

**PENGGALIAN WAWASAN DENGAN VISUALISASI  
DATA DAN ALGORITMA FP-GROWTH  
(STUDI KASUS NOBLE COFFEE)**



Disusun Oleh:

N a m a : Daffa Nayaka Bagaskara Jusuf  
NIM : 19523096

**PROGRAM STUDI INFORMATIKA – PROGRAM SARJANA  
FAKULTAS TEKNOLOGI INDUSTRI  
UNIVERSITAS ISLAM INDONESIA**

**2023**



## HALAMAN PENGESAHAN DOSEN PENGUJI

**PENGGALIAN WAWASAN DENGAN VISUALISASI  
DATA DAN ALGORITMA FP-GROWTH  
(STUDI KASUS NOBLE COFFEE)**

**TUGAS AKHIR**

Telah dipertahankan di depan sidang pengujian sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer dari Program Studi Informatika – Program Sarjana di Fakultas Teknologi Industri Universitas Islam Indonesia

Yogyakarta,

Tim Penguji

Dr. Syarif Hidayat, S.Kom., M.I.T.

**Anggota 1**

Andhika Giri Persada, S.Kom., M.Eng.

**Anggota 2**

Sheila Nurul Huda, S.Kom., M.Cs.

Mengetahui,

Ketua Program Studi Informatika – Program Sarjana

Fakultas Teknologi Industri

Universitas Islam Indonesia



(Dhomas Hatta Fudholi, S.T., M.Eng., Ph.D.)

**HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR**

Yang bertanda tangan di bawah ini:

Nama : Daffa Nayaka Bagaskara Jusuf

NIM : 19523096

Tugas akhir dengan judul:

**PENGGALIAN WAWASAN DENGAN VISUALISASI  
DATA DAN ALGORITMA FP-GROWTH  
(STUDI KASUS NOBLE COFFEE)**

Menyatakan bahwa seluruh komponen dan isi dalam tugas akhir ini adalah hasil karya saya sendiri. Apabila di kemudian hari terbukti ada beberapa bagian dari karya ini adalah bukan hasil karya sendiri, tugas akhir yang diajukan sebagai hasil karya sendiri ini siap ditarik kembali dan siap menanggung risiko dan konsekuensi apapun.

Demikian surat pernyataan ini dibuat, semoga dapat dipergunakan sebagaimana mestinya.

Yogyakarta, 2 November 2023



(Daffa Nayaka Bagaskara Jusuf)

## **HALAMAN PERSEMBAHAN**

Tugas akhir ini saya persembahkan untuk memenuhi kebutuhan orang-orang yang ingin mengetahui dan mempelajari pengolahan data lebih mendalam.

**HALAMAN MOTO**

*“Karena sesungguhnya sesudah kesulitan itu ada kemudahan”*

Q.S Al-Insyirah Ayat 5

*“Never ruin an apology with an excuse”*

Benjamin Franklin

*“To be, or not to be, that is the question”*

William Shakespeare

## KATA PENGANTAR

*Assalamu'alaikum Wr. Wb.*

Alhamdulillah, segala puji dan syukur kepada Allah SWT yang tak henti memberikan kenikmatan serta rahmat kepada seluruh hamba-Nya. Shalawat serta salam kepada junjungan kita Nabi Muhammad SAW.

Dengan rahmat dan hidayah Allah SWT tugas akhir dengan judul "PENGALIAN WAWASAN DENGAN VISUALISASI DATA DAN ALGORITMA FP-GROWTH (STUDI KASUS NOBLE COFFEE)" dapat terselesaikan. Tugas akhir ini disusun sebagai syarat yang harus dipenuhi untuk menyelesaikan studi Strata-1 pada program studi Informatika Fakultas Teknologi Industri Universitas Islam Indonesia.

Dalam proses menyelesaikan tugas akhir ini tidak lepas dari dukungan, bantuan, serta bimbingan dari berbagai pihak. Pada dasarnya saya ingin mengucapkan banyak terima kasih terhadap pihak yang telah memberikan dukungan secara langsung maupun tidak langsung, oleh karena itu dengan penuh rasa syukur saya ingin mengucapkan terimakasih kepada:


1. Bapak Dthomas Hatta Fudholi, S.T., M.Sc., Ph.D. selaku Ketua Program Studi Informatika Universitas Islam Indonesia yang telah membantu saya mendapatkan wawasan tentang dunia informatika secara luas.
2. Bapak Dr Syarif Hidayat, S.T., M.I.T sebagai dosen pembimbing dalam penjaluran penelitian saya yang selalu memberikan arahan, bimbingan, serta dukungannya dalam mengerjakan tugas akhir ini.
3. Ibu Elyza Gustri Wahyuni, S.T., M.Cs. sebagai dosen pembimbing akademik yang tidak henti-hentinya merespon kebingungan saya saat menempuh masa perkuliahan
4. Kedua orang tua saya yang telah memberikan segala bentuk kasih sayang sejak saya dilahirkan sampai waktu yang tak terhingga dan tak terhitung lamanya.
5. Kedua sahabat terdekat saya, Ghazy dan Farras, yang telah menemani saya mendapatkan berbagai jenis kenangan manis maupun pahit semasa perkuliahan.
6. Teman teman kos Martha Putra dan Santoso Dalem yang telah menemani masa perkuliahan saya dan membantu penyelesaian tugas akhir saya.
7. Rekan kerja serta atasan saya pada saat saya bekerja di Forus Coffee dan Noble Coffee yang telah memberikan dukungan dan pandangan terkait pengembangan visualisasi data penjualan.
8. Serta masih banyak lagi pihak lainnya yang tidak bisa disebutkan namanya.

Saya menyadari bahwa tugas akhir ini masih jauh dari sempurna. Oleh karena itu, saya sangat mengharapkan kritik, saran, dan masukan yang membangun dari pembaca agar penulisan ilmiah ini dapat lebih baik lagi di masa yang akan datang.

Akhir kata, saya berharap penulisan skripsi ini dapat memberikan manfaat dan kontribusi yang positif bagi perkembangan ilmu pengetahuan di bidang analisis data. Semoga tugas akhir ini dapat membuka wawasan dan memberikan inspirasi kepada pembaca yang mengkaji topik serupa di masa mendatang. Akhirnya, saya mengakhiri kata pengantar ini dengan memohon maaf jika terdapat kekurangan dan kesalahan di dalamnya. Semoga penulisan ini dapat menjadi pijakan untuk penelitian lebih lanjut di masa yang akan datang.

*Wassalamu'alaikum Wr. Wb.*

Yogyakarta, 2 November 2023



( Daffa Nayaka Bagaskara Jusuf )



## SARI

Perkembangan teknologi digital telah menciptakan transformasi besar dalam cara kita berkomunikasi dan berbisnis. Fenomena ini memberikan peluang baru bagi perusahaan untuk memahami pelanggan dengan lebih baik dan meningkatkan daya saing mereka. Dalam konteks ini, analisis data menjadi kunci, memungkinkan bisnis untuk mengoptimalkan operasional, mengelola persediaan, dan meningkatkan penjualan.

Studi ini mencoba menggali perilaku pembelian pelanggan di Noble Coffee, sebuah kafe di Yogyakarta. Dengan menggunakan algoritma FP-Growth, analisis data dilakukan untuk mengidentifikasi aturan asosiasi antar produk dan kategori dengan penjualan tertinggi. Visualisasi data dilakukan melalui platform Power BI, memungkinkan pemahaman mendalam tentang perilaku pembelian pelanggan.

Hasil analisis menunjukkan beberapa aturan asosiasi yang signifikan, seperti hubungan antara produk "King" dengan berbagai menu lainnya seperti "Mix Plater" dan "Café Latte". Analisis ini memberikan wawasan tentang produk yang sering dibeli bersamaan, memungkinkan Noble Coffee untuk merancang strategi penjualan yang lebih efektif.

Penelitian ini menyoroti pentingnya analisis data dan visualisasi dalam mendukung keputusan bisnis. Melalui pendekatan ini, perusahaan dapat memahami kebutuhan pelanggan dengan lebih baik, meningkatkan kepuasan pelanggan, dan mengoptimalkan operasional mereka. Kesimpulannya, analisis data berbasis algoritma FP-Growth dan visualisasi menggunakan Power BI dapat menjadi instrumen yang efektif dalam memahami perilaku pelanggan dan meningkatkan kinerja bisnis kafe seperti Noble Coffee.

Kata kunci: *FP-Growth*, Aturan Asosiasi, Analisis Data, Visualisasi Data, Power BI

## GLOSARIUM

Algoritma	Ilmu pengembangan algoritma yang memanfaatkan konsep matematis dan statistik dalam menjalankan tugas tertentu tanpa instruksi eksplisit.
Analisis	Proses pengolahan data dengan tujuan untuk menemukan informasi yang berguna
Asosiasi	Sifat saling terhubung
Data Warehouse	Sebuah sistem yang bertugas mengarsipkan sekaligus melakukan analisis data historis
Dataset	Kumpulan data diatur dalam format tertentu.
Informasi	Data yang memberi nilai tambah pada pemahaman suatu subjek. Adapun definisi lainnya bahwa data informasi adalah data yang telah dibentuk menjadi bentuk yang lebih berarti dan berguna bagi manusia.
Komputasi	Daya pemrosesan, memori, jaringan, penyimpanan, dan sumber daya lainnya
Machine Learning	Ilmu pengembangan algoritma yang memanfaatkan konsep matematis dan statistik dalam menjalankan tugas tertentu tanpa instruksi eksplisit.
Rantai Pasok	Serangkaian proses bisnis yang menghubungkan beberapa aktor untuk peningkatan nilai tambah bahan baku/produk dan mendistribusikannya kepada konsumen
Repository	Tempat penyimpanan

## DAFTAR ISI

HALAMAN JUDUL .....	i
HALAMAN PENGESAHAN DOSEN PEMBIMBING.....	ii
HALAMAN PENGESAHAN DOSEN PENGUJI.....	iii
HALAMAN PERNYATAAN KEASLIAN TUGAS AKHIR.....	iv
HALAMAN PERSEMBAHAN .....	v
HALAMAN MOTO.....	vi
KATA PENGANTAR .....	vii
SARI .....	ix
GLOSARIUM.....	x
DAFTAR ISI.....	xi
DAFTAR TABEL .....	xii
DAFTAR GAMBAR.....	xiii
BAB I PENDAHULUAN.....	2
1.1 Latar Belakang .....	2
1.2 Rumusan Masalah .....	4
1.3 Batasan Masalah.....	4
1.4 Tujuan Penelitian.....	4
1.5 Manfaat Penelitian.....	4
1.6 Sistematika Penulisan.....	5
BAB II LANDASAN TEORI.....	6
2.1 Penelitian Terdahulu .....	6
2.2 Intelijen Bisnis.....	7
2.3 Power BI.....	8
2.4 <i>Machine Learning</i> .....	10
2.5 Penambangan Aturan Asosiasi.....	11
2.6 FP-Growth.....	13
BAB III METODOLOGI PENELITIAN .....	15
3.1 Tahap Penelitian .....	15
3.2 Pengumpulan Data .....	16
3.3 Pembersihan Data.....	16
3.4 Implementasi Algoritma FP-Growth.....	16
3.5 ETL.....	17
3.6 Visualisasi Data.....	17
BAB IV HASIL DAN PEMBAHASAN.....	18
4.1 Pengumpulan Data .....	18
4.2 Pembersihan Data.....	20
4.3 Implementasi FP-Growth .....	23
4.4 ETL.....	27
4.5 Visualisasi Data.....	31
BAB V KESIMPULAN DAN SARAN .....	44
5.1 Kesimpulan.....	44
5.2 Saran.....	45
DAFTAR PUSTAKA .....	47
LAMPIRAN .....	52

**DAFTAR TABEL**

Tabel 4.1 Penjelasan Fungsi Kolom .....	19
---	----

## DAFTAR GAMBAR

Gambar 3.1 Tahap Penelitian .....	15
Gambar 4.1 Halaman Laporan Transaksi.....	18
Gambar 4.2 Pilihan Ekspor Data.....	19
Gambar 4.3 Dataset csv.....	19
Gambar 4.4 Inkonsistensi Penulisan Tanggal .....	21
Gambar 4.5 Kolom Kosong .....	21
Gambar 4.6 Kolom Nilai 0.....	21
Gambar 4.7 Kode Pembersihan Data .....	23
Gambar 4.8 Kode Pengelompokan Itemset.....	24
Gambar 4.9 Kode Pencarian Frekuensi Kemunculan Itemset.....	25
Gambar 4.10 Frekuensi Kemunculan Itemset.....	26
Gambar 4.11 Kode Pembentukan Aturan Asosiasi .....	26
Gambar 4.12 Aturan Asosiasi .....	27
Gambar 4.13 Connect Data .....	28
Gambar 4.14 Tipe Data Pada Kolom .....	28
Gambar 4.15 Duplikat Kolom Konsekuen .....	29
Gambar 4.16 Transformasi Data Aturan Asosiasi Item .....	30
Gambar 4.17 Relasi Antar Tabel.....	30
Gambar 4.18 Hubungan Antar Tabel .....	31
Gambar 4.19 Card .....	32
Gambar 4.20 Jam Sibuk .....	33
Gambar 4.21 Kuantitas Penjualan Produk Tertinggi.....	34
Gambar 4.22 Kuantitas Penjualan Kategori Tertinggi .....	35
Gambar 4.23 Penjualan Harian .....	36
Gambar 4.24 Tabel Detail .....	37
Gambar 4.25 Grafik Penjualan Minuman .....	38
Gambar 4.26 Visualisasi Aturan Asosiasi Minuman .....	39
Gambar 4.27 Grafik Penjualan Makanan .....	40
Gambar 4.28 Visualisasi Aturan Asosiasi Makanan .....	41
Gambar 4.29 Visualisasi Aturan Asosiasi Kategori Produk .....	42

## **BAB I**

### **PENDAHULUAN**

#### **1.1 Latar Belakang**

Perkembangan teknologi secara digital yang telah sampai pada titik ini terjadi secara pesat dalam beberapa tahun terakhir (Danuri, 2019). Peningkatan teknologi telah mengubah cara kita saling berkomunikasi. Kemampuan komunikasi manusia telah berkembang hingga menjadikan satu sama lain dapat tetap terhubung sehingga memperkuat hubungan sosial dan bisnis. Terlebih lagi dengan berkembangnya teknologi, transformasi digital telah membukakan peluang baru dalam dunia bisnis (Yana Siregar & Irwan Padli Nasution, 2020). Penggunaan data, analisis kecerdasan buatan, dan pengembangan aplikasi mengubah cara perusahaan melakukan interaksi dengan pelanggan, manajemen, serta rantai pasok.

Perkembangan teknologi digital telah mengubah tata kelola bisnis dengan menghadirkan efisiensi operasional yang lebih tinggi, peningkatan aksesibilitas informasi, dan peluang baru untuk inovasi produk dan layanan. Hal ini telah memungkinkan perusahaan untuk merespons pasar dengan lebih cepat dan lebih akurat, serta menghasilkan peningkatan daya saing yang signifikan (Indrayani, 2012). Penerapan teknologi digital juga berdampak pada model bisnis, memungkinkan peralihan dari penjualan produk menjadi model berlangganan atau layanan berbasis langganan, yang dapat meningkatkan loyalitas pelanggan dan pendapatan jangka panjang (Rifai et al., 2022).

Salah satu bidang bisnis yang terkena dampak dari perkembangan teknologi adalah kafe. Pengumpulan dan analisis data pelanggan memungkinkan kafe untuk memahami perilaku pembelian pelanggan, yang pada akhirnya memungkinkan untuk dilakukan penyusunan menu yang lebih sesuai dan promosi yang lebih efektif (Andhika Putra et al., 2020). Dengan mengamati pola pemesanan dan konsumsi, kafe dapat mengoptimalkan persediaan dan mengurangi pemborosan (Lenk et al., 2018). Penerapan analisis data juga dapat membantu kafe dalam pengelolaan tenaga kerja, dengan memprediksi jam-jam sibuk dan mengatur jadwal karyawan dengan lebih efisien (Ahn & Sohn, 2009). Dengan memahami tren pasar, kafe dapat merancang kampanye pemasaran yang lebih efektif dan tepat sasaran (Verma et al., 2012).

Penerapan analisis data memungkinkan kafe untuk meningkatkan penjualan, mengelola persediaan secara efisien, melakukan pemasaran dengan lebih efektif, dan memberikan pengalaman lebih memuaskan terhadap pelanggan (W. S. Lee et al., 2018). Salah satu alat analisis data yang dapat digunakan pada kasus ini adalah penerapan algoritma FP-

Growth(Wijaya & Pratama, 2022). FP-Growth adalah salah satu algoritma analisis data untuk mendapatkan frequent-pattern ataupun frequent itemset(Borgelt, 2005). Penerapan algoritma ini memungkinkan pengelola bisnis untuk mengidentifikasi produk paling populer dan menentukan aturan asosiasi yang terjadi didalamnya(Wei Zhang et al., 2008). Informasi yang didapatkan tersebut akhirnya dapat menciptakan paket promosi yang menarik untuk meningkatkan penjualan(Kurniawan & Yusuf, 2023).

Noble Coffee adalah café yang menghadirkan minuman coffee dan non-coffee serta makanan bertajuk makanan asia dan eropa. Noble Coffee merupakan cafe yang beralamatkan pada jalan Anggajaya 2 Condongcatur Yogyakarta. Untuk memertahankan sales dari bisnis tersebut diperlukan langkah agar dapat menjaga pelanggannya untuk tetap melakukan transaksi pada Noble Coffee. Berdasarkan beberapa penjelasan diatas, Noble Coffee perlu memahami kebiasaan perilaku transaksi para pelanggannya dengan memahami perilaku transaksi pelanggan secara mendalam. Kebiasaan perilaku transaksi pelanggan yang dimaksud adalah isi item dari satu struk penjualan. Item yang dibeli oleh pelanggan kebanyakan tidak hanya satu item saja. Item yang dibeli dapat berupa minuman saja, makanan saja, minuman dan makanan, makanan dan minuman, serta minuman dan makanan. Apabila mengidentifikasi perilaku transaksi tersebut dilakukan secara manual akan memakan waktu yang sangat lama serta terdapat kemungkinan terjadinya human error(D. Lee et al., 2013). Solusi yang dapat ditawarkan, seperti yang sudah dijelaskan diatas, dapat menggunakan algoritma FP-Growth untuk menyelesaikan masalah tersebut.

Banyaknya data transaksi yang terjadi pada Noble Coffee dapat diubah menjadi itemset yang bisa digunakan sebagai variabel untuk mencari pola transaksi pelanggan. Algoritma FP-Growth melakukan pencarian frekuensi kemunculan item dalam itemset dengan menggunakan association rules. Adapun cara untuk menampilkan wawasan kemungkinan kemunculan item dengan melihat minimum *support* dan minimum *confidence*.

Noble Coffee ini hanya memiliki 1 jenis produk *bundling* yang mengalami penurunan penjualan dalam beberapa bulan terakhir. Promo *bundling* tersebut adalah promo makan pagi yang dimulai pada pukul 10.00 sampai 14.00. Akan tetapi promo tersebut tidak dapat meningkatkan penjualan secara maksimal. Oleh karena itu diperlukan analisis ulang terkait penjualan yang ada dalam bulan tersebut. Hasil dari data transaksi yang nantinya dianalisis dapat diimplementasikan juga dengan menggunakan algoritma FP-Growth serta nantinya akan divisualisasikan untuk memudahkan dalam pembacaannya.

Visualisasi data dalam penelitian ini menggunakan aplikasi Power BI. Power BI adalah platform visualisasi data yang memungkinkan pengguna untuk mengubah data yang kompleks menjadi visualisasi yang mudah dipahami (Becker & Gould, 2019). Dengan Power BI, pengguna dapat membuat dashboard interaktif, laporan, dan diagram yang memungkinkan mereka untuk dengan cepat mengidentifikasi tren, pola, dan wawasan penting dari data mereka. Pengguna juga dapat menggabungkan data dari berbagai sumber yang berbeda, memberikan gambaran holistik tentang bisnis atau situasi tertentu. Dengan visualisasi yang menarik dan mudah diakses, Power BI membantu pengguna dalam pengambilan keputusan yang lebih informasional dan lebih baik.

## 1.2 Rumusan Masalah

Bagaimana cara mendapatkan pemahaman terkait perilaku transaksi yang terjadi untuk dapat meningkatkan penjualan setelah terjadi penurunan penjualan terhadap promosi *item bundling* yang telah berlaku?

## 1.3 Batasan Masalah

Batasan masalah pada penelitian ini adalah:

- a. Data yang diambil berasal dari data transaksi Noble Coffee yang diperoleh dari Moka POS pada bulan Juni dan Juli 2023
- b. Data diolah menggunakan *python pyspark*.
- c. Data yang digunakan adalah data pembelian.
- d. Algoritma yang digunakan untuk mendapatkan hubungan asosiasi produk terbeli oleh pelanggan pada penelitian ini adalah algoritma FP-Growth

## 1.4 Tujuan Penelitian

Mendapatkan rekomendasi *item bundling* yang dapat diterapkan sebagai promosi dengan melihat hasil visualisasi data dengan hasil implementasi algoritma FP-Growth.

## 1.5 Manfaat Penelitian

- a. Membantu pihak seperti pengelola bisnis untuk mendapatkan pemahaman terkait perilaku transaksi, meningkatkan penjualan, meramalkan perilaku pasar, serta dapat mengambil keputusan lebih baik.
- b. Membantu sebagai rujukan untuk penelitian dengan topik yang berkaitan.



## **1.6 Sistematika Penulisan**

Sistematika penulisan dalam penyusunan tugas akhir ini mencakup beberapa bab. Bab-bab tersebut memberikan gambaran menyeluruh tentang masalah yang dihadapi dan cara penyelesaiannya. Setiap bab memberikan pemahaman mendalam tentang aspek-aspek tertentu dalam penelitian ini. Berikut sistematika penulisan yang terbagi dalam 5 bab:

### **BAB I PENDAHULUAN**

Bab ini membahas latar belakang masalah, rumusan masalah, batasan masalah, tujuan penelitian, manfaat penelitian, dan sistematika penulisan.

### **BAB II LANDASAN TEORI**

Bab ini membahas konsep, teori, dan studi terdahulu yang relevan dengan topik penelitian. Landasan teori membantu pembacanya memahami relevansi dan urgensi topik penelitian dengan membandingkan hasil penelitian terdahulu, teori yang telah ada, dan temuan terkini tentang bidang tersebut untuk memberikan dasar perumusan kerangka penelitian.

### **BAB III METODOLOGI PENELITIAN**

Bab ini membahas tahapan penelitian yang akan dilakukan dengan memaparkan pengambilan data, pemrosesan data, implementasi algoritma FP-Growth, dan visualisasi data.

### **BAB IV HASIL DAN PEMBAHASAN**

Bab ini membahas tentang hasil pengimplementasian metode serta pembahasannya.

### **BAB V KESIMPULAN DAN SARAN**

Bab ini adalah bab terakhir yang membahas kesimpulan dan saran dari penelitian yang telah dilakukan.

## **BAB II**

### **LANDASAN TEORI**

#### **2.1 Penelitian Terdahulu**

Pada bagian ini peneliti akan memaparkan hasil dari beberapa penelitian yang telah ada dengan topik yang relevan. Pemaparan hasil penelitian terdahulu merupakan hal penting dan dapat digunakan sebagai referensi. Berikut ini merupakan beberapa penelitian terdahulu yang penulis jadikan referensi.

Penelitian pertama dilakukan oleh (SAPTOMO, 2022) berjudul “Implementasi *Business Intelligence* Dalam Membuat Keputusan Penjualan Properti Dengan Metode Online Analytical Processing (OLAP)”. Pada penelitian ini didapatkan hasil berupa performa yang didapatkan dari implementasi intelijen bisnis. Perusahaan mampu mendapatkan wawasan informasi bermanfaat dari data yang telah dimiliki serta dari luar perusahaan dengan bentuk visualisasi. Wawasan yang didapatkan berupa informasi berguna untuk membantu pengambilan keputusan berupa hubungan antar faktor yang mempengaruhi pembelian seperti umur, golongan pekerjaan, dan upah. Dari informasi yang didapatkan, didapatkan fakta bahwa penjualan yang terjadi mengalami penurunan dikarenakan kondisi daya beli masyarakat yang terkena dampak covid-19.

Penelitian kedua dilakukan oleh (QISTHANI, 2019) berjudul “Perancangan Sistem Business Intelligence Untuk Mengetahui Perilaku Penjualan Produk”. Penelitian ini terfokuskan pada analisis perilaku penjualan produk. Dari penelitian ini didapatkan fakta bahwa hasil dari visualisasi data dapat membantu mendapatkan sejumlah fakta yang mempengaruhi suatu penjualan. Penelitian ini juga memberikan fakta bahwa dengan implementasi intelijen bisnis, pelaku bisnis dapat memantau apa yang terjadi apabila perusahaannya memiliki cabang lebih dari 1 outlet.

Penelitian ketiga dilakukan oleh (Zikra, 2022) berjudul “Implementasi *Business Intelligence* pada ACCAbsensi Menggunakan Aplikasi Power BI”. Penelitian tersebut melakukan perhitungan kinerja karyawan menggunakan metode ETL. Metode tersebut membuat perhitungan kinerja lebih efisien dengan update yang dapat dilakukan secara berkala. Visualisasi yang dilakukan membudahkan monitoring karyawan serta memberikan wawasan tambahan mengenai kinerja karyawan.

Penelitian keempat dilakukan oleh (Nabilla, 2021) berjudul “Pengembangan Business Intelligence Pada Sistem Informasi Distributor”. Dari penelitian ini didapatkan hasil berupa distributor masker dapat mengetahui pendapatan dan profitnya sesuai dengan periode waktu yang diinginkan. Dari pengembangan intelijen bisnis pada sistem ini, disebutkan bahwa dengan mengetahui varian masker mana saja yang memberikan profit paling banyak serta paling sedikit distributor dapat mengambil langkah dan strategi untuk mengambil keputusan marketing. Distributor masker ini juga dapat mengambil wawasan reseller dengan penjualan tertinggi untuk nantinya dapat dijadikan mitra.

Penelitian kelima dilakukan oleh (Edhya, 2022) berjudul “Business Intelligence Data Marketing Menggunakan Metode Kimball Dan ETL Dengan Power BI”. Penelitian ini menghasilkan informasi berupa terjadi peningkatan pendapatan, jenis produk dengan penjualan tertinggi, dan produk dengan penjualan terendah. Selain itu ada juga peta yang menampilkan jumlah pembelian tertinggi dalam satu daerah. Visualisasi yang dilakukan menggunakan Power BI juga membuat pembacaan data lebih menarik dan mudah dipahami.

## **2.2 Intelijen Bisnis**

Intelijen bisnis merupakan sebuah pendekatan terstruktur yang memanfaatkan teknologi dan berbagai metode analisis data untuk merubah data mentah menjadi informasi yang bermakna dan bermanfaat (Chaudhuri et al., 2011). Dalam era di mana data menjadi aset terbesar, intelijen bisnis menjadi landasan penting bagi keberhasilan organisasi dan perusahaan modern (Foley & Guillemette, 2010). Pendekatan ini melibatkan sejumlah teknik dan alat, termasuk analisis bisnis yang mendalam, data mining, visualisasi data yang kreatif, alat pengolahan data canggih, dan infrastruktur data yang efisien. Semua ini bertujuan untuk membantu organisasi memahami data mereka dengan lebih baik, membimbing pengambilan keputusan bisnis yang cerdas, serta meningkatkan kinerja dan daya saing mereka di pasar yang kompetitif.

Salah satu peran utama intelijen bisnis adalah menyediakan metrik kinerja dan perbandingan standar (Watson & Wixom, 2007). Dalam dunia bisnis yang terus berubah, memiliki informasi yang tepat waktu mengenai kinerja perusahaan adalah kunci. Dengan menggunakan intelijen bisnis, pimpinan bisnis dapat memantau kemajuan organisasi mereka terhadap tujuan bisnis yang telah ditetapkan. Informasi ini bukan hanya mencakup aspek finansial, melainkan juga menilai performa operasional dan strategis perusahaan. Keputusan bisnis yang diambil berdasarkan pemahaman mendalam mengenai data tersebut dapat

membimbing organisasi menuju arah yang benar, meningkatkan keunggulan kompetitif, dan memberikan nilai tambah kepada pelanggan.

Intelijen bisnis juga membuka pintu bagi organisasi untuk memahami perilaku pembelian pelanggan dengan lebih baik. Dalam dunia yang dipenuhi dengan data transaksi, intelijen bisnis membantu dalam mengurai pola-pola yang tersembunyi di balik sekumpulan angka dan fakta. Dengan memahami perilaku pembelian pelanggan, bisnis dapat menyusun strategi pemasaran yang lebih efektif dan personal (LEE & PARK, 2005). Analisis pelanggan ini tidak hanya sebatas mengetahui produk atau layanan yang paling diminati, tetapi juga memahami konteks di balik setiap keputusan pembelian. Hal ini membantu bisnis membangun hubungan yang kuat dengan pelanggan, menciptakan pengalaman yang lebih memuaskan, dan memperluas pangsa pasar.

Dalam dunia bisnis yang dinamis, mengidentifikasi peluang baru adalah kunci keberhasilan (Jenkins, 2009). Intelijen bisnis memberikan kemampuan untuk menggali data secara mendalam, memahami tren pasar, dan mengidentifikasi peluang pertumbuhan yang mungkin terlewatkan oleh mata manusia. Alat data mining dan analisis prediktif dalam intelijen bisnis membantu organisasi melihat ke depan, meramalkan perilaku pasar, dan merespons dengan cepat terhadap perubahan tren (Bose & Mahapatra, 2001). Dengan memanfaatkan teknologi ini, bisnis dapat beradaptasi dengan lingkungan bisnis yang dinamis dan meraih peluang pertumbuhan bahkan di tengah persaingan yang ketat.

Intelijen bisnis tidak hanya menjadi sekadar alat analisis data, melainkan menjadi kekuatan pengubah permainan di dunia bisnis. Organisasi dan perusahaan yang mampu mengoptimalkan potensi intelijen bisnis tidak hanya mampu bertahan di pasar yang kompetitif, tetapi juga mampu memimpin dengan inovasi, memahami konsumen dengan lebih baik, dan menciptakan solusi bisnis yang relevan serta berdaya saing (Gessner & Scott, 2009). Oleh karena itu, investasi dalam intelijen bisnis tidak hanya merupakan kebutuhan, tetapi juga merupakan strategi bisnis yang bijak dalam menghadapi tantangan dan peluang yang ada di era digital ini

### **2.3 Power BI**

Dilansir dari halaman utama Microsoft, Power Business Intelligence adalah software yang dikembangkan oleh Microsoft sebagai kumpulan layanan perangkat lunak, aplikasi, dan konektor yang bekerja bersamaan untuk mengubah berbagai sumber data menjadi visualisasi data yang bersifat statis dan interaktif. Power BI menyediakan layanan *cloud-based* bernama "Power BI Services", serta layanan *desktop* bernama "Power BI Desktop". Power BI

menawarkan *data warehouse* dengan kemampuan meliputi data preparation, data discovery, dan dashboard interaktif.

Power BI merupakan alat intelijen bisnis serbaguna yang dapat digunakan dalam skala industri yang luas. Berikut beberapa contoh penggunaan Power BI dalam industri yang berbeda:

### **2.3.1 Manufaktur**

Power BI dalam segi manufaktur dapat digunakan untuk analisis data produksi, penjualan, dan pendapatan dengan aman (Lima et al., 2021). Selain itu dapat digunakan untuk memantau peralatan, serta memberikan pemeliharaan preventif yang proaktif. Fokus lainnya lagi ditujukan pada pembentukan rantai pasok yang lebih tangguh. Dengan pemahaman mendalam terhadap data, perusahaan mampu merancang strategi pasok yang lebih efisien dan responsif terhadap perubahan pasar. Penggunaan Power BI dalam konteks manufaktur memiliki potensi besar untuk memberdayakan semua karyawan dengan memungkinkan mereka mengakses, berkolaborasi, dan menganalisis data di dalam organisasi dengan lebih efektif.

### **2.3.2 Ritel**

Dalam konteks ritel, Power BI bukan hanya sekadar alat, melainkan juga solusi cerdas untuk manajemen rantai pasok yang efektif (Iliashenko et al., 2019). Dengan pemanfaatan Power BI, para pelaku usaha ritel dapat mengoptimalkan strategi pemasaran mereka melalui analisis data yang mendalam. Melalui analisis prediktif yang canggih, pelaku usaha ritel dapat memahami perilaku pelanggan secara lebih baik, memungkinkan pengambilan keputusan yang lebih tepat. Pengelolaan data yang efisien juga merupakan keunggulan Power BI, membantu pelaku usaha ritel dalam menyusun strategi berbasis data yang kuat dan berdaya saing tinggi.

### **2.3.3 Kesehatan**

Dalam bidang kesehatan, Power BI menjadi alat yang memungkinkan untuk berbagi wawasan dan pengetahuan dengan para ahli kesehatan di seluruh dunia yang memungkinkan mereka untuk memperoleh pendekatan terbaik dalam merawat pasien (Sharma & Sarkar, 2022). Selain itu, Power BI juga memungkinkan pemantauan penggunaan energi dalam periode waktu tertentu dengan akurasi tinggi. Dengan kemampuan prediktifnya yang canggih, Power BI dapat mendeteksi potensi risiko seperti gangguan pasokan listrik secara dini, dan memberikan solusi secara real-time (Mosadeghrad et al., 2021). Data yang diperoleh juga dapat digunakan untuk mencapai efisiensi tinggi dalam penggunaan energi, memberikan dampak positif pada lingkungan dan efektivitas biaya di sektor kesehatan.

## 2.4 Machine Learning

*Machine Learning* merupakan cabang dari kecerdasan buatan (AI) dan ilmu komputer yang terfokuskan pada penggunaan data dan algoritma untuk meniru cara manusia belajar. Algoritma Machine Learning menggunakan metode komputasional untuk "mempelajari" informasi langsung dari data tanpa bergantung pada persamaan yang telah ditetapkan sebelumnya sebagai model (Jordan & Mitchell, 2015). Terdapat dua jenis antara lain:

1. Pembelajaran Terawasi (Supervised Learning)

Pembelajaran terawasi adalah jenis *machine learning* yang menggunakan data berlabel untuk melatih model (Jiang et al., 2020). Dalam pembelajaran terawasi, data pelatihan yang mencakup input dan output diinginkan diproses untuk membangun fungsi yang dapat memetakan data baru ke nilai output yang diharapkan. Ada dua jenis masalah utama dalam pembelajaran terawasi: klasifikasi dan regresi (Burkart & Huber, 2021). Pengguna harus menentukan jenis data yang akan digunakan sebagai set pelatihan dan mengumpulkan set pelatihan yang mewakili penggunaan fungsi di dunia nyata. Representasi input dan struktur fungsi dipelajari juga harus ditentukan.

Algoritma pembelajaran terawasi seperti jaringan saraf tiruan, naïve Bayes, dan mesin vektor pendukung (SVM) digunakan untuk mencapai ini. Sistem dilatih dengan data berlabel, yang kemudian diuji dengan data yang telah diberi label tetapi tidak diungkapkan kepada algoritma. Model pembelajaran terawasi dapat memprediksi hasil di masa depan berdasarkan data pelatihan (Saravanan & Sujatha, 2018). Ini membantu menghilangkan pekerjaan klasifikasi manual dan mendukung pengambilan keputusan dengan akurasi. Pentingnya pembelajaran terawasi semakin terlihat dalam aplikasinya di berbagai bidang termasuk pengenalan gambar, analisis teks, dan prediksi pasar keuangan.

2. Pembelajaran Tidak Terawasi (Unsupervised Learning)

Pembelajaran tak terawasi adalah jenis *machine learning* di mana algoritma menemukan pola dari data yang tidak berlabel (Bhatia et al., 2019). Tanpa membutuhkan data berlabel sebelumnya, pembelajaran tak terawasi mengidentifikasi pola tersembunyi dan struktur intrinsik dalam data. Tujuannya adalah mengelompokkan data berdasarkan kesamaan, pola, dan perbedaan tanpa petunjuk awal. Model tak terawasi digunakan untuk tugas pengelompokan, asosiasi, dan reduksi dimensionalitas. Pengelompokan menggabungkan dataset dengan atribut serupa, asosiasi menemukan hubungan antar variabel, dan reduksi dimensionalitas mengurangi jumlah variabel dengan mempertahankan informasi.

Contoh algoritma tak terawasi termasuk analisis komponen utama, analisis kluster, dan deteksi anomali. Pembelajaran ini secara komputasi tergolong mahal, terutama untuk dataset besar atau algoritma kompleks. Meskipun demikian, model ini lebih ekonomis karena tidak memerlukan data berlabel yang sulit dan mahal didapatkan.

*Machine learning* memiliki dua tujuan utama: mengklasifikasikan data berdasarkan model yang telah dikembangkan, dan membuat prediksi untuk hasil di masa depan berdasarkan model tersebut (Carbonell et al., 1983). Sebagai contoh, algoritma *machine learning* yang khusus untuk mengklasifikasikan data mungkin menggunakan visi komputer pada tahi lalat, dikombinasikan dengan pembelajaran terawasi untuk melatihnya mengklasifikasikan tahi lalat yang bersifat kanker. Sebuah algoritma *machine learning* untuk perdagangan saham dapat memberi informasi pedagang tentang prediksi potensial di masa depan.

*Machine learning* merupakan komponen penting dalam bidang *data science* yang terus berkembang. Melalui penggunaan metode statistika, algoritma dilatih untuk membuat klasifikasi atau prediksi, serta menemukan wawasan kunci dalam proyek penambangan data. Algoritma *machine learning* menemukan pola alami dalam data yang menghasilkan wawasan dan membantu membuat keputusan dan prediksi yang lebih baik. Mereka digunakan setiap hari untuk membuat keputusan kritis dalam diagnosis medis, perdagangan saham, peramalan beban energi, dan banyak lagi.

## 2.5 Penambangan Aturan Asosiasi

Penambangan aturan asosiasi adalah teknik yang digunakan untuk mengidentifikasi pola dan hubungan antara kumpulan data yang besar (Agrawal & Shafer, 1996). Hal ini melibatkan penemuan hubungan antara variabel dalam data dan menggunakan hubungan tersebut untuk membuat prediksi atau keputusan. Contoh klasik dari penambangan aturan asosiasi adalah *market basket analysis*, di mana seorang pembeli menaruh barang dari toko ke dalam keranjang, dan mungkin merekomendasikan barang terkait yang tersedia di toko kepada pembeli (Kaur & Kang, 2016). Contoh lainnya adalah kumpulan data transaksi di toko kelontong, di mana penambangan aturan asosiasi dapat digunakan untuk mengidentifikasi hubungan antara barang-barang yang sering dibeli bersama. Sebagai contoh, aturan "Jika pelanggan membeli roti, mereka juga mungkin membeli susu" adalah aturan asosiasi yang bisa ditambang dari kumpulan data ini. Aturan asosiasi berguna untuk menganalisis dan

memprediksi perilaku pelanggan, analisis keranjang belanja, pengelompokan produk, desain katalog, tata letak toko, dan merekomendasikan barang di situs web e-commerce.

Aturan asosiasi terdiri dari dua bagian: antisedan (jika) dan konsekuen (maka). Antisedan adalah item yang ditemukan dalam data, dan konsekuen adalah item yang ditemukan bersama dengan antisedan. Kekuatan dari asosiasi ini diukur menggunakan keyakinan (*confidence*) aturan, yaitu probabilitas bahwa item B hadir jika item A hadir. Banyak algoritma untuk menghasilkan aturan asosiasi telah diajukan, seperti Apriori, Eclat, dan FP-Growth. Algoritma Apriori adalah pendekatan paling sederhana dan langsung dalam penambangan aturan asosiasi, akan tetapi terdapat banyak peningkatan yang dapat diterapkan padanya.

Untuk mengevaluasi kekuatan hubungan antar variabel dalam data diperlukan tolak ukur yang disebut dengan *support*, *confidence*, dan *lift* (McNicholas et al., 2008). Penjelasan dari tolak ukur tersebut adalah sebagai berikut:

### 1. *Support*

*Support* didefinisikan sebagai proporsi kasus di mana aturan asosiasi berlaku, atau dengan kata lain, persentase kemunculan item dalam bagian antisedan (bagian "jika" dari aturan) muncul dalam transaksi yang sama dengan item dalam konsekuen (bagian "maka" dari aturan). *Support* adalah ukuran seberapa sering *itemset* muncul secara bersamaan sebagai persentase dari keseluruhan transaksi. Nilai *support* dihitung dengan membagi jumlah transaksi dari data yang mengandung *itemset* dengan jumlah total transaksi. Nilai *support* yang tinggi menunjukkan bahwa *itemset* sering muncul dalam data.

$$Support(itemset) = \frac{Jumlah\ transaksi\ itemset}{Jumlah\ transaksi\ dataset}$$

### 2. *Confidence*

*Confidence* adalah ukuran seberapa mungkin suatu *itemset* akan muncul jika *itemset* lain muncul. *Confidence* adalah probabilitas bahwa konsekuen (bagian "maka" dari aturan) muncul dalam suatu transaksi jika antisedan (bagian "jika" dari aturan) muncul dalam transaksi yang sama. *Confidence* dihitung dengan membagi jumlah transaksi yang mengandung baik antisedan dan konsekuen dengan jumlah transaksi yang mengandung antisedan. Nilai *confidence* yang tinggi menunjukkan bahwa aturan tersebut benar untuk sebagian besar transaksi yang mengandung antisedan.

$$Confidence(A \rightarrow B) = \frac{Support(A \cup B)}{Support(A)}$$

### 3. *Lift*



Lift adalah ukuran kekuatan hubungan antara antisedan dan konsekuen. Lift mengukur rasio antara *support* yang diamati dengan *support* yang diharapkan jika antisedan dan konsekuen bersifat independen. Nilai lift yang lebih besar dari 1 menunjukkan bahwa antisedan dan konsekuen berkorelasi positif, sementara nilai lift yang kurang dari 1 menunjukkan bahwa mereka berkorelasi negatif. Nilai lift 1 menunjukkan bahwa antisedan dan konsekuen bersifat independen.

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

## 2.6 FP-Growth

FP-Growth adalah algoritma yang digunakan dalam penambangan data untuk menemukan pola-pola yang sering muncul tanpa melalui proses pembentukan calon. Algoritma ini merupakan salah satu perkembangan dari algoritma Apriori, yang menghasilkan itemset dan menemukan itemset yang paling sering muncul dengan memiliki kelemahan tersendiri (Hunyadi, 2011). Algoritma FP-Growth membangun FP-Tree, yang menggambarkan basis data dalam bentuk pohon, bukan dengan menggunakan strategi buat dan uji milik algoritma Apriori. Fokus dari algoritma FP-Growth adalah menjabarkan jalur-jalur item dan penambangan pola.

Terdapat beberapa keunggulan dari algoritma FP-Growth apabila dibandingkan dengan algoritma Apriori, antara lain:

1. Algoritma FP-Growth hanya membutuhkan dua kali pemindaian basis data dibandingkan dengan algoritma Apriori yang memindai transaksi untuk setiap iterasi.
2. Itemset tidak dibentuk di dalam algoritma ini, sehingga mempercepat proses pembuatan.
3. Basis data disimpan dalam versi yang lebih padat di dalam memori.
4. Skalabilitas algoritma ini efisien untuk digunakan dalam menambang *frequent patterns* yang Panjang maupun pendek.

Pada sisi lainnya algoritma Apriori memiliki keunggulannya sendiri, seperti:

1. Bekerja lebih cepat pada dataset yang lebih besar.
2. Pembentukan kandidat dapat diparalelkan
3. Melakukan scan database dengan lebih mendalam

Namun algoritma apriori memiliki beberapa kelemahan, seperti:

1. Apriori perlu dilakukan pembentukan kandidat itemset terlebih dahulu, sehingga memiliki kemungkinan jumlah itemset akan sangat banyak apabila datasetnya besar.
2. Apriori perlu melakukan pemindaian secara berulang kali terhadap database untuk memastikan *support* tiap itemset telah dibuat. Hal ini membuat komputasi menjadi mahal.
3. Pembentukan kandidat itemset yang sangat lama mengakibatkan *runtime* meningkat drastis dengan mengacu pada jumlah item yang berbeda.

Oleh karena itu, penggunaan algoritma FP-Growth lebih efisien dan cepat dibandingkan dengan algoritma Apriori. Akan tetapi performa kedua algoritma ini bergantung pada beberapa faktor seperti ukuran dataset, *minimum support*, dan jumlah itemset.

Algoritma FP-Growth memiliki dua fase utama yaitu konstruksi FP-Tree dan penambahan itemset yang sering muncul(Wang et al., 2017). Berikut adalah rincian dari fase konstruksi FP-Tree:

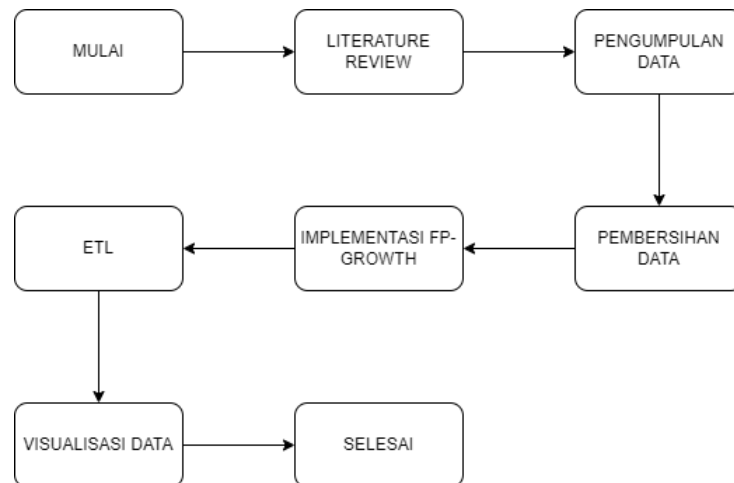
1. Pemindaian basis data dengan menghitung frekuensi kemunculan setiap item.
2. Item yang tidak sering muncul, diukur dengan *minimum support*, akan dihapus dari basis data untuk mengurangi ukuran dataset.
3. Item yang tersisa diurutkan berdasarkan frekuensinya secara dari yang tertinggi sampai yang terbawah.
4. Node akar dibuat untuk FP-Tree dengan nilai kosong.
5. Untuk tiap transaksi dalam basis data, jalur dibuat dalam FP-Tree dengan menambahkan item dalam transaksi satu persatu.
6. Jika item sudah ada dalam jalur maka jumlah item ditambahkan.
7. Namun jika item tidak ada dalam jalur, maka node baru akan dibuat untuk item tersebut.

Setelah FP-Tree telah dibuat tahapan selanjutnya adalah melakukan penambahan itemset yang sering muncul. Dalam penambahan itemset ini, terdapat 3 tahapan, antara lain:

1. Itemset yang sering muncul di-*extract* dari FP-Tree.
2. Untuk tiap itemset yang sering muncul, basis pola dasar dibuat dengan meng-*extract* jalur yang muncul di dalam FP-Tree.
3. Basis pola dasar tersebut digunakan untuk membuat FP-Tree baru yang nantinya dilakukan penambahan lagi secara berulang untuk mendapatkan itemset yang sering muncul.

## BAB III METODOLOGI PENELITIAN

### 3.1 Tahap Penelitian



Gambar 3.1 Tahap Penelitian

Ilustrasi skema pada Gambar 3.1 menggambarkan urutan tahapan yang dilakukan selama penelitian. Proses analisis data dan visualisasi dilakukan setelah data telah dikumpulkan dan dibersihkan. Dalam penelitian ini, tahapan pembersihan data dan implementasi algoritma FP-Growth menggunakan bahasa pemrograman *Python* dengan tambahan implementasi *PySpark*.

Penelitian ini dimulai dengan proses pengumpulan data transaksi sebagai langkah awal dalam pengembangan dataset penelitian. Data tersebut kemudian menjalani tahap pembersihan, di mana data-data yang tidak relevan dan inkonsisten dihapus agar mendapatkan keakuratan dan konsistensi yang diperlukan. Setelah mendapatkan dataset yang akurat, langkah berikutnya melibatkan penerapan algoritma FP-Growth untuk mengidentifikasi serangkaian itemset yang sering diambil, membentuk hubungan asosiasi yang signifikan.

Proses selanjutnya melibatkan tahap ETL (Extract, Transform, Load) guna menghubungkan data utama dengan aturan asosiasi, membangun relasi yang diperlukan. Setelah proses ETL diselesaikan, data siap untuk divisualisasikan menggunakan alat intelijen bisnis Power BI. Visualisasi ini memberikan pemahaman intuitif tentang perilaku pembelian pelanggan dan tren yang teridentifikasi. Tahapan terakhir penelitian ini adalah analisis hasil

temuan yang terungkap melalui dashboard visualisasi, memberikan wawasan mendalam tentang perilaku pembelian yang dapat diambil sebagai kesimpulan penelitian.

### 3.2 Pengumpulan Data

Data yang digunakan dalam penelitian ini berasal dari hasil ekstraksi data transaksi data oleh Noble Coffee yang tersimpan dalam backoffice Moka POS. Saat ekstraksi data transaksi ini, akan diperoleh data dengan format CSV.

### 3.3 Pembersihan Data

Pembersihan data atau *data cleaning* adalah tahapan krusial dalam proses *machine learning* dengan mengidentifikasi serta menghapus data yang hilang, duplikat, dan tidak relevan. Tujuan utama dari proses pembersihan data adalah memastikan bahwa data yang digunakan dalam analisis tersebut akurat, konsisten, dan bebas dari kesalahan. Ketika pembersihan data diabaikan, kinerja *machine learning* dapat terpengaruh oleh data yang tidak akurat dan inkonsisten.

Contoh dari pembersihan data antara lain:

1. Menghapus data duplikat atau tidak relevan.
2. Memastikan konsistensi data dengan memastikan format tiap dataset berbentuk sama. Seperti melihat format dari penulisan tanggal dan waktu, penulisan tersebut harus terstandar agar tiap dataset yang ada menjadi konsisten.
3. Menghapus data dengan nilai yang hilang mengidentifikasi data yang kosong.

Keberhasilan pembersihan data sangat penting karena data mentah sering kali tidak lengkap dan tidak konsisten sehingga dapat merugikan hasil analisis. Dengan memiliki data yang bersih, analisis yang dilakukan memiliki dasar yang kuat sehingga memungkinkan untuk dilakukan interpretasi dan prediksi yang akurat. Sebaliknya apabila data tidak bersih, model *machine learning* dapat memberikan hasil yang salah dan merugikan proses pengambilan keputusan. Oleh karena itu, memperhatikan pembersihan data dengan cermat adalah langkah yang sangat penting dalam pengembangan solusi *machine learning*.

### 3.4 Implementasi Algoritma FP-Growth

Setelah data telah dibersihkan, barulah algoritma FP-Growth dapat diimplementasikan. Dari dataset yang telah dibersihkan, dilanjutkan untuk membuat itemset sesuai dengan nomor resi atau id transaksi yang bersifat unik. Implementasi algoritma FP-Growth ini dilakukan

dengan memanfaatkan *class* yang telah tersedia dalam *package* oleh *PySpark*. *Package* tersebut ada didalam class machine learning *PySpark* yang dapat diakses melalui 'pyspark.ml.fpm.FPGrowth'. Dalam implementaisnya diperlukan untuk memenuhi parameter *minimum support* dan *minimum confidence*.

### 3.5 ETL

Sebelum data divisualisasikan, perlu untuk dilakukan *extract*, *transform*, dan *load* terlebih dahulu. ETL berfungsi sebagai integrasi antar data untuk disimpan dalam satu *repository* agar memiliki format yang sama sebagai persiapan visualisasi. ETL meningkatkan hasil analisis data dan intelijen bisnis dengan kemampuannya untuk memproses data secara efisien.

### 3.6 Visualisasi Data

Setelah ETL dilakukan, tahapan selanjutnya adalah visualisasi data. Visualisasi data adalah representasi data menggunakan elemen visual seperti grafik, diagram, dan peta berbentuk informasi. Visualisasi data memudahkan pembacanya mendapatkan informasi, wawasan, dan cerita dengan merubah lembaran angka dalam tabel menjadi grafik dan diagram. Pentingnya visualisasi data adalah dapat membantu pembacanya melihat, berinteraksi, dan memahami data dengan lebih baik. Baik itu sederhana maupun kompleks, visualisasi yang tepat mampu membawa semua pembacanya ke arah yang sama tanpa melihat tingkat keahlian mereka.

Nantinya visualisasi data dapat menggunakan jenis grafik berupa diagram garis, batang, *pie*, hash tree, maupun *network*. Tiap visualisasi memiliki keunggulannya masing masing dalam memudahkan pembacanya mengartikan suatu data.

## BAB IV

### HASIL DAN PEMBAHASAN

#### 4.1 Pengumpulan Data

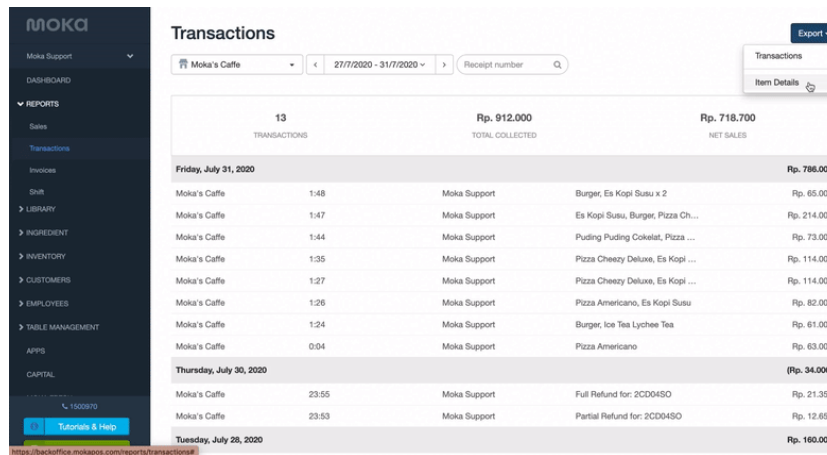
Data yang digunakan pada penelitian ini diperoleh melalui ekspor data penjualan yang tersedia dalam platform backoffice Moka POS. Informasi terkait penjualan dapat diakses melalui menu transaksi yang tertera di *sidebar* sebagaimana ditunjukkan dalam ilustrasi pada Gambar 4.1.

13		Rp. 912.000		Rp. 718.700	
TRANSACTIONS		TOTAL COLLECTED		NET SALES	
<b>Friday, July 31, 2020</b>					
Moka's Caffe	1:48	Moka Support	Burger, Es Kopi Susu x 2		Rp. 65.000
Moka's Caffe	1:47	Moka Support	Es Kopi Susu, Burger, Pizza Ch...		Rp. 214.000
Moka's Caffe	1:44	Moka Support	Puding Puding Cokelat, Pizza ...		Rp. 73.000
Moka's Caffe	1:35	Moka Support	Pizza Cheesy Deluxe, Es Kopi ...		Rp. 114.000
Moka's Caffe	1:27	Moka Support	Pizza Cheesy Deluxe, Es Kopi ...		Rp. 114.000
Moka's Caffe	1:26	Moka Support	Pizza Americano, Es Kopi Susu		Rp. 82.000
Moka's Caffe	1:24	Moka Support	Burger, Ice Tea Lychee Tea		Rp. 61.000
Moka's Caffe	0:04	Moka Support	Pizza Americano		Rp. 63.000
<b>Thursday, July 30, 2020</b>					
Moka's Caffe	23:55	Moka Support	Full Refund for: 2CD04SO		Rp. 21.350
Moka's Caffe	23:53	Moka Support	Partial Refund for: 2CD04SO		Rp. 12.650
<b>Tuesday, July 28, 2020</b>					
<b>Rp. 160.000</b>					

Gambar 4.1 Halaman Laporan Transaksi

Dalam menu laporan tersebut, terdapat opsi untuk memilih outlet yang diinginkan, menentukan jangka waktu data yang akan diekspor, dan memilih jenis data yang akan diekspor.

Gambar 4.2 mengilustrasikan langkah untuk melakukan ekspor data penjualan yang tersimpan dalam aplikasi Moka POS.



Gambar 4.2 Pilihan Ekspor Data

Hasilnya dari ekspor data ini adalah file csv yang memuat rincian penjualan item selama 2 periode bulan, yaitu bulan Juni dan Juli 2023, pada outlet Noble Coffee. Dataset ini menjadi dataset utama dalam pelaksanaan penelitian. Dataset penjualan Noble Coffee memuat 5.763 baris data yang berisikan jenis item yang terjual beserta rinciannya yang terletak pada 25 kolom berbeda. Detail data ini terlampir pada Gambar 4.3.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	AA	AB
1	Outlet	Receipt N Date	Time	Category	Brand	Items	Variant	SKU	Quantity	Modifier / Discount	Gross Sale	Discounts	Refunds	Net Sales	Gratuity	Tax	Sales Type	Collected	Served By	Customer	Payment	Event	Type				
2	Coffee by 3BPBV0	31/7/2023	23:21:47	Main Cour	Unbrande	Japanese Curry Katsu			1	Staff Appr	32000	11200	0	20800	0	2080	Dine In	Kasir Noble			Cash	Payment					
3	Coffee by 3BPBV0	31/7/2023	23:07:54	Salad	Unbrande	Signature:Thousand Island			1		26000	0	0	26000	0	2600	Dine In	Kasir Noble			GrabFood	Payment					
4	Coffee by 3BPBV0	31/7/2023	22:17:39	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			BCA QR	Payment					
5	Coffee by 3BPBV0	31/7/2023	22:17:39	Milk Base	Unbrande	Red Velvet Ice			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
6	Coffee by 3BPBV0	31/7/2023	22:16:46	Coffee Tw	Unbrande	Hazelnut Ice			1		28000	0	0	28000	0	2800	Dine In	Kasir Noble			Cash	Payment					
7	Coffee by 3BPBV0	31/7/2023	22:09:43	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			BCA QR	Payment					
8	Coffee by 3BPBV0	31/7/2023	21:56:34	Milk Base	Unbrande	Matcha Gr Ice			1		26000	0	0	26000	0	2600	Dine In	Kasir Noble			BCA QR	Payment					
9	Coffee by 3BPBV0	31/7/2023	21:56:34	Coffee MI	Unbrande	Ace			1		28000	0	0	28000	0	2800	Dine In	Kasir Noble			BCA QR	Payment					
10	Coffee by 3BPBV0	31/7/2023	21:54:44	Ricebowl	Unbrande	Crispy Chi BBQ			1	Disc	30000	3000	0	27000	0	2700	Dine In	Kasir Noble			BCA QR	Payment					
11	Coffee by 3BPBV0	31/7/2023	21:54:44	Coffee MI	Unbrande	Queen			1	Disc	24000	2400	0	21600	0	2160	Dine In	Kasir Noble			BCA QR	Payment					
12	Coffee by 3BPBV0	31/7/2023	21:54:44	White	Unbrande	Pinky no Bara			1	Disc	30000	3000	0	27000	0	2700	Dine In	Kasir Noble			BCA QR	Payment					
13	Coffee by 3BPBV0	31/7/2023	21:54:44	Coffee MI	Unbrande	Queen			2		48000	4800	0	43200	0	4320	Dine In	Kasir Noble			BCA QR	Payment					
14	Coffee by 3BPBV0	31/7/2023	21:53:29	Classic Co	Unbrande	American-Hot			1		22000	0	0	22000	0	2200	Dine In	Kasir Noble			Cash	Payment					
15	Coffee by 3BPBV0	31/7/2023	21:53:29	Snacks	Unbrande	Chicken Popporn			1		23000	0	0	23000	0	2300	Dine In	Kasir Noble			Cash	Payment					
16	Coffee by 3BPBV0	31/7/2023	21:44:43	Coffee MI	Unbrande	King			2		50000	0	0	50000	0	5000	Dine In	Kasir Noble			Cash	Payment					
17	Coffee by 3BPBV0	31/7/2023	21:08:13	White	Unbrande	Blooming in Tokyo			1		30000	0	0	30000	0	3000	Dine In	Kasir Noble			Cash	Payment					
18	Coffee by 3BPBV0	31/7/2023	21:00:17	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			Cash	Payment					
19	Coffee by 3BPBV0	31/7/2023	20:45:13	Sparkles	Unbrande	Lychee Addict			1		32000	0	0	32000	0	3200	Dine In	Kasir Noble			BCA QR	Payment					
20	Coffee by 3BPBV0	31/7/2023	19:27:48	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			BCA QR	Payment					
21	Coffee by 3BPBV0	31/7/2023	19:36:46	Tea	Unbrande	Strawberry Tea			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
22	Coffee by 3BPBV0	31/7/2023	19:34:39	Classic Co	Unbrande	Cafe Latte Ice			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
23	Coffee by 3BPBV0	31/7/2023	19:34:39	Coffee MI	Unbrande	King			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
24	Coffee by 3BPBV0	31/7/2023	19:21:47	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			BCA QR	Payment					
25	Coffee by 3BPBV0	31/7/2023	18:27:28	Tea	Unbrande	Lychee Tea			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
26	Coffee by 3BPBV0	31/7/2023	18:27:28	Snacks	Unbrande	Onion Ring			1		20000	0	0	20000	0	2000	Dine In	Kasir Noble			BCA QR	Payment					
27	Coffee by 3BPBV0	31/7/2023	18:20:26	Pasta	Unbrande	Aglio Olio Spaghetti			1		30000	0	0	30000	0	3000	Dine In	Kasir Noble			BCA QR	Payment					
28	Coffee by 3BPBV0	31/7/2023	18:20:26	Classic Co	Unbrande	American Ice			1		22000	0	0	22000	0	2200	Dine In	Kasir Noble			BCA QR	Payment					
29	Coffee by 3BPBV0	31/7/2023	16:52:18	Sparkles	Unbrande	Kyoto Sunrise			1		32000	0	0	32000	0	3200	Dine In	Kasir Noble			BCA QR	Payment					
30	Coffee by 3BPBV0	31/7/2023	16:52:18	Main Cour	Unbrande	Japanese Curry Katsu			1		32000	0	0	32000	0	3200	Dine In	Kasir Noble			BCA QR	Payment					
31	Coffee by 3BPBV0	31/7/2023	16:45:06	Coffee MI	Unbrande	Queen			1		24000	0	0	24000	0	2400	Dine In	Kasir Noble			BCA QR	Payment					
32	Coffee by 3BPBV0	31/7/2023	16:45:06	Coffee MI	Unbrande	King			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					
33	Coffee by 3BPBV0	31/7/2023	16:45:06	Ricebowl	Unbrande	Crispy Chi Mushroom			1		30000	0	0	30000	0	3000	Dine In	Kasir Noble			BCA QR	Payment					
34	Coffee by 3BPBV0	31/7/2023	16:45:06	Main Cour	Unbrande	Hongkong Fried Rice			1		30000	0	0	30000	0	3000	Dine In	Kasir Noble			BCA QR	Payment					
35	Coffee by 3BPBV0	31/7/2023	15:49:49	Cold Bottl	Unbrande	Mineral Water			1		8000	0	0	8000	0	800	Dine In	Kasir Noble			Cash	Payment					
36	Coffee by 3BPBV0	31/7/2023	15:49:49	Tea	Unbrande	Lychee Tea			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			Cash	Payment					
37	Coffee by 3BPBV0	31/7/2023	14:31:54	Tea	Unbrande	Strawberry Tea			1		25000	0	0	25000	0	2500	Dine In	Kasir Noble			BCA QR	Payment					

Gambar 4.3 Dataset csv

Setiap kolom dalam data tersebut memiliki fungsi dan makna tersendiri yang dijelaskan secara terperinci pada Tabel 4.1.

Tabel 4.1 Penjelasan Fungsi Kolom

Kolom	Deskripsi
Outlet	Menjelaskan nama cabang toko
Receipt Number	Menjelaskan nomor resi pembelian
Date	Menjelaskan tanggal pembelian

Time	Menjelaskan waktu pembelian
Category	Menjelaskan kategori produk
Brand	Menjelaskan jenis kelompok produk
Items	Menjelaskan nama produk
Variant	Menjelaskan tipe produk
SKU	Menjelaskan barkode produk
Quantity	Menjelaskan jumlah produk terbeli
Modifier Applied	Menjelaskan permintaan tambahan produk
Discount Applied	Menjelaskan jenis diskon terpakai
Gross Sales	Menjelaskan penjualan kotor
Discounts	Menjelaskan nominal diskon
Refunds	Menjelaskan nominal retur
Net Sales	Menjelaskan nominal pendapatan bersih
Gratuity	Menjelaskan nominal biaya tambahan
Tax	Menjelaskan nominal pajak
Sales Type	Menjelaskan tipe pembelian
Collected By	Menjelaskan nama device yang digunakan
Served By	Menjelaskan nama penjaga kasir
Customer	Menjelaskan profil pembeli
Payment Method	Menjelaskan metode pembayaran
Event Type	Menjelaskan tipe kejadian (pembelian / retur)
Reason of Refund	Menjelaskan alasan retur

#### 4.2 Pembersihan Data

Proses pembersihan data dilakukan untuk memastikan bahwa data yang digunakan dalam analisis tersebut akurat, konsisten, dan bebas dari kesalahan. Hasil dari pengambilan data melalui ekspor data dari aplikasi Moka POS memiliki beberapa kejanggalan, seperti format tanggal yang tidak sesuai, kolom yang kosong, dan data yang tidak sesuai dengan batasan penelitian. Oleh karena itu, tahapan pembersihan data diperlukan untuk menjadikan dataset lebih konsisten dan memiliki dasar yang kuat untuk dilakukan analisis. Contoh dari inkonsistensi data yang telah disebutkan diatas terdapat pada Gambar 4.4.



	C
1	Date
2	31/7/2023
3	31/7/2023
4	31/7/2023
5	31/7/2023
6	31/7/2023
7	31/7/2023
64	30-07-2023
65	30-07-2023
66	30-07-2023
67	30-07-2023
68	30-07-2023
69	30-07-2023

Gambar 4.4 Inkonsistensi Penulisan Tanggal

Sedangkan contoh dari data yang tidak memiliki nilai ada pada Gambar 4.5 dan yang memiliki nilai 0 saja ada pada Gambar 4.6.

	I	U	V
1	SKU	Served By	Customer
2			
3			
4			
5			
6			
7			
8			
9			
10			
11			
12			
13			
14			

Gambar 4.5 Kolom Kosong

Q
Gratuity
0
0
0
0
0
0
0
0
0
0
0
0
0
0

Gambar 4.6 Kolom Nilai 0

Pembersihan data dapat langsung menggunakan python pyspark dikarenakan setelah data dibersihkan akan langsung dilaksanakan implementasi algoritma FP-Growth. Dengan menggunakan bantuan pyspark dataset dapat diolah dalam bentuk dataframe dengan proses pembersihan data yang tertera pada Gambar 4.7.

```
from pyspark.sql import functions as F

# Buat dataframe dari dataset csv
df = spark\
    .read\
    .format('csv')\
    .option("inferSchema","true")\
    .option("header","true")\
    .load("Sales Data.csv")

#Pisahkan kolom date dengan delimiter '/' dan '-'
split_regex = '[-/]+'
split_col = F.split(df['Date'], split_regex)

#Tempatkan pada variabel baru
df = df.withColumn('Day', split_col.getItem(0)) \
    .withColumn('Month', split_col.getItem(1)) \
    .withColumn('Year', split_col.getItem(2))

#Satukan lagi dengan urutan MM/dd/yyyy
df = df.withColumn(
    'Date', F.concat(F.col("Month"), F.lit("/"), F.col("Day"), F.lit("/"),
    F.col("Year"))
)

#Pastikan kolom Date memiliki format yang benar
df = df.withColumn("Date", F.to_date(F.col("Date"), "M/d/y"))

#Pastikan kolom Time memiliki format yang benar
df = df.withColumn('Time', F.date_format('Time', 'HH:mm:ss'))

#Hapus kolom yang tidak digunakan
df = df.drop("SKU", "Served By", "Customer", "Gratuity", "Day", "Month",
"Year")

#Filter data agar memenuhi event type pembelian
df = df.where(F.col("Event Type") == "Payment")
```

```
#Cek total baris data setelah dibersihkan
print(df.count())

#Export menjadi csv dengan nama Clean datasales
cleaned = df.select("*").toPandas()
cleaned.to_csv('Clean Datasales.csv')
```

Gambar 4.7 Kode Pembersihan Data

Setelah data dilakukan pembersihan, terdapat beberapa baris menghilang dikarenakan sesuai dengan batasan penelitian yaitu data yang digunakan hanya bersifat data pembelian, bukan data refund. Total baris data setelah dilakukan pembersihan menjadi 5.719 yang mana sebelumnya total data sesaat setelah diekspor dari Moka POS berjumlah 5.763 data. Kolom yang tidak digunakan dan dilakukan pembersihan data mengurangi total kolom data menjadi 21 kolom dengan menghilangkan kolom SKU, Served By, Customer, dan Gratuity.

### 4.3 Implementasi FP-Growth

Setelah menyelesaikan tahap pembersihan data, langkah selanjutnya adalah melakukan implementasi algoritma FP-Growth. Pada tahap ini kita mencari aturan asosiasi yang dapat digunakan untuk mendapatkan itemset yang sering muncul dari perilaku pembelian pelanggan. Aturan asosiasi yang dicari meliputi aturan asosiasi dari produk dan kategori. Tahapan dari implementasi algoritma ini dimulai dari pengelompokan itemset, pencarian frekuensi kemunculan itemset, dan pembentukan aturan asosiasi.

#### 4.3.1 Pengelompokan Itemset

Tahapan yang dilakukan pada saat mengelompokkan produk dan kategori menjadi itemset dilakukan dengan memberikan id baru berdasarkan kolom Receipt Number. Id baru tersebut nantinya dilakukan pengelompokkan lagi untuk membuat kolom baru yang mengelompokkan antara Items dan Category dengan id yang sama. Kode eksekusi dari pengelompokan itemset dapat dilihat dari Gambar 4.8.

```
from pyspark.sql.window import Window

#Pendeclarasian id baru dari Receipt Number yang sama
itemset = df.withColumn("id", F.dense_rank()\
                        .over(Window.orderBy(F.col("Receipt\
Number")))).sort("id")

#Menghapus id duplikat, lalu
byCategory = itemset.dropDuplicates(['id', 'Category']).sort('id')
```

```

byItems = itemset.dropDuplicates(['id', 'Items']).sort('id')

#Mengelompokkan itemset berdasarkan nama item dan kategorinya
byCategory = byCategory.groupBy("id")\
    .agg(F.collect_list("Category")).sort(F.col('id').asc())
byItems = byItems.groupBy("id")\
    .agg(F.collect_list("Items")).sort(F.col('id').asc())

#Menampilkan itemset
byCategory.show(10)
byItems.show(10)

```

Gambar 4.8 Kode Pengelompokan Itemset

### 4.3.2 Pencarian Frekuensi Kemunculan Itemset

Tahapan implemmentasi algoritma FP-Growth ini menghasilkan data tabel berupa frekuensi kemunculan itemset dalam keseluruhan transaksi. Batasan yang diperlukan dalam mencari frekuensi kemunculan itemset adalah minimum support yaitu seberapa sering itemset muncul, dan minimum confidence yaitu seberapa benar pasangan itemset tersebut benar muncul. Dalam penelitian ini dikarenakan jumlah dataset yang besar, maka minimum support diberikan pada angka 0.001 dengan pertimbangan akan banyak itemset yang tidak sering muncul dalam transaksi akan didapatkan. Kekurangan ini nantinya akan ditutup dengan menentukan minimum confidence, yaitu pada angka 0.5 untuk memastikan bahwa kebenaran kemunculan pasangan itemset ini bernilai 50%. Berikut baris kode untuk mencari frekuensi kemunculan itemset tertera pada Gambar 4.9

```

from pyspark.ml.fpm import FPGrowth

#Penentuan minimum support dan confidence
minSup = 0.001
minConf = 0.5

#Algoritma FP-Growth yang telah tersedia dalam salah satu library machine
learning oleh pyspark
fpCategory = FPGrowth(itemsCol="collect_list(Category)", \
                      minSupport=minSup, minConfidence=minConf)
modelCategory = fpCategory.fit(byCategory)

fpItems = FPGrowth(itemsCol="collect_list(Items)", \
                  minSupport=minSup, minConfidence=minConf)
modelItems = fpItems.fit(byItems)

#Menyimpan hasil temuan pada variabel baru
category = modelCategory.freqItemsets
items = modelItems.freqItemsets

#Menampilkan frekuensi kemunculan itemset dengan frekuensi tertinggi
category.sort(F.col("freq").desc()).show()
items.sort(F.col("freq").desc()).show()

```

Gambar 4.9 Kode Pencarian Frekuensi Kemunculan Itemset

Setelah Itemset dikelompokkan, didapatkan fakta bahwa produk yang memiliki frekuensi pada itemset tertinggi ada pada produk bernama King, dan kategori tertinggi yaitu *Coffee Milk*. Berikut adalah hasil dari pencarian frekuensi kemunculan itemset dari kode diatas yang tertera pada Gambar 4.10

items freq	items freq
[Coffee Milk] 1201	[King]  658
[Milk Based]  717	[Queen]  463
[Classic Coffee]  547	[Americano]  322
[Snacks]  421	[Mineral Water]  302
[Tea]  373	[Chocolate]  251
[Cold Bottle]  302	[Red Velvet]  225
[Frappe]  246	[Cafe Latte]  196
[Main Course]  219	[Ace]  186
[Ricebowl]  206	[Matcha Greentea]  177
[Sparkles]  197	[Lychee Tea]  174
[White]  191	[French Fries]  146
[Coffee Twist]  166	[Sunny Taro]  143
[Snacks, Coffee M...]  142	[Crispy Chicken R...]  132
[Milk Based, Coff...]  128	[Black Tea]  100
[Pasta]  121	[Mix Platter]  89
[Mocktails]  109	[Japanese Curry K...]  87
[Classic Coffee, ...]  88	[Garlic Beef Rice...]  83
[Snacks, Milk Based]  87	[Hongkong Fried R...]  83
[Snacks, Classic ...]  69	[Chicken Popcorn]  81
[Manual Brew]  68	[Pinky no Bara]  75

only showing top 20 rows

Gambar 4.10 Frekuensi Kemunculan Itemset

### 4.3.3 Pembentukan Aturan Asosiasi

Dengan implementasi library algoritma diatas, aturan asosiasi bisa langsung didapatkan dengan mengakses `associationRules` seperti pada Gambar 4.11 dibawah.

```
#Menyimpan aturan asosiasi dari kategori dan items pada variabel baru
rulesCategory = modelCategory.associationRules
rulesItems = modelItems.associationRules

#Menampilkan aturan asosiasi tersebut
rulesCategory.show()
rulesItems.show()

#Export menjadi csv
rulesC = rulesCategory.select("*").toPandas()
rulesI = rulesItems.select("*").toPandas()

rulesC.to_csv('Category Rules.csv')
rulesI.to_csv('Items Rules.csv')
```

Gambar 4.11 Kode Pembentukan Aturan Asosiasi

Aturan asosiasi yang telah didapatkan nantinya akan divisualisasikan menggunakan aplikasi intelijen bisnis Power BI. Nantinya dengan menggunakan pengetahuan akan aturan asosiasi ini akan didapatkan wawasan berupa hubungan asosiasi pembelian produk oleh pelanggan yang terjadi selama 2 bulan di Noble Coffee. Berikut adalah hasil dari pembentukan aturan asosiasi dari kode diatas yang disertakan pada Gambar 4.12

antecedent	consequent	confidence	lift	support
[Frappe, Classic ...]	[Snacks]	0.5	4.065320665083135	0.001752848378615...
[White, Cold Bott...]	[Snacks]	0.8333333333333334	6.775534441805226	0.001460706982179375
[White, Cold Bott...]	[Classic Coffee]	0.8333333333333334	5.214008043875686	0.001460706982179375
[Main Course, Col...]	[Coffee Milk]	0.6666666666666666	1.900083263946711	0.001168565585743...
[White, Classic C...]	[Cold Bottle]	0.75	8.500827814569536	0.001752848378615...
[White, Classic C...]	[Snacks]	0.75	6.097980997624703	0.001752848378615...
[White, Cold Bott...]	[Coffee Milk]	0.75	2.137593671940085	0.001752848378615...
[White, Cold Bott...]	[Snacks]	0.75	6.097980997624703	0.001752848378615...
[White, Milk Base...]	[Frappe]	0.8	11.13170731707317	0.001168565585743...
[White, Snacks, C...]	[Cold Bottle]	0.5	5.667218543046357	0.001752848378615...
[White, Snacks, C...]	[Classic Coffee]	0.5	3.1288848263254114	0.001752848378615...
[White, Cold Bott...]	[Classic Coffee]	0.5454545454545454	3.413328901445903	0.001752848378615...
[White, Cold Bott...]	[Coffee Milk]	0.5454545454545454	1.5546135795927636	0.001752848378615...
[White, Cold Bottle]	[Classic Coffee]	0.5	3.1288848263254114	0.00237131171486...
[White, Cold Bottle]	[Snacks]	0.6875	5.589815914489311	0.003213555360794...
[Frappe, Milk Bas...]	[White]	0.5	8.960732984293193	0.001168565585743...
[White, Frappe, M...]	[Coffee Milk]	0.8	2.2800999167360536	0.001168565585743...
[White, Frappe, C...]	[Snacks]	0.8	6.504513064133016	0.001168565585743...
[White, Snacks, C...]	[Cold Bottle]	0.75	8.500827814569536	0.001752848378615...
[White, Snacks, C...]	[Coffee Milk]	0.75	2.137593671940085	0.001752848378615...

only showing top 20 rows

antecedent	consequent	confidence	lift	support
[Alpine Breeze, F...]	[Ace]	0.5555555555555556	10.224014336917563	0.001460706982179375
[Mix Platter, Caf...]	[Mineral Water]	0.8	9.067549668874172	0.001168565585743...
[Alpine Breeze, Ace]	[French Fries]	0.8333333333333334	19.537671232876715	0.001460706982179375
[Brown Brulee, Mi...]	[Mix Platter]	0.5714285714285714	21.977528089887638	0.001168565585743...
[Brown Brulee, Mi...]	[Cafe Latte]	0.5714285714285714	9.979591836734693	0.001168565585743...
[Cireng, French F...]	[King]	0.5	2.6010638297872344	0.001168565585743...
[Cireng, Chocolate]	[King]	0.5714285714285714	2.972644376899696	0.001168565585743...
[Mix Platter, Caf...]	[Queen]	0.5	3.696544276457883	0.001460706982179375
[Mix Platter, Caf...]	[Mineral Water]	0.6	6.800662251655629	0.001752848378615...
[Mix Platter, Caf...]	[King]	0.5	2.6010638297872344	0.001460706982179375
[Mix Platter, Caf...]	[Queen]	0.6666666666666666	4.928725701943844	0.001168565585743...
[Mix Platter, Caf...]	[King]	0.6666666666666666	3.4680851063829787	0.001168565585743...
[Chicken Popcorn, ...]	[Mineral Water]	1.0	11.334437086092715	0.001168565585743...
[Mix Platter, Caf...]	[Mineral Water]	0.8	9.067549668874172	0.001168565585743...
[Cireng, Mix Plat...]	[King]	0.5	2.6010638297872344	0.001168565585743...
[Brown Brulee, Ca...]	[Mineral Water]	0.6666666666666666	7.556291390728476	0.001168565585743...
[Brown Brulee, Mi...]	[Mineral Water]	1.0	11.334437086092715	0.001168565585743...
[Kyoto Sunrise, J...]	[Mineral Water]	0.5	5.667218543046357	0.001168565585743...
[Mix Platter, Min...]	[Cafe Latte]	0.5	8.732142857142858	0.001168565585743...
[Cafe Latte, Mine...]	[Mix Platter]	0.8	30.768539325842696	0.001168565585743...

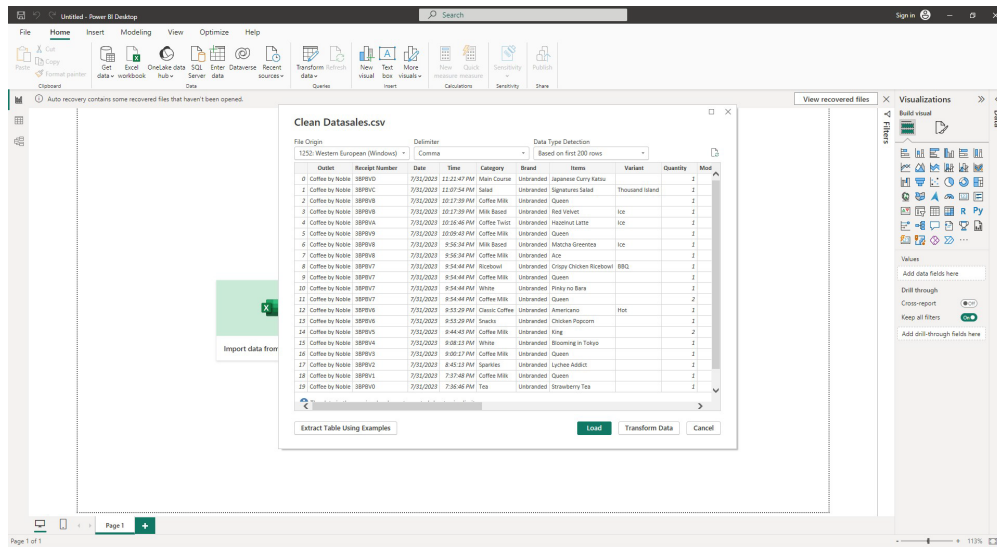
only showing top 20 rows

Gambar 4.12 Aturan Asosiasi

#### 4.4 ETL

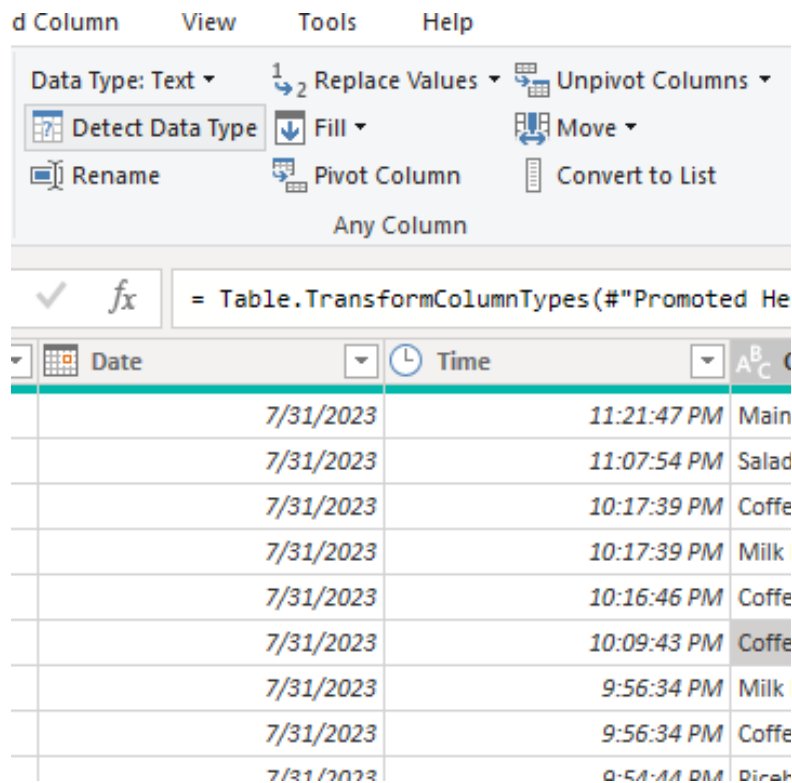
ETL atau extract, transform, load merupakan salah satu tahapan sebelum dapat dilakukan visualisasi data. Untuk mempersiapkan data agar dapat dianalisis visualisasinya secara maksimal, perlu untuk menghubungkan data utama dengan aturan asosiasi yang telah didapatkan dari implementasi algoritma FP-Growth serta membangun relasi antar items dan category.

Pertama tama aplikasi Power BI dibuka terlebih dahulu, setelah itu dilanjutkan dengan memilih interface *get data*. Setelah itu pilih dataset, dengan melakukan *connect* dengan pilihan 'text/csv', yang telah dibersihkan sebelumnya dengan menggunakan baris kode pypark. Gambar 4.13 menampilkan tampilan yang muncul setelah dataset telah dibuka lalu akan dihadirkan dengan pilihan load atau transform.



Gambar 4.13 Connect Data

Langkah selanjutnya adalah memilih *Transform Data* untuk memastikan dataset yang digunakan siap untuk divisualisasikan. Untuk tetap memastikan bahwa dataset dalam penelitian ini bersifat historis maka perlu dipastikan bahwa kolom *Date* memiliki jenis data 'Date' dan *Time* memiliki jenis data 'Time' pada menu *Transform*. Hal ini dapat dipastikan dengan melihat symbol disamping *header* kolom serta melihat *data type* pada Gambar 4.14.



Gambar 4.14 Tipe Data Pada Kolom



Selanjutnya adalah aktivasi dataset agar masuk kedalam pilihan data Power BI. Setelah dataset utama telah diketahui oleh Power BI yang selanjutnya dilakukan adalah memasukkan data aturan asosiasi yang telah didapatkan sebelumnya. Dengan melakukan langkah yang sama yaitu melakukan transform setelah memilih koneksi data, kolom konsekuen diduplikat dan diubah *value*-nya agar dapat dilakukan relasi antar tabel. Untuk menduplikat dan merubah *value* dapat dilakukan pada *interface* transform Power BI dalam Power Query dengan klik kanan nama kolom dan pilih *duplicate column* seperti pada Gambar 4.15.

Table.DuplicateColumn("#Changed Type", "consequent", "consequent - Copy")

	A <sub>0</sub> antecedent	A <sub>0</sub> consequent	1.2 confidence	1.2 lift	1.2 support	A <sub>0</sub> consequent - Copy
0	['Alpine Breeze', 'French Fries']	['Ace']	0.555555556	10.22401434	0.001460707	['Ace']
1	['Mix Platter', 'Cafe Latte', 'Queen']	['Mineral Water']	0.8	9.067549669	0.001168566	['Mineral Water']
2	['Alpine Breeze', 'Ace']	['French Fries']	0.833333333	19.53767123	0.001460707	['French Fries']
3	['Brown Brulee', 'Mineral Water']	['Mix Platter']	0.571428571	21.97752809	0.001168566	['Mix Platter']
4	['Brown Brulee', 'Mineral Water']	['Cafe Latte']	0.571428571	9.97959184	0.001168566	['Cafe Latte']
5	['Cireng', 'French Fries']	['King']	0.5	2.60106383	0.001168566	['King']
6	['Cireng', 'Chocolate']	['King']	0.571428571	2.972644377	0.001168566	['King']
7	['Mix Platter', 'Cafe Latte']	['Queen']	0.5	3.696544276	0.001460707	['Queen']
8	['Mix Platter', 'Cafe Latte']	['Mineral Water']	0.6	6.800662252	0.001752848	['Mineral Water']
9	['Mix Platter', 'Cafe Latte']	['King']	0.5	2.60106383	0.001460707	['King']
10	['Mix Platter', 'Cafe Latte', 'Mineral Water']	['Queen']	0.666666667	4.928725702	0.001168566	['Queen']
11	['Mix Platter', 'Cafe Latte', 'Mineral Water']	['King']	0.666666667	3.468085106	0.001168566	['King']
12	['Chicken Popcorn', 'Cafe Latte']	['Mineral Water']	1	11.33443709	0.001168566	['Mineral Water']
13	['Mix Platter', 'Cafe Latte', 'King']	['Mineral Water']	0.8	9.067549669	0.001168566	['Mineral Water']
14	['Cireng', 'Mix Platter']	['King']	0.5	2.60106383	0.001168566	['King']
15	['Brown Brulee', 'Cafe Latte']	['Mineral Water']	0.666666667	7.556291391	0.001168566	['Mineral Water']
16	['Brown Brulee', 'Mix Platter']	['Mineral Water']	1	11.33443709	0.001168566	['Mineral Water']
17	['Kyoto Sunrise', 'Japanese Curry Katsu']	['Mineral Water']	0.5	5.667218543	0.001168566	['Mineral Water']
18	['Mix Platter', 'Mineral Water', 'King']	['Cafe Latte']	0.5	8.732142857	0.001168566	['Cafe Latte']
19	['Cafe Latte', 'Mineral Water', 'Queen']	['Mix Platter']	0.8	30.76853933	0.001168566	['Mix Platter']
20	['Cafe Latte', 'Mineral Water', 'King']	['Mix Platter']	0.571428571	21.97752809	0.001168566	['Mix Platter']
21	['Chicken Popcorn', 'Ace']	['French Fries']	0.666666667	15.63013699	0.001168566	['French Fries']
22	['Nasi Kulit Sambal Bawang', 'Cafe Latte']	['Mineral Water']	0.666666667	7.556291391	0.001168566	['Mineral Water']
23	['Mix Platter', 'French Fries']	['Mineral Water']	0.666666667	7.556291391	0.001168566	['Mineral Water']
24	['Mix Platter', 'French Fries']	['King']	0.666666667	3.468085106	0.001168566	['King']
25	['Kyoto Sunrise', 'Mineral Water']	['Japanese Curry Katsu']	0.8	31.47586207	0.001168566	['Japanese Curry Katsu']
26	['Mix Platter', 'Mineral Water', 'Queen']	['Cafe Latte']	0.8	13.97142857	0.001168566	['Cafe Latte']

Gambar 4.15 Duplikat Kolom Konsekuen

Selanjutnya setelah kolom konsekuen telah terduplikasi, dilakukan rename nama kolom agar memudahkan dalam mengakses data dengan klik kanan kolom konsekuen duplikat lalu pilih '*Rename*' dan nama kolom diganti menjadi 'item'. Setelah itu *value* dari kolom item diselaraskan dengan apa yang ada pada kolom *Items* dataset utama, dengan melakukan *replace value* pada pilihan transformasi kolom setelah klik kanan. Hal yang perlu diubah adalah “[” dan “]” menjadi kosong. Hasil yang didapatkan adalah kolom item pada data aturan asosiasi yang memiliki *value* sama dengan kolom *Items* pada dataset utama. Gambar 4.16 adalah hasil setelah dilakukan transformasi pada data aturan asosiasi.

item
Ace
Mineral Water
French Fries
Mix Platter
Cafe Latte
King
King
Queen
Mineral Water

Gambar 4.16 Transformasi Data Aturan Asosiasi Item

Hal yang sama dilakukan kepada data aturan asosiasi kategori yang didapatkan sebelumnya. Setelah proses transformasi usai, maka data siap digunakan, akan tetapi data perlu diberi hubungan antar tabel dengan memilih pilihan *Manage Relationship* pada fitur Power BI yaitu *Model View*. Model View memungkinkan penggunaanya untuk mendapatkan visualisasi tabel serta hubungannya. Fitur *Manage Relationship* digunakan untuk menentukan hubungan antar tabel untuk nantinya dapat digunakan sebagai pembantu visualisasi data agar lebih interaktif. Saat mengatur hubungan antar tabel, perlu dilakukan pemilihan untuk memperjelas kolom mana yang saling memiliki hubungan seperti pada Gambar 4.17.

Create relationship

Select tables and columns that are related.

Clean Datasales

Outlet	Receipt Number	Date	Time	Category	Brand	Items	Variant
Coffee by Noble	3BPBUJ	Monday, July 31, 2023	4:45:06 PM	Coffee Milk	Unbranded	King	
Coffee by Noble	3BPBUS	Monday, July 31, 2023	2:31:54 PM	Coffee Milk	Unbranded	King	
Coffee by Noble	3BPBU1	Sunday, July 30, 2023	10:33:00 PM	Coffee Milk	Unbranded	King	

Items Rules

antecedent	consequent	confidence	lift	support	Item
reeze', 'French Fries]	['Ace]	0.555555555555556	10.2240143369176	0.00146070698217937	Ace
ter', 'Cafe Latte', 'Queen]	['Mineral Water]	0.8	9.06754966887417	0.0011685655857435	Mineral Water
reeze', 'Ace]	['French Fries]	0.833333333333333	19.5376712328767	0.00146070698217937	French Fries

Cardinality: Many to many (\*)

Cross filter direction: Both

Make this relationship active

Assume referential integrity

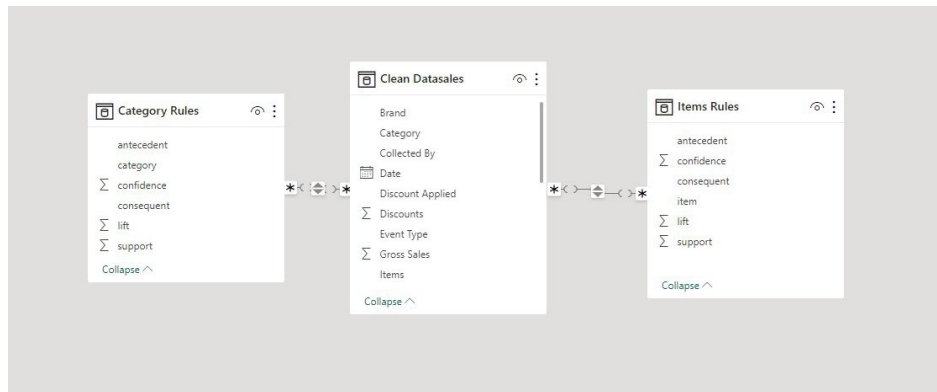
Apply security filter in both directions

This relationship has cardinality Many-Many. This should only be used if it is expected that neither column (Items and Item) contains unique values, and that the significantly different behavior of Many-many relationships is understood. [Learn more](#)

OK Cancel

Gambar 4.17 Relasi Antar Tabel

Hal yang sama berlaku untuk membuat relasi antar tabel dari dataset dengan aturan asosiasi kategori. Dengan melakukan tahapan yang sama, didapatkan model data yang berbentuk seperti pada Gambar 4.18 dengan garis hubungannya.



Gambar 4.18 Hubungan Antar Tabel

## 4.5 Visualisasi Data

### 4.5.1 Visualisasi

Dalam kerangka penelitian ini, visualisasi data disajikan melalui tiga halaman berbeda, masing-masing memberikan wawasan yang mendalam tentang berbagai aspek bisnis. Pada halaman pertama, terdapat sejumlah visualisasi yang merangkum analisis keseluruhan bisnis. Data-data ini memberikan gambaran menyeluruh tentang performa bisnis, menggambarkan pola penjualan, tren harian, serta variasi dalam kategori produk.

Halaman kedua dari visualisasi data ini didedikasikan untuk memberikan pemahaman yang lebih mendalam mengenai minuman yang ditawarkan oleh bisnis ini. Visualisasi yang disajikan di halaman ini mengungkapkan detail-detail spesifik tentang perilaku transaksi pelanggan terhadap minuman, mencakup informasi mengenai produk-produk minuman yang paling populer, pola konsumsi harian, hubungan asosiasi yang terjadi dan perubahan perilaku pembelian selama dua bulan terakhir.

Sementara itu, halaman ketiga dari visualisasi data ini memfokuskan diri pada aspek makanan dalam bisnis tersebut. Visualisasi pada halaman ini memberikan gambaran terperinci tentang produk-produk makanan yang paling diminati oleh pelanggan, hubungan asosiasi pembelian makanan oleh pelanggan, serta variasi penjualan dalam kategori makanan. Data-data ini memberikan gambaran menyeluruh tentang sejauh mana produk-produk makanan ini menarik minat pelanggan dan bagaimana perilaku pembelian produk oleh pelanggan berubah seiring waktu.

Dengan membagi visualisasi data menjadi tiga halaman terpisah, penelitian ini menyajikan gambaran yang komprehensif dan terperinci tentang bisnis tersebut. Pendekatan ini memungkinkan pembaca untuk menggali informasi dengan lebih terfokus, membantu dalam identifikasi tren-tren spesifik yang dapat membimbing keputusan strategis bisnis. Dengan menggabungkan analisis keseluruhan dengan pemahaman yang mendalam tentang penjualan

minuman dan makanan, bisnis memiliki kerangka kerja yang kuat untuk merumuskan strategi pemasaran yang lebih cerdas, meningkatkan kepuasan pelanggan, dan meraih kesuksesan dalam industri yang kompetitif.

Pada halaman pertama, visualisasi akan lebih difokuskan kepada pemahaman mendalam tentang pilihan pelanggan terhadap berbagai menu yang tersedia di Noble Coffee. Analisis ini akan didasarkan pada data yang telah dibersihkan dan dikelola dengan cermat. Salah satu visualisasi data yang pertama adalah mengenai wawasan jumlah pendapatan kotor selama 2 bulan terakhir, yang disajikan dengan menggunakan card visual. Visualisasi ini memiliki keunggulan dalam mempermudah pembacaan laporan keuangan, karena dalam konteks bisnis, informasi yang paling dicari adalah seberapa besar pendapatan yang berhasil dihasilkan.

Oleh karena itu, Gambar 4.19 memberikan gambaran visual pertama untuk menjawab pertanyaan inti dalam bisnis ini. Analisis yang teliti terhadap data ini memberikan pemahaman mendalam tentang sejauh mana pilihan-pilihan menu yang ditawarkan oleh Noble Coffee berkontribusi terhadap pendapatan keseluruhan. Dengan memahami perilaku pembelian pelanggan terhadap menu-menu tertentu, bisnis dapat mengambil langkah-langkah strategis untuk meningkatkan penjualan, mengoptimalkan persediaan, dan bahkan mengembangkan menu-menu baru yang sesuai dengan keinginan pelanggan.

Dengan demikian, melalui visualisasi data yang teliti dan relevan, bisnis dapat menggali informasi berharga yang membantu mereka mengarahkan kebijakan dan strategi bisnis mereka ke arah yang lebih menguntungkan. Analisis ini bukan hanya sekadar memberikan gambaran tentang pendapatan, tetapi juga merupakan langkah awal yang sangat penting dalam memahami perilaku pelanggan dan menghasilkan rekomendasi yang konkret untuk meningkatkan pengalaman pelanggan dan memperluas pangsa pasar.



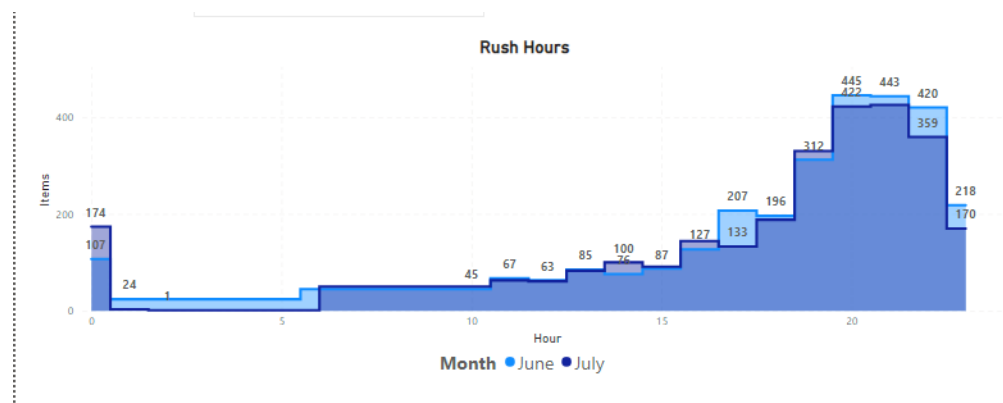
Gambar 4.19 Card

Dalam analisis data tersebut, tidak hanya memberikan wawasan berupa seberapa besar pendapatan dalam 2 bulan terakhir, card disampingnya memvisualisasikan jumlah item terjual dalam periode waktu tersebut yaitu sejumlah 6074 produk terjual. Visualisasi selanjutnya adalah area chart yang menjelaskan jam sibuk dimana Noble Coffee ramai kedatangan pengunjung.

Jam sibuk atau *Rush Hour* dapat menjadi patokan terkait penjualan yang memiliki kuantitas terbanyak tersebar pada jam-jam tertentu. Seperti pada Gambar 4.20, grafik ini dapat memberikan insight berupa kenaikan jumlah pembelian produk oleh pelanggan dalam jam-jam tertentu. Dalam hal ini, pukul 20.00 dan 21.00 adalah contoh jam dengan jumlah pembeli tertinggi. Hal ini terlihat dari data yang dipaparkan, menunjukkan bahwa 2 jam tersebut merupakan jam dengan jumlah pembeli tertinggi.

Namun, tidak hanya itu yang dapat disimpulkan dari visualisasi tersebut. Analisis juga menyoroti jam yang memiliki penjualan paling rendah, yaitu pada pagi hari. Selama bulan Juni dan Juli, terdapat hanya 45 penjualan pada pukul 10 pagi. Hal ini menunjukkan bahwa waktu tersebut bukanlah waktu yang efektif untuk menargetkan pelanggan. Oleh karena itu, untuk meningkatkan penjualan, bisa dipertimbangkan untuk mengoptimalkan strategi pemasaran pada jam-jam tersebut.

Dengan demikian, analisis visual tersebut memberikan pemahaman mendalam tentang perilaku pembelian pelanggan sepanjang hari. Dari hasil ini, pemilik bisnis dapat membuat keputusan yang lebih cerdas, seperti mengatur staf sesuai dengan jam-jam sibuk atau mengadakan promosi khusus pada jam-jam dengan penjualan rendah untuk meningkatkan minat pelanggan maupun mengadakan promosi pada jam-jam dengan penjualan yang tinggi untuk mendapatkan pendapatan yang lebih.



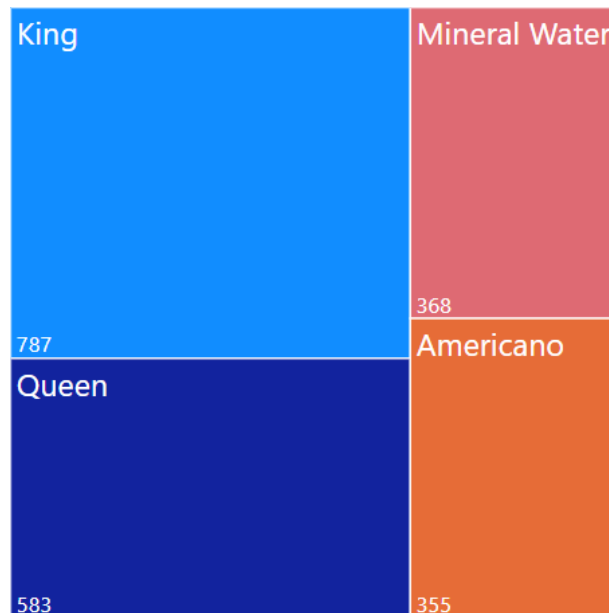
Gambar 4.20 Jam Sibuk

Dalam mengolah data yang divisualisasikan, penting untuk menyediakan informasi perbandingan yang jelas sehingga dapat memudahkan analisis dan pemahaman. Salah satu cara yang efektif adalah dengan menambahkan legenda data yang memungkinkan komparasi antar elemen. Contohnya, pada grafik jam sibuk di atas, bulan Juni direpresentasikan dengan warna biru muda sementara bulan Juli diwakili oleh warna biru tua. Dengan menggunakan perbedaan

warna ini, pembaca dapat dengan mudah membandingkan pola jam sibuk di kedua bulan tersebut.

Setelah memahami pola jam sibuk, langkah selanjutnya adalah mengidentifikasi produk yang mencatat penjualan tertinggi selama periode dua bulan tersebut. Untuk tujuan ini, tree map digunakan sebagai grafik yang memvisualisasikan data dengan memperlihatkan perbedaan ukuran area. Dalam konteks visualisasi data ini, tree map digunakan untuk memberikan wawasan mendalam mengenai produk-produk yang memiliki penjualan tertinggi. Grafik ini, seperti yang terlihat pada Gambar 4.21, memberikan representasi visual yang jelas tentang distribusi penjualan produk, memungkinkan pembaca untuk mengidentifikasi dengan mudah produk-produk yang paling diminati oleh pelanggan selama dua bulan tersebut.

Dengan memanfaatkan teknik visualisasi ini, analisis data menjadi lebih mendalam dan informatif. Penggunaan legenda dan grafik tree map tidak hanya mempermudah pemahaman, tetapi juga membantu bisnis dalam mengambil keputusan strategis, seperti mengatur stok produk, merencanakan promosi, dan mengidentifikasi tren pembelian pelanggan. Dengan memahami perilaku pembelian pelanggan secara lebih rinci, bisnis dapat mengoptimalkan strategi pemasaran mereka, meningkatkan kepuasan pelanggan, dan pada akhirnya, meningkatkan hasil keuangan bisnis mereka.



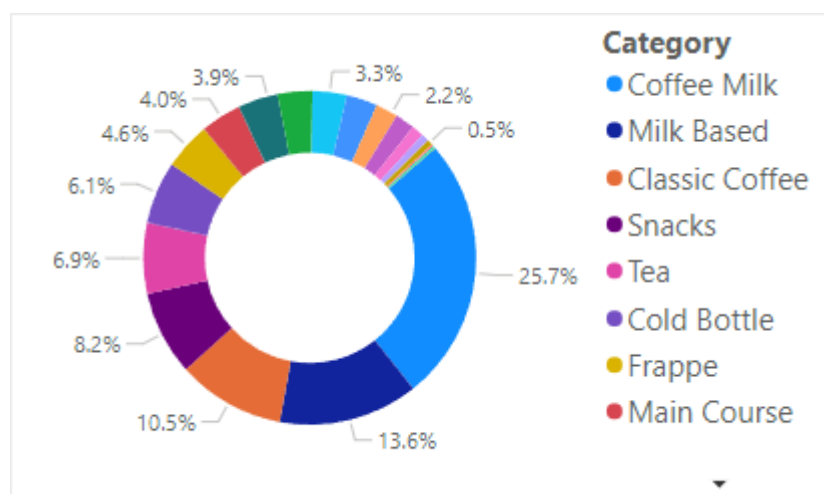
Gambar 4.21 Kuantitas Penjualan Produk Tertinggi

Dari visualisasi di atas, dapat disimpulkan bahwa terdapat dua menu yang mencatatkan penjualan tertinggi, yaitu produk bernama King dengan total pembelian sebanyak 787 unit dan Queen dengan 583 pembelian. Melalui analisis lebih lanjut, dapat ditemukan bahwa kedua

produk tersebut termasuk dalam kategori Coffee Milk. Menariknya, Coffee Milk sendiri merupakan kategori yang menduduki peringkat pertama dalam hal penjualan terbanyak, berdasarkan data yang tercantum pada Gambar 4.22.

Penting untuk dicatat bahwa keberhasilan penjualan King dan Queen, serta dominasi Coffee Milk dalam tabel penjualan, menunjukkan tren pasar yang jelas terhadap jenis produk ini. Analisis mendalam terhadap tren ini memberikan wawasan yang berharga bagi bisnis, memungkinkan mereka untuk fokus pada strategi pemasaran yang memaksimalkan potensi penjualan produk-produk sejenis atau bahkan mengembangkan variasi menu Coffee Milk yang dapat menarik lebih banyak pelanggan.

Dengan demikian, pemahaman yang mendalam tentang perilaku pembelian pelanggan terhadap kategori produk tertentu, seperti Coffee Milk, memberikan landasan yang kuat bagi pengambilan keputusan bisnis yang cerdas. Dalam konteks ini, bisnis dapat merancang strategi pemasaran yang lebih terarah, mengoptimalkan persediaan produk, dan meningkatkan kualitas layanan pelanggan, semuanya didasarkan pada data dan tren yang jelas dari perilaku pembelian pelanggan. Analisis ini tidak hanya membantu bisnis untuk tetap bersaing dalam pasar, tetapi juga membangun hubungan yang lebih kuat dengan pelanggan, menciptakan dasar yang kokoh untuk pertumbuhan bisnis yang berkelanjutan.



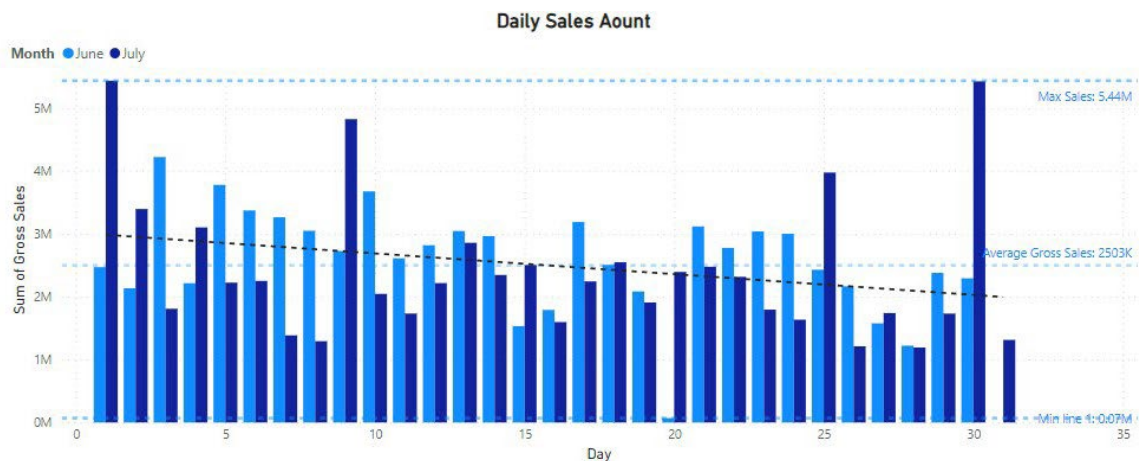
Gambar 4.22 Kuantitas Penjualan Kategori Tertinggi

Dalam analisis data, terdapat beberapa temuan yang menarik. Pertama-tama, persentase penjualan produk dengan kategori Coffee Milk ternyata mencapai tingkat tertinggi, yakni sekitar 25,7%. Temuan ini menunjukkan bahwa produk-produk dalam kategori Coffee Milk sangat diminati oleh pelanggan, menggambarkan kepopuleran minuman tersebut di pasar. Fakta kedua yang ditemukan adalah bahwa persentase penjualan produk non-coffee yang

termasuk dalam kategori Milk Based mencapai angka sekitar 13,6% dari total penjualan. Ini menunjukkan bahwa meskipun Coffee Milk mendominasi, produk-produk non-kopi juga memiliki pangsa pasar yang signifikan.

Selain itu, dalam kategori makanan, Snack menonjol dengan total penjualan sebesar 8,2% dari keseluruhan. Temuan ini mengindikasikan bahwa produk-produk snack memiliki daya tarik tersendiri bagi pelanggan, mungkin karena kepraktisan dan variasi yang ditawarkan.

Selanjutnya, analisis juga menyoroti aspek penjualan harian, di mana penjualan bisa mencapai angka luar biasa sebesar 5,44 juta rupiah dalam sehari. Visualisasi data melalui bar chart pada Gambar 4.23 memberikan gambaran yang jelas tentang fluktuasi penjualan sepanjang hari, membantu bisnis untuk mengidentifikasi jam-jam sibuk dan potensial untuk meningkatkan layanan pelanggan selama periode-periode tersebut.



Gambar 4.23 Penjualan Harian

Terdapat data menarik terkait penjualan harian yang perlu diperhatikan. Angka minimum penjualan pada beberapa hari berada di bawah angka 1 juta dalam satu hari, menunjukkan variasi yang signifikan dalam kinerja penjualan harian. Meskipun demikian, rata-rata penjualan selama dua bulan mencapai angka sekitar 2,5 juta per hari, menggambarkan gambaran umum tentang performa penjualan bisnis tersebut.

Menariknya, bulan Juli menunjukkan pola penjualan yang cukup dinamis. Beberapa hari, seperti pada tanggal 1, 9, dan 30, mencatatkan peningkatan penjualan yang signifikan. Namun, perlu dicatat bahwa meskipun terdapat peningkatan pada beberapa hari tertentu, penjualan harian secara keseluruhan mengalami penurunan seiring berjalannya waktu. Hal ini ditegaskan dengan adanya trend line berwarna hitam putus-putus yang terletak di tengah grafik, mencerminkan pola penurunan secara konsisten dalam periode waktu tersebut.



Visualisasi data dalam penelitian ini juga dirancang dengan cermat, mengintegrasikan tabel untuk memberikan pemahaman yang lebih lengkap. Tabel ini terfokus pada nama produk, jumlah terjual, dan pendapatannya, menyediakan informasi rinci yang memperkaya analisis. Gambar 4.24 memperlihatkan visualisasi dalam bentuk tabel, menciptakan landasan yang kokoh untuk memahami detail produk secara lebih mendalam.

Dengan mempertimbangkan informasi ini, bisnis memiliki landasan yang kuat untuk mengidentifikasi tren penjualan, menanggapi fluktuasi harian, dan mengambil tindakan yang diperlukan untuk meningkatkan penjualan. Analisis yang cermat terhadap pola penjualan harian dan penggunaan visualisasi data yang komprehensif membantu bisnis membuat keputusan yang informasional dan strategis, membawa mereka menuju kesuksesan yang berkelanjutan di pasar yang kompetitif.

Items	Sum of Quantity	Sum of Gross Sales
King	787	19.675M
Queen	583	13.992M
Mineral Water	368	2.944M
Americano	355	7.810M
Chocolate	257	6.425M
Red Velvet	232	5.800M
Cafe Latte	212	5.300M
Ace	192	5.376M
Matcha Greentea	188	4.888M
Lychee Tea	187	4.675M
French Fries	151	3.020M
Sunny Taro	147	3.675M
<b>Total</b>	<b>6074</b>	<b>152.663M</b>

Gambar 4.24 Tabel Detail

Halaman kedua dari penelitian ini menampilkan sejumlah grafik yang memberikan pemahaman yang mendalam tentang perilaku pembelian minuman oleh pelanggan di Noble Coffee. Gambar 4.25 menghadirkan visualisasi data dalam bentuk grafik pie dan area yang membahas secara rinci tentang perilaku pembelian pelanggan terhadap minuman-minuman yang ditawarkan di kafe tersebut.

Grafik pie pada Gambar 4.25 memberikan gambaran yang jelas tentang pembagian persentase penjualan minuman berdasarkan kategorinya. Dengan warna-warna yang berbeda, grafik pie ini mencerminkan sejauh mana pelanggan tertarik pada berbagai jenis minuman yang ada di menu. Dalam hal ini, penggunaan warna dan proporsi yang tepat membantu pembaca memahami dengan mudah perbandingan antar pilihan minuman, membimbing mereka menuju pemahaman yang lebih dalam tentang perilaku pembelian pelanggan.

Sementara itu, grafik area yang disajikan pada halaman ini memberikan gambaran kronologis tentang perilaku pembelian minuman oleh pelanggan di Noble Coffee. Dengan sumbu waktu sebagai acuan, grafik area ini mengilustrasikan fluktuasi dalam jumlah penjualan minuman harian. Perubahan warna dan area pada grafik ini memberikan informasi tambahan tentang variasi dalam pembelian minuman sepanjang waktu, mengidentifikasi pola-pola tertentu yang mungkin berkaitan dengan faktor-faktor tertentu, seperti musim atau promosi khusus.

Melalui visualisasi data yang kompleks ini, penelitian ini tidak hanya memberikan gambaran yang kaya tentang perilaku pembelian pelanggan, tetapi juga membantu mengidentifikasi tren dan pola yang mungkin tidak langsung terlihat.



Gambar 4.25 Grafik Penjualan Minuman

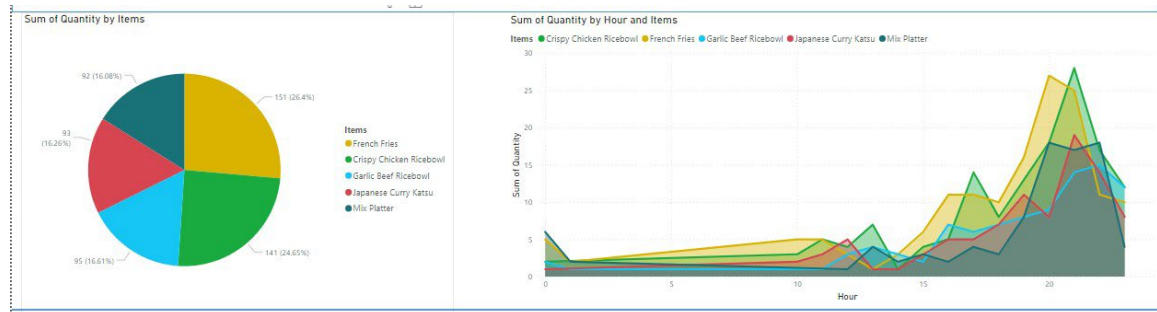
Dalam analisis minuman di Noble Coffee, terdapat lima jenis minuman yang mencatat penjualan tertinggi, yaitu King, Queen, Mineral Water, Americano, dan Chocolate. Masing-masing minuman ini memiliki waktu penjualan yang spesifik, menunjukkan perilaku pembelian pelanggan terkait konsumsi minuman di berbagai jam.

Sebagai contoh, dari grafik di sebelah kanan, terlihat bahwa Americano cenderung dinikmati pada pagi hari, lebih banyak diminati daripada keempat produk lainnya pada waktu tersebut. Namun, yang menarik, minuman ini juga sangat diminati saat tengah malam, mengindikasikan bahwa ada kelompok pelanggan yang menikmati minuman ini sebagai bagian dari rutinitas malam mereka.

Dari temuan-temuan ini, dapat dilakukan analisis lebih lanjut dengan memperoleh wawasan dari hasil visualisasi. Langkah selanjutnya adalah melibatkan data mengenai aturan asosiasi yang sebelumnya diperoleh, sebagaimana yang terlihat pada Gambar 4.26. Analisis aturan asosiasi memberikan gambaran yang lebih mendalam tentang kecenderungan pembelian bersamaan antara minuman dan produk lainnya. Dengan demikian, kombinasi analisis visual dan data asosiasi memberikan bisnis pemahaman yang komprehensif tentang perilaku pelanggan, membimbing mereka dalam mengatur waktu jualan dan menawarkan promosi yang



Pada halaman selanjutnya adalah visualisasi yang tidak jauh berbeda dari visualisasi minuman. Terdapat pie chart serta *hourly sales* terhadap tiap makanan yang dijual di Noble Coffee pada Gambar 4.27.



Gambar 4.27 Grafik Penjualan Makanan

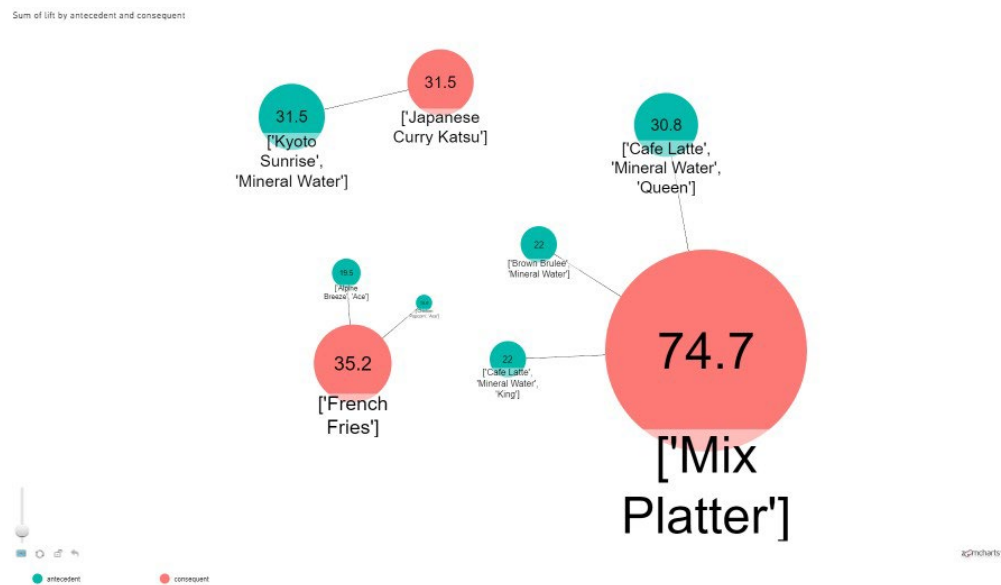
Dalam konteks penjualan makanan di Noble Coffee, terdapat lima jenis makanan yang menonjol dengan kuantitas penjualan tertinggi. Makanan-makanan ini mencakup French Fries, Crispy Chicken Ricebowl, Garlic Beef Ricebowl, Japanese Curry Katsu, dan Mix Plater. Meskipun demikian, menarik untuk dicatat bahwa penjualan makanan di coffee shop ini tidak sefokus dengan minuman, terutama karena pelanggan cenderung membeli makanan pada waktu-waktu tertentu.

Perilaku pembelian makanan oleh pelanggan ini menunjukkan bahwa pelanggan cenderung lebih suka membeli makanan berat pada jam makan siang, jam pulang kerja, dan jam makan malam. Oleh karena itu, penjualan makanan berat, terutama yang berbasis ayam seperti Crispy Chicken Ricebowl dan Japanese Curry Katsu, cenderung tinggi pada jam-jam tersebut. Salah satu makanan yang paling diminati oleh pelanggan adalah French Fries, yang memiliki pola penjualan yang stabil hampir sepanjang hari, menunjukkan popularitas konstan sepanjang waktu.

Dengan memperhatikan perilaku pembelian pelanggan yang cenderung memilih makanan pada waktu tertentu, Noble Coffee dapat mengoptimalkan stok dan menyusun menu spesial pada jam-jam tersebut, meningkatkan daya tarik menu yang paling diminati pada saat-saat sibuk tersebut.

Selanjutnya, hasil visualisasi dari aturan asosiasi yang didapatkan dari produk-produk teratas ini memberikan gambaran yang lebih rinci tentang hubungan pembelian bersamaan antara minuman dan makanan tersebut. Detail-detail ini dapat memberikan wawasan berharga mengenai kombinasi minuman dan makanan yang sering dibeli bersamaan oleh pelanggan. Visualisasi aturan asosiasi ini, yang dijelaskan dengan lebih detil pada Gambar 4.28, memberikan landasan yang solid untuk merencanakan promosi dan penawaran yang cerdas,

menghasilkan pengalaman pelanggan yang lebih baik, dan meningkatkan penjualan secara keseluruhan.

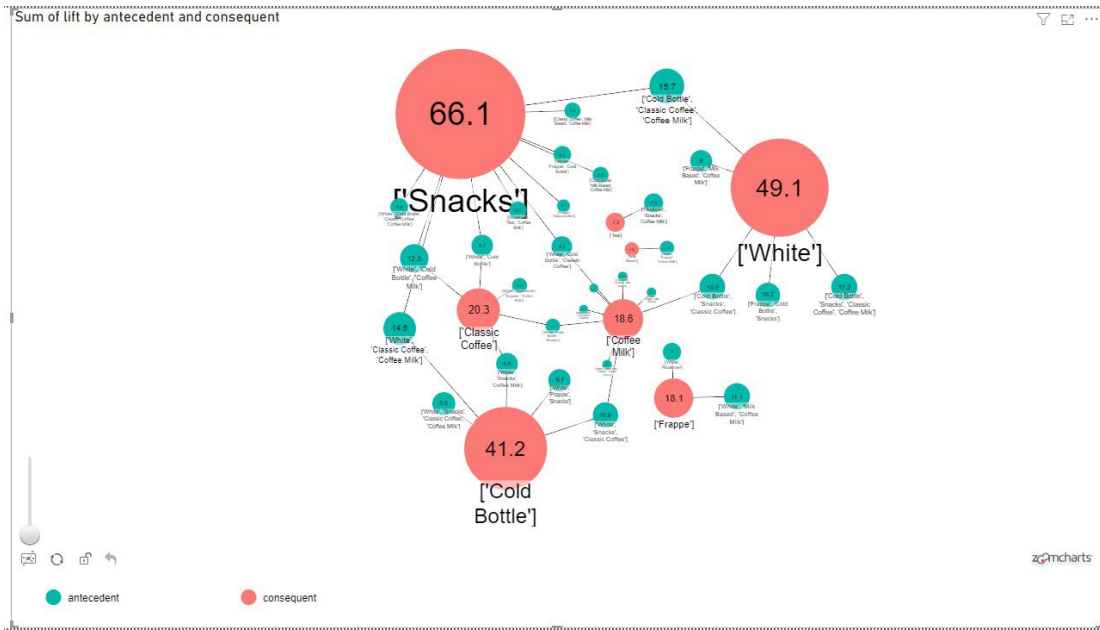


Gambar 4.28 Visualisasi Aturan Asosiasi Makanan

Hasil analisis dari visualisasi di atas mengungkap beberapa informasi yang signifikan. Salah satu temuan utama adalah bahwa Mix Plater merupakan makanan yang paling sering memiliki pasangan itemset, menunjukkan bahwa makanan ini seringkali dibeli bersamaan dengan berbagai jenis minuman atau makanan lainnya di Noble Coffee.

Visualisasi berikutnya yang dihasilkan adalah mengenai kategori produk. Dari visualisasi ini, dapat dinyatakan bahwa kombinasi pembelian kopi susu sebagai minuman dan snack sebagai makanan adalah yang paling umum terjadi, dengan pasangan itemset paling banyak dibandingkan dengan kombinasi produk lainnya. Temuan ini mencerminkan kecenderungan pelanggan untuk menikmati kopi susu bersamaan dengan camilan ringan atau snack, menciptakan pengalaman santai dan menyenangkan di Noble Coffee.

Gambar 4.29 memperlihatkan visualisasi yang menggambarkan itemset kategori yang saling berkesinambungan dengan sangat jelas. Penggunaan visualisasi ini memberikan gambaran yang komprehensif tentang perilaku pembelian pelanggan, memungkinkan bisnis untuk merencanakan tawaran promosi yang lebih cerdas dan merespons kebutuhan pelanggan dengan lebih tepat.



Gambar 4.29 Visualisasi Aturan Asosiasi Kategori Produk

#### 4.5.2 Rekomendasi

Setelah didapatkan visualisasi terkait penjualan serta visualisasi hasil implementasi algoritma FP-Growth, didapatkan beberapa rekomendasi yang dapat dijadikan acuan dalam mengoptimalkan strategi bisnis. Menilik dari visualisasi *rush hour* didapatkan wawasan berupa jam sibuk seperti pada pukul 17.00 sampai 22.00 yang merupakan jam banyak terjadi transaksi. Apabila dibandingkan dengan promo yang telah ada pada Noble Coffee, jam sibuk ini dapat dijadikan pilihan untuk dijadikan jam pengadaan promo yang baru.

Melihat visualisasi selanjutnya yaitu *pie chart* untuk memvisualisasikan total penjualan secara kategori menu didapatkan pilihan 2 kategori makanan dan minuman teratas. Kategori tersebut adalah *Coffee Milk* dan *Snacks*. Oleh karena itu seluruh menu yang memenuhi kategori ini dapat dijadikan acuan untuk dijadikan rekomendasi menu dalam promo. Produk yang memenuhi kategori *Coffee Milk* adalah Ace, King, dan Queen. Sedangkan produk yang memenuhi kategori *Snacks* adalah French Fries, Potato Wedges, Chicken Popcorn, BBQ Chicken Wings, Onion Ring, Bolognese Garlic Bruschetta, dan Mix Plater. Dengan memperhatikan visualisasi jumlah penjualan produk tertinggi yang telah terjadi, produk King dan Queen termasuk pada salah satu produk dalam daftar diatas. Oleh karena itu visualisasi hubungan asosiasi dapat dilihat lebih lanjut.

Dengan melakukan filter pada Power BI untuk memperlihatkan kedua item tersebut dalam visualisasi hubungan asosiasi didapatkan beberapa pasangan produk untuk dijadikan sebagai promo *item bundling*, pasangan produk tersebut antara lain:

1. King → Mix Plater, French Fries
2. King → Mix Plater, Café Latte, Mineral Water
3. King → Cireng, Coklat
4. King → Mix Plater, Café Latte
5. King → Cireng, French Fries
6. Queen → Mix Plater, Café Latte, Mineral Water
7. Queen → Mix Plater, Café Latte

## **BAB V**

### **KESIMPULAN DAN SARAN**

Bab ini akan membahas dan merangkum hasil penelitian ini, berdasarkan pemrosesan, implementasi, dan visualisasi. Saran akan membahas catatan yang bisa dijadikan acuan dalam pengembangan lebih lanjut terkait penelitian ini untuk waktu yang akan datang.

#### **5.1 Kesimpulan**

Dalam penelitian ini, data penjualan dan perilaku pembelian pelanggan di Noble Coffee dianalisis secara mendalam. Penggunaan berbagai teknik visualisasi data memberikan wawasan komprehensif tentang perilaku pembelian pelanggan, pola jam sibuk, preferensi minuman, dan makanan yang paling diminati. Selain itu, hubungan antara minuman dan makanan yang sering dibeli bersamaan juga teridentifikasi.

Hasil analisis visualisasi data menunjukkan minuman dalam kategori Coffee Milk, khususnya produk King dan Queen, sangat diminati oleh pelanggan. Pemahaman mendalam tentang perilaku pembelian oleh pelanggan ini memberikan peluang bagi bisnis untuk merancang strategi pemasaran yang lebih terarah. Dengan pengetahuan ini, bisnis dapat mengoptimalkan stok produk, meningkatkan kualitas layanan pelanggan, dan merencanakan promosi yang lebih efektif.

Analisis aturan asosiasi juga membantu mengidentifikasi kombinasi produk yang sering dibeli bersamaan. Informasi ini membimbing bisnis dalam merencanakan promosi dan penawaran yang cerdas, menciptakan strategi yang dapat meningkatkan penjualan dengan efisien.

Selain itu, analisis perilaku pembelian harian dan jam sibuk memberikan peluang bagi bisnis untuk mengoptimalkan waktu jualan dan menawarkan promosi yang relevan. Dengan menyesuaikan strategi pemasaran dan penawaran produk pada jam-jam tertentu, Noble Coffee dapat meningkatkan kepuasan pelanggan dan penjualan bisnisnya.

Hasil analisis visual penjualan dan penerapan algoritma FP-Growth memberikan beberapa rekomendasi strategi bisnis yang dapat dioptimalkan. Dari visualisasi, terlihat pola waktu sibuk transaksi antara pukul 17.00 hingga 22.00, menjadi potensi waktu yang ideal untuk peluncuran promosi baru di Noble Coffee. Fokus pada kategori penjualan tertinggi seperti Coffee Milk dan Snacks, terutama produk spesifik seperti Ace, King, Queen, dan beberapa menu Snacks,



menawarkan peluang promosi yang menjanjikan. Selain itu, analisis asosiasi produk mengidentifikasi pasangan produk untuk paket bundling, seperti King dengan Mix Plater dan French Fries, memberikan landasan yang kuat untuk strategi promosi bundling yang dapat meningkatkan penjualan secara signifikan.

Dengan pemahaman mendalam tentang perilaku pembelian pelanggan, Noble Coffee memiliki dasar yang kuat untuk merancang strategi pemasaran yang lebih cerdas dan efektif. Bisnis dapat mengatur promosi, mengoptimalkan persediaan produk, dan merespons kebutuhan pelanggan dengan lebih tepat. Penggunaan teknik visualisasi data, seperti grafik pie, grafik area, dan aturan asosiasi, memberikan pandangan yang holistik dan membantu bisnis membuat keputusan yang informasional dan strategis. Semua temuan ini memberikan peluang besar bagi Noble Coffee untuk berkembang dan meningkatkan keberhasilannya di pasar.

## 5.2 Saran

Harapannya adalah bahwa penelitian mendalam yang dilakukan ini akan menjadi kontribusi yang berharga bagi komunitas ilmiah dan industri. Dengan menyajikan temuan-temuan yang mendalam dan analisis yang teliti, penelitian ini diharapkan dapat memberikan landasan yang kokoh bagi penelitian lanjutan dalam bidang ini. Keberhasilan penelitian ini diharapkan juga dapat mengilhami peneliti lain untuk menjelajahi aspek-aspek baru dalam topik ini, membuka jalan bagi penemuan-penemuan baru dan inovatif di masa mendatang.

Selain itu, diharapkan bahwa hasil penelitian ini akan memotivasi institusi pendidikan, organisasi industri, dan pemangku kepentingan lainnya untuk mendukung penelitian lebih lanjut di bidang ini. Dengan demikian, penelitian ini diharapkan tidak hanya menjadi dokumen tertulis, tetapi juga sebuah pijakan yang menggerakkan perkembangan pengetahuan dan memajukan industri ke depannya.

Tidak hanya itu, harapannya adalah bahwa penelitian ini dapat menjadi rujukan utama bagi para pelaku bisnis, pengambil kebijakan, dan pemangku kepentingan lainnya. Dengan menggali wawasan mendalam tentang preferensi pelanggan dan perilaku pasar, penelitian ini memiliki potensi untuk memberikan pedoman berharga bagi strategi bisnis di masa depan. Dengan begitu, penelitian ini diharapkan dapat membantu bisnis untuk lebih memahami pasar dan mengambil keputusan yang lebih cerdas, meningkatkan daya saing mereka di pasar yang semakin kompleks ini.

Dengan menjadikan penelitian ini sebagai pedoman dan sumber referensi yang andal, diharapkan bahwa para peneliti, akademisi, dan praktisi di masa mendatang akan dapat

membangun pengetahuan lebih lanjut, membuka dialog intelektual yang produktif, dan memperluas wawasan dalam bidang ini. Sebagai hasilnya, harapannya adalah bahwa penelitian ini akan menjadi landasan yang kuat untuk inovasi dan pengembangan di masa depan, membawa manfaat positif bagi industri, masyarakat, dan ekonomi secara keseluruhan.

Adapun beberapa poin yang bisa dijadikan saran penelitian lanjutan yang serupa adalah sebagai berikut:

1. Menggunakan dataset dengan periode waktu lebih panjang.
2. Menggunakan dataset pada bidang bisnis yang lain.
3. Menggunakan algoritma lain untuk digunakan sebagai perbandingan.
4. Merencanakan dashboard visualisasi yang lebih menarik.

## DAFTAR PUSTAKA

- Agrawal, R., & Shafer, J. C. (1996). Parallel mining of association rules. *IEEE Transactions on Knowledge and Data Engineering*, 8(6), 962–969. <https://doi.org/10.1109/69.553164>
- Ahn, J. S., & Sohn, S. Y. (2009). Customer pattern search for after-sales service in manufacturing. *Expert Systems with Applications*, 36(3), 5371–5375. <https://doi.org/10.1016/j.eswa.2008.06.061>
- Andhika Putra, R., Bayu Putra, R., & Fitri, H. (2020). ANALISIS SWOT DALAM PENENTUAN STRATEGI BISNIS KAFE 165 UNIVERSITAS PUTRA INDONESIA YPTK PADANG. *Jurnal Ilmu Manajemen Terapan*, 1(5), 464–476. <https://doi.org/10.31933/jimt.v1i5.206>
- Becker, L. T., & Gould, E. M. (2019). Microsoft Power BI: Extending Excel to Manipulate, Analyze, and Visualize Diverse Data. *Serials Review*, 45(3), 184–188. <https://doi.org/10.1080/00987913.2019.1644891>
- Bhatia, R., Benno, S., Esteban, J., Lakshman, T. V., & Grogan, J. (2019). Unsupervised machine learning for network-centric anomaly detection in IoT. *Proceedings of the 3rd ACM CoNEXT Workshop on Big Data, Machine Learning and Artificial Intelligence for Data Communication Networks*, 42–48. <https://doi.org/10.1145/3359992.3366641>
- Borgelt, C. (2005). An implementation of the FP-growth algorithm. *Proceedings of the 1st International Workshop on Open Source Data Mining: Frequent Pattern Mining Implementations*, 1–5. <https://doi.org/10.1145/1133905.1133907>
- Bose, I., & Mahapatra, R. K. (2001). Business data mining — a machine learning perspective. *Information & Management*, 39(3), 211–225. [https://doi.org/10.1016/S0378-7206\(01\)00091-X](https://doi.org/10.1016/S0378-7206(01)00091-X)
- Burkart, N., & Huber, M. F. (2021). A Survey on the Explainability of Supervised Machine Learning. *Journal of Artificial Intelligence Research*, 70, 245–317. <https://doi.org/10.1613/jair.1.12228>
- Carbonell, J. G., Michalski, R. S., & Mitchell, T. M. (1983). AN OVERVIEW OF MACHINE LEARNING. In *Machine Learning* (pp. 3–23). Elsevier. <https://doi.org/10.1016/B978-0-08-051054-5.50005-4>

- Chaudhuri, S., Dayal, U., & Narasayya, V. (2011). An overview of business intelligence technology. *Communications of the ACM*, 54(8), 88–98. <https://doi.org/10.1145/1978542.1978562>
- Danuri, M. (2019). *PERKEMBANGAN DAN TRANSFORMASI TEKNOLOGI DIGITAL*. <https://doi.org/https://doi.org/10.53845/infokam.v15i2.178>
- Edhya, B. F. P. (2022). BUSINESS INTELLIGENCE DATA MARKETING MENGGUNAKAN METODE KIMBALL DAN ETL DENGAN POWER BI. *Kurawal - Jurnal Teknologi, Informasi Dan Industri*, 5(2), 87–97. <https://doi.org/10.33479/kurawal.v5i2.642>
- Foley, É., & Guillemette, M. G. (2010). What is Business Intelligence? *International Journal of Business Intelligence Research*, 1(4), 1–28. <https://doi.org/10.4018/jbir.2010100101>
- Gessner, G., & Scott, R. A. (2009). Using Business Intelligence Tools to Help Manage Costs and Effectiveness of Business-to-Business Inside-Sales Programs. *Information Systems Management*, 26(2), 199–208. <https://doi.org/10.1080/10580530902797623>
- Hunyadi, D. (2011). Performance Comparison of Apriori and FP-Growth Algorithms in Generating Association Rules. *Proceedings of the 5th European Conference on European Computing Conference*, 376–381.
- Iliashenko, O., Iliashenko, V., & Esser, M. (2019). BI systems implementation for supply chain sector in retail companies. *Proceedings of the International Conference on Digital Technologies in Logistics and Infrastructure (ICDTLI 2019)*. <https://doi.org/10.2991/icdtli-19.2019.53>
- Indrayani, H. (2012). PENERAPAN TEKNOLOGI INFORMASI DALAM PENINGKATAN EFEKTIVITAS, EFISIENSI DAN PRODUKTIVITAS PERUSAHAAN. *Jurnal EL-RIYASAH*, 3(1), 48. <https://doi.org/10.24014/jel.v3i1.664>
- Jenkins, H. (2009). A 'business opportunity' model of corporate social responsibility for small- and medium-sized enterprises. *Business Ethics: A European Review*, 18(1), 21–36. <https://doi.org/10.1111/j.1467-8608.2009.01546.x>
- Jiang, T., Gradus, J. L., & Rosellini, A. J. (2020). Supervised Machine Learning: A Brief Primer. *Behavior Therapy*, 51(5), 675–687. <https://doi.org/10.1016/j.beth.2020.05.002>
- Jordan, M. I., & Mitchell, T. M. (2015). Machine learning: Trends, perspectives, and prospects. *Science*, 349(6245), 255–260. <https://doi.org/10.1126/science.aaa8415>

- Kaur, M., & Kang, S. (2016). Market Basket Analysis: Identify the Changing Trends of Market Data Using Association Rule Mining. *Procedia Computer Science*, 85, 78–85. <https://doi.org/10.1016/j.procs.2016.05.180>
- Kurniawan, R., & Yusuf, R. (2023). PENERAPAN METODE ASOSIASI DENGAN ALGORITMA APRIORI UNTUK MENDUKUNG STRATEGI PROMOSI. *EDUSAINTEK: Jurnal Pendidikan, Sains Dan Teknologi*, 10(2), 514–528. <https://doi.org/10.47668/edusaintek.v10i2.786>
- Lee, D., Park, S.-H., & Moon, S. (2013). Utility-based association rule mining: A marketing solution for cross-selling. *Expert Systems with Applications*, 40(7), 2715–2725. <https://doi.org/10.1016/j.eswa.2012.11.021>
- LEE, J., & PARK, S. (2005). Intelligent profitable customers segmentation system based on business intelligence tools. *Expert Systems with Applications*, 29(1), 145–152. <https://doi.org/10.1016/j.eswa.2005.01.013>
- Lee, W. S., Moon, J., & Song, M. (2018). Attributes of the coffee shop business related to customer satisfaction. *Journal of Foodservice Business Research*, 21(6), 628–641. <https://doi.org/10.1080/15378020.2018.1524227>
- Lenk, K. M., Caspi, C. E., Harnack, L., & Laska, M. N. (2018). Customer Characteristics and Shopping Patterns Associated with Healthy and Unhealthy Purchases at Small and Non-traditional Food Stores. *Journal of Community Health*, 43(1), 70–78. <https://doi.org/10.1007/s10900-017-0389-5>
- Lima, E., Bayot, R., Brito, P., Rodrigues, N., Ribeiro, B., & Lopes, N. (2021). Business Analytical Framework for the Manufacturing Industry. *2021 IEEE 19th International Conference on Industrial Informatics (INDIN)*, 1–8. <https://doi.org/10.1109/INDIN45523.2021.9557370>
- McNicholas, P. D., Murphy, T. B., & O'Regan, M. (2008). Standardising the lift of an association rule. *Computational Statistics & Data Analysis*, 52(10), 4712–4721. <https://doi.org/10.1016/j.csda.2008.03.013>
- Mosadeghrad, A., janababie, ghasem, Kalantari, B., Abbasi, M., & Dehnavi, H. (2021). Evaluation of efficiency of general hospitals in Iran. *Scientific Journal of Kurdistan University of Medical Sciences*, 26(3), 151–169. <https://doi.org/10.52547/sjku.26.3.151>
- Nabilla, N. U. (2021). Pengembangan Business Intelligence pada Sistem Informasi Distributor. *Automata Vol 2 No 2*.

- QISTHANI, N. N. (2019). *PERANCANGAN SISTEM BUSINESS INTELLIGENCE UNTUK MENGETAHUI PERILAKU PENJUALAN PRODUK*.
- Rifai, D., Fitri, S., & Ramadhan, I. N. (2022). Perkembangan Ekonomi Digital Mengenai Perilaku Pengguna Media Sosial Dalam Melakukan Transaksi. *ADI Bisnis Digital Interdisiplin Jurnal*, 3(1), 49–52. <https://doi.org/10.34306/abdi.v3i1.752>
- SAPTOMO, A. B. (2022). *IMPLEMENTASI BUSINESS INTELLIGENCE DALAM MEMBUAT KEPUTUSAN PENJUALAN PROPERTI DENGAN METODE ONLINE ANALYTICAL PROCESSING (OLAP)*.
- Saravanan, R., & Sujatha, P. (2018). A State of Art Techniques on Machine Learning Algorithms: A Perspective of Supervised Learning Approaches in Data Classification. *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 945–949. <https://doi.org/10.1109/ICCONS.2018.8663155>
- Sharma, N., & Sarkar, D. (2022). Healthcare Data Analytics Using Power BI. *International Journal of Software Innovation*, 10(1), 1–10. <https://doi.org/10.4018/IJSI.293267>
- Verma, R., Stock, D., & McCarthy, L. (2012). Customer Preferences for Online, Social Media, and Mobile Innovations in the Hospitality Industry. *Cornell Hospitality Quarterly*, 53(3), 183–186. <https://doi.org/10.1177/1938965512445161>
- Wang, B., Chen, D., Shi, B., Zhang, J., Duan, Y., Chen, J., & Hu, R. (2017). Comprehensive Association Rules Mining of Health Examination Data with an Extended FP-Growth Method. *Mobile Networks and Applications*, 22(2), 267–274. <https://doi.org/10.1007/s11036-016-0793-6>
- Watson, H. J., & Wixom, B. H. (2007). The Current State of Business Intelligence. *Computer*, 40(9), 96–99. <https://doi.org/10.1109/MC.2007.331>
- Wei Zhang, Hongzhi Liao, & Na Zhao. (2008). Research on the FP Growth Algorithm about Association Rule Mining. *2008 International Seminar on Business and Information Management*, 315–318. <https://doi.org/10.1109/ISBIM.2008.177>
- Wijaya, K. T., & Pratama, I. (2022). Penerapan Algoritma FP-Growth Untuk Analisis Data Transaksi Penjualan Di Internet Learning Cafe Kaliurang. *Jurnal Nasional Komputasi Dan Teknologi Informasi (JNKTI)*, 5(4), 642–651. <https://doi.org/10.32672/jnkti.v5i4.4585>
- Yana Siregar, L., & Irwan Padli Nasution, M. (2020). Perkembangan Teknologi Informasi Terhadap Peningkatan Bisnis Online. *Hirarki : Jurnal Ilmiah Manajemen Dan Bisnis*, 2(1), 71–75. <https://journal.upp.ac.id/index.php/Hirarki/article/view/331>

Zikra, A. A. (2022). Implementasi Business Intelligence pada ACCAbsensi Menggunakan Aplikasi Power BI. *Automata Vol 3 No 2*.

## LAMPIRAN