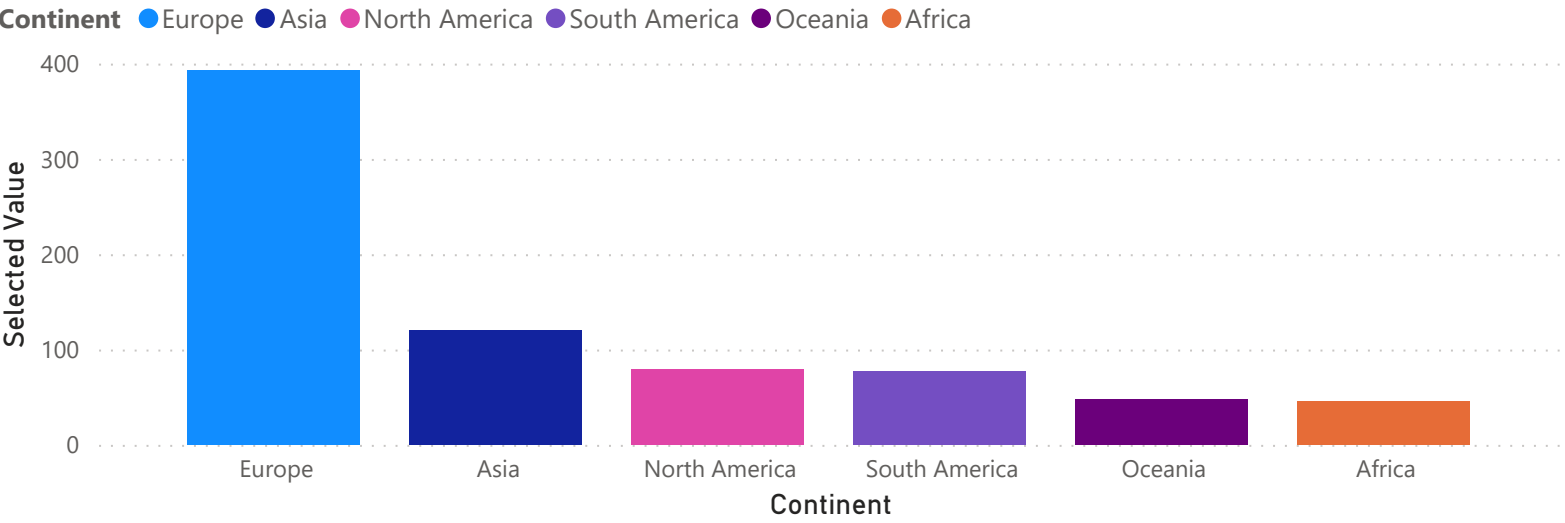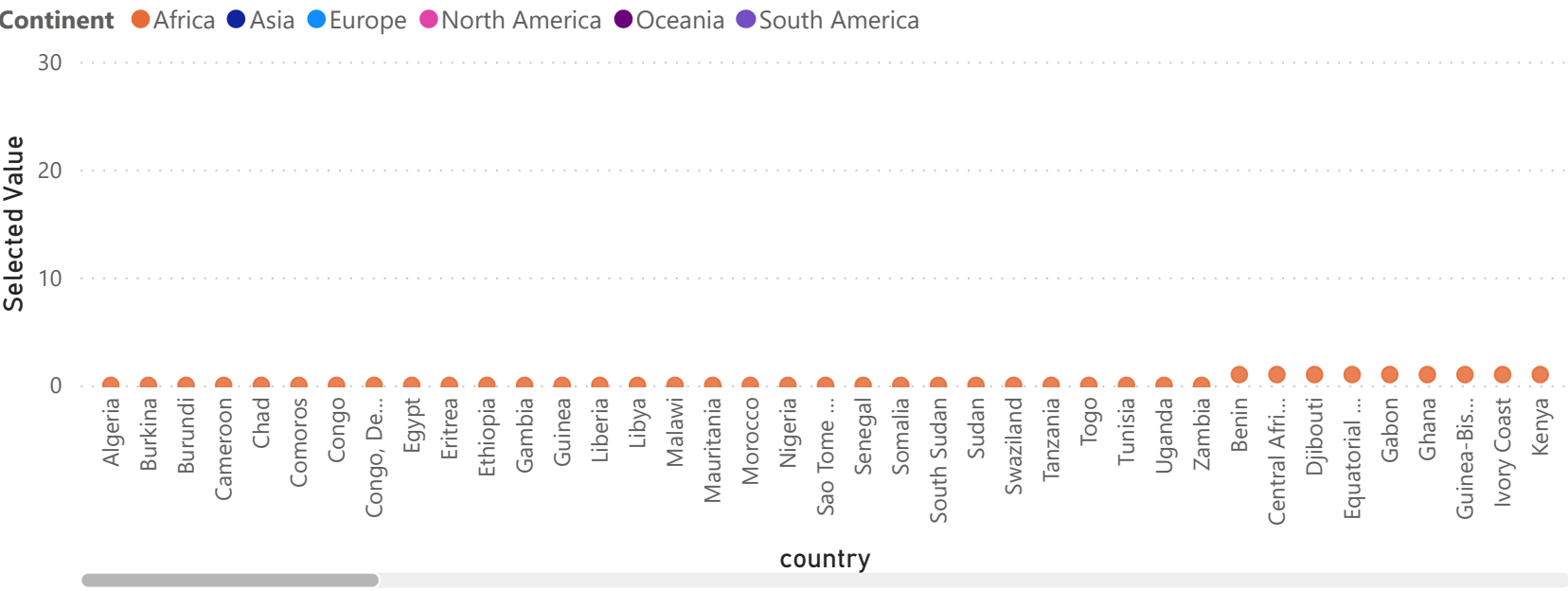# Country social factor clustering

Ivan Shamoon

Goal: Can a machine learning clustering algorithm accurately characterize countries into their respective regions based on a variety of factors.

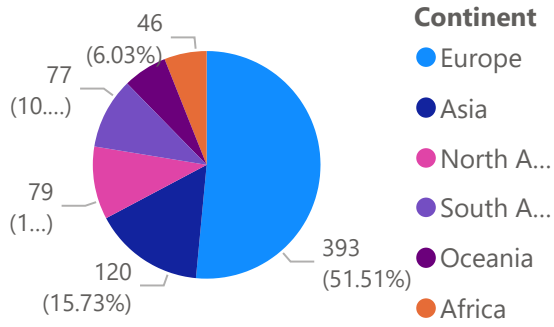## Initial observations

Selected Value by Continent and Continent

Continent ● Europe ● Asia ● North America ● South America ● Oceania ● Africa
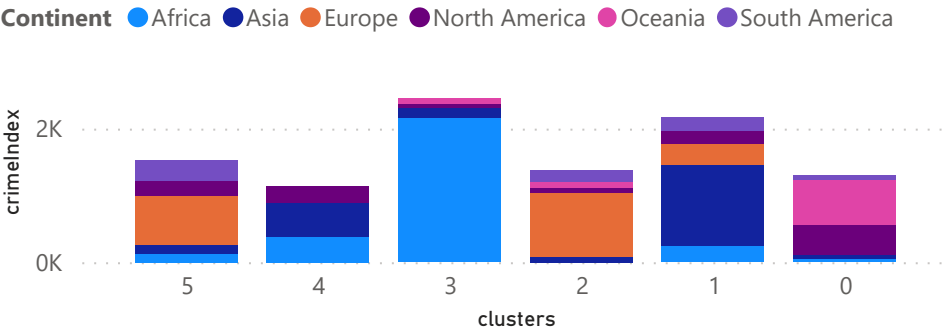
Name: Homosexuality Tolerance

☐ Average Schooling
☐ Crime Index
☐ Female Life Expectancy
☐ Food Scarcity
☒ Homosexuality Tolerance
☐ Male Life Expectancy
☐ Obesity
☐ Political Stability
☐ Water Score

Selected Value by Continent and country

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

country: Algeria, Burkina, Burundi, Cameroon, Chad, Comoros, Congo, Congo, De..., Egypt, Eritrea, Ethiopia, Gambia, Guinea, Liberia, Libya, Malawi, Mauritania, Morocco, Nigeria, Sao Tome..., Senegal, Somalia, South Sudan, Sudan, Swaziland, Tanzania, Togo, Tunisia, Uganda, Zambia, Benin, Central Afri..., Djibouti, Equatorial..., Gabon, Ghana, Guinea-Bis..., Ivory Coast, Kenya

Continent
● Europe — 393 (51.51%)
● Asia — 120 (15.73%)
● North A... — 79 (1...)
● South A... — 77 (10....)
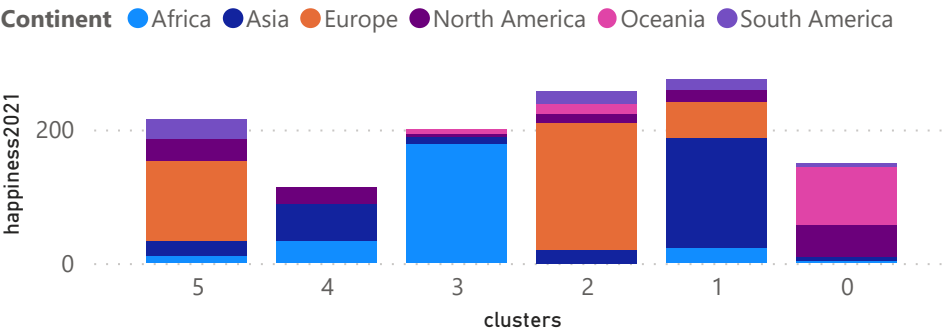● Oceania — 46 (6.03%)
● Africa

Clustering Analysis

Clustering Analysis allows us to gather interesting insights about their characterization. Each country falls under one cluster and by surveying its category we can view its behavior in the scope of its social factors.
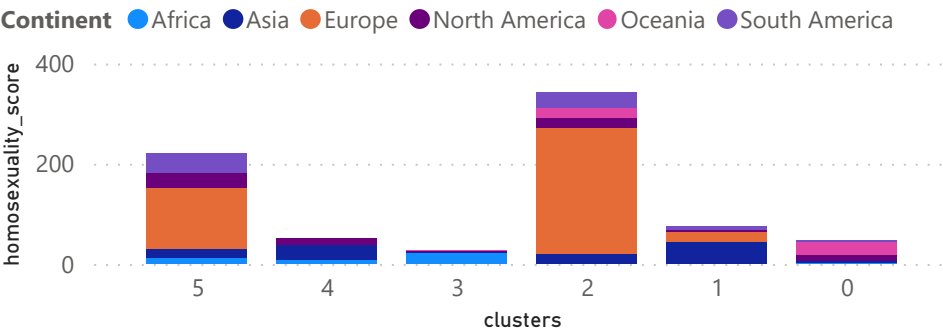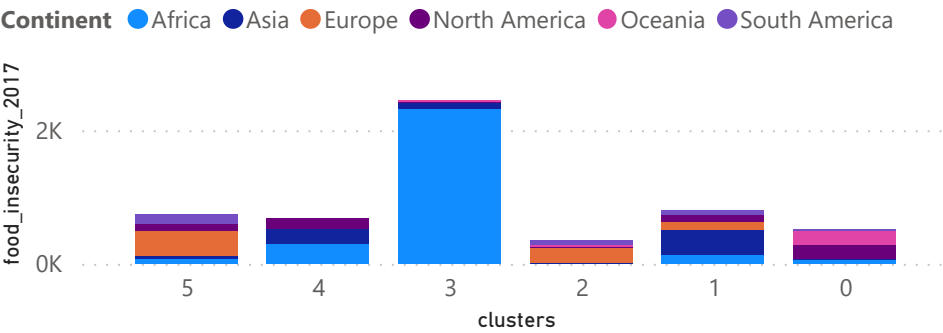
crimeIndex by clusters and Continent

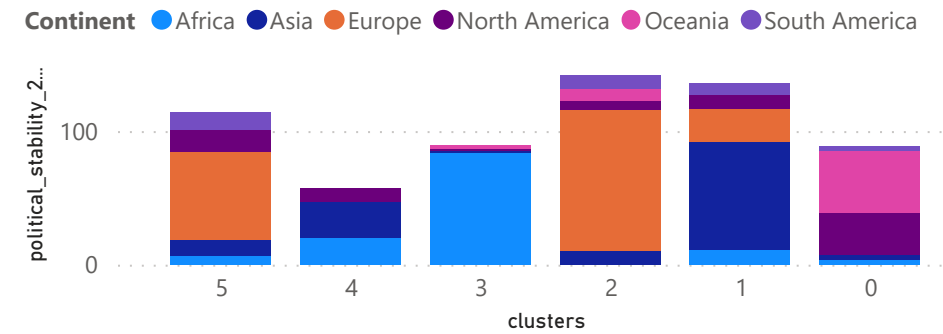**Continent** ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	20.80	219.80		557.23	192.23	90.10	**1,080.15**
Asia	43.10	1,652.54	281.00	96.83	364.93	258.60	**2,697.01**
Europe		550.39	2,591.19			1,320.04	**4,461.61**
North America	368.49	152.70	174.10	14.00	144.40	263.90	**1,117.59**
Oceania	403.18		167.40	15.50			**586.08**
South America	39.30	129.60	203.60			235.10	**607.60**
Total	**874.87**	**2,705.02**	**3,417.29**	**683.56**	**701.56**	**2,167.74**	**10,550.04**

happiness2021 by clusters and Continent

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	3.47	22.97		178.06	33.52	11.01	**249.02**
Asia	5.68	164.29	19.47	11.12	56.20	22.58	**279.34**
Europe		53.73	190.50			120.20	**364.43**
North America	48.41	18.56	14.05	3.62	24.39	31.81	**140.83**
Oceania	86.76		14.46	7.23			**108.45**
South America	5.87	16.42	18.53			29.66	**70.49**
Total	**150.18**	**275.96**	**257.02**	**200.02**	**114.11**	**215.26**	**1,212.56**

homosexuality_score by clusters and Continent

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	3	0	22	8	13		**46**
Asia	4	45	20	3	30	18	**120**
Europe		20	252		121		**393**
North America	11	3	20	1	13	31	**79**
Oceania	27		20	1			**48**
South America	3	6	30		38		**77**
Total	**48**	**74**	**342**	**27**	**51**	**221**	**763**

food_insecurity_2017 by clusters and Continent

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	53.40	132.00		2,326.54	304.07	70.99	**2,887.00**
Asia	27.10	371.59	19.60	81.92	219.36	51.22	**770.80**
Europe		126.87	217.82			365.48	**710.18**
North America	208.17	100.88	13.90	26.99	153.19	117.68	**620.81**
Oceania	205.47		26.00	16.53			**248.00**
South America	25.37	66.34	71.10			141.64	**304.46**
Total	**519.50**	**797.69**	**348.42**	**2,451.98**	**676.63**	**747.02**	**5,541.25**

political_stability_2020 by clusters and Continent

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	4.09	10.76		83.58	20.10	6.65	**125.18**
Asia	3.79	81.36	10.68	1.75	26.68	11.96	**136.22**
Europe		24.38	105.32			65.66	**195.36**
North America	30.74	10.38	7.09	1.96	10.36	16.87	**77.40**
Oceania	46.80		8.34	2.26			**57.40**
South America	3.42	8.84	10.16			13.23	**35.65**
Total	**88.84**	**135.72**	**141.59**	**89.55**	**57.14**	**114.37**	**627.20**

obesity_2016 by clusters and Continent

Continent ● Africa ● Asia ● Europe ● North America ● Oceania ● South America

Continent	0	1	2	3	4	5	Total
Africa	18.90	144.40		366.67	115.50	39.10	**684.57**
Asia	20.60	650.50	36.50	31.20	103.44	33.64	**875.88**
Europe		207.16	630.80			475.29	**1,313.25**
North America	198.80	62.20	65.60	22.70	90.90	125.00	**565.20**
Oceania	520.40		59.80	21.30			**601.50**
South America	26.40	64.01	84.20			107.51	**282.12**
Total	**785.10**	**1,128.26**	**876.90**	**441.87**	**309.84**	**780.55**	**4,322.53**

Insights about the Clusters

Cluster 0:
* Lacks obesity
* Poltically stable
* Good protein supply
* High avg schooling
* Strong LGBT tolerant
* High life span in both sexes compared
* High emissions

Cluster 1:
* Somewhat obease
* Politically unstable
* Low protein supply
* Low emissions
* Lacks any representation in Europe
* Moderately malnourished

Cluster 2:
* Not Oceania
* Very high emissions
* Not LGBT tolerant
* High protein supply
* A little obease
* Malnourished

Cluster 3:
* Very malnourished
* Not Tolerant of LGBT
* Low emissions
* No representation in Europe
* No representation in South America
* No representation in Africa
* Middle ground political stability

Cluster 4:
* Not representation in Oceania
* Low emissions
* High obesity
* Politically stable
* High protein supply
* Malnourished
* Tolerant of LGBT
* High water score

Cluster 5:
* Not Europe or South America
* Somewhat malnourished
* Not LGBT tolerant
* Low water score
* Low emissions
* Low avergae schooling
* Low protein supply