

HELP INTERNATIONAL

Final Project – Krisna Sandy Pribadi



ABOUT US

HELP International adalah LSM kemanusiaan internasional yang berkomitmen untuk memerangi kemiskinan dan menyediakan fasilitas dan bantuan dasar bagi masyarakat di negara-negara terbelakang saat terjadi bencana dan bencana alam.

OUR OBJECTIVE

Untuk mengkategorikan negara menggunakan faktor sosial ekonomi dan kesehatan yang menentukan pembangunan negara secara keseluruhan.

OUR PROBLEM

HELP International telah berhasil mengumpulkan sekitar \$ 10 juta. Saat ini, CEO LSM perlu memutuskan bagaimana menggunakan uang ini secara strategis dan efektif. Jadi, CEO harus mengambil keputusan untuk memilih negara yang paling membutuhkan bantuan. Oleh karena itu, Tugas teman-teman adalah mengkategorikan negara menggunakan beberapa faktor sosial ekonomi dan kesehatan yang menentukan perkembangan negara secara keseluruhan. Kemudian kalian perlu menyarankan negara mana saja yang paling perlu menjadi fokus CEO.

DATASET

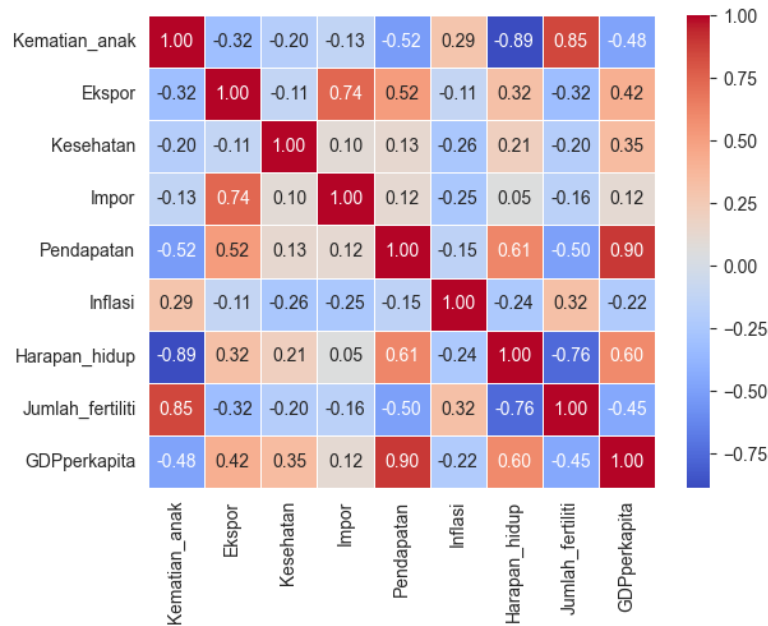
Dataset yang akan diolah terdiri dari 167 baris data negara dan 10 kolom dengan penjelasan sebagai berikut:

- **Negara** : Nama negara
- **Kematian_anak**: Kematian anak di bawah usia 5 tahun per 1000 kelahiran
- **Ekspor** : Ekspor barang dan jasa perkapita
- **Kesehatan**: Total pengeluaran kesehatan perkapita
- **Impor**: Impor barang dan jasa perkapita
- **Pendapatan**: Penghasilan bersih perorang
- **Inflasi**: Pengukuran tingkat pertumbuhan tahunan dari Total GDP
- **Harapan_hidup**: Jumlah tahun rata-rata seorang anak yang baru lahir akan hidup jika pola kematian saat ini tetap sama
- **Jumlah_fertiliti**: Jumlah anak yang akan lahir dari setiap wanita jika tingkat kesuburan usia saat ini tetap sama
- **GDPperkapita**: GDP per kapita. Dihitung sebagai Total GDP dibagi dengan total populasi.

```
<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
```

No Missing Values

MULTIVARIATE ANALYSIS



Berdasarkan hasil analisis multivariat tersebut kita dapatkan variable dengan nilai korelasi tertinggi adalah sebagai berikut:

- GDPperkapita dengan Pendapatan (0.90)
- Jumlah_fertiliti dengan Kematian_anak (0.85)

Dari data tersebut, variable dengan korelasi yang tinggi merupakan variable yang berada dalam satu kategori, GDPperkapita dan Pendapatan termasuk dalam kategori faktor ekonomi dan Jumlah_fertiliti dan Kematian_anak termasuk dalam faktor kesehatan.

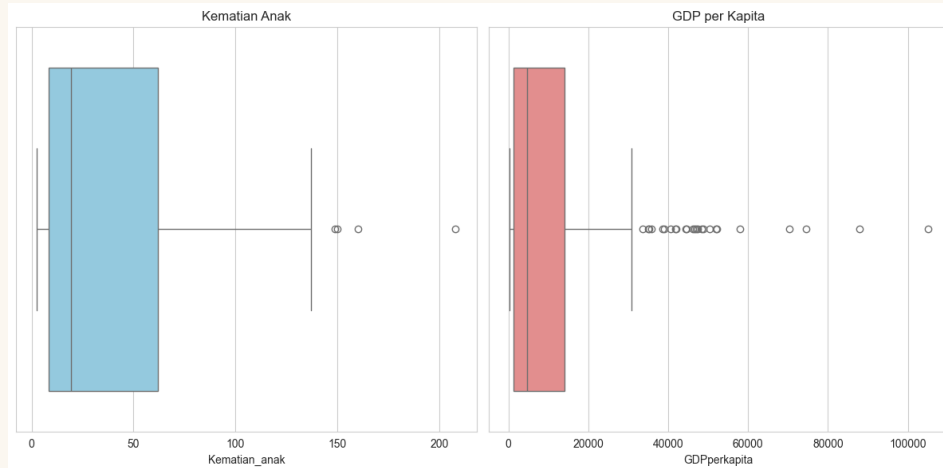
Diputuskan untuk menggunakan 2 variable sebagai dasar analisis dan clustering selanjutnya, yaitu, GDPperkapita dan Kematian_anak.

Dua data ini dipilih karena:

- Keduanya merupakan indikator yang komprehensif dan mencakup berbagai aspek penting dari kesehatan dan ekonomi.
- Keduanya dapat diakses dengan mudah dan akurat.
- Keduanya dapat digunakan untuk membandingkan kesejahteraan berbagai negara.

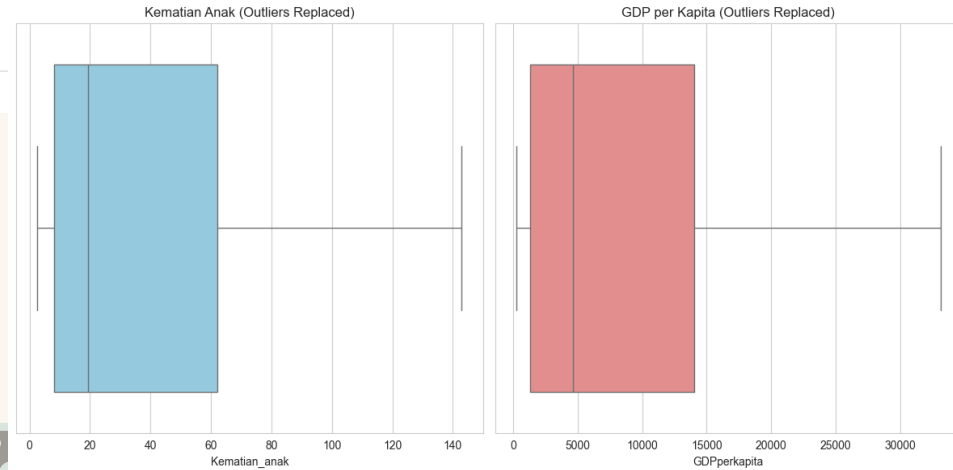
OUTLIERS

Before



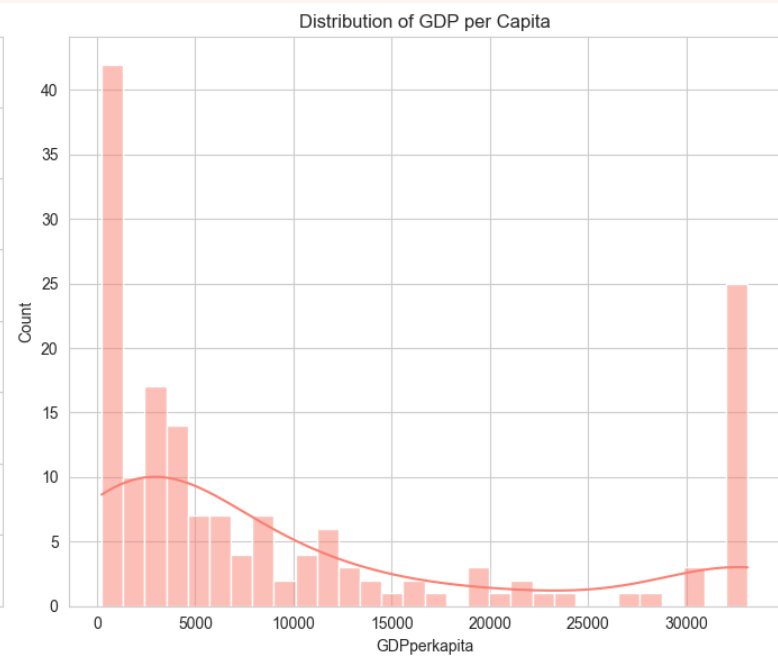
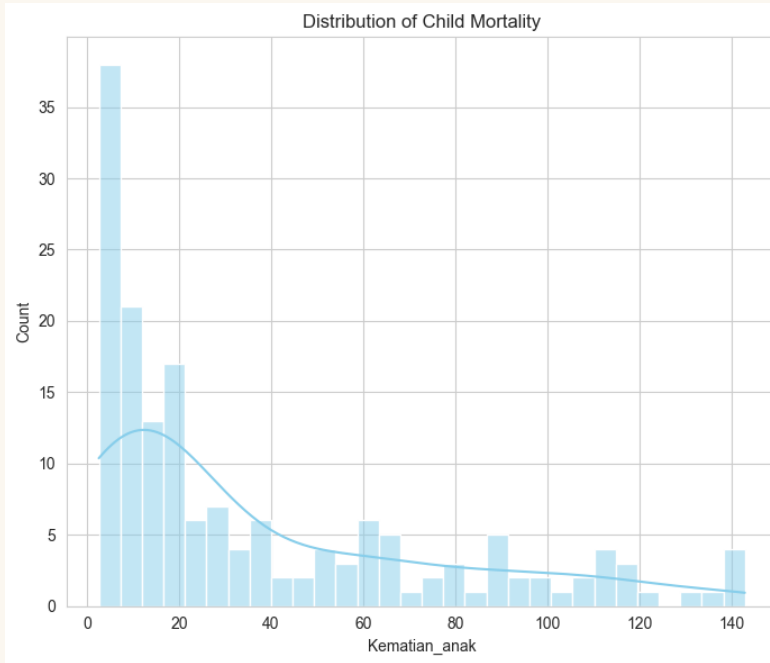
Untuk menangani outliers tersebut, akan dilakukan penggantian data menggunakan upper dan lower bound. Metode ini dilakukan karena hasil analisis sangat sensitif, yaitu untuk menentukan keputusan pemberian bantuan kepada negara yang membutuhkan.

After



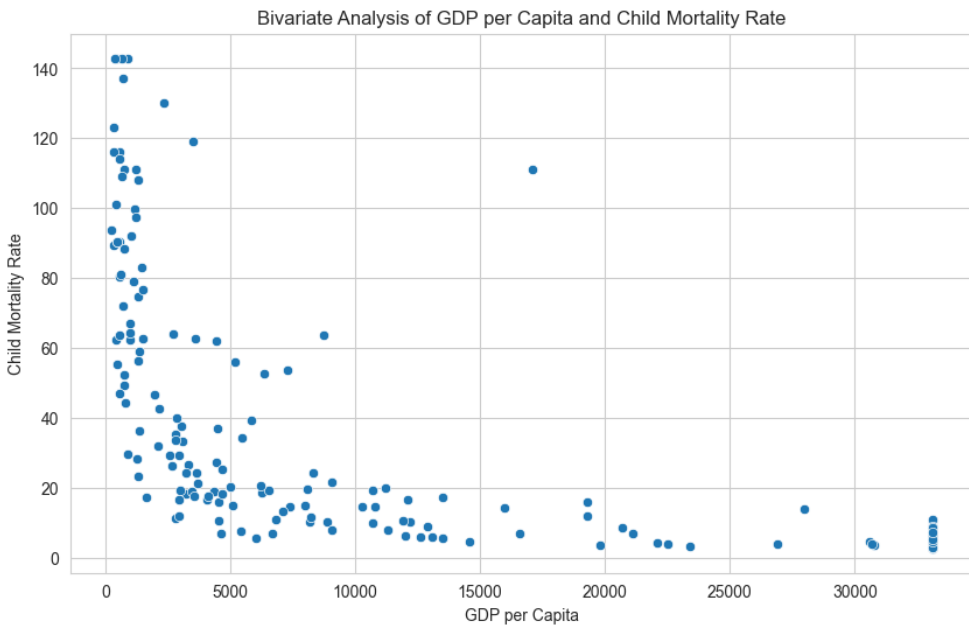
Jika dilakukan penghapusan outliers, dikhawatirkan akan membuat hasil analisis tidak valid.

UNIVARIATE ANALYSIS



- Distribusi Kematian_anak "skewed" ke kanan, yang menunjukkan bahwa meskipun sebagian besar negara memiliki angka kematian anak yang rendah, terdapat banyak negara dengan angka kematian anak yang lebih tinggi.
- Distribusi PDB per kapita juga "skewed" ke kanan, yang menunjukkan bahwa sebagian besar negara mempunyai PDB per kapita yang relatif rendah, dan hanya sedikit negara yang mempunyai PDB per kapita yang jauh lebih tinggi.

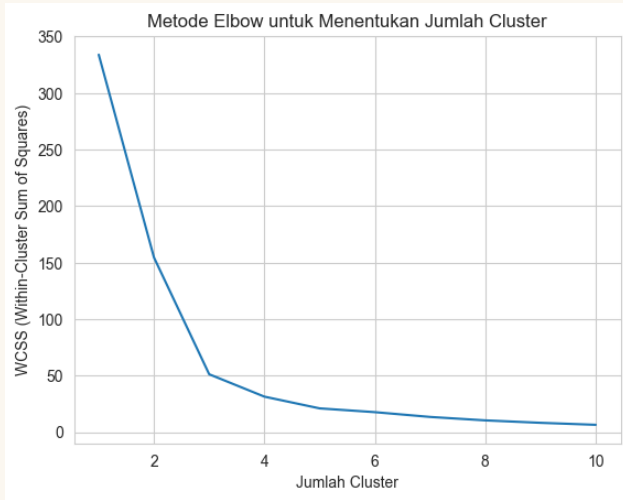
BIVARIATE ANALYSIS



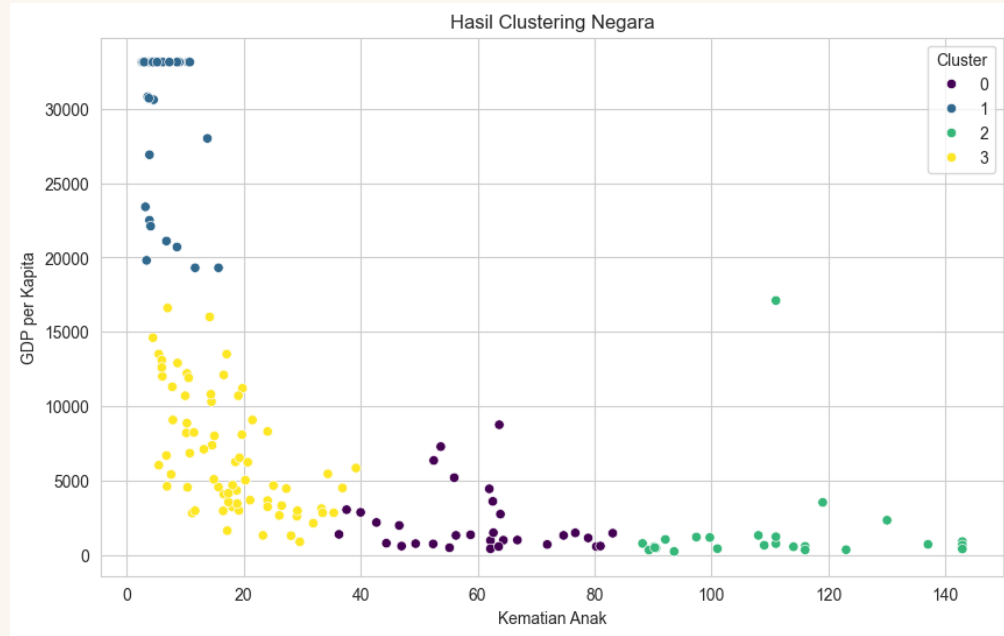
Analisis bivariat PDB per Kapita (PDBperkapita) dan Angka Kematian Anak (Kematian_anak) menunjukkan adanya korelasi negatif antara kedua variabel. Ketika PDB per kapita meningkat, angka kematian anak cenderung menurun.

Hal ini menunjukkan bahwa tingkat pendapatan yang lebih tinggi di suatu negara berkaitan dengan tingkat kesehatan anak yang lebih baik, kemungkinan besar disebabkan oleh peningkatan layanan kesehatan, nutrisi, dan kondisi lingkungan.

CLUSTERING



Membagi data menggunakan 4 cluster



Berdasarkan hasil clustering tersebut, dapat dilihat bahwa negara yang berada pada cluster warna hijau, yaitu cluster 2, memiliki angka kematian anak yang tinggi dan GDP per kapita yang rendah.

RECOMMENDATION

	Kematian_anak	GDPperkapita	Harapan_hidup
Negara			
Sierra Leone	142.875	399	55.0
Central African Republic	142.875	446	47.5
Haiti	142.875	662	32.1
Chad	142.875	897	56.5
Mali	137.000	708	59.5

Saya hanya merekomendasikan 5 negara dikarenakan angka US\$ 10.000.000 merupakan angka yang tidak banyak untuk membantu sebuah negara. Dengan mengerucutkan menjadi 5 penerima bantuan diharapkan bantuan tersebut mencukupi dan dapat dimaksimalkan penggunaannya. Kelima negara tersebut direkomendasikan untuk menjadi prioritas penerima bantuan dengan alasan berikut:

- **Tingkat Kematian Anak:** Tingkat kematian anak yang tinggi menunjukkan adanya masalah serius dalam sistem kesehatan dan sanitasi negara tersebut. Bantuan dapat digunakan untuk meningkatkan infrastruktur kesehatan, memberikan akses ke layanan kesehatan yang lebih baik, dan mengurangi faktor-faktor risiko kematian anak.
 - **GDP per Kapita:** Pendapatan per kapita yang rendah menandakan bahwa penduduk di negara tersebut mungkin menghadapi kesulitan ekonomi dan kehidupan yang sulit. Bantuan dapat digunakan untuk meningkatkan keberlanjutan ekonomi, memberikan pelatihan keterampilan, dan mengembangkan sektor-sektor ekonomi yang berpotensi.
- Saya menambahkan satu kriteria lagi yang dapat membantu untuk menentukan prioritas, yaitu angka harapan hidup.
- **Harapan Hidup:** Harapan hidup yang rendah mencerminkan adanya tantangan kesehatan dan lingkungan hidup yang serius. Bantuan dapat digunakan untuk meningkatkan sistem kesehatan, menyediakan air bersih, sanitasi, dan pendidikan kesehatan kepada masyarakat.