

0 0 0

KOMPARASI EFFEKTIVITAS PENANGANAN BANJIR PEMPROV DKI JAKARTA

MAHABHARATA

ANGGOTA KELOMPOK

1. Anthony Edbert Feriyanto
2. Guntara Hambali
3. Rafka Athallah Putra Hasyim

SUMBER



[https://data.jakarta.go.id/organization
/badan-penanggulangan-bencana-
daerah](https://data.jakarta.go.id/organization/badan-penanggulangan-bencana-daerah)

TABLE OF CONTENT

Latar Belakang

Pengenalan Masalah

Hipotesis

Metodologi

Metode

Variabel

Analisis

Kesimpulan

Rekomendasi

LATAR BELAKANG

Permasalahan banjir di Jakarta merupakan topik yang selalu dibicarakan oleh seluruh masyarakat Indonesia. Hal ini disebabkan oleh salah satu faktor yaitu Kota Jakarta sebagai Ibu Kota Negara Indonesia. Dengan adanya bencana banjir di ibu kota, timbulah stigma negatif bahwa banjir merupakan penyakit yang akan selalu melekat pada nama Jakarta. Hal ini menimbulkan citra buruk terhadap nama Kota Jakarta.



LATAR BELAKANG



Setiap tahunnya, Pemprov DKI Jakarta selalu berusaha untuk menjalankan berbagai macam program demi menanggulangi banjir. Pada tahun 2015-2017, Pemprov DKI Jakarta berusaha untuk mengimplementasikan program yang sudah direncanakan seperti manajemen daerah aliran sungai, membuat bendungan dan waduk, kanalisasi dan saluran pengalihan, dan masih banyak yang lainnya. Tak juga lupa dengan tahun 2018-2020, Pemprov DKI yang melaksanakan program-program lain seperti drainase dan perbaikan saluran, retensi air, serta penataan ruang terbuka hijau.

LATAR BELAKANG

Dengan banyaknya metode yang sudah dilakukan oleh Pemprov DKI Jakarta, kami ingin melakukan komparasi data mengenai efektivitas penanganan banjir sehingga di masa yang akan datang, Pemprov DKI Jakarta dapat melanjutkan program dengan tingkat keberhasilan yang tinggi dan mengganti metode yang kurang berdampak pada penanganan banjir yang sudah meradang ini.



PENGENALAN MASALAH

Berdasarkan tujuan utama kelompok kami untuk melakukan komparasi efektivitas penanganan banjir, berikut permasalahan yang kami hadapi untuk membuat model prediksi :

PENGENALAN MASALAH

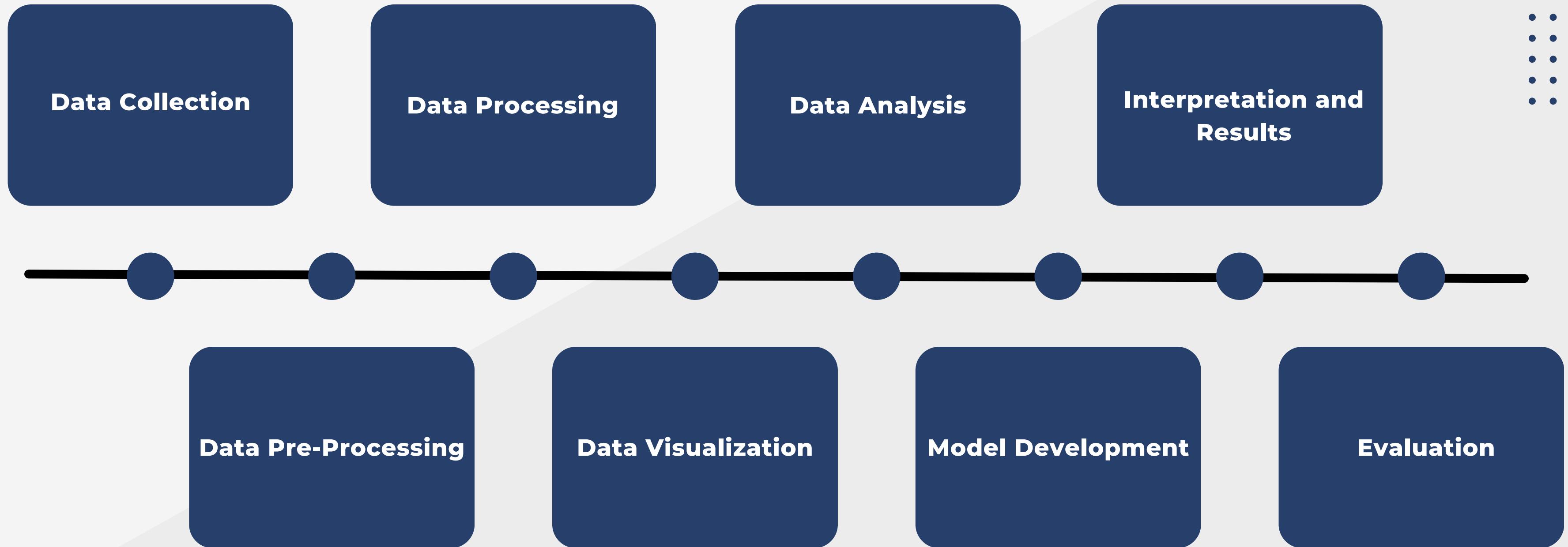
1. Periode tahun manakah yang menjalankan programnya secara lebih efektif?
2. Apakah banjir yang terjadi di Jakarta berkaitan dengan kalender cuaca di Indonesia?
3. Bagaimana cara untuk menjelaskan data yang sudah dianalisis sehingga mudah dipahami oleh semua kalangan masyarakat?
4. Bagaimana cara menyajikan data menjadi simpel sehingga dapat membantu masyarakat umum untuk menjawab pertanyaan mereka yang belum terjawab selama ini?
5. Hal apa yang berkaitan erat dengan banjir yang sudah meradang di Ibu Kota DKI Jakarta?

HIPOTESIS

Metode K-Means dapat membantu para awam untuk memahami data mengenai efektivitas Pemprov DKI Jakarta dalam menangani banjir

Program yang dijalankan oleh Pemprov DKI Jakarta untuk menangani banjir dari rentang tahun 2015 - 2017 lebih efektif dibandingkan program pada tahun 2018-2020.

METODOLOGI



METODE

Data Collection



Data collection adalah prosedur mengumpulkan, mengukur, dan menganalisis berbagai tipe informasi menggunakan teknik yang berstandar.

METODE

Data Pre-
Processing



Data pre-processing adalah proses mengubah data mentah ke dalam bentuk yang lebih mudah dipahami.

METODE

Data Processing



Data processing adalah proses pengumpulan data mentah dan mentransformasi data tersebut menjadi informasi.

METODE

Data Visualization



Data visualization adalah proses menggunakan elemen visual seperti diagram, grafik, atau peta untuk merepresentasikan data.

METODE

Data Analysis



Data analysis adalah proses pengolahan data yang mencakup memeriksa data, membersihkan data, mengubah data, dan membuat model yang sesuai sehingga hasil dari proses tersebut dapat terdefinisi dan menemukan informasinya sebagai dasar dari pengambilan keputusan suatu masalah

METODE

Model Development



Model development adalah pemodelan data pada hasil eksplorasi pola yang telah dilakukan pada tahap-tahap sebelumnya.

METODE

Interpretation and Results



Pada tahap ini, semua komponen yang telah dikembangkan sebelumnya sudah mencapai puncak penggunaannya dan dapat ditemukan hasil-hasil yang sesuai dengan masalah yang ingin dipecahkan

METODE

Evaluation



Evaluation adalah proses menggunakan metrik evaluasi yang berbeda-beda untuk memahami kinerja model *machine learning* serta kekuatan dan kelemahan dari model yang telah dibuat.

VARIABEL



Bulan



Kecamatan



Kelurahan



Lama



RW



KK

Terdampak



Jiwa

Terdampak



Max

Ketinggian Air
Maksimal



Min

Ketinggian Air
Minimal



Rata-Rata

Ketinggian Air

ANALISIS AWAL

PRE-PROCESSING

1. Import Library

Library

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
import re
from sklearn.cluster import KMeans
import random
from sklearn.metrics import silhouette_score
import seaborn as sns
```

[8]

Python

PRE-PROCESSING

2. Import Dataset

Dataset

```
#daftar dataset banjir jakarta 2015-2017
dataset1 = [
    "https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/b335b232-d54d-4ae3-b089-0e37c6c7f5f",
    "https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/5c8159de-db10-40e5-937b-f5f66ff5fb7"
]
```

[239] Python

```
#dataset referensi Januari 2015
df_ref1 = pd.read_csv("https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/31e34d71-d568-4f9c-8f55-c")
```

[240] Python

Dataset

```
#import dataset banjir tahun 2019
df_banjir_2019 = pd.read_csv("https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/06748d424f3311503b")
```

[410] Python

```
#import dataset banjir tahun 2020
df_banjir_2020 = pd.read_csv("https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/8ab1b8ccbed6e2555e")
```

[411] Python

▶ ▾

```
#base dataset Januari 2018
df_ref2 = pd.read_csv("https://data.jakarta.go.id/dataset/46a1ef1a-4143-4c6d-b91d-fa20044660b2/resource/d4cdcb64-d992-4741-843c-2")
```

[412] Python

+ Code + Markdown

PRE-PROCESSING

3. Variable Selection

```
#menghapus variabel yang tidak diperlukan  
df1 = df_ref1.drop(["korban_meninggal", "korban_hilang", "korban_luka_berat", "korban_luka_ringan", "jumlah_jiwa_pengungsi_ter
```

Python

```
#menghapus variabel yang tidak diperlukan pada base dataset  
df_ref2 = df_ref2.drop(["korban_meninggal", "korban_hilang", "korban_luka_berat", "korban_luka_ringan", "jumlah_jiwa_pengungsi_
```

Python

PRE-PROCESSING

4. Pre-Processing (1)

```
#isi variabel kk terdampak yang kosong dengan rata-rata
df1.loc[23, "kk_terdampak"] = "0"
df1.loc[30, "kk_terdampak"] = "0"
df1.loc[31, "kk_terdampak"] = "0"
df1["kk_terdampak"] = df1["kk_terdampak"].astype("int64")

kk_terdampak_mean = df1["kk_terdampak"].mean()
df1.loc[23, "kk_terdampak"] = kk_terdampak_mean
df1.loc[30, "kk_terdampak"] = kk_terdampak_mean
df1.loc[31, "kk_terdampak"] = kk_terdampak_mean
df1["kk_terdampak"] = df1["kk_terdampak"].astype("int64")
```

Python

```
#isi variabel jiwa terdampak yang kosong dengan rata-rata
df1.loc[23, "jiwa_terdampak"] = "0"
df1.loc[30, "jiwa_terdampak"] = "0"
df1.loc[31, "jiwa_terdampak"] = "0"
df1.loc[32, "jiwa_terdampak"] = "0"
df1["jiwa_terdampak"] = df1["jiwa_terdampak"].astype("int64")

jiwa_terdampak_mean = df1["jiwa_terdampak"].mean()
df1.loc[23, "jiwa_terdampak"] = jiwa_terdampak_mean
df1.loc[30, "jiwa_terdampak"] = jiwa_terdampak_mean
df1.loc[31, "jiwa_terdampak"] = jiwa_terdampak_mean
df1.loc[32, "jiwa_terdampak"] = jiwa_terdampak_mean
df1["jiwa_terdampak"] = df1["jiwa_terdampak"].astype("int64")
```

PRE-PROCESSING

4. Pre-Processing (2)

```
#mengconvert column yang memiliki value dengan format yang berbeda
df_ref2["ketinggian_air_max"] = df_ref2["rata_rata_ketinggian_air"].apply(lambda st: st[st.find("-")+1:]).astype("int64")
df_ref2["ketinggian_air_min"] = df_ref2["rata_rata_ketinggian_air"].apply(lambda st: st[:st.find("-")])
df_ref2.loc[6, "ketinggian_air_min"] = "0"
df_ref2.loc[7, "ketinggian_air_min"] = "0"
df_ref2.loc[8, "ketinggian_air_min"] = "0"

df_ref2["ketinggian_air_min"] = df_ref2["ketinggian_air_min"].astype("int64")

df_ref2["ketinggian_air_avg"] = ((df_ref2["ketinggian_air_min"]+df_ref2["ketinggian_air_max"]))/2
```

414]

Python

```
#mengconvert column yang memiliki value dengan format yang berbeda
df_ref2.loc[0, "lama_genangan"] = "2"
df_ref2.loc[1, "lama_genangan"] = "6"
df_ref2["lama_genangan"] = df_ref2["lama_genangan"].astype("int64")
```

415]

Python

PRE-PROCESSING

5. Concatenating Data

```
#gabung ketiga dataset
for i in range(len(dataset1)):
    data = pd.read_csv(dataset1[i])
    df_ref1 = pd.concat([df_ref1, data])

df_ref1.reset_index(drop=True, inplace=True)
df_ref1.loc[12, "rata_rata_ketinggian_air"]= "10 - 80"
df_ref1.loc[17, "rata_rata_ketinggian_air"]= "5 - 90"
```

Python

EXPLORATORY DATA ANALYSIS

Exploratory Data Analysis (EDA) adalah proses analisis awal yang dilakukan pada dataset untuk memahami karakteristik, pola, dan hubungan antara variabel-variabelnya. Tujuannya adalah untuk mendapatkan wawasan yang mendalam tentang data, mengidentifikasi anomali atau kekurangan dalam data, serta merumuskan hipotesis yang dapat diuji lebih lanjut.

EDA 2015-2017

Bulan	Waktu Asli Kejadian
bulan 1	Januari 2015
bulan 2	Februari 2015
bulan 3	Maret 2015
bulan 4	April 2015
bulan 5	Mei 2015
bulan 6	Juni 2015
bulan 7	Juli 2015
bulan 8	Agustus 2015
bulan 9	September 2015
bulan 10	Okttober 2015
bulan 11	November 2015
bulan 12	Desember 2015
bulan 13	Januari 2016
bulan 14	Februari 2016
bulan 15	Maret 2016
bulan 16	April 2016
bulan 17	Mei 2016
bulan 18	Juni 2016

bulan 19	Juli 2016
bulan 20	Agustus 2016
bulan 21	September 2016
Bulan 22	Okttober 2016
bulan 23	November 2016
bulan 24	Desember 2016
bulan 25	Januari 2017
bulan 26	Februari 2017
bulan 27	Maret 2017
bulan 28	April 2017
bulan 29	Mei 2017
bulan 30	Juni 2017
bulan 31	Juli 2017
bulan 32	Agustus 2017
bulan 33	September 2017
bulan 34	Okttober 2017
bulan 35	November 2017
bulan 36	Desember 2017

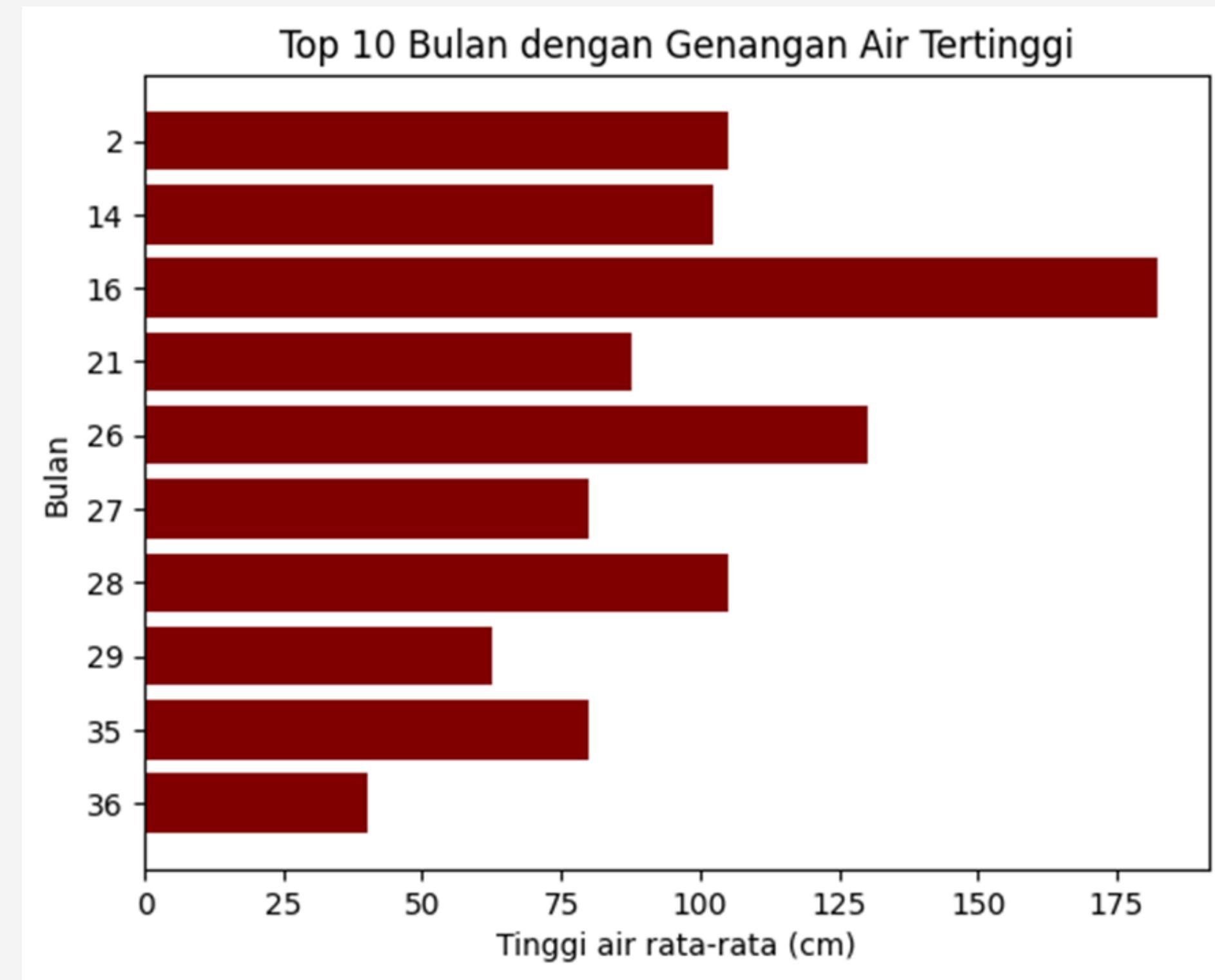
EDA 2015-2017



EDA 2015-2017



EDA 2015-2017



EDA 2015-2017



EDA 2015-2017

- Pada bulan ke-2 (Februari 2015), terjadi fenomena banjir dengan angka kecamatan terdampak tertinggi (35) selama periode ini. Namun, di bulan selanjutnya (Maret 2015), tidak ada kategori yang masuk ke “top 10 ketiga kategori tersebut”. Artinya penanganan banjir sudah menjadi lebih baik.
- Pada bulan ke-14 (Februari 2016), terjadi kembali banjir yang mengerogoti 25 kecamatan, lalu diikuti bulan ke-16 (April 2016), dan bulan ke-21 (September 2016). Berdasarkan data, pada tahun 2016 terjadi penurunan efektivitas penanganan banjir yang pada tahun sebelumnya dinilai sudah berhasil, sekarang menjadi parah kembali. Hal ini bisa disimpulkan dari data yang tercatat pada “top 10 bulan dengan kecamatan terdampak terbanyak”, muncul 3 bulan pada tahun 2016.
- Pada tahun 2017, banjir meradang wilayah DKI Jakarta kembali. Dapat diamati terdapat 6 bulan di tahun 2017 yang tercantum dalam semua grafik. Jika mengaitkan data dengan fenomena yang terjadi pada tahun 2017, dapat disimpulkan bahwa Pilkada DKI Jakarta merupakan salah satu faktor yang menyebabkan banjir kembali meradang di DKI Jakarta.

EDA 2015-2017

- Walaupun banjir di tahun 2016 meningkat, program yang dijalankan oleh Pemprov DKI Jakarta saat itu berhasil menekan angka kecamatan terdampak. Hal tersebut dapat dilihat dari bulan ke-2 (Februari 2015) yang tadinya terdapat lebih dari 35 kecamatan terdampak menjadi 25 kecamatan terdampak pada bulan ke-14 (Februari 2016) dan secara bertahap semakin turun angkanya hingga naik kembali pada bulan ke-26 (Februari 2017).
- Dari data, terdapat pola yang menarik, yaitu angka kecamatan terdampak yang tinggi terletak pada bulan februari di tiap tahunnya. Setelah mengkaji lebih lanjut, musim penghujan biasa terjadi dari rentang bulan Oktober-Maret.
- Kesimpulan dari ketiga grafik tersebut adalah tren tingkat keparahan banjir tiap tahun pada 2015-2017 cenderung meningkat.
- Selain itu, bisa kita lihat, bulan-bulan yang masuk ke top 10 adalah sama persis di tiap variabel (kecamatan, lama genangan, tinggi rata-rata). Oleh karena itu, diduga terdapat korelasi yang cukup kuat antarvariabel ini. Di samping itu, dapat diduga juga bahwa dapat dilakukan clustering atau pengelompokan data.

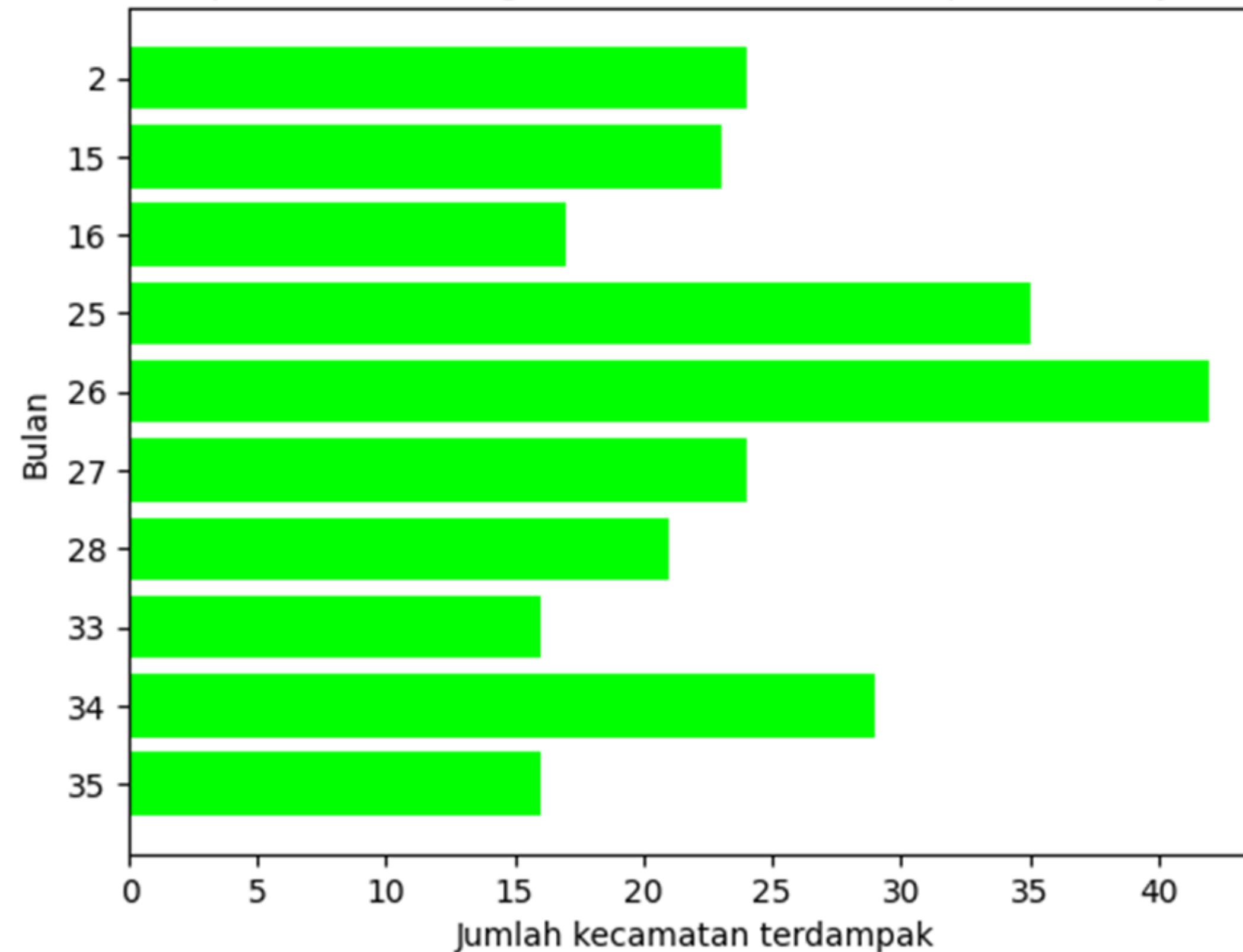
EDA 2018-2020

Bulan	Waktu Asli Kejadian
bulan 1	Januari 2018
bulan 2	Februari 2018
bulan 3	Maret 2018
bulan 4	April 2018
bulan 5	Mei 2018
bulan 6	Juni 2018
bulan 7	Juli 2018
bulan 8	Agustus 2018
bulan 9	September 2018
bulan 10	Oktober 2018
bulan 11	November 2018
bulan 12	Desember 2018
bulan 13	Januari 2019
bulan 14	Februari 2019
bulan 15	Maret 2019
bulan 16	April 2019
bulan 17	Mei 2019
bulan 18	Juni 2019

bulan 19	Juli 2019
bulan 20	Agustus 2019
bulan 21	September 2019
Bulan 22	Oktober 2019
bulan 23	November 2019
bulan 24	Desember 2019
bulan 25	Januari 2020
bulan 26	Februari 2020
bulan 27	Maret 2020
bulan 28	April 2020
bulan 29	Mei 2020
bulan 30	Juni 2020
bulan 31	Juli 2020
bulan 32	Agustus 2020
bulan 33	September 2020
bulan 34	Oktober 2020
bulan 35	November 2020
bulan 36	Desember 2020

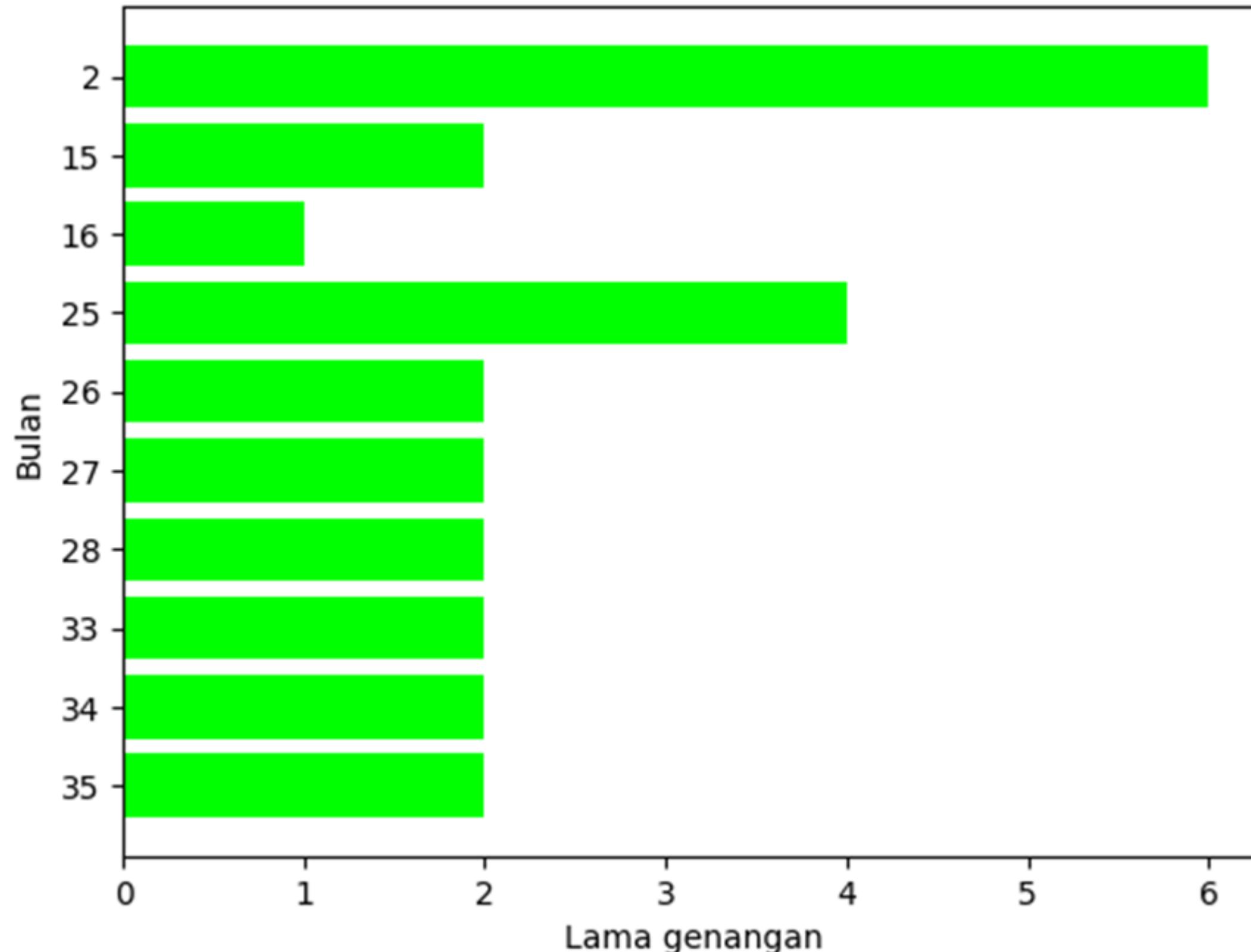
EDA 2018-2020

Top 10 Bulan dengan Kecamatan Terdampak Terbanyak



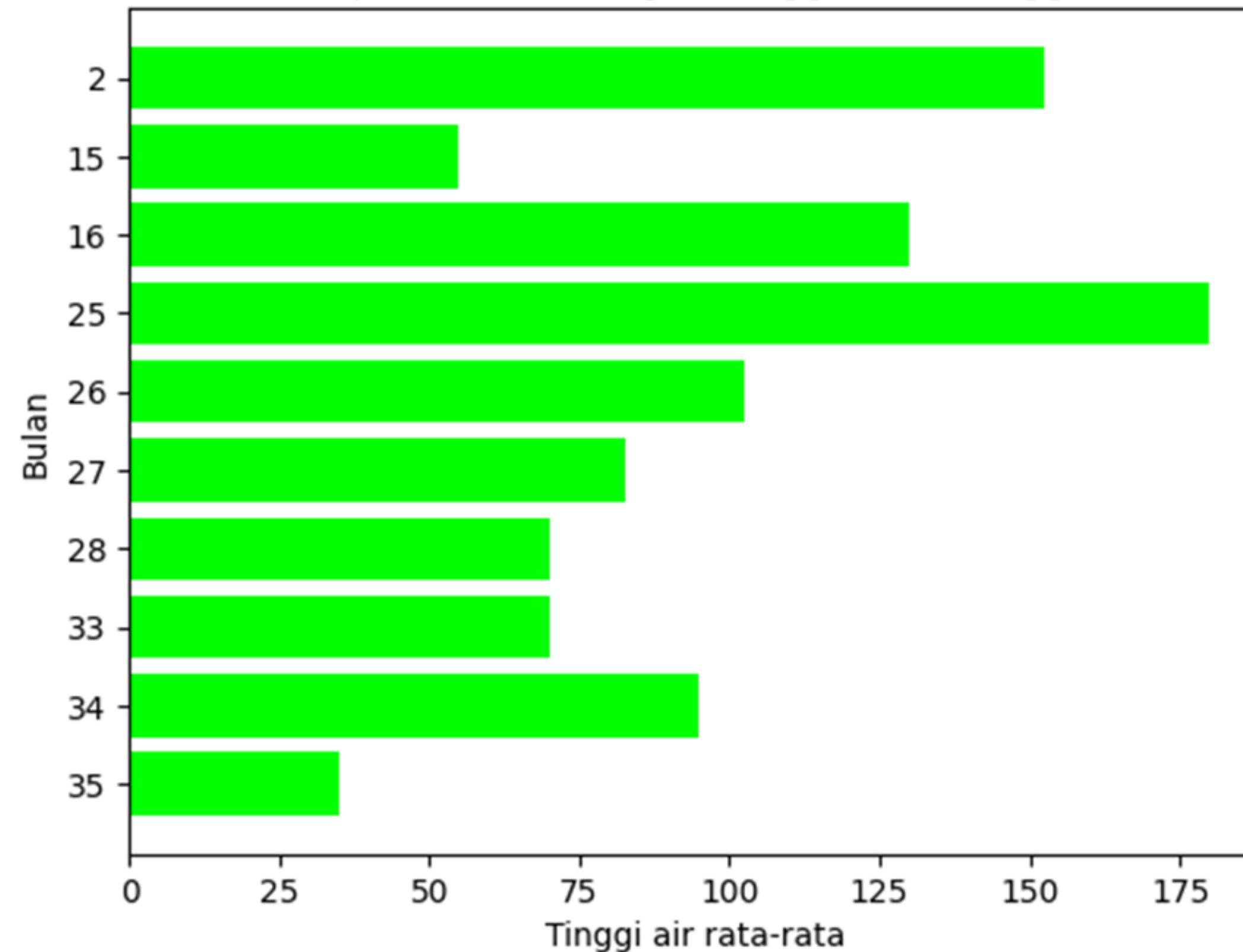
EDA 2018-2020

Top 10 Bulan dengan Genangan Terlama

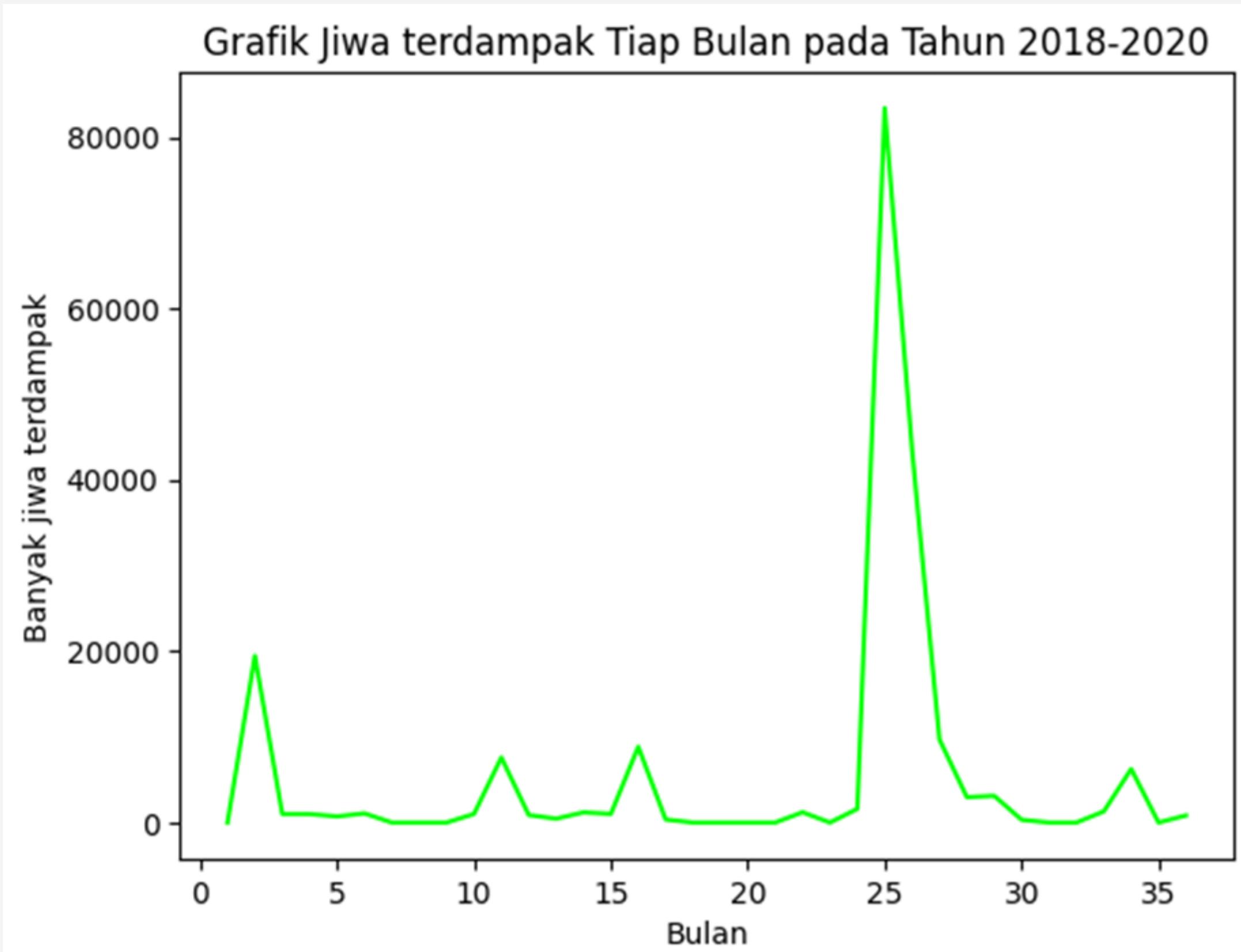


EDA 2018-2020

Top 10 Bulan dengan Tinggi Air Tertinggi



EDA 2018-2020



EDA 2018-2020

- Dari ke-4 grafik mengenai “top 10 bulan dengan kecamatan terdampak terbanyak, genangan terlama, tinggi air tertinggi, jiwa terdampak” bisa diamati bahwa angka kasus banjir di DKI Jakarta cenderung meningkat dari tahun 2018-2020. Hal tersebut dapat simpulan dari banyaknya bulan pada tahun tertentu yang masuk ke kategori “top 10”.
- Pada tahun 2018, terdapat 1 bulan dengan angka yang tergolong tinggi berdasarkan grafik yaitu bulan ke-2 (Februari 2018). Lalu pada tahun 2018, Terdapat 2 bulan yaitu bulan ke 15 (Maret 2019) dan ke-16 (April 2019). Walaupun bulan yang masuk ke kategori “top 10” angka tertinggi lebih banyak dibandingkan tahun sebelumnya, namun dari segi kuantitas, angka kasus yang terjadi lebih sedikit dibandingkan tahun sebelumnya. Namun, di bulan ke-25 (Januari 2020), kasus banjir di Jakarta melonjak naik kembali. Berdasarkan grafik dapat dilihat bahwa angka jiwa terdampak banjir tertinggi terjadi pada bulan ke-25 (Januari 2020).
- Terdapat 7 bulan yang masuk ke kategori “top 10” yang artinya penanganan banjir di tahun 2020 tergolong menurun efektifitasnya dibandingkan tahun sebelumnya.

EXPLORATORY DATA ANALYSIS

EDA melibatkan penggunaan berbagai metode statistik dan teknik visualisasi data untuk menganalisis setiap variabel dalam dataset serta hubungannya dengan variabel lain. Beberapa teknik yang umum digunakan dalam EDA antara lain:

1. STATISTIKA DESKRIPTIF

Apabila dikomparasikan, kejadian banjir di Jakarta pada tahun 2018-2020 relatif sedikit lebih ringan dibandingkan dengan pada tahun 2015-2017. Ini dapat dilihat dari *statistic mean*, kuartil, dan nilai maksimal dari empat variabel. Perbedaan paling signifikan dapat dilihat dari rata-rata jumlah jiwa yang terdampak dimana pada tahun 2018-2020 terdapat 5519 jiwa terdampak dibandingkan dengan pada tahun 2015-2017 yang terdapat 17392 jiwa terdampak.

2015-2017

	bulan	kecamatan	lama_genangan	jiwa_terdampak	ketinggian_air_avg
count	36.000000	36.000000	36.000000	36.000000	36.000000
mean	18.500000	10.944444	1.944444	17392.611111	67.013889
std	10.535654	9.089851	1.510808	39513.384832	46.100897
min	1.000000	0.000000	0.000000	0.000000	0.000000
25%	9.750000	4.000000	1.000000	2120.250000	43.750000
50%	18.500000	11.000000	2.000000	7205.000000	60.000000
75%	27.250000	15.250000	2.000000	16054.000000	100.625000
max	36.000000	38.000000	7.000000	231566.000000	182.500000

2018-2020

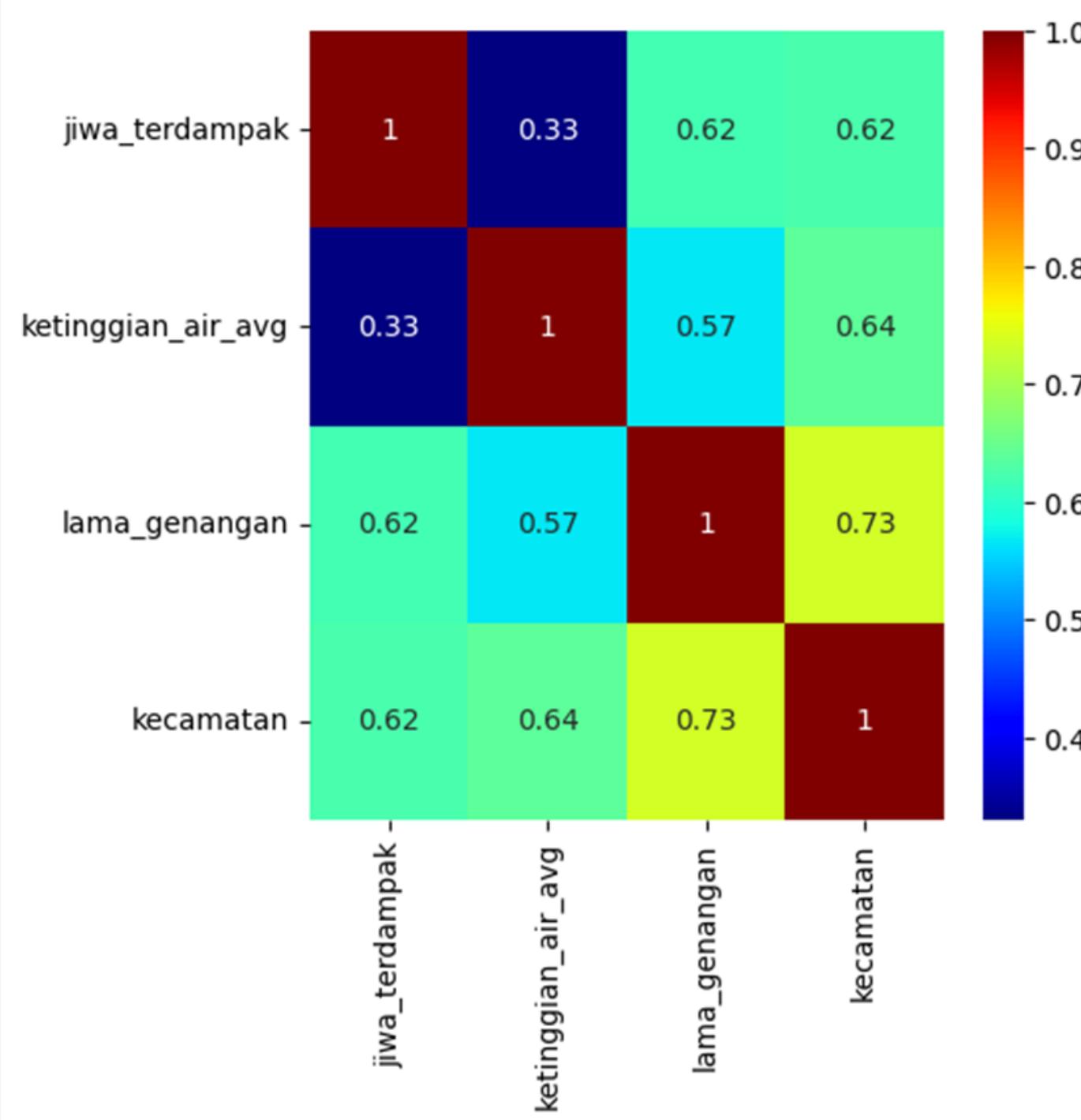
	bulan	kecamatan	lama_genangan	jiwa_terdampak	ketinggian_air_avg
count	36.000000	36.000000	36.000000	36.000000	36.000000
mean	18.500000	11.000000	1.277778	5519.416667	54.277778
std	10.535654	10.53701	1.185896	15519.229218	43.408269
min	1.000000	0.000000	0.000000	0.000000	0.000000
25%	9.750000	3.500000	1.000000	0.000000	31.875000
50%	18.500000	8.500000	1.000000	932.500000	53.750000
75%	27.250000	16.000000	2.000000	1942.250000	70.625000
max	36.000000	42.000000	6.000000	83406.000000	180.000000

2. KORELASI

Korelasi adalah analisis hubungan antara variabel menggunakan yang membantu dalam mengidentifikasi hubungan positif, negatif, atau tidak ada hubungan antara variabel-variabel. Kami menggunakan *Heat Map* untuk menggambarkan korelasi antara 2 variabel. Korelasi dapat dikatakan kuat jika memiliki nilai lebih dari 0,6. Nilai tertinggi dari suatu korelasi antar variable adalah 1 atau -1. Semakin dekat dengan 1 atau -1, maka kedua variabel tersebut saling berkaitan.

HEAT MAP

2015-2017

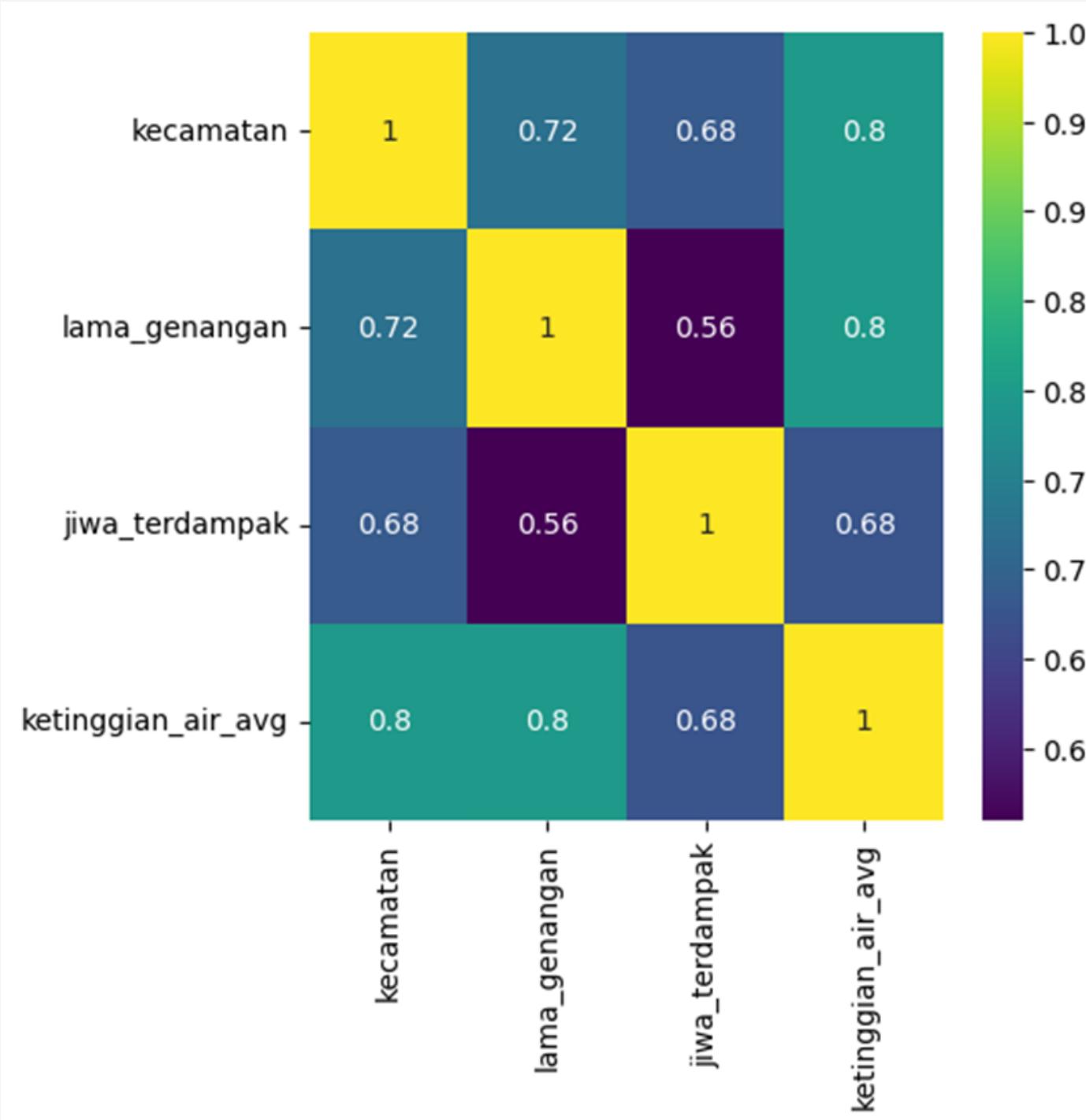


Variabel yang memiliki korelasi tertinggi pada data 2015-2017 adalah :

- Kecamatan dan lama genangan -> 0.73 -> good correlation
- Kecamatan dan ketinggian air avg -> 0.64 -> good correlation
- Lama genangan dan jiwa terdampak -> 0.62 -> good correlation

HEAT MAP

2018-2020



Variabel yang memiliki korelasi tertinggi pada data 2018-2020 adalah :

- Ketinggian air avg dan lama genangan -> 0.8 -> good correlation
- Kecamatan dan ketinggian air avg -> 0.8 -> good correlation
- Kecamatan dan lama genangan -> 0.72 -> good correlation

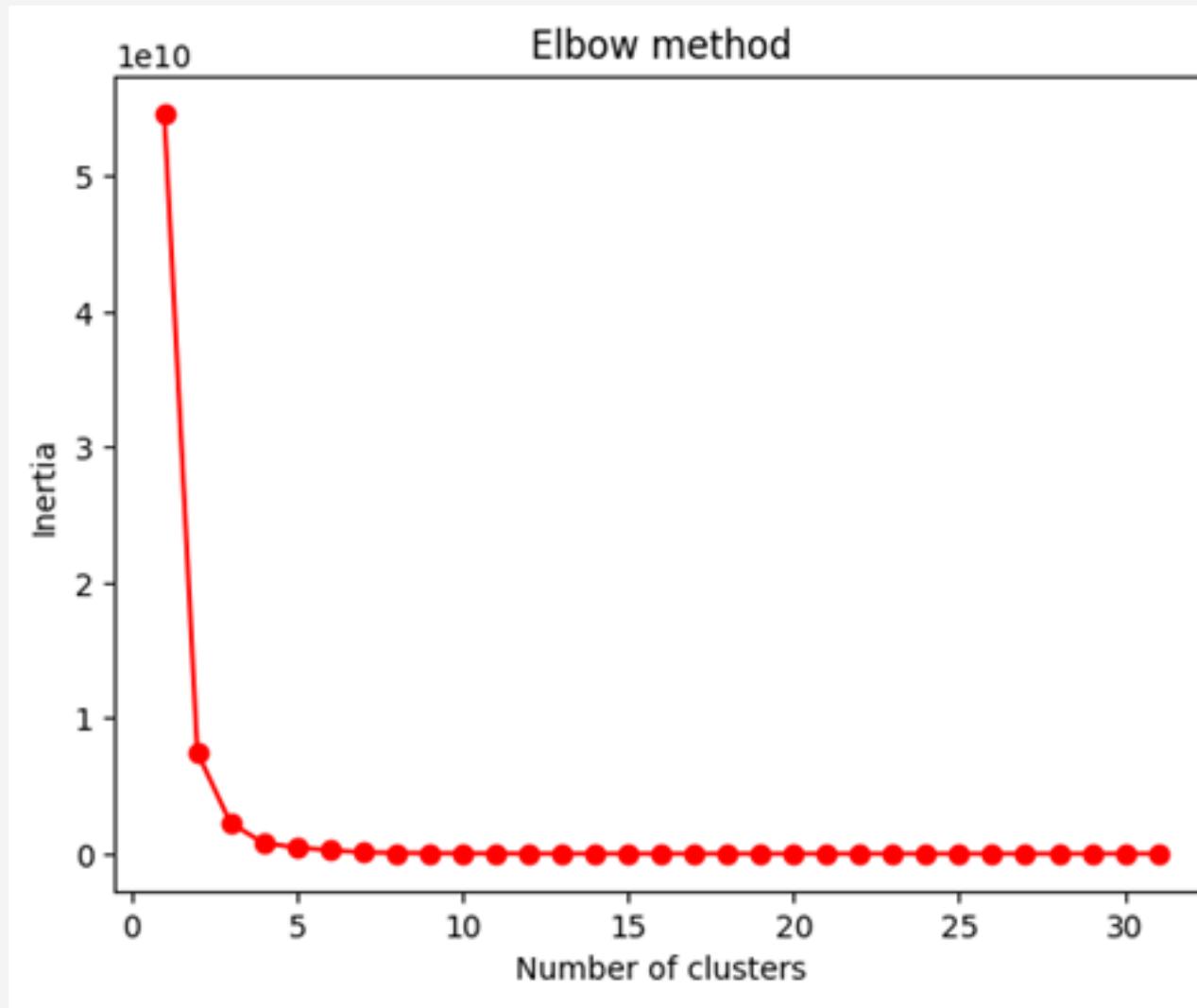
ANALISIS MENDALAM

PENGANTAR ANALISIS

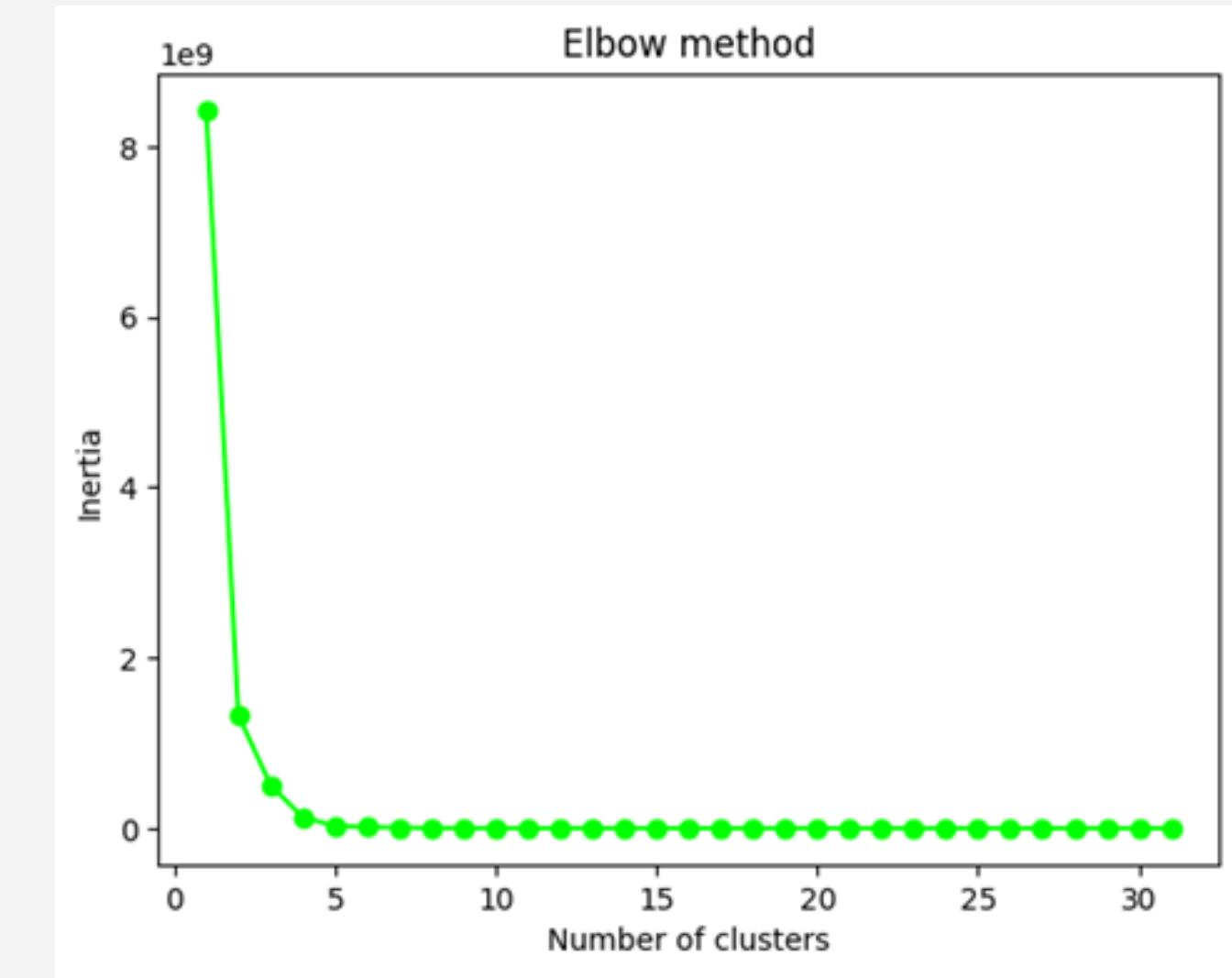
Dari analisis awal yang dilakukan (korelasi dan eksplorasi spasial), ditemukan bahwa terdapat korelasi antarvariabel dalam dataset ini. Namun, apabila digunakan model regresi, korelasi yang ada kurang bagus dan memiliki *r2 score* yang kurang baik. Akan tetapi, korelasi yang ada cukup bagus untuk dilakukan clustering. Oleh karena itu, kami melakukan metode *K-Means Clustering*.

ELBOW METHOD

2015-2017



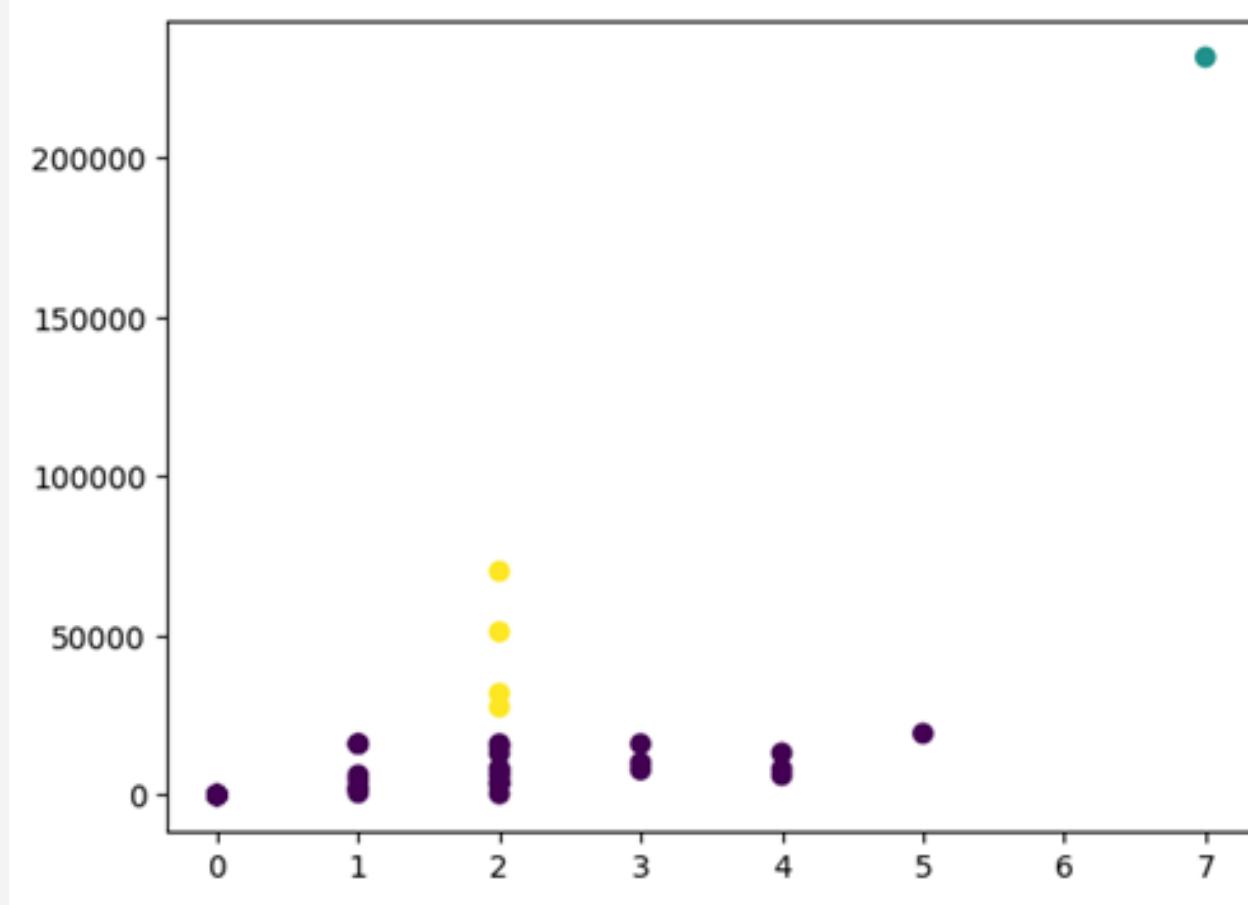
2018-2020



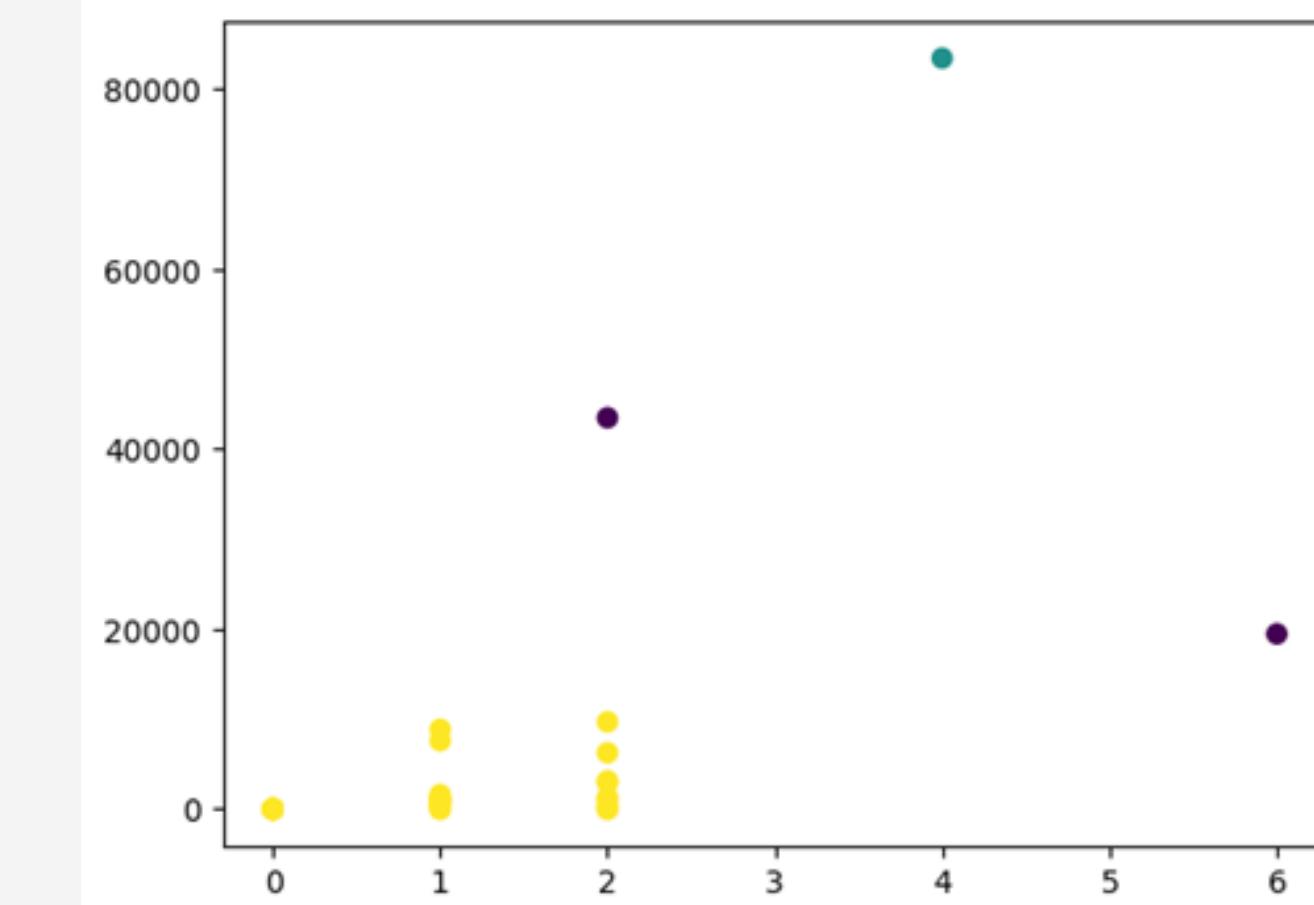
Elbow method digunakan untuk menentukan jumlah cluster yang optimal. Dari kedua plot inersia kedua data di atas, didapatkan jumlah cluster yang optimal adalah 3.

K-MEANS CLUSTERING

2015-2017

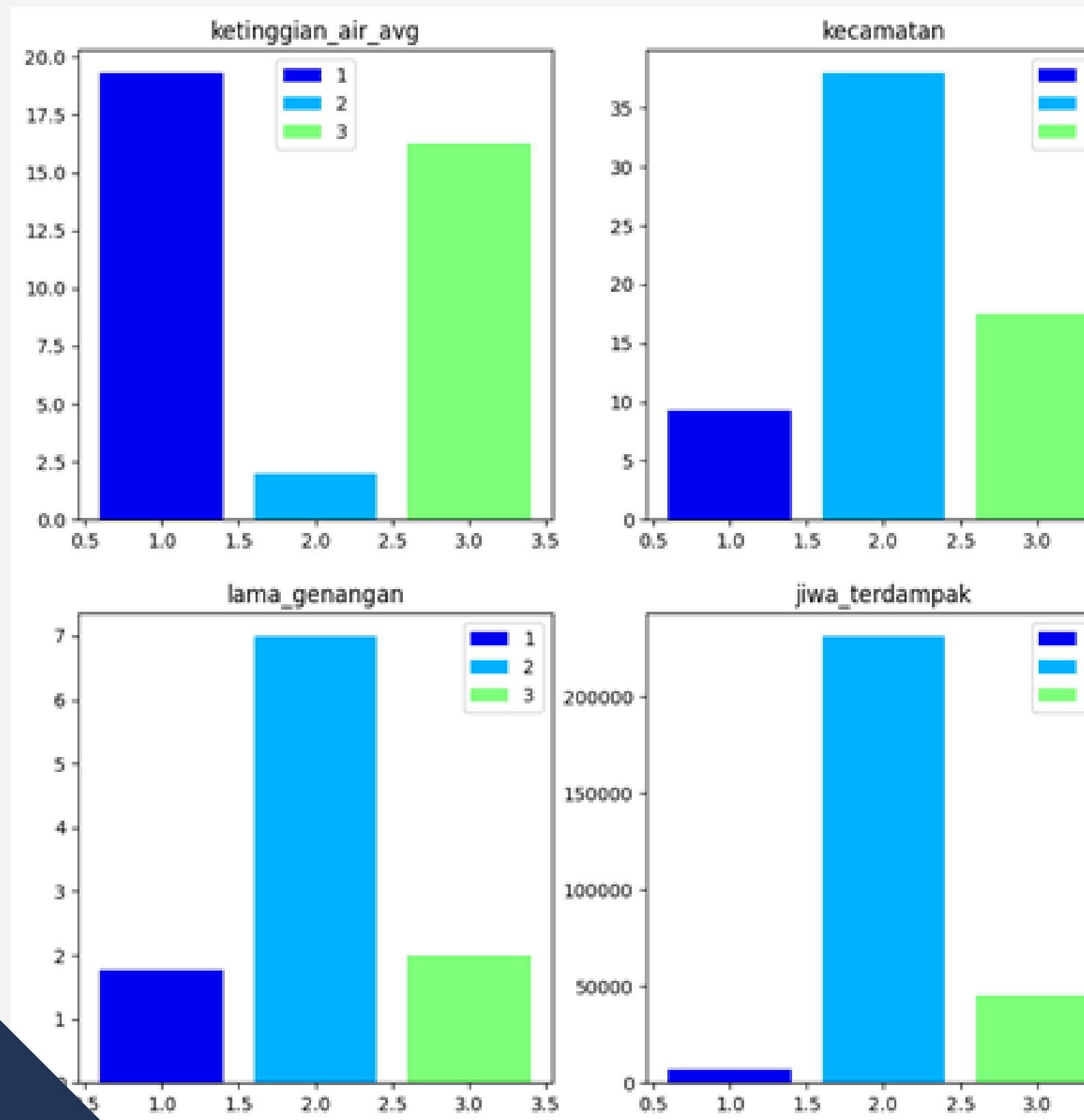


2018-2020



Gambar di atas merupakan plot jiwa terdampak terhadap lama genangan dari dua rentang tahun. Terlihat bahwa hasil clustering cukup bagus (data-data yang berdekatan berada pada satu cluster).

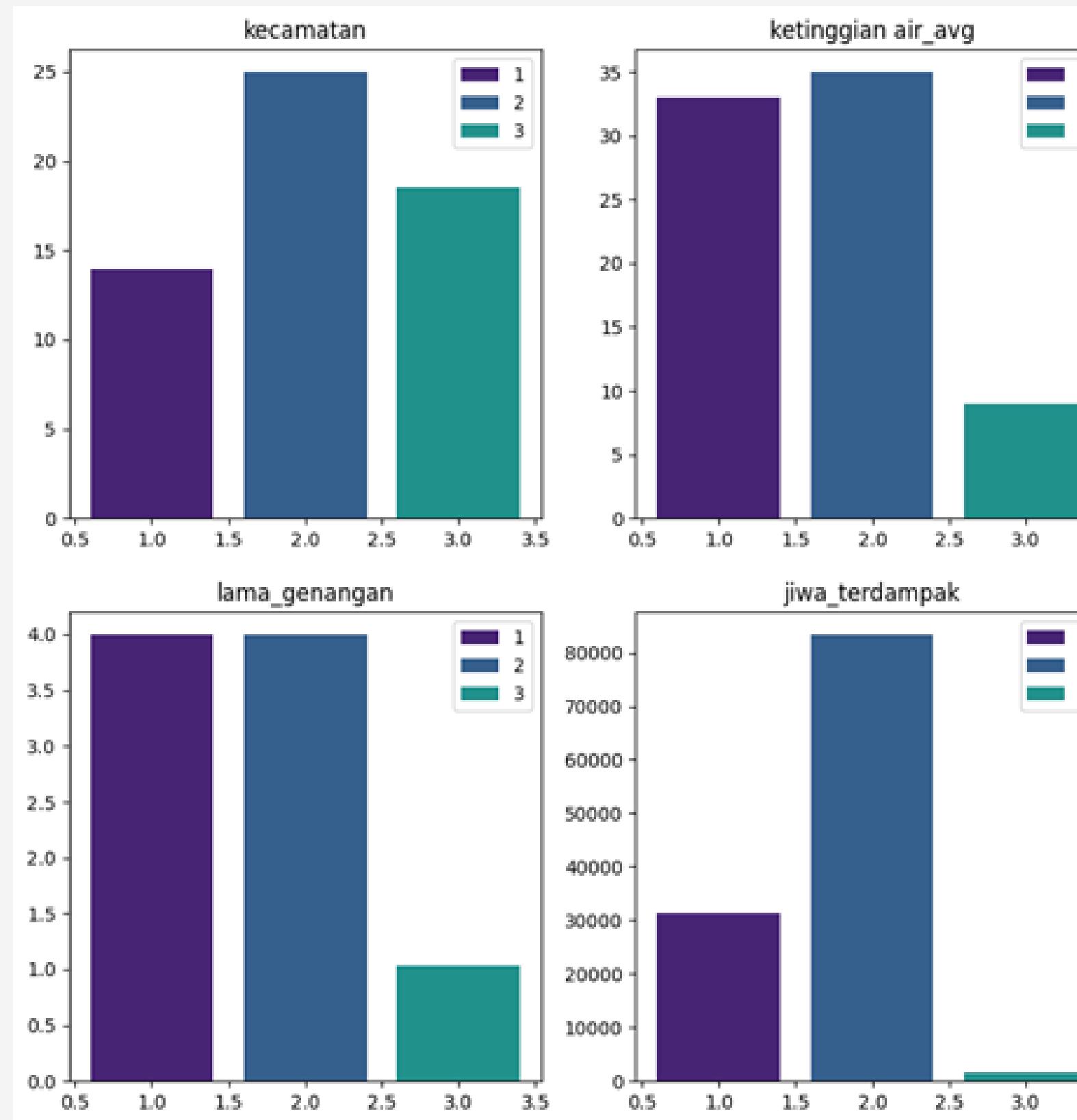
DIAGRAM BATANG CLUSTER 2015-2017



Berdasarkan pengamatan secara visual, dapat kita lihat dari keempat variable, secara rata-rata, intensitas banjir pada tiap-tiap cluster adalah sebagai berikut:

- Berat: Cluster 2
- Sedang: Cluster 3
- Ringan: Cluster 1

DIAGRAM BATANG CLUSTER 2018-2020

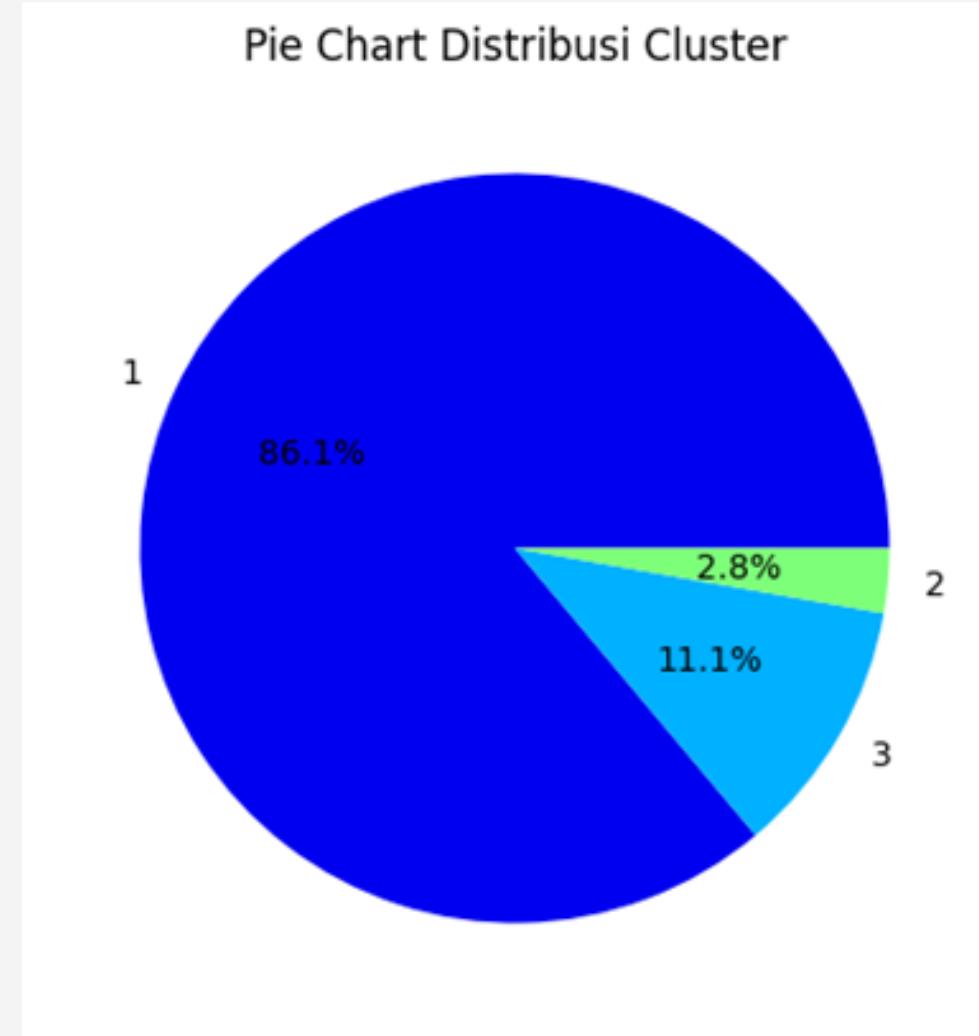


Sama seperti sebelumnya, dapat kita lihat dari keempat variabel, secara rata-rata, intensitas banjir pada tiap-tiap cluster adalah sebagai berikut:

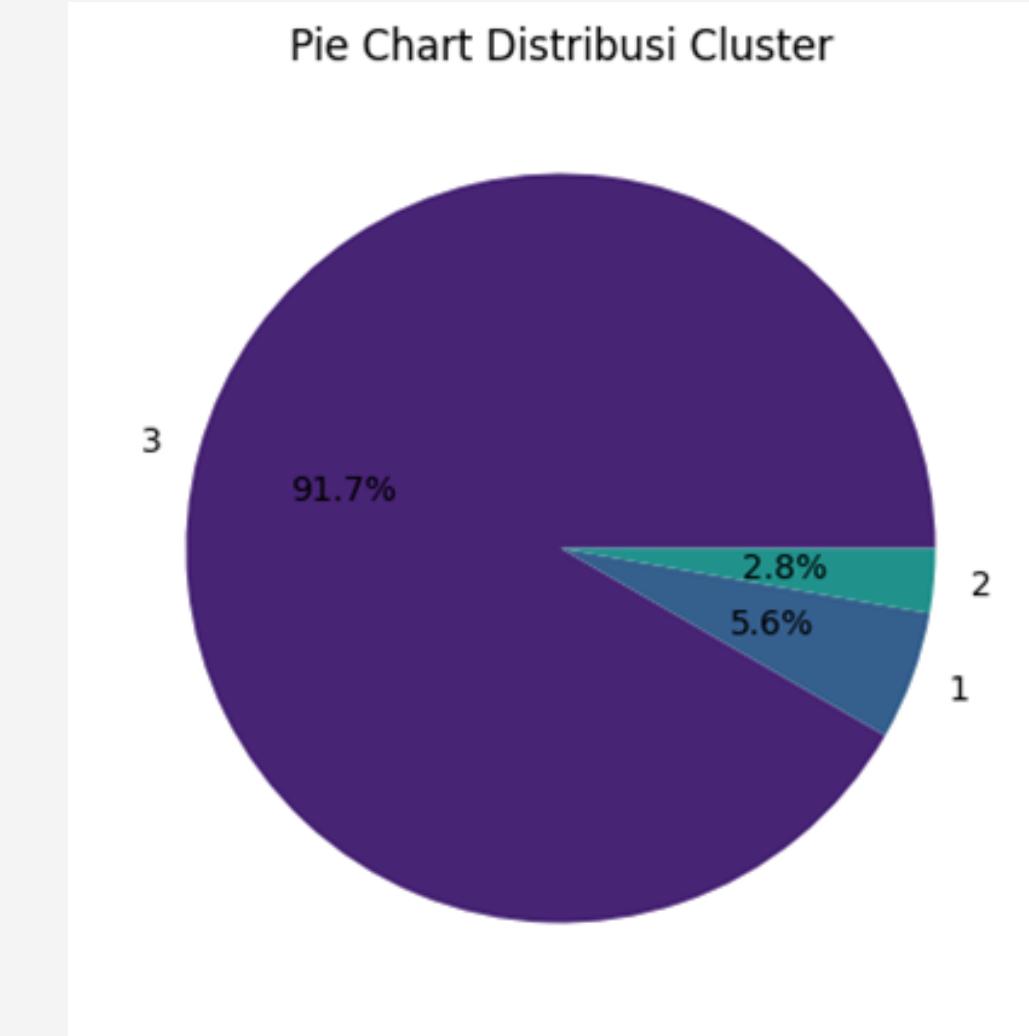
- Berat: Cluster 2
- Sedang: Cluster 1
- Ringan: Cluster 3

DISTRIBUSI CLUSTER

2015-2017



2018-2020



Pada tahun 2015-2017, cluster ringan (cluster 1) mendapatkan distribusi 86.1%. Sedangkan pada tahun 2018-2020, cluster ringan (cluster 3) mendapatkan distribusi 91.7%. Pada tahun 2015-2017, cluster sedang (cluster 3) mendapatkan distribusi 11.1%. Sedangkan pada tahun 2018-2020, cluster sedang (cluster 1) mendapatkan distribusi 5.6%. Untuk cluster berat, kedua rentang tahun mendapatkan distribusi yang sama yaitu 2.8%.

EVALUASI MODEL

2015-2017

```
silhouette_score = silhouette_score(df1_drop, y_kmeans)
print("Silhouette Score:", silhouette_score)
```

```
Silhouette Score: 0.7253941224796568
```

Python

2018-2020

```
silhouette_score = silhouette_score(df2, y_kmeans)
print("Silhouette Score:", silhouette_score)
```

```
[442]
```

```
... Silhouette Score: 0.8433746596286208
```

Python

Silhouette score adalah cara mengukur seberapa mirip suatu titik data dalam suatu *cluster* dibandingkan dengan klaster lain. Nilai koefisien *silhouette* ada di antara titik [-1, 1]. Skor 1 menunjukkan hasil yang terbaik, yang berarti bahwa titik data i sangat padat di dalam *cluster* tempat data itu berada dan jauh dari *cluster* lainnya.

KESIMPULAN

Berdasarkan rangkaian analisis data yang telah dilaksanakan di atas, didapat poin-poin kesimpulan berikut:

- Berdasarkan statistika deskriptif (Mean, median, stdev, dll.) didapatkan bahwa kejadian banjir di Jakarta pada tahun 2018-2020 relatif sedikit lebih ringan daripada tahun 2015-2017.
- Berdasarkan eksplorasi spasial dan visualisasi data, pada rentang tahun 2015-2017 terjadi tren kenaikan jumlah bulan yang mengalami banjir intensitas tinggi yang konstan. Sedangkan pada rentang tahun 2018-2020, kenaikan terjadi dengan melonjak pada tahun 2020
- Berdasarkan model clustering distribusi cluster pada rentang 2015-2017 dengan 2018-2020 cukup seimbang. Namun, rentang tahun yang kedua sedikit lebih unggul karena memiliki proporsi cluster ringan yang sedikit lebih besar.

REKOMENDASI

Pemprov DKI Jakarta perlu mengkaji lebih lanjut mengenai program yang mereka jalankan untuk menekan angka bencana banjir sehingga program yang dijalankan bersifat efektif dan efisien. Selain itu, berdasarkan analisis data, Pemprov DKI Jakarta dapat melanjutkan program yang dilaksanakan pada tahun 2019 dan tahun 2016 karena pada tahun tersebut angka kejadian banjir menurun dibandingkan tahun sebelumnya. Pemprov DKI dapat melanjutkan program seperti Pintu Air, Pembuatan Waduk, dan program lainnya yang dapat membuat irigasi di Kota Jakarta lancar secara skala besar. Tak hanya Pemprov DKI saja yang harus bertanggung jawab terhadap banjir, namun juga seluruh masyarakat Indonesia terutama warga DKI Jakarta dapat berpartisipasi melalui hal sederhana seperti selalu membuang sampah pada tempatnya sehingga saluran perairan di Jakarta tidak tersumbat.

TERIMAKASIH