

Analisis Prediksi Harga Rumah Sesuai Spesifikasi Menggunakan Metode Regresi Linear Berganda Berbasis Shiny R

Rolland Roy Hakiki¹, Renaldi Riyandi², Dealmus³, Yansensius Oktavianus Reins Dima⁴, Noviyanti P⁵
Teknologi Informasi Institut Shanti Bhuna
rolland@shantibhuana.ac.id¹, renaldiriyandi@shantibhuana.ac.id², dealmus@shantibhuana.ac.id³,
yansensius@shantibhuana.ac.id⁴, noviyanti@shantibhuana.ac.id⁵

Abstrak. Harga rumah dipengaruhi oleh berbagai faktor, seperti jumlah tempat tidur, ukuran rumah, dan kamar mandi. Penelitian ini bertujuan untuk menganalisis pengaruh faktor-faktor tersebut terhadap harga rumah menggunakan metode Regresi Linear Berganda. Permasalahan yang diangkat adalah bagaimana faktor jumlah tempat tidur dan ukuran rumah mempengaruhi harga rumah. Data diperoleh 505 data dari dataset Kaggle, dengan 271 data yang relevan. Penelitian ini mengembangkan model prediksi berbasis Regresi Linear Berganda dengan evaluasi akurasi model serta aplikasi berbasis Shiny R untuk analisis interaktif. Hasil penelitian menunjukkan bahwa faktor jumlah tempat tidur dan ukuran rumah berpengaruh signifikan terhadap harga rumah, dengan nilai koefisien regresi B_0 (5.33157), B_1 (-0.02481), dan B_2 (0.22060), konsisten pada perhitungan manual, RStudio, dan Shiny R. Studi ini menyadari adanya keterbatasan seperti asumsi linearitas dan variabel eksternal yang tidak dimasukkan. Penelitian ini merekomendasikan pengembangan model lebih lanjut dengan menambahkan variabel tambahan seperti lokasi dan usia bangunan, serta penggunaan metode yang lebih kompleks untuk meningkatkan akurasi. Temuan ini diharapkan dapat membantu pembeli dan penjual dalam pengambilan keputusan properti serta berkontribusi pada literatur ilmiah di bidang analisis harga properti.

Kata Kunci: Prediksi Harga Rumah, Regresi Linear Berganda, RStudio, Shiny R

1. Pendahuluan

Analisis harga rumah menjadi isu utama dalam industri *real estate*, terutama dengan adanya fluktuasi harga yang signifikan dalam beberapa tahun terakhir. Misalnya, menurut laporan dari Badan Pusat Statistik (BPS) Indonesia (2023), harga rumah di kota besar Indonesia mengalami kenaikan 10% dalam dua tahun terakhir, yang memperlihatkan volatilitas harga yang tinggi. Permintaan rumah yang terus meningkat didorong oleh faktor-faktor seperti urbanisasi, pertumbuhan penduduk, dan daya beli yang berbeda-beda antar wilayah. Dalam hal ini, ukuran rumah, jumlah kamar, dan kondisi pasar lokal memainkan peranan penting dalam menentukan harga rumah. Oleh karena itu, penting untuk memiliki model prediksi harga rumah yang akurat untuk membantu pengambilan keputusan, baik bagi pengembang, pembeli, maupun penjual properti. [1]

Penelitian ini bertujuan untuk mengembangkan model prediksi harga rumah dengan menggunakan Regresi Linear Berganda, dengan fokus pada dua variabel utama: jumlah kamar (*beds*) dan ukuran rumah (*size*). Pemilihan kedua variabel ini didasarkan pada bukti empiris yang menunjukkan adanya hubungan signifikan antara ukuran rumah dan jumlah kamar dengan harga rumah, yang sudah dibuktikan dalam berbagai penelitian sebelumnya. Regresi Linear Berganda memungkinkan analisis hubungan antara variabel dependen (harga rumah) dan beberapa variabel independen (jumlah kamar dan ukuran rumah) secara kuantitatif, sehingga memberikan hasil yang mudah dipahami dan fleksibel dalam aplikasinya. [2]

Salah satu faktor yang mempengaruhi harga rumah di kota besar seperti ukuran rumah dan lokasi geografis memiliki pengaruh signifikan. [3] Sementara itu, sumber lain menunjukkan bahwa jumlah kamar memiliki hubungan positif dengan harga rumah di kawasan pinggiran kota. [4] Namun, penelitian ini memfokuskan pada dua faktor utama tersebut dan mengabaikan faktor lokasi yang dianggap lebih kompleks dan memerlukan analisis lebih mendalam. [5]

Melalui pendekatan sistematis yang mencakup pengumpulan data, analisis, dan transformasi data, serta Pembangunan model regresi, penelitian ini bertujuan untuk menghasilkan model prediksi harga rumah yang praktis dan aplikatif. Kontribusi utama dari penelitian ini adalah mengembangkan aplikasi berbasis Shiny R yang memungkinkan pengguna untuk melakukan analisis interaktif, dengan menggabungkan visualisasi dan perhitungan model regresi yang dinamis, memberikan kemudahan dalam eksplorasi hasil prediksi harga rumah.[6]

Penelitian ini juga berfokus pada penerapan metode yang lebih kompleks dan memasukkan variabel tambahan seperti lokasi geografis dan kondisi ekonomi yang lebih mendalam dalam studi lanjutan, yang diharapkan dapat meningkatkan akurasi prediksi harga rumah.[7]

2. Landasan Teori

2.1 Rumah

Rumah umumnya dianggap sebagai tempat berlindung dan beristirahat yang memberikan rasa nyaman dan kenyamanan bagi penghuni. Selain fungsinya sebagai tempat tinggal, rumah juga memiliki peran penting dalam meningkatkan kualitas hidup dan mencerminkan status sosial penghuninya. Sebagai tempat untuk berkumpul dan bekerja, rumah berperan sebagai sarana untuk memulihkan kondisi fisik dan mental setelah aktivitas sehari-hari. Sebagai tambahan, Turner menyatakan bahwa rumah memiliki fungsi utama yang meliputi penunjang rasa aman, kesempatan, dan identitas keluarga. Dari sisi psikologi, rumah juga menjadi tempat yang memberi ketenangan dan kenyamanan bagi seluruh penghuni. Penelitian ini menggunakan Aplikasi RStudio untuk menganalisis harga rumah berdasarkan variabel-variabel seperti jumlah kamar dan ukuran rumah. Model prediksi harga rumah diharapkan memberikan informasi yang akurat untuk pengambilan keputusan yang lebih baik dalam transaksi jual beli properti. [8]

2.2 Harga Rumah

Penentuan harga rumah sangat dipengaruhi oleh beberapa faktor, di antaranya adalah luas bangunan, jumlah kamar, kondisi rumah, serta lokasi. Beberapa faktor eksternal, seperti harga NJOP (Nilai Jual Objek Pajak) dan tren pasar, juga turut memengaruhi harga jual rumah.[9] Pembeli biasanya memperhatikan berbagai faktor seperti jumlah kamar tidur, kamar mandi, serta kondisi bangunan untuk menentukan apakah rumah tersebut memenuhi kebutuhan mereka. Semakin strategis lokasi rumah dan semakin baik kondisinya, semakin tinggi harga rumah tersebut. Faktor-faktor ini menunjukkan bahwa harga rumah tidak hanya bergantung pada satu variabel, melainkan beberapa variabel yang saling berhubungan.[10]

2.3 Metode Regresi Linear Berganda

Regresi Linier Berganda adalah metode statistik yang digunakan untuk memodelkan hubungan antara satu variabel dependen (seperti harga rumah) dengan dua atau lebih variabel independen (seperti jumlah kamar dan ukuran rumah). Metode ini dapat digunakan untuk memprediksi nilai variabel dependen berdasarkan nilai variabel independen yang ada. Regresi Linear Berganda sangat efektif dalam melihat bagaimana faktor-faktor yang berbeda secara bersama-sama mempengaruhi harga rumah. Model regresi ini dapat digambarkan dengan persamaan sebagai berikut: [11]

$$Y = B_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n$$

Keterangan:

Y = Variabel Dependen (Harga Rumah)

X = Variabel Independen (Jumlah Kamar dan Ukuran Rumah)

B₀ = Konstanta (Intercept)

β = Koefisien yang menunjukkan seberapa besar pengaruh setiap variabel independent terhadap variabel dependen.

2.4 RStudio

RStudio adalah *Integrated Development Environment* (IDE) yang digunakan untuk pemrograman dalam bahasa R. Dengan RStudio, pengguna dapat dengan mudah melakukan analisis data, statistik, dan visualisasi hasil. RStudio menyediakan berbagai fitur, seperti paket analisis statistik, visualisasi grafis, dan pengolahan data yang memungkinkan pengguna untuk bekerja dengan data secara lebih efisien. Penggunaan RStudio dalam penelitian ini mempermudah dalam mengolah dan menganalisis data, serta menghasilkan laporan yang mudah dipahami.[12]

Untuk memulai, pengguna perlu menginstal R terlebih dahulu sebelum mengunduh dan menginstal RStudio dari situs resminya. Setelah itu, proyek baru dapat dibuat dengan menulis skrip dalam file berekstensi . R dan mengeksekusinya di *Source Panel*. Pengguna juga dapat menginstal paket tambahan dengan perintah `install.packages("nama_paket")`. Dengan kemudahan penggunaan dan fitur yang lengkap, RStudio menjadi alat yang sangat bermanfaat bagi siapa saja yang ingin mengolah dan menganalisis data dalam berbagai bidang.[13]

2.5 Shiny R

Shiny adalah *framework* yang memungkinkan pengguna untuk membangun aplikasi web interaktif menggunakan bahasa pemrograman R.[14] Dengan Shiny, aplikasi seperti *dashboard*, alat visualisasi, dan analisis data dapat dibangun tanpa memerlukan keterampilan HTML, CSS, atau *JavaScript* yang mendalam. Shiny sangat berguna untuk memvisualisasikan dan menganalisis data secara interaktif, memungkinkan pengguna untuk membuat aplikasi berbasis data yang dapat diakses oleh banyak orang.[15] Dalam penelitian ini, Shiny digunakan untuk membuat aplikasi yang memungkinkan analisis harga rumah secara dinamis, dengan visualisasi seperti scatter plot dan residual plot yang dapat diperbarui sesuai dengan inputan pengguna.

2.6 Metode Penelitian

Metode penelitian yang digunakan dalam penelitian ini adalah pendekatan kuantitatif dengan menggunakan analisis regresi linear berganda untuk memprediksi harga rumah berdasarkan variabel jumlah kamar dan ukuran rumah. Proses penelitian dimulai dengan memperoleh data yang relevan dari sumber yang dapat dipercaya yakni Kaggle. Data yang digunakan dalam penelitian ini hanya data sekunder karena sudah tersedia secara terbuka dan relevan untuk analisis harga rumah, dan tidak melibatkan data primer dikarenakan tujuan penelitian adalah untuk menganalisis hubungan antara harga rumah dan faktor-faktor yang telah dicatat dalam dataset sekunder. Variabel yang mencakup terdiri dari harga rumah (sebagai variabel dependen), jumlah kamar, dan ukuran rumah (sebagai variabel independen). Data ini kemudian dibersihkan dan diproses menggunakan aplikasi RStudio.

Tahapan penelitian meliputi beberapa langkah utama, yakni:

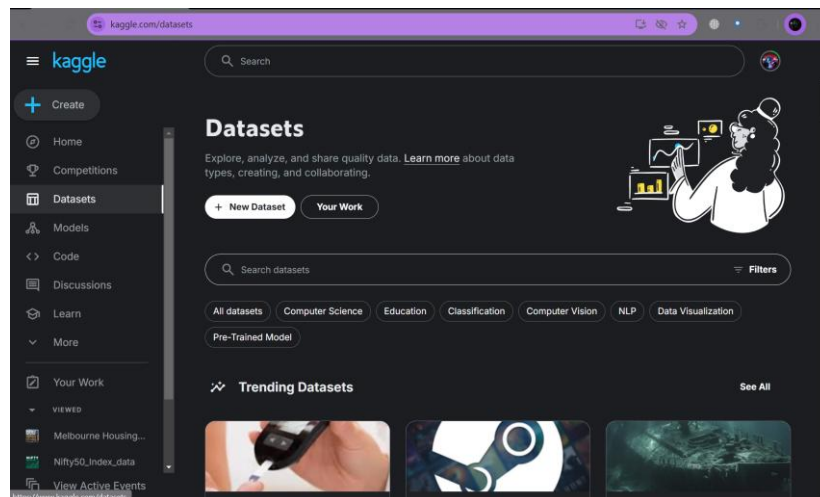
1. Pengumpulan Data
2. Pengolahan Data
3. Transformasi Data
4. Visualisasi Hubungan Antarvariabel
5. Analisis Penelitian
6. Validasi Hasil Model
7. Evaluasi Hasil
8. Implementasi Shiny R

Metode penelitian ini diharapkan dapat memberikan wawasan yang mendalam tentang faktor-faktor yang mempengaruhi harga rumah, serta memberikan solusi praktis melalui aplikasi berbasis data yang dapat digunakan oleh pemangku kepentingan dalam industry properti.

3. Pembahasan

Pengumpulan Data

Penelitian ini menggunakan dataset yang diambil dari Kaggle <https://www.kaggle.com/datasets>, yang mencakup berbagai atribut yang relevan dengan analisis harga rumah, termasuk jumlah kamar tidur (*beds*), ukuran rumah (*size*), dan harga rumah (*price*).



Gambar 1. Kaggle.com

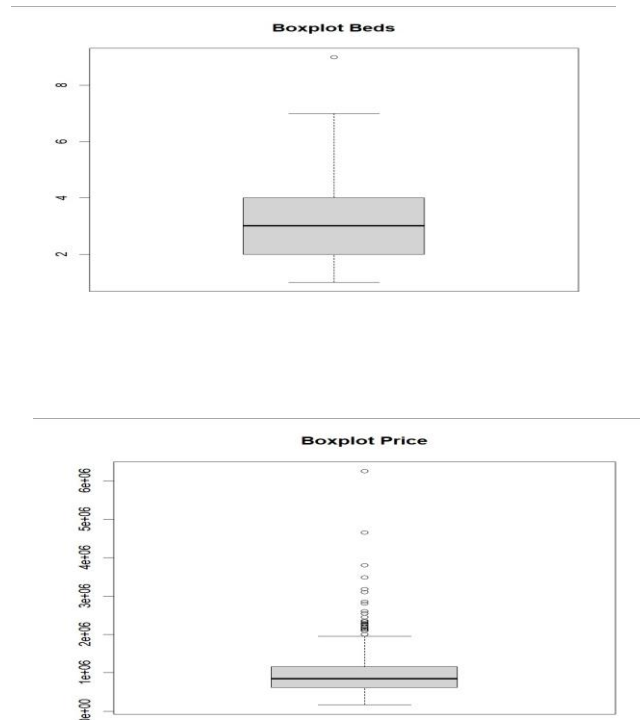
Gambar 1, menunjukkan tampilan dari dashboard awal website Kaggle, yakni dimana dataset yang belum diolah tersedia untuk didownload dan digunakan. Platform ini dimiliki oleh Google dan merupakan salah satu sumber utama bagi ilmuwan data yang mencari dataset untuk penelitian lebih lanjut.

Pengolahan Data

Gambar 2, di bawah ini menunjukkan data awal sebelum dilakukan *pre-processing* dan transformasi data. Dataset ini memiliki jumlah data sebanyak 505. Data awal ini mengandung nilai yang tidak teratur, *outlier*, skala yang tidak seragam, atau distribusi yang tidak sesuai untuk analisis tertentu. Selain itu, data ini juga dapat mengandung nilai yang hilang, duplikasi, atau inkonsistensi yang dapat mempengaruhi hasil analisis. Oleh karena itu, diperlukan proses pembersihan, standarisasi, dan normalisasi agar data lebih siap digunakan dalam pemodelan dan prediksi.

	beds	baths	size	price
1	3	3.0	2850	1175000
2	4	5.0	3040	1057500
3	3	1.0	1290	799000
4	3	2.0	2360	565000
5	3	3.5	1942	1187000
6	2	2.0	963	701000
7	1	1.0	756	480000
8	4	6.0	3300	1795000
9	4	2.0	2060	1025000
10	1	1.0	672	450000
11	3	2.5	1760	1135000
12	1	1.5	924	568000
13	3	2.0	2500	648000

Gambar 2. Data Awal



Gambar 3. Boxplot data awal

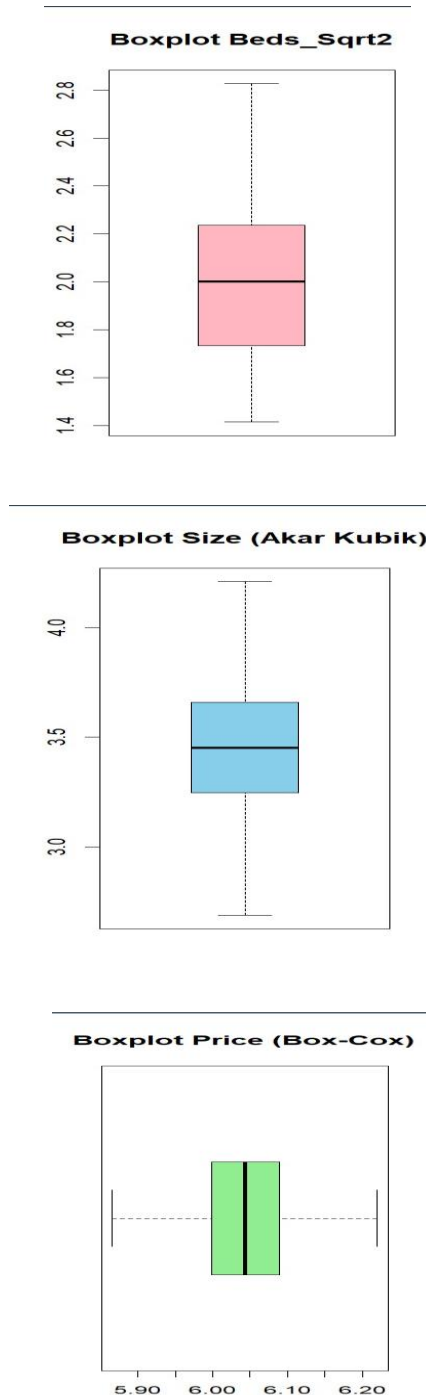
Gambar 3, menunjukkan visualisasi boxplot awal dari semua kolom yang ada di dalam data. Visualisasi ini menunjukkan bahwa data masih belum terdistribusi normal, yang terlihat dari adanya ketidakseimbangan dalam penyebaran nilai. Selain itu, masih terdapat banyak outlier (pencilan) pada hampir semua kolom, yang menandakan bahwa data mengandung nilai ekstrem dan belum memenuhi asumsi distribusi normal. Oleh karena itu, diperlukan proses *pre-processing* lebih lanjut untuk menangani outlier dan menyesuaikan distribusi data agar lebih optimal untuk analisis dan pemodelan.

Transformasi Data

	beds_sqrt2	size_cbrt	price_boxcox
1	2.000000	3.765368	6.091672
2	2.236068	3.806089	6.076965
3	2.000000	3.299375	6.036754
4	2.000000	3.648816	5.984787
5	2.000000	3.532170	6.093079
6	1.732051	3.142469	6.017430
7	1.414214	3.018239	5.959443
8	2.236068	3.858504	6.148657
9	2.236068	3.567066	6.072566
10	1.414214	2.959567	5.949247
11	2.000000	3.474712	6.086862
12	1.414214	3.120891	5.985601
13	2.000000	3.684031	6.005647

Gambar 4. Transformasi Data

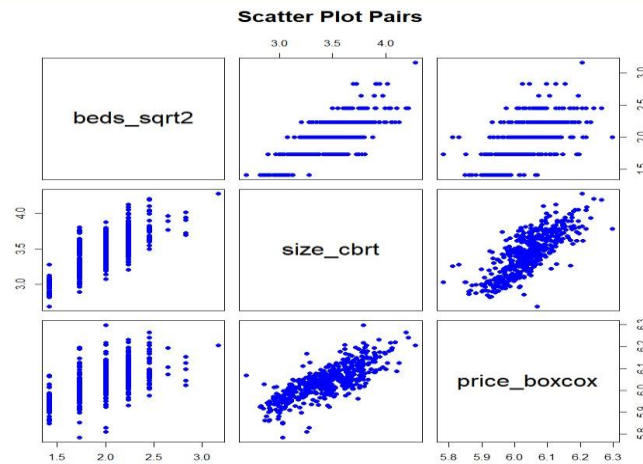
Gambar 4, menunjukkan tampilan data setelah dilakukan transformasi pada beberapa variabel untuk meningkatkan normalitas dan stabilitas varians. Variabel *beds* telah diubah ke dalam bentuk akar kuadrat (*beds_sqrt2*) guna mengurangi skewness dan memperhalus distribusi data. Variabel *size* ditransformasikan menggunakan akar kubik (*size_cbrt*) agar penyebaran data lebih merata dan tidak terlalu ekstrem. Sementara itu, variabel *price* mengalami transformasi Box-Cox (*price_boxcox*) untuk menyesuaikan distribusi dengan asumsi normalitas, sehingga analisis prediktif dapat memberikan hasil yang lebih akurat dan stabil.



Gambar 5. Boxplot setelah transformasi

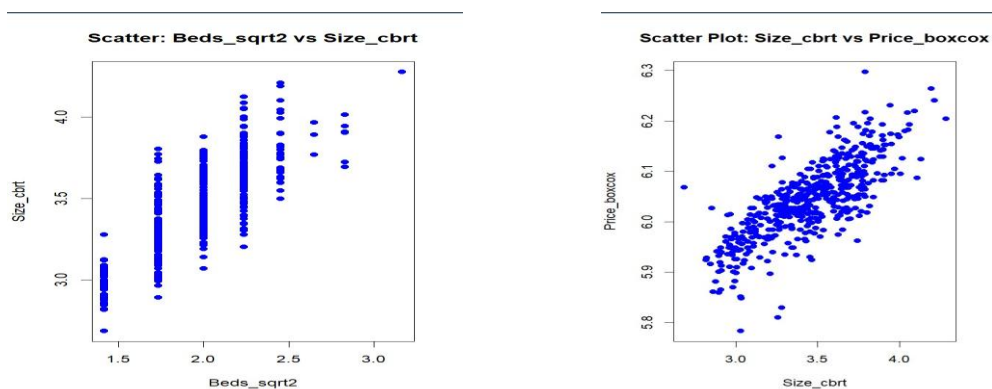
Gambar 5, di atas menunjukkan tampilan boxplot dari data yang telah melalui proses transformasi, di mana distribusi data telah mendekati normal. Pada setiap posisi, tidak terdapat outlier yang mencolok, menandakan bahwa transformasi yang dilakukan berhasil mengurangi pencilan dalam data. Dengan demikian, data yang telah ditransformasi ini lebih stabil dan siap digunakan untuk analisis lebih lanjut dalam penelitian selanjutnya.

Visualisasi Hubungan Antarvariabel



Gambar 6. Scatter Plot Pairs antara variabel

Gambar 6, menunjukkan visualisasi *scatter plot pairs* yang menggambarkan hubungan, korelasi, serta potensi adanya outlier dalam data. Dalam plot ini, titik-titik berwarna biru yang padat menunjukkan tingkat hubungan antara variabel yang digunakan. Pola penyebaran titik memberikan gambaran awal mengenai kekuatan dan arah korelasi antarvariabel, yang dapat membantu dalam mengidentifikasi hubungan linier di antara mereka. Visualisasi ini digunakan sebagai dasar sebelum memasukkan variabel-variabel tersebut ke dalam model regresi linier berganda untuk analisis lebih lanjut.

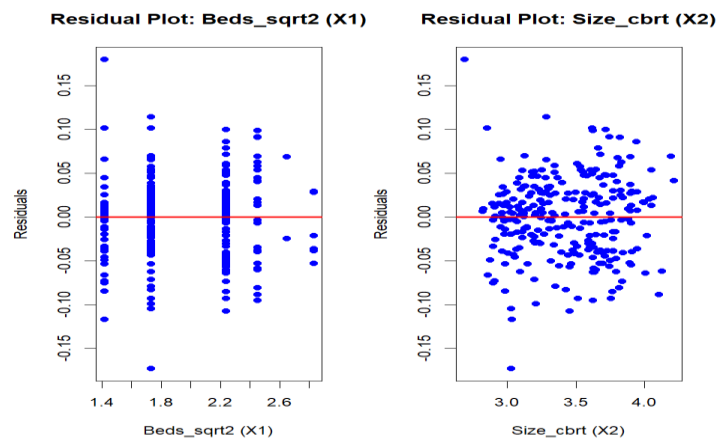


Gambar 7. Scatter Plot perbandingan antar variabel.

Gambar 7, menunjukkan *scatter plot*, di mana setiap titik dalam plot ini mewakili satu pengamatan dalam data. Sumbu horizontal (X) menunjukkan nilai dari satu variabel, sedangkan sumbu vertikal (Y) menunjukkan nilai dari variabel lainnya. Scatter plot digunakan untuk mengidentifikasi pola hubungan

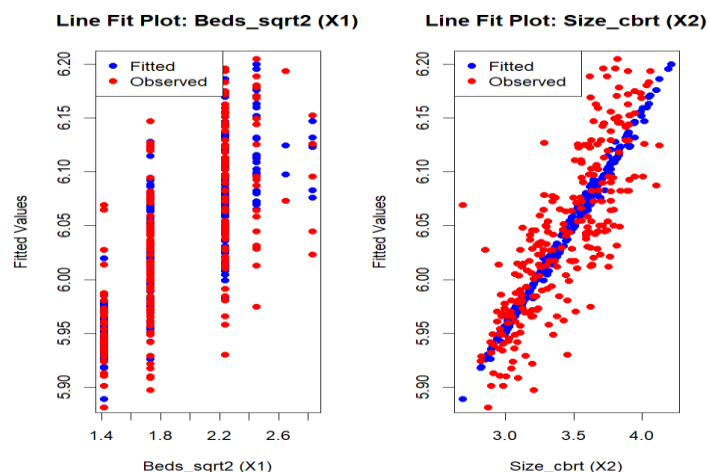
antarvariabel, baik itu hubungan linier maupun non-linier. Selain itu, visualisasi ini juga membantu dalam menentukan jenis korelasi yang terjadi, apakah positif, negatif, atau tidak ada hubungan sama sekali. Dengan demikian, *scatter plot* menjadi alat yang penting dalam eksplorasi data sebelum melakukan analisis lebih lanjut, seperti regresi atau pemodelan statistik lainnya.

Analisis Penelitian



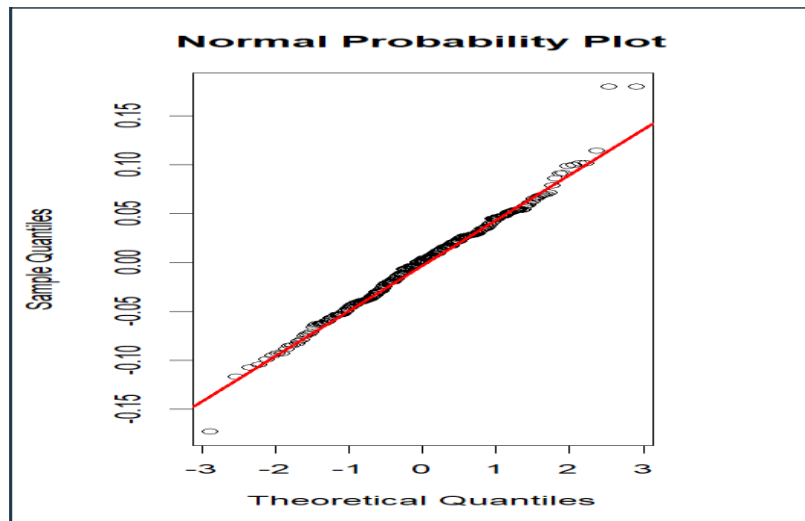
Gambar 8. Residuals Plot X1 dan X2

Gambar 8, menunjukkan visualisasi *residual plot* yang menggambarkan selisih antara nilai aktual dan nilai prediksi model (*residuals*) terhadap masing-masing variabel prediktor, yaitu *X1* dan *X2*. Tujuan utama dari visualisasi ini adalah untuk memeriksa apakah residual tersebar secara acak di sekitar garis nol, yang ditunjukkan oleh garis horizontal merah. Jika residual tersebar secara acak tanpa pola tertentu, maka asumsi linearitas dan homoskedastisitas dalam regresi terpenuhi. Namun, jika terlihat pola tertentu, seperti tren melengkung atau penyebaran residual yang semakin melebar atau menyempit, hal ini dapat mengindikasikan pelanggaran terhadap asumsi model regresi, yang mungkin memerlukan transformasi variabel atau pendekatan pemodelan alternatif



Gambar 9. Visualisasi Line Fit Plot X1 dan X2

Gambar 9, di atas menampilkan visualisasi *Line Fit Plot*, yang digunakan untuk memvisualisasikan sejauh mana model regresi sesuai dengan data aktual. Plot pertama menunjukkan hubungan antara *beds_sqrt2* (sumbu X) dan nilai prediksi model (*fitted values*, sumbu Y), sementara plot kedua menunjukkan hubungan serupa untuk *size_cbrt*. Dalam plot ini, titik biru merepresentasikan nilai yang diestimasi oleh model berdasarkan variabel-variabel prediktor, sedangkan titik merah menunjukkan nilai aktual dari *price_boxcox*. Legend di pojok kiri atas memberikan informasi yang jelas untuk membedakan antara nilai prediksi (titik biru) dan nilai aktual (titik merah).



Gambar 10. Normal Probability Plot

Gambar 10, merupakan Normal Probability Plot (QQ plot) untuk memeriksa apakah residual dari model (*resid(model)*) mengikuti distribusi normal. Jika titik-titik pada plot mendekati garis merah, residual dianggap distribusi normal, yang merupakan asumsi penting dalam banyak analisis statistik. Grafik ini juga membantu mengevaluasi kualitas model regresi dengan melihat seberapa baik data menyebar di sekitar garis regresi dan mengidentifikasi pola residu untuk memastikan bahwa pola penyimpangan acak.

Validasi Hasil Model

SUMMARY OUTPUT		Model 1	Model 2	Model 3	Model 4	Model 5
R Square: 0.8977		0.8977	0.8977	0.8977	0.8977	0.8977
Adjusted R Square: 0.8955		0.8955	0.8955	0.8955	0.8955	0.8955
F-Statistic: 15.77		15.77	15.77	15.77	15.77	15.77
P-Value: 0.000000		0.000000	0.000000	0.000000	0.000000	0.000000
Standard Error: 0.000000		0.000000	0.000000	0.000000	0.000000	0.000000
Observations: 270		270	270	270	270	270
ANOVA						
Regression		1	1	1	1	1
Residual		268	268	268	268	268
Total		269	269	269	269	269
Coefficients						
Intercept		5.33157	5.33157	5.33157	5.33157	5.33157
X1		-0.02481	-0.02481	-0.02481	-0.02481	-0.02481
X2		0.00000	0.00000	0.00000	0.00000	0.00000
X3		0.00000	0.00000	0.00000	0.00000	0.00000
X4		0.00000	0.00000	0.00000	0.00000	0.00000
X5		0.00000	0.00000	0.00000	0.00000	0.00000

Rumus perhitungan manual regresi linier berganda

$$Y = b_0 + b_1x_1 + b_2x_2$$

Diketahui :

$$b_0 = 5,33157$$

$$b_1 = -0,02481$$

$$b_2 = 0,22060$$

$$x_1 = 1.414213562$$

$$x_2 = 3.018239042$$

$$Y = 5,33157 + (-0,02481 \cdot 1.414213562) + (0,22060 \cdot 3.018239042)$$

$$Y = 5,33157 + (-0,03508663847) + (0,66582353266)$$

$$Y = 5,96230689419$$

Gambar 11. Perhitungan Manual Excel dan Perhitungan Manual menggunakan Data

Gambar 11, merupakan validasi hasil model perhitungan manual menggunakan Excel dan mendapatkan hasil yang sama dengan data di RStudio. Perhitungan manual tersebut menghasilkan data yaitu b_0

= 5.331569267, $b_1 = -0.02480694$, dan $b_2 = 0.220598578$.. Dilakukan juga perhitungan manual menggunakan data yang terdistribusi normal, dengan nilai x_1 dan x_2 , masing masing 1.414213562 dan 3.018239042, menghasilkan nilai prediksi sebesar $Y = 5,96230689419$, yang dapat disimpulkan bahwa hasil perhitungan manual berdasarkan data terdistribusi ini menghasilkan nilai prediksi rumah yang hampir sama, pada data transformasi senilai 5.959442603 sedangkan hasil perhitungan manual senilai $Y = 5,96230689419$.

```
Call:
lm(formula = price_boxcox ~ beds_sqrt2 + size_cbrt, data = data)

Residuals:
    Min       1Q   Median       3Q      Max
-0.17283 -0.03402  0.00316  0.02837  0.17993

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5.33157    0.03447   154.695  <2e-16 ***
beds_sqrt2   -0.02481    0.01464   -1.694   0.0914 .
size_cbrt     0.22060    0.01602   13.771  <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.04766 on 267 degrees of freedom
Multiple R-squared:  0.6676,    Adjusted R-squared:  0.6651
F-statistic: 268.1 on 2 and 267 DF,  p-value: < 2.2e-16
```

Gambar 12. Hasil RStudio

Gambar 12, dapat dilihat hasil perbandingan antara persamaan yang dihasilkan oleh RStudio dan Excel yang ditampilkan pada **Gambar 11**. Nilai yang didapatkan untuk Intercept (b_0), koefisien b_1 (beds_sqrt2), b_2 (size_cbrt), serta nilai *R-Square* menunjukkan kesamaan yang signifikan. Hal ini menandakan bahwa transformasi data yang dilakukan telah berhasil, karena hasil yang diperoleh dari kedua perangkat lunak tersebut konsisten. Dengan demikian, data yang telah ditransformasi dapat digunakan sepenuhnya untuk analisis lebih lanjut, mengingat distribusi data sudah mendekati normal dan memenuhi asumsi model regresi.

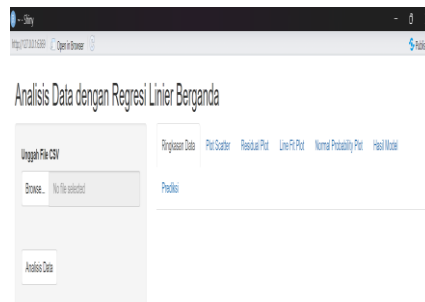
Evaluasi Hasil

Perbandingan Hasil Koefisien

Sumber	Intercept (B0)	beds_sqrt2 (B1)	size_cbrt (B2)
Manual	5.33157	-0.02481	0.22060
RStudio	5.33157	-0.02481	0.22060
Excel	5.33156	-0.02481	0.22060

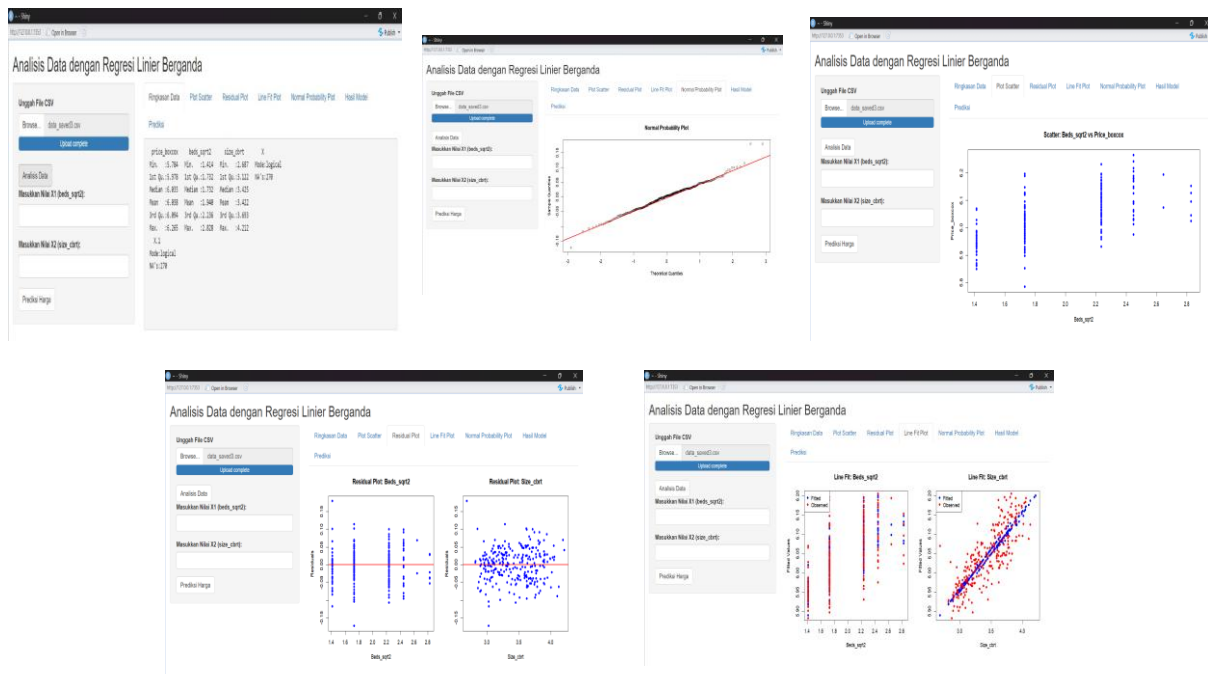
Dapat dilihat B0, B1 dan B2 adalah nilai intercept dan koefisien dari ketiga sumber sangat konsisten dan hampir sama persis disemua metode, bedanya hanya sangat kecil di Intercept (5.33156) dibandingkan Manual dan RStudio (5.33157). Jadi bisa dikatakan Model regresi linear berganda yang dihitung secara manual, dengan RStudio, maupun Excel menghasilkan koefisien yang sama, membuktikan bahwa proses analisis data dilakukan dengan benar dan akurat serta proses perhitungannya konsisten dan valid.

Implementasi Shiny R



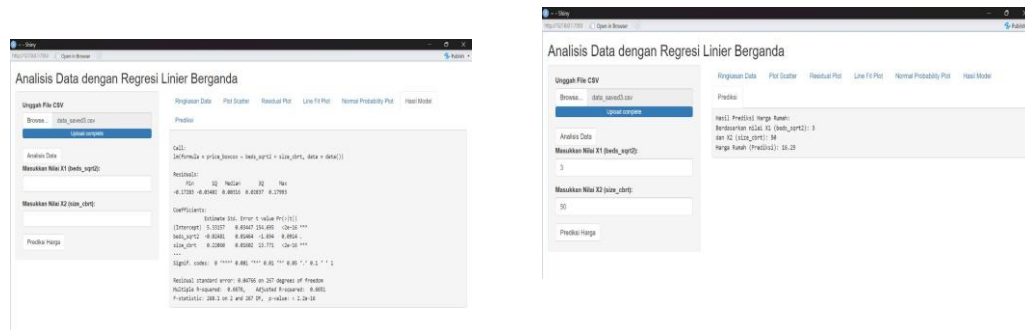
Gambar 13. Tampilan Awal Shiny R

Gambar 13, menunjukkan tampilan awal dari program Shiny R, yang menyediakan berbagai menu interaktif untuk mempermudah analisis data. Menu-menu yang tersedia meliputi: *Unggah File (CSV)* untuk memuat data, *Tombol Analisis Data* untuk memulai proses analisis, serta berbagai pilihan visualisasi dan hasil analisis seperti *Ringkasan Data*, *Plot Scatter*, *Residual Plot*, *Line Fit Plot*, *Normal Probability Plot*, serta *Hasil Model* dan *Prediksi*. Program ini dirancang untuk memberikan antarmuka yang mudah digunakan bagi pengguna dalam melakukan eksplorasi dan analisis data secara komprehensif.



Gambar 14. Visualisasi hasil Shiny R.

Gambar 14, menampilkan visualisasi dari program Shiny R yang dihasilkan setelah implementasi data yang telah ditransformasi. Hasil dari implementasi ini meliputi berbagai visualisasi interaktif seperti *Ringkasan Data*, *Plot Scatter*, *Residual Plot*, *Line Fit Plot*, *Normal Probability Plot*, serta *Hasil Model* dan *Prediksi*. Hasil visualisasi yang ditampilkan menggunakan Shiny R ini konsisten dengan hasil yang sebelumnya diperoleh di RStudio, menunjukkan bahwa kedua platform menghasilkan output yang serupa dan dapat digunakan untuk analisis lebih lanjut.



Gambar 15. Hasil Model dan Hasil Prediksi

Gambar 15, merupakan hasil model perhitungan versi Shiny R, didapatkan hasil yang sama antara perhitungan di RStudio dan perhitungan manual menggunakan Excel. Dan terdapat menu untuk memprediksi harga rumah sesuai dengan variabel beds dan size yang ingin diprediksi, angka yang digunakan untuk prediksi di setarakan dengan hasil data yang sudah terdistribusi normal pada summary data. Setelah data di prediksi menggunakan Shiny R dan di prediksi menggunakan perhitungan manual, menghasilkan nilai yang sama, maka penelitian ini berhasil.

4. Kesimpulan

Penelitian ini berhasil menganalisis dan memprediksi harga rumah menggunakan Metode Regresi Linear Berganda berbasis Shiny R. Dengan memanfaatkan dataset dari Kaggle yang terdiri dari 299 data, penelitian ini meneliti hubungan antara variabel independen, yaitu jumlah tempat tidur (beds_sqrt2) dan ukuran rumah (size_cbrt), terhadap variabel dependen, yaitu harga rumah yang telah ditransformasi (price_boxcox).

Hasil analisis menunjukkan bahwa nilai yang diperoleh dari perhitungan Manual, perhitungan menggunakan RStudio dan Shiny R terdapat kesamaan sebagai berikut nilai B0 (5.33157), B1 (-0.02481) dan B2 (0.22060) dapat dikatakan bahwa menggunakan model Regresi linear berganda dapat memodelkan pengaruh kedua variabel tersebut terhadap harga rumah dengan tingkat akurasi yang memadai. Secara spesifik, penelitian ini mengkonfirmasi bahwa jumlah tempat tidur dan ukuran rumah memiliki pengaruh signifikan terhadap harga rumah. Meningkatnya jumlah tempat tidur cenderung meningkatkan harga rumah, karena lebih banyak tempat tidur biasanya menunjukkan kapasitas dan kenyamanan yang lebih baik. Selain itu, ukuran rumah yang lebih besar juga berkontribusi pada peningkatan harga, karena memberikan lebih banyak ruang bagi penghuni. Proses analisis mencakup pengumpulan data, transformasi data, dan evaluasi model melalui berbagai visualisasi, termasuk scatter plot, residual plot, dan normal probability plot. Meskipun penelitian ini memberikan wawasan yang berharga bagi pembeli dan penjual properti dalam pengambilan keputusan, terdapat beberapa keterbatasan yang perlu diperhatikan, seperti asumsi linearitas dan variabel eksternal yang tidak dimasukkan dalam model. Oleh karena itu, disarankan untuk penelitian lebih lanjut untuk memasukkan variabel tambahan dan menerapkan metode analisis yang lebih kompleks guna meningkatkan akurasi prediksi.

5. Referensi

- [1] G. K. Uyanık and N. Güler, "A Study on Multiple Linear Regression Analysis," *Procedia - Soc. Behav. Sci.*, vol. 106, pp. 234–240, 2013, doi: 10.1016/j.sbspro.2013.12.027.
- [2] Sudariana and Yoedani, "Analisis Statistik Regresi Linier Berganda," *Seniman Trans.*, vol. 2, no. 2, pp. 1–11, 2022.
- [3] H. Hakim, D. Kamil, and B. Alatas, "Pendekatan Machine Learning untuk Estimasi Harga Rumah dengan Regresi Linier," vol. 1, no. 1, pp. 18–22, 2025.
- [4] P. Pemberian *et al.*, "Jurnal Riset Ilmiah," vol. 1, no. 7, pp. 565–570, 2024.
- [5] M. Regresi Linier Berganda Kualitas Super Member Supermall Terhadap Peningkatan Jumlah Pengunjung Pada Supermall Karawang Popy Purnamasari Wahid Suyitno, P. Studi Sistem Informasi, S. Nusa Mandiri Jakarta JIDamai no, and W. Jati Barat Jakarta-Selatan, "Copyright@2015. P2M STMIK BINA INSANI," *Bina Insa. Ict J.*, vol. 2, no. 2, pp. 101–116, 2015.
- [6] T. N. Padilah and R. I. Adam, "Analisis Regresi Linier Berganda Dalam Estimasi Produktivitas Tanaman Padi Di Kabupaten Karawang," *FIBONACCI J. Pendidik. Mat. dan Mat.*, vol. 5, no. 2, p. 117, 2019, doi: 10.24853/fbc.5.2.117-128.
- [7] A. F. L. Putra, J. Junaidi, Z. Situmorang, and A. H. Nasyuha, "Regresi Linier Berganda Untuk Memprediksi Jumlah Nasabah," *J. Sci. Soc. Res.*, vol. 5, no. 2, p. 236, 2022, doi: 10.54314/jssr.v5i2.915.
- [8] A. K. Khotimah, A. A. Rahman, M. Z. Alam, Y. H. Nur, and T. R. Afi, "Analisis Regresi Linier Berganda Dalam Estimasi Indeks Pembangunan Manusia di Indonesia Multiple Linear Regression Analysis In Estimating The Human Development Index In Indonesia," vol. 15, no. November, pp. 90–99, 2024, doi: 10.30872/eksponensial.v15i2.1318.
- [9] A. Bintang and N. Agustina, "Analisis Variabel yang Memengaruhi Harga Properti Residensial Tipe Kecil di 16 Kota Indonesia Tahun 2015-2019," *Semin. Nas. Off. Stat.*, vol. 2021, no. 1, pp. 343–352, 2021, doi: 10.34123/semnasoffstat.v2021i1.871.
- [10] G. Khalda Rifdan, N. Rahaningsih, A. Bahtiar, I. Ali, and N. Dienwati Nuris, "Ramalan Penjualan Rumah Menggunakan Algoritma Linear Regresi Di Tebet Jakarta Selatan," *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 8, no. 2, pp. 1847–1851, 2024, doi: 10.36040/jati.v8i2.9022.
- [11] * Rafif *et al.*, "Analisis Prediksi Harga Rumah di Bandung Menggunakan Regresi Linear Berganda," *J. Creat. Student Res.*, vol. 1, no. 6, pp. 395–404, 2023.
- [12] F. M. Puri, S. Tri, A. Ramadhani, D. A. Istiqomah, V. A. Windarni, and K. R. Hidayat, "Analisis Rekomendasi Pembuatan Produk Menggunakan RStudio Dan Twitter (Studi Kasus : Git Solution) rekomendasi produk dan membantu perusahaan lebih memahami kebutuhan konsumen . Data Penelitian berjudul " Event Recommendation using Social Media " menyimpulkan bahwa peneliti berhasil memprediksi popularitas acara di masa depan dan merekomendasikan Trend Based on Social Media Big Data Analysis Using K-Mean Clustering Algorithm (A Case Study on Yogyakarta Culinary Industry) ", menunjukkan hasil kesimpulan , penggalan akurasi data 70 % -80 % dibandingkan dengan penjualan pada beberapa restoran di Yogyakarta penelitian , meliputi jenis".
- [13] A. S. Amir, M. A. Tiro, and Ruliana, "Development of R Package for Regression Analysis with User Friendly Interface," *ARRUS J. Math. Appl. Sci.*, vol. 2, no. 1, pp. 23–35, 2022, doi: 10.35877/mathscience728.
- [14] S. Ma'arif, I. Hafidz, and R. P. K., "Framework Visualisasi Geografi Yang Dengan Bahasa Pemrograman R," *J. Tek. Its*, vol. 4, no. 1, pp. 1–4, 2015, [Online]. Available: <http://personalpages.manchester.ac.uk/mbs/julia.handl/v>
- [15] H. Dan, L. Di, and K. Barat, "UNTUK MEMONITOR POTENSI AREA RAWAN KEBAKARAN," vol. 8, no. 4, pp. 7801–7808, 2024.