Feature Impact on Global Happiness Scores

Ross Brinkerhoff
Capstone II | May 2024

Kawin Harasai via Unsplash

“The *World Happiness Report* reflects a worldwide demand for more attention to happiness and well-being as criteria for government policy.”

—WHR website

Raw Data from WHR

- Data for country and year, ranging from 2005 to 2022
- 165 of 195 countries
- 10 global regions
- 2,199 Rows
- 13 features including the target

Data Features

- Country Name
- Region
- Year
- Log GDP
- Healthy Life Expectancy

- Social Support
- Freedom to make life choices
- Generosity
- Perceptions of corruption
- Positive Affect
- Negative Affect
- Confidence in National Government

Target Variable

- Life Ladder Score 0-10

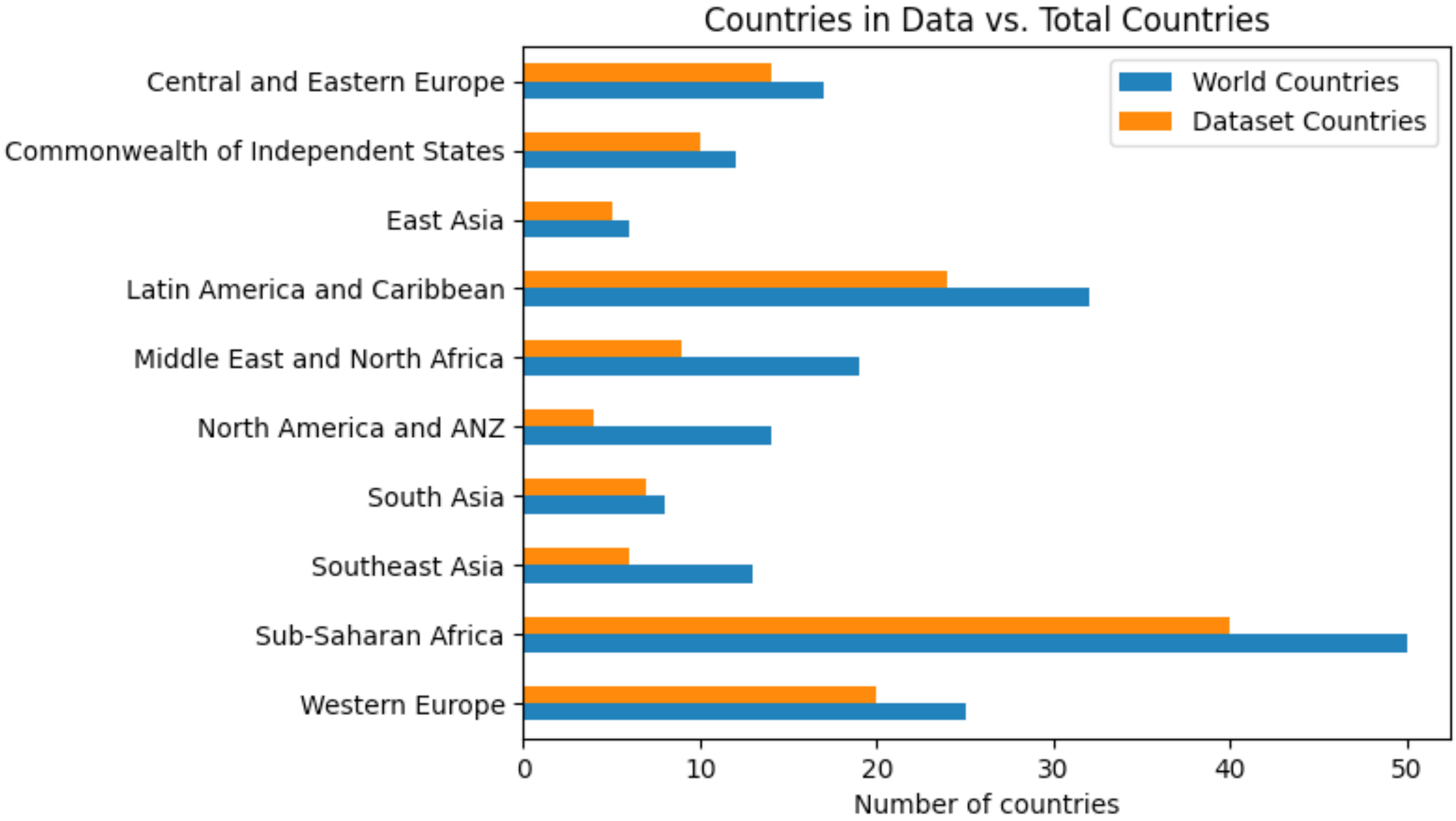
Why feature impact?

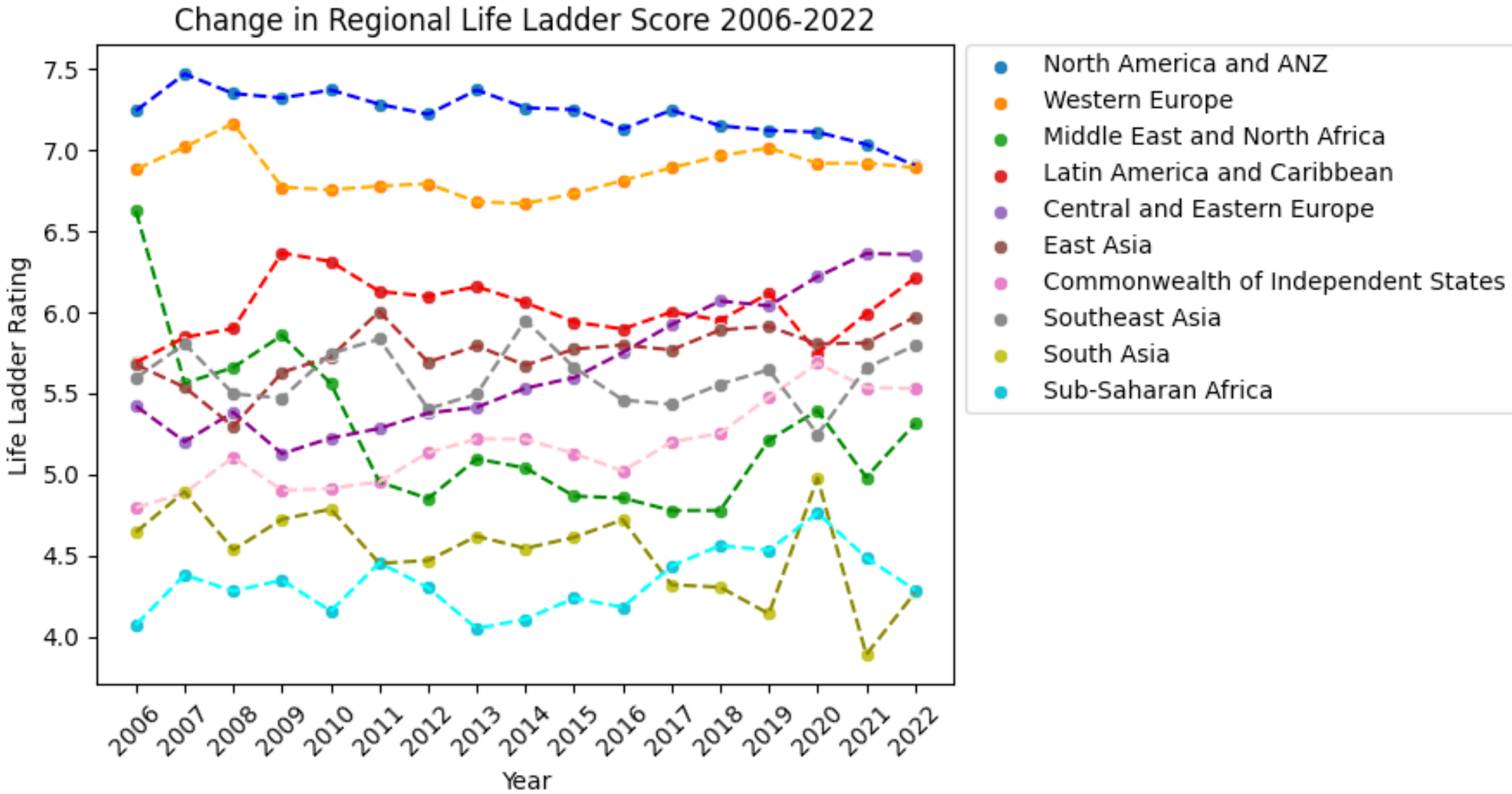
- Governments / stakeholders always have limited attention and resources

- Which features have the most impact and can serve as a focus for action?

Data Wrangling – Missing Survey Data

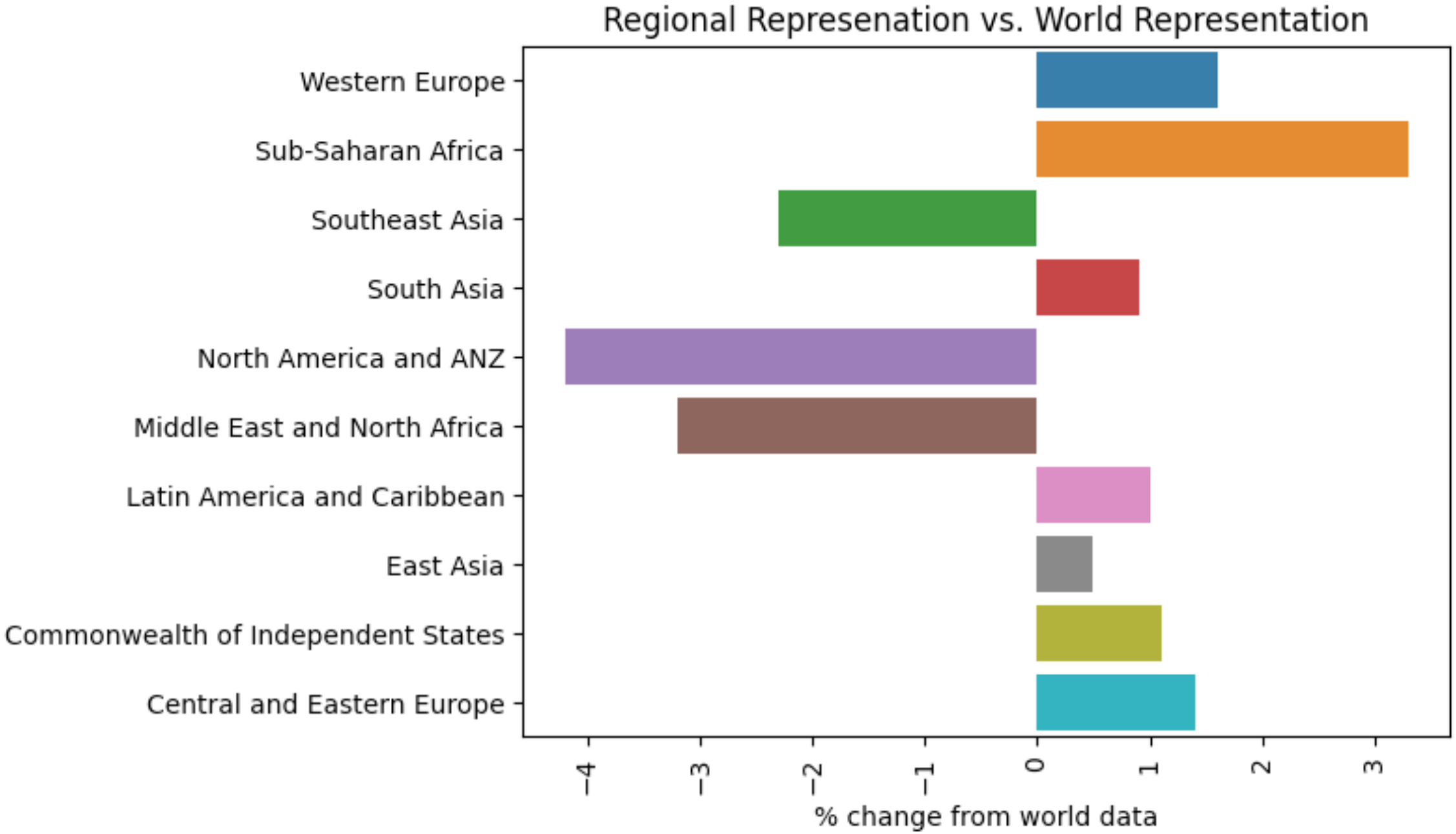
Country	Year	Life Ladder	Perceptions of Corruption	Confidence in Government
China	2006	4.560495	NaN	NaN
China	2007	4.862862	NaN	NaN
China	2008	4.846295	NaN	NaN
China	2009	4.454361	NaN	NaN
China	2010	4.652737	NaN	NaN

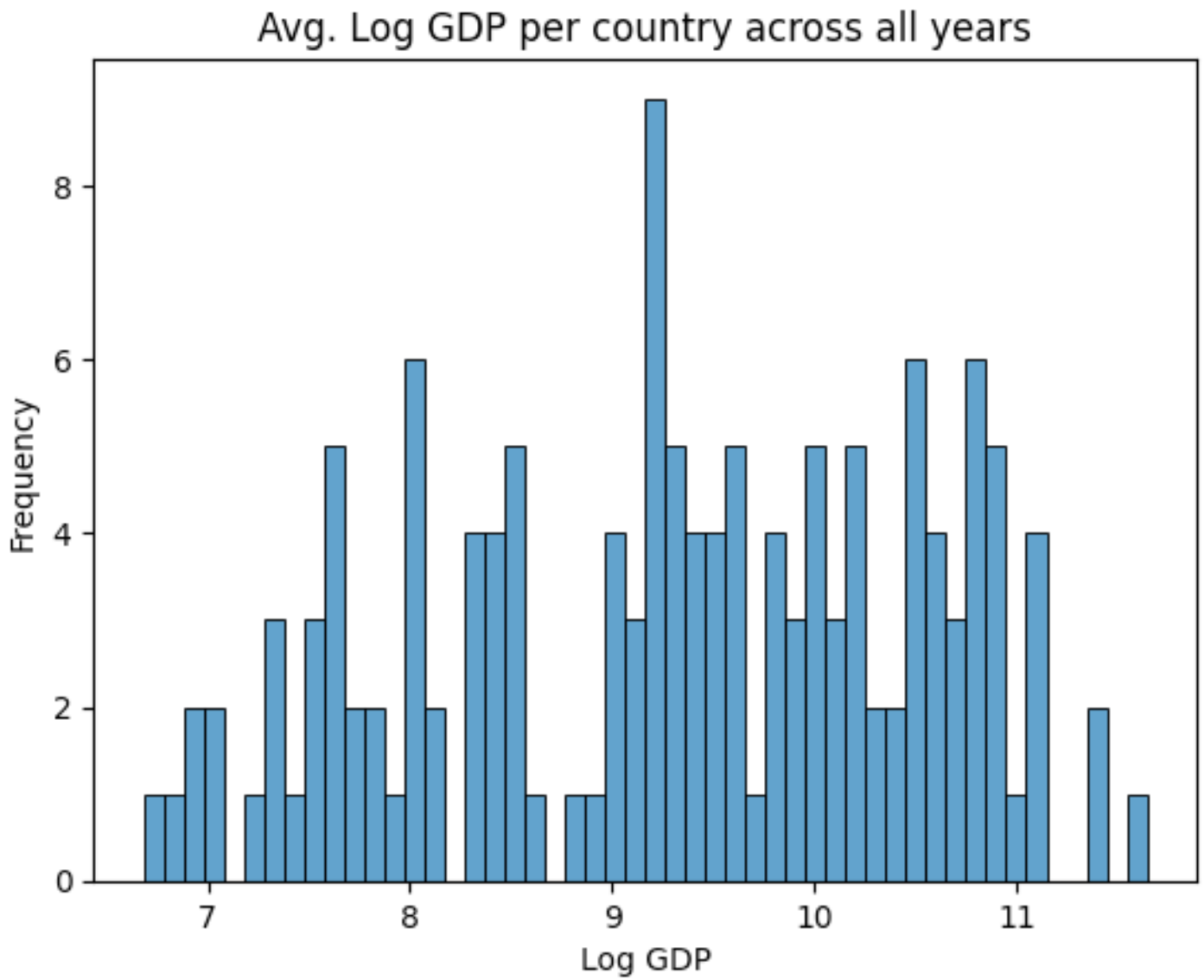

Countries in Data vs. Total Countries
World Countries
Dataset Countries
Central and Eastern Europe
Commonwealth of Independent States
East Asia
Latin America and Caribbean
Middle East and North Africa
North America and ANZ
South Asia
Southeast Asia
Sub-Saharan Africa
Western Europe
0
10
20
30
40
50
Number of countries


Change in Regional Life Ladder Score 2006-2022
Life Ladder Rating
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
Year
North America and ANZ
Western Europe
Middle East and North Africa
Latin America and Caribbean
Central and Eastern Europe
East Asia
Commonwealth of Independent States
Southeast Asia
South Asia
Sub-Saharan Africa

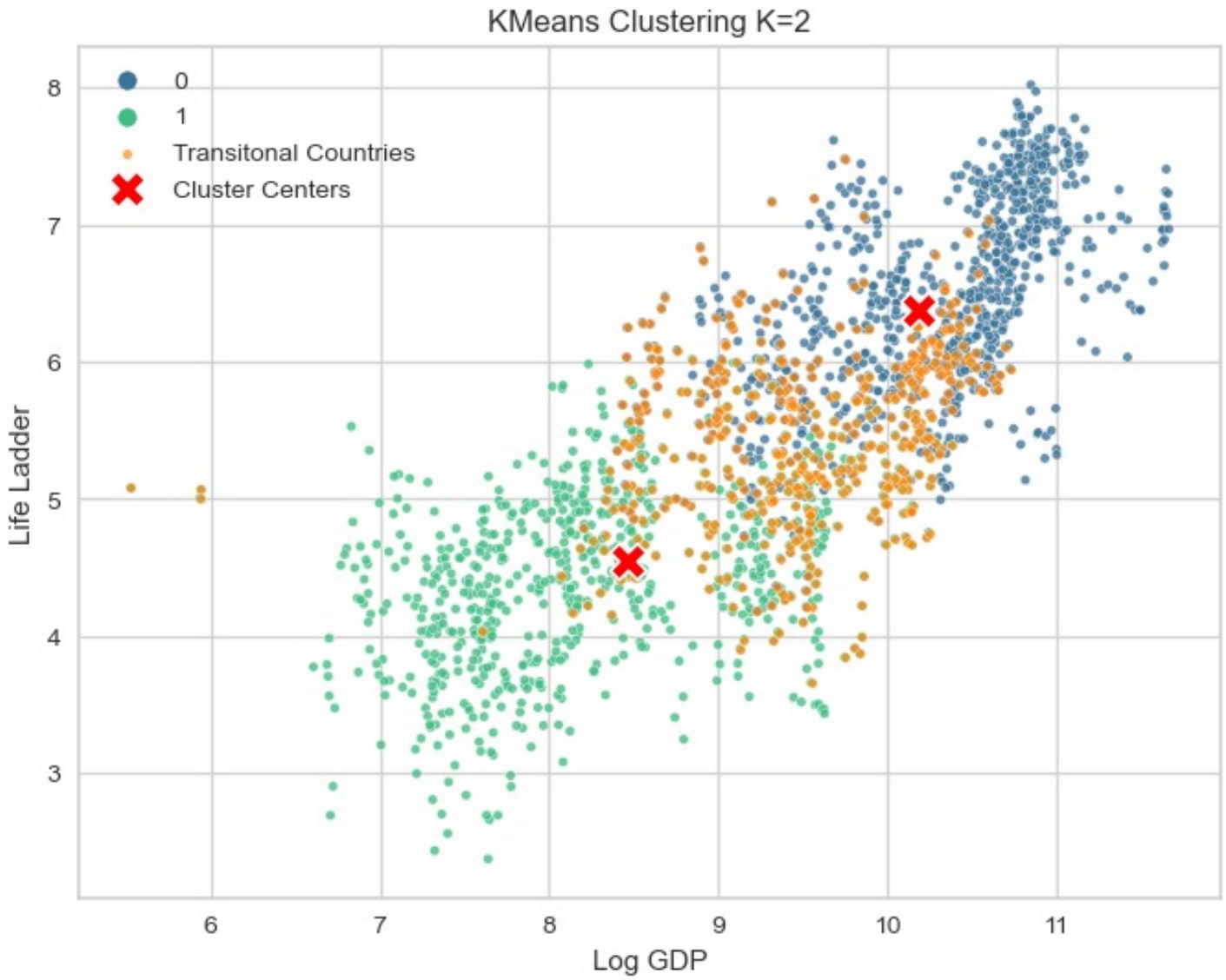
Feature Correlation

Feature	Correlation with Target
Log GDP Per Capita	0.8
Social Support	0.72
Healthy Life Expectancy	0.74
Freedom to Make Life choices	0.57
Positive Affect	0.52

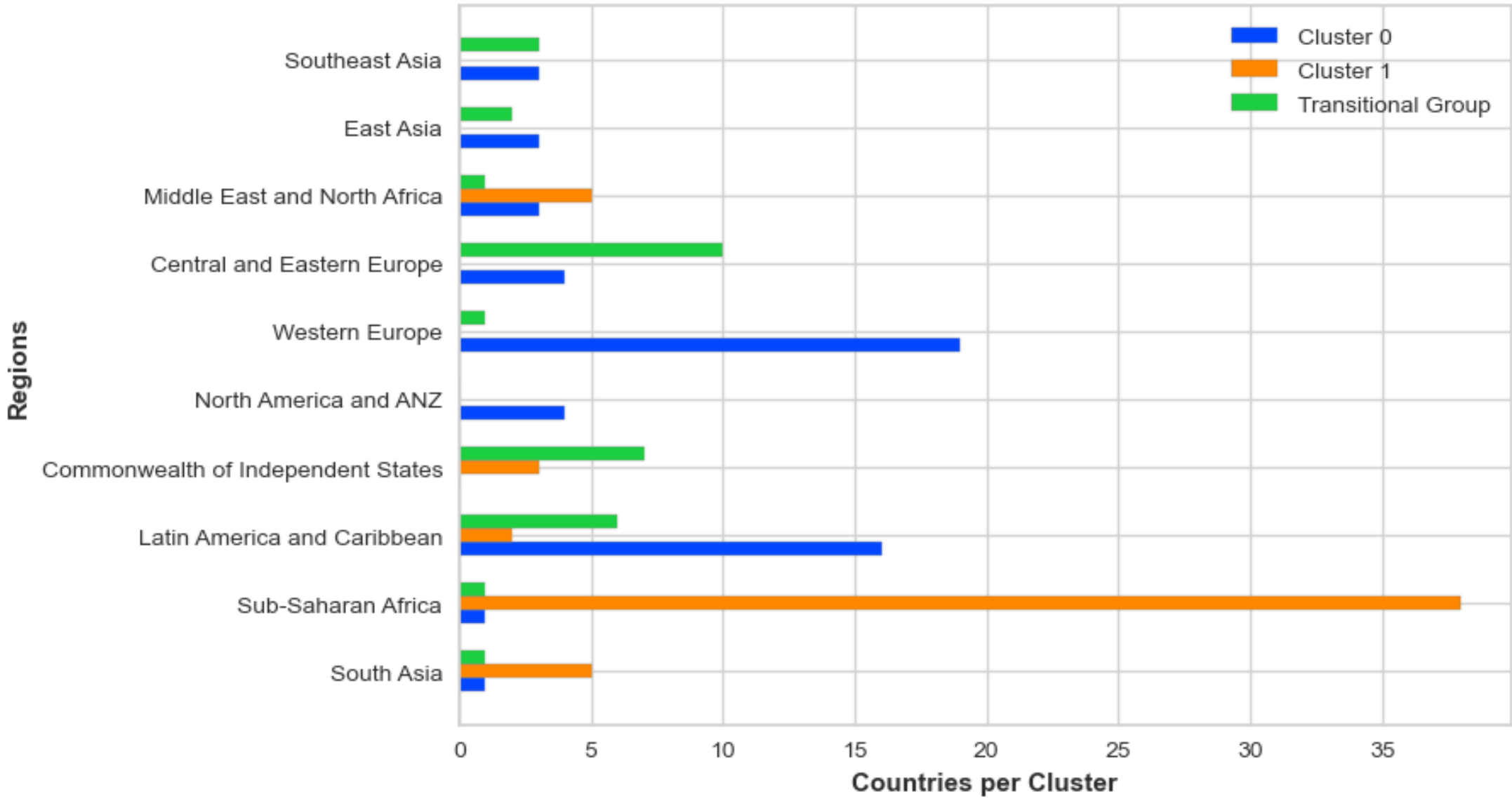

Regional Represenation vs. World Representation
Western Europe
Sub-Saharan Africa
Southeast Asia
South Asia
North America and ANZ
Middle East and North Africa
Latin America and Caribbean
East Asia
Commonwealth of Independent States
Central and Eastern Europe
−4
−3
−2
−1
0
1
2
3
% change from world data

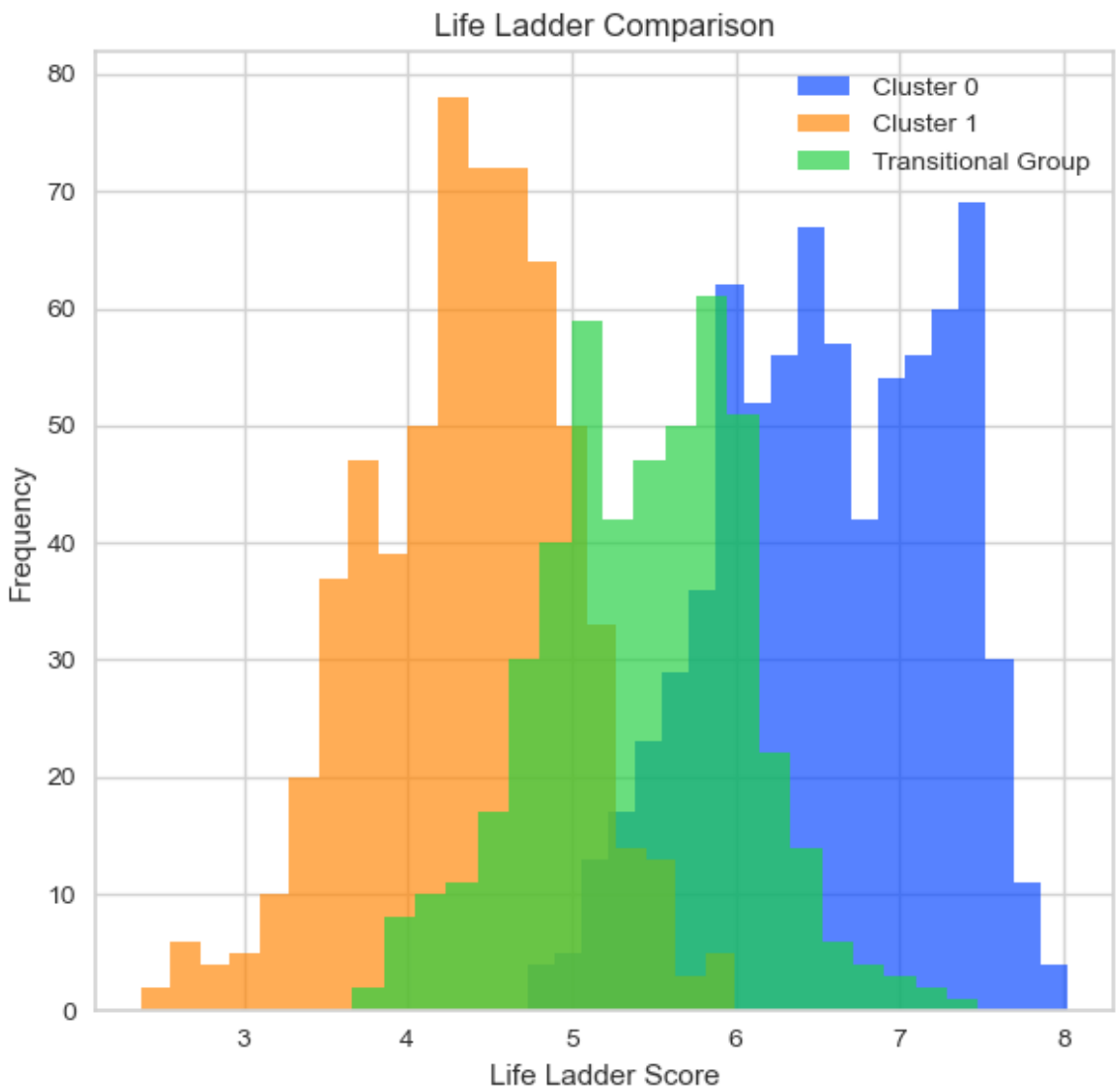

Avg. Log GDP per country across all years
Frequency
0
2
4
6
8
7
8
9
10
11
Log GDP

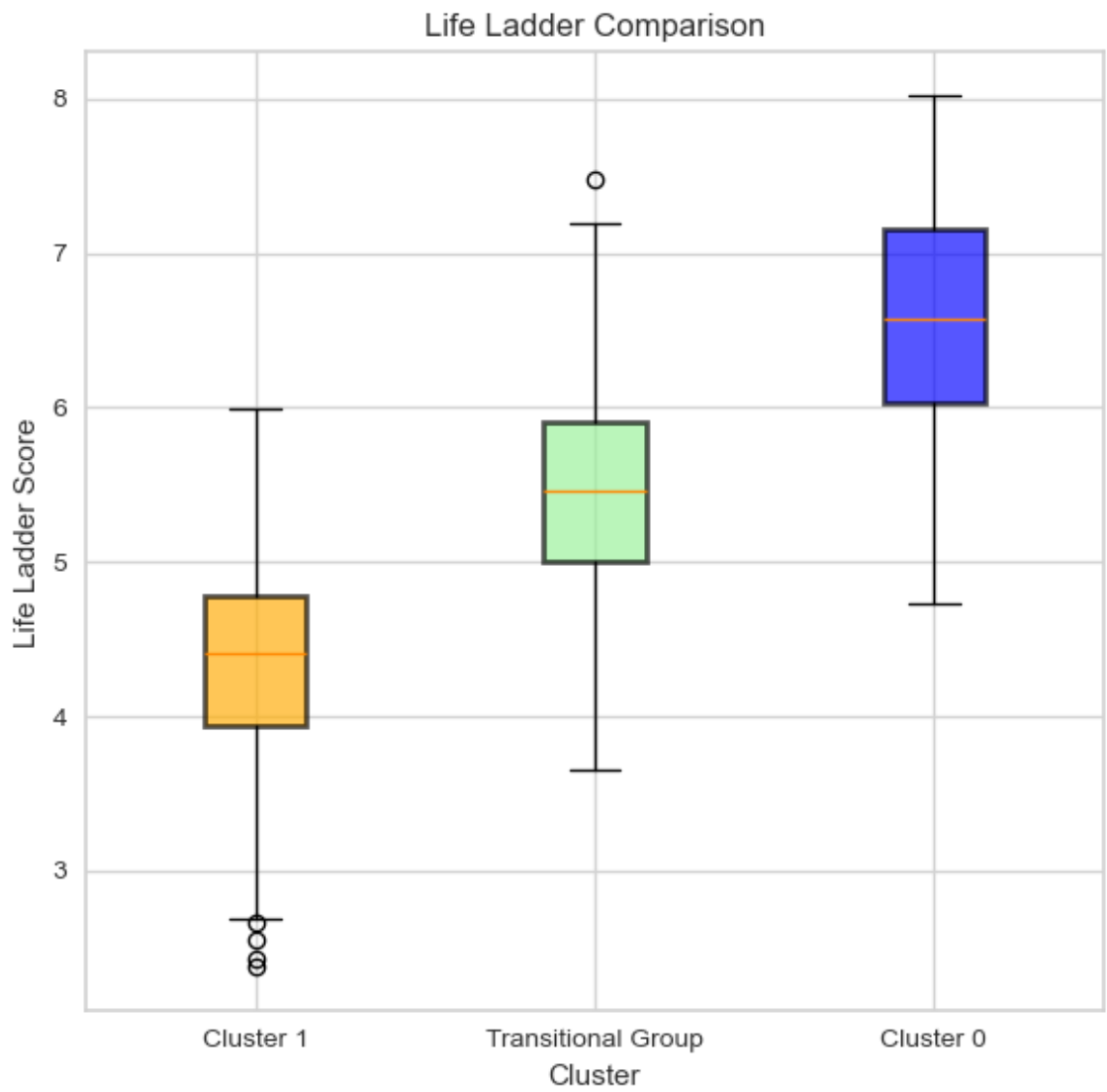
Cluster Analysis with K-Means

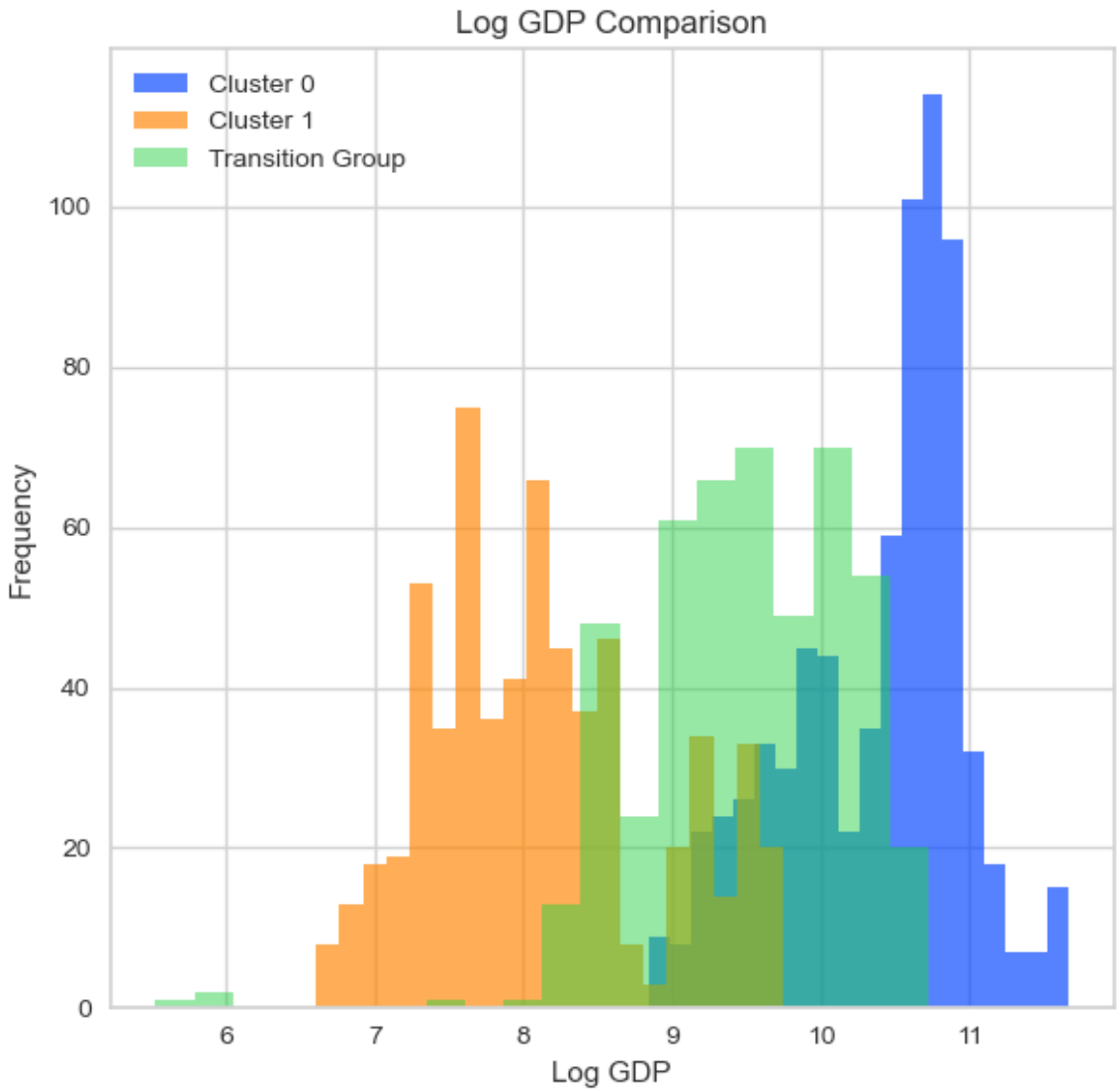


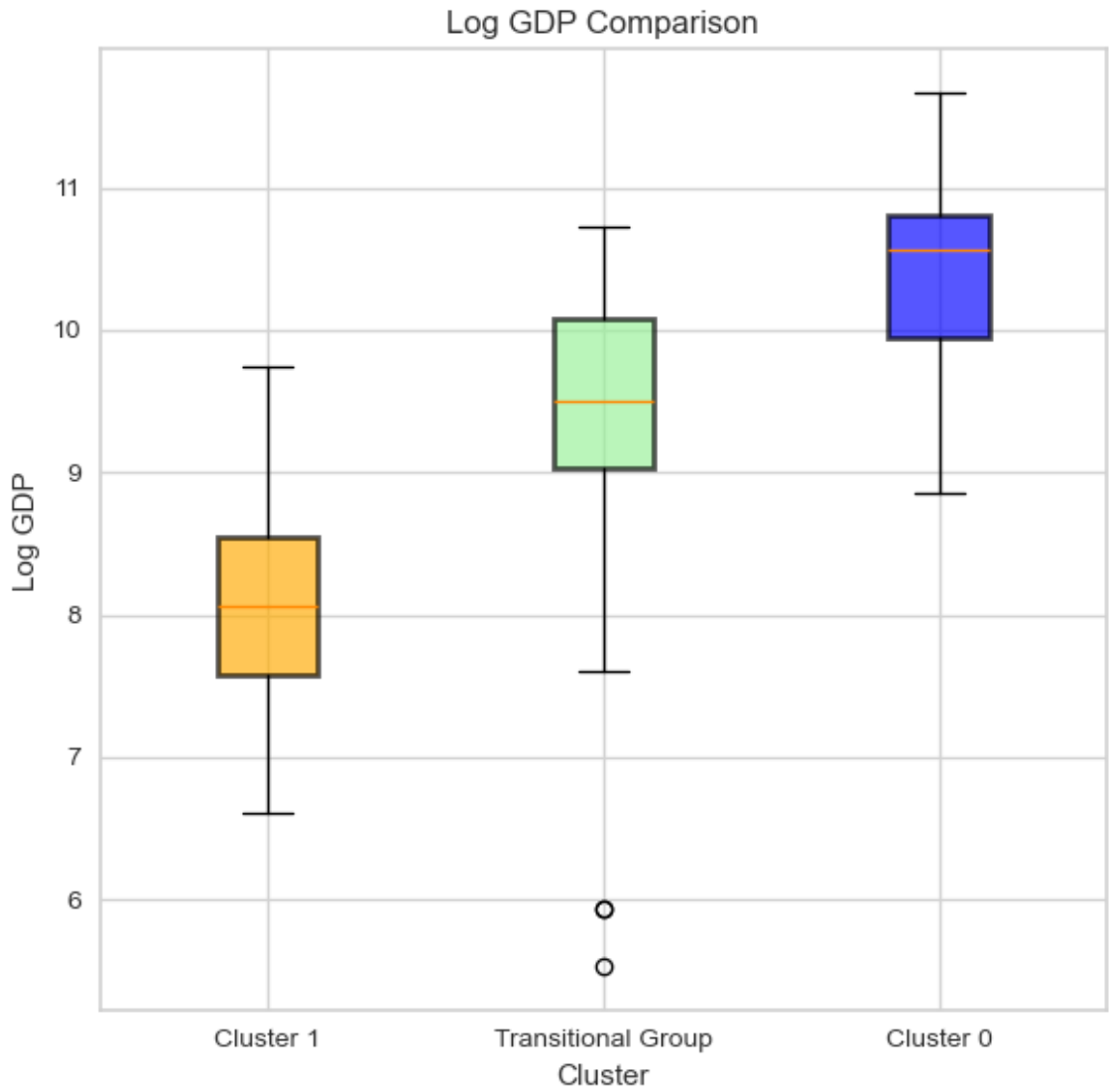
Feature	Cluster 0 Mean	Cluster 1 Mean
Life Ladder	6.4	4.5
Log GDP Per Capita	10.2	8.5
Social Support	0.89	0.72
Healthy Life Expectancy	68.4	57.7
Freedom to Make Life Choices	0.82	0.67
Generosity	0.03	-0.02
Perceptions of Corruption	0.69	0.81
Positive Affect	0.7	0.6
Negative Affect	0.24	0.3
Confidence in National Government	0.46	0.5

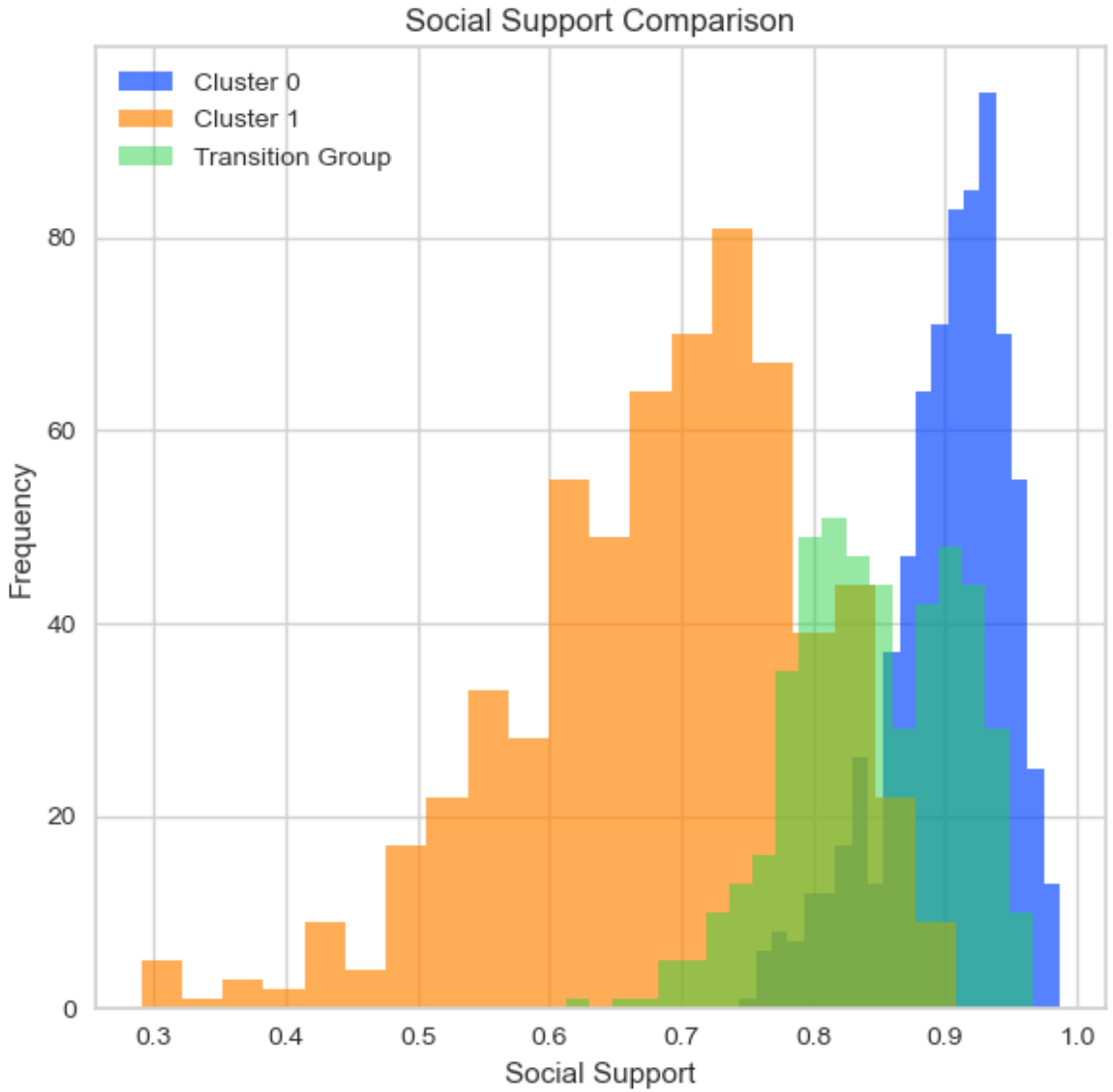

Cluster 0
Cluster 1
Transitional Group
Southeast Asia
East Asia
Middle East and North Africa
Central and Eastern Europe
Western Europe
North America and ANZ
Commonwealth of Independent States
Latin America and Caribbean
Sub-Saharan Africa
South Asia
Regions
0
5
10
15
20
25
30
35
Countries per Cluster

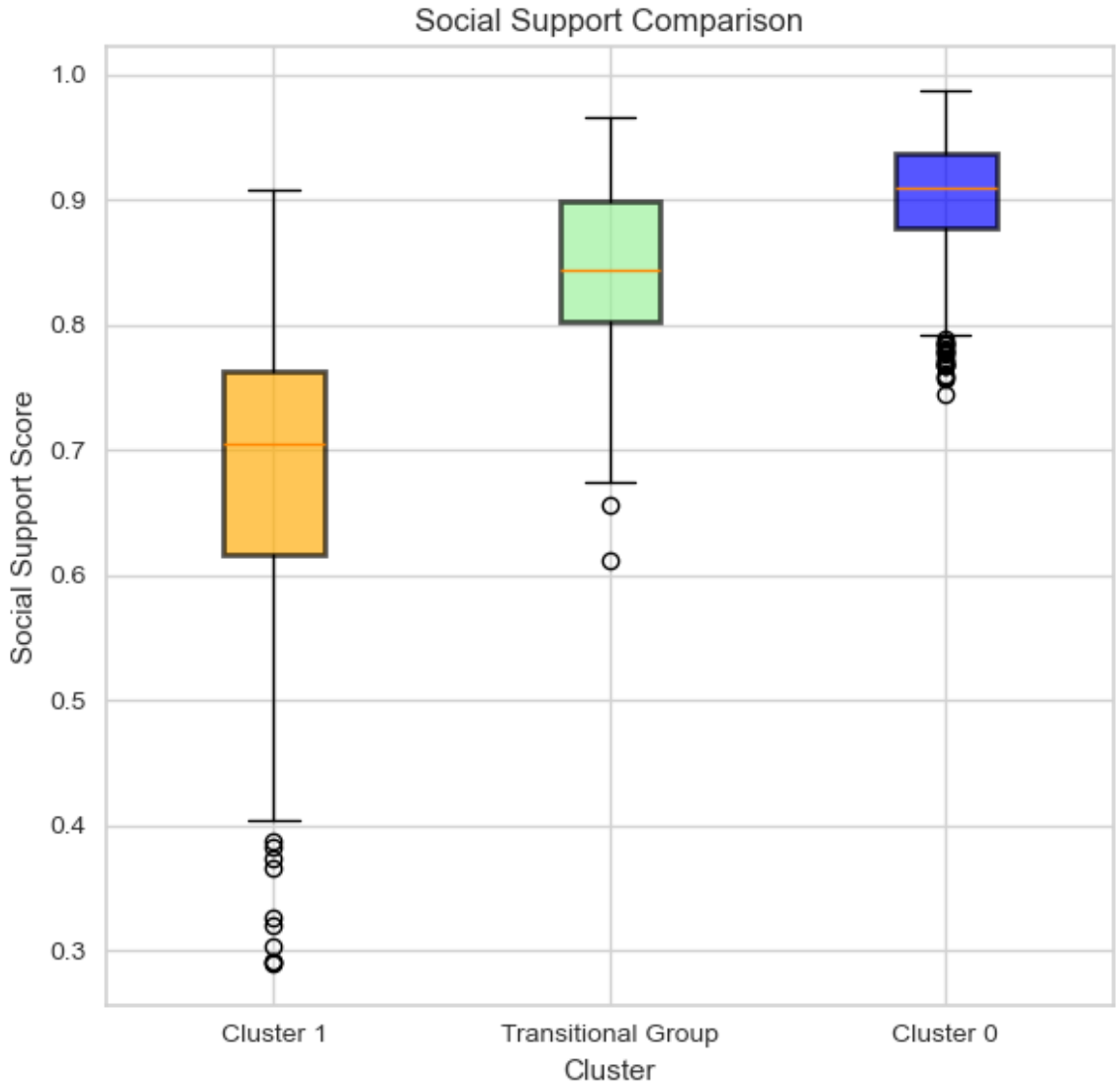

Life Ladder Comparison
Cluster 0
Cluster 1
Transitional Group
Frequency
Life Ladder Score
0
10
20
30
40
50
60
70
80
3
4
5
6
7
8


Life Ladder Comparison
Life Ladder Score
Cluster
Cluster 1
Transitional Group
Cluster 0
3
4
5
6
7
8


Log GDP Comparison
Cluster 0
Cluster 1
Transition Group
Frequency
0
20
40
60
80
100
6
7
8
9
10
11
Log GDP

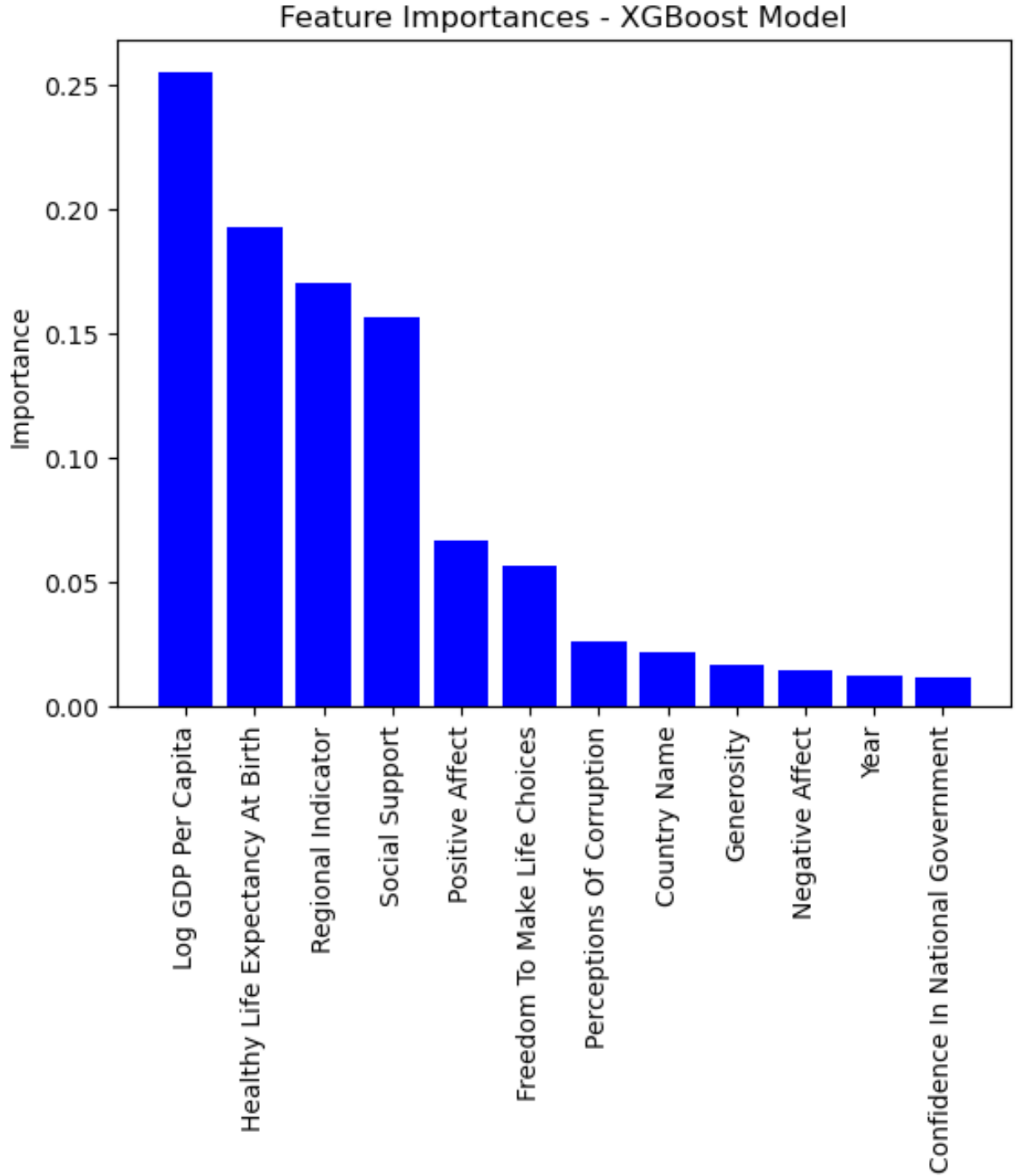

Log GDP Comparison
Log GDP
6
7
8
9
10
11
Cluster 1
Transitional Group
Cluster 0
Cluster

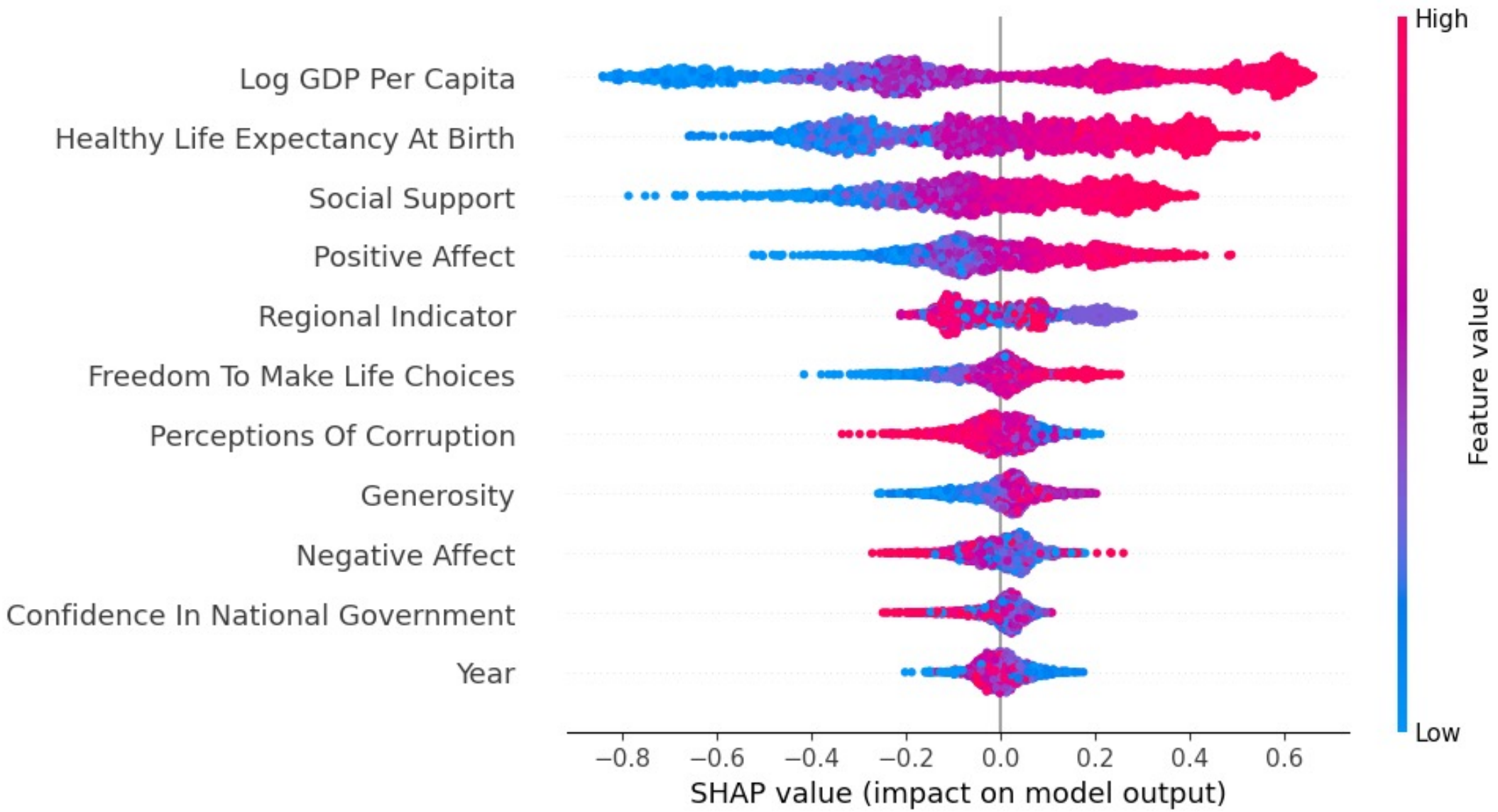

Social Support Comparison
Cluster 0
Cluster 1
Transition Group
Frequency
0
20
40
60
80
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Social Support

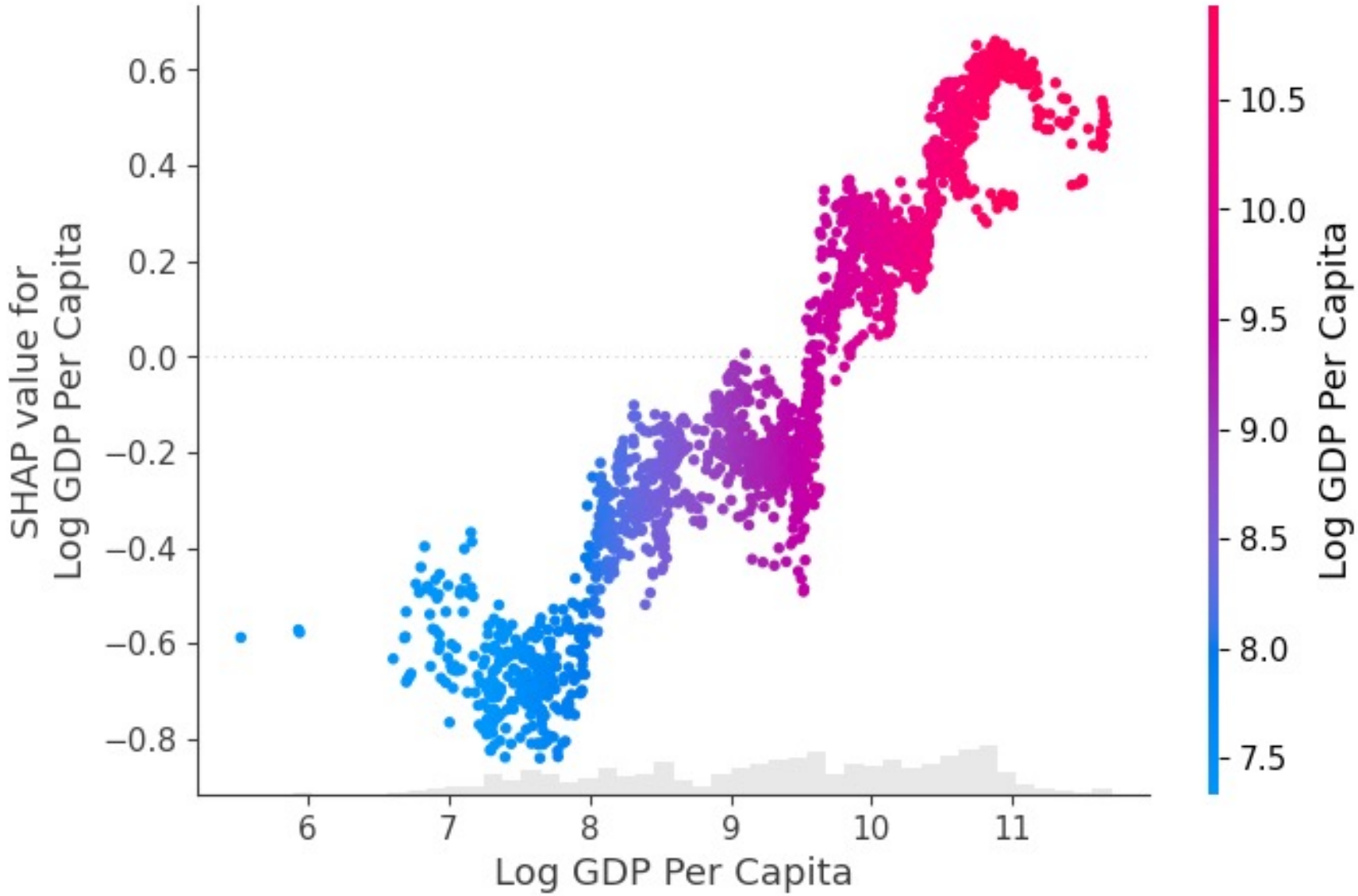

Social Support Comparison
Social Support Score
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Cluster 1
Transitional Group
Cluster 0
Cluster

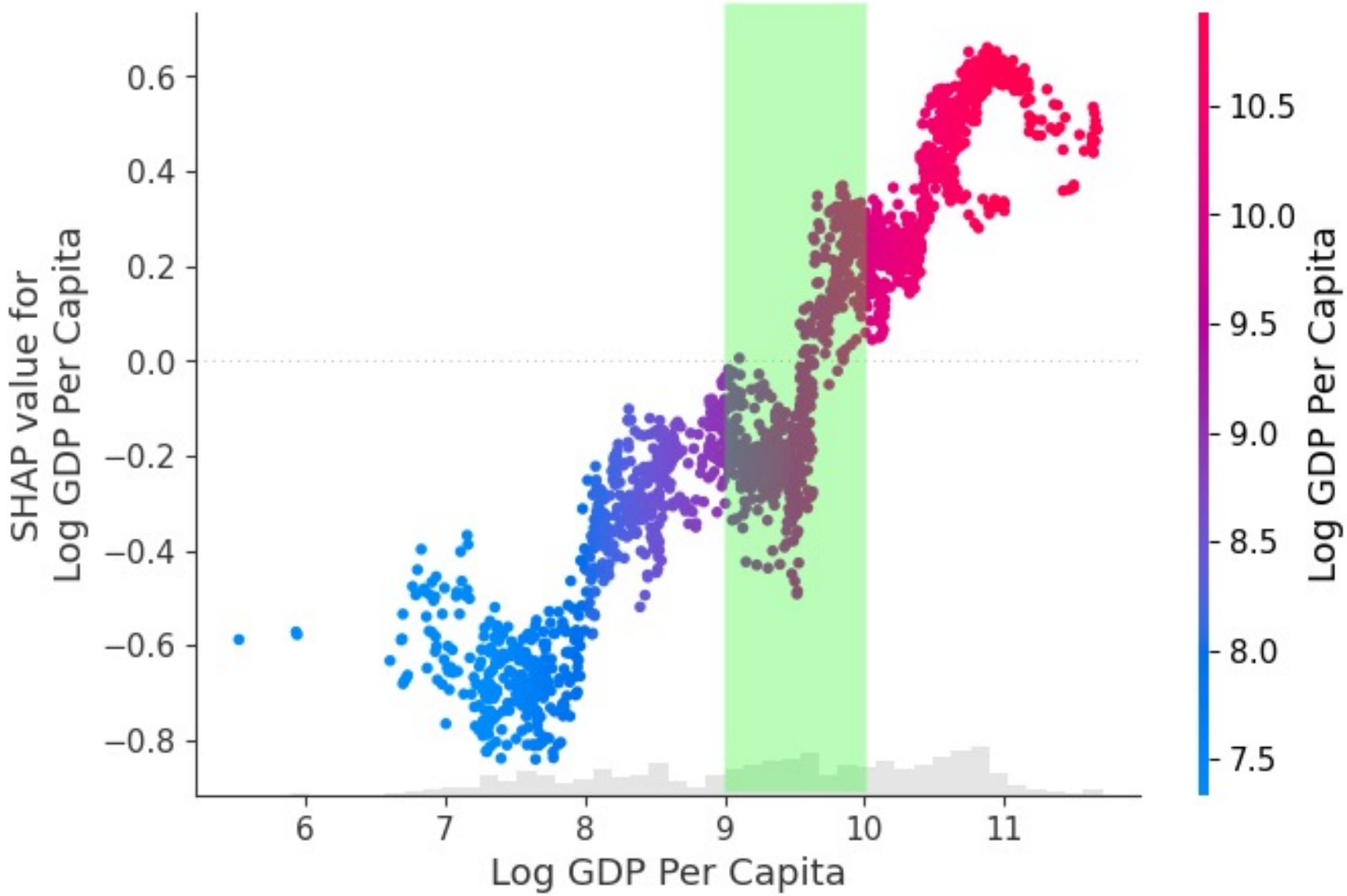
Predictive Modeling

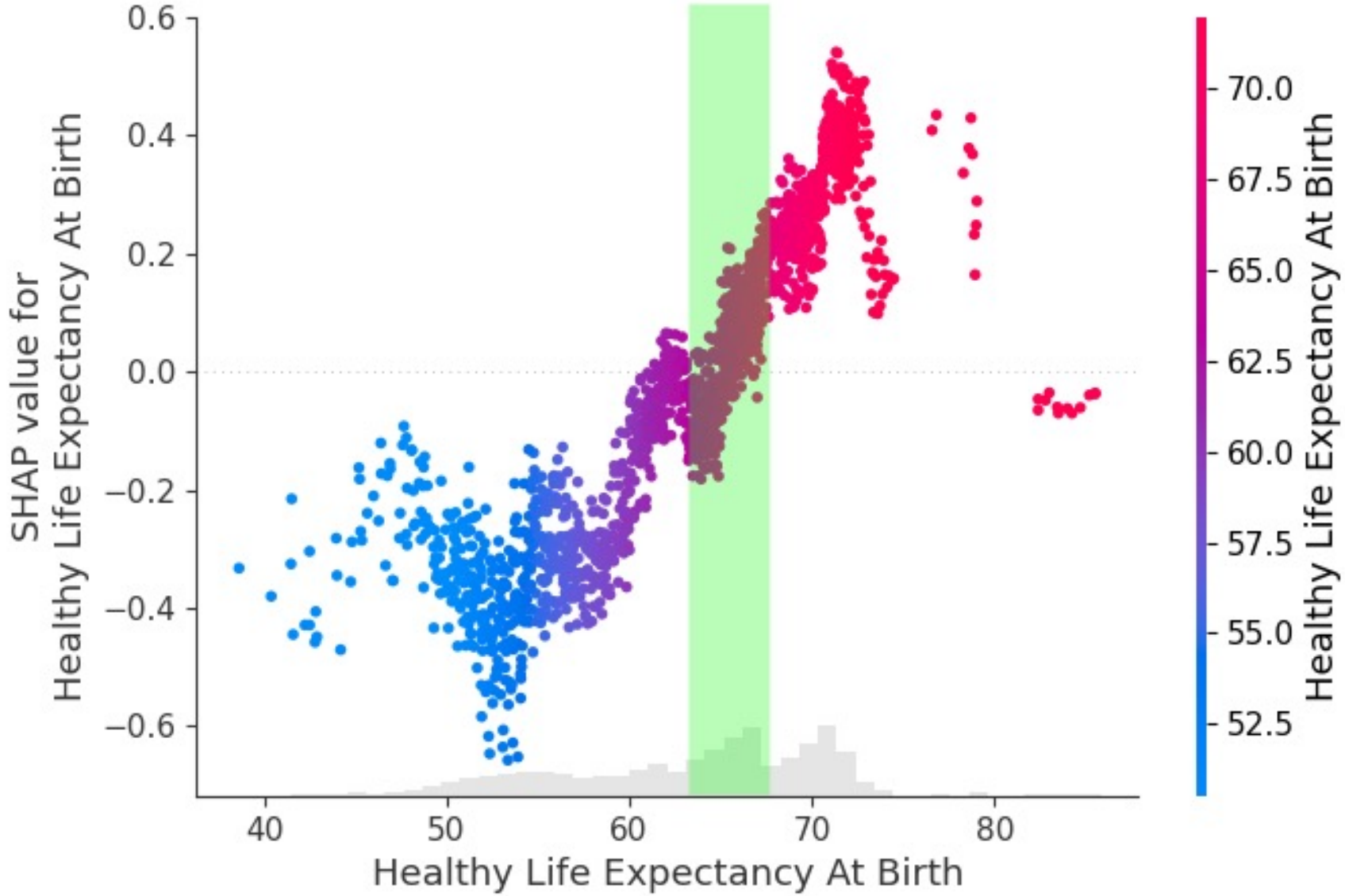
Model	r2	MAE	MAPE	Improvement over Baseline
RF Regression	0.8978	0.2828	0.056	0.78%
XGBoost	0.9103	0.2675	0.0526	7.91%
LGBM	0.9107	0.2701	0.0533	3.67%

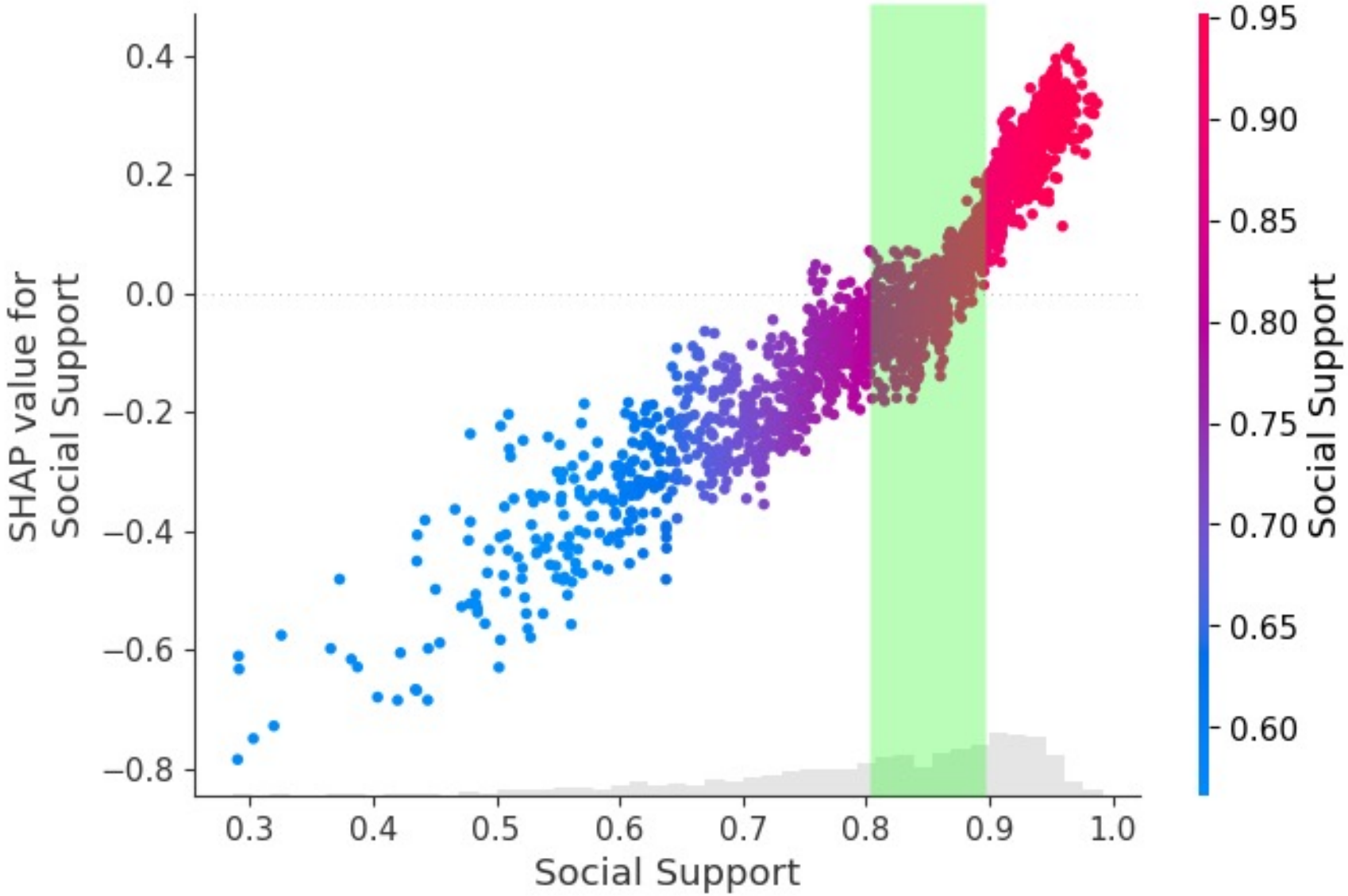

Feature Importances - XGBoost Model
Importance
0.25
0.20
0.15
0.10
0.05
0.00
Log GDP Per Capita
Healthy Life Expectancy At Birth
Regional Indicator
Social Support
Positive Affect
Freedom To Make Life Choices
Perceptions Of Corruption
Country Name
Generosity
Negative Affect
Year
Confidence In National Government

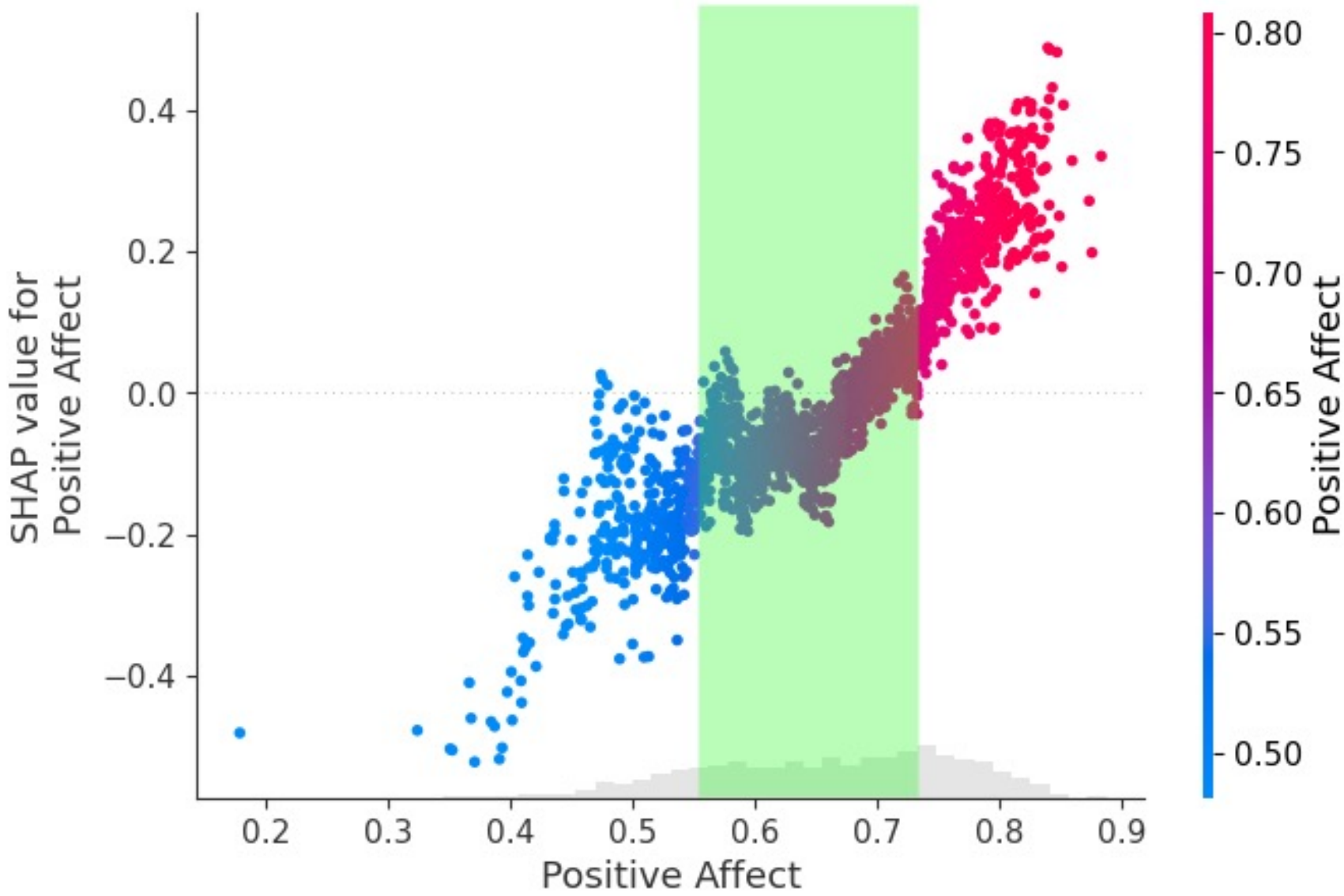

Log GDP Per Capita
Healthy Life Expectancy At Birth
Social Support
Positive Affect
Regional Indicator
Freedom To Make Life Choices
Perceptions Of Corruption
Generosity
Negative Affect
Confidence In National Government
Year
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
SHAP value (impact on model output)
High
Feature value
Low

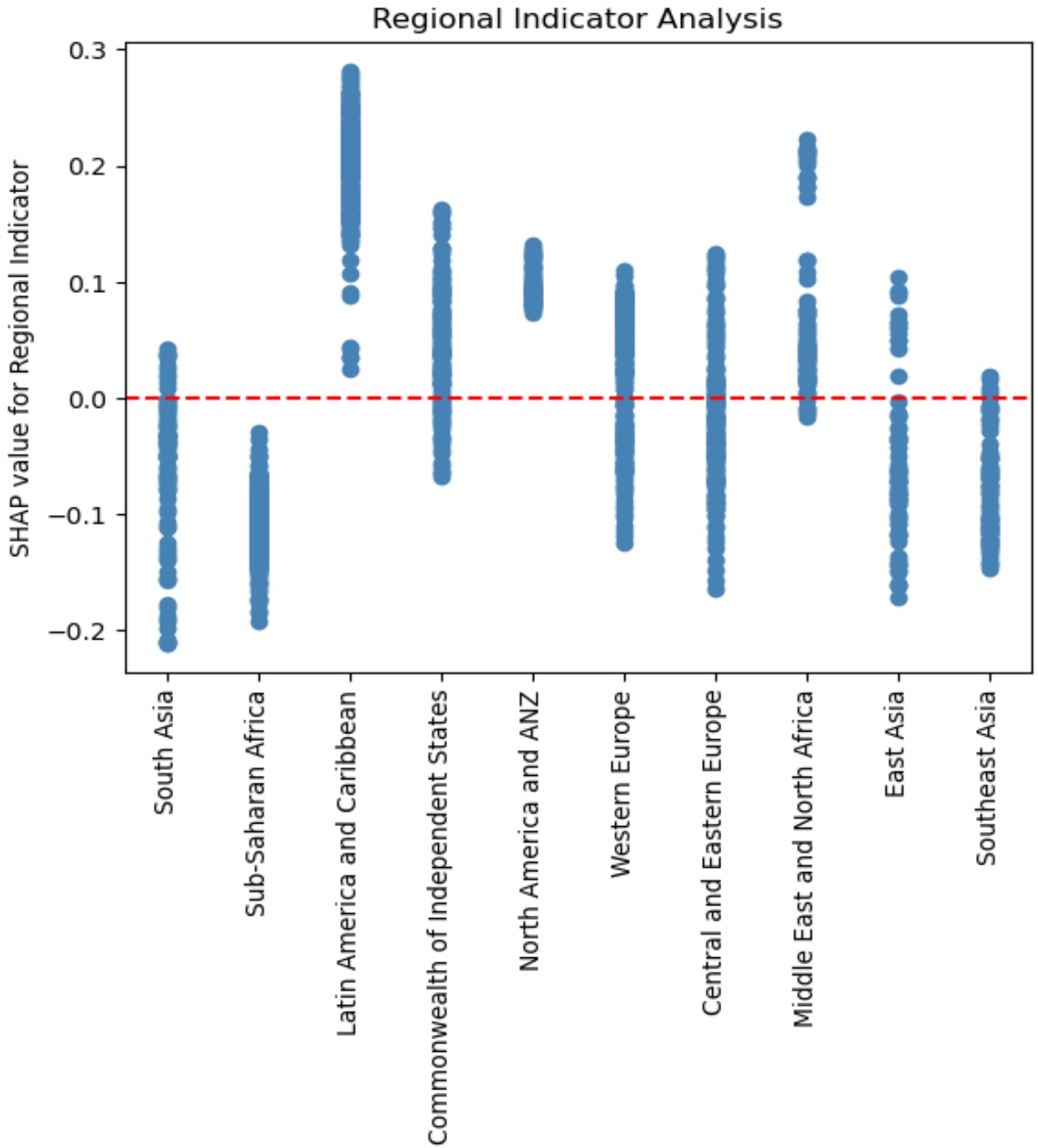

SHAP value for
Log GDP Per Capita
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
6
7
8
9
10
11
Log GDP Per Capita
10.5
10.0
9.5
9.0
8.5
8.0
7.5
Log GDP Per Capita


SHAP value for
Log GDP Per Capita
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
6
7
8
9
10
11
Log GDP Per Capita
10.5
10.0
9.5
9.0
8.5
8.0
7.5
Log GDP Per Capita


SHAP value for
Healthy Life Expectancy At Birth
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
40
50
60
70
80
Healthy Life Expectancy At Birth
70.0
67.5
65.0
62.5
60.0
57.5
55.0
52.5
Healthy Life Expectancy At Birth


SHAP value for
Social Support
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
Social Support
0.95
0.90
0.85
0.80
0.75
0.70
0.65
0.60
Social Support


SHAP value for
Positive Affect
0.4
0.2
0.0
−0.2
−0.4
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
Positive Affect
0.80
0.75
0.70
0.65
0.60
0.55
0.50
Positive Affect


Regional Indicator Analysis
SHAP value for Regional Indicator
0.3
0.2
0.1
0.0
−0.1
−0.2
South Asia
Sub-Saharan Africa
Latin America and Caribbean
Commonwealth of Independent States
North America and ANZ
Western Europe
Central and Eastern Europe
Middle East and North Africa
East Asia
Southeast Asia

Major Feature Target Zones

Feature	General Measurement	Approximate Threshold
LOG GDP	Economic Strength	9.8
Healthy Life Expectancy	Life Span	67 years
Social Support	Strength of Personal Social Networks	0.9
Positive Affect	Previous days enjoyment	0.7

Recommendation 1:

Deeper analysis of the most impactful features using new and actionable metrics.

Recommendation 2:

Regional analysis

- Are the regional differences real?

Country level analysis

- Policy is made at the country level

Recommendation 3:

Develop proxy metrics for missing survey responses:

- Perceptions of corruption
- Confidence in government