

Contexto del El Banco Mundial

El Banco Mundial es una organización internacional que se dedica a brindar apoyo financiero y técnico a países en desarrollo mediante préstamos para proyectos de infraestructura, sectoriales, institucionales, entre otros, buscando mejorar el bienestar económico y social de su población. Históricamente, el Producto Interno Bruto (PIB) ha sido uno de los indicadores clave para evaluar la capacidad de pago de un país. Sin embargo, si bien existen reportes anuales del PIB de los países, aún no se entiende a la perfección con qué otros indicadores como tasa de empleo, acceso a internet, expectativa de vida, entre otros, se encuentra relacionado.

ENTENDIMIENTO Y PREPARACIÓN DE LOS DATOS

a. Carga de datos:

```
In [23]: gapminder = pd.read_csv('gapminder_final.csv', sep=";", encoding='ISO-8859-1')
```

```
In [24]: gapminder.head()
```

Out[24]:

	country	incomeperperson	alcoholconsumption	armedforcesrate	breastcancerper100th	co2emissions	femaleemployrate	hivrate	internet userate	lifeex
0	pais	Ingreso por persona	Consumo de alcohol	tasa de fuerzas armadas	NaN	NaN	tasa de empleo femenino	Tasa de VIH	Usuarios de Internet	Espe
1	Ireland	2.759.509.135	14.92	0.4500244	74.9	1633778667.0	51.0	0.2	6.977.039.441	
2	Albania	1.914.996.551	7.29	10.247.361	57.4	223747333.3	4.209.999.847	NaN	4.498.994.696	
3	Central African Rep.	2.395.187.494	3.17	0.1486587	16.5	8338000.0	6.290.000.153	4.7	2.300.026.653	
4	Barbados	9.243.587.053	6.42	0.6639555	62.5	36160666.67	6.029.999.924	1.4	7.002.859.927	

lifeexpectancy	oilperperson	polityscore	relectricperperson	suicideper100th	employrate	urbanrate	Unnamed: 16
Esperanza de vida	Petróleo por persona	Puntuación de política	consumo electricidad	suicidio	tasa empleo	tasa urbana	NaN
80.557	1.700.261.751	10.0	2.051.802.338	1.036.507	5.990.000.153	61.34	NaN
76.918	NaN	9.0	6.363.413.834	7.699.329.853	5.140.000.153	46.72	NaN
48.398	NaN	-1.0	NaN	1.354.841.995	7.130.000.305	38.58	NaN
76.835	NaN	NaN	NaN	3.108.602.524	6.690.000.153	39.84	NaN

b. Exploración de los datos: por Columnas

```
In [25]: gapminder.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 179 entries, 0 to 178
Data columns (total 17 columns):
#   Column                Non-Null Count  Dtype
---  ---
0   country                179 non-null    object
1   incomeperperson        176 non-null    object
2   alconsumption          179 non-null    object
3   armedforcesrate        175 non-null    object
4   breastcancerper100th   177 non-null    object
5   co2emissions           176 non-null    object
6   femaleemployrate       178 non-null    object
7   hivrate                160 non-null    object
8   internet userate       171 non-null    object
9   lifeexpectancy         179 non-null    object
10  oilperperson           73 non-null     object
11  polityscore            162 non-null    object
12  relectricperperson      144 non-null    object
13  suicideper100th        177 non-null    object
14  employrate             178 non-null    object
15  urbanrate              179 non-null    object
16  Unnamed: 16            6 non-null      float64
dtypes: float64(1), object(16)
memory usage: 23.9+ KB
```

c. Exploración datos: Tamaño

```
In [5]: print('Cantidad de Filas y columnas:',gapminder.shape)
print('Nombre columnas:',gapminder.columns)

Cantidad de Filas y columnas: (179, 17)
Nombre columnas: Index(['country', 'incomeperperson', 'alconsumption', 'armedforcesrate',
                        'breastcancerper100th', 'co2emissions', 'femaleemployrate', 'hivrate',
                        'internet userate', 'lifeexpectancy', 'oilperperson', 'polityscore',
                        'relectricperperson', 'suicideper100th', 'employrate', 'urbanrate',
                        'Unnamed: 16'],
                        dtype='object')
```

d. Datos NULOS

```
In [14]: gapminder.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 179 entries, 0 to 178
Data columns (total 1 columns):
#   Column
Non-Null Count  Dtype
---  ---
0   country;incomeperperson;alcoconsumption;armedforcesrate;breastcancerper100th;co2emissions;femaleemployrate;hivrate;internet use
rate;lifexpectancy;oilperperson;polityscore;relectricperperson;suicideper100th;employrate;urbanrate; 179 non-null    object
dtypes: object(1)
memory usage: 1.5+ KB
```

e. Distribución de las Variables Numéricas

```
In [7]: gapminder.describe()
```

Out[7]:

Unnamed: 16	
count	6.000000
mean	56.866667
std	20.192730
min	30.640000
25%	40.805000
50%	64.120000
75%	66.900000
max	81.460000

f. Variables categóricas

```
In [12]: gapminder.describe(include=['O'])
```

Out[12]:

	country	incomeperperson	alcoconsumption	armedforcesrate	breastcancerper100th	co2emissions	femaleemployrate	hivrate	internet userate	lit
count	179	176	179	175	177	176	178	160	171	
unique	166	162	161	163	135	164	146	52	161	
top	Lithuania	Rep.	5.56	11.421.267	28.1	1633778667.0	4.209.999.847	0.1	563.000.337	
freq	2	3	2	2	7	2	3	29	2	

lifeexpectancy	oilperperson	polityscore	relectricperperson	suicideper100th	employrate	urbanrate
179	73	162	144	177	178	179
165	66	22	131	166	134	165
72.974	1.487.412.187	10.0	6.149.072.872	3.334.186	4.729.999.924	82.44
4	2	35	2	2	4	2

g. Tipos de columnas

h.

```
In [13]: gapminder.dtypes.value_counts()

Out[13]: object      16
         float64     1
         dtype: int64
```

i. Datos faltantes

```
In [17]: (
         gapminder
         .isnull()
         .sum()
         )

Out[17]: country      0
         incomeperperson  3
         alccconsumption  0
         armedforcesrate  4
         breastcancerper100th  2
         co2emissions  3
         femaleemployrate  1
         hivrate      19
         internet userate  8
         lifeexpectancy  0
         oilperperson  106
         polityscore    17
         relectricperperson  35
         suicideper100th  2
         employrate     1
         urbanrate      0
         Unnamed: 16    173
         dtype: int64
```

- **Top 5 de los indicadores más Importantes:**

1. Ingreso por persona (incomeperperson)
2. tasa empleo (employrate)
3. tasa de empleo femenino (femaleemployrate)
4. Petróleo por persona (oilperperson)
5. Usuarios de Internet (internet userate)

```
gapminder.describe(include=object)
```

	country	incomeperperson	alcoconsumption	armedforcesrate	breastcancerper100th	co2emissions	femaleemployrate
count	179	178	179	175	177	178	
unique	166	162	161	163	135	164	
top	Lithuania	Rep.	5.56	11.421.267	28.1	1633778667.0	4.209.99
freq	2	3	2	2	7	2	

```
gapminder.describe(include=object)
```

	hivrate	internet userate	lifeexpectancy	oilperperson	polityscore	relectricperperson	suicideper100th	employrate
	160	171	179	73	162	144	177	178
	52	161	165	66	22	131	166	134
	0.1	563.000.337	72.974	1.487.412.187	10.0	6.149.072.872	3.334.186	4.729.999.924
	29	2	4	2	35	2	2	4

MODELO DE REGRESIÓN:

Creación de modelo

```
gapminder.sample(5)
```

	country	incomeperperson	alcoholconsumption	armedforcesrate	breastcancerper100th	co2emissions	femaleemployrate	hivrate	internet userate
140	Malawi	1.841.417.988	1.44	0.1145925	10.5	28490000.0	69.0	11.0	2.259.975.885
96	Togo	2.852.244.493	1.92	0.3400308	28.1	32233666.67	4.840.000.153	3.2	5.379.819.715
63	Bangladesh	5.580.628.766	0.17	0.2878915	16.6	598774000.0	5.359.999.847	0.06	370.000.326
76	Rwanda	3.382.663.912	9.99	0.7268554	8.8	20152000.0	80.5	2.9	1.300.000.612
40	Oman	1.119.181.101	0.92	44.659.976	13.2	487993000.0	2.260.000.038	0.1	6.198.741.286

lifeexpectancy	oilperperson	polityscore	relectricperperson	suicideper100th	employrate	urbanrate	Unnamed: 16
54.21	NaN	6.0	NaN	1.017.173.481	7.180.000.305	18.8	NaN
57.062	NaN	-4.0	6.623.852.193	6.369.887.829	6.390.000.153	42.0	NaN
68.944	0.032281466	5.0	6.811.522.858	1.453.835.678	6.809.999.847	27.14	NaN
55.442	NaN	-3.0	NaN	1.311.794.853	7.980.000.305	18.34	NaN
72.974	NaN	-8.0	2.124.608.816	3.716.739.416	5.070.000.076	71.62	NaN

Perfilando datos

```
gapminder["country"].describe(percentiles=[.25, .5, .75, .95, .99])
```

```
count      179
unique      166
top      Lithuania
freq         2
Name: country, dtype: object
```

```
gapminder["incomeperperson"].describe(percentiles=[.25, .5, .75, .95, .99])
```

```
count      176
unique      162
top      Rep.
freq         3
Name: incomeperperson, dtype: object
```

```
gapminder["femaleemployrate"].describe(percentiles=[.25, .5, .75, .95, .99])
```

```
count      178
unique      146
top      4.209.999.847
freq         3
Name: femaleemployrate, dtype: object
```

```
gapminder["oilperperson"].describe(percentiles=[.25, .5, .75, .95, .99])
```

```
count      73
unique      66
top      1.487.412.187
freq         2
Name: oilperperson, dtype: object
```

```
gapminder["employrate"].describe(percentiles=[.25, .5, .75, .95, .99])
```

```
count      178
unique      134
top      4.729.999.924
freq         4
Name: employrate, dtype: object
```

```
gapminder["internet userate"].describe(percentiles=[.25, .5, .75, .95, .99])
```

count	171
unique	161
top	563.000.337
freq	2

Name: internet userate, dtype: object

GENERACIÓN DE RESULTADOS

Dado que un alto PIB puede representar un buen comportamiento de pago de un país, a partir de las relaciones más importantes que, según el modelo de regresión, existe entre el PIB y otros indicadores,

¿qué conjunto de políticas públicas recomendaría implementar, a partir de la premisa de que la mejora en estas áreas indicaría al Banco Mundial que el país es estable, está en una trayectoria de desarrollo sostenible y tiene la capacidad de administrar y reembolsar préstamos de manera efectiva?

Para lograr la transición hacia un desarrollo sostenible, se necesita iniciar transformaciones económicas y sociales, empezando por una reestructuración productiva que satisfaga el triple criterio de elevar la competitividad, disminuir los rezagos sociales. Para ello se necesita incrementar el ahorro interno, hasta ahora insuficiente para sostener una acumulación endógena de capital en un volumen que permita revertir la situación de pobreza que aqueja a un alto porcentaje de la población. Además, se hace imprescindible aumentar el gasto social, sobre todo en educación y salud, y crear empleos de calidad, otorgando especial atención a la equidad de género y a una mejor inserción social de los jóvenes.