
Implementation of Linear Regression to Predict Product Sales at Nursila Bakery Store

Muhammad Meikel Yola ¹⁾, Hanifah Ekawati ²⁾, dan Ahmad Abul Khoir ³⁾

^{1,3}Teknik Informatika, STMIK Widya Cipta Dharma

²Bisnis Digital, STMIK Widya Cipta Dharma

^{1,2,3}Jl. M. Yamin, Gn. Kelua, Kec. Samarinda Ulu, Kota Samarinda, Kalimantan Timur 75123

E-mail: 2043125@wicida.ac.id¹⁾, hanifah@wicida.ac.id²⁾, khair@wicida.ac.id³⁾

ABSTRACT

Uncertainty in consumer demand is a major challenge for Micro, Small, and Medium Enterprises (MSMEs), including Nursila Bakery. This study aims to develop a web-based cookie sales prediction system using the Linear Regression algorithm and the CRISP-DM methodology. Data used are derived from historical sales records during the Idul Fitri period from 2021 to 2025. Analysis was conducted using Simple Linear Regression for single-product prediction and Multiple Linear Regression with interaction variables to compare products based on size categories. Results indicate that all regression models are statistically significant (Sig. < 0.05). The models produced strong coefficients of determination (R^2), ranging from 0.863 for medium-sized Nastar to 0.990 for small-sized Nastar. In the multiple model, the combination of Semprit Susu and Putri Salju Vanilla (Medium) achieved the highest accuracy with an R^2 of 0.936. The implementation is expected to assist management in planning production and inventory strategies more accurately.

Keywords: *Linear Regression, Sales Prediction, CRISP-DM, SME.*

Implementasi Regresi Linear Untuk Memprediksi Penjualan Produk Pada Toko Nursila Bakery

ABSTRAK

Ketidakpastian permintaan konsumen merupakan tantangan utama bagi UMKM, termasuk Toko Nursila Bakery. Penelitian ini bertujuan membangun sistem prediksi penjualan kue kering berbasis web menggunakan algoritma Regresi Linear dengan metodologi CRISP-DM. Data yang digunakan berasal dari pencatatan penjualan historis periode Idul Fitri tahun 2021-2025. Analisis dilakukan melalui dua pendekatan: Regresi Linear Sederhana untuk produk tunggal dan Regresi Linear Berganda dengan variabel interaksi untuk membandingkan produk berdasarkan kategori ukuran. Hasil analisis menunjukkan seluruh model regresi signifikan secara statistik (Sig. < 0.05). Model menghasilkan koefisien determinasi (R^2) yang kuat, berkisar antara 0,863 pada produk Nastar Sedang hingga 0,990 pada Nastar Kecil. Pada model berganda, kombinasi Semprit Susu dan Putri Salju Vanilla (Sedang) memiliki akurasi tertinggi dengan nilai R^2 sebesar 0,936. Implementasi sistem ini diharapkan membantu manajemen dalam merencanakan strategi produksi dan persediaan secara lebih akurat.

Kata Kunci: Regresi Linear, Prediksi Penjualan, CRISP-DM, UMKM

1. PENDAHULUAN

Perkembangan teknologi informasi serta meningkatnya ketersediaan data telah mendorong pelaku usaha untuk mengadopsi pendekatan berbasis analisis kuantitatif dalam pengambilan keputusan bisnis. Pada sektor Usaha Mikro, Kecil, dan Menengah (UMKM), khususnya di bidang kuliner, data penjualan umumnya telah tersedia dalam jumlah yang cukup, namun pemanfaatannya sebagai dasar perencanaan bisnis masih belum optimal. Padahal, data historis penjualan memiliki potensi besar untuk dianalisis guna memprediksi permintaan produk di masa mendatang, sehingga dapat

mendukung pengelolaan persediaan, penentuan strategi pemasaran, serta peningkatan efisiensi operasional.

Seiring dengan meningkatnya konsumsi produk bakery, pelaku usaha di sektor ini dihadapkan pada tantangan utama berupa ketidakpastian permintaan konsumen. Kesalahan dalam memperkirakan jumlah penjualan berpotensi menimbulkan permasalahan serius, seperti kelebihan stok yang menyebabkan kerugian akibat produk tidak terjual dan mengalami penurunan kualitas, serta kekurangan stok yang berdampak pada menurunnya tingkat kepuasan pelanggan dan hilangnya peluang keuntungan.

Toko Nursila Bakery sebagai salah satu UMKM yang bergerak di bidang produksi dan penjualan produk bakery menghadapi permasalahan serupa, khususnya dalam memprediksi permintaan kue kering pada periode hari raya. Selama ini, proses peramalan penjualan masih banyak bergantung pada intuisi dan pengalaman pemilik usaha tanpa didukung oleh analisis data yang sistematis. Pendekatan tersebut sering kali menimbulkan ketidaksesuaian antara jumlah produksi dan permintaan aktual pasar, yang pada akhirnya berdampak pada inefisiensi operasional serta penurunan tingkat profitabilitas usaha.

Permasalahan utama yang dihadapi Nursila Bakery dalam konteks penjualan musiman berkaitan dengan optimalisasi strategi produk untuk mencapai target penjualan yang maksimal. Penjualan kue kering dipengaruhi oleh berbagai faktor eksternal, seperti perubahan preferensi konsumen dari tahun ke tahun, serta karakteristik produk yang ditawarkan, termasuk jenis dan ukuran kemasan. Fokus penelitian diarahkan pada analisis pengaruh kombinasi variabel waktu dan atribut produk terhadap jumlah penjualan.

Berbagai penelitian terdahulu menunjukkan bahwa penerapan metode analisis data mampu meningkatkan akurasi pengambilan keputusan bisnis. Penelitian yang dilakukan oleh Abul Khair dkk. (2024) di lingkungan STMIK Widya Cipta Dharma, misalnya, membuktikan bahwa pemanfaatan algoritma data mining berperan penting dalam meningkatkan efisiensi dan ketepatan analisis dalam mendukung keputusan manajerial. Temuan tersebut mengindikasikan bahwa adopsi teknologi analitik, termasuk metode regresi, relevan dan aplikatif untuk diterapkan dalam penyelesaian permasalahan bisnis pada skala UMKM.

Regresi linear merupakan salah satu metode statistik yang efektif untuk memodelkan hubungan antara satu variabel dependen dengan variabel independen. Metode ini memungkinkan analisis pengaruh simultan terhadap jumlah penjualan, sehingga mampu menghasilkan model prediksi yang lebih akurat dan objektif. Melalui penerapan regresi linear, diharapkan Nursila Bakery dapat memperoleh informasi prediktif yang bermanfaat dalam menyusun perencanaan produksi, pengelolaan persediaan, serta perumusan strategi pemasaran yang lebih tepat sasaran.

Dengan demikian, penelitian ini diharapkan tidak hanya memberikan solusi praktis bagi permasalahan prediksi penjualan pada Toko Nursila Bakery, tetapi juga memberikan kontribusi akademis dalam memperkaya kajian mengenai pemanfaatan analisis data dan metode regresi linear untuk meningkatkan daya saing dan keberlanjutan UMKM di industri bakery Indonesia.

2. RUANG LINGKUP

2.1 Rumusan Masalah

Berdasarkan latar belakang diatas, maka rumusan masalah yang dikemukakan adalah “Bagaimana Mengimplementasikan Algoritma Regresi Linear dalam

memprediksi penjualan produk pada Toko Nursila Bakery”.

2.2 Batasan Masalah

Agar riset dapat dilaksanakan secara efisien dan memiliki fokus yang terarah, lingkup pembahasan perlu dibatasi, beberapa batasannya yaitu:

1. Penelitian ini hanya berfokus pada penerapan algoritma Regresi Linear untuk memprediksi penjualan produk di Nursila Bakery.
2. Regresi Linear yang digunakan adalah Regresi Linear Sederhana dan Regresi Linear Berganda.
3. Data yang dimanfaatkan dalam penelitian ini berasal dari catatan penjualan kue kering yang di jual di hari raya idul fitri, tidak mencakup semua data penjualan yang diambil melalui sistem pencatatan transaksi Nursila Bakery.
4. Sistem yang dikembangkan pada penelitian ini berbasis web.

2.3 Tujuan Penelitian

Penelitian ini bertujuan untuk membangun sistem prediksi penjualan menggunakan algoritma Regresi Linear pada Toko Nursila Bakery.

2.4 Manfaat Penelitian

Secara praktis, temuan dalam studi ini diharapkan dapat dimanfaatkan oleh berbagai pihak, mencakup:

1. Bagi Nursila Bakery

Dengan diterapkannya Regresi Linear ini diharapkan dapat membantu dalam pengambilan keputusan terkait strategi berdasarkan prediksi penjualan dan mengurangi risiko kelebihan atau kekurangan stok melalui prediksi permintaan yang akurat

2. Bagi Perguruan Tinggi

Hasil dari penulisan ini bagi perguruan tinggi berfungsi sebagai contoh atau referensi baru untuk kepentingan mahasiswa. Mereka dapat menggunakannya sebagai bahan penelitian tambahan saat mereka menyelesaikan tugas akhir atau skripsi mereka.

3. BAHAN DAN METODE

3.1 Implementasi

Menurut Nugroho (2020) Implementasi adalah tahapan lanjutan yang dilakukan setelah perancangan aplikasi, dan dapat diartikan sebagai upaya aktualisasi sistem yang telah didesain. Proses implementasi mencakup serangkaian langkah prosedural yang berurutan, mulai dari kegiatan inisiasi hingga tahap akhir, yang harus dieksekusi secara sistematis untuk mewujudkan sistem yang telah dirancang agar dapat berfungsi optimal.

3.2 Regresi Linear

Menurut Hasanah dkk. (2022) Metode regresi linear merupakan sebuah teknik statistika yang diimplementasikan untuk mengonstruksi model hubungan matematis antara satu atau lebih variabel bebas (independen; X) dengan satu variabel terikat (dependen; Y). Secara umum, analisis ini bertujuan untuk memodelkan bagaimana variabel prediktor memengaruhi variabel respons. Berdasarkan jumlah variabel

independen yang digunakan, analisis regresi diklasifikasikan menjadi dua jenis:

1. Regresi Linear Sederhana: Digunakan apabila hanya terdapat satu variabel independen yang memengaruhi variabel dependen.
2. Regresi Linear Berganda: Digunakan jika model melibatkan lebih dari satu variabel independen untuk memprediksi nilai variabel dependen.

Dalam pemodelan regresi, terdapat dua peran variabel yang sangat krusial. Variabel Independen (Variabel Bebas) merupakan variabel yang bertindak sebagai stimulus, penyebab, atau faktor risiko yang memengaruhi variabel lain. Variabel ini sering disebut sebagai prediktor atau perlakuan (*treatment*). Variabel Dependen (Variabel Terikat) merupakan variabel yang nilainya dipengaruhi atau menjadi akibat dari adanya variabel bebas. Dalam analisis regresi, variabel ini harus berskala data numerik (interval atau rasio) dan hanya berjumlah satu variabel.

3.3 Prediksi

Menurut Roza, dkk. (2020) Prediksi atau peramalan merupakan suatu proses ilmiah yang digunakan untuk memperkirakan nilai, kondisi, atau kejadian di masa depan dengan cara menganalisis data historis serta informasi yang tersedia saat ini. Dalam konteks sistem informasi, prediksi tidak dipandang sebagai sebuah tebakan semata, melainkan hasil dari pengolahan data masa lalu yang disusun secara sistematis menggunakan model matematis tertentu untuk memetakan kemungkinan-kemungkinan yang akan terjadi pada periode mendatang.

Tujuan utama dari penerapan fungsi prediksi ini adalah untuk meminimalkan tingkat ketidakpastian (*uncertainty*) yang sering dihadapi oleh para pengambil keputusan. Dengan adanya gambaran mengenai masa depan yang berbasis pada data, suatu organisasi atau perusahaan dapat melakukan perencanaan strategis yang lebih terukur dan antisipatif. Hal ini mencakup berbagai aspek manajerial, seperti pengelolaan sumber daya, penentuan target layanan, hingga mitigasi risiko guna menghindari kerugian yang mungkin timbul akibat perubahan situasi yang tidak terduga.

Dalam pelaksanaannya, efektivitas sebuah prediksi sangat bergantung pada kualitas data masa lalu dan ketepatan metode statistik yang digunakan, seperti metode Regresi Linear Berganda. Integrasi metode ini ke dalam sebuah sistem informasi berbasis web bertujuan untuk mengotomatisasi perhitungan matematis yang kompleks agar menjadi lebih cepat, akurat, dan minim kesalahan manusia (*human error*). Dengan demikian, sistem prediksi mampu mentransformasi kumpulan data mentah menjadi informasi strategis yang memudahkan manajemen dalam memantau tren dan menentukan arah kebijakan perusahaan secara lebih efektif.

3.4 Penjualan

Menurut Satria (2020) Penjualan merupakan proses persetujuan transaksional antara penjual dan pembeli, di mana terjadi pertukaran kepemilikan produk atau jasa

dari penjual kepada pembeli dengan imbalan sejumlah kompensasi finansial yang diserahkan oleh pembeli.

Mekanisme pembayaran dalam penjualan secara umum terbagi menjadi dua kategori utama, yaitu:

1. Penjualan Tunai (*Cash Sales*) Penjualan tunai didefinisikan sebagai transaksi langsung di mana pembeli melakukan pembayaran penuh atas harga barang sebelum perusahaan menyerahkan produk atau jasa kepada konsumen.

2. Penjualan Kredit (*Credit Sales*) Penjualan kredit adalah mekanisme pembayaran tertunda, di mana perusahaan menyerahkan barang terlebih dahulu, sementara pembeli hanya membayar sebagian kecil dari harga jual, dan sisanya akan dilunasi melalui pembayaran angsuran sesuai kesepakatan.

3.5 Produk

Menurut Budiono (2022) Produk dalam arti luas memiliki makna yang lebih kompleks, di antaranya meliputi produk fisik dan non-fisik, service atau layanan jasa, individu, organisasi, lokasi, ide, rights (hak paten). Adapun beberapa jenis produk, di antaranya adalah sebagai berikut :

1. *Consumers goods*, merupakan produk yang diterima oleh konsumen tertentu demi memenuhi kebutuhan penggunaan secara pribadi, dan pada akhirnya jenis produk ini akan hasil dalam penggunaan tersebut.
2. *Industrial goods*, merupakan produk yang diterima atau dibeli untuk kemudian diolah menjadi barang dalam wujud yang lain dan kemudian dijual kembali untuk memperoleh keuntungan industriasi.

3.6 Machine Learning

Menurut Roihan dkk. (2020) *Machine Learning* merupakan implementasi algoritma matematika yang bertujuan untuk mengenali pola data sebagai dasar prediksi masa depan melalui siklus pelatihan (*training*) dan pengujian (*testing*). Secara garis besar, bidang ini terbagi menjadi tiga kategori utama: *Supervised Learning* yang menitikberatkan pada klasifikasi menggunakan data berlabel, *Unsupervised Learning* yang menerapkan teknik pengelompokan (*clustering*) pada data tanpa label, serta Reinforcement Learning yang berfokus pada proses belajar mandiri dalam lingkungan dinamis tanpa umpan balik eksplisit. Ketiga paradigma ini merepresentasikan upaya sistem dalam memperoleh kecerdasan buatan melalui pemrosesan data yang sistematis.

3.7 Data Mining

Menurut Sudarsono dkk. (2021) Data mining didefinisikan sebagai proses penarikan informasi penting dari data. Informasi ini diekstrak melalui proses kompleks, sering kali melibatkan kecerdasan buatan, teknik statistik, matematika, dan pembelajaran mesin. Teknik-teknik kompleks ini dimanfaatkan untuk menemukan serta mengekstraksi informasi yang relevan dari kumpulan data berukuran besar. Data mining, yang berawal dari bidang pembelajaran mesin dan statistika,

kini telah berkembang menjadi cabang ilmu yang luas dalam ranah kecerdasan buatan dan rekayasa pengetahuan, serta mencakup berbagai disiplin seperti ilmu komputer, biologi, keuangan, dan bidang lainnya.

3.8 Variabel Regresi

Menurut Hayati & Andre Saputra (2023) variabel bebas (independen) didefinisikan sebagai elemen yang memberikan pengaruh atau menjadi determinan utama atas perubahan yang terjadi pada variabel lain. Sebaliknya variabel terikat (dependen) merupakan variabel yang posisinya dipengaruhi oleh keberadaan variabel bebas tersebut. Intinya, variabel independen adalah faktor penyebab yang diamati, sedangkan variabel dependen adalah dampak atau respons yang diukur dalam sebuah penelitian atau model analisis.

3.9 SPSS

Menurut Miftahul Janna & Herianto (2021) SPSS merupakan singkatan dari Statistical Package for the Social Sciences, yaitu sebuah perangkat lunak yang digunakan untuk mengolah dan menganalisis data statistik, baik secara interaktif maupun dalam bentuk pemrosesan batch. Dikenal karena tampilannya yang mudah dan cara kerjanya yang sederhana, SPSS digunakan untuk mengolah dan menganalisis data dengan kemampuan statistik dan manajemen data yang didukung antarmuka grafis. Awalnya digunakan di bidang ilmu sosial, kini SPSS telah diadopsi di berbagai bidang ilmu oleh peneliti pasar, bidang kesehatan, pemerintah, dan lainnya. Fitur dasarnya mencakup manajemen data (pemilihan kasus, pengolahan file, pembuatan data baru) serta dokumentasi data.

3.10 CRISP-DM

Menurut Lestiwati dkk. (2025) Cross-Industry Standard Process for Data Mining (CRISP-DM) adalah salah satu metodologi yang banyak digunakan dalam pengembangan perangkat lunak. Metodologi ini menyediakan pendekatan yang terstruktur dan sistematis untuk menjalankan proyek data mining, mulai dari pemahaman kebutuhan bisnis hingga penerapan solusi di lingkungan operasional.

Menurut konsepnya, data mining sendiri adalah proses pengumpulan dan pengolahan data dalam jumlah besar untuk mengekstrak informasi penting, menemukan pola tersembunyi, hubungan, serta wawasan yang dapat digunakan untuk mendukung pengambilan keputusan strategis. CRISP-DM berfungsi sebagai kerangka kerja yang fleksibel dan interaktif untuk mengubah data mentah menjadi wawasan berharga dan keunggulan kompetitif bagi organisasi. CRISP-DM terdiri dari enam tahapan utama yang bersifat iteratif dan dapat dilalui beberapa kali untuk perbaikan berkelanjutan.

1. Business Understanding (Pemahaman Bisnis)

Tahap ini merupakan fase krusial yang menekankan pentingnya memahami konteks dan tujuan bisnis secara mendalam. Kegiatan utama meliputi menentukan tujuan bisnis (misalnya, peningkatan penjualan, efisiensi operasional), menilai situasi ketersediaan sumber daya

yang ada, mendefinisikan tujuan pengumpulan data, dan menghasilkan rencana proyek secara komprehensif.

2. Data Understanding (Pemahaman Data)

Tahap ini berfokus pada pengidentifikasian, pengumpulan, dan analisis kumpulan data yang relevan untuk mencapai tujuan proyek. Aktivitas yang dilakukan meliputi pengumpulan data awal, deskripsi data (misalnya, melalui statistik ringkasan), eksplorasi data untuk mendapatkan wawasan awal, dan verifikasi kualitas data (misalnya, mengidentifikasi data yang hilang atau outlier).

3. Data Preparation (Persiapan Data)

Tahap ini sering disebut sebagai fase intensif yang tujuan utamanya adalah menyiapkan kumpulan data akhir yang siap untuk proses pemodelan. Kegiatan yang dilakukan adalah memperbaiki kualitas data agar sesuai dengan proses selanjutnya. Ini mencakup pembersihan data (*data cleaning*), penanganan nilai yang hilang, penghapusan duplikasi, penanganan outlier, transformasi data (misalnya, normalisasi), dan konstruksi fitur (*feature engineering*) dari berbagai sumber data.

4. Modeling (Pemodelan)

Pada fase ini, berbagai teknik pemodelan dibangun dan dinilai berdasarkan kumpulan data yang telah dipersiapkan. Kegiatan meliputi memilih teknik pemodelan yang sesuai (misalnya, regresi, clustering, atau klasifikasi), menghasilkan desain pengujian, membangun model menggunakan data training, dan menilai model tersebut dengan teknik validasi untuk menghindari overfitting.

5. Evaluation (Evaluasi)

Fase evaluasi melihat lebih luas apakah model yang dihasilkan sudah sesuai dengan tujuan bisnis dan memberikan wawasan yang bermanfaat serta dapat diterapkan. Kegiatan pada tahap ini mencakup evaluasi hasil model berdasarkan metrik yang relevan, peninjauan kembali seluruh proses yang telah dilakukan, dan penentuan langkah selanjutnya, termasuk melakukan iterasi ulang jika hasil model belum memuaskan.

6. Deployment (Penerapan)

Ini adalah tahap terakhir dan paling penting, di mana model yang sudah tervalidasi diimplementasikan dalam sistem operasional. Perencanaan untuk penerapan (deployment) dimulai sejak fase Pemahaman Bisnis. Aktivitasnya meliputi mengatur infrastruktur untuk penerapan, menghasilkan laporan dan visualisasi yang mudah dipahami, serta merancang strategi untuk pemantauan dan pemeliharaan model secara berkelanjutan.

4. PEMBAHASAN

4.1 Business Understanding

Nursila Bakery, yang telah beroperasi sejak tahun 2016, sering menghadapi tantangan dalam menyelaraskan kapasitas produksi dengan tingginya permintaan pasar. Permasalahan utama terletak pada kurangnya perencanaan yang terukur, sehingga potensi penjualan sering kali tidak terserap secara maksimal.

Tidak adanya sistem estimasi yang akurat menyebabkan pengambilan keputusan produksi masih bersifat konvensional. Hal ini sering kali mengakibatkan ketidaksiapan stok saat permintaan melonjak, yang pada akhirnya menghambat optimalisasi keuntungan toko.

Sebagai langkah strategis untuk mengatasi kendala tersebut, diperlukan implementasi sistem prediksi penjualan guna memperkuat perencanaan bisnis. Algoritma Regresi Linear merupakan metode yang sangat sesuai untuk kebutuhan ini karena mampu memproyeksikan volume penjualan di masa depan dengan mengolah variabel data historis.

4.2 Data Understanding

Tahap Data Understanding dilakukan untuk memberikan gambaran komprehensif mengenai dataset yang akan digunakan sebelum masuk ke proses pemodelan. Fokus utama pada tahap ini adalah memahami variabel-variabel yang tersedia serta melakukan pengecekan awal terhadap kualitas data.

4.2.1 Deskripsi Data

Data yang digunakan dalam penelitian ini berasal dari pencatatan penjualan historis Toko Nursila Bakery yang mencakup transaksi musiman pada periode hari raya Idulfitri dan Iduladha dalam rentang tahun 2021 hingga 2025. Berikut adalah data penjualan dari Toko Nursila Bakery.

Tabel 1. Data Penjualan

No	Year	Period	Cake	Size	Sold
1	2021	Fitri	Almond Cookies	Small	3
2	2021	Fitri	Almond Cookies	Medium	3
3	2021	Fitri	Beng-beng	Small	3
...
369	2025	Fitri	Brownies Cookies	Small	8
370	2025	Fitri	Brownies Cookies	Big	7
371	2025	Fitri	Choco Chips	Small	3

Tabel 1 menampilkan data penjualan dari 2021-2025 dengan jumlah data sebanyak 371 baris. Dataset tersebut terdiri dari beberapa variabel penting sebagai berikut:

1. Tahun: Keterangan tahun dalam catatan penjualan.
2. Periode: Keterangan Periode Idul Fitri dan Idul Adha.
3. Jenis Kue: Nama-nama jenis kue yang terjual.
4. Ukuran: Ukuran toples untuk kue yang di jual.
5. Jumlah Terjual: Jumlah Kue yang terjual (variabel target).

4.2.2 Eksplorasi Data

Dari data penjualan yang tercatat, dilakukan eksplorasi untuk mengetahui apa saja jenis kue apa saja yang selalu terjual dan bagaimana penjualannya dari tahun ke tahun. Hasil eksplorasi ini diharapkan dapat mengidentifikasi tren penjualan tiap jenis kue serta pola permintaan konsumen sepanjang tahun.

Terdapat 37 jenis kue yang tercatat di dalam data penjualan yaitu Almond Cookies, Beng-beng, Brownies Cookies, Choco Chips, Choco Mede, Choco Nut, Choco Stick, Coklat Filling, Crackel, Good Time, Kacang Skippy, Kanji Gula Merah, Kastangel, Keripik Keju, Lidah Kucing, London Almond, Marie Mente, Nastar, Nutella, Oatmeal, Palm Cheese, Putri Salju Pandan, Putri Salju Vanilla, Rambutan, Rempeyek, Sagu Gula Merah, Sagu Keju, Semprit Coklat, Semprit Susu, Semprit Vanilla, Sweet Orange, Thumbprint Blueberry, Thumbprint Campur, Thumbprint Coklat, Thumbprint Nut, Thumbprint Strawberry, dan Vinse. Dari 37 jenis kue yang tercatat, hanya ada 9 jenis kue yang selalu terjual dari tahun 2021-2025:

1. Almond Cookies,
2. Beng-beng,
3. Lidah Kucing,
4. Marie Mente,
5. Nastar,
6. Putri Salju Vanilla,
7. Rambutan,
8. Sagu Keju, dan
9. Semprit Susu.

Dari 9 jenis kue yang selalu terjual, akan dilakukan analisis terhadap penjualan dari tahun 2021-2025. terdapat beberapa jenis kue dan ukuran yang menampilkan tren positif dan linear, diantaranya adalah:

- 1) Almond Cookies (Besar),
- 2) Beng-beng (Kecil),
- 3) Lidah Kucing (Kecil),
- 4) Nastar (Kecil),
- 5) Nastar (Sedang),
- 6) Nastar (Besar),
- 7) Putri Salju Vanilla (Sedang),
- 8) Rambutan (Sedang),
- 9) Rambutan (Besar),
- 10) Sagu Keju (Besar), dan
- 11) Semprit Susu (Sedang).

4.2.3 Data Preparation

Tahap persiapan data (Data Preparation) merupakan fase yang paling intensif dalam penelitian ini. Tujuan utama dari fase ini adalah untuk menghasilkan kumpulan data akhir yang ideal bagi proses modeling dengan algoritma Regresi Linear. Berbagai langkah kunci diterapkan, meliputi pembersihan, transformasi, dan normalisasi data agar seluruh variabel memenuhi asumsi dasar regresi linear.

Adapun rincian langkah-langkah dalam tahap Data Preparation adalah sebagai berikut:

1. Eliminasi Variabel Produk Tidak Aktif: Dilakukan penghapusan terhadap jenis-jenis kue yang tidak mencatatkan aktivitas penjualan selama lima tahun berturut-turut. Langkah ini diperlukan untuk memastikan bahwa dataset hanya mengandung produk yang memiliki relevansi data berkelanjutan, sehingga meminimalkan potensi bias dari data yang stagnan..
2. Standardisasi Rentang Waktu (Time Series): Data penjualan pada periode Idul Adha dikeluarkan dari dataset penelitian. Hal ini disebabkan oleh keterbatasan ketersediaan data yang hanya tercatat pada rentang tahun 2021-2024 (4 tahun). Mengingat syarat minimal analisis time series dalam penelitian ini adalah 5 tahun, maka data tersebut dinilai tidak memenuhi kriteria konsistensi waktu yang ditetapkan.
3. Validasi Tren dan Linearitas: Data yang tidak menunjukkan tren positif atau pola linear pada grafik penjualan dieliminasi dari proses analisis. Penyaringan ini dilakukan untuk menjamin bahwa data yang diproses benar-benar memiliki distribusi yang sesuai dengan karakteristik algoritma Regresi Linear, sehingga tingkat akurasi prediksi dapat dipertanggungjawabkan.
4. Data Modeling: Data yang akan di gunakan dalam modeling adalah data-data yang memenuhi kriteria pada saat eksplorasi data dilakukan. Data tersebut dapat dilihat pada tabel 2 di bawah.

Tabel 2. Data Modeling

Cake	Size	Penjualan				
		1	2	3	4	5
Almond Cookies	Big	0	3	1	6	7
Beng-beng	Small	11	8	20	24	49
Lidah Kucing	Small	7	7	0	3	1
Nastar	Small	5	10	15	19	22
Nastar	Medium	10	17	43	44	48
Nastar	Big	3	4	25	44	47
Putri Salju Vanilla	Medium	2	12	19	44	45
Rambutan	Medium	7	7	17	34	40
Rambutan	Big	3	3	14	18	40
Sagu Keju	Big	2	4	12	12	13

Sempurit	Medium	6	7	7	18	19
Susu						

4.2.4 Modeling

Pada tahap ini, dilakukan modeling pada setiap jenis kue yang terjual 5 tahun terakhir yang menunjukkan tren positif atau linear menggunakan SPSS. Berikut adalah detailnya.

4.2.4.1 Regresi Linear Sederhana

1. Nastar (Kecil)

	Tahun	Penjualan
1	1	5
2	2	10
3	3	15
4	4	19
5	5	22

Gambar 1 Data Penjualan Nastar(Kecil)

Gambar 1 menampilkan rincian data penjualan Nastar (Kecil) yang digunakan dalam penelitian ini. Berikut detail outputnya:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.995 ^a	.990	.986	.796
a. Predictors: (Constant), Tahun				

Gambar 2 Model Summary Nastar(Kecil)

Gambar 2 menampilkan *Model Summary* dari kue Nastar dengan ukuran kecil. R = 0,995 menunjukkan hubungan yang sangat kuat dan positif antara Tahun dan Penjualan. R Square = 0,990 berarti 99,0% variasi naik-turunnya penjualan dipengaruhi oleh faktor waktu (tahun). Std. Error = 0,796 artinya rata-rata kesalahan prediksi dari model sebanyak 0,796 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	184.900	1	184.900	291.947	<.001 ^b
	Residual	1.900	3	.633		
	Total	186.800	4			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Tahun						

Gambar 3 ANOVA Nastar(Kecil)

Gambar 4.13 menampilkan *ANOVA* dari kue Nastar berukuran kecil. Sig. = 0,001 menunjukkan nilai yang lebih kecil dari 0,05 (0,001 < 0,05), maka model regresi ini signifikan secara statistik.

Hasil perhitungan Nastar (Kecil) adalah:

1. $\sum x = 15$
2. $\sum y = 71$
3. $\sum x^2 = 55$
4. $\sum xy = 256$

Rumus perhitungannya adalah: $y = 1,3 + 4,3(x)$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

$$y = 1,3 + 4,3(6) = 27,1$$

Jadi, prediksi penjualan untuk tahun ke-6 atau 2026 adalah 27 unit toples kecil.

2. Nastar (Sedang)

	Tahun	Penjualan
1	1	10
2	2	17
3	3	43
4	4	44
5	5	48

Gambar 4 Data Penjualan Nastar(Sedang)

Gambar 4 menampilkan rincian data penjualan Nastar (Sedang) yang digunakan dalam penelitian ini. Berikut detail outputnya:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.929 ^a	.863	.817	7.490
a. Predictors: (Constant), Tahun				

Gambar 5 Model Summary Nastar (Sedang)

Gambar 5 menampilkan *Model Summary* dari kue Nastar dengan ukuran sedang. R = 0,929 menunjukkan hubungan yang sangat kuat dan positif antara Tahun dan Penjualan. R Square = 0,863 berarti 86,3% variasi naik-turunnya penjualan dipengaruhi oleh faktor waktu (tahun). Std. Error = 7,490 artinya rata-rata kesalahan prediksi dari model sebanyak 7,490 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1060.900	1	1060.900	18.911	.022 ^b
	Residual	168.300	3	56.100		
	Total	1229.200	4			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Tahun						

Gambar 6 ANOVA Nastar(Sedang)

Gambar 6 menampilkan ANOVA dari kue Nastar berukuran sedang. Sig. = 0,022 menunjukkan nilai yang lebih kecil dari 0,05 (0,022 < 0,05), maka model regresi ini signifikan secara statistik.

Hasil perhitungan Nastar (Sedang) adalah:

1. $\sum x = 15$
2. $\sum y = 162$
3. $\sum x^2 = 55$
4. $\sum xy = 589$

Maka rumusnya adalah: $y = 1,5 + 10,3(x)$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

$$y = 1,5 + 10,3(6) = 63,3$$

Jadi, prediksi penjualan untuk tahun ke-6 atau 2026 adalah 63 unit toples sedang.

3. Nastar (Besar)

	Tahun	Penjualan
1	1	3
2	2	4
3	3	25
4	4	44
5	5	47

Gambar 7 Data Penjualan Nastar(Besar)

Gambar 7 menampilkan rincian data penjualan Nastar (Besar) yang digunakan dalam penelitian ini. Berikut detail outputnya:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.962 ^a	.926	.901	6.603
a. Predictors: (Constant), Tahun				

Gambar 8 Model Summary Nastar (Besar)

Gambar 8 menampilkan *Model Summary* dari kue Nastar dengan ukuran besar. R = 0,962 menunjukkan hubungan yang sangat kuat dan positif antara Tahun dan Penjualan. R Square = 0,926 berarti 92,6% variasi naik-turunnya penjualan dipengaruhi oleh faktor waktu (tahun). Std. Error = 6,603 artinya rata-rata kesalahan prediksi dari model sebanyak 6,603 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1638.400	1	1638.400	37.578	.009 ^b
	Residual	130.800	3	43.600		
	Total	1769.200	4			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Tahun						

Gambar 9 ANOVA Nastar (Besar)

Gambar 9 menampilkan *ANOVA* dari kue Nastar berukuran besar. Sig. = 0,009 menunjukkan nilai yang lebih kecil dari 0,05 (0,009 < 0,05), maka model regresi ini signifikan secara statistik.

Hasil perhitungan Nastar (Besar) adalah:

1. $\sum x = 15$
2. $\sum y = 123$
3. $\sum x^2 = 55$
4. $\sum xy = 497$

Maka rumusnya adalah: $y = -13,8 + 12,8(x)$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

$$y = -13,8 + 12,8(6) = 63$$

Jadi, prediksi penjualan untuk tahun ke-6 atau 2026 adalah 63 unit toples besar.

4. Rambutan (Sedang)

	Tahun	Penjualan
1	1	7
2	2	7
3	3	17
4	4	34
5	5	40

Gambar 10 Data Penjualan Rambutan (Sedang)

Gambar 10 menampilkan rincian data penjualan Rambutan (Sedang) yang digunakan dalam penelitian ini. Berikut detail outputnya:

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.960 ^a	.922	.896	4.936
a. Predictors: (Constant), Tahun				

Gambar 11 Model Summary Rambutan (Sedang)

Gambar 11 menampilkan Model Summary dari kue Rambutan dengan ukuran sedang. $R = 0,960$ menunjukkan hubungan yang sangat kuat dan positif antara Tahun dan Penjualan. $R^2 = 0,922$ berarti 92,2% variasi naik-turunnya penjualan dipengaruhi oleh faktor waktu (tahun). Std. Error = 4,936 artinya rata-rata kesalahan prediksi dari model sebanyak 4,936 unit.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	864.900	1	864.900	35.495	.009 ^b
	Residual	73.100	3	24.367		
	Total	938.000	4			

a. Dependent Variable: Penjualan
b. Predictors: (Constant), Tahun

Gambar 12 ANOVA Rambutan (Sedang)

Gambar 12 menampilkan ANOVA dari kue Rambutan berukuran sedang. Sig. = 0,009 menunjukkan nilai yang lebih kecil dari 0,05 ($0,009 < 0,05$), maka model regresi ini signifikan secara statistik.

Hasil perhitungan Putri Salju Vanilla (Sedang) adalah:

1. $\sum x = 15$
2. $\sum y = 105$
3. $\sum x^2 = 55$
4. $\sum xy = 408$

Maka rumusnya adalah: $y = -6,9 + 9,3(x)$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

$$y = -6,9 + 9,3(6) = 48,9$$

Jadi, prediksi penjualan untuk tahun ke-6 atau 2026 adalah 49 unit toples sedang.

5. Rambutan (Besar)

	Tahun	Penjualan
1	1	3
2	2	3
3	3	14
4	4	18
5	5	40

Gambar 13 Data Penjualan Rambutan (Besar)

Gambar 13 menampilkan rincian data penjualan Rambutan (Besar) yang digunakan dalam penelitian ini. Berikut detail outputnya:

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.927 ^a	.860	.813	6.560

a. Predictors: (Constant), Tahun

Gambar 14 Model Summary Rambutan (Besar)

Gambar 14 menampilkan Model Summary dari kue Rambutan dengan ukuran besar. $R = 0,927$ menunjukkan hubungan yang sangat kuat dan positif antara Tahun dan Penjualan. $R^2 = 0,860$ berarti 86,0% variasi naik-turunnya penjualan dipengaruhi oleh faktor waktu (tahun). Std. Error = 6,560 artinya rata-rata kesalahan prediksi dari model sebanyak 6,560 unit.

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	792.100	1	792.100	18.407	.023 ^b
	Residual	129.100	3	43.033		
	Total	921.200	4			

a. Dependent Variable: Penjualan
b. Predictors: (Constant), Tahun

Gambar 15 ANOVA Rambutan (Besar)

Gambar 15 menampilkan ANOVA dari kue Rambutan berukuran besar. Sig. = 0,023 menunjukkan nilai yang lebih kecil dari 0,05 ($0,023 < 0,05$), maka model regresi ini signifikan secara statistik.

Hasil perhitungan Rambutan (Besar) adalah:

1. $\sum x = 15$
2. $\sum y = 78$
3. $\sum x^2 = 55$
4. $\sum xy = 323$

Maka rumusnya adalah: $y = -11,1 + 8,9(x)$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

$$y = -11,1 + 8,9(6) = 42,3$$

Jadi, prediksi penjualan untuk tahun ke-6 atau 2026 adalah 42 unit toples besar.

4.2.4.2 Regresi Linear Berganda

1. Ukuran Kecil

Untuk memprediksi volume penjualan berdasarkan variabel Tahun dan jenis produk, digunakan model regresi linear berganda dengan variabel interaksi agar perbedaan tren penjualan antara produk Beng-beng (Kecil) dan Lidah Kucing (Kecil) dapat ditangkap secara spesifik.

Persamaan regresi yang digunakan adalah:

$$Y = a + b_1X_1 + b_2X_2 + b_3(X_1 \cdot X_2) + e$$

Keterangan Variabel

- 1) Y : Penjualan (Variabel Dependen)
- 2) X₁ : Tahun (Variabel Independen 1)
- 3) X₂ : Jenis Kue (Variabel dummy: 1 = Beng-beng, 0 = Lidah Kucing)
- 4) X₁X₂ : Variabel Interaksi antara Tahun dan Jenis Kue.

Hasil Perhitungannya:

- 1) $a = 8,4$
- 2) $b_1 = -1,6$
- 3) $b_2 = -13,6$
- 4) $b_3 = 10,8$

Berdasarkan hasil perhitungan, diperoleh model regresi akhir:

$$Y = 8,4 - 1,6X_1 - 13,6X_2 + 10,8(X_1X_2)$$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

- 1) Beng-beng ($X_2 = 1$): $Y = 8,4 - 1,6(6) - 13,6(1) + 10,8(6)$
 $Y = 8,4 - 9,6 - 13,6 + 64,8 = 50$ toples
- 2) Lidah Kucing ($X_2 = 0$): $Y = 8,4 - 1,6(6) = -1,2$ (dibulatkan menjadi 0 toples)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.942 ^a	.887	.830	6.116
a. Predictors: (Constant), Interaksi, Tahun, JenisKue				

Gambar 16 Model Summary Kue Ukuran Kecil

Gambar 16 menampilkan Model Summary dari jenis kue yang berukuran kecil (Beng-beng dan Lidah Kucing). $R = 0,942$ menunjukkan hubungan yang sangat kuat antar variabel independen. $R\text{ Square} = 0,887$ berarti variasi Penjualan dapat dijelaskan oleh variabel Tahun, Jenis Kue, dan Interaksi 88,7%. Std. Error = 6,116 artinya rata-rata kesalahan prediksi dari model sebanyak 6,116 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1755.600	3	585.200	15.647	.003 ^b
	Residual	224.400	6	37.400		
	Total	1980.000	9			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Interaksi, Tahun, JenisKue						

Gambar 17 ANOVA Kue Ukuran Kecil

Gambar 17 menampilkan ANOVA dari jenis kue yang berukuran kecil. Sig. = 0,003 menunjukkan nilai yang lebih kecil dari 0,05 ($0,003 < 0,05$), maka model regresi ini signifikan secara statistik.

2. Ukuran Sedang

Berdasarkan data penjualan, ditetapkan Semprit Susu sebagai ($X_2=0$) dan Putri Salju Vanilla sebagai ($X_2=1$).

Hasil akhir :

- 1) $a = 0,3$
- 2) $b_1 = 3,7$
- 3) $b_2 = -11,3$
- 4) $b_3 = 8,1$

Berdasarkan hasil perhitungan, diperoleh model regresi akhir:

$$Y = 0,3 - 3,7X_1 - 11,3X_2 + 8,1(X_1X_2)$$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

- 1) Putri Salju Vanilla ($X_2 = 1$): $Y = 0,3 - 3,7(6) - 11,3(1) + 8,1(6) = 0,3 - 22,2 - 11,3 + 48,6 = 59,8$ (dibulatkan menjadi 60 toples)
- 2) Semprit Susu ($X_2 = 0$): $Y = 0,3 + 3,7(6) = 22,5$ (dibulatkan menjadi 23 toples)

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.968 ^a	.936	.904	4.710
a. Predictors: (Constant), Interaksi, Tahun, JenisKue				

Gambar 18 Model Summary Kue Ukuran Sedang

Gambar 18 menampilkan Model Summary dari jenis kue yang berukuran sedang (Putri Salju Vanilla dan Semprit Susu). $R = 0,968$ menunjukkan hubungan yang sangat kuat antar variabel independen. $R\text{ Square} = 0,936$ berarti variasi Penjualan dapat dijelaskan oleh variabel Tahun, Jenis Kue, dan Interaksi 93,6%. Std. Error =

4,710 artinya rata-rata kesalahan prediksi dari model sebanyak 4,710 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	1951.800	3	650.600	29.328	<.001 ^b
	Residual	133.100	6	22.183		
	Total	2084.900	9			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Interaksi, Tahun, JenisKue						

Gambar 19 ANOVA Kue Ukuran Sedang

Gambar 19 menampilkan ANOVA dari jenis kue berukuran sedang. Sig. = 0,001 menunjukkan nilai yang lebih kecil dari 0,05 ($0,001 < 0,05$), maka model regresi ini signifikan secara statistik.

3. Ukuran Besar

Berdasarkan data penjualan, ditetapkan Sagu Keju sebagai ($X_2 = 0$) dan Almond Cookies sebagai ($X_2 = 1$).

Hasil akhir :

- 1) $a = -0,4$
- 2) $b_1 = 3$
- 3) $b_2 = -1,3$
- 4) $b_3 = -1,3$

Berdasarkan hasil perhitungan, diperoleh model regresi akhir:

$$Y = -0,4 - 3X_1 - 1,3X_2 - 1,3(X_1X_2)$$

Jika ingin memprediksi penjualan untuk tahun ke-6 (2026) maka perhitungannya:

- 1) Sagu Keju ($X_2 = 0$): $Y = -0,4 - 3(6) - 1,3(1) - 1,1(6) = -0,4 + 3(6) = 17,6$ (dibulatkan menjadi 18 toples).
- 2) Almond Cookies ($X_2 = 1$): $Y = -1,7 + 1,7(6) = 8,5$ (dibulatkan menjadi 9 toples).

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.938 ^a	.880	.820	2.062
a. Predictors: (Constant), Interaksi, Tahun, JenisKue				

Gambar 20 Model Summary Kue Ukuran Besar

Gambar 20 menampilkan Model Summary dari jenis kue yang berukuran besar (Sagu Keju dan Almond Cookies). $R = 0,938$ menunjukkan hubungan yang sangat kuat antar variabel independen. $R\text{ Square} = 0,880$ berarti variasi Penjualan dapat dijelaskan oleh variabel Tahun, Jenis Kue, dan Interaksi 88,0%. Std. Error = 2,062 artinya rata-rata kesalahan prediksi dari model sebanyak 2,062 unit.

ANOVA ^a						
Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	186.500	3	62.167	14.627	.004 ^b
	Residual	25.500	6	4.250		
	Total	212.000	9			
a. Dependent Variable: Penjualan						
b. Predictors: (Constant), Interaksi, Tahun, JenisKue						

Gambar 21 ANOVA Kue Ukuran Besar

Gambar 4.31 menampilkan ANOVA dari jenis kue berukuran besar. Sig. = 0,004 menunjukkan nilai yang

lebih kecil dari 0,05 ($0,004 < 0,05$), maka model regresi ini signifikan secara statistik.

4.2.5 Evaluasi Model

Fase evaluasi bertujuan untuk menilai sejauh mana model regresi yang telah dibangun memenuhi tujuan bisnis dan memiliki tingkat akurasi yang dapat dipertanggungjawabkan sebelum diimplementasikan secara luas. Berdasarkan hasil.

Pengujian yang dilakukan terhadap data penjualan Toko Nursila Bakery, berikut adalah hasil evaluasi mendalamnya:

Tabel 3 Evaluasi Model Regresi Linear Sederhana

Cake	Size	R	R ²	Std. E	Sig.
Nastar	Small	0,995	0,990	0,796	0,001
Nastar	Medium	0,929	0,863	7,490	0,022
Nastar	Big	0,962	0,926	6,603	0,009
Rambutan	Medium	0,960	0,922	4,936	0,009
Rambutan	Big	0,927	0,860	6,560	0,023

Berdasarkan Tabel 3, hasil evaluasi menunjukkan bahwa model regresi linear sederhana untuk kue Nastar dan Rambutan dengan berbagai macam ukuran memiliki tingkat akurasi yang sangat tinggi. Hal ini terlihat dari nilai korelasi (R) yang seluruhnya berada di atas 0,92 dan nilai koefisien determinasi (R²) yang kuat, di mana Nastar Kecil mencapai angka tertinggi sebesar 0,990. Selain itu, seluruh model dinyatakan valid secara statistik karena nilai signifikansinya berada di bawah 0,05, sehingga model ini sangat layak digunakan untuk memprediksi data penjualan.

Tabel 4 Evaluasi Model Regresi Linear Berganda

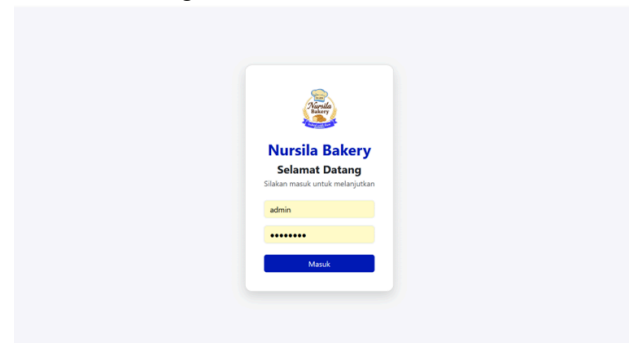
Cake	R	R ²	Std. E	Sig.
Beng-beng dan Lidah Kucing	0,942	0,887	6,116	0,003
Semprit Susu dan Putri Salju Vanilla	0,968	0,936	4,710	0,001
Almond Cookies dan Sagu Keju	0,938	0,880	2,062	0,004

Tabel 4 menyajikan hasil evaluasi untuk model regresi linear berganda pada kelompok kombinasi jenis kue. Secara keseluruhan, model ini menunjukkan performa yang sangat baik dengan nilai R² di atas 0,88, yang berarti variabel-variabel yang diuji mampu menjelaskan variasi data dengan sangat baik. Kombinasi Semprit Susu dan Putri Salju Vanilla memiliki akurasi tertinggi dengan nilai R² sebesar 0,936. Dengan tingkat signifikansi yang sangat rendah (jauh di bawah 0,05), model berganda ini terbukti akurat dan dapat dipertanggungjawabkan hasilnya.

4.2.6 Deployment

Tahap ini merupakan fase terakhir pada model regresi linear ini. Sistem ini dirancang untuk memudahkan pihak Toko Nursila Bakery dalam melakukan prediksi penjualan secara mandiri melalui antarmuka pengguna yang intuitif.

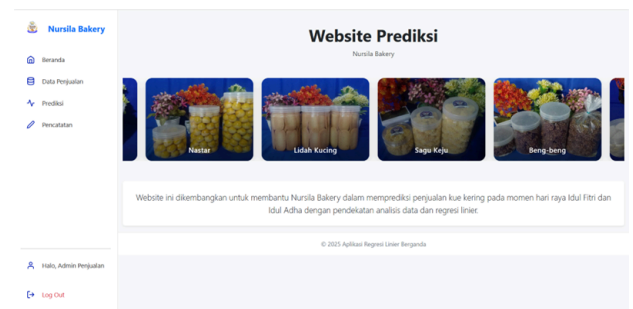
1. Halaman Login



Gambar 22 Halaman Login

Gambar 22 menampilkan halaman login dari website. Halaman ini merupakan pintu masuk utama untuk mengakses fungsi-fungsi internal sistem. Pengguna diwajibkan memasukkan username dan password yang valid untuk memastikan keamanan data penjualan dan model prediksi yang tersimpan di sistem.

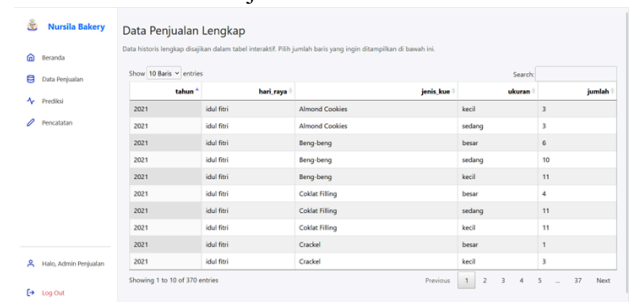
2. Halaman Utama



Gambar 23 Halaman Utama

Gambar 23 merupakan tampilan awal website. Setelah berhasil melakukan login, pengguna akan diarahkan ke halaman utama yang berfungsi sebagai dasbor navigasi. Halaman ini menyediakan menu navigasi di sisi kiri untuk mengakses menu lainnya. Selain sebagai pusat navigasi, halaman utama juga menampilkan ringkasan informasi penting seperti jenis kue unggulan.

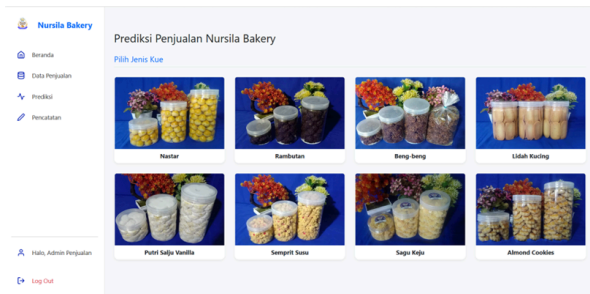
3. Halaman Data Penjualan



Gambar 24 Halaman Data Penjualan

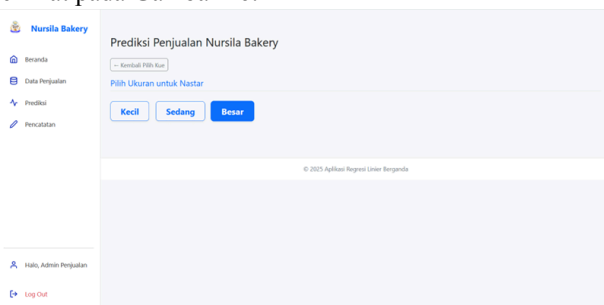
Gambar 24 menyajikan "Data Penjualan Lengkap" yang bersumber dari basis data historis Toko Nursila Bakery. Data ditampilkan dalam bentuk tabel yang mencakup kolom tahun, periode hari raya, jenis kue, ukuran, dan jumlah terjual. Fitur ini memudahkan admin untuk meninjau kembali tren penjualan dari tahun 2021 hingga 2025. Tabel ini memungkinkan pengguna untuk memantau performa penjualan tiap jenis kue secara detail.

4. Halaman Prediksi



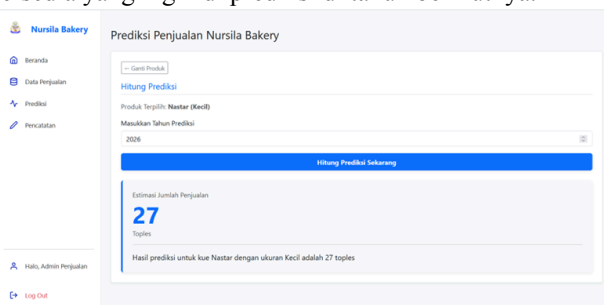
Gambar 25 Halaman Prediksi

Gambar 25 menampilkan antarmuka utama dari menu prediksi pada aplikasi Nursila Bakery. Halaman ini merupakan gerbang utama di mana implementasi algoritma Regresi Linear dimulai. Pengguna disuguhkan dengan katalog visual berupa kartu produk yang mencakup berbagai jenis kue, seperti Almond Cookies, Nastar, hingga Semprit Susu. Fungsi utama dari halaman ini adalah untuk memudahkan pengguna dalam memilih komoditas spesifik yang akan dianalisis tren penjualannya pada tahun berikutnya. Setelah pengguna memilih jenis kue tertentu, sistem akan mengarahkan ke halaman pemilihan parameter ukuran seperti yang terlihat pada Gambar 26.



Gambar 26 Halaman Pilih Ukuran

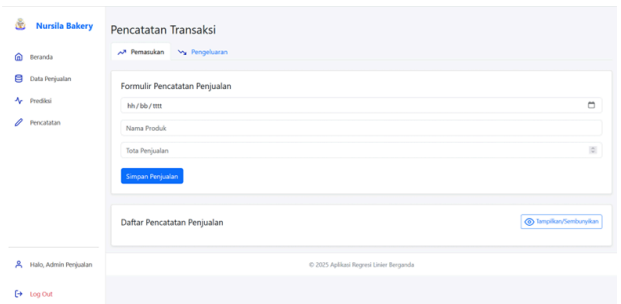
Gambar 26 menampilkan ukuran jenis kue yang tersedia yang ingin di prediksi di tahun berikutnya.



Gambar 27 Halaman Hasil Prediksi

Gambar 27 menyajikan antarmuka hasil akhir di mana sistem menampilkan estimasi angka penjualan berdasarkan input tahun dari pengguna. Setelah memasukkan tahun target dan menekan tombol hitung, sistem memproses data menggunakan algoritma Regresi Linier untuk memprediksi bahwa penjualan produk Nastar (Kecil) pada tahun 2026 adalah sebanyak 27 toples. Output ini memberikan gambaran prediktif yang memudahkan pihak manajemen dalam pengambilan keputusan terkait perencanaan stok dan strategi produksi pada periode mendatang secara lebih terukur.

5. Halaman Pencatatan Transaksi



Gambar 28 Halaman Pencatatan Transaksi

Gambar 4.38 menampilkan halaman yang digunakan untuk mencatat pemasukan dan pengeluaran dari Toko. Admin dapat mengisi formulir pencatatan penjualan yang mencakup tanggal transaksi, nama produk, dan jumlah penjualan. Admin juga dapat mencatat pengeluaran dari Toko mencakup tanggal pengeluaran, deskripsi pengeluaran (misalnya gaji karyawan), dan total pengeluaran.

5. KESIMPULAN

Berdasarkan penelitian yang telah dilakukan mengenai “bagaimana mengimplementasikan regresi linear dalam memprediksi penjualan produk pada toko Nursila Bakery”, maka dapat disimpulkan beberapa hal sebagai berikut:

1. Pengolahan Data Historis: Penelitian ini berhasil mengolah data transaksi penjualan musiman periode Idulfitri tahun 2021–2025 dengan fokus pada 11 varian produk yang menunjukkan tren linear positif. Data periode Iduladha dieliminasi karena belum memenuhi kriteria minimal time series selama 5 tahun.
2. Pemodelan Regresi: Model prediksi diimplementasikan melalui dua pendekatan utama, yaitu Regresi Linear Sederhana untuk menganalisis tren produk tunggal, serta Regresi Linear Berganda yang menggunakan variabel interaksi (Tahun, Jenis Kue, dan Interaksi) untuk membandingkan pola perkembangan penjualan antar produk dalam kategori ukuran yang sama.
3. Akurasi Model Regresi Sederhana: Hasil pengujian menggunakan SPSS menunjukkan tingkat akurasi yang sangat kuat. Nilai koefisien determinasi (R^2) tertinggi dicapai oleh produk Nastar Kecil sebesar

0,990, diikuti oleh Nastar Besar sebesar 0,926, dan Rambutan Sedang sebesar 0,922.

4. Akurasi Model Regresi Berganda: Model regresi berganda juga menunjukkan performa yang sangat baik, di mana kombinasi produk ukuran sedang (Semprit Susu dan Putri Salju Vanilla) menghasilkan nilai R^2 tertinggi sebesar 0,936.
5. Signifikansi Statistik: Seluruh model prediksi yang dibangun dinyatakan valid dan signifikan secara statistik karena memiliki nilai Sig. < 0,05, sehingga layak digunakan sebagai dasar pengambilan keputusan..
6. Sistem Prediksi Berbasis Web: Implementasi sistem berbasis web menggunakan bahasa pemrograman Python dan kerangka kerja CRISP-DM berhasil menyediakan solusi praktis bagi pemilik toko. Sistem ini tidak hanya mampu menghasilkan angka prediksi penjualan, tetapi juga memfasilitasi pencatatan transaksi pemasukan dan pengeluaran secara

6. SARAN

Untuk pengembangan sistem dan penelitian lebih lanjut di masa mendatang, penulis memberikan beberapa saran sebagai berikut:

1. Penambahan Rentang Waktu Data: Disarankan bagi pemilik toko untuk mengumpulkan data historis penjualan dalam jangka waktu yang lebih panjang (minimal 10 tahun) guna menangkap tren pasar jangka panjang secara lebih stabil.
2. Pengembangan Variabel Bebas: Peneliti selanjutnya dapat mengembangkan model ini dengan menambahkan variabel independen lain yang potensial memengaruhi penjualan, seperti harga bahan baku, daya beli konsumen, atau tren media sosial.
3. Pemeliharaan Model (*Retraining*): Mengingat dinamisnya data penjualan, model regresi dalam sistem perlu dilatih ulang (*retraining*) setiap kali terdapat data transaksi tahunan terbaru agar akurasi prediksi tetap relevan.
4. Integrasi Sistem: Disarankan untuk membangun integrasi antara sistem prediksi ini dengan platform penjualan daring (*e-commerce*) atau aplikasi kasir agar data transaksi dapat terekam secara otomatis ke dalam basis data.

7. REFERENSI

Abul Khair, A., Pratiwi, H., & Juli Saputra, N. (2024). PENERAPAN ALGORITMA K-NEAREST NEIGHBOR UNTUK KLASIFIKASI PENERIMA BEASISWA PADA STMIK WIDYA CIPTA DHARMA.

Anggrawan, A., Hairani, H., & Azmi, N. (2022). Prediksi Penjualan Produk Unilever Menggunakan Metode Regresi Linear. *Jurnal Bumigora Information Technology (BITe)*, 4(2), 123–132. <https://doi.org/10.30812/bite.v4i2.2416>

Budiono, I. N. (2022). *Soft File Buku MPBS* (A. Arifin, Ed.). IAIN Parepare Nusantara Press.

Dewi, S. P., Nurwati, N., & Rahayu, E. (2022). Penerapan Data Mining Untuk Prediksi Penjualan Produk Terlaris Menggunakan Metode K-Nearest Neighbor. *Building of Informatics, Technology and Science (BITS)*, 3(4), 639–648. <https://doi.org/10.47065/bits.v3i4.1408>

Handayani, M., Jayadilaga, Y., Fitri, A. U., Rachman, D. A., Fajriah Istiqamah, N., Diah, T., Pratiwi, A. P., & Kas, R. (2023). Sosialisasi dan Pengenalan Aplikasi Pengolahan Data SPSS pada Mahasiswa Socialization and Introduction of the SPSS Data Processing Application to Health Administration Students of the Faculty of Sports and Health Sciences. *JIPM : Jurnal Informasi Pengabdian Masyarakat*, 1(2), 24–32. <https://e-journal.nalanda.ac.id/index.php/jipm>

Hasanah, M., Harani, N. H., & Riza, N. (2022). Implementasi Barcode Dan Algoritma Regresi Linear Untuk Memprediksi Data Persediaan Barang (R. M. Awangga, Ed.). Kreatif Industri Nusantara.

Hasibuan, E., & Karim, A. (2022). Implementasi Machine Learning untuk Prediksi Harga Mobil Bekas dengan Algoritma Regresi Linear berbasis Web. *Jurnal Ilmiah Komputasi*, 21(4). <https://doi.org/10.32409/jikstik.21.4.3327>

Hayati, S., & Andre Saputra, L. (2023). Pengaruh Motivasi Kerja Terhadap Kinerja Karyawan Dengan Kepuasan Kerja Sebagai Variabel Intervening Pada Cv. Jaya Anugrah. *Business Management*, 2(1), 49–53. <https://ejournal.mandalanursa.org/index.php/Bisnis>

Kartiningrum, E. D., Notobroto, H. B., Otok, B. W., Kumarijati, E. N., & Yuswatiningsih, E. (2022). Aplikasi Regresi dan Korelasi Dalam Analisis Data Hasil Penelitian (M. F.-Klin. A. Dr. Rifaatul Laila Mahmudah, Ed.). STIKes Majapahit Mojokerto.

Kumalasari, I., Saputra, A. A., Pakpahan, A. G. S., Kurtubi, A., Amiruddin, A., Fridaniarta, B., Wicaksono, E. Y., Saputra, H., Putra, M. Y. A., & Azahra, R. Y. (2023). PELATIHAN DAN PEMBUATAN WEBSITE MENGGUNAKAN HTML DAN CSS. *Beujroh : Jurnal Pemberdayaan dan Pengabdian pada Masyarakat*, 1(1), 119–125. <https://doi.org/10.61579/beujroh.v1i1.41>

Lestawati, T. B., Ekawati, H., & Arriyanti, E. (2025). Modeling the Prediction of Palm Oil Prices Using Multiple Linear Regression in Rantau Pulung

-
- District, East Kutai Regency. Sebatik, 29(1).
<https://doi.org/10.46984/sebatik.v29i1.0000>
- Maisaroh, S., & Fiska, R. R. (2025). Implementasi Algoritma Regresi Linier Berganda untuk Prediksi Penjualan di D'Kopikap. *Jurnal Teknik Informatika dan Teknologi Informasi*, 5(1), 114–131.
<https://doi.org/10.55606/jutiti.v5i1.5149>
- Manao, I. K. (2024). PENGARUH KUALITAS PELAYANAN TERHADAP VOLUME PENJUALAN DI UD. ANIS BERKAT KECAMATAN TELUKDALAM. *Jurnal Ilmiah Mahasiswa Nias Selatan*, 7(2), 204–215.
- Miftahul Janna, N., & Herianto. (2021). KONSEP UJI VALIDITAS DAN RELIABILITAS DENGAN MENGGUNAKAN SPSS.
- Munada, M. S., & Kurniawan, R. (2024). PENGEMBANGAN APLIKASI SALES ORDER MENGGUNAKAN BAHASA PEMROGRAMAN PHP PADA PT. SINAR SOSRO (STUDI KASUS DI KANTOR PEMASARAN KUNINGAN). *JATI (Jurnal Mahasiswa Teknik Informatika)*, 8(2), 2236–2242.
- Muthohir, M. (2021). Mudah Membuat Web Bagi Pemula. Yayasan Prima Agus Teknik.
- Nabila, D., & Haryanto, E. V. (2024). Penerapan Metode Regresi Linear Berganda Dalam Data Mining Prediksi Penjualan Produk Pada PT Cipta Mebelindo Berbasis Web. Dalam *JID*.
<http://kti.potensi-utama.ac.id/index.php/JID>
- Nurhalizah, R. S., Ardianto, R., & Purwono, P. (2024). Analisis Supervised dan Unsupervised Learning pada Machine Learning: Systematic Literature Review. *Jurnal Ilmu Komputer dan Informatika*, 4(1), 61–72. <https://doi.org/10.54082/jiki.168>
- Panggabean, D. S. O., Buulolo, E., & Silalahi, N. (2020a). Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 56–62.
<https://doi.org/10.30865/jurikom.v7i1.1947>
- Panggabean, D. S. O., Buulolo, E., & Silalahi, N. (2020b). Penerapan Data Mining Untuk Memprediksi Pemesanan Bibit Pohon Dengan Regresi Linear Berganda. *JURIKOM (Jurnal Riset Komputer)*, 7(1), 56.
<https://doi.org/10.30865/jurikom.v7i1.1947>
- Rahman, S., Sembiring, A., Siregar, D., Khair, H., Prahmana, G., Puspadini, R., & Zen, M. (2023). PYTHON : DASAR DAN PEMROGRAMAN BERORIENTASI OBJEK TAHTA MEDIA GROUP (Tahta Media, Ed.). Tahta Media Group.
- Roihan, A., Abas Sunarya, P., & Rafika, A. S. (2020). Pemanfaatan Machine Learning dalam Berbagai Bidang: Review paper. Dalam *IJCIT (Indonesian Journal on Computer and Information Technology)* (Vol. 5, Nomor 1).
- Saputra, W., & Ramadhan, M. H. (2024). Metode Regresi Linear Berganda Dalam Prediksi Penjualan Produk Pada PT. Bumi Menara Internusa Multiple Linear Regression Method in Predicting Product Sales at PT. Bumi Menara Internusa. *Jurnal Rekayasa Sistem*, 2(2), 1842–1855.
<https://kti.potensi-utama.ac.id/index.php/>
- Satria, W. (2020). JARINGAN SYARAF TIRUAN BACKPROPAGATION UNTUK PERAMALAN PENJUALAN PRODUK (STUDI KASUS DI METRO ELECTRONIC DAN FURNITURE). *Djtechno : Journal of Information Technology Research*, 1(1), 14–19.
- Selay, A., Andgha, G. D., Alfari, M. A., Izdhihar, M., Wahyudi, B., Falah, M. N., Encep, M., & Khaira, M. (2023). SISTEM INFORMASI PENJUALAN. Dalam *Karimah Tauhid* (Vol. 2, Nomor 1).
- Sholikhan, M. (2022). HTML, CSS dan Javascript. Penerbit Yayasan Prima Agus Teknik.
- Sudarsono, B. G., Leo, M. I., Santoso, A., & Hendrawan, F. (2021). ANALISIS DATA MINING DATA NETFLIX MENGGUNAKAN APLIKASI RAPID MINER. *JBASE - Journal of Business and Audit Information Systems*, 4(1).
<https://doi.org/10.30813/jbase.v4i1.2729>
- Surya Ningsih, K., Jamilah Aruan, N., & Taufik Al Afkari Siahaan, A. (2022). Yayasan Insan Cipta Medan APLIKASI BUKU TAMU MENGGUNAKAN FITUR KAMERA DAN AJAX BERBASIS WEBSITE PADA KANTOR DISPORA KOTA MEDAN. *SITek*, 1(3), 94–99.
- Wahyudi, J., Asbari, M., Sasono, I., Pramono, T., & Novitasari, D. (2022). Database Management in MYSQL. *Edumaspul - Jurnal Pendidikan*, 6(2).