LUND UNIVERSITY
School of Economics and Management

Department of Economics

Human Rights Violations and Machine Learning

Cluster Analysis of Countries using the CIRIGHTS Data set

Master Essay in Data Analytics and Business Economics (DABN01, 15 credits)

Authors: Emilie Haubrich and Emilie Stahle

Supervisor: Simon Reese

May 2023

Abstract

This master's thesis explores the use of unsupervised machine learning techniques to cluster countries based on their degree of human rights violations. Accordingly, the study evaluates the performance of two clustering methods, K-Means clustering and Latent Class Analysis (LCA), using two cluster validation metrics (Silhouette Coefficient and Dunn Index), as well as an Accuracy measure using the Human Rights index. It analyses the characteristics of clusters and the assignments thereof over four decades to provide compact insights for policymakers. The results, in turn, show that both clustering methods perform equally well, however, LCA is chosen for the bulk of the analysis out of respect for the categorical nature of the data. Consequently, cluster profiling identifies three clusters with varying levels of human rights scores, although, looking at each variable and decade individually, we see that they do not all follow the same order of magnitude that the overall cluster scores suggest. Furthermore, the probability transition matrix shows that, generally, countries do not change significantly over time, in terms of their level of respect for human rights. Finally, policy advice for stable countries involves using cluster 1 as a "gold standard", incentivizing cluster 2, and taking a proactive approach for cluster 3. In turn, for unstable countries, advice includes incentivizing further improvements for countries that have shown positive progress, understanding reasons for decline, and stabilising and monitoring closely those that have shown fluctuating tendencies. The paper concludes that unsupervised machine learning for detecting human rights violations is useful, efficient, and provides insights into patterns that are not immediately apparent. Furthermore, it is a useful instrument to summarise these patterns in a clear and interpretable way.

Keywords: Human Rights Violation, Clustering, K-Means, Latent Class Analysis, Transition Matrix.

Table of Contents

List of Figures

List of Tables

1. Introduction

The protection of human rights is a crucial issue for all nations around the world, as it ensures that individuals are treated with dignity, respect, and fairness. However, the extent to which human rights are respected and upheld varies significantly across continents (Adewusi & Kocadal, 2022). While there are international human rights regimes in place, such as treaties, laws, and human rights institutions, their effectiveness in promoting human rights protection remains a matter of debate (Guzman & Linos, 2013).

Unfortunately, human rights violations continue to occur globally, with more countries experiencing violations than those where they are effectively protected (Robertson & Merrills, 1996). To address this issue, measuring human rights becomes important to understand their global variation and to find solutions to improve their protection in the future (Landman, 2004).

Empirical research on human rights has traditionally relied on qualitative research methods and expert opinions to identify patterns of human rights violation (Poe, Tate & Keith, 1999; Gartner & Regan, 1996; Adewusi & Kocadal, 2022; Grewal & Voeten, 2015). However, the use of new methods can provide more accurate and comprehensive insights into patterns of such matters (Emerson, Satterthwaite & Pandey, 2018).

This paper aims to explore the use of unsupervised machine learning techniques to cluster countries based on their degree of human rights violations, identifying patterns and trends across different regions and countries. Hence, the thesis also aims to evaluate the performance of these clustering methods, using two internal cluster validation metrics, the Silhouette Coefficient and the Dunn Index, to measure cluster structure and stability over time. The goal of the analysis is to identify larger groups of countries and to assess their trends over four decades to provide insights for policymakers, a matter of great importance in their efforts to address human rights violations. Furthermore, the thesis will explore the use of data visualisation and other visual features in human rights communication and advocacy to enable policymakers to make data-based decisions (Rall, Satterthwaite, Pandey, Emerson, Boy, Nov & Bertini, 2016).

The motivation for this research is the need to fill a research gap by conducting a comprehensive analysis of the patterns and trends of human rights violations in the CIRIGHTS data set (Cingranelli, Richards & Clay, 2021), using the machine learning algorithms K-Means clustering and Latent Class Analysis (LCA). Furthermore, using clustering analysis allows us to summarise patterns of a large number of disaggregated human rights violation indices in a clear and interpretable manner. Specifically, our work will contribute to this field by condensing these topic-specific measures into a simple framework, making it informative and easily conveyed to policymakers Hence, the research questions guiding this study include:

1. Which clustering method (K-Means or LCA) performs better in terms of cluster structure and stability over time?

2. What are the characteristics (i.e., feature values) of each cluster and how do these characteristics change over time?

3. How do cluster assignments change over four decades and how can these trends be communicated to policymakers to help guide their decisions?

The thesis is structured as follows: in order to introduce and describe the related scientific contributions, section 2 contains a literature review focusing on human rights and clustering algorithms. In section 3, we focus on the data source, pre-processing, and the exploratory data analysis, where we also give a short description of the selected features. The theoretical part, which focuses on machine learning for unsupervised clustering and cluster validity methods, is described in section 4. Then, we move onto the empirical analysis in section 5, where we describe the methodology, compare the algorithms (LCA and K-means), and then analyse the LCA algorithm further in terms of cluster profiling, transition probabilities, and feature development trends over four decades. Finally, our summary and implications, limitations, and future work are presented in section 6.

2. Literature Review

Recent empirical research on the prevalence of human rights violations often relies on the Cingranelli-Richards Human Rights Dataset (CIRIGHTS). The article by Cingranelli, Richards and Clay (2021) introduces the aforementioned data set, which is a collection of data on human rights practices across different countries of the world. The dataset is publicly available online through Dataverse, and it covers a wide range of human rights topics, including civil liberties, political rights, torture, extrajudicial killings, and disappearances. The data was collected using a standardised coding system, and the article provides information on the methodology and sources used to compile the data set (Cingranelli, Richards & Clay, 2021).

The CIRIGHTS dataset has already been exploited for research that focused on specific human rights violations, such as in Cingranelli, Mark, Gibney, Haschke, Wood and Arnon (2019), who conducted comparative cross-country data analysis using econometric methods to determine whether human rights violations are on the rise. The authors also examined whether human rights abuses make violent internal conflicts, such as terrorism, civil wars, and protests more likely. Accordingly, the study showed that governments that engage in human rights violations, such as torture, political imprisonment, and killings, are more likely to cause such violent internal conflicts.

Furthermore, Roser (2016) is committed to making human rights measurable and to alerting policymakers about such matters on the *Our World in Data* homepage. Namely, he attempts to provide an overall measure of human rights, in contrast to the topic-specific indicators in the CIRIGHTS data set. To this end, Roser uses the so-called *V-Dem's Civil Liberties Index scores,* which rates each country on a spectrum, with some countries protecting human rights more than others. Correspondingly, the spectrum ranges from 0 ('fewest rights') to 1 ('most rights'). Furthermore, V-Dem covers 202 countries going back to 1789 and uses expert knowledge on each country, year, and topic to make an assessment of the index, which takes place through a survey every year. In this survey, experts answer very specific questions on entirely explained scales about characteristics of human rights, in order to avoid subjective impressions and to make the data transparent and comparable. Therefore, this homepage gives us another benchmark when testing the CIRIGHTS index and clustering.

In the article by Cope, Crabtree and Fariss (2020) patterns of disagreement were discussed as indicators of state repression. That is, the authors argue that there is significant disagreement among existing measures of state repression, which can lead to discrepancies in research findings. Specifically, they compared the Cingranelli-Richards (CIRIGHTS) index with the Varieties of Democracy (V-Dem) index. Consequently, they propose a method for identifying patterns of disagreement among different indicators of state repression, using a latent variable model, which they apply to a sample of countries. Hence, the results show that there are distinct patterns of disagreement among the indicators, which the article highlights as being important to consider when analysing state repression (Cope, Crabtree & Fariss, 2020).

Therefore, it is important to consider that the findings in this thesis may only be applicable to the CIRIGHTS data set.

To the best of our knowledge, the CIRIGHTS dataset has not yet been used for cluster analysis. However, a number of studies employ clustering algorithms on different data sets and for different research questions, in a way that inspires the investigation in this thesis. Therefore, for our study, we consider the K-Means algorithm and Latent Class Analysis.

The K-Means algorithm, the oldest partitional method and one of the most popular, has been studied widely with various extensions in the literature and applied to a wide range of areas (Jain & Dubes, 1988). Therefore, it can be seen as the most used model for comparison, when it comes to clustering (Lin & Ng, 2012; Lv, Liu, Shi, Benediktsson & Du, 2019).

The difficulties in estimating panel data models with parameter heterogeneity when group membership is uncertain are discussed by Lin and Ng (2012). Against this background, it is interesting to note that the article mentions the effectiveness of conditional group clustering and the evaluation of K-Means, both of which may be used to conduct an analysis of different decades in the CIRIGHTS dataset.

Considering the nature of the CIRIGHTS data set, we are going to perform Latent Class Analysis (LCA), as a second algorithm. LCA is used for data sets that consist of categorical variables and assume to have, within each group, variables independent of one another (Lazarsfeld & Henry, 1968; Goodman, 1974). Furthermore, LCA is most popularly used in social and psychological studies (Nylund, Bellmore, Nishina & Graham, 2007), in health and clinical research (Pence, Miller & Gaynes, 2009), as well as in educational research (Roussos, Templin & Henson, 2007; Xu, 2011).

Using LCA in the research by Nosetti et al. (2020), the analysis of data related to paediatricians treating sleep-disordered breathing in Italy revealed underlying response patterns, allowing the identification of clusters of respondents with similar awareness, attitude, practice, and satisfaction. The characteristics of the two recognised groups were then further analysed and used to develop new specific concepts. In other words, Nosetti et al. (2020) provide decisive information on how education and training may be designed for different clusters. This makes the article particularly important for our research, as it illustrates what the use of LCA may look like in practice and how it was presented.

The research paper by Magidson and Vermouth (2022), which compares K-Means with LCA in a supervised learning setting (i.e., labelled classes), shows that LCA outperforms the K-Means algorithm. Using discriminant analysis, those results can be seen as the gold standard in determining the best possible outcome of each clustering technique. Hence, it gives us an indication of which algorithm should be used and may perform better.

In the dissertation by Xu (2011), LCA was compared to the K-Means algorithm with three kinds of preprocessed data. The three resulting data sets were thus used in comparing the two aforementioned algorithms and did so by considering clustering results, i.e. intra- and inter-cluster distances, Davies-Bouldin index, and cluster evolution. The results in turn show that LCA surpasses the K-Means approach in detecting clusters and linkages within them, as well as in obtaining consistent clustering results in different scenarios. Thus, this paper provides insight into the importance of using inter-cluster or intra-cluster distances to evaluate the cluster solution.

3. Data

To cluster the world with a focus on human rights violations and identify trends, we use data from the CIRIGHTS Human Rights Dataset (Cingranelli, Richards & Clay, 2021). The reason for selecting this data collection is that it offers quantitative measures for the degree to which every country on Earth upholds certain types of human rights and is the largest of its kind (Cingranelli, Richards & Clay, 2021). Accordingly, it contains data on more than 72 internationally recognized human rights and 6251 observations, which provide information on how human rights are assessed categorically in each country and year (1981-2020).

After accessing the data set, the first step in preparing it was to remove variables containing more than 50% of missing values, after which we proceeded to remove the remainder of missing values. Next, we removed countries that do not exist anymore, i.e., Yugoslavia, Czechoslovakia, and renamed others for the purpose of clarity, i.e., South/North Korea. Finally, we converted our "year" variable from integer to string and dropped country identifier variables, apart from country name and initials.

The variables used in our dataset, excluding variables contained in the original data set but omitted from ours, can be separated into three categories of violations, as done so in the initial CIRIGHTS project (see Appendix Table A1 for how the variables are coded) (Cingranelli, Richards & Clay, 2021). Furthermore, it is important to note that despite being numerically coded, the measures below are all categorical and not continuous, generally 0-1-2.

Physical integrity rights:

1) "disap": represents **disappearances** where agents of the state are likely responsible.

2) "kill": represents **extrajudicial killings** in the sense of killings by government officials without due process of law.

3) "polpris": represents the **political imprisonment** of people by government officials because of e.g., political, religious or other beliefs.

4) "tort": represents **torture** in the sense of purposefully inflicting extreme pain (includes rape and beatings).

5) "physint_sum": represents the additional **physical integrity rights index**, constructed from the torture, extrajudicial killings, political imprisonment, and disappearance indicators.

6) **brutality-based mass atrocity**: represents the "widespread extrajudicial killing of non-combatant members of society…". This is separated into two variables in our data set: "bbatrocity" and "bbatrocity_intensity"

Empowerment rights and freedom:

1) "assn": represents the right to **freedom of assembly and association** with other people in e.g., political parties or trade unions.

2) "formov": represents **freedom of foreign movement** and travel (i.e., to *another* country).

3) "dommov": represents **freedom of domestic movement** (i.e., *within* one's country).

4) "speech": represents **freedom of speech** and press in the sense of government censorship/ownership of the media.

5) "elecsd": represents the right to freely determine one's own political system and leadership, known as **electoral self-determination.**

6) "rel_free": represents **freedom of religion** in the sense of exercising and practising one's religious beliefs.

7) "wecon": represents **women's economic rights** including, inter alia, equal pay for equal work, free choice of profession without husband's consent, job security, no discrimination by employers etc.

8) "wopol": represents **women's political rights** including, inter alia, the right to vote, the right to run for political office etc.

Justice rights:

1) "injud": represents **independence of the judiciary** in the sense of independence of control from other sources, such as a government branch or the military.

The distribution of variable values across years and countries, as illustrated in Figure 1 below, reveals important stylized facts. In particular, 75% of the variables considered for our study (i.e., 12 out of 16) are skewed towards values that represent good scores in terms of violations, i.e., a higher score (apart from the brutality-based mass atrocity variables, in which case a low score is better). There are four variables that do not follow the same distribution: namely, torture which is skewed towards bad scores and freedom of speech, women's economic rights, and women's political rights, whose observations are more concentrated in the middle region (neither the best nor the worst score). More precisely, women's economic rights have a slight tendency towards "bad" values and women's political rights have a tendency towards "good" values.

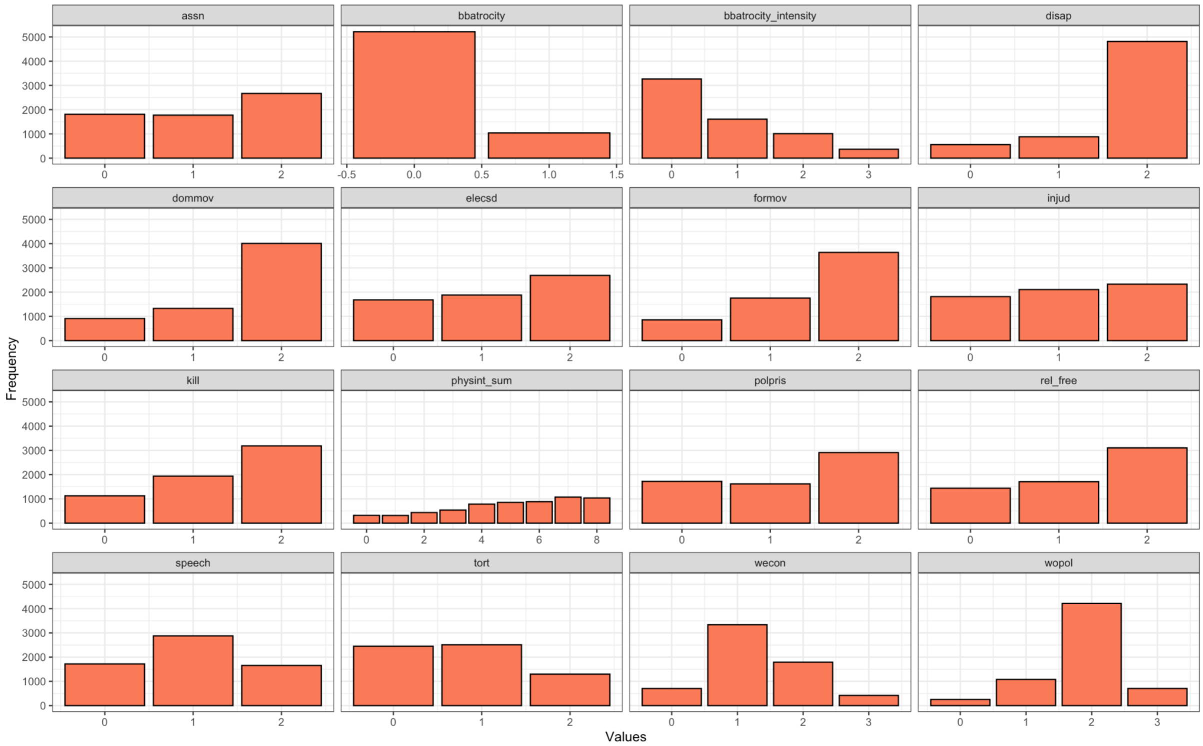


Figure 1. Distribution of each variable

In addition to the distribution plot above, we take a deeper look at descriptive statistics, which reveal additional, more detailed, information about each variable.

Table 1. Descriptive statistics for physical integrity rights

	mean	st. dev.	median	min	max	skew	kurtosis
physint_sum	5.02	2.35	5	0	8	-0.51	-0.73
polpris	1.19	0.84	1	0	2	-0.37	-1.48
disap	1.68	0.63	2	0	2	-1.78	1.78
tort	0.82	0.75	1	0	2	0.32	-1.18
kill	1.33	0.76	2	0	2	-0.63	-1.02
bbatrocity	0.17	0.37	0	0	1	1.79	1.21
bbatrocity_intensity	0.76	0.93	0	0	3	0.94	-0.23

As mentioned previously, the only variable within the physical integrity rights to follow a different trend is "torture", which can once again be confirmed by looking at the mean value (Table 1). That is, considering that all variables, apart from the brutality-based mass atrocity variables, have values that can be interpreted in the same way (i.e., the higher the score, the better), torture has a much lower value, indicating that countries are generally worse at respecting this type of human right. Furthermore, the variable with the highest degree of skewness (left-skew) is “disappearances”, which belongs to the physical integrity rights, indicating that there is a much larger number of countries or years with no reported disappearances compared to occasionally or frequently occurring disappearances.

Table 2. Descriptive statistics for empowerment rights and freedom

	mean	st. dev.	median	min	max	skew	kurtosis
speech	0.99	0.73	1	0	2	0.02	-1.15
rel_free	1.27	0.81	1	0	2	-0.52	-1.29
dommov	1.50	0.74	2	0	2	-1.08	-0.33
formov	1.45	0.72	2	0	2	-0.90	-0.56
assn	1.14	0.84	1	0	2	-0.26	-1.52
elecsd	1.16	0.82	1	0	2	-0.31	-1.45
wopol	1.86	0.65	2	0	3	-0.72	1.20
wecon	1.31	0.76	1	0	3	0.35	-0.10

Freedom of speech, which belongs to empowerment rights and freedom, has perhaps the most symmetric distribution out of all variables in our data set. That is, it has an almost null skew and its mean and median are almost identical (Table 2). Looking at the distribution plot described previously, one can see that most of the observations lie in the centre, with a very similar number of countries or years with either complete censorship/ownership of the media or, on the contrary, complete freedom in that regard.

Table 3. Descriptive statistics for justice rights

	mean	st. dev.	median	min	max	skew	kurtosis
injud	1.08	0.81	1	0	2	-0.15	-1.46

The sole variable included in our analysis that belongs to the justice rights category is the independence of the judiciary (injud). Looking at the descriptive statistics above, this variable seems to be fairly evenly distributed across all three possible values (Table 3). That is, with a mean lying around 1 and a standard deviation that also tends towards 1 (approximately ± 1 from a score of 1), this indicates that both a score of 0 and 2 have fairly equal weight in terms of assigned observations).

Finally, looking at the variables altogether, we can confirm that the entirety of the variables' range has been covered. That is, each score that is possible to assign to a country or a year has been assigned at least once.

Looking at Figure 2 below, it is evident that, overall, there is a reasonable amount of (positive) correlation between all variables, apart from wopol, which seems to be the least correlated with all other variables. The brutality-based mass atrocity variables, bbatrocity, and bbatrocity_intensity, are fairly strongly correlated (even more so for the latter variable) with all other variables belonging to the physical integrity rights, which makes sense because they were constructed from these indices. Moreover, the negative correlation does not come as a surprise, since, as mentioned earlier they move in opposite directions. That is, for the aforementioned variables, a high score is bad, whereas, for the other variables within that group, a high score is good. Furthermore, albeit only slightly, there is a fairly negative correlation with the other variables belonging to the remaining two groups. Another interesting observation is the fact that between bbatrocity and bbatrocity_intensity, the latter consistently seems to be more strongly correlated with the other variables in our data set. On another note, we observe that physint_sum is strongly positively correlated with the variables from which it was constructed, which seems logical.

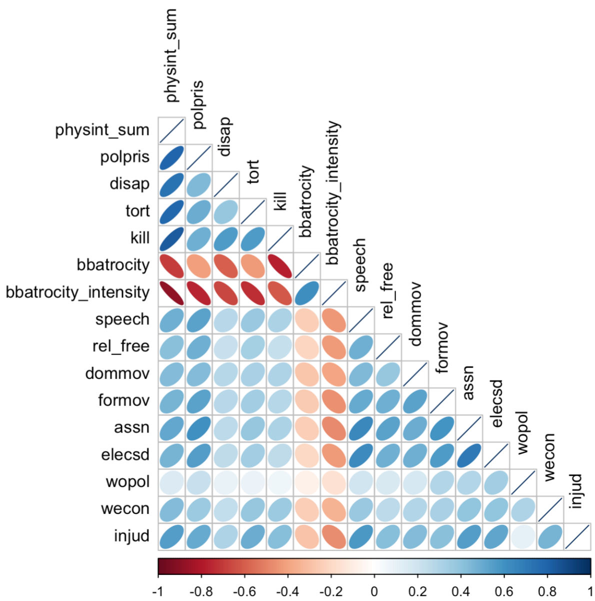


Figure 2. Correlation matrix of CIRIGHTS data set

4. Theoretical Analysis

4.1 Machine Learning for Unsupervised Clustering

Within the framework of our thesis, we take a deeper look at unsupervised learning, as the CIRIGHTS data set we are using falls within the scope of unlabelled data. As manual labelling can be arbitrary and hence difficult to justify, unsupervised learning gives us the opportunity to find hidden structures and enables a technique that seeks to summarise and explain key features. Therefore, clustering is a data-driven way of identifying groups for the CIRIGHTS data set, whose labels can be defined in a second step (LeCun, Bengio & Hinton, 2015). Furthermore, clustering can play an important role in analysing big data sets and grouping big data in a compact format that remains informative to outstanding persons and is easily conveyed to, in our case, policymakers (Fahad et al., 2014).

Accordingly, the algorithms selected for our analysis are K-Means, which we decide to employ because of the numeric data we are presented with, as well as LCA which we use in order to keep the, albeit numeric, categorical nature of our data intact.

The first clustering algorithm we will present that we have used to identify country groupings is K-Means, where "K" stands for the pre-specified number of clusters. The main difference between this algorithm and a Gaussian Mixture Model, a common alternative algorithm that we have not included in our analysis, is that K-Means makes "hard" cluster assignments as opposed to building a probabilistic model (Lindholm, 2022). That is, assigning, in our case, a country to a cluster depending on a certain probability of belonging to this cluster. Unlike this, the main objective of K-Means is to group data points (i.e., countries) together based on their degree of similarity, measured with, more specifically, the pairwise squared Euclidean distances.

$$\arg \min_{R_1, R_2, \ldots, R_K} \sum_{k=1}^{K} \frac{1}{|R_k|} \sum_{\mathbf{x}, \mathbf{x}' \in R_k} \|\mathbf{x} - \mathbf{x}'\|_2^2$$

More specifically, the algorithm initiates the iterative process by randomly selecting K cluster centres, "centroids", $\hat{\mu}_1, \hat{\mu}_2, \cdots, \hat{\mu}_K$ and adapts their position until the distances between them and their respective data points are minimised.

$$\arg \min_{R_1, R_2, \ldots, R_K} \sum_{k=1}^{K} \sum_{\mathbf{x} \in R_k} \|\mathbf{x} - \widehat{\boldsymbol{\mu}}_k\|_2^2$$

When it comes to choosing the number of clusters, "K", one of the most commonly-used methods is the so-called "Elbow method" (Figure 3) and it essentially amounts to making a subjective, "visual" judgement. That is, one runs the algorithm with a range of values for K and plots the objective function, i.e., the intra-cluster distance defined above, graphically with each respective K. The optimal number of clusters will then correspond to the point on the graph where the curve seems to flatten, i.e., the decrease in intra-cluster distance appears to diminish, and further increasing the number of clusters is insignificant (Lindholm, 2022).

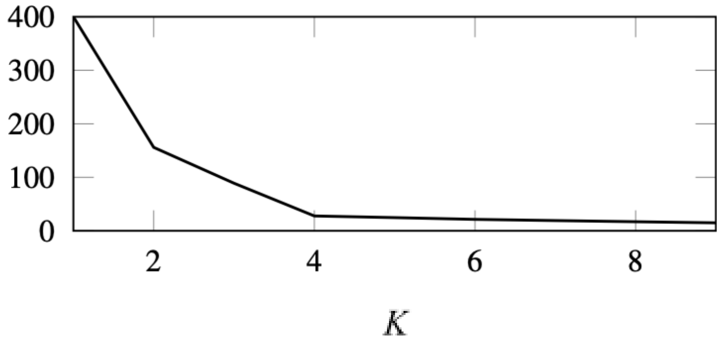


Figure 3. Elbow Curve (Lindholm, 2022)

The second algorithm considered for this thesis, as a special case of a finite mixture model, is the Latent Class Analysis (LCA), an algorithm used to analyse complex data sets with categorical variables. It is similar to cluster analysis, as it examines class membership and the probability of a given data point belonging to each group (Hagenaars & McCutcheon, 2002; Linzer & Lewis, 2011).

The critical implication of LCA is that the clusters or latent classes identified by the analysis will differ from each other in terms of the patterns of response probabilities across the observed variables. In other words, the clusters will be distinct from each other because they represent different combinations of response probabilities across the observed variables (Lazarsfeld & Henry, 1968; Linzer & Lewis, 2011).

For ease of understanding the following section, one may refer to the following legend:

R	Number of classes (fixed in advance)
J	Number of categorical indicator variables
K_j	Number of possible values for each variable
N	Number of observations
Y_{ijk}	Potential outcomes of the J variables: Transformation from original J categorical variables into $\sum_{j=1}^{J} K_j$ indicator variables for individual categories
p_r	Unconditional prior probability of individual being in a class
π_{jrk}	Class conditional probability that the k*th* result on the j*th* variable is produced by an observation in class r = 1,...,R

In the latent class model, the class conditional probability π_{jrk} is formally represented as a vector containing all conditional class probabilities for a given class r:

$$f(Y_i; \pi_r) = \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}$$

In addition, the probability distribution function across all classes is the weighted sum seen as follows,

$$\mathrm{P}(Y_i \mid \pi, p) = \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}$$

where p_r and π_{jrk} are estimated from the data.

Considering that the clustering algorithm is based on a probabilistic model of the observed data, it is appropriate to use maximum likelihood for estimation. However, there are also unobserved components, namely, cluster memberships. For this reason, LCA employs the expectation-maximisation (EM) algorithm, which is designed for models where some characteristics of the data are latent. Hence, the EM algorithm is used to update the parameters by computing an expected log-likelihood (see below) and then maximising the expectation (Lazarsfeld & Henry, 1968; Hagenaars & McCutcheon, 2002; Linzer& Lewis, 2011).

$$\log L = \sum_{i=1}^{N} \ln \sum_{r=1}^{R} p_r \prod_{j=1}^{J} \prod_{k=1}^{K_j} (\pi_{jrk})^{Y_{ijk}}$$

1. **Expectation step:** The “missing” class membership probabilities are calculated while substituting in $\hat{p}_r^{old}$ and $\hat{\pi}_{jrk}^{old}$(the initial values of $\hat{p}_r$ and $\hat{\pi}_{jrk}$, respectively). This step is formally stated as:

$$\widehat{\mathrm{P}}(r_i \mid Y_i) = \frac{\hat{p}_r f(Y_i; \hat{\pi}_r)}{\sum_{q=1}^{R} \hat{p}_q f(Y_i; \hat{\pi}_q)}$$

2. **Maximisation step:** The parameter estimates are updated by maximising the log-likelihood function, given the posterior $\widehat{P}(r_i \mid Y_i)$, with

$$\hat{p}_r^{new} = \frac{1}{N} \sum_{i=1}^{N} \widehat{\mathrm{P}}(r_i \mid Y_i)$$

as the new prior probabilities and

$$\hat{\pi}_{jr}^{new} = \frac{\sum_{i=1}^{N} Y_{ij} \widehat{\mathrm{P}}(r_i \mid Y_i)}{\sum_{i=1}^{N} \widehat{\mathrm{P}}(r_i \mid Y_i)}$$

as the new class-conditional outcome probabilities. These probabilities denote the probability of the observation belonging to each of the classes, where the class with the highest posterior probability is the one to which the observation is assigned (Hagenaars & McCutcheon, 2002; Linzer & Lewis, 2011).

3. The algorithm replicates these steps until the overall log-likelihood reaches a maximum and no longer increases from an arbitrarily defined small value.

To choose the optimal number of clusters for LCA, one can consider the Akaike information and Bayesian information criteria (AIC and BIC). The BIC is a criterion for model selection, where models with smaller BIC are preferred (Schwarz, 1978; Findley, 1991), and is formally stated as follows:

$$\mathrm{BIC} = k \ln(n) - 2 \ln(\widehat{L})$$

Here, $\widehat{L}$ is the maximised value of the likelihood function, x is the observed data, n the number of data points in x, and k the number of parameters estimated.

Furthermore, BIC is closely related to AIC, which is an estimator of prediction error and estimates the quality of each model, relative to each of the other models (Akaike, 1974), and is formally stated as follows:

$$\mathrm{AIC} = 2k - 2 \ln(\hat{L})$$

4.2 Cluster Validity Methods

Difficulties may arise when applying clustering techniques and particularly when trying to evaluate which clustering method is the most suitable. Hence, when it comes to unlabeled data, there is no easy evaluation method, seeing as no mistake signal can be implemented to evaluate a potential solution (Shirkhorshidi, Aghabozorgi, Wah & Herawan, 2014; Nguyen, Dinh, Sriboonchitta & Huynh, 2019).

Therefore, choosing the right metric for evaluating the performance of an algorithm is crucial, as it determines the validity of the results and the extent to which they can be relied upon. The literature thus provides various metrics that can be used to evaluate the performance. Namely, Liu, Li, Xiong, Gao and Wu (2010) recommend using *Compactness* to evaluate the performance of an algorithm, a metric that evaluates the intra-cluster distances between data objects in the same cluster, based on a similarity measure. Furthermore, the compactness and the degree of separation of clusters can be described by the *Dunn Index*, where a high score indicates good clustering (Dunn, 1974; Ben, Hamza & Bouaguel, 2021). Considering intrinsic measures, which do not require truth labels, as in the case of the CIRIGHTS data set, one can evaluate clustering performance using the Silhouette Coefficient (Chen & Gopalakrishnan, 1998; Rousseeuw, 1987; Alexander, Alexander, Barkhof & Denaxas, 2021) or the Davies-Bouldin Index (Davies & Bouldin, 1979).

In order to narrow down the scope of our clustering quality evaluation, we have decided to use two of the aforementioned indices, namely, the Silhouette Coefficient and Dunn Index. This, in turn, helps us decide which algorithm to focus on in the empirical analysis later on.

The first cluster validation method we want to introduce is the *Silhouette Coefficient*, which is a widely accepted metric for evaluating the quality of a clustering technique. This measure ranges from -1 to 1, and provides an indication of the goodness of the identified clusters, where a score of 1 may be an indication of “good” clustering and a score of -1, “misclassification” (Rousseeuw, 1987).

The Silhouette Coefficient assigns a score to each observation and aggregates the scores at a cluster level and at the dataset level, for a *dissimilarity d* (often selected to be the squared Euclidean distance) and a *point x* in a *cluster C*. The scores are defined as *b(x)* and *a(x)*, where the former is the average distance between all clusters and the latter is the average intra-cluster distance, or the average distance between any two locations inside a cluster (see Figure 4).

$$b(\mathbf{x}) = \min_{C' \neq C} \{ \mathrm{Avg}_{\mathbf{y} \in C'} \{ d(\mathbf{x}, \mathbf{y}) \} \}, \quad a(\mathbf{x}) = \mathrm{Avg}_{\mathbf{w} \in \mathbf{w_x}} \{ d(\mathbf{x}, \mathbf{w}) \}$$

$$\text{silhouette } (\mathbf{x}) = \frac{b(\mathbf{x}) - a(\mathbf{x})}{\max\{a(\mathbf{x}), b(\mathbf{x})\}} = \begin{cases} 1 - a(\mathbf{x})/b(\mathbf{x}) & \text{if } a(\mathbf{x}) < b(\mathbf{x}) \\ 0 & \text{if } a(\mathbf{x}) = b(\mathbf{x}) \\ b(\mathbf{x})/a(\mathbf{x}) - 1 & \text{if } a(\mathbf{x}) > b(\mathbf{x}) \end{cases}$$

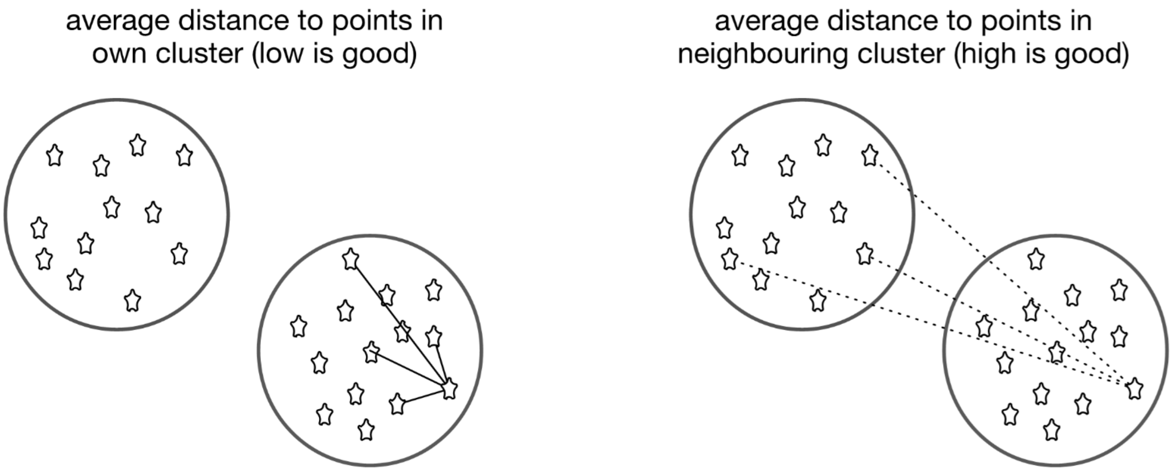


Figure 4. Silhouette Coefficient Visualisation (Boily, n.d.)

The second validation measure is the *Dunn index,* which is a measure of cluster compactness and separation. This measure has a range between *zero and infinity* and is *maximised* to achieve optimal results. The Dunn index is especially useful in evaluating largely unstructured, freely-formatted data that cannot be easily evaluated by other means (Dunn, 1974; Ben, Hamza & Bouaguel, 2021).

In order to calculate the Dunn index, each cluster's objects are first examined, where the distance between each object in the cluster and the objects in the other clusters is computed (see Figure 5). The minimum of this distance is taken as the *inter*-cluster separation, while the *intra*-cluster compactness is then determined by the maximal distance within each cluster, which is the maximum diameter of the cluster (Dunn, 1974; Ben, Hamza & Bouaguel, 2021).

$$DI_m = \frac{\min_{1 \leqslant i < j \leqslant m} \delta(C_i, C_j)}{\max_{1 \leqslant k \leqslant m} \Delta_k}$$

The Dunn index is obtained by dividing the shortest distance between clusters by the largest diameter within a cluster, once these two values (Ci and Cj) have been recorded. In Figure 5, the relationship between the shortest distance between points in different clusters (left side) and the largest distance within a cluster (right side) is quantified using the Dunn index. Accordingly, the highest value of the Dunn index indicates the best data partitions with evenly distributed, densely packed clusters. In contrast, a lower value is indicative of less compact or less well-separated clusters (Dunn, 1974; Ben, Hamza & Bouaguel, 2021).

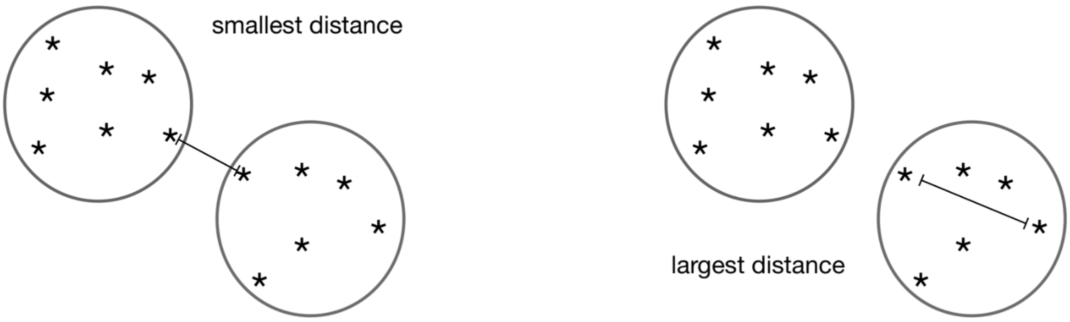


Figure 5. Dunn Index Visualisation (Boily, n.d.)

5. Empirical Analysis

5.1 Methodology

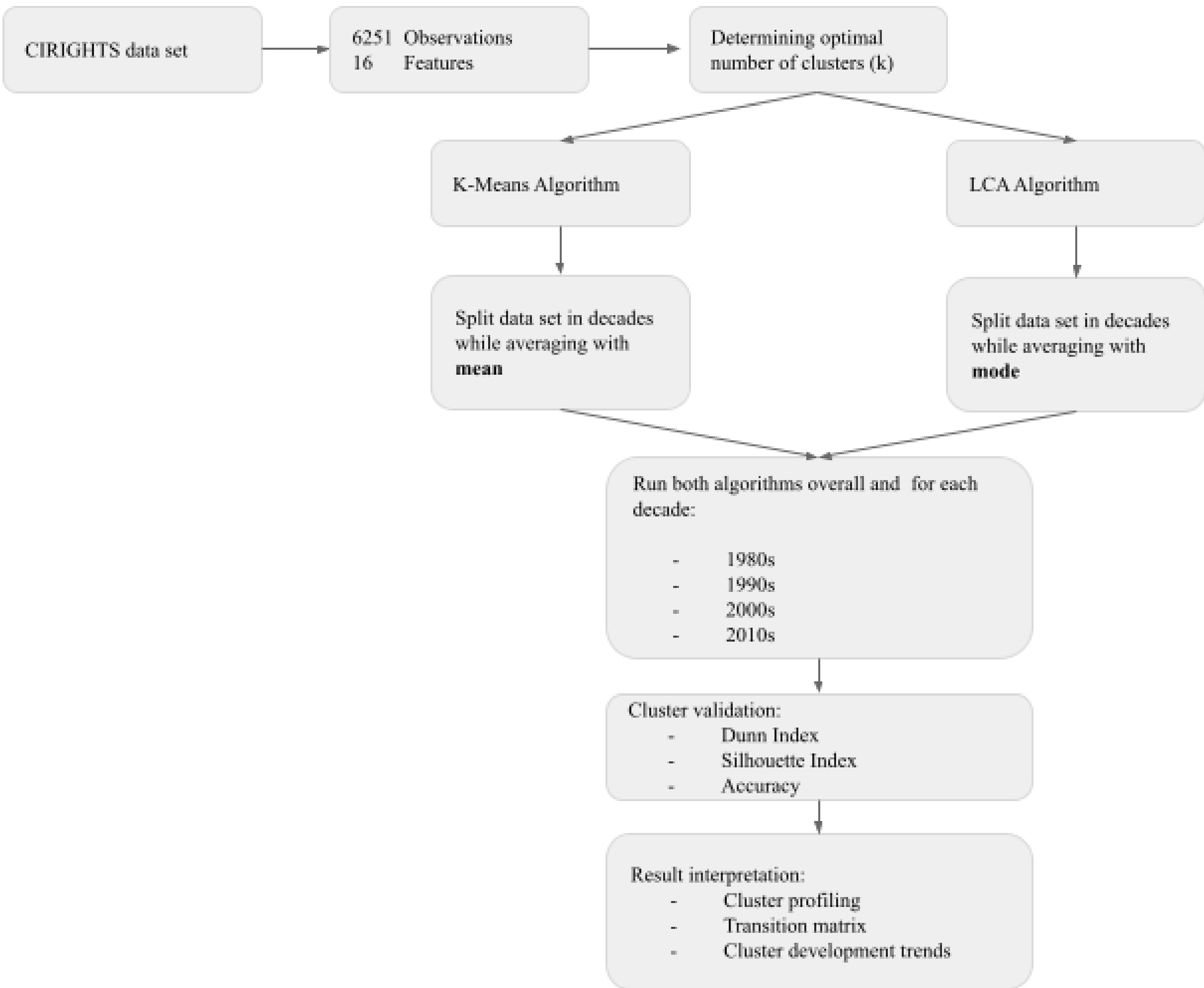


Figure 6. Workflow Diagram

For a general overview of the methodology described in more detail below, please refer to Figure 6 above.

As preparation to run our algorithms, we scaled our data using statistical standardisation (only for K-Means) and then created a new data set in which we aggregated our values by country (using the **mean** for K-Means and **mode** for LCA), in order for each country only to appear in one unique cluster. As such, we went from panel to cross-sectional data. Furthermore, we removed the variable containing country names, as, in order for the algorithms to run properly, the data set can only contain numeric (for k-means) or categorical data (for LCA). However, the downside of aggregating is that one loses time trends in the data, which is why we additionally separated our original data set into four separate ones, each representing different decades. That is, we accounted for variability in cluster assignments over time. Subsequently, we aggregated the four data sets, as done previously.

To choose the optimal number of clusters for the CIRIGHTS data set, we considered the literature-recommended elbow technique (Lindholm, 2022; Tibshirani, Walther & Hastie, 2001) and concluded that 3 is the optimal number of clusters for both K-Means and LCA (Figure 7).

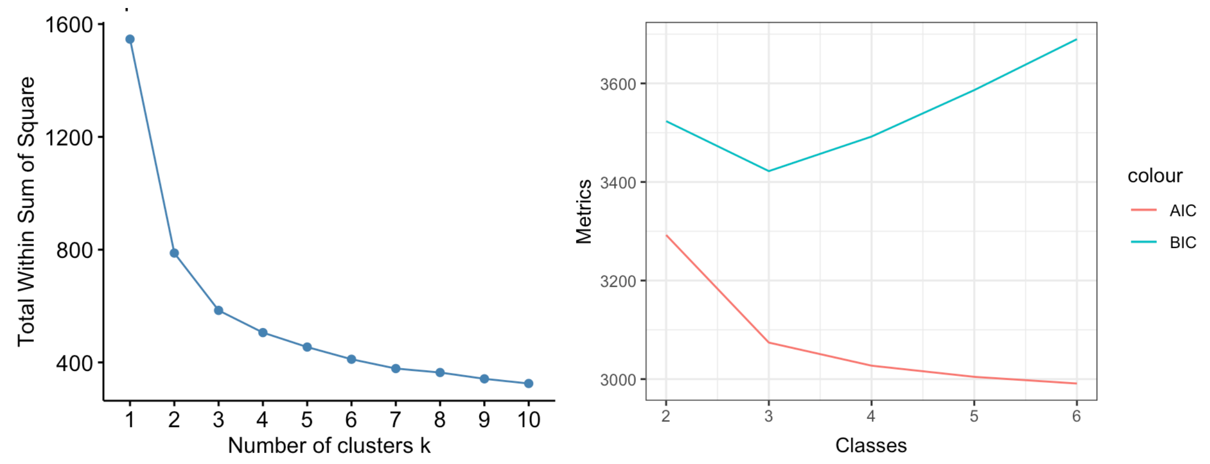


Figure 7. Optimal number of clusters for K-Means (left) and LCA (right)

When evaluating the algorithms in terms of performance and significance, we start off by mentioning that values larger than 0 are achieved for both algorithms, indicating that clustering is successful and can be used for further investigation.

Table 4. Dunn Index comparison for LCA and K-Means

	Overall	1980s	1990s	2000s	2010s
LCA	0.20	0.26	0.23	0.18	0.25
K-Means	0.14	0.28	0.22	0.23	0.26

The Dunn Index, which measures compactness and the degree of separation, shows a better overall performance for LCA. Furthermore, it can be observed in Table 4 that both algorithms cluster almost equally well. Therefore, no superior algorithm can be determined at this point.

Table 5. Silhouette Index Comparison for LCA and K-Means

	Overall	1980s	1990s	2000s	2010s
LCA	0.32	0.35	0.31	0.31	0.31
K-Means	0.33	0.36	0.33	0.34	0.36

The Silhouette Index, which focuses on cluster structure and measures how separated and distinct the clusters are, shows that K-Means performs minimally better than LCA, as seen in Table 5. However, the deviations are so small that no superior algorithm can be determined here either.

Table 6. Accuracy calculated with Human Rights Index and cluster assignments

	Overall	Cluster 1	Cluster 2	Cluster 3
LCA	81.18%	90.91%	70.73%	86.84%
K-Means	81.18%	96.97%	69.51%	78.95%

Since both algorithms perform equally well at this point, we have added another benchmark (see Table 6). This is the Human Rights index from the CIRIGHTS data set, which draws an average score from all available features in the dataset (see Appendix B). Our procedure for this was as follows: first, we divided the maximum score of 100 into 3 groups 66-100 (equal to cluster 1), 33-65 (equal to cluster 2), and 0-32 (equal to cluster 3). From this, we can determine an *accuracy* score for our particular cluster assignments, which corresponds to the share of “correctly” classified countries. Accordingly, our findings are that the K-Means algorithm is better at identifying "good" cases and LCA is better at identifying "bad" cases.

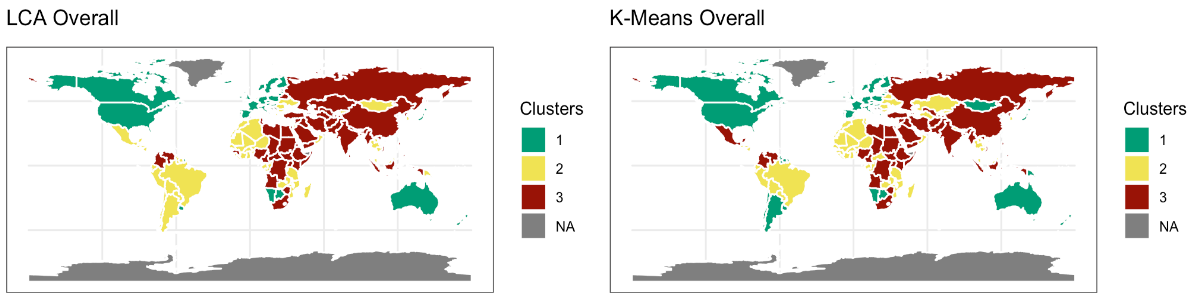


Figure 8. Comparison of LCA and K-Means Clustering Map

As a means of visually comparing the two algorithms we have plotted the map of country clusters averaged over the whole time period (Figure 8). As we can see, the clustering outcome is similar for both algorithms, with the exception of a small number of countries. What can be observed is that LCA seems to cluster certain countries more "strictly" than K-Means, e.g., Mongolia, Kazakhstan, Argentina, etc. However, as a whole and following the evaluation metrics above, we consider the two algorithms to perform equally well. Therefore, in order to respect the categorical nature of our data, we decide to focus more extensively on LCA for the upcoming clustering analysis.

5.2 Latent Class Analysis

We start off by looking at the static clustering output of the LCA algorithm. Namely, we analyse the cluster profiles, by using the entire, aggregated data set. More specifically, we would like to gain insight into and an overview of the median values of each feature in each cluster, in order to identify which patterns the cluster formation and allocation of LCA follow (see Table 7).

At this point, it is important to reiterate that the scores are based on government respect for a variety of internationally recognized human rights and on a scale from 0 (poor) to 2 or 8 (good). However, only the brutality-based mass atrocity variables, bbatrocity and bbtracity_intensity, follow a logic of descending order.

Table 7. Cluster Medians

Feature	Cluster 1	Cluster 2	Cluster 3
physint_sum	7	5	2
polpris	2	1	0
disap	2	2	2
tort	1	1	0
kill	2	1	0
bbatrocity	0	0	0
bbatrocity_intensity	0	0	2
speech	2	1	0
rel_free	2	2	0
dommov	2	2	1
formov	2	2	1
assn	2	1	0
elecsd	2	1	0
wopol	2	2	2
wecon	2	1	1
injud	2	1	0
Total	32	23	7

Cluster 1 (36% of countries): *Best government respect for human rights* - Contains the highest and accordingly the best values across all features. The only variables that don't have the highest score as their median are physint_sum (physical integrity rights index), tort (torture), wopol (women's political rights), and wecon (women's economic rights). Note that both variables that represent women's rights don't have the highest score.

Cluster 2 (37% of countries): *Middle government respect for human rights* - Predominantly the cluster which is in the middle and thus neither contains bad nor particularly good values. Overall, cluster 2 has a total median value (sum of all median values minus bbatrocity and bbatrocity_intensity) closer to that of cluster 1 than 3 (see Table 7; Figure 9). That is, the difference between cluster 3 and 2 is 14 and between 2 and 1 is 9. More particularly, the features rel_free (freedom of religion), dommov (freedom of domestic movement), and formov (freedom of foreign movement) have a tendency towards values of cluster 1.

Cluster 3 (27% of countries): *Worst government respect for human rights* - Here we see the lowest and correspondingly worst scores in terms of respect for human rights. Although the values are not as “extreme” as those of cluster 1, which almost only had the highest possible values for each feature, cluster 3 has the lowest possible value for almost each variable in the physical integrity rights that represents only itself (tort (torture), kill (extrajudicial killings), polpris (political imprisonment)).

Another interesting finding is that all 3 clusters have the best possible value for wopol and disap and are therefore rated the same.

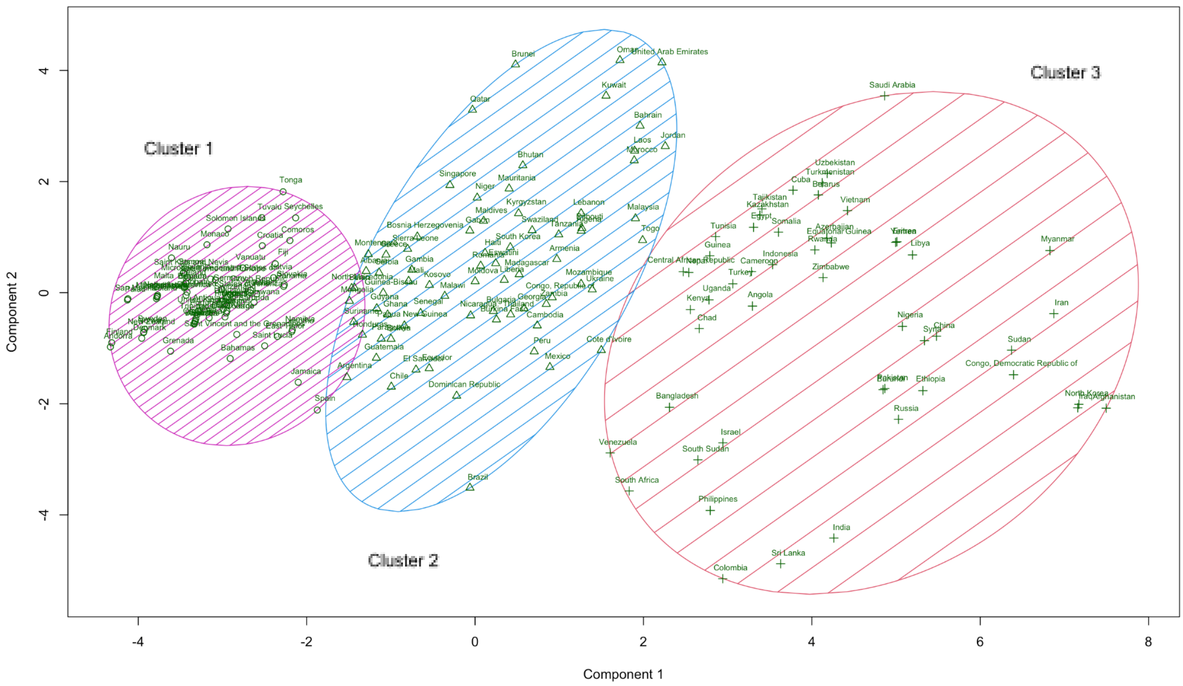


Figure 9. Cluster plot with first two principal components

From this section on, we look at the *dynamic* clustering output of LCA, namely, the cluster allocations over four decades. Hence, below, we have depicted a transition matrix, as defined and inspired by Markov Chains (Rogers & Girolami, 2020), in which we have calculated conditional probabilities of switching from one cluster to another in the next decade (Table 8). That is, each value in the table represents the probability of going from cluster x to cluster y in the transition period *t - t + 1*, conditional on being in cluster x in period t.

Table 8. Transition Matrix of LCA over decades

	1980s-1990s	1990s-2000s	2000s-2010s
1-1	89.47%	67.19%	87.67%
1-2	5.26%	32.81%	12.33%
1-3	5.26%	0.00%	0.00%
2-1	29.41%	2.00%	7.46%
2-2	44.12%	48.00%	76.12%
2-3	25.00%	50.00%	16.42%
3-1	11.11%	2.56%	0.00%
3-2	22.22%	35.90%	13.73%
3-3	62.96%	61.54%	86.27%

Looking at the transition probabilities above, we have focused on three main findings. The first observation, and perhaps the most apparent one, is that the highest probabilities are staying in the same cluster over time. More specifically, the highest overall is found when staying in cluster 1, then 3, and finally 2. Furthermore, for cluster 1 and 3, the lowest probability of staying in the same cluster is when going from the 1990s to the 2000s, whereas for cluster 2, the probability increases with time. As such, in the two earliest transition periods, it was fairly uncertain that a country belonging to cluster 2 would stay in this cluster for the following decade. However, overall, it seems that the probability of staying in the same cluster is highest in the transition period 2000s-2010s. Therefore, one may wonder whether this implies that these later decades were more similar in terms of what affects the degree of human rights violations. Next, we have observed that there are no countries that belonged to cluster 1 in the 1990s (2000s) and radically switched to cluster 3 in the 2000s (2010s). Equivalently, no countries that belonged to cluster 3 in the 2000s radically changed to cluster 1 in the 2010s. This could, once again, be connected with the fact that the highest chances of being in the same cluster were between the 2000s and 2010s, lowering the chance of a radical change. Last but not least, there was an extremely low probability of going from cluster 2 in the 1990s to cluster 1 in the 2000s, whereas the probability of moving towards bad levels of human rights violations (cluster 3) was as high as 50%.

To dig deeper into how the feature values have developed over time within each cluster, we are now going to take a look at the cluster averages for each feature in each decade. For this purpose, we first scaled the data to represent the means of each variable on a comparable scale and then visualised them in Figure 10.

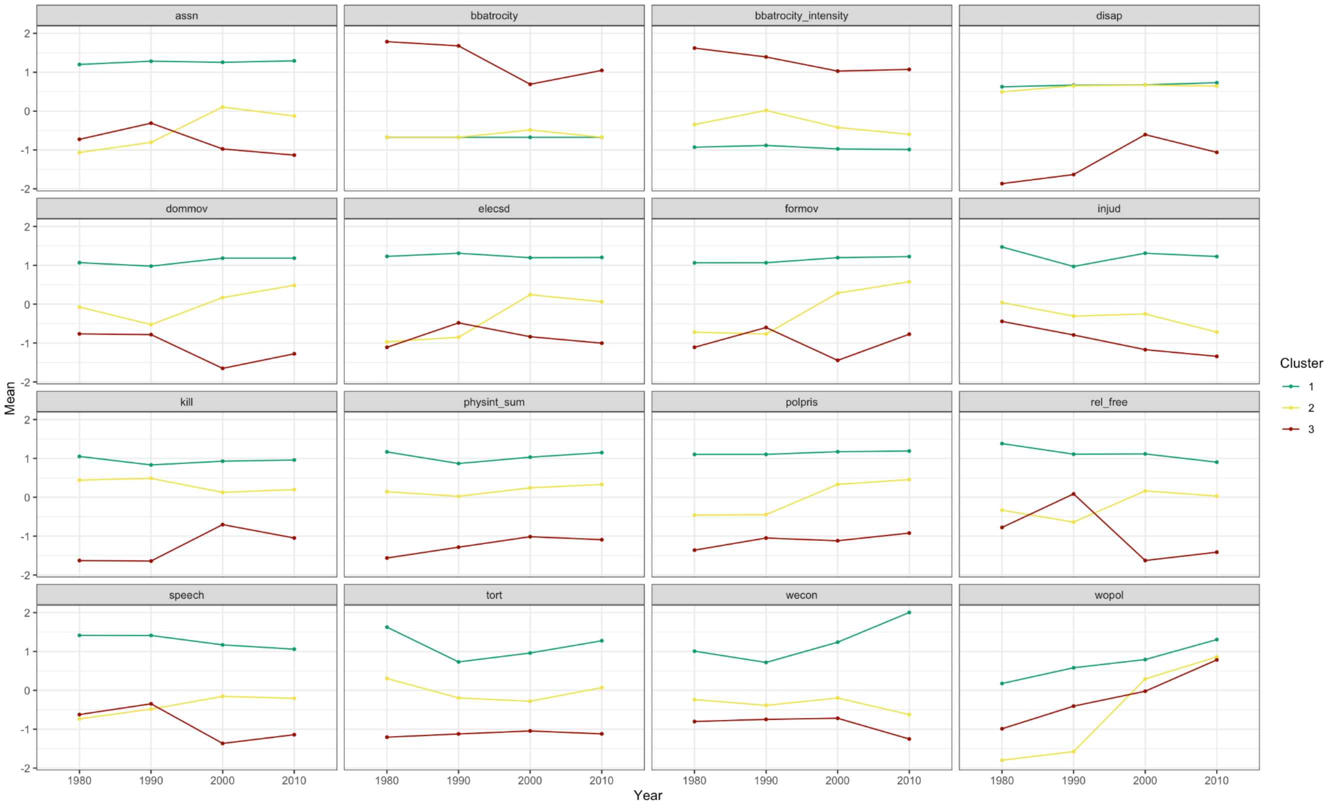


Figure 10. Cluster means development trends across features

Looking at the figure above, we have mainly noticed three interesting trends. First of all, we can observe that, over time, the variables *bbatrocity* and *disap* have mean values that are almost identical for cluster 1 and 2, despite the latter representing "worse" values across all features. The next interesting finding is that the variable *assn* actually had lower average (hence worse) scores in cluster 2 than in cluster 3, in the 1980s and 1990s. As such, it is only from the 2000s onwards that cluster 2's mean increased sharply and that of cluster 3 decreased, reversing the order. A similar trend can also be seen from other variables as well (i.e., *dommov*, *elecsd*, *formov*, *polpris*, *rel_free*, *speech*), where the gap between cluster 2 and 3 widens only in the later decades. Finally, the third feature that stands out is that of wopol, which, not only, has the sharpest increase in mean values over time for every cluster, but also shows a convergence of them in the later decades.

Next, in order to communicate our results as clearly as possible to policymakers, we have created the following maps that have two different emphases.

Clustering Stable: Describes the countries that have remained in the same cluster over time (Figure 11).

Clustering Unstable: Describes the countries that have switched clusters over time. In particular, the visualisation shows improvements, deteriorations, and fluctuations (Figure 12).

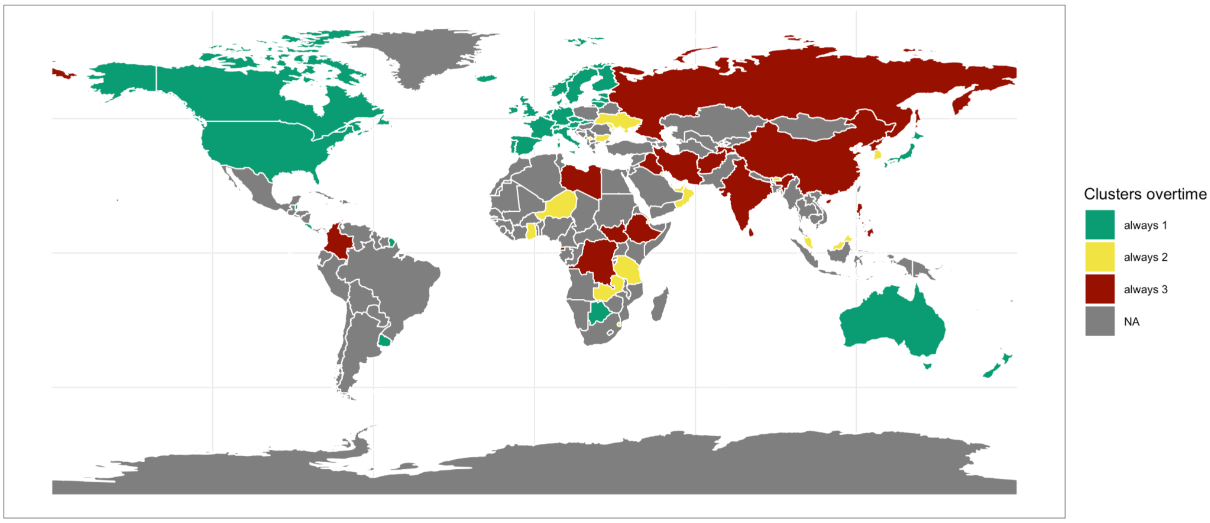


Figure 11. Clustering with stable countries over four decades

Always 1:

Sample countries/regions: North America, Western/Northern Europe, Australia

Advice: These countries should be considered the "gold standard" internationally and do not need any further monitoring in terms of human rights violations.

Always 2:

Sample countries/regions: Ukraine, Bulgaria, Niger, Tanzania, Oman, Malaysia

Advice: Here, a focus should be placed on which individual features still need to be improved in order to e.g., move from cluster 2 to cluster 1.

Always 3:

Sample countries/regions: Numerous Asian countries (Russia, China, India, Iran, Afghanistan…), Colombia, Libya, Dem. Rep. of Congo

Advice: The countries in cluster 3, which are marked in red on the map, should be questioned in terms of why they continue to receive such poor scores concerning human rights. Questions policymakers should ask themselves are: Why have there not been any changes? Why are so many Asian countries associated with poor human rights scores? When it comes to steps to take, there should be a direct and proactive focus on making these countries “behave” better, such as sanctions (to the extent possible), rather than monitoring and awaiting potential improvements.

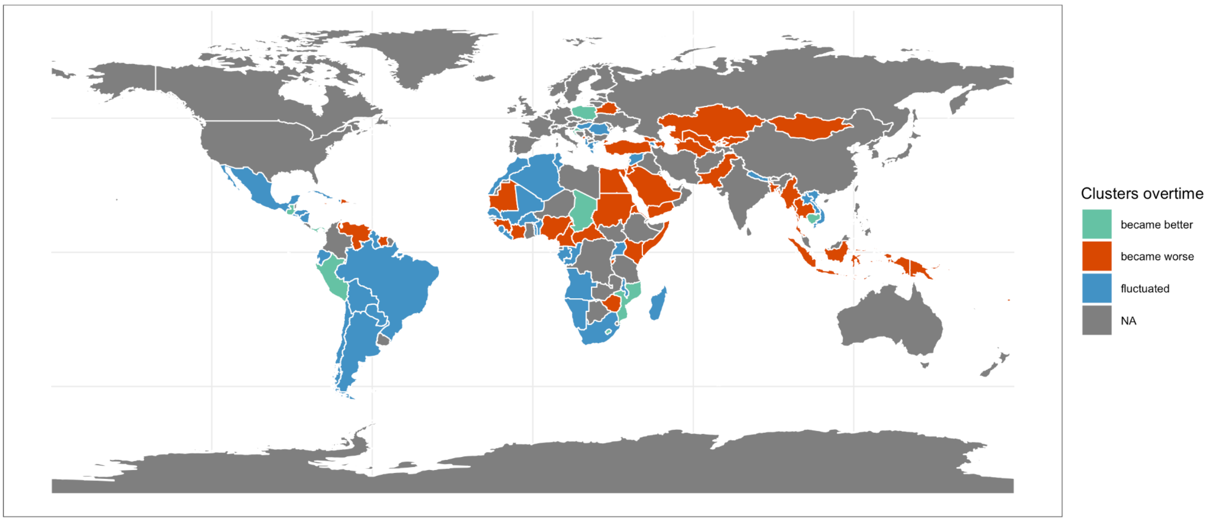


Figure 12. Clustering with unstable countries over four decades

Became better:

Sample countries/regions: Peru, Mozambique, Poland

Advice: Here, we depict the countries that have shown positive development over time. As such, it is important to continue to support them and not let them fall into old structures. As an incentive, policymakers should use the “gold-standard” mentioned above as an example for countries that are comparable structurally, demographically etc., and have the potential and ability to further better themselves (e.g., Poland can be compared to other European countries). Furthermore, countries that belong to this category may in turn serve as best practice for countries that have remained in cluster 2 or 3 over time.

Became worse:

Sample countries/regions: Middle East, Kazakhstan, Mongolia, Belarus, Turkey, Indonesia

Advice: The map shows that some countries are becoming more vulnerable in terms of human rights. Hence, policymakers should try to pick up on common characteristics between these countries, more specifically, their political/governmental and social structure (e.g., similarities between Middle Eastern countries), and dig deeper into what is influencing the deterioration of these countries in terms of human rights. Furthermore, this group of countries is particularly important to focus on, as they are not going in the right direction over time and, therefore, efforts to reverse the trend should be of high priority to policymakers.

Fluctuated:

Sample countries/regions: South/Central America, Southern/Western Africa, Romania, Hungary

Advice: The remainder of countries in our data set are those that do not fall into one of the previous categories and, hence, have shown a behaviour of alternating between clusters over time. Therefore, as these countries are considered very unstable, it is important for policymakers to monitor these more closely and observe whether the trends lead to a worse or more positive cluster assignment, overall. Accordingly, policymakers should identify stabilising factors and focus on what has led these countries to, at some point in time, attain good levels of human rights violations.

6. Conclusion

6.1 Summary and Implication

The aim of this thesis is to as effectively as possible cluster countries based on their degree of human rights violations, using 2 different clustering algorithms, namely, K-Means and LCA. After having run both algorithms, using the entire data set as well as with 4 separate decades to get a better insight into the evolution of cluster assignments over time, we can confidently say that both methods performed equally well for our study. That is, as mentioned previously, we found that the existing literature often claimed that LCA outperforms K-Means as a clustering algorithm. However, our three performance indicators, Dunn index, Silhouette coefficient, and Accuracy (using the Human Rights index as labels), argue differently.

Nevertheless, we decided to focus on LCA for the bulk of our analysis in order to respect the categorical nature of our data. Hence, when looking at the characteristics of each cluster, we found a clear order of magnitude when it comes to the degree of human rights violations. That is, cluster 1 showed the highest/best overall scores and cluster 3 the lowest/worst scores, with cluster 2, as expected, taking second place. However, when looking into the cluster development trends, we saw that not each feature and each decade respected the best-to-worst order of cluster 1 to 3. Therefore, this shows that the clustering was conducted taking into account the variables altogether and not individually.

Remaining on the topic of dynamic clustering, we also computed a transition matrix, which showed us that the highest probabilities were for countries to stay in the same cluster over time. Hence, this may point towards the fact that looking at the very big picture, countries do not change that much over time in terms of their level of respect for human rights.

Finally, we considered yet another implication of our results, namely, potential policy advice. That is, we created groups of “stable” (always cluster 1/2/3) and “unstable” (became better/worse, fluctuated) countries and suggested a policy strategy for each. Namely, countries that remained in cluster 1 should be used as the “gold-standard”, those that remained in cluster 2 should be incentivised and finally, policymakers should take a proactive approach for countries remaining in cluster 3. Next, policymakers should incentivise countries that have become better with time, understand the reasons for those that have become worse and focus on reversing the trend, and finally, they should try to stabilise and monitor more closely countries that have been extra unstable, i.e., fluctuated over time.

Hence the implication of using unsupervised machine learning is that it can provide us with valuable and prompt insights by uncovering patterns that may not be immediately visible. This approach is comparable to using a benchmark, such as the HR index, to evaluate performance. Furthermore, unsupervised ML offers the potential to identify trends and developments within clusters, giving us a deeper understanding of the data, and summarising a wide range of variables into an easily interpretable framework.

6.2 Limitations and Future Work

There are several limitations that need to be considered when it comes to unsupervised learning with the CIRIGHTS data set. One major drawback is the lack of accuracy, which makes it difficult to evaluate the performance of the model. Instead, evaluation is mostly based on subjective views. Additionally, choosing the appropriate number of clusters for latent class analysis can be challenging and often relies on the subjective elbow curve method.

Another limitation arises from the fact that aggregating, due to the panel structure of our data set, may lead to a loss of preciseness. Furthermore, there may be a bias in clustering the countries, as the dataset and indices were created in the USA, a well-developed, First World country.

To address these limitations, future work could include focusing on one particular violation and clustering countries based on this, individually, to narrow down the scope of the analysis. Furthermore, for the same purpose, one could look at each year individually, rather than aggregating, leading to a more precise and dynamic clustering. On the other hand, to broaden the scope of the analysis and take into account more country characteristics, it may be useful to add socioeconomic indicators to the data set. Finally, an additional step that one may take is to take advantage of the time series nature of the data and use past information to predict future clusters, using other, supervised, machine learning techniques.

References

Adewusi, E. O. & Kocadal, Ö. (2022). A Comparative Analysis of Human Rights Protection in European Union and African Union Countries: An fsQCA Approach, *Uluslararası İlişkiler Dergisi*, pp.1–21

Alexander, N., Alexander, D. C., Barkhof, F. & Denaxas, S. (2021). Identifying and Evaluating Clinical Subtypes of Alzheimer's Disease in Care Electronic Health Records Using Unsupervised Machine Learning, *BMC Medical Informatics and Decision Making*, vol. 21, no. 1, p.343

Akaike, H. (1974). A New Look at the Statistical Model Identification, *IEEE Transactions on Automatic Control*, vol. 19, no. 6, pp.716–723

Ben, C.-E., Hamza, A. & Bouaguel, W. (2021). Parallel and Scalable Dunn Index for the Validation of Big Data Clusters, *Parallel Computing*, vol. 102, p.102751

Boily, P. (n.d.). 14.3 Clustering Evaluation | Data Understanding, Data Analysis, and Data Science - Course Notes (DRAFT), *Www.idlewyldanalytics.com*, Available Online: https://www.idlewyldanalytics.com/clust-ev [Accessed 8 May 2023]

Chen, S. S. & Gopalakrishnan, P. S. (1998). Clustering via the Bayesian Information Criterion with Applications in Speech Recognition, in *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP '98 (Cat. No.98CH36181)*, Vol. 2, 1998 IEEE International Conference on Acoustics, Speech, and Signal Processing, Seattle, WA, USA, 1998, Seattle, WA, USA: IEEE, pp.645–648, Available Online: http://ieeexplore.ieee.org/document/675347/ [Accessed 15 May 2023]

Cingranelli, D., Mark, S., Gibney, M., Haschke, P., Wood, R. & Arnon, D. (2019). Human Rights Violations and Violent Internal Conflict, *Social Sciences*, vol. 8, no. 2, p.41

Cingranelli, D. L., Richards, D. L. & Clay, K. C. (2021). The CIRI Human Rights Dataset, Version 2014.04.14., Available Online: https://dataverse.harvard.edu/citation?persistentId=doi:10.7910/DVN/UKCPXT [Accessed 15 May 2023]

CIRIGHTS. (2022). *CIRIGHTS*, Available Online: https://cirights.com [Accessed 23 April 2023]

Cope, K. L., Crabtree, C. & Fariss, C. J. (2020). Patterns of Disagreement in Indicators of State Repression, *Political Science Research and Methods*, vol. 8, no. 1, pp.178–187

Davies, D. L. & Bouldin, D. W. (1979). A Cluster Separation Measure, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-1, no. 2, pp.224–227

Dunn, J. C. (1974). Well-Separated Clusters and Optimal Fuzzy Partitions, *Journal of Cybernetics*, vol. 4, no. 1, pp.95–104

Emerson, J., Satterthwaite, M. L. & Pandey, A. V. (2018). The Challenging Power of Data Visualization for Human Rights Advocacy, in M. K. Land & J. D. Aronson (eds), *New Technologies for Human Rights Law and Practice*, 1st edn, [e-book] Cambridge University Press, pp.162–187, Available Online: https://www.cambridge.org/core/product/identifier/9781316838952%23CN-bp-8/type/book_part [Accessed 9 May 2023]

Fahad, A., Alshatri, N., Tari, Z., Alamri, A., Khalil, I., Zomaya, A. Y., Foufou, S. & Bouras, A. (2014). A Survey of Clustering Algorithms for Big Data: Taxonomy and Empirical Analysis, *IEEE Transactions on Emerging Topics in Computing*, vol. 2, no. 3, pp.267–279

Findley, D. F. (1991). Counterexamples to Parsimony and BIC, *Annals of the Institute of Statistical Mathematics*, vol. 43, no. 3, pp.505–514

Gartner, S. S. & Regan, P. M. (1996). Threat and Repression: The Non-Linear Relationship between Government and Opposition Violence, *Journal of Peace Research*, vol. 33, no. 3, pp.273–287

Goodman, L. A. (1974). Exploratory Latent Structure Analysis Using Both Identifiable and Unidentifiable Models, *Biometrika*, vol. 61, no. 2, pp.215–231

Grewal, S. & Voeten, E. (2015). Are New Democracies Better Human Rights Compliers?, *International Organization*, vol. 69, no. 2, pp.497–518

Guzman, A. T. & Linos, K. (2013). Human Rights Backsliding, *SSRN Electronic Journal*, [e-journal], Available Online: http://www.ssrn.com/abstract=2229072 [Accessed 15 May 2023]

Hagenaars, J. A. & McCutcheon, A. L. (eds). (2002). Applied Latent Class Analysis, Cambridge ; New York: Cambridge University Press

Jain, A. K. & Dubes, R. C. (1988). Algorithms for Clustering Data, Englewood Cliffs, N.J: Prentice Hall

Landman, T. (2004). Measuring Human Rights: Principle, Practice and Policy, *Human Rights Quarterly*, vol. 26, no. 4, pp.906–931

Lazarsfeld, P. F. & Henry, N. W. (1968). Latent Structure Analysis, New York: Houghton, Mifflin

LeCun, Y., Bengio, Y. & Hinton, G. (2015). Deep Learning, *Nature*, vol. 521, no. 7553, pp.436–444

Lin, C.-C. & Ng, S. (2012). Estimation of Panel Data Models with Parameter Heterogeneity When Group Membership Is Unknown, *Journal of Econometric Methods*, [e-journal] vol. 1, no. 1, Available Online: https://www.degruyter.com/document/doi/10.1515/2156-6674.1000/html [Accessed 15 May 2023]

Lindholm, A. (2022). Machine Learning: A First Course for Engineers and Scientists, Cambridge, UK ; New York, NY: Cambridge University Press

Linzer, D. A. & Lewis, J. B. (2011). PoLCA : An *R* Package for Polytomous Variable Latent Class Analysis, *Journal of Statistical Software*, [e-journal] vol. 42, no. 10, Available Online: http://www.jstatsoft.org/v42/i10/ [Accessed 18 May 2023]

Liu, Y., Li, Z., Xiong, H., Gao, X. & Wu, J. (2010). Understanding of Internal Clustering Validation Measures, in *2010 IEEE International Conference on Data Mining*, 2010 IEEE 10th International Conference on Data Mining (ICDM), Sydney, Australia, December 2010, Sydney, Australia: IEEE, pp.911–916, Available Online: http://ieeexplore.ieee.org/document/5694060/ [Accessed 15 May 2023]

Lv, Z., Liu, T., Shi, C., Benediktsson, J. A. & Du, H. (2019). Novel Land Cover Change Detection Method Based on K-Means Clustering and Adaptive Majority Voting Using Bitemporal Remote Sensing Images, *IEEE Access*, vol. 7, pp.34425–34437

Magidson, J. & Vermunt, J. K. (2002). Latent Class Models for Clustering : A Comparison with K-Means, *Canadian journal of marketing research*, vol. 20, no. 1, pp.36–43

Nguyen, T.-H. T., Dinh, D.-T., Sriboonchitta, S. & Huynh, V.-N. (2019). A Method for K-Means-like Clustering of Categorical Data, *Journal of Ambient Intelligence and Humanized Computing*, [e-journal], Available Online: http://link.springer.com/10.1007/s12652-019-01445-5 [Accessed 14 April 2023]

Nosetti, L., Paglietti, M. G., Brunetti, L., Masini, L., La Grutta, S., Cilluffo, G., Ferrante, G., Zaffanello, M., Verrillo, E., Pavone, M., Niespolo, A. C., Broggi, G., Cutrera, R., & on behalf of the IPRS (SIMRI)Task Force on Sleep Disordered Breathing (SDB-TF). (2020). Application of Latent Class Analysis in Assessing the Awareness, Attitude, Practice and Satisfaction of Paediatricians on Sleep Disorder Management in Children in Italy, *PLOS ONE*, vol. 15, no. 2, p.e0228377

Nylund, K., Bellmore, A., Nishina, A. & Graham, S. (2007). Subtypes, Severity, and Structural Stability of Peer Victimization: What Does Latent Class Analysis Say?, *Child Development*, vol. 78, no. 6, pp.1706–1722

Pence, B. W., Miller, W. C. & Gaynes, B. N. (2009). Prevalence Estimation and Validation of New Instruments in Psychiatric Research: An Application of Latent Class Analysis and Sensitivity Analysis., *Psychological Assessment*, vol. 21, no. 2, pp.235–239

Poe, S. C., Tate, C. N. & Keith, L. C. (1999). Repression of the Human Right to Personal Integrity Revisited: A Global Cross-National Study Covering the Years 1976-1993, *International Studies Quarterly*, vol. 43, no. 2, pp.291–313

Rall, K., Satterthwaite, M. L., Pandey, A. V., Emerson, J., Boy, J., Nov, O. & Bertini, E. (2016). Data Visualization for Human Rights Advocacy, *Journal of Human Rights Practice*, vol. 8, no. 2, pp.171–197

Robertson, A. H. & Merrills, J. G. (1996). Human Rights in the World: An Introduction to the Study of the International Protection of Human Rights, 4th ed., Manchester [England] ; New York : New York, NY, USA: Manchester University Press ; Distributed exclusively in the USA by St. Martin's Press

Rogers, S. & Girolami, M. (2020). A First Course in Machine Learning, Second edition, first issued in paperback., Boca Raton London New York: CRC Press, Taylor & Francis Group

Roser, M. (2016). Human Rights, *Our World in Data*, Available Online: https://ourworldindata.org/human-rights

Rousseeuw, P. J. (1987). Silhouettes: A Graphical Aid to the Interpretation and Validation of Cluster Analysis, *Journal of Computational and Applied Mathematics*, vol. 20, pp.53–65

Roussos, L. A., Templin, J. L. & Henson, R. A. (2007). Skills Diagnosis Using IRT-Based Latent Class Models, *Journal of Educational Measurement*, vol. 44, no. 4, pp.293–311

Schwarz, G. (1978). Estimating the Dimension of a Model, *The Annals of Statistics*, [e-journal] vol. 6, no. 2, Available Online: https://projecteuclid.org/journals/annals-of-statistics/volume-6/issue-2/Estimating-the-Dimension-of-a-Model/10.1214/aos/1176344136.full [Accessed 23 May 2023]

Shirkhorshidi, A. S., Aghabozorgi, S., Wah, T. Y. & Herawan, T. (2014). Big Data Clustering: A Review, in B. Murgante, S. Misra, A. M. A. C. Rocha, C. Torre, J. G. Rocha, M. I. Falcão, D. Taniar, B. O. Apduhan, & O. Gervasi (eds), *Computational Science and Its Applications – ICCSA 2014*, Vol. 8583, [e-book] Cham: Springer International Publishing, pp.707–720, Available Online: http://link.springer.com/10.1007/978-3-319-09156-3_49 [Accessed 15 May 2023]

Tibshirani, R., Walther, G. & Hastie, T. (2001). Estimating the Number of Clusters in a Data Set Via the Gap Statistic, *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 63, no. 2, pp.411–423

Xu, B. (2011). Clustering Educational Digital Library Usage Data: Comparisons of Latent Class Analysis and K-Means Algorithms, [e-journal], Available Online: https://digitalcommons.usu.edu/etd/954 [Accessed 15 May 2023]

Appendices

Appendix A: Variable description

Table A1. Coding of our variables

Variable	Coding
“disap”	0: occurred frequently 1: occurred occasionally 2: never occurred/unreported
“kill”	0: practised frequently 1: practised occasionally 2: not practised/unreported
“polpris”	0: many political prisoners 1: few political prisoners 2: none
“tort”	0: practised frequently 1: practised occasionally 2: not practised/unreported
“physint_sum”	0: no government respect for these four rights to 8: full respect
“bbatrocity”	1: if a country scores 0 (i.e., bad) on extrajudicial killings *and* a score of 0 on disappearances, torture, and/or political imprisonment. 0: otherwise
“bbatrocity_intensity”	1, 2, or 3, representing the number of *extra* physical integrity rights that score 0
“assn”	0: completely denied or severely restricted for all citizens 1: limited for all citizens or severely denied/restricted for certain groups 2: virtually unrestricted
“formov”	0: severely restricted 1: somewhat restricted

	2: unrestricted
“dommov”	0: severely restricted 1: somewhat restricted 2: unrestricted
“speech”	0: complete censorship/ownership 1: some 2: none
“elecsd”	0: free and fair elections not respected 1: moderately respected 2: generally respected
“rel_free”	0: severe and widespread government restrictions 1: moderate 2: practically absent
“wecon”	0: no economic rights for women under law and government tolerates high level of discrimination 1: some economic rights under law, however not enforced in practice and moderate level of discrimination tolerated 2: some economic rights under law *and* enforced in practice, but still low level of discrimination tolerated 3: virtually all rights guaranteed by law and no discrimination tolerated
“wopol”	0: laws don’t guarantee any of women’s and completely restrict women in this regard 1: political equality guaranteed by law but women hold < 5% of seats in high-ranking government positions 2: political equality guaranteed by law but women only hold between 5 and 30% of seats in high-ranking government positions 3: political equality guaranteed by law and in practice, women hold > 30% of seats in high-ranking government positions
“injud”	0: not independent 1: partially independent

	2: generally independent

Table A2. Variables omitted from our data set due to missing values

Empowerment rights and freedom	Worker rights	Justice rights
• Women's social rights	• Right to form workers' unions, • Right to bargain collectively • Reasonable limitation on working hours • Right to be free from forced labour • Children's rights • Right to a minimum wage • Occupational health and safety rights • Human trafficking	• Right to a fair trial

Appendix B: Model validation

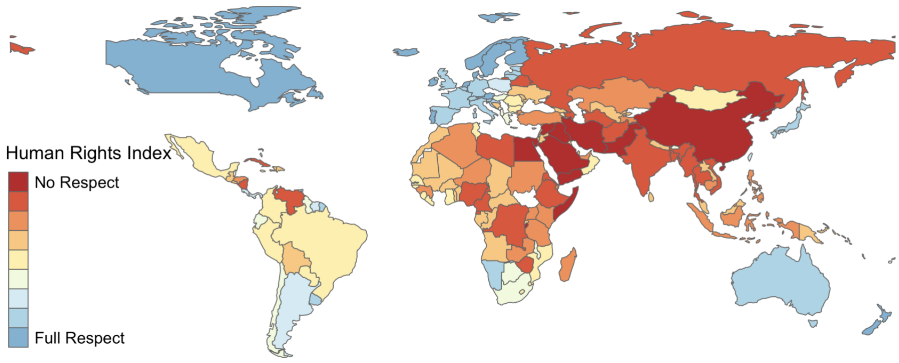


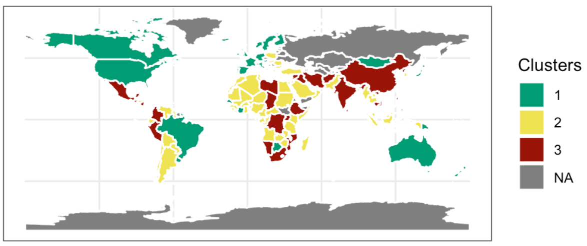
Figure B1. Map representing the Human rights index of the CIRIGHTS project (CIRIGHTS, 2022)

Appendix C: Clustering results

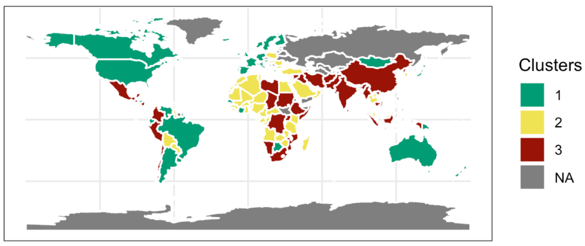
Table C1. Transition matrix for K-Means

	1980s-1990s	1990s-2000s	2000s-2010s
1-1	88.37%	93.65%	87.00%
1-2	4.65%	4.76%	13.00%
1-3	6.98%	1.59%	0.00%
2-1	22.81%	10.00%	7.55%
2-2	56.14%	64.00%	86.79%
2-3	21.05%	26.00%	5.66%
3-1	12.12%	12.50%	0.00%
3-2	18.18%	30.00%	39.47%
3-3	63.64%	57.50%	60.53%

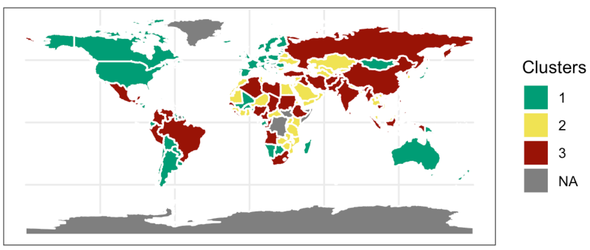
LCA clusters 1980s



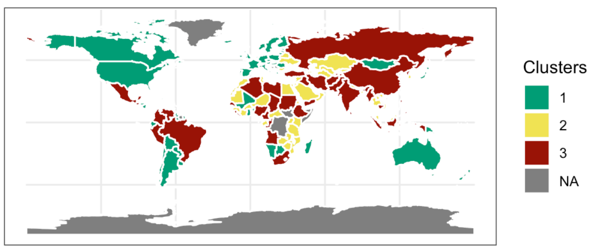
K-Means clusters 1980s



LCA clusters 1990s



K-Means clusters 1990s



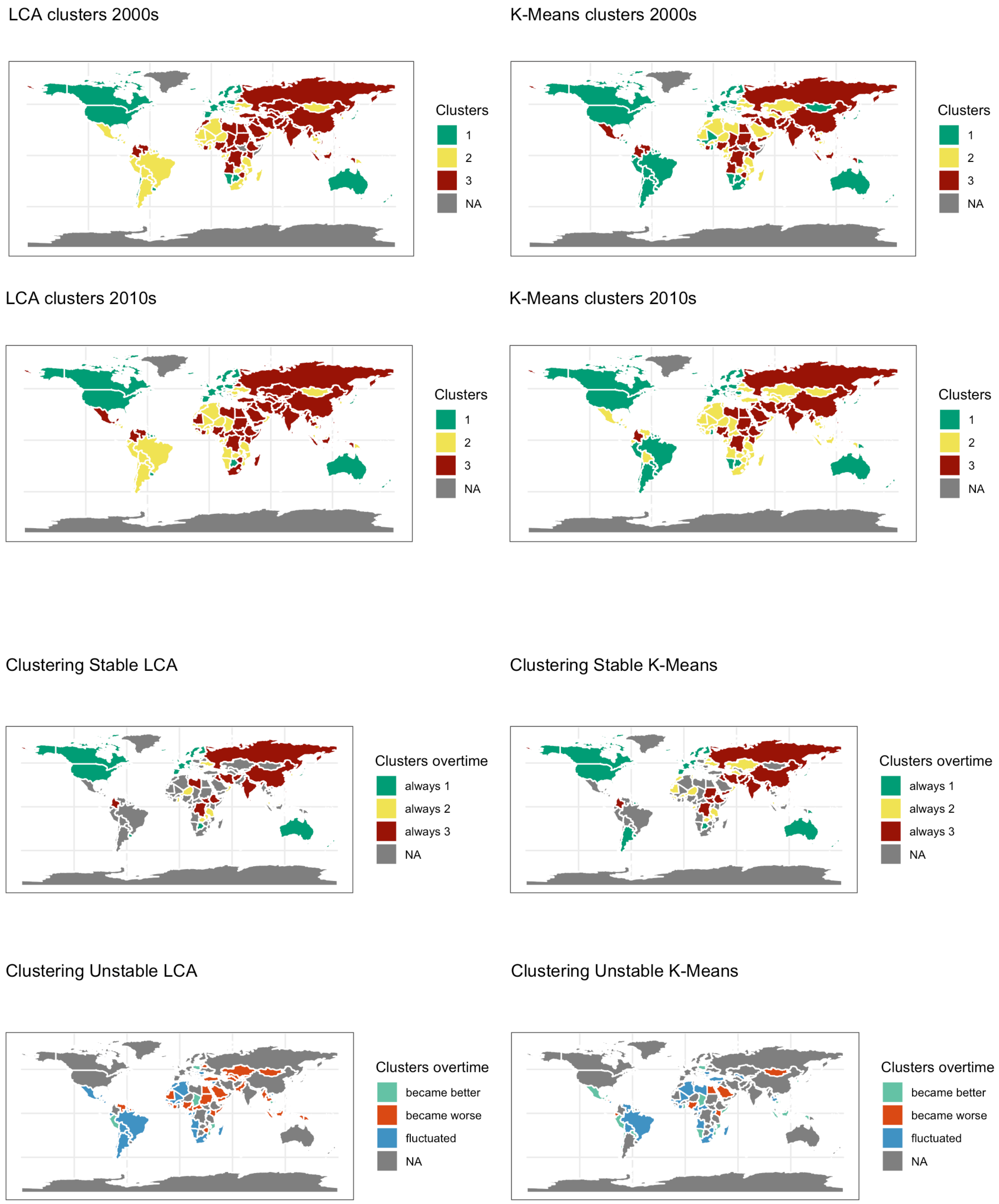


Figure C1. Clustering maps for LCA and K-Means