

ANALYSIS OF CUSTOMER WATER BILL PAYMENTS PDAM SURYA SEMBADA

Trimitha Karina Ningsih
3321600026

Lecture Advisor :

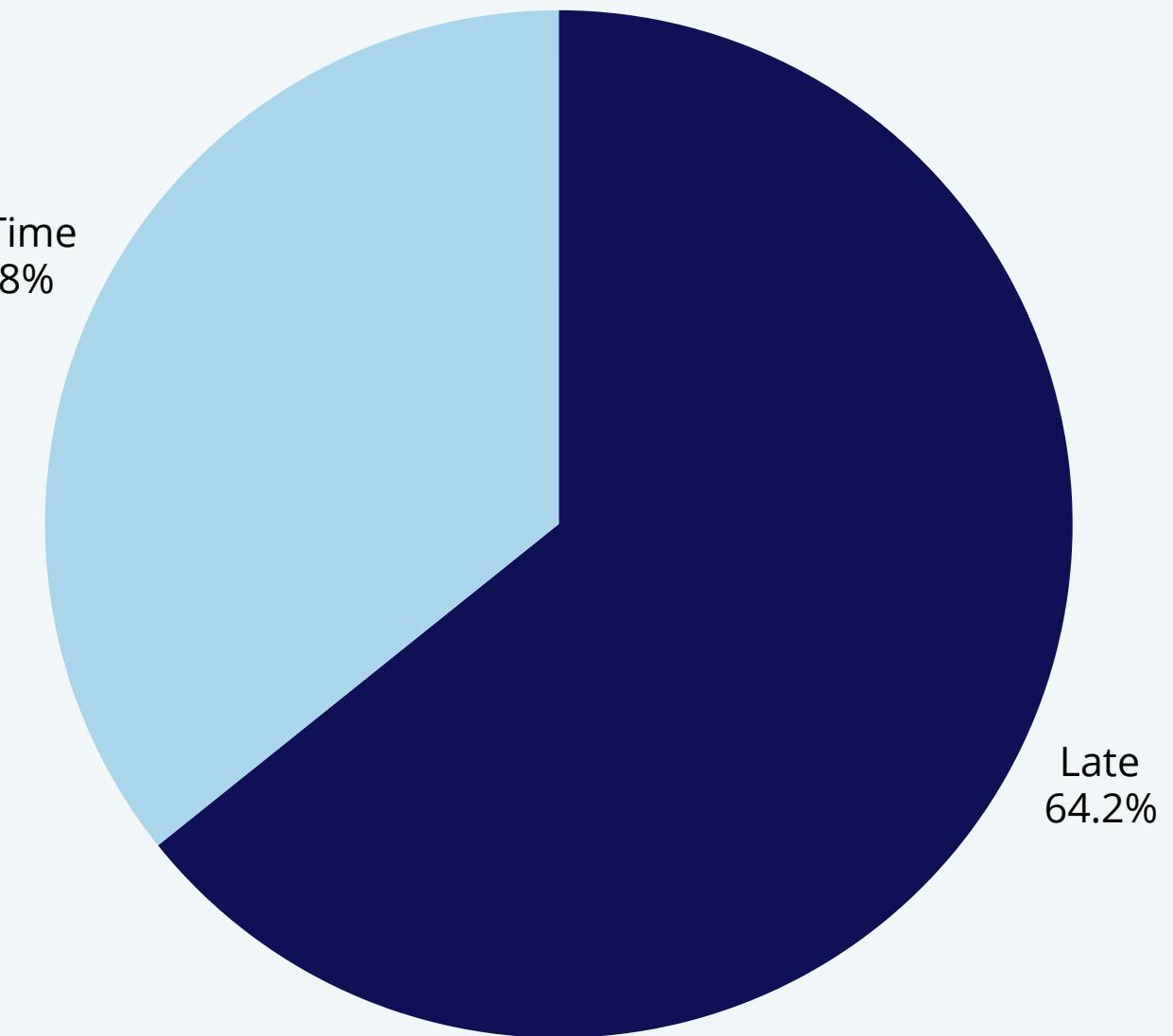
Ira Prastyaningrum,S.Si, MT
198005292008122005

Entin Martiana Kusumaningtyas,S.Kom, M.Kom
197403122000122001

Research Background

- PDAM Surya Sembada is a public company that provides clean water for Surabaya. It has a major challenge: 65.4% of customers are late in paying their water bills (January 2023 – March 2024).
- This delay has an impact on the company's working capital and operational smoothness.
- Currently, the company have only records customers who are in arrears, without the ability to predict potential arrears.
- There is no early warning, making it difficult to prevent delays.

customer payment status percentage in January 2023 - April 2024



Problem



High Delays

65.4% of customers delay their water bill payments.



Limited Existing System

Only records customers who have already defaulted, without future predictions.



Early Warning

PDAM lacks a system to proactively anticipate late payments.

Solution



Payment Analysis System

A system designed to analyze historical customer payment data to identify trends

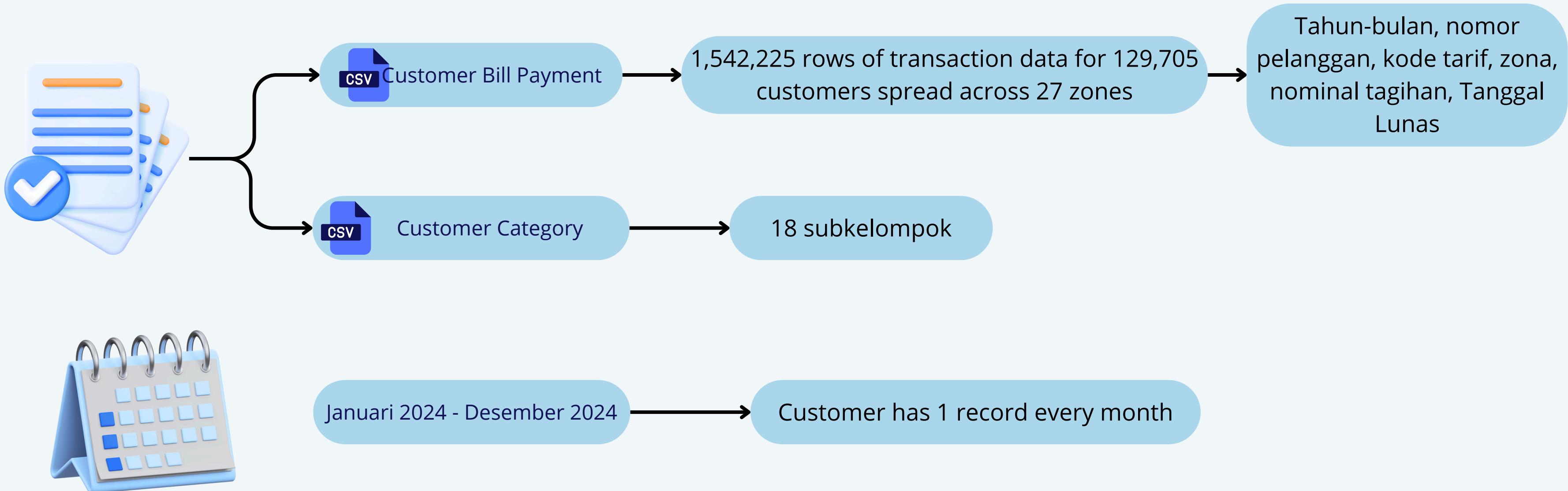


Predict Future Payment

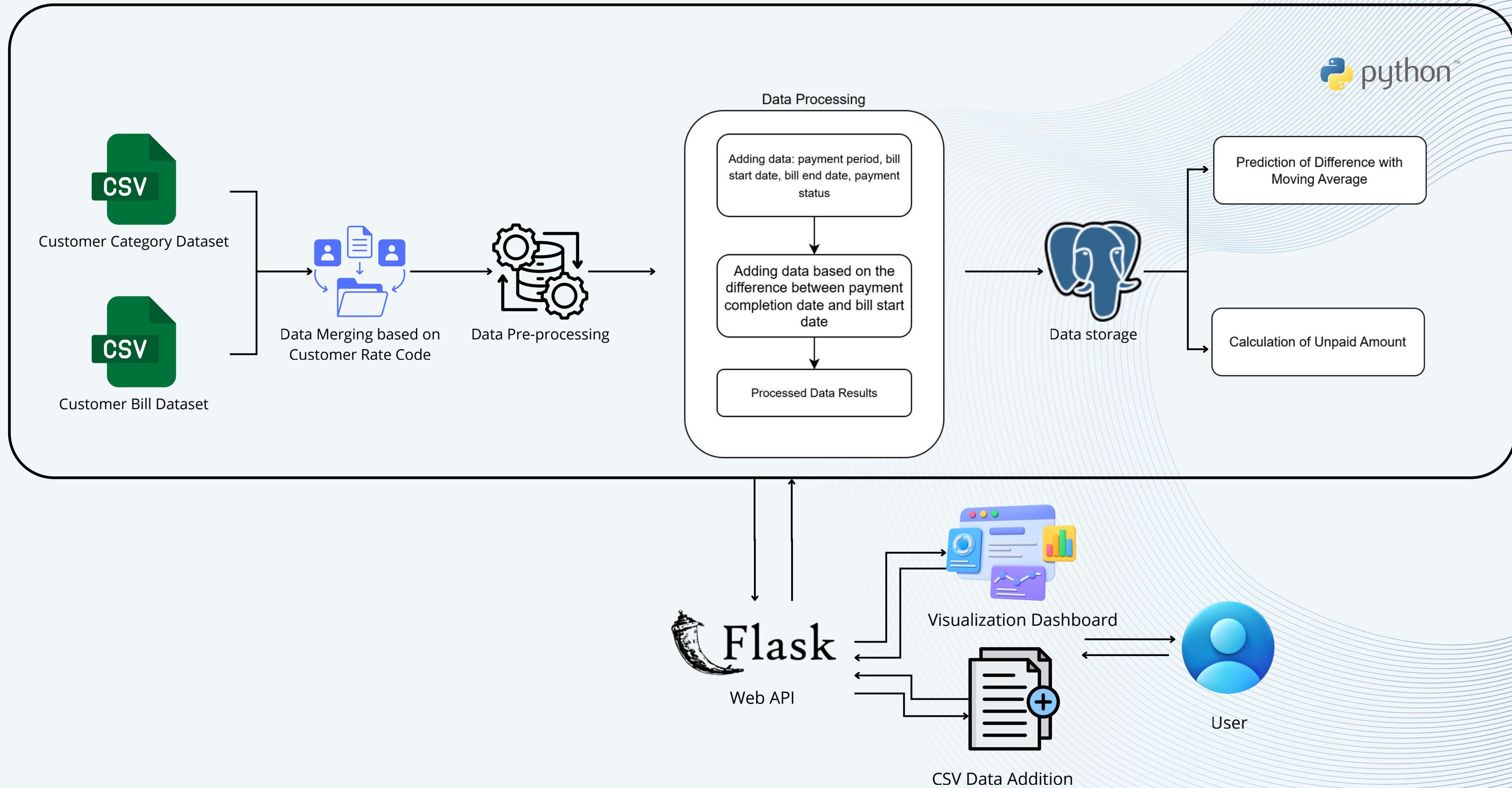
Predict future payment status

Dataset Information

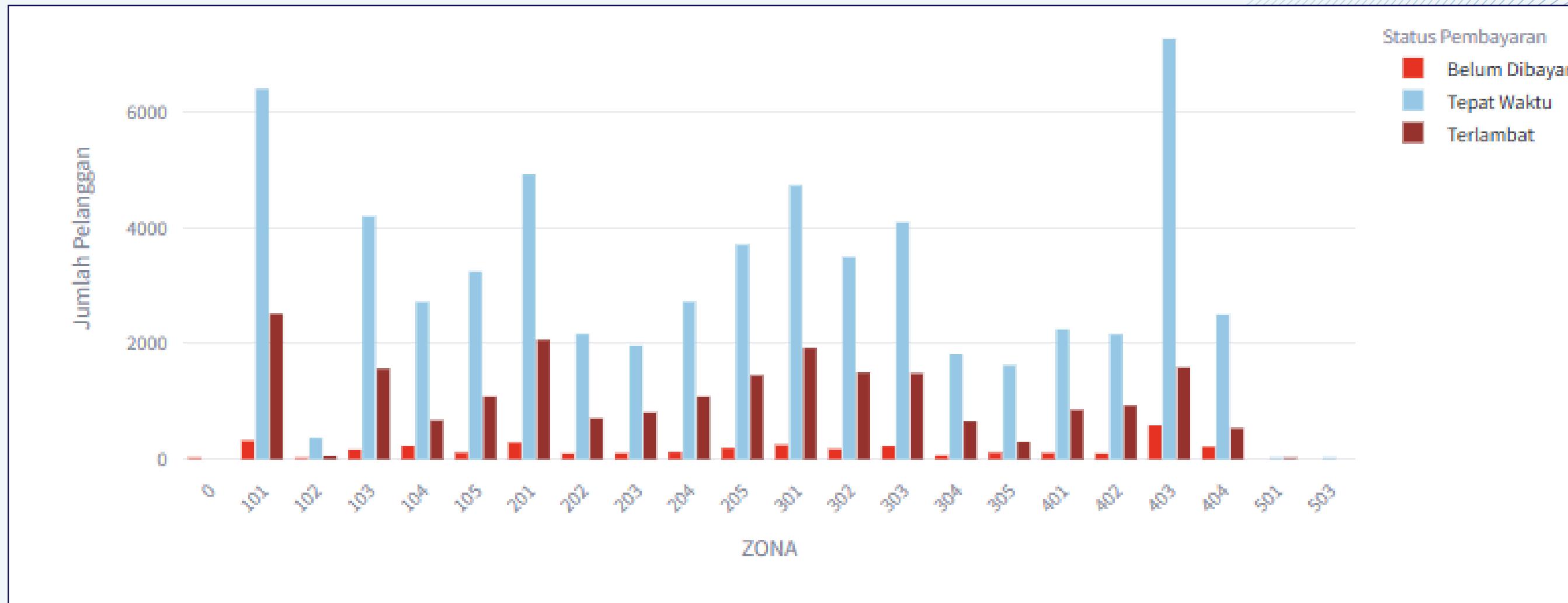
csv file data format obtained from PDAM Surya Sembada



Design System

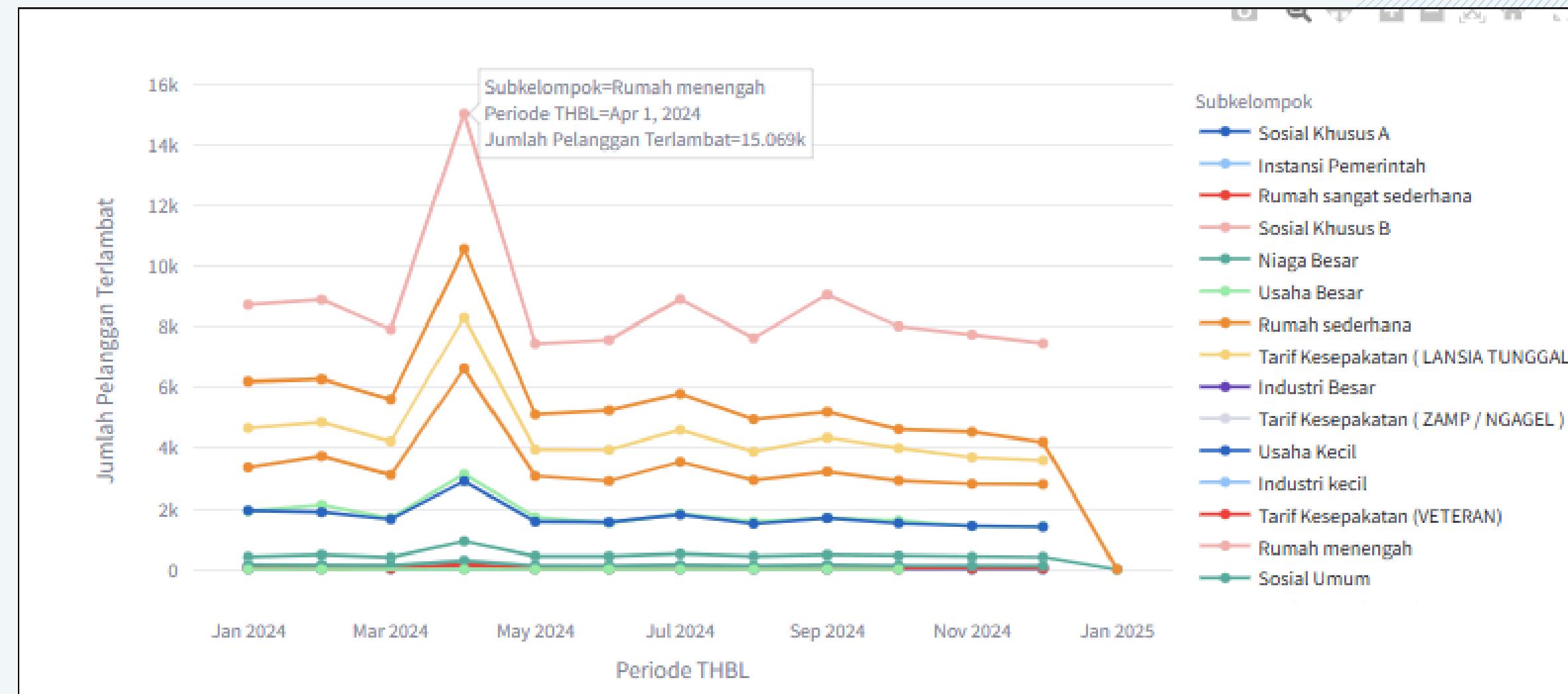


Dataset Analysis



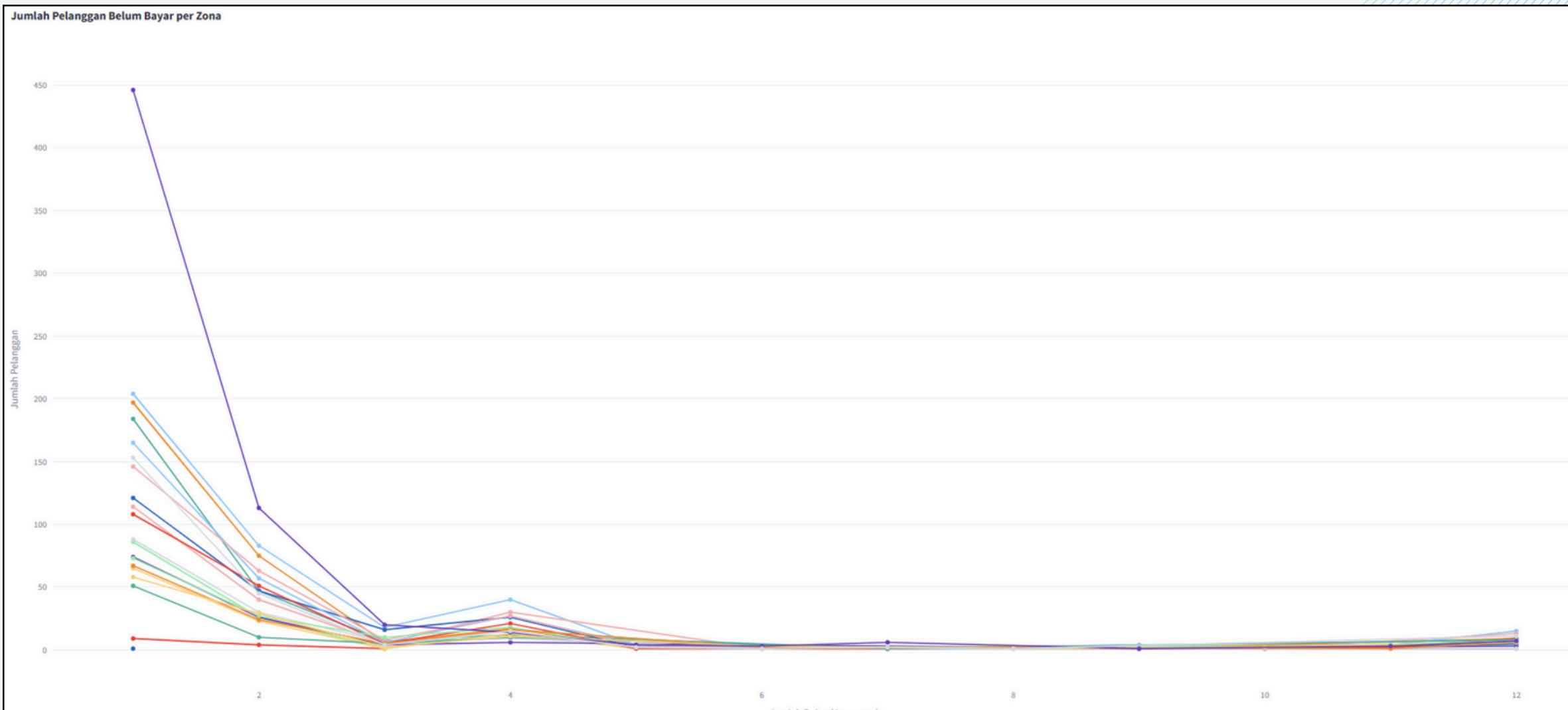
In December 2024, the number of late payments and customers who have not paid their bills is still quite high.

Dataset Analysis



The “**Rumah Menengah**” subgroup dominates the highest number of delays each month, which is in line with the largest number of customers coming from this group.

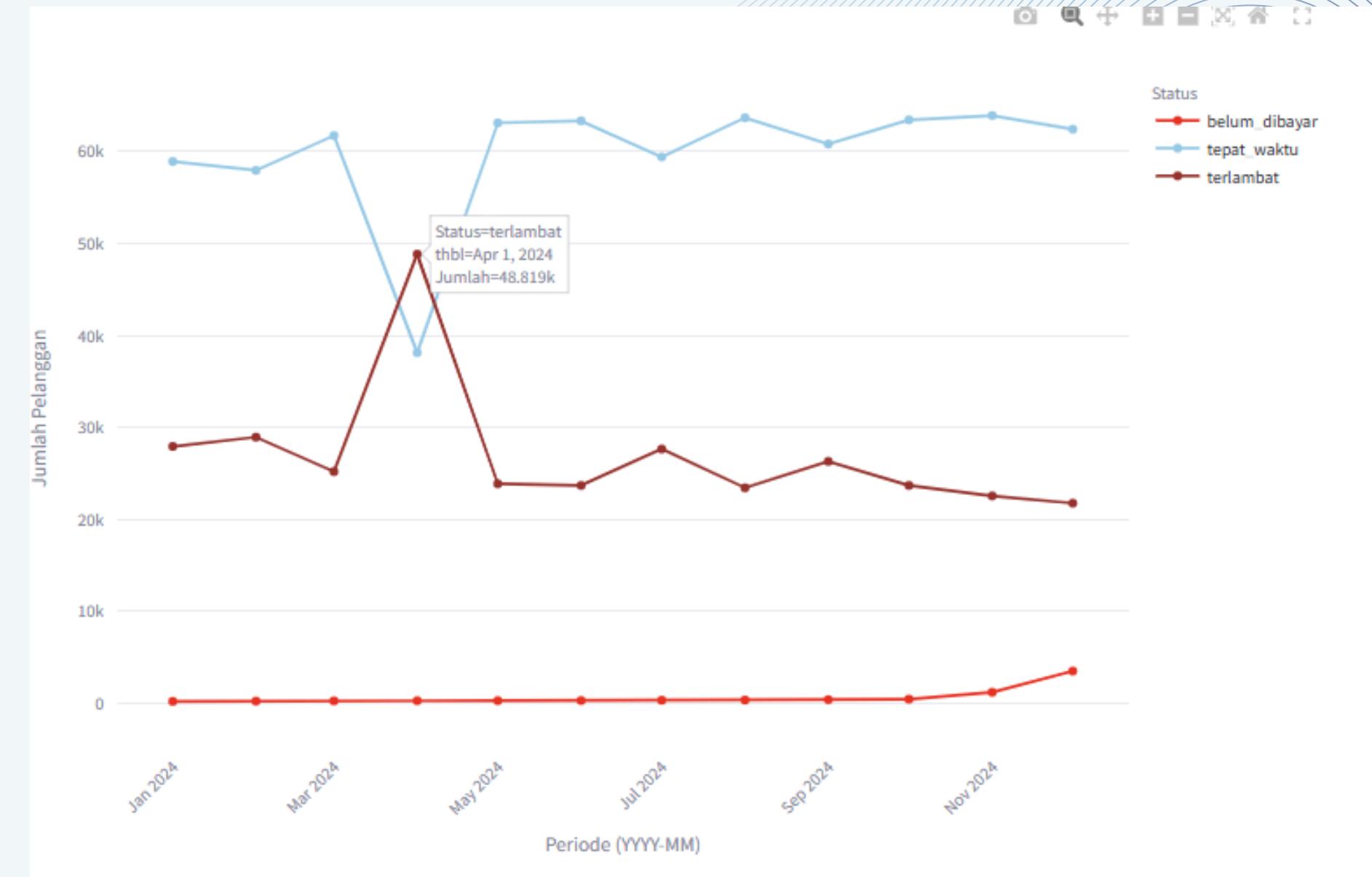
Dataset Analysis



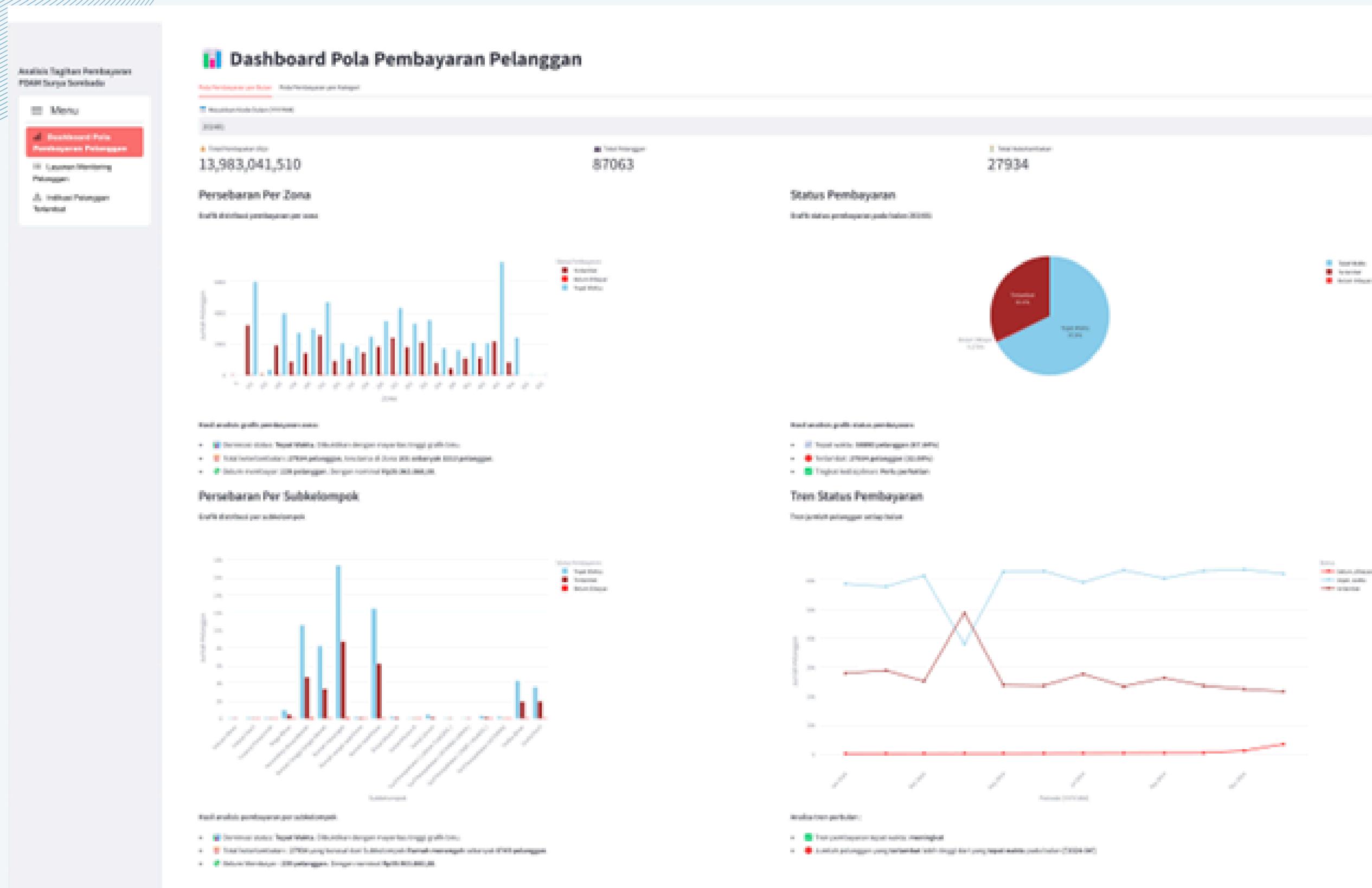
There are still customers who have not paid their bills for more than three months.

Dataset Analysis

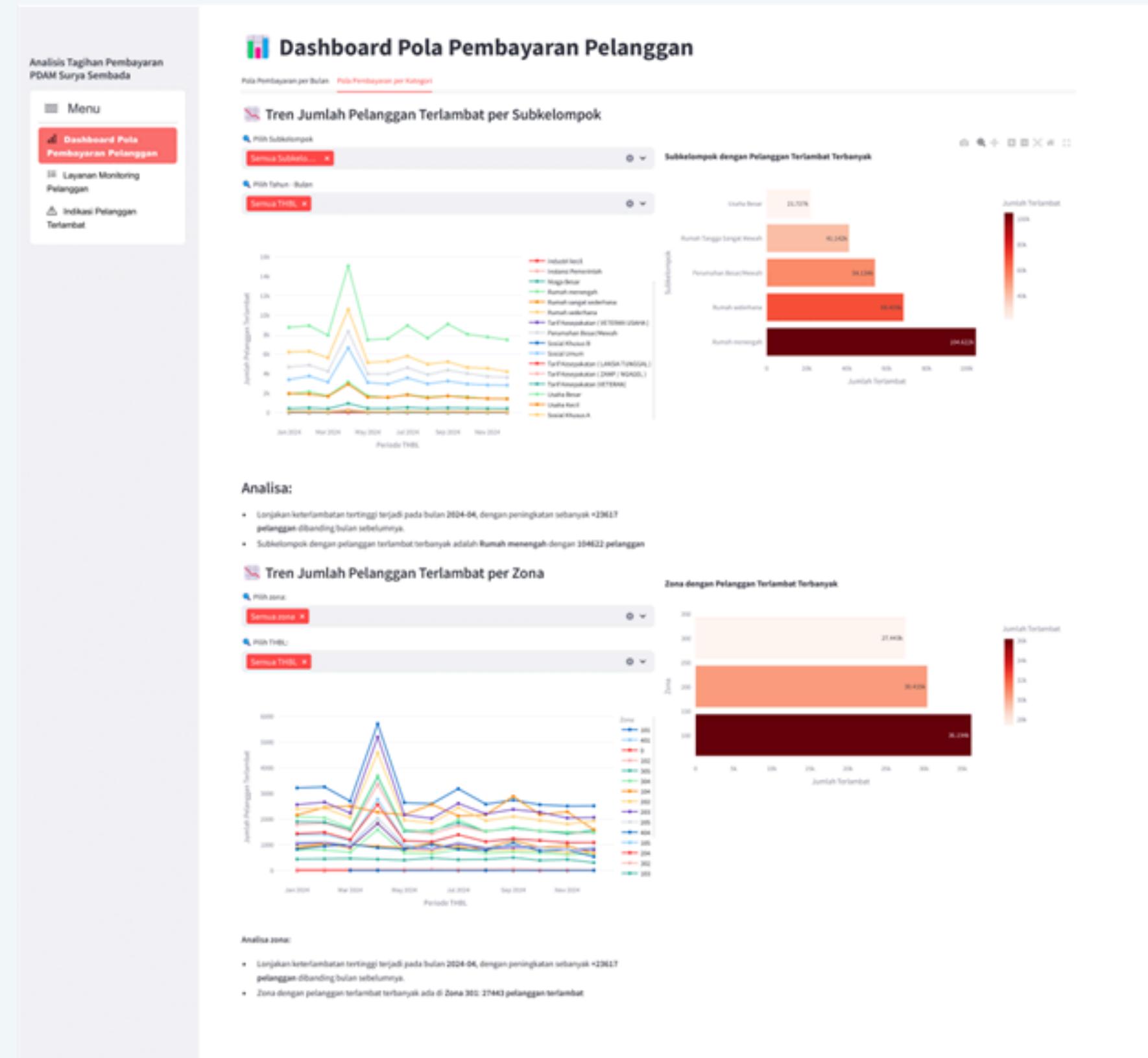
There was a spike in the number of delays in **April 2024** due to the long holiday "Eid Mubarak" which occurred in the beginning to the middle of the month. The majority of payments also occurred in period 1, namely the 1st - 15th of the month.



Dashboard Analysis Information



Dashboard Analysis Information



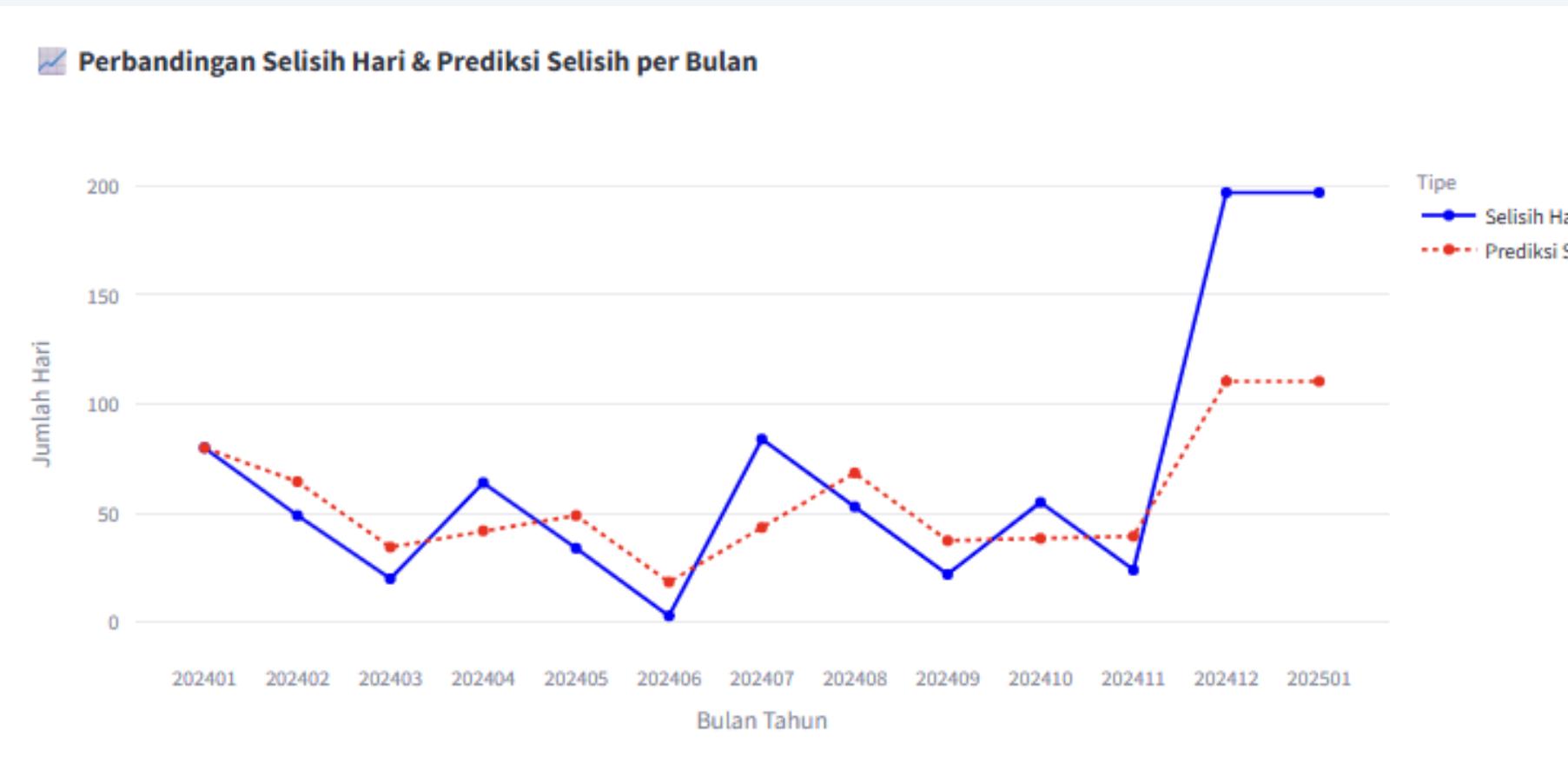
Prediction Method Using Moving Average

This prediction is made on the difference in days between the initial billing date and the payment date as the main target variable. This difference in days is taken as the main variable because it reflects the timeliness of payment by customers, which is an important indicator in analyzing payment behavior.

<i>Windows Size</i>	<i>MSE</i>	<i>MAE</i>	<i>MAPE</i>
2	31,2963	2.7182	36.06%
3	39,9360	2.8719	42.81%
4	39,5484	2.7768	43.90%
5	38,2084	2.5616	40.76%
6	35,0929	2.3029	37.25%

Window Size 2 gives the best performance: MSE=30.45, MAE=2.67, MAPE=37.19%, RMSE=5.52. Window size 2 is optimal because it is able to capture relatively stable short-term customer behavior patterns. The moving average model with window size 2 is more sensitive to dynamic changes or recurring delay patterns.

Implementation method for customer 11859AA



Early Year Stable Pattern: This customer exhibits a pattern of on-time payments with a low and stable day lag from January through August 2024.

Significant Late Spike: There is a large spike in late payments in September, peaking in October 2024.

Prediction Model Poorly Predicts Extreme Fluctuations: The moving average model is less able to predict this drastic spike, although it returns closer to the actual value once the payment pattern stabilizes.

Monitoring Service with an Indication of the Customer Being Late

Analisis Tagihan Pembayaran PDAM Surya Sembada

Menu

- Dashboard Pola Pembayaran Pelanggan
- Layanan Monitoring Pelanggan**
- Pelanggan Belum Membayar Tagihan
- Tambahkan Data

Pelanggan Potensial Terlambat Bayar Monitoring Tagihan Pelanggan

Pilih Tahun-Bulan

202501

Layanan Monitoring Pelanggan

Pelanggan Terlambat **87,165** orang

Zona Terbanyak **Zona 403** 9,402 pelanggan

Subkelompok Terbanyak **Rumah menengah** 28,319 orang

Detail Grafik

Pilih tampilan grafik:

Zona Subkelompok

Jumlah Pelanggan per Zona

Analisis Grafik :

Sebanyak 87165 pelanggan diprediksi akan mengalami keterlambatan dalam pembayaran tagihan air. Dari seluruh zona yang dianalisa, Zona 403 mencatat jumlah tertinggi dengan 9402 pelanggan yang diprediksi terlambat. Hal ini menunjukkan bahwa Zona 403 memiliki tingkat potensi keterlambatan yang paling tinggi dan dapat menjadi prioritas dalam penanganan lebih lanjut.

Lihat Detail Daftar Pelanggan

Zona	Jumlah Pelanggan
0	~8,500
101	~8,500
102	~500
103	~5,800
104	~3,800
105	~4,500
201	~7,200
202	~3,200
203	~3,000
204	~4,000
205	~5,300
301	~6,800
302	~5,200
303	~5,800
304	~2,500
305	~2,000
401	~3,200
402	~3,500
403	~9,400
404	~3,500

Monitoring Service with an Indication of the Customer Being Late

Pelanggan Potensial Terlambat Bayar [Monitoring Tagihan Pelanggan](#)

Masukkan Nomor Pelanggan:

[Search](#)

Bulan Tahun	Zona	Kode Tarif	Subkelompok	Periode	Awal Tagihan	Tanggal Lunas	Tanggal Tenggat	Jumlah Tagihan	Status Pembayaran	Selisih Hari	Prediksi Selisih
202401	0	3.4	Usaha Besar	1	2024-01-01	2024-03-21	2024-01-15	124740	Terlambat	80	80
202402	0	3.4	Usaha Besar	1	2024-02-01	2024-03-21	2024-02-15	133540	Terlambat	49	64.5
202403	0	3.4	Usaha Besar	1	2024-03-01	2024-03-21	2024-03-15	168740	Terlambat	20	34.5
202404	0	3.4	Usaha Besar	1	2024-04-01	2024-06-04	2024-04-15	216440	Terlambat	64	42
202405	0	3.4	Usaha Besar	1	2024-05-01	2024-06-04	2024-05-15	115940	Terlambat	34	49
202406	0	3.4	Usaha Besar	1	2024-06-01	2024-06-04	2024-06-15	115940	Tepat Waktu	3	18.5
202407	0	3.4	Usaha Besar	1	2024-07-01	2024-09-23	2024-07-15	115940	Terlambat	84	43.5
202408	0	3.4	Usaha Besar	1	2024-08-01	2024-09-23	2024-08-15	115940	Terlambat	53	68.5
202409	0	3.4	Usaha Besar	1	2024-09-01	2024-09-23	2024-09-15	115940	Terlambat	22	37.5
202410	0	3.4	Usaha Besar	1	2024-10-01	2024-11-25	2024-10-15	115940	Terlambat	55	38.5

Perbandingan Selisih Hari & Prediksi Selisih per Bulan

Analisis Grafik :

Berdasarkan hasil prediksi pada bulan terakhir yang tersedia, pelanggan diperkirakan akan melakukan pembayaran secara **terlambat**. Nilai rata-rata prediksi selisih hari mencapai **110.50** hari, yang menjadi indikator potensi keterlambatan pembayaran pada bulan berikutnya.

Pelanggan **11859AA** tidak mengalami belum bayar lebih dari satu bulan secara berturut-turut.

Jumlah Hari

Type

- Selisih Hari
- Prediksi Selisih

Outstanding customer information



Pelanggan Belum Membayar Tagihan

Detail Grafik Pelanggan Belum Bayar

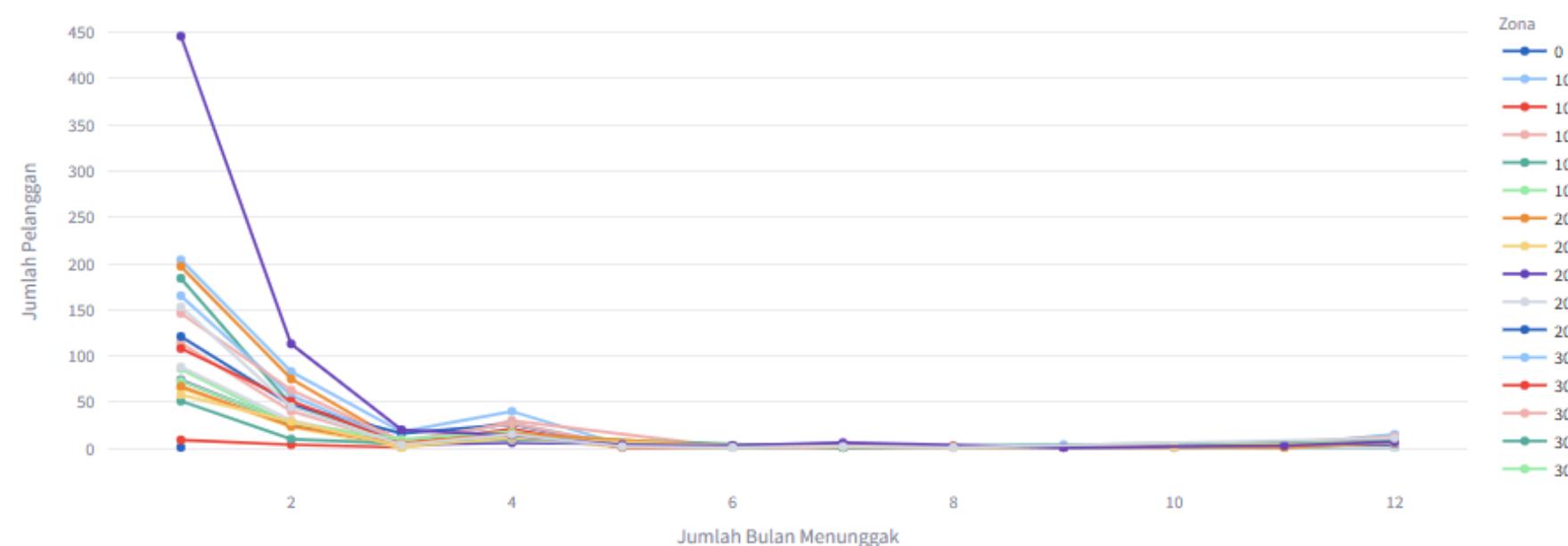
Pilih tampilan grafik:

Zona Subkelompok

Pilih zona:

Semua zona X

Jumlah Pelanggan Belum Bayar per Zona



Analisis Grafik :

Pelanggan dengan tunggakan paling lama tercatat memiliki keterlambatan pembayaran hingga 12 bulan, yang mencerminkan adanya risiko keterlambatan jangka panjang dalam sistem pembayaran.

Pada kategori tunggakan selama 12 bulan ini, Zona 301 menjadi wilayah dengan jumlah pelanggan belum membayar terbanyak.

Secara keseluruhan, terdapat sebanyak 15 pelanggan dalam kategori ini.

Temuan ini menunjukkan bahwa Zona 301 memerlukan perhatian khusus, karena menjadi pusat akumulasi pelanggan dengan potensi keterlambatan yang tinggi dan berkepanjangan.

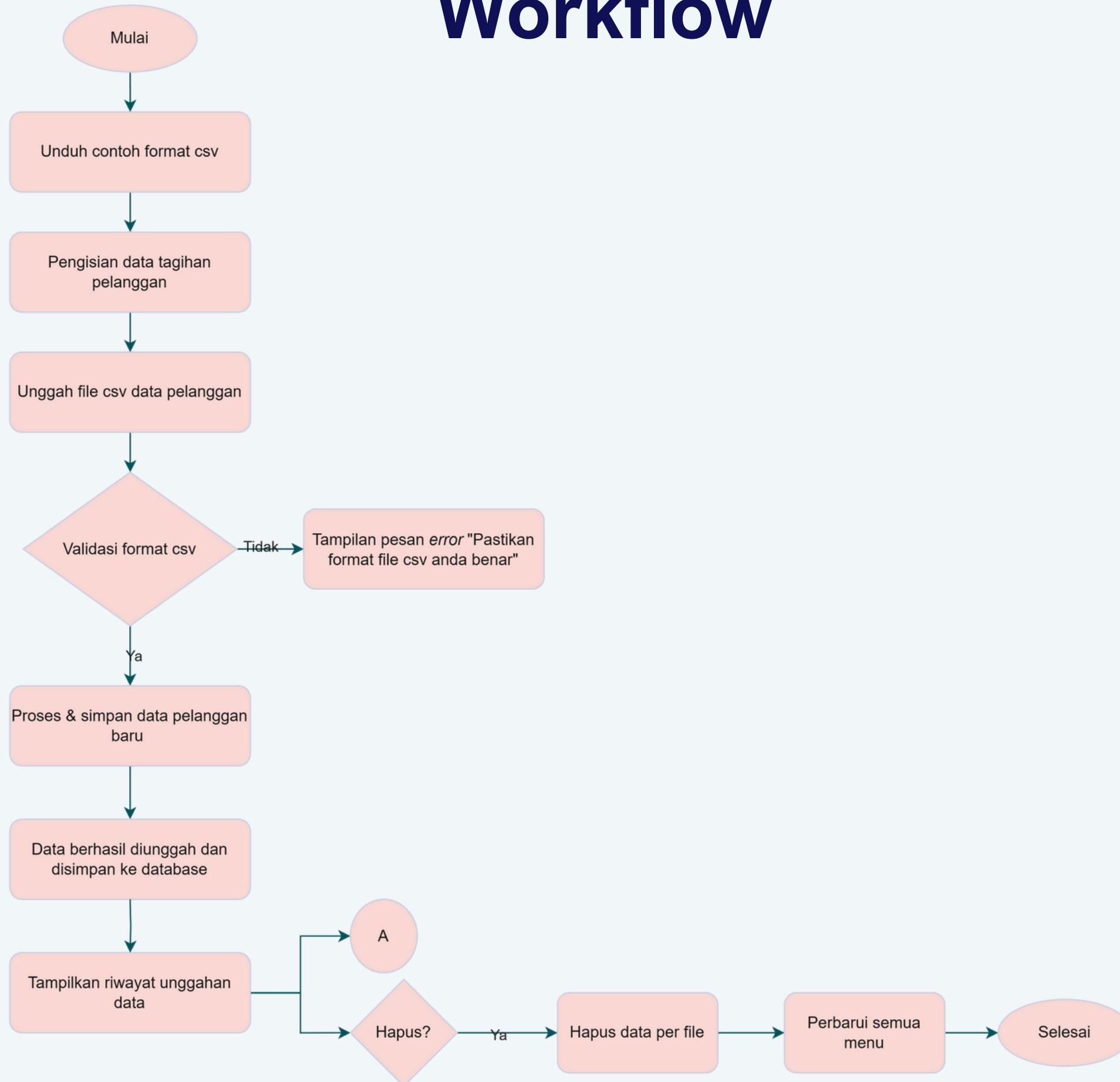
Lihat Detail Daftar Pelanggan Belum Bayar

Add Data Page

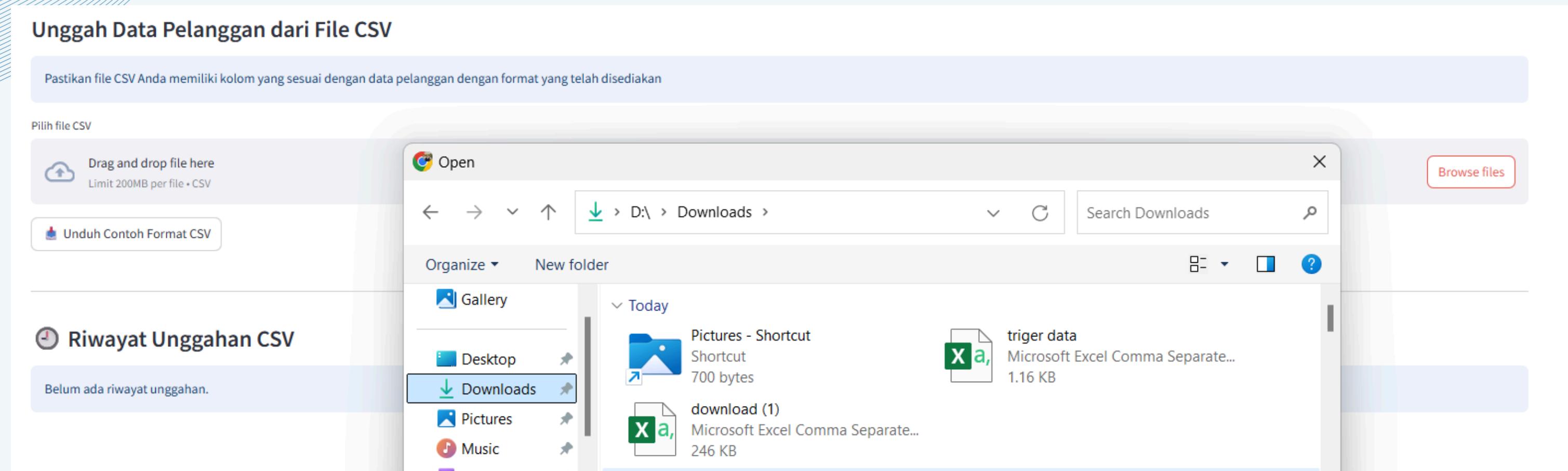
Purpose

- Enables users to upload new data into the system.
- Provides a user-friendly interface for updating customer and payment records.
- Ensures integration of new data with prediction and visualization modules

Workflow



CSV File Upload Form



Used to add customer billing or payment data in csv format

Data Preview and Format Validation



The screenshot shows a user interface for uploading a CSV file named "contoh_data_pelanggan_pdam.csv". A green success message indicates "File CSV berhasil diunggah!". Below this, a preview table titled "Pratinjau Data:" displays three rows of data with columns: kode_tagihan, thbl, no_plg, kd_tarif, subkelompok, zona, periode, awal_tagihan, tgl_lunas, tgl_tenggat, and rp_tagihan. The data rows are:

	kode_tagihan	thbl	no_plg	kd_tarif	subkelompok	zona	periode	awal_tagihan	tgl_lunas	tgl_tenggat	rp_tagihan
0	2025012259AA	202501	2259AA	3.1	Rumah menengah	101	1	2025-01-01	2025-01-14	2025-01-15	120000
1	2025012260BB	202501	2260BB	2.2	Niaga Besar	102	2	2025-01-01	2025-02-20	2025-01-15	150000
2	2025012261CC	202501	2261CC	1.1	Rumah sederhana	103	3	2025-01-01	2025-03-18	2025-01-15	100000

At the bottom left is a button labeled "Proses & Simpan ke Database".

The system only accepts files with a specific format and column structure and displays a sample of the uploaded data for user confirmation before saving.

Data Preview and Format Validation



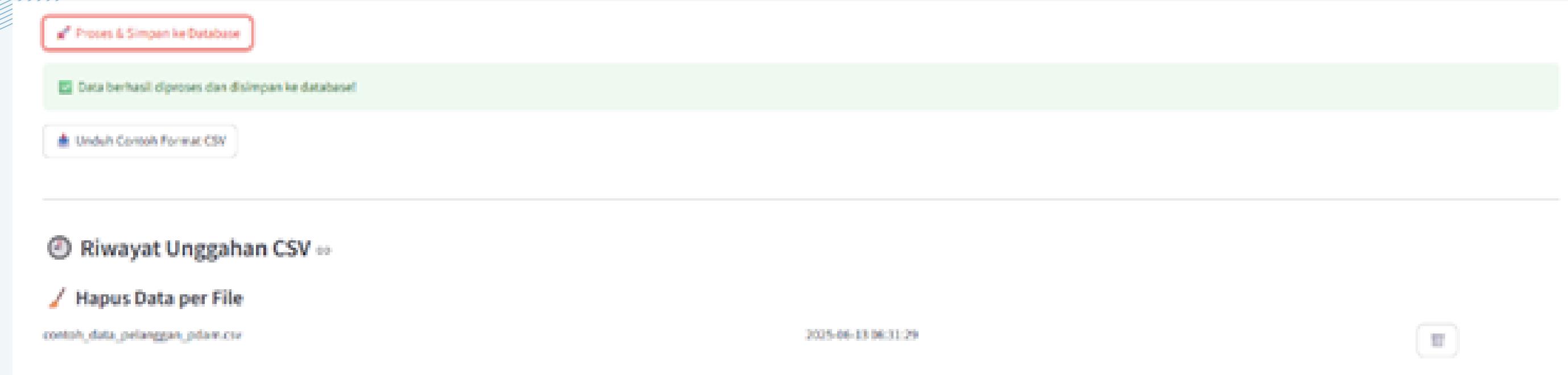
The screenshot shows a user interface for uploading a CSV file named "contoh_data_pelanggan_pdam.csv". A green success message indicates "File CSV berhasil diunggah!". Below this, a preview table titled "Pratinjau Data:" displays three rows of data with columns: kode_tagihan, thbl, no_plg, kd_tarif, subkelompok, zona, periode, awal_tagihan, tgl_lunas, tgl_tenggat, and rp_tagihan. The data rows are:

	kode_tagihan	thbl	no_plg	kd_tarif	subkelompok	zona	periode	awal_tagihan	tgl_lunas	tgl_tenggat	rp_tagihan
0	2025012259AA	202501	2259AA	3.1	Rumah menengah	101	1	2025-01-01	2025-01-14	2025-01-15	120000
1	2025012260BB	202501	2260BB	2.2	Niaga Besar	102	2	2025-01-01	2025-02-20	2025-01-15	150000
2	2025012261CC	202501	2261CC	1.1	Rumah sederhana	103	3	2025-01-01	2025-03-18	2025-01-15	100000

At the bottom left is a button labeled "Proses & Simpan ke Database".

The system only accepts files with a specific format and column structure and displays a sample of the uploaded data for user confirmation before saving.

Save Data Button



Commits the validated data to the database. Data is saved to the database, and the system automatically updates predictions and visualizations.

Daftar Pustaka

- Surabaya, P. S. S. K. (n.d.). Sejarah & Status PDAM Surya Sembada Surabaya | PDAM Surya Sembada Kota Surabaya. <https://www.pdam-sby.go.id/read/sejarah-status-pdam-surya-sembada-surabaya>
- PERBANDINGAN ALGORITMA K-MEANS DAN NAÏVE BAYES UNTUK MEMPREDIKSI PRIORITAS PEMBAYARAN TAGIHAN RUMAH SAKIT BERDASARKAN TINGKAT KEPENTINGAN PADA PT. PERTAMINA (PERSERO). (n.d.). <https://ejurnal.ulbi.ac.id/index.php/informatika/article/view/1383/809>
- Ilmi R.H. Zer, P. F., Batubara, E. R., Alkhairi, P., Tambunan, F. N., & Rosnelly, R. (2022). PENGENALAN POLA KEMAMPUAN PELANGGAN DALAM MEMBAYAR AIR PDAM MENGGUNAKAN ALGORITMA NAÏVE BAYES. Jurnal TIMES, 10(2), 29–38. <https://doi.org/10.51351/jtm.10.2.2021656>
- Bogdanchikov, A., Zhaparov, M., & Sulihev, R. (2013). Python to learn programming. Journal of Physics. Conference Series, 423, 012027. <https://doi.org/10.1088/1742-6596/423/1/012027>
- M. R. S. Alfarizi, M. Z. Al-farish, M. Taufiqurrahman, G. Ardiansah, and M. Elgar, “Penggunaan Python Sebagai Bahasa Pemrograman untuk Machine Learning dan Deep Learning,” Karimah Tauhid, vol. 2, no. 1, pp. 1–6, Jan. 2023, doi: 10.30997/KARIMAHTAUHID.V2I1.7518.
- Kaggle: your machine learning and data science community. (n.d.). <https://www.kaggle.com/>
- Relan, K. (2019). Building REST APIs with Flask. In Apress eBooks. <https://doi.org/10.1007/978-1-4842-5022-8>
- Momjian, B. (2000). PostgreSQL: Introduction and Concepts.
- Yalamarthi, R. N. S. V., Shaik, S., Singh, D., & Rakhra, M. (2022). Real-Time Face Mask Detection Using Streamlit, TensorFlow, Keras and Open-CV. 2022 IEEE International Conference on Data Science and Information System (ICDSIS). <https://doi.org/10.1109/icdsis55133.2022.9915817>
- Appiah, K.O., Chizema, A., & Arthur, J. (2015). Predicting corporate failure: a systematic literature review of methodological issues. International Journal of Law and Management, 57, 461-485.
- Al-Taie, M. Z., Kadry, S., & Lucas, J. P. (2019). Online Data Preprocessing: A Case Study Approach. International Journal of Electrical and Computer Engineering (IJECE), 9(4), 2620. <https://doi.org/10.11591/ijece.v9i4.pp2620-2626>
- Rachman, R. (2018). Penerapan metode moving average dan exponential smoothing pada peramalan produksi industri garment. Jurnal Informatika, 5(2), 211–220. <https://doi.org/10.31294/ji.v5i2.3309>
- Moreno, J. J. M., Pol, A. L. P., Abad, A. J. S., & Blasco, B. C. (2013b). Using the R-MAPE index as a resistant measure of forecast accuracy. PubMed, 25(4), 500–506. <https://doi.org/10.7334/psicothema2013.23>
- Intan, S., Sapulette, W., & Soukotta, R. C. (2020, July 1). Analisis Keterlambatan Waktu Pelaksanaan Proyek Konstruksi Di Kota Ambon: Klasifikasi Dan Peringkat Dari Penyebab-Penyebabnya. Intan | Manumata: Jurnal Ilmu Teknik. <http://ojs.ukim.ac.id/index.php/manumata/article/view/441/334>
- Subagyo, P. (1986). Forecasting: Konsep dan aplikasi. Yogyakarta: BPPE UGM.
- Stewart, & Ken. (2025, April 13). Mean squared error (MSE) | Definition, Formula, Interpretation, & Facts. Encyclopedia Britannica. <https://www.britannica.com/science/mean-squared-error>
- Hodson, T. O. (2022). Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. Geoscientific Model Development, 15, 5481–5487. [Commits the validated data to the database. Data is saved to the database, and the system automatically updates predictions and visualizations.](#)
- Putra, I. K. G. A., & Ariwijaya, M. (2023). Analisis dan visualisasi data untuk meningkatkan penjualan menggunakan Exploratory Data Analysis (EDA). Jurnal Ilmu Komputer, 12(3), 227-234. <https://doi.org/10.24843/JLK.2023.v12.i03.p07>
- Qian, H. (2023). Stock price prediction: Moving average and Markov chain. Advances in Economics Management and Political Sciences, 36(1), 66–71. <https://doi.org/10.54254/2754-1169/36/20231786>
- An Ensemble Approach to Stock Price Prediction Using Deep Learning and Time Series Models. (2024). Journal Name, Volume (Issue), Page range. <https://doi.org/10.4108/eai.5-1-2024.2342607>
- Intan | Manumata: Jurnal Ilmu Teknik. <http://ojs.ukim.ac.id/index.php/manumata/article/view/441/334>
- Stock Price Prediction using Moving Average Time Series. (2023). Journal of Financial Data Science , 5 (2), 45-60. <https://doi.org/10.4108/eai.5-1-2024.2342607>

Thank You

