

**LAPORAN TUGAS AKHIR
REGRESI TERAPAN**

**PEMODELAN OPTIMAL UNTUK PREDIKSI KUAT TEKAN
BETON MENGGUNAKAN PENDEKATAN REGRESI**



Kelompok 11

Shofia Ishma Najiyya	162112133087
Chelsea Dheirranaya Sitinjak	164231051
Cuthbert Young	164231052
Nadra Cinta Ruth Simarmata	164231053

**PROGRAM STUDI S1 TEKNOLOGI SAINS DATA
FAKULTAS TEKNOLOGI MAJU DAN MULTIDISIPLIN
UNIVERSITAS AIRLANGGA
GASAL 2024/2025**

DAFTAR ISI

DAFTAR ISI	ii
DAFTAR TABEL	iii
DAFTAR GAMBAR.....	iv
BAB I PENDAHULUAN	5
1.1 Latar Belakang	5
1.2 Rumusan Masalah.....	5
1.3 Tujuan Penelitian	6
BAB II TINJAUAN PUSTAKA	7
2.1 Tinjauan Pustaka Statistika	7
2.1.1 Statistika	7
2.1.2 Uji Asumsi Klasik	7
2.1.3 Outlier.....	7
2.1.4 Analisis Regresi.....	7
2.1.5 Residual Standard Error (RSE)	8
2.2 Tinjauan Pustaka Non-Statistika.....	8
2.2.1 Beton	8
2.2.2 Agregat	8
2.2.3 Air.....	8
2.2.4 <i>Superplasticizer</i>	8
2.2.5 Kekuatan tekan beton	9
BAB III METODOLOGI	10
3.1 Sumber Data.....	10
3.2 Variabel Penelitian.....	10
3.3 Metode Penelitian	11
3.3.1 <i>Exploratory Data Analysis (EDA)</i>	11
3.3.2 Ordinary Least Squares (OLS).....	11
3.3.3 Uji Hipotesis.....	11
3.3.4 <i>Robust Regression</i>	12
3.3.5 <i>Generalized Least Squares (GLS)</i>	12
3.4 Prosedur Analisis	12
BAB IV ANALISIS DAN PEMBAHASAN	13
4.1 Exploratory Data Analysis (EDA).....	13
4.1.1 Statistika Deskriptif.....	13
4.1.2 Pengecekan Outlier	13
4.1.3 Visualisasi Data.....	15
4.2 Analisis Regresi	17
4.2.1 Model Regresi OLS (<i>Ordinary Least Square</i>)	17
4.2.2 Pengujian Hipotesis.....	19
4.2.3 Pengujian Asumsi.....	20
4.2.4 Alternatif Model Regresi.....	22
4.3 Pemilihan Model Regresi Terbaik	24
BAB V KESIMPULAN DAN SARAN	26
5.1 Kesimpulan	26
5.2 Saran	26
DAFTAR PUSTAKA.....	28
LAMPIRAN	29

DAFTAR TABEL

Tabel 1. Variabel Penelitian	10
Tabel 2. Statistika Deskriptif	13
Tabel 3. Variabel <i>Outliers</i>	14
Tabel 4. Model Regresi Terbaik	25

DAFTAR GAMBAR

Gambar 1. <i>Boxplot</i>	14
Gambar 2. Histogram	15
Gambar 3. Scatter Plot.....	16
Gambar 4. <i>Heatmap</i>	17
Gambar 5. <i>OLS</i>	18
Gambar 6. Uji F	19
Gambar 7. Uji t	19
Gambar 8. Uji Normalitas Residual: <i>Shapiro-Wilk</i>	20
Gambar 9. Uji Homoskedastisitas: <i>Breusch-Pagan</i>	21
Gambar 10. Uji Multikolinearitas: <i>VIF</i>	21
Gambar 11. Uji Autokorelasi: <i>Durbin-Watson</i>	22
Gambar 12. Model Regresi <i>Robust</i>	23
Gambar 13. Model Regresi GLS	24

BAB I

PENDAHULUAN

1.1 Latar Belakang

Kekuatan tekan beton adalah salah satu indikator utama untuk menilai kualitas beton, terutama dalam konstruksi besar seperti jembatan, gedung bertingkat tinggi, dan infrastruktur berat lainnya. Beton yang berkualitas tinggi mampu menahan beban berat dan menjamin stabilitas struktur. Kekuatan tekan beton dipengaruhi oleh banyak faktor, seperti jenis dan proporsi material yang digunakan (semen, agregat kasar dan halus, air, serta bahan tambahan seperti superplasticizer) hingga usia beton selama proses hidrasi. Usia beton memainkan peran penting karena kekuatan tekan meningkat seiring waktu, menjadikan proses prediksi kekuatan tekan semakin kompleks.

Kesalahan dalam memprediksi kekuatan tekan dapat menyebabkan kegagalan struktur yang berbahaya dan mahal. Oleh karena itu, pendekatan analisis data modern, seperti Ridge Regression, Robust Regression, dan Weighted Least Squares Regression, menjadi solusi yang semakin populer. Metode ini memungkinkan pemodelan hubungan yang kompleks antara variabel serta mengatasi tantangan seperti multikolinearitas dan heteroskedastisitas dalam data beton.

Dalam konstruksi modern, prediksi kekuatan tekan beton yang akurat sangat diperlukan untuk memenuhi kebutuhan efisiensi dan inovasi. Model prediksi yang andal dapat membantu menentukan komposisi material yang optimal, mengurangi pemborosan, dan memantau kualitas beton di berbagai tahap pembangunan. Hal ini tidak hanya meningkatkan efisiensi biaya tetapi juga memastikan standar keamanan yang tinggi.

Penelitian ini bertujuan untuk mengembangkan model regresi optimal yang mampu memprediksi kekuatan tekan beton secara akurat dengan mempertimbangkan interaksi antara komposisi material, bahan tambahan, dan usia beton. Dengan pendekatan ini, diharapkan dapat memberikan kontribusi signifikan dalam pengambilan keputusan yang lebih baik di bidang konstruksi.

1.2 Rumusan Masalah

1. Bagaimana pengaruh semua variabel dalam data (semen, agregat kasar, agregat halus, air, superplasticizer, dan usia beton) terhadap kekuatan tekan beton?

2. Metode regresi mana yang memberikan hasil prediksi paling optimal untuk kekuatan tekan beton?
3. Bagaimana memilih model regresi terbaik untuk meningkatkan akurasi prediksi kekuatan tekan beton?

1.3 Tujuan Penelitian

1. Menganalisis pengaruh semua variabel dalam data (komposisi material dan usia beton) terhadap kekuatan tekan beton.
2. Menerapkan berbagai metode regresi (Ridge, Robust, Generalized Least Squares, Weighted Least Squares) untuk memprediksi kekuatan tekan beton.
3. Memilih model regresi terbaik untuk optimasi prediksi kekuatan tekan beton.

BAB II

TINJAUAN PUSTAKA

2.1 Tinjauan Pustaka Statistika

2.1.1 Statistika

Statistika adalah cabang ilmu matematika yang berfokus pada pengumpulan, pengolahan, analisis, dan interpretasi data untuk menarik kesimpulan atau membuat keputusan. Ghazali (2011) juga menyatakan bahwa dengan menggunakan metode statistik yang tepat, kita dapat membuat keputusan yang lebih objektif dan akurat berdasarkan data yang ada, yang sangat penting dalam penelitian ilmiah, ekonomi, dan berbagai disiplin ilmu lainnya.

2.1.2 Uji Asumsi Klasik

Uji Asumsi klasik meliputi normalitas residual, tidak adanya multikolinearitas antar variabel independen, homoskedastisitas (konstannya varians residual), dan tidak adanya autokorelasi pada residual. Ketika salah satu asumsi tersebut tidak terpenuhi, maka hasil regresi dapat menjadi tidak akurat atau bias. (Gujarati, 2003).

2.1.3 Outlier

Outlier adalah data yang memiliki nilai yang jauh berbeda atau ekstrem dibandingkan dengan sebagian besar data dalam suatu set data. Penting untuk mendeteksi dan menangani outlier, baik dengan menghapusnya, melakukan transformasi data, atau menggunakan teknik statistik yang robust terhadap outlier.

2.1.4 Analisis Regresi

Analisis regresi adalah metode statistik yang digunakan untuk mengidentifikasi hubungan antara variabel dependen dengan satu atau lebih variabel independen.

2.1.4.1 Ordinary Least Squares (OLS)

Ordinary Least Squares (OLS) adalah metode statistik yang digunakan dalam analisis regresi linier untuk mengestimasi parameter model regresi. OLS bekerja dengan meminimalkan jumlah kuadrat dari selisih antara nilai yang diamati dan nilai yang diprediksi oleh model.

2.1.4.2 Robust Regression

Robust regression adalah metode dalam analisis regresi yang digunakan untuk mengatasi masalah data yang memiliki pengaruh besar, data pencilan (outliers), atau ketika asumsi-asumsi dasar regresi linier tidak terpenuhi.

Metode ini lebih tahan terhadap data yang tidak mengikuti distribusi normal atau memiliki variabilitas yang heterogen (heteroskedastisitas).

2.1.4.3 Generalized Least Squares (GLS)

Generalized Least Squares (GLS) adalah metode statistik yang digunakan dalam analisis regresi untuk mengatasi masalah heteroskedastisitas atau autokorelasi dalam data. Tidak seperti metode *Ordinary Least Squares* (OLS) yang mengasumsikan bahwa variabilitas residual adalah konstan dan residual tidak berkorelasi, GLS memungkinkan estimasi parameter regresi yang efisien dengan mengakomodasi situasi di mana asumsi-asumsi ini tidak terpenuhi.

2.1.5 Residual Standard Error (RSE)

Residual Standard Error (RSE) adalah ukuran yang digunakan dalam analisis regresi untuk mengevaluasi seberapa baik model regresi mencocokkan data yang diamati. RSE mengukur rata-rata penyimpangan atau kesalahan prediksi yang dibuat oleh model regresi dibandingkan dengan nilai sebenarnya.

2.2 Tinjauan Pustaka Non-Statistika

2.2.1 Beton

Beton adalah bahan konstruksi komposit yang terdiri dari campuran semen, air, agregat kasar (seperti pasir dan kerikil), dan bahan tambahan lainnya. Beton digunakan secara luas dalam berbagai jenis konstruksi, seperti gedung, jalan, jembatan, dan infrastruktur lainnya, karena kemampuannya untuk dibentuk sesuai dengan kebutuhan desain dan daya tahannya terhadap kondisi lingkungan.

2.2.2 Agregat

Agregat adalah material seperti pasir, kerikil, dan batu pecah yang digunakan dalam campuran beton atau aspal. Menurut Murdock (1960), agregat menyumbang sekitar 60-75% dari volume beton, sehingga kualitasnya sangat mempengaruhi kekuatan dan daya tahan struktur.

2.2.3 Air

Air adalah salah satu komponen penting dalam campuran beton dan berbagai aplikasi lainnya dalam konstruksi. Air digunakan untuk menghidrasi semen, memungkinkan proses pengerasan dan pengikatan yang membentuk kekuatan beton.

2.2.4 Superplasticizer

Superplasticizer, juga dikenal sebagai penurun air jangkauan tinggi, adalah bahan tambahan kimia yang digunakan dalam campuran beton untuk meningkatkan kualitas beton dengan mengurangi jumlah air yang diperlukan

2.2.5 Kekuatan tekan beton

Kekuatan tekan beton adalah salah satu parameter kunci yang menentukan kualitas dan kinerja beton. Kekuatan tekan ini mengacu pada kemampuan beton untuk menahan beban yang diterapkan secara merata pada permukaannya tanpa mengalami keruntuhan.

BAB III

METODOLOGI

3.1 Sumber Data

Sumber data yang digunakan dalam *project* ini merupakan data sekunder yang berasal dari dosen pengampu mata kuliah Regresi Terapan SD-A1 Program Studi Teknologi Sains Data, Fakultas Teknologi Maju dan Multidisiplin, Universitas Airlangga. Data yang digunakan berisi informasi tentang kekuatan tekan beton (*concrete compressive strength*) dengan bahan utama dalam campuran beton yaitu *cement*, *blast furnace slag*, *fly ash*, *water*, *superplasticizer*, *coarse aggregate*, dan *fine aggregate*, serta terdapat informasi tambahan yaitu *age* yang mana adalah usia beton. Jumlah observasi dalam data ini berjumlah 1030. Data dapat diakses pada *link* berikut: <https://bit.ly/DataKelompok11RegresiTerapan>.

3.2 Variabel Penelitian

Variabel yang digunakan pada *project* ini terdiri dari variabel dependen atau variabel respons (Y) dan variabel independen atau variabel prediktor (X). Informasi terkait variabel yang digunakan dalam *project* ini adalah sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Nama Variabel	Deskripsi	Satuan
Y	<i>Concrete Compressive Strength</i>	Kekuatan tekan beton, properti target untuk diprediksi	MPa
X_1	<i>Cement</i>	Komponen utama dalam beton untuk memberikan kekuatan pada struktur beton	kg/m^3
X_2	<i>Blast Furniture Slag</i>	Produk sampingan pembuatan baja, digunakan untuk meningkatkan durabilitas beton	kg/m^3
X_3	<i>Fly Ash</i>	Sisa pembakaran batubara, digunakan sebagai bahan tambahan untuk beton	kg/m^3
X_4	<i>Water</i>	Digunakan untuk memicu reaksi hidrasi semen dan mempengaruhi konsistensi beton	kg/m^3
X_5	<i>Superplasticizer</i>	Bahan tambahan untuk meningkatkan keenceran beton tanpa perlu menambah air	kg/m^3

X_6	<i>Coarse Aggregate</i>	Agregat kasar (batu kerikil) untuk menambah kekuatan beton	kg/m^3
X_7	<i>Fine Aggregate</i>	Agregat halus (pasir) untuk memberikan stabilitas campuran beton	kg/m^3
X_8	<i>Age (day)</i>	Usia beton dalam hari	Hari (1-365)

3.3 Metode Penelitian

Metode yang digunakan untuk menganalisis data dalam *project* ini adalah sebagai berikut.

3.3.1 Exploratory Data Analysis (EDA)

Exploratory Data Analysis (EDA) adalah analisis yang dilakukan untuk memberikan informasi tentang karakteristik data yang akan diamati. Berikut beberapa tindakan yang dilakukan dalam *exploratory data analysis*:

- Menghitung statistika deskriptif, seperti rata-rata, median, standar deviasi, nilai minimum, dan nilai maksimum dari seluruh variabel penelitian.
- Melakukan pemeriksaan *outlier* pada data.
- Membuat visualisasi data seperti histogram, *scatter plot*, dan *heatmap*.

3.3.2 Ordinary Least Squares (OLS)

Penelitian ini menggunakan metode OLS untuk memperoleh model regresi dari hubungan variabel dependen (*concrete compressive strength*) dengan variabel independen (*cement, blast furnace slag, fly ash, water, superplasticizer, coarse aggregate, fine aggregate, dan age*). Model yang akan diperoleh yaitu:

$$Y = \beta + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_8 X_8$$

Model ini diharapkan memenuhi asumsi yang bersifat linier, tidak bias, dan varians minimum atau BLUE (*Best Linear Unbiased Estimator*).

3.3.3 Uji Hipotesis

Pengujian hipotesis mencakup uji signifikansi koefisien regresi dan pengujian asumsi. Uji signifikansi koefisien regresi dilakukan secara simultan menggunakan uji-F dan secara parsial menggunakan uji-t. Pengujian asumsi meliputi uji normalitas, autokorelasi, heteroskedastisitas, dan multikolinearitas. Uji normalitas dilakukan menggunakan *Shapiro-Wilk*. Uji autokorelasi menggunakan *Durbin-Watson*. Uji heteroskedastisitas dilakukan menggunakan *Breusch Pagan*. Kemudian uji multikolinearitas dilakukan dengan menghitung nilai VIF (*Variance Inflation*

Factor) untuk memastikan bahwa setiap variabel independen tidak saling bergantung satu sama lain.

3.3.4 Robust Regression

Robust regression digunakan untuk mengatasi masalah pada data yang tidak memenuhi asumsi klasik *Ordinary Least Squares* (OLS) terutama apabila data terdapat *outlier*. *Robust regression* juga dapat digunakan jika asumsi normalitas residual dan homoskedastisitas tidak terpenuhi.

3.3.5 Generalized Least Squares (GLS)

Generalized Least Squares (GLS) digunakan ketika asumsi homoskedastisitas (variansi residual yang konstan) dan independensi residual (tidak ada autokorelasi) pada OLS dilanggar.

3.4 Prosedur Analisis

Analisis dilakukan menggunakan *tools* Python dan Rstudio dengan prosedur sebagai berikut:

- a. Melakukan *Exploratory Data Analysis* (EDA)
- b. Membuat model regresi menggunakan *Ordinary Least Square* (OLS)
- c. Pengujian asumsi pada model regresi
- d. Membuat model regresi menggunakan *Robust Regression* dan *Generalized Least Square*
- e. Pengujian signifikansi koefisien regresi
- f. Perbandingan model dan mencari model terbaik
- g. Membuat kesimpulan

BAB IV

ANALISIS DAN PEMBAHASAN

4.1 Exploratory Data Analysis (EDA)

4.1.1 Statistika Deskriptif

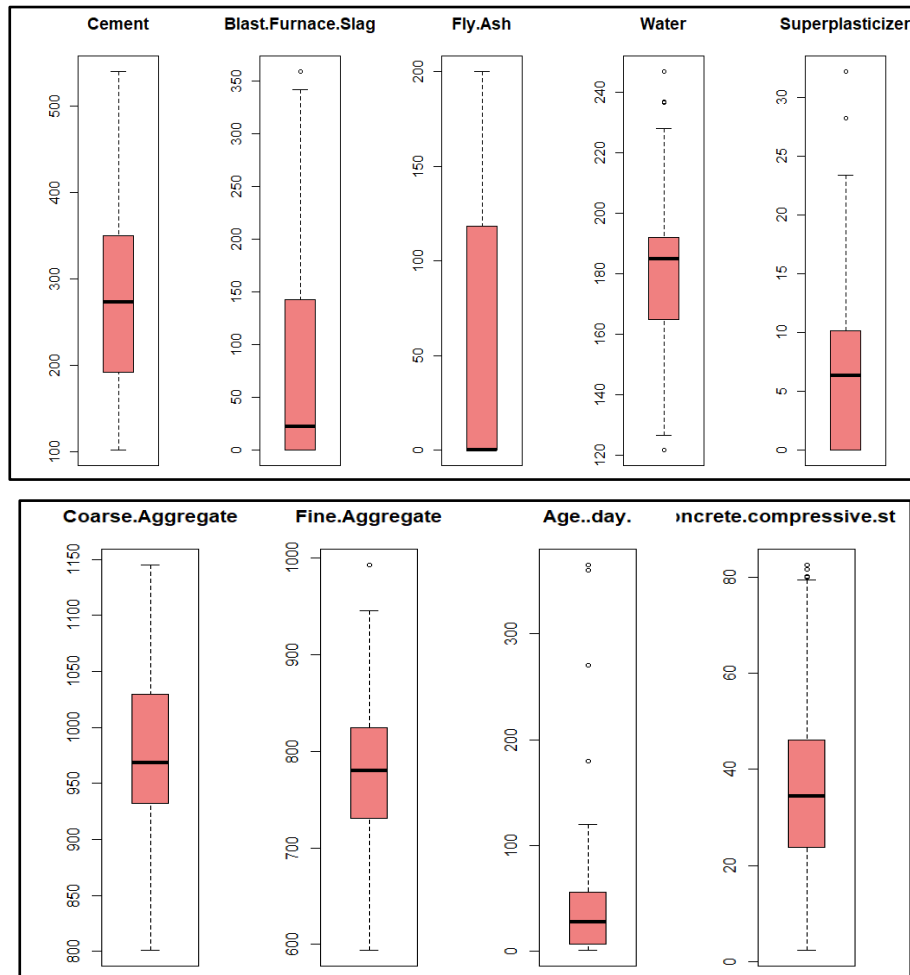
Tabel 2. Statistika Deskriptif

Var	Min	1st Qu.	Median	Mean	3rd Qu.	Max
Y	2.332	23.707	34.443	35.818	46.136	82.599
X_1	102.0	192.4	272.9	281.2	350.0	540.0
X_2	0.0	0.0	22.0	73.9	142.9	359.4
X_3	0.0	0.0	0.0	54.19	118.27	200.10
X_4	121.8	164.9	185.0	181.6	192.0	247.0
X_5	0.0	0.0	6.35	6.203	10.16	32.2
X_6	801.0	932.0	968.0	972.9	1029.4	1145.0
X_7	594.0	731.0	779.5	773.6	824.0	992.6
X_8	1.0	7.0	28.0	45.66	56.0	365.0

Pada tahap ini, terdapat statistika deskriptif berupa *summary* dataset yang menunjukkan variabilitas signifikan di berbagai komponen campuran beton. Variabel seperti fly ash, superplasticizer, dan water memiliki distribusi yang cukup terpusat, sementara variabel lain seperti cement dan blast furnace slag menunjukkan variasi yang lebih besar. Kekuatan kompresif beton memiliki variasi yang cukup luas, yang mungkin berkaitan dengan variasi dalam komposisi campuran beton dan umur beton yang diuji.

4.1.2 Pengecekan Outlier

Penulis menggunakan pendekatan berbasis Interquartile Range (IQR) untuk mendeteksi outlier pada setiap variabel dalam data ini. Pendekatan ini melibatkan perhitungan kuartil pertama (Q1), kuartil ketiga (Q3), dan rentang antar kuartil (IQR), yang kemudian digunakan untuk menentukan batas bawah dan batas atas untuk mengidentifikasi nilai-nilai yang berada di luar rentang normal, yang dianggap sebagai outlier. Selain pendekatan IQR, penulis juga membuat boxplot untuk memvisualisasikan distribusi data dan mendeteksi outlier secara grafis. Hasil visualisasi boxplot dapat dilihat pada gambar berikut:



Gambar 1. Boxplot

Hasil pengecekan outlier tersebut dapat dilihat pada tabel berikut :

Tabel 3. Variabel Outliers

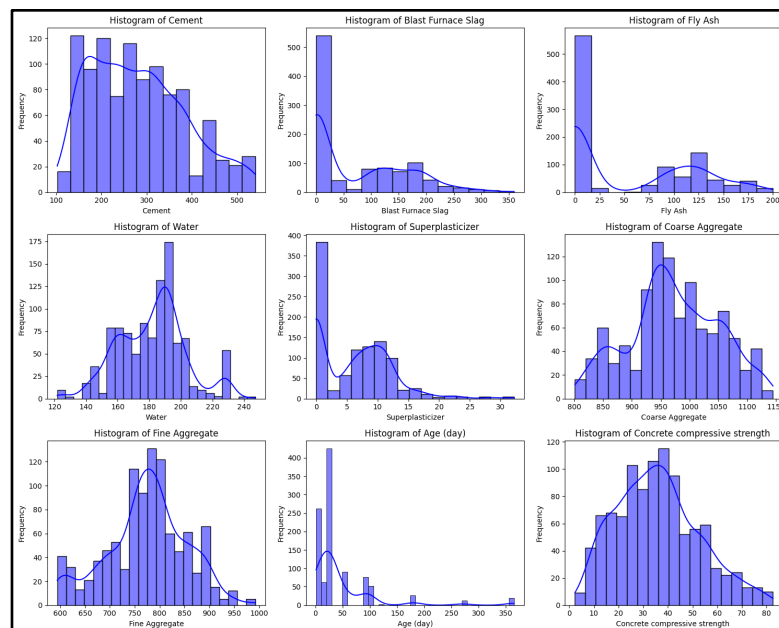
Variabel	Nama Variabel	Outliers
Y	<i>Concrete Compressive Strength</i>	4
X_1	<i>Cement</i>	0
X_2	<i>Blast Furniture Slag</i>	2
X_3	<i>Fly Ash</i>	0
X_4	<i>Water</i>	9
X_5	<i>Superplasticizer</i>	10
X_6	<i>Coarse Aggregate</i>	0
X_7	<i>Fine Aggregate</i>	5
X_8	<i>Age (day)</i>	59

Hasil pengecekan outlier menunjukkan bahwa sebagian besar variabel, seperti *Cement*, *Fly Ash*, dan *Coarse Aggregate*, tidak memiliki outlier. Namun, variabel

Water dan Superplasticizer memiliki jumlah *outlier* yang signifikan, masing-masing 9 dan 10, yang menunjukkan adanya sampel dengan nilai ekstrem dalam penggunaan air dan *superplasticizer*. Selain itu, *Fine Aggregate* dan *Age (days)* juga menunjukkan beberapa *outlier*, dengan *Age (days)* memiliki jumlah *outlier* yang cukup besar (59), yang mungkin mencerminkan variasi besar dalam usia beton yang diuji. *Concrete Compressive Strength* memiliki 4 *outlier*, menandakan variasi besar dalam kekuatan beton. Keberadaan *outlier* ini perlu diperhatikan karena dapat memengaruhi analisis lebih lanjut dan hubungan antara komposisi beton dan kekuatannya.

4.1.3 Visualisasi Data

4.1.3.1 Histogram

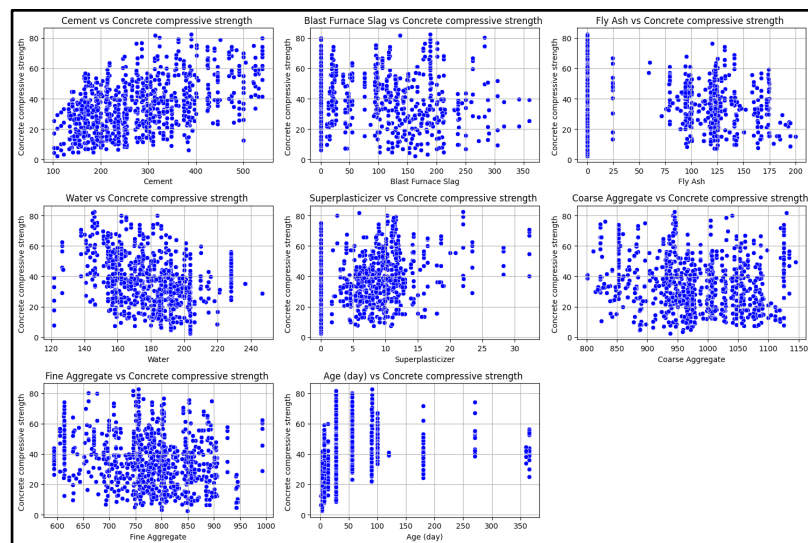


Gambar 2. Histogram

Gambar diatas merupakan visualisasi histogram yang digunakan untuk menjelaskan frekuensi dan distribusi pada data. Untuk variabel *cement*, dapat dilihat terdapat penyebaran yang relatif merata di kedua sisi puncak dengan puncak di sekitar nilai 200-300. Kemudian untuk variabel *blast furnace slag*, sebagian besar nilai berada di sekitar 0-50 dengan distribusi yang menurun drastis setelahnya, dan dapat dilihat distribusi data miring ke kanan (*positively skewed*). Lalu untuk variabel *fly ash*, hampir semua nilai berada di rentang 0-25 dan distribusi data miring ke kanan (*positively skewed*). Kemudian untuk variabel *water*, distribusi data cenderung normal dengan penyebaran data

cenderung merata di kedua sisi. Lalu untuk variabel *superplasticizer*, sebagian besar data berada di kisaran 0-5 dengan penurunan yang signifikan, lalu distribusi data miring ke kanan (*positively skewed*). Kemudian untuk variabel *coarse aggregate*, distribusi cenderung normal dengan penyebaran data merata di kedua sisi puncak. Lalu untuk variabel *fine aggregate*, terlihat distribusi hampir normal namun di akhir terdapat sedikit penyimpangan di ekor kiri. Kemudian untuk variabel *age*, dapat dilihat distribusi data tidak merata dari hari ke-0 hingga hari ke-365. Kemudian untuk visualisasi variabel *concrete compressive strength*, distribusi data cenderung normal dan penyebaran data cukup merata di kedua sisi puncak.

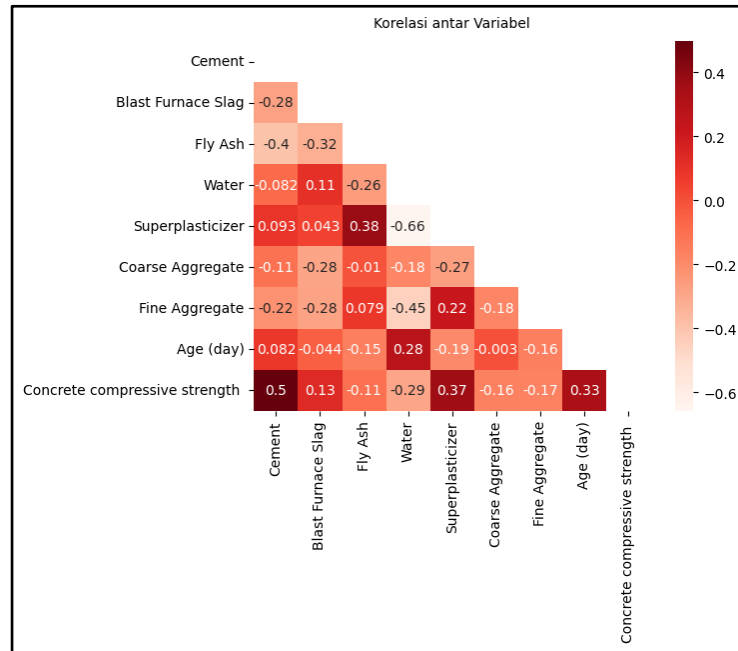
4.1.3.2 Scatter Plot



Gambar 3. Scatter Plot

Gambar diatas merupakan visualisasi *scatter plot* dari hubungan antara variabel dependen yaitu *concrete compressive strength* dengan variabel independen yaitu *cement*, *blast furnace slag*, *fly ash*, *water*, *superplasticizer*, *coarse aggregate*, *fine aggregate*, dan *age*. Dapat dilihat penyebaran titik data variabel *cement* dan *age* dengan variabel dependen menunjukkan hubungan positif karena penyebaran data cenderung membentuk titik ke atas. Kemudian untuk variabel *blast furnace slag*, *fly ash*, *water*, *superplasticizer*, *coarse aggregate*, dan *fine aggregate* tidak terlihat pola yang jelas karena penyebaran titik data yang merata.

4.1.3.3 Heatmap



Gambar 4. Heatmap

Gambar diatas merupakan visualisasi *heatmap* antara setiap variabel dengan variabel lainnya. Warna yang semakin gelap menandakan korelasi positif yang kuat antar variabel dan warna yang semakin memudar menunjukkan korelasi negatif antar variabel. Dapat diketahui untuk variabel dengan korelasi positif terkuat yaitu variabel *fly ash* dan *water*, kemudian variabel *superplasticizer* dengan *concrete compressive strength*. Lalu untuk variabel dengan korelasi negatif terkuat dimiliki oleh variabel *water* dan *superplasticizer*.

4.2 Analisis Regresi

4.2.1 Model Regresi OLS (*Ordinary Least Square*)

Residuals:				
Min	1Q	Median	3Q	Max
-28.653	-6.303	0.704	6.562	34.446
Coefficients:				
	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-23.163756	26.588421	-0.871	0.383851
Cement	0.119785	0.008489	14.110	< 2e-16 ***
Blast.Furnace.Slag	0.103847	0.010136	10.245	< 2e-16 ***
Fly.Ash	0.087943	0.012585	6.988	5.03e-12 ***
Water	-0.150298	0.040179	-3.741	0.000194 ***
Superplasticizer	0.290687	0.093460	3.110	0.001921 **
Coarse.Aggregate	0.018030	0.009394	1.919	0.055227 .
Fine.Aggregate	0.020154	0.010703	1.883	0.059968 .
Age..day.	0.114226	0.005427	21.046	< 2e-16 ***
--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 10.4 on 1021 degrees of freedom				
Multiple R-squared: 0.6155, Adjusted R-squared: 0.6125				
F-statistic: 204.3 on 8 and 1021 DF, p-value: < 2.2e-16				

Gambar 5. OLS

Model regresi OLS bertujuan untuk memahami hubungan linier antara variabel independen (Cement, Blast Furnace Slag, Fly Ash, Water, Superplasticizer, Coarse Aggregate, Fine Aggregate, Age) terhadap **variabel dependen** (Concrete Compressive Strength).

Dari hasil output R, didapatkan persamaan OLS Regression dari output tersebut adalah :

- Concrete Compressive Strength = $-35.5047 + 0.1264 (\text{Cement}) + 0.1116 (\text{Blast Furnace Slag}) + 0.0968 (\text{Fly Ash}) - 0.1417 (\text{Water}) + 0.2744 (\text{Superplasticizer}) + 0.0205 (\text{Coarse Aggregate}) + 0.0267 (\text{Fine Aggregate}) + 0.1326 (\text{Age})$.
- R-squared sebesar 0.6155 menunjukkan bahwa model ini mampu menjelaskan sekitar 61.55% variansi dalam variabel target (Concrete Compressive Strength).
- Adjusted R-squared sebesar 0.6125 menunjukkan bahwa model ini mampu menjelaskan sekitar 61.25% variansi target setelah mempertimbangkan jumlah variabel independen dalam model.
- Residual Standard Error (RSE) sebesar 10.4 MPa, yang menunjukkan bahwa rata-rata kesalahan prediksi model adalah sekitar 10.4 MPa dari nilai aktual.

Model OLS ini menunjukkan bahwa variabel independen seperti Cement, Blast Furnace Slag, dan Age memiliki pengaruh signifikan terhadap kekuatan beton, dengan Cement dan Age memberikan dampak terbesar. Namun, model ini rentan terhadap outlier, yang terlihat dari rentang residual yang cukup besar.

4.2.2 Pengujian Hipotesis

4.2.2.1 Uji Serentak (Uji F)

Hipotesis :

$$H_0 : \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$$

$$H_1 : \text{setidaknya ada satu } \beta_i \neq 0, i = 1, 2, 3, 4$$

$$\alpha : 0.05$$

Daerah kritis : $p\text{-value} < \alpha$

Uji Statistik :

F.hitung	p.value	alpha	Kesimpulan
value 204.2691	6.76158e-206	0.05	Tolak H0

Gambar 6. Uji F

Hasil uji serentak (Overall F-Test) menunjukkan bahwa nilai F-hitung sebesar 204.2691 dengan p-value sebesar 6.76×10^{-206} , yang jauh lebih kecil dari taraf signifikansi ($\alpha = 0.05$). Dengan demikian, hipotesis nol H_0 , yang menyatakan bahwa semua koefisien regresi sama dengan nol, **ditolak**. Hal ini mengindikasikan bahwa model regresi secara keseluruhan **signifikan**, sehingga terdapat setidaknya satu variabel independen yang memiliki pengaruh signifikan terhadap variabel dependen (Concrete Compressive Strength).

4.2.2.2 Uji Parsial (Uji t)

Hipotesis :

$$H_0 : \beta_i = 0, i = 1, 2, 3, 4$$

$$H_1 : \beta_i \neq 0, i = 1, 2, 3, 4$$

$$\alpha : 0.05$$

Daerah kritis : $p\text{-value} < \alpha$

Uji Statistik :

	Variable	Estimate	Std..Error	t.value	p.value	Significance
(Intercept)	(Intercept)	-23.1638	26.5884	-0.8712	3.839e-01	Not Significant
Cement	Cement	0.1198	0.0085	14.1100	1.963e-41	Significant
Blast.Furnace.Slag	Blast.Furnace.Slag	0.1038	0.0101	10.2452	1.633e-23	Significant
Fly.Ash	Fly.Ash	0.0879	0.0126	6.9879	5.030e-12	Significant
Water	Water	-0.1503	0.0402	-3.7407	1.937e-04	Significant
Superplasticizer	Superplasticizer	0.2907	0.0935	3.1103	1.921e-03	Significant
Coarse.Aggregate	Coarse.Aggregate	0.0180	0.0094	1.9193	5.523e-02	Not Significant
Fine.Aggregate	Fine.Aggregate	0.0202	0.0107	1.8831	5.997e-02	Not Significant
Age..day.	Age..day.	0.1142	0.0054	21.0457	5.841e-82	Significant

Gambar 7. Uji t

Hasil uji parsial (t-Test) menunjukkan bahwa:

- Variabel Cement, Blast.Furnace.Slag, Fly.Ash, Water, Superplasticizer, dan Age..day. memiliki pengaruh signifikan terhadap Concrete Compressive Strength ($p\text{-value} < 0.05$).
- Variabel Cement, Blast.Furnace.Slag, Fly.Ash, Superplasticizer, dan Age..day. berpengaruh positif, menunjukkan bahwa peningkatan jumlah bahan ini meningkatkan kekuatan beton.
- Sebaliknya, variabel Water memiliki pengaruh negatif, artinya penambahan air cenderung menurunkan kekuatan beton.
- Variabel lain seperti Coarse.Aggregate dan Fine.Aggregate tidak signifikan ($p\text{-value} \geq 0.05$), sehingga kontribusinya terhadap kekuatan beton tidak terlalu terlihat.

4.2.3 Pengujian Asumsi

4.2.3.1 Uji Normalitas Residual

Hipotesis :

H_0 : Error berdistribusi normal

H_1 : Error tidak berdistribusi normal

$\alpha : 0.05$

Daerah kritis : $p\text{-value} < \alpha$

Shapiro-Wilk normality test
data: residuals_ols
W = 0.99532, p-value = 0.002986

Gambar 8. Uji Normalitas Residual: *Shapiro-Wilk*

Hasil uji normalitas residual menggunakan Shapiro-Wilk Test menunjukkan nilai $W = 0.99532$ dan $p\text{-value}$ sebesar 0.002986. Karena $p\text{-value} < 0.05$, maka hipotesis nol (H_0) yang menyatakan bahwa residual berdistribusi normal ditolak. Dengan demikian, residual **tidak berdistribusi normal**, yang mengindikasikan pelanggaran asumsi normalitas pada model regresi.

4.2.3.2 Uji Homoskedastisitas

Hipotesis :

H_0 : Variansi Error Konstan (homoskedastisitas)

H_1 : Variansi Error Tidak Konstan (heteroskedastisitas)

$\alpha : 0.05$

Daerah kritis : $p\text{-value} < \alpha$

```
studentized Breusch-Pagan test  
  
data:  ols_model  
BP = 137.17, df = 8, p-value < 2.2e-16
```

Gambar 9. Uji Homoskedastisitas: *Breusch-Pagan*

Hasil uji homoskedastisitas menggunakan Breusch-Pagan Test menunjukkan nilai statistik $BP = 137.17$, derajat kebebasan (df) = 8, dan $p\text{-value} < 2.2e-16$. Karena $p\text{-value}$ jauh lebih kecil dari 0.05, maka hipotesis nol (H_0) yang menyatakan bahwa residual memiliki varians konstan **ditolak**. Dengan demikian, terdapat indikasi kuat **adanya heteroskedastisitas** dalam model, yang berarti varians residual tidak konstan.

4.2.3.3 Uji Multikolinearitas

Hipotesis :

H_0 : $VIF < 10$ (tidak terjadi multikolinearitas)

H_1 : $VIF > 10$ (terjadi multikolinearitas)

		Variable	VIF
Cement		Cement	7.49
Blast.Furnace.Slag		Blast.Furnace.Slag	7.28
Fly.Ash		Fly.Ash	6.17
Water		Water	7.00
Superplasticizer		Superplasticizer	2.97
Coarse.Aggregate		Coarse.Aggregate	5.08
Fine.Aggregate		Fine.Aggregate	7.01
Age..day.		Age..day.	1.12

Gambar 10. Uji Multikolinearitas: *VIF*

Berdasarkan hasil VIF, tidak ada variabel dengan $VIF > 10$, sehingga hipotesis nol yang menyatakan tidak terjadi multikolinearitas gagal ditolak. Hasil uji ini menunjukkan tidak terdapat multikolinearitas.

4.2.3.4 Uji Autokorelasi

Hipotesis:

H_0 : Tidak ada autokorelasi pada residual.

H_1 : Terdapat autokorelasi pada residual.

$\alpha : 0.05$

Daerah kritis : $p\text{-value} < \alpha$

```
Durbin-Watson test  
  
data: ols_model  
DW = 1.2815, p-value < 2.2e-16  
alternative hypothesis: true autocorrelation is not 0
```

Gambar 11. Uji Autokorelasi: *Durbin-Watson*

Hasil uji independensi residual menggunakan Durbin-Watson Test menunjukkan nilai statistik $DW = 1.2815$ dengan $p\text{-value} < 2.2e-16$. Karena $p\text{-value} < 0.05$, maka hipotesis nol (H_0) yang menyatakan tidak ada autokorelasi **ditolak**. Nilai $DW = 1.2815$, yang jauh dari nilai ideal 2, mengindikasikan adanya autokorelasi positif pada residual model regresi. Hal ini menunjukkan **pelanggaran asumsi independensi residual**.

4.2.4 Alternatif Model Regresi

Berdasarkan hasil uji asumsi pada model OLS, terdapat beberapa permasalahan utama:

- Residual tidak berdistribusi normal ($p\text{-value} < 0.05$ pada Shapiro-Wilk Test)
- Varians residual tidak konstan (heteroskedastisitas, $p\text{-value} < 0.05$ pada Breusch-Pagan Test)
- Autokorelasi positif ($p\text{-value} < 0.05$ pada Durbin-Watson Test).

Untuk mengatasi pelanggaran asumsi ini, model alternatif yang direkomendasikan adalah:

- Robust Regression, yang dapat menangani heteroskedastisitas dan ketidaknormalan residual secara bersamaan.
- Generalized Least Squares (GLS) dapat digunakan untuk menangani autokorelasi dan heteroskedastisitas.

4.2.4.1 Model Regresi Robust

Residuals:				
Min	1Q	Median	3Q	Max
-33.7679	-6.6741	0.5478	6.4248	34.3113
Coefficients:				
	Value	Std. Error	t value	
(Intercept)	-35.5047	26.1004	-1.3603	
Cement	0.1264	0.0083	15.1697	
Blast.Furnace.Slag	0.1116	0.0100	11.2142	
Fly.Ash	0.0968	0.0124	7.8318	
Water	-0.1417	0.0394	-3.5932	
Superplasticizer	0.2744	0.0917	2.9907	
Coarse.Aggregate	0.0205	0.0092	2.2237	
Fine.Aggregate	0.0267	0.0105	2.5401	
Age..day.	0.1326	0.0053	24.8913	
Residual standard error: 9.586 on 1021 degrees of freedom				

Gambar 12. Model Regresi *Robust*

Dari hasil *output R*, didapatkan persamaan *Robust Regression* dari output tersebut adalah:

- Concrete Compressive Strength = $-35.5047 + 0.1264 (\text{Cement}) + 0.1116 (\text{Blast Furnace Slag}) + 0.0968 (\text{Fly Ash}) - 0.1417 (\text{Water}) + 0.2744 (\text{Superplasticizer}) + 0.0205 (\text{Coarse Aggregate}) + 0.0267 (\text{Fine Aggregate}) + 0.1326 (\text{Age})$.
- Residual Standard Error (RSE) sebesar 9.586 MPa, yang lebih kecil dibandingkan model OLS (10.4 MPa), menunjukkan bahwa Robust Regression memberikan prediksi yang lebih akurat dengan kesalahan prediksi yang lebih rendah.
- Robust Regression tidak memiliki nilai R-squared, karena metode ini tidak berfokus pada menjelaskan total variansi tetapi lebih pada menghasilkan estimasi koefisien yang tahan terhadap outlier.

Model Robust Regression menunjukkan bahwa variabel independen memiliki pengaruh yang sama seperti pada model OLS. Namun, model ini lebih stabil karena mampu mengurangi pengaruh nilai outlier, menghasilkan estimasi yang lebih akurat dan residual yang lebih kecil.

4.2.4.2 Model Regresi GLS (*Generalized Least Squares*)

Coefficients:				
	Value	Std.Error	t-value	p-value
(Intercept)	-36.30154975	30.042740569	-1.208330	2.272001e-01
Cement	0.12588463	0.009854089	12.774862	8.907434e-35
Blast.Furnace.Slag	0.10863675	0.011009709	9.867359	5.356039e-22
Fly.Ash	0.08052679	0.014147240	5.692050	1.639625e-08
Water	-0.12464285	0.043111545	-2.891171	3.919440e-03
Superplasticizer	0.16951413	0.101093544	1.676805	9.388679e-02
Coarse.Aggregate	0.01854129	0.010937100	1.695266	9.032990e-02
Fine.Aggregate	0.02956206	0.011877821	2.488845	1.297427e-02
Age..day.	0.10998461	0.005454603	20.163633	2.261130e-76
Residual Standard Error (RSE):				
10.52057				
Model Fit Statistics:				
AIC: 7666.75				
BIC: 7720.964				
Log-likelihood: -3822.375				

Gambar 13. Model Regresi GLS

Dari hasil *output R*, didapatkan persamaan *GLS* dari *output* tersebut adalah:

- Concrete Compressive Strength = -36.3016 + 0.1259 (Cement) + 0.1086 (Blast Furnace Slag) + 0.0805 (Fly Ash) - 0.1246 (Water) + 0.1695 (Superplasticizer) + 0.0185 (Coarse Aggregate) + 0.0296 (Fine Aggregate) + 0.1100 (Age).
- Residual Standard Error (RSE) sebesar 10.52 MPa, yang sedikit lebih tinggi dibandingkan Robust Regression (9.586 MPa) tetapi lebih kecil dibandingkan OLS (10.4 MPa). Hal ini menunjukkan bahwa GLS memberikan prediksi yang lebih stabil dengan menangani autokorelasi antar residual.

4.3 Pemilihan Model Regresi Terbaik

Pemilihan model regresi terbaik dilakukan dengan membandingkan *Residual Standard Error* (RSE), R^2 , dan *Adjusted R²* dari model OLS, *Robust Regression*, dan GLS. Berikut adalah hasil perbandingannya:

Tabel 4. Model Regresi Terbaik

Pembanding	<i>OLS Regression</i>	<i>Robust Regression</i>	<i>GLS Regression</i>
Nilai RSE	10.4	9.586	10.52
<i>R-Squared</i>	0.6155	-	0.6050
<i>Adjusted R-Squared</i>	0.6125	-	0.6050

Berdasarkan tabel, RSE pada *Robust Regression* lebih kecil dibandingkan OLS dan GLS. RSE yang lebih rendah menunjukkan bahwa model ini menghasilkan prediksi yang lebih akurat dengan kesalahan rata-rata yang lebih kecil. Artinya, Robust regression memberikan prediksi dengan kesalahan lebih kecil. *Robust regression* tahan terhadap outlier, sehingga estimasi koefisien lebih stabil dibandingkan OLS. Sehingga dapat disimpulkan bahwa model *Robust regression* adalah model regresi yang terbaik untuk dataset yang digunakan.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Analisis Pengaruh Variabel:

- Variabel independen seperti Cement, Blast Furnace Slag, Fly Ash, Water, Superplasticizer, dan Age memiliki pengaruh signifikan terhadap Concrete Compressive Strength. Variabel Cement dan Age memberikan kontribusi positif terbesar terhadap kekuatan beton, sedangkan variabel Water memiliki pengaruh negatif yang signifikan.
- Variabel Coarse Aggregate dan Fine Aggregate tidak menunjukkan pengaruh signifikan terhadap kekuatan beton.

Perbandingan Model Regresi:

- Model OLS memberikan nilai R^2 sebesar 0.6155 dan Adjusted R^2 sebesar 0.6125, menunjukkan kemampuan menjelaskan 61.25% variansi data. Namun, model ini melanggar asumsi normalitas, homoskedastisitas, dan autokorelasi, sehingga hasil estimasi kurang andal.
- Model Robust Regression menghasilkan nilai Residual Standard Error (RSE) sebesar 9.586 MPa, lebih kecil dibandingkan OLS (10.4 MPa) dan GLS (10.52 MPa), menunjukkan akurasi prediksi yang lebih tinggi. Model ini juga lebih stabil terhadap pengaruh outlier dan pelanggaran asumsi.
- Model GLS mampu menangani autokorelasi residual, tetapi nilai RSE-nya lebih besar dibandingkan Robust Regression.

Pemilihan Model Terbaik:

Berdasarkan perbandingan statistik, dalam memilih model regresi terbaik digunakan perbandingan RSE setiap model. Robust Regression dipilih sebagai model terbaik untuk memprediksi kekuatan tekan beton. Model ini mampu memberikan hasil prediksi yang akurat dengan stabilitas estimasi koefisien yang lebih baik, terutama dalam kondisi data dengan outlier dan pelanggaran asumsi regresi.

5.2 Saran

Untuk prediksi kekuatan tekan beton di masa depan, disarankan menggunakan Robust Regression, terutama jika data memiliki outlier atau variansi residual yang tidak konstan. Selain itu, penambahan variabel seperti suhu saat beton dicetak atau kandungan

mineral dalam bahan dapat dipertimbangkan untuk meningkatkan kualitas model prediksi. Dalam pengumpulan data beton di masa depan, distribusi variabel yang lebih merata perlu dipastikan untuk mengurangi dampak outlier dan meningkatkan akurasi model. Model prediksi yang dihasilkan juga dapat diimplementasikan oleh industri konstruksi untuk menentukan komposisi material beton yang optimal, mengurangi pemborosan, serta meningkatkan efisiensi biaya dalam proyek pembangunan.

DAFTAR PUSTAKA

- Box, G. E. P. (1971). *The Use of Statistics in the Analysis of Outliers*. Journal of the Royal Statistical Society, Series B (Methodological), 33(2), 191-211.
- Chatterjee, S., & Hadi, A. S. (2006). *Regression Analysis by Example* (4th ed.). Wiley-Interscience.
- Gujarati, D. N. (2003). *Basic Econometrics* (4th ed.). McGraw-Hill.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2010). *Multivariate Data Analysis* (7th ed.). Pearson.
- Mehta, P. K., & Monteiro, P. J. M. (2014). *Concrete: Microstructure, Properties, and Materials*. McGraw-Hill Education.
- Montgomery, D. C., Peck, E. A., & Vining, G. G. (2021). *Introduction to Linear Regression Analysis*. Wiley.
- Murdock, H. L., & Rushing, M. A. (2014). *Concrete Technology: Theory and Practice* (2nd ed.). Prentice Hall.
- Neville, A. M. (2012). *Properties of Concrete* (5th ed.). Longman.
- Wooldridge, J. M. (2016). *Introductory Econometrics: A Modern Approach* (6th ed.). Cengage Learning.

LAMPIRAN

Daftar lampiran dapat diakses pada folder berikut:

[Daftar Lampiran](#)