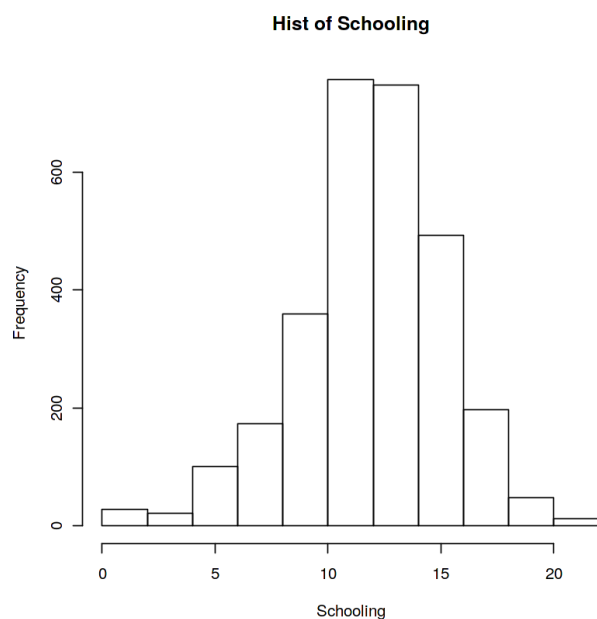
Hist of Adult Mortality
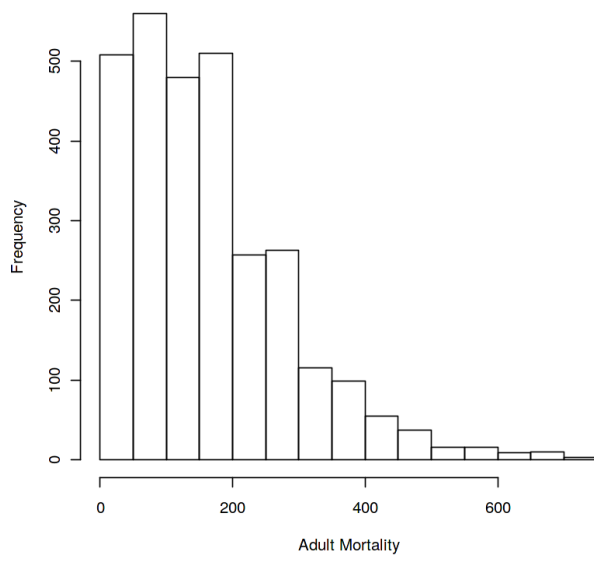
Hist of Alcohol
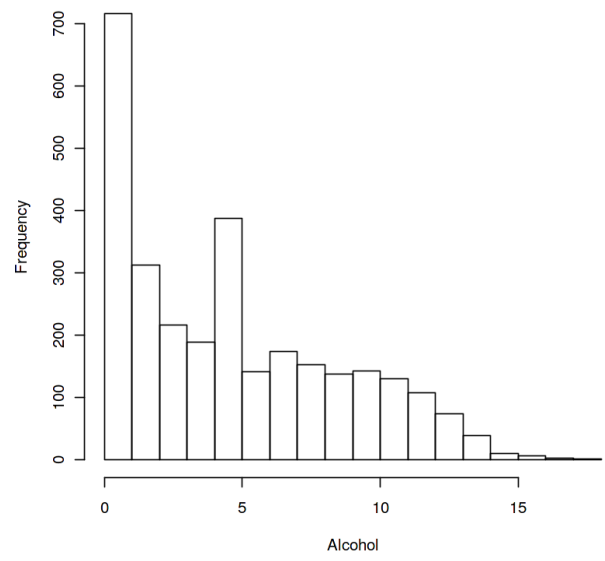
Hist of infant.deaths
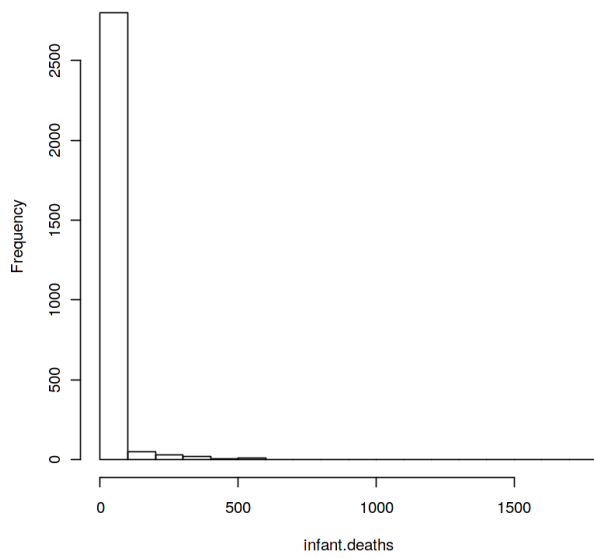
Hist of percentage.expenditure

Hist of Schooling

Here, we can observe that Life Expectancy distribution has gradually increased over the years.

In high-income countries, life expectancy at age 60 years has increased in recent decades. Falling tobacco use (for men only) and cardiovascular disease mortality (for both men and women) are the main factors contributing to this rise. In high-income countries, avoidable male mortality has fallen since 1980 because of decreases in avoidable cardiovascular deaths. For men in Latin America, the Caribbean, Europe, and central Asia, and for women in all regions, avoidable mortality has changed little or increased since 1980. As yet, no evidence exists that the rate of improvement in older age

mortality (60 years and older) is slowing down or that older age deaths are being compressed into a narrow age band as they approach a hypothesised upper limit to longevity.
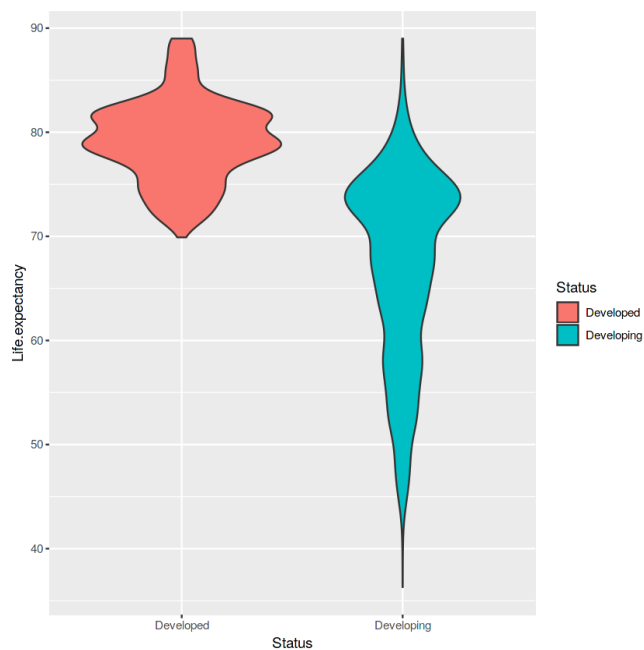
The life expectancy in developing countries is less than that of Developed countries.

This is because improvements in health and welfare increase **life expectancy**. The higher the **life expectancy**, the better shape a **country** is in. As you can see from the plot, more **developed** regions of the world generally have higher **life expectancies** than less **developed** regions with lower **life expectancies**.

Insights:

In developed countries adult mortality rate,prevelance of thinness of child between 1 to 19 years,infants deaths and deaths from HIV.AIDS are less comparing to developing countries which ultimately leads higher life expectancy in developed countries.

- On the other hand the other variables such as Income composition of resources ,number of years of schooling,percentage expenditure and Total expenditure on health are more in developed countries than developing countries which again increase the life expectancy in developed countries.
- Also by the increased numbers of immunization coverage against Hepatitis B, Polio and Diphtheria in developed countries results in high life expectancy in developed countries in contrast to developing countries.
- Life expectancy is high in developed countries compared to developing countries.