

LAMPIRAN

Lampiran 1: Source Code

```
# 1. IMPORT LIBRARY

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

from sklearn.preprocessing import StandardScaler

from scipy.cluster.hierarchy import linkage, dendrogram

from sklearn.cluster import AgglomerativeClustering

from sklearn.metrics import silhouette_score

from google.colab import files

# 2. UPLOAD & LOAD DATASET

uploaded = files.upload()

filename = list(uploaded.keys())[0]

df = pd.read_excel(filename)

# 3. CEK STRUKTUR DATA & MISSING VALUE

df.info()

print(df.isnull().sum())

df.describe()
```

```
# 4. KONVERSI Usage dari Byte ke GB
df['Usage (GB)'] = df['Usage (Byte)'] / 1e9

# 5. FILTERING DATA
df_cleaned = df[(df['Tagihan (Rp)'] > 0) &
                (df['Usage (GB)'] > 0) &
                (df['Abonemen (Rp)'] < 5_000_000) &
                (df['Durasi Langganan (Bulan)'] <= 120)].copy()

print(f'✅ Data setelah difilter: {df_cleaned.shape[0]} pelanggan")

# 6. FEATURE ENGINEERING TAMBAHAN
df_cleaned['Rp_per_GB'] = df_cleaned['Tagihan (Rp)'] /
(df_cleaned['Usage (GB)'] + 1e-6)
df_cleaned['Session_Intensity'] = df_cleaned['Jumlah Session (kali)'] /
(df_cleaned['Durasi Langganan (Bulan)'] + 1)

# 7. TRANSFORMASI LOG UNTUK FITUR SKEWED
df_cleaned['Log_Usage_GB'] = np.log1p(df_cleaned['Usage (GB)'])
df_cleaned['Log_Session'] = np.log1p(df_cleaned['Jumlah Session (kali)'])

# 8. PILIH FITUR FINAL UNTUK CLUSTERING
features = ['Abonemen (Rp)', 'Rp_per_GB', 'Session_Intensity',
```

```
'Tagihan (Rp)', 'Layanan TV (1=Ya, 0=Tidak)',  
'Kategori (Encoded)', 'Durasi Langganan (Bulan)',  
'Log_Usage_GB', 'Log_Session']  
  
print(f'✅ Fitur yang digunakan: {features}')  
# 9. IMPUTASI MISSING VALUE DENGAN MEDIAN (jika ada)  
df_cleaned[features] =  
df_cleaned[features].fillna(df_cleaned[features].median())  
print("✅ Missing values diimputasi dengan median.")  
  
# 10. NORMALISASI FITUR UNTUK CLUSTERING  
scaler = StandardScaler()  
scaled_data = scaler.fit_transform(df_cleaned[features])  
print("✅ Data berhasil dinormalisasi.")  
  
# 11. DENDROGRAM VISUALIZATION  
linked = linkage(scaled_data, method='ward')  
plt.figure(figsize=(12, 6))  
dendrogram(linked, truncate_mode='lastp', p=10,  
            leaf_rotation=45., leaf_font_size=12., show_contracted=True)  
plt.title('🇮🇩 Dendrogram Pelanggan IndiHome')
```

```
plt.xlabel('Cluster')

plt.ylabel('Distance')

plt.grid()

plt.show()

# 12. EVALUASI SILHOUETTE SCORE UNTUK CLUSTER 2-5

scores = {}

print(" 🚩 Silhouette Scores per jumlah cluster:")

for n_clusters in range(2, 6):

    model = AgglomerativeClustering(n_clusters=n_clusters,

linkage='ward')

    labels = model.fit_predict(scaled_data)

    score = silhouette_score(scaled_data, labels)

    scores[n_clusters] = score

    print(f" - Cluster = {n_clusters}, Silhouette Score = {score:.4f}")

# 13. TENTUKAN JUMLAH CLUSTER TERBAIK

best_n = max(scores, key=scores.get)

print(f" ✅ Jumlah cluster terbaik: {best_n}")

# 14. FINAL CLUSTERING

final_model = AgglomerativeClustering(n_clusters=best_n, linkage='ward')
```

```
df_cleaned['Cluster'] = final_model.fit_predict(scaled_data)

print("✅ Clustering final selesai.")

# 15. RANGKUMAN HASIL PER CLUSTER

cluster_summary =

df_cleaned.groupby('Cluster')[features].mean().round(2)

print("🔥 Rata-rata per Cluster:")

display(cluster_summary)

print("🔥 Jumlah Pelanggan per Cluster:")

print(df_cleaned['Cluster'].value_counts().sort_index())

# 16. EKSPOR HASIL

df_cleaned.to_csv('hasil_segmentasi_indihome.csv', index=False)

files.download('hasil_segmentasi_indihome.csv')
```

Lampiran 2: Dokumentasi



DAFTAR RIWAYAT HIDUP



Farid Rizky Sukamto lahir pada tanggal 07 Februari 2001 di Bontang Provinsi Kalimantan Timur. Penulis merupakan anak kedua dari 3 bersaudara pasangan Bapak Sukamto dan Ibu Umi Maesaroh. Penulis memulai Pendidikan formal di SDN 4 Karangtengah Jl. Maluku No.14, Karangtengah, Kec. Kec. Sananwetan, Kota Blitar tahun 2008 dan lulus di tahun 2014. Pada tahun yang sama, penulis melanjutkan Sekolah Menengah Pertama di SMPN 5 Kota Blitar lulus pada tahun 2017. Setelah lulus dari SMPN 5 Kota Blitar penulis melanjutkan Pendidikan di SMKN 1 BLITAR jurusan Teknik Komputer dan Jaringan dan lulus di tahun 2020. Pendidikan berikutnya di Universitas Islam Blitar dengan mengambil program studi Teknik Informatika.