CUSTOMER SEGMENTATION

Yodi Ramadhani Alfariz

github.com/yodialfa

Bussines Problem

- Offering, diskon dan strategi marketing yang kurang terfokus memungkinkan memakan biaya yang sangat besar.
- Diperlukannya segmentasi untuk setiap customer agar dapat melihat pola dari setiap customers.
- Segmentasi adalah langkah awal untuk Customer Relationship Management.

Objectives

Clustering

Mengcluster Customer Berdasarkan RFM Analisis sehingga kelas tiap customer terfokus ke feature tertentu

Recommendations

Improvement terhadap kekurangan pada transaksi seperti Marketing Campaign pada cluster tertentu

Customer Segmentation

Data Understanding

Dataset consist of : 541.909 rows and 8 columns

No.	Columns	Details
1	InvoiceNo	transaction number code
2	StockCode	item code
3	Description	item description
4	Quantity	number of items purchased
5	InvoiceDate	purchase date
6	UnitPrice	item price
7	CustomerID	customers code
8	Country	country of customers

Data Cleansing

```
InvoiceNo            0
StockCode            0
Description       1454
Quantity             0
InvoiceDate          0
UnitPrice            0
CustomerID      135080
Country              0
dtype: int64
```

Filling Description with same stockcode

For CustomerID we have to focus for customer segmentation, so we drop it

RFM Analysis

Recency

Variabel untuk mengukur nilai pelanggan berdasarkan rentang waktu (tanggal, bulan, tahun) transaksi terakhir pelanggan sampai saat ini

Frequency

Mengacu pada berapa kali pelanggan berinteraksi dengan produk atau melakukan transaksi dalam periode waktu tertentu.

Monetary

Menunjukkan jumlah total yang dihabiskan oleh pelanggan untuk membeli produk Anda dalam periode waktu tertentu.

Data Preprocessing with RFM

Recency

	CustomerID	LastPurchaseDate
0	12346.0	2011-01-18
1	12347.0	2011-12-07
2	12348.0	2011-09-25
3	12349.0	2011-11-21
4	12350.0	2011-02-02

	CustomerID	recency
0	12346.0	325
1	12347.0	2
2	12348.0	75
3	12349.0	18
4	12350.0	310

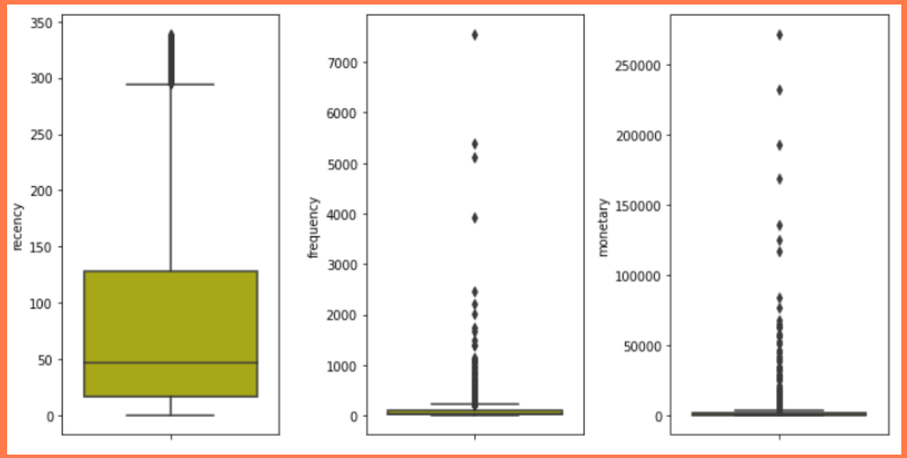
Frequency

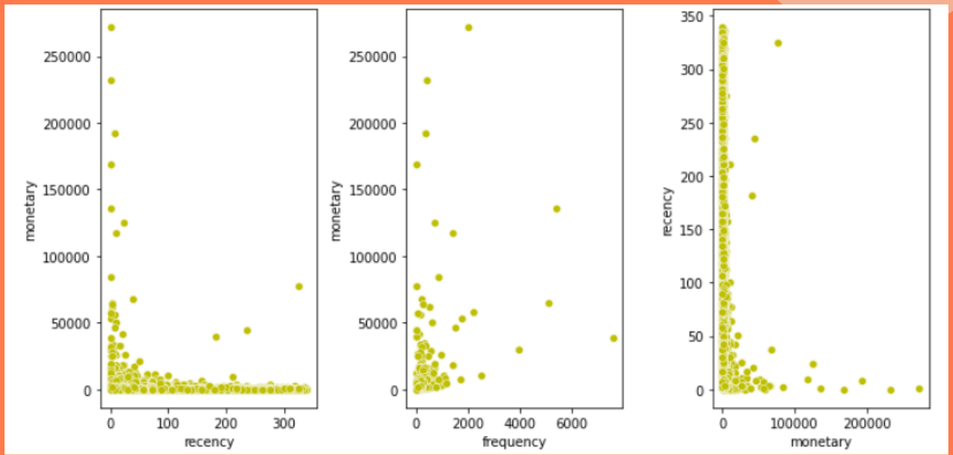
	CustomerID	frequency
0	12346.0	1
1	12347.0	151
2	12348.0	14
3	12349.0	73
4	12350.0	17

Monetary

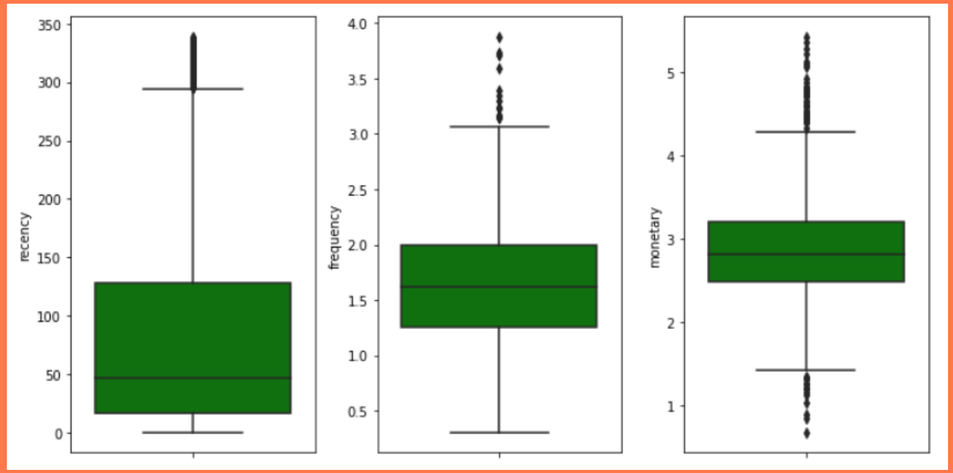
	CustomerID	monetary
0	12346.0	77183.60
1	12347.0	3598.21
2	12348.0	904.44
3	12349.0	1757.55
4	12350.0	334.40

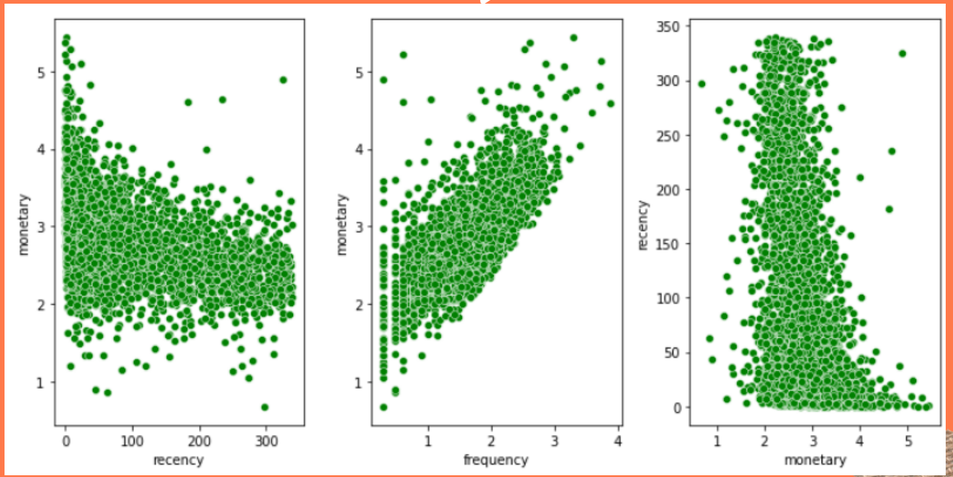
Data Distribution



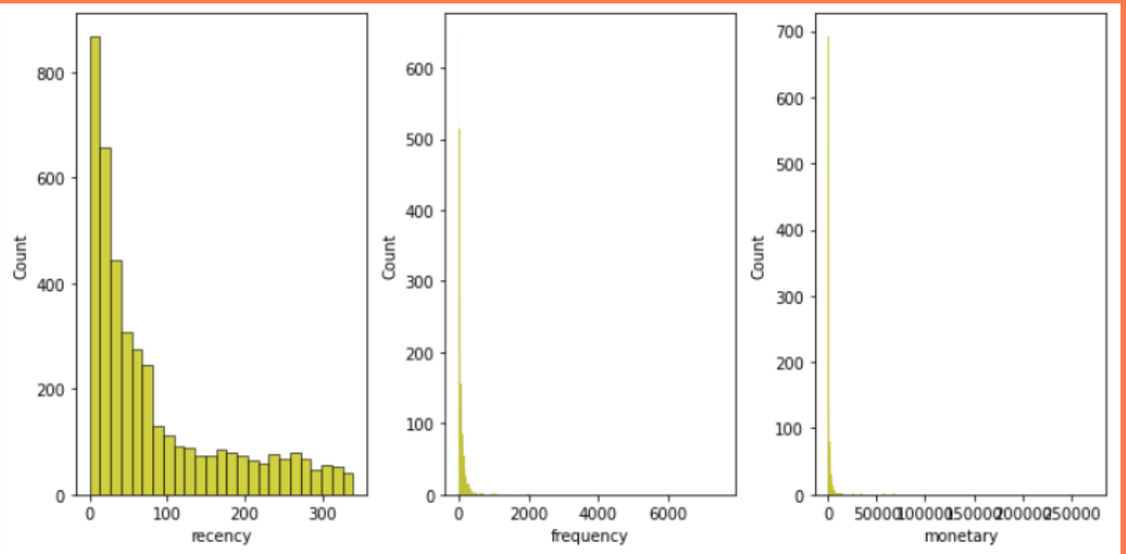


Transform data to reduce outlier

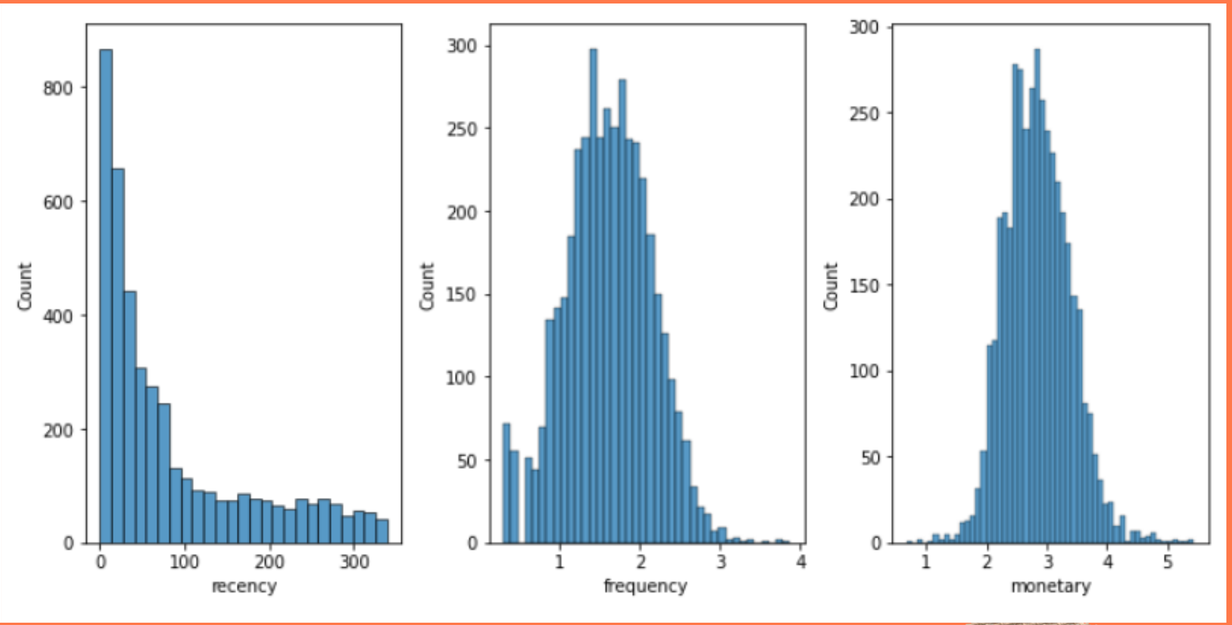




Data Distribution

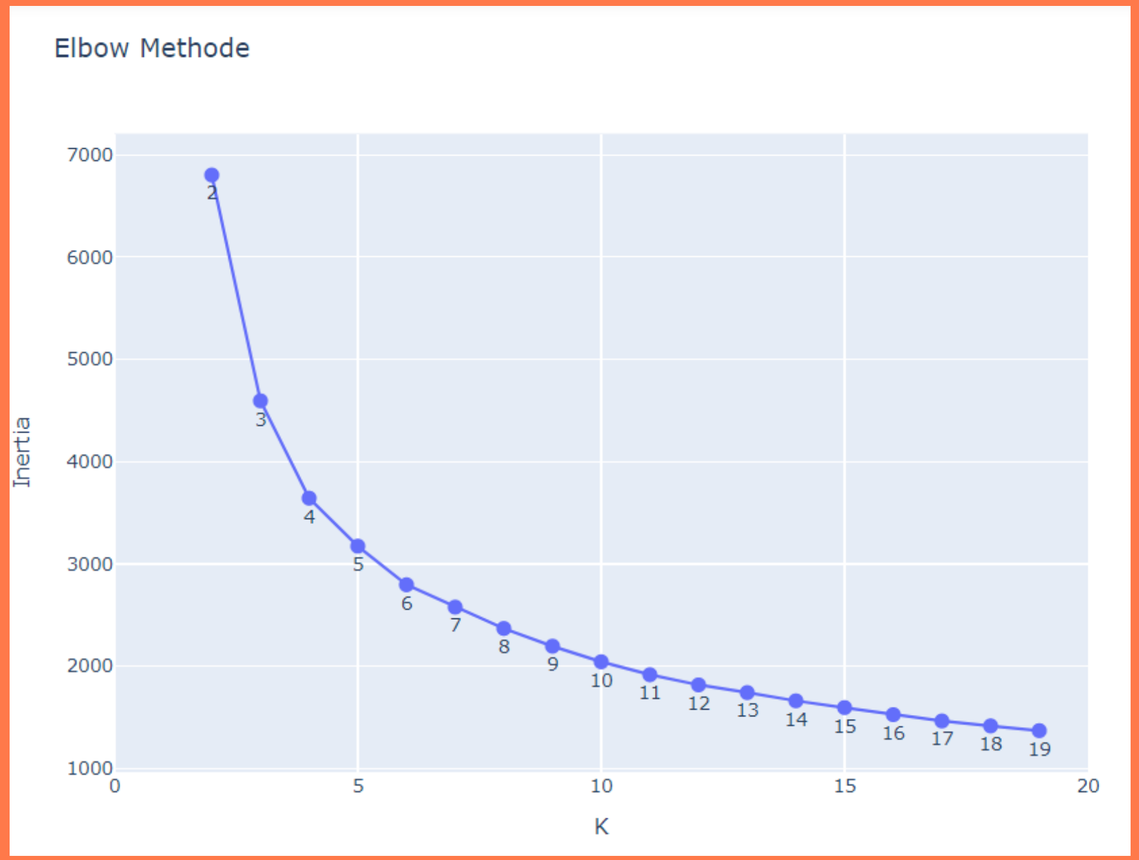


Transform data to reduce outlier

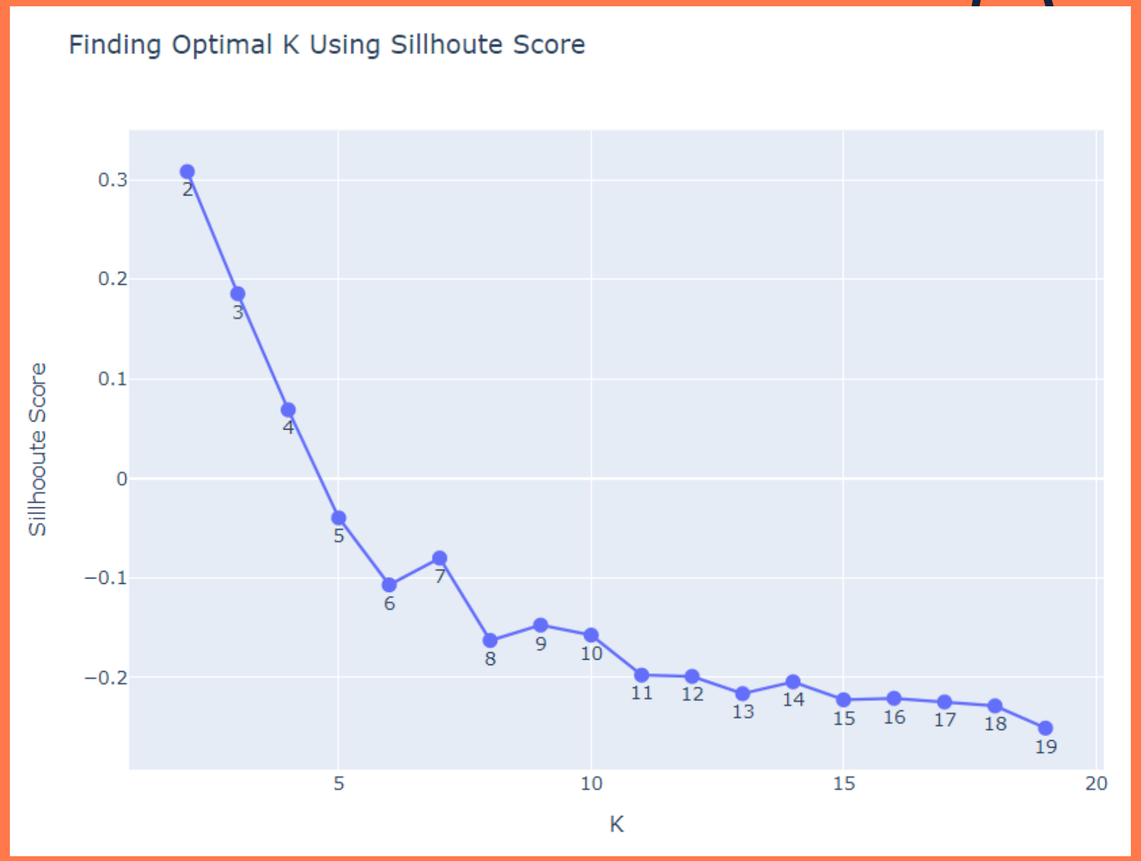


Clustering With K-Means

Mencari Potensial K Dengan Elbow Method Pada RFM DataFrame



Mencari Potensial K Degan Sillhoute Score Menggunakan Euclidean



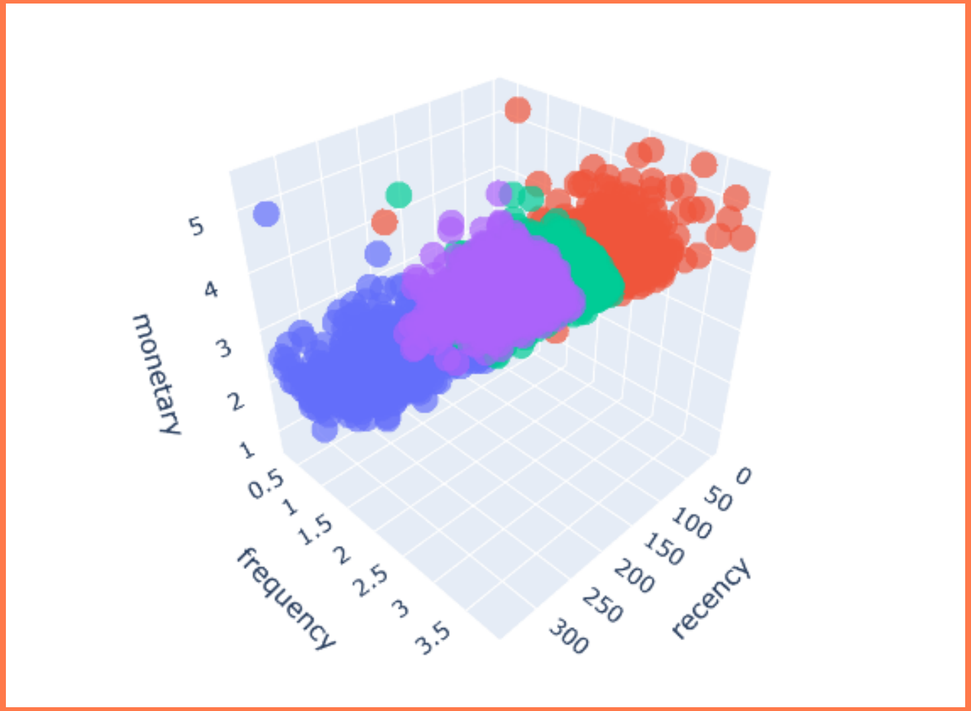
K = 4 diambil dikarenakan pada elbow method menjukan nilai yang tidak begitu signifikan dengan nilia k setelahnya, begitupun dengan sillhoute method, pada k yang tertinggi adalah 2,3,4 dan 5 , dikarenakan pada elbow method k potensial adalah 4 maka kita akan ambil 4 untuk nilai k.

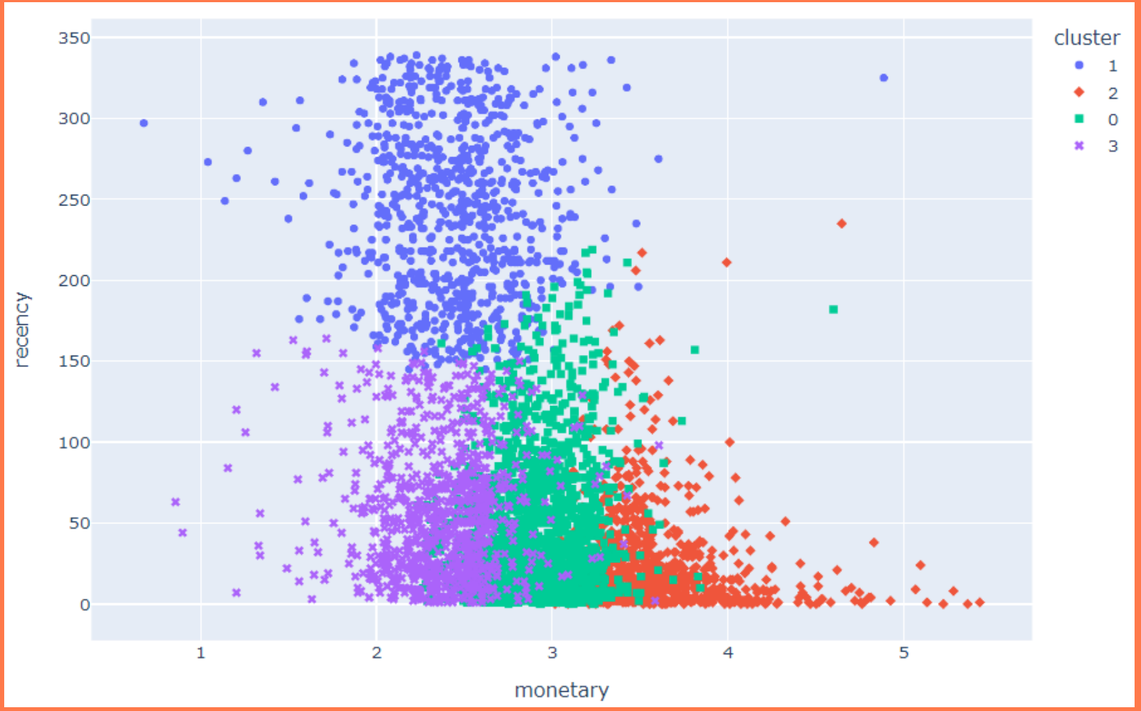
Clustering With K-Means

Sebaran Data Dengan Scatterplot
Setelah Pengklusteran

Clustering With K-Means

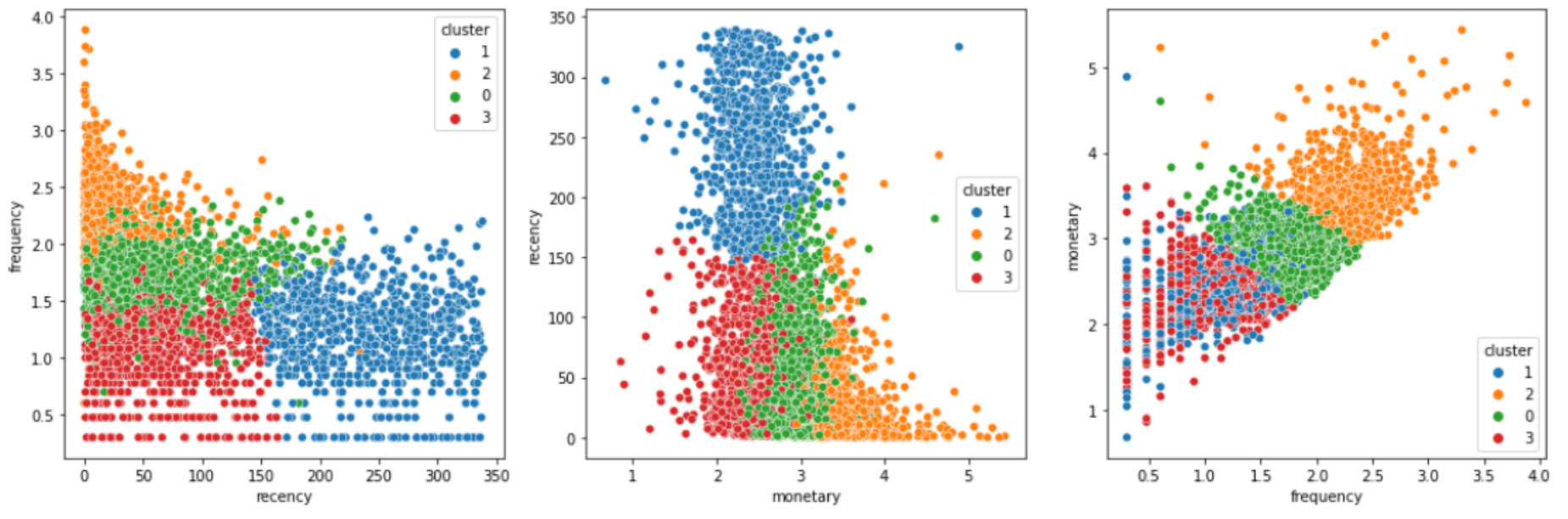
Sebaran Data dengan Scatterplot
Setelah Pengclusteran





Clustering With K-Means

Scatterplot
After Clustering



Clustering With K-Means

Mapping Class From Clusters

Middle Spending
Active Customers

0

New Customers

3

Low Spending
Old Customers

1

Best Customers

2

Customer Segemntation

Recommendation

Best Customers

- Penawaran Product Baru atau Product yang mungkin belum dibeli sebelumnya
- Memberikan reward untuk customer relationship management

Low Spending Pasif Customers

- Melakukan follow up dengan marketing campaign terkait bundling product dan memberikan rekomendasi product yang paling diminati.
- Memberikan penawaran potongan harga agar customer spent lebih banyak.

Middle Spending Active Customers

- Memberikan penawaran potongan harga agar menarik customer untuk bertransaksi lagi.
- Memberikan poin sebagai reward dan dapat ditukarkan dengan product atau potongan harga

New Customers

- Memberikan Rekomendasi Product berdasarkan product yang diminati customer lainnya.
- Memberikan poin agar customer bertransaksi lebih intens.

Note : Product Reccomendation sangat disarankan untuk tahap selanjutnya

THANK YOU