

Desain Dashboard Deskriptif dan Prediktif Analitik untuk Frequent Itemset Mining serta Cross-selling pada Produk Bakery

3 April 2023

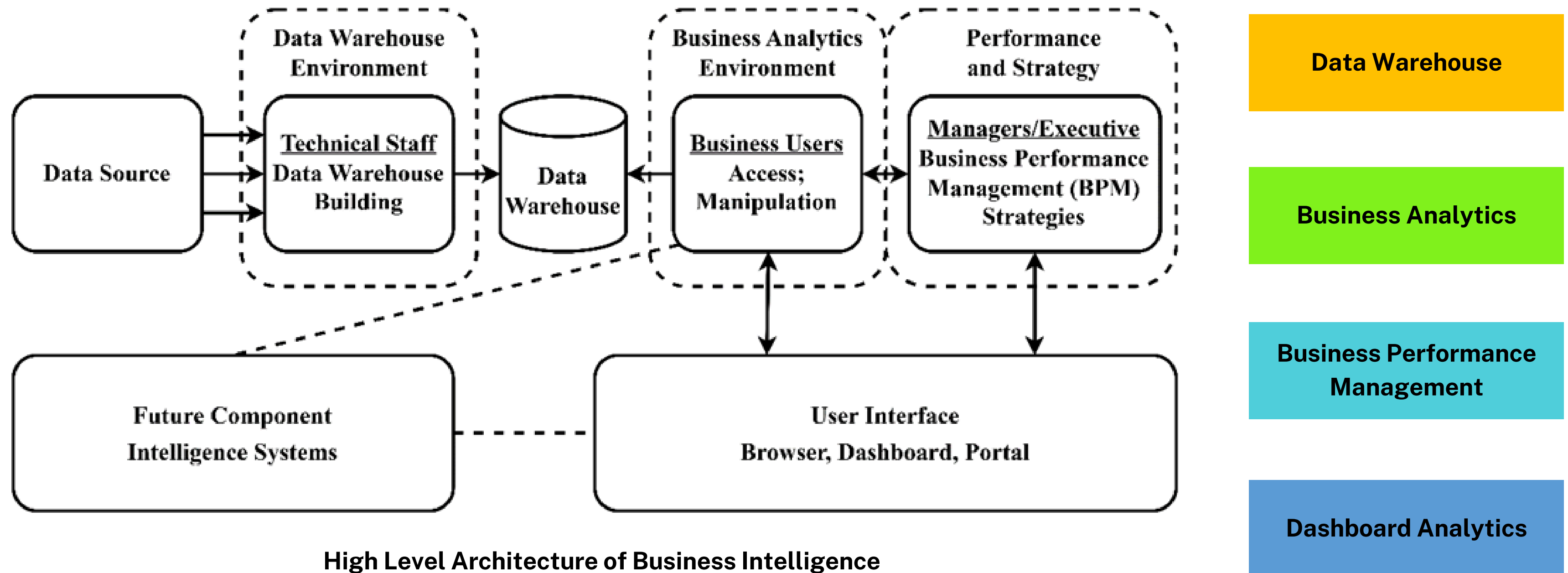
Presented by
Hary Yusuf 119190044

Content



- | | |
|--------------------------------|-------------------------------------|
| 1 Latar Belakang | 7 Informasi Dataset |
| 2 Rumusan Masalah | 8 Pemodelan SC 1 dan 2 |
| 3 Tujuan Penelitian | 9 Desain Dashboard |
| 4 Batasan Penelitian | 10 Hasil Ekstraksi Informasi |
| 5 Penelitian Terdahulu | 11 Dashboard Analitik |
| 6 Metodologi Penelitian | 12 Kesimpulan dan Saran |

Latar Belakang



[3] W. Eckerson, "Smart Companies in the 21st Century: The Secrets of Creating Successful Business Intelligent Solutions," Seattle, 2003.

Latar Belakang (2)

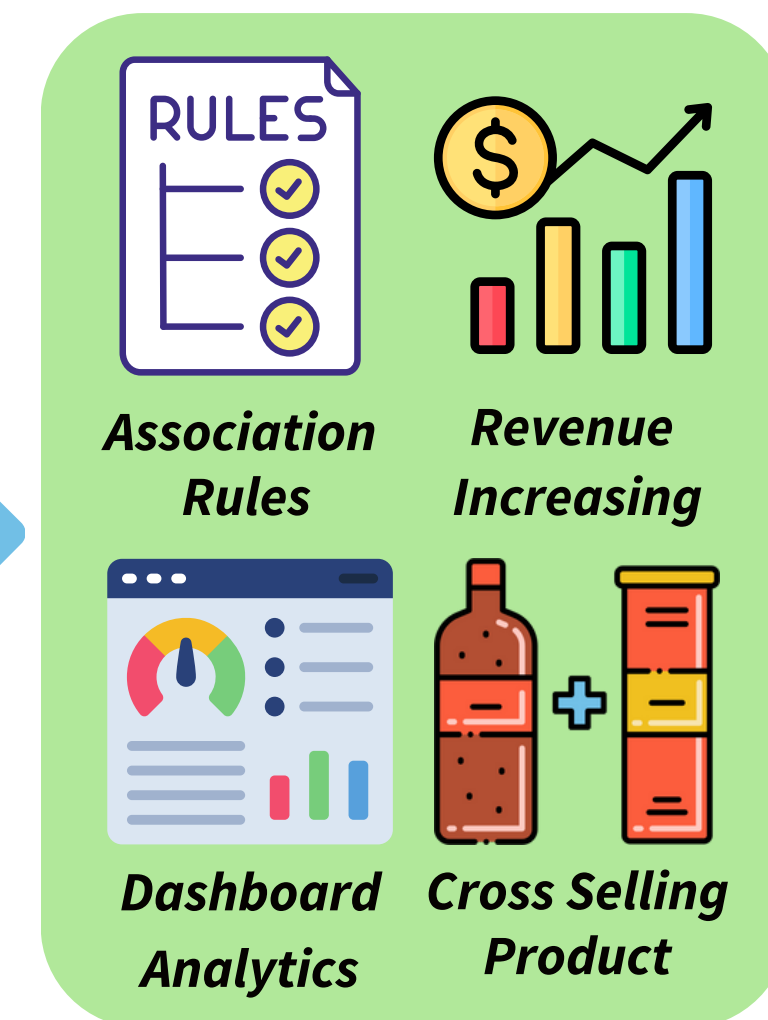
Input



Process



Output



Penelitian ini berfokus pada **pengembangan sistem pemasaran produk bakery**, dengan **memanfaatkan data transaksi** untuk **mendukung perumusan strategi pemasaran** yang disajikan dalam bentuk **dashboard analitik**

Rumusan Masalah

- 1 Perlunya **analisis entitas sistem** yang akan dikembangkan
- 2 Perlunya **ekstraksi frequent itemset** berdasarkan pola penjualan yang terjadi pada dataset produk bakery
- 3 Perlunya **penerapan aturan asosiasi** yang dihasilkan dari proses frequent itemset mining terhadap strategi pemasaran produk bakery
- 4 Perlunya **rancangan dashboard analitik** untuk mendukung kinerja pemasaran produk bakery

Tujuan Penelitian

- 1 **Menganalisis kebutuhan dan proses bisnis** dalam pengembangan sistem pemasaran produk bakery
- 2 Mengekstraksi pengetahuan (knowledge extraction) berupa **aturan asosiasi berbasis frequent itemset mining** dan **mengukur kinerja rule** tersebut
- 3 **Merancang paket cross-selling** dari hasil association rules sebagai pendukung pengambilan keputusan dalam penerapan strategi pemasaran
- 4 **Merancang dashboard deskriptif dan prediktif analitik** dari hasil knowledge extraction berbasis frequent itemset mining

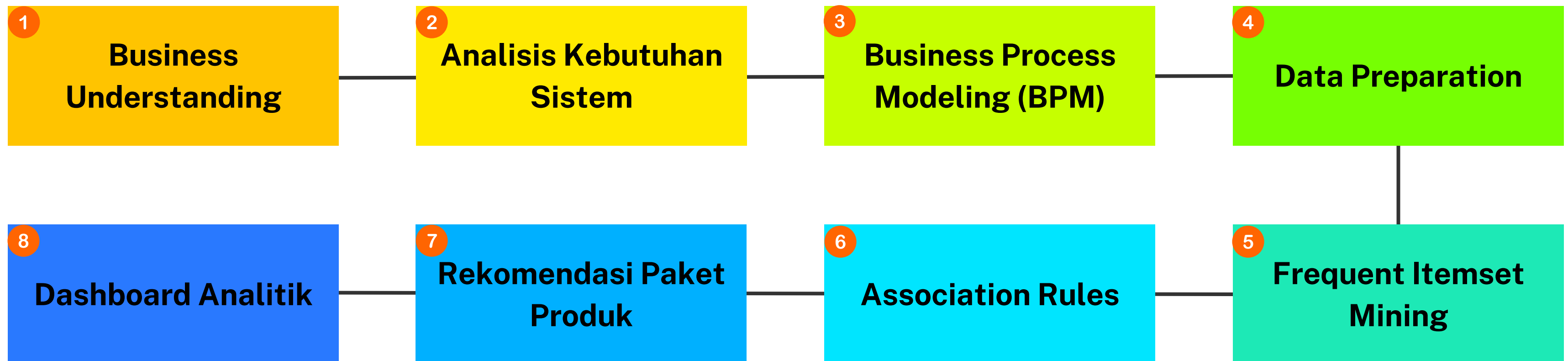
Batasan Penelitian

- 1 Penelitian menggunakan **dataset produk bakery** yang berasal dari **situs kaggle** (data source)
- 2 Penelitian yang dilakukan berfokus pada **analisis dan perhitungan data transaksi konsumen** untuk memperoleh **frequent itemset** dan **association rules**
- 3 Penelitian ini **tidak membutuhkan variabel kuantitas produk** yang dibeli per transaksi yang dilakukan konsumen **dalam memperoleh aturan asosiasi**
- 4 Hasil **visualisasi dashboard analitik** hanya menjabarkan **hasil analisis deskriptif dan prediktif** produk bakery

Penelitian Terdahulu

No	Judul	Tahun	Metode	Hasil
1	Penerapan <i>Market Basket Analysis</i> dengan Menggunakan Metode <i>Association Rule</i> Untuk Pengenalan Pola Perilaku Konsumen [6]	2022	<i>Association Rules, Algoritma Apriori</i>	Rule {Gula→Teh} dapat menjadi rekomendasi bagi Kasimura <i>Supermarket</i> karena memiliki nilai support ≥ 0.2 dan nilai confidence = 0.667
2	Analisis Pola Pembelian Konsumen pada Transaksi Penjualan Menggunakan Teknik <i>Association Rules</i> dengan Algoritma <i>Apriori</i> Studi Kasus: Ghiyas <i>Cake & Bakery</i> [10]	2022	<i>Association Rules, Algoritma Apriori</i>	Terdapat 4 <i>rules</i> dengan min_support 15% dan min_conf 70%, yaitu {soes → kue talam}, {kue talam, buka ambon → soes}, {bika ambon, soes → kue talam}, {brownies, burger → karamel}
3	Implementasi Algoritma <i>Apriori</i> pada <i>Market Basket Analysis</i> (MBA) Aksesoris Telepon Seluler [19]	2022	<i>Association Rules, Algoritma Apriori</i>	Rule {Spiral Cord, Kepala Casan Samsung→Kabel USB} memiliki nilai confidence tertinggi, yaitu 100%
4	<i>Market Basket Analysis with Equivalence Class Transformation Algorithm</i> (ECLAT) for Inventory Management Using Economic Order Quantity [17]	2022	<i>Association Rules, ECLAT Algorithm</i>	Terdapat 20 <i>rules</i> yang dihasilkan dari penggunaan algoritma ECLAT pada 212 <i>Mart</i> dengan nilai min_supp = 0.1% dan min_conf = 30%
5	<i>Analysis of Contributory Factors of Fatal Pedestrian Crashes by Mixed Logit Model and Association Rules</i> [22]	2022	<i>Association Rules</i>	Aturan asosiasi menyediakan 119 aturan valid berupa faktor penyebab kecelakaan dengan menganalisis 101,032 kasus kecelakaan pejalan kaki di Italia
6	Analisis <i>Frequent Itemset</i> dan Desain Rekomendasi Paket Produk pada Adiksi <i>Coffee</i> berbasis <i>Market Basket Analysis</i> [30]	2022	<i>Association Rules, Algoritma Apriori</i>	Terdapat 7 <i>rules</i> yang dihasilkan pada Adiksi <i>Coffee</i> berdasarkan pengelompokan data <i>all transaction</i> (semua periode waktu) dengan min_supp = 0.01, min_conf = 0.2, dan lift > 1

Metodologi Penelitian



Catatan:

- Tujuan Penelitian 1 diperoleh melalui **step 1, 2, dan 3**
- Tujuan Penelitian 2 diperoleh melalui **step 4, 5, dan 6**
- Tujuan Penelitian 3 diperoleh melalui **step 7**
- Tujuan Penelitian 4 diperoleh melalui **step 8**

Analisis Kebutuhan dan Proses Bisnis

1 Aktor yang terlibat dalam pengembangan sistem pemasaran produk bakery:

Technical Staff

Data Analyst

Manajer

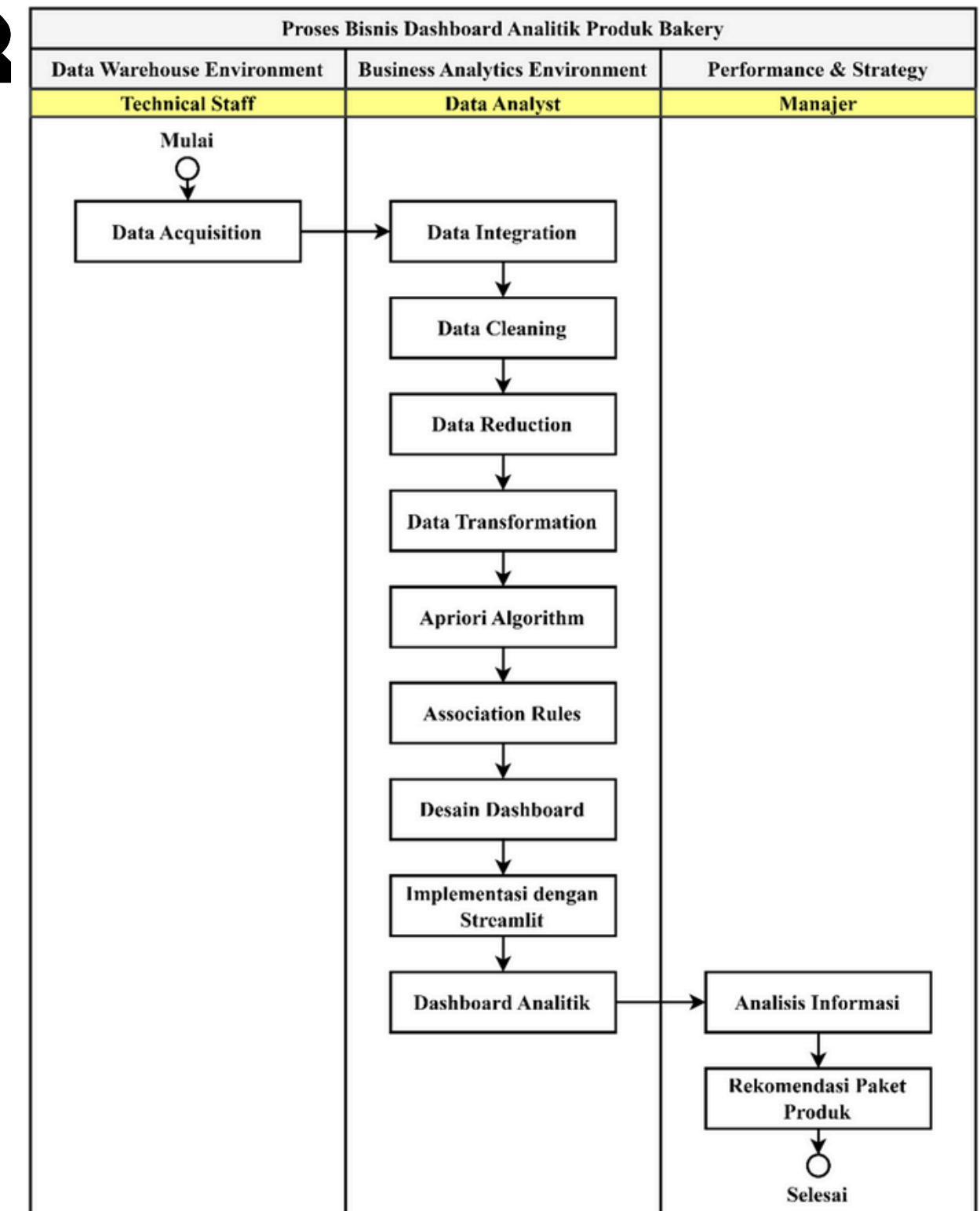
Analisis Entitas Sistem Dashboard Analitik

No	Input	Process	Output
1	Dataset produk bakery	Mengakumulasi total transaksi, total <i>item</i> yang terjual, dan <i>item</i> yang memiliki penjualan terbanyak	KPI (<i>key performance indicator</i>) berupa total transaksi, total <i>item</i> terjual, dan <i>item best seller</i>
2	Tabel <i>frequent itemset</i>	Menentukan variabel yang menjadi nilai x dan y berturut-turut, yaitu atribut <i>item</i> dan total	Total penjualan per <i>item</i> dalam bentuk <i>bar chart</i>
3	Dataset produk bakery dan dataset hasil transformasi atribut <i>Item</i>	Menampilkan dataset produk bakery dan dataset hasil transformasi	Tabel dataset penjualan dan tabel transaksi
4	Dataset <i>association rules</i>	Menentukan variabel yang menjadi nilai x, y, dan z berturut-turut, yaitu atribut <i>support</i> , <i>confidence</i> , dan <i>lift ratio</i>	Aturan asosiasi dalam bentuk <i>scatter plot</i>
5	Dataset <i>association rules</i>	Menampilkan dataset <i>association rules</i> dari hasil <i>knowledge extraction</i>	Tabel basis aturan asosiasi dan tabel aturan yang terpilih

3

terdapat 5 fitur yang menyajikan hasil analisis deskriptif (1, 2, 3) dan prediktif (4, 5) pada dataset produk bakery

2



Informasi Dataset

Tampilan dataset produk bakery

<i>Date</i>	<i>Time</i>	<i>Transaction</i>	<i>Item</i>
10/30/2016	9:58:11	1	Bread
10/30/2016	10:05:34	2	Scandinavian
10/30/2016	10:05:34	2	Scandinavian
10/30/2016	10:07:57	3	Hot chocolate
...
4/9/2017	14:32:58	9682	Coffee
4/9/2017	14:32:58	9682	Tea
4/9/2017	14:57:06	9683	Coffee
4/9/2017	14:57:06	9683	Pastry
4/9/2017	15:04:24	9684	Smoothies

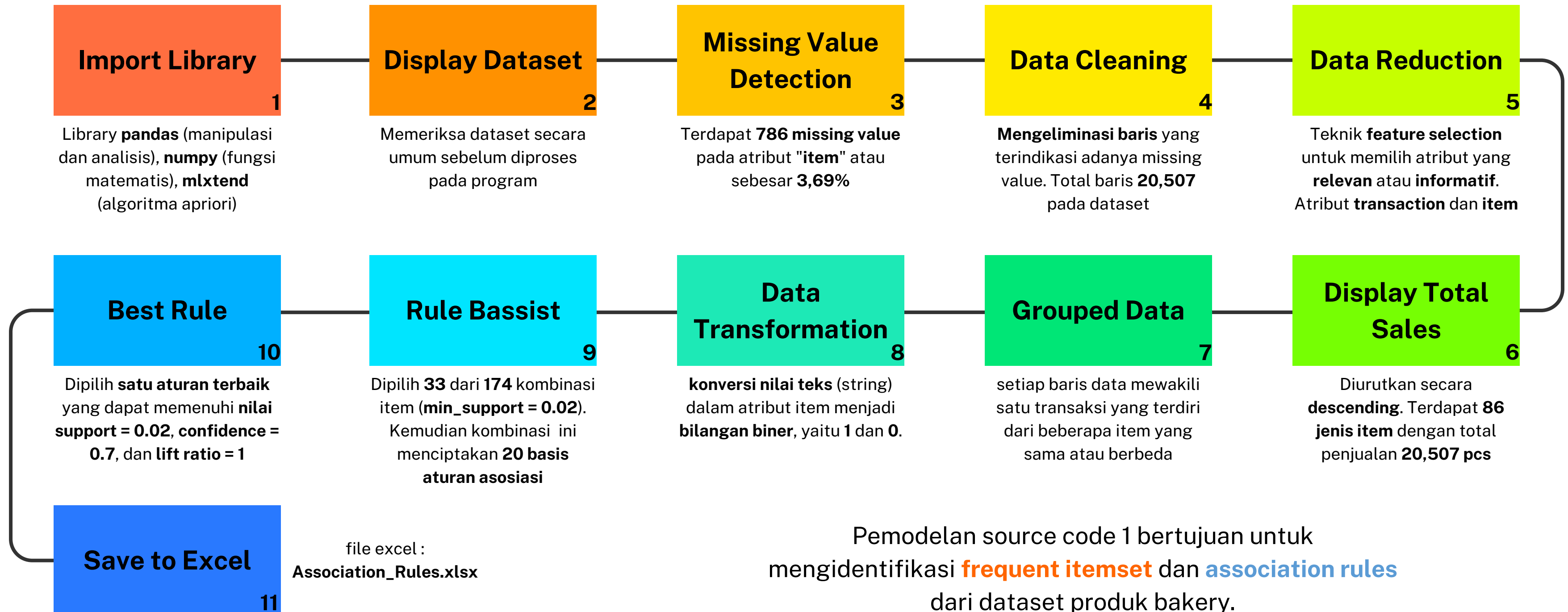
Dataset produk bakery berasal dari situs **kaggle** yang diunggah oleh **Sulman Sarwar** pada tahun **2019**.

Deskripsi Atribut

No	Atribut	Deskripsi
1	<i>Date</i>	Merupakan tanggal dari pencatatan setiap <i>item</i> yang terjual, memiliki format yang dimulai dari <i>month</i> , <i>day</i> , dan <i>year</i> (mm/dd/yyyy)
2	<i>Time</i>	Berisi waktu terjualnya setiap <i>item</i> , memiliki format AM/PM yang dimulai dari <i>hour</i> , <i>minute</i> , dan <i>second</i> (hh:mm:ss)
3	<i>Transaction</i>	Merupakan identitas nomor transaksi untuk setiap <i>item</i> yang terjual
4	<i>Item</i>	Berisi nama produk yang terjual per <i>item</i>

Memiliki **21,293 baris**, **4 kolom** (atribut), **9,684 transaksi**, dan waktu pencatatan selama **162 hari**.

Proses Pemodelan Source Code 1



Pemodelan source code 1 bertujuan untuk mengidentifikasi **frequent itemset** dan **association rules** dari dataset produk bakery.

Desain Dashboard Analitik

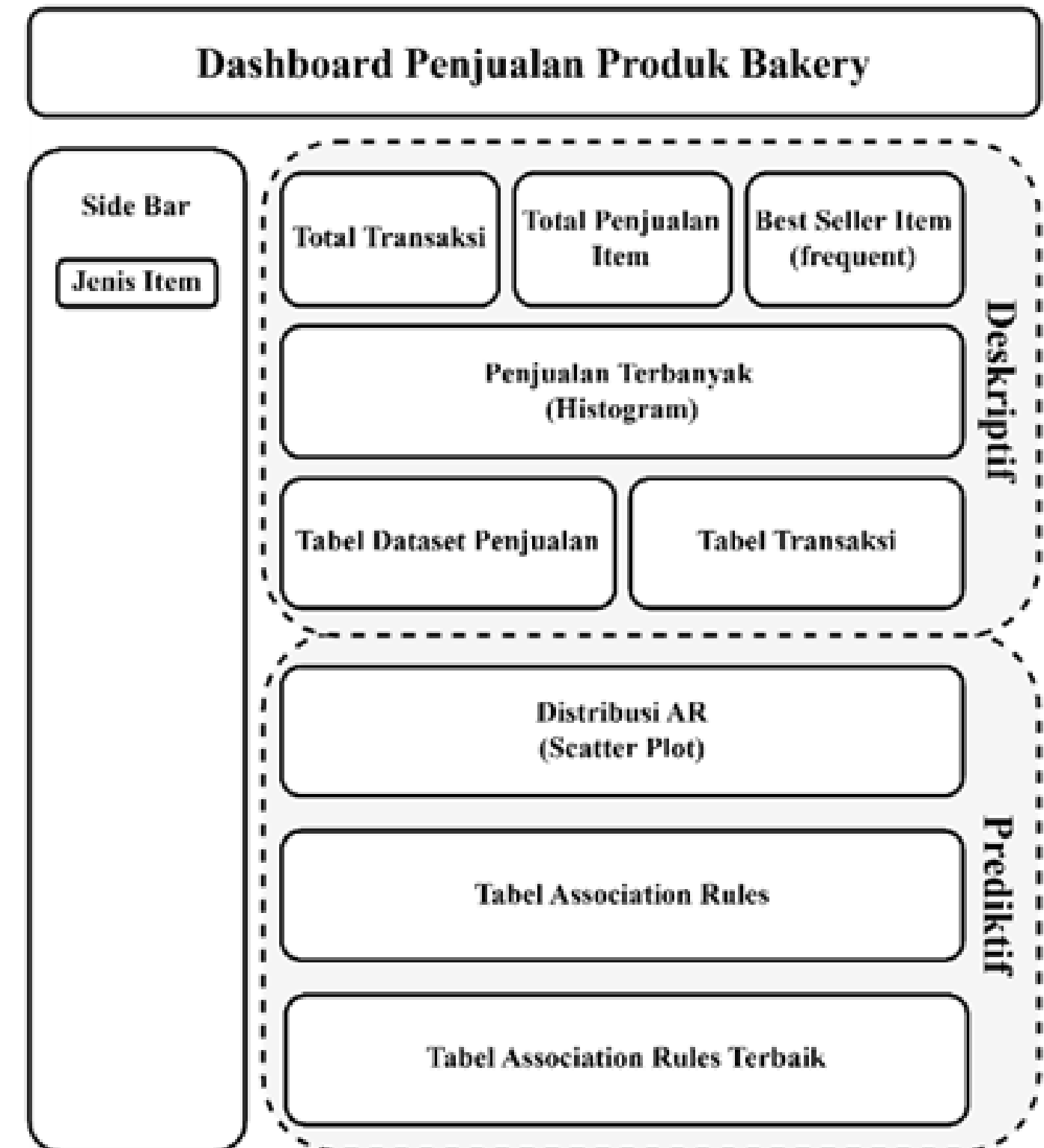
Menurut [40] terdapat elemen penting pada desain dashboard seperti **data, structure, visual representation, page layout, screen space, interaction, meta data, dan color**

Elemen tersebut **diintegrasikan** untuk menjadi satu kesatuan desain yang baik untuk **meningkatkan penggunaannya** oleh user

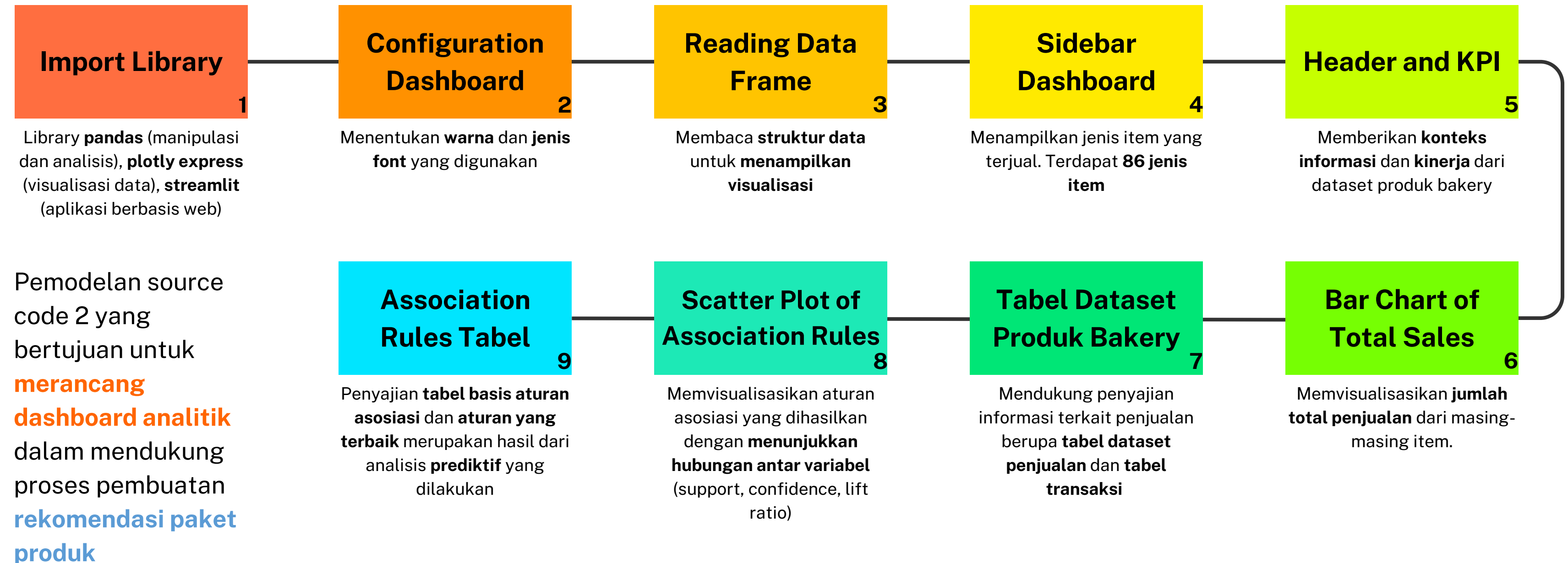
Pemilihan warna yang digunakan berasal dari **situs colorhunt.co**, yaitu memilih palet warna yang memiliki **popularitas tertinggi**

Jenis font yang digunakan adalah "**sans-serif**" karena memiliki tampilan yang jelas, mudah dibaca, modern, profesional, dan fleksibel

No	Warna	Kode Warna (HEX)	Keterangan
1		#222831	Background color
2		#393E46	Secondary background color
3		#00ADB5	Primary color
4		#EEEEEE	Text color



Proses Pemodelan Source Code 2



No	Itemsets	Support
1	{'coffee'}	0.478
2	{'bread'}	0.327
3	{'tea'}	0.143
4	{'cake'}	0.104
5	{'coffee', 'bread'}	0.090
6	{'pastry'}	0.086
7	{'sandwich'}	0.072
8	{'medialuna'}	0.062
9	{'hot chocolate'}	0.058
10	{'cake', 'coffee'}	0.055
11	{'cookies'}	0.054
12	{'tea', 'coffee'}	0.050
13	{'coffee', 'pastry'}	0.048
14	{'brownie'}	0.040
15	{'farm house'}	0.039
16	{'juice'}	0.039
17	{'muffin'}	0.038
18	{'coffee', 'sandwich'}	0.038
19	{'alfajores'}	0.036
20	{'coffee', 'medialuna'}	0.035
21	{'scone'}	0.035
22	{'soup'}	0.034
23	{'toast'}	0.034
24	{'coffee', 'hot chocolate'}	0.030
25	{'bread', 'pastry'}	0.029
26	{'scandinavian'}	0.029
27	{'cookies', 'coffee'}	0.028
28	{'tea', 'bread'}	0.028
29	{'cake', 'tea'}	0.024
30	{'coffee', 'toast'}	0.024
31	{'cake', 'bread'}	0.023
32	{'coffee', 'juice'}	0.021
33	{'truffles'}	0.020

Frequent Itemset

Frequent Itemset
dihasilkan melalui
penggunaan
algoritma apriori
dengan menetapkan
nilai minimum
support = 0.02

Penentuan aturan
asosiasi terbaik
dipilih dari perolehan
nilai confidence =
0.7, dan **nilai lift**
ratio = 1.

Rule {**toast** → **coffee**}
menjadi dasar
penentuan dari
rekomendasi paket
penjualan produk
bakery

Hasil Ekstraksi Informasi

No	Antecedents	Consequents	Support	Confidence	Lift Ratio
1	{'toast'}	{'coffee'}	0.024	0.704	1.472
2	{'medialuna'}	{'coffee'}	0.035	0.569	1.190
3	{'pastry'}	{'coffee'}	0.048	0.552	1.154
4	{'juice'}	{'coffee'}	0.021	0.534	1.117
5	{'sandwich'}	{'coffee'}	0.038	0.532	1.113
6	{'cake'}	{'coffee'}	0.055	0.527	1.102
7	{'cookies'}	{'coffee'}	0.028	0.518	1.084
8	{'hot chocolate'}	{'coffee'}	0.030	0.507	1.060
9	{'pastry'}	{'bread'}	0.029	0.339	1.035
10	{'cake'}	{'tea'}	0.024	0.229	1.605
11	{'tea'}	{'cake'}	0.024	0.167	1.605
12	{'coffee'}	{'cake'}	0.055	0.114	1.102
13	{'coffee'}	{'pastry'}	0.048	0.099	1.154
14	{'bread'}	{'pastry'}	0.029	0.089	1.035
15	{'coffee'}	{'sandwich'}	0.038	0.080	1.113
16	{'coffee'}	{'medialuna'}	0.035	0.074	1.190
17	{'coffee'}	{'hot chocolate'}	0.030	0.062	1.060
18	{'coffee'}	{'cookies'}	0.028	0.059	1.084
19	{'coffee'}	{'toast'}	0.024	0.049	1.472
20	{'coffee'}	{'juice'}	0.021	0.043	1.117
Keterangan					
			Aturan asosiasi yang terbaik		

Association Rules

Dashboard Analytics

Jenis Item:

- bread ×
- scandinavian ×
- hot chocolate ×
- jam ×
- cookies ×
- muffin ×
- coffee ×
- pastry ×
- medialuna ×
- tea ×
- tartine ×
- chocolates ×
- mineral water ×
- farm house ×
- fudge ×
- juice ×
- raspberry shortb... ×
- victorian sponge ×
- frittata ×
- hearty & seasonal ×
- soup ×

Dashboard Penjualan *Produk Bakery*

Total Transaksi

9,684

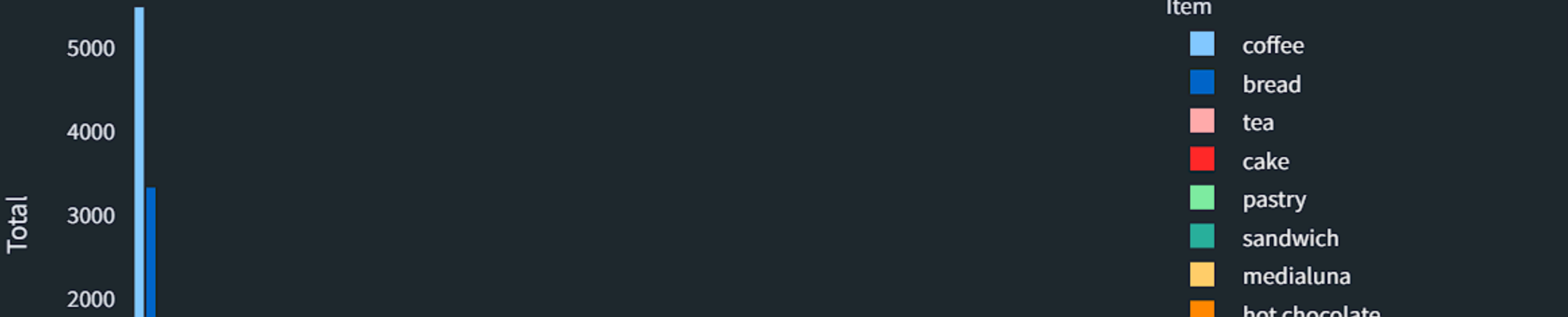
Total Item Terjual

20,507 pcs

Item Best Seller

Coffee

Total Penjualan per Item



frittata x

hearty & seasonal x

soup x

pick and mix bowls x

smoothies x cake x

mighty protein x

chicken sand x coke x

my-5 fruit shoot x

focaccia x sandwich x

alfajores x eggs x

brownie x

dulce de leche x

honey x the bart x

granola x fairy doors x

empanadas x

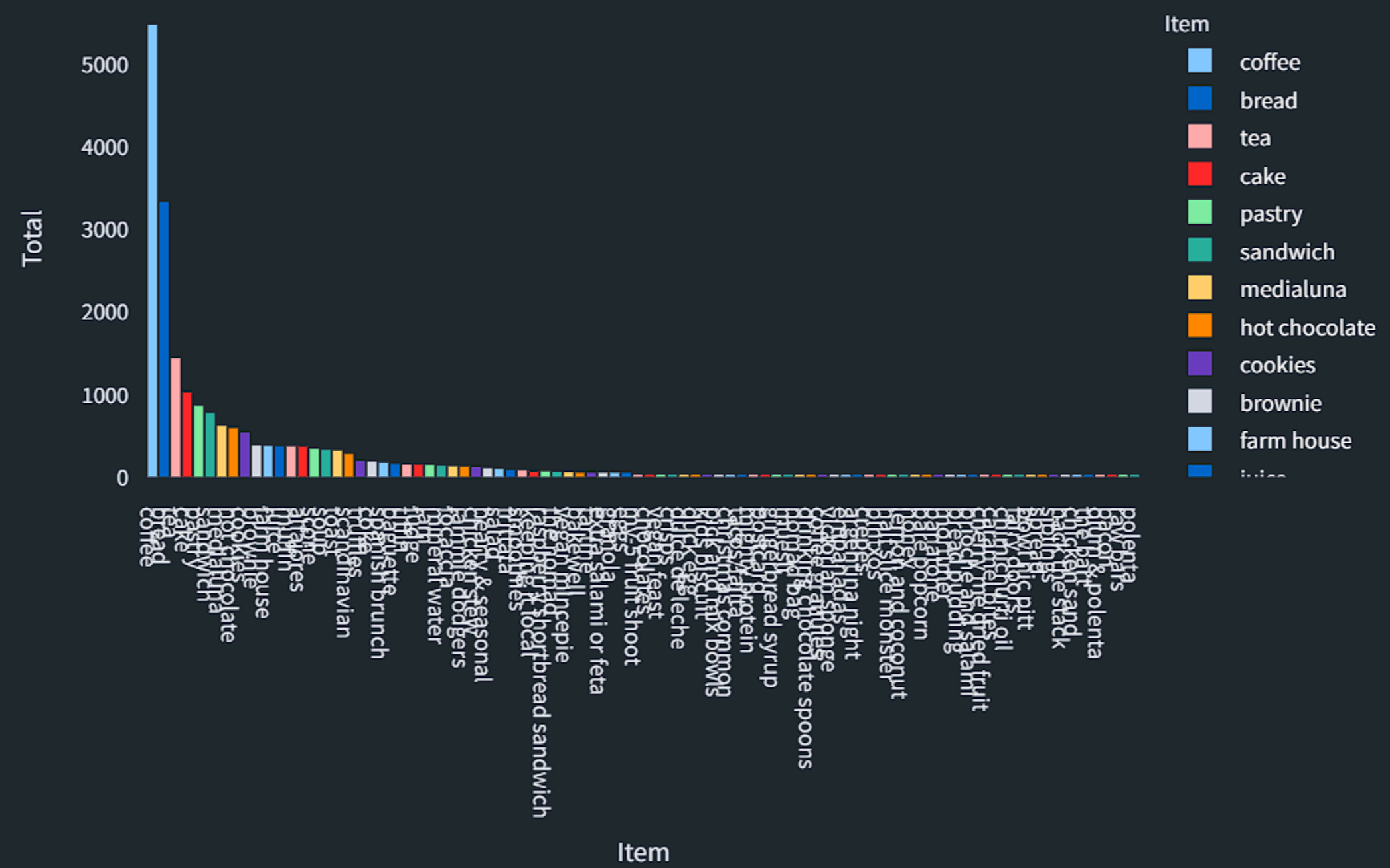
keeping it local x

bowl nic pitt x

bread pudding x

truffles x

Total Penjualan per Item



#Note:

-Pencatatan data transaksi dimulai dari tanggal 10/30/2016 sampai 4/9/2017, yang berjumlah 162 hari

-Memiliki 86 jenis item yang dijual



granola x fairy doors x

empanadas x

keeping it local x

bowl nic pitt x

bread pudding x

truffles x

chimichurri oil x

bacon x spread x

kids biscuit x siblings x x v

caramel bites x

jammie dodgers x

tiffin x

olum & polenta x

polenta x the nomad x

hack the stack x

bakewell x

lemon and coco... x

toast x scone x

crepes x

#Note:

-Pencatatan data transaksi dimulai dari tanggal 10/30/2016 sampai 4/9/2017, yang berjumlah 162 hari

-Memiliki 86 jenis item yang dijual

Tabel Dataset Penjualan

	No	Date	Time	Transaction	Item
0	1	10/30/2016	9:58:11	1	bread
1	2	10/30/2016	10:05:34	2	scandi
2	3	10/30/2016	10:05:34	2	scandi
3	4	10/30/2016	10:07:57	3	hot cho
4	5	10/30/2016	10:07:57	3	jam
5	6	10/30/2016	10:07:57	3	cookie
6	7	10/30/2016	10:08:41	4	muffin
7	8	10/30/2016	10:13:03	5	coffee
8	9	10/30/2016	10:13:03	5	pastry
9	10	10/30/2016	10:13:03	5	bread

Tabel Transaksi

	No	Transaction	Item
0	1	1	bread
1	2	2	scandinavian,scandinavian
2	3	3	hot chocolate,jam,cookies
3	4	4	muffin
4	5	5	coffee,pastry,bread
5	6	6	medialuna,pastry,muffin
6	7	7	medialuna,pastry,coffee,tea
7	8	8	pastry,bread
8	9	9	bread,muffin
9	10	10	scandinavian,medialuna



granola x fairy doors x

empanadas x

keeping it local x

bowl nic pitt x

bread pudding x

truffles x

chimichurri oil x

bacon x spread x

kids biscuit x siblings x x v

caramel bites x

jammie dodgers x

tiffin x

olum & polenta x

polenta x the nomad x

hack the stack x

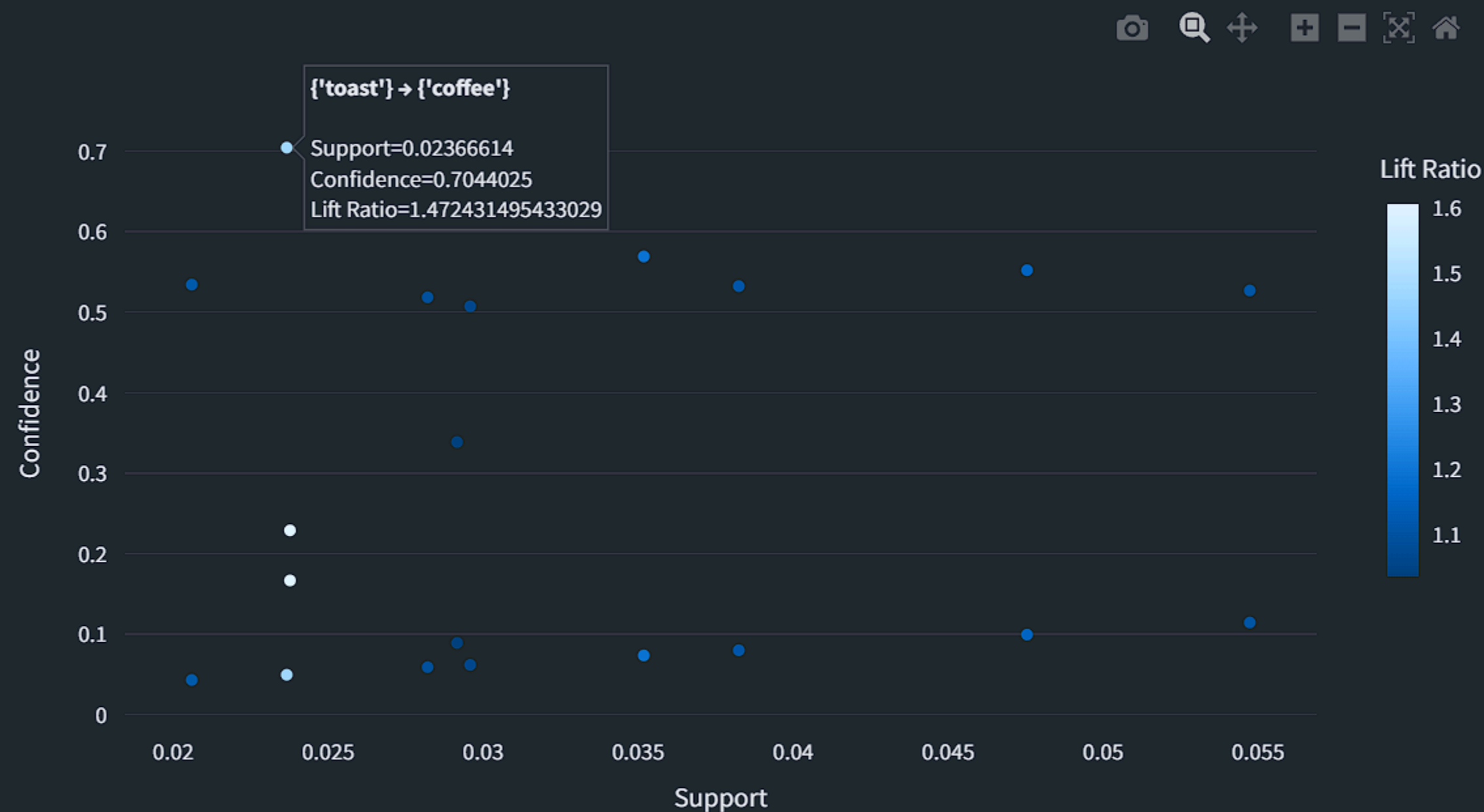
bakewell x

lemon and coco... x

toast x scone x

crepes x

Scatter Plot Association Rules



#Note:

-Nilai minimum support = 0.02

-Nilai minimum lift ratio = 1

-Nilai minimum confidence = 0.7



brioche and salami x

salad x

chicken stew x

spanish brunch x

extra salami or feta x

duck egg x

baguette x

vegan feast x

postcard x

nomad bag x

coffee granules x

drinking chocola... x

christmas comm... x

argentina night x

half slice monster x

cherry me dried ... x

mortimer x

raw bars x

tacos/fajita x

1	2	{'pastry'}	{'bread'}	0.0861	0.3272	0.0292	0.3387	1.035
2	3	{'cake'}	{'coffee'}	0.1039	0.4784	0.0547	0.527	1.1015
3	4	{'coffee'}	{'cake'}	0.4784	0.1039	0.0547	0.1144	1.1015
4	5	{'cake'}	{'tea'}	0.1039	0.1426	0.0238	0.2289	1.6048
5	6	{'tea'}	{'cake'}	0.1426	0.1039	0.0238	0.1667	1.6048
6	7	{'cookies'}	{'coffee'}	0.0544	0.4784	0.0282	0.5184	1.0837
7	8	{'coffee'}	{'cookies'}	0.4784	0.0544	0.0282	0.059	1.0837
8	9	{'coffee'}	{'hot chocolate'}	0.4784	0.0583	0.0296	0.0618	1.0603
9	10	{'hot chocolat	{'coffee'}	0.0583	0.4784	0.0296	0.5072	1.0603

Tabel Association Rules Terbaik

	No	Antecedents	Consequents	Antecedent Support	Consequent Support	Support	Confidence	Lift Ratio
0	20	{'toast'}	{'coffee'}	0.0336	0.4784	0.0237	0.7044	1.4724

#Note:

-Antecedents = Item A

-Consequents = Item B

-Rules = Jika membeli A, maka akan membeli B



Kesimpulan

- 1 Pengembangan sistem pemasaran produk bakery terletak pada **pemanfaatan data transaksi** untuk mendukung **pengambilan keputusan** yang disajikan pada **dashboard analitik**
- 2 Dataset produk bakery hanya **menghasilkan satu aturan asosiasi terbaik** yang memenuhi ukuran kinerja rule dari total 20 aturan yang telah diciptakan, yaitu **jika membeli toast maka akan membeli coffee**
- 3 Rule {toast → coffee} menjadi **dasar penentuan** dari **rekomendasi paket penjualan produk bakery**
- 4 Dashboard analitik produk bakery menyajikan informasi terkait hasil **analisis deskriptif** dan **prediktif**.

Saran

Penelitian selanjutnya dapat menggunakan dataset yang berasal dari industri bakery yang memungkinkan untuk dilakukannya **penerapan hasil penelitian secara realistis**

Penelitian selanjutnya diharapkan mampu untuk **mengintegrasikan API** dalam **penggunaan dashboard analitik**



CONTACT ME | CLICKABLE



Hary Yusuf



Hary Yusuf



+62 813 8652 2811



haryyusuf11@gmail.com