

# **Unsupervised Machine Learning: K-means Clustering and PCA on Country Data**

Factors of Child mortality

(refers to the probability of dying between birth and exactly  
five years of age expressed per 1,000 live births)

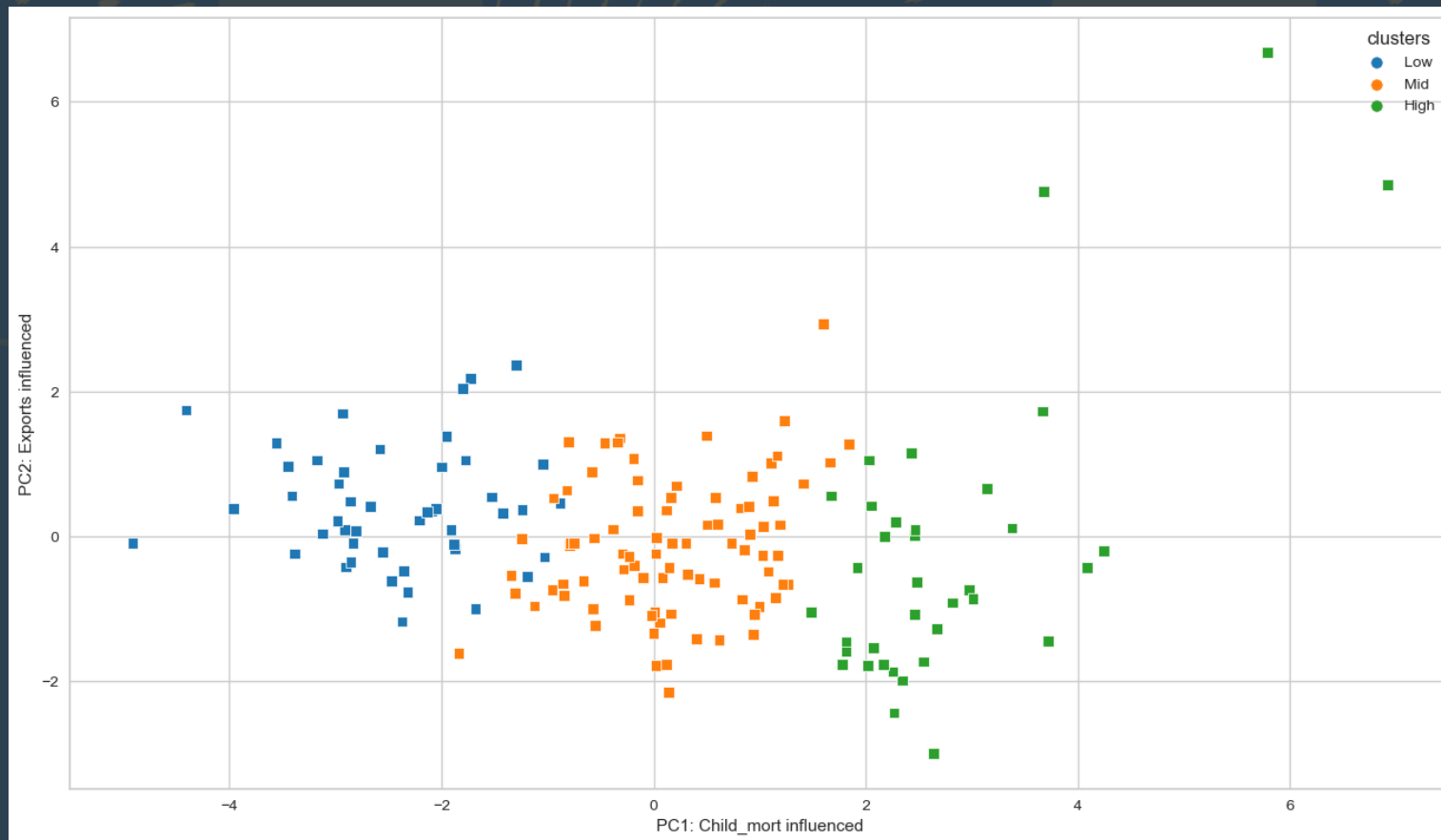
# Meet the 'Country-data.csv'

	country	child_mort	exports	health	imports	income	inflation	life_expec	total_fer	gdpp
0	Afghanistan	90.2	10.0	7.58	44.9	1610	9.44	56.2	5.82	553
1	Albania	16.6	28.0	6.55	48.6	9930	4.49	76.3	1.65	4090
2	Algeria	27.3	38.4	4.17	31.4	12900	16.10	76.5	2.89	4460
3	Angola	119.0	62.3	2.85	42.9	5900	22.40	60.1	6.16	3530
4	Antigua and Barbuda	10.3	45.5	6.03	58.9	19100	1.44	76.8	2.13	12200
...	...	...	...	...	...	...	...	...	...	...
162	Vanuatu	29.2	46.6	5.25	52.7	2950	2.62	63.0	3.50	2970
163	Venezuela	17.1	28.5	4.91	17.6	16500	45.90	75.4	2.47	13500
164	Vietnam	23.3	72.0	6.84	80.2	4490	12.10	73.1	1.95	1310
165	Yemen	56.3	30.0	5.18	34.4	4480	23.60	67.5	4.67	1310
166	Zambia	83.1	37.0	5.89	30.9	3280	14.00	52.0	5.40	1460

167 rows × 10 columns

# Principal component analysis - visual

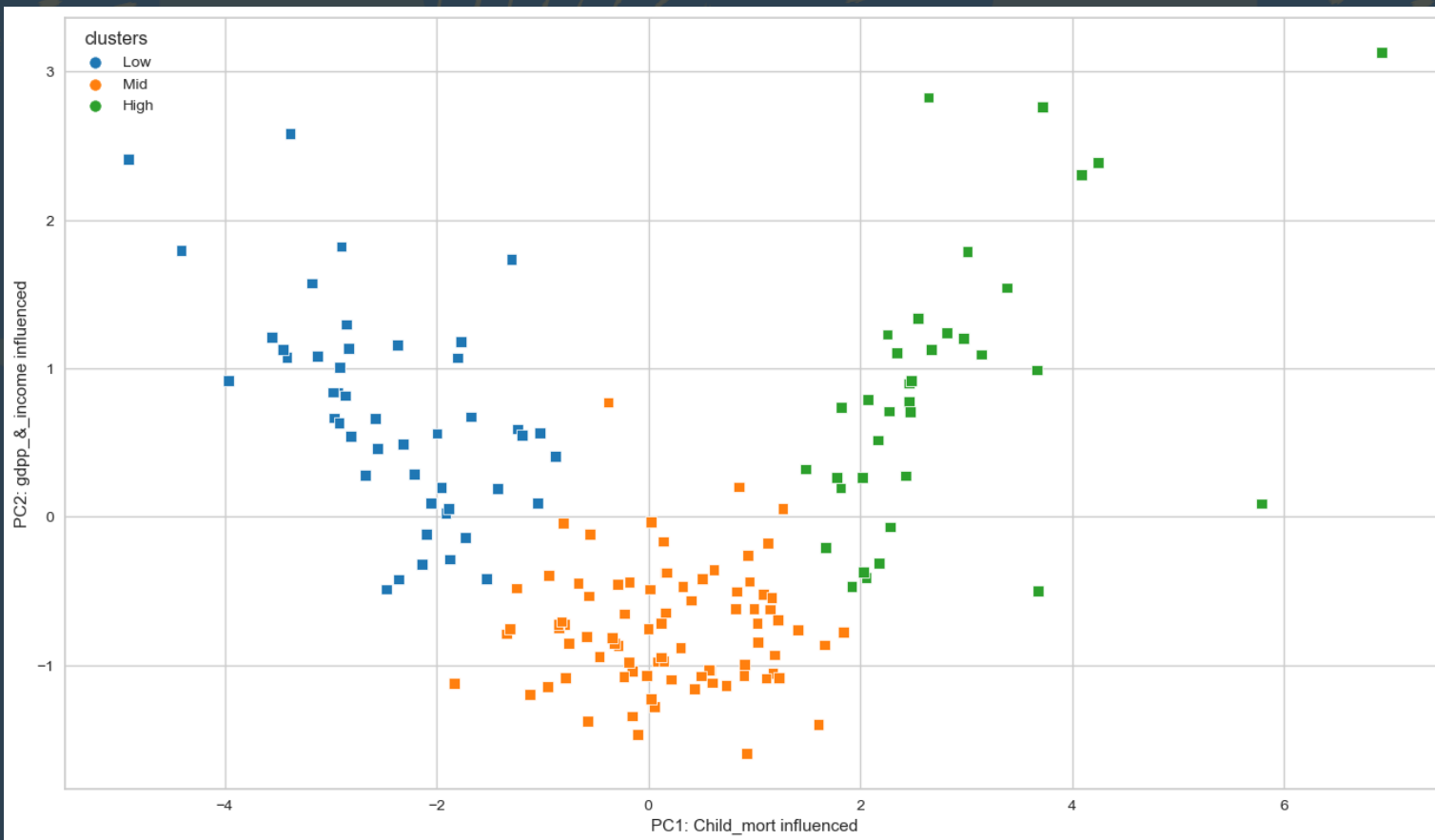
## PC1:PC2



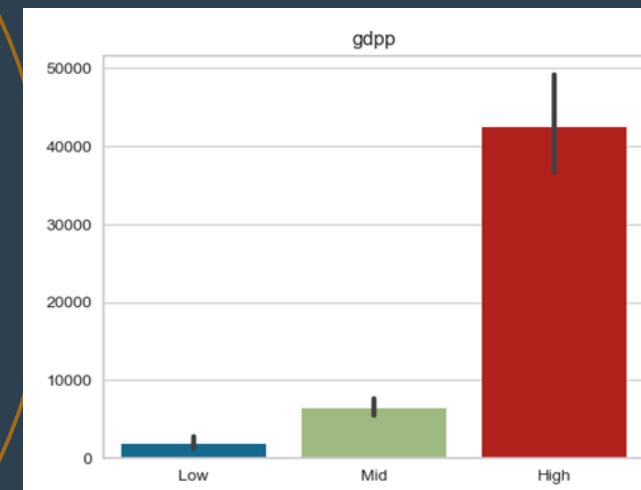
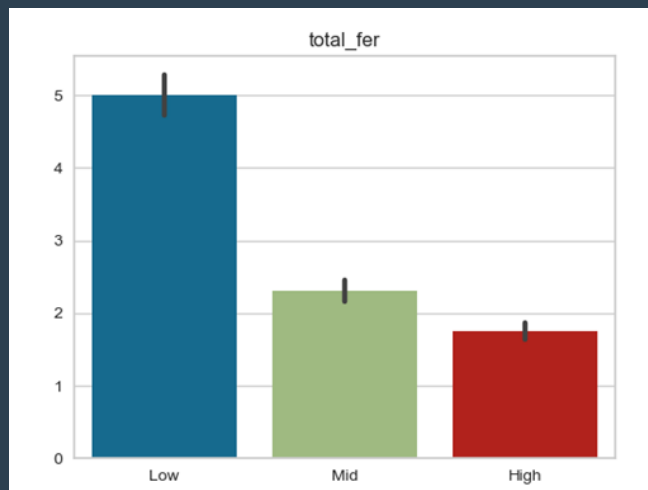
Cluster Interpreter  
Low: Less prosperous countries  
Mid: In-between  
High: Prosperous countries

# Principal component analysis - visual

## PC1:PC4

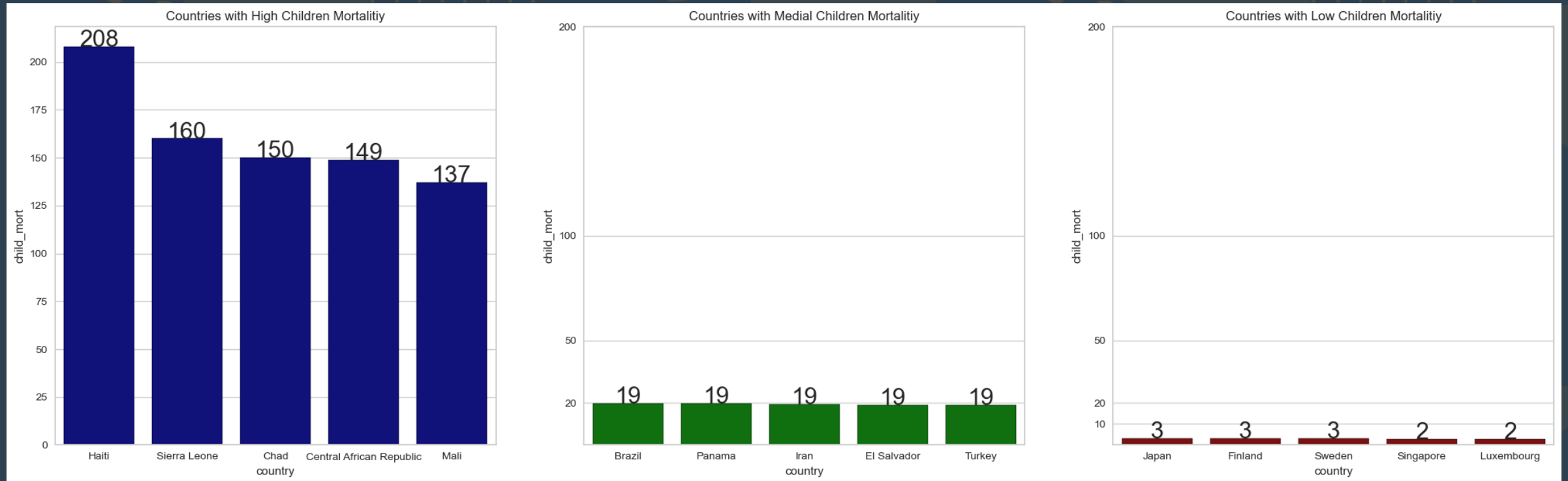


Cluster Interpreter  
Low: Less prosperous countries  
Mid: In-between  
High: Prosperous countries

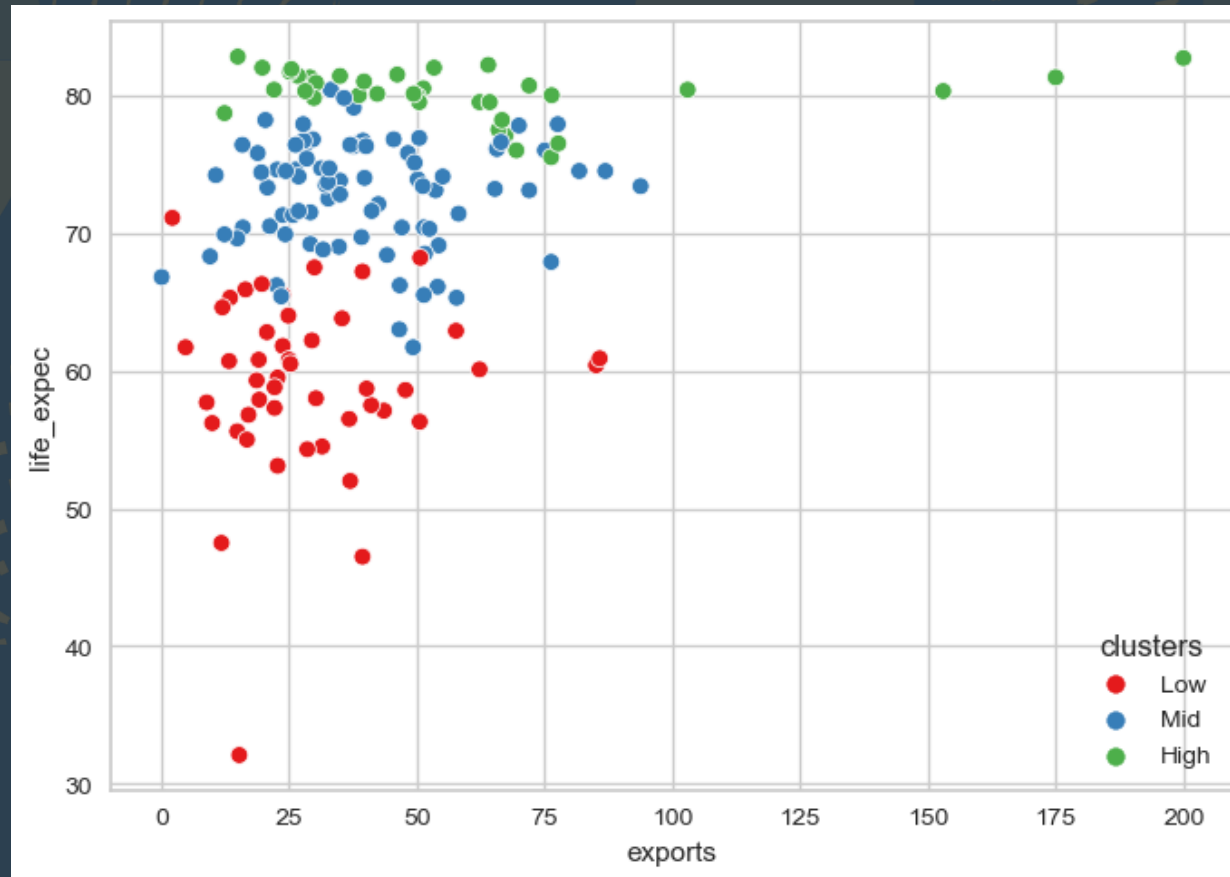


Low: Less prosperous countries  
Mid: In-between  
High: Prosperous countries

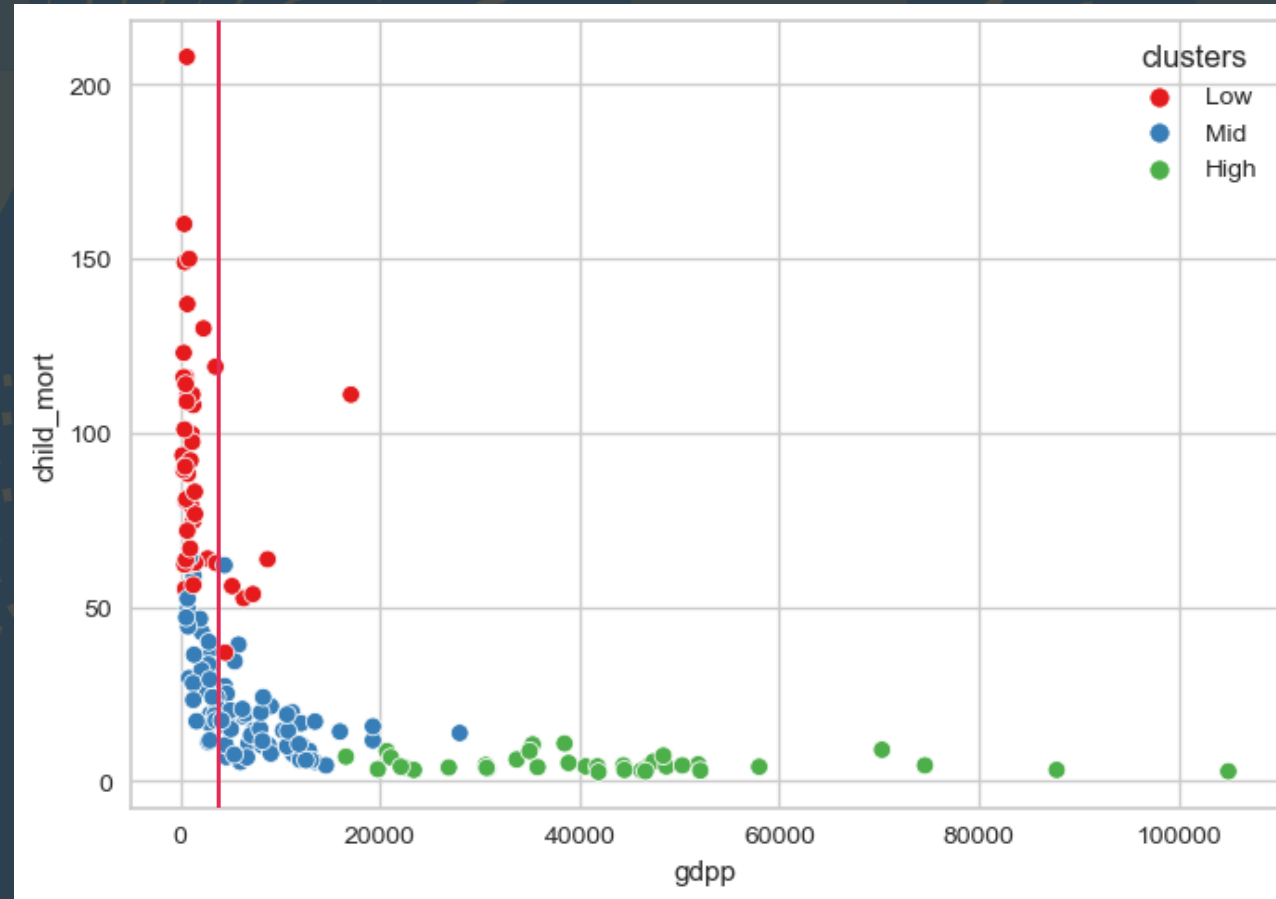
# The extremes of each cluster...



# Life expectancy vs. exports

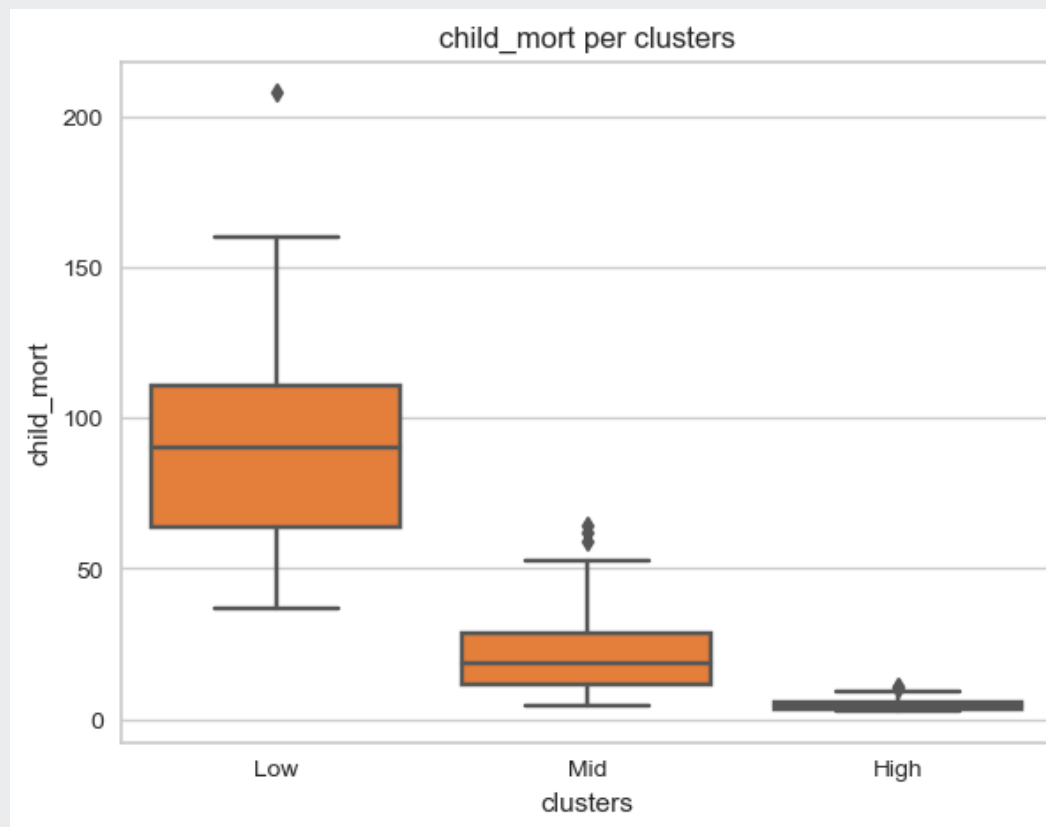


# Child mortality vs. Gdp(perc)

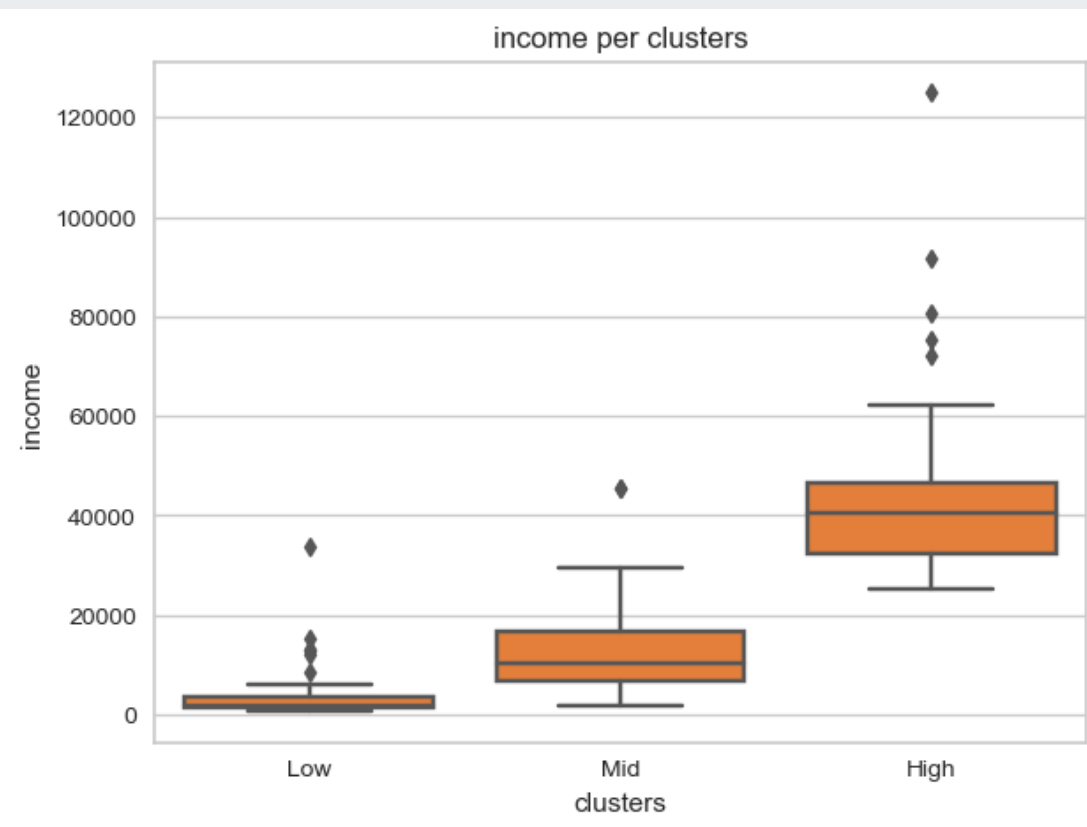




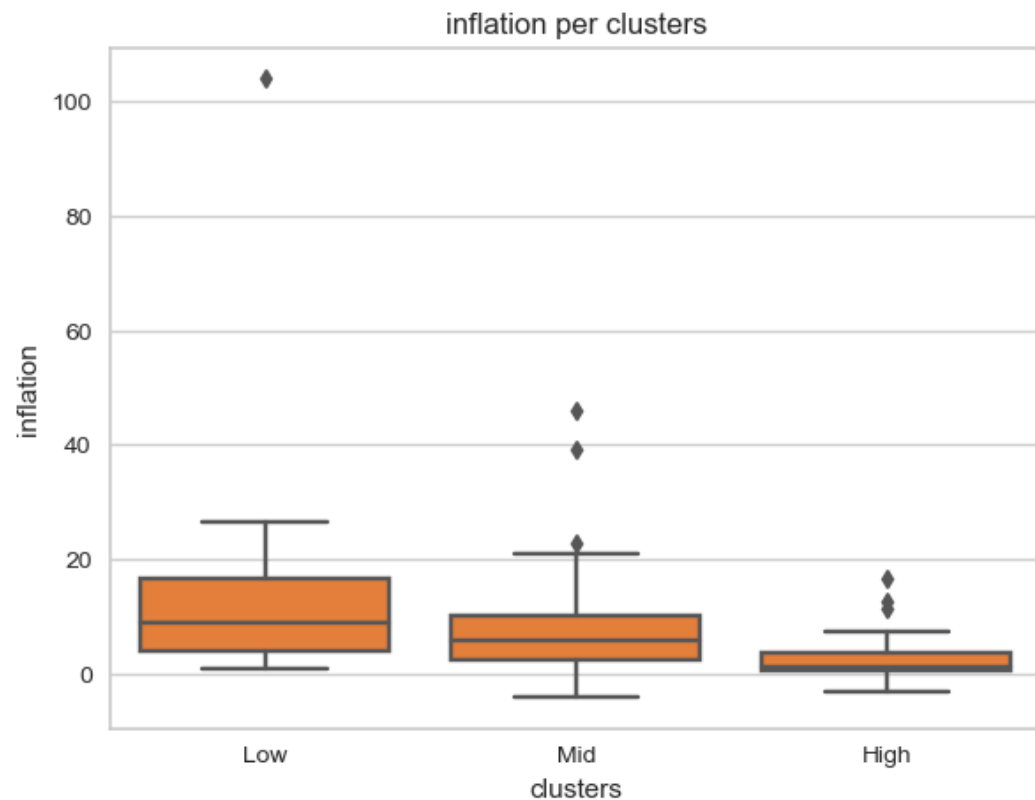
# Child mortality



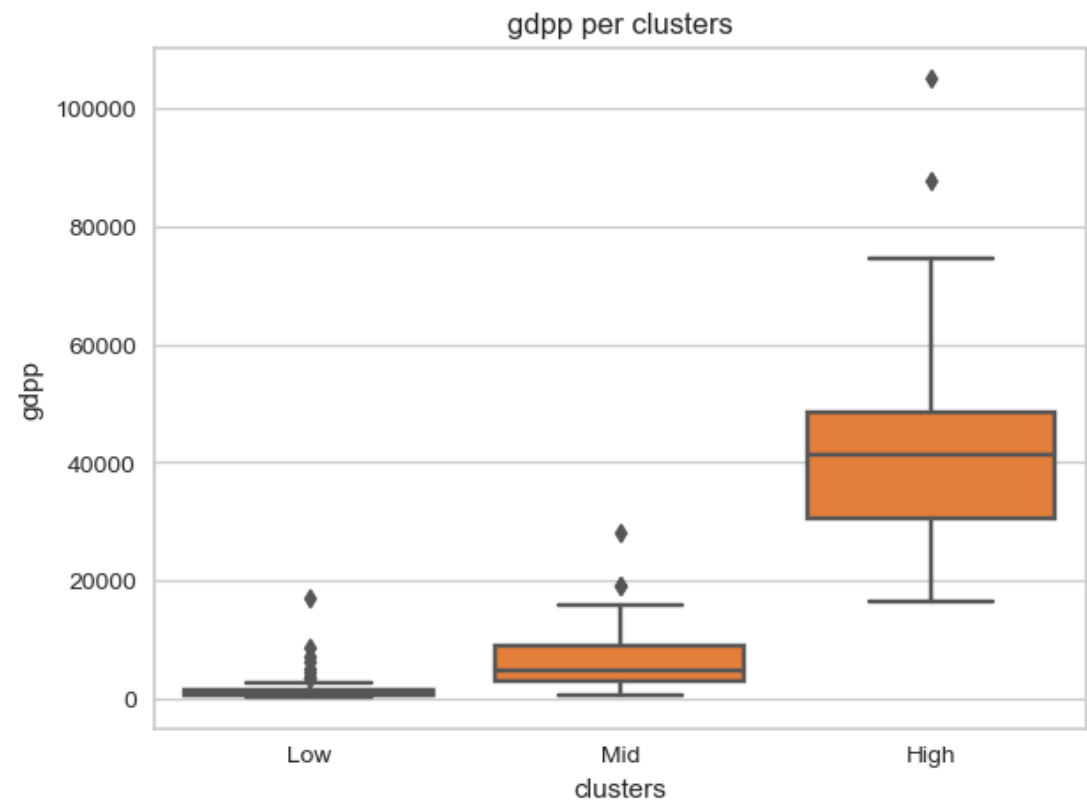
# Income



# Inflation



# Gdp(perc)



**Thanks for your attention!**  
**Questions?**

