BAB 4
HASIL PENELITIAN

4.1 RINGKASAN HASIL PENELITIAN

Penelitian ini berfokus pada analisis efektivitas media promosi di Prodi Sistem Informasi Universitas Jenderal Achmad Yani Yogyakarta dengan menggunakan algoritma *K-Means*. Penelitian ini dilakukan dengan mengumpulkan data dari Biro Kerjasama Promosi dan PMB berdasarkan data SICAMA khusus pendaftar Prodi Sistem Informasi yang diperoleh dari tahun 2019 hingga tahun 2022 yaitu jumlah total keseluruhan ada 338 baris. Setelah itu dilakukan *pre-processing data* yaitu data dengan dengan *data cleaning*, dan *feature selection*. Dilanjutkan dengan mencari titik optimal kluster dengan metode elbow yaitu 2 klaster yang terbentuk, selanjutnya membuat *modeling* dengan menerapkan Algoritma *K-Means* mengelompokkan data berdasarkan kemiripan data. *Clustering* dilakukan 2 kali, pertama dengan mengklasterkan wilayah dan kedua dengan mengklasterkan media promosi di tiap *cluster* wilayah. Lalu dilakukan visualisasi data hasil *cluster* menggunakan *scatter plot* dan analisis cluster untuk memudahkan dalam membaca data yang dimiliki. Langkah terakhir dengan membuat *dashboard* yang diperlukan untuk mempresentasikan data yang telah diolah dengan menggunakan streamlit. Berikut adalah hasil penerapannya:

4.2 HASIL PENGUMPULAN DATA

Data diperoleh dari Biro Kerjasama Promosi dan Admisi Unjaya berdasarkan data SICAMA yaitu pendaftar Prodi Sistem Informasi dari tahun 2019 dengan 61 data, 2020 dengan 40 data, 2021 dengan 150 data dan tahun 2022 dengan 87 data. Data yang diproses yaitu penggabungan data dari 2019 hingga 2022 dengan jumlah total keseluruhan ada 338 data.

Tabel 4.1 Data SICAMA

No	Tahun	ID	Nama	Prodi Pil 1	Prodi Pil 2	Provinsi	Asal Info
1	2019	20190114	Dita Yanti Fransisca Hasugian	Sistem Informasi (S-1)	Manajemen (S-1)	Sumatera Utara	Media Sosial
2	2019	20190161	Michael Hosea Nugraha	Sistem Informasi (S-1)	Sistem Informasi (S-1)	Jawa Barat	Keluarga
3	2019	20190191	Oktavila Nauli Sinurat	Sistem Informasi (S-1)	Rekam Medis dan Informasi Kesehatan (D-3)	Kalimantan Barat	Website
4	2019	20190195	Azkiyyatun Nufus Azzahra	Sistem Informasi (S-1)	Informatika (S-1)	Jawa Tengah	Pameran Pendidikan
5	2019	20190209	Avita Dwi Febryanti	Sistem Informasi (S-1)	Psikologi (S-1)	Jawa Timur	Brosur
6	2019	20190237	Ahmad Ilham Nur Febriyanto	Sistem Informasi (S-1)	Teknik Industri (S-1)	Yogyakarta	Keluarga
7	2019	20191512	Muhammad Zulfikar Reyhan	Sistem Informasi (S-1)	Informatika (S-1)	Jakarta	Website
8	2019	20191522	Jihan Fadhil Setyoabdi	Sistem Informasi (S-1)	Informatika (S-1)	Jawa Tengah	Satuan TNI AD
9	2019	20192122	Davit ahmad zikri	Sistem Informasi (S-1)	Informatika (S-1)	Sumatera Barat	Media Sosial
10	2019	20192602	Anodya Larasati	Sistem Informasi (S-1)	Teknik Industri (S-1)	Jambi	Brosur

Tabel 4.1 adalah hasil ekspor data SICAMA khususnya prodi Sistem Informasi.

4.3 HASIL PRE-PROCESSING DATA

Tahapan ini digunakan untuk eksplorasi data yang telah diperoleh, dalam memvisualisasikan data dilakukan tahapan memproses dataset agar terbaca pada sistem dengan menggunakan alat *google collaboratory* dalam pengerjaan.

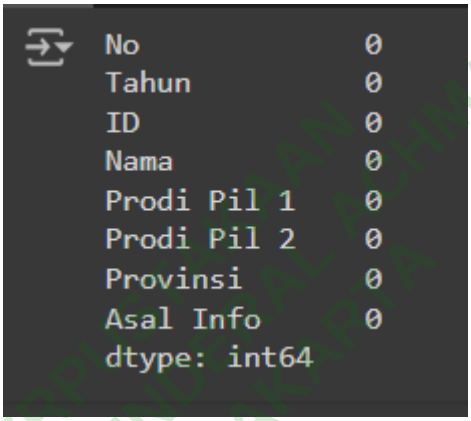
Berikut adalah tahapan yang dilakukan saat *pre-processing data*:

1. Data Cleaning

Data cleaning bertujuan untuk mengecek data, apabila terdapat missing value, setelah itu jika ada maka data akan dihapus atau *drop*.

```
print(data.isnull().sum())
data.dropna(inplace=True)
```

Kode diatas menghasilkan *output* seperti pada gambar 4.1



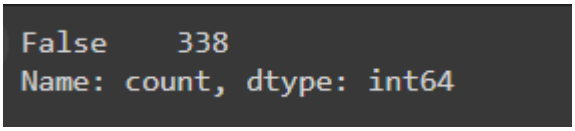
Gambar 4.1 Missing Value

2. Data Duplikat

Data duplikat bertujuan untuk mengecek data, apabila terdapat data yang terduplikasi.

```
# Menampilkan duplikat
duplicates = data.duplicated()
duplicates.value_counts()
```

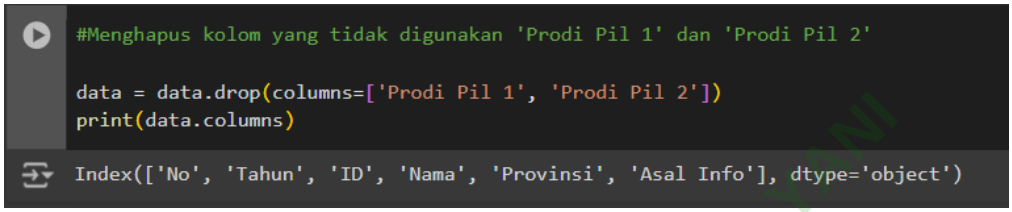
Hasil output dari pengecekan data duplikat yaitu tidak ada nilai yang terduplikat pada data yang di ekspor, terlihat pada gambar 4.2



Gambar 4.2 Data Duplikat

3. Menghapus Kolom yang Tidak Relevan

Tahap ini dilakukan untuk menghapus kolom yang ada pada *dataframe* karena kolom tersebut tidak relevan pada saat pembuatan model dan tidak digunakan dalam penganalisisan data. Kode dan hasil penghapusan kolom terlihat pada gambar 4.3.



Gambar 4.3 Penghapusan Kolom

4. Feature Selection

Pada tahap ini digunakan untuk memilih variabel yang akan digunakan yaitu pendaftar per provinsi berdasarkan jumlah pendaftarnya.

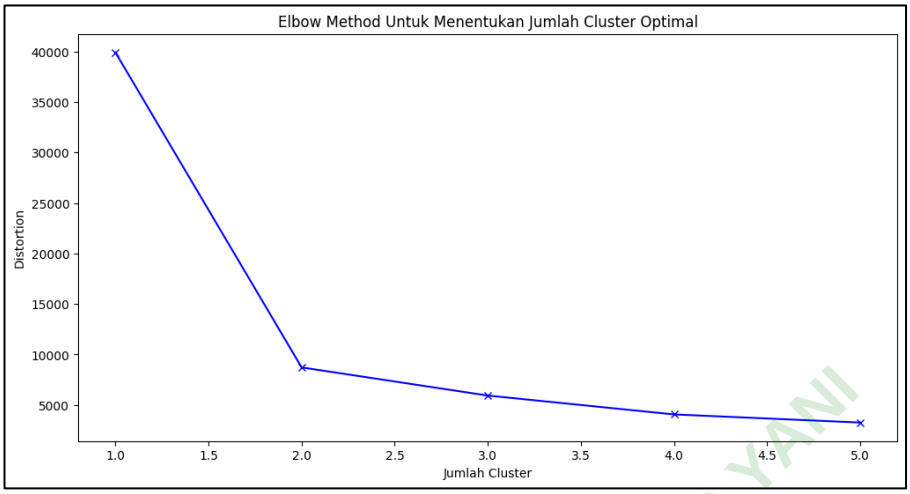
```
pendaftar_per_provinsi =
data['Provinsi'].value_counts().reset_index()
X = pendaftar_per_provinsi[['Jumlah_Pendaftar']]
```

4.4 HASIL MODELING DAN EVALUATION

Tahapan ini digunakan untuk membuat model *clustering K-Means* dengan mengklasterkan wilayah setelah itu dilanjutkan dengan mengklasterkan media promosi berdasarkan wilayah cluster yang terbentuk. Berikut adalah tahapan dalam modeling dengan *algoritma K-Means*:

1. Hasil Metode Elbow

Elbow adalah metode yang digunakan dalam menentukan jumlah *cluster* ideal yang akan digunakan untuk menentukan titik kluster. Jumlah *cluster* optimal yang diambil adalah titik 2 dilihat dari titik siku yang ada pada gambar 4.4



Gambar 4.4 Hasil Metode Elbow Clustering Wilayah

Pada gambar 4.4 menunjukkan bahwa titik *cluster* yang optimal adalah 2 *cluster*.

2. Standarisasi Data

Pada tahap ini dilakukan standarisasi data menggunakan *standar scaler* yang bertujuan untuk menempatkan data ke dalam rentang data yang tidak terlalu jauh. Berikut adalah kode yang digunakan:

```
from sklearn.preprocessing import StandardScaler

scaler = StandardScaler()
X_scaled = scaler.fit_transform(X)
```

3. Hasil *Clustering K-Means Wilayah*

Clustering K-Means dilakukan dengan mengklasterkan wilayah berdasarkan jumlah pendaftar. Pada tahap ini menggunakan *library* dari *sklearncluster* dengan *import KMeans*.

Berikut ini adalah kode yang digunakan saat membuat model:

```
from sklearn.cluster import KMeans
kmeans = KMeans(n_clusters=2, random_state=42)
kmeans.fit(X_scaled)

# Menambahkan label KMeans cluster ke dalam dataframe
pendaftar_per_provinsi['Cluster'] = kmeans.labels_
```

```
# Menggabungkan hasil clustering ke dalam dataframe asli
data = data.merge(pendaftar_per_provinsi[['Provinsi',
'Cluster']], on='Provinsi', how='left')
```

Kode tersebut membuat 2 cluster sesuai dari hasil metode elbow dan random state 42 yang digunakan untuk memastikan hasil yang konsisten setiap kali kode dijalankan. Lalu model akan mengelompokkan data cluster dan menggabungkan hasil *cluster* pada kolom baru yang berisi label *cluster* dari setiap *cluster* yang terbentuk dapat dilihat pada tabel 4.2

Tabel 4.2 Hasil Clustering K-Means berdasarkan Wilayah

No	Tahun	ID	Nama	Provinsi	Asal Info	Cluster
1	2019	20190114	Dita Yanti Fransisca	Sumatera Utara	Media Sosial	1
2	2019	20190161	Michael Hosea N	Jawa Barat	Keluarga	1
3	2019	20190191	Oktavila Nauli Sinurat	Kalimantan Barat	Website	0
4	2019	20190195	Azkiyyatun Nufus Azzahra	Jawa Tengah	Pameran Pendidikan	1
5	2019	20190209	Avita Dwi Febryanti	Jawa Timur	Brosur	1

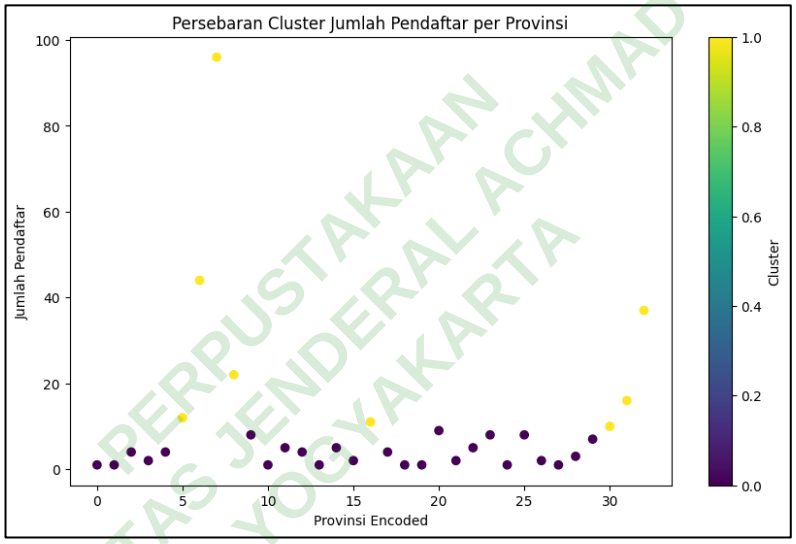
Hasil dari model algoritma KMeans terbentuk 2 *cluster* yaitu 90 data di *cluster* 0 dan *cluster* 1 sebanyak 248 data. Berikut adalah hasil dari masing-masing *cluster*:

a. *Cluster* 0 merupakan wilayah dengan jumlah pendaftar tersedikit (<10 pendaftar) terdiri dari wilayah Nusa Tenggara Timur, Sulawesi Selatan, Riau, Kalimantan Barat, Sumatera Barat, Papua Barat, Kalimantan Tengah, Kepulauan Bangka Belitung, Maluku, Kalimantan Timur, Jakarta, Banten, Sulawesi Utara, Sulawesi Tengah, Papua, Bengkulu, Kepulauan Riau, Aceh, Bali, Kalimantan

Selatan, Nusa Tenggara Barat, Sulawesi Tenggara, Sulawesi Barat, Maluku Utara, dan Kalimantan Utara.

b. *Cluster* 1 merupakan wilayah dengan jumlah pendaftar terbanyak (>= 10 pendaftar) terdiri dari wilayah Jawa Tengah, Jawa Barat, Yogyakarta, Jawa Timur, Jakarta, Sumatera Utara, Jambi, Lampung, dan Sumatera Selatan.

Langkah selanjutnya yaitu membuat visualisasi cluster dengan scatter plot yang digunakan untuk melihat persebaran data. *Cluster* 0 berwarna ungu dan *cluster* 1 berwarna kuning yang terlihat pada gambar 4.5.



Gambar 4.5 Scatter Plot Clustering Berdasarkan Wilayah

Dilanjutkan dengan mengevaluasi model menggunakan *Silhouette Score*. Hasil dari *Silhouette Score* menunjukkan angka 0,83 dimana hasil tersebut mendekati nilai 1 menunjukkan bahwa titik data yang terkelompok sudah baik. Kode dan hasil keluaran dari evaluasi dari *Silhouette Score* dapat dilihat pada gambar 4.6.

```
from sklearn.metrics import silhouette_score
# Menambahkan label cluster ke dataframe
pendaftar_per_provinsi['KMeans_Cluster'] = kmeans.labels_

silhouette_avg = silhouette_score(X_scaled, pendaftar_per_provinsi['KMeans_Cluster'])
print(f'Silhouette Score: {silhouette_avg}')
```

```
Silhouette Score: 0.8323912494766573
```

Gambar 4.6 Silhoutte Score Clustering Berdasarkan Wilayah

Tahapan berikutnya dengan menyimpan hasil dari *clustering* wilayah dengan kode dibawah.

```
# Menyimpan setiap kluster ke file excel
for cluster in df['Cluster_Media'].unique():
    cluster_df = df[df['Cluster_Media'] == cluster]
    cluster_df.to_excel(f'cluswil{cluster}.xlsx,
index=False)
```

Menyimpan model data *cluster* untuk digunakan pada tahap modeling kedua berdasarkan media promosi per *cluster*. Saat menyimpan hasil cluster, dilakukan membuat variabel ”output” yang berisi nama judul, setelah itu men-*convert* data menjadi excel.

4. Hasil *Clustering K-Means* Media Promosi per Klaster Wilayah

Sesuai dengan tujuan penelitian, pemodelan wilayah digunakan untuk mengklasterkan media promosi pada tiap klaster wilayah. Berikut adalah tahapan pengklusteran media promosi pada tiap klaster wilayah:

a. *Clustering* K-Means Media Promosi Klaster 1

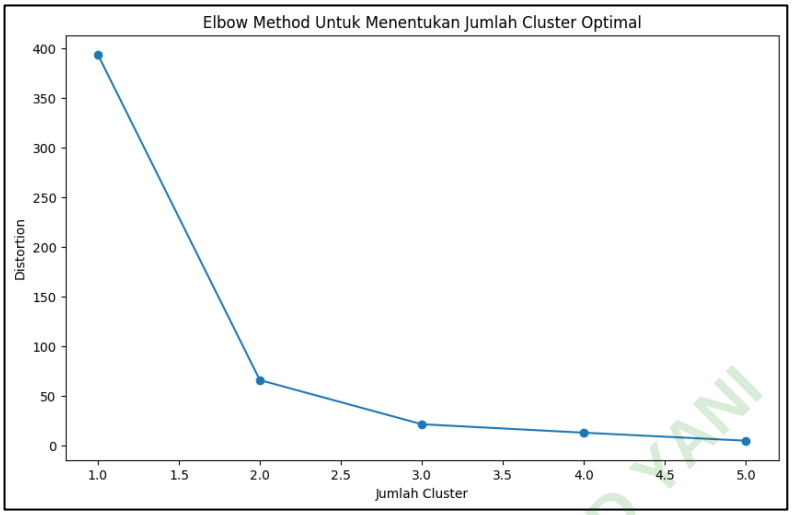
Tahap pertama dilakukan dengan menginisiasi hasil data cluster wilayah 1 untuk dianalisis.

```
w1 = "drive/My Drive/DataSet/NEW/clust1wil.xlsx"
data = pd.read_excel(w1)
```

Selanjutnya dilakukan pemilihan variabel yang akan digunakan untuk *clustering* media promosi, pada tahap ini target variable X adalah Asal_Info, dan dilakukan standarisasi data agar klaster tidak terlalu jauh berkelompok.

```
X = data[['Asal Info']]
scaler = StandardScaler()
X_scaled1 = scaler.fit_transform(X)
```

Tahapan berikutnya yaitu menentukan jumlah *cluster* ideal yang akan digunakan untuk menentukan titik kluster dengan metode Elbow. Jumlah *cluster* optimal yang diambil adalah titik 2 dilihat dari titik siku yang ada pada gambar 4.7



Gambar 4.7 Hasil Metode Elbow Clustering Media Promosi 1

Setelah menentukan titik kluster yang ada yaitu 2, selanjutnya dengan mengklasterkan media promosi sesuai dengan variabel yang telah ditentukan dengan menggunakan *K-M*eans, dan menyimpan data klaster ke kolom Cluster_Promosi.

Kode dibawah digunakan untuk mengklasterkan media promosi

```
from sklearn.cluster import KMeans
optimal_clusters = 2
kmeans = KMeans(n_clusters=optimal_clusters,
random_state=42)
df[Cluster_Promosi'] = kmeans.fit_predict(X_scaled1)
```

Tahapan berikutnya yaitu menampilkan data yang sudah di kluster dengan membuat variabel clustered_data1 yang berisi kolom yang akan ditampilkan.

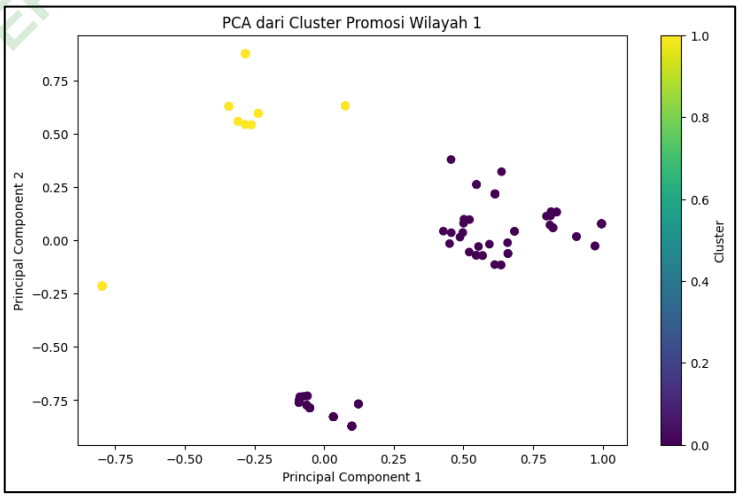
```
# Menampilkan data yang sudah dikluster
clustered_data1 =
data[['No','ID','Tahun','Nama','Provinsi', 'Asal
Info', 'Cluster_Promosi']]
```

Dari kode diatas menghasilkan keluaran tabel kluster media promosi berdasarkan kluster wilayah 1 yang terlihat pada tabel 4.3

Tabel 4.3 Hasil Clustering Media Promosi Wilayah 1

No	ID	Tahun	Nama	Provinsi	Asal Info	Cluster_Promosi
1	20190114	2019	Dita Yanti Fransisca Hasugian	Sumatera Utara	Media Sosial	0
2	20190161	2019	Michael Hosea Nugraha	Jawa Barat	Keluarga	1
3	20190195	2019	Azkiyyatun Nufus Azzahra	Jawa Tengah	Pameran Pendidikan	1
4	20190209	2019	Avita Dwi Febryanti	Jawa Timur	Brosur	1
5	20190237	2019	Ahmad Ilham Nur Febriyanto	Yogyakarta	Keluarga	1

Langkah selanjutnya yaitu membuat visualisasi cluster promosi wilayah 0 dengan scatter plot. *Cluster* 0 berwarna ungu dan *cluster* 1 berwarna kuning yang terlihat pada gambar 4.8.



Gambar 4.8 Scatter Plot Clustering Media Promosi Wilayah 1

Berdasarkan hasil *clustering K-Means* di wilayah 1 dari daerah yang mayoritas pendaftarnya banyak, diperoleh 2 kluster yaitu:

1) *Cluster* 0 media promosi terdiri dari 130 data dengan sumber media promosi yang paling banyak yaitu website (34 data), teman (25 data), keluarga (25 data), satuan, TNI AD (8 data), lainnya (7 data), brosur (6 data), guru BK (4 data), spanduk/baliho (4 data), sosialisasi di sekolah (4 data), kakak kelas (2 data), dosen Unjaya (2 data), datang ke kampus (2 data), alumni Unjaya (2 data), mahasiswa Unjaya (1 data), dan pameran pendidikan (1 data).
2) *Cluster* 1 media promosi terdiri dari 118 data didalamnya hanya ada media sosial.

Dari hasil tersebut dapat disimpulkan bahwa *platform* yang paling efektif yaitu penggunaan media sosial, informasi dan artikel dari website, melalui relasi teman dan informasi keluarga sangat berpengaruh dalam meningkatkan minat pendaftar dan dapat digunakan Staf PMB saat promosi.

Tahap selanjutnya adalah evaluasi dari hasil cluster media promosi pada wilayah 1 dengan *Silhoutte Score* yang menunjukkan nilai 0,71 dimana hasil tersebut mendekati nilai 1 mengindikasikan bahwa titik data yang terkelompok sudah cukup baik. Kode dan hasil keluaran dari evaluasi dari *Silhouette Score* dapat dilihat pada gambar 4.9.

```
silhouette_avg = silhouette_score(X_scaled1, df['Cluster_Promosi'])
print(f'Silhouette Score: {silhouette_avg}')
```

```
Silhouette Score: 0.7107975481586696
```

Gambar 4.9 Silhoutte Score Media Promosi Wilayah 1

b. Clustering K-Means Media Promosi Klaster 0

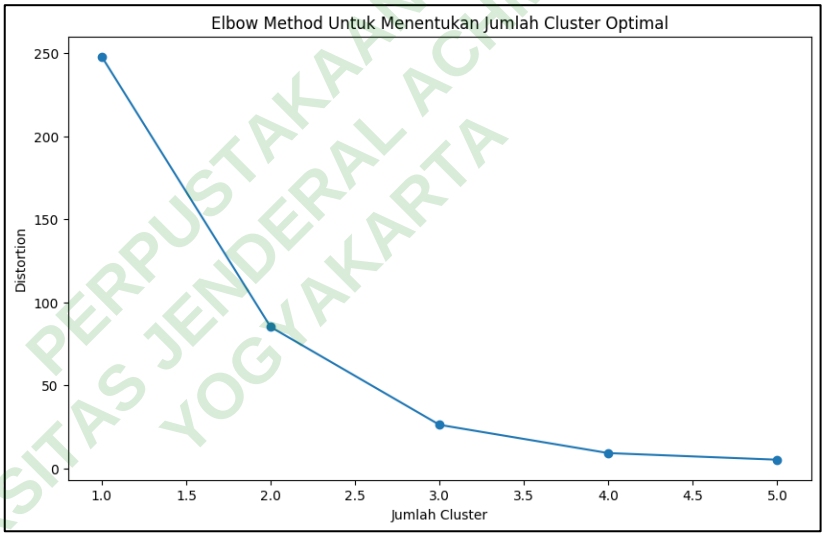
Tahap pertama dilakukan dengan menginisiasi hasil data cluster wilayah 0 untuk dianalisis.

```
w2 = "drive/My Drive/DataSet/NEW/clust0wil.xlsx"
df = pd.read_excel(w2)
```

Selanjutnya dilakukan pemilihan variabel yang akan digunakan untuk *clustering* media promosi, pada tahap ini target variable X adalah Asal_Info, dan dilakukan standarisasi data agar klaster tidak terlalu jauh berkelompok.

```
X = df[['Asal Info']]
scaler = StandardScaler()
X_scaled2 = scaler.fit_transform(X)
```

Tahapan berikutnya yaitu menentukan jumlah *cluster* ideal yang akan digunakan untuk menentukan titik kluster dengan metode Elbow. Jumlah *cluster* optimal yang diambil adalah titik 2 dilihat dari titik siku yang ada pada gambar 4.10



Gambar 4.10 Hasil Metode Elbow Clustering Media Promosi 0

Setelah menentukan titik kluster yang ada yaitu 2, selanjutnya dengan mengklasterkan media promosi sesuai dengan variabel yang telah ditentukan dengan menggunakan *K-M*eans, dan menyimpan data klaster ke kolom Cluster_Promosi.

```
from sklearn.cluster import KMeans
optimal_clusters = 2
kmeans = KMeans(n_clusters=optimal_clusters,
random_state=42)
df[' Cluster_Promosi'] = kmeans.fit_predict(X_scaled2)
```

Tahapan berikutnya yaitu menampilkan data yang sudah di kluster dengan membuat variabel clustered_data1 yang berisi kolom spesifik yang akan ditampilkan saat *running code*.

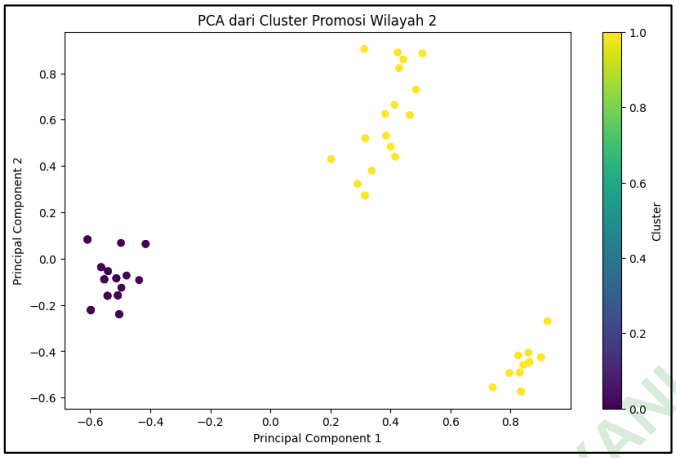
```
clustered_data2 =
df[['No','ID','Tahun','Nama','Provinsi', 'Asal Info',
'Cluster_Promosi']]
```

Dari kode diatas menghasilkan keluaran tabel kluster media promosi berdasarkan kluster wilayah 0 yang terlihat pada tabel 4.4

Tabel 4.4 Hasil Clustering Media Promosi Wilayah 0

No	ID	Tahun	Nama	Provinsi	Asal Info	Cluster_Promosi
1	20190191	2019	Oktavila Nauli Sinurat	Kalimantan Barat	Website	1
2	20191512	2019	Muhammad Zulfikar Reyhan	Jakarta	Website	1
3	20192122	2019	Davit ahmad zikri	Sumatera Barat	Media Sosial	0
4	20195095	2019	Afifa Nabila Nurrosni	Bali	Teman	1
5	20195119	2019	Sri Wahyuni	Kalimantan Timur	Media Sosial	0

Langkah selanjutnya yaitu membuat visualisasi cluster promosi wilayah 0 dengan scatter plot. *Cluster* 0 berwarna ungu dan *cluster* 1 berwarna kuning yang terlihat pada gambar 4.11.



Gambar 4.11 Scatter Plot Clustering Media Promosi Wilayah 0

Berdasarkan hasil *clustering K-Means* di wilayah 0 dari daerah yang mayoritas pendaftarnya sedikit, diperoleh 2 kluster yaitu:

1) *Cluster* 1 media promosi terdiri dari 45 data dengan sumber media informasi paling banyak yaitu website (20 data), teman (9 data), keluarga (8 data), datang ke kampus (2 data), alumni Unjaya (2 data), dosen Unjaya (2 data), pameran pendidikan (1 data), satuan TNI AD (1 data).
2) *Cluster* 0 media promosi terdiri dari 45 data didalamnya hanya ada media sosial.

Dari hasil tersebut dapat disimpulkan bahwa penggunaan media sosial banyak digunakan oleh calon pendaftar, dilanjutkan dengan website, melalui relasi teman dan informasi dari keluarga sangat berpengaruh dalam meningkatkan minat pendaftar dan dapat digunakan Staf PMB saat promosi.

Tahap selanjutnya adalah evaluasi menggunakan *Silhoutte Score* yang menunjukkan nilai 0,83 dimana nilai mendekati 1 sudah terkelompok dengan baik. Kode dan hasil keluaran dari evaluasi dari *Silhouette Score* dapat dilihat pada gambar 4.12.

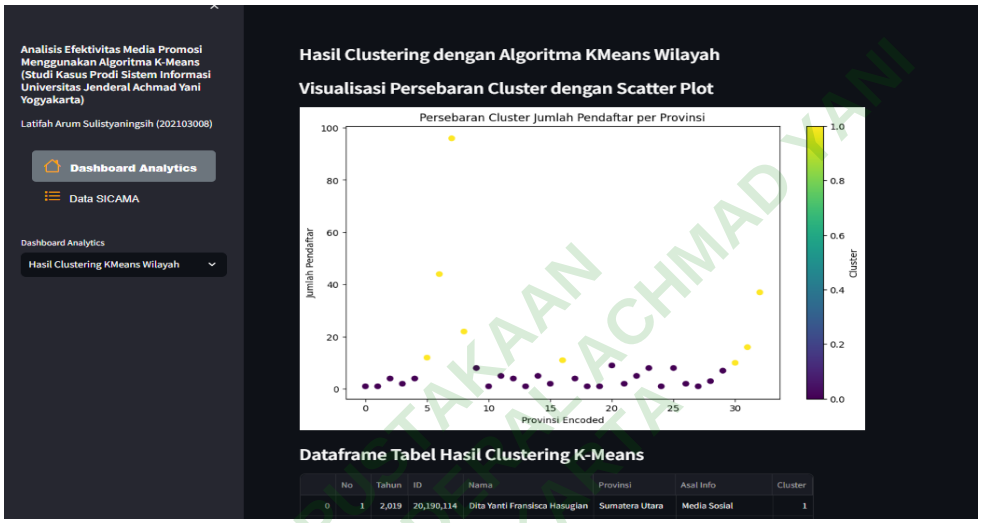
```
silhouette_avg = silhouette_score(X_scaled2, df['Cluster_Promosi'])
print(f'Silhouette Score: {silhouette_avg}')
```

```
Silhouette Score: 0.8369218092438727
```

Gambar 4.12 Silhoutte Score Media Promosi Wilayah 0

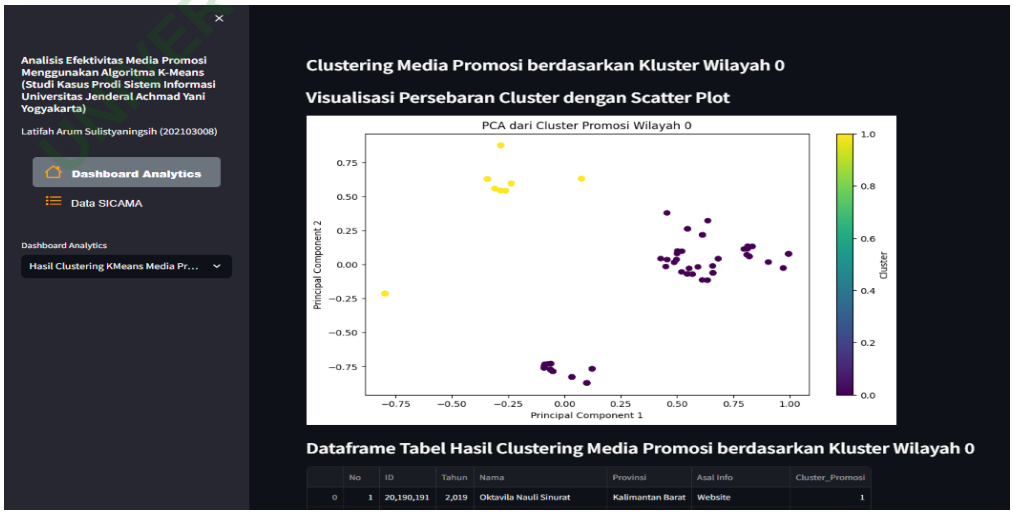
4.5 HASIL DEPLOYEMENT

Dashboard analytics dibuat untuk menampilkan hasil dari penelitian yang telah dilakukan agar mudah dimengerti oleh orang awam yang tidak begitu paham dengan pemrograman. Halaman pertama *dashboard* berisi hasil *clustering K-Means*, visualisasi data yang berisi *scatter plot* persebaran cluster, *dataframe*, dan hasil *clustering* yang terlihat pada gambar 4.13.



Gambar 4.13 Dashboard Halaman Pertama

Pada gambar 4.14 adalah bagian hasil *clustering* media promosi tiap wilayah *cluster*. Pada halaman ini ditampilkan hasil *scatter plot* berdasarkan media promosi di tiap wilayah cluster serta hasil analisisnya.



Gambar 4.14 Dashboard Halaman Kedua

Halaman terakhir *dashboard* berisi tabel *dataframe* pendaftar Prodi Sistem Informasi pada tahun 2019, 2020, 2021, dan 2022. *Dashboard* dapat dilihat pada gambar 4.15.

Gambar 4.15 Dashboard Halaman Terakhir

4.6 HASIL PEMBAHASAN

Berdasarkan *clustering* wilayah terdapat 2 kelompok, berikut adalah hasil dari *clustering* menggunakan algoritma *K-Means*:

1. *Cluster* 0 sebanyak 248 data yang merupakan wilayah dengan jumlah pendaftar tersedikit, terdiri dari wilayah Nusa Tenggara Timur, Sulawesi Selatan, Riau, Kalimantan Barat, Sumatera Barat, Papua Barat, Kalimantan Tengah, Kepulauan Bangka Belitung, Maluku, Kalimantan Timur, Jakarta, Banten, Sulawesi Utara, Sulawesi Tengah, Papua, Bengkulu, Kepulauan Riau, Aceh, Bali, Kalimantan Selatan, Nusa Tenggara Barat, Sulawesi Tenggara, Sulawesi Barat, Maluku Utara, dan Kalimantan Utara.
2. *Cluster* 1 sebanyak 90 data yang merupakan wilayah dengan jumlah pendaftar terbanyak, terdiri dari wilayah Jawa Tengah, Jawa Barat, Yogyakarta, Jawa Timur, Jakarta, Sumatera Utara, Jambi, Lampung, dan Sumatera Selatan.

Setelah itu dilanjutkan dengan pengelompokan media promosi di masing-masing wilayah. Hasil dari *clustering* di wilayah 1 dari daerah yang pendaftarnya banyak, diperoleh 2 *cluster* yaitu:

1. *Cluster* 0 media promosi terdiri dari 130 data dengan sumber media promosi yang paling banyak yaitu website (34 data), teman (25 data), keluarga (25 data), satuan, TNI AD (8 data), lainnya (7 data), brosur (6 data), guru BK (4 data), spanduk/baliho (4 data), sosialisasi di sekolah (4 data), kakak kelas (2 data), dosen Unjaya (2 data), datang ke kampus (2 data), alumni Unjaya (2 data), mahasiswa Unjaya (1 data), dan pameran pendidikan (1 data).
2. *Cluster* 1 media promosi terdiri dari 118 data didalamnya hanya ada media sosial.

Hasil *clustering K-Means* di wilayah 0 dari daerah yang pendaftarnya sedikit, diperoleh 2 *cluster* yaitu:

1. *Cluster* 1 media promosi terdiri dari 45 data dengan sumber media informasi paling banyak yaitu website (20 data), teman (9 data), keluarga (8 data), datang ke kampus (2 data), alumni Unjaya (2 data), dosen Unjaya (2 data), pameran pendidikan (1 data), satuan TNI AD (1 data).
2. *Cluster* 0 media promosi terdiri dari 45 data didalamnya hanya ada media sosial.

Berdasarkan pola yang sama diantara *cluster* media promosi, menunjukkan bahwa *platform* yang paling efektif adalah media sosial, website, teman, dan keluarga. Informasi yang disampaikan melalui media promosi tersebut sangat berpengaruh dalam menarik calon mahasiswa. Sementara itu, kesesuaian model dan visualisasi pada *dashboard* yang dibuat sudah sesuai. *Dashboard* menunjukkan bahwa pada penerapan algoritma *K-Means* ini dapat memberikan informasi terkait hasil dari penelitian dan analisisnya.