**Hypothesis: The strength of the rule of law is positively associated with the educational developments （school life expectancy）**

1. central tendency

By describing the average or most prevalent values of the variables, the central tendency of variables aids in the summary of the data points.

In the case of Rule of Law: estimates, the median is -0.18, the mode is -0.04, and the mean is 0, indicating a slightly skewed distribution.

The Duration of Compulsory Education has a mean of 9.92 years, a median of 10 years, and a mode of 9 years, indicating a more symmetrical distribution around the average.

Government Expenditure on Education shows an average of US$ 11,959.06 million, a median of US$ 1,306.43 million, and no mode available, indicating a highly skewed distribution.

The School Life Expectancy variable has a mean of 11.52 years, a median of 11.9 years, and a mode of 5.54 years, which indicates some degree of asymmetry.

The GDP per capita has a mean value of US$ 16,188.23, a median value of US$ 5,949.11, and no available mode, indicating a skewed distribution.

The expenditure on education as a percentage of total government expenditure has a mean of 15.08%, a median of 14.1%, and a mode of 3.28%, indicating variability in the data.

For the other 5 variables regarding government, we can see that all of them have a mean score of 0, indicating that on average, countries scored in the middle of the distribution for each indicator.

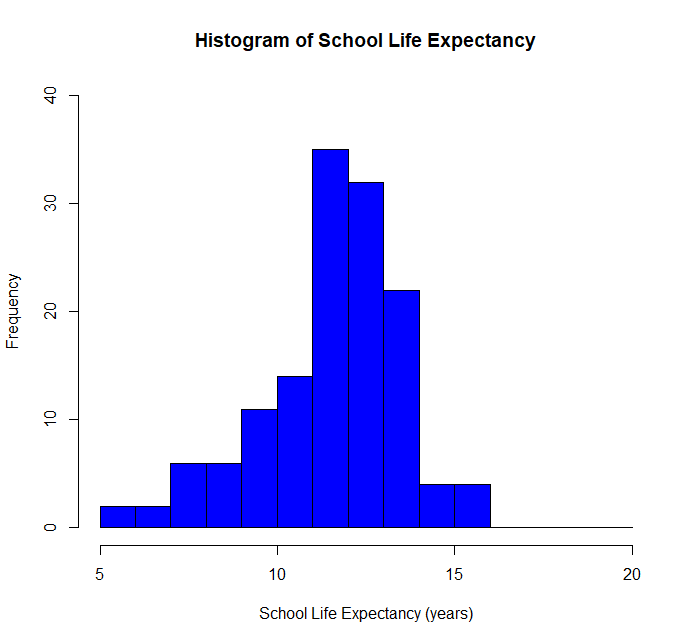
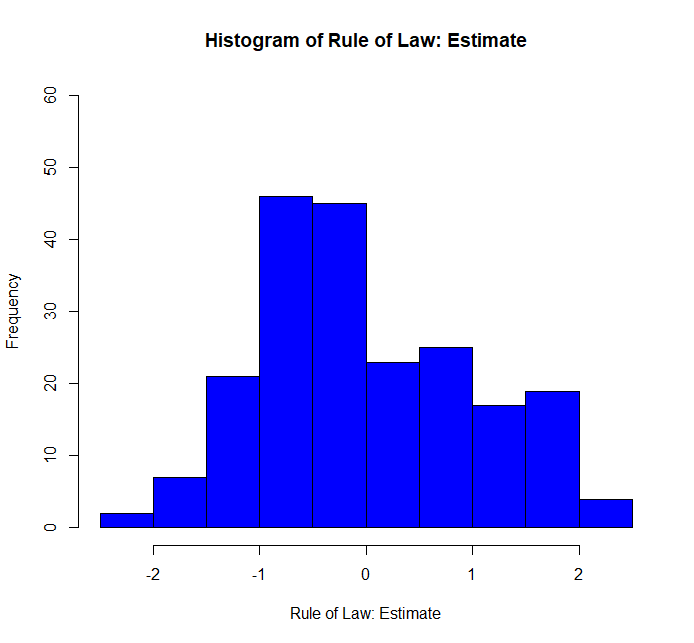
Regarding the variable "Control of Corruption" (Coc), the median score is -0.26, which is lower than the mean, suggesting that the distribution may be slightly negatively skewed. The mode is 1.22, which is much higher than both the mean and median, indicating that there may be a small number of countries with very high scores for control of corruption, skewing the distribution.

The variable "Government Effectiveness" (Ge) has a median score of -0.11, which is slightly lower than the mean. The mode is 1, which is the highest possible score for this variable, indicating that a significant number of countries have a high level of perceived government effectiveness.

For "Political Stability and Absence of Violence/Terrorism" (Psaaov), the median score is 0.07, which is higher than the mean. The mode is 0.92, which is relatively close to the median and indicates that there may be a cluster of countries with scores around this value.

The variable "Regulatory Quality" (Rq) has a median score of -0.08, which is lower than the mean. The mode is 0.91, which is similar to the mode for Psaaov and may suggest a similar distribution.

Finally, for "Voice and Accountability" (Vaa), the median score is 0.1, which is higher than the mean. The mode is 1.27, which is much higher than both the mean and median and indicates that there may be a small number of countries with very high scores for voice and accountability, skewing the distribution.



2. Dispersion

Each variable's dispersion offers important context for understanding the data by revealing information about the distribution and departure of data points from the central tendency.

The Rule of Law (rol) estimates have an interquartile range (IQR) of 1.52, a variance of 1, a standard deviation (SD) of 1, and a range of 4.39, indicating that there is moderate dispersion and variability in the rule of law across nations.

The Duration of Compulsory Education (dec) has a range of 11, an IQR of 3, a variance of 5.6, and an SD of 2.37, indicating that there is a smaller spread in the duration of compulsory education across nations.

Government Expenditure on Education (govex) displays a significant range of 170,217.5, an IQR of 9,302.68, a large variance of 791,588,736, and an SD of 28,135.19, highlighting considerable variability in education expenditure among countries, which is likely influenced by their respective economic capacities.

The Expenditure on Education as a percentage of total government expenditure (ex) has a range of 23.82, an IQR of 6.67, a variance of 23.34, and an SD of 4.83, indicating moderate dispersion and implying that education priorities differ among countries.

The School Life Expectancy (sl) variable exhibits a range of 10.42, an IQR of 2.16, a variance of 3.98, and an SD of 1.99, suggesting a relatively narrow spread and a more consistent school life expectancy across countries.

The GDP per capita (gdpper) presents a substantial range of 166,997.5, an IQR of 16,154.99, a large variance of 598,242,849, and an SD of 24,459, reflecting significant economic disparities among countries, with both developed and developing economies included in the dataset.

Regarding the 5 other variables of government, we can see that they all have the same variance and standard deviation of 1, indicating that they have the same level of spread around the mean. However, they differ in their range and interquartile range.

Control of Corruption (Coc) has the smallest range of 3.98, indicating that the values are more tightly clustered around the mean. Its interquartile range of 1.48 is also the same as the other four variables, indicating that the middle 50% of the values are spread out similarly.

Government Effectiveness (Ge) has the second smallest range of 4.42, indicating a similar level of clustering around the mean as Coc. Its interquartile range is also the same as the other four variables.

Political Stability and Absence of Violence/Terrorism (Psaaov) has the largest range of 4.92, indicating a wider spread of values around the mean. Its interquartile range of 1.49 is similar to the other four variables.

Regulatory Quality (Rq) has a range of 4.58, which is between the ranges of Ge and Psaaov. Its interquartile range of 1.55 is the largest among the five variables, indicating that the middle 50% of the values are spread out more widely.

Voice and Accountability (Vaa) has the smallest range of 3.93, indicating a tighter clustering of values around the mean than the other four variables. Its interquartile range of 1.72 is the largest among the five variables, indicating that the middle 50% of the values are spread out more widely than the other four variables.

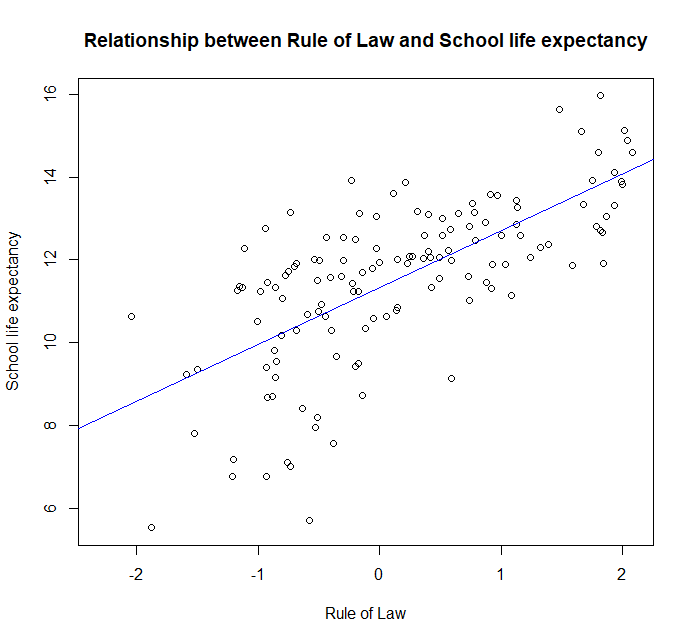
3. Association

The Pearson product-moment correlation coefficient is used to calculate the degree of association between the primary dependent variable, School Life Expectancy (sl), and the primary independent variable, Rule of Law (rol). With a correlation coefficient of 0.68, the analysis demonstrates a significant positive correlation between the two variables. This suggests that longer life expectancies for students are linked to higher rule of law estimates. Indicating that the association between the two variables is statistically significant and unlikely to have happened by chance, the p-value of 7.08832e-20 is much lower than the threshold of 0.05. The true correlation coefficient is likely to fall within this range, according to the 95 percent confidence interval of the correlation coefficient, which ranges from 0.583 to 0.766.

4. Bivariate hypothesis testing

The main dependent variable, School Life Expectancy (sl), and the main independent variable, Rule of Law, are examined using the bivariate hypothesis test. (rol). The analysis uses a linear regression model to determine if there is a significant association between the two variables. The model shows that there is a strong positive linear relationship between the variables, with a coefficient estimate of 1.3724 and a p-value of <2.2e-16, which is much smaller than the significance level of 0.05, indicating a highly significant relationship between the two variables.

The intercept of the model is 11.3296, which represents the predicted value of sl when rol is equal to 0. The R-squared value of 0.4691 indicates that approximately 47% of the variance in sl can be explained by the variation in rol. The F-statistic of 116.7 on 1 and 132 degrees of freedom is also highly significant, further supporting the association between the two variables.



5. Multivariate hypothesis testing

Model 1:

In this model, the dependent variable is school life expectancy, and the independent variable is the rule of law. The coefficient estimate of the rule of law is 1.372, and it is statistically significant at the 99% level (p<0.01). This suggests that higher levels of the rule of law are associated with higher levels of school life expectancy. The R-squared value of the model is 0.469, indicating that approximately 47% of the variation in school life expectancy can be explained by the rule of law.

Model 2:

In this model, the dependent variable remains school life expectancy, and there are four independent variables: government expenditure on education (in US millions), expenditure on education as a percentage of total government expenditure, GDP per capita (in current US dollars), and the duration of compulsory education (in years). Only the rule of law is statistically significant with the coefficient 1.128, and it is statistically significant at the 99% level (p<0.01).The R-squared value of the model is 0.544, which is higher than in Model 1, and the adjusted R-squared value is only 0.512.

Model 3:

This model includes the same independent variables as Model 2, but it also includes five governance indicators: control of corruption, government effectiveness, political stability and absence of violence/terrorism, regulatory quality, and voice and accountability. Of these variables, only government effectiveness and Political Stability and Absence of Violence/Terrorism are statistically significant at the 99% level, with a coefficient estimate of 1.401 and 0.578. This suggests that higher levels of government effectiveness and Political Stability and Absence of Violence/Terrorism are associated with higher levels of school life expectancy. The R-squared value of the model is 0.564, indicating that it explains more of the variation in school life expectancy than Model 2. The adjusted R-squared value is 0.543, indicating that the model could still be improved. Notable that the main independent variable the rule of law is not statistically significant with a coefficient -0.384 which is against the hypothesis.

Model 4:

In this final model, the dependent variable remains school life expectancy, and the independent variables includes Model 2 and Model 3. In this case, three of the governance indicators are statistically significant: government effectiveness (at the 1% level),Voice and Accountability(at the 5% level) and regulatory quality (at the 5% level). The coefficients for the other governance indicators are not statistically significant. The R-squared value of the model is 0.664, indicating that it explains a substantial portion of the variation in school life expectancy.

