PCA EXERCISE

SIA - 2022

CONSIGNA

El conjunto de datos correspondiente a características económicas, sociales y geográficas de 28 países de Europa. Las variables son:

- Country: Nombre del país.
- Área: área.
- GDP: producto bruto interno.
- Inflation: inflación anual.
- Life.expect: Expectativa de vida media en años.
- Military
- Pop.growth: tasa de crecimiento poblacional.
- Unemployment: tasa de desempleo.

LIBRERÍAS

- Lenguaje de programación Python.
- Cálculo de **PCA** con la librería Sklearn.
- Gráficos utilizando Plotly.
- Manejo de datos con Pandas.

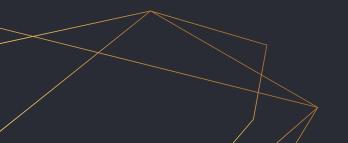


CONSIDERACIONES

Se estandarizó la información de los países de la siguiente manera:

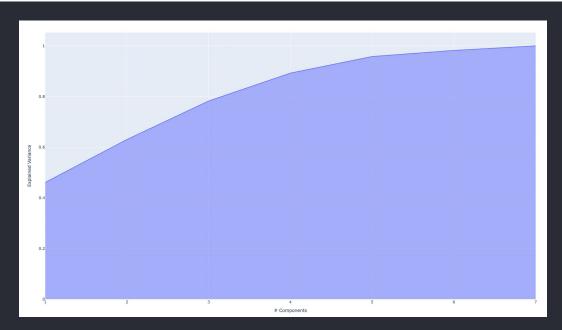
$$\tilde{X}_i = \frac{X_i - \bar{X}_i}{s}$$

RESULTADOS



VARIABILIDAD PCA

# Component	Component 1	Component 2	Component 3	Component 4	Component 5	Component 6	Component 7
Variance	0.4610	0.1696	0.1519	0.1101	0.0654	0.0241	0.0179
Accumulated Variance	0.4610	0.6306	0.7825	0.8925	0.9580	0.9821	1.0000



Se alcanza una variabilidad mayor al 95% a partir de la componente 5.

PRIMER COMPONENTE

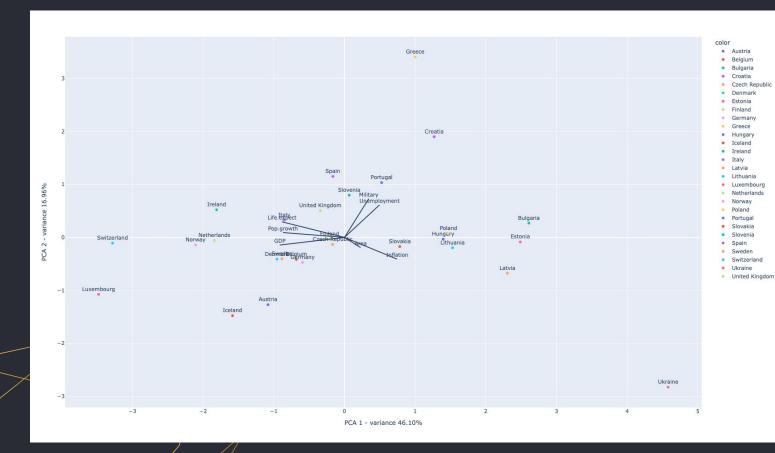
Field	Area	GDP	Inflation	Life.expect	Military	Pop.growth	Unemployment
Value	0.1249	-0.1729	0.8983	0.0449	-0.3240	0.1901	0.0666
Country				Value			
Austria				-1.0817			
Belgium				-0.6811			
Bulgaria				2.6099			
Croatia				1.2701			
Czech Republic				-0.1672			
Denmark				-0.9552			
Estonia				2.4877			
Finland				-0.2106			
Germany				-0.5924			
Greece				1.0005			
Hungary				1.3969			
Iceland				-1.5837			
Ireland				-1.8089			
Italy				-0.8532			
Latvia				2.3061			
Lithuania				1.5301			
Luxembourg				-3.4784			
Netherlands				-1.8401			
Norway				-2.1065			
Poland				1.4718			
Portugal				0.5265			
Slovakia				0.7830			
Slovenia				0.0675			
Spain				-0.1638			
				100000000000000000000000000000000000000			

Ejemplo: Austria tiene alto PIB y baja inflación.

-0.8851

Sweden

PCA BIPLOT



¡GRACIAS!

