

Analytics for Applied Economics and Business

Investigating Life Satisfaction vs. Freedom of Expression using Econometric Methods in R

Word Count Excl. Title, Headings, References, Appendix, Figures: 2,720



20/11/2023

IMPERIAL COLLEGE LONDON
IMPERIAL COLLEGE BUSINESS SCHOOL

Table of Contents:

I.	Introduction & Motivation.....	2
II.	Data and Data Limitations.....	2
III.	Analysis Specifications.....	3
IV.	Empirical Analysis.....	5
V.	Graphical Analysis.....	11
VI.	Conclusion.....	14
VII.	References.....	15
VIII.	Appendix.....	16

I. Introduction and Motivation

This report aims to regress freedom of self-expression on life satisfaction, and in addition understand the role that other less-explored socio-economic factors may or may not have on life satisfaction and freedom of expression.

Our investigation focuses on socio-economic factors which are parallel to freedom of expression, and we find these factors, notably perceptions of social support mechanisms, significantly impact life satisfaction. We then extend our study to global sub-regions to investigate the heterogeneity of our model across different geographic locations and cultural groupings.

In today's society, freedom of expression underpins political dissent, cultural diversity, creative innovation and improves self-development. It helps create a dialogue and an understanding between different communities. It could be argued that the absence of this freedom carries a negative impact on democracy, society and mental health, and this is what we aim to investigate: does increased freedom of expression have a significant positive impact on life satisfaction?

II. Data

Our dataset comprises 156 countries categorized into 11 sub-regions, spanning the years 2011 to 2022. Self-reported life satisfaction is measured utilizing the Cantril Ladder score, prompting respondents to envision their satisfaction on a scale from 0 to 10.

Freedom of Expression is evaluated using the "central estimate" from the V-Dem dataset, considering political discussions, press and media freedom, and academic and cultural expression, with scores ranging from 0 to 1 (1 denoting the highest freedom).

Social support is determined as the national average of binary responses regarding having reliable relatives or friends during challenging times. Corruption perception is derived from survey responses gauging the widespread nature of corruption in government and businesses.

The Gender Inequality Index (GII) reflects gender-based disadvantage in reproductive health, empowerment, and the labour market, ranging from 0 (gender equality) to 1 (maximum gender-based disadvantage). After merging the data, a dummy matrix was created for the "Income Group" variable, categorizing it into high income, upper middle, lower middle, and lower, all NA values were removed.

II.a. Data Limitations

Recognizing the potential for incompleteness and omitted variable bias, note that our study interprets available information while being mindful of the inherent limitations present in the data. Inconsistencies in country representation across years and potential bias in certain regions underscore data constraints. Variables, drawn from respondents' subjective experiences, introduce limitations. While self-reporting is a valuable measure, it's crucial to acknowledge associated biases. Oswald (2007) makes the case that self-reporting is often the only measure available to quantify certain concepts, while Helliwel et al. (2008) find that individuals across the world tend to value the same basket of variables similarly, furthering the case for the legitimate use of self-reporting data. Nevertheless, the nature of subjective data introduces the possibility of bias, impacting the precision of our findings.

Despite these considerations, potential bias remains, impacting result precision. We don't foresee significant reverse causality issues, as it's implausible for higher life satisfaction to cause improved explanatory variables, emphasizing the intuitive influence of chosen factors on satisfaction.

III. Analysis Specifications

Regression 1: Primary Simple Linear Regression Equation

$$\text{Life Satisfaction} = \beta_0 + \beta_1 \text{Freedom of Expression} + \epsilon$$

Our primary regression equation would be to study the effect of freedom of expression, our key explanatory variable, on life satisfaction.

Regression 2: Regression Equations with each Control Variable

$$\text{Life Satisfaction} = \beta_0 + \beta_1 \text{Freedom of Expression} + \beta_2 \text{Social Support} + \epsilon$$

$$\text{Life Satisfaction} = \beta_0 + \beta_1 \text{Freedom of Expression} + \beta_2 \text{GII} + \epsilon$$

$$\text{Life Satisfaction} = \beta_0 + \beta_1 \text{Freedom of Expression} + \beta_2 \text{Perceptions of Corruption} + \epsilon$$

We assessed how individual control variables acted in simple regressions, examining their connections with freedom of expression and their roles in shaping satisfaction.

Regression 3: Regression Equation Accounting for Control Variables

$$\text{Life Satisfaction} = \beta_0 + \beta_1 \text{Freedom of Expression} + \beta_2 \text{Social Support} + \beta_3 \text{GII} + \beta_4 \text{Perceptions of Corruption} + \beta_5 \text{Income Dummies} + \beta_6 \text{Year} + \epsilon$$

Our third regression incorporates control variables beyond freedom of expression, including Social Support, recognizing its impact on life satisfaction and self-expression through interpersonal interactions.

The Gender Inequality Index (GII) accounts for gender dynamics, ensuring diverse perspectives, especially the female viewpoint, influence life satisfaction.

Perceptions of Corruption assess the societal level of corruption, shedding light on conditions that either facilitate or hinder diverse viewpoints, impacting life satisfaction.

Income levels categorize countries, providing insights into their effect on satisfaction. The consideration of years adds a temporal dimension. This aims to understand factors contributing to overall well-being, encompassing interpersonal relationships, gender dynamics, corruption perceptions and economic disparities.

Regression 4: Regression Adding Sub-Region

$$\begin{aligned} \text{Life Satisfaction} = & \beta_0 + \beta_1 \text{Freedom of Expression} + \beta_2 \text{Social Support} + \beta_3 \text{GII} + \\ & \beta_4 \text{Perceptions of Corruption} + \beta_5 \text{Income Dummies} + \beta_6 \text{Year} + \\ & \beta_7 \text{Sub-Region} + \epsilon \end{aligned}$$

Acknowledging the inherent implications that a country and its culture may have on all the variables; we have further accounted for sub-region.

We have laid out all of the divisions and criteria along which we will base our investigation and discussions.

We now move to an analysis of our findings using the aforementioned specifications.

IV. Empirical Analysis

Given our 10-year data span, we conducted an examination to verify that there are no time-related trends. The variables are stationary; an Augmented Dickey Fuller Test was performed on each variable, revealing p-values below 0.05 for all. (Additional test results are available in the appendix for reference.)

Augmented Dickey-Fuller Test

```
data: df$freedom.of.expression
Dickey-Fuller = -8.6251, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary
```

Freedom of Expression on Life Satisfaction: SLR1

```
Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression)

Residuals:
    Min       1Q   Median       3Q      Max
-3.2983 -0.8155  0.0863  0.8580  2.4621

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    3.9458     0.1078   36.59  <2e-16 ***
df$freedom.of.expression 2.0778     0.1374   15.13  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.058 on 1272 degrees of freedom
Multiple R-squared:  0.1525,    Adjusted R-squared:  0.1518
F-statistic: 228.8 on 1 and 1272 DF,  p-value: < 2.2e-16
```

In this simple model we regress freedom of expression on life satisfaction. The intercept $\hat{\beta}_0$ at 3.9458 tells us that when there is “0” freedom of expression, life satisfaction for an individual would be 3.9458 out of 10. This could be due to factors other than freedom of expression affecting life satisfaction such as social support, perceptions of corruption etc.

The coefficient estimates for freedom of expression $\hat{\beta}_1$ at a 2.0778 tells us that with each additional 1 unit increase on the freedom of expression index, there would be an equivalent 2.0778 units of increase in life satisfaction. This seems like a strong positive relationship implied by our model, and we will investigate the strength of this relationship by adding other explanatory variables. We can see that the p-values of both the intercept and $\hat{\beta}_1$ are below the 0.05 threshold and hence our results are statistically significant.

Equation of Freedom of Expression and GII on Satisfaction: SLR2

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.95732 -0.53179  0.06706  0.54978  2.32803

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.19329    0.09936  62.330 < 2e-16 ***
df$freedom.of.expression  0.85741    0.10334   8.297 2.7e-16 ***
df$gii_value   -3.90990    0.11030 -35.449 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7506 on 1271 degrees of freedom
Multiple R-squared:  0.5738,    Adjusted R-squared:  0.5732
F-statistic: 855.7 on 2 and 1271 DF,  p-value: < 2.2e-16

```

By accounting for freedom of expression and controlling for gender inequalities, we can see that gender inequality has a much greater impact on life satisfaction in comparison to freedom of expression. Adding gender inequality to our investigation shifts our intercept $\hat{\beta}_0$ to 6.19329, effectively setting a higher base life satisfaction as reference against which to subtract the negative impact of gender inequality; this will apply to all countries to some degree as no country has a perfect GII. Interestingly, we can see that gender inequality is a much more important component of life satisfaction in this context, and that the effects of freedom of expression “diminish” as the base level of life satisfaction increases.

Equation of Freedom of Expression and Social Support on Satisfaction: SLR3

```

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$Social.support)

Residuals:
    Min       1Q   Median       3Q      Max
-2.37032 -0.45359  0.06603  0.57202  2.32001

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   -0.5771    0.1457   -3.96 7.9e-05 ***
df$freedom.of.expression  1.0639    0.1002  10.61 < 2e-16 ***
df$Social.support    6.5056    0.1792  36.31 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7417 on 1271 degrees of freedom
Multiple R-squared:  0.5839,    Adjusted R-squared:  0.5833
F-statistic: 891.9 on 2 and 1271 DF,  p-value: < 2.2e-16

```

In this case the intercept is nearly 0, demonstrating a negative effect of -0.57, attributed to social support, which depresses the intercept. This suggests that in the absence of social support or freedom of expression, life satisfaction could be negative. Notably, the estimated impact of Social Support is substantial at 6.5056, signifying that a 1-unit increase in Social Support corresponds to a 6.5056-unit increase in Life Satisfaction. It appears that Social Support significantly contributes to Life Satisfaction, consequently decreasing the estimated impact of freedom of expression from 2.0778 to 1.0639.

Equation of Freedom of Expression and Perception of Corruption on Satisfaction: SLR4

```

Residuals:
    Min       1Q   Median       3Q      Max
-2.8726 -0.5948  0.1595  0.6418  2.7957

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      6.1583    0.1518  40.56  <2e-16 ***
df$freedom.of.expression  1.7921    0.1226  14.62  <2e-16 ***
df$Perceptions.of.corruption -2.6817    0.1431 -18.74  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9369 on 1271 degrees of freedom
Multiple R-squared:  0.336,    Adjusted R-squared:  0.335
F-statistic: 321.6 on 2 and 1271 DF,  p-value: < 2.2e-16

```

The coefficient estimate for perceptions of corruption reveals a substantial -2.6817 effect on life satisfaction, affirming that the higher the perception of corruption citizens have in a country the lesser their life satisfaction. Perceptions of corruption also have a bigger impact than that of freedom of expression on life satisfaction.

The relationship between democracy and subjective well-being has been well documented in civic realms, although there are opposing views. According to Altman et al. (2017) and Bavetta et al. (2017), a higher level of political autonomy and freedom is associated with an increase in people's life satisfaction. However, some scholars have found no or even a negative link between life satisfaction and democracy (Sujarwoto and Tampubolon, 2015). We will discuss autonomy again later.

Freedom of Expression on Life Satisfaction Controlled for GII, Social Support, Perception of Corruption, Income groups and Years: MLR1

```

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$gii_value +
    df$Social.support + df$Perceptions.of.corruption + df$LowerMiddle +
    df$Lower + df$UpperMiddle + df$Higher + factor(df$year),
    data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.13483 -0.35059  0.01447  0.38169  2.16848

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.00140    0.22007  18.182  < 2e-16 ***
df$freedom.of.expression  0.62641    0.09062   6.913  7.56e-12 ***
df$gii_value      -0.74537    0.17009  -4.382  1.27e-05 ***
df$Social.support   3.47979    0.19939  17.452  < 2e-16 ***
df$Perceptions.of.corruption -1.48117    0.10248 -14.453  < 2e-16 ***
df$LowerMiddle     -0.49784    0.07910  -6.294  4.27e-10 ***
df$Lower           -0.94021    0.09864  -9.532  < 2e-16 ***
df$UpperMiddle     -0.26625    0.06317  -4.215  2.68e-05 ***
df$Higher          NA           NA       NA       NA
factor(df$year)2012  -0.06094    0.07872  -0.774  0.4390
factor(df$year)2013  -0.10766    0.07825  -1.376  0.1691
factor(df$year)2014  -0.17237    0.07746  -2.225  0.0262 *
factor(df$year)2015  -0.13204    0.07755  -1.703  0.0889 .
factor(df$year)2016  -0.17998    0.07784  -2.312  0.0209 *
factor(df$year)2017  -0.09136    0.07703  -1.186  0.2359
factor(df$year)2018  -0.07114    0.07779  -0.914  0.3606
factor(df$year)2019  -0.08104    0.07790  -1.040  0.2984
factor(df$year)2020  -0.05188    0.08257  -0.628  0.5299
factor(df$year)2021  -0.05315    0.08232  -0.646  0.5186
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5936 on 1256 degrees of freedom
Multiple R-squared:  0.7366,    Adjusted R-squared:  0.733
F-statistic: 206.6 on 17 and 1256 DF,  p-value: < 2.2e-16

```

We include dummy variables for income levels and time as factor variables. This equation estimates the impact of freedom of expression when controlling for key variables on the life satisfaction of an average individual. Our results show the explanatory variables are statistically significant with small standard errors, all p-values below 0.05, F-test $p < 2.2e-16$ and adjusted R squared of 0.733, demonstrating a good explanation of the variance in Y, also implying the causality of our explanatory variables. The intercept represents a base life satisfaction of 4.0014 (out of 10) in a

high-income country in 2011 (indicated by the NA entries “higher-income” dummy variable and omission of 2011 as a factor variable). The results are consistent with our inferences; there exists a positive relationship between freedom of expression and life satisfaction, but this relationship is not as strong as suggested by **SLR1** when controlling for other factors, increasing life satisfaction by only 0.6261 units instead of 2.0778 units for every additional unit improvement in the index for freedom of expression.

Social support is the most impactful factor influencing life satisfaction, increasing it by 3.4797 units for every additional unit increase in the proportion of individuals. Perceptions of corruption (-1.48117) and gender inequality (-0.74537) have greater impacts than freedom of expression on life satisfaction, indicating that a ‘fair’ society may matter more to wellbeing. Income groupings follow intuitively, with lower income categories exhibiting correspondingly larger negative impacts on life satisfaction. Years exhibit some statistical significance but still have notably low t and high p values relative to key explanatory variables, indicating that there is no significant bearing of years on life satisfaction.

Our set of results could be interpreted as a ranking of socio-economic priorities, in which more rudimentary prerequisites such as social support, just governance, and gender equality take priority over freedom of speech for having a satisfying life, indicating that freedom of expression may be a ‘luxury’ concern.

Adding Sub-Regions to the equation: MLR2

```
Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$gii_value +
    df$Social.support + df$Perceptions.of.corruption + df$LowerMiddle +
    df$Lower + df$UpperMiddle + df$Higher + factor(df$year) +
    factor(df$SubRegion.Name), data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-1.98285 -0.31716  0.02583  0.32199  2.08800

Coefficients: (1 not defined because of singularities)
(Intercept)                    5.14022    0.25058   20.513   < 2e-16   ***
df$freedom.of.expression       0.42288    0.09390    4.503   7.31e-06   ***
df$gii_value                   -1.55491    0.23024   -6.753   2.21e-11   ***
df$Social.support               2.40130    0.20364   11.792   < 2e-16   ***
df$Perceptions.of.corruption   -0.89870    0.12545   -7.164   1.34e-12   ***
df$LowerMiddle                 -0.49225    0.07485   -6.576   7.09e-11   ***
df$Lower                       -0.75862    0.09376   -8.091   1.39e-15   ***
df$UpperMiddle                 -0.38018    0.05889   -6.456   1.54e-10   ***
df$Higher                      NA         NA         NA         NA
factor(df$year)2012            -0.05879    0.07032   -0.836   0.403299   .
factor(df$year)2013            -0.11541    0.06990   -1.651   0.098960   .
factor(df$year)2014            -0.14646    0.06922   -2.116   0.034565   *
factor(df$year)2015            -0.12052    0.06930   -1.739   0.082256   .
factor(df$year)2016            -0.16154    0.06965   -2.319   0.020531   *
factor(df$year)2017            -0.08268    0.06903   -1.198   0.231197   .
factor(df$year)2018            -0.03805    0.06976   -0.545   0.585535   .
factor(df$year)2019            -0.05089    0.07009   -0.726   0.467956   .
factor(df$year)2020            -0.02261    0.07418   -0.305   0.760552   .
factor(df$year)2021            -0.05283    0.07398   -0.714   0.475272   .
factor(df$SubRegion.Name)Central Asia    -0.09485    0.16680   -0.569   0.569684   .
factor(df$SubRegion.Name)Eastern Asia    -0.65389    0.15719   -4.160   3.40e-05   ***
factor(df$SubRegion.Name)Eastern Europe  -0.47369    0.14466   -3.274   0.001088   **
factor(df$SubRegion.Name)Latin America and the Caribbean  0.29368    0.13714    2.142   0.032424   *
factor(df$SubRegion.Name)Northern Africa -0.58522    0.18502   -3.163   0.001599   **
factor(df$SubRegion.Name)Northern America 0.20455    0.16216    1.261   0.207383   .
factor(df$SubRegion.Name)Northern Europe -0.24760    0.12613   -1.963   0.049869   *
factor(df$SubRegion.Name)South-eastern Asia -0.16413    0.14860   -1.104   0.269593   .
factor(df$SubRegion.Name)Southern Asia    -0.33109    0.15606   -2.122   0.034076   *
factor(df$SubRegion.Name)Southern Europe  -0.62491    0.14409   -4.337   1.56e-05   ***
factor(df$SubRegion.Name)Sub-Saharan Africa -0.51526    0.14688   -3.508   0.000467   ***
factor(df$SubRegion.Name)Western Asia    -0.47677    0.14555   -3.276   0.001083   **
factor(df$SubRegion.Name)Western Europe  -0.04525    0.13156   -0.344   0.730959   .

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5299 on 1243 degrees of freedom
Multiple R-squared:  0.7923,    Adjusted R-squared:  0.7872
F-statistic: 158 on 30 and 1243 DF, p-value: < 2.2e-16
```


Controlling for sub-regions, we see differences compared to our previous results. Our intercept is closer to the world mean life satisfaction of 5.51, representing a base life satisfaction of 5.14022 in a high-income country (in Oceania in 2011). Our new model's adjusted R squared increased to 0.7872, indicating it more accurately captures variance. Social support remains the biggest influence on life satisfaction but with a decreased coefficient of 2.40130, while the coefficient of gender inequality doubled to -1.55491 and now has a greater impact on life satisfaction than perceptions of corruption, whose coefficient has nearly halved. This may be because of larger impacts of gender inequality on life satisfaction in some sub-regions and tolerance for corruption. Year and time effects remain largely insignificant. Some sub-regions such as Eastern Asia exhibit statistical significance, but these effects are weak for most regions with relatively low t-values compared to key variables and thus may not have a significant impact.

The impact of freedom of expression is further dampened, down to 0.42288. A possible explanation is that there are sub-regions in which life satisfaction is high in countries which exhibit low freedom of expression, mostly in Asia. These countries often exhibit autocracy; Saudi Arabia and China, where we can posit that increases in economic wealth and basic societal functioning are much more highly valued relative to freedom of expression. This points to the “luxury” aspect of freedom of expression; Steckermeier (2020) finds a strong positive relationship between perceived autonomy and life satisfaction in Europe, but more interestingly finds that less autonomy increases the impact proportion of ‘*basic functionings*’ on life satisfaction.

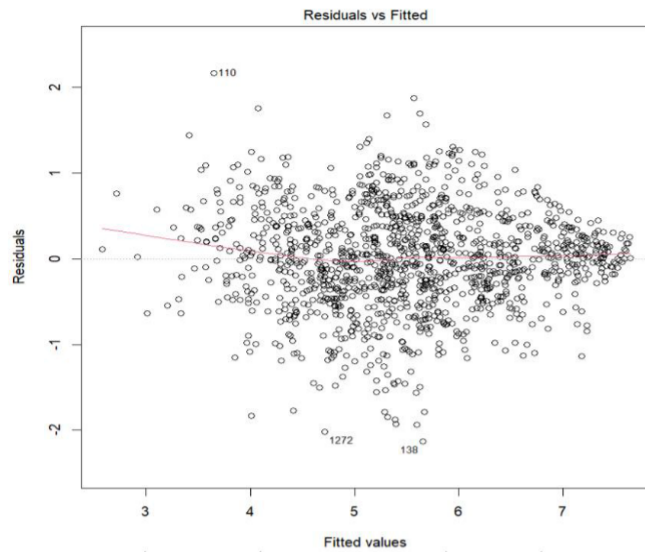
There are sub regions in which there is moderately high freedom of expression but low life satisfaction, mostly in Africa. Where this is due to low-income and absence of support, it implies that poverty, lack of opportunity and lack of support having an outsized impact on life satisfaction. Overall, this additional model upholds our findings that while there exists a positive relationship between freedom of expression and life satisfaction, it is a weak relationship which is superseded by other factors which have greater influence on life satisfaction.

Tests for Unbiasedness:

VIF – Variance Inflation Factor

The model undergoes scrutiny for multicollinearity among predictor variables through a Variance Inflation Factor (VIF) test, typically compared against a threshold of 10. Examination of the table below reveals that all Generalized VIF (GVIF) values for our core variables are below 10 (Vittinghoff E, 2012), satisfying the established threshold and signifying an absence of collinearity among the variables.

	GVIF	Df	GVIF ^{1/(2*Df)}
df\$freedom.of.expression	1.863205	1	1.364993
df\$gii_value	9.833721	1	3.135876
df\$Social.support	2.742579	1	1.656073
df\$Perceptions.of.corruption	2.440962	1	1.562358
factor(df\$income_group)	8.431617	3	1.426653
factor(df\$year)	1.112757	10	1.005356
factor(df\$SubRegion.Name)	29.596596	13	1.139164

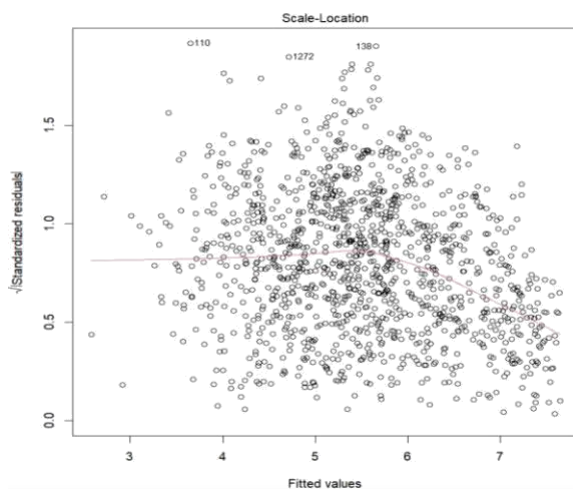
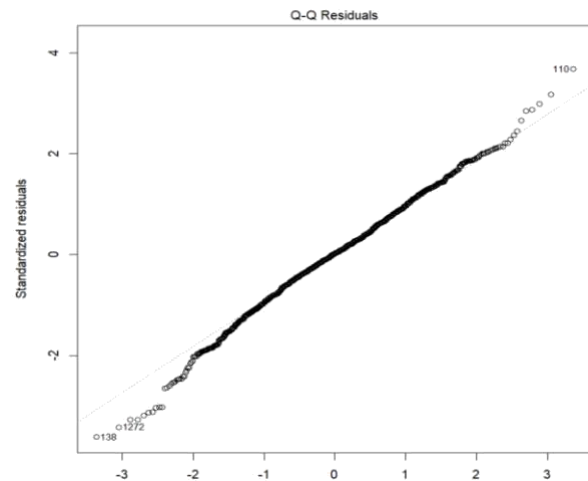


Linearity:

From the plot we can see that the red line is nearly straight indicating a high degree on linearity in our parameters.

Normality:

This plot shows that the residuals / errors terms are normally distributed as a high majority of the points fall on the straight line.



Homoscedasticity:

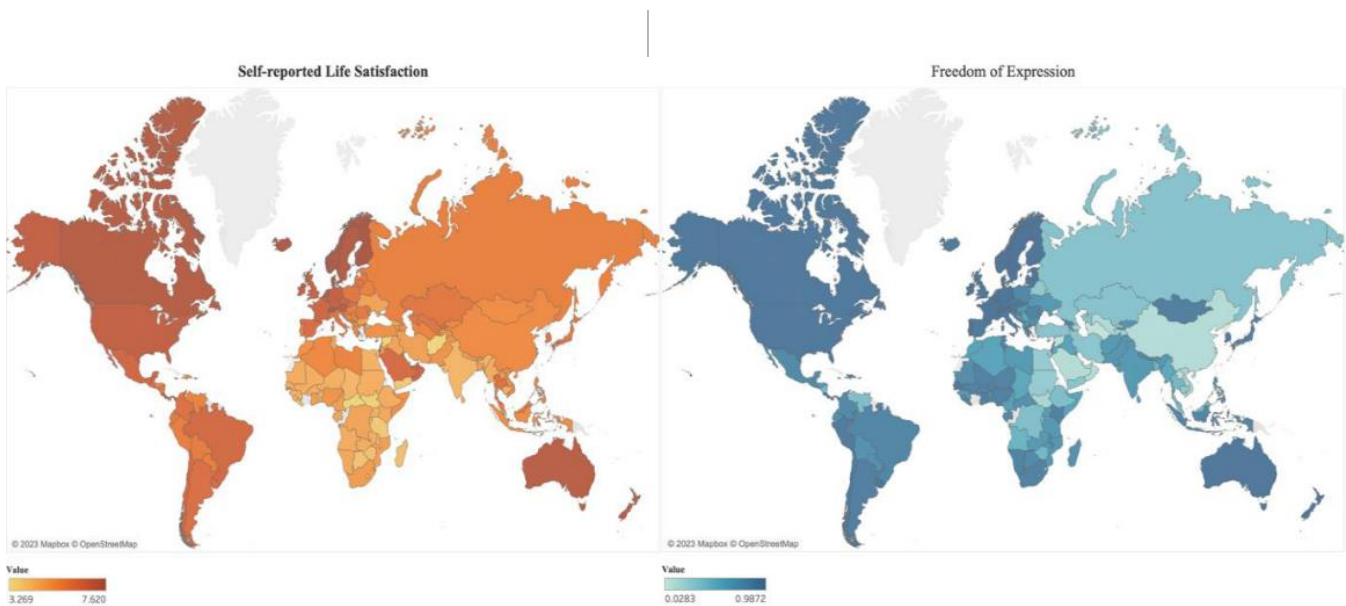
The plot tests for a constant variance, from the below graph we can say that the data is generally homoscedastic.

Additionally, the mean of the residuals is 2.580 which is nearly 0 and so is the variance. Proving our model to be a good fit.

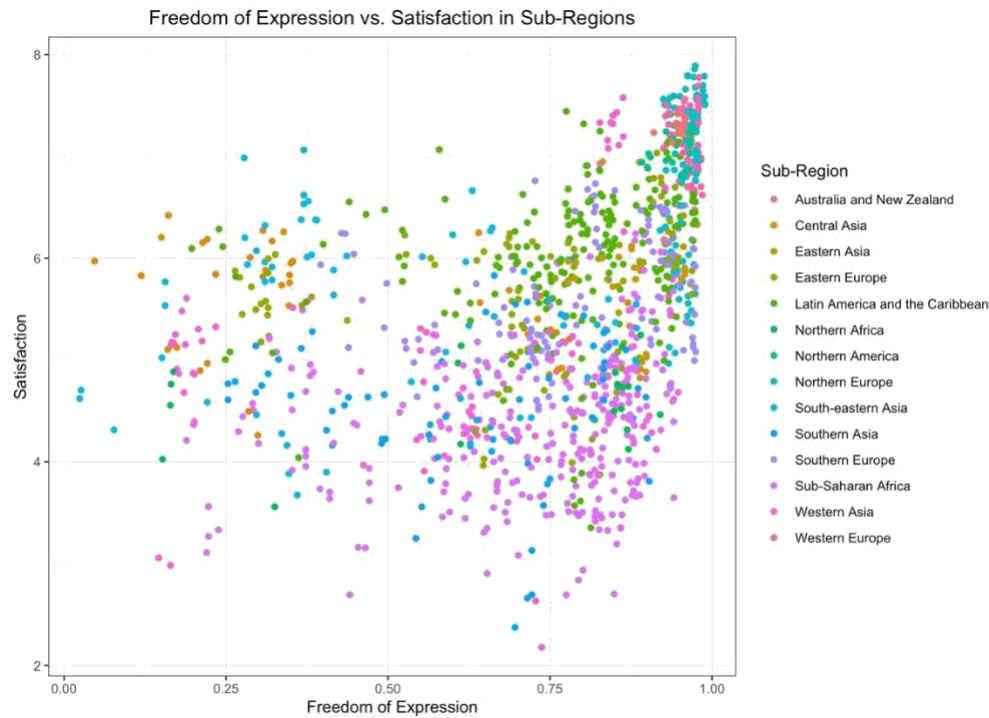
```
> mean(lm2$residuals)
[1] 2.580469e-18
> var(lm2$residuals)
[1] 0.2742251
```

V. Graphical Analysis:

Life Satisfaction vs. Freedom of Expression

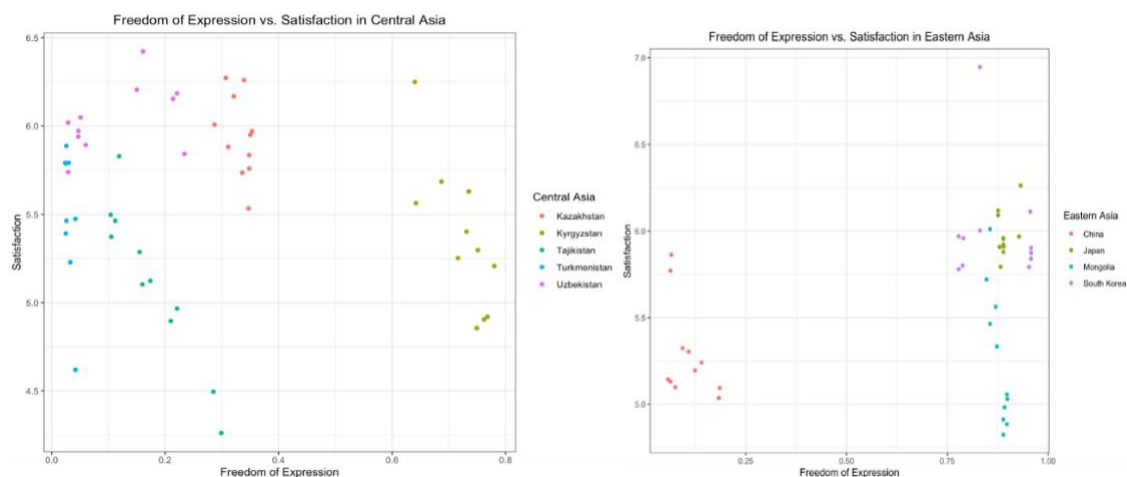


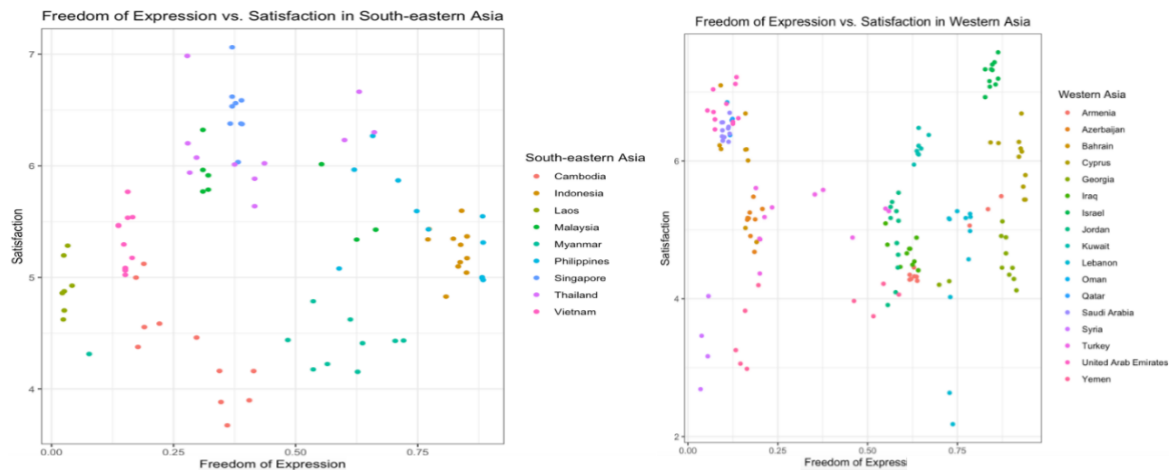
Generally, a positive relationship is observed between life satisfaction and freedom of expression. People in Oceania, Europe, and Americas typically experience higher levels of freedom of expression and life satisfaction, in contrast to those in Asia and Africa, where these levels are often lower. In the Middle East, life satisfaction seems to be high, but freedom of expression is notably low. For contrast, sub-Saharan Africa seems to have a higher freedom of expression but lower life satisfaction.



The scatterplot above illustrates the relationship we are investigating by subregion. For most subregions, there is a positive relationship between satisfaction and freedom of expression, stronger in western countries. According to Pittau, Zelli, and Gelman (2010), in poorer regions, basic economic factors primarily influence life satisfaction, whereas in wealthier regions, cultural aspects play a more significant role in determining life satisfaction.

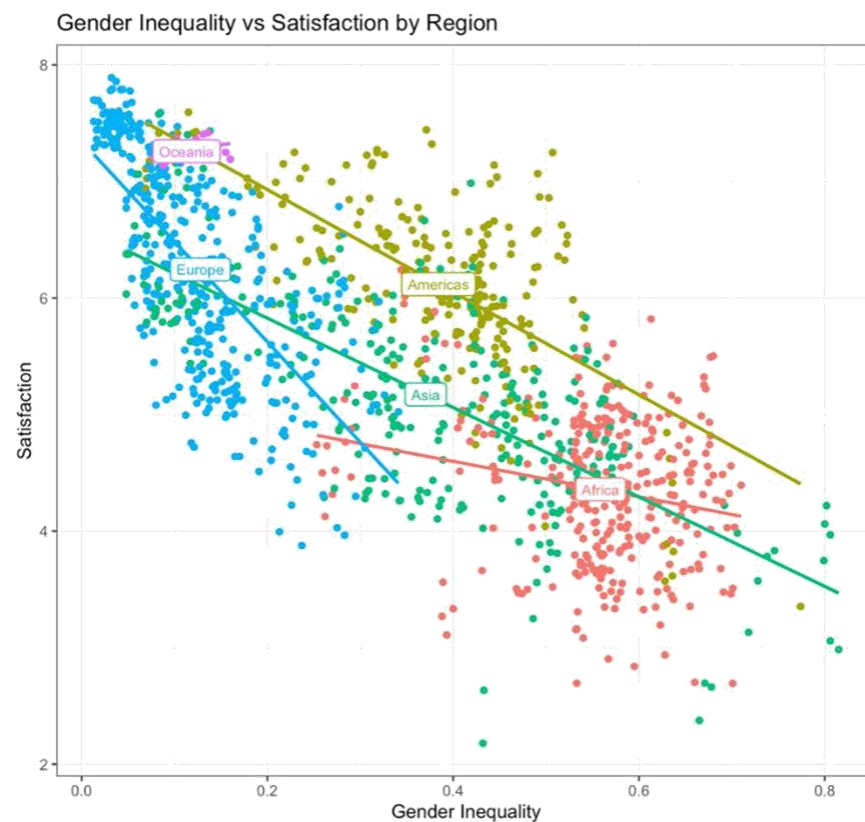
Since Oceania, Europe and America are already economically advanced, freedom of expression as an aspect of non-materialistic issues is an important factor influencing life satisfaction. In contrast, most countries in Asia and Africa focus more on economic development than on non-material aspects such as freedom of expression. And this may be why some countries in Asia have low levels of freedom of expression but high levels of satisfaction. These regions are shown in the four charts below.





The charts highlight Asian countries with high satisfaction but low freedom of expression, including Uzbekistan, Turkmenistan, China, Laos, Vietnam, Saudi Arabia, Qatar, and the UAE. The Middle East exhibits high satisfaction, likely influenced by their substantial natural resource wealth, and perhaps even by their highly religious cultures. Marital status and standard of living strongly impact life satisfaction in Central, East, and Southeast Asia, reflecting the region's emphasis on family values, suggesting that freedom of expression may not be the foremost factor influencing life satisfaction in certain Asian countries (Bayomi and Fernandez, 2019; Ngoo, Tey, and Tan, 2015).

GII vs. Freedom of Expression



The scatterplot illustrates the inverse relationship between gender inequality and life satisfaction across regions. This aligns with Fischer, Bjørnskov, and Dreher's (2007) findings that

increased societal equality correlates with higher life satisfaction for both genders. Regional variations exist, with Europe exhibiting the steepest slope, indicating greater sensitivity to changes in gender inequality. Surveyed European women's strong concerns about gender equality contribute to this heightened sensitivity. Additionally, the Americas demonstrate higher life satisfaction than Africa and Asia at the same GII level, attributed to favourable policies, laws, and educational resources in North American countries.

VI. Conclusion

Our study of the relationship between freedom of expression, socio-economic factors, and life satisfaction, shows us that although a positive correlation between freedom of expression and life satisfaction is evident, our findings highlight the importance of other variables with greater impact.

Further research into global sub-regions unravelled variations in the strength and directionality of relationships. Notably, the impact of freedom of expression, while positive, was heterogenous in different sub regions. The concept of freedom of expression, often considered paramount, revealed itself as a 'luxury' concern, with more fundamental factors such as social support, governance, and gender equality assuming precedence in influencing life satisfaction.

An avenue for substantial exploration lies in delving deeper into social support factors. Examining the interactions of social support within distinct sub-regions could provide an interesting insight into other determinants influencing life satisfaction.

VII. References

- Başlevent, C. & Kirmanoğlu, H., 2017. Gender Inequality in Europe and the Life Satisfaction of Working and Non-Working Women. *Journal of Happiness Studies*, 18, pp.107-124.
- Bayomi, N. & Fernandez, J.E., 2019. Towards Sustainable Energy Trends in the Middle East: A Study of Four Major Emitters. *Energies*, 12(9), 1615.
- Chung, K.H. and Im, T. (2021). Happiness in developing countries: can government competitiveness substitute for formal institutions? *International Review of Administrative Sciences*, p.002085232110004.
- Fischer, J.A., Bjørnskov, C. & Dreher, A., 2007. On Gender Inequality and Life Satisfaction: Does Discrimination Matter?. University of St. Gallen, Economics Discussion Paper.
- Helliwell, J.F., Barrington-Leigh, C. P., Harris, A., Huang, H. (2008). International evidence on the social context of well-being. NBER Working Paper Series, No. 14720. <https://www.nber.org/papers/w14720/>
- Pittau, M.G., Zelli, R. & Gelman, A., 2010. Economic Disparities and Life Satisfaction in European Regions. *Social Indicators Research*, 96(2), pp.339–361.
- Ngoo, Y.T., Tey, N.P. & Tan, E.C., 2015. Determinants of Life Satisfaction in Asia. *Social Indicators Research*, 124(1), pp.141–156.
- Oswald, A. J. (1997). Happiness and Economic Performance. *The Economic Journal*, 107(445), 1815– 1831.
- Our World in Data. (n.d.). World Bank income groups. [online] Available at: <https://ourworldindata.org/grapher/world-banks-income-groups>
- Our World in Data. (n.d.). Self-reported life satisfaction. [online] Available at: <https://ourworldindata.org/grapher/happiness-cantril-ladder?tab=table>
- Steckermeier, L.C. (2020) The Value of Autonomy for the Good Life. An Empirical Investigation of Autonomy and Life Satisfaction in Europe. *Soc Indic Res* 154, 693–723
- Sujarwoto, S., Tampubolon, G. Decentralisation and Citizen Happiness: A Multilevel Analysis of Self-rated Happiness in Indonesia. *J Happiness Stud* 16, 455–475 (2015).
- V-dem.net. (2022). The V-Dem Dataset – V-Dem. [online] Available at: <https://v-dem.net/data/the-v-dem-dataset>
- Vittinghoff, E. (2012). Regression methods in biostatistics : linear, logistic, survival, and repeated measures models. New York: Springer.

VIII. Appendix

R Code Input :

```
>
> dfn = data26
> dfn$Higher = ifelse( dfn$income_group == "High-income countries", 1, 0)
> dfn$LowerMiddle = ifelse( dfn$income_group == "Lower-middle-income countries", 1, 0)
> dfn$Lower = ifelse( dfn$income_group == "Low-income countries", 1, 0)
> dfn$UpperMiddle = ifelse( dfn$income_group == "Upper-middle-income countries", 1, 0)
> df = na.omit(dfn)
> ## first lm
> lm1= lm(df$Satisfaction ~ df$freedom.of.expression, data= df)
> summary(lm1)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression, data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-3.2983 -0.8155  0.0863  0.8580  2.4621

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      3.9458     0.1078   36.59 <2e-16 ***
df$freedom.of.expression  2.0778     0.1374   15.13 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.058 on 1272 degrees of freedom
Multiple R-squared:  0.1525,    Adjusted R-squared:  0.1518
F-statistic: 228.8 on 1 and 1272 DF,  p-value: < 2.2e-16

> ##social support
> lms = lm(df$Satisfaction ~ df$freedom.of.expression +df$Social.support)
> summary(lms)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$Social.support)

Residuals:
    Min       1Q   Median       3Q      Max
-2.37032 -0.45359  0.06603  0.57202  2.32001

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    -0.5771     0.1457   -3.96 7.9e-05 ***
df$freedom.of.expression  1.0639     0.1002   10.61 < 2e-16 ***
df$Social.support    6.5056     0.1792   36.31 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7417 on 1271 degrees of freedom
Multiple R-squared:  0.5839,    Adjusted R-squared:  0.5833
F-statistic: 891.9 on 2 and 1271 DF,  p-value: < 2.2e-16

> ##GII
> lmgii = lm( df$Satisfaction~ df$freedom.of.expression+df$gii_value)
> summary(lmgii)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$gii_value)

Residuals:
    Min       1Q   Median       3Q      Max
-2.95732 -0.53179  0.06706  0.54978  2.32803

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      6.19329     0.09936   62.330 < 2e-16 ***
df$freedom.of.expression  0.85741     0.10334    8.297 2.7e-16 ***
df$gii_value     -3.90990     0.11030  -35.449 < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.7506 on 1271 degrees of freedom
Multiple R-squared:  0.5738,    Adjusted R-squared:  0.5732
F-statistic: 855.7 on 2 and 1271 DF,  p-value: < 2.2e-16

> ## per of cor
> lmr = lm(df$Satisfaction~ df$freedom.of.expression + df$Perceptions.of.corruption)
> summary(lmr)
```

```

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$Perceptions.of.corruption)

Residuals:
    Min       1Q   Median       3Q      Max
-2.8726 -0.5948  0.1595  0.6418  2.7957

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      6.1583    0.1518   40.56 <2e-16 ***
df$freedom.of.expression  1.7921    0.1226   14.62 <2e-16 ***
df$Perceptions.of.corruption -2.6817    0.1431  -18.74 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9369 on 1271 degrees of freedom
Multiple R-squared:  0.336,    Adjusted R-squared:  0.335
F-statistic: 321.6 on 2 and 1271 DF,  p-value: < 2.2e-16

> ##second lm + control - SUB REG
> lm2 = lm(df$Satisfaction ~ df$freedom.of.expression + df$gii_value + df$Social.support + df$Perceptions.of.corruption+df$LowerMiddle
+ df$Lower+df$UpperMiddle + df$Higher + factor(df$year), data= df )
> summary(lm2)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$gii_value +
  df$Social.support + df$Perceptions.of.corruption + df$LowerMiddle +
  df$Lower + df$UpperMiddle + df$Higher + factor(df$year),
  data = df)

Residuals:
    Min       1Q   Median       3Q      Max
-2.13483 -0.35059  0.01447  0.38169  2.16848

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)      4.00140    0.22007   18.182 < 2e-16 ***
df$freedom.of.expression  0.62641    0.09062    6.913 7.56e-12 ***
df$gii_value     -0.74537    0.17009   -4.382 1.27e-05 ***
df$Social.support  3.47979    0.19939   17.452 < 2e-16 ***
df$Perceptions.of.corruption -1.48117    0.10248  -14.453 < 2e-16 ***
df$LowerMiddle    -0.49784    0.07910   -6.294 4.27e-10 ***
df$Lower         -0.94021    0.09864   -9.532 < 2e-16 ***
df$UpperMiddle    -0.26625    0.06317   -4.215 2.68e-05 ***
df$Higher         NA           NA       NA      NA
factor(df$year)2012  -0.06094    0.07872   -0.774  0.4390
factor(df$year)2013  -0.10766    0.07825   -1.376  0.1691
factor(df$year)2014  -0.17237    0.07746   -2.225  0.0262 *
factor(df$year)2015  -0.13204    0.07755   -1.703  0.0889 .
factor(df$year)2016  -0.17998    0.07784   -2.312  0.0209 *
factor(df$year)2017  -0.09136    0.07703   -1.186  0.2359
factor(df$year)2018  -0.07114    0.07779   -0.914  0.3606
factor(df$year)2019  -0.08104    0.07790   -1.040  0.2984
factor(df$year)2020  -0.05188    0.08257   -0.628  0.5299
factor(df$year)2021  -0.05315    0.08232   -0.646  0.5186
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5936 on 1256 degrees of freedom
Multiple R-squared:  0.7366,    Adjusted R-squared:  0.733
F-statistic: 206.6 on 17 and 1256 DF,  p-value: < 2.2e-16

> ##second lm + control + SUB REG
> lm3 = lm(df$Satisfaction ~ df$freedom.of.expression + df$gii_value + df$Social.support + df$Perceptions.of.corruption+df$LowerMiddle
+ df$Lower+df$UpperMiddle + df$Higher +factor(df$year) + factor(df$SubRegion.Name), data= df)
> summary(lm3)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$gii_value +
  df$Social.support + df$Perceptions.of.corruption + df$LowerMiddle +
  df$Lower + df$UpperMiddle + df$Higher + factor(df$year) +
  factor(df$SubRegion.Name), data = df)

```

```

Residuals:
    Min       1Q   Median       3Q      Max
-1.98285 -0.31716  0.02583  0.32199  2.08800

Coefficients: (1 not defined because of singularities)
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.14022    0.25058   20.513 < 2e-16 ***
df$freedom.of.expression  0.42288    0.09390    4.503 7.31e-06 ***
df$gii_value   -1.55491    0.23024   -6.753 2.21e-11 ***
df$Social.support  2.40130    0.20364   11.792 < 2e-16 ***
df$Perceptions.of.corruption -0.89870    0.12545   -7.164 1.34e-12 ***
df$LowerMiddle -0.49225    0.07485   -6.576 7.09e-11 ***
df$Lower      -0.75862    0.09376   -8.091 1.39e-15 ***
df$UpperMiddle -0.38018    0.05889   -6.456 1.54e-10 ***
df$Higher      NA           NA         NA      NA
factor(df$year)2012    -0.05879    0.07032   -0.836 0.403299
factor(df$year)2013    -0.11541    0.06990   -1.651 0.098960 .
factor(df$year)2014    -0.14646    0.06922   -2.116 0.034565 *
factor(df$year)2015    -0.12052    0.06930   -1.739 0.082256 .
factor(df$year)2016    -0.16154    0.06965   -2.319 0.020531 *
factor(df$year)2017    -0.08268    0.06903   -1.198 0.231197
factor(df$year)2018    -0.03805    0.06976   -0.545 0.585535
factor(df$year)2019    -0.05089    0.07009   -0.726 0.467956
factor(df$year)2020    -0.02261    0.07418   -0.305 0.760552
factor(df$year)2021    -0.05283    0.07398   -0.714 0.475272
factor(df$SubRegion.Name)Central Asia    -0.09485    0.16680   -0.569 0.569684
factor(df$SubRegion.Name)Eastern Asia    -0.65389    0.15719   -4.160 3.40e-05 ***
factor(df$SubRegion.Name)Eastern Europe  -0.47369    0.14466   -3.274 0.001088 **
factor(df$SubRegion.Name)Latin America and the Caribbean  0.29368    0.13714    2.142 0.032424 *
factor(df$SubRegion.Name)Northern Africa -0.58522    0.18502   -3.163 0.001599 **
factor(df$SubRegion.Name)Northern America  0.20455    0.16216    1.261 0.207383
factor(df$SubRegion.Name)Northern Europe  -0.24760    0.12613   -1.963 0.049869 *
factor(df$SubRegion.Name)South-eastern Asia -0.16413    0.14860   -1.104 0.269593
factor(df$SubRegion.Name)Southern Asia    -0.33109    0.15606   -2.122 0.034076 *
factor(df$SubRegion.Name)Southern Europe  -0.62491    0.14409   -4.337 1.56e-05 ***
factor(df$SubRegion.Name)Sub-Saharan Africa -0.51526    0.14688   -3.508 0.000467 ***
factor(df$SubRegion.Name)Western Asia    -0.47677    0.14555   -3.276 0.001083 **
factor(df$SubRegion.Name)Western Europe  -0.04525    0.13156   -0.344 0.730959
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.5299 on 1243 degrees of freedom
Multiple R-squared:  0.7923,    Adjusted R-squared:  0.7872
F-statistic: 158 on 30 and 1243 DF,  p-value: < 2.2e-16

> ## per of cor
> lmr = lm(df$Satisfaction~ df$freedom.of.expression + df$Perceptions.of.corruption)
> summary(lmr)

Call:
lm(formula = df$Satisfaction ~ df$freedom.of.expression + df$Perceptions.of.corruption)

Residuals:
    Min       1Q   Median       3Q      Max
-2.8726 -0.5948  0.1595  0.6418  2.7957

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.1583    0.1518   40.56 <2e-16 ***
df$freedom.of.expression  1.7921    0.1226   14.62 <2e-16 ***
df$Perceptions.of.corruption -2.6817    0.1431  -18.74 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.9369 on 1271 degrees of freedom
Multiple R-squared:  0.336,    Adjusted R-squared:  0.335
F-statistic: 321.6 on 2 and 1271 DF,  p-value: < 2.2e-16

> adf.test(df$Satisfaction)

Augmented Dickey-Fuller Test

data: df$Satisfaction
Dickey-Fuller = -8.6864, Lag order = 10, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(df$Satisfaction) : p-value smaller than printed p-value
> adf.test(df$freedom.of.expression)

```