UNIVERSITY OF PENNSYLVANIA

SCHOOL OF SOCIAL POLICY AND PRACTICE

MASTER OF SCIENCE IN SOCIAL POLICY AND DATA ANALYTICS

FINAL POLICY RESEARCH REPORT 4

AWARENESS OF CLIMATE CHANGE AND ENVIRONMENTAL IMPACT

Authors:

Rose Barragan

YuQin Li

Professor: Dr. Momin Malik

Course: MSSP 8970

Applied Linear Modeling

Fall 2023

December 21st, 2023

INDEX

1. INTRODUCTION AND RELEVANCE OF THE STUDY

Climate change is a global phenomenon affecting millions of people around the world significantly. It is notorious that climate change is causing alterations in weather patterns, a rise in sea levels, and increasingly frequent extreme weather events.

For example, according to data from the World Meteorological Organization (WMO) –the UN agency responsible for monitoring and protecting the climate and the environment [1] - the global average temperature in 2019 has already exceeded that of the pre-industrial period by 1.1ºC. And 2023 stands out for being the hottest year in history [2] due to the retreat of ice and increased emissions of greenhouse gases of human origin.

As the Six Assessment Report launched by the Intergovernmental Panel on Climate Change (IPCC) [3] points out, the impacts of this temperature increase are already evident today, manifesting themselves in extreme weather events such as heat waves, droughts, floods, hurricanes, and uncontrolled wildfires.

Those most affected by this crisis are the most vulnerable and impoverished communities and the natural environment. If no action is taken, the average temperature of the Earth's surface is projected to rise during this century, exceeding 3ºC, with some regions experiencing even more pronounced warming [4].

Therefore, it is critical to emphasize that this is a global problem that needs to be addressed from an international perspective, as set out in the United Nations (UN) Sustainable Development Goals (SDGs) [5]. Global collaboration and countries' commitment are fundamental to developing effective strategies and lasting solutions from the perspective of sustainability and the environment.

In this critical context, this project aims to explore people's levels of Knowledge, Awareness, caring, concern, and actions towards environmental protection. Based on this data, the project seeks to create a series of recommendations that can contribute to increasing people's knowledge and design actions to help mitigate the effects of climate change at a global level.

2. SCOPE AND ALIGNMENT WITH THE SDGs

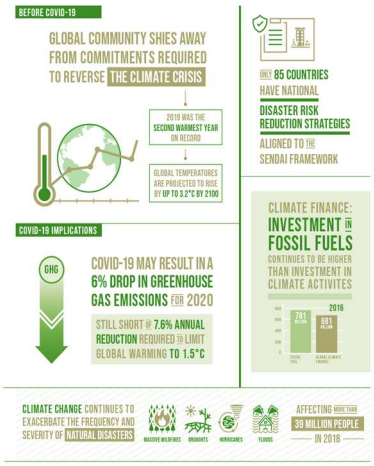
As mentioned in the previous section, it is essential to acknowledge that climate change and its harmful consequences for humans and the planet involves a joint international perspective through the Sustainable Development Goals (SDGs) of the United Nations [5].

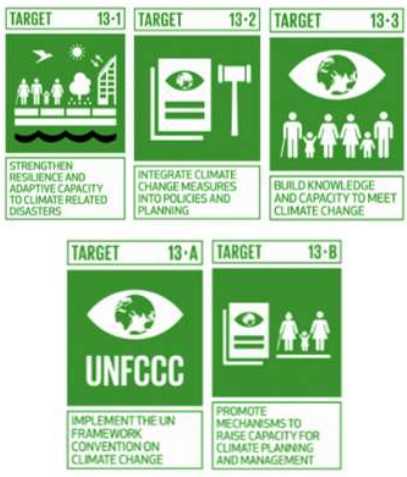
The SDGs set a universal agenda structured in 17 interconnected global goals established in 2015 to jointly address social and environmental challenges to achieve a more equitable, prosperous, and sustainable future.

Tackling climate change and environmental protection aligns closely with several of the SDGs, particularly Goal #13, *"urgent action to combat climate change and its impacts,"* which has a total of five milestones or targets [6]:

- ***Strengthen*** *resilience and adaptive capacity to climate-related disasters*
- ***Integrate*** *climate change measures into policies and planning*
- ***Build*** *Knowledge and capacity to meet climate change*
- ***Build*** *Knowledge and capacity to meet climate change*
- ***Promote*** *mechanisms to raise capacity for planning and management*

To address climate change, countries adopted the Paris Agreement at COP21 [7] in Paris on December 12, 2015, to limit global temperature rise to below 2 degrees Celsius. However, the climate crisis remains unstopped, as the international community is reluctant to commit fully to its unwinding.





Source: [13] University of Saskatchewan. (n.d.).

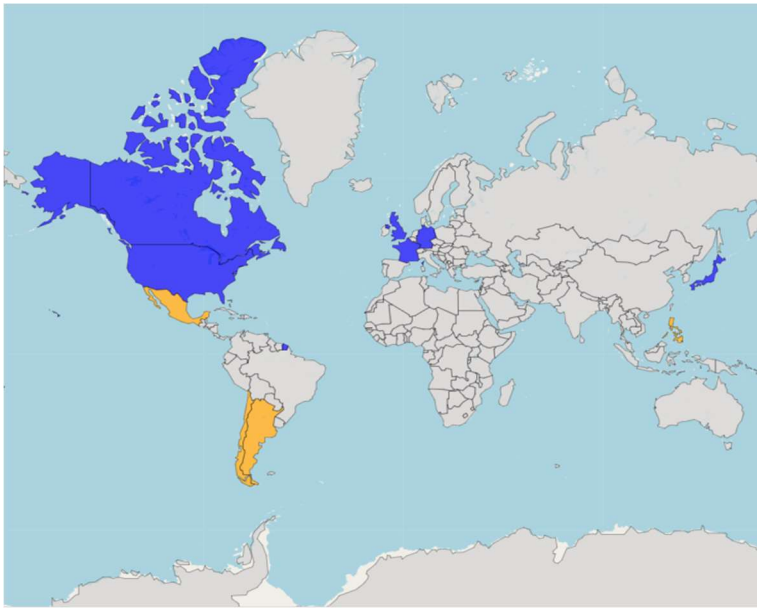
Source: [6] United Nations. (n.d.).

3. PROJECT'S GOALS AND METHODS FROM A DATA ANALYTICS APPROACH

As mentioned in the previous chapters, climate change requires urgent and coordinated Action through robust government policies and effective individual behaviors. That is why the main objective of this project is to address this challenge through a data analysis-based approach.

To achieve this goal, the research will start analyzing the answers collected in the 2010 Environment International Social Survey Programme (ISSP) given by ten countries, dividing them into two groups:

- For the developed nations (in blue), there are the United States, the United Kingdom, Japan, Germany, France, and Canada.
- For the developing countries (in orange), there are Mexico, the Philippines, Argentina, and Chile.



Source: Barragan,R., Li,Y. (December, 2023) (own).

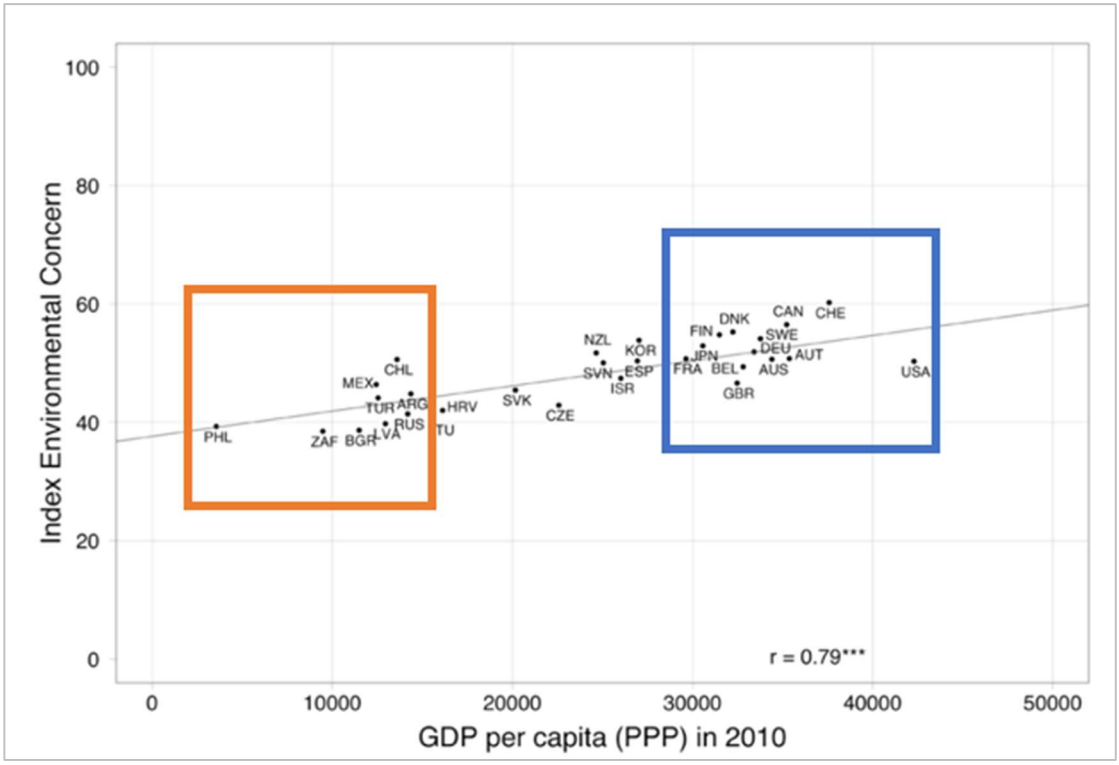
The reason why we perform this partitioning is because we refer to the insights gathered in a study conducted in October 2013 by Axel Franzen and Dominikus Vogl, called, "*Two Decades of Measuring Environmental Attitudes: A Comparative Analysis of 33 Countries*" [8]. The study concluded that there is a positive correlation between people's level of environmental concern and a country's prosperity (DGP).

While it is true that over the last 20 years, the environmental concern of the population in many countries has declined slightly, in countries where the economic situation has improved, the decrease in the level of concern is relatively smaller.

In other words, this correlation between the GDP and environmental concern level suggests that economic development may contribute to maintaining higher levels of environmental concern.

This is why our study starts from this premise, selecting countries in the top 10 at each extreme [Figure #1], as they are either economically strong (developed) or experiencing rapid economic growth (developing).

Figure #1 – GDP per capita and Environmental Concern levels



Source: [8] Franzen, A., & Vogl, D. (2013).

Therefore, we intend to examine and compare the attitudes and actions of these two groups of nations in four key areas: environmental Awareness, Willingness to contribute to environmental protection, daily actions in favor of the environment, and level of environmental Knowledge. In other words, these four areas of questioning refer to the cognitive aspect (whether there is a rational insight) or the conative aspect (whether there is a willingness to take Action based on that insight) [8].

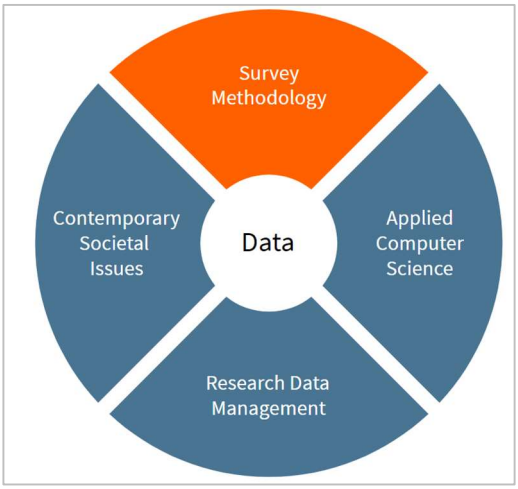
That is why the focus on data analysis is crucial to obtaining solid, evidence-based information on how these countries address climate change. We hope to better understand how individual and collective attitudes relate to environmental protection by analyzing data gathered for patterns and correlations.

Our ultimate goal is to use this information to design best practices that can contribute significantly to tackling climate change and promoting environmental protection globally.

4. DATASET OVERVIEW

4.1. DESCRIPTION OF THE DATASET: SOURCES

The dataset on which this study is based is provided by GESIS, which is an interdisciplinary social science research institution working in four broad areas [Figure #2] to research and develop methods for the use of digital behavioral data, optimize the usability of social science data, and advance answers to today's big social questions [9].



Source: [9] GESIS. (n.d.).

GESIS leads the International Social Survey Program (ISSP) [10]. The ISSP was born in 1984 as an ongoing program of cross-national collaboration to conduct annual surveys on topics of importance to the social sciences and now has nearly 50 member countries around the world.

Each ISSP module focuses on a specific topic. In our case, the questions/dataset revolve around the environment, climate change, and environmental protection [11]. The surveys are repeated at regular intervals, collecting random samples of citizens from participating nations, and are used to produce a ranking of a country's average environmental concern.

The ISSP surveys contain a series of questions in which respondents can indicate their agreement or disagreement on a five-point scale (Likert scale) [12] ranging from "strongly agree" to "strongly disagree" (or from very willing to very unwilling).

4.2. DESCRIPTION OF THE DATASET: VARIABLES

4.2.1. Independent Variables

Independent variables are understood as those that influence or predict the dependent variable. We generated three composite variables as our independent variables, denoted as ① Environmental Knowledge Level (Knowledge), ② Environmental Awareness (Awareness), and ③ Daily Environmental Action (Action), respectively. These variables were created by averaging participants' responses to a set of survey questions, all measured on a five-point Likert scale.

The creation of these variables will support the following statistical and linear regression analysis to represent the respondents' answers in a clearer and more understandable way, thus facilitating the analysis and interpretation of the data. The specific composition of the variables is as follows:

1) Environmental Knowledge Level

- V18: Level of Knowledge and Awareness of the causes of various environmental issues
- V19: Level of Knowledge and Awareness of the Solutions to various environmental issues

2) Environmental Awareness

- V39: Environmental pollution caused by cars
- V40: Environmental pollution caused by industries
- V41: Pollution caused by the use of pesticides and chemicals
- V42: Pollution of rivers and lakes
- V43: Rise in the world's temperature caused by climate change
- V44: Modifying the genes of certain crops
- V45: Nuclear power stations

3) Daily Environmental Actions

- V55: Recycling plastic and newspapers
- V56: Purchasing fruits and vegetables without pesticides and fertilizers
- V57: Reducing driving for environmental protection
- V58: Reducing household energy fuel usage for environmental protection
- V59: Storing or reusing water resources for environmental protection
- V60: Avoiding the purchase of specific products for environmental protection

4.2.2. Dependent Variable

We understand dependent variables as those that seek to explain the independent variables. In its analysis, the dependent variable is a generated composite variable, denoted as "The Willingness to Contribute to Environmental Protection (Willingness)." This variable is constructed based on the following questions:

- V29: Willing to incur higher costs for environmental protection
- V30: Willing to bear higher taxes for environmental protection
- V31: Willing to reduce living expenses for environmental protection

5. PRELIMINARY DATA WORK: DATA CLEANING

5.1. OBSERVATIONS

The raw dataset has a total of **50437 observations** corresponding to the 28 countries whose responses have been collected. However, as explained in Section 3, this study will focus on the performance of 10 countries (US, GB, JP, DE, FR, CA, MX, PH, AR, CL). Thus, the analysis considers only the data from the chosen countries, leading to a considerable reduction in the number of observations, coming down to 13713 after the filtering. It is worth noting that, from the **13713 observations**, 8310 (60.6%) belong to developed countries, and 5402 (39.4%) refer to developing countries.

Furthermore, we created a new column for binary codification where the value of "1" refers to the developed countries (with the specified codes: US, GB, JP, DE, FR, CA) and "0" for the developing ones (MX, PH, AR, CL).

5.1.1. Handling Missing Values

Another data cleaning technique implemented before moving on to linear regression analysis is the removal of missing values (NAs) from our dataset. After this operation, we will also remove the values 8 ("can't choose") and 9 ("No answer"). So, implementing all these filtering and cleaning techniques will result in a complete dataset with a total of **7443 observations**.

5.2. REASSIGNMENT OF VALUES

The reason we reassigned the value of the columns (29, 30, 31, 39, 40, 41, 42, 43, 44, 45, 55, 56, 57, 58, 59, 60) is to achieve a better understanding and clarity of the variables, thus contributing to our upcoming linear regression analysis work.

With this reassignment of values, we seek to reflect that respondents who express higher levels of agreement also have higher levels of environmental concern. Such reallocation is done by subtracting each original value in these columns out of 6.

In other words, the original scale is being inverted. After this transformation, higher values in these columns (6) indicate higher agreement or environmental Awareness, and 0 represents the minimum, achieving greater clarity and comprehension of the responses.

5.3. VARIABLE CENTERING

The centering of the variables is essential to better compare them and prepare the field for the following linear regression analysis work.

We have achieved the centering of the variables by subtracting 3. After the centering, the new mean of these variables will be closer to zero, which will facilitate the interpretation of the results.

Table #1 – All variables after reassignment and centering

Variable	Min	1st Qu	Median	Mean	3rd Qu	Max
V18C	-2	0	0	0.1955	1	2
V19C	-2	-1	0	0.04662	1	2
V29rC	-2	-1	0	-0.04662	1	2
V30rC	-2	-2	0	-0.4099	1	2
V31rC	-2	-1	0	-0.1364	1	2
V39rC	-2	0	1	0.6276	1	2
V40rC	-2	0	1	1.041	2	2
V41rC	-2	0	1	0.9078	2	2
C42rC	-2	0	1	0.9927	2	2
V43rC	-2	0	1	0.8166	2	2
V44rC	-2	0	1	0.5165	1	2
V45rC	-2	0	1	0.6487	2	2
V55rC	-1	1	2	1.253	2	2
V56rC	-1	0	0	0.3581	1	1
V57rC	-1	-1	0	0.08061	1	2
V58rC	-1	0	1	0.5362	1	2
V59rC	-1	0	0	0.4498	1	2
V60rC	-1	0	0	0.3767	1	2

Source: Barragan,R., Li,Y. (December, 2023) (own).

5.4. CRONBACH ALPHA TEST

Before proceeding to the statistical and linear regression analysis, we will conduct an analysis of internal consistency and reliability through the Cronbach Alpha Test applied to our set of variables in a questionnaire [14]. Cronbach's alpha values range from 0 to 1, so the higher values indicate higher reliability. Recall that section 4.2.1. identifies the four groups of independent variables:

1) Willingness (contains variables related to Willingness to contribute to environmental protection), including v29rC, v30rC, and v31rC, with an alpha score of 0.82.

2) Awareness (contains variables related to environmental Awareness), which includes v39rC, v40rC, v41rC, v42rC, v43rC, v44rC, and v45rC, with the alpha result being 0.81.

3) Action (contains variables related to environmental actions), which includes v55rC, v56rC, v57rC, v58rC, v59rC, and v60rC, the alpha result being 0.79.

4) Knowledge (contains the variables related to environmental Knowledge), which includes v18C and v19C, with an alpha score of 0.77.

These results show that Cronbach's alpha values range between 0.77 and 0.82, indicating excellent consistency and reliability in each set.

6. DESCRIPTIVE STATISTICS

Table #2 – Descriptive statistics

Variables	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
C_flag	2	7443	0.72	0.45	1.00	0.78	0.00	0	1	1	-1.01	-0.99	0.01
v18C	37	7443	0.20	1.01	0.00	0.19	1.48	-2	2	4	-0.13	-0.28	0.01
v19C	38	7443	-0.12	1.01	0.00	-0.10	1.48	-2	2	4	0.01	-0.32	0.01
v29rC	39	7443	-0.05	1.19	0.00	-0.01	1.48	-2	2	4	-0.24	-1.00	0.01
v30rC	40	7443	-0.41	1.22	0.00	-0.44	1.48	-2	2	4	0.14	-1.16	0.01
v31rC	41	7443	-0.14	1.21	0.00	-0.12	1.48	-2	2	4	-0.14	-1.08	0.01
v39rC	42	7443	0.63	0.88	1.00	0.64	1.48	-2	2	4	-0.01	-0.56	0.01
v40rC	43	7443	1.04	0.82	1.00	1.09	1.48	-2	2	4	-0.52	-0.13	0.01
v41rC	44	7443	0.91	0.91	1.00	0.97	1.48	-2	2	4	-0.46	-0.39	0.01
v42rC	45	7443	0.99	0.88	1.00	1.06	1.48	-2	2	4	-0.58	-0.10	0.01
v43rC	46	7443	0.82	1.01	1.00	0.92	1.48	-2	2	4	-0.53	-0.35	0.01
v44rC	47	7443	0.52	1.08	1.00	0.56	1.48	-2	2	4	-0.28	-0.70	0.01
v45rC	48	7443	0.65	1.13	1.00	0.72	1.48	-2	2	4	-0.36	-0.84	0.01
v55rC	49	7443	1.25	0.99	2.00	1.42	0.00	-1	2	3	-1.03	-0.23	0.01
v56rC	50	7443	0.36	0.96	0.00	0.32	1.48	-1	2	3	0.11	-0.96	0.01
v57rC	51	7443	0.08	0.89	0.00	0.01	1.48	-1	2	3	0.44	-0.61	0.01
v58rC	52	7443	0.54	0.91	1.00	0.55	1.48	-1	2	3	-0.05	-0.81	0.01
v59rC	53	7443	0.45	0.96	0.00	0.44	1.48	-1	2	3	0.01	-0.95	0.01
v60rC	54	7443	0.38	0.89	0.00	0.35	1.48	-1	2	3	0.09	-0.74	0.01

Variables	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
awareness	55	7443	0.79	0.66	0.86	0.80	0.64	-2	2	4	-0.23	-0.23	0.01
willingness	56	7443	-0.20	1.04	0.00	-0.17	0.99	-2	2	4	-0.11	-0.80	0.01
action	57	7443	0.51	0.65	0.50	0.53	0.74	-1	2	3	-0.19	-0.37	0.01
knowledge	58	7443	0.04	0.91	0.00	0.04	0.74	-2	2	4	-0.06	-0.09	0.01

Source: Barragan,R., Li,Y. (December, 2023) (own).

The **awareness variable** presents a mean of 0.79, indicating a high consciousness level for protecting the environment. In addition, the median is slightly higher than the mean (0.86), suggesting that more than half of the respondents reflected a high level of environmental protection awareness.

Although the collected responses range between -2 and 2, there is a moderate variability in the responses reflected through a standard deviation equal to 0.66; it should be noted that there is a slight propensity towards the lower values (skew of -0.23). In addition, the distribution is slightly less steep than the normal distribution (Kurosis of -0.23) with shorter tails.

The **willingness variable** presents a mean of -0.20, indicating that people reflect a rather poor intention to protect the environment. This data is supported by the median value (0.00), indicating that respondents remain neutral regarding Willingness.

As with the awareness variable, willingness responses range between -2 and 2, but there is a high variability and dispersion in the responses reflected through a standard deviation equal to 1.04. Although the values suggest a reasonably symmetrical distribution (skew equal to -0.11), the distribution is flatter than usual (-0.80).

The **Action variable** presents a mean of 0.51, indicating a moderate level of proactivity in protecting the environment. In line with the mean, the median is 0.50, suggesting a perfect symmetry in the responses, where exactly half of the respondents have taken Action in favor of environmental protection.

A difference in the action variable with respect the others is that the collected responses range between -1 and 2, suggesting greater positive responses than negative ones, with a standard deviation equal to 0.74. As commented in the previous points, the data is closer to symmetry with a slight propensity towards the lower values (skew of -0.19). In addition, the distribution is slightly less steep than the normal distribution (Kurosis of -0.37).

The **knowledge variable** presents a mean of 0.04, indicating a good level of Knowledge regarding environmental protection. In line with the mean, the median is 0.00, suggesting it is very close to the mean and an almost perfect symmetry in the responses, and with a moderate standard deviation equal to 0.94.

As commented, the data is closer to symmetry with a slight propensity towards the lower values (skew of -0.06). In addition, the distribution is slightly less steep than the normal distribution (Kurosis of -0.09).

Overall, the statistical analysis indicates that there is a generally positive attitude towards the level of environmental Knowledge, Awareness, and Action, with moderate variability in these measures. The willingness variable has a slightly negative mean but high variability, suggesting that there are strongly divided opinions on this aspect.

7. CORRELATION MATRIX

For C_flag (indicate whether a country is classified as developed or developing), it is interesting that it has a moderate negative correlation with Awareness (-0.29). This suggests that developed countries tend to have lower environmental awareness levels than developing countries.

Additionally, there's a weak negative correlation between the C_flag and Willingness (-0.01), indicating that the status of being a developed country doesn't strongly influence the Willingness to take environmental Action.

However, it's worth noting that the C_flag positively correlates with Knowledge and Action, with values of 0.01 and 0.20, respectively. This implies a relatively weak connection between developed countries and environmental Knowledge. Still, interestingly, developed countries tend to take more environmental actions than developing nations.

Shifting to Awareness, it has a weak positive correlation with Action (0.19) and with Willingness (0.17). This suggests that as environmental awareness increases, there's a tendency for more environmental actions to be taken, and people are more willing to participate in environmental efforts.

Additionally, Awareness has a very weak positive correlation with Knowledge (0.05), indicating that there's almost no discernible relationship between the level of Awareness and environmental Knowledge. In other words, being more aware of environmental issues doesn't necessarily mean having more Knowledge about them.

Regarding people's Willingness to help the environment, it's interesting to note that it has a moderate positive correlation with taking Action (with a correlation of 0.22). This suggests that people who are more willing to make a difference in environmental matters tend to be more actively involved in environmental actions.

Furthermore, Willingness also shows a moderate positive correlation with Knowledge (with a correlation of 0.16). This implies that having a better understanding of environmental issues might slightly boost a person's eagerness to contribute to environmental protection efforts.

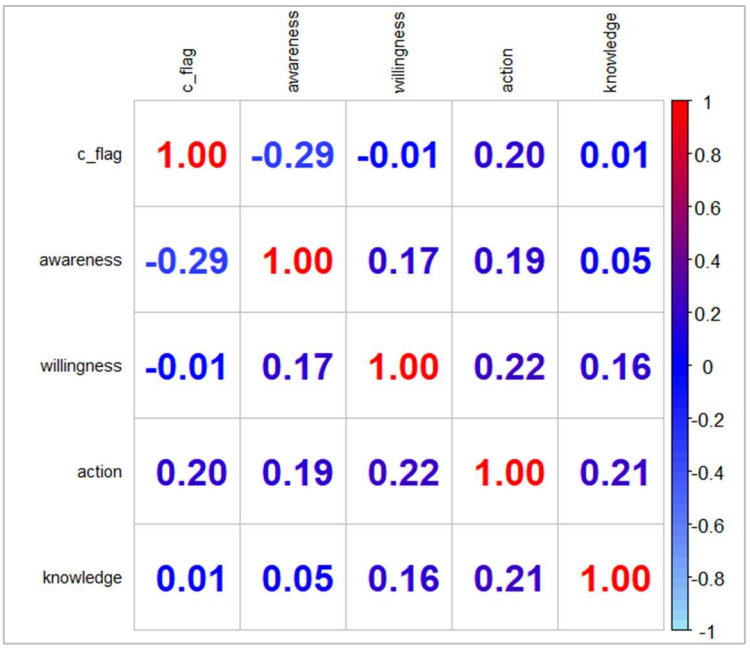
Shifting our focus to Knowledge, it appears to be moderately positively correlated with taking Action (with a correlation of 0.21). In other words, more Knowledge about the environment could lead to greater involvement in environmental actions.

Overall, it's important to acknowledge that these correlations move along the range between weak and moderate. This means that while there are correlations between these variables, they're not particularly strong.

However, it's worth noting that the correlation matrix brings out an interesting observation regarding environmental Awareness. It appears that environmental Awareness might be lower in developed countries compared to developing ones.

In addition, there's a moderately positive correlation between Willingness and Action, suggesting that those who are more inclined to contribute to environmental protection are also more likely to participate in environmental actions. Lastly, Knowledge seems to have a moderate correlation with taking Action, implying that it could influence how often someone gets involved in environmental actions.

Table #3 – Correlation matrix



Source: Barragan,R., Li,Y. (December, 2023) (own).

8. LINEAR REGRESSION ANALYSIS

For this linear regression model [Table #4], the dependent variable is Willingness, and the independent variables are c_flag, Awareness, Action, and Knowledge.

After controlling for Awareness, Action, and Knowledge, the coefficient for c_flag is –0.018. This coefficient indicates that when participants came from developed countries (c_flag = 1), the predicted willingness score decreased by 0.018 units compared to those from developing countries. However, this result is not statistically significant since the p-value is greater than 0.1.

On the other hand, the other three coefficients of Awareness (0.199), Action (0.279), and Knowledge (0.133) are all statistically highly significant for the linear regression model since they have a p-value of less than <0.01. This means that an increase in Awareness, as for Action, leads to a higher level of Willingness. In the same way, a higher level of Knowledge is positively associated with higher Willingness.

In summary, the regression analysis results indicate that, in this model, Awareness, Action, and Knowledge have a positive and statistically significant impact on Willingness, meaning that as these factors increase, Willingness to take Action also tends to increase. However, the variable C_flag (developed countries) does not appear to significantly affect Willingness in this model.

Table #4 – Stragarzer linear model

	Dependent variable:
	willingness
c_flag	-0.018
	(0.028)
awareness	0.199***
	(0.019)
action	0.279***
	(0.019)
knowledge	0.133***
	(0.013)
Constant	-0.490***
	(0.030)
Observations	7,443
R2	0.080
Adjusted R2	0.079
Residual Std. Error	0.993 (df = 7438)
F Statistic	161.560*** (df = 4; 7438)
Note:	*p<0.1; **p<0.05; ***p<0.01

Source: Barragan,R., Li,Y. (December, 2023) (own).

9. ASSUMPTIONS OF HOMOSCEDASTICITY AND LINEARITY

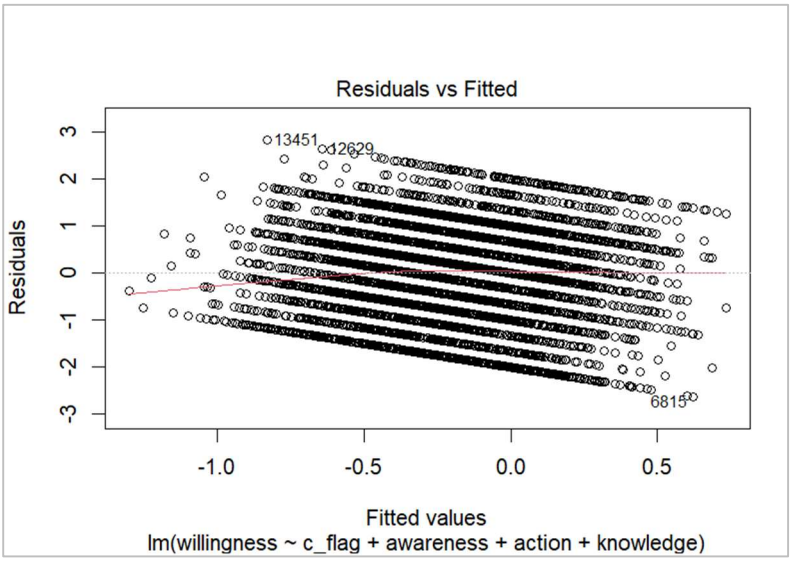
9.1. RESIDUALS VS FITTED

Under an ideal regression model, we would expect the residuals to be randomly scattered around the horizontal line in the plot, i.e., close to zero and showing no specific pattern. However, looking at the plot [Plot #1], we note the presence of distinct and nearly parallel lines of points, which is unusual and suggests that the data may not be continuous or that they may have discrete characteristics.

In addition, the dispersion of the residuals appears to hold consistently at different fitted values, which tells us that there is homogeneity in the variance, i.e., homoscedasticity. However, the presence of these discrete lines might suggest that the response variable could be categorical rather than continuous, as numerical responses correspond to a category.

Furthermore, the points that are labeled with numbers represent observations that could be influential or outliers. These points stand out because they have significantly large residuals and depart considerably from the values predicted by the model.

Plot #1 – Residuals vs Fitted



Source: Barragan,R., Li,Y. (December, 2023) (own).

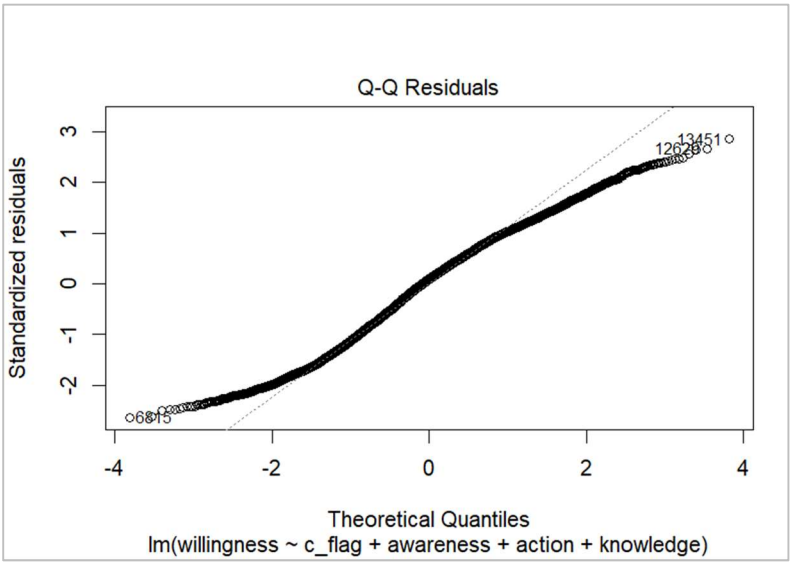
9.2. NORMALITY: Q-Q PLOT

Ideally, the points should fall approximately along a straight line if the data is normally distributed in a Q-Q plot. However, in this plot [Plot #2], the points deviate from the line systematically, suggesting that the residuals do not follow a normal distribution.

Furthermore, this plot shows that the residuals have a distribution with thicker tails than a normal distribution, which is evident from the curve at the ends of the plot. This could indicate the presence of outliers.

In other words, the assumptions of normality and linearity are both violated.

Plot #2 – Q-Q Residuals



Source: Barragan,R., Li,Y. (December, 2023) (own).

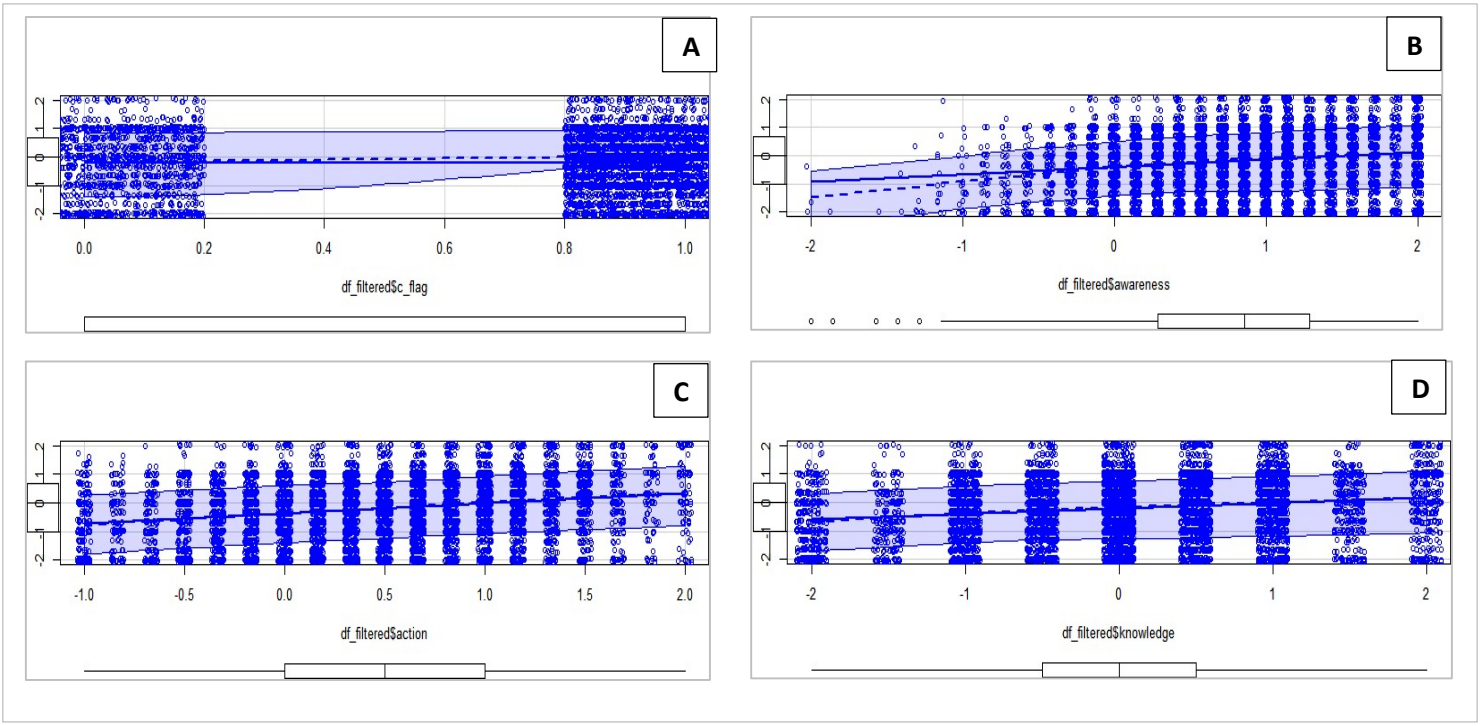
9.3. SCATTERPLOTS

The scatterplots of residuals against each independent variable, and against the fitted values of Willingness are displayed in plot #3. The assumptions of homoscedasticity are met for Willingness ~ action and Willingness ~ knowledge, as the line of best fit on these two plots is flat, and its residuals show a constant spread, no clear patterns, and randomness.

However, in plot #3.B (Willingness ~ Awareness), the line of best fit on this plot is not flat, indicating its heteroskedasticity.

For the plot #3.A, displaying the relationship between c_flag and Willingness, since c_flag is a binary factor, the interpretability of the scatter plot may be somewhat limited. Nonetheless, we observe that the average residual values for the two groups remain stable around the fitted values, suggesting that the model satisfies the assumption of homoscedasticity across different factor levels.

Plot #3 – Combo of Scatterplots of Residuals



Source: Barragan,R., Li,Y. (December, 2023) (own).

10. INTERACTIONS EFFECTS

Table #5 – Interaction effect: Awareness

```
> lm2 <- lm(willingness ~ c_flag + awareness + Caware, data = df_filtered)

================================================
                        Dependent variable:
                    ---------------------------
                              willingness
------------------------------------------------
c_flag                         -0.130**
                                (0.050)

awareness                      0.111***
                                (0.037)

Caware                         0.235***
                                (0.043)

Constant                       -0.307***
                                (0.046)

------------------------------------------------
Observations                     7,443
R2                               0.034
Adjusted R2                      0.034
Residual Std. Error       1.018 (df = 7439)
F Statistic           88.260*** (df = 3; 7439)
================================================
Note:                *p<0.1; **p<0.05; ***p<0.01
```

Source: Barragan,R., Li,Y. (December, 2023) (own).

The linear model #2 [Table #5] predicts Willingness (the dependent variable) based on three factors: C_flag, Awareness, and their interaction term called "Caware," which is the product of Awareness and c_flag.

For C_flag, the interaction coefficient is -0.130. This means that, while keeping the other variables constant, the level of Willingness is 0.130 units lower in developed countries compared to developing countries. In other words, the negative effect indicates that developed countries might have a lower willingness to protect the environment.

As for Awareness, the interaction coefficient is 0.111. This implies that a one-unit increase in Awareness corresponds to a 0.111-unit increase in Willingness. Now, regarding the interaction term "Caware," its coefficient is 0.235, which is a crucial finding. This positive and

significant coefficient indicates that the positive impact of Awareness on Willingness is more striking in developed countries. In other words, in developed countries, both the main effect of Awareness and the interaction effect contribute to a higher willingness.

In summary, the coefficients are -0.130 for C_flag, 0.111 for Awareness, and 0.235 for "Caware." All three coefficients are highly significant for the model, as indicated by p-values less than 0.05, 0.01, and 0.01, respectively.

Overall, even though developed countries may start with a lower baseline level of Willingness, their level of Awareness has a stronger positive influence. This suggests that efforts should be in the line to raise Awareness to protect the environment, resulting in a more pronounced effect on Willingness in developed nations.

Table #6 – Interaction effect: Action

```
> lm3 <- lm(willingness ~ c_flag + action + Cact, data = df_filtered)

================================================
                        Dependent variable:
                    ---------------------------
                            willingness
------------------------------------------------
c_flag                       -0.218***
                              (0.031)

action                       0.225***
                              (0.030)

Cact                         0.228***
                              (0.038)

Constant                     -0.252***
                              (0.024)

------------------------------------------------
Observations                    7,443
R2                              0.058
Adjusted R2                     0.057
Residual Std. Error       1.005 (df = 7439)
F Statistic          151.576*** (df = 3; 7439)
================================================
Note:                *p<0.1; **p<0.05; ***p<0.01
```

Source: Barragan,R., Li,Y. (December, 2023) (own).

The linear model #3 [Table #6] explores the factors influencing Willingness, assessing the impact of C_flag, Action, and their interaction term "Cact," which is the product of Action and C_flag.

For C_flag, the interaction coefficient is -0.218. This means that keeping the other variables constant, the level of Willingness is 0.218 units lower in developed countries. In other words, the negative effect indicates that developed countries might be less proactive in taking Action to protect the environment.

As for Action, the interaction coefficient is 0.225. This implies that a one-unit increase in Action corresponds to a 0.225 increase in Willingness. Now, regarding the interaction term "Cact," its coefficient is 0.228, which is a crucial finding. This positive and significant coefficient indicates that the positive impact of Action on Willingness is more striking in developed countries. In other words, in developed countries, both the main effect of Action and the Cact lead to a higher effect on Willingness (0.225 + 0.228 = 0.453).

Furthermore, the coefficients are as follows: -0.218 for C_flag, 0.225 for Action, and 0.228 for "Cact." All three coefficients are highly relevant for the model, as indicated by p-values lower than 0.01 for all three.

Overall, even though developed countries may start with a lower baseline level of Willingness, their level of Action has a stronger positive influence. This suggests that efforts should be in the line to raise proactiveness to take Action to protect the environment, resulting in a more pronounced effect on Willingness in developed nations.

Table #7 – Interaction effect: Knowledge

```
> lm4 <- lm(willingness ~ c_flag + knowledge + Cknow, data = df_filtered)

================================================
                        Dependent variable:
                    ---------------------------
                              willingness
------------------------------------------------
c_flag                         -0.299***
                                (0.088)

knowledge                      0.122***
                                (0.023)

Cknow                          0.091***
                                (0.028)

Constant                       -0.555***
                                (0.072)

------------------------------------------------
Observations                     7,443
R2                               0.028
Adjusted R2                      0.027
Residual Std. Error        1.021 (df = 7439)
F Statistic           70.527*** (df = 3; 7439)
================================================
Note:                *p<0.1; **p<0.05; ***p<0.01
```

Source: Barragan,R., Li,Y. (December, 2023) (own).

The linear model #4 [Table #7] explores the factors influencing Willingness, assessing the impact of C_flag, Knowledge, and their interaction term "CKnow," which is the product of Knowledge and C_flag.

For C_flag, the interaction coefficient is -0.299. This means that keeping the other variables constant, the level of Willingness is 0.299 units lower in developed countries. In other words, the negative effect indicates that developed countries might have a lower Willingness regarding environmental protection.

As for Knowledge, the interaction coefficient is 0.122. This implies that a one-unit increase in Knowledge corresponds to a 0.122-unit increase in Willingness. Now, regarding the interaction term "CKnow," its coefficient is 0.091. This positive and significant coefficient indicates that the positive impact of Knowledge on Willingness is higher in developed countries.

In other words, in developed countries, both the main effect of Knowledge and the CKnow lead to a higher effect on Willingness (0.122 + 0.091 = 0.0.213).

Furthermore, the coefficients are as follows: -0.299 for C_flag, 0.122 for Knowledge, and 0.091 for "CKnow." All three coefficients are highly relevant for the model, as indicated by p-values lower than 0.01 for all three.

Overall, even though developed countries may start with a lower baseline level of Willingness, their level of Knowledge has a stronger positive influence. This suggests that efforts should be made to raise Knowledge to protect the environment, resulting in a more pronounced effect on Willingness in developed nations.

R-Squared comparison

There is one last comparison that is worth commenting on among all three models, which is the R squared. Lm2 (Awareness) has an R2 equal to 0.034, lm3 (Action) has an R2 equal to 0.058, and lm4 (Knowledge) has an R2 equal to 0.028. All three R squares show the importance of design-tailored strategies that focus on fostering Awareness, Action, and Knowledge to increase Willingness levels.

However, among the three values, lm3 has the highest R-squared value (0.058) overall. This suggests that lm3, which includes Action and its interaction with C_flag, explains more of the variability in Willingness than the other two models. Therefore, among the three models, lm3 would be considered the "best" in terms of explaining the dependent variable based on the R-squared criterion.

11. CONCLUSIONS AND RECOMMENDATIONS

11.1. CONCLUSIONS

In the current global context, one of the major challenges facing society is protecting the environment and addressing issues related to climate change. However, in order to design a series of effective recommendations, this study has analyzed the factors (C-flag, Awareness, Action, and Knowledge) that most influence people's Willingness to contribute to the protection of the environment from an understanding of the variables, correlation analysis, linear regression models and interaction effects.

After all this, we conclude that the variable that presents a greater impact on Willingness is the Action variable.

- In section #7 (*Correlation Matrix*), we conclude that the variable "Action" has the highest correlation coefficient with "Willingness," being equal to 0.22, indicating a moderate positive correlation. Hence, people who are more willing to make a difference in environmental issues tend to be more actively involved in environmentally-related actions.

- In section #8 (*Linear Modeling*), we see that of the three coefficients of Awareness (0.199), Action (0.279), and Knowledge (0.133), the coefficient of Action has a higher weight and relevance to the model with a p-value of less than <0.01, making it statistically highly significant for the linear regression model. This means that a higher level of Knowledge is positively associated with higher Willingness.

- In section #10 (*Interaction Effect*), we have seen that lm3 (Action) is the model with the highest interaction coefficient, which is equal to 0.225. This positive and significant coefficient indicates that the positive impact of Action on Willingness is higher in developed countries). In addition, the combination of the main effect of Action and its interaction with the variable 'Cact' results in a significant increase in Willingness, adding up to a total effect of 0.453.

- On top of that, the lm3 also has the highest R-squared value (0.058) overall, which includes Action and its interaction with C_flag, explaining more of the variability in Willingness than the other two models.

All in all, it demonstrates the existence of a strong action-will synergy in developed countries. These findings underline the importance of designing personalized action-oriented strategies in developed nations.

11.2. RECOMMENDATIONS

Based on the conclusions, the recommendations for governments to promote "active action" and thereby increase the Willingness of citizens, especially in developed nations, to contribute to environmental conservation are the following:

I. **Create** incentive programs, like tax reductions, that target companies in the private sector, praising their proactivity in adopting sustainable actions or practices, such as using renewable energy, recycling, or energy saving.

II. **Increase** the level of investment in green infrastructures, for example, efficient public transportation systems, recycling facilities, or urban green spaces.

III. **Foster** citizen participation in environmental projects such as reforestation, beach and park clean-ups, and other environmental conservation initiatives. In this line, participation in volunteering programs and experiences promoted by environmental NGOs that work on environmental projects can also increase the action-will levels.

IV. **Develop** effective environmental education campaigns that proactively inform about the consequences of climate change and environmental degradation, including precise information on how people's proactive actions can make a significant difference.

12. REFERENCES

[1] Overview. World Meteorological Organization. (2023, December 7). https://wmo.int/about-wmo/overview

[2] Peralta, P. (2023, October 5). *2023 se perfila como el año Más Caluroso de la Historia, con septiembre con el calor más intenso*. France 24. https://www.france24.com/es/medio-ambiente/20231005-2023-se-perfila-como-el-a%C3%B1o-m%C3%A1s-caluroso-de-la-historia-con-septiembre-con-el-calor-m%C3%A1s-intenso

[3] Unep. (n.d.). *Facts about the climate emergency*. United Nations Environment Program. https://www.unep.org/facts-about-climate-emergency?gad_source=1&gclid=CjwKCAiA1fqrBhA1EiwAMU5m_yDkX_qzRjBPbxHdPc1zyX3ujueZ4AbvzPQq3uEw_LD4b_DRo1waQRoC4kcQAvD_BwE

[4] IPCC. (2023, August). *IPCC: Sixth Assessement Report*. IPCC Intergovernmental Panel on Climate Change. https://www.ipcc.ch/report/ar6/wg1/

[5] Department of Economic and Social Affairs. (n.d.). The 17 goals |sustainable development. United Nations. https://sdgs.un.org/goals

[6] *Climate change - united nations sustainable development*. United Nations. https://www.un.org/sustainabledevelopment/climate-change/

[7] United Nations. (2015, December 12). *The Paris Agreement*. United Nations. https://www.un.org/en/climatechange/paris-agreement

[8] Franzen, A., & Vogl, D. (2013). Two decades of measuring environmental attitudes: A comparative analysis of 33 countries. *Global Environmental Change*, *23*(5), 1001–1008. https://doi.org/10.1016/j.gloenvcha.2013.03.009

[9] GESIS. (n.d.). *Leibniz Institute for the Social Sciences*. GESIS. https://www.gesis.org/en/research

[10] GESIS. (n.d.). *The International Social Survey Programme*. ISSP - International Social Survey Programme – ISSP. https://www.gesis.org/en/issp/home

[11] GESIS. (2019). *Gesis - Leibniz Institute for the Social Sciences*. 2010 – ISSP. https://www.gesis.org/en/issp/modules/issp-modules-by-topic/environment/2010

[12] Survey Monkey. (n.d.). *Likert scale: Examples and how to use it*. Survey Monkey. https://www.surveymonkey.com/mp/likert-scale/

[13] University of Saskatchewan. (n.d.). *SDG spotlight: Goal 13*. USask Health Sciences. https://healthsciences.usask.ca/news-and-announcements/announcements/2020/nov-sdg-spotlight-goal-13.php

[14] *Using and interpreting Cronbach's alpha*. Using and Interpreting Cronbach's Alpha | UVA Library. (n.d.). https://library.virginia.edu/data/articles/using-and-interpreting-cronbachs-alpha

13. R SCRIPT

```
### set working directory

setwd()

getwd()


###Section #5. Data cleaning

setwd()

getwd()

data <- read.csv("ISSP_Environment_2010.csv")

nrow(data)

# total 50437 observations


install.packages("dplyr")

library(dplyr)


### Section 5.1 Observations

#only use the data from our chosen countries

df <- data %>%

 filter(c_alphan %in% c("US", "GB-GBN", "JP", "DE", "FR", "CA", "MX", "PH", "AR", "CL"))

nrow(df)

# 13713 left after filtering
```

```
df <- df %>%
  mutate(c_flag = ifelse(c_alphan %in% c("US", "GB-GBN", "JP", "DE", "FR", "CA"), 1, 0))
summary(df$c_flag) # 60.6% observations are from developed countries


variables <- c("c_alphan", "c_flag", "v18", "v19", "v29", "v30", "v31",
               "v39", "v40", "v41", "v42", "v43", "v44", "v45", "v55", "v56",
               "v57", "v58", "v59", "v60")
df <- na.omit(df[variables])
#since 8,9 = don't want to answer, we also omit this number
df_filtered <- df[!apply(df, 1, function(row) any(row %in% c(8, 9))), , drop = FALSE
]#now 7443 observations left


### Section 5.2. Reassignment of Values
#In order to reflect that respondents who express higher agreement also have
#higher levels of environmental concern, a reassignment of values was applied to
#the "Willingness to Contribute to Environmental Protection", "environment
#awareness measurement" and "Daily Environmental Actions" indicators in the questionnaire.
df_filtered$v29r <- 6 - df_filtered$v29
df_filtered$v30r <- 6 - df_filtered$v30
df_filtered$v31r <- 6 - df_filtered$v31
df_filtered$v39r <- 6 - df_filtered$v39
df_filtered$v40r <- 6 - df_filtered$v40
df_filtered$v41r <- 6 - df_filtered$v41
df_filtered$v42r <- 6 - df_filtered$v42
df_filtered$v43r <- 6 - df_filtered$v43
df_filtered$v44r <- 6 - df_filtered$v44
```

```
df_filtered$v45r <- 6 - df_filtered$v45
df_filtered$v55r <- 6 - df_filtered$v55
df_filtered$v56r <- 6 - df_filtered$v56
df_filtered$v57r <- 6 - df_filtered$v57
df_filtered$v58r <- 6 - df_filtered$v58
df_filtered$v59r <- 6 - df_filtered$v59
df_filtered$v60r <- 6 - df_filtered$v60

summary(df_filtered)

#center the variable
df_filtered$v18C <- df_filtered$v18 -3
df_filtered$v19C <- df_filtered$v19 -3
df_filtered$v29rC <- df_filtered$v29r -3
df_filtered$v30rC <- df_filtered$v30r -3
df_filtered$v31rC <- df_filtered$v31r -3
df_filtered$v39rC <- df_filtered$v39r -3
df_filtered$v40rC <- df_filtered$v40r -3
df_filtered$v41rC <- df_filtered$v41r -3
df_filtered$v42rC <- df_filtered$v42r -3
df_filtered$v43rC <- df_filtered$v43r -3
df_filtered$v44rC <- df_filtered$v44r -3
df_filtered$v45rC <- df_filtered$v45r -3
df_filtered$v55rC <- df_filtered$v55r -3
df_filtered$v56rC <- df_filtered$v56r -3
df_filtered$v57rC <- df_filtered$v57r -3
df_filtered$v58rC <- df_filtered$v58r -3
df_filtered$v59rC <- df_filtered$v59r -3
df_filtered$v60rC <- df_filtered$v60r -3
```

```
summary(df_filtered)


### Section 5.4. Cronbrach Alpha Test
#test the Cronbach's α. Cronbach’s alpha is a measure of internal consistency,
#assessing how well a set of items in a questionnaire or scale reliably measures
#a single underlying construct. It ranges from 0 to 1, with higher values
#indicating greater reliability. An ideal alpha value should be greater than 0.7.
install.packages("psych")
library(psych)

awareness <- c("v39rC", "v40rC", "v41rC", "v42rC", "v43rC", "v44rC", "v45rC")
result1 <- alpha(df_filtered[, awareness])
result1 # alpha = 0.81

willingness <- c("v29rC", "v30rC", "v31rC")
result2 <- alpha(df_filtered[, willingness])
result2 # alpha = 0.82

action <- c("v55rC", "v56rC", "v57rC", "v58rC", "v59rC", "v60rC")
result3 <- alpha(df_filtered[, action])
result3 # alpha = 0.79

knowledge <- c("v18C", "v19C")
result4 <- alpha(df_filtered[, knowledge])
result4 # alpha = 0.77

#### create IV & DV
df_filtered$awareness <- rowMeans(df_filtered[,awareness])
df_filtered$willingness <- rowMeans(df_filtered[,willingness])
```

```
df_filtered$action <- rowMeans(df_filtered[,action])
df_filtered$knowledge <- rowMeans(df_filtered[,knowledge])
##################analysis start!############################
### Section 6. Descriptive statistics
describe(df_filtered)
summary(df_filtered)

### Section 7. Correlation matrix
vars <- c("c_flag", "awareness", "willingness", "action", "knowledge")
cor_matrix <- cor(df_filtered[, vars])
cor_matrix

install.packages("corrplot")
library(corrplot)
corrplot(cor_matrix, method = "color", type = "upper", order = "hclust", tl.cex = 0.7)

corrplot(cor_matrix,
         method = "number",
         type = "full",
         tl.cex = 0.7,
         tl.col = "black",
         addCoef.col = "black",
         col = colorRampPalette(c("lightblue", "blue", "red"))(50),
         tl.pos = "lt",
         number.cex = 1.5)

### Section 8. Linear Regression Analysis
library(stargazer)
lm1 <- lm(willingness ~ c_flag + awareness + action + knowledge, data = df_filtered)
stargazer::stargazer(lm1, type="text")
```

```
### Section 9. Assumptions
library(ggplot2)
library(car)
plot(lm1, which = 1)
plot(lm1, which = 2)

scatterplot(fitted(lm(willingness ~ c_flag, data = df_filtered)),
        resid(lm(willingness ~ c_flag, data = df_filtered)))
scatterplot(fitted(lm(willingness ~ awareness, data = df_filtered)),
        resid(lm(willingness ~ awareness, data = df_filtered)))
scatterplot(fitted(lm(willingness ~ action, data = df_filtered)),
        resid(lm(willingness ~ action, data = df_filtered)))
scatterplot(fitted(lm(willingness ~ knowledge, data = df_filtered)),
        resid(lm(willingness ~ knowledge, data = df_filtered)))

scatterplot(df_filtered$awareness, df_filtered$willingness, jitter = list(x = 1, y = 1))
scatterplot(df_filtered$c_flag, df_filtered$willingness, jitter = list(x = 1, y = 1))
scatterplot(df_filtered$action, df_filtered$willingness, jitter = list(x = 1, y = 1))
scatterplot(df_filtered$knowledge, df_filtered$willingness, jitter = list(x = 1, y = 1))


### Section 10. Interactions effects
df_filtered$Caware <- df_filtered$awareness * df_filtered$c_flag
df_filtered$Cact <- df_filtered$action * df_filtered$c_flag
df_filtered$Cknow <- df_filtered$knowledge * df_filtered$c_flag

lm2 <- lm(willingness ~ c_flag + awareness + Caware, data = df_filtered)
stargazer(lm2, type = "text")
```

```
lm3 <- lm(willingness ~ c_flag + action + Cact, data = df_filtered)
stargazer(lm3, type = "text")

lm4 <- lm(willingness ~ c_flag + knowledge + Cknow, data = df_filtered)
stargazer(lm4, type = "text")
```