BAB III
METODE PENELITIAN

3.1 Objek Penelitian

Objek dalam penelitian ini adalah aktivitas penjualan dan data transaksi yang tercatat melalui sistem *Point of Sale* (DjePOS) pada Djejama Street Coffee. Sistem ini digunakan secara langsung dalam proses pelayanan pelanggan sehari-hari, sehingga setiap transaksi yang terjadi di kedai terekam secara terstruktur dan dapat dijadikan sumber data utama untuk analisis pola pembelian menggunakan algoritma Apriori. Dengan menjadikan DjePOS dan data transaksinya sebagai objek, penelitian ini berupaya menggali keterkaitan antarpoduk yang benar-benar muncul dalam praktik operasional, bukan sekadar skenario simulasi.

3.1.1 Peralatan dan Perangkat Lunak Penelitian

Pelaksanaan penelitian ini didukung oleh perangkat keras dan perangkat lunak yang digunakan sebagai media utama dalam proses pengolahan data transaksi serta penerapan algoritma Apriori. Spesifikasi perangkat yang digunakan adalah sebagai berikut:

1) Perangkat Keras
   a) Laptop
   b) CPU: AMD Ryzen 7 6800H with Radeon Graphics (3,20 GHz)
   c) RAM: 16 GB
   d) Arsitektur Sistem: 64-bit
2) Perangkat Lunak
   a) Sistem Operasi: Windows 11
   b) Bahasa Pemrograman: Python
   c) Library Utama:
      i. Pandas – digunakan untuk memuat, membersihkan, dan mengolah data transaksi dalam format CSV
      ii. Mlxtend – digunakan untuk proses pembentukan *frequent itemset* dan aturan asosiasi menggunakan algoritma Apriori

d) Aplikasi Pendukung:
   i. Aplikasi pengolah lembar kerja (spreadsheet) untuk eksplorasi awal dan verifikasi struktur data

3.1.2 Lokasi dan Waktu Penelitian

Penelitian ini dilaksanakan pada lingkungan operasional Djejama Street Coffee yang berlokasi di Jl. HS. Ronggo Waluyo No.187, Sukaharja, Telukjambe Timur, Karawang, Jawa Barat 41361. Waktu pelaksanaan penelitian direncanakan berlangsung pada rentang Oktober 2025 sampai dengan Juni 2026, mencakup kegiatan pengumpulan data transaksi, pengolahan dan analisis data menggunakan algoritma Apriori, hingga penyusunan laporan penelitian.

3.1.3 Sistem DjePOS

Pada awal berdirinya, pencatatan transaksi di Djejama Street Coffee dilakukan secara manual, sehingga pemilik usaha mengalami kesulitan dalam menelusuri riwayat penjualan, melakukan rekapitulasi pendapatan, dan memantau pergerakan produk. Untuk mengatasi permasalahan tersebut dikembangkan DjePOS, yaitu aplikasi *Point of Sale* yang digunakan kasir untuk mencatat transaksi penjualan, mengelola data produk beserta kategorinya, serta memasukkan informasi pengeluaran secara *realtime*. Dalam penelitian ini, DjePOS diposisikan sebagai objek teknis karena seluruh data yang dianalisis mulai dari identitas transaksi, daftar item yang dibeli, hingga kategori produk berasal langsung dari *log* yang dihasilkan sistem pada saat proses penjualan berlangsung.

3.1.4 Dataset dan Ruang Lingkup Penelitian

Data yang digunakan dalam penelitian ini merupakan data transaksi penjualan yang diekspor dari DjePOS dalam bentuk berkas CSV dengan struktur kolom sebagaimana ditunjukkan pada Tabel 3.1 Setiap baris data merepresentasikan kombinasi Order ID, tanggal transaksi, metode pembayaran, kategori produk, nama produk, jumlah yang terjual, harga satuan, dan total nilai penjualan, sehingga informasi yang tersedia cukup lengkap untuk dibentuk menjadi transaksi per nota dalam analisis Apriori. Pada tahap awal, data yang telah terkumpul mencakup periode sekitar dua bulan, yaitu Oktober hingga November, dan direncanakan untuk

diperluas menjadi kurang lebih lima bulan (hingga Februari 2026) agar pola pembelian yang diperoleh lebih stabil dan mencerminkan perilaku pelanggan yang berulang.

Tabel 3.1 Struktur Data Transaksi DjePOS

Order ID	Tanggal	Metode Pembayaran	Kategori	Produk	Qty	Harga
ID-001	13/10/2025	Qris	Coffee	Silky	2	Rp18.000
ID-004	13/10/2025	Qris	*Powdery*	Coklat	1	Rp20.000
ID-011	13/10/2025	Qris	Coffee	Silky	1	Rp18.000
ID-011	13/10/2025	Qris	Beverages	Sosis	1	Rp15.000
ID-011	13/10/2025	Qris	*Hot Menu*	*Latte Hot*	1	Rp15.000
ID-017	14/10/2025	Cash	Coffee	Silky	2	Rp18.000
ID-017	14/10/2025	Cash	Beverages	Kentang	1	Rp15.000
ID-020	14/10/2025	Cash	Coffee	Silky	1	Rp18.000
ID-020	14/10/2025	Cash	Beverages	Mie Rakyat	1	Rp13.000
ID-042	16/10/2025	Qris	*Powdery*	Matcha	1	Rp20.000

Ruang lingkup penelitian dibatasi hanya pada transaksi penjualan yang tercatat di DjePOS, sehingga jenis transaksi lain seperti retur, penyesuaian stok, maupun pencatatan pengeluaran tidak disertakan dalam proses penambangan pola. Fokus analisis diarahkan pada produk F&B yang dibeli oleh pelanggan, meliputi minuman kopi (misalnya *Silky, Latte, Legendary, Classic, Americano Ice*), minuman non-kopi dalam kategori *Powdery* (seperti Coklat dan Matcha), serta makanan pendamping dalam kategori *Beverages* dan *Hot Menu* (seperti Kentang, Mie Rakyat, Telur, Sosis, dan *Latte Hot*). Batasan ini ditetapkan agar aturan asosiasi yang dihasilkan benar-benar merepresentasikan hubungan antarpoduk dalam keranjang belanja pelanggan di Djejama Street Coffee dan dapat dimanfaatkan sebagai dasar perumusan paket *bundling*, rekomendasi menu, maupun pengelolaan stok di kedai tersebut.

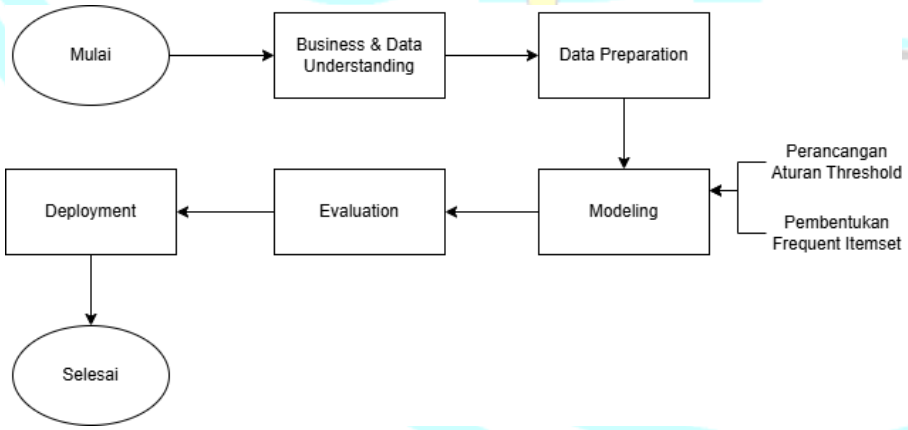
3.2 Prosedur Penelitian

Penelitian ini termasuk dalam kategori penelitian terapan *applied research* dengan pendekatan kuantitatif, karena memanfaatkan teknik *data mining* untuk menganalisis data transaksi penjualan yang terekam dalam bentuk numerik dan

kategorikal pada sistem DjePOS di Djejama Street Coffee. Metode yang digunakan berupa eksperimen terhadap data transaksi dengan menerapkan algoritma Apriori menggunakan beberapa skenario nilai *minimum support* dan *minimum confidence*, sehingga dapat dianalisis pengaruh pengaturan *threshold* terhadap jumlah dan kualitas aturan asosiasi yang dihasilkan.

Prosedur penelitian disusun secara sistematis dengan mengacu pada kerangka kerja CRISP-DM (*Cross Industry Standard Process for Data Mining*), yang meliputi tahapan *business understanding*, *data understanding*, *data preparation*, *modeling*, *evaluation*, dan *deployment*. Penggunaan framework ini bertujuan untuk memastikan bahwa proses penelitian dilakukan secara terstruktur dan sesuai dengan standar dalam pengembangan model *data mining*.

Selain itu, pada tahap *modeling* dilakukan pendekatan penentuan *threshold* berbasis aturan (*rule-based threshold*) yang memanfaatkan karakteristik statistik data transaksi, sehingga nilai *threshold* yang digunakan tidak ditentukan secara subjektif, melainkan berdasarkan hasil analisis data. Alur keseluruhan penelitian divisualisasikan dalam bentuk bagan alir pada Gambar 3.1 yang menunjukkan tahapan proses mulai dari pengumpulan data hingga evaluasi hasil.



Gambar 3.1 Prosedur Penelitian

A. *Business & Data Understanding*

Data penelitian diperoleh dari transaksi penjualan yang tercatat pada sistem DjePOS dalam bentuk berkas CSV yang memuat atribut Order ID, tanggal transaksi, metode pembayaran, kategori produk, nama produk, jumlah terjual, harga satuan, dan total nilai transaksi. Periode data direncanakan mencakup rentang

Oktober 2025 hingga Februari 2026 sehingga pola pembelian yang dianalisis merepresentasikan perilaku pelanggan yang berlangsung secara berulang dalam kondisi operasional normal di Djejama Street Coffee. Pada tahap ini juga dilakukan pemahaman awal terhadap konteks permasalahan, khususnya terkait kebutuhan analisis pola pembelian untuk mendukung penyusunan rekomendasi *bundling* dan pengelolaan stok.

B. *Data Preparation*

Tahap *data preparation* merupakan proses penting dalam penelitian ini karena berperan dalam memastikan bahwa data transaksi yang digunakan telah berada dalam kondisi bersih, konsisten, dan sesuai dengan kebutuhan algoritma Apriori. Data mentah hasil ekspor dari sistem DjePOS pada umumnya masih memiliki berbagai karakteristik yang perlu disesuaikan, seperti perbedaan format penulisan, keberadaan data yang tidak valid, serta representasi transaksi yang belum mencerminkan struktur pembelian secara utuh.

Oleh karena itu, dilakukan serangkaian proses pengolahan data untuk mentransformasikan data mentah menjadi bentuk *basket transaction*, yaitu representasi transaksi yang menunjukkan kumpulan item yang dibeli secara bersamaan dalam satu transaksi. Proses yang dilakukan dalam tahap ini meliputi:

1) Pembersihan Data

   Tahap ini dilakukan untuk menghapus data yang tidak valid serta memastikan tidak terdapat nilai kosong pada atribut utama seperti *orderKey* dan *productName*, sehingga tidak mengganggu proses analisis selanjutnya.

2) Normalisasi nama produk

   Dilakukan penyeragaman penulisan nama produk untuk menghindari perbedaan representasi item yang sama, misalnya akibat penggunaan huruf besar dan kecil atau variasi penamaan, sehingga setiap produk memiliki identitas yang konsisten.

3) Pemisahan produk kombinasi

   Pada beberapa data ditemukan satu entri yang merepresentasikan lebih dari satu produk. Oleh karena itu, dilakukan pemisahan sehingga setiap item dapat diperlakukan sebagai entitas yang berdiri sendiri dalam proses analisis.

4) Pengelompokan transaksi

Data dikelompokkan berdasarkan *orderKey* untuk membentuk satu transaksi utuh yang merepresentasikan kumpulan produk yang dibeli secara bersamaan oleh pelanggan.

5) Penghapusan duplikasi item

Item yang muncul lebih dari satu kali dalam satu transaksi dihilangkan duplikasinya untuk menjaga keakuratan perhitungan nilai *support*.

6) Penyaringan transaksi

Hanya transaksi yang memiliki lebih dari satu item yang dipertahankan, karena transaksi tunggal tidak dapat menghasilkan hubungan asosiasi antar item.

Melalui rangkaian proses tersebut, dihasilkan dataset dalam bentuk *basket transaction* yang telah siap digunakan dalam penerapan algoritma Apriori. Kualitas hasil pada tahap ini sangat berpengaruh terhadap pola asosiasi yang akan dihasilkan, sehingga proses *data preparation* dilakukan secara cermat untuk memastikan representasi data yang akurat dan relevan.

C. *Modeling*

Tahap *modeling* merupakan inti dari penelitian ini, yang bertujuan untuk menemukan pola asosiasi antar produk berdasarkan data transaksi yang telah dipersiapkan. Pada tahap ini digunakan algoritma Apriori untuk membentuk *frequent itemset* serta menghasilkan aturan asosiasi *association rules* yang menggambarkan keterkaitan antar item dalam transaksi pelanggan.

Sebelum proses pembentukan aturan dilakukan, terlebih dahulu ditentukan nilai *minimum support* dan *minimum confidence* yang akan digunakan sebagai parameter dalam algoritma. Berbeda dengan pendekatan umum yang menetapkan nilai *threshold* secara tetap, penelitian ini menggunakan pendekatan *rule-based threshold* yang mempertimbangkan karakteristik distribusi data transaksi. Parameter seperti nilai rata-rata (*mean*), *median*, dan kuartil digunakan sebagai dasar dalam menentukan nilai *threshold*, sehingga lebih adaptif dan tidak bergantung pada metode *trial-and-error*.

Setelah nilai *threshold* ditentukan, algoritma Apriori diterapkan untuk membentuk pola asosiasi secara bertahap. Proses ini dimulai dari pembentukan *frequent 1-itemset*, kemudian dilanjutkan dengan pembentukan kombinasi item yang lebih besar hingga tidak ditemukan lagi kombinasi yang memenuhi nilai *minimum support*. Dari *frequent itemset* yang terbentuk, selanjutnya dihasilkan aturan asosiasi yang dievaluasi menggunakan nilai *confidence* dan *lift*

Secara umum, tahapan penerapan algoritma Apriori dalam penelitian ini dapat dijelaskan melalui pseudocode berikut:

1) Hitung nilai *support* setiap item dalam dataset transaksi
2) Bentuk *frequent 1-itemset* dengan memilih item yang memiliki *support* ≥ *minimum support*
3) Untuk k = 2 hingga tidak ada kandidat:
    a) Bentuk kandidat k-*itemset* dari kombinasi *frequent (k-1)-itemset*
    b) Hitung nilai *support* setiap kandidat
    c) Pilih kandidat yang memenuhi *support* ≥ *minimum support* sebagai *frequent itemset*
4) Bentuk aturan asosiasi dari setiap *frequent itemset*:
    a) Tentukan *antecedent* dan *consequent*
    b) Hitung nilai *confidence* untuk setiap aturan
    c) Simpan aturan dengan *confidence* ≥ *minimum confidence*
5) Hitung nilai *lift* untuk mengukur kekuatan hubungan antar item

Melalui tahapan tersebut, dihasilkan aturan asosiasi yang tidak hanya memenuhi kriteria *minimum support* dan *minimum confidence*, tetapi juga memiliki kekuatan hubungan yang signifikan. Hasil dari tahap *modeling* ini kemudian digunakan sebagai dasar dalam proses evaluasi untuk menilai kualitas aturan yang dihasilkan serta pengaruh variasi *threshold* terhadap pola asosiasi yang terbentuk.

D. *Evaluation*

Tahap evaluasi dilakukan untuk menilai kualitas aturan asosiasi yang dihasilkan dari algoritma Apriori serta melihat pengaruh variasi nilai *threshold* terhadap hasil yang terbentuk. Evaluasi dilakukan menggunakan metrik *support*, *confidence*, dan

lift untuk mengetahui frekuensi kemunculan pola, tingkat kepercayaan hubungan antaritem, serta kekuatan asosiasi yang dihasilkan.

Selain itu, dilakukan pengujian beberapa skenario kombinasi nilai *minimum support* dan *minimum confidence*, termasuk nilai *threshold* yang diperoleh melalui pendekatan *rule-based*. Setiap skenario diterapkan pada dataset yang sama agar hasilnya dapat dibandingkan secara objektif. Hasil evaluasi kemudian dianalisis berdasarkan jumlah aturan, kualitas metrik, cakupan pola, dan relevansi aturan terhadap kebutuhan operasional DjePOS, sehingga dapat diketahui konfigurasi threshold yang paling seimbang antara jumlah dan kualitas aturan asosiasi.

E. *Deployment*

Tahap *deployment* dalam penelitian ini dilakukan dengan mengarahkan hasil analisis ke dalam bentuk rancangan pemanfaatan pada sistem operasional, khususnya pada sistem *Point of Sale* (POS) DjePOS yang digunakan di Djejama Street Coffee. Tahap ini bertujuan untuk memastikan bahwa hasil penambangan pola asosiasi tidak hanya berhenti pada tahap analisis, tetapi dapat dimanfaatkan sebagai dasar dalam mendukung pengambilan keputusan.

Rancangan pemanfaatan tersebut berupa konsep fitur *dashboard* analitik pada sistem DjePOS yang menyajikan hasil aturan asosiasi dalam bentuk *top rules*. Aturan yang ditampilkan merupakan hasil seleksi berdasarkan nilai *support*, *confidence*, dan *lift* tertinggi yang telah memenuhi *threshold* yang ditentukan menggunakan pendekatan *rule-based*. Penyajian dalam bentuk ini bertujuan untuk mempermudah pengguna dalam mengidentifikasi pola pembelian yang paling signifikan tanpa harus menelusuri seluruh aturan yang dihasilkan.

Selain itu, *dashboard* juga dirancang untuk menyajikan informasi terkait parameter yang digunakan dalam proses analisis, seperti nilai *minimum support* dan *minimum confidence* yang diperoleh dari perhitungan berbasis karakteristik data. Dengan demikian, pengguna tidak hanya memperoleh hasil akhir berupa aturan asosiasi, tetapi juga dapat memahami dasar penentuan parameter yang digunakan.