

IMPLEMENTASI INTELIJEN BISNIS UNTUK VISUALISASI POLA PEMBELIAN DENGAN ALGORITMA *FP-GROWTH*

(Studi Kasus: Noble Coffee Yogyakarta)

Daffa Nayaka Bagaskara Jusuf

Universitas Islam Indonesia, Indonesia

E-mail: bagaskara2010@gmail.com

Abstract: This research aims to explore customer purchasing patterns at Noble Coffee, a cafe in Yogyakarta. Utilizing the *FP-Growth* algorithm, data analysis was conducted to identify association rules among products and categories with the highest sales. Data visualization was performed through the *Power BI* platform, enabling a deep understanding of customer purchasing behavior. The analysis revealed several significant association rules, such as the relationship between the product 'King' and various other items like 'Mix Plater' and 'Café Latte'. This analysis provides insights into frequently co-purchased products, allowing Noble Coffee to design more effective sales strategies. The study highlights the importance of data analysis and visualization in supporting business decisions. Through this approach, companies can better understand customer needs, enhance customer satisfaction, and optimize their operations. In conclusion, data analysis using the *FP-Growth* algorithm and visualization through *Power BI* can be effective instruments in understanding customer behavior and improving the performance of cafes, such as Noble Coffee.

Keywords: Data Analysis, Association Rules, Data Visualization, *FP-Growth*, *Power BI*, Business, Purchasing Patterns.

Abstrak: Penelitian ini mencoba menggali pola 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: Data Analisa, Aturan Asosiasi, Visualisasi Data, *FP-Growth*, *Power BI*, Bisnis, Pola Pembelian.

Copyright (c) 2024 The Authors. This is an open access article under the CC BY-SA 4.0 license (<https://creativecommons.org/licenses/by-sa/4.0/>)

PENDAHULUAN

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 merespon 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 kebiasaan dan preferensi 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 dan preferensi, kafe dapat merancang kampanye pemasaran yang lebih efektif dan tepat sasaran (Verma et al., 2012). Namun, mengadopsi analisis data juga memerlukan investasi dalam infrastruktur dan keterampilan analisis.

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 di jalan Anggajaya 2 Condongcatur Yogyakarta. Untuk mempertahankan *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 preferensi 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 minuman, makanan dan makanan, 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. 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.

METODE

Dalam pelaksanaan penelitian ini metode yang digunakan adalah metode analisis deskriptif yang nantinya divisualisasikan menggunakan aplikasi intelijen bisnis. Analisis deskriptif adalah tipe analisis data yang membantu menggambarkan, menunjukkan, atau merangkum titik-titik data secara konstruktif sehingga pola-pola dapat muncul yang memenuhi setiap kondisi data. Analisis deskriptif dapat dikategorikan menjadi empat jenis, yaitu ukuran frekuensi, kecenderungan pusat, dispersi atau variasi, dan posisi. Metode ini optimal untuk satu variabel pada satu waktu (Lawless & Heymann, 2010). Dalam dunia bisnis, analisis deskriptif memungkinkan perusahaan untuk memahami karakteristik dan perilaku pelanggan mereka dengan lebih baik (Campbell & Cunningham, 1983). Dengan merangkum data penjualan, preferensi pelanggan, dan tren pembelian, bisnis dapat mengidentifikasi produk atau layanan yang paling diminati oleh pelanggan. Informasi ini memungkinkan perusahaan untuk mengoptimalkan stok produk, merencanakan strategi pemasaran yang lebih efektif, dan mengembangkan produk atau layanan baru yang sesuai dengan kebutuhan pelanggan (Palepu & Healy, 2013). Maka perlu dilakukan visualisasi menggunakan aplikasi intelijen bisnis bertujuan untuk memudahkan pembacanya dalam membaca hasil olahan data (Waskom, 2021). Tahapan dari penelitian ini antara lain:

1. Pengumpulan Data

Data dikumpulkan dari hasil transaksi Noble Coffee bulan Juni dan Juli 2023 yang tersimpan dalam *backoffice* Moka POS. Moka POS adalah cloud-based point of sales (POS) yang didesain untuk membantu bisnis mengatur transaksi harian secara real-time (Susilowaty & Rukismono, 2021).

2. Pemrosesan Data

Pemrosesan data dilakukan menggunakan *Apache Spark*. *Apache Spark* adalah alat berguna dalam analisis data yang memberikan beberapa keunggulan (Shanahan & Dai, 2015). Berikut adalah keunggulan dari *Apache Spark* seperti:

- a. Cepat dan efisien dengan *in-memory caching* dan eksekusi *query* teroptimasi untuk analisis data semua ukuran dengan cepat.
- b. Fleksibel dalam mengatasi beban kerja aktivitas data analisis.
- c. Ketersediaan API untuk SQL, *batch processing*, dan *machine learning*.
- d. Komunitas *open-source* yang besar dengan lebih dari 1.000 developer dari 200 organisasi berbeda yang memberikan kontribusi dalam pengembangannya.
- e. Ketersediaan untuk menyokong analisis kompleks untuk dataset yang besar.

3. Implementasi Algoritma *FP-Growth*

FP-Growth (*Frequent Pattern Growth*) adalah algoritma data mining yang digunakan untuk menemukan pola yang sering muncul dalam *dataset* (Wang et al., 2002). Algoritma ini berguna dalam tugas analisis asosiasi seperti *market basket analysis*. Tujuan dari algoritma ini adalah menemukan hubungan antar item yang sering dibeli bersamaan dalam transaksi. Tahapan kerja *FP-Growth* adalah melakukan konstruksi *FP-Tree* dengan memindai *dataset* untuk mengidentifikasi item yang sering muncul dengan menggunakan ambang batas yang disebut nilai *support*. Kemudian *FP-Tree* dibangun dari item-item yang sering muncul. *FP-Tree* struktur data yang menggambarkan hubungan antar item serta frekuensi kemunculannya.

Konsep penting yang digunakan dalam aturan asosiasi adalah *support*, *confidence*, dan *lift* (Lin et al., 2002). *Support* adalah persentase kemunculan suatu *itemset* dari *dataset* yang mengukur seberapa sering suatu kombinasi produk muncul dalam transaksi. *Support* dihitung dengan membagi jumlah total transaksi yang berisi *itemset* tersebut dengan total jumlah transaksi dalam *dataset*.

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

Confidence mengukur seberapa sering aturan asosiasi ditemukan benar, sebagai contoh adalah seberapa sering item B dibeli ketika item A juga dibeli. *Confidence*

dinyatakan sebagai persentase dan dihitung sebagai rasio antara *support* dari *itemset* gabungan (*itemset* A dan B) dengan *support* dari *itemset* A.

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

Lift adalah ukuran untuk mengetahui sejauh mana dua item (A dan B) berkorelasi. *Lift* lebih dari 1 menunjukkan bahwa pembelian item A memiliki efek positif pada pembelian item B. Sementara *lift* kurang dari 1 menunjukkan efek negatif. *Lift* 1 menunjukkan bahwa pembelian item A dan B bersifat independen.

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

4. Visualisasi Data

Visualisasi data dalam penelitian ini menggunakan platform analisis bisnis yang dikembangkan oleh *Microsoft* bernama *Power BI*. *Power BI* berguna untuk menggabungkan, menganalisis, dan menggambarkan data dari berbagai sumber dengan cara interaktif dan mudah dipahami. *Power BI* memiliki alat pada platformnya untuk melakukan transformasi data, analisis visual, dan berbagi wawasan bisnis. Fungsi utama dari *Power BI* dalam visualisasi data adalah sebagai berikut:

- a. *Power BI* memungkinkan penggunaannya untuk menghubungkan dan menggabungkan data dari berbagai sumber, termasuk basis data relasional, file excel, data cloud, layanan web, dan masih banyak sumber data lainnya. Oleh karena itu sumber yang berbeda dan dari berbagai tempat dalam terintegrasi.
- b. Alat transformasi data *Power BI* memungkinkan untuk dilakukan membersihkan dan merapikan data. Operasi seperti menghapus kolom yang tidak perlu, mengganti nilai yang hilang maupun kosong, maupun menggabungkan data dari berbagai tabel menjadi satu.
- c. *Power BI* memiliki kemampuan untuk membuat model data yang kompleks dengan menggabungkan tabel, mendefinisikan hubungan antar tabel, dan membuat kolom kalkulasi dengan formula DAX (*Data Analysis Expression*)
- d. Visualisasi yang dilakukan pada *Power BI* dapat dilakukan filter dengan cara yang interaktif pada visualisasinya secara langsung untuk menyelidiki data dengan cara yang lebih mendalam. Interaksi antar visual dapat dilakukan pada

visualisasinya langsung dimana perubahan satu visualisasi mempengaruhi visualisasi lainnya secara real-time.

- e. *Power BI* dapat digunakan untuk menyesuaikan visualisasi dan menggambarkan data sesuai dengan kebutuhan merek atau perusahaan. Selain itu, *Power BI* dapat diintegrasikan dengan berbagai aplikasi bisnis seperti *Microsoft Office 365*, *SharePoint*, dan *Azure* untuk meningkatkan fungsionalitas dan kolaborasi.

HASIL DAN PEMBAHASAN

Hasil

1. Pengumpulan Data

Data yang digunakan pada penelitian ini diperoleh melalui ekspor data penjualan yang tersedia dalam *platform backoffice* Moka POS. 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.

13 TRANSACTIONS		Rp. 912.000 TOTAL COLLECTED		Rp. 718.700 NET SALES	
Friday, July 31, 2020					
Rp. 786.000					
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
Thursday, July 30, 2020					
(Rp. 34.000)					
Moka's Caffe	23:55	Moka Support	Full Refund for: 2CDD4SO		Rp. 21.350
Moka's Caffe	23:53	Moka Support	Partial Refund for: 2CDD4SO		Rp. 12.650
Tuesday, July 28, 2020					
Rp. 160.000					

Gambar 1. Pilihan Ekspor Data.

Hasil 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 2.

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
Outlet	Receipt N	Date	Time	Category	Brand	Items	Variant	SKU	Quantity	Modifier	J	Discount	Gross Sale	Discounts	Refunds	Net Sales	Gratuity	Tax	Sales Type	Collected	Served By	Customer	Payment	Event	Type		
2	38PBVD	31/7/2023	23:23:44	Main Cou	Unbranded	Japanese Curry Katsu			1		Staff	Appr	32000	11200	0	20800	0	20800	0	20800	Dine In	Kasir Noble		Cash	Payment		
3	38PBVC	31/7/2023	23:07:54	Salad	Unbranded	Signature Thousand Island			1				26000	0	0	26000	0	26000	0	26000	Dine In	Kasir Noble		GrabFood	Payment		
4	38PBVB	31/7/2023	22:17:39	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		BCA QR	Payment		
5	38PBVB	31/7/2023	22:17:39	Milk Base	Unbranded	Red Velvet Ice			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA QR	Payment		
6	38PBVA	31/7/2023	22:16:46	Coffee Tl	Unbranded	Hazelnut Iced			1				28000	0	0	28000	0	28000	0	28000	Dine In	Kasir Noble		Cash	Payment		
7	38PBV9	31/7/2023	22:09:43	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		BCA QR	Payment		
8	38PBVB	31/7/2023	21:56:34	Milk Base	Unbranded	Matcha Gr Ice			1				26000	0	0	26000	0	26000	0	26000	Dine In	Kasir Noble		BCA QR	Payment		
9	38PBVB	31/7/2023	21:56:34	Coffee Mi	Unbranded	Ace			1				28000	0	0	28000	0	28000	0	28000	Dine In	Kasir Noble		BCA QR	Payment		
10	38PBV7	31/7/2023	21:54:44	Ricebowl	Unbranded	Crispy Chi BBQ			1		Disc		30000	3000	0	27000	0	27000	0	27000	Dine In	Kasir Noble		BCA QR	Payment		
11	38PBV7	31/7/2023	21:54:44	Coffee Mi	Unbranded	Queen			1		Disc		24000	2400	0	21600	0	21600	0	21600	Dine In	Kasir Noble		BCA QR	Payment		
12	38PBV7	31/7/2023	21:54:44	White	Unbranded	Pinkiy no Bara			1		Disc		30000	3000	0	27000	0	27000	0	27000	Dine In	Kasir Noble		BCA QR	Payment		
13	38PBV7	31/7/2023	21:54:44	Coffee Mi	Unbranded	Queen			2		Disc		48000	4800	0	43200	0	43200	0	43200	Dine In	Kasir Noble		BCA QR	Payment		
14	38PBVB	31/7/2023	21:53:29	Classic Co	Unbranded	American Hot			1				22000	0	0	22000	0	22000	0	22000	Dine In	Kasir Noble		Cash	Payment		
15	38PBVB	31/7/2023	21:53:29	Snacks	Unbranded	Chicken Popcorn			1				23000	0	0	23000	0	23000	0	23000	Dine In	Kasir Noble		Cash	Payment		
16	38PBV5	31/7/2023	21:44:43	Coffee Mi	Unbranded	King			2				50000	0	0	50000	0	50000	0	50000	Dine In	Kasir Noble		Cash	Payment		
17	38PBV4	31/7/2023	21:08:13	White	Unbranded	Blooming In Tokyo			1				30000	0	0	30000	0	30000	0	30000	Dine In	Kasir Noble		Cash	Payment		
18	38PBV3	31/7/2023	21:08:17	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		Cash	Payment		
19	38PBV2	31/7/2023	20:45:23	Sparkles	Unbranded	Lychee Adict			1				32000	0	0	32000	0	32000	0	32000	Dine In	Kasir Noble		BCA QR	Payment		
20	38PBV1	31/7/2023	19:37:48	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		BCA	Payment		
21	38PBV0	31/7/2023	19:36:46	Tea	Unbranded	Strawberry Tea			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA QR	Payment		
22	38PBU2	31/7/2023	19:34:39	Classic Co	Unbranded	Cafe Latte Ice			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA	Payment		
23	38PBUI	31/7/2023	19:34:39	Coffee Mi	Unbranded	King			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA	Payment		
24	38PBUI	31/7/2023	19:21:47	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		BCA QR	Payment		
25	38PBUX	31/7/2023	18:27:28	Tea	Unbranded	Lychee Tea			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA QR	Payment		
26	38PBUX	31/7/2023	18:27:28	Snacks	Unbranded	Onion Ring			1				20000	0	0	20000	0	20000	0	20000	Dine In	Kasir Noble		BCA QR	Payment		
27	38PBUIW	31/7/2023	18:20:25	Pasta	Unbranded	Aglio Olio Spaghetti			1				30000	0	0	30000	0	30000	0	30000	Dine In	Kasir Noble		BCA QR	Payment		
28	38PBUIW	31/7/2023	18:20:26	Classic Co	Unbranded	American Ice			1				22000	0	0	22000	0	22000	0	22000	Dine In	Kasir Noble		BCA QR	Payment		
29	38PBUIV	31/7/2023	16:52:18	Sparkles	Unbranded	Kyoto Sunrise			1				32000	0	0	32000	0	32000	0	32000	Dine In	Kasir Noble		BCA QR	Payment		
30	38PBUIV	31/7/2023	16:52:18	Main Cou	Unbranded	Japanese Curry Katsu			1				32000	0	0	32000	0	32000	0	32000	Dine In	Kasir Noble		BCA QR	Payment		
31	38PBUIU	31/7/2023	16:45:06	Coffee Mi	Unbranded	Queen			1				24000	0	0	24000	0	24000	0	24000	Dine In	Kasir Noble		BCA QR	Payment		
32	38PBUIU	31/7/2023	16:45:06	Coffee Mi	Unbranded	King			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA QR	Payment		
33	38PBUIU	31/7/2023	16:45:06	Ricebowl	Unbranded	Crispy Chi Mushroom			1				30000	0	0	30000	0	30000	0	30000	Dine In	Kasir Noble		BCA QR	Payment		
34	38PBUIU	31/7/2023	16:45:06	Main Cou	Unbranded	Hongkong Fried Rice			1				30000	0	0	30000	0	30000	0	30000	Dine In	Kasir Noble		BCA QR	Payment		
35	38PBUII	31/7/2023	15:49:49	Colo Bottl	Unbranded	Mineral Water			1				8000	0	0	8000	0	8000	0	8000	Dine In	Kasir Noble		Cash	Payment		
36	38PBUII	31/7/2023	15:49:49	Tea	Unbranded	Lychee Tea			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		Cash	Payment		
37	38PBUI5	31/7/2023	14:31:54	Tea	Unbranded	Strawberry Tea			1				25000	0	0	25000	0	25000	0	25000	Dine In	Kasir Noble		BCA QR	Payment		

Gambar 2. Dataset csv.

2. Pembersihan Data

Proses pembersihan data dilakukan untuk memastikan bahwa data yang digunakan dalam analisis tersebut akurat, konsisten, dan bebas dari kesalahan. Tahapan pembersihan data diperlukan untuk menjadikan dataset lebih konsisten dan memiliki dasar yang kuat untuk dilakukan analisis. Pembersihan data menggunakan *python pyspark* dikarenakan setelah data dibersihkan akan langsung dilaksanakan implementasi algoritma *FP-Growth*. Dengan menggunakan bantuan *pyspark* dataset dapat diolah dalam bentuk data *frame* dengan proses pembersihan data yang tertera pada Gambar 3.

```

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 3. Kode Pembersihan Data.

3. Implementasi FP-Growth

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.

```

from pyspark.sql.window import Window

#Pendeklarasian 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. Kode Pengelompokan Itemset.

Tahapan implementasi 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 minimum *support* diberikan pada angka 0.001 dengan pertimbangan akan banyak itemset yang tidak sering muncul dalam transaksi akan didapatkan. Akan tetapi dengan mengecilkan nilai minimum *support* dengan jumlah data setelah dibersihkan ada pada angka 3.423 akan mendapatkan banyak pasangan item. Kekurangan ini nantinya akan ditutup dengan menentukan minimum *confidence*, yaitu pada angka 0.5 untuk memastikan bahwa kebenaran kemunculan pasangan itemset ini bernilai 50%.

```

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 5. 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*. Dengan implementasi *library* algoritma diatas, aturan asosiasi bisa langsung didapatkan dengan mengakses *associationRules* seperti pada Gambar 6 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 6. 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 pola pembelian pelanggan yang terjadi selama 2 bulan di Noble Coffee.

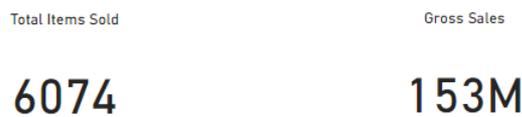
antecedent	consequent	confidence	lift	support
[Frappe, Classic C...]	[Snacks]	0.8333333333333334	4.065328665803135	0.001752848378615...
[White, Cold Bott...]	[Snacks]	0.8333333333333334	6.775534441805226	0.001468706982179375
[White, Cold Bott...]	[Classic Coffee]	0.8333333333333334	5.214880843875886	0.001468706982179375
[Main Course, Col...]	[Coffee Milk]	0.6666666666666666	1.968002329467114	0.001168565585743...
[White, Classic C...]	[Cold Bottle]	0.75	0.508827814569536	0.001752848378615...
[White, Classic C...]	[Snacks]	0.75	0.897808997624703	0.001752848378615...
[White, Cold Bott...]	[Coffee Milk]	0.75	2.13759367194085	0.001752848378615...
[White, Cold Bott...]	[Snacks]	0.75	6.897808997624703	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	3.2289482623254114	0.001752848378615...
[White, Cold Bott...]	[Classic Coffee]	0.5454545454545454	3.41328901445903	0.001752848378615...
[White, Cold Bott...]	[Coffee Milk]	0.5454545454545454	1.5546135795927636	0.001752848378615...
[White, Cold Bottle]	[Classic Coffee]	0.5	3.1280482623254114	0.00233713174486...
[White, Cold Bottle]	[Snacks]	0.6875	5.59815914489311	0.002113555369794...
[Frappe, Milk Bas...]	[White]	0.5	0.96873294293193	0.001168565585743...
[White, Frappe, M...]	[Coffee Milk]	0.8	2.2880999167369536	0.001168565585743...
[White, Frappe, C...]	[Snacks]	0.8	6.598412064239185	0.001168565585743...
[White, Snacks, C...]	[Cold Bottle]	0.75	0.508827814569536	0.001752848378615...
[White, Snacks, C...]	[Coffee Milk]	0.75	2.13759367194085	0.001752848378615...

only showing top 20 rows

Gambar 7. Aturan Asosiasi.

4. Visualisasi Data

Pada halaman pertama, visualisasi akan lebih difokuskan kepada pemahaman mendalam tentang pilihan pelanggan terhadap berbagai menu yang tersedia di Noble Coffee. 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.

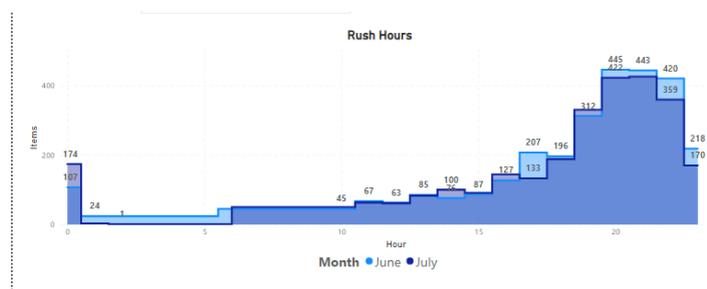


Gambar 8. Card.

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 9, 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.

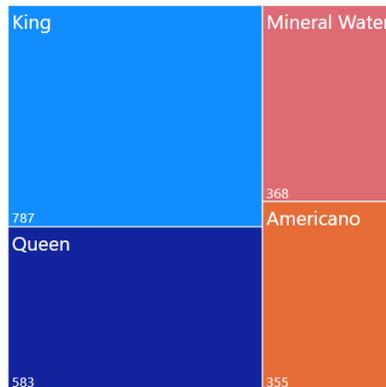


Gambar 9. Jam Sibuk.

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.

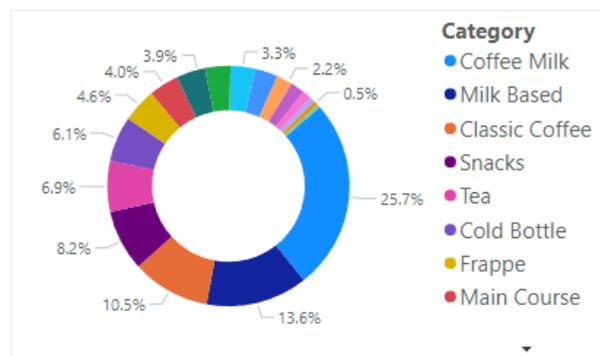
Setelah memahami pola jam sibuk, langkah selanjutnya adalah mengidentifikasi produk yang mencatat penjualan tertinggi selama periode dua bulan tersebut. Untuk tujuan ini, tree map digunakan sebagai grafik yang memvisualisasikan data dengan memperlihatkan perbedaan ukuran area. Dalam konteks visualisasi data ini, tree map digunakan untuk memberikan wawasan mendalam mengenai produk-produk yang memiliki penjualan tertinggi. Grafik ini, seperti yang terlihat pada Gambar 10, 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.



Gambar 10.1 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 11.

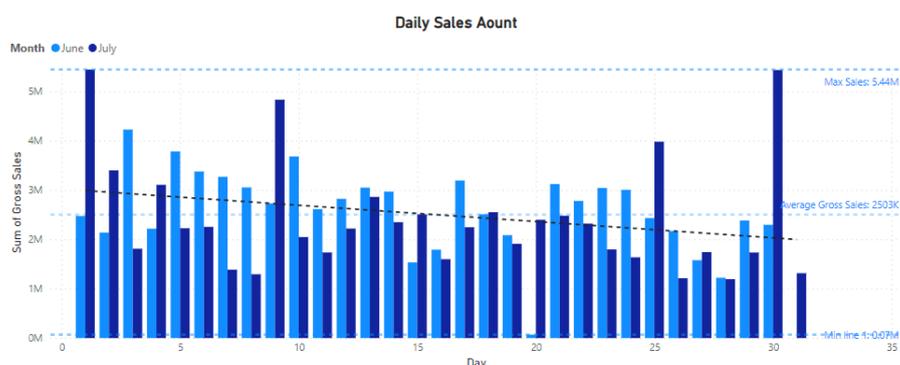


Gambar 11. Kuantitas Penjualan Kategori Tertinggi.

Dalam analisis data ini 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 12 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 12. 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.

Tabel 1. Item, Jumlah Terjual, dan Pendapatan.

Item	Sun Of Quantity	Sun Of Gross Sales
King	787	19.675M
Queen	583	13.992M
Mineral Water	368	2.944M
Amiricano	355	7.810M
Chocolate	257	6.425M
Red Velvet	232	5.800M
Café 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
Total	6074	152.663M

Halaman kedua dari penelitian ini menampilkan sejumlah grafik yang memberikan pemahaman yang mendalam tentang pola pembelian minuman pelanggan di Noble Coffee. Gambar 25 menghadirkan visualisasi data dalam bentuk grafik pie dan area yang membahas secara rinci tentang preferensi pelanggan terhadap minuman-minuman yang ditawarkan di kafe tersebut.

Grafik pie pada Gambar 13 memberikan gambaran tentang pembagian persentase penjualan minuman berdasarkan kategorinya. Sementara itu, grafik area yang disajikan pada halaman ini memberikan gambaran kronologis tentang pola pembelian minuman 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.



Gambar 13. Grafik Penjualan Minuman.

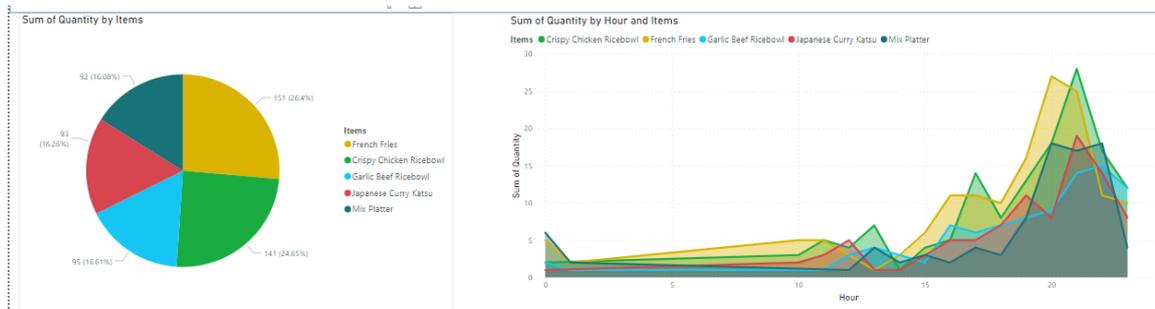
Dalam analisis minuman di Noble Coffee, terdapat lima jenis minuman yang mencatat penjualan tertinggi, yaitu *King*, *Queen*, *Mineral Water*, *Americano*, dan *Chocolate*. Masing-masing minuman ini memiliki waktu penjualan yang spesifik, menunjukkan preferensi pelanggan terkait konsumsi minuman di berbagai jam. Sebagai contoh, dari grafik di sebelah kanan, terlihat bahwa *Americano* cenderung dinikmati pada pagi hari, lebih banyak diminati dari pada 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 14. 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 relevan, yang dapat meningkatkan pengalaman pelanggan serta penjualan bisnis. Dengan pendekatan ini, Noble Coffee dapat memahami pola pembelian pelanggan secara lebih mendalam dan merancang strategi pemasaran yang lebih cerdas dan efektif.

8. *Queen* → *Mix Plater, Café Latte*

9. *Ace* → *Alpine Breeze, French Fries*

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 15.



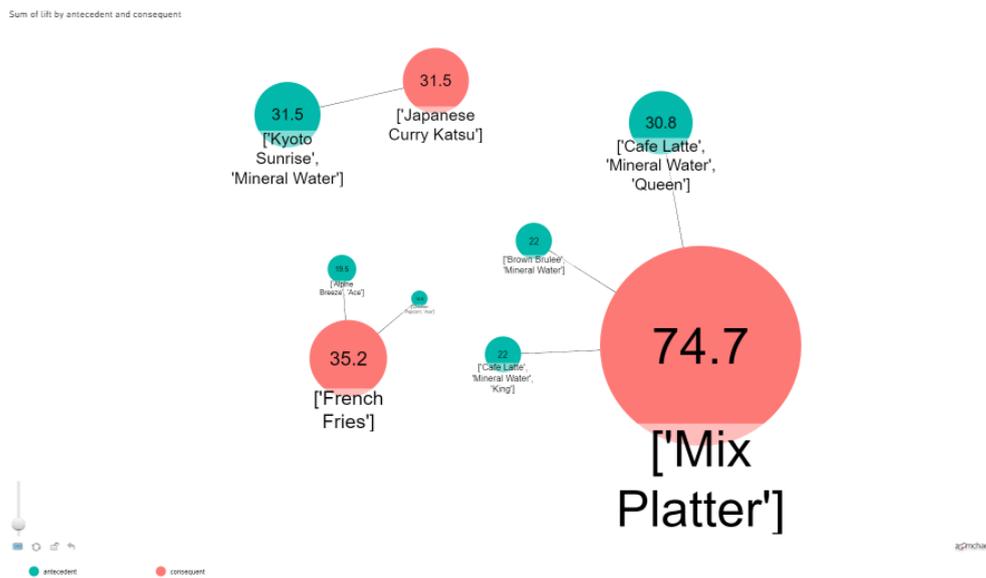
Gambar 15. Grafik Penjualan Makanan.

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.

Pola pembelian makanan 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.

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 detail pada Gambar 16, memberikan landasan yang solid untuk merencanakan promosi

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



Gambar 16. 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.

Pembahasan

Penelitian ini memiliki tujuan untuk mendapatkan informasi yang dapat membantu Noble Coffee memahami pola pembelian pelanggan. Produk diberlakukan data *mining* menggunakan algoritma *FP-Growth* untuk mendapatkan aturan asosiasi antar produknya. Algoritma *FP-Growth* adalah algoritma yang digunakan dalam *data mining* untuk *frequent pattern mining*. *FP-Growth* adalah algoritma yang didapatkan dari hasil improvisasi algoritma Apriori yang menghasilkan *itemset* dan mendapatkan *most frequent itemset*, sedangkan algoritma *FP-Growth* membuat struktur berbentuk pohon yang merepresentasikan database dari pada membuat kandidat *itemset*.

Dalam jurnal berjudul “*Data Visualization as a Communication Tool*” oleh (Archambault et al., 2015), dijelaskan bahwa setiap visualisasi memiliki ciri khas komunikasinya masing-masing. Visualisasi yang ditujukan untuk membantu pembacaan

data transaksi berisikan beragam jenis grafik mulai dari *card*, *pie chart*, *bar chart*, *table*, dan *network chart*.

Gross sales atau yang sering dikenal dengan istilah “Penjualan Kotor” diberlakukan analisis tren untuk memberikan pemahaman mendalam tentang transaksi dalam rentang waktu satu bulan pada bulan Juni dan Juli. Data ini memberikan gambaran jelas tentang perubahan perilaku pembelian pelanggan selama waktu tersebut. Pengamatan untuk mendapatkan faktor yang jelas dalam penurunan penjualan pada 2 minggu akhir pada bulan tersebut dapat dilakukan lebih lanjut.

Rush hour atau yang sering dikenal dengan istilah “Jam Sibuk” data transaksi merujuk pada informasi yang diperoleh dari data transaksi yang terjadi pada jam-jam sibuk dalam suatu hari. Agar dapat memahami dan merencanakan strategi pada jam-jam tersebut. Kenaikan jam sibuk terjadi pada jam 4 sore sampai jam 9 malam yang mana menjadi wawasan untuk dapat dijadikan strategi penjualan seperti memaksimalkan penjualan pada jam tersebut.

Adapun cara untuk memahami pola pembelian pelanggan selain dengan implementasi algoritma *FP-Growth* yaitu dengan membaca data penjualan. Dengan memvisualisasikan data penjualan, didapatkan insight berupa 6074 item terjual dengan omzet 153 juta rupiah dalam dua bulan tersebut. Akan tetapi bila dilihat secara harian, terdapat penurunan omzet harian yang dijelaskan dalam *trend line* grafik *daily sales*. Jam sibuk yang didapatkan dari data tersebut adalah pada jam 4 sore sampai 9 malam. Terdapat 4 item dengan jumlah penjualan tertinggi yaitu menu bernama *King*, *Queen*, *Mineral Water*, dan *Americano*. Serta terdapat 4 kategori dengan penjualan tertinggi yaitu kategori *Coffee Milk*, *Milk Based*, *Classic Coffee*, dan *Snacks*.

Dari hasil analisis data transaksi tersebut, ada beberapa item dan kategori yang dapat dijadikan acuan untuk dikelompokkan lebih lanjut. Acuan tersebut digunakan untuk membaca aturan asosiasi yang dihasilkan dari algoritma *FP-Growth* sebelumnya dengan melihat menu dan kategori dengan penjualan tertinggi. Dengan merujuk pada kategori penjualan tertinggi yaitu *Coffee Milk* serta mengacu pada nilai yang *support* di atas 0.001, *confidence* diatas 0.5, dan *lift* diatas 1 akan didapatkan aturan asosiasi sebagai berikut:

1. *King* → *Mix Plater*, *French Fries*
2. *King* → *Mix Plater*, *Café Latte*, *Mineral Water*

3. *King* → *Cireng, Chocolate*
4. *King* → *Mix Plater, Café Latte*
5. *King* → *Cireng, Mix Plater*
6. *King* → *Cireng, French Fries*
7. *Queen* → *Mix Plater, Café Latte, Mineral Water*
8. *Queen* → *Mix Plater, Café Latte*
9. *Ace* → *Alpine Breeze, French Fries*

SIMPULAN

Dalam penelitian ini, telah dilakukan analisis mendalam terhadap data penjualan dan preferensi pelanggan di Noble Coffee. Melalui penggunaan berbagai teknik visualisasi data, penelitian ini memberikan wawasan yang 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 telah diidentifikasi.

Hasil analisis visualisasi data menunjukkan bahwa minuman dalam kategori *Coffee Milk*, terutama produk *King* dan *Queen*, sangat diminati oleh pelanggan. Pemahaman mendalam terhadap preferensi ini memberi 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 telah 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 pola 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.

Dengan pemahaman mendalam tentang preferensi pelanggan dan pola pembelian, 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.

DAFTAR RUJUKAN

- 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>
- 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>
- Campbell, N. C. G., & Cunningham, M. T. (1983). Customer analysis for strategy development in industrial markets. *Strategic Management Journal*, 4(4), 369–380. <https://doi.org/10.1002/smj.4250040407>
- Danuri, M. (2019). *PERKEMBANGAN DAN TRANSFORMASI TEKNOLOGI DIGITAL*. <https://doi.org/https://doi.org/10.53845/infokam.v15i2.178>
- 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>
- 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>
- Lawless, H. T., & Heymann, H. (2010). *Descriptive Analysis* (pp. 227–257).
https://doi.org/10.1007/978-1-4419-6488-5_10
- 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, 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>
- Lin, W.-Y., Tseng, M.-C., & Su, J.-H. (2002). *A Confidence-Lift Support Specification for Interesting Associations Mining* (pp. 148–158). https://doi.org/10.1007/3-540-47887-6_14
- Palepu, K. G., & Healy, P. M. (2013). *Business analysis and valuation*. South-Western, Cengage Learning.
- 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>
- Shanahan, J. G., & Dai, L. (2015). Large Scale Distributed Data Science using Apache Spark. *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2323–2324. <https://doi.org/10.1145/2783258.2789993>
- Susilowaty, O. A. I., & Rukismono, M. (2021). Faktor–Faktor Yang Dipertimbangkan Dalam Memilih Moka POS Sebagai Sistem Point Of Sale di Surabaya. *Seminar Nasional Ilmu Terapan*, 5(1), B08–B08.

- 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, K., Tang, L., Han, J., & Liu, J. (2002). *Top Down FP-Growth for Association Rule Mining* (pp. 334–340). https://doi.org/10.1007/3-540-47887-6_34
- Waskom, M. (2021). seaborn: statistical data visualization. *Journal of Open Source Software*, 6(60), 3021. <https://doi.org/10.21105/joss.03021>
- 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>