Final Project Presentation

```
Kelompok GaNi
```

Nama:

Daniel Andrew Ketaren (211402062)

Sarmida Uli Sinaga (211402071)

Muhammad Hatta Abdillah (211402110)

Luthfi Muzhaffar Lubis (211402119)

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: Singapore House & Development Board Resale Price

Problem : **Regression**

Tujuan:

- Analisis data Singapore HDB Resale Price 2016 2020 untuk perspektif kontemporer.
- Predictive Modeling; menggunakan analisis regresi untuk mengidentifikasi faktor-faktor penentu utama yang mempengaruhi harga jual kembali.
- Geospatial Analysis; menemukan kawasan properti bernilai tinggi untuk investasi yang terinformasi
- Property Features; mengkaji pengaruh tipe rumah, luas lantai, dan jangka waktu sewa terhadap harga jual kembali.
- Trend Analysis; melihat perubahan volume transaksi untuk mengidentifikasi tren pasar.

Explorasi Data dan Visualisasi

Business Understanding

Dataset ini mencakup informasi mengenai harga jual kembali rumah Dewan Pengembangan Perumahan (HDB) di Singapura dari tahun 2015 hingga 2020. Singapura, sebagai negara dengan lahan yang terbatas, menghadapi tantangan dalam penyediaan perumahan bagi masyarakatnya. Harga bersaing dan mahalnya rumah baru mendorong banyak individu untuk mempertimbangkan rumah susun HDB yang dijual kembali.

Faktor lokasi, tipe rumah, luas, dan infrastruktur menjadi fokus utama analisis. Analisis harga jual kembali ini tidak hanya penting bagi calon pembeli, tetapi juga bagi pihak ekonomi dan pengembang properti yang berkepentingan memahami dinamika pasar perumahan.

Guna memenuhi instruksi spesifik serta optimisasi model, **butuh dilakukan** *data cleansing* **pada dataset ini** (setelah digabungkan)

Setelah penggabungan 2 dataset yang ada, kami mendapati dimensi dari data ialah sebanyak 117527 rows × 11 columns. Kemudian kami melakukan filtering dimana data yang digunakan hanya dalam rentang tahun 2016 - 2020. Setelah dilakukan tahapan-tahapan data cleansing, kami mendapati dimensi dari dataset yang siap ditelusuri pada tahapan EDA ialah 99501 rows × 11 columns.

| 117524 | 2020-09 | YISHUN | 5 ROOM | 835 | | | | | | |
|--------------------------|---------|--------|-----------|-----|--|--|--|--|--|--|
| 117525 | 2020-09 | YISHUN | EXECUTIVE | 791 | | | | | | |
| 117526 | 2020-09 | YISHUN | EXECUTIVE | 387 | | | | | | |
| 117527 rows × 11 columns | | | | | | | | | | |



| 99744 | 2020-09 | YISHUN | 5 ROOM | 83 |
|----------|----------------|--------|-----------|----|
| 99745 | 2020-09 | YISHUN | EXECUTIVE | 79 |
| 99746 | 2020-09 | YISHUN | EXECUTIVE | 38 |
| 99527 ro | ws × 11 column | s | | |

1) Kendala 1: duplikasi data sejumlah 220 data

Solusi 1 : menghapus data duplikat menggunakan fungsi *drop*()





2) Kendala 2 : **perbedaan format value pada kolom 'remaining_lease'**Solusi 2 : mendefinisikan suatu fungsi untuk menyamakan format value dengan dilakukannya pembulatan, apabila hitungan bulan lebih dari 6 bulan maka akan dinaikkan menjadi tahun berikutnya dan sebaliknya apabila kurang dari 6 bulan.





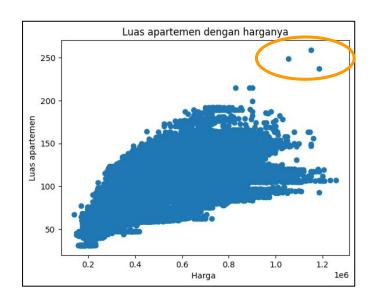
| ate | remaining_lease | resale_p |
|-----|-----------------|----------|
| 79 | 62 | 2300 |
| 78 | 61 | 2400 |
| 86 | 69 | 2760 |
| 980 | 63 | 2800 |
| 977 | 60 | 2850 |
| | | |
| 87 | 66 | 4400 |
| 987 | 65 | 4580 |
| 987 | 66 | 4900 |
| | | |

3) Kendala 3:

anomali data yang diindikasikan sebagai outlier

Solusi 3

Melakukan pengecekan data yang outlier dengan menggunakan metode IQR (*Interquartile Range*) atau rentang akar kuartil dari sekumpulan data.



| | month | town | flat_type | block | street_name | storey_range | floor_area_sqm | flat_model | lease_commence_date | remaining_lease | resale_price |
|-------|---------|-----------------|-----------|-------|-------------|--------------|----------------|------------|---------------------|-----------------|--------------|
| 18710 | 2016-12 | KALLANG/WHAMPOA | 3 ROOM | 57 | JLN MA'MOR | 01 TO 03 | 259.0 | Terrace | 1972 | 54 | 1150000.0 |
| 39066 | 2017-12 | KALLANG/WHAMPOA | 3 ROOM | 65 | JLN MA'MOR | 01 TO 03 | 249.0 | Terrace | 1972 | 53 | 1053888.0 |
| 55149 | 2018-09 | KALLANG/WHAMPOA | 3 ROOM | 41 | JLN BAHAGIA | 01 TO 03 | 237.0 | Terrace | 1972 | 52 | 1185000.0 |
| | | | | | <u> </u> | + Code + N | Markdown | | | | |

- <jelaskan apakah data butuh dibersihkan' atau tidak usah>
-
 <beri pemaparan dimensi dari data, brp baris, berapa kolom, berapa yang missing, dst>
- dan bagaimana solusi Anda terhadap problem tersebut>
- <misal: terdapat missing value pada kolom ____, solusi dari kami adalah ____>
- <bol><boleh didukung dengan visualisasi>

Pada dataset ini terdapat 3 insight menarik yang dapat kami sajikan, antara lain :

- 1) Analisis lokasi properti di Singapura untuk menentukan nilai yang sepadan.
- 2) Identifikasi faktor-faktor penentu harga properti yang paling signifikan.
- 3) Eksplorasi tren dan volume transaksi properti dari tahun 2016 hingga 2020.

1) Analisis lokasi properti di Singapura untuk menentukan nilai yang sepadan.

Adapun ketentuan daerah yang paling sepadan disini merujuk dari indikator:

- 1. Mahal atau tidaknya suatu apartemen di daerah tersebut
- 2. Luas atau tidaknya bangunan tersebut
- 3. Berapa harga rata rata dari bangunan berdasarkan flat_typenya
- 4. Berapa harga rata rata dari bangunan berdasarkan flat_modelnya

Dari data di atas, kita bisa memakai beberapa kolom antara lain ; resale price, flat type, town, floor_area_sqm, flat_model

1) Analisis lokasi properti di Singapura untuk menentukan nilai yang sepadan.

| Heatmap kota dengan persebaran banya | k apartem | en dengai | n rentang re | esale_price |
|--|------------------|-------------------|--------------------|-------------|
| ANG MO KIO - 18 2787 992 | 559 | 190 | 1 | - 500 |
| BEDOK - 6 3284 1485 | 717 | 79 | 0 | 500 |
| BISHAN - 0 164 706 | 653 | 332 | 49 | |
| BUKIT BATOK - 0 2307 1133 | 262 | 7 | 0 | |
| BUKIT MERAH - 44 1117 879 | 1215 | 536 | 18 | |
| BUKIT PANJANG - 0 1756 1854 | 311 | 3 | 0 | - 400 |
| BUKIT TIMAH - 0 18 56 | 97 | 95 | 8 | 400 |
| CENTRAL AREA - 0 170 312 | 91 | 199 | 70 | |
| CHOA CHU KANG - 0 2682 1568 | 47 | 3 | 0 | |
| CLEMENTI - 0 1113 573 | 424 | 140 | 7 | |
| GEYLANG - 47 1303 527 | 455 | 108 | 0 | - 300 |
| HOUGANG - 1 2384 1770 | 610 | 63 | 1 | - 300 |
| IURONG EAST - 0 1164 736 | 292 | 28 | 0 | |
| JURONG WEST - 58 3919 3244 | 199 | 0 | 0 | |
| ALLANG/WHAMPOA - 16 1200 720 | 694 | 256 | 23 | |
| MARINE PARADE - 1 180 259 | 58 | 77 | 0 | 100 |
| PASIR RIS - 1 505 1993 | 484 | 5 | o | - 200 |
| PUNGGOL - 0 1318 4359 | 327 | 10 | 0 | |
| QUEENSTOWN - 18 917 485 | 769 | 394 | 36 | |
| SEMBAWANG - 1 1602 1010 | 1 | 0 | 0 | |
| SENGKANG - 1 2474 5123 | 226 | 0 | 0 | |
| SERANGOON - 1 705 792 | 436 | 94 | 1 | - 100 |
| TAMPINES - 0 1937 3590 | 981 | 68 | 0 | |
| TOA PAYOH - 53 1559 562 | 603 | 393 | 27 | |
| WOODLANDS - 19 4747 2288 | 311 | 15 | 0 | |
| | 176 | | | |
| the state of the s | - 1 | 4 | 0 | - 0 |
| 6 6 6 | 0 | 0 | 0 | |
| 8 8 8 | 00 | 9 | 00 | |
| ğ ğ ğ | 00 | 00 | 00 | |
| (0, 200000) | 3, | 7 | 7 | |
| [0, 200000) | [600000, 800000) | 0, | 90 | |
| 00 00 | 000 | 000 | 00 | |
| [24 |)9] | [800000, 1000000) | (1000000, 1200000) | |
| resale | _price | | _ | |

| | | [0, 50) | [50, 100) | [100, 150) floor area sqn | [150, 200) | [200, 250) | - |
|------|-------------------|---------|-----------|------------------------------|------------|------------|-----------|
| | YISHUN - | 77 | 4343 | 2251 | 79 | 0 | - 0 |
| | WOODLANDS - | 53 | 2580 | 4391 | 356 | 0 | |
| | TOA PAYOH - | 90 | 2051 | 1025 | 31 | 0 | - 500 |
| | TAMPINES - | 38 | 2154 | 4262 | 122 | 0 | |
| | SERANGOON - | 12 | 956 | 995 | 66 | 0 | |
| | SENGKANG - | 140 | 4171 | 3513 | 0 | 0 | - 1000 |
| | SEMBAWANG - | 59 | 972 | 1583 | 0 | 0 | 44.500 |
| | QUEENSTOWN - | 78 | 2044 | 483 | 14 | 0 | |
| | PUNGGOL - | 102 | 3873 | 2039 | 0 | 0 | - 1500 |
| | PASIR RIS - | 17 | 209 | 2637 | 125 | 0 | |
| | MARINE PARADE - | 2 | 444 | 129 | 0 | 0 | 100157338 |
| _ | KALLANG/WHAMPOA - | 40 | 1983 | 853 | 30 | 3 | - 2000 |
| COWI | JURONG WEST - | 46 | 2771 | 4464 | 139 | 0 | |
| = | JURONG EAST - | 42 | 1202 | 949 | 27 | 0 | |
| | HOUGANG - | 35 | 2088 | 2605 | 101 | 0 | - 2500 |
| | GEYLANG - | 87 | 1745 | 588 | 20 | 0 | |
| | CLEMENTI - | 16 | 1797 | 401 | 43 | 0 | 5000 |
| | CHOA CHU KANG - | 34 | 762 | 3450 | 52 | 2 | - 3000 |
| | CENTRAL AREA - | 3 | 687 | 156 | 0 | 0 | |
| | BUKIT TIMAH - | 0 | 79 | 186 | 9 | 0 | 3300 |
| | BUKIT PANJANG - | 49 | 1593 | 2260 | 22 | 0 | - 3500 |
| | BUKIT MERAH - | 160 | 2363 | 1282 | 5 | 0 | |
| | BUKIT BATOK - | 0 | 2156 | 1455 | 98 | 0 | - 4000 |
| | BISHAN - | 0 | 572 | 1275 | 57 | 0 | |
| | BEDOK - | 82 | 3875 | 1554 | 60 | 0 | |
| | ANG MO KIO - | 87 | 3669 | 754 | 37 | 0 | |

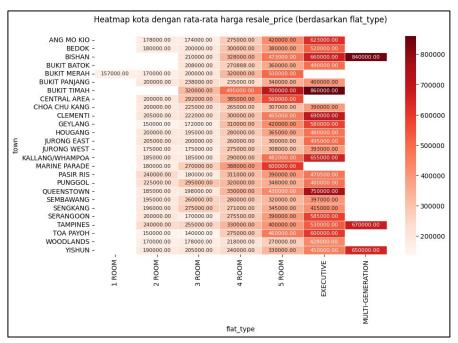
2) Identifikasi faktor-faktor penentu harga properti yang paling signifikan.

dapat divisualisasikan dengan kolom kolom lain untuk melihat keterkaitan antara kolom kolom tersebut dengan resale_price. Hal ini juga dapat membantu menemukan model yang tepat kedepannya.

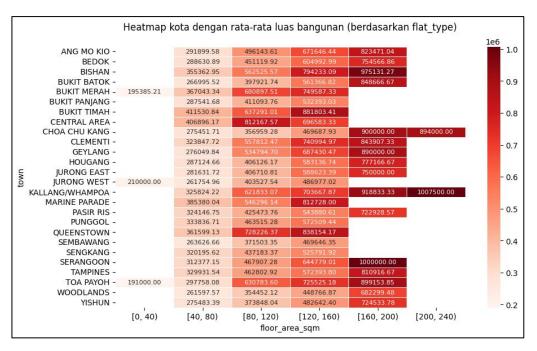
Kolom kolom yang akan dicek antara lain pengaruhnya antara lain:

- flat_type terhadap resale_price
- 2. floor_area terhadap resale_price
- 3. flat_model terhadap resale_price

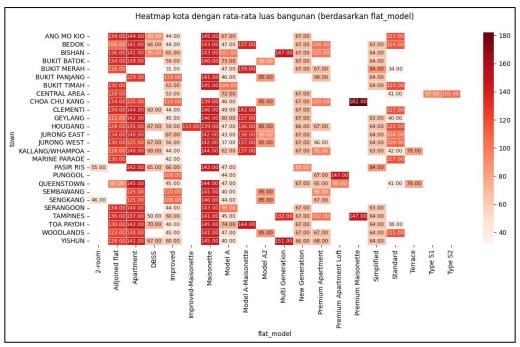
- 2) Identifikasi faktor-faktor penentu harga properti yang paling signifikan.
- Pengaruh flat_type terhadap resale_price



- 2) Identifikasi faktor-faktor penentu harga properti yang paling signifikan.
- Pengaruh floor_area terhadap resale_price



- 2) Identifikasi faktor-faktor penentu harga properti yang paling signifikan.
- Pengaruh flat_model terhadap resale_price

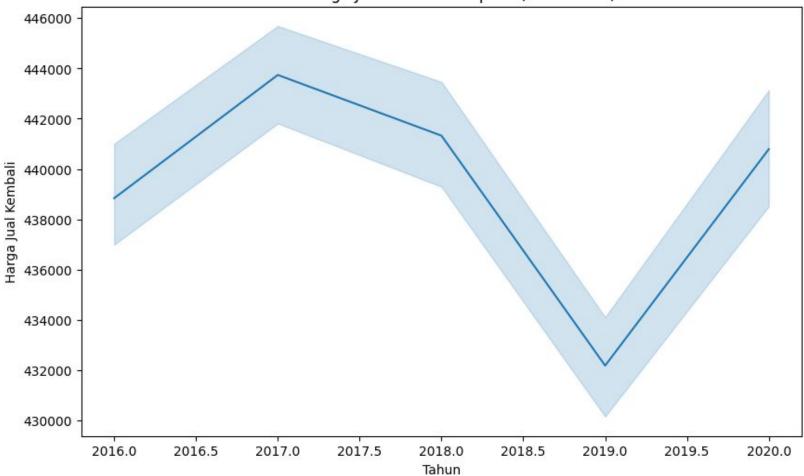


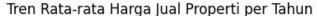
3) Eksplorasi tren dan volume transaksi properti dari tahun 2016 hingga 2020.

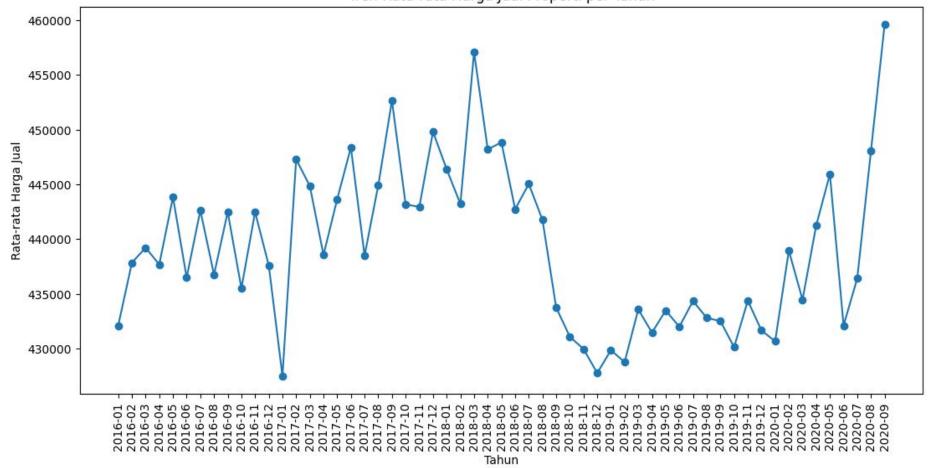
Heatmap kota dengan rata-rata harga setiap bulannya

| | | | | | | | H | leatmap | kota den | gan rata- | rata harga | setiap bu | lannya | | | | | | | | | |
|-------------------------------|-----------|--------------|--------------------------|---|---------------------------|-------------------|----------------|--|--|--------------|----------------|--------------------|---------------|---|--------------------------|-------------------|-------------|--------------|----------|--------------|----------|----------|
| 2016-01 -413153.40 420411.9 | 0 644086. | 406360.31 | 511390.14 39 | 98003.42 | 45000.00 730 | 0070.55 | 82854.82 4067 | 61.23 391620 | 0.33 416109.5 | 97 405155.52 | 403972.67 4917 | 41.70 496000 | 00 469743.00 | 448488.00 | 5760.00 40 | 5030.30 450080.0 | 0 514711.54 | 449120.89 49 | 50495.33 | 382052.24 30 | 67686.09 | |
| 2016-02 -420209.63 439066.9 | | 396306.12 | 79078 67 40 | 01762.00 | 84000 00 643 | 3197 71 3 | 88505.42 3970 | 83.33 41740 | 3.03 410481.4 | 42 393863.16 | 425613.69 4911 | 84.21 570857 | 459068.97 | 434995.23 53 | 31251.37 41 | 6090.37 455153.7 | 1 503451.15 | 454983.9046 | 52224.74 | 391953.20 3 | 46311.94 | |
| 2016-03 -398818.15 405699.7 | | | 18043.76 41 | | 41500 00 618 | | 02494.61 4143 | | | | | 82.88555680 | | 434946.63 | | 4303.71 435963.3 | | | | | | |
| 2016-04 -430451.96 393820.2 | | 9 379857.48 | 59653 11 38 | 35118.76 | 76714 29 511 | 1740.67 3 | 94304.33 4359 | 57.45 476197 | 7.67 414838.4 | 49 402860.35 | 402602.40 5528 | | 00 467753.49 | 448030.44 51 | 9601.31.41 | 5025.66 440758.2 | 8 473367.65 | 450974.62 48 | 83853.51 | 403972.08 30 | 66117.87 | |
| 2016-05 -431903.45 411593.8 | | 392378.26 | | 25747.307 | 31833 33 643 | | 93496.61 4732 | | | | | 76 46 548188 | 10 465344 15 | | | 2407.37 446898.6 | | | | | | |
| 2016-06 -397840.72 405817.8 | | 396068.00 | | 09935.88 | 34428 57 603 | | | | | | 399321.67 5296 | 54 04 482490 | 67 491037 04 | 445996 08 5 | | 4548.95 437427.5 | | | | | | |
| 2016-07 -388784.42 388311.6 | | | 32108.65 41 | | 21500 00 622 | | | | | | 401630.13 5271 | | 00 473385 57 | | | 3463.87 444822.4 | | | | | | |
| 2016-08 -421375.62 408985.3 | | 7 395272.10 | | | 04114.29.702 | | | | | | 405321.57 4999 | | | | | 4617.96 435865.2 | | | | | | |
| 2016-09 -419146.62 425786.1 | | 0 380372.00 | | | 62666.67 710 | | | | | | 406273.92 5332 | | 44 481969.78 | | | 5883.50 436645.4 | | | | | | - 800000 |
| 2016-10 -436764.82 394342.0 | | 4 390081.88 | | | 33629 33 550 | | | | | | 407179.10 4672 | | | | | 9712.86 429852.8 | | | | | | |
| | | 376764.87 | | | 33629.33 550 | | | | | | | | | | | | | | | | | |
| 2016-11 -431546.96 432697.5 | | | | 31651.31 | 59500.00 679 | | | | | | 418088.94 4755 | | .00 492341.13 | 465483.65 49 | | 0114.46 435867.8 | | | | | | |
| 2016-12 -439659.40 407832.0 | | | 32351.09 40 | | | | | | | | 390771.61 4875 | | 56 448444.67 | 473883.65 | | 4720.86 426313.2 | | | | | | |
| 2017-01 -413971.21 410027.3 | | | 513036.36 40 | | 12296.00 605 | | | | | | 401508.68 4737 | | 00 455888.86 | 469878.74 47 | | 0043.48 422845.8 | | | | | | |
| 2017-02 -398604.29 401326.8 | | 223301.43 | | 23724.23 | 29000.00 581 | | | | | | 390337.93 4680 | | 35 494166.67 | 468588.42 | | 2367.97 441181.7 | | | | | | |
| 2017-03 -435649.03 411079.7 | | 34 385267.84 | | 26614.81 | 67666.67 687 | | | | | | 394964.37 5097 | | | 453644.27 | | 1697.11 430020.8 | | | | | | |
| 2017-04 -421959.91 405896.7 | | 9 373550.20 | | | 98777.60 571 | | | | | | 389073.34 5284 | | | | | 7514.33 432914.0 | | | | | | |
| 2017-05 -444063.31 400064.3 | | | | | 73333.33 644 | 1400.00 | 883575.94 4858 | 805.88 403601 | 1.66 431254. | 30 397696.77 | 404385.49 5064 | 89.80 577571 | 43 492413.15 | 455050.42 | 4257.76 40 | 0131.03 429708.2 | 8 466617.82 | 489845.18 44 | 43555.37 | 380086.55 34 | 47867.53 | |
| 2017-06 -456342.43 417128.8 | 5 676963. | 388508.49 | 89296.00 40 | 00542.04 | 29600.00 708 | 3464.29 | 85710.95 5312 | 99.33 415816 | 5.70 453699. | 56 453550.00 | 389184.55 4940 | 50.00 465983 | 33 485125.00 | 438556.44 | 5474.73 40 | 7863.71 426278.5 | 1 515593.75 | 455961.95 49 | 95625.19 | 392861.84 35 | 55631.14 | |
| 2017-07 -429436.51 412494.1 | 0 716802 | 75 386222.21 | 39127 73 44 | 47426.45 | | 3181 82 3 | 88810.52 4151 | 65.03 39274 | 1.68 404458. | 09429944.41 | 391248.44 5094 | 88.89 537088 | 89 479286.88 | 438578.69 | 6101.83 38 | 9005.52 431669.2 | 2 465562.50 | 467005.38 52 | 29983.33 | 384421.79 34 | 40099.17 | - 700000 |
| 2017-08 -422186.71 455711.0 | 8 635197. | 383080.41 | 526142.46 42 | 22718.10 | 68666.67 664 | 1482.67 3 | 96740.69 4956 | 98.33 413954 | 1.26 446229.9 | 94 433890.14 | 399396.39 4737 | 26.34 557777 | 78 471687.74 | 447752.27 5 | 0673.28 38: | 3950.53 431187.79 | 9 500114.34 | 468453.50 52 | 20831.70 | 407002.77 34 | 43220.16 | |
| 2017-09 -421210.41 419811.1 | | 413339.22 | | | 93866 40 731 | | 96283.75 | | | | 394579.25 5311 | | 71 471517.76 | | | 7942.92 425452.2 | | | | 389183.94 35 | | |
| 2017-10 -393771.33 430626.7 | | 386752.03 | | 37440.06 4 | | | | | | | 392418.67 4881 | | | | | 2988.10 432347.6 | | | | | | |
| 2017-11 -411593.23 442565.1 | | 6 354127.67 | | 39348.75 | 45750 00 507 | | | | | | 392464.95 4936 | | 491348.91 | | | 6997.27 432800.3 | | | | 374274.61 35 | | |
| 2017-11 -411593.23 442565.1 | | | | | 77666 67 578 | | | | | | 398430.49 4950 | | | | | 1625.30 450814.8 | | | | 388449.15 35 | | |
| | | 7 430044.97 | | 43915.74 7 | | | | | | | | | | | | 4841.85 421893.0 | | | | | | |
| 2018-01 -402677.75 436464.3 | | | | | | | | | | | | | 00 488017.24 | 452734.34 48 | | | | | | | | |
| 2018-02 -381908.62 439376.8 | | | | | | | | | | | 384079.66 4808 | | 33 492785.71 | 453106.24 | | 1264.13 430835.1 | | | | | | |
| 2018-03 -435370.62 435416.6 | | | | | | | | | | | 395596.39 4856 | | | 441662.70 | | 5095.43 430308.6 | | | | 378927.03 35 | | |
| E 2018-04 -446903.49 381602.4 | | | | 300,0.23 | 46648.00 630 | | | | | | 392101.41 4803 | | 82 514243.86 | | | 2506.00 418332.5 | | | | 376000.41 36 | | 600004 |
| 2018-05 -439163.82 457459.4 | | 379764.89 | | 16116.10 | 50000 00 641 | 1375.00 3 | 374241.26 4734 | 158.67 444393 | 3.94 418517. | 22 421648.65 | 403491.29 5352 | 60.98 516428 | 57 506758.33 | 449995.04 | 1153.07 39 | 0091.11 426432.5 | 5 485111.09 | 514304.76 49 | 94320.33 | 368052.96 35 | 50689.29 | - 600000 |
| 2018-06 -430614.59 389536.4 | 3 677734. | 379441.50 | 57429.55 <mark>42</mark> | 21827.57 | 64555.43 686 | 5187.50 | 89327.82 5137 | 57.89 458052 | 2.90 409642. | 80 396184.21 | 384011.01 5246 | 82.84 581642 | 86 491262.52 | 435134.87 60 | 08064.56 <mark>38</mark> | 7090.98 421011.7 | 6 468321.12 | 484861.53 53 | 13306.70 | 377796.19 34 | 48313.71 | |
| 2018-07 -442292.92 409447.1 | 0 580388. | 389240.71 | 87636.58 44 | 44786.39 | 69769.23 562 | 2307.69 | 84288.29 4752 | 93.73 386773 | 3.52 450693.5 | 95 456044.75 | 388691.05 4700 | 45.80 456180 | 00 502152.12 | 451280.57 | 39720.59 38! | 5149.73 432326.3 | 6 491936.00 | 498947.07 | 51630.46 | 374568.59 34 | 44083.94 | |
| 2018-08 -427889.44 396668.7 | 3 628366 | 0 386151.99 | 87946.47 42 | 29731.45 | 72029.33 536 | 464.00 3 | 883947.60 5159 | 28.57 452973 | 3.91 460803. | 63 450368.92 | 362457.06 4996 | 71.24 594125 | 87 501264.43 | 441636.00 5 | 7533.29 38: | 3905.18 427383.5 | 7 476664.40 | 471775.85 53 | 30926.67 | 383554.74 33 | 39403.24 | |
| 2018-09 -404265.12 390869.0 | 7 668823. | 370852.21 | 68675.37 44 | 40022.18 | 42500.00 671 | 1200.00 3 | 81077.48 4308 | 312.04 440096 | 5.85 443176. | 73 412611.50 | 373097.54 4416 | 33.53 628376 | 00 511122.19 | 459984.23 51 | 7894.62 37 | 5566.62 429470.0 | 3 533481.44 | 447548.49 43 | 73079.30 | 361538.51 36 | 64262.18 | |
| 2018-10 -414557.06 428600.7 | 9 673729 | 350059.43 | 53804 44 42 | 22821.07 | 64129 33 532 | | | | | | 373679.30 5103 | | 00 497042.91 | 453544.69 53 | 34623.52 37 | 7643.31 424050.1 | 0 498906.13 | 463033.12 4 | 71698.51 | 367566.64 34 | 48710.29 | |
| 2018-11 -410242.23 419795.0 | | 349421.46 | | 27460.73 | 53472 00 616 | | | | | | 387431.22 4605 | | | 442241 98 49 | | 2615.80 420742.5 | | | | | | |
| 2018-12 -443345.06 422670.0 | | 355154.67 | | | 50125 00 603 | | | | | | 367882.28 4419 | | 479866.38 | 472814 02 5 | | 8227.90 423994.3 | | | | | | |
| 2019-01 -386977.03 407493.9 | | 8 400673.00 | | | 03750 00 672 | | | | | | 382783.65 5147 | | | | | 5326.75 424744.7 | | | | | | |
| 2019-02 -399927.86 402436.8 | | 6 375689.03 | | 01244.58 | 19333 33 548 | | | | | | 379486.08 4538 | | | | | 9059.14 420356.1 | | | | | | |
| 2019-03 -381252.79 393315.7 | | 378786.03 | | 28882.34 | | | | | | | 384645.57 5097 | | 00 506527 07 | | | 5061.06 442130.3 | | | | | | - 500000 |
| 2019-04 -370753.07 415530.8 | | 397009.88 | | CONTRACTOR OF THE PARTY OF THE | 07000 00 695 | | | | | | | | | Maria Control of the | | | | | | | | 300000 |
| | | | | | | | | | | | 378757.27 4757 | | | | | 6115.94 447871.3 | | | | | | |
| 2019-05 -400962.47 390378.9 | | 374568.97 | | | 88225.00 557 | | | | | | 384097.56 4971 | | | | | 3673.13 439481.0 | | | | | | |
| 2019-06 -355098.56 393382.7 | | 356538.07 | | | 85592.00 609 | | | | | | 383020.21 5065 | | | | | 0843.66 443737.1 | | | | | | |
| 2019-07 -371321.18 388786.8 | | | | 10090.57 | | | 391473.15 4764 | | | | | | 50 485917.12 | | | 0746.95 445402.6 | | | | | | |
| 2019-08 -379545.80 389035.9 | | | | | 28333.33 <mark>669</mark> | | | | | | | | | | | 9132.88 452758.6 | | | | | | |
| 2019-09 -405394.46 399824.6 | 5 664048. | 391971.44 | 95565.11 42 | 29393.76 5 | 33333.33 602 | 2324.00 3 | 85402.96 4611 | 97.20 414104 | 4.60 416374.0 | 09 381081.08 | 380006.18 5089 | 51.49 426660 | 00 488492.18 | 464903.90 | 2363.62 38 | 7728.22 432242.9 | 1 516115.38 | 464394.62 44 | 43362.29 | 377112.55 37 | 71285.82 | |
| 2019-10 -395020.59 392133.9 | 3 584863 | 385091.65 | 504052.76 42 | 20752.24 | 99750.00 574 | 1055.50 3 | 89388.23 4144 | 18.38 429840 | 0.14 430025. | 50 415363.23 | 381436.64 5052 | 37.82 593250 | 00 500704.23 | 443143.67 50 | 7593.28 36 | 8809.72 449703.4 | 2 488766.52 | 485408.26 47 | 74899.55 | 370948.68 37 | 72335.16 | |
| 2019-11 -419996.14 390630.4 | | 9 364272.29 | 56296.00 43 | 39577.59 | 03000.00 691 | | 03730.88 4411 | | | | | 97 42 452444 | 36 508857.14 | | | 6069.93 451433.74 | | | | | | |
| 2019-12 -370797.54 389023.3 | | 375067.85 | | 20496.70 | 57377.60 571 | | | | | | 401474.97 5059 | 27.73 463583 | 33 489926.57 | | | 7518.89 442139.6 | | | | | | |
| 2020-01 -392976.74 410321.2 | | 382491.87 | 02787.74 41 | 16497.59 | 27236 50 472 | | 80553.82 540 | | | | 397387.21 5085 | | | | | 5753.73 436321.1 | | | | | | - 400000 |
| 2020-02 -419725.66 408863.2 | | 6 380575.71 | | 18640.32 | 500 | | | | | | 385627.33 4753 | | | | | 0823.80 451462.8 | | | | | | 100000 |
| 2020-03 -397885.28 385361.6 | | 0 374043.44 | | 11111.04 | 04500 00 526 | | | | | | | | | | | 4956.98 450382.0 | | | | | | |
| 2020-03 -357863.28 383361.6 | | 4 368400.00 | | 32603.83 | 85000.00 320 | | | | | | 427237 50 4851 | | | 487022 13 57 | | 9400.00 457893.5 | | | | | | |
| 2020-04 -432058.82 334163.1 | | | | 73066.67 5 | 05000.00 | aria di Kamanan 💆 | 131176.47 4330 | No. of Concession, Name of Street, or other party of the Concession, Name of Street, or other pa | No. of Concession, Name of | | 456209 78 4421 | | Section 1997 | 464206 29 | | 0833.33 432589.7 | | 450662.52 45 | | | | |
| | | | | | 200.00 407 | | | | | | | | | | | | | | | | | |
| 2020-06 -377749.59 395686.1 | | 383263.72 | | | 98000.00 582 | | | | | | 393080.91 4557 | | | 438286.93 | | 4757.81 445601.5 | | | | | | |
| 2020-07 -370383.19 390219.6 | | 399128.47 | | | 29333.33 637 | | | | | | 404325.51 4558 | | | | | 5769.16 453800.8 | | | | | | |
| 2020-08 -405964.99 407061.6 | | | | | 50000.00 600 | | | | | | 404240.37 4658 | | | | | 4821.76 452164.0 | | | | | | |
| 2020-09 -412595.35 430138.2 | 7 602486 | 424908.56 | 94658.58 43 | 32826.24 | 13000.00 610 | 0181.824 | 15576.46 4929 | 62.29 51796 | 1.29 451606. | 15 421519.76 | 427647.38 4866 | 33.70 482000 | 00 495790.00 | 469911.90 | 32617.23 39 | 4082.09 456242.6 | 7 490942.45 | 487930.54 4 | 37803.18 | 406594.66 37 | /6154.50 | |
| - : | _ | J | ÷ | (2) | + | + | (0 | 1 1 | 15 | | L | | 10 | | - | (5) | _ | 10 | + | 10 | - | |
| MO KIO | Ā | BATOK | Ā | ž | TIMAH | AREA | Š | 5 9 | 9 | EAST | S: | 5 5 | £ | 70 | 5 | S S | 6 | ŭ | РАУОН | DS | 5 | |
| ā ă | 主 | A | 8 | ₹ | Σ | A. | 3 | A A | < | B | Ne se | ₹ ≨ | 00 | 9 | 6 | ₹ ₹ | ō | ≥ | ₹ | 3 | 돐 | |
| | BISHAN | and and | MERAH | PANJANG | - | | ā | GEYLANG | 2 | 5 | 5 | PARADE | PASIR RIS | N | ST | δ ξ | 9 | AMPINE | <u>a</u> | 2 | YISHUN | |
| | | Ħ | = | 9 | BUKIT | ≨ | 丢 | GEYLANG | HOUGANG | N N | ž | E | Æ | PUNGGOL | Z | SEMBAWANG | SERANGOON | ₹ | OA | WOODLANDS | 389 | |
| | | | | | = | E | 0 | | 10 | JURONG | JURONG WEST | ₹ Ĭ | | - | 出 | ≥ Si | E | 9.500 | Ĕ | ŏ | | |
| | | 5 | 5 | ~ | | | | | | _ | | | | | | | | | | | | |
| ANG M | | BUKIT | BUKIT | Z | B | | ò | | | = | ⊇ : | 3 6 | | | 2 | S | S | | | ĕ | | |
| | | BU | BU | BUKIT | В | CENTRAL | AOH. | | | <u> </u> | 2, | MAR ANC | | | QUEENSTOWN | 22 | S | | | š | | |
| | | BU | BUI | BUKI | В | CEN | CHOA CHU KANG | | | = | ਕੁ | LLANG/WH MARINE | | | D | ŭ | S | | | W | | |
| | | BU | BUI | BUKI | B | CEN | СНО | | | Ξ. | 크 | MAR | | | No. | 55 | v | | | W | | |
| | | BO | BU | BUKI | М | CEN | СНО | | | tov | | KALLANG/WHAMPOA | | | O | 35 | v | | | W | | |

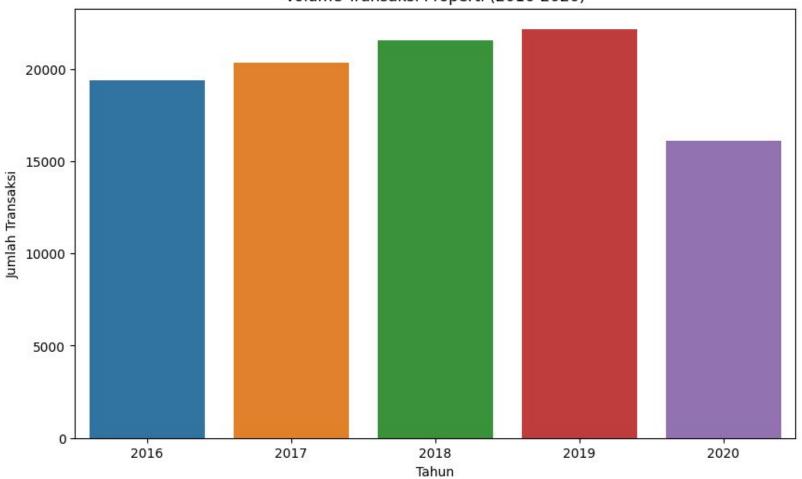








Volume Transaksi Properti (2016-2020)



Modelling

Modelling

- → Regression Modelling
- → Metode train test split :
 - ◆ Untuk nilai variable dari **x** ialah seluruh kolom pada dataset **kecuali** *resale price*
 - Untuk nilai variable dari y ialah kolom resale price (sebagai target)
 - Dilakukan pembagian terhadap dataset dimana 80% data train dan 20% data test.
 Adapun untuk random state: 42.
- → Model dievaluasi dengan melihat berapa besar masing masing **skor** untuk masing masing **MAE, MSE, dan RMSE**, dengan tujuan mencari selisih terkecil pada model.

- → Linear Regression, menjadi model pertama yang dicoba
- → Selanjutnya XGBoost, Random Forest, dan Ridge
- → Untuk improvement terhadap model awal, memakai *Feature Engineering* (*One Hot Encoding*), dan *Hyperparameter Tuning*

 \rightarrow

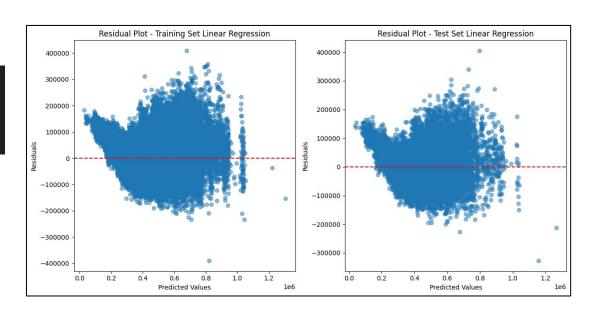
→ Model ; Linear Regression

Evaluation Metrics for Linear Regression: Training Mean Squared Error: 3648145211.614126 Test Mean Squared Error: 3633060933.009406

0.8394169683097157

Linear Regression:

MAE_Linear: 46335.65032444496 MSE_Linear: 3633060933.009406 RMSE Linear: 60274.87812521404



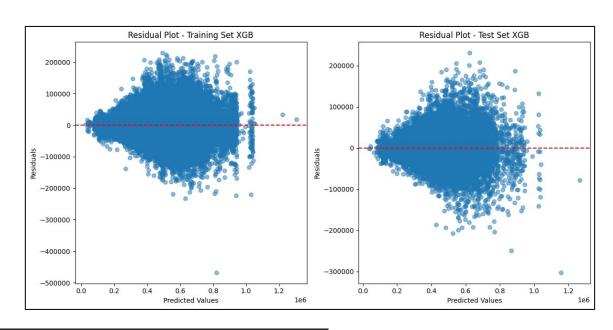
Linear Regression Test Score: 0.8394169683097157 Linear Regression Train Score: 0.8397787052710466

→ Model ; XGBoost

Evaluation Metrics for XGBoost Regressor: Training Mean Squared Error: 1336036003.3288689 Test Mean Squared Error: 1455572736.0569136 0.935662988562065

XGBoost:

MAE_xgb: 27787.537198978695 MSE_xgb: 1455572736.0569136 RMSE_xgb: 38151.96896697356



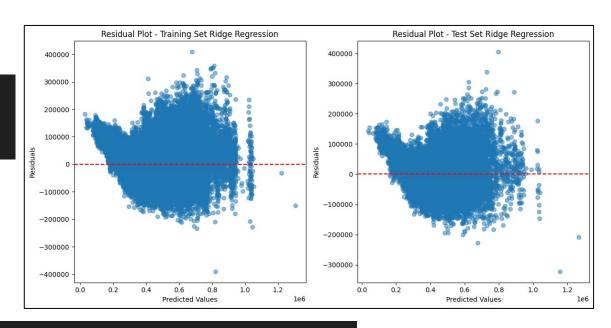
XGBoost Test Score : 0.935662988562065 XGBoost Train Score : 0.9413232188301145

→ Model ; Ridge Regression

Evaluation Metrics for Ridge Regression: Training Mean Squared Error: 3648201070.528127 Test Mean Squared Error: 3632830248.1277447 0.8394271646918698

Ridge Regression:

MAE_ridge: 46335.16036608969 MSE_ridge: 3632830248.1277447 RMSE_ridge: 60272.96448763529



Ridge Regression Test Score : 0.8394271646918698 Ridge Regression Train Score : 0.8397762520278218

Model ; Random Forest FINAL MODEL



Residual Plot - Training Set Random Forest

Evaluation Metrics for Random Forest Regressor: Training Mean Squared Error: 793096948.8641349 Test Mean Squared Error: 1322375793.6049771

0.9415503571543364

Random Forest:

MAE rf: 25866.844407430945 MSE_rf: 1320332194.8125448 RMSE rf: 36336.37564222036

200000 200000 100000 100000 -100000 -100000-200000 -300000 -200000 -400000-300000 1.0 1.2 0.6 1.2 0.2 1.0 Predicted Values Predicted Values

Residual Plot - Test Set Random Forest

Random Forest Test Score: 0.9415503571543364 Random Forest Train Score: 0.9651683218123953

Hyperparameter Tuning terhadap Final Model

Setelah mendapatkan **Final Model** (Random Forest), maka kita akan melakukan improvement kedua untuk mendapatkan akurasi yang lebih tinggi terhadap model dengan **Hyperparameter Tuning**. Algoritme yang kami gunakan adalah **GridSearchCV()**.

```
Hyperparameter terbaik: {'max_depth': None, 'min_samples_leaf': 1, 'min_samples_split': 10}
Mean Squared Error: 1268429164.4677923
R-squared: 0.9439348239761315
```

Mean Squared Error turun dari 1,3 miliar menjadi 1,2 miliar. R-Squared meningkat dari 0.941 menjadi 0.943

Conclusion

Conclusion

Interesting Insights

- 1. Daerah **Woodlands** dan **Yishun** terdapat banyak sekali apartemen yang cukup murah dengan kisaran harga 200000 400000
- 2. **Bukit Merah** merupakan daerah yang didominasi oleh apartemen berharga menengah ke atas.
- 3. Untuk apartemen berukuran antara 50-100 sqm, pilihan terbanyak terdapat di 5 daerah, yaitu **Yishun** (terbanyak), **Sengkang**, **Punggol**, **Bedok**, dan **Ang Mo Kio**
- Untuk apartemen berukuran sedang (100-150 sqm), pilihan terbanyak terdapat di 3 daerah, yaitu Woodlands, Yishun, dan Jurong West
- Adapula Woodlands juga mendominasi apartemen besar dengan ukuran antara 150-200 sqm
- 6. Semakin banyak jumlah room pada apartemen, maka harga resale price semakin tinggi, terlebih lagi terhadap apartemen dengan tipe executive dan multi-generation

Rekomendasi

- Strategi Pemasaran yang Disesuaikan; Sesuaikan strategi pemasaran untuk setiap daerah berdasarkan pada preferensi khusus yang diidentifikasi. Misalnya, beri penekanan pada harga terjangkau dalam materi promosi untuk Woodlands dan Yishun, dan soroti fitur desain minimalis untuk Yishun dan Sengkang.
- Apartemen Terjangkau ; Pertimbangkan untuk mempromosikan apartemen di Woodlands dan Yishun kepada pelanggan yang mencari opsi terjangkau. Sorot keefisienan biaya di daerah ini dan ketersediaan apartemen berukuran standar.
- 3. Apartemen Minimalis ; Tekankan Yishun dan Sengkang sebagai pilihan ideal bagi pelanggan yang mencari apartemen minimalis. Sediakan informasi tentang beragam pilihan yang tersedia di daerah ini sesuai dengan gaya hidup minimalis.

Terima kasih!

Ada pertanyaan?