

# Pengaruh Keluarga Terhadap Banyaknya Sampah Yang Ada Di Lingkungan RT01 Banten Indah Permai Dengan Metode Regresi Linear Sederhana Menggunakan Orange Data Mining

Muhamad Hasbi As'siddiq

Program Studi Teknik Informatika , Universitas Serang Raya

e-mail: mhasbiraz@gmail.com

## ABSTRAKSI

Sampah merupakan kumpulan sisa-sisa pembuangan yang berasal dari manusia, beragam jenis sampah yang dihasilkan oleh rumah tangga yaitu organik, non-organik dan sampah domestik yang biasa dihasilkan oleh rumah tangga. Sampah ini dihasilkan setiap harinya oleh keluarga di lingkungan Rt 01 Rw 027 di Perumahan Banten Indah Permai yang sudah dicatat oleh ketua rt untuk rekapan agar pembayaran iuran kepada petugas yang bertanggung jawab dapat disesuaikan tiap harinya, karena pembayaran petugas sampah ini dibayar perhari sesuai jumlah banyak(KG) nya sampah yang dia angkut tiap harinya. Dengan metode prediksi dan menggunakan regresi linear sederhana dengan bantuan Orange Mining kita akan mencari tahu pengaruh banyaknya keluarga terhadap banyak nya sampah yang terkumpul. Hasil analisis menyoroiti hubungan signifikan antara faktor-faktor seperti ukuran keluarga, pola konsumsi, dan kesadaran lingkungan dengan tingkat akumulasi sampah. Penemuan ini dapat memberikan wawasan yang berharga bagi upaya pengelolaan sampah di lingkungan tersebut dengan menekankan pentingnya peran keluarga dalam mengurangi dan mengelola sampah secara efektif.

**Kata Kunci:** keluarga, orange mining, prediksi, regresi linear sederhana, sampah

## ABSTRACT

Waste is the accumulation of refuse generated by humans, which in the household context is categorized into organic, inorganic, and domestic waste. In the RT 01 RW 027 community of Perumahan Banten Indah Permai, each household produces waste daily, and these quantities are recorded by the neighborhood head to calculate the daily fee for waste collectors—where the tariff is based on the weight of waste (in kilograms) transported. By employing predictive methods and simple linear regression on the Orange platform, this study aims to measure the influence of family size on the daily volume of waste collected. The analysis results demonstrate a positive and significant relationship between family size, consumption patterns, and levels of environmental awareness with the rate of waste accumulation. These findings provide valuable insights for designing more effective waste management strategies, emphasizing the active role of each family in waste reduction and organization efforts.

**Keywords:** family, orange mining, prediction, simple linear regression, waste

## Penulis Korespondensi

Muhamad Hasbi As'siddiq

Tanggal Submit : 26/01/2024

Tanggal Diterima : 11/07/2024

Tanggal Terbit : 30/08/2024

This is an open access article under the [CC-BY-NC-SA](https://creativecommons.org/licenses/by-nc-sa/4.0/) license



**Copyright:** © 2023 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 (CC BY-NC-SA 4.0) International License (<https://creativecommons.org/licenses/by-nc-sa/4.0/>).

Publisher's Note: JPPM stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

## I. PENDAHULUAN

Pengelolaan sampah telah menjadi tantangan global yang membutuhkan perhatian serius dalam upaya menjaga keberlanjutan lingkungan. Sampah adalah sisa kegiatan sehari – hari atau proses alam yang berbentuk

padat[1]. Sampah memiliki dampak yang sangat besar, berdampak negatif terhadap kesehatan dan kelangsungan hidup masyarakat. Bila sampah dibuang secara sembarangan atau ditumpuk tanpa ada pengelolaan yang baik, maka akan menimbulkan

berbagai dampak kesehatan yang serius[2]. Salah satu faktor yang memengaruhi tingkat akumulasi sampah adalah peran keluarga dalam mengelola limbah domestik. Lingkungan RT01 Banten Indah Permai menjadi subjek penting dalam penelitian ini, karena memungkinkan untuk memahami secara mendalam pengaruh faktor-faktor keluarga terhadap kuantitas sampah yang terakumulasi[3].

Studi ini bertujuan untuk menganalisis dampak yang ditimbulkan oleh karakteristik keluarga terhadap banyaknya sampah yang ada di lingkungan tersebut. Metode yang digunakan dalam penelitian ini adalah Regresi Linear Sederhana dengan dukungan dari alat analisis aplikasi data mining Orange Mining[4]. Metode regresi linear adalah alat statistik yang dipergunakan untuk mengetahui pengaruh antara satu atau beberapa variabel terhadap satu buah variabel[5]. Aplikasi data mining ini yang mampu menggali pola-pola tersembunyi dalam data yang dikumpulkan. Metode prediksi ini memungkinkan kita untuk dapat mengetahui pengaruhnya[6]. Mengatur prediksi yang tepat dalam algoritma yang benar dapat dilakukan dengan bantuan Orange data Mining[7].

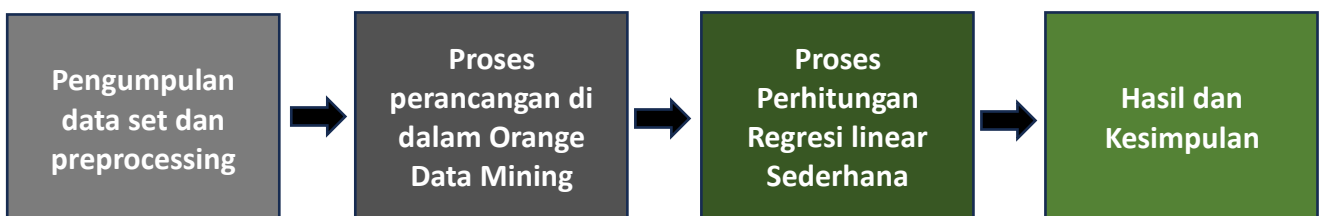
Penelitian ini mengumpulkan data melalui survei[8] yang merinci informasi tentang struktur keluarga[9], kebiasaan konsumsi, praktik pengelolaan sampah, dan bertanya langsung ke ketua Rt01. Analisis ini memperhatikan hubungan antara variabel-variabel ini dan jumlah sampah yang dihasilkan oleh setiap keluarga sebanyak 34 kepala keluarga. Sebagai dasar

penelitian, metodologi ini mengandalkan data primer yang dikumpulkan melalui survei dan wawancara langsung dengan pemimpin masyarakat setempat. Dengan mendekati ketua RT01, penelitian ini dapat memanfaatkan pengetahuan lokal yang mendalam, yang kemudian dianalisis untuk mendapatkan pemahaman yang lebih komprehensif tentang isu pengelolaan sampah di tingkat rumah tangga. Dan juga ini memberi impact ke pada petugas pembuang sampah karena di bayar perhari dan di berikan bonus dengan banyak nya sampah yang dia angkut (kg).

Diharapkan hasil dari penelitian ini dapat memberikan wawasan yang lebih jelas mengenai bagaimana peran keluarga memengaruhi jumlah sampah di lingkungan RT01 Banten Indah Permai. Implikasi dari temuan ini diharapkan dapat menjadi landasan bagi kebijakan-kebijakan lokal yang bertujuan untuk meningkatkan kesadaran lingkungan dan pengelolaan sampah yang lebih efektif di tingkat rumah tangga.

## II. METODE PENELITIAN

Tujuan dari penelitian ini untuk menganalisa seberapa berpengaruhnya banyak nya keluarga yang buang sampah terhadap banyak nya sampah yang diangkut oleh petugas setiap harinya bertujuan untuk memprediksi banyak nya sampah harian dengan menggunakan aplikasi Orange Mining untuk simulasinya.



Gambar 1. Tahapan Penelitian

### 1. Pengumpulan data set dan preprocessing

Tahapan penelitian ini menggunakan metode survey dan kuisisioner yang langsung bertanya ke ketua RT dan penulis langsung mencatat dan mendata yang sudah di sediakan oleh ketua RT dan melakukan pre processing yang mana melihat langsung data dan merangkai data agar dapat di jadikan menjadi informasi yang berguna sehingga terbentuk judul “Pengaruh Keluarga Terhadap Banyaknya Sampah Yang Ada Di Lingkungan Rt01 Banten Indah Permai Dengan Metode Regresi Linear Sederhana Menggunakan Orange Mining”

### 2. Proses perancangan di dalam Orang Data Mining

Orange data mining adalah perangkat visualisasi data, machine learning, dan data mining open source. Ini menampilkan front-end pemrograman visual untuk analisis data kualitatif eksploratif dan visualisasi data

interaktif[10]. Tahapan ini kita mulai merancang alur data yang akan di gunakan.

### 3. Proses Perhitungan Regresi Linear Sederhana

Tahapan ini adalah tahapan inti yang mana penulis dapat menentukan seberapa besar pengaruh nya dalam perhitungan regresi linear sederhana. Tahapan ini memiliki 2 proses yaitu perhitungan dengan orange data mining dan perhitungan secara manual.

### 4. Hasil dan Evaluasi

Tahapan ini adalah tahapan terakhir, penulis akan menampilkan hasil dari perhitungan dan membuat kesimpulan dari seluruh tahapan yang telah di lakukan.

#### 2.1. Data Penelitian

Data – data di bawah ini di dapat dari buku catatan harian ketua RT.01 lingkungan Banten Indah Permai. Data yang akan di gunakan di dalam perancangan Orange mining sebanyak 100 hari dan yang akan di tampil kan hanya sampling per 10 hari.

Tabel 1. Sample Data

No.	Hari ke-	Banyaknya Keluarga Yang Buang Sampah	Banyaknya Sampah (kg)
1	31/08/2023	13	66
2	01/09/2023	7	21
3	02/09/2023	22	122
4	03/09/2023	5	16
5	04/09/2023	7	29
6	05/09/2023	17	38
7	06/09/2023	21	86
8	07/09/2023	29	92
9	08/09/2023	10	35
10	09/09/2023	9	19
...	.....	.....	.....
...	.....	.....	.....
99	07/12/2023	10	24
100	08/12/2023	11	18

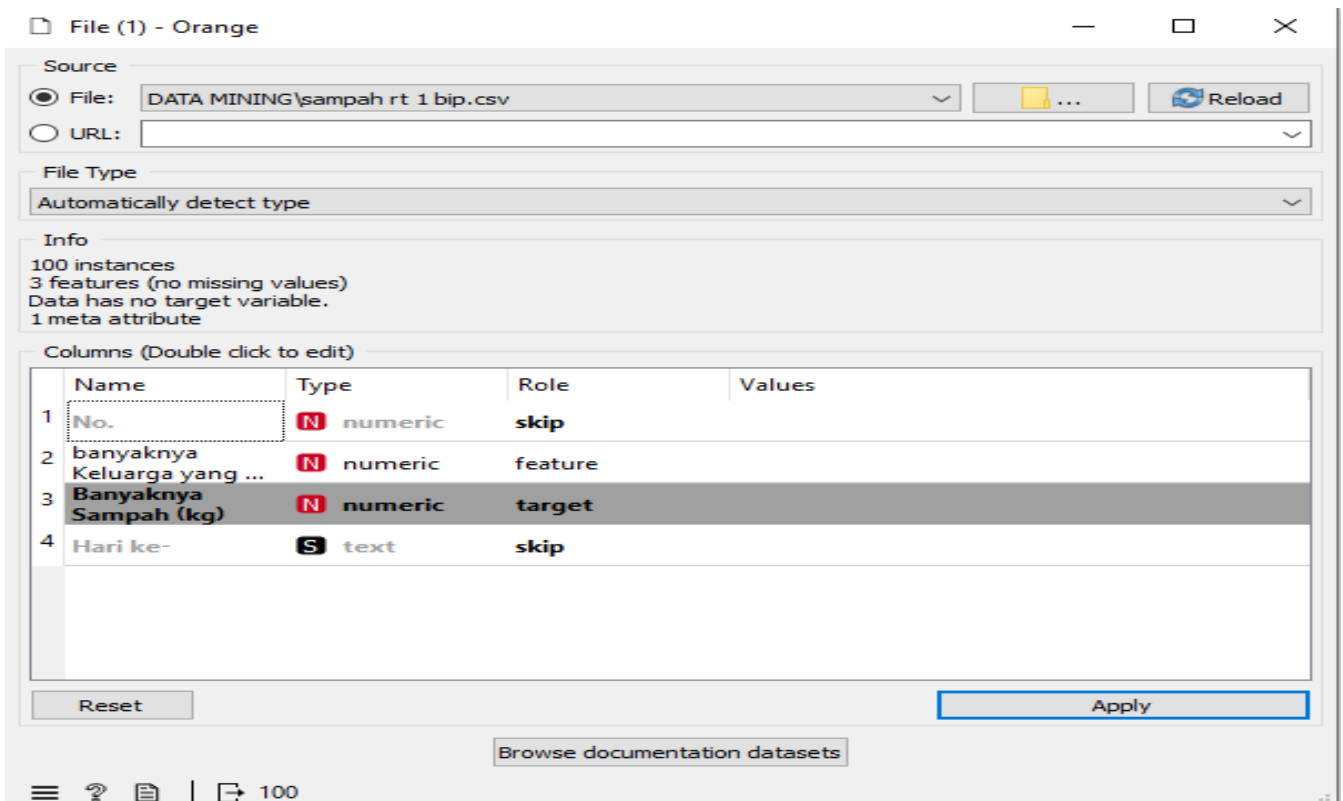
## 2.2. Perancangan Data Mining

Model prediksi untuk regresi linear dapat dilakukan di dalam Orange Mining dengan pembuatan variabel sebagai berikut :

1. Variabel X ( banyaknya keluarga yang buang sampah) yang menjadi feature

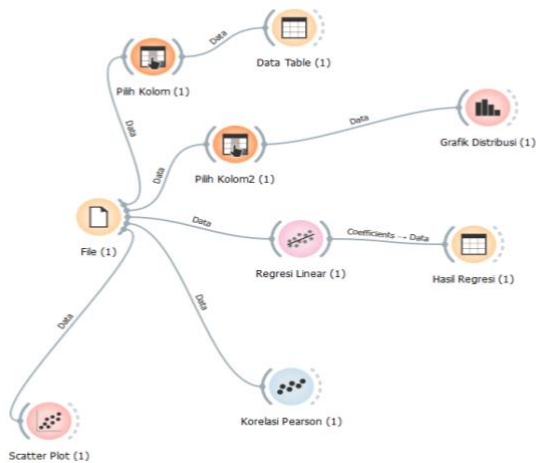
2. Variabel Y ( banyak nya sampah (KG) ) yang akan menjadi target

3. Di implementasikan di dalam Orange Mining di bagian file :



Gambar 2. Variabel di Orange Data Mining

Penggambaran model yang akan di gunakan untuk mengolah data di dalam orange data mining dengan menampilkan beberapa fitur yang ada di dalam aplikasi tersebut. Pada Gambar 3 akan memperlihatkan alur proses yang berawal dari data menjadi hasil regresi dan beberapa tambahan untuk menjadi pendukung hasil perhitungan.

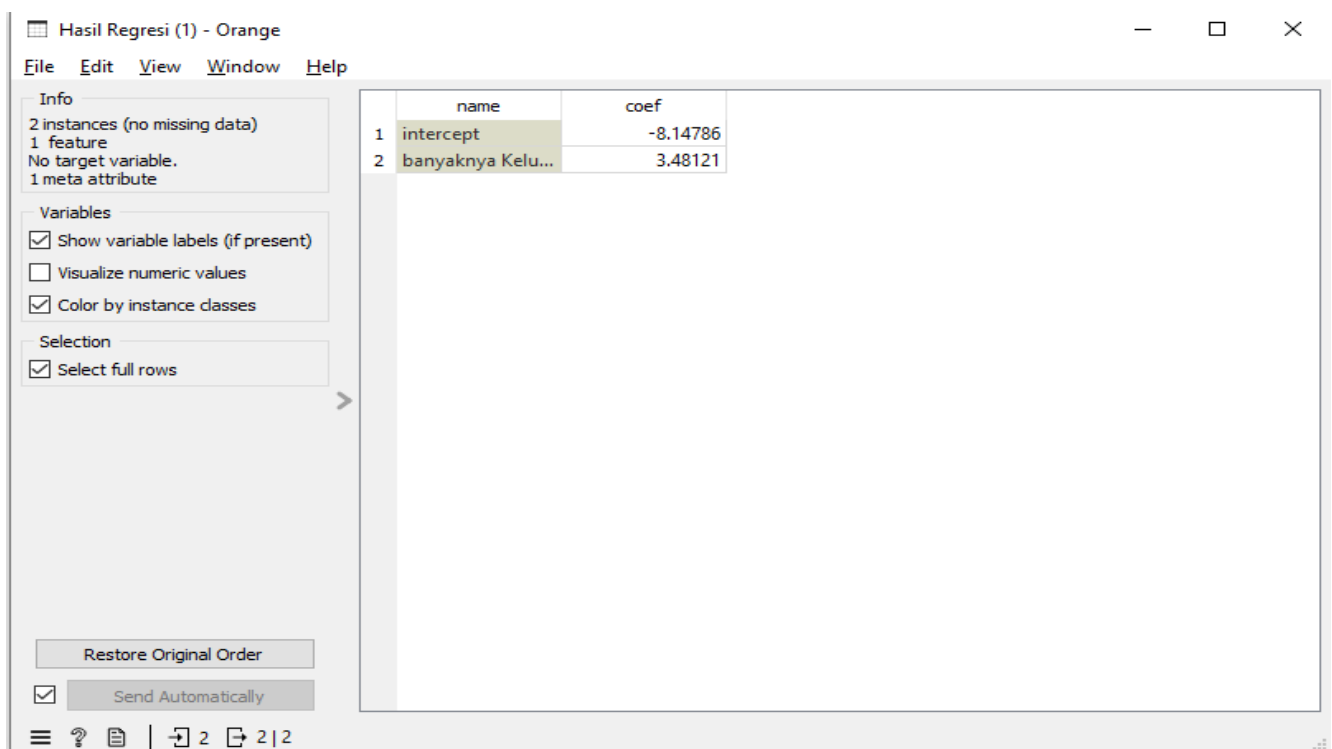


Gambar 3. Model perancangan di Orange data Mining

### III. HASIL DAN PEMBAHASAN

Analisis data dalam Orange Data Mining dimulai dengan mempersiapkan dataset berisi variabel jumlah keluarga dan total sampah (kg) yang diangkut. Penulis menggunakan tiga widget utama untuk memverifikasi hasil regresi linier sederhana serta kekuatan hubungan antar variabel, yakni widget Regression Table, Scatter Plot, dan Pearson Correlation. Tujuan utama adalah memperoleh nilai koefisien regresi, melihat pola sebaran data, dan mengukur derajat korelasi antara banyaknya keluarga dan akumulasi sampah.

Widget **Regression Table** (Gambar 4) di Orange menyajikan ringkasan model regresi linier sederhana yang telah dibuat. Intercept (konstanta) tercatat sebesar  $-8,14786$ , sedangkan koefisien untuk variabel “banyaknya keluarga” tercatat  $3,48121$ . Nilai intercept negatif ini menunjukkan bahwa pada kondisi ideal “0 keluarga,” prediksi sampah negatif, yang secara praktis berarti model tidak valid pada rentang usia atau kondisi di luar data. Koefisien  $3,48121$  menandakan bahwa setiap penambahan satu keluarga diprediksi akan menambah  $3,48121$  kg sampah harian.



Gambar 4. Widget icon Hasil Regresi (1)

Persamaan regresi linier sederhana yang terbentuk adalah:

$$\hat{Y} = -8.147 + 3.481X$$

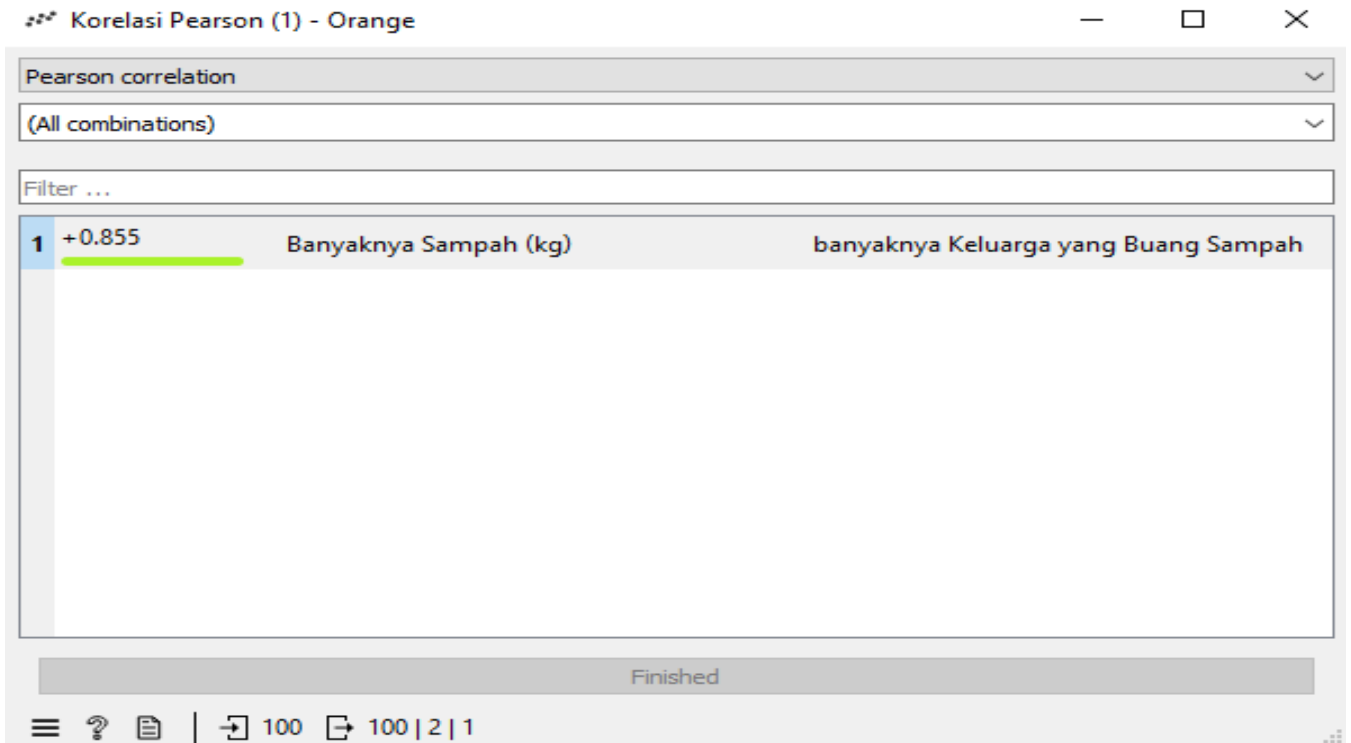
Di mana  $\hat{Y}$  adalah prediksi berat sampah (kg) dan  $X$  adalah jumlah keluarga. Persamaan ini dapat digunakan untuk memproyeksikan volume sampah berdasarkan jumlah keluarga, misalnya untuk 5

keluarga prediksi sampah harian sebesar  $-8,14786 + 3,48121 \times 5 = 9,20619$  kg. Interpretasi praktisnya, kebijakan pengelolaan sampah harus mempertimbangkan bahwa wilayah dengan lebih banyak keluarga akan menghasilkan sampah yang jauh lebih besar.

Widget Pearson Correlation (Gambar 5) memberikan nilai koefisien korelasi  $r=0,855$ . Angka ini

mengindikasikan kekuatan hubungan yang sangat kuat dan positif antara jumlah keluarga dan berat sampah harian. Semakin besar nilai  $r$  mendekati 1, semakin kuat hubungan linear antar variabel. Dalam konteks ini,

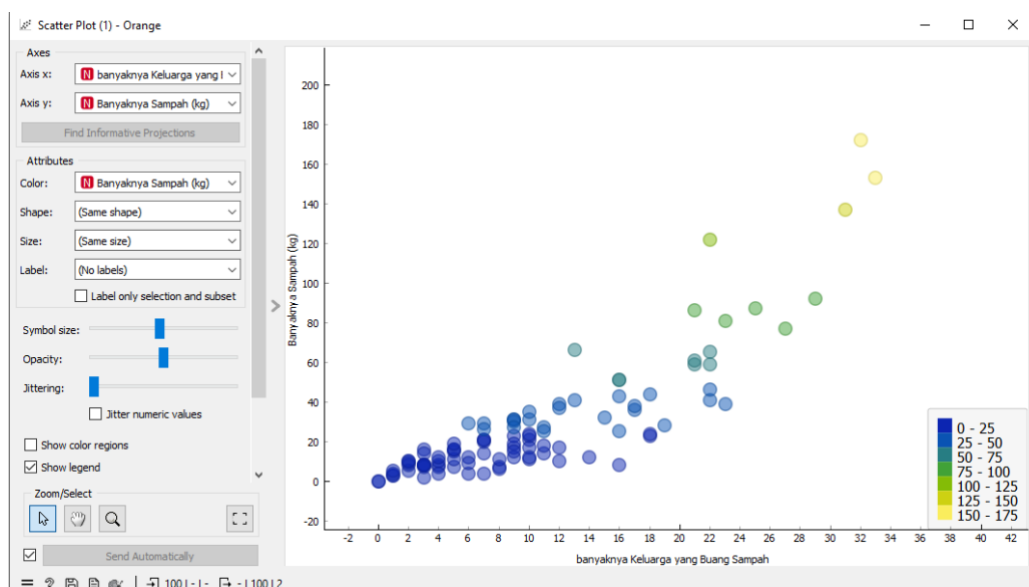
nilai 0,855 menunjukkan bahwa variabel jumlah keluarga menjelaskan sebagian besar variasi dalam total sampah.



Gambar 5. Widget icon Korelasi Pearson(1)

Nilai korelasi sebesar 0,855 juga memenuhi kriteria “korelasi sangat kuat” menurut pedoman umum (nilai di atas 0,7). Hal ini menegaskan bahwa peningkatan jumlah keluarga hampir selalu diikuti oleh peningkatan berat sampah harian. Oleh karena itu, intervensi pengurangan sampah di lingkungan RT 01 RW 027 sebaiknya fokus pada strategi penurunan produksi sampah per keluarga, misalnya melalui program pengomposan atau daur ulang.

Widget Scatter Plot (Gambar 6) memvisualisasikan setiap observasi sebagai titik pada bidang dua dimensi: sumbu horizontal menampilkan jumlah keluarga dan sumbu vertikal menampilkan berat sampah. Pola titik-titik tersebut mengikuti arah garis regresi naik, sehingga mempermudah identifikasi outlier dan pola umum. Beberapa titik yang jauh dari garis tren dapat menunjukkan kasus perlakuan atau kebiasaan pembuangan sampah yang tidak biasa.



Gambar 6. Widget icon Scatter Plot (1)

Visualisasi melalui scatter plot membantu mengevaluasi kesesuaian model regresi: semakin rapat titik-titik di sekitar garis, semakin baik model menjelaskan data. Dalam hasil ini, sebagian besar titik berada dekat dengan garis, menandakan model regresi linier sederhana sudah cocok. Titik-titik outlier perlu dianalisis lebih lanjut untuk memahami faktor-faktor unik, seperti acara masyarakat atau gangguan operasional petugas sampah.

Secara keseluruhan, kombinasi hasil regresi—dengan koefisien yang signifikan—nilai korelasi tinggi, dan pola scatter plot yang konsisten, menunjukkan bahwa model regresi linier sederhana di Orange valid untuk memprediksi akumulasi sampah berdasarkan jumlah keluarga. Rekomendasi selanjutnya adalah melakukan analisis residual untuk mengecek asumsi model dan memperluas variabel prediktor (misalnya

pendapatan atau kesadaran lingkungan) guna meningkatkan akurasi prediksi. Dengan demikian, hasil ini dapat menjadi dasar kebijakan penentuan tarif iuran sampah dan strategi pengelolaan lingkungan di Perumahan Banten Indah Permai.

### Perhitungan Manual Regresi Linier Sederhana

Pada perhitungan manual ini, variabel bebas X (jumlah keluarga) dan variabel terikat Y (berat sampah harian dalam kg) ditetapkan terlebih dahulu. Untuk mendukung perhitungan, dibuat kolom tambahan pada tabel:  $X^2$ ,  $Y^2$ , dan  $XY$ . Kemudian, dihitung total  $\sum X$ ,  $\sum Y$ ,  $\sum x^2$ ,  $\sum Y^2$ ,  $\sum XY$ , berdasarkan 100 observasi (Tabel 2).

Table 1. Tabel Perhitungan  $\sum X, Y, X^2, Y^2, XY$

No.	Hari ke-	(X)	(Y)	$X^2$	$Y^2$	XY
1	31/08/2023	13	66	169	4356	858
2	01/09/2023	7	21	49	441	147
3	02/09/2023	22	122	484	14884	2684
4	03/09/2023	5	16	25	256	80
5	04/09/2023	7	29	49	841	203
6	05/09/2023	17	38	289	1444	646
7	06/09/2023	21	86	441	7396	1806
8	07/09/2023	29	92	841	8464	2668
9	08/09/2023	10	35	100	1225	350
10	09/09/2023	9	19	81	361	171
...	...	...	...	...	...	...
...	...	...	...	...	...	...
...	...	...	...	...	...	...
99	07/12/2023	10	24	100	576	240
100	08/12/2023	11	18	121	324	198
Total	$\sum$	1076	2931	17614	185969	52551

Setelah nilai-nilai  $\sum$  ditemukan, langkah berikutnya adalah menghitung koefisien regresi  $b_1$  (slope) dan intersep  $b_0$  menggunakan rumus kuadrat terkecil:

$$a = \frac{(\sum Y)(\sum X^2) - (\sum X)(\sum XY)}{(n)(\sum X^2) - (\sum X)^2}$$

$$b = \frac{(n)(\sum XY) - (\sum X)(\sum Y)}{(n)(\sum X^2) - (\sum X)^2}$$

$$b_0 = \frac{(2931)(17614) - (1076)(52551)}{(100)(17614) - (1076)^2}$$

$$b_0 = \frac{(2931)(17614) - (1076)(52551)}{(100)(17614) - (1157776)}$$

$$b_0 = \frac{51626634 - 56544876}{1761400 - 1157776}$$

$$b_0 = -8.147$$

Dengan demikian, persamaan regresi linier sederhana adalah:

$$Y = -8.147 + 3.481X$$

Selanjutnya, koefisien determinasi  $R^2$  dihitung untuk mengukur seberapa besar variasi Y yang dijelaskan oleh X:

$$R^2 = \frac{((n)(\sum XY) - (\sum X)(\sum Y))^2}{(n)(\sum X^2) - (\sum X)^2(n)(\sum Y^2) - (\sum Y)^2}$$

Perhitungan intersep  $b_0$  diperoleh sebagai berikut:

$$b_1 = \frac{(100)(52551) - (1076)(2931)}{(100)(17614) - (1076)^2}$$

$$b_1 = \frac{(100)(52551) - (1076)(2931)}{(100)(17614) - (1157776)}$$

$$b_1 = \frac{5255100 - 3153756}{1761400 - 1157776}$$

$$b_1 = 3.481$$

$$R^2 = \frac{((100)(52551) - (1076)(2931))^2}{(100(17614) - (1076)^2)(100(185969) - (2931)^2)}$$

$$R^2 = \frac{(2101344)^2}{(1761400 - 1076^2)(18596900 - 2931^2)}$$

$$R^2 = +0.855$$

Nilai  $R^2=0,855$  mengindikasikan bahwa 85,5 % variasi berat sampah harian di RT 01 RW 027 dapat dijelaskan oleh jumlah keluarga. Hasil ini konsisten dengan perhitungan pada Orange Data Mining, sehingga model manual dan model aplikasi terbukti sama validnya.

## V. KESIMPULAN

Berdasarkan hasil analisis regresi linier sederhana dan korelasi Pearson, terdapat hubungan positif yang sangat kuat ( $b_1 = 3,481$ ,  $R^2 = 0,855R$ ) antara jumlah keluarga dan volume sampah harian di lingkungan RT 01 RW 027 Banten Indah Permai. Model regresi yang diperoleh ( $\hat{Y} = -8,147 + 3,481X$ ) dapat memprediksi berat sampah (kg) berdasarkan banyaknya keluarga, dengan tingkat kewajaran yang tinggi sesuai validasi di Orange Data Mining. Temuan ini menegaskan bahwa peningkatan jumlah keluarga secara linier meningkatkan akumulasi sampah di wilayah tersebut.

Hasil penelitian ini memberikan dasar kuantitatif bagi pengelola lingkungan dan petugas kebersihan untuk merencanakan frekuensi dan kapasitas pengangkutan sampah. Dengan mengetahui estimasi beban sampah harian—misalnya sekitar 110 kg untuk 34 kepala keluarga—pengambilan kebijakan tarif iuran dan penjadwalan kendaraan pengangkut dapat dioptimalkan. Selain itu, model ini juga dapat digunakan sebagai alat pemantauan dini untuk mendeteksi potensi penumpukan sampah dan merancang kampanye kesadaran agar setiap keluarga berperan aktif dalam pengurangan sampah.

Untuk meningkatkan akurasi dan cakupan temuan, penelitian mendatang disarankan melibatkan variabel tambahan seperti pendapatan per kapita, tingkat pendidikan, pola konsumsi, dan kesadaran lingkungan. Penggunaan metode regresi linier berganda atau algoritma machine learning (misalnya Random Forest atau Gradient Boosting) dapat mengeksplorasi interaksi kompleks antarvariabel. Selain itu, data longitudinal—mencakup beberapa periode pengukuran—dapat membantu memahami dinamika musim atau peristiwa khusus yang memengaruhi produksi sampah. Perlu pula dilakukan analisis residual dan pengujian asumsi model lebih mendalam untuk memastikan kesahihan inferensi statistik.

## DAFTAR PUSTAKA

- [1] A. Ramon and A. Afriyanto, "KARAKTERISTIK PENANGANAN SAMPAH RUMAH TANGGA DI KOTA BENGKULU," *Jurnal Kesehatan Masyarakat Andalas*, vol. 10, no. 1, 2017, doi: 10.24893/jkma.10.1.24-31.2015.
- [2] R. Riswan, H. R. Sunoko, and A. Hadiyanto, "PENGELOLAAN SAMPAH RUMAH TANGGA DI KECAMATAN DAHA SELATAN," *Jurnal Ilmu Lingkungan*, vol. 9, no. 1, 2012, doi: 10.14710/jil.9.1.31-38.
- [3] Salawati *et al.*, "Pemanfaatan Sampah Organik Rumah Tangga Metode Ember Tumpuk Menjadi Pupuk Organik Cair dan Padat," *Jurnal Pengabdian Masyarakat*, vol. 4, no. 3, 2021.
- [4] I. Indriyanti, N. Ichsan, H. Fatah, T. Wahyuni, and E. Ermawati, "IMPLEMENTASI ORANGE DATA MINING UNTUK PREDIKSI HARGA BITCOIN," *Jurnal Responsif: Riset Sains dan Informatika*, vol. 4, no. 2, 2022, doi: 10.51977/jti.v4i2.762.
- [5] P. Katemba and R. K. Djoh, "PREDIKSI TINGKAT PRODUKSI KOPI MENGGUNAKAN REGRESI LINEAR," *Jurnal Ilmiah Flash*, vol. 3, no. 1, 2017, doi: 10.32511/flash.v3i1.136.
- [6] C. Santoso, "IMPLEMENTASI ORANGE DATA MINING UNTUK PREDIKSI BIAYA ASURANSI," *Aisyah Journal Of Informatics and Electrical Engineering (A.J.I.E.E)*, vol. 5, no. 1, 2023, doi: 10.30604/jti.v5i1.180.
- [7] D. Safitri, S. S. Hilabi, and F. Nurapriani, "ANALISIS PENGGUNAAN ALGORITMA KLASIFIKASI DALAM PREDIKSI KELULUSAN MENGGUNAKAN ORANGE DATA MINING," *Rabit: Jurnal Teknologi dan Sistem Informasi Univrab*, vol. 8, no. 1, 2023, doi: 10.36341/rabit.v8i1.3009.
- [8] P. S. Brenner, *Understanding Survey Methodology*, vol. 4. 2020.
- [9] I. Sutomo and Budihardjo, "The rejection of religious nationalism towards the secular state and the Islamic caliphate: Indonesian religious figures perspective," *Indonesian Journal of Islam and Muslim Societies*, vol. 11, no. 1, 2021, doi: 10.18326/IJIMS.V11I1.115-137.
- [10] "Orange (software) - Wikipedia." Accessed: Jan. 19, 2024. [Online]. Available: [https://en.wikipedia.org/wiki/Orange\\_\(software\)](https://en.wikipedia.org/wiki/Orange_(software))