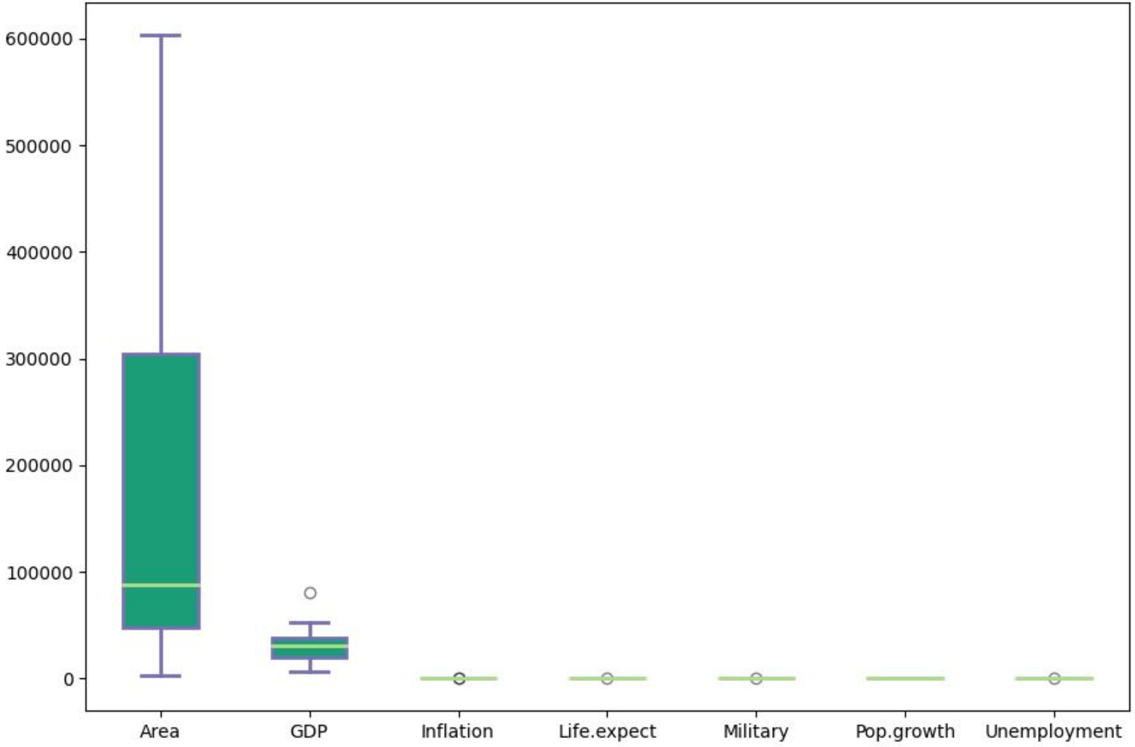
# Librerías utilizadas

- scikit-learn
- numpy
- pandas
- matplotlib
- seaborn

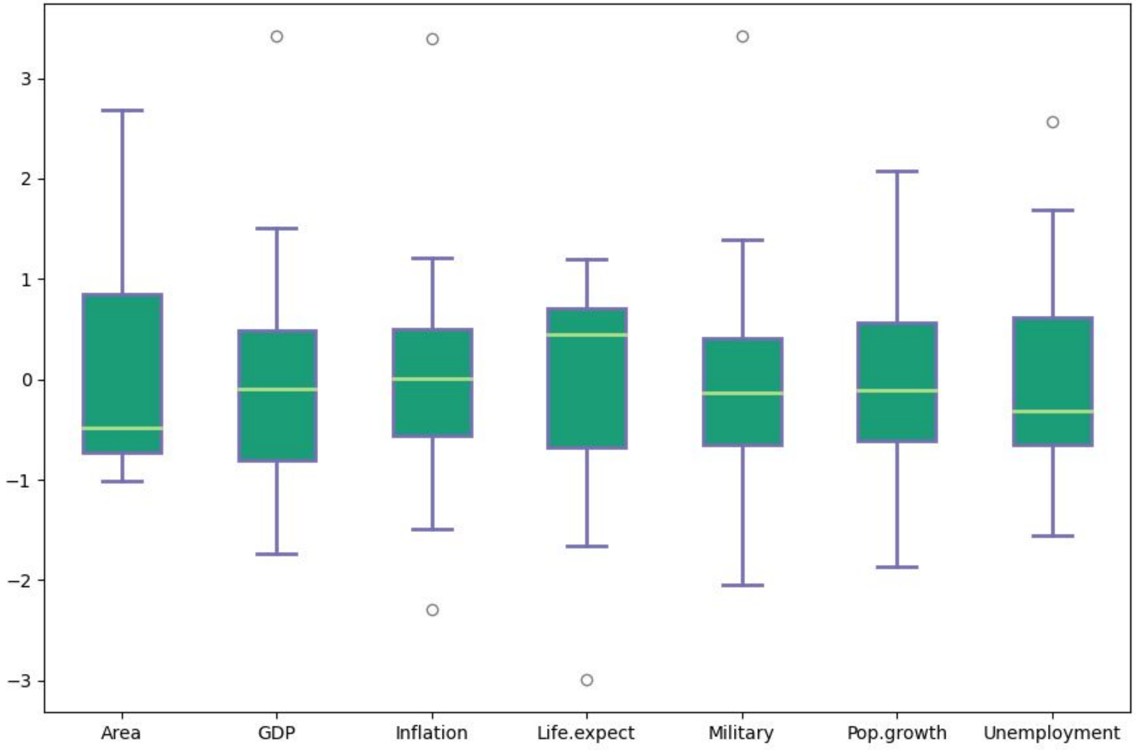
# Variables del conjunto de datos

- 8 variables
- 7 son numéricas **( Area, GDP, Inflation, Life.expect, Military, Pop.growth, Unemployment)**
- 1 es de texto **(Country)**

# Boxplot - Datos no estandarizados



# Boxplot - Datos estandarizados



# Heatmap - Matriz de correlación (7x7)

# Matriz de covarianza (no estandarizada)

```
[[ 2.74030532e+10 -3.32532112e+08  7.40899280e+04 -1.14405584e+04  1.34898400e+04 -7.36391063e+03  1.97070511e+04]
 [-3.32532112e+08  2.10311362e+08 -9.99195767e+03  3.24198796e+0  -3.30582672e+03  5.53555026e+03 -3.58024603e+04]
 [ 7.40899280e+04 -9.99195767e+03  1.95513228e+00 -3.02870370e+00  5.41005291e-02 -3.36021164e-01  1.30015873e+00]
 [-1.14405584e+04  3.24198796e+04 -3.02870370e+00  1.01707157e+01  -1.61641667e-01  1.23525926e+00 -3.66957407e+00]
 [ 1.34898400e+04 -3.30582672e+03  5.41005291e-02 -1.61641667e-01  6.42165476e-01 -1.13574603e-01  1.09564286e+00]
 [-7.36391063e+03  5.53555026e+03 -3.36021164e-01  1.23525926e+00  -1.13574603e-01  2.51969312e-01 -4.10380952e-01]
 [ 1.97070511e+04 -3.58024603e+04  1.30015873e+00 -3.66957407e+00  1.09564286e+00 -4.10380952e-01  2.18824868e+01]]
```

# Matriz de covarianza (estandarizada)

```
[[ 1.03703704 -0.14364715  0.33194482 -0.02247327  0.10545766 -0.09190295  0.02639174]
 [-0.14364715  1.03703704 -0.51100531  0.72693995 -0.2949983   0.78858747  -0.5473025 ]
 [ 0.33194482 -0.51100531  1.03703704 -0.7043479   0.05007077 -0.49647685  0.20613618]
 [-0.02247327  0.72693995 -0.7043479   1.03703704 -0.06559162  0.80020784  -0.25508553]
 [ 0.10545766 -0.2949983   0.05007077 -0.06559162  1.03703704 -0.2928048   0.30310386]
 [-0.09190295  0.78858747 -0.49647685  0.80020784 -0.2928048   1.03703704  -0.18124219]
 [ 0.02639174 -0.5473025   0.20613618 -0.25508553  0.30310386 -0.18124219   1.03703704]]
```

# Matriz de correlación - Autovectores y autovalores

**Autovectores (7x7):**

[ 1.24873902e-01 -1.72872202e-01 8.98296740e-01 -3.24016926e-01 -6.66428246e-02 1.90118083e-01 4.48503976e-02]
[-5.00505858e-01 -1.30139553e-01 8.39557607e-02 3.90632444e-01 3.97408435e-01 6.38657073e-01 -8.42554739e-02]
[ 4.06518155e-01 -3.69657243e-01 1.98194675e-01 6.89500539e-01 2.26700295e-01 -3.23867263e-01 1.64685649e-01]
[-4.82873325e-01 2.65247797e-01 2.46082460e-01 -1.01786561e-01 5.07031305e-01 -6.06434187e-01 2.67714373e-02]
[ 1.88111616e-01 6.58266888e-01 2.43679433e-01 3.68147581e-01 -1.37309597e-01 3.55960680e-02 -5.62374796e-01]
[-4.75703554e-01 8.26219831e-02 1.63697207e-01 3.47867772e-01 -6.71146682e-01 -1.20855625e-01 3.92462767e-01]
[ 2.71655820e-01 5.53203705e-01 5.00135736e-04 1.01587422e-02 2.44662434e-01 2.59704965e-01 7.01967912e-01]

**Autovalores:**

[3.22716568 1.18712341 1.06319053 0.45784862 0.12564189 0.16867389 0.77035598]

# Matriz de covarianza - Autovectores y autovalores

**Autovectores (7x7):**

```
[ 1.24873902e-01 -1.72872202e-01  8.98296740e-01 -3.24016926e-01  -6.66428246e-02  1.90118083e-01  4.48503976e-02]
 [-5.00505858e-01 -1.30139553e-01  8.39557607e-02  3.90632444e-01   3.97408435e-01  6.38657073e-01 -8.42554739e-02]
 [ 4.06518155e-01 -3.69657243e-01  1.98194675e-01  6.89500539e-01   2.26700295e-01 -3.23867263e-01  1.64685649e-01]
 [-4.82873325e-01  2.65247797e-01  2.46082460e-01 -1.01786561e-01   5.07031305e-01 -6.06434187e-01  2.67714373e-02]
 [ 1.88111616e-01  6.58266888e-01  2.43679433e-01  3.68147581e-01  -1.37309597e-01  3.55960680e-02 -5.62374796e-01]
 [-4.75703554e-01  8.26219831e-02  1.63697207e-01  3.47867772e-01  -6.71146682e-01 -1.20855625e-01  3.92462767e-01]
 [ 2.71655820e-01  5.53203705e-01  5.00135736e-04  1.01587422e-02   2.44662434e-01  2.59704965e-01  7.01967912e-01]
```

**Autovalores:**

[3.34669033 1.23109094 1.10256796 0.47480597 0.13029529 0.17492107 0.79888768]

# PCA - Componentes

```
[[ 1.24873902e-01 -5.00505858e-01  4.06518155e-01 -4.82873325e-01  1.88111616e-01 -4.75703554e-01  2.71655820e-01]
 [-1.72872202e-01 -1.30139553e-01 -3.69657243e-01  2.65247797e-01  6.58266888e-01  8.26219831e-02  5.53203705e-01]
 [ 8.98296740e-01  8.39557607e-02  1.98194675e-01  2.46082460e-01  2.43679433e-01  1.63697207e-01  5.00135736e-04]
 [ 4.48503976e-02 -8.42554739e-02  1.64685649e-01  2.67714373e-02 -5.62374796e-01  3.92462767e-01  7.01967912e-01]
 [-3.24016926e-01  3.90632444e-01  6.89500539e-01 -1.01786561e-01  3.68147581e-01  3.47867772e-01  1.01587422e-02]
 [ 1.90118083e-01  6.38657073e-01 -3.23867263e-01 -6.06434187e-01  3.55960680e-02 -1.20855625e-01  2.59704965e-01]
 [ 6.66428246e-02 -3.97408435e-01 -2.26700295e-01 -5.07031305e-01  1.37309597e-01  6.71146682e-01 -2.44662434e-01]]
```

# PCA - Distribución de datos

Primera componente → 46.10%
Segunda componente → 16.96%
Tercera componente → 15.19%
Cuarta componente → 11%
Quinta componente → 6.54%
Sexta componente → 2.41%
Séptima componente → 1.80%

# PCA - Distribución de datos [ACUMULADO]

Primera componente → 46.10%
Segunda componente → 63.06%
Tercera componente → 78.25%
Cuarta componente → 89.25%
Quinta componente → 95.79%
Sexta componente → 98.2%
Séptima componente → 100%

# PCA - Cargas de la primer componente

| AREA | GDP | Inflation | Life.expect | Military | Pop.growth | Unemployement |
|---|---|---|---|---|---|---|
| [ 0.1248739 | -0.50050586 | 0.40651815 | -0.48287333 | 0.18811162 | -0.47570355 | 0.27165582] |

Valores “extremos” positivos (x > 0.25) → características “negativas” de un país (INFLATION, UNEMPLOYEMENT)
Valores “neutros” ( -0.25 =< x <= 0.25)→ características “neutras” de un país (AREA, MILITARY)
Valores “extremos” negativos (x < -0.25) → características “positivas” de un país (GDP, LIFE.EXPECT, POP.GROWTH)