

BAB III

METODE PENELITIAN

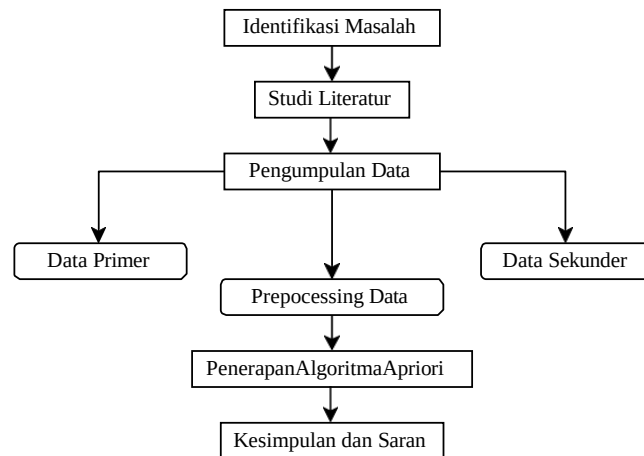
Metode penelitian yang digunakan adalah metode kuantitatif, karena analisis dilakukan berdasarkan data numerik berupa transaksi dan frekuensi kunjungan. Pendekatan data mining diterapkan untuk menemukan pola keterkaitan antarproduk yang sering muncul bersamaan dalam transaksi, dengan menggunakan teknik *association rule mining* melalui algoritma Apriori. Pola kunjungan dianalisis menggunakan statistik deskriptif melalui pengelompokan data transaksi berdasarkan jam dan hari, sedangkan preferensi menu dianalisis menggunakan algoritma Apriori untuk menemukan keterkaitan antar menu berdasarkan nilai *support*, *confidence*, dan *lift*. Algoritma ini digunakan untuk mengidentifikasi *frequent itemset* dan membentuk aturan asosiasi yang menggambarkan hubungan antar item secara sistematis.

Sebelum dianalisis, data melalui proses pengumpulan, pembersihan, dan penyesuaian format. Analisis dilakukan dengan menentukan nilai minimum *support* dan *confidence* untuk menemukan pola pembelian dan kecenderungan. Hasil aturan asosiasi memberikan gambaran mengenai menu yang sering dibeli bersama, pola kunjungan pada waktu tertentu, dan kombinasi produk yang diminati. Secara keseluruhan, metode ini membantu memahami kebiasaan pelanggan dan menjadi dasar bagi penataan menu, penyusunan paket, serta strategi promosi.

3.1 Kerangka Penelitian

Kerangka penelitian menggambarkan alur langkah yang digunakan dalam penelitian ini, mulai dari pengumpulan data hingga penarikan kesimpulan.

Kerangka ini menjelaskan tahapan analisis pola kunjungan serta penerapan algoritma Apriori untuk mengetahui preferensi menu di Waroeng *Fake*. Adapun kerangka penelitian yang digunakan adalah sebagai berikut:



Gambar 3.1 Kerangka Penelitian

Gambar 3.1 ditunjukkan tahapan alur kerja dari penelitian ini. Proses penelitian diawali dengan tahap identifikasi masalah, yaitu mengamati permasalahan yang terjadi di Waroeng *Fake* Tembilahan terkait pengelolaan data penjualan yang belum dimanfaatkan untuk mengetahui pola pembelian konsumen. Tahap selanjutnya adalah studi literatur yang bertujuan untuk mengumpulkan teori-teori yang berkaitan dengan data mining, algoritma Apriori, serta penelitian terdahulu yang relevan.

Setelah itu, dilakukan pengumpulan data yang terdiri dari data primer berupa observasi, wawancara, dan dokumentasi data penjualan, serta data sekunder berupa artikel teori dan hasil penelitian sebelumnya. Data yang telah terkumpul kemudian melalui tahap *preprocessing* untuk membersihkan, menyeleksi, dan menyesuaikan format data agar siap digunakan dalam proses analisis.

Tahap berikutnya adalah penerapan algoritma Apriori untuk menemukan pola asosiasi atau keterkaitan antar produk yang sering dibeli secara bersamaan oleh konsumen. Hasil dari proses analisis tersebut kemudian dievaluasi untuk menilai kesesuaian dan manfaat pola yang dihasilkan. Tahapan penelitian diakhiri dengan penarikan kesimpulan dan pemberian saran berdasarkan hasil evaluasi yang diperoleh.

3.2 Teknik Pengumpulan Data

Teknik pengumpulan data dilakukan untuk memperoleh data kunjungan dan preferensi pemilihan menu yang akan digunakan dalam proses analisis serta penerapan algoritma Apriori. Adapun teknik yang digunakan dalam penelitian ini terdiri dari data primer dan data sekunder

3.2.1 Data Primer

Data primer adalah data yang diperoleh secara langsung dari sumber asli untuk keperluan penelitian ini. Data primer dalam penelitian ini diperoleh melalui:

1. Observasi: Peneliti mengamati secara langsung aktivitas yang berlangsung di Waroeng *Fake* guna mengetahui kebiasaan pelanggan dan situasi lapangan secara nyata.
2. Wawancara: Wawancara dilakukan dengan pemilik Waroeng *Fake* untuk memperoleh informasi yang lebih mendalam mengenai kendala dalam pengelolaan data transaksi, pola kunjungan, serta pemahaman pemilik terkait menu-menu yang paling diminati. Wawancara ini juga bertujuan untuk mengidentifikasi kebutuhan pemilik usaha terhadap analisis yang mampu membantu menemukan keterkaitan antarmenu secara otomatis dan akurat melalui penerapan algoritma Apriori.

3. Dokumentasi/Pengumpulan Data: Data diperoleh dari catatan transaksi penjualan yang kemudian diubah ke dalam format digital (*Excel*). Data yang dikumpulkan meliputi daftar menu yang dibeli, jumlah transaksi, waktu kunjungan, serta kombinasi item yang muncul dalam satu pembelian. Data yang digunakan dalam penelitian ini merupakan data transaksi yang tercatat selama periode November 2025 hingga Mei 2026. Total transaksi yang memenuhi kelengkapan atribut dan digunakan dalam analisis sebanyak 1.143 transaksi. Seluruh data transaksi yang telah dikumpulkan selanjutnya digunakan sebagai dasar dalam tahap *preprocessing*, yang meliputi proses pembersihan, seleksi, dan transformasi data, sebelum dianalisis menggunakan algoritma Apriori untuk menemukan pola asosiasi.

1.2.2 Data Sekunder

Data sekunder adalah data yang diperoleh dari sumber yang sudah ada. Data sekunder dalam penelitian ini seperti artikel teori dan penelitian terdahulu yang relevan dengan penerapan apriori dalam analisis pola kunjungan dan preferensi menu.

3.3 Variabel dan Indikator Penelitian

Variabel penelitian merupakan segala sesuatu yang menjadi objek pengamatan dalam penelitian ini. Berdasarkan judul dan rumusan masalah yang telah ditetapkan, penelitian ini memiliki dua variabel utama yaitu daripola kunjungan aktivitas dan preferensi menu pelanggan. Masing-masing variabel dijabarkan ke dalam indikator yang dapat diukur secara langsung dari data transaksi sebagaimana disajikan pada tabel berikut.

Tabel 3.1 Variabel dan Indikator Penelitian

variabel	Indikator	Sumber Data	Alat Ukur	Hasil yang Diharapkan
Pola Kunjungan Pelanggan	Frekuensi transaksi per jam	Data transaksi	Perhitungan dan Pengelompokan Frekuensi Transaksi	Distribusi transaksi per jam (<i>peak/off-peak</i>)
Pola Kunjungan Pelanggan	Frekuensi transaksi per hari	Data transaksi	Perhitungan dan Pengelompokan Frekuensi Transaksi	Distribusi transaksi per hari (<i>weekday/weekend</i>)
Preferensi Menu Pelanggan	Nilai <i>support</i> item menu	Data transaksi	Algoritma Apriori	Item/kombinasi menu yang sering dibeli (<i>frequent itemset</i>)
Preferensi Menu Pelanggan	Nilai <i>confidence</i> aturan asosiasi	Data transaksi	Algoritma Apriori	Kecenderungan pembelian menu secara bersamaan
Preferensi Menu Pelanggan	Nilai <i>lift ratio</i>	Data transaksi	Algoritma Apriori	Kekuatan hubungan antar menu dalam transaksi

Variabel pola kunjungan pelanggan diukur berdasarkan dua indikator, yaitu frekuensi transaksi per jam dan per hari. Kedua indikator ini digunakan untuk mengidentifikasi waktu kunjungan yang paling ramai dan paling sepi selama periode penelitian.

Variabel preferensi menu pelanggan diukur menggunakan algoritma Apriori dengan tiga ukuran evaluasi utama, yaitu *support* untuk mengetahui tingkat kemunculan item menu, *confidence* untuk mengukur kecenderungan pembelian menu secara bersamaan, dan *lift ratio* untuk menilai kekuatan hubungan antar menu dalam transaksi.

3.4 Preprocessing Data

Tahap *preprocessing* merupakan tahap penting karena memastikan data yang akan dianalisis berada dalam kondisi yang bersih, rapi, dan layak olah. Proses yang dilakukan meliputi:

1. *Data Cleaning*: Menghapus transaksi yang tidak lengkap (misalnya tanggal atau daftar menu kosong), menghilangkan data duplikat, menghapus transaksi *void*, pembatalan, atau koreksi kasir, menyamakan format penulisan menu (contoh: “Kopi Susu” vs “kopi susu”), memastikan format tanggal dan jam sudah konsisten
2. *Data Selection*: Memilih atribut yang relevan seperti tanggal transaksi untuk pola kunjungan, jam transaksi untuk menentukan pola ramai atau sepi, dan daftar menu yang dibeli dalam satu kali transaksi digunakan untuk apriori.
3. *Data Transformation*: Mengubah format data mentah menjadi data yang lebih terstruktur agar dapat diolah oleh algoritma.
4. Pembentukan Dataset Apriori: Setelah data dibersihkan dan ditransformasikan, dataset khusus Apriori disiapkan. Dataset ini biasanya dalam format: CSV (*Comma Separated Values*), Setiap baris berisi satu transaksi lengkap, tidak menggunakan angka penjualan, tetapi daftar item.

Setelah tahap *preprocessing* data selesai, proses selanjutnya adalah penerapan algoritma Apriori untuk menganalisis pola keterkaitan antar menu berdasarkan data transaksi penjualan. Algoritma Apriori digunakan untuk menemukan *frequent itemset*, yaitu kombinasi menu yang sering muncul secara bersamaan dalam satu transaksi, yang kemudian membentuk aturan asosiasi (*association rules*).

Proses penerapan algoritma Apriori diawali dengan penentuan nilai *minimum support* dan *minimum confidence* sebagai ambang batas dalam pemilihan itemset yang dianggap signifikan. Nilai *minimum support* yang digunakan dalam penelitian ini ditetapkan sebesar 1% dan *minimum confidence* sebesar 20%. Penetapan nilai tersebut didasarkan pada pertimbangan jumlah transaksi dan variasi menu yang tersedia, sehingga diharapkan dapat menghasilkan *frequent itemset* dan aturan asosiasi yang cukup representatif namun tetap bermakna. Nilai *support* digunakan untuk mengukur tingkat kemunculan suatu item atau kombinasi item dalam keseluruhan transaksi, yang dirumuskan sebagai berikut:

Perhitungan *Support*, *Confidence*, dan *Lift*:

a. Perhitungan *Support*

Support digunakan untuk mengukur seberapa sering suatu item atau kombinasi item muncul dalam keseluruhan transaksi. Nilai *support* menjadi dasar dalam menentukan apakah suatu itemset memenuhi nilai *minimum support* sehingga dapat diproses pada tahap berikutnya. Rumus perhitungan *support* untuk satu item adalah sebagai berikut:

$$Support(A) = \frac{Frekuensi(A)}{Total Transaksi} \times 100\% \quad (1)$$

Rumus (1) digunakan untuk menghitung nilai *support* suatu item, yaitu persentase kemunculan item terhadap seluruh transaksi yang dianalisis. Semakin besar nilai *support*, semakin sering item tersebut muncul dalam data transaksi.

Keterangan:

- a) *Support(A)* : Nilai persentase kemunculan item A.
- b) *Frekuensi(A)* : Jumlah transaksi yang mengandung item A.

c) Total Transaksi : Jumlah seluruh transaksi yang dianalisis.

b. Perhitungan Support untuk dua item (A dan B)

Untuk mengetahui frekuensi kemunculan dua item secara bersamaan, digunakan rumus support dua item sebagai berikut:

$$Support(A \cap B) = \frac{Frekuensi(A \cap B)}{Total Transaksi} \times 100\% \quad (2)$$

Rumus (2) digunakan untuk menghitung nilai *support* dari kombinasi dua item, yaitu persentase transaksi yang memuat item A dan item B secara bersamaan terhadap seluruh transaksi. Keterangan:

- a) $Support(A \cap B)$: Nilai persentase kemunculan item A dan B secara bersamaan.
- b) $Frekuensi(A \cap B)$: Jumlah transaksi yang mengandung item A dan item B.
- c) Total Transaksi : Jumlah seluruh transaksi yang dianalisis.

c. Perhitungan Confidence

Confidence digunakan untuk mengukur tingkat kepercayaan suatu aturan asosiasi, yaitu kemungkinan item B dibeli ketika pelanggan membeli item A. Rumus *confidence* adalah sebagai berikut:

$$Confidence(A \rightarrow B) = \frac{Frekuensi(A \cap B)}{Frekuensi(A)} \times 100\% \quad (3)$$

Rumus (3) digunakan untuk menghitung nilai *confidence*, yaitu persentase peluang munculnya item B setelah item A muncul dalam suatu transaksi. Semakin tinggi nilai *confidence*, semakin kuat hubungan antara kedua item. Keterangan:

- a) $Confidence(A \rightarrow B)$: Tingkat kepercayaan aturan asosiasi dari item A ke item B

- b) Frekuensi($A \cap B$) : Jumlah transaksi yang mengandung item A dan item B.
- c) Frekuensi(A) : Jumlah transaksi yang mengandung item A.

d. Perhitungan Lift

Lift digunakan untuk mengukur kekuatan hubungan antara dua item dengan membandingkan nilai confidence terhadap nilai support konsekuensi. Rumus lift adalah sebagai berikut:

$$Lift(A \rightarrow B) = \frac{Confidence(A \rightarrow B)}{Support(B)} \quad (4)$$

Rumus (4) digunakan untuk mengetahui apakah hubungan antara item A dan item B benar-benar memiliki keterkaitan atau hanya terjadi secara kebetulan. Keterangan:

- a) $Lift(A \rightarrow B)$: Nilai kekuatan hubungan antara item A dan item B.
- b) $Confidence(A \rightarrow B)$: Nilai confidence aturan asosiasi.
- c) $Support(B)$: Nilai support item B.

Support dan *confidence* digunakan sebagai filter awal untuk menentukan aturan yang layak, sedangkan *lift* digunakan untuk mengukur seberapa kuat hubungan antar menu. Hasil analisis Apriori membantu Waroeng *Fake* dalam strategi *bundling* menu, promosi, dan pengembangan produk[19].

5. Membentuk *Dataset* Akhir: Setelah seluruh proses *preprocessing* selesai, data dibagi menjadi dua dataset akhir sesuai kebutuhan:

- a. Dataset pola kunjungan

Berisi tanggal, hari, jam, dan jumlah transaksi pada setiap periode.

- b. Dataset Apriori (itemset)

Berisi daftar menu dalam setiap transaksi untuk menghasilkan *association rule*.

Kedua dataset ini kemudian digunakan dalam tahap analisis pola kunjungan dan penerapan algoritma Apriori.

Setelah tahap preprocessing selesai, data yang telah siap selanjutnya dianalisis menggunakan algoritma Apriori. Penelitian ini menggunakan *Google Colab* dengan pustaka *Python mlxtend* sebagai alat bantu untuk memvalidasi hasil association rules yang diperoleh dari proses analisis algoritma Apriori. Proses validasi dilakukan dengan tahapan berikut: (1) data transaksi diunggah dan dinormalisasi nama itemnya; (2) data ditransformasi ke format boolean menggunakan modul *TransactionEncoder*; (3) frequent itemset dibentuk menggunakan fungsi *apriori()* dengan parameter minimum support 0,01 (1%); (4) aturan asosiasi dibentuk menggunakan fungsi *association_rules()* dengan parameter minimum confidence 0,20 (20%) dan disaring dengan nilai lift > 1.

3.5 Analisis Pola Kunjungan

Analisis pola kunjungan dilakukan dengan menghitung dan mengelompokkan frekuensi transaksi berdasarkan jam dan hari, untuk mengetahui kecenderungan waktu kedatangan pelanggan. Analisis ini dilakukan secara terpisah dari algoritma Apriori, karena bertujuan menggambarkan distribusi frekuensi kunjungan, bukan menemukan keterkaitan antar menu.

a. Perhitungan Persentase Kunjungan

Persentase kunjungan pada setiap jam dan hari dihitung dengan membandingkan jumlah transaksi pada kelompok waktu tertentu terhadap total keseluruhan transaksi, menggunakan rumus berikut:

$$Persentase(t) = \frac{Frekuensi(t)}{Total\ Transaksi} \times 100\%$$

Keterangan:

- a) Persentase(t) : Persentase jumlah transaksi pada jam/hari ke-t
- b) Frekuensi(t) : Jumlah transaksi yang tercatat pada jam/hari ke-t
- c) Total Transaksi : Jumlah seluruh transaksi yang dianalisis (1.143 transaksi)

b. Pengelompokan Status Kunjungan Berdasarkan Jam

Status kunjungan per jam dikelompokkan berdasarkan jumlah transaksi ke dalam tiga kategori sebagai berikut:

- a) Sepi: jumlah transaksi kurang dari 30
- b) Normal: jumlah transaksi 30 sampai dengan 100
- c) Ramai: jumlah transaksi lebih dari 100

c. Pengelompokan Status Kunjungan Berdasarkan Hari

Status kunjungan per hari dikelompokkan menggunakan pendekatan nilai tertinggi dan terendah (*min-max*): hari dengan jumlah transaksi tertinggi dikelompokkan sebagai "Tertinggi", hari dengan jumlah transaksi terendah dikelompokkan sebagai "Terendah", dan hari lainnya dikelompokkan sebagai "Normal".

d. Langkah Kerja Analisis di *Microsoft Excel*

Proses analisis pola kunjungan dilakukan menggunakan *Microsoft Excel* dengan tahapan sebagai berikut:

- a) Mengekstraksi jam transaksi dari kolom waktu menggunakan fungsi =*HOUR()* untuk memisahkan komponen jam dari data tanggal-waktu transaksi.
- b) Mengelompokkan dan menghitung jumlah transaksi pada setiap jam/hari menggunakan fungsi =*COUNTIFS()* atau fitur *PivotTable*.

- c) Menghitung persentase masing-masing kelompok jam/hari terhadap total transaksi menggunakan rumus pada poin (a).
- d) Mengelompokkan status kunjungan pada setiap kelompok berdasarkan kategori pada poin (b) dan (c).
- e) Mengurutkan hasil perhitungan untuk mengidentifikasi jam dan hari dengan tingkat kunjungan tertinggi dan terendah selama periode penelitian.

Hasil dari tahapan ini kemudian disajikan dalam bentuk tabel distribusi kunjungan berdasarkan jam dan hari, sebagaimana ditampilkan pada Bab IV.