

zenius

Kampus
Merdeka
INDONESIA JAYA

Final Project Presentation

Nomor Kelompok: 5

Nama Mentor: Rachmadio Noval Lazuardi

Nama:

- Azri Ansori Rahman
- Bisma Adhira

Machine Learning Class

Program Studi Independen Bersertifikat
Zenius Bersama Kampus Merdeka



Petunjuk

- Waktu presentasi adalah 5 menit (tentatif, tergantung dari banyaknya kelompok yang mendaftarkan diri)
- Waktu tanya jawab adalah 5 menit
- Silakan menambahkan gambar/visualisasi pada slide presentasi
- Upayakan agar tetap dalam format poin-poin (ingat, ini presentasi, bukan esai)
- Jangan masukkan *code* ke dalam slide presentasi (tidak usah memasukan screenshot jupyter notebook)

- 1. Latar Belakang**
- 2. Explorasi Data dan Visualisasi**
- 3. Modelling**
- 4. Kesimpulan**

Latar Belakang

Latar Belakang Project

Sumber Data:

<https://www.kaggle.com/datasets/teyang/singapore-hdb-flat-resale-prices-19902020>

Problem: **regression**

Tujuan:

- Mengidentifikasi faktor tinggi rendahnya harga jual flat pada di Singapura
- Memprediksi harga jual kembali flat HDB

Explorasi Data dan Visualisasi



Business Understanding

Terbatasnya lahan di Singapura serta harga rumah baru dan apartemen yang mahal mendorong masyarakat Singapura untuk membeli apartemen HDB.

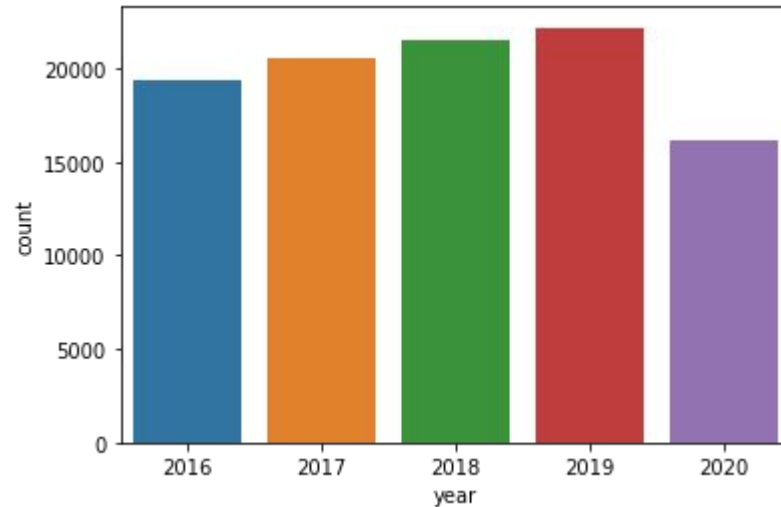
HDB merupakan sebuah lembaga negara yang berada di bawah Kementrian Pembangunan Negara yang bertugas di bidang perumahan publik di Singapura.

Penting meneliti harga dan penggerak harga HDB, bukan hanya untuk individu atau keluarga yang ingin membeli rumah baru, tetapi juga bagi para analis dan ekonom yang mencoba memahami stabilitas dan kinerja sektor ini.

Data Cleansing

Data terdiri dari 11 kolom, 99.7474 baris, tidak ada missing value.

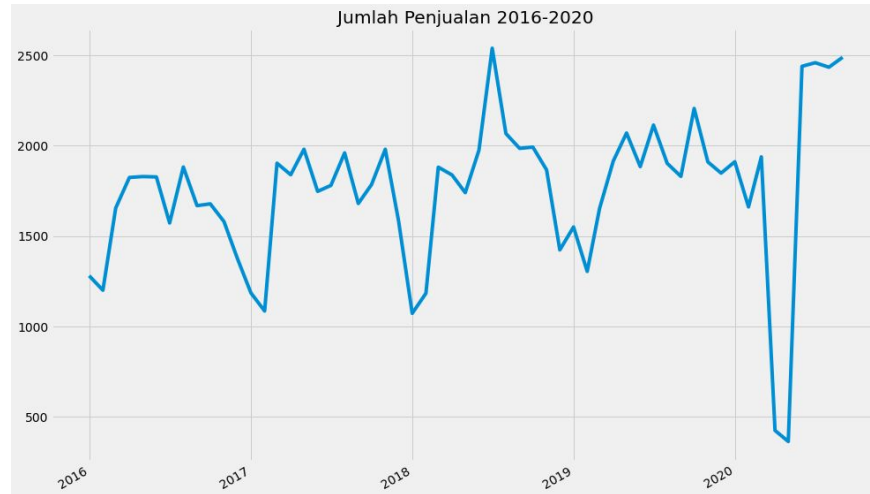
Year to Resale Price



Dapat terlihat pada tahun 2020 jumlah resale price paling rendah dibanding dengan tahun-tahun sebelumnya, karena mungkin data 2020 tidak sepenuhnya hingga bulan Desember

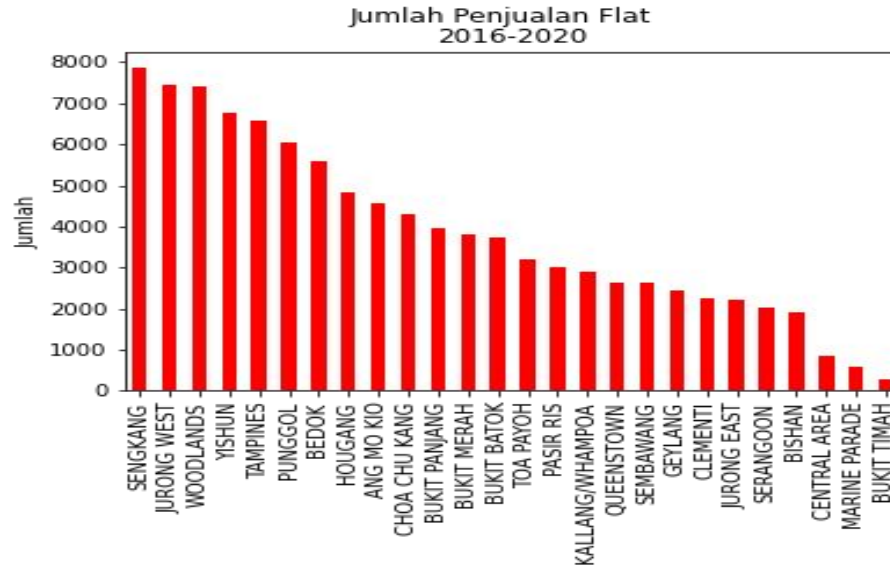
**Apa saja faktor tinggi rendahnya
resale flat ?**

Year to Resale Price



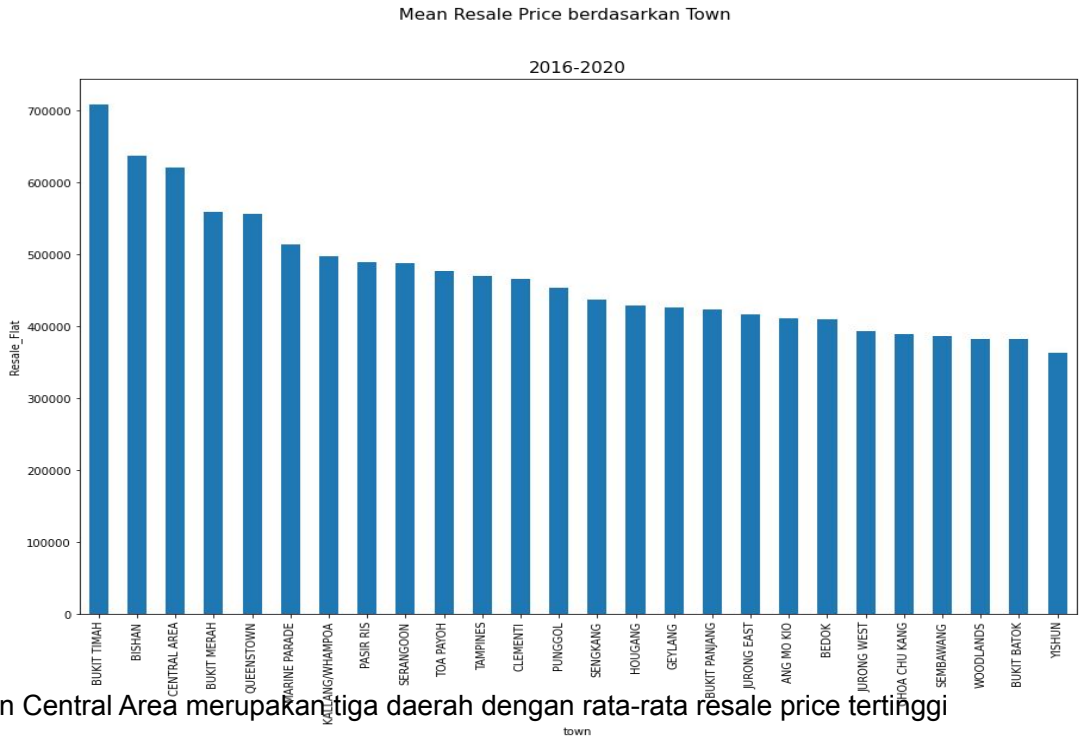
Jumlah Penjualan pada flat selalu menurun pada awal tahun, khususnya pada tahun 2020

Town



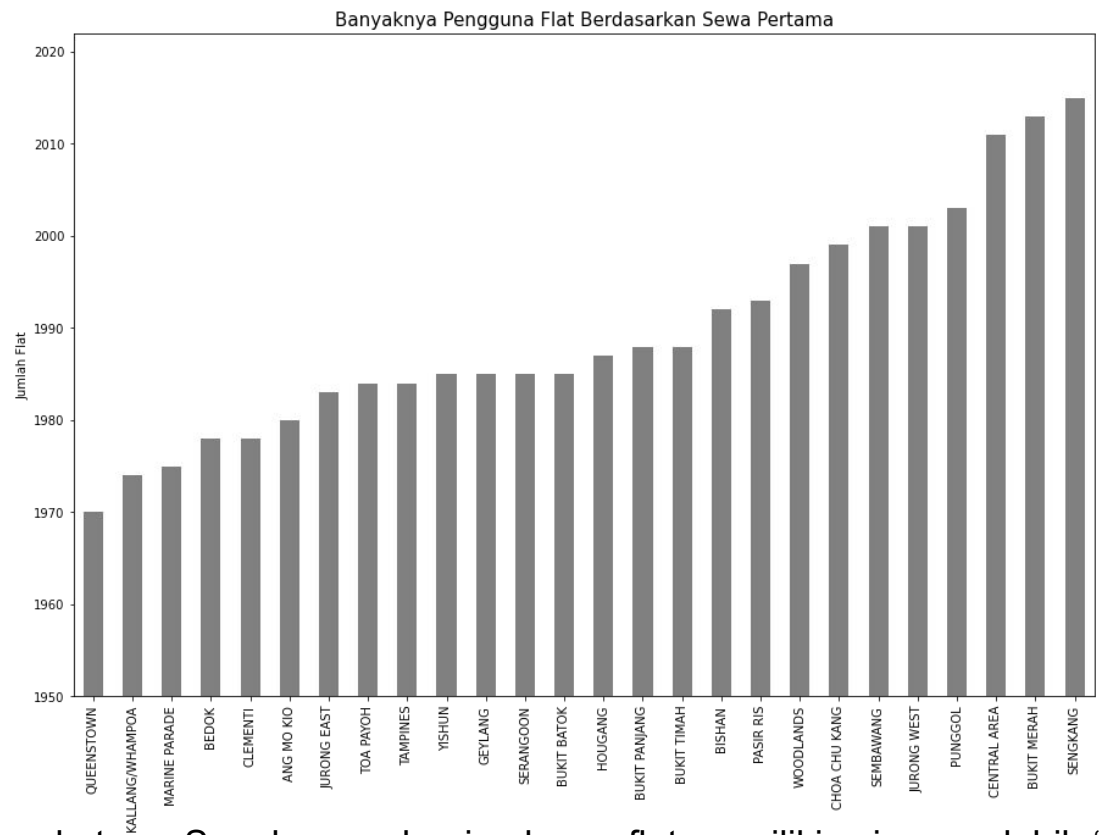
Flat yang paling banyak terjual adalah SENGKANG, JURONG WEST, dan WOODLANDS

Town to Resale Price

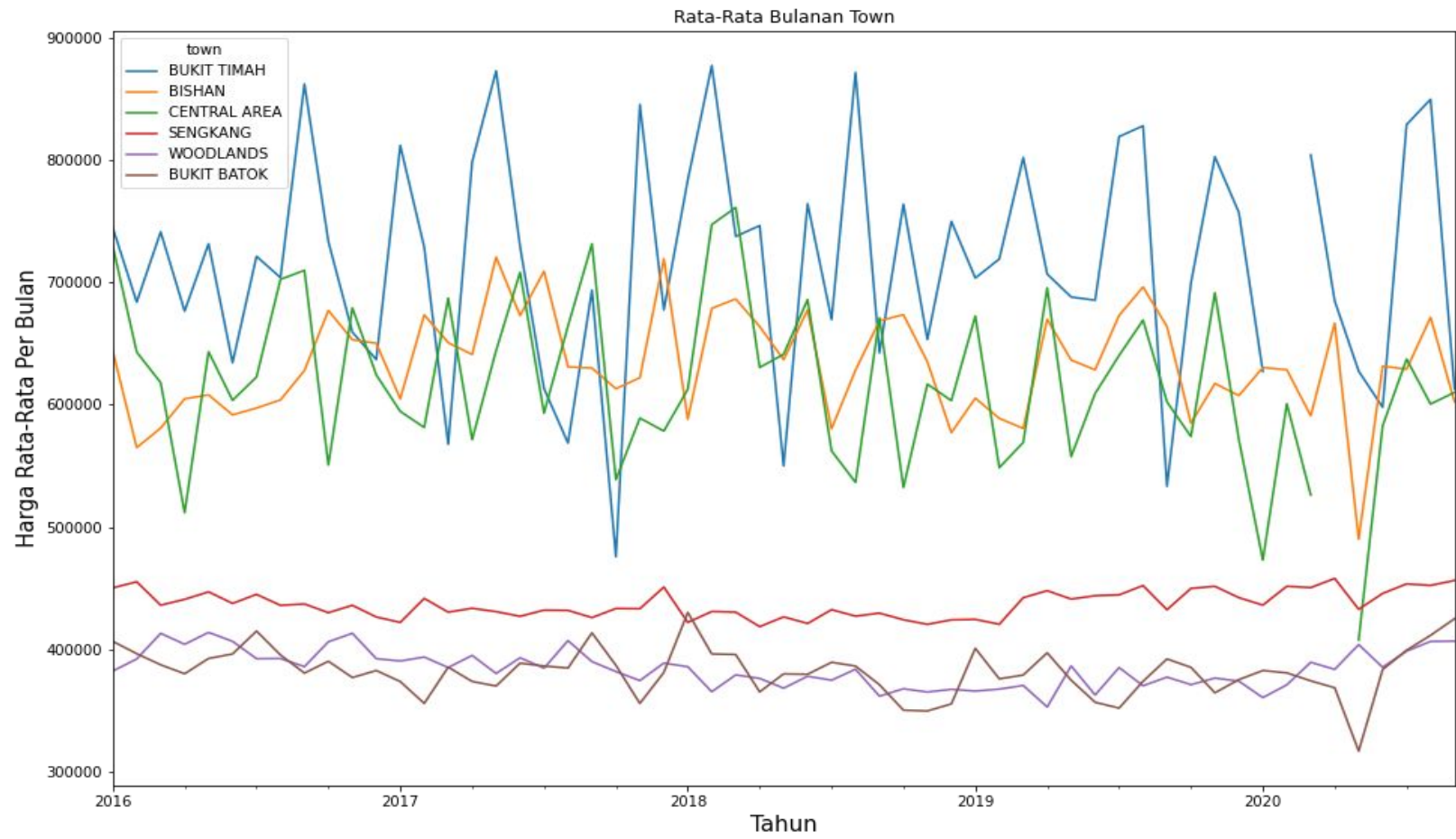


Bukit Timah, Bishan, dan Central Area merupakan tiga daerah dengan rata-rata resale price tertinggi

Sedangkan Woodlands, Bukit Batok, dan Yishun merupakan apartement dengan rata-rata resale flat terendah



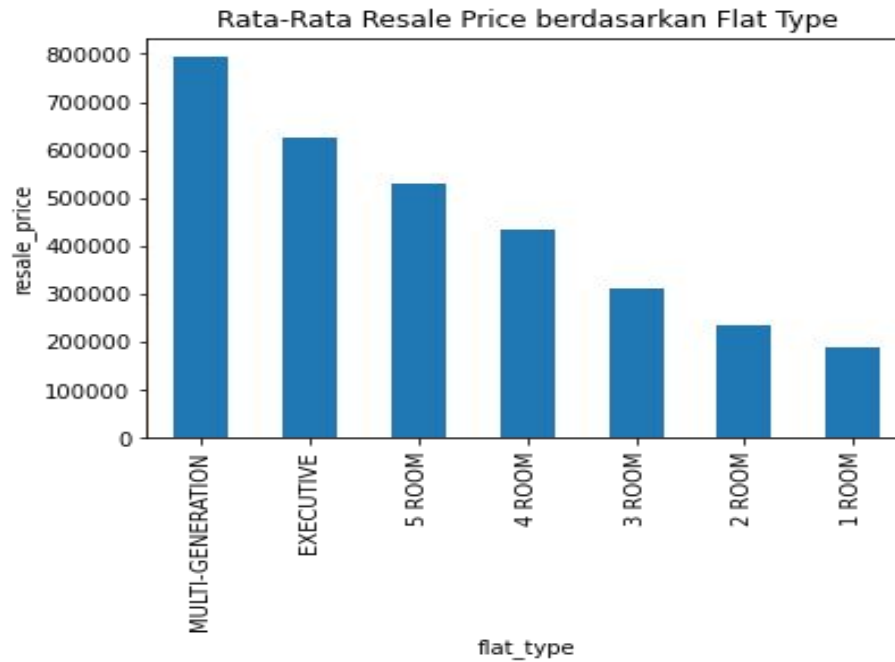
Namun, pada town Sengkang sebagian besar flat memiliki usia yang lebih ‘muda’ yaitu dengan waktu sewa pertama diatas tahun 2010



Berdasarkan Grafik :

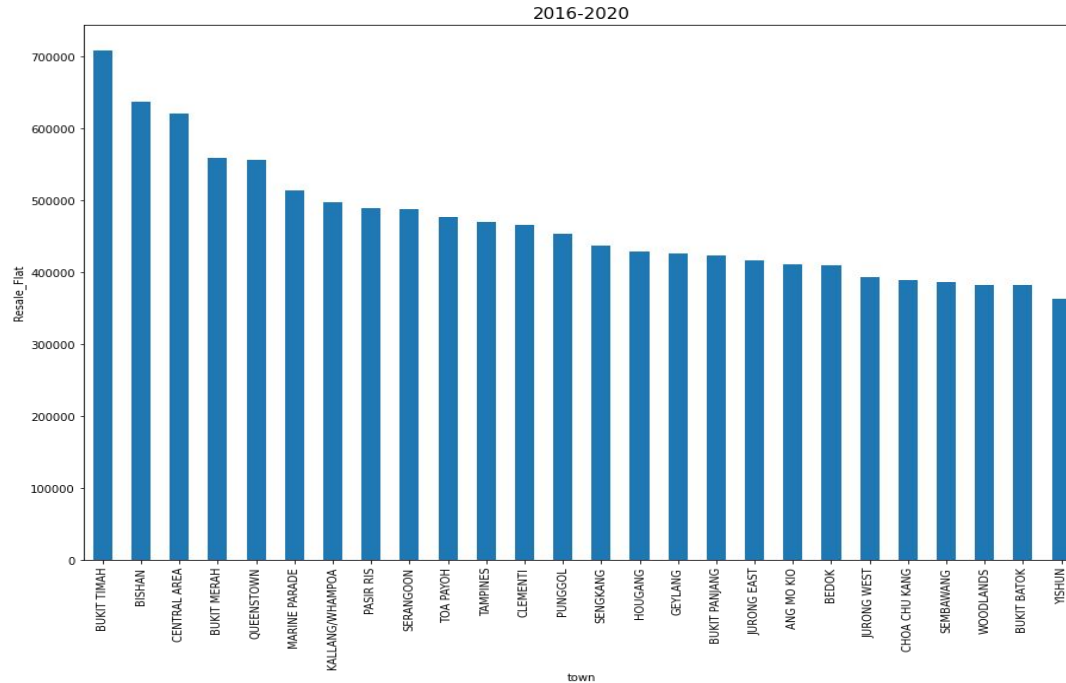
- Pada town Sengkang, flat harga cenderung stabil.
- Sengkang memiliki jangka waktu penggunaan flat yang lebih lama juga. Sehingga, Sengkang bisa menjadi salah satu flat dari HDB yang 'worth it' untuk dibeli kembali.
- Central Area memiliki harga jual yang tinggi dan usia flat yang lebih 'muda'
- Rata-rata flat pada town Bukit Timah, Bishan, dan Central Area cenderung tidak stabil, dimana pada periode tertentu mengalami kenaikan dan penurunan rata-rata penjualan

Flat_Type

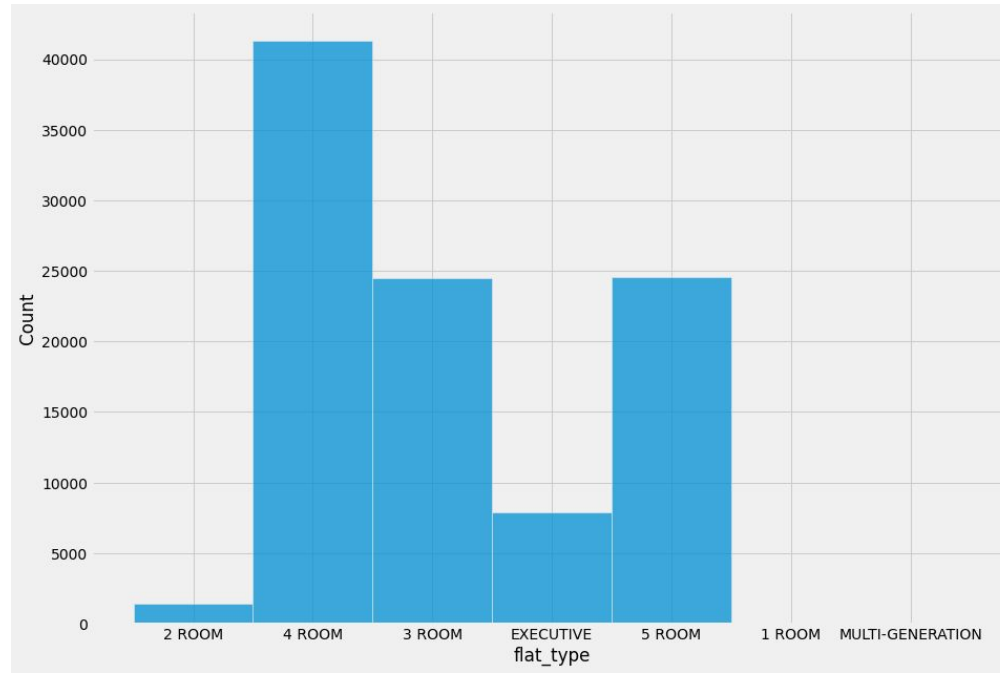


Flat Type

Mean Resale Price berdasarkan Town

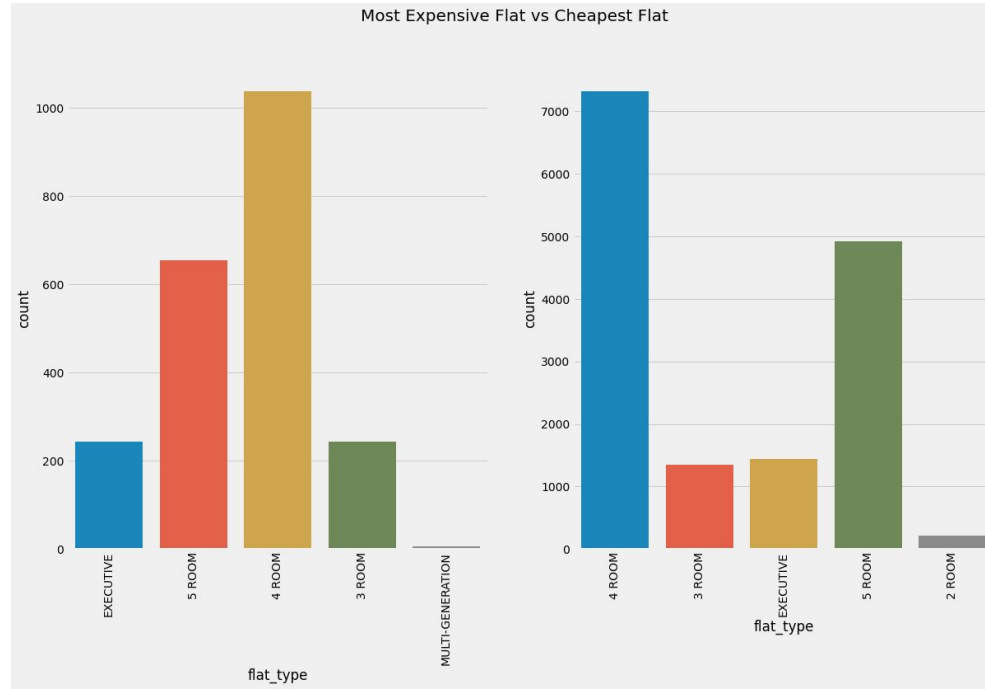


Total Pengguna Flat Type



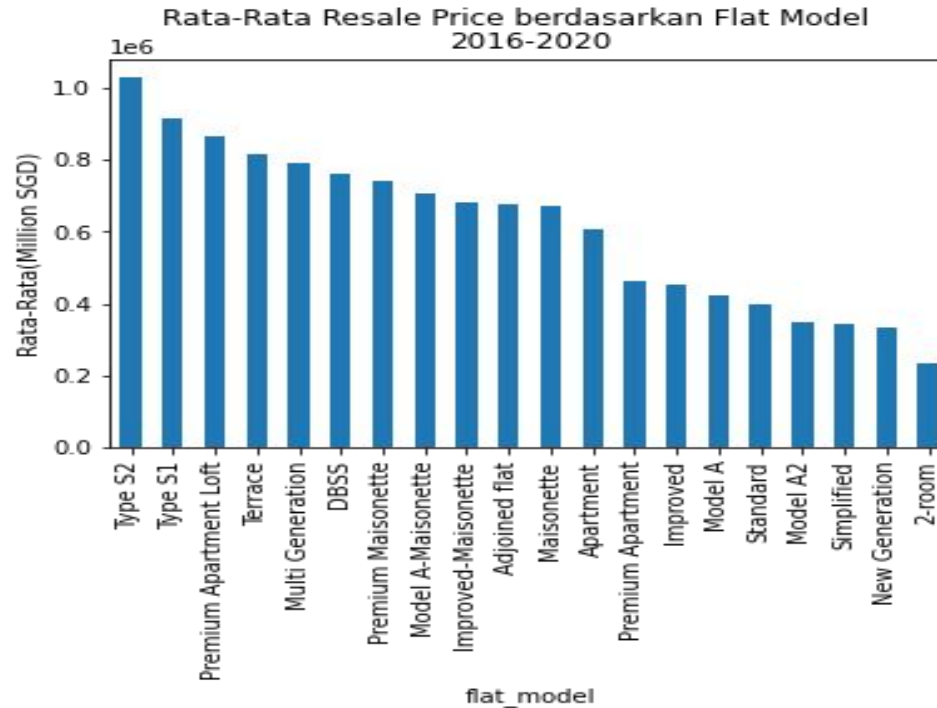
Sepertinya penduduk Singapura tidak begitu memiliki minat terhadap rumah dengan type 1 room, 2 room, dan Multi Generation

Flat Type



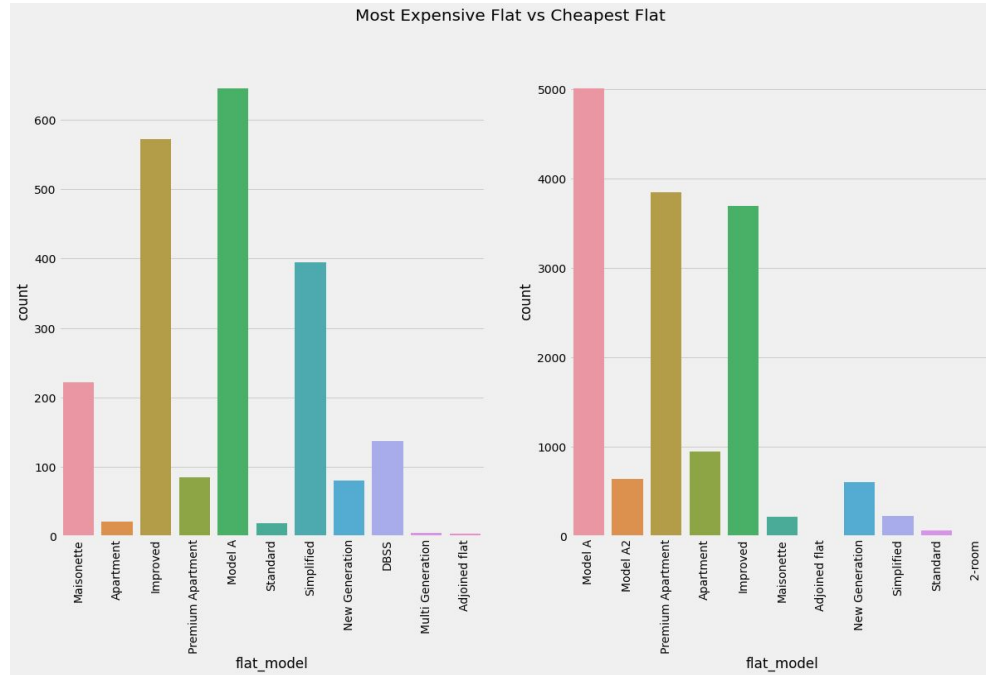
- Most Expensive Flat & Cheapest Flat diambil dari 3 town tertinggi dan terendah berdasarkan mean.
- Flat type favorit pada tiap jenis adalah 4 ROOM dan 5 ROOM

Mean for Resale Price Based on Flat Model



Type S2, Type S1, Premium Apartment Loft merupakan model apartement yang paling mahal menurut rata-rata resale price nya.

Mean for Resale Price Based on Flat Model



- Most Expensive Flat & Cheapest Flat diambil dari 3 town tertinggi dan terendah berdasarkan mean.

Type S2, Type S1, Premium Apartment Loft merupakan model apartement yang paling mahal menurut rata-rata resale price nya.

Mean for Resale Price Based on Flat Model

Pada grafik diatas, untuk banyaknya penjualan flat 2016-2020 untuk Most Expensive Flat adalah

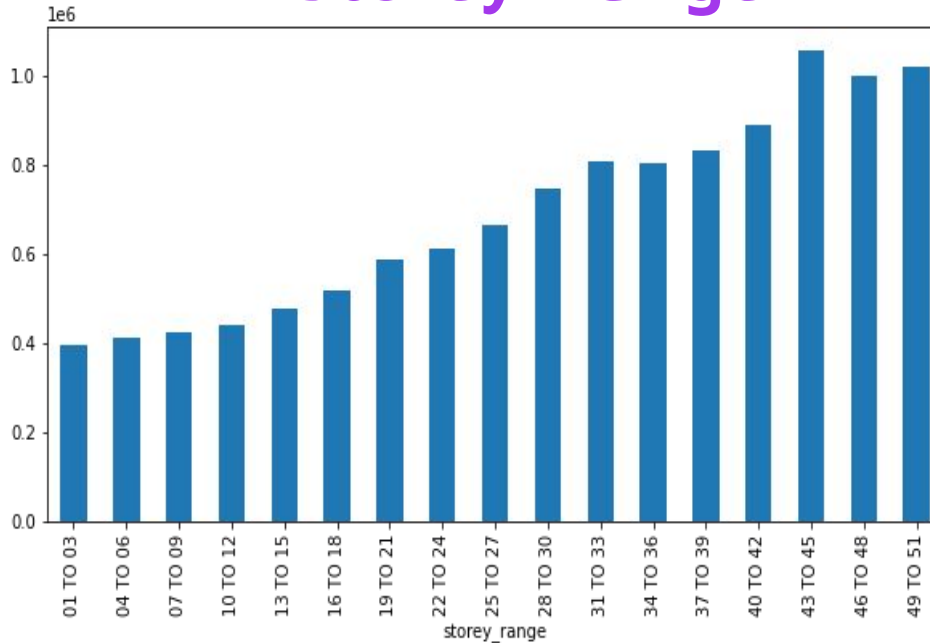
1. Model A
2. Improved
3. Simplified

Sedangkan pada Cheapest Flat adalah

1. Model A
2. Improved
3. New Generation

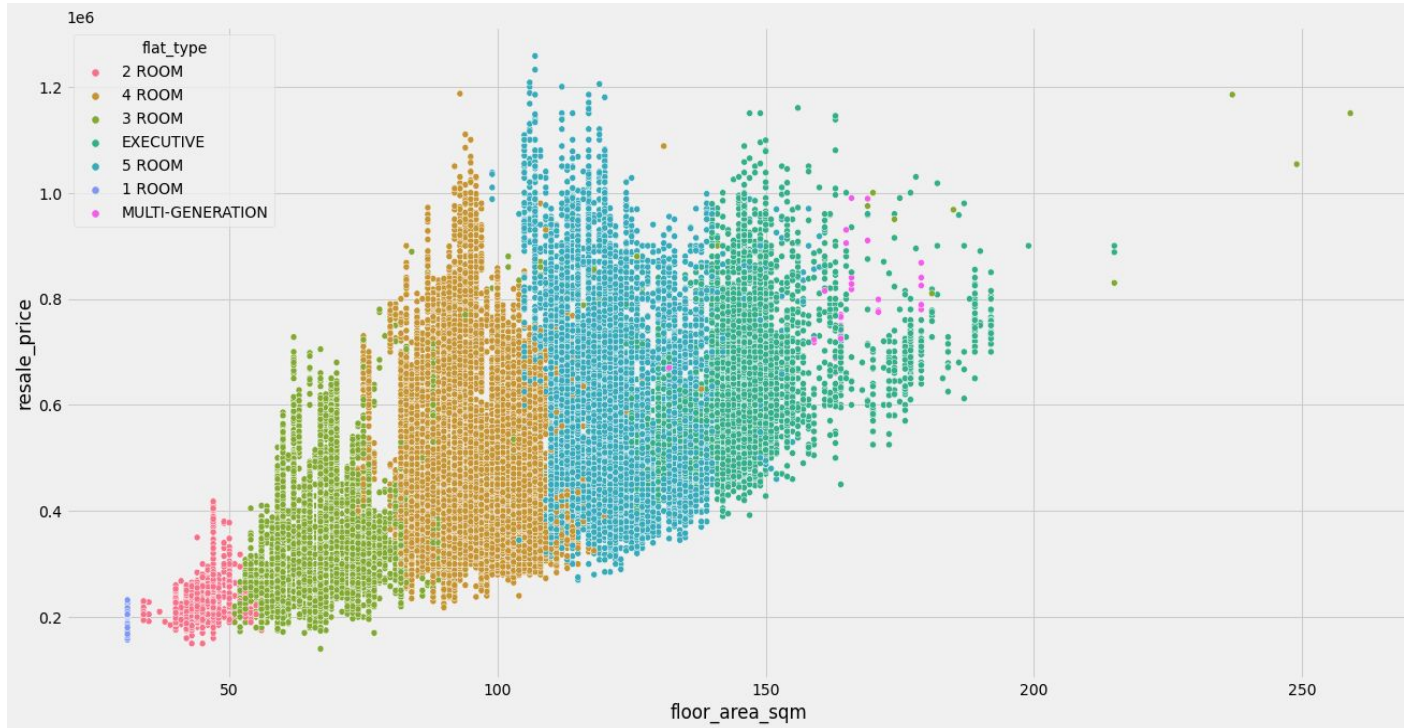
Dengan flat model yang sama, tetapi karena perbedaan daerah menyebabkan perbedaan harga flat diantara beberapa town. Sehingga, town dapat mempengaruhi resale price dari suatu flat

Storey Range

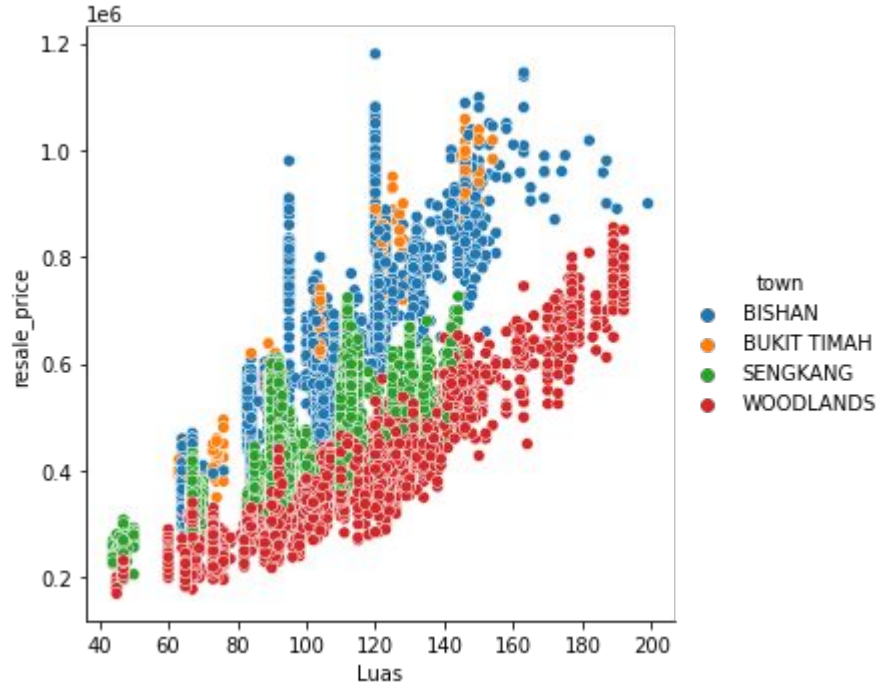


Ternyata, semakin tinggi dari tingkat flat, maka harga semakin mahal, yang dapat terlihat pada rata-rata resale price diatas. Namun, pada lantai 43 to 45 lah rata-rata harga flat paling mahal

Floor Area Sqm

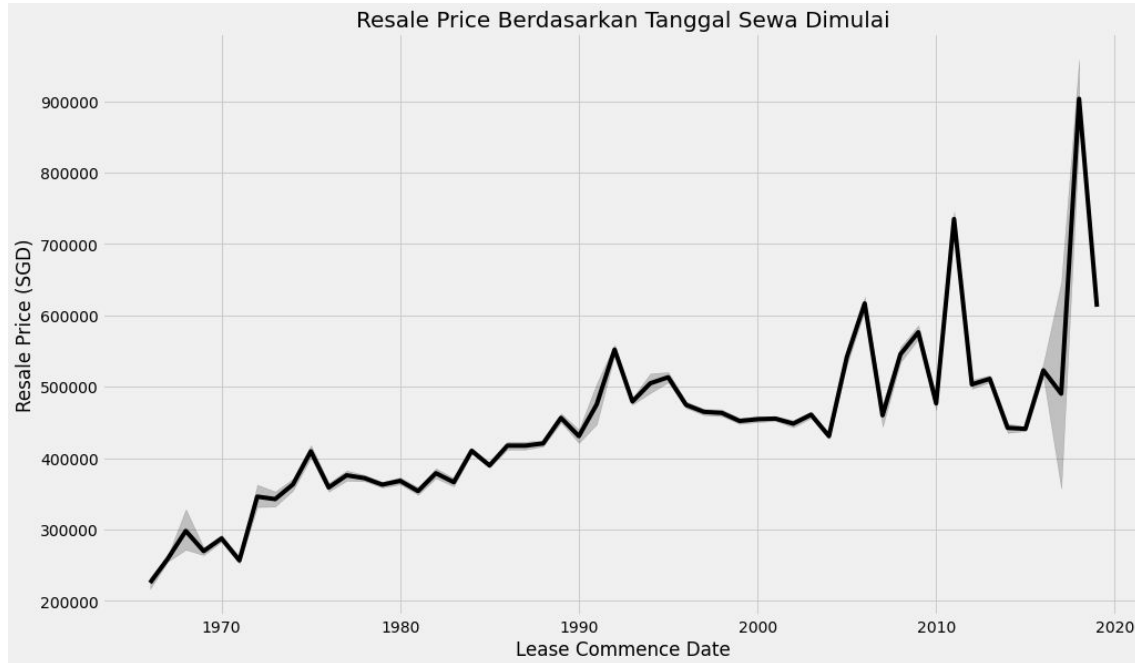


Persebaran Flat Berdasarkan Floor Area & Resale Price



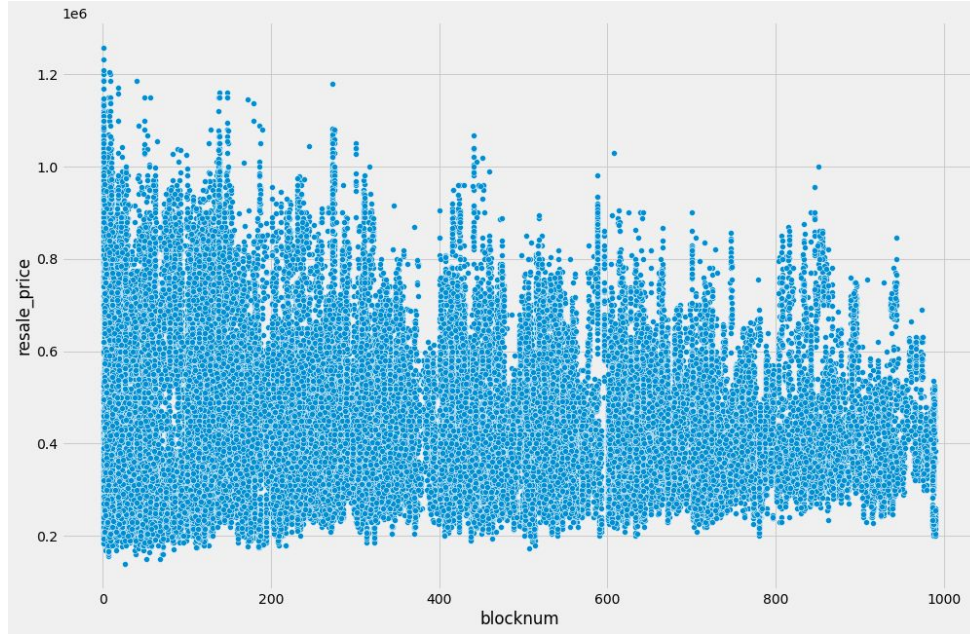
- Terlihat pada gambar disamping bahwa beberapa flat dengan luas yang relatif sama, namun memiliki harga yang berbeda
- Kemungkinan resale price tersebut memiliki harga yang berbeda karena mungkin faktor 'town'

Lease Commence Date & Remaining Lease



Seiring dengan 'muda' nya flat, nilai jual akan cenderung lebih mahal

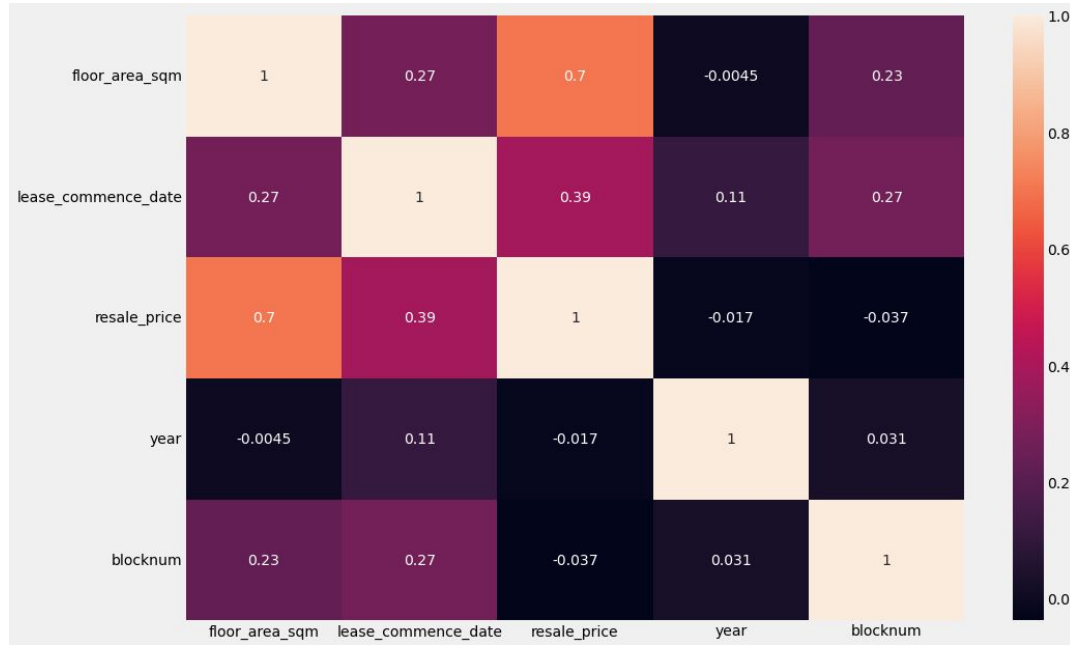
Block



Block tidak memiliki pola tertentu untuk bagi resale price

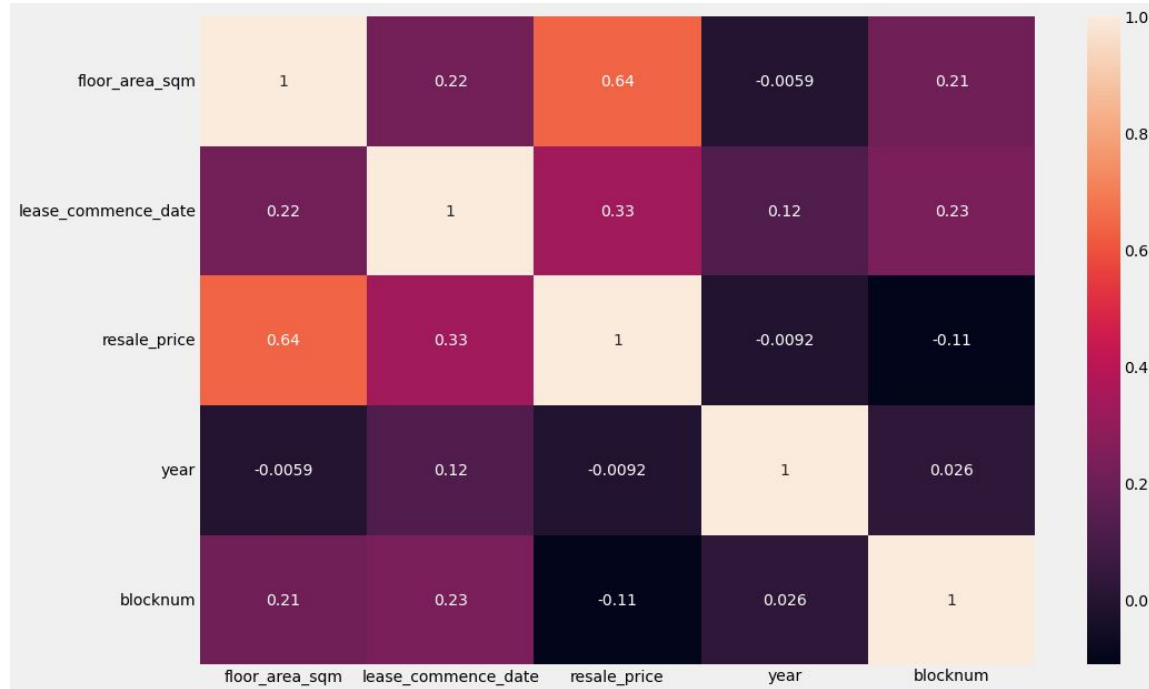
Analisis Korelasi

- Dilakukan dengan metode spearman



Analisis Korelasi

- Dilakukan dengan metode Pearson



Modelling



Multivariate Linier Regression

- Training data = 80%, Test data = 20%
- Metrik yang digunakan : Mean Absolute Error (MAE), Root Mean Square, R-Square
- Model yang digunakan adalah Multivariate Linier Regression
- Berdasarkan evaluasi metrik, didapatkan nilai
 - $MAE = 42584.26361757749$
 - $MSE = 3071854627.2131996$
 - $RMSE = 55424.314404539095$
 - $R\text{-Squared} = 86,29978 \%$

Hyperparameter Tuning for Linear Regression

- Training data = 80%, Test data = 20%
- Metrik yang digunakan : Mean Absolute Error (MAE), Root Mean Square, R-Square
- Model yang digunakan adalah Ridge Regression, Lasso dan Elastic Net

Hyperparameter Tuning for Linear Regression

Model	Metric			
	MAE	MSE	RMSE	R-Squared
Ridge	42584.72078637992	3071879015.601394	55424.534419347125	0.8629968118750111
Lasso	42584.46705067491	3071701602.2396073	55422.9339014059	0.863004724359864
Elastic Net	76274.3800453787	11072874480.929977	105227.72676880356	0.50615922766143

Random Forest Regressor

- Random Forest bisa digunakan dalam model klasifikasi maupun regresi.
 - Training data = 80%, Test data = 20%
 - Kolom yang digunakan adalah town, flat_type, storey_range, floor_area_sqm, lease_commence_date, dan resale_price
 - Metrik yang digunakan : Root Mean Square, R-Square
 - Model yang digunakan adalah Multivariate Linier Regression
 - Berdasarkan evaluasi metrik, didapatkan nilai
- RMSE = 33554.817874284396
- R-Squared = 94,9784668%

Hyperparameter Tuning Random Forest Regressor

- Didapatkan beberapa nilai untuk model random forest optimal seperti:
 - 1) Maksimal depth yaitu 10,23,36,50
 - 2) Minimul leaf atau daun adalah 1,2,4
 - 3) Banyaknya n yaitu 10,30,50
 - 4) Minimum split 2,5, dan 10

Conclusion

Conclusion

1. Peminat flat menuju akhir tahun hingga awal tahun selalu menurun.
2. Harga flat di beberapa town memiliki harga yang lebih tinggi dibanding beberapa town yang memiliki type dan model yang sama serta luas flat yang relatif sama, mungkin ini terpengaruh oleh lokasi town yang memiliki daya tarik tersendiri
3. Flat dengan tipe 1 Room dan Multi Generation memiliki peminat yang sangat sedikit, dibandingkan dengan flat type yang lainnya dari tahun 2016-2020.
4. Fitur atau kolom town, luas flat, flat model, flat type, waktu sewa awal serta tahun memiliki pengaruh terhadap resale price berdasarkan proses Exploratory Data Analysis. Namun beberapa fitur seperti block tidak terlalu berpengaruh terhadap tinggi rendahnya resale price
5. Model Machine Learning yang terbaik adalah Random Forest Regressor berdasarkan perbandingan nilai RMSE dan R-Squared dibandingkan dengan model Multivariate Regression dan Ridge Regression yang memiliki nilai error(RMSE) yang lebih tinggi dan nilai R-Squared yang lebih rendah

Rekomendasi Untuk HDB Resale Price

Recommendation

1. Pihak pengelola bisa memberikan harga menarik (misal diskon awal tahun) untuk meningkatkan peminat flat pada akhir tahun dan awal tahun, karena menurunnya peminat pada periode tersebut
2. Flat dengan type 1 room tidak terlalu diminati, diharapkan pengelola bisa mengganti type flat ini atau memperbaiki fitur-fitur didalamnya agar memiliki peminat yang tinggi
3. Perlu meningkatkan peminat pada daerah Central Area, Marine Parade, dan Bukit Timah karena peminat yang tidak begitu banyak akibat tingginya harga dari flat tersebut. Perlu mengadakan promo untuk meningkatkan peminat serta menyeimbangkan penjualan flat di tiap town-nya
4. Untuk flat type Multi Generation, sebaiknya diadakan promo karena calon pembeli flat ini bisa dari berbagai kalangan dari berbagai generasi.

Terima kasih!

Ada pertanyaan?

zenius



Kampus
Merdeka
INDONESIA JAYA