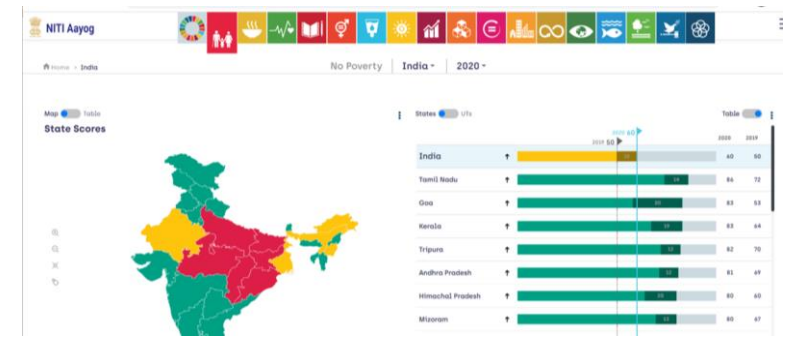


IES 2022



SDGs India Index Analysis using SHAP

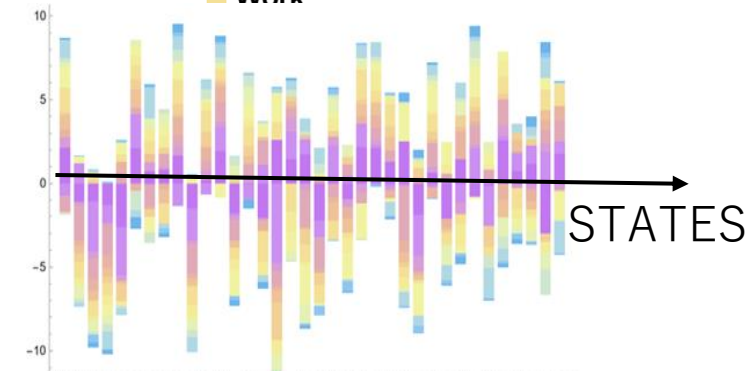
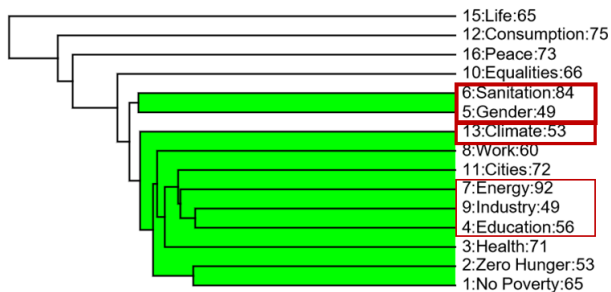
2022/08/11

Prof Takako Hashimoto, Chiba University of Commerce


Prof Yukari Shirota, Gakushuin University

Prof Basabi Chakraborty, Iwate Prefectural University

- Peace
- Life on Land
- Climate Action
- Consumption
- Cities
- Reduced Inequalities
- Industry
- Work
- Energy
- Sanitation
- Gender
- Education
- Health
- Zero Hunger
- No Poverty

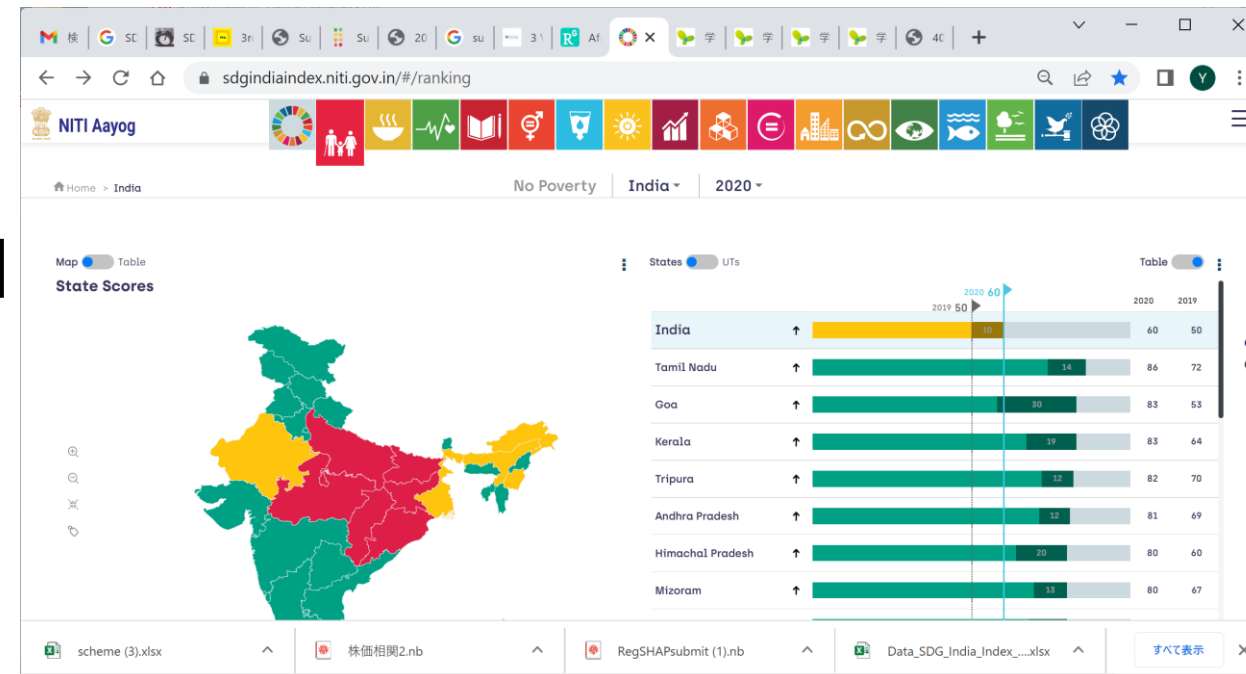


Contents

- 
- 1. Research Objective:
India SDGs Score Analysis by States**
 - 2. State Distribution Similarity by Clustering**
 - 3. Regression Plus SHAP Analysis**
 - 4. Comparison 2018/2022 by SHAP**
 - 5. Discussion**
 - 6. Conclusions**

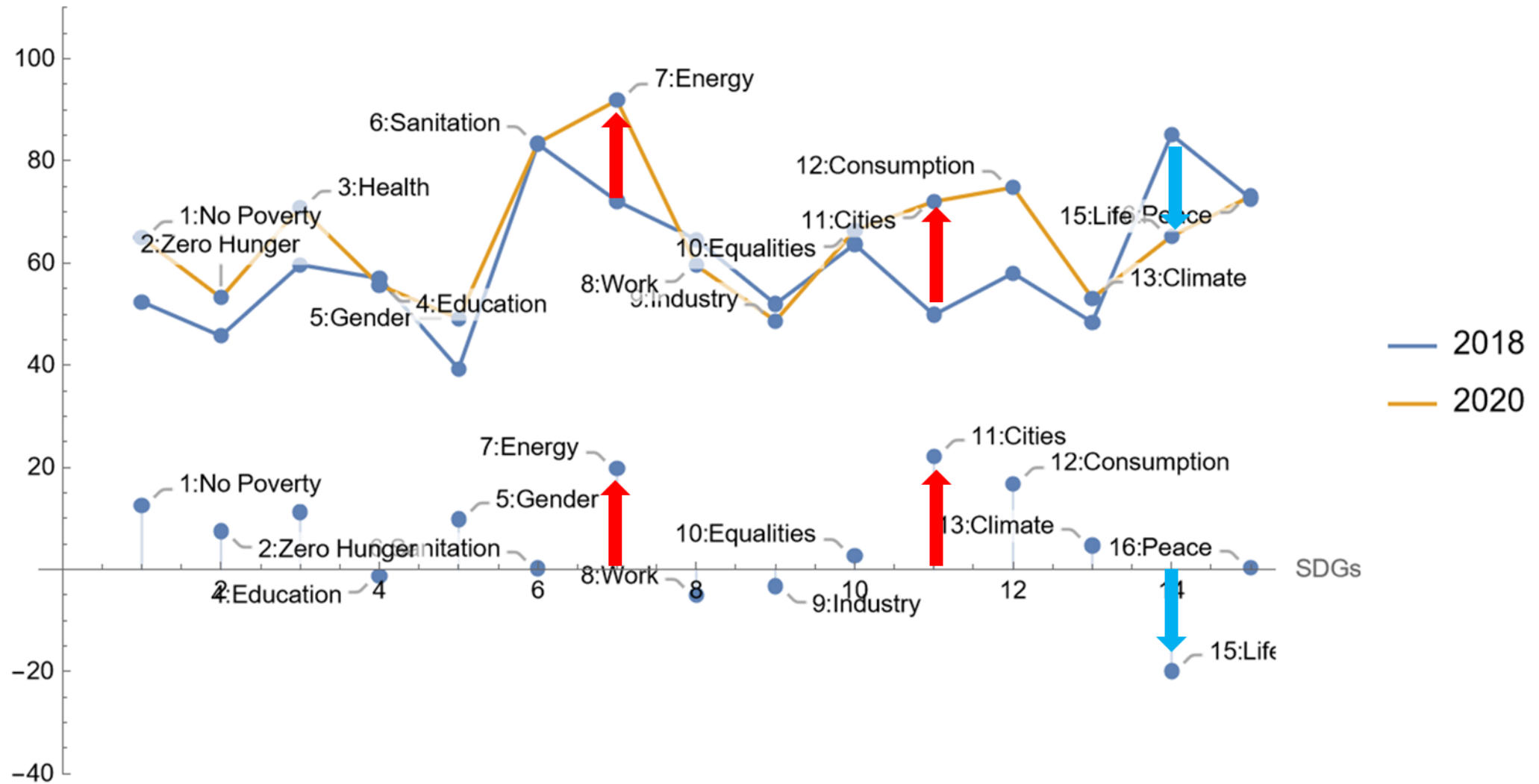
Data **NITI** Aayog's SDG India Index <https://sdgindiaindex.niti.gov.in/#/ranking>

- India **National Institution for Transforming India**
- 2020/2018
- **15 SDGs Indexed** scores except for SDGs14:LifeBelowWater and SDGs17:Partnership.
- SDG1 to SDG17
- 28 States' Data



Comparison between 2018/2020 by Objective Measurement

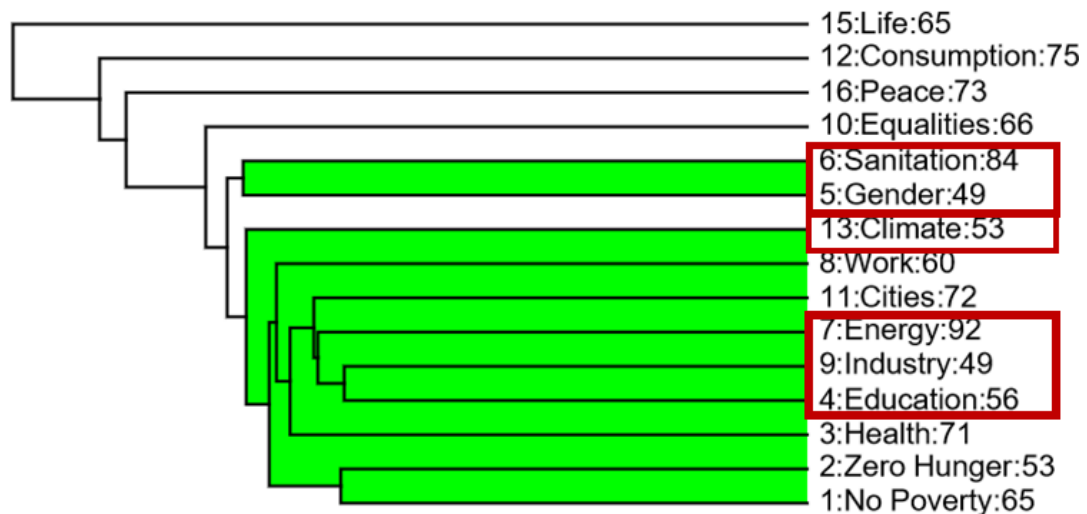
MeanScore



Objective Measurement

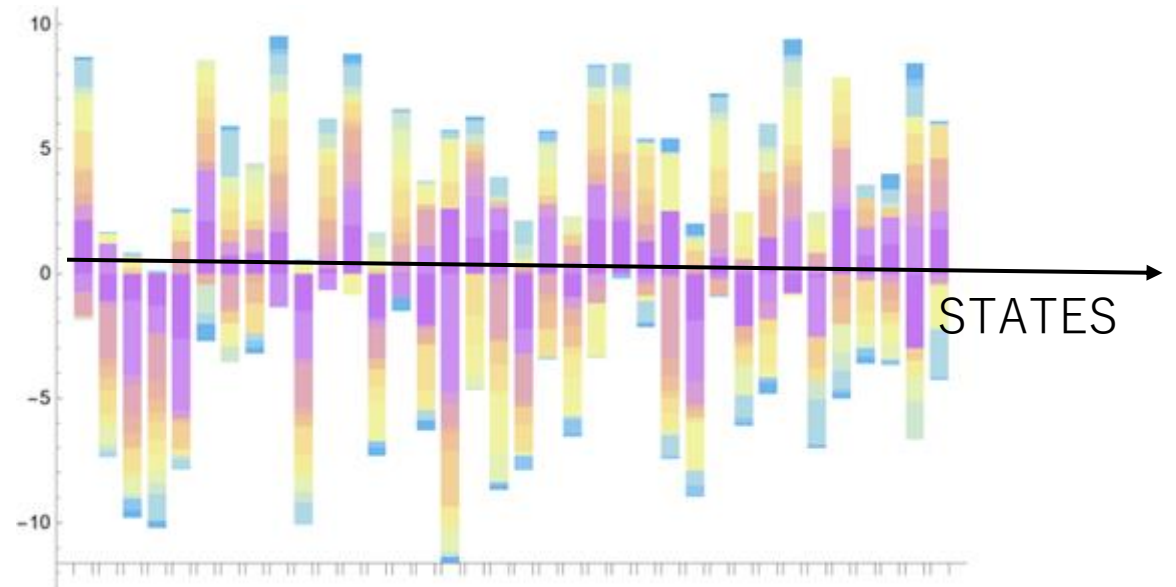
Distribution Similarity among States

- Which SDGs scores have similarity ?
- Hierarchical clustering




Investigate Characteristics of Each State

- Regression + SHAP
- Each state's time series change analysis by SHAP

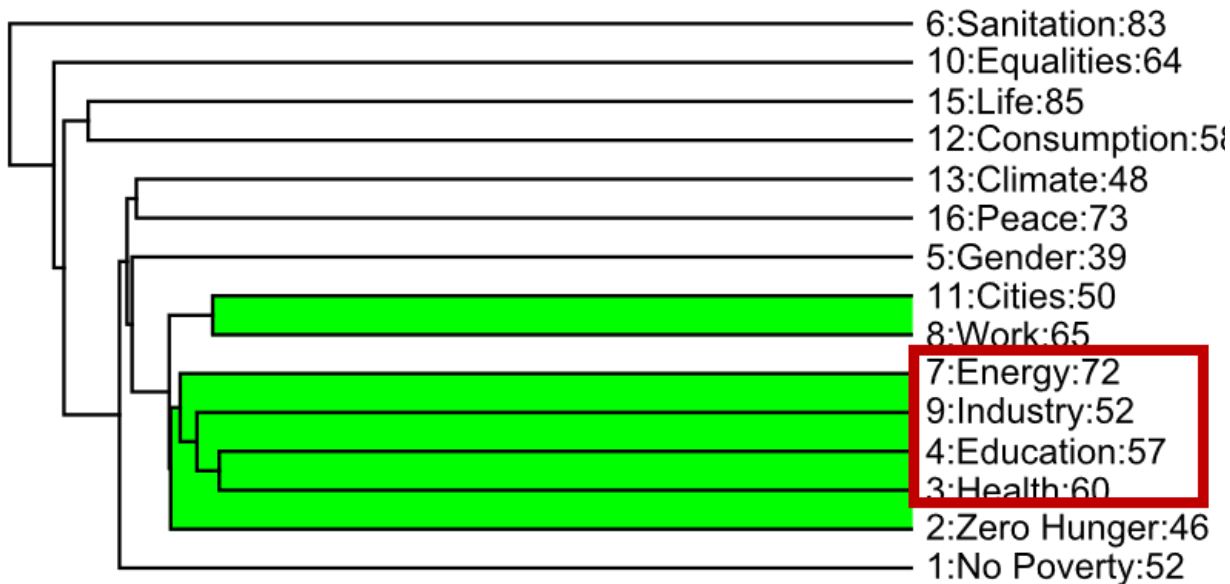


Contents

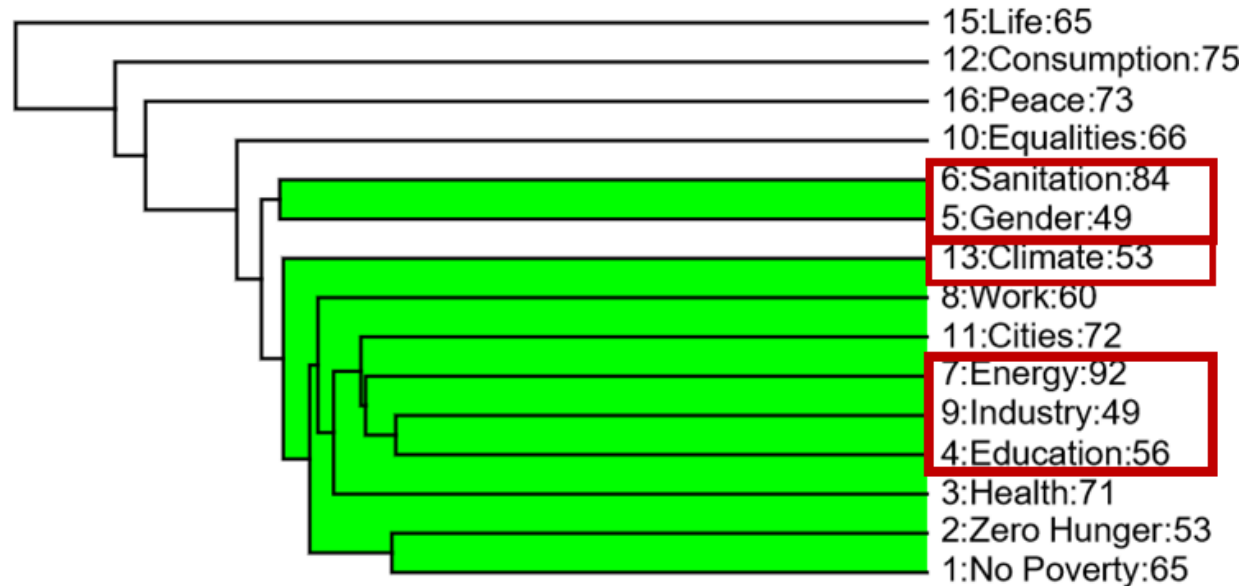
1. Research Objective:
India SDGs Score Analysis by States
-  2. State Distribution Similarity by Clustering
3. Regression Plus SHAP Analysis
4. Comparison 2018/2022 by SHAP
5. Discussion
6. Conclusions

Hierarchical Clustering for Distribution Similarity of 28 States

2018



2020

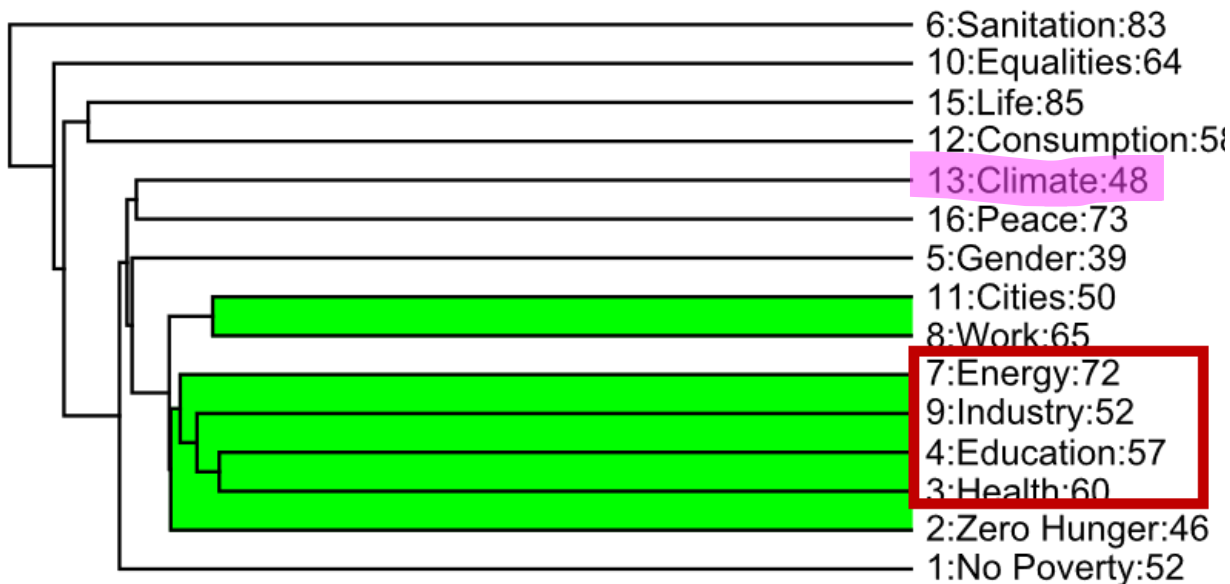


The most similarity set in both 2018 and 2020 is {4:Education, 9:Industry, and 7:Energy}.

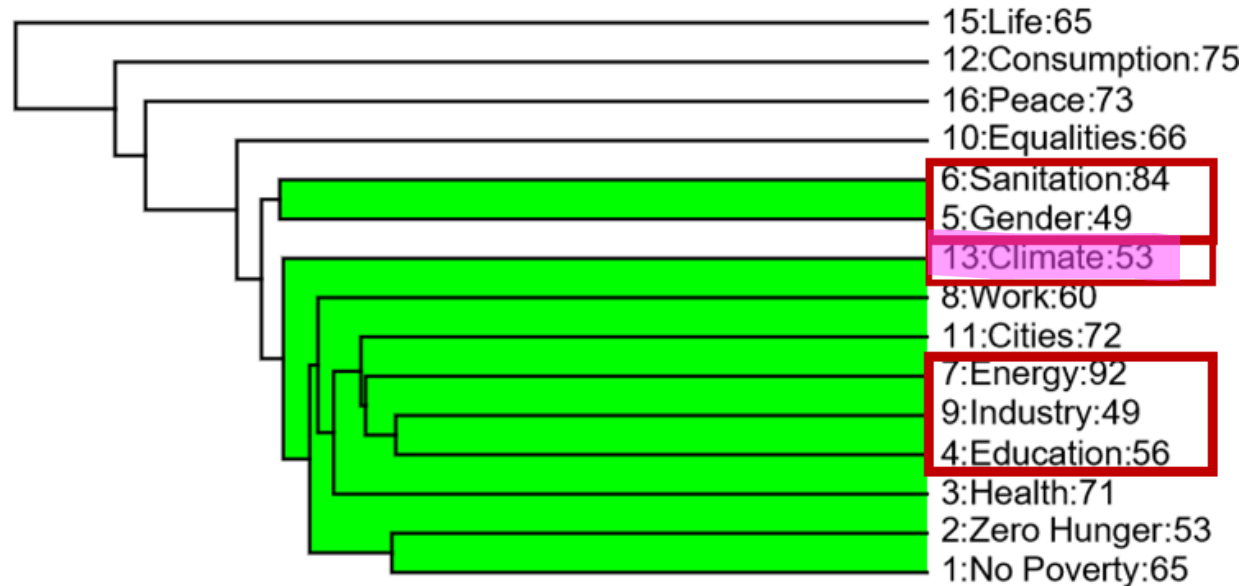
Which SDGs were changed ?

13:Climate became of importance; the member of the large cluster of {4:Education, 9:Industry, and 7:Energy}.

2018



2020




The most impressive difference was **6:Sanitation** became the nearest one to **5:Gender** in **2020**, although there was no similarity of this in 2018

Relation between Sanitation and Gender score increase

- Sanitation improvement especially toilet increase must improve the gender related things.
- If there is no girls' toilets, girls cannot easily go to school.

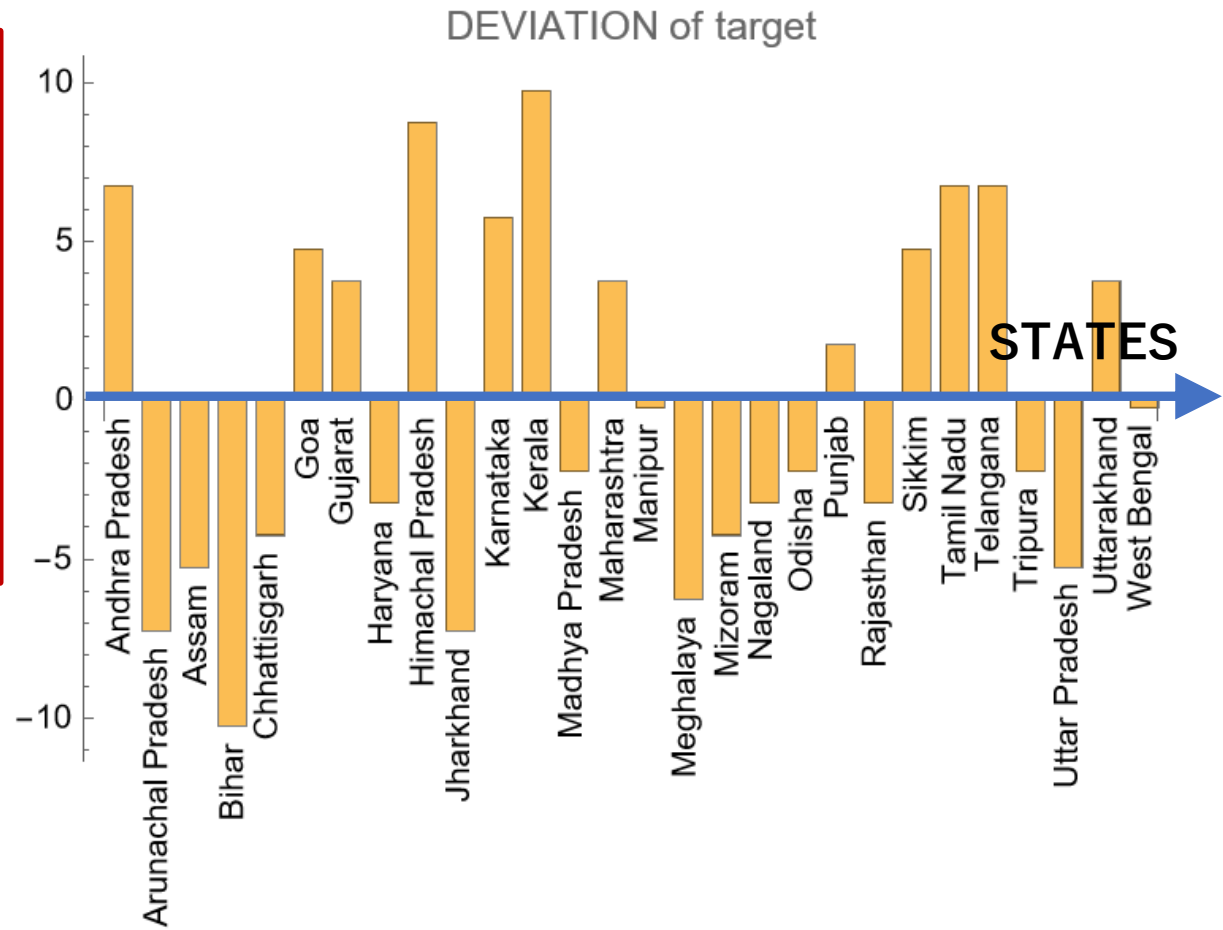


Contents

1. Research Objective:
India SDGs Score Analysis by States
2. State Distribution Similarity by Clustering
-  3. Regression Plus SHAP Analysis
4. Comparison 2018/2022 by SHAP
5. Discussion
6. Conclusions

Regression Data

- **Target** variable: SDGs Index score (composite score) of each state
- **Predictor** variables: 15 SDGs scores of each state

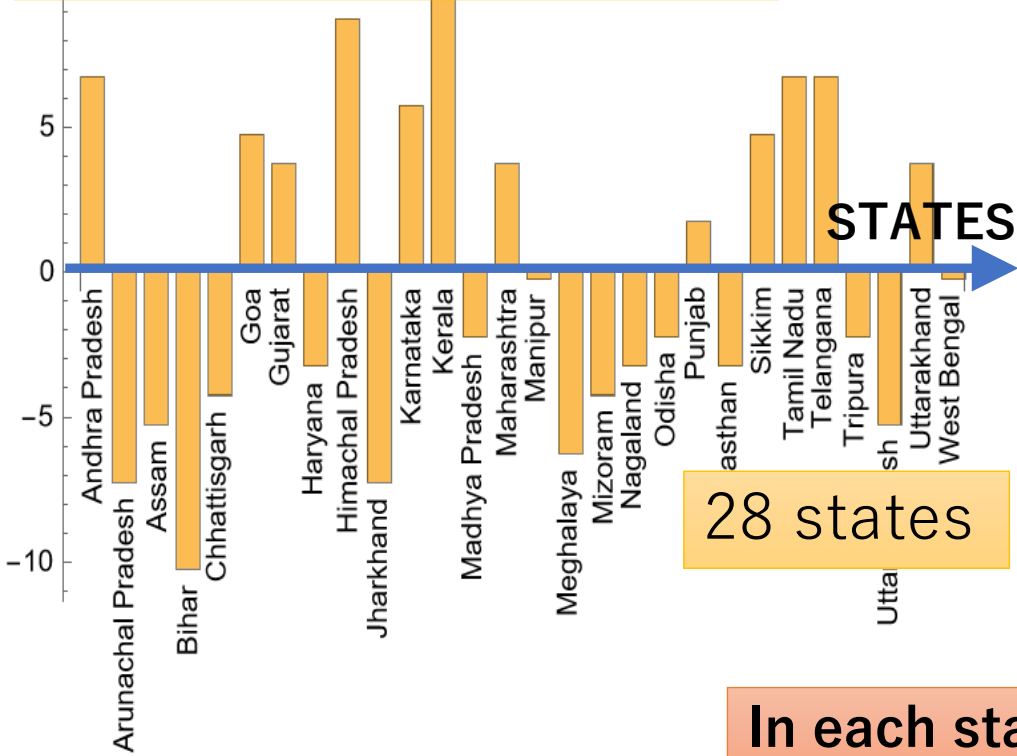


Deviation divided into SHAP values

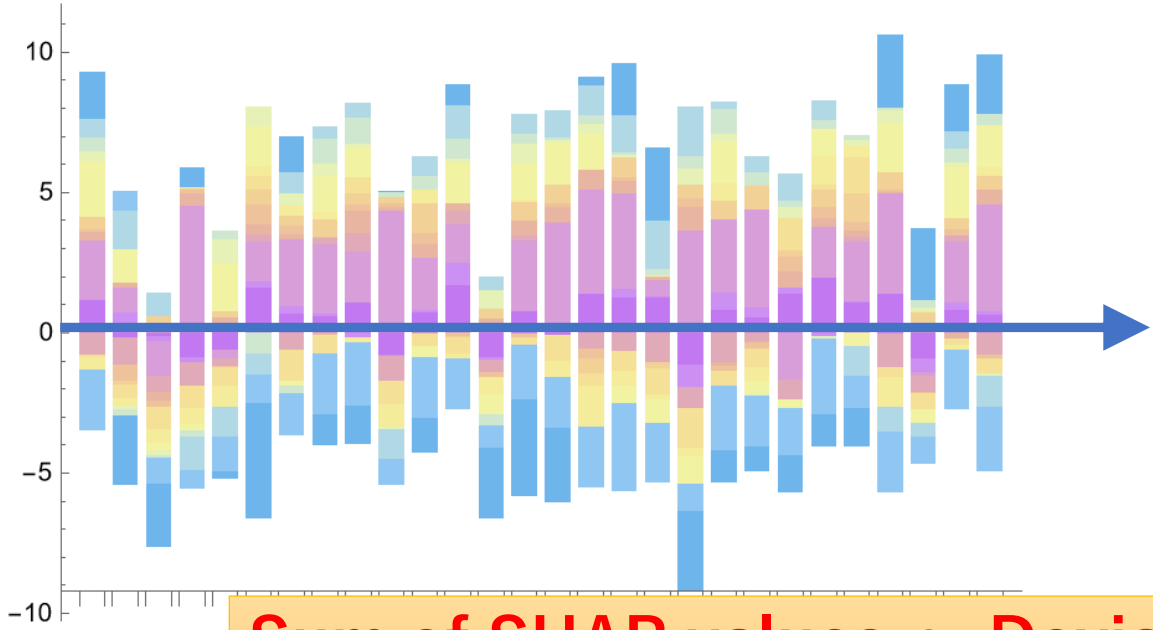
- 16:Peace
- 15:Life
- 13:Climate
- 12:Consumption
- 11:Cities
- 10:Equalities
- 9:Industry
- 8:Work
- 7:Energy
- 6:Sanitation
- 5:Gender
- 4:Education
- 3:Health
- 2:Zero Hunger
- 1:No Poverty

Deviation of target values

Deviation is **target - average**



A state has 15 predictor_SHAP values

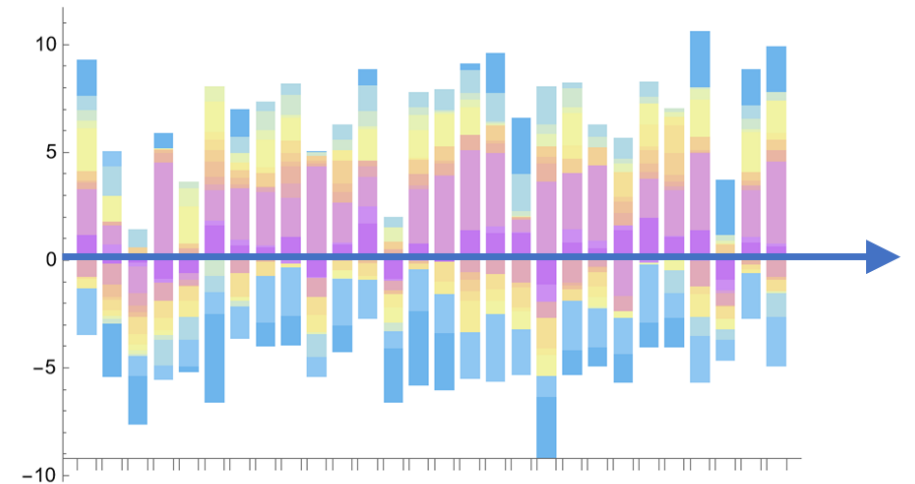
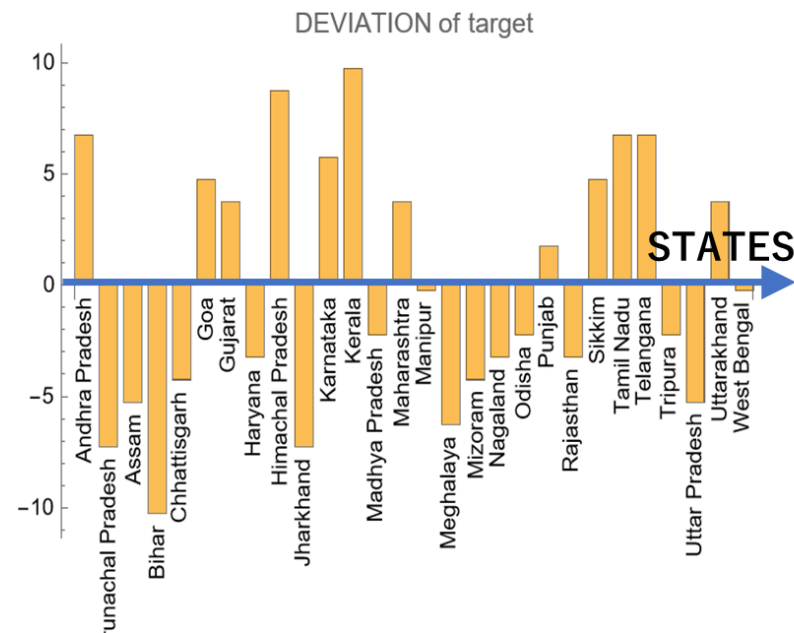


Sum of SHAP values \approx Deviation

In each state, sum of 15 SHAP values becomes target deviation

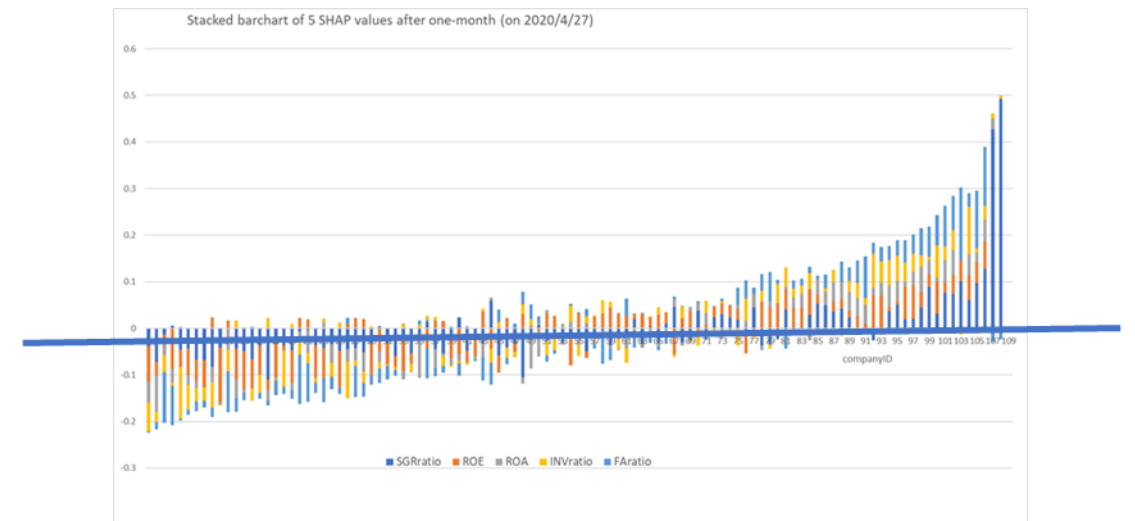
What is SHAP

- Each predictor's contribution to target value
- What is meaning of deviation ?
 - > Deviation from the average of the India states.
UPPER/LOWER than the average score
- SHAP value may be negative.




Advantage of SHAP

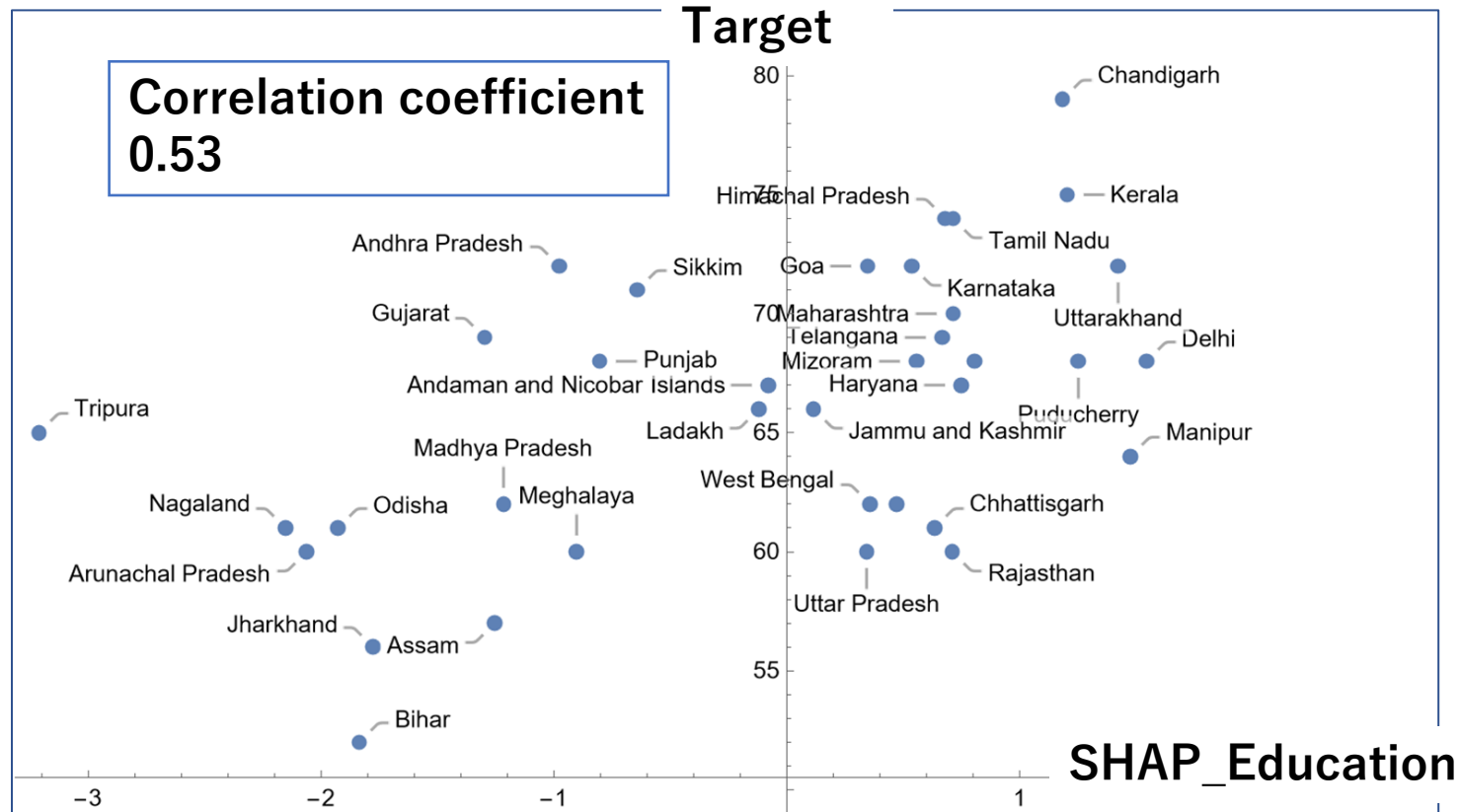
- Each state's **characteristics** evaluated
- In **the** state's behavioral structure, each predictor's contribution to target can be evaluated
- In AI regressions, SHAP approach widely used
- Applicable to all other application fields after regression analysis



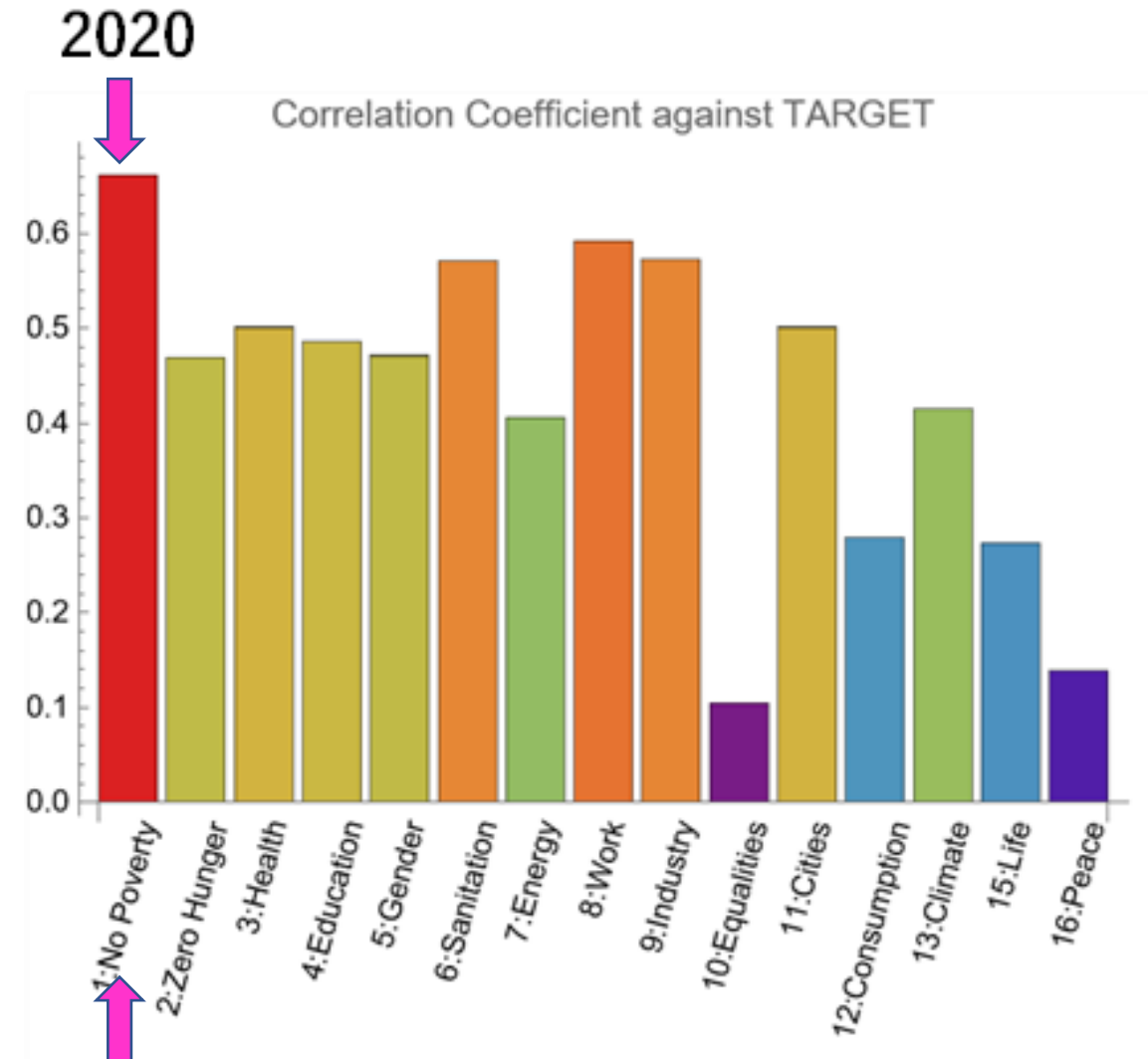
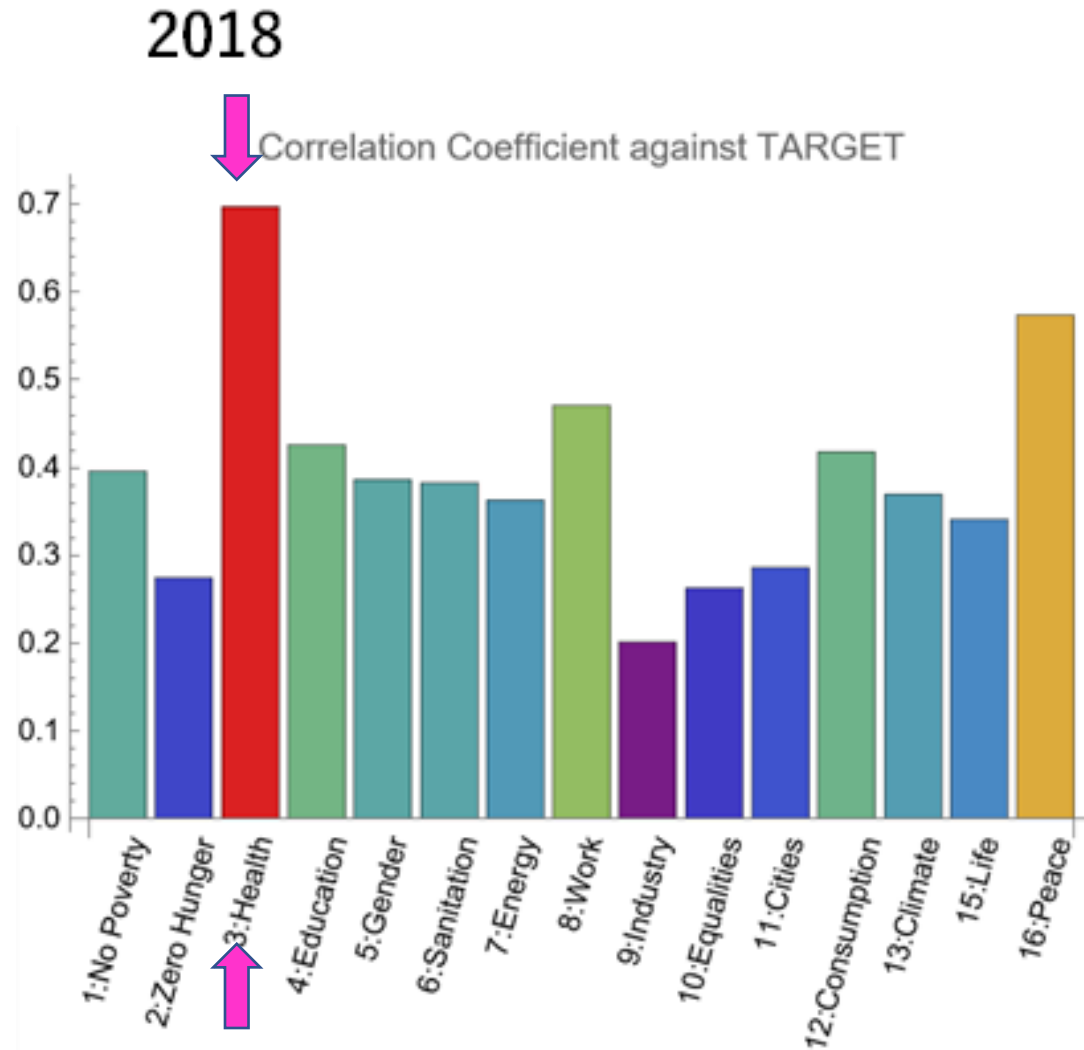
Contents

1. Research Objective:
India SDGs Score Analysis by States
2. State Distribution Similarity by Clustering
3. Regression Plus SHAP Analysis
-  4. Comparison 2018/2022 by SHAP
5. Discussion
6. Conclusions

- Which predictor is the **most important factor** to increase the target ?
- Which SHAP predictor is the **highest correlation** to the target ?

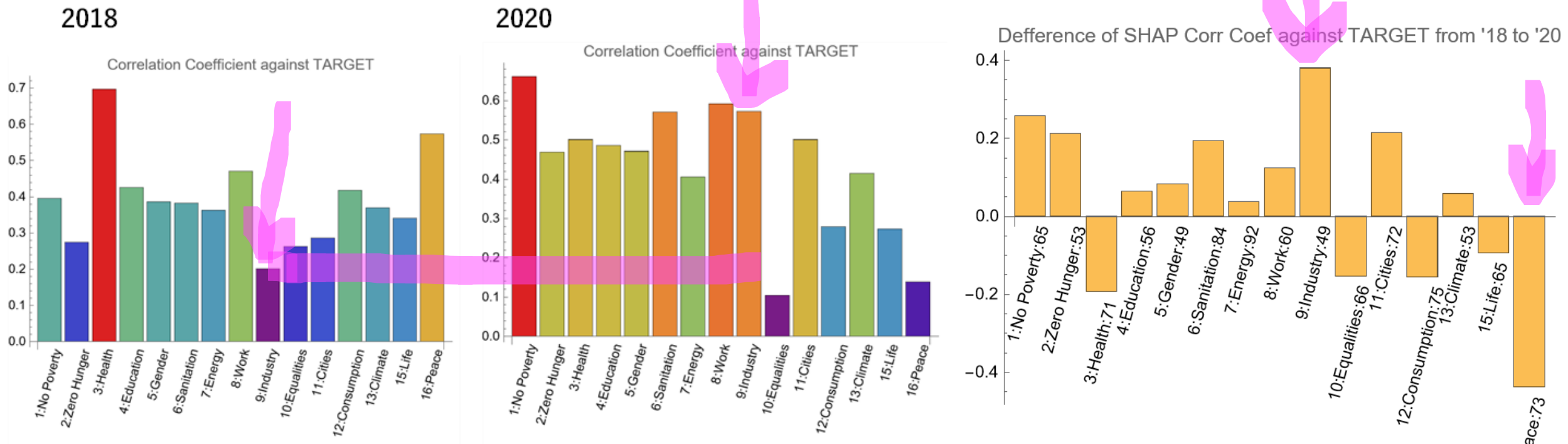


Correlation with 15 SHAP values and target Comparison 2018/2020




Correlation with 15 SHAP values and target Comparison 2018/2020

- In 2020, coefficients became higher as a whole, because the composite score increased
- The highest increased one is 9: **Industry**.
- The largest decreased one is 16: **Peace**.

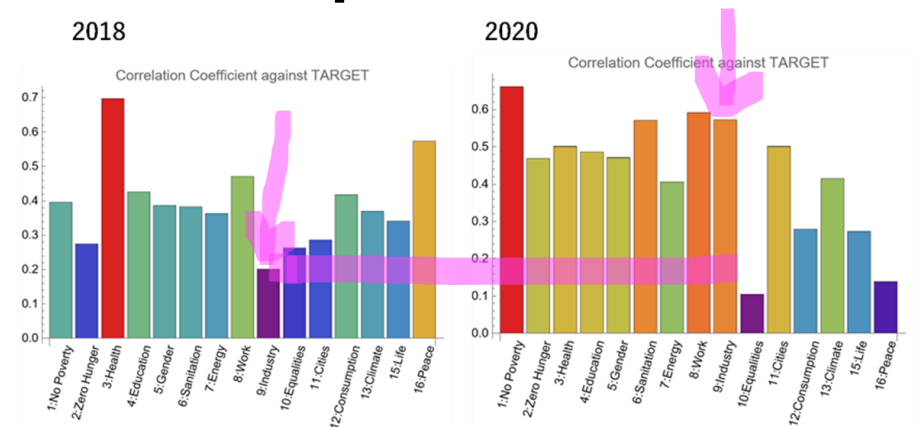


Contents


1. Research Objective:
India SDGs Score Analysis by States
2. State Distribution Similarity by Clustering
3. Regression Plus SHAP Analysis
4. Comparison 2018/2022 by SHAP
-  5. Discussion
6. Conclusions

Important factor to increase target value

- SHAP value reflects its state characteristics
 - intrinsic contribution of each SDGs score in each state
- In 2020, correlation coefficients of **16:Peace** and **3:Health** **decreased**, which does not mean a decline in the raw data, but rather a decrease, compared to **other** predictor variables.
- Even if the raw SDGs score increased, other predictors' contributions were still larger.



Contents

1. Research Objective:
India SDGs Score Analysis by States
2. State Distribution Similarity by Clustering
3. Regression Plus SHAP Analysis
4. Comparison 2018/2022 by SHAP
5. Discussion
-  6. Conclusions

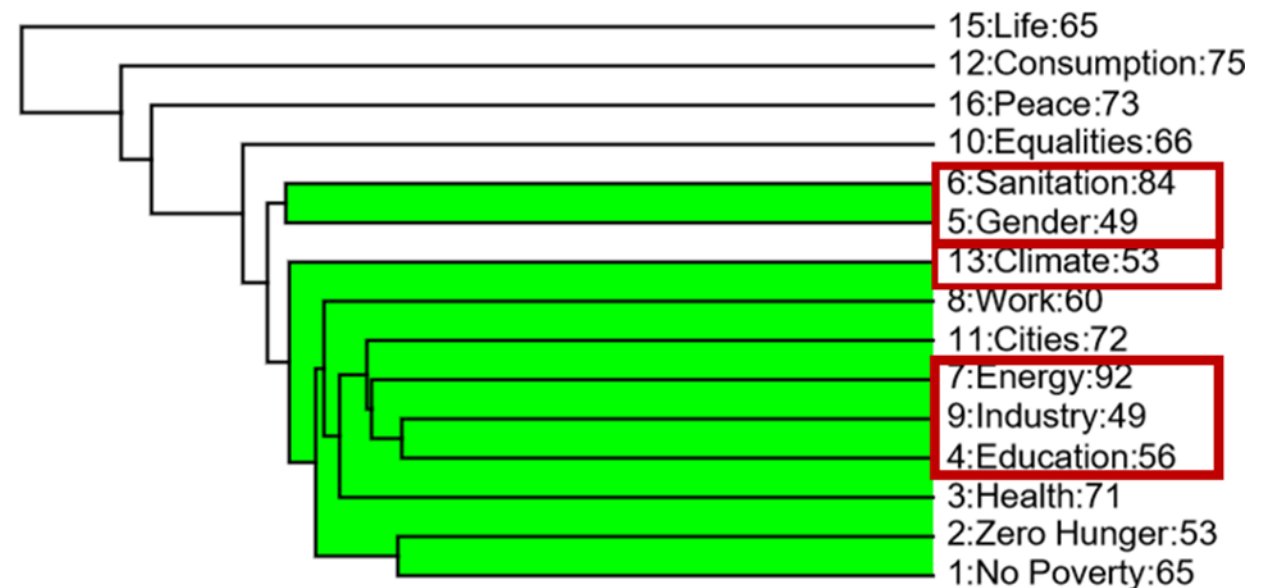
Conclusions

- India SDGs Index analysis by 28 states in 2018/2020

State Distribution Similarity by Clustering

We found

1. Tight relationship of **Energy/Industry/Education**
In the energy-higher state, industry is also higher
2. Relationship between **Sanitation/Gender** appeared in 2020
3. **Climate** score joined in INDUSTRY cluster in 2020



Conclusions

Regression Plus SHAP Analysis

In 2020, important factors to increase SDGs Index (target)

1. No Poverty
2. Work
3. Industry
4. Sanitation

- SHAP can measure each state's characteristics
- In the state, the score's intrinsic contribution to the target

Correlation Coefficient between target and 15 SDGs scores

