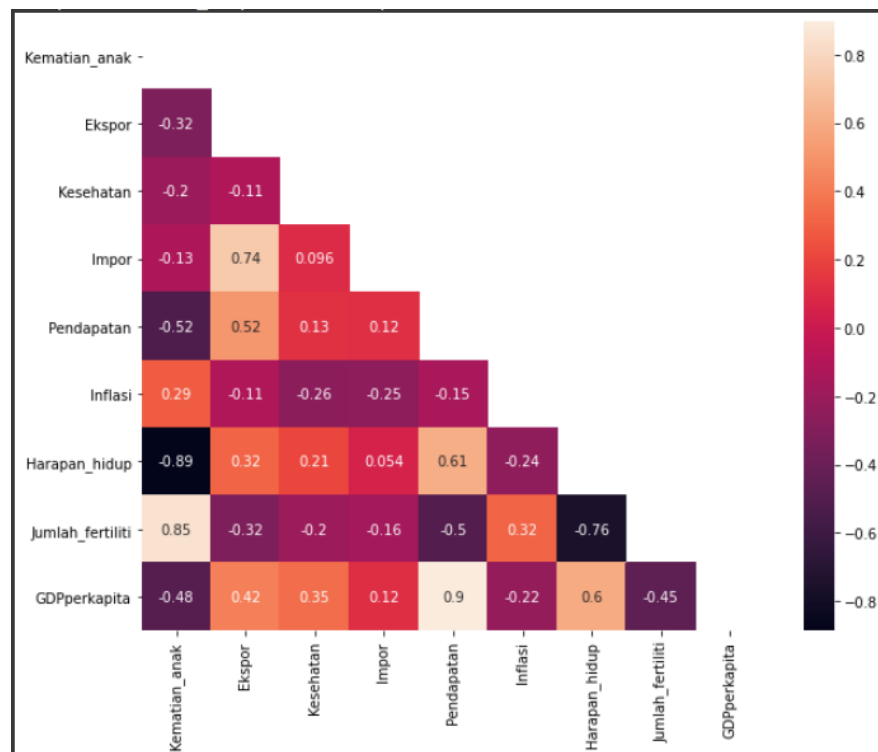


Laporan Final Project

Pada *final project clustering* ini saya akan menyelesaikan permasalahan LSM HELP International untuk memberikan bantuan agar CEO dapat menentukan negara mana yang paling membutuhkan bantuan. Hal ini dilakukan agar CEO dapat mengeluarkan anggaran secara strategis dan efektif.

- *Data Exploration:*

1. Korelasi heat map untuk melihat korelasi atau ketergantungan data satu sama lain.

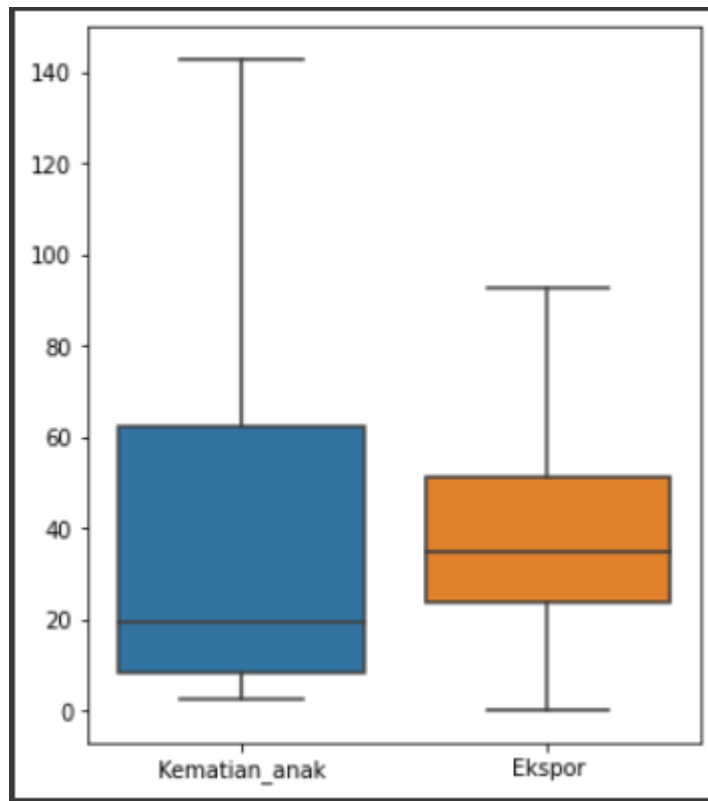


2. Menggunakan *feature selection* untuk mendapatkan *feature* yang optimal dengan mencari data yang korelasinya di bawah -0.2. Hal itu dilakukan karena ingin mendapatkan 2 *feature* namun dengan korelasi kecil. Itu saya lakukan karena menurut sumber yang saya baca, bahwa semakin besar korelasinya, itu memiliki efek yang sama dengan variabel dependen / *target*, maka saya memilih *feature* dengan korelasi di bawah -0.2.

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 167 entries, 0 to 166
Data columns (total 2 columns):
#   Column          Non-Null Count  Dtype
---  -
0   Kematian_anak    167 non-null    float64
1   Ekspor           167 non-null    float64
dtypes: float64(2)
memory usage: 2.7 KB
```

- *Pre-Processing:*

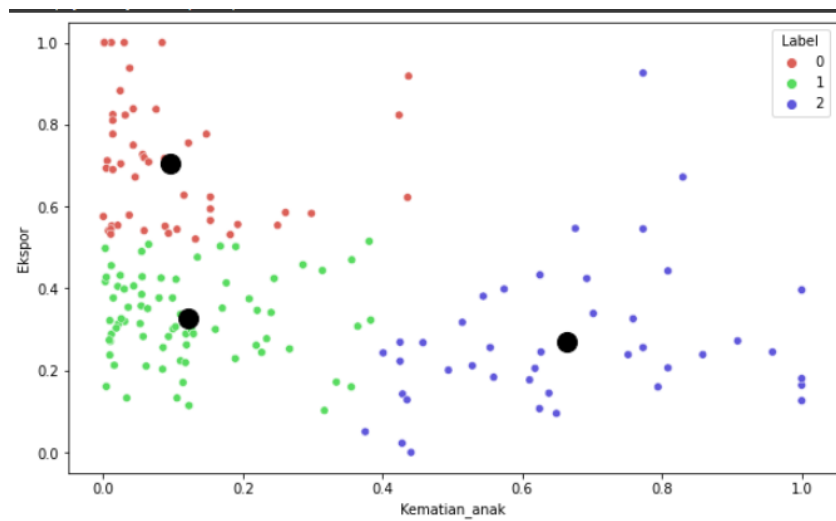
1. Menghilangkan *outliers* pada *feature* yang sudah dipilih agar data yang digunakan dapat semakin akurat dan tidak semakin melenceng dari data-data lainnya.



2. Melakukan *scaling* menggunakan *MinMax*.

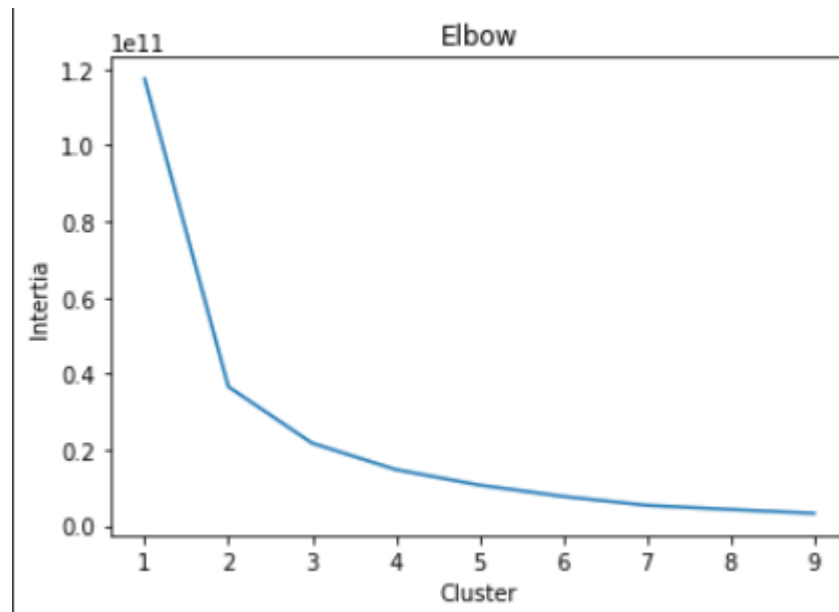
- *Modelling:*

1. Pertama saya melakukan pemodelan k-Means data yang sudah di *scale* di atas dengan jumlah *cluster* 3 (random karena belum dilakukannya evaluasi).

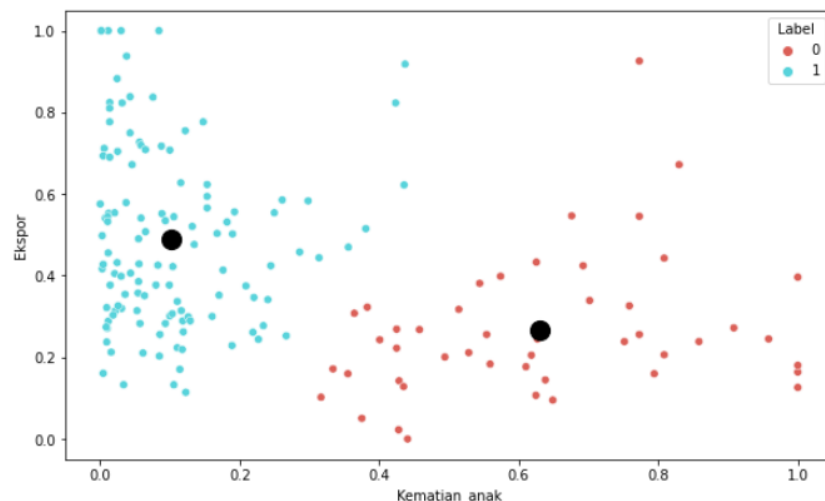


- *Evaluation:*

1. Menggunakan metode *elbow* untuk mendapatkan jumlah *cluster* terbaik dengan rumus *SSE (inertia)*.



Didapatkan bahwa dengan *elbow*, *cluster* yang optimal adalah 2.



2.

- Kesimpulan: didapatkan bahwa *cluster* negara dengan label 0 yang lebih layak mendapatkan bantuan dari LMS HELP. Kenapa label 0? Karena, dilihat dari pengelompokannya bahwa *cluster* 0 ini adalah negara2 yang kematian anaknya tinggi, serta ekspor yang sangat rendah juga. Jadi, kalau dilihat dari ‘kacamata’ sosial-ekonomi dan kesehatan, negara-negara ini yang pantas / layak mendapatkan bantuan. Berikut list-list dari negara yang membutuhkan bantuan:

```
array(['Afghanistan', 'Angola', 'Bangladesh', 'Benin', 'Burkina Faso',
      'Burundi', 'Cameroon', 'Central African Republic', 'Chad',
      'Comoros', 'Congo, Dem. Rep.', 'Cote d'Ivoire',
      'Equatorial Guinea', 'Eritrea', 'Gambia', 'Ghana', 'Guinea',
      'Guinea-Bissau', 'Haiti', 'India', 'Kenya', 'Kiribati', 'Lao',
      'Lesotho', 'Liberia', 'Madagascar', 'Malawi', 'Mali', 'Mauritania',
      'Mozambique', 'Myanmar', 'Nepal', 'Niger', 'Nigeria', 'Pakistan',
      'Rwanda', 'Senegal', 'Sierra Leone', 'South Africa', 'Sudan',
      'Tajikistan', 'Tanzania', 'Timor-Leste', 'Togo', 'Uganda', 'Yemen',
      'Zambia'], dtype=object)
```