

Analisis Habitabilitas (Kelayakhunian)

Kelompok DUAR - KASDD B



ANGGOTA KELOMPOK



(2006596573)

Marvel Krent



(2006596522)

Gitan Sahl
Tazakha
Wijaya



(2006473945)

Muhammad
Iqbal Asrif

Table of contents

01

Deskripsi Data

Deskripsi dari Data Habitability Analysis

02

Visualisasi Data

Visualisasi dari Data Habitability Analysis

03

Preprocessing Data

Melakukan pembersihan data

04

Modelling

Klasifikasi, Regresi, Clustering

A stylized illustration of a woman with dark hair pulled back, wearing a purple long-sleeved shirt. She is holding a black smartphone in her left hand and a black pen in her right hand, pointing it towards the phone's screen. She has a small hoop earring and a thin necklace. The background behind her features abstract geometric shapes like diamonds and triangles in shades of purple, teal, and grey.

01

Deskripsi Data

Deskripsi dari Data Habitability
Analysis

Informasi Fitur

NAME	TYPE	DESCRIPTION
Property_ID	int64	ID unik dari property
Property_Type	Object	Jenis dari properti (Apartment, Bungalow, etc)
Property_Area	int64	Area properti dalam feet persegi
Number_of_Windows	float64	Jumlah Jendela

Informasi Fitur [2]

NAME	TYPE	DESCRIPTION
Number_of_Doors	int64	Jumlah Pintu
Furnishing	Object	Jenis dari Furnish (Fully Furnished, Semi Furnished, atau Unfurnished)
Frequency_of_Powercuts	float64	Rata-rata mati lampu dalam seminggu
Power_Backup	Object	Kemampuan untuk menyediakan listrik backup

Informasi Fitur [3]

NAME	TYPE	DESCRIPTION
Water_Supply	int64	Kesediaan air (All time, Once in a day - Morning, Once in a day - Evening, and Once in two days)
Traffic_Density_Score	Object	Tingkat kepadatan lalu lintas/kemacetan
Crime_Rate	float64	Tingkat kejahatan di daerah properti (Well below average, Slightly below average, Slightly above average, and Well above average)
Dust_and_Noise	Object	Tingkat debu dan suara (High, Medium, Low)

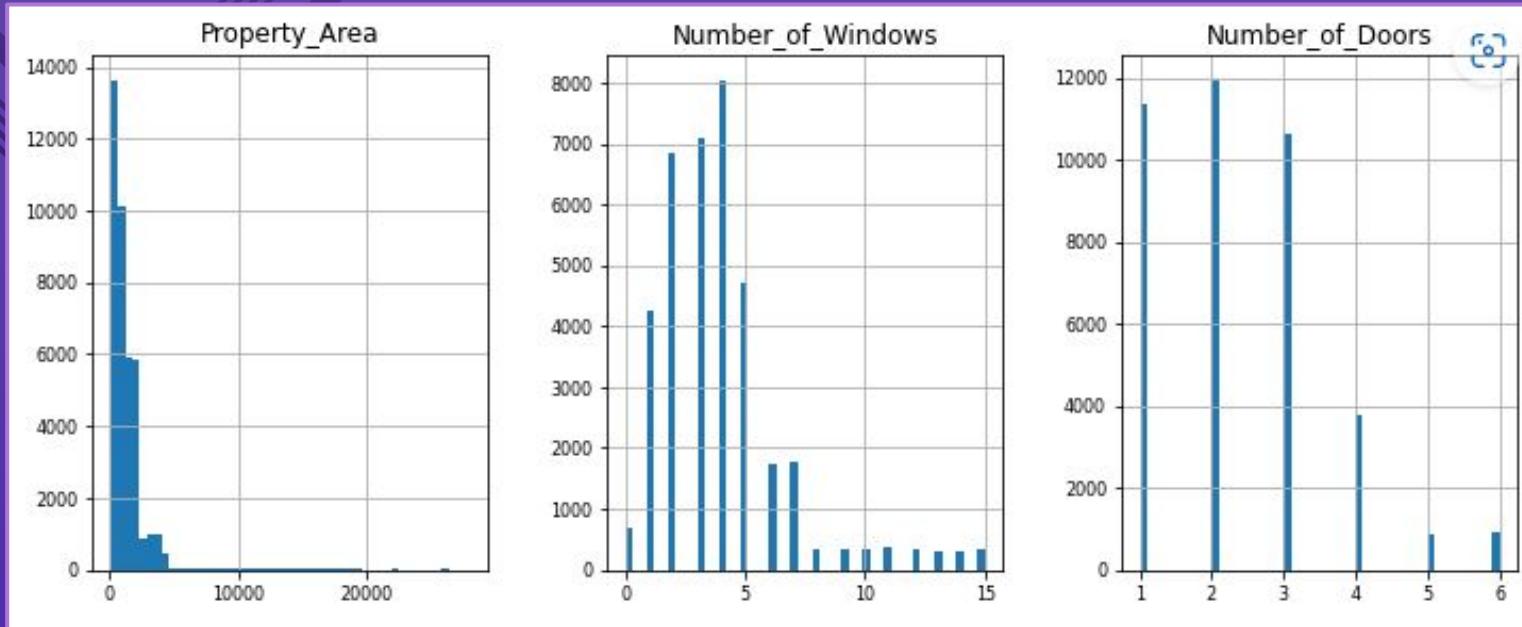
Informasi Fitur [4]

NAME	TYPE	DESCRIPTION
Air_Quality_Index	Object	Tingkat kualitas udara
Neighborhood_Review	float64	Penilaian daerah oleh masyarakat
Habitability_category	Object	Kategori kelayakhunian suatu properti

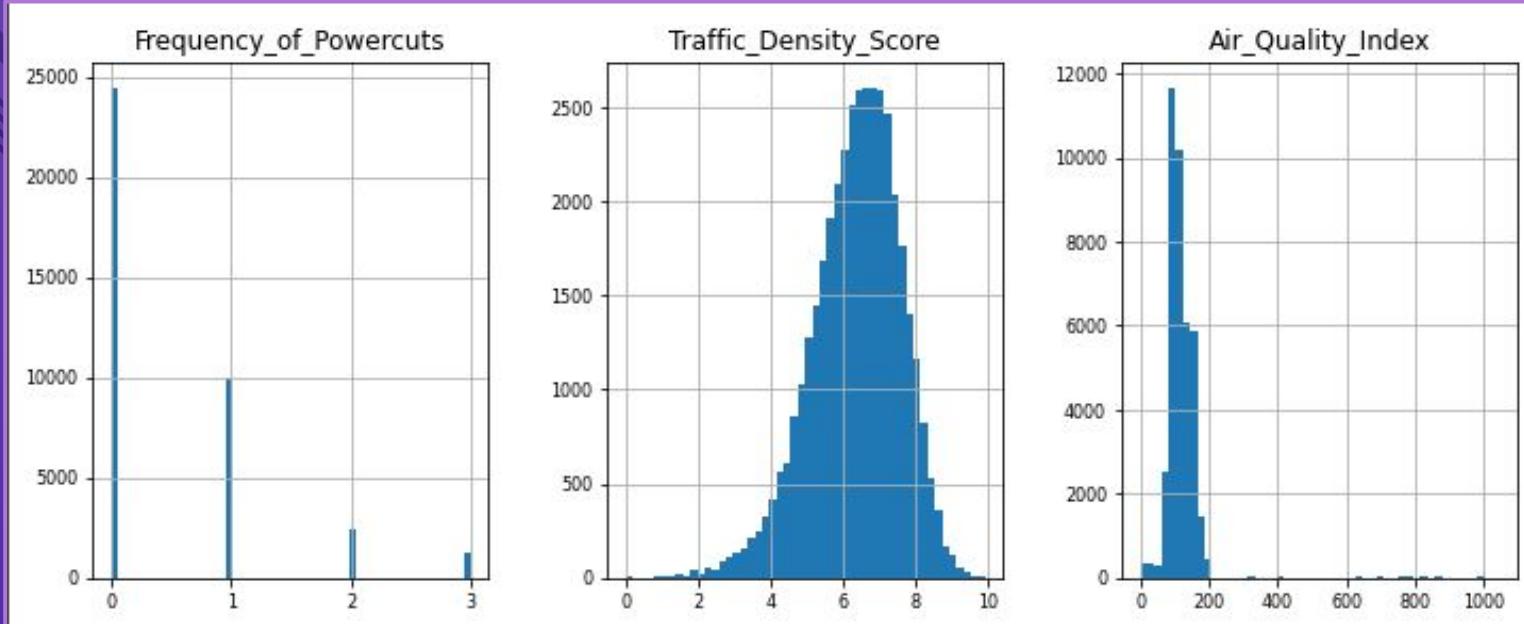
Deskripsi Statistik

	Property_Area	Number_of_Windows	Number_of_Doors	Frequency_of_Powercuts	Traffic_Density_Score	Air_Quality_Index	Neighborhood_Review
count	39499.000000	37845.000000	39499.000000	38116.000000	39499.000000	39499.000000	39499.000000
mean	1391.715638	3.923768	2.331375	0.484678	6.354772	121.352566	3.735945
std	1953.287544	2.690992	1.179964	0.752207	1.219681	62.384669	0.796843
min	100.000000	0.000000	1.000000	0.000000	0.000000	0.000000	0.000000
25%	494.000000	2.000000	1.000000	0.000000	5.600000	95.000000	3.220000
50%	932.000000	4.000000	2.000000	0.000000	6.470000	113.000000	3.850000
75%	1770.000000	5.000000	3.000000	1.000000	7.220000	140.000000	4.260000
max	28064.000000	15.000000	6.000000	3.000000	9.910000	1044.000000	5.000000

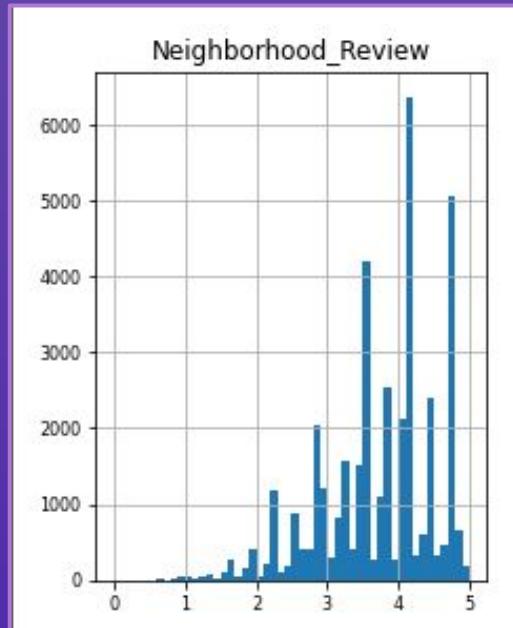
Distribusi Data



Distribusi Data [2]



Distribusi Data [3]



02

Visualisasi Data

Visualisasi dari Data Habitability Analysis

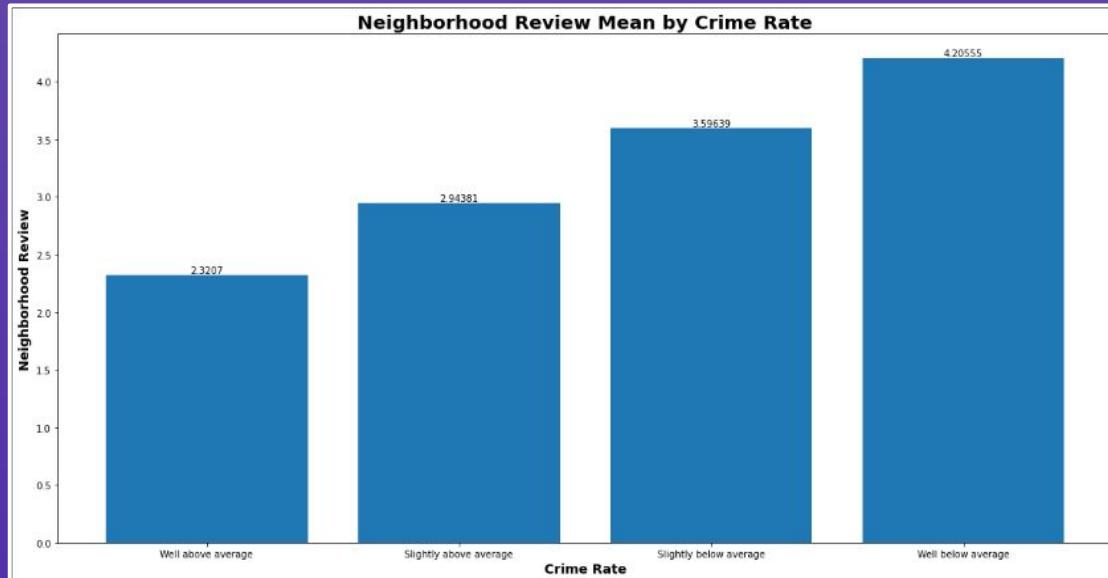


Rata-rata Nilai Neighborhood_Review untuk setiap Kategori Crime_Rate

Problem Guideline Nomor 1A



Visualisasi



Well Above Avg

2,32

Slightly Above

2,94

Slightly Below

3,59

Well Below Avg

4,20

Rata-rata nilai review dari pemukiman berdasarkan Crime Rate mengikuti tingkat kriminalitasnya.

Ciri-ciri Properti yang Memiliki Habitability_category Bernilai “High”

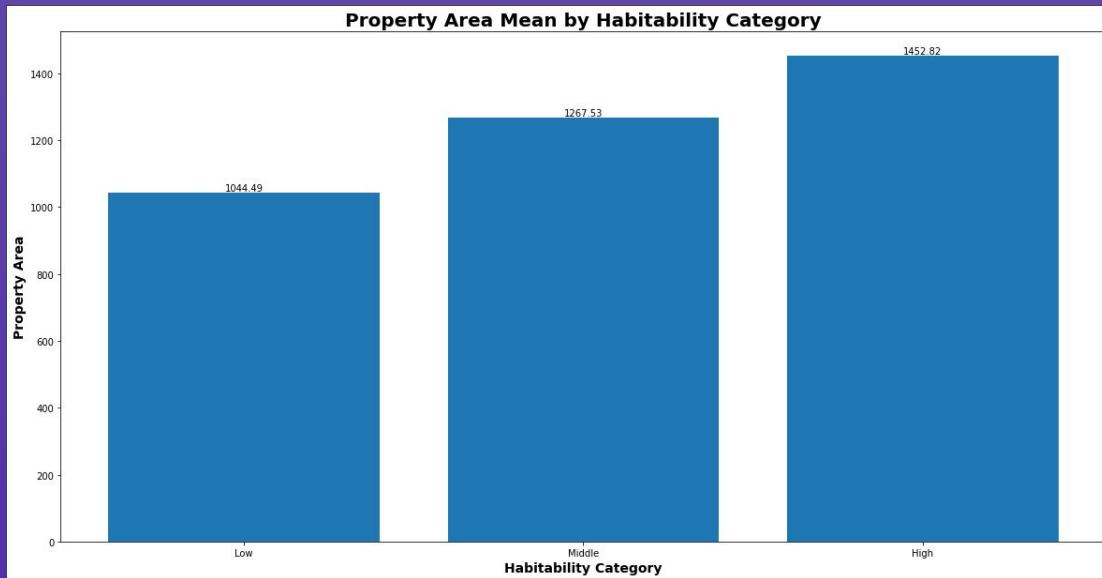
Problem Guideline Nomor 1B





Berdasarkan Fitur Kuantitatif

Visualisasi Rata-rata Luas Area



Habitability High

1452,8

Habitability Medium

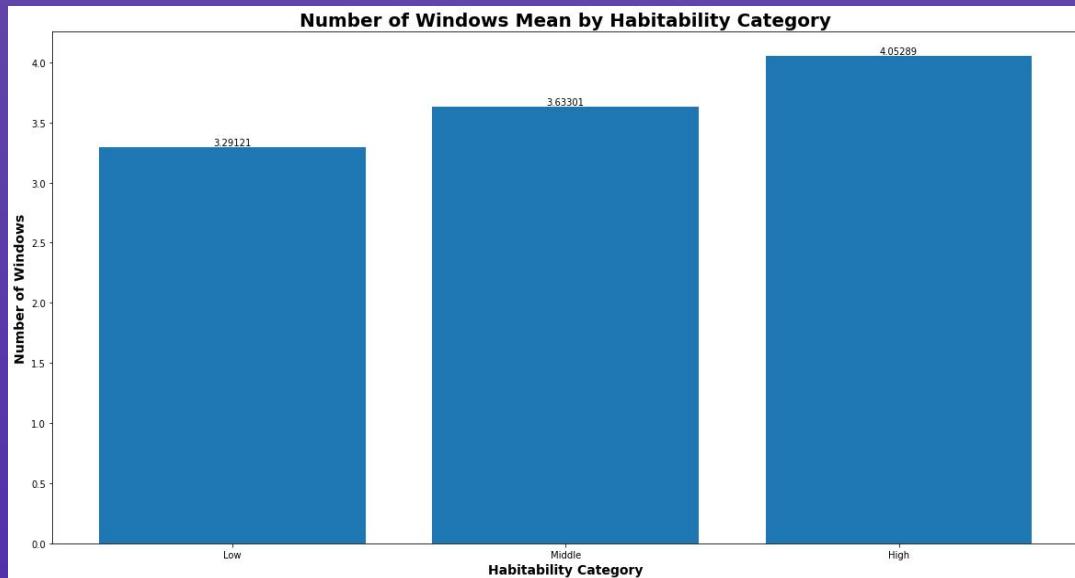
1267,5

Habitability Low

1044,5

Luas area tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High" memiliki rata-rata yang paling **tinggi** dibandingkan kategori lainnya.

Visualisasi Rata-rata Jumlah Jendela



Habitability High

4,05

Habitability
Medium

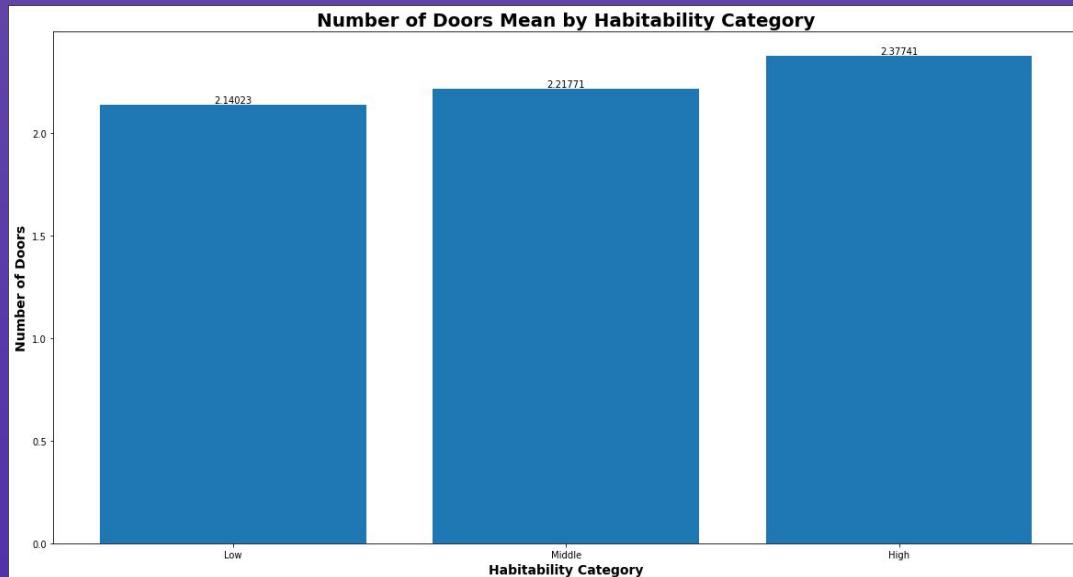
3,63

Habitability Low

3,29

Jumlah jendela tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High" memiliki rata-rata yang paling **tinggi** dibandingkan kategori lainnya.

Visualisasi Rata-rata Jumlah Pintu



Habitability High

2,37

Habitability Medium

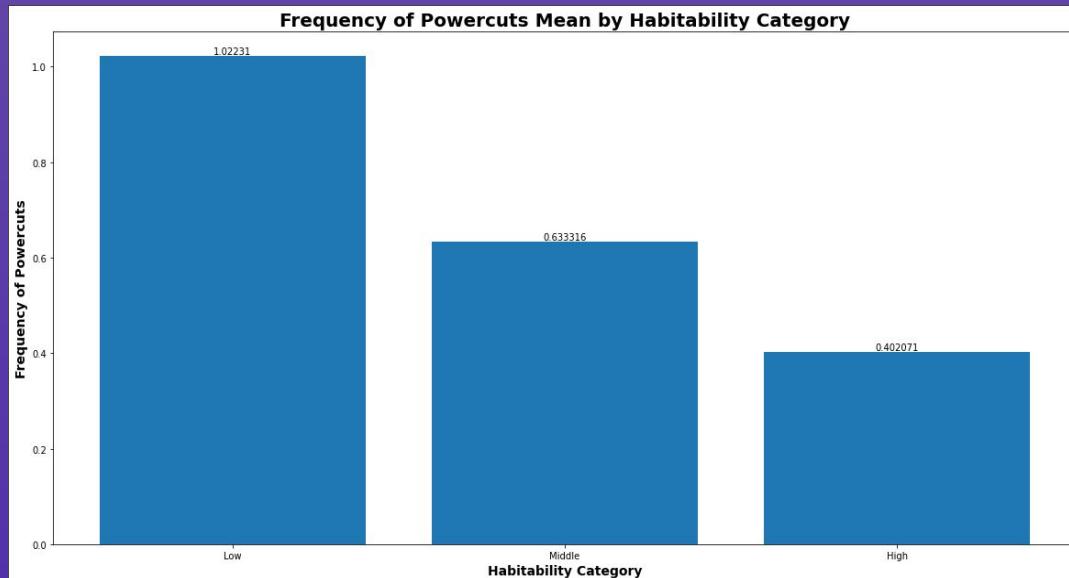
2,21

Habitability Low

2,14

Jumlah pintu tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High" memiliki rata-rata yang paling **tinggi** dibandingkan kategori lainnya.

Visualisasi Rata-rata Frekuensi Powercuts



Habitability High

0,40

Habitability
Medium

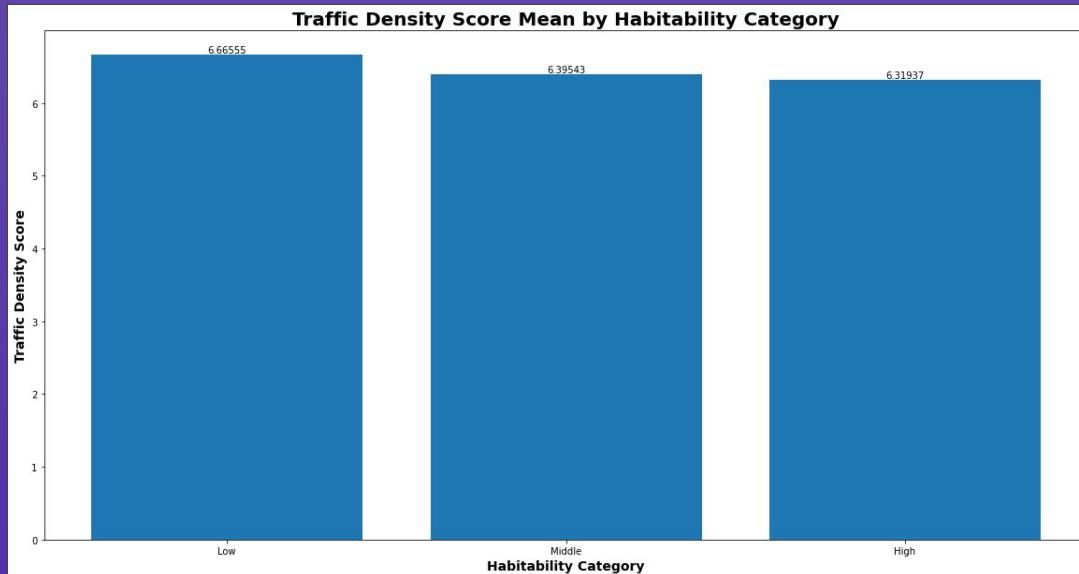
0,63

Habitability Low

1,02

Frekuensi *powercuts* tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High” memiliki rata-rata yang paling **rendah** dibandingkan kategori lainnya.

Visualisasi Rata-rata Traffic Density



Habitability High

Habitability
Medium

Habitability Low

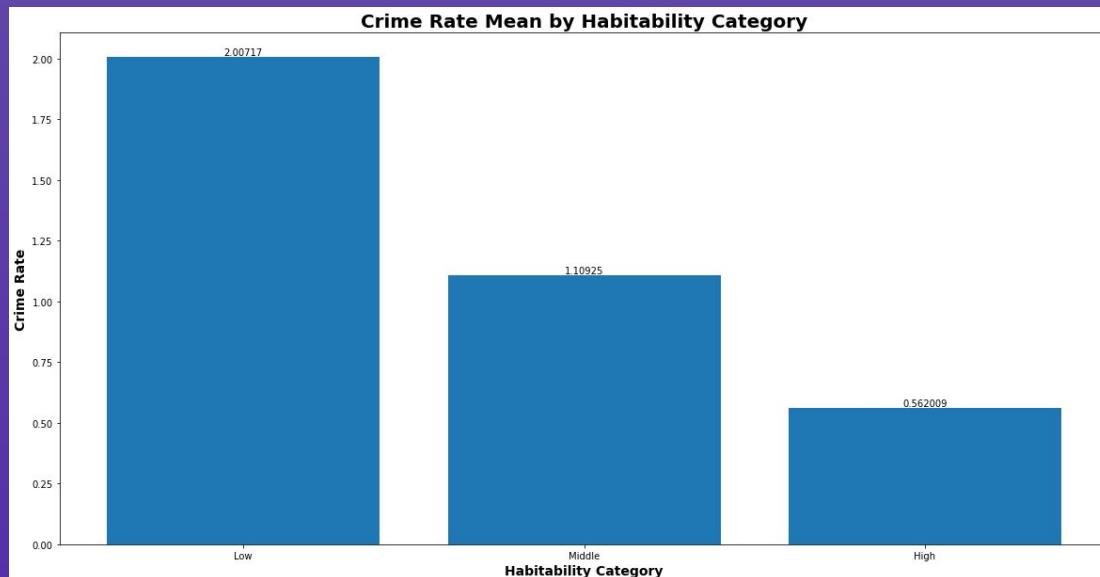
6,31

6,39

6,67

Kepadatan lalu lintas pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High” memiliki rata-rata yang paling **rendah** dibandingkan kategori lainnya.

Visualisasi Rata-rata Crime Rate



Habitability High

0,56

Habitability Medium

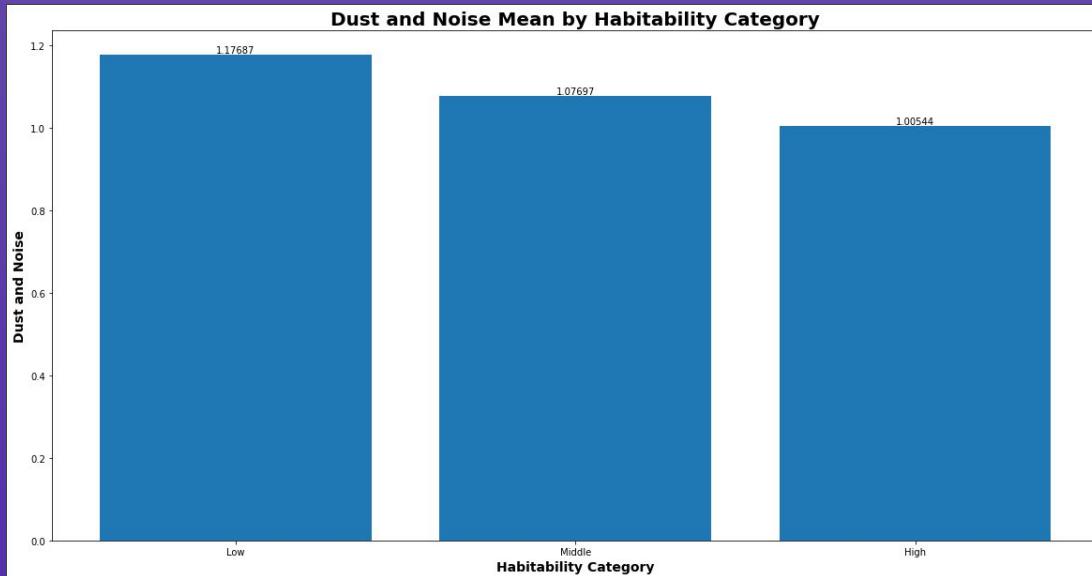
1,10

Habitability Low

2,00

Tingkat kriminalitas pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High" memiliki rata-rata yang paling **rendah** dibandingkan kategori lainnya.

Visualisasi Rata-rata Dust and Noise



Habitability High
Habitability Medium
Habitability Low

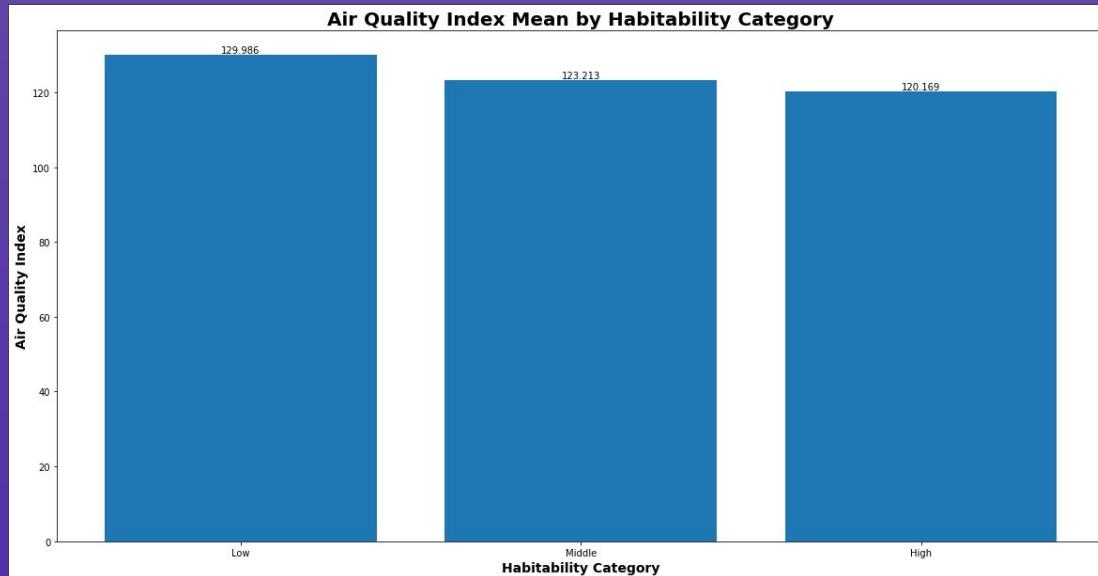
1,00

1,07

1,17

Debu dan kebisingan pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High” memiliki rata-rata yang paling **rendah** dibandingkan kategori lainnya.

Visualisasi Rata-rata Air Quality Index



Habitability High

120,17

Habitability Medium

123,21

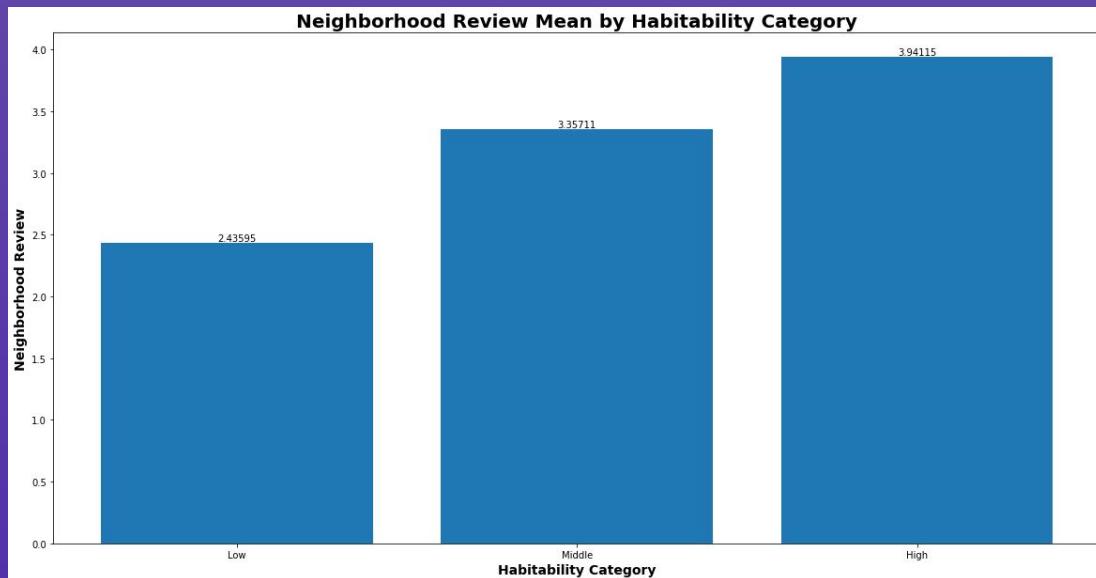
Habitability Low

129,98

Index kualitas udara pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High” memiliki rata-rata yang paling **rendah*** dibandingkan kategori lainnya.

*Index kualitas udara semakin kecil semakin baik

Visualisasi Rata-rata Neighborhood Review



Habitability High

3,94

Habitability Medium

3,35

Habitability Low

2,43

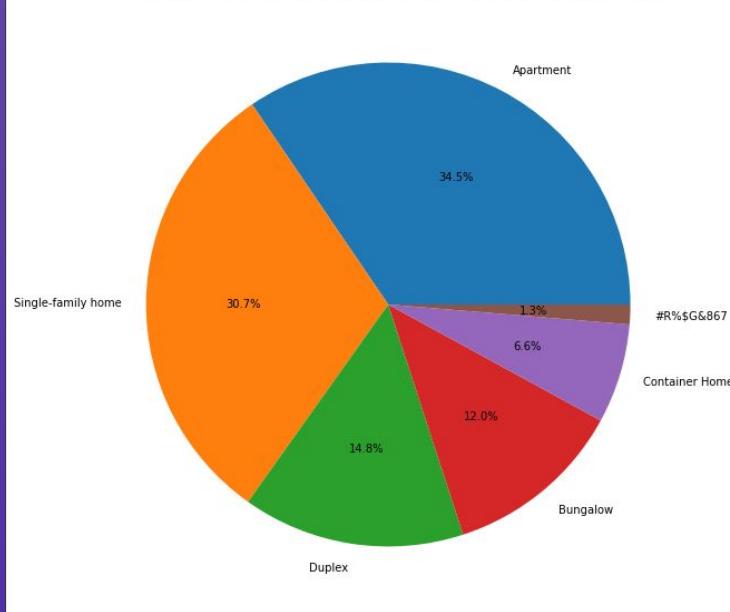
Neighborhood review pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High” memiliki rata-rata yang paling **tinggi** dibandingkan kategori lainnya.



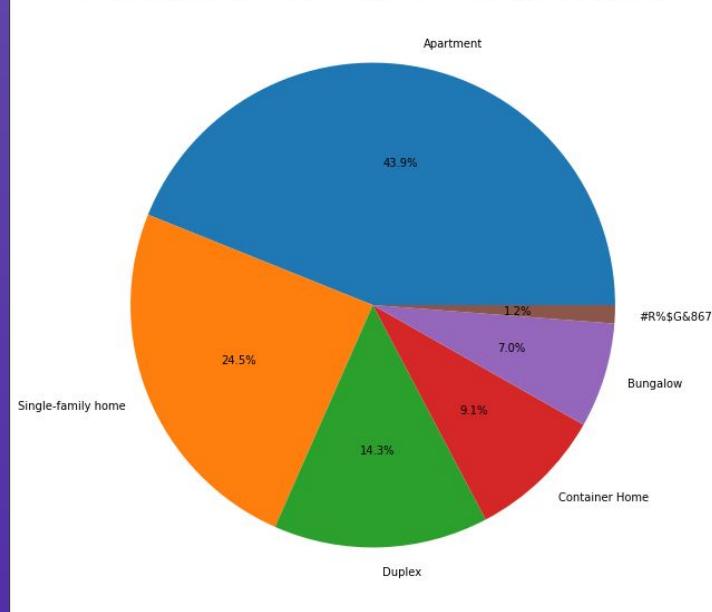
Berdasarkan Fitur Kualitatif

Visualisasi Tipe Properti

Property Type Proportion On High Habitability Category

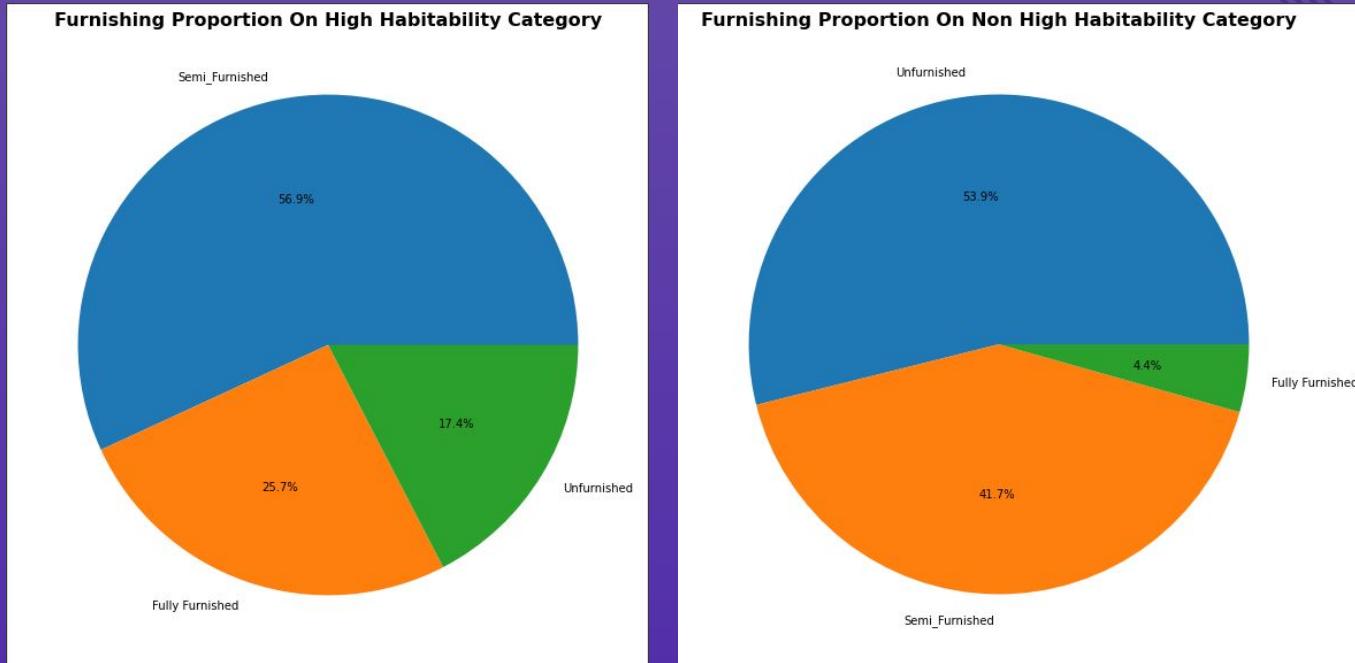


Property Type Proportion On Non High Habitability Category



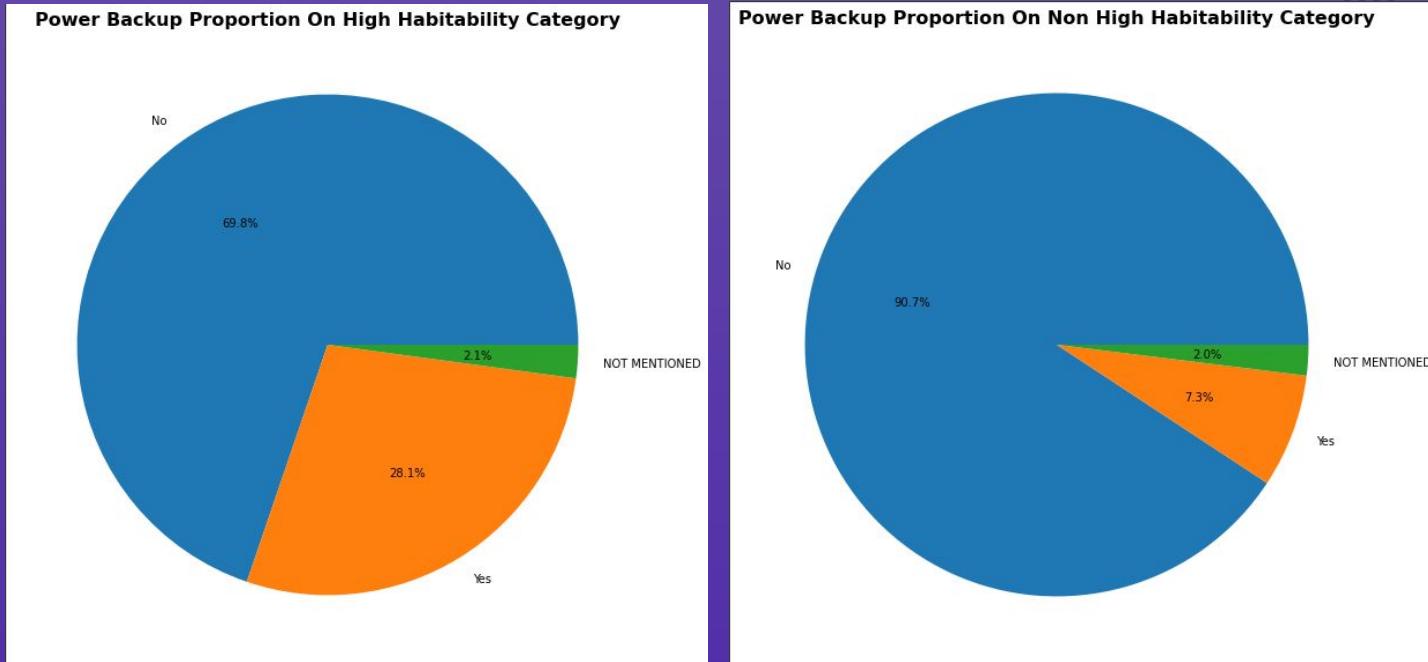
Pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High", mayoritas bertipe apartemen atau rumah keluarga tunggal dengan proporsi 34.5% dan 30.7%.

Visualisasi Furnishing



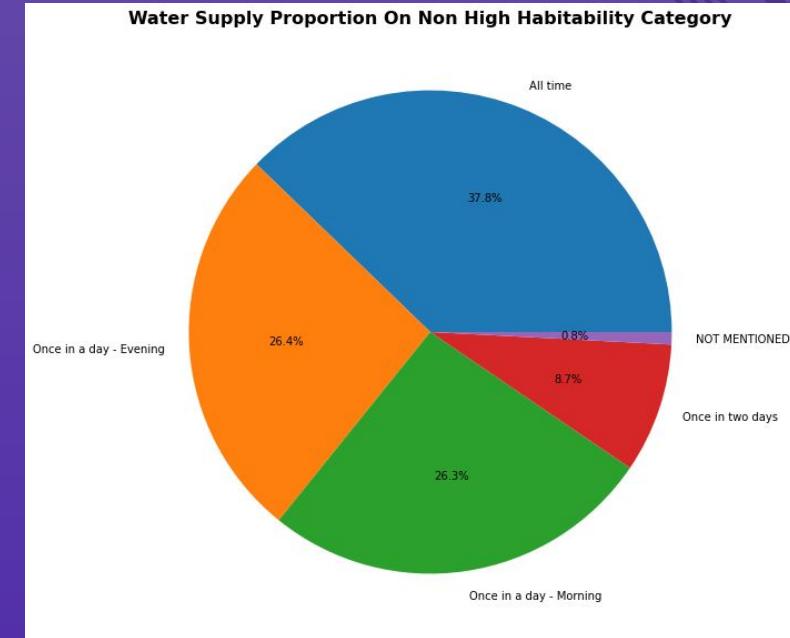
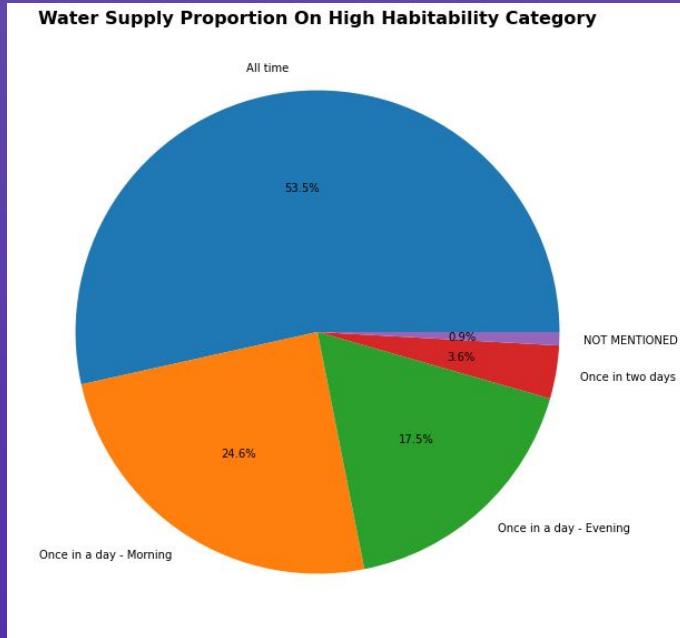
Pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian "High", mayoritas bersifat semi furnished.

Visualisasi Power Backup



Pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High”, mayoritas tidak memiliki power backup. Namun, terdapat lebih banyak rumah yang memiliki power backup dibanding dengan tempat tinggal yang tidak termasuk kategori kelayakhunian “High”

Visualisasi Water Supply



Pada tempat tinggal yang termasuk ke dalam kategori kelayakhunian “High”, mayoritas pasokan airnya sepanjang waktu (*all time*).

Kesimpulan

Ciri-ciri Properti yang Memiliki Habitability_category Bernilai “High”?

- Memiliki rata-rata luas area, jumlah jendela, jumlah pintu, dan *neighborhood review* yang **tinggi**.
- Memiliki rata-rata frekuensi *powercuts*, *traffic density*, *crime rate*, dan *dust & noise* yang **rendah**.
- Mayoritas bersifat **semi furnished**.
- Memiliki pasokan air **sepanjang waktu**.
- Mayoritas **tidak memiliki** power backup
- Mayoritas bertipe **apartemen** atau **rumah keluarga tunggal**

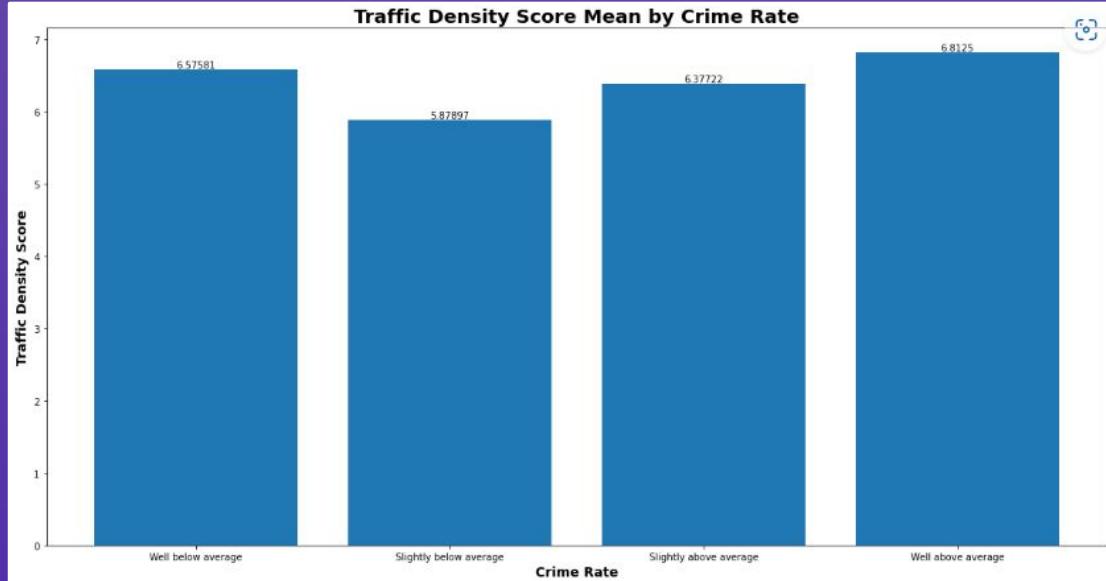


Hubungan antara Traffic_Density_Score dan Crime_Rate

Problem Guideline Nomor 1C



Visualisasi Bar Chart



Well Above Avg

6,81

Slightly Above

6,37

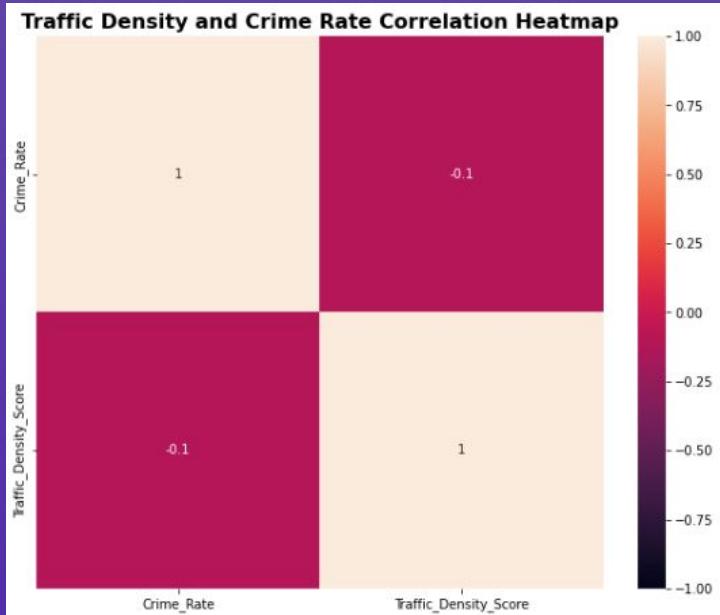
Slightly Below

5,87

Well Below Avg

6,57

Visualisasi Korelasi



Kami juga melakukan analisis korelasi dengan metode kendall setelah memberikan nilai ordinal pada kolom Crime_Rate dan memperoleh hasil $-0,1$

Kesimpulan

Berdasarkan visualisasi sebelumnya, terlihat tingkat kriminalitas memiliki korelasi positif terhadap tingkat kepadatan lalu lintas pada tingkat “Slightly below average” hingga “Well above average”, yaitu semakin tinggi tingkat kriminalitas, semakin tinggi pula rata-rata tingkat kepadatan lalu lintas. Namun, tingkat kriminalitas “Well below average” memiliki tingkat kepadatan lalu lintas yang tinggi sehingga membuat nilai korelasi antara kedua fitur ini lemah.

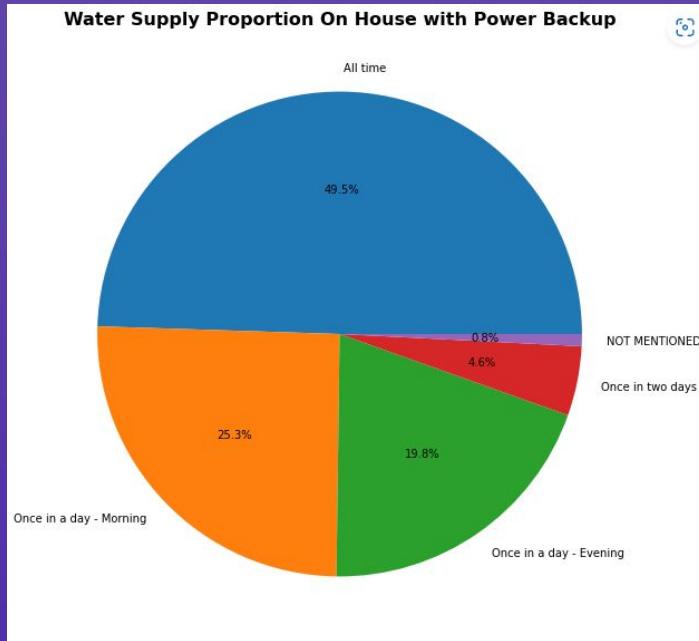


Apakah bagi rumah yang memiliki Power_Backup, terdapat juga persediaan air yang lancar?

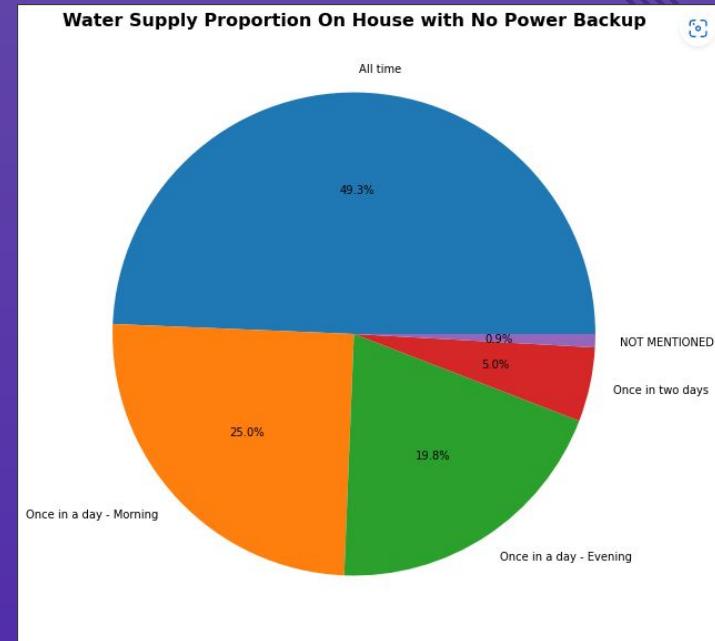
Problem Guideline Nomor 1D



Visualisasi Pasokan Air



Dengan Power Backup



Tanpa Power Backup

Kesimpulan

Berdasarkan visualisasi sebelumnya, terlihat rumah yang memiliki cadangan listrik terdapat persediaan air yang lancar. Namun, apabila dilihat perbedaannya dengan rumah yang tidak memiliki cadangan listrik, tidak terlalu signifikan. Terlihat rumah dengan cadangan listrik 49,5% dan tanpa cadangan listrik 49,3%.

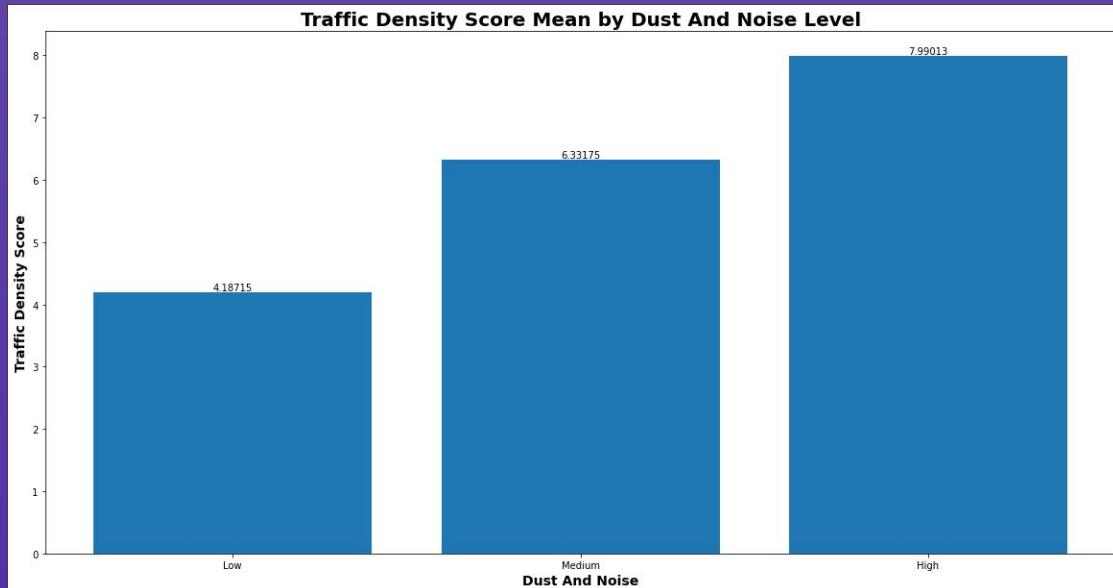


Hubungan antara Traffic_Density_Score dan Dust_and_Noise

Eksplorasi Mandiri 1



Visualisasi Traffic Density Score



Dust & Noise High

7,99

Dust & Noise
Medium

6,33

Dust & Noise Low

4,18

Rata-rata traffic density score dari pemukiman berdasarkan Dust and Noise Level mengikuti tingkat keparahannya.

Kesimpulan

Berdasarkan visualisasi sebelumnya, terlihat semakin tinggi rata-rata kepadatan lalu lintas, membuat debu dan kebisingan pada daerah tersebut termasuk ke dalam kategori high. Sedangkan, semakin rendah rata-rata kepadatan lalu lintas, membuat debu dan kebisingan pada daerah tersebut termasuk ke dalam kategori low. Dengan demikian, terdapat hubungan antara traffic density score dan dust & noise.

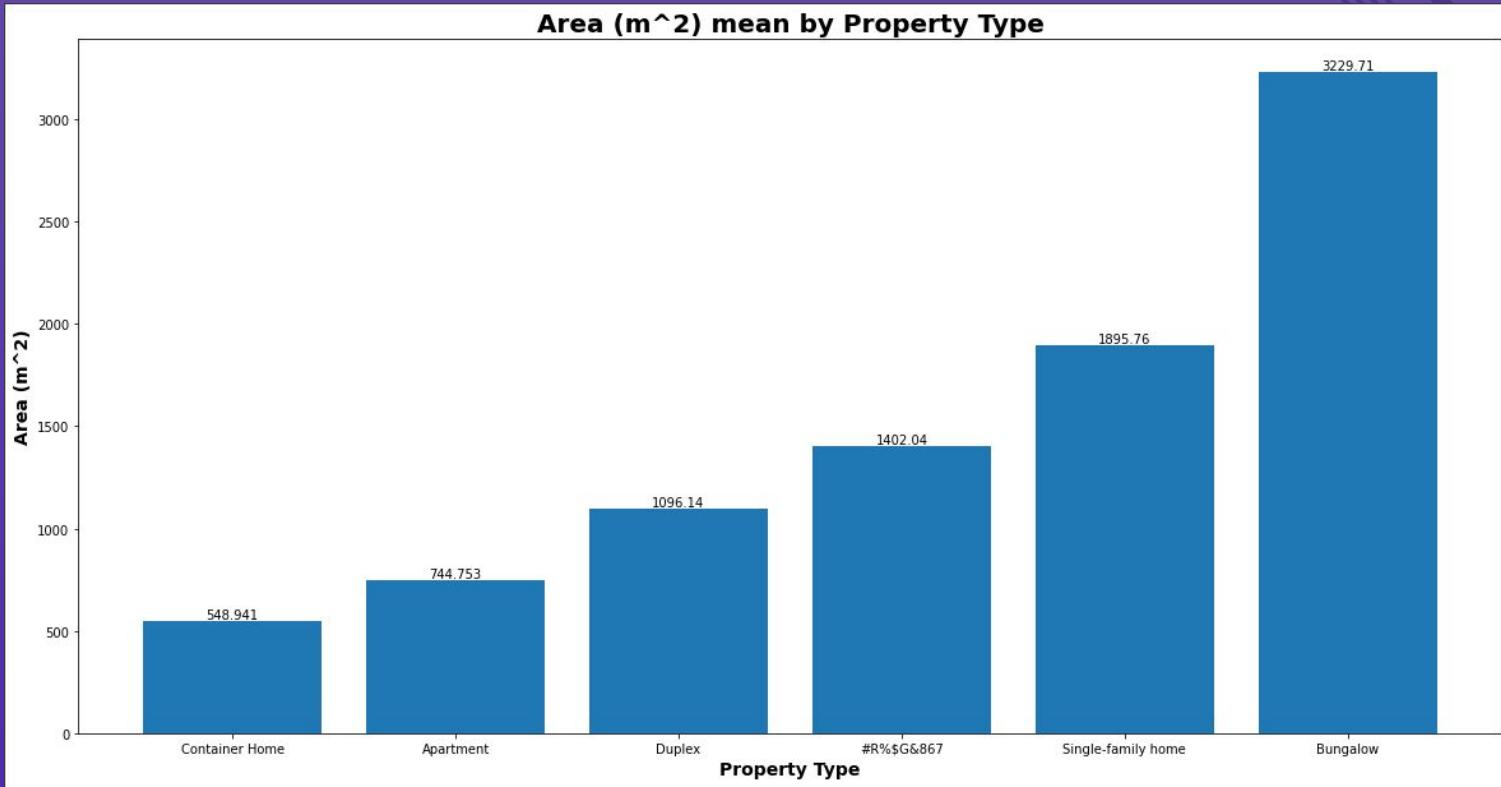


Apakah tipe properti dengan rata-rata luas area tertinggi?

Eksplorasi Mandiri 2



Visualisasi Luas Area Tipe Properti



Kesimpulan

Berdasarkan visualisasi sebelumnya, tipe properti dengan luas area tertinggi adalah Bungalow yang kemudian disusul dengan Single-Family Home. Dengan demikian, apabila dalam pencarian tempat tinggal seseorang lebih mengutamakan luas area, maka pilihan yang tepat adalah Bungalow.

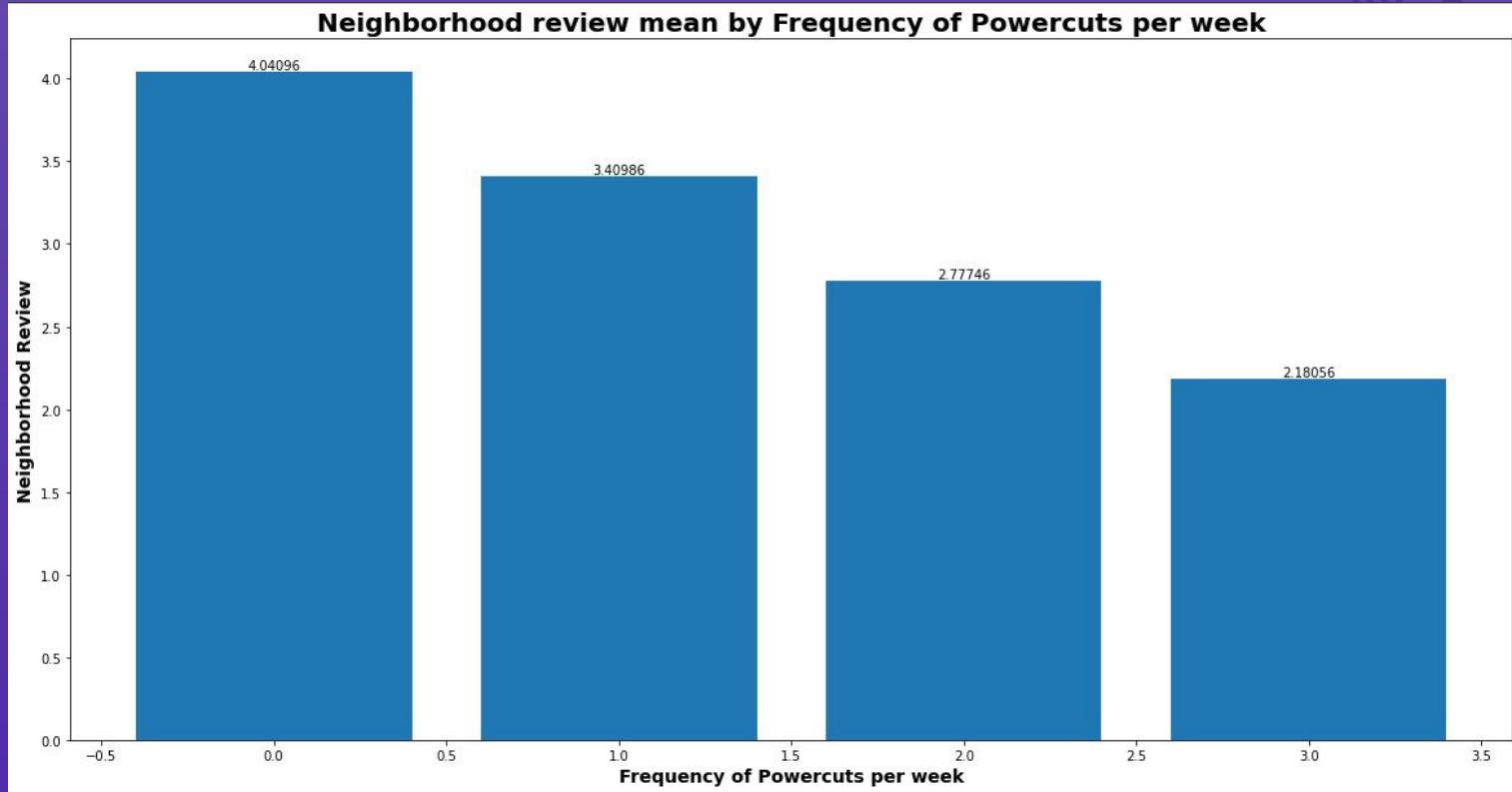


Apakah terdapat hubungan antara Neighborhood_Review dan Frequency_of_Powercuts

Eksplorasi Mandiri 3



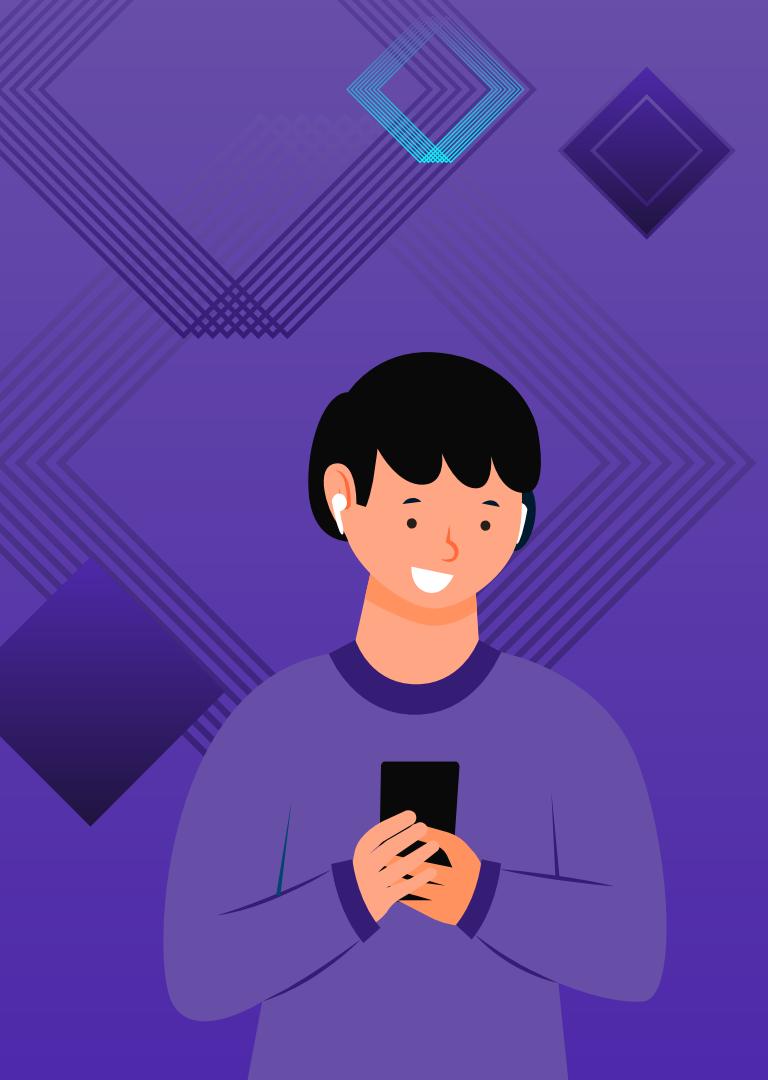
Visualisasi Neighborhood Review



Kesimpulan

Berdasarkan visualisasi sebelumnya, terdapat hubungan antara neighborhood review dan frekuensi dari mati listrik. Hubungan tersebut menyatakan bahwa semakin **jarang** suatu tempat tinggal mati listrik, maka nilai neighborhood review semakin **tinggi**. Sedangkan, semakin **sering** suatu tempat tinggal mati listrik, maka nilai neighborhood review semakin **rendah**.





03

Preprocessing Data

Melakukan pembersihan data

Missing & Inconsistent Values



Inconsistent Value

Terdapat inconsistent value pada kolom Property_Type, Power_Backup, dan Water_Supply yang diubah menjadi null.

```
data['Property_Type'].replace({'#R%$G&867': np.NaN}, inplace=True)
data['Power_Backup'].replace({'NOT MENTIONED': np.NaN}, inplace=True)
data['Water_Supply'].replace({'NOT MENTIONED': np.NaN}, inplace=True)
```

Pemeriksaan Missing Value pada Setiap Fitur

Terdapat **7139 baris** yang mengandung missing value

	Total	Percent
Number_of_Windows	1654	0.041874
Frequency_of_Powercuts	1383	0.035014
Dust_and_Noise	1219	0.030862
Furnishing	1042	0.026380
Power_Backup	828	0.020963
Crime_Rate	787	0.019925
Property_Type	510	0.012912
Water_Supply	345	0.008734



Penanganan Missing Values

39.499

Jumlah Baris Keseluruhan

32.360

Jumlah Baris Setelah Dikurangi

7.139

Jumlah Baris Dengan Missing Value

Jumlah row tersebut masih cukup banyak sebagai data masukan. Dengan demikian, kami memutuskan untuk melakukan drop pada missing value



Pemeriksaan Outliers

Pemeriksaan Outliers pada Setiap Fitur

```
Air_Quality_Index      623
Crime_Rate             0
Dust_and_Noise         0
Frequency_of_Powercuts 1009
Furnishing              0
Habitability_category   0
Neighborhood_Review     552
Number_of_Doors          0
Number_of_Windows        1717
Power_Backup              0
Property_Area            1370
Property_Type              0
Traffic_Density_Score     431
Water_Supply                0
dtype: int64
```

Penanganan

- **Air_Quality_Index:** Normal karena kualitas udara suatu wilayah bisa saja sangat baik dan sangat buruk, terutama pada wilayah dengan tingkat polusi tinggi
- **Frequency_of_Powercuts:** Data diskrit dengan unique value 4 sehingga bukan merupakan outlier
- **Neighborhood_Review:** Normal jika masyarakat memberikan nilai rendah atau tinggi pada suatu wilayah.
- **Number_of_Windows:** Normal sebuah tempat tinggal memiliki banyak jendela, terutama pada tempat tinggal dengan area luas
- **Property_Area:** Normal terdapat perbedaan area yang cukup jauh karena perbedaan tipe properti
- **Traffic_Density_Area:** Normal sebuah wilayah memiliki kepadatan lalu lintas yang sangat tinggi karena mobilitas penduduknya.





Pemeriksaan Duplikasi Data

0

Tidak Ditemukan Duplikasi Data

```
print('Jumlah duplikasi pada data adalah', df.duplicated(keep=False).sum())
```

```
Jumlah duplikasi pada data adalah 0
```



Penanganan Data Kategorikal

Ordinal Encoding

```
furnishing = {"Unfurnished" : 0, "Semi_Furnished" : 1, "Fully Furnished": 2}
Crime_Rate = {"Well below average" : 0, "Slightly below average" : 1, "Slightly above average" : 2, "Well above average" : 3}
Dust_and_Noise = {"Low" : 0, "Medium" : 1, "High" : 2}
Habitability_category = {"Low" : 0, "Middle" : 1, "High" : 2}

df_ordinal = df.copy(deep=True)

df['Furnishing'].replace(furnishing, inplace=True)
df['Crime_Rate'].replace(Crime_Rate,inplace=True)
df['Dust_and_Noise'].replace(Dust_and_Noise,inplace=True)
df['Habitability_category'].replace(Habitability_category,inplace=True)
```

```
df = pd.get_dummies(df, drop_first=True)
```

One hot encoding

Kami melakukan encoding untuk setiap fitur yang bersifat kategorikal.

04

Modelling

Klasifikasi, Regresi, Clustering





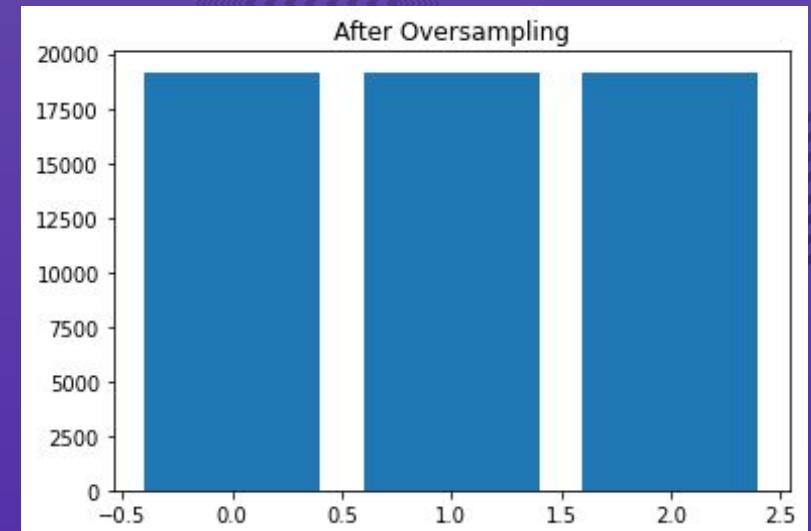
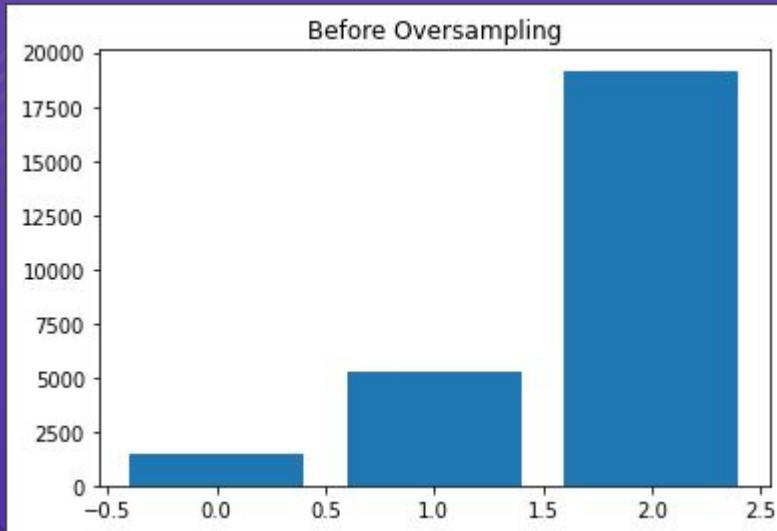
Model Klasifikasi

Scaling

```
scaler = StandardScaler()  
X_train_scaled = scaler.fit_transform(X_train)  
X_test_scaled = scaler.transform(X_test)  
  
X_train = pd.DataFrame(X_train_scaled, index=X_train.index, columns=X_train.columns)  
X_test = pd.DataFrame(X_test_scaled, index=X_test.index, columns=X_test.columns)
```

Kami melakukan scaling menggunakan StandardScaler untuk setiap fitur yang bersifat numerikal.

Data tidak balance oleh karena itu dilakukan **balancing data** menggunakan oversampling **SVM SMOTE**.



Model Selection



Random Forest

Menghasilkan F1 Score Mean sebesar 0,919 dan Standard Variance 0,030



Logistic Regression

Menghasilkan F1 Score Mean sebesar 0,828 dan Standard Variance 0,024



SVC

Menghasilkan F1 Score Mean sebesar 0,849 dan Standard Variance 0,031

Evaluasi Model Menggunakan Laporan Klasifikasi

Model **Random Forest Classifier**

Hasil Evaluasi berdasarkan classification report

	precision	recall	f1-score	support
0	0.74	0.82	0.78	369
1	0.58	0.63	0.61	1312
2	0.92	0.89	0.90	4791
accuracy			0.83	6472
macro avg	0.75	0.78	0.76	6472
weighted avg	0.84	0.83	0.84	6472

Confusion Matrix

prediction		0	1	2
actual	0	303	66	0
1	108	828	376	
2	0	531	4260	

Butuh informasi lebih lengkap? silakan simak di bawah ini :

F1 Macro Average: 0.7619173628439878

F1 Micro Average: 0.832972805933251

Precision Macro Average: 0.7457248362020282

Precision Micro Average: 0.832972805933251

Recall Macro Average: 0.7804676536120402

Recall Micro Average: 0.832972805933251

Evaluasi Model Menggunakan Laporan Klasifikasi

Model **Logistic Regression**

Hasil Evaluasi berdasarkan classification report

	precision	recall	f1-score	support
0	0.71	0.94	0.81	369
1	0.52	0.75	0.61	1312
2	0.95	0.82	0.88	4791
accuracy			0.81	6472
macro avg	0.73	0.83	0.77	6472
weighted avg	0.85	0.81	0.82	6472

Confusion Matrix

prediction	0	1	2	
actual	0	346	23	0
0	346	23	0	
1	139	982	191	
2	1	878	3912	

Butuh informasi lebih lengkap? silakan simak di bawah ini :

F1 Macro Average: 0.767920462047202

F1 Micro Average: 0.8096415327564895

Precision Macro Average: 0.728963694666683

Precision Micro Average: 0.8096415327564895

Recall Macro Average: 0.8342253273555486

Recall Micro Average: 0.8096415327564895

Evaluasi Model Menggunakan Laporan Klasifikasi

Model **SVC**

Hasil Evaluasi berdasarkan classification report

	precision	recall	f1-score	support
0	0.70	0.90	0.79	369
1	0.54	0.68	0.60	1312
2	0.93	0.85	0.89	4791
accuracy			0.82	6472
macro avg	0.73	0.81	0.76	6472
weighted avg	0.84	0.82	0.83	6472

Confusion Matrix

prediction	0	1	2	
actual	0	333	36	0
0	333	36	0	
1	137	888	287	
2	5	714	4072	

Butuh informasi lebih lengkap? silakan simak di bawah ini :

F1 Macro Average: 0.7603960230600594

F1 Micro Average: 0.8178306551297899

Precision Macro Average: 0.7257787948438871

Precision Micro Average: 0.8178306551297899

Recall Macro Average: 0.8097317463468937

Recall Micro Average: 0.8178306551297899

Kesimpulan

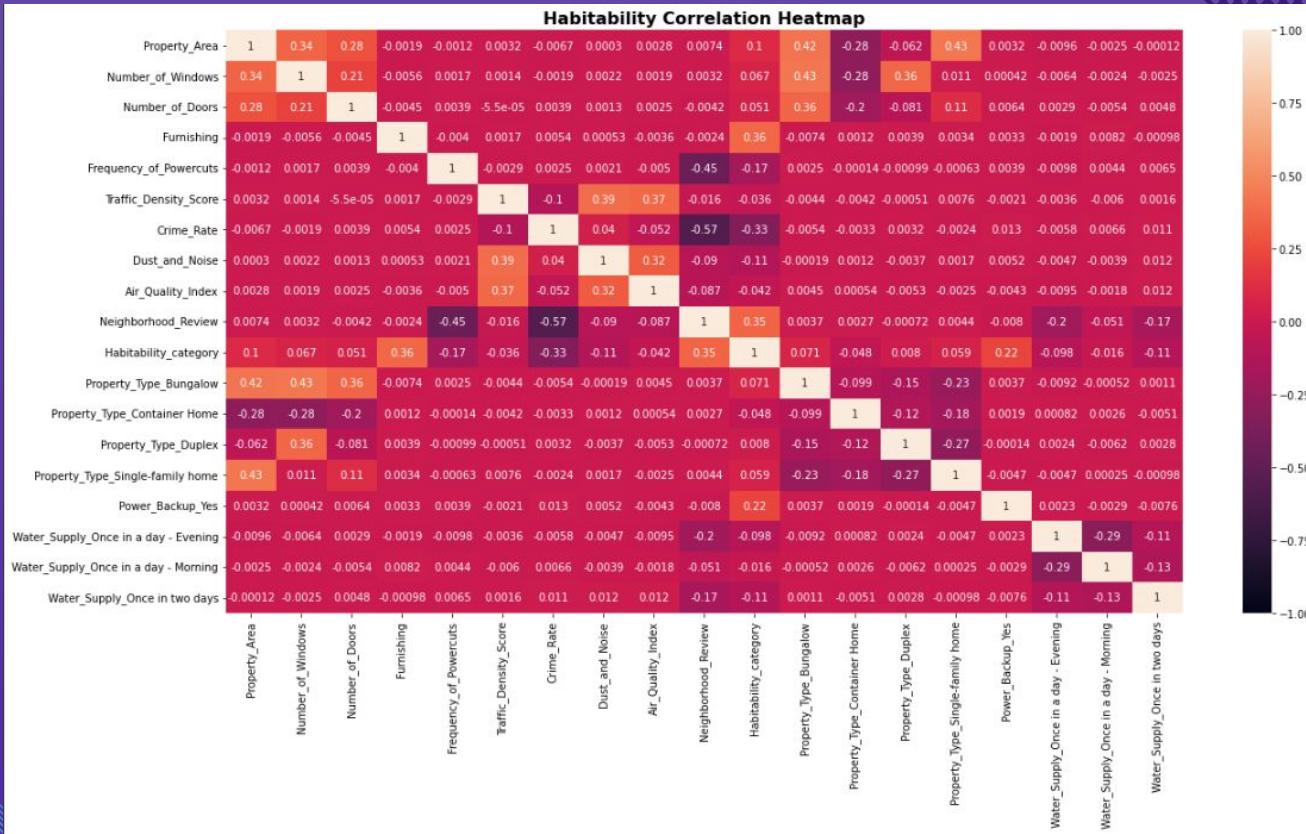
Berdasarkan evaluasi yang telah dilakukan, kami memilih **logistic regression** sebagai model klasifikasi yang paling sesuai untuk data ini dengan memberikan asumsi setiap label memiliki tingkat urgensi yang sama. Hal ini dikarenakan logistic regression memiliki variansi paling rendah pada f1-score setiap label kelas, yaitu **0.019633**.



Model Regresi



Korelasi antar Fitur (Heatmap)



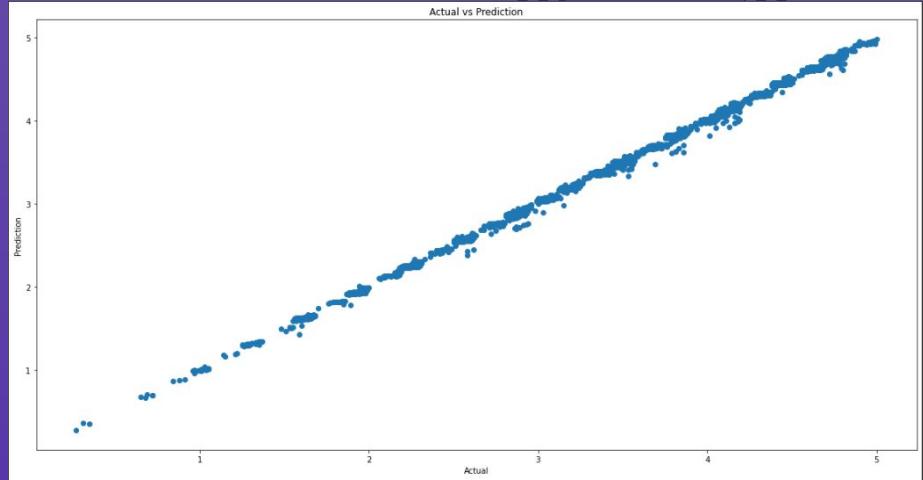
Tidak ada fitur yang memiliki korelasi yang sangat tinggi.

Menggunakan model **LinearRegression**
menghasilkan R Square Mean **0,999** dan
Standard Variance **0,000**

```
model = LinearRegression()
scores = evaluate_classification(X_train, y_train, model)
print('R Square Mean and Standard Variance: %.3f (%.3f)' % (mean(scores), std(scores)))

R Square Mean and Standard Variance: 0.999 (0.000)
```

Evaluasi Model Logistic Regression



Score MAE, MSE, RMSE, R²

```
MAE: 0.021529759136287602
MSE: 0.000810364323871086
RMSE: 0.02846689873995912
R_squared: 0.998740539765254
```

Kesimpulan

Meskipun **tidak** memiliki fitur dengan korelasi yang sangat tinggi terhadap label, model berhasil memprediksi dengan sangat baik. Bahkan, berhasil mencapai nilai **r square 0.99** pada data training dan testing.



Model Clustering



Memilih Fitur Kontinu dan Standarisasi Fitur

Standarisasi menggunakan **StandardScaler**

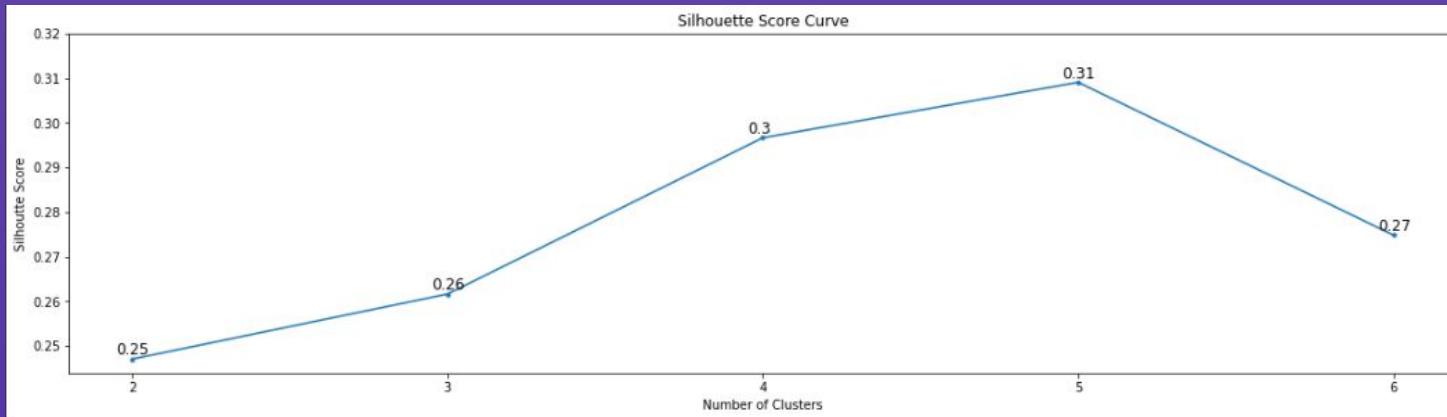
```
scaler = StandardScaler()
df_scaled = scaler.fit_transform(df)

df_scaled = pd.DataFrame(df_scaled, index = df.index, columns = df.columns)
```

```
list_of_cont = []
for i in df.columns:
    if len(df[i].value_counts()) > 20:
        list_of_cont.append(i)
list_of_cont

['Property_Area',
 'Traffic_Density_Score',
 'Air_Quality_Index',
 'Neighborhood_Review']
```

Mencari Jumlah Kluster

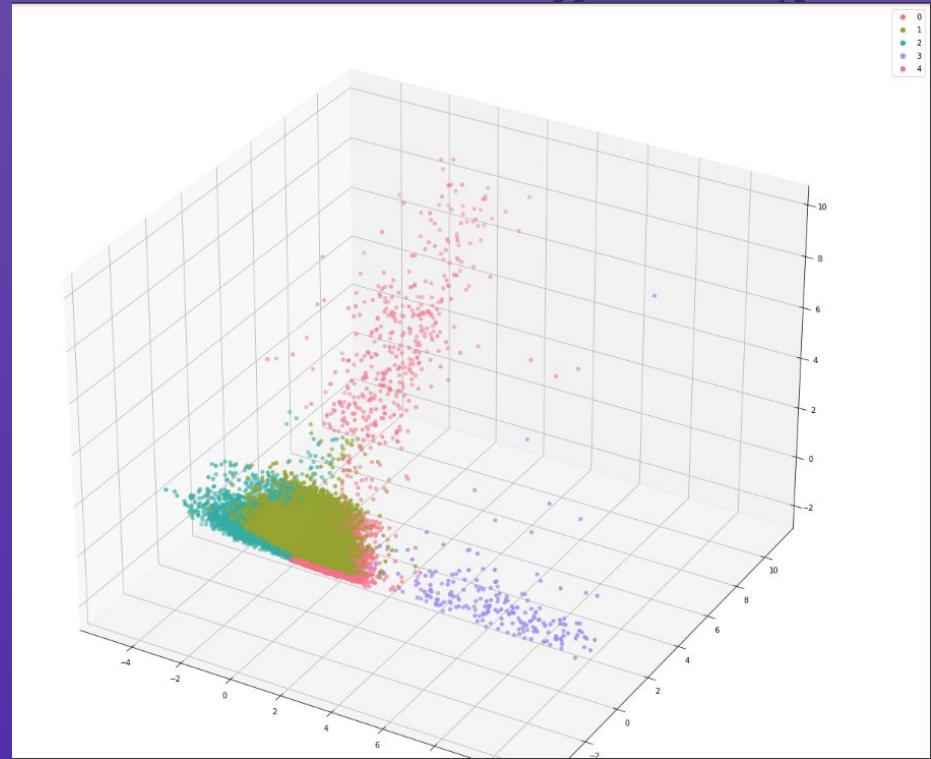


Berdasarkan visualisasi di atas, skor Silhouette **terbaik** didapat dengan **skor 0.31** pada **5 kluster**

Fitting menggunakan **KMeans** dengan jumlah kluster **5**

```
KMean= KMeans(n_clusters=5)
KMean.fit(df_scaled[list_of_cont])
label=KMean.predict(df_scaled[list_of_cont])
```

Visualisasi Klustering menggunakan PCA



Proporsi Kluster

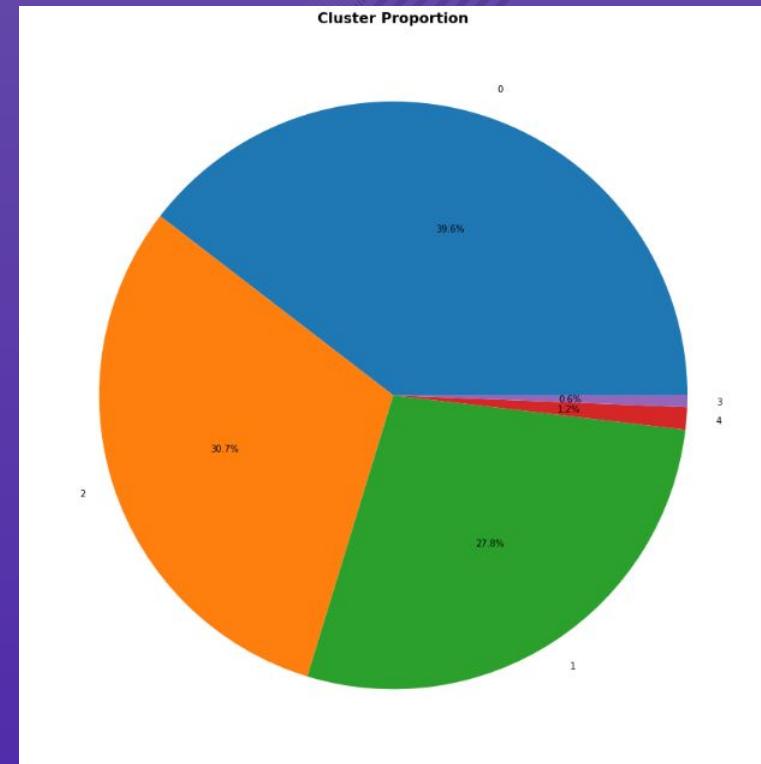
Kluster 0 **39,6%**

Kluster 2 **30,7%**

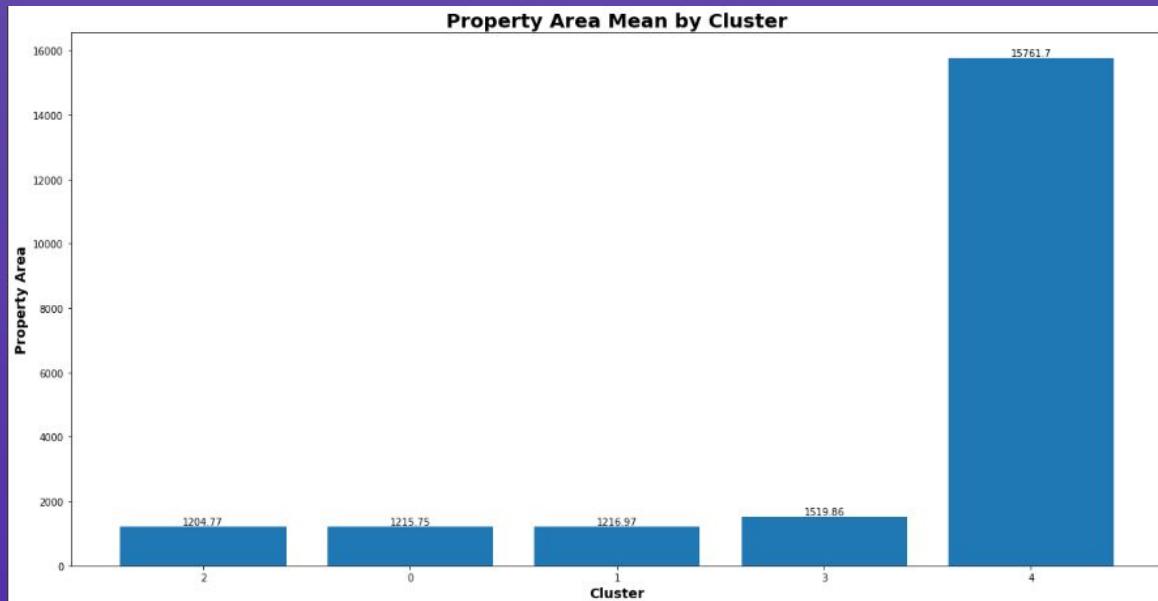
Kluster 1 **27,8%**

Kluster 4 **1,2%**

Kluster 3 **0,6%**



Interpretasi Kluster

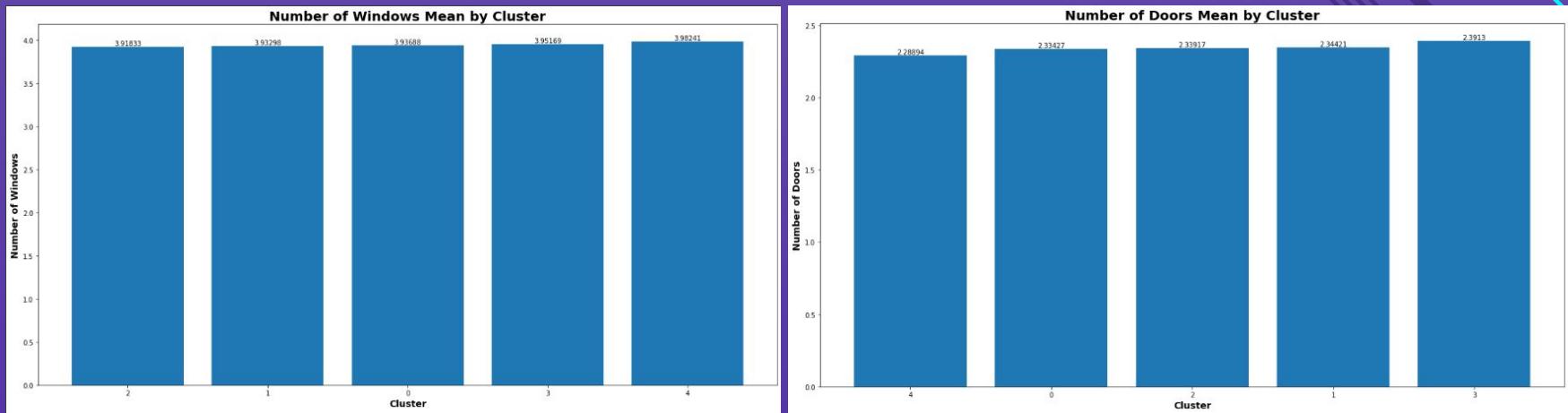


Properti Terluas
Kluster 4

15761,7

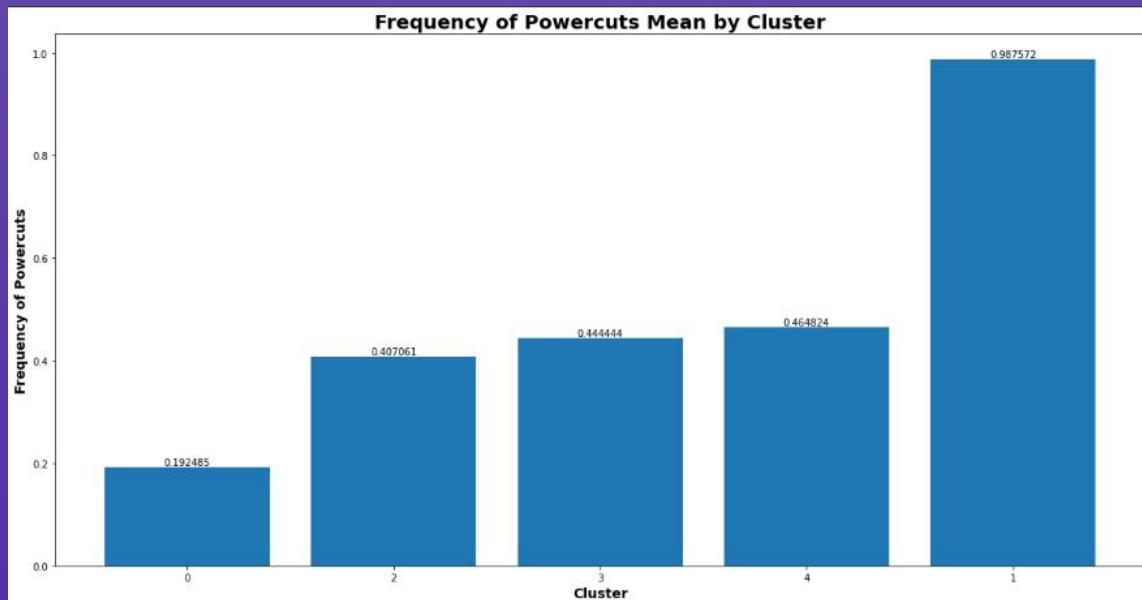
Berdasarkan visualisasi di atas, properti terluas merupakan properti pada kluster 4 dengan rata-rata luas properti 15761,7.

Interpretasi Kluster



Berdasarkan visualisasi di atas, jumlah pintu dan jendela di setiap kluster menunjukkan hasil yang kurang lebih sama.

Interpretasi Kluster



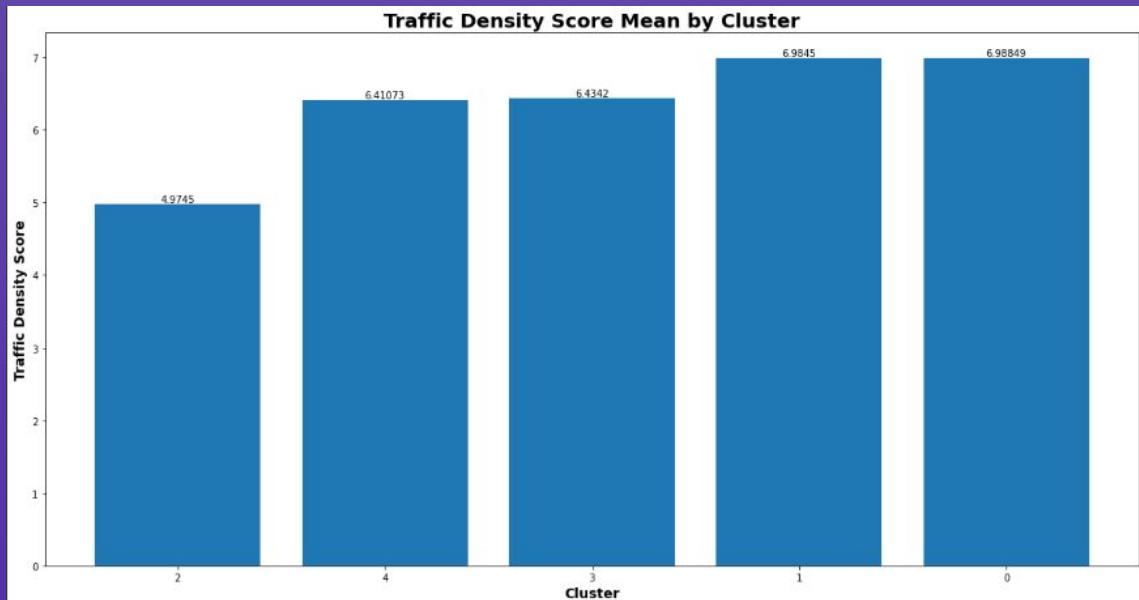
Sering Mati Listrik

Kluster 1

0,987

Berdasarkan visualisasi di atas, properti dengan frekuensi mati listrik tersering adalah properti pada kluster 1 dan terjarang adalah kluster 0.

Interpretasi Kluster



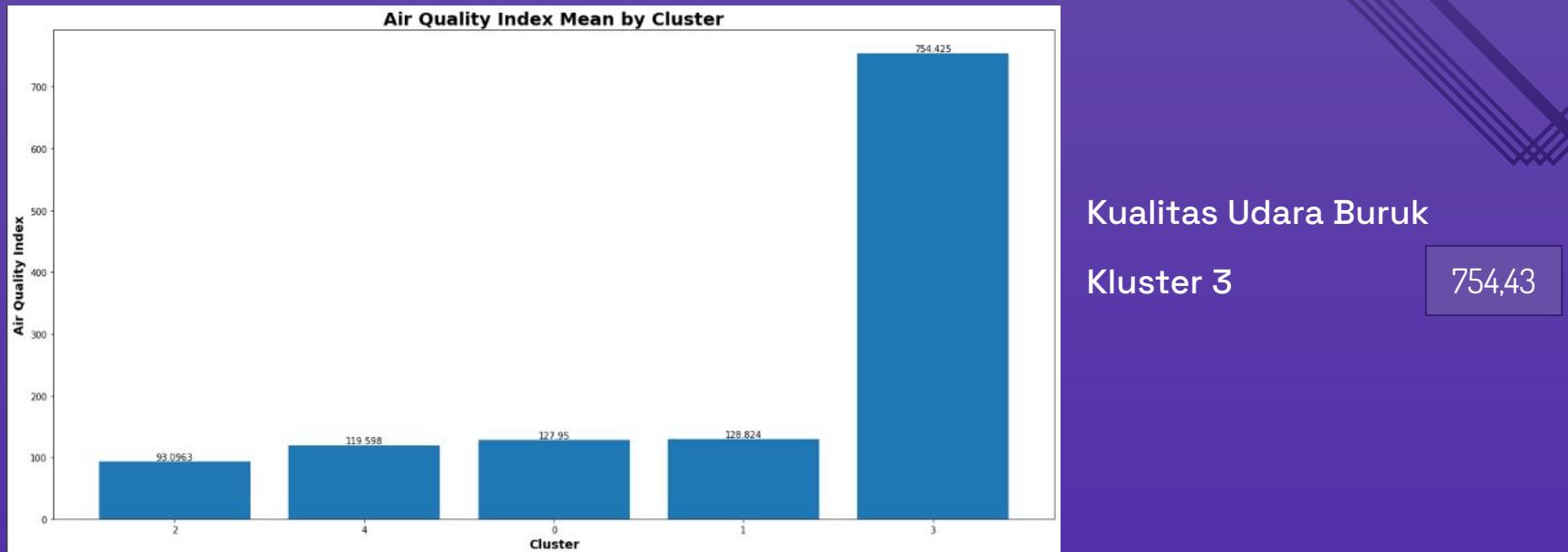
Tidak Padat Lalu Lintas

Kluster 2

4,97

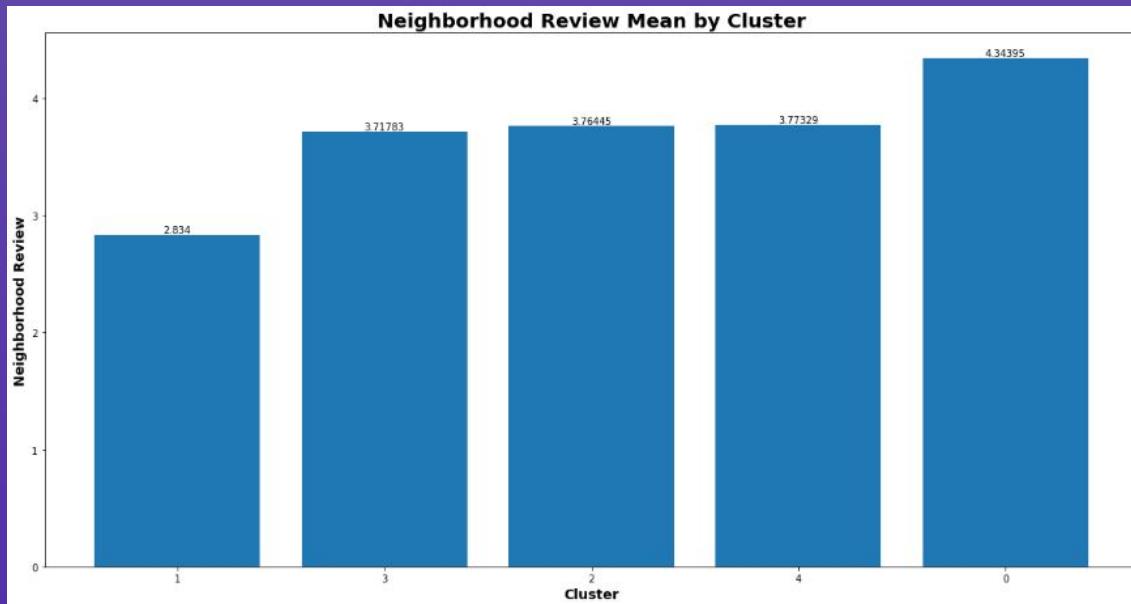
Berdasarkan visualisasi di atas, properti dengan kepadatan lalu lintas terendah adalah properti pada kluster 2.

Interpretasi Kluster



Berdasarkan visualisasi di atas, properti dengan kualitas udara terburuk adalah properti pada kluster 3 dan terbaik properti pada kluster 2.

Interpretasi Kluster

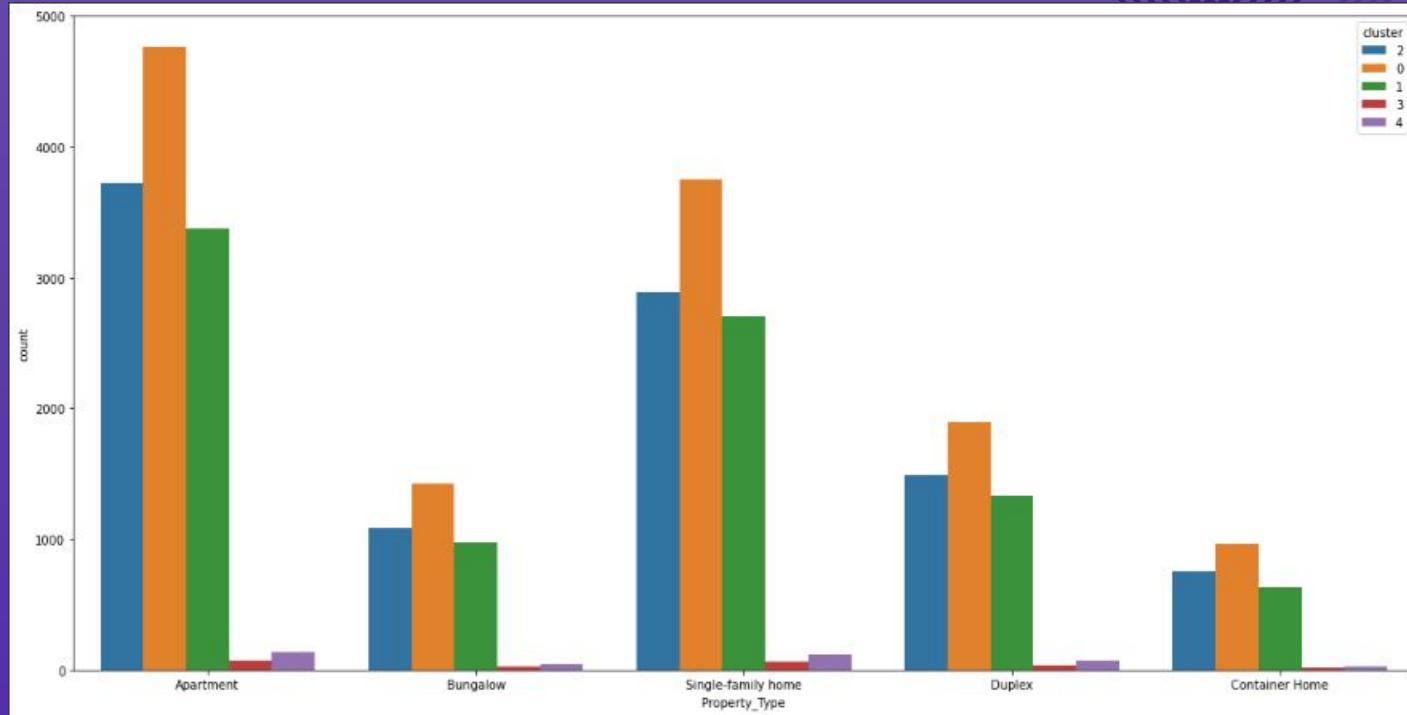


Review Terbaik
Kluster 0

4,34

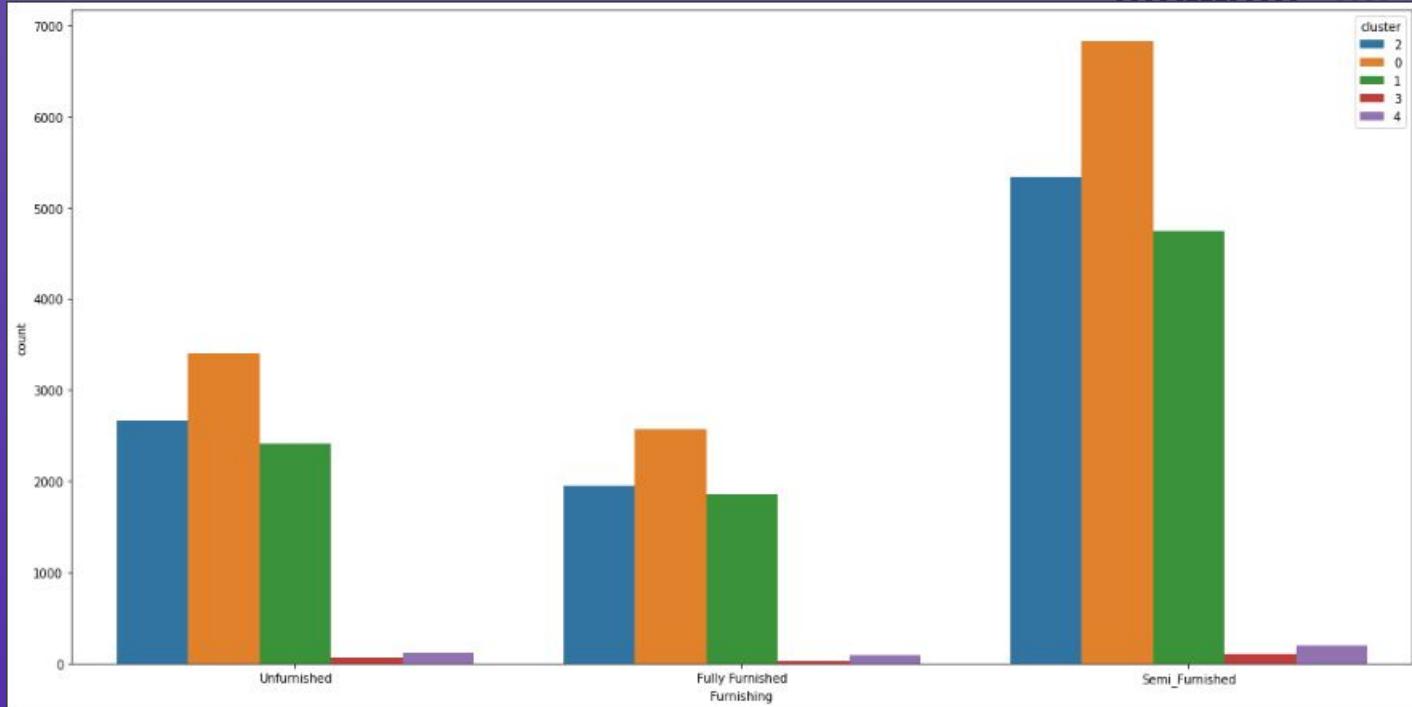
Berdasarkan visualisasi di atas, properti dengan neighborhood review terbaik adalah properti pada kluster 0 dan terburuk properti pada kluster 1.

Interpretasi Kluster Fitur Kategorikal



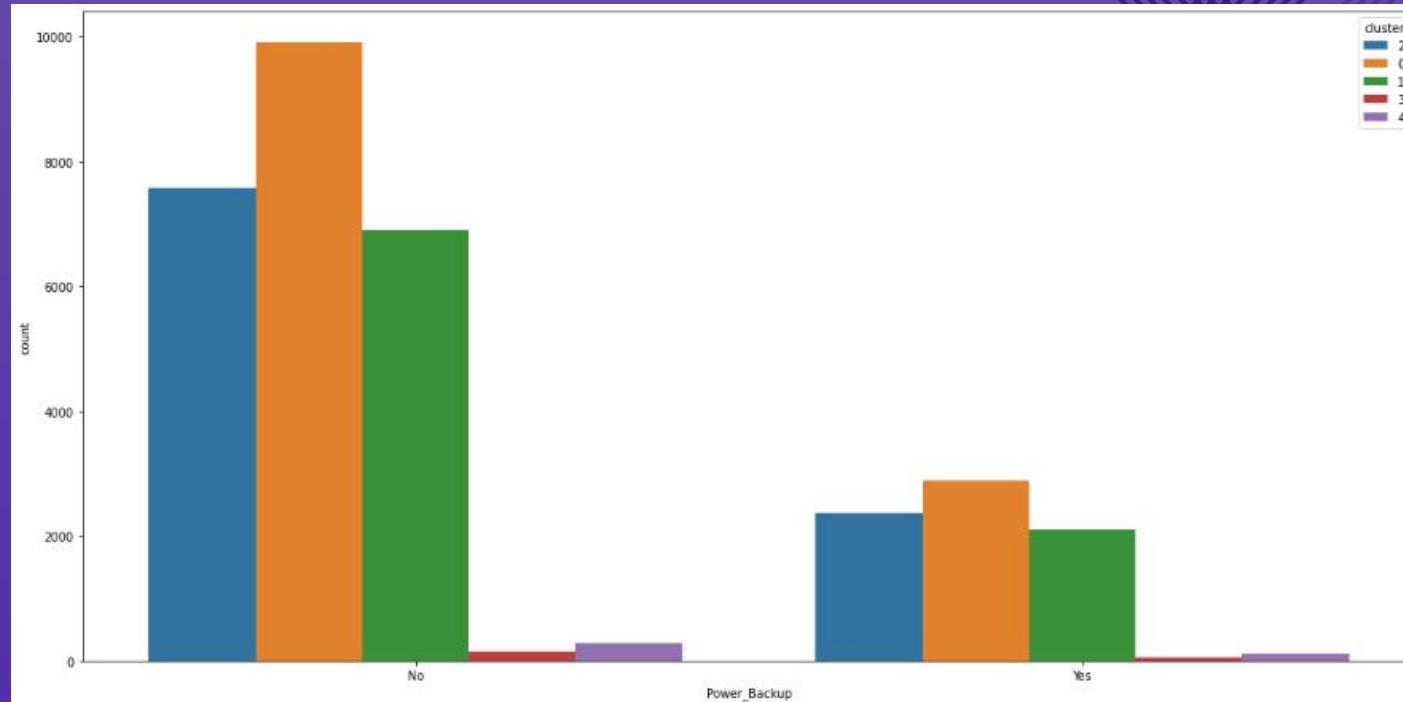
Berdasarkan visualisasi di atas cluster tidak memiliki karakteristik suatu tipe properti karena proporsi cluster untuk tiap tipe properti mirip.

Interpretasi Kluster Fitur Kategorikal



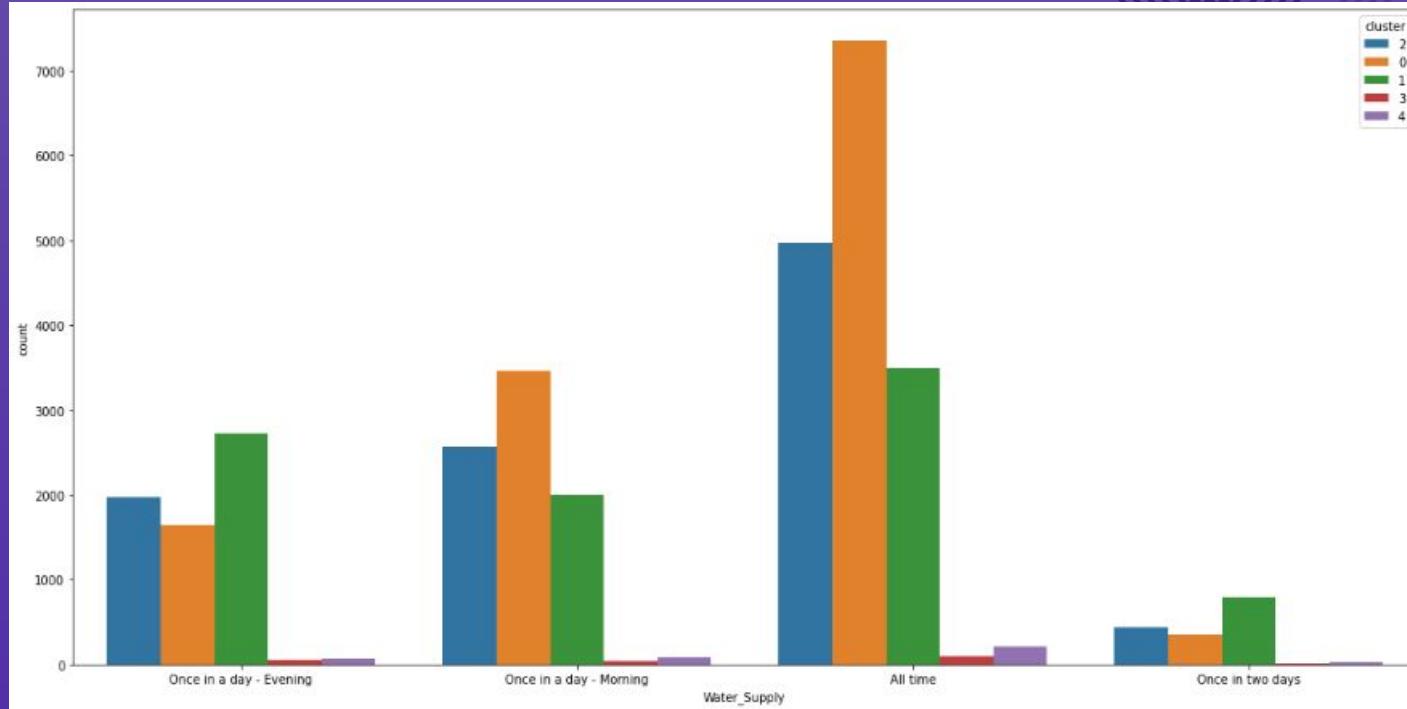
Berdasarkan visualisasi di atas cluster tidak memiliki karakteristik suatu kategori furnishing karena proporsi cluster untuk tiap kategori furnishing mirip.

Interpretasi Kluster Fitur Kategorikal



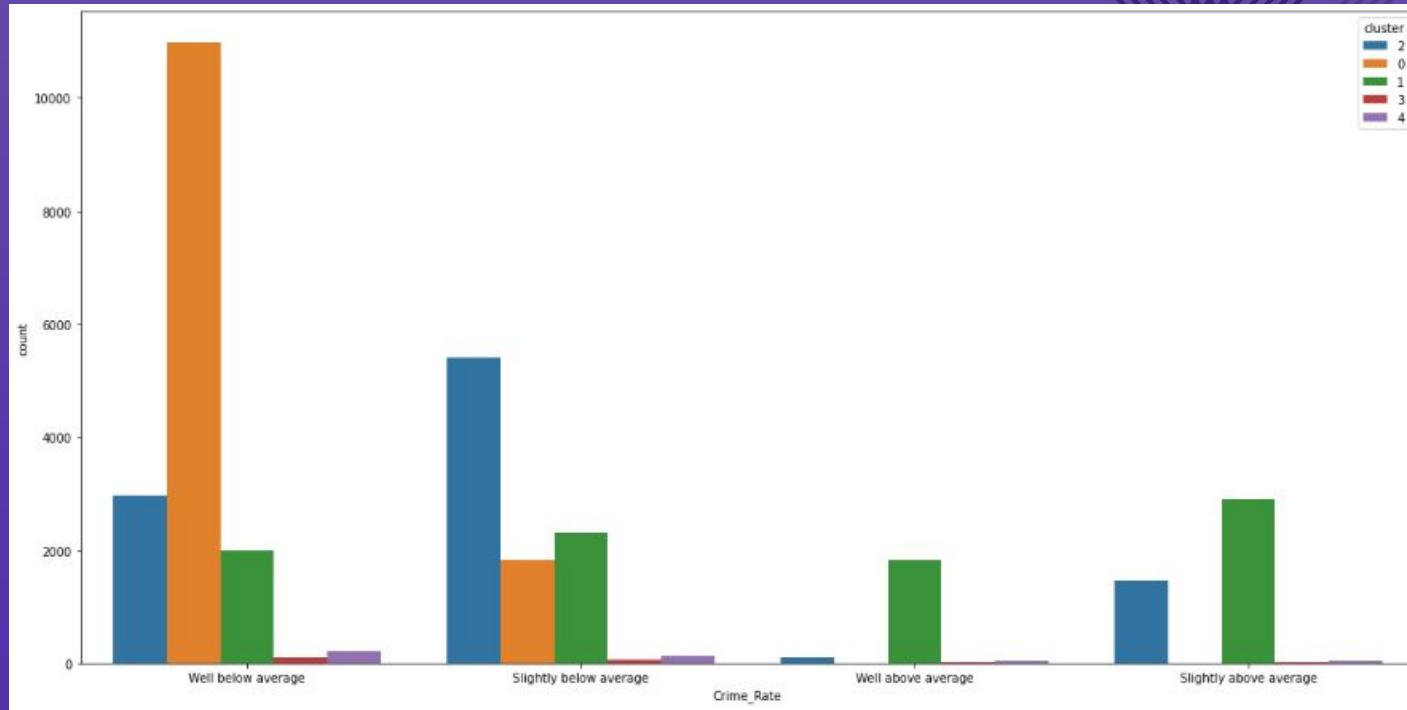
Berdasarkan visualisasi di atas cluster tidak memiliki karakteristik suatu ketersediaan power backup karena proporsi cluster untuk tiap ketersediaan power backup mirip.

Interpretasi Kluster Fitur Kategorikal



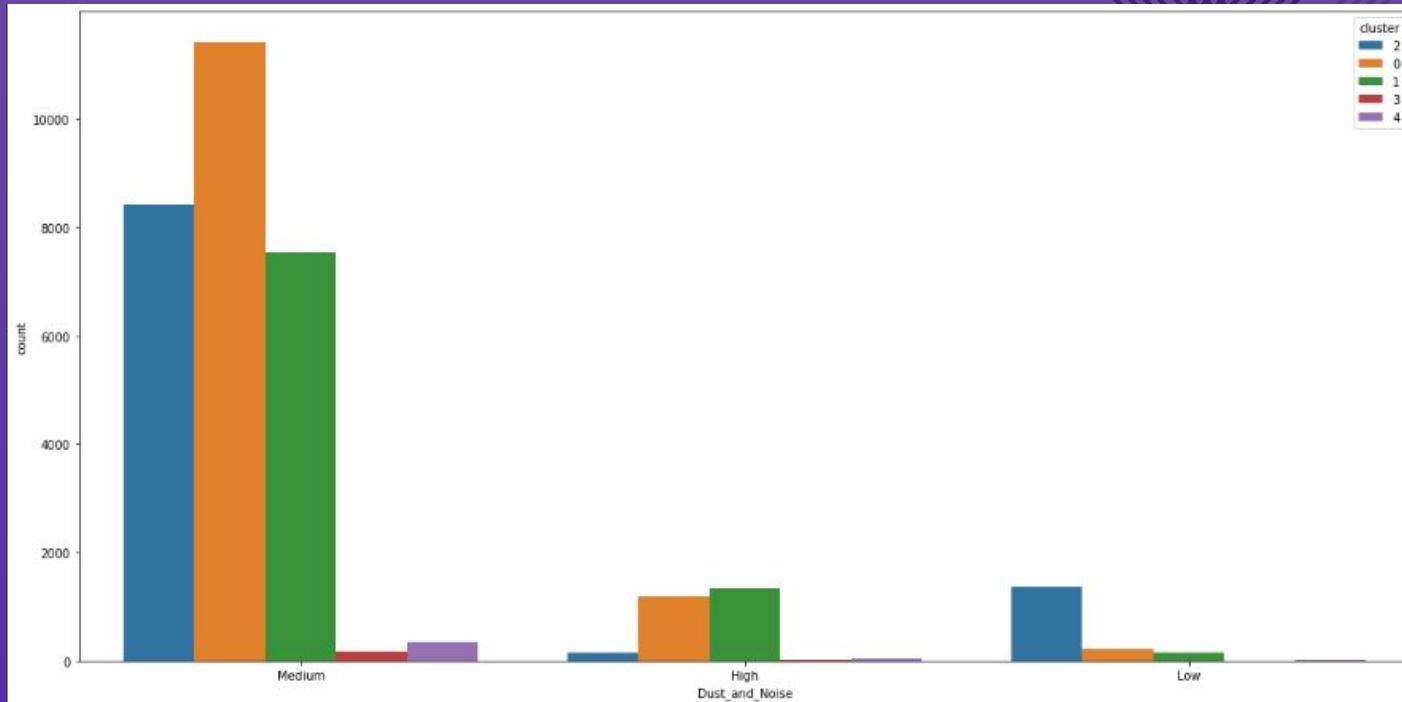
Berdasarkan visualisasi di atas mayoritas kluster 1 memiliki pasokan air sepanjang waktu, tetapi mayoritas kategori pasokan sekali sehari pada malam hari dan sekali tiap dua hari merupakan kluster 1

Interpretasi Kluster Fitur Kategorikal



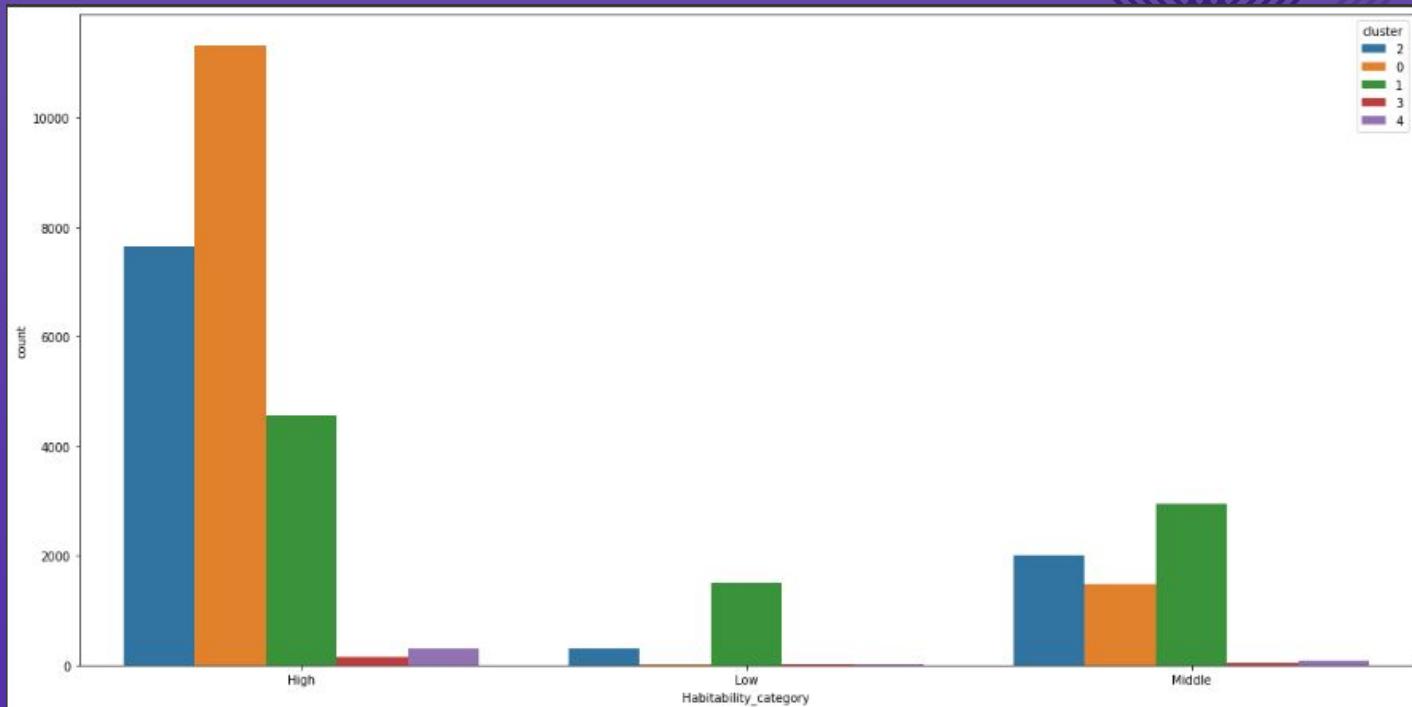
Berdasarkan visualisasi di atas mayoritas tingkat kriminalitas "Well below average" merupakan cluster 0, "Slightly below average" merupakan cluster 2, "Well above average" dan "Slightly above average" merupakan cluster 1

Interpretasi Kluster Fitur Kategorikal



Berdasarkan visualisasi di atas mayoritas tingkat debu dan bising “Medium” merupakan cluster 0, “High” merupakan cluster 1, “Low” merupakan cluster 2

Interpretasi Kluster Fitur Kategorikal



Berdasarkan visualisasi di atas mayoritas kategori kelayakhunian “High” merupakan cluster 0, “Low” dan “Middle” merupakan cluster 1

Thanks!



Do you have any questions?

Kelompok DUAR - KASDD B
Fakultas Ilmu Komputer
Universitas Indonesia



CREDITS: This presentation template was created by Slidesgo, including icons by Flaticon and infographics & images by Freepik

Please keep this slide for attribution