Stats 140XP: Final Project

Ethan Allavarpu · Raymond Bai · Jaclyn Chiu · Ariel Chow · Carlie Lin · Dara Tan
Explore the GLOBE

1 December 2021

Contents

Abstract

We wanted to answer two main questions: which leadership qualities do countries tend to view similarly and if countries align their perceptions of societal practices and values. For determining leadership qualities and similar countries, we used principal component analysis (PCA) then k-means clustering to create four "clusters" of countries with similar leadership beliefs. For determining the alignment of perceptions of societal practices and values, we used linear regression models for each pair. Then, we performed t-tests for each dimension to check for significant slopes. We found that leadership characteristics are generally clustered by if the trait is negative or positive and that leadership values are clustered by region, except for one interesting case. On the perceptions side, we shockingly discovered that societal practices and societal values of countries do not align and are instead opposing. In the future, we recommend looking into causes behind the interesting disconnect between practices and values. Some limitations to our project are the data silences due to the lack of information from many countries.

Problem Statements

1. Which characteristics or traits do countries tend to group together when determining "good" leadership values?
    - Which countries have similar perceptions of these leadership values?
2. Do societal practices and societal values align?
    - If they do not, which practices and values deviate most significantly?

Description of Dataset

The data set we have chosen to analyze is the **Dana Landis Leadership** dataset, which comes from the GLOBE Research Survey. The data provided in the folder had survey results for (1) leadership and (2) societal and culture data and a PDF describing the nature of the survey, but nothing more. To glean more information, we found the two questionnaires (alpha and beta) described in the informational PDF to get the original questions asked in the surveys. While we do not have a "codebook" in a traditional sense, the original questions asked may help guide us in understanding what each variable means and how the survey represents respondents answers numerically. The survey is on a 1 to 7 scale, with 1 being a negative response, 4 a "neutral" score, and 7 positive.

The GLOBE (Global Leadership & Organizational Behavior Effectiveness) research program is an interdisciplinary study aiming to identify the interrelationships between societal culture, societal effectiveness and organizational leadership. Surveying from over 17,000 middle managers from 62 cultures, the 2004 research survey provides data that measures the leadership and societal culture practices of each country (House et al, 2004).

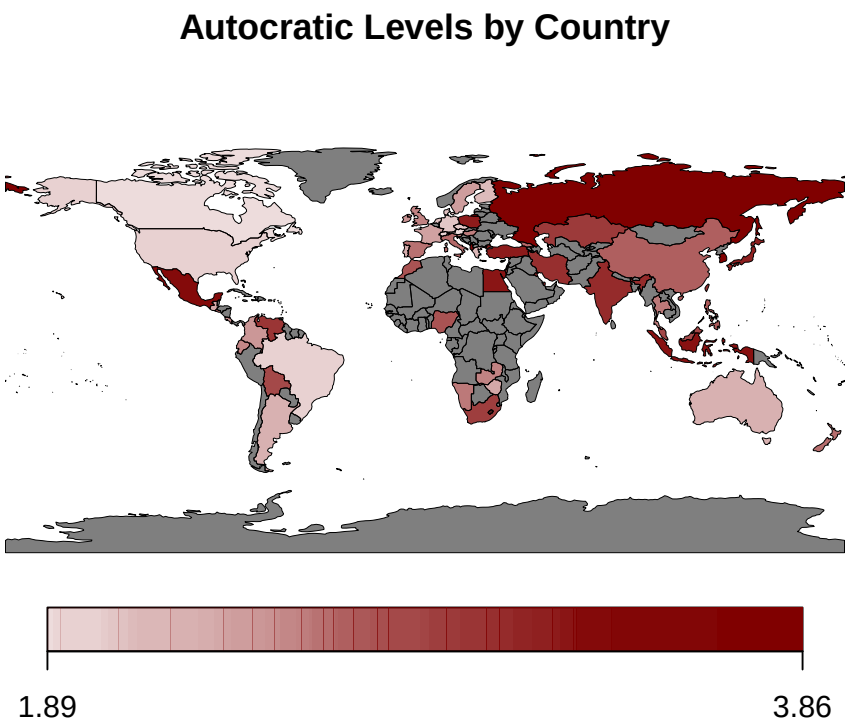
A single, complete observation from the leadership survey is about a specific culture (either a country or a subcountry) about the following characteristics: Country; Country Name; Performance Oriented; Autocratic; Modesty; Charismatic 3 Self sacrifice; Team 1 Collaborative Team Orientation; Decisive; Diplomatic; Face saver; Charismatic 1 Visionary; Humane oriented; Integrity; Bureaucratic Originally Labeled *Procedural*; Administratively Competent; Self centred; Autonomous; Status Conscious; Charismatic 2 Inspirational; Malevolent; Team 2 Team Integrator; Internally Competitive Originally Labeled *Conflict Inducer*; Participative; Charismatic Value based Global Leadership Dimension; Team Oriented Global Leadership Dimension; Self Protective Global Leadership Dimension; Participative Global Leadership Dimension; Humane Oriented Global Leadership Dimension; Autonomous Global Leadership Dimension; Country Cluster.

A single, complete observation from the society and culture survey has a similar structure as the leadership survey: each observation is for a specific culture (either a country or a subcountry) about the following characteristics: Country; Country Name; Uncertainty Avoidance Societal Practices; Future Orientation Societal

Practices; Power Distance Societal Practices; Collectivism I Societal Practices Institutional Collectivism ; Humane Orientation Societal Practices; Performance Orientation Societal Practices; Collectivism II Societal Practices In group Collectivism ; Gender Egalitarianism Societal Practices; Assertiveness Societal Practices; Uncertainty Avoidance Societal Values; Future Orientation Societal Values; Power Distance Societal Values; Collectivism I Societal Values Institutional Collectivism ; Human Orientation Societal Values; Performance Orientation Societal Values; Collectivism II Societal Values In group Collectivism ; Gender Egalitarianism Societal Values; Assertiveness Societal Values; Country Cluster.

Visualization and Exploratory Data Analysis

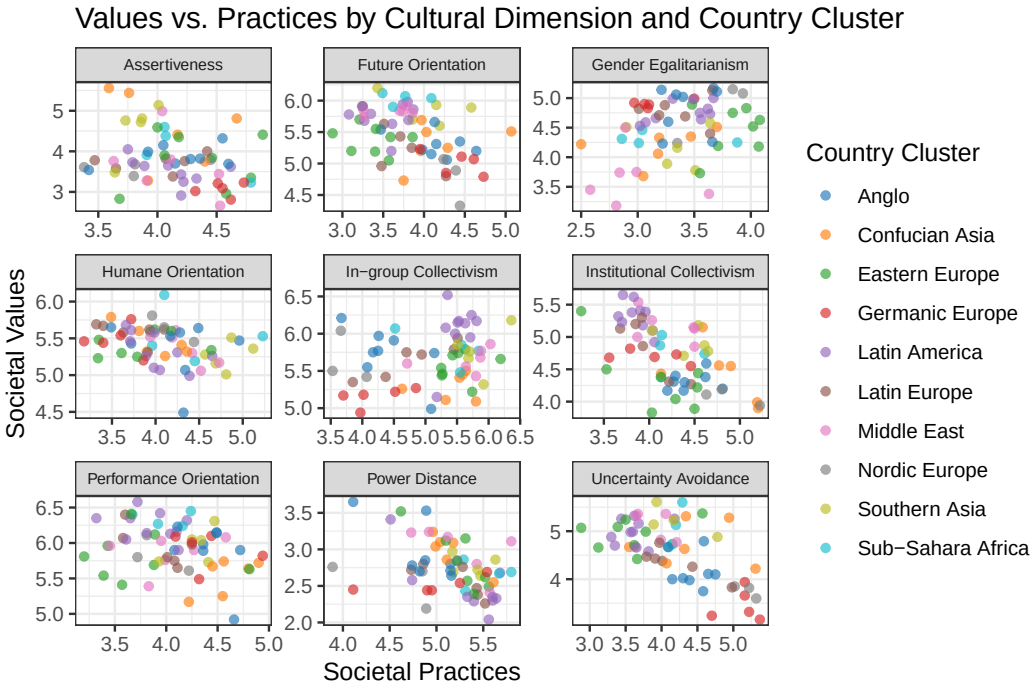
Leadership



The above map shows autocratic levels by country. The lighter colors correspond with lower levels of autocracy while the darker colors correspond with higher levels of autocracy. For example, we see that Russia, a well-known autocratic country, is shaded darkly. Notably, there do seem to be clusters of autocratic countries in the East and clusters of less autocratic countries in the West. The East is known for being more collectivist while the West is known for being more individualistic. As we delve into later, beliefs and cultures do seem to vary by region.

From boxplots we generated (not pictured), the difference by country cluster (as defined by the dataset) seems visible, prompting us to consider clustering countries to determine how they may be separated. We do want to note that there are only 62 observations–this totals to fewer than 62 countries because Germany and South Africa have two observations each (West vs. East and White vs. Black, respecively). The limited number of observations may limit us in the scope of our analysis because we have less than a third of the total countries (demonstrated by the vast swaths of grey on the world map for which there is no data) and clustering would further reduce block/group sizes.

Societal Values and Practices



Note: Countries with missing country clusters were excluded.

The plot grid demonstrates how, contrary to what *should* happen, the societal values of a country do not correlate well with the societal practices (what *should be* does not align with what *is*). This may be a source of further exploration, as we may want to see whether there is a significant difference between the values and practices. When we look at the per-region divisions though (by color), the trend (or lack thereof) seems to persist for these concepts.

Analysis

Leadership Values

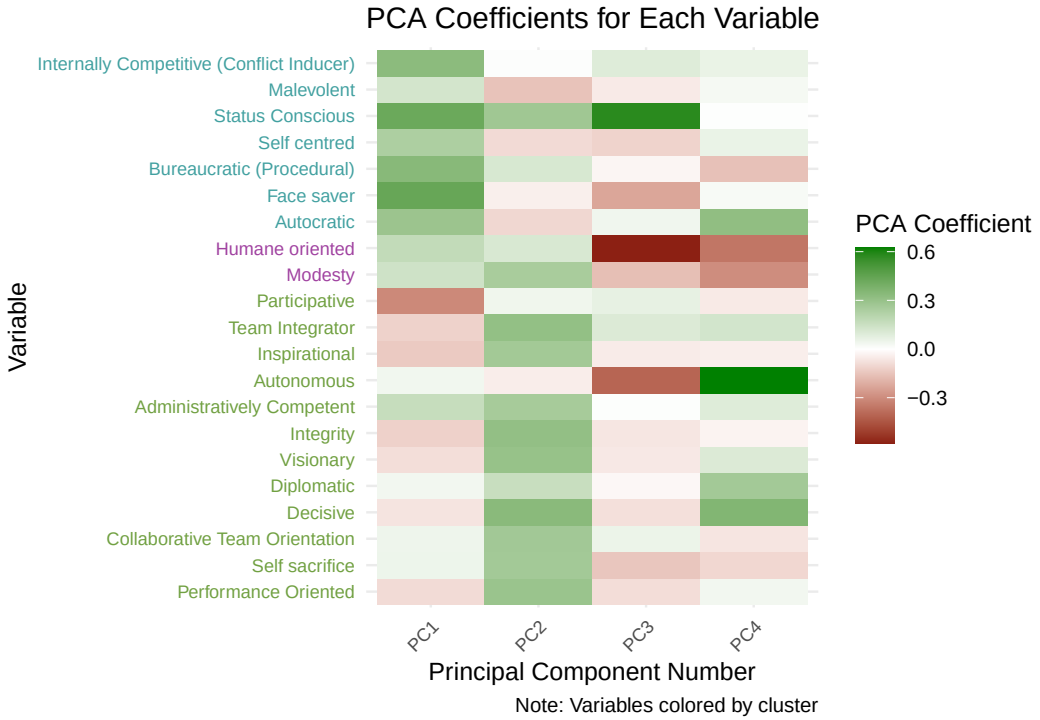
For the leadership values problem statement, the first objective is to collapse the data into the first four principal components through principal component analysis (PCA). PCA finds the directions which capture the most variability (spread) in the data–the first four account for the maximum variation. PCA allows us to visualize trends in the leadership values: countries tend to have similar sentiments about variables that "point" in the same direction (had principal component values that aligned).

Before performing PCA, though, we remove the second-order factor analysis variables due to the heavy correlation with the original predictor variables and a more difficult interpretation of these variables. Since our goal is to understand the relationship between certain leadership characteristics, keeping these complicated variables might reduce our understanding and make interpretation more difficult.

After performing PCA, we perform k-means clustering on the first four principal components to determine the "groups" of leadership characteristics that have similar perceptions. K-means clustering creates k groups that are the "most similar" to the other observations in their groups, minimizing the between-group variability:

1. Internally competitive, malevolent, status conscious, self-centred, bureaucratic, face saver, autocratic
2. Humane-oriented, modesty

3. Participative, team integrator, inspirational, autonomous, administratively competent, integrity, visionary, diplomatic, decisive, collaborative team orientation, self-sacrifice, performance-oriented



The green cluster seems to describe positive characteristics (e.g., participative, inspirational), while the blue cluster appears to describe negative characteristics (e.g., malevolent, self-centered). From this, we note that these "positive" and "negative" characteristics tend to occur together. The third "cluster" (i.e., humane oriented and modesty) did not closely fit with the other groups, leading us to believe that they may be considered separately from the other characteristics.

After considering the leadership characteristics, we cluster countries based on similar perceptions. We do this by using the variables of the countries transformed into the first four principal components, then running k-means clustering on those components. The clustering results segregate the countries into the following segments:

Country Clusters with Similar Leadership Values

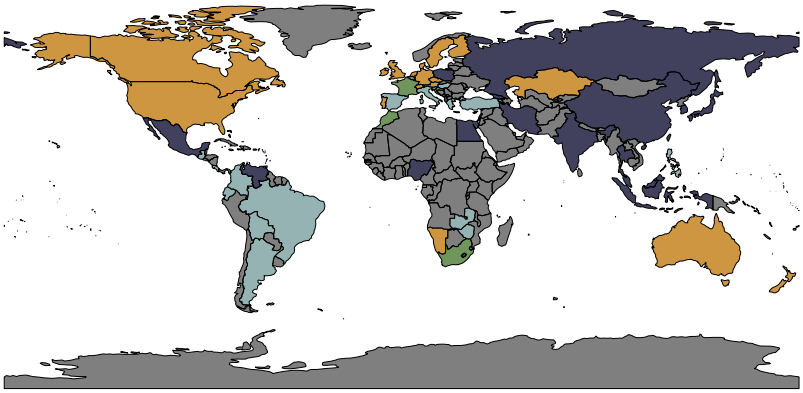


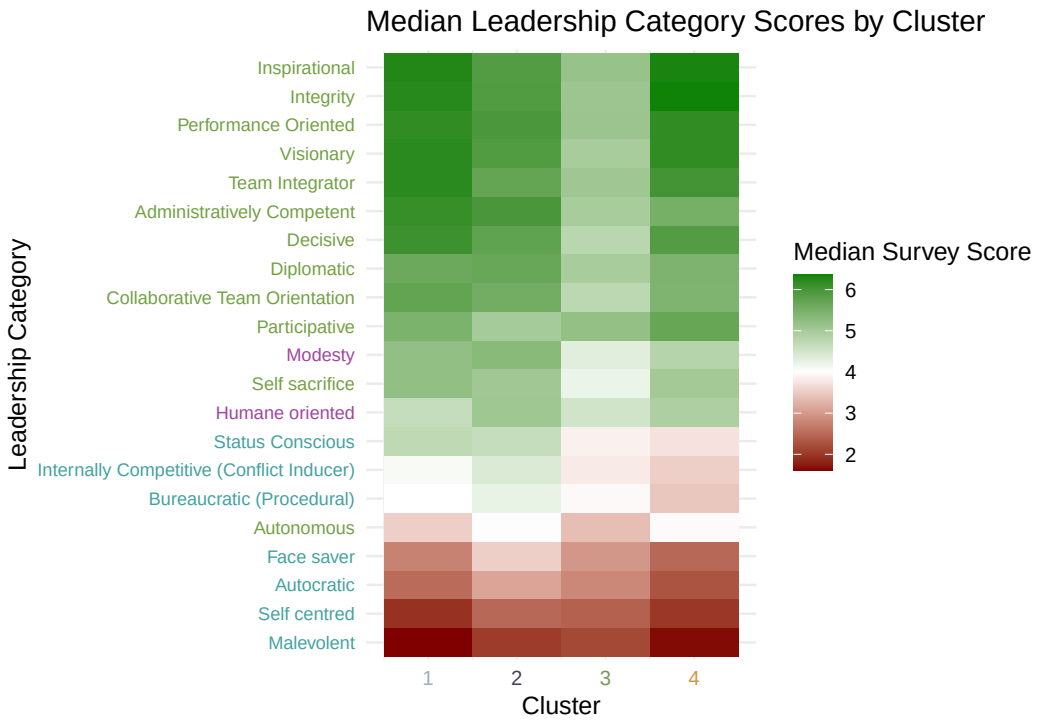
Table 1: Regions with Similar Leadership Perceptions

Cluster 1	Cluster 2	Cluster 3	Cluster 4
Argentina	Albania	France	Australia
Bolivia	China	Morocco	Austria
Brazil	Egypt	Qatar	Canada (English-speaking)
Colombia	Georgia	South Africa (Black Sample)	Czech Republic
Costa Rica	Hong Kong		Denmark
Ecuador	India		England
El Salvador	Indonesia		Finland
Greece	Iran		French Switzerland
Guatemala	Japan		Germany (EAST)
Hungary	Kuwait		Germany (WEST)
Israel	Malaysia		Ireland
Italy	Mexico		Kazakhstan
Philippines	Nigeria		Namibia
Slovenia	Poland		Netherlands
Spain	Russia		New Zealand
Turkey	South Korea		Portugal
Zambia	Taiwan		Singapore
Zimbabwe	Thailand		South Africa (White Sample)
	Venezuela		Sweden
			Switzerland
			USA

Note: East and West Germany belong to the same cluster and are colored as such. South Africa is colored

by the cluster the Black sample belongs in, rather than the White sample. Countries that are not included in the data are colored gray.

From the above grouping, we see that regionality (generally) determines a country's respective cluster and leadership perspectives. Cluster 1 (light blue) tends to describe Latin American and the Mediterranean, Cluster 2 (dark blue) generally includes Asia, and Cluster 4 (gold) describes the Anglo regions and northern Europe. The one cluster that appears to be "out there" is Cluster 3 (green), which does not have a particular region associated with it. To visualize the differences between these clusters, we create a heatmap of the median variable values for each cluster:



From here, we see that, despite the clustering, the countries tend to agree on beneficial and detrimental leadership qualities; the main difference is in the intensity: clusters 1 and 4 seem to have the strongest responses (most positive and most negative), while cluster 3 does not seem to feel terribly strongly about positive leadership characteristics. The few qualities of intrigue, though, are the ones for which some clusters viewed positively (green) while others viewed negatively (red):

- Status Conscious
- Internally Competitive (Conflict Inducer)
- Bureaucratic (Procedural)
- Autonomous

Societal Practices and Values

From the scatterplots above, we note that societal practices and societal values do not always align. For instance, in the case of the Humane Orientation cultural dimension, there does not appear to be any correlation between practices and values at all, with societal practices in this dimension varying from less than 3.5 to more than 5.0 on the survey's seven-point scale even though societal values were rated relatively similarly across all of the surveyed countries. In some other cases, such as for the Uncertainty Avoidance cultural dimension, there appears to be a negative correlation between practices and values, as indicated by

the general trend of scores for practices in this dimension decreasing even as scores for values increase. This negative correlation might suggest that, rather than aligning, societal practices and societal values are in fact in conflict.

To quantify the observations we made from our scatterplots above, we fit simple linear regression models for each of the nine cultural dimensions, using the societal values rating as the predictor and the societal practices rating as the response. In other words, we fit the model:

$$y = \beta_0 + \beta_1 x$$

Where x is the societal values rating and y is the societal practices rating for each of the nine cultural dimensions. Additionally, we performed t-tests for each dimension, to test the hypotheses $H_0 : \beta_1 = 0$ and $H_1 : \beta_1 \neq 0$. For this, we employed the *Bonferroni correction* to change the significance threshold from $\alpha = 0.05$ to $\alpha = \frac{0.05}{9} \approx 0.0056$ since maintaining a significance level of $\alpha = 0.05$ would increase the experiment-wise error rate: $P(\text{Any False Positive}) = 1 - P(\text{No False Positives}) = 1 - 0.95^9 \approx 0.3698$. The table below shows the values of β_1 that we obtained and the corresponding p-values. Cultural dimensions for which the p-value is lower than the corrected $\alpha \approx 0.0056$ are indicated with green shading.

Table 2: Results of Simple Linear Regression by Cultural Dimension

Cultural Dimension	Coefficient Value	p-value
Uncertainty Avoidance	-0.6199	0.0000
Institutional Collectivism	-0.5251	0.0000
Power Distance	-0.4991	0.0006
Future Orientation	-0.4725	0.0009
Humane Orientation	-0.5944	0.0116
Gender Egalitarianism	0.2437	0.0124
Performance Orientation	-0.3459	0.0268
Assertiveness	-0.1507	0.0414
In-group Collectivism	0.4393	0.0991

As shown above, most of the β_1 values obtained were negative, including all four values that were significant at the $\alpha = 0.56\%$ significant level as calculated using the Bonferroni correction. This means that, as the rating for the societal values of a cultural dimension increase, the rating for the societal practices of that cultural dimension actually decrease. In other words, rather than being aligned, societal practices in fact run counter to societal values.

Conclusions

Leadership characteristics are largely clustered based on whether each trait is negative or positive. For example, the traits malevolent and self centered have similar ratings while the traits integrity and inspiration also have similar ratings. Additionally, leadership values tend to be clustered based on their geographic region. This suggests countries within smaller distances of one another have similar more related leadership perspectives.

Our findings indicate that the societal practices and societal values of countries do not align and are instead opposing. Thus, we conclude that a country's societal value does not entail the same practices being practiced among its constituents. This analysis can be useful for government officials trying to develop policies that promote cultural values. If a country places a high value on uncertainty avoidance (the desire to rely on social norms and procedures to avoid unpredictable events), then the same country may undergo unexpected and unconventional events in the future.

Suggestions for Further Research

Regarding the leadership investigation, some further research we may consider is investigating *why* Cluster 3 differs so greatly from the other three clusters. In Cluster 3 we have France, Qatar, Morocco, and South Africa (Black Sample). Understanding why these countries are so different may help us see if there is an underlying link. Similarly, we may want to consider the differences between South Africa's Black and White samples and Germany's East and West samples. Since the survey was conducted in 2006, Apartheid and the Berlin Wall, respectively, may play a role in any potential differences. Ultimately, the above two suggestions relate to us considering the country's framework: looking at other indices (such as a freedom index, government approval ratings, or a happiness index) could shift us from an unsupervised learning situation (just understanding relationships between variables) to a model-building mindset in which we might determine which characteristics seem to have associations with more freedom/approval/happiness.

For both leadership and the societal analyses, further research may be conducting more surveys to look at the potential changes in ideology and perspectives over time to understanding the changing dynamics at both the national and global level. Researchers should also look into the potential causes behind the inverse relationship between social values and practices.

Limitations

In the dataset, there are only 62 countries observed, limiting the significance of the analysis and the confidence we have in clustering countries by geographical similarities. In addition, there is no explanation as to why only certain countries were surveyed. We also recognize that at the time of data collection, the Apartheid and the Berlin Wall potentially played significant factors in the differing valuations of social and leadership values for citizens in South Africa and Germany, respectively. It is notable that East and West Germany are split up and South Africa is also split up by race.