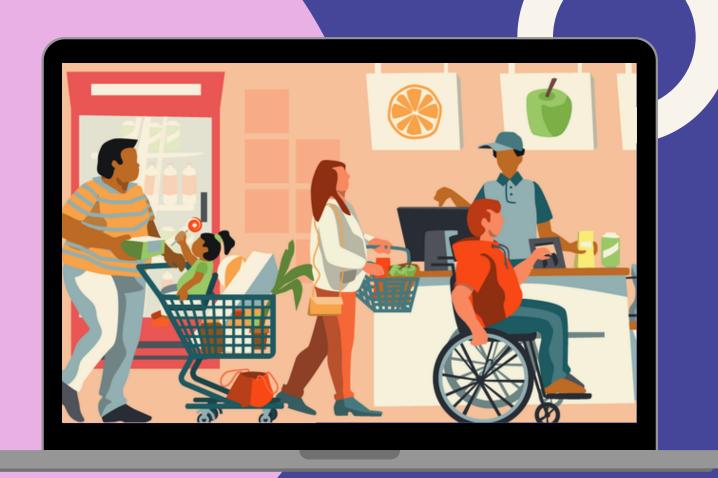


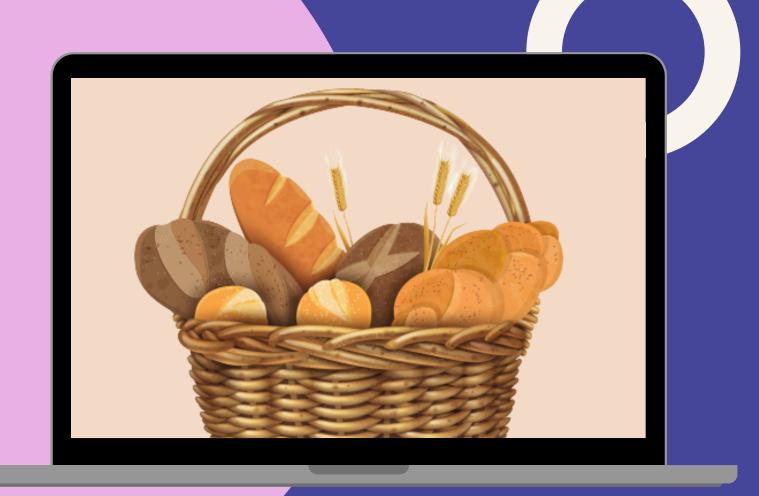
#### Anggota:

Daniel Bernard Sahala Simamora
Muhammad Satrio Pinoto Negoro
Nadhif Ramadhan
Ridwan Dwi Adani
Rudi Suwandi Putra



### Background

Indonesia is recorded to have a total of 64.2 million MSMEs in 2021 (Ministry of Finance of the Republic of Indonesia). However, most of the developing MSMEs have not used data as a weapon to market their products. MSME actors often experience a situation where the products sold are unequal, some products are quickly sold out and sold out, but not a few are also unsold products. For this reason, it is important for MSMEs to make a mature strategy for their products to be sold in order to maximize revenue.



### Business Understanding

The Bread is a business engaged in Food & Baverage. At this time The Bread is a cafe that is experiencing growth. However, The Bread has the problem of uneven sales of its products. There are products that sell out quickly and run out, but not a few products that never sell.

Our Data Science team will help analyze which products are often purchased together, so The Bread can sell them in bundling and can increase sales and profit.



### Problem 1

At what time does the maximum number of transactions occur?

### Problem 2

What products do customers often buy at the same time?

### PROBLEM



### Objective

Determine which products customers buy at the same time

### Benefits

Maximizing sales and revenue from The Bread

# OBJECTIVES AND BENEFITS



# DATA UNDERSTANDING

From the data provided, the data collection, description and exploration process will be carried out (Exploratory Data Analysis)





	Transaction	ltem	date_time	period_day	weekday_weekend
0	1	Bread	30-10-2016 09:58	morning	weekend
1	2	Scandinavian	30-10-2016 10:05	morning	weekend
2	2	Scandinavian	30-10-2016 10:05	morning	weekend
3	3	Hot chocolate	30-10-2016 10:07	morning	weekend
4	3	Jam	30-10-2016 10:07	morning	weekend

#### The dataset consists of columns containing:

- Transaction: a unique order code for each customer who makes a purchase.
- Item: products in the form of food and beverages, such as coffee and sandwiches.
- date\_time: when the order is placed, in the format dd-mm-yyyy, hh: mm.

- period\_day: shows the time background (morning, afternoon, evening, and night)
- weekday\_weekend: indicates weekdays or weekends

```
Data columns (total 5 columns):
     Column
                      Non-Null Count
 #
                                       Dtype
     Transaction
                      20507 non-null
                                       int64
 0
                                       object
     Item
                      20507 non-null
                                       object
     date time
                      20507 non-null
     period day
                                       object
                      20507 non-null
     weekday weekend
                                       object
 4
                      20507 non-null
```

It can be seen that the dataset only has one numeric feature, namely Transaction. In the dataset too, there is no missing data (NaN), so there is no need to do this missing value handling.

	Transaction
count	20507.000000
mean	4976.202370
std	2796.203001
min	1.000000
25%	2552.000000
50%	5137.000000
75%	7357.000000
max	9684.000000

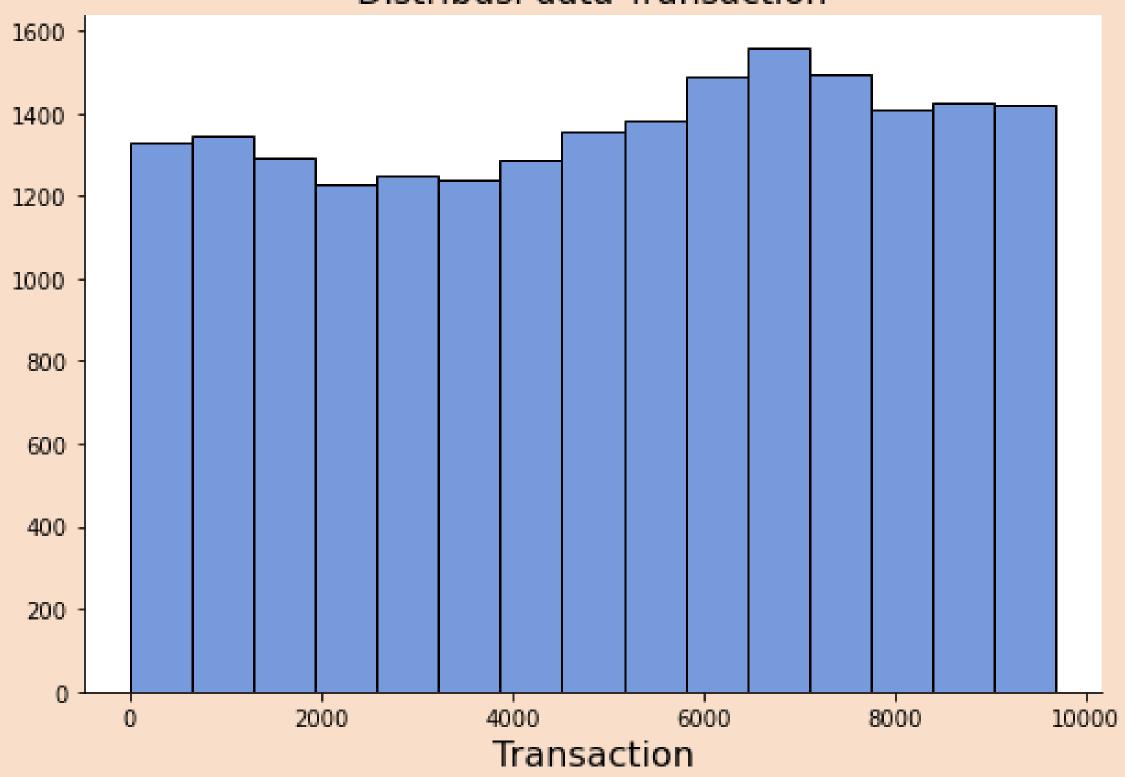
It can be seen that the Transaction data has a fairly good distribution because there is no significant spike between the quartile values, but the standard deviation still tends to be too large, which indicates that many values are far from the mean.

### DATA VISUALIZATION

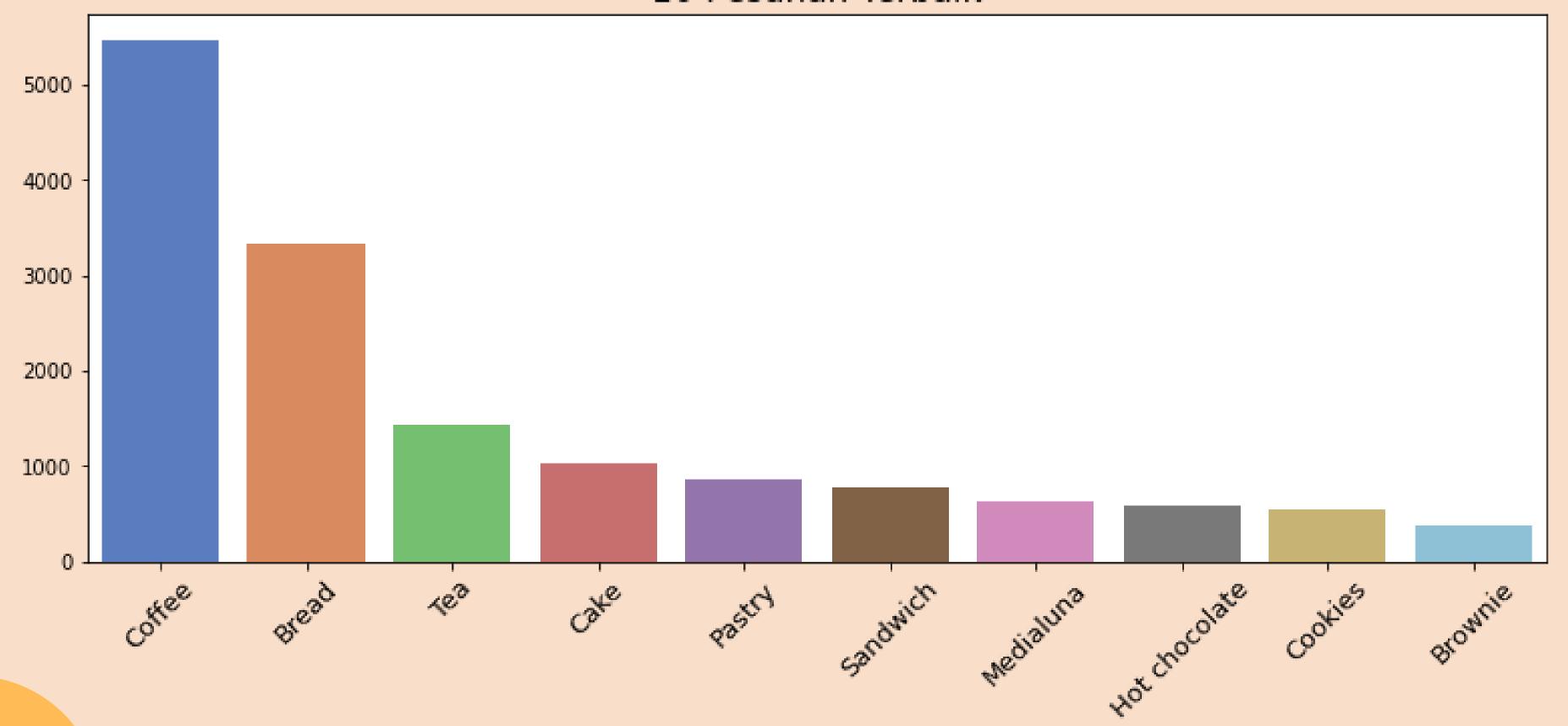
- Transaction data distribution
- Top 10 Orders
- Number of Transactions Each Month
- Total Orders Per Day
- Hourly Order Quantity
- Top 10 Orders in the Morning, Afternoon, Evening, and Evening



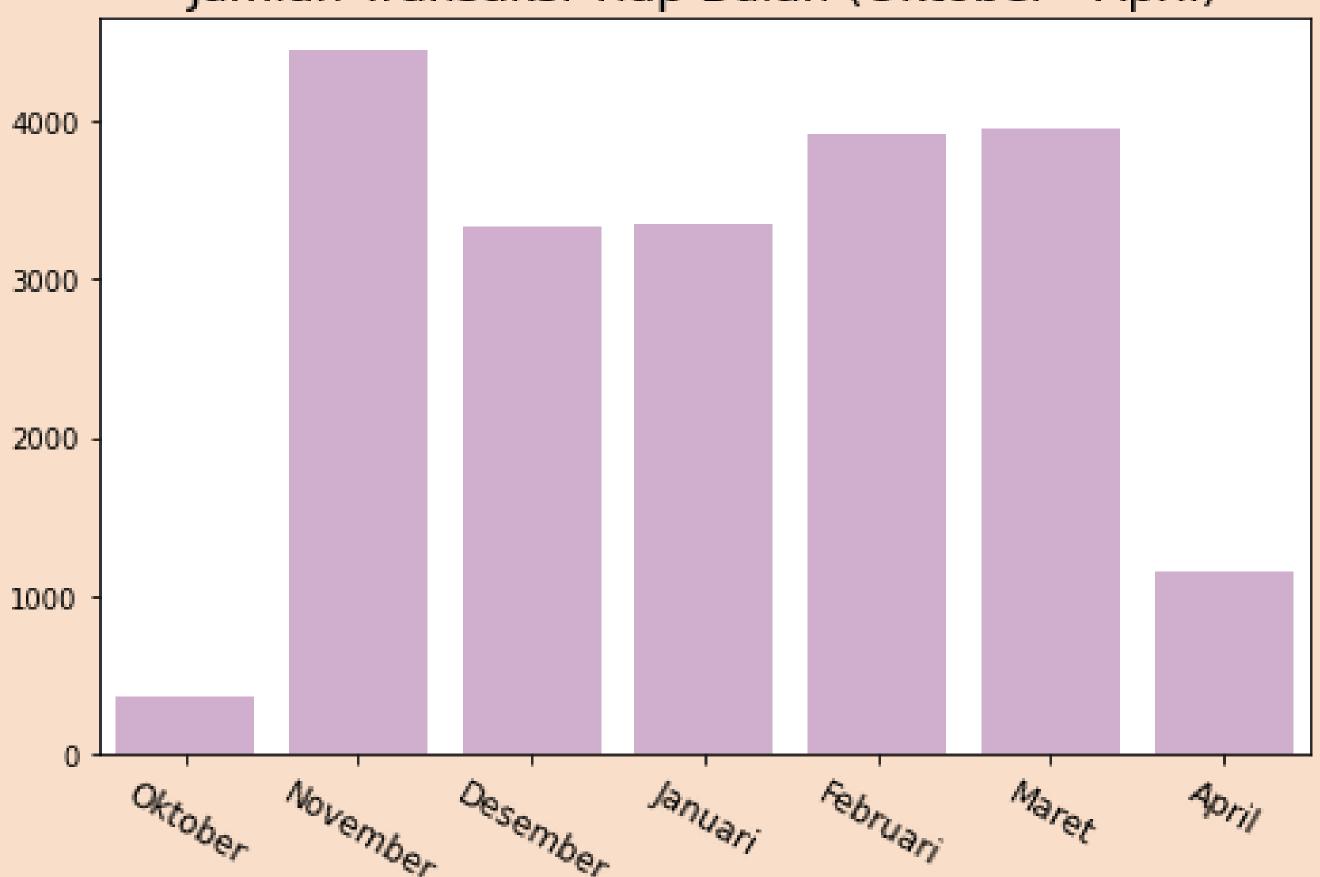
#### Distribusi data Transaction



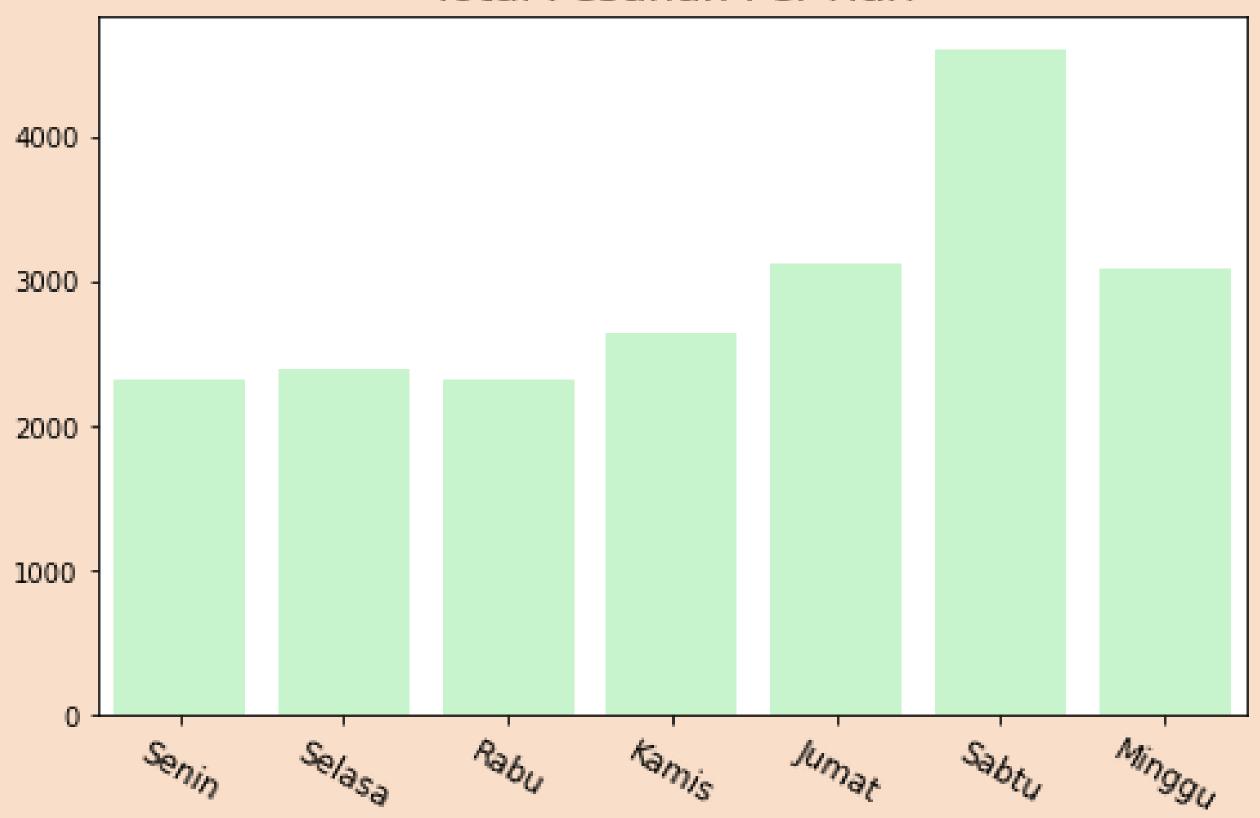
#### 10 Pesanan Terbaik



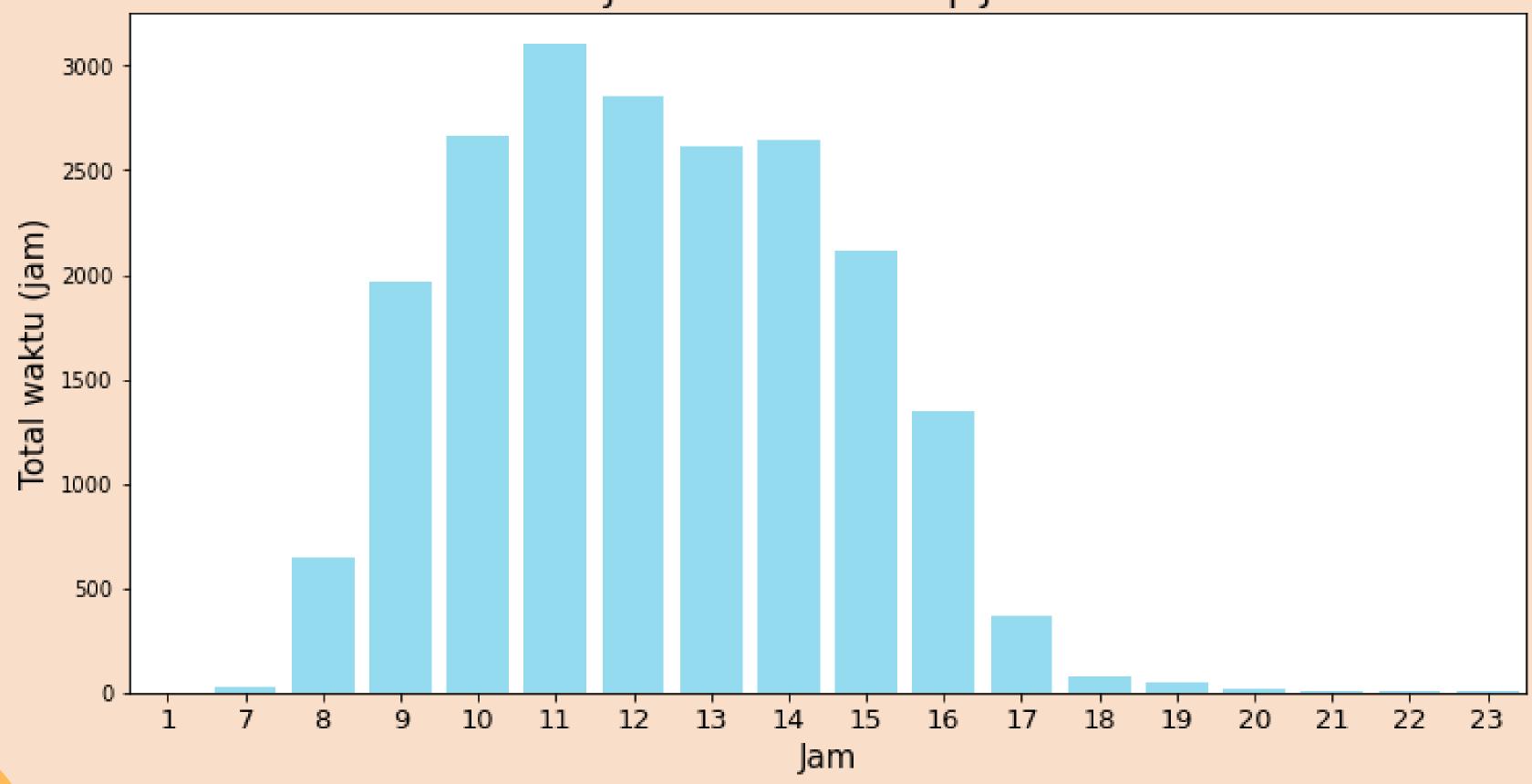
Jumlah Transaksi Tiap Bulan (Oktober - April)



#### Total Pesanan Per Hari



Jumlah Pesanan Tiap Jam





# DATA PREPARATION

Furthermore, in Data Preparation, a little data integration will be carried out





### Data Integration to be carried out:

- Remove whitespace
- Apply each string in lowercase form.

	Transaction	ltem
0	1	bread
1	2	scandinavian
2	2	scandinavian
3	3	hot chocolate
4	3	jam
5	3	cookies
6	4	muffin
7	5	coffee
8	5	pastry
9	5	bread

# MODELING



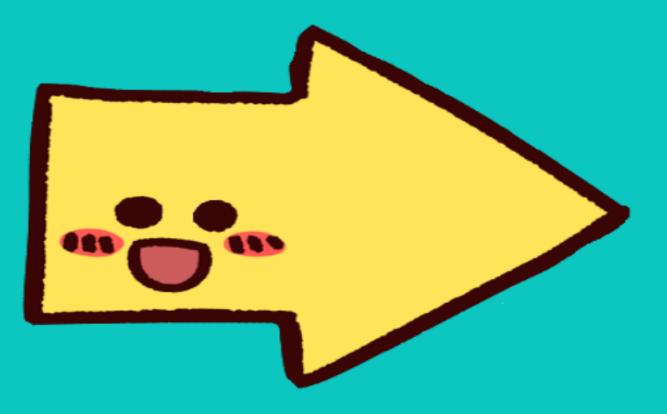


# Apriori Algorithm

#### 3 Metrics used:

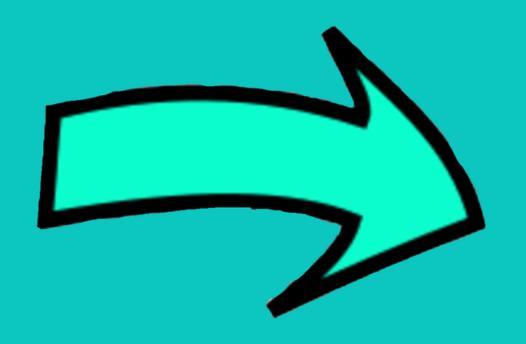
- Support
- Confidence
- Lift

# Build model!



# Group data against transaction number and item type:

	Transaction	ltem
0	1	bread
1	2	scandinavian
2	2	scandinavian
3	3	hot chocolate
4	3	jam
20502	9682	coffee
20503	9682	tea
20504	9683	coffee
20505	9683	pastry
20506	9684	smoothies



	Transaction	ltem	Count
13072	6850	coffee	4
13170	6887	coffee	4
12457	6560	coffee	4
10179	5455	coffee	3
7501	4035	bread	3
			***
6496	3487	bread	1
6497	3487	juice	1
6498	3488	coffee	1
6499	3489	cake	1
18886	9684	smoothies	1

Pivot table against transaction numbers and items:

	Transaction	Item	Count
13072	6850	coffee	4
13170	6887	coffee	4
12457	6560	coffee	4
10179	5455	coffee	3
7501	4035	bread	3
6496	3487	bread	1
6497	3487	juice	1
6498	3488	coffee	1
6499	3489	cake	1
18886	9684	smoothies	1

ltem	adjustment	afternoon with the baker	alfajores	argentina night	art tray	bacon	baguette	bakewell	bare popcorn	basket	 the bart	the nomad	tiffin	toast	truffles	tshirt
Transaction																
1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
2	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
3	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
9680	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
9681	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	1.0	0.0
9682	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
9683	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0
9684	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	 0.0	0.0	0.0	0.0	0.0	0.0

9465 rows × 94 columns

## Perform encoding

Item	adjustment	afternoon with the baker	alfajores	argentina night	art tray	bacon	baguette	bakewell	bare popcorn	basket	 the bart	the nomad	tiffin	toast	truffles	tshirt
Transaction											 					
1	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
9680	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
9681	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	1	0
9682	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
9683	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0
9684	0	0	0	0	0	0	0	0	0	0	 0	0	0	0	0	0

9465 rows × 94 columns

## Apriori model initialization

### Support = 1%

	support	itemsets
6	0.478394	(coffee)
2	0.327205	(bread)
26	0.142631	(tea)
4	0.103856	(cake)
34	0.090016	(coffee, bread)
19	0.086107	(pastry)
21	0.071844	(sandwich)
16	0.061807	(medialuna)
12	0.058320	(hot chocolate)
42	0.054728	(coffee, cake)

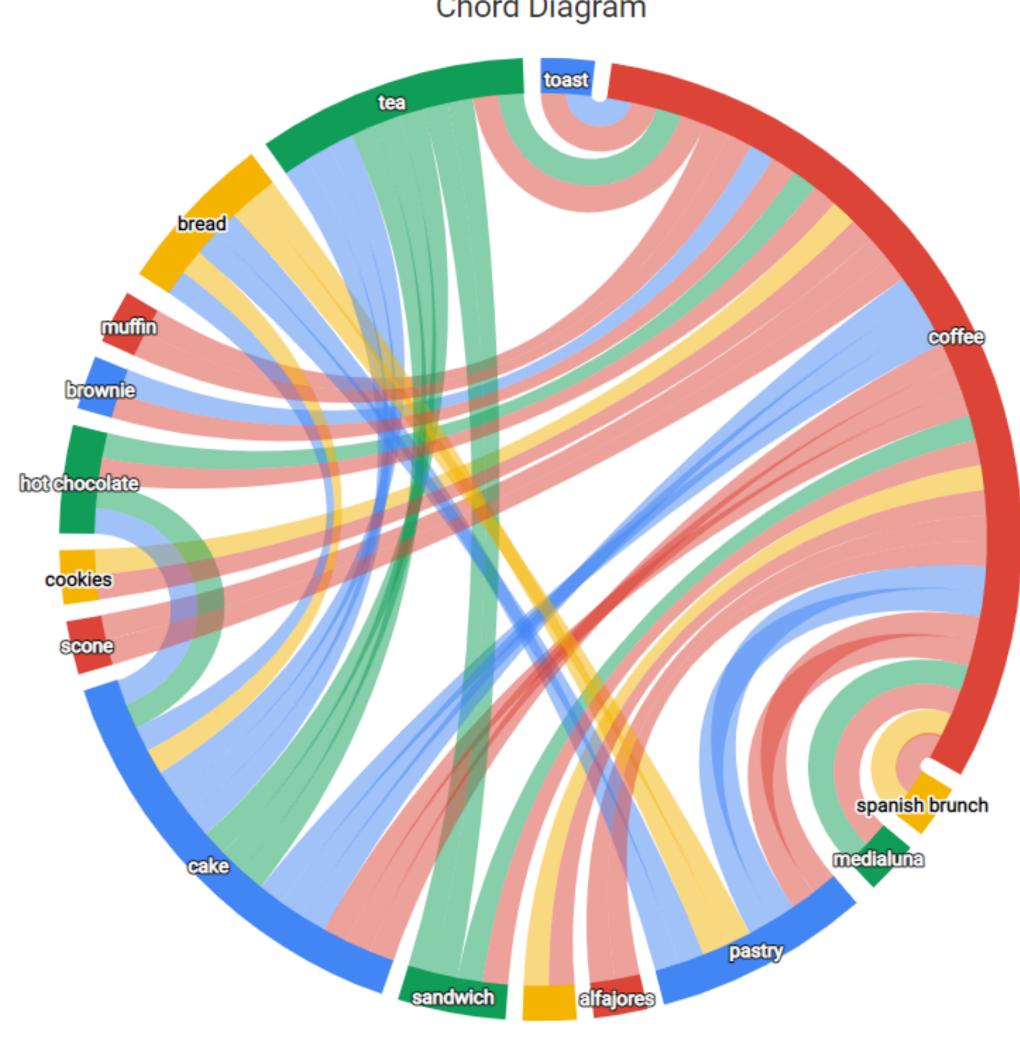
# Group by Association Rules

Metric: lift
Minimum Threshold: 1

	antecedents	consequents	support	confidence	lift
31	(toast)	(coffee)	0.023666	0.704403	1.472431
29	(spanish brunch)	(coffee)	0.010882	0.598837	1.251766
19	(medialuna)	(coffee)	0.035182	0.569231	1.189878
23	(pastry)	(coffee)	0.047544	0.552147	1.154168
1	(alfajores)	(coffee)	0.019651	0.540698	1.130235
16	(juice)	(coffee)	0.020602	0.534247	1.116750
25	(sandwich)	(coffee)	0.038246	0.532353	1.112792
7	(cake)	(coffee)	0.054728	0.526958	1.101515
27	(scone)	(coffee)	0.018067	0.522936	1.093107
13	(cookies)	(coffee)	0.028209	0.518447	1.083723
15	(hot chocolate)	(coffee)	0.029583	0.507246	1.060311
4	(brownie)	(coffee)	0.019651	0.490765	1.025860
21	(muffin)	(coffee)	0.018806	0.489011	1.022193
3	(pastry)	(bread)	0.029160	0.338650	1.034977
10	(cake)	(tea)	0.023772	0.228891	1.604781
39	(coffee, tea)	(cake)	0.010037	0.201271	1.937977

# EVALUATION

#### **Chord Diagram**



### Recommendations

### **Coffee Bundling**

Toast + Coffee

### Tea Bundling

Cake + Tea

### **Drink Bundling**

Coffe + Hot Chocolate

### **Snack Bundling**

Pastry + Bread

