

## ONU - Análise de Objetivo de Desenvolvimento Sustentável

### Estática de Blocos Econômicos e Desenvolvimento

Integrantes:

- Eduardo Mendes Vaz
  - João Lucas de Moraes Barros Cadorniga
- 

### Introdução

Na primeira análise do desenvolvimento mundial em relação ao Objetivo de Desenvolvimento Sustentável IV da ONU — o de investimento e melhoria das oportunidades educacionais —, utilizamos o fator Investimento Governamental em Educação, como % do PIB<sup>1</sup> ao longo dos anos. Agora, com probabilidade e estatística, temos **o objetivo de analisar e classificar dois blocos econômicos** em relação a esses dados.

### 1. Mercosul

O primeiro bloco é o **Mercado Comum do Sul**, um dos maiores do hemisfério sul. Ele surgiu em 1991 como um processo de integração regional, inicialmente entre Argentina, Brasil, Paraguai e Uruguai — hoje admitindo direitos a outras nações sul-americanas.

Desde sua criação com o Tratado de Assunção, seu maior foco foram alicerces de **desenvolvimento econômico** e democrático. No entanto, também vigoram acordos migratórios, trabalhistas, entre outros. O bloco surgiu principalmente pela proximidade geográfica, mas também pela convergência cultural e econômica.

**Principais benefícios do bloco<sup>2</sup>:**

1. Residência legal com igualdade de direitos;
  2. Turismo sem passaporte;
  3. Transferência de educação pelos países;
- **A diminuição ou até isenção de barreiras tarifárias sobre o comércio de bens.** Além de acordos facilitados entre os membros da considerada **quinta economia do mundo**.

### 2. União Européia

Analogamente, o outro bloco analisado é a **União Europeia**, que é considerada um dos mais poderosos do mundo — tanto por possuir inúmeros grandes países como membros, quanto por sua longa história.

Composto por 27 países-membros, o grupo surgiu especialmente por proximidade geográfica no continente, como uma maneira de unir as economias da Europa pelo Tratado de Maastricht, em 1992. Apesar de **foco econômico**, a UE tem grande **solidez em objetivos sustentáveis e sociais**.<sup>3</sup>

### Principais benefícios do bloco<sup>3</sup>:

1. Livre circulação de pessoas sem visto;
2. Acordos de diminuição de taxas e livre circulação de mercadorias;
3. Proteção conjunta dos países;
- **A participação no maior mercado único do mundo, com moeda unificada, o Euro, e serviços telefônicos e digitais unificados.**

<sup>1</sup> Fonte: [World Bank](#)

<sup>2</sup> Fonte: [Mercosul](#)

<sup>3</sup> Fonte: [European Union](#)

## Análise Probabilística, Classificador e Conclusões

Levando em considerações as probabilidades calculadas, criamos um **sistema de classificação**:

- Probabilidade de ser bloco, dado avançado: Mercosul 28.57% x 71.43% UE
- Probabilidade de ser bloco, dado atrasado: Mercosul 30% x 70% UE

Assim, percebemos que nosso classificador terá um **problema**: ao mesmo tempo que classificará países avançados como da UE — probabilidade dominante 71.43% —, também colocará países atrasados no bloco — probabilidade 70%.<sup>1</sup>

Ou seja, ao alimentar os países do estudo, **todos serão classificados como membros da União Européia**.

Portanto podemos concluir que o indicador utilizado para análise é **enviesado**, pois não nos permite usá-lo para criar um classificador que *consiga distinguir os blocos*.

Apesar disso, os **resultados de acurácia** foram interessantes:

- Acertos **Mercosul: 0%**
- Acertos **UE: 100%**
- Acertos **TOTAIS: 71.05%**

São explicadas pelo *viés do modelo*, que causou acerto completo da UE, mas claro, nenhum do Mercosul.

Por fim, conclui-se que esse indicador **não é um bom preditor para esses blocos econômicos** (o que não descarta sua importância na análise das ODEs, ele simplesmente não nos ajudou para criar um modelo entre esses blocos).

---

<sup>1</sup> Cálculos completos abaixo no PDF

**Observação:** Gráfico de distribuição disponível no **final do PDF**

## Classificação dos Países & Probabilidades Iniciais

```
import pandas as pd
import numpy as np
```

```
df =
pd.read_csv('dados/API_SE.XPD.TOTL.GD.ZS_DS2_en_csv_v2_4411877.csv',
skiprows=4)
```

*#OBS: Venezuela está suspensa do Mercosur, portanto, não foi incluída no cálculo*

```
mercosul = df[df['Country Name'].isin(['Uruguay', 'Argentina',
'Paraguay', 'Brazil', 'Bolivia', 'Chile', 'Peru', 'Colombia',
'Ecuador', 'Guyana', 'Suriname'])]
eu = df[df['Country Name'].isin(['Austria', 'Belgium', 'Bulgaria',
'Denmark', 'Finland', 'Germany', 'France', 'Greece', 'Ireland',
'Italy', 'Luxembourg', 'Netherlands', 'Portugal', 'Romania', 'Spain',
'Sweden', 'United Kingdom', 'Cyprus', 'Croatia', 'Czech Republic',
'Estonia', 'Hungary', 'Latvia', 'Lithuania', 'Malta', 'Poland',
'Slovakia', 'Slovenia'])]
```

*# Fator foi calculado pela mediana da % do PIB investido em educação nos anos analisados na APS1 (2004 a 2018), e então calculada a média com todos os anos. O fator foi aplicado para que os valores de cada país fossem comparáveis.*

*# Acima do fator, o país é considerado avançado em relação a ODS, e, abaixo, é considerado atrasado.*

```
fatorx = df[[str(i) for i in range(2004, 2019)]].median().mean() # ~ 4.15
```

*# cria coluna chamada Avançado no dataframe mercosul com True se a mediana dos anos 2004 a 2018 for maior que fatorx*

```
mercosul['Avançado'] = mercosul[[str(i) for i in range(2004, 2019)]].median(axis=1) > fatorx
```

*# cria coluna chamada Avançado no dataframe eu com True se a mediana dos anos 2004 a 2018 for maior que fatorx*

```
eu['Avançado'] = eu[[str(i) for i in range(2004, 2019)]].median(axis=1) > fatorx
```

### # Probabilidades Condicionais

```
p_avancado_dado_mercosul = len(mercosul[mercosul['Avançado'] == True])/len(mercosul)
p_avancado_dado_eu = len(eu[eu['Avançado'] == True])/len(eu)
print(f'Probabilidade de país ser avançado, dado que faz parte do Mercosul: {p_avancado_dado_mercosul * 100 :.2f}%')
print(f'Probabilidade de país ser avançado, dado que faz parte da União Europeia: {p_avancado_dado_eu * 100 :.2f}%')
```

Probabilidade de país ser avançado, dado que faz parte do Mercosul: 72.73%

Probabilidade de país ser avançado, dado que faz parte da União Europeia: 74.07%

```
C:\Users\neodr\AppData\Local\Temp\ipykernel_7140\569763842.py:19:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

See the caveats in the documentation:

[https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
mercosul['Avançado'] = mercosul[[str(i) for i in range(2004, 2019)]].median(axis=1) > fatorx
```

```
C:\Users\neodr\AppData\Local\Temp\ipykernel_7140\569763842.py:22:
SettingWithCopyWarning:
```

A value is trying to be set on a copy of a slice from a DataFrame.  
Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation:

[https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)

```
eu['Avançado'] = eu[[str(i) for i in range(2004, 2019)]].median(axis=1) > fatorx
```

### Considerações:

Para a análise do Mercosul, foram levados em consideração todos os países oficiais do Mercosul (ARG, BRA, PAR, URU), mais os estados associados (BOL, CHI, COL, EQU, GUI, PER, SUR), totalizando **11 países**.

*OBS: a Venezuela, por estar atualmente bloqueada do Mercosul, não foi levada em consideração.*

No estudo da UE, foram levados em consideração todos os **27 países** do bloco.

### Observação:

Para classificar os países de ambos blocos econômicos em **Avançados** ou **Atrasados**, criamos um fator de comparação, calculado pela mediana da porcentagem do PIB

investido em educação nos anos analisados na 1ª entrega (2004 a 2018), e então tirada a média do período. Para ser considerado avançado, **a mediana do investimento de uma nação durante esses anos deve estar acima do fator**.<sup>1</sup>

## Probabilidades Condicionais utilizando o Teorema de Bayes

```
# Dataframe de todos os países analisados
países_analisados = df[df['Country Name'].isin(['Uruguay',
'Argentina', 'Paraguay', 'Brazil', 'Bolivia', 'Chile', 'Peru',
'Colombia', 'Ecuador', 'Guyana', 'Suriname', 'Austria', 'Belgium',
'Bulgaria', 'Denmark', 'Finland', 'Germany', 'France', 'Greece',
'Ireland', 'Italy', 'Luxembourg', 'Netherlands', 'Portugal',
'Romania', 'Spain', 'Sweden', 'United Kingdom', 'Cyprus', 'Croatia',
'Czech Republic', 'Estonia', 'Hungary', 'Latvia', 'Lithuania',
'Malta', 'Poland', 'Slovakia', 'Slovenia'])]
```

```
# Coluna de Avançado
países_analisados['Avançado'] = países_analisados[[str(i) for i in
range(2004, 2019)]] .median(axis=1) > fatorx
```

```
# Probabilidade de país ser avançado no nosso universo de países
analisados
p_avancado = len(países_analisados[países_analisados['Avançado'] ==
True])/len(países_analisados)
print(f'Probabilidade de ser avançado: {p_avancado * 100 :.2f}%')
```

```
# Probabilidade de país fazer parte do Mercosul
p_mercosul = len(mercosul)/len(países_analisados)
print(f'Probabilidade de fazer parte do Mercosul: {p_mercosul * 100
:.2f}%')
```

```
# Probabilidade de país fazer parte da União Europeia
p_eu = len(eu)/len(países_analisados)
print(f'Probabilidade de fazer parte da União Européia: {p_eu * 100
:.2f}%')
```

```
# Probabilidade de país fazer parte do Mercosul, dado que é avançado
p_mercosul_dado_avancado = (p_avancado_dado_mercosul *
p_mercosul)/p_avancado
print(f'Probabilidade de ser do Mercosul, dado que é avançado:
{p_mercosul_dado_avancado * 100 :.2f}%')
```

```
# Probabilidade de país fazer parte da União Europeia, dado que é
avançado
```

```
p_eu_dado_avancado = (p_avancado_dado_eu * p_eu)/p_avancado
print(f'Probabilidade de ser da UE, dado que é avançado::
{p_eu_dado_avancado * 100 :.2f}%')
```

Probabilidade de ser avançado: 73.68%  
 Probabilidade de fazer parte do Mercosul: 28.95%  
 Probabilidade de fazer parte da União Européia: 71.05%  
 Probabilidade de ser do Mercosul, dado que é avançado: 28.57%  
 Probabilidade de ser da UE, dado que é avançado:: 71.43%

C:\Users\neodr\AppData\Local\Temp\ipykernel\_7140\2821323730.py:5:  
 SettingWithCopyWarning:  
 A value is trying to be set on a copy of a slice from a DataFrame.  
 Try using .loc[row\_indexer,col\_indexer] = value instead

See the caveats in the documentation:  
[https://pandas.pydata.org/pandas-docs/stable/user\\_guide/indexing.html#returning-a-view-versus-a-copy](https://pandas.pydata.org/pandas-docs/stable/user_guide/indexing.html#returning-a-view-versus-a-copy)  
 pais\_analisados['Avançado'] = pais\_analisados[[str(i) for i in  
 range(2004, 2019)]] .median(axis=1) > fatorx

### Importante:

A partir da divisão, foram calculadas algumas **probabilidades**.<sup>1</sup> No entanto, quatro delas são as mais importantes, pois ditam o **sistema de classificação desenvolvido**:

País ser...

- do Mercosul, dado que é avançado: 28.57%
- da UE, dado que é avançado: 71.43%
- do Mercosul, dado que é atrasado: 30%
- da UE, dado que é atrasado: 70%

*# Probabilidade de país do universo analisado ser atrasado em relação ao investimento em educação*

```
p_atrasado = len(paises_analisados[paises_analisados['Avançado'] == False])/len(paises_analisados)
print(f'Probabilidade de ser avançado: {p_atrasado * 100 :.2f}%')
```

*# Probabilidade de país do universo analisado ser atrasado, dado que faz parte do Mercosul*

```
p_atrasado_dado_mercosul = len(mercosul[mercosul['Avançado'] == False])/len(mercosul)
```

*# Probabilidade de país do universo analisado ser atrasado, dado que faz parte da União Europeia*

```
p_atrasado_dado_eu = len(eu[eu['Avançado'] == False])/len(eu)
```

*# Probabilidade de país ser do Mercosul, dado que é atrasado*

```
p_mercosul_dado_atrasado = (p_atrasado_dado_mercosul * p_mercosul)/p_atrasado
```

```
print(f'Probabilidade de ser do Mercosul, dado que é atrasado:
{p_mercosul_dado_atrasado * 100 :.2f}%')
```

```
# Probabilidade de país ser da União Europeia, dado que é atrasado
p_eu_dado_atrasado = (p_atrasado_dado_eu * p_eu)/p_atrasado
print(f'Probabilidade de ser da UE, dado que é atrasado:
{p_eu_dado_atrasado * 100 :.2f}%')
```

Probabilidade de ser avançado: 26.32%

Probabilidade de ser do Mercosul, dado que é atrasado: 30.00%

Probabilidade de ser da UE, dado que é atrasado: 70.00%

## Classificador de Países

```
# Classificador enviesado criado pelas probabilidades encontradas de
ser avançado ou atrasado
```

```
def classificador(eh_avancado):
    if eh_avancado:
        return 'UE'
    return 'Mercosul'
```

```
# Cria coluna de classificação aplicando o classificador em todos os
países analisados
```

```
paises_analisados['Classificação'] =
paises_analisados['Avançado'].apply(classificador)
```

```
corretos_mercosul = 0
```

```
corretos_eu = 0
```

```
# itera sobre todos os países analisados e confere se o classificador
acertou
```

```
for index, row in pais_analisados.iterrows():
    if row['Classificação'] == 'UE' and row['Country Name'] in
eu['Country Name'].values:
        corretos_eu += 1
    elif row['Classificação'] == 'Mercosul' and row['Country Name'] in
mercosul['Country Name'].values:
        corretos_mercosul += 1
```

```
print(f'Porcentagem de acertos TOTAL: {(corretos_mercosul +
corretos_eu)/len(paises_analisados) * 100 :.2f}%')
print(f'Porcentagem de acertos para o Mercosul:
{corretos_mercosul/len(mercosul) * 100 :.2f}%')
print(f'Porcentagem de acertos para a UE: {corretos_eu/len(eu) * 100
:.2f}%')
```

```
Porcentagem de acertos TOTAL: 71.05%
Porcentagem de acertos para o Mercosul: 0.00%
Porcentagem de acertos para a UE: 100.00%
```

```
C:\Users\neodr\AppData\Local\Temp\ipykernel_7140\2126375194.py:9:
SettingWithCopyWarning:
A value is trying to be set on a copy of a slice from a DataFrame.
Try using .loc[row_indexer,col_indexer] = value instead
```

```
See the caveats in the documentation:
https://pandas.pydata.org/pandas-docs/stable/user\_guide/indexing.html#
returning-a-view-versus-a-copy
    paises_analisados['Classificação'] =
paises_analisados['Avançado'].apply(classificador)
```

## Extra: Hipóteses sobre os Resultados

(não coube na contagem de palavras :) )

Esse fenômeno de viés do classificador pode ser explicado por algumas hipóteses:

- **Matemática:** o número de países membros da União Europeia é mais do que o dobro do Mercosul (como pode ser observado no gráfico desenvolvido<sup>1</sup>), o que pode causar que os números da UE "engulam" os do Mercosul.
- **Geo-política:** os países da União Europeia são desenvolvidos, enquanto o Mercosul é sub-desenvolvidos. Ao mesmo tempo, a UE também possui grande **disparidade interna**, com países altamente reconhecidos mundialmente, até pequenos países que possuem pouca contribuição econômica mundial, e que então muito provavelmente **são atrasados em uma meta de investimento em educação**. Enquanto isso, o Mercosul possui menor disparidade.

## Distribuição dos Países dos Blocos

```
import matplotlib.pyplot as plt
import seaborn as sns

plt.figure(figsize=(16, 9))

sns.set_theme()
sns.set_context("notebook", font_scale=1.2, rc={"lines.linewidth":
2.5})
sns.axes_style("whitegrid")

dados = pd.DataFrame()

dados['Blocos'] = ['Mercosul', 'União Européia']
dados['Avançado'] = [len(mercocul[mercocul['Avançado'] == True]),
```



```

len(eu[eu['Avançado'] == True]))
dados['Atrasado'] = [len(mercosul[mercosul['Avançado'] == False]),
len(eu[eu['Avançado'] == False])]

dados.plot.bar(x='Blocos', stacked=True, figsize=(16, 9), rot=0)

plt.title('Um país avançado possui média de 4.15% do PIB anual
investido em educação de 2003 a 2018', fontsize=14,
fontweight='light')
plt.suptitle('Distribuição de países avançados e atrasados por bloco
econômico em relação à Investimento em Educação', fontsize=20,
fontweight='bold')
plt.xlabel('Blocos Econômicos', fontsize=14, fontweight='bold')
plt.ylabel('Quantidade de países', fontsize=14, fontweight='bold')

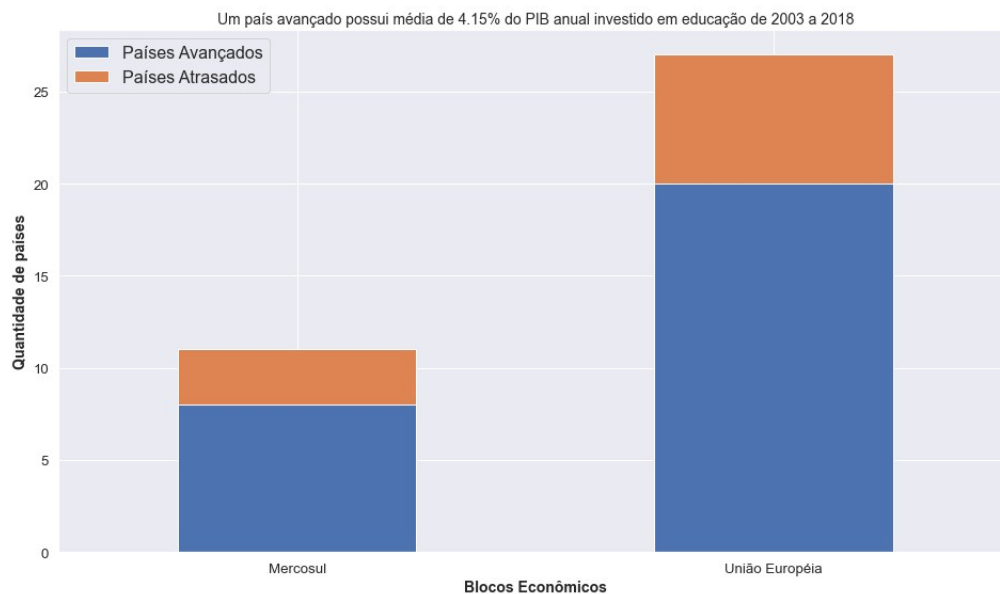
plt.legend(['Países Avançados', 'Países Atrasados'], loc='upper left',
fontsize=16)

plt.show()

```

<Figure size 1152x648 with 0 Axes>

**Distribuição de países avançados e atrasados por bloco econômico em relação à Investimento em Educação**



### Atualizações de Feedback no Trabalho Anterior:

- Pedido: "Faltou a legenda no gráfico da direita."
- Correção: legenda inserida, como visto na figura abaixo.

APS2-Corrigiada