

Kelompok 1

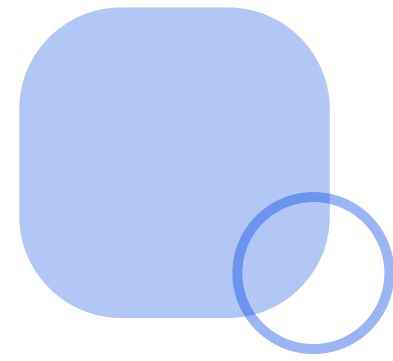
SEGMENTASI PELANGGAN PRODUK SARI ROTI DENGAN DBSCAN CLUSTERING ANALYSIS

ANGGOTA KELOMPOK

- 01 Muhammad Nouval Habibie (2211521020)
- 02 Nabil Rizki Navisa (2211522018)
- 03 Ilham Nofaldi (2211522028)
- 04 Mustafa Fathur Rahman (2211522036)
- 05 Syauqi Nabiih Marwa (2211523012)



DISTRIBUSI SARI ROTI

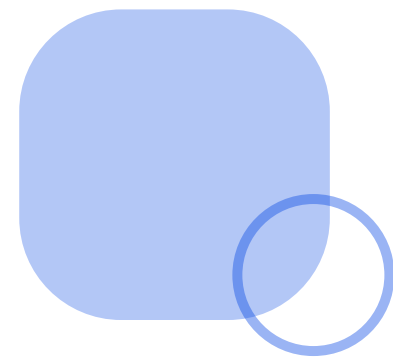


Perusahaan Distribusi Sari Roti adalah perusahaan yang melakukan proses bisnis mendistribusikan berbagai macam produk Sari Roti ke berbagai outlet. Perusahaan Distribusi Sari Roti kemudian membuat laporan penjualan produk Sari Roti di seluruh outlet yang berisi jumlah produk yang diletakkan pada outlet, jumlah produk yang terjual, dan jumlah produk yang tersisa, serta beberapa keterangan lain mengenai penjualan produk.





MASALAH

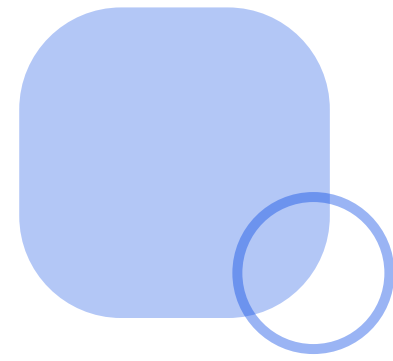


Tantangan muncul dalam mengelompokkan pelanggan berdasarkan pola pembelian karena data penjualan yang ada belum tersegmentasi dengan baik. Hal ini menyulitkan perusahaan dalam mengidentifikasi tren dan membuat prediksi permintaan yang akurat, berpotensi menyebabkan masalah stok dan ketidakpuasan pelanggan.





SOLUSI



Solusi yang diberikan untuk mengatasi masalah tersebut adalah penggunaan Algoritma DBSCAN, di mana:

- DBSCAN (*Density-Based Spatial Clustering of Applications with Noise*): Algoritma clustering yang digunakan untuk mengelompokkan data berdasarkan kepadatan titik-titik data dalam ruang.
- Fokus: Mengelompokkan outlet berdasarkan waktu pembelian dan jenis produk yang dibeli oleh pelanggan.



KENAPA DBSCAN?



Algoritma DBSCAN dipilih karena dapat mengidentifikasi cluster dengan bentuk arbitrer dan menangani data dengan noise. DBSCAN juga efektif untuk data dengan kepadatan berbeda, cocok untuk pola pembelian yang beragam di berbagai outlet Sari Roti. Keunggulan-keunggulan ini membuat DBSCAN ideal untuk mengelompokkan outlet berdasarkan waktu dan jenis produk yang dibeli.

DATASET

Dataset distribusi Sari Roti yang didapatkan dari laporan distribusi mulai dari tanggal 1 hingga 9 Juni 2022 .

SALESMAN ID	SALESMAN NIK	SALESMAN NAME	DROPPING NUMBER	DROPPING DATE	OUTLET BRAND	OUTLET CODE	OUTLET NAME	PRODUCT ID	PRODUCT CODE
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	19	FG101031
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	20	FG101032
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	137	FG101069
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	120	FG101074
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	166	FG101083
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	165	FG101084
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	126	FG203022
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	72	FG204001
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	153	FG206030
71299205	135441871299205	REZKI SAPUTRA	407134420522060115234830	2022-06-01	Sari Roti	1354390796	(HWK) ANDI KEC. PASAMAN	178	FG206032

DATASET (2)

SHORT NAME	QTY ESTIMASI	QTY DROPPING	DROPPING CBP	DISC. DROPPING	DROPPING RBP	QTY RETUR BAIK	RETUR BAIK CBP	RETUR BAIK RBP	DISC RETUR BAIK	QTY RETUR BS	RETUR BS CBP	RETUR BS RBP
RTSII	0	2	29000	2900	26100	0	0	0	0	0	0	0
RTGII	0	3	64500	6450	58050	0	0	0	0	0	0	0
SCK2	0	12	60000	6000	54000	0	0	0	0	0	0	0
SSM2	0	2	11000	1100	9900	0	0	0	0	0	0	0
RJTS2	0	1	18000	1800	16200	0	0	0	0	0	0	0
RJKU II	0	2	39000	3900	35100	0	0	0	0	0	0	0
SRM2	0	1	6000	600	5400	0	0	0	0	0	0	0
DIC	0	1	5500	550	4950	0	0	0	0	0	0	0
RKS II	0	1	10500	1050	9450	0	0	0	0	0	0	0
TOC II SS	0	6	105000	10500	94500	0	0	0	0	0	0	0

DATASET (3)

[illegible]

PENERAPAN ALGORITMA DBSCAN DENGAN PERHITUNGAN MANUAL.

Iterasi 1

Pada iterasi I titik pusat awal ditentukan secara acak dengan demikian titik pusat adalah titik A dengan $x_p=8$, $y_p=1$, $z_p=2$

$$\begin{aligned} AB &= \sqrt{(8-5)^2 + (1-1)^2 + (2-1)^2} \\ &= 3.16 \end{aligned}$$

$$\begin{aligned} AC &= \sqrt{(8-6)^2 + (1-43)^2 + (2-1)^2} \\ &= 42.05 \end{aligned}$$

$$\begin{aligned} AD &= \sqrt{(8-9)^2 + (1-26)^2 + (2-2)^2} \\ &= 25.10 \end{aligned}$$

$$\begin{aligned} AE &= \sqrt{(8-8)^2 + (1-38)^2 + (2-1)^2} \\ &= 37.01 \end{aligned}$$

$$\begin{aligned} AF &= \sqrt{(8-3)^2 + (1-26)^2 + (2-2)^2} \\ &= 25.49 \end{aligned}$$

$$\begin{aligned} AG &= \sqrt{(8-6)^2 + (1-37)^2 + (2-1)^2} \\ &= 36.06 \end{aligned}$$

$$\begin{aligned} AH &= \sqrt{(8-4)^2 + (1-6)^2 + (2-3)^2} \\ &= 6.48 \end{aligned}$$

$$\begin{aligned} AI &= \sqrt{(8-4)^2 + (1-1)^2 + (2-3)^2} \\ &= 4.12 \end{aligned}$$

$$\begin{aligned} AJ &= \sqrt{(8-5)^2 + (1-33)^2 + (2-1)^2} \\ &= 32.15 \end{aligned}$$

$$\begin{aligned} AK &= \sqrt{(8-5)^2 + (1-3)^2 + (2-1)^2} \\ &= 3.74 \end{aligned}$$

PENERAPAN ALGORITMA DBSCAN DENGAN PERHITUNGAN MANUAL.

Iterasi 1

Titik	Jarak
AB	3.16227766
AC	42.05948169
AD	25.01999201
AE	37.01351105
AF	25.49509757
AG	36.06937759
AH	6.480740698
AI	4.123105626
AJ	32.15587038
AK	3.741657387

Eps = 5

Titik	Name_Code	QTY SELLS	Date	Jarak Ke Titik A
A	8	1	2	0
B	5	1	1	3.16227766
C	6	43	1	42.05948169
D	9	26	2	25.01999201
E	8	38	1	37.01351105
F	3	26	2	25.49509757
G	6	37	1	36.06937759
H	4	6	3	6.480740698
I	4	1	3	4.123105626
J	5	33	1	32.15587038
K	5	3	1	3.741657387

PENERAPAN ALGORITMA DBSCAN DENGAN PERHITUNGAN MANUAL.

Iterasi 2

Pada iterasi II titik pusat didapat dari iterasi I yaitu titik G dengan $x_p=6$, $y_p=37$, $z_p=2$.

Titik	Jarak
GA	36.06937759
GB	36.01388621
GC	6
GD	11.44552314
GE	2.236067977
GF	11.44552314
GG	0
GH	31.12876483
GI	36.11094017
GJ	4.123105626
GK	34.0147027

Titik	Name_Code	QTY SELLS	Date	Jarak Ke Titik G
A	8	1	2	36.06937759
B	5	1	1	36.01388621
C	6	43	1	6
D	9	26	2	11.44552314
E	8	38	1	2.236067977
F	3	26	2	11.44552314
G	6	37	1	0
H	4	6	3	31.12876483
I	4	1	3	36.11094017
J	5	33	1	4.123105626
K	5	3	1	34.0147027



PENERAPAN ALGORITMA DBSCAN DENGAN PERHITUNGAN MANUAL.

Iterasi 3

Pada iterasi III titik pusat didapat dari iterasi II yaitu titik I dengan $x_p=4$, $y_p=1$, $z_p=3$.

Titik	Jarak
IA	4.123105626
IB	2.236067977
IC	42.09513036
ID	25.51470164
IE	37.26929031
IF	25.03996805
IG	36.11094017
IH	5
II	0
IJ	32.07802986
IK	3

Titik	Name_Code	QTY SELLS	Date	Jarak Ke Titik I
A	8	1	2	4.123105626
B	5	1	1	2.236067977
C	6	43	1	42.09513036
D	9	26	2	25.51470164
E	8	38	1	37.26929031
F	3	26	2	25.03996805
G	6	37	1	36.11094017
H	4	6	3	5
I	4	1	3	0
J	5	33	1	32.07802986
K	5	3	1	3



PENERAPAN ALGORITMA DBSCAN DENGAN PERHITUNGAN MANUAL.

Iterasi

Pembagian cluster sementara dari core point yang berbeda. Syarat untuk menjadi 1 cluster penuh adalah $\text{min_samples} = 5$. untuk menghitung seluruh data maka diperlukan otomatisasi menggunakan pemrograman python DBSCAN

Iterasi 1	Iterasi 2	Iterasi 3
AB	GE	IH
AI	GJ	
AK		

CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Mengimport Library dan menampilkan data

```
import pandas as pd
import numpy as np
from sklearn.cluster import DBSCAN
import matplotlib.pyplot as plt
from mpl_toolkits.mplot3d import
Axes3D

data =
pd.read_excel('sarirotidata2.xlsx')
df = pd.DataFrame(data)
data.head(20)
```




CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Date dipilih dari DataFrame dan diubah menjadi array numpy X. Objek DBSCAN kemudian dibuat dengan parameter `eps=5` dan `min_samples=5`

```
X = df[['Name_Code', 'QTY SELLS', 'Date']].values
dbscan = DBSCAN(eps=5, min_samples=5)
clusters = dbscan.fit_predict(X)
df['Cluster'] = clusters
```

CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

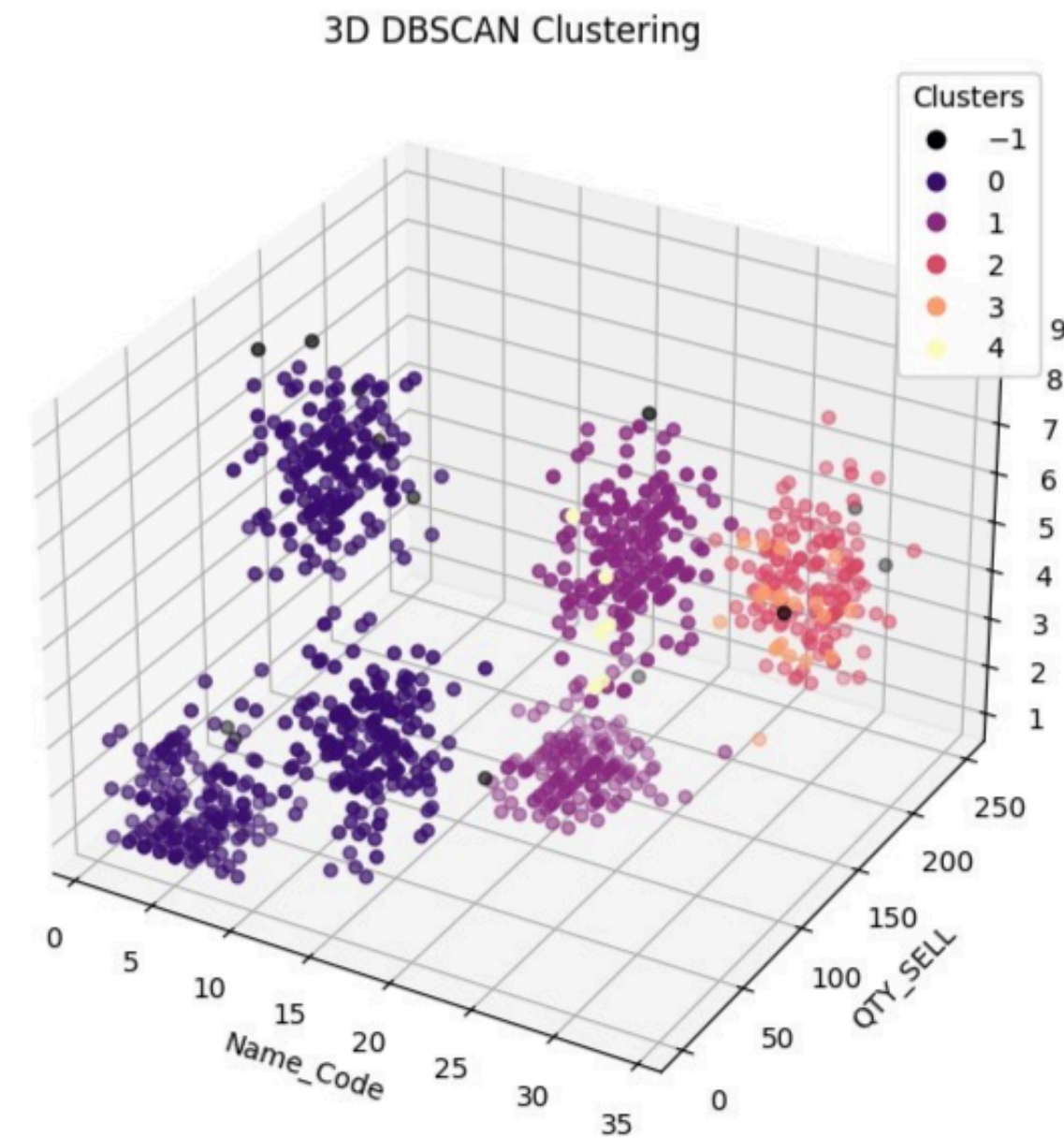
Plot 3D dibuat menggunakan matplotlib

```
fig = plt.figure(figsize=(20, 7))
ax = fig.add_subplot(111, projection='3d')
scatter = ax.scatter(df['Name_Code'], df['QTY_SELLS'], df['Date'], c=df['Cluster'], cmap='magma')
ax.set_xlabel('Name_Code')
ax.set_ylabel('QTY_SELL')
ax.set_zlabel('Date')
ax.set_title('3D DBSCAN Clustering')
legend1 = ax.legend(*scatter.legend_elements(), title="Clusters")
ax.add_artist(legend1)
plt.show()
```

CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Hasil visualisasi



CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Defaultdict dari pustaka collections diimpor untuk membuat dictionary dengan nilai default berupa list.

```
from collections import defaultdict

dbscan = DBSCAN(eps=5, min_samples=5)
clusters = dbscan.fit_predict(X)

cluster_members = defaultdict(list)
for i, label in enumerate(clusters):
    cluster_members[label].append(i)

for label, members in
    cluster_members.items():
        if label != -1 and len(members) >= 5:
            print(f'Cluster {label}:
{members}')
```



CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Menampilkan jumlah anggota tiap kluster

```
cluster_counts = {label: len(members) for label, members in  
cluster_members.items() if label != -1}  
  
for label, count in cluster_counts.items():  
    print(f'Cluster {label} memiliki {count} anggota')
```



CLUSTERING DENGAN ALGORITMA DBSCAN

Penerapan Algoritma DBScan

Hasil jumlah anggota tiap klaster

```
Cluster 0 memiliki 488 anggota  
Cluster 1 memiliki 324 anggota  
Cluster 2 memiliki 141 anggota  
Cluster 3 memiliki 21 anggota  
Cluster 4 memiliki 6 anggota
```

KESIMPULAN



Penelitian ini berhasil menerapkan algoritma DBSCAN untuk segmentasi pelanggan Sari Roti berdasarkan pola pembelian, mengungkapkan kelompok pelanggan dengan preferensi produk dan waktu pembelian yang berbeda. Informasi ini memungkinkan perusahaan untuk mengoptimalkan manajemen stok, distribusi, dan strategi pemasaran untuk meningkatkan efisiensi operasional dan kepuasan pelanggan.





THANK YOU