

# Final Project



**Data Science**



**Nama : Yong Richard Sriwijaya**

# Latar Belakang



HELP International

Dana \$ 10 juta



Memberikan  
Bantuan



Poor Country

- Social and Economy
- Health

Diperlukan Data Analysis:

Pengkategorian Negara yang  
membutuhkan Dana

# Reading and Checking Data

## Membaca File Data\_Negara\_HELP.csv



	Negara	Kematian_anak	Ekspor	Kesehatan	Impor	Pendapatan	Inflasi	Harapan_hidup	Jumlah_fertiliti	GDPperkapita
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

# Reading and Checking Data

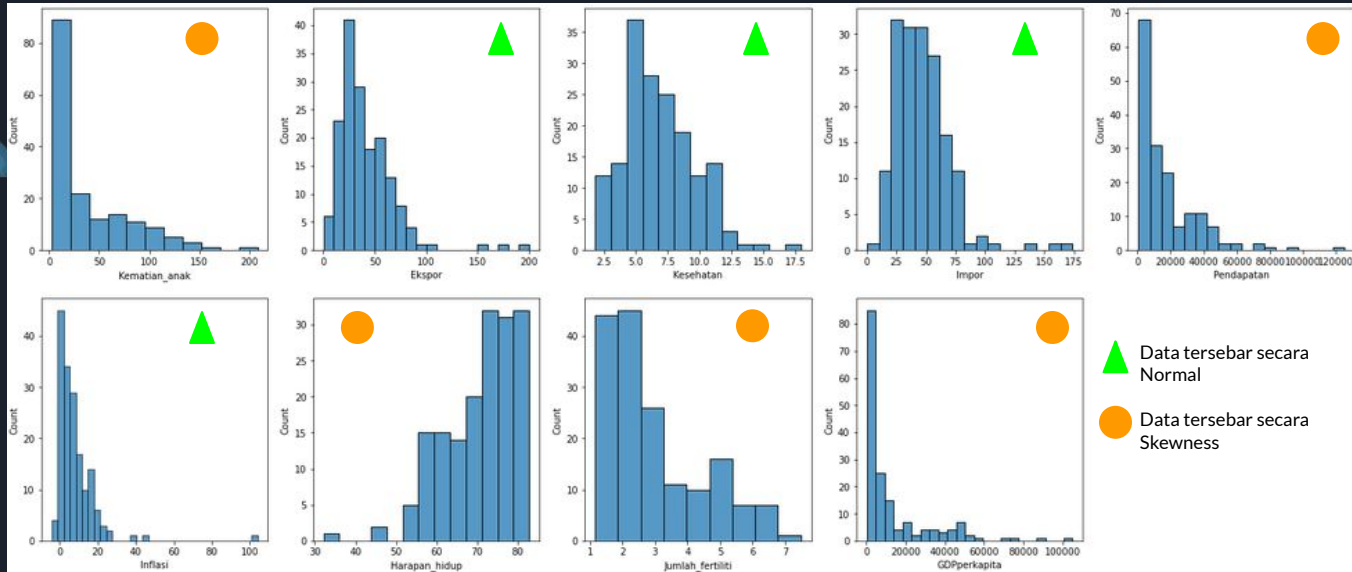
Memeriksa data “NaN” ataupun berbeda format

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 167 entries, 0 to 166
Data columns (total 10 columns):
#   Column                Non-Null Count  Dtype
---  -
0   Negara                167 non-null   object
1   Kematian_anak         167 non-null   float64
2   Ekspor                167 non-null   float64
3   Kesehatan             167 non-null   float64
4   Impor                 167 non-null   float64
5   Pendapatan            167 non-null   int64
6   Inflasi               167 non-null   float64
7   Harapan_hidup         167 non-null   float64
8   Jumlah_fertiliti     167 non-null   float64
9   GDPperkapita          167 non-null   int64
dtypes: float64(7), int64(2), object(1)
memory usage: 13.2+ KB
```

Tidak ada data yang hilang dan berada dalam satu format

# Exploratory Data Analysis (EDA)

## Univariate Analysis



“Sebaran Data secara Skewness ini Menunjukkan Bahwa Terjadi Perbedaan yang CUKUP besar pada Negara Tertentu”.

# Exploratory Data Analysis (EDA)

## Univariate Analysis

Berdasarkan Data Sebelumnya  
Data skewness meliputi:

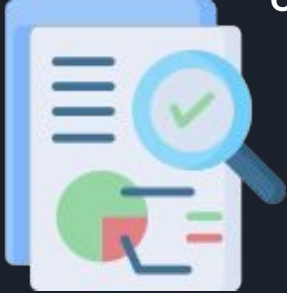
1. Kematian Anak
2. Harapan Hidup
3. Fertiliti

Health

4. Pendapatan
5. GDP Perkapita

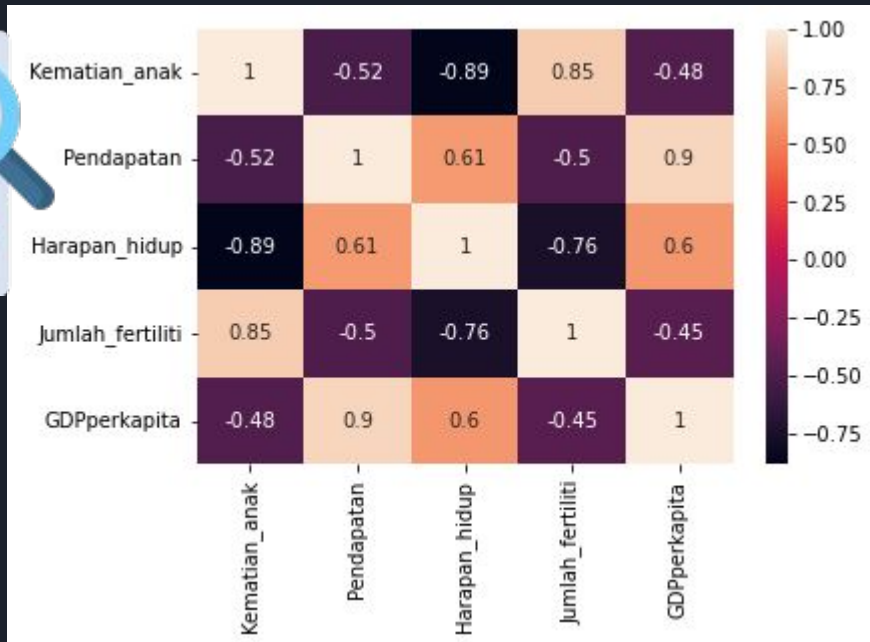
Economy

Data Skewness sudah mendukung  
prasyarat kategori yang ditetapkan  
HELP International

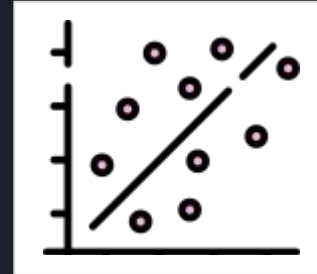


# Exploratory Data Analysis (EDA)

## Multivariate Analysis



## Korelasi Antara Data:



### A. Economy:

1. GDP Perkapita : Pendapatan = 0.9 (Saling Terkait)

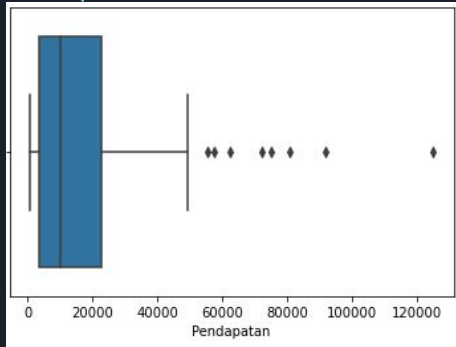
### B. Health:

1. Harapan Hidup : Jumlah Fertiliti = -0.76 (Saling Terkait)
2. Harapan Hidup : Kematian Anak = -0.89 (Saling Terkait)
3. Jumlah Fertiliti : Kematian Anak = 0.85 (Saling Terkait)

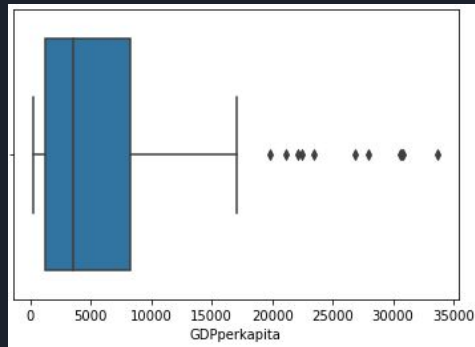
# Outlier Treatment

## Data Economy

Data Asli

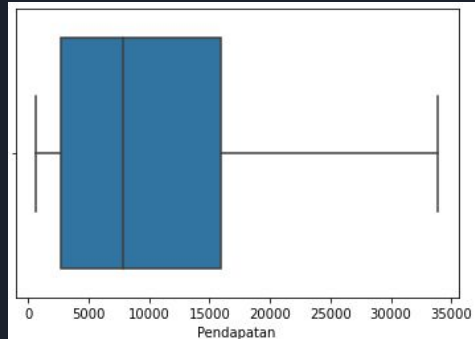
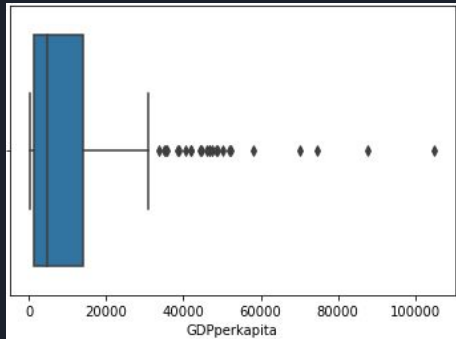


Data Tanpa Outlier



Menghindari hilangnya data.  
Maka hanya dilakukan 2x penghilangan data outlier

Remove Outliers  
Pengulangan 2X

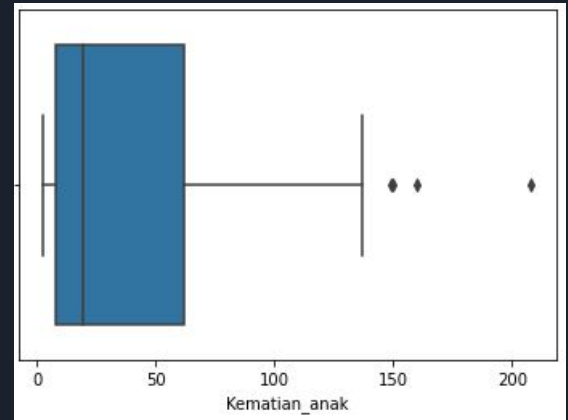
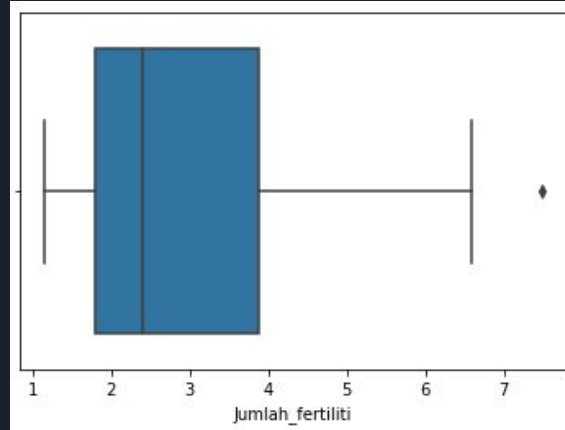
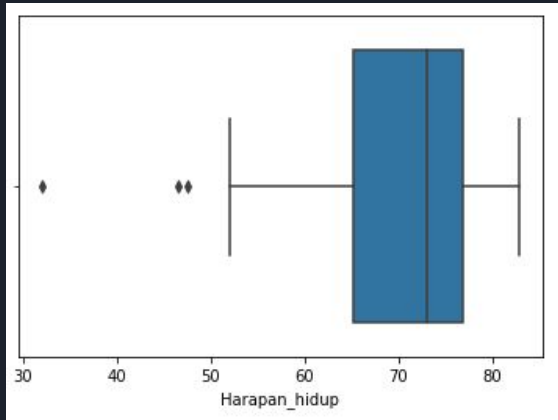




# Outlier Treatment

## Data Health

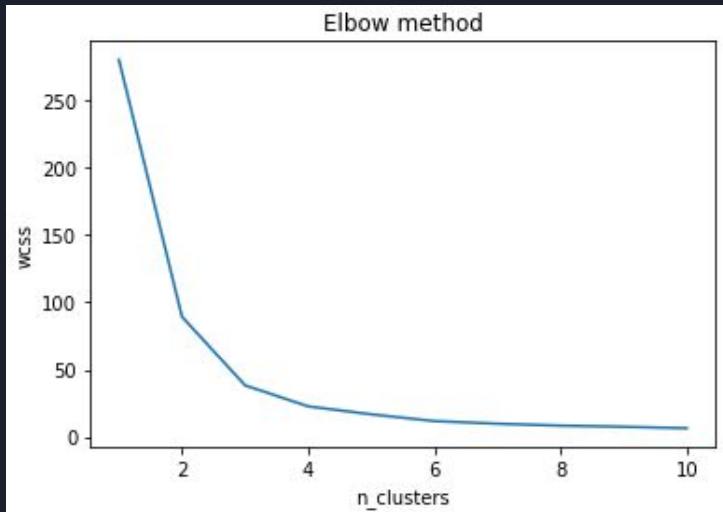
Data Asli



**Data Outliers = Data Penting.  
Tidak dilakukan Outlier Treatment**

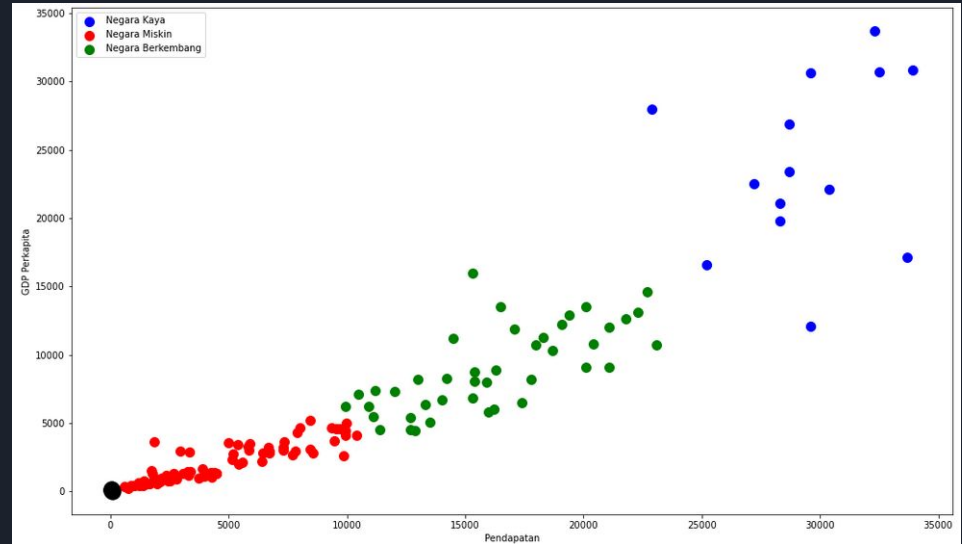
# Clustering Data - KMeans Data Economy

## Elbow Method



Membentuk Siku pada N Cluster = 3

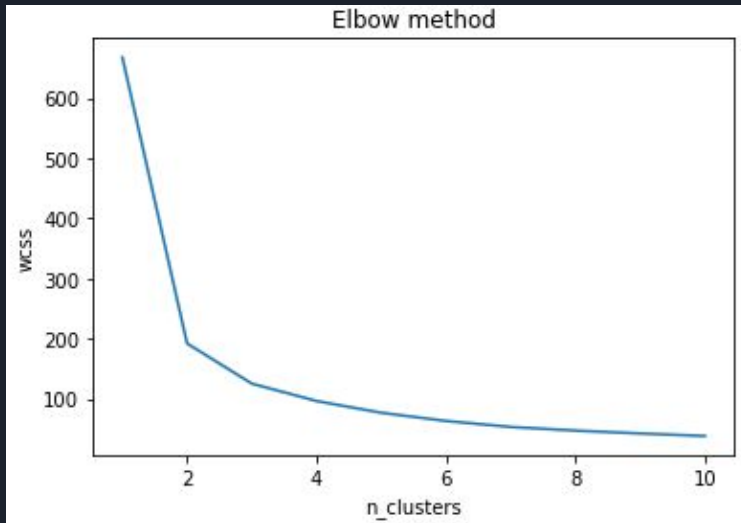
## Data Hasil Pengkelompokan



Pengelompokan Data menjadi 3 Bagian

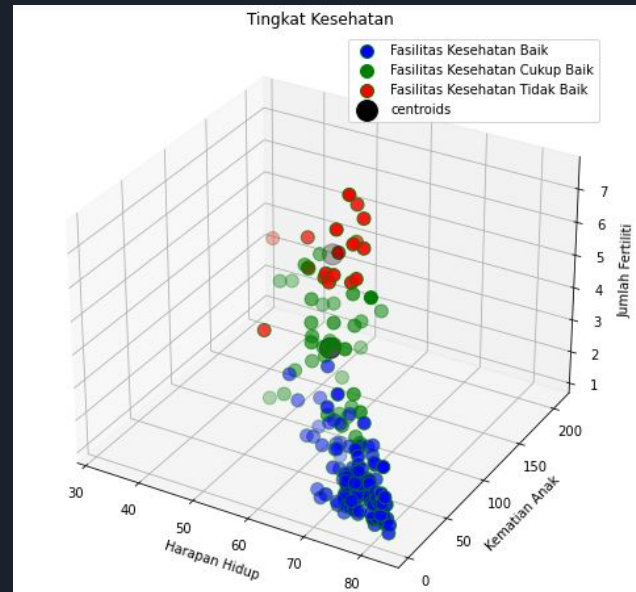
# Clustering Data - KMeans Data Health

## Elbow Method



Membentuk Siku pada N Cluster = 3

## Data Hasil Pengkelompokan



Pengelompokan Data menjadi 3 Bagian

# Combine Data

## Data Economy

### Data Awal + Label (Tingkat Ekonomi)

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Ekonomi
0	Afghanistan	90.2	1610	56.2	5.82	553	1.0
1	Albania	16.6	9930	76.3	1.65	4090	1.0
2	Algeria	27.3	12900	76.5	2.89	4460	2.0
3	Angola	119.0	5900	60.1	6.16	3530	1.0
4	Antigua and Barbuda	10.3	19100	76.8	2.13	12200	2.0
...	...	...	...	...	...	...	...
162	Vanuatu	29.2	2950	63.0	3.50	2970	1.0
163	Venezuela	17.1	16500	75.4	2.47	13500	2.0
164	Vietnam	23.3	4490	73.1	1.95	1310	1.0
165	Yemen	56.3	4480	67.5	4.67	1310	1.0
166	Zambia	83.1	3280	52.0	5.40	1460	1.0

167 rows x 7 columns

**Tingkat Ekonomi = 1**  
**Artinya Negara Miskin**

### Data Negara dengan Tingkat Ekonomi Rendah

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Ekonomi
0	Afghanistan	90.2	1610	56.2	5.82	553	1.0
1	Albania	16.6	9930	76.3	1.65	4090	1.0
3	Angola	119.0	5900	60.1	6.16	3530	1.0
6	Armenia	18.1	6700	73.3	1.69	3220	1.0
12	Bangladesh	49.4	2440	70.4	2.33	758	1.0
...	...	...	...	...	...	...	...
161	Uzbekistan	36.3	4240	68.8	2.34	1380	1.0
162	Vanuatu	29.2	2950	63.0	3.50	2970	1.0
164	Vietnam	23.3	4490	73.1	1.95	1310	1.0
165	Yemen	56.3	4480	67.5	4.67	1310	1.0
166	Zambia	83.1	3280	52.0	5.40	1460	1.0

84 rows x 7 columns

**Daftar Negara Miskin yang Berjumlah 84**

# Combine Data

## Data Economy

### Daftar Nama Negara Miskin Secara Lengkap:

```
['Afghanistan' 'Albania' 'Angola' 'Armenia' 'Bangladesh' 'Belize' 'Benin'  
'Bhutan' 'Bolivia' 'Bosnia and Herzegovina' 'Burkina Faso' 'Burundi'  
'Cambodia' 'Cameroon' 'Cape Verde' 'Central African Republic' 'Chad'  
'China' 'Comoros' 'Congo, Dem. Rep.' 'Congo, Rep.' 'Cote d'Ivoire'  
'Ecuador' 'Egypt' 'El Salvador' 'Eritrea' 'Fiji' 'Gambia' 'Georgia'  
'Ghana' 'Guatemala' 'Guinea' 'Guinea-Bissau' 'Guyana' 'Haiti' 'India'  
'Indonesia' 'Jamaica' 'Jordan' 'Kenya' 'Kiribati' 'Kyrgyz Republic' 'Lao'  
'Lesotho' 'Liberia' 'Madagascar' 'Malawi' 'Mali' 'Mauritania'  
'Micronesia, Fed. Sts.' 'Moldova' 'Mongolia' 'Morocco' 'Mozambique'  
'Myanmar' 'Namibia' 'Nepal' 'Niger' 'Nigeria' 'Pakistan' 'Paraguay'  
'Peru' 'Philippines' 'Rwanda' 'Samoa' 'Senegal' 'Sierra Leone'  
'Solomon Islands' 'Sri Lanka' 'Sudan' 'Tajikistan' 'Tanzania'  
'Timor-Leste' 'Togo' 'Tonga' 'Tunisia' 'Turkmenistan' 'Uganda' 'Ukraine'  
'Uzbekistan' 'Vanuatu' 'Vietnam' 'Yemen' 'Zambia']
```

**Karena Jumlah Negara ini MASIH BANYAK  
Perlu ada nya pembandingan data antara data Economy  
dan Health**

# Combine Data

## Data Health

### Data Awal + Label (Tingkat Kesehatan)

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Kesehatan
0	Afghanistan	90.2	1610	56.2	5.82	553	1
1	Albania	16.6	9930	76.3	1.65	4090	0
2	Algeria	27.3	12900	76.5	2.89	4460	0
3	Angola	119.0	5900	60.1	6.16	3530	2
4	Antigua and Barbuda	10.3	19100	76.8	2.13	12200	0
...	...	...	...	...	...	...	...
162	Vanuatu	29.2	2950	63.0	3.50	2970	0
163	Venezuela	17.1	16500	75.4	2.47	13500	0
164	Vietnam	23.3	4490	73.1	1.95	1310	0
165	Yemen	56.3	4480	67.5	4.67	1310	1
166	Zambia	83.1	3280	52.0	5.40	1460	1

167 rows x 7 columns

**Tingkat Kesehatan = 2**  
**Artinya Negara dengan Fasilitas Kesehatan Buruk**

### Data Negara dengan Tingkat Fasilitas Kesehatan Rendah

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Kesehatan
3	Angola	119.0	5900	60.1	6.16	3530	2
17	Benin	111.0	1820	61.8	5.36	758	2
25	Burkina Faso	116.0	1430	57.9	5.87	575	2
28	Cameroon	108.0	2660	57.3	5.11	1310	2
31	Central African Republic	149.0	888	47.5	5.21	446	2
32	Chad	150.0	1930	56.5	6.59	897	2
37	Congo, Dem. Rep.	116.0	609	57.5	6.54	334	2
40	Cote d'Ivoire	111.0	2690	56.3	5.27	1220	2
49	Equatorial Guinea	111.0	33700	60.9	5.21	17100	2
63	Guinea	109.0	1190	58.0	5.34	648	2
64	Guinea-Bissau	114.0	1390	55.6	5.05	547	2
66	Haiti	208.0	1500	32.1	3.33	662	2
87	Lesotho	99.7	2380	46.5	3.30	1170	2
97	Mali	137.0	1870	59.5	6.55	708	2
106	Mozambique	101.0	918	54.5	5.56	419	2
112	Niger	123.0	814	58.8	7.49	348	2
113	Nigeria	130.0	5150	60.5	5.84	2330	2
132	Sierra Leone	160.0	1220	55.0	5.20	399	2

**Terdapat 18 Negara**

# Combine Data

## Data Health-Economy

### Penggabungan Data Health-Economy

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Ekonomi	Tingkat_Kesehatan
0	Afghanistan	90.2	1610	56.2	5.82	553	1.0	1
1	Albania	16.6	9930	76.3	1.65	4090	1.0	0
2	Algeria	27.3	12900	76.5	2.89	4460	2.0	0
3	Angola	119.0	5900	60.1	6.16	3530	1.0	2
4	Antigua and Barbuda	10.3	19100	76.8	2.13	12200	2.0	0
...	...	...	...	...	...	...	...	...
162	Vanuatu	29.2	2950	63.0	3.50	2970	1.0	0
163	Venezuela	17.1	16500	75.4	2.47	13500	2.0	0
164	Vietnam	23.3	4490	73.1	1.95	1310	1.0	0
165	Yemen	56.3	4480	67.5	4.67	1310	1.0	1
166	Zambia	83.1	3280	52.0	5.40	1460	1.0	1

167 rows x 8 columns

1. **Tingkat Ekonomi = 1**  
Artinya Negara Miskin
2. **Tingkat Kesehatan = 2**  
Artinya Negara dengan Fasilitas Kesehatan Buruk

### Data Negara Miskin dengan Tingkat Fasilitas Kesehatan Rendah

	Negara	Kematian_anak	Pendapatan	Harapan_hidup	Jumlah_fertiliti	GDPperkapita	Tingkat_Ekonomi	Tingkat_Kesehatan
3	Angola	119.0	5900	60.1	6.16	3530	1.0	2
17	Benin	111.0	1820	61.8	5.36	758	1.0	2
25	Burkina Faso	116.0	1430	57.9	5.87	575	1.0	2
28	Cameroon	108.0	2660	57.3	5.11	1310	1.0	2
31	Central African Republic	149.0	888	47.5	5.21	446	1.0	2
32	Chad	150.0	1930	56.5	6.59	897	1.0	2
37	Congo, Dem. Rep.	116.0	609	57.5	6.54	334	1.0	2
40	Cote d'Ivoire	111.0	2690	56.3	5.27	1220	1.0	2
63	Guinea	109.0	1190	58.0	5.34	648	1.0	2
64	Guinea-Bissau	114.0	1390	55.6	5.05	547	1.0	2
66	Haiti	208.0	1500	32.1	3.33	662	1.0	2
87	Lesotho	99.7	2380	46.5	3.30	1170	1.0	2
97	Mali	137.0	1870	59.5	6.55	708	1.0	2
106	Mozambique	101.0	918	54.5	5.56	419	1.0	2
112	Niger	123.0	814	58.8	7.49	348	1.0	2
113	Nigeria	130.0	5150	60.5	5.84	2330	1.0	2
132	Sierra Leone	160.0	1220	55.0	5.20	399	1.0	2

Terdapat 17 Negara

Negara Equatorial Guinea Mempunyai Tingkat Ekonomi yang cukup Tinggi = Tidak Masuk Kriteria

# Kesimpulan



HELP International

Dana \$ 10 juta



Memberikan  
Bantuan



Ke 17 Negara