

# LOYALTY CUSTOMER SEGMENTATION



Kelompok 12

M. Ali Ma'sum - Meliana R - Niken Larasati

# Agenda

01

## Introduction

Background and Objective

02

## Data Collection

Data Survey

03

## Data Understanding

Understanding data.

04

## Data Preprocessing

EDA and Feature  
Engineering

05

## Modeling

Model Comparison

06

## Insight

What we get?

07

## Recommendation

What should we do?

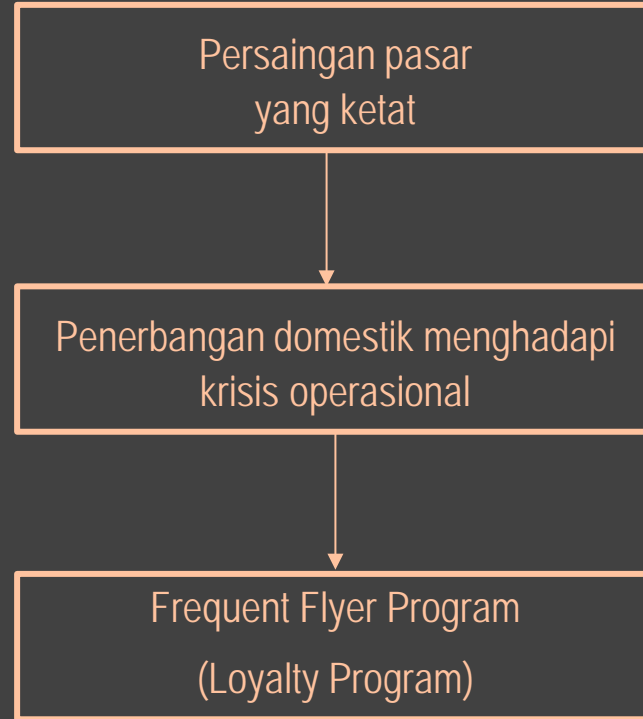




# **INTRO- 01.DUCTION**



## Background



# Introduction



## Frequent Flyer Program (FFP)

Program loyalitas membership yang ditawarkan oleh maskapai

- Sebagai salah satu faktor untuk membedakan mana customer yang loyal dan tidak (Dolnicar *et al.*, 2011)
- Untuk mempertahankan loyalitas customer, dan
- Sebagai salah satu cara marketing untuk membantu meningkatkan penjualan pada maskapai (Alisanta, 2014)



# Objective



Menganalisis karakteristik pelanggan dari beberapa jenis customer



Menyediakan solusi untuk berbagai jenis customer.



Membuat modeling dari data airline customer segmentation dengan unsupervised learning



Merumuskan strategi pemasaran yang sesuai





02.

# DATA COLLECTION

# DATA COLLECTION

## Dataset:

<https://www.kaggle.com/c/sa-customer-segmentation/data>

## Review Columns

Di kolom **work country** masih ada beberapa aksara china. Dan ada data yang out range (atau tidak sesuai dengan kalender) yaitu 29/2/2014 dikolom **last flight date**

## Type data

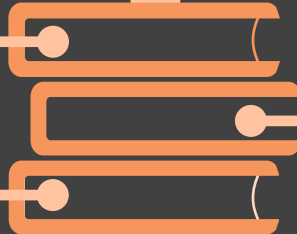
Int, float, Object, DateTime

## Data Shape:

Rows = 62988 , columns = 23

## Missing Values Detected

**Work country** 0.04%, **work city** 4 % , **work province** 7 % ,  
**age** 0.67 % , **sum yr 1** 0.87%, **sum yr 2** 0.22%, **gender** 0.0% (3 rows)





# 03

---

## DATA UNDERSTANDING



# DATA DESCRIPTION

Deskripsi data dari berbagai variabel adalah sebagai berikut,

Tabel 1. Deskripsi Variabel-Variabel pada Dataset

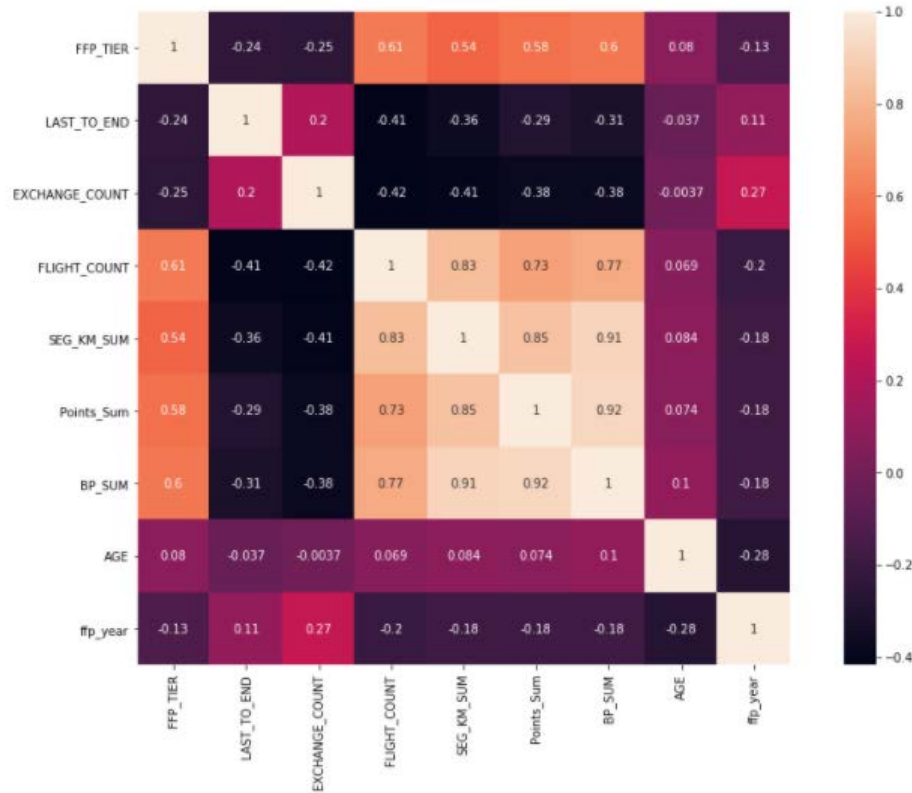
No.	Variabel	Deskripsi
1.	MEMBER_NO	Nomor keanggotaan
2.	FFP_DATE	Tanggal join membership
3.	FIRST_FLIGHT_DATE	Tanggal penerbangan pertama
4.	GENDER	Jenis kelamin customer
5.	FFP_TIER	Level dari membership (4,5,6)
6.	WORK_CITY	Kota customer bekerja
7.	WORK_PROVINCE	Provinsi customer bekerja
8.	WORK_COUNTRY	Negara customer bekerja
9.	AGE	Umur customer
10.	LOAD_TIME	Waktu/batas pengaksesan data penerbangan
11.	FLIGHT_COUNT	Jumlah berapa kali melakukan penerbangan
12.	BP_SUM	Total BP (British Petroleum) bahan bakar
13.	SUM_YR_1	Sum yield revenue 1
14.	SUM_YR_2	Sum yield revenue 2
15.	SEG_KM_SUM	Akumulasi dari jumlah jarak (km) penerbangan yang sudah ditempuh
16.	LAST_FLIGHT_DATE	Tanggal penerbangan terakhir
17.	LAST_TO_END	Jarak waktu antara penerbangan terakhir kali dengan LOAD_TIME
18.	AVG_INTERVAL	Rata rata jangka waktu.
19.	MAX_INTERVAL	Maksimal jangka waktu.
20.	EXCHANGE_COUNT	Jumlah pertukaran point
21.	avg_discount	Rata-rata nilai diskon yang sesuai dengan level membership customer
22.	Points_Sum	Jumlah point yang didapatkan customer setiap melakukan penerbangan
23.	Point_NotFlight	Jumlah point yang didapatkan customer ketika membatalkan penerbangan

# 04

---

## Data Pre-processing

# EXPLORATORY DATA



The Strong Correlation FFP TIER :

Flight count

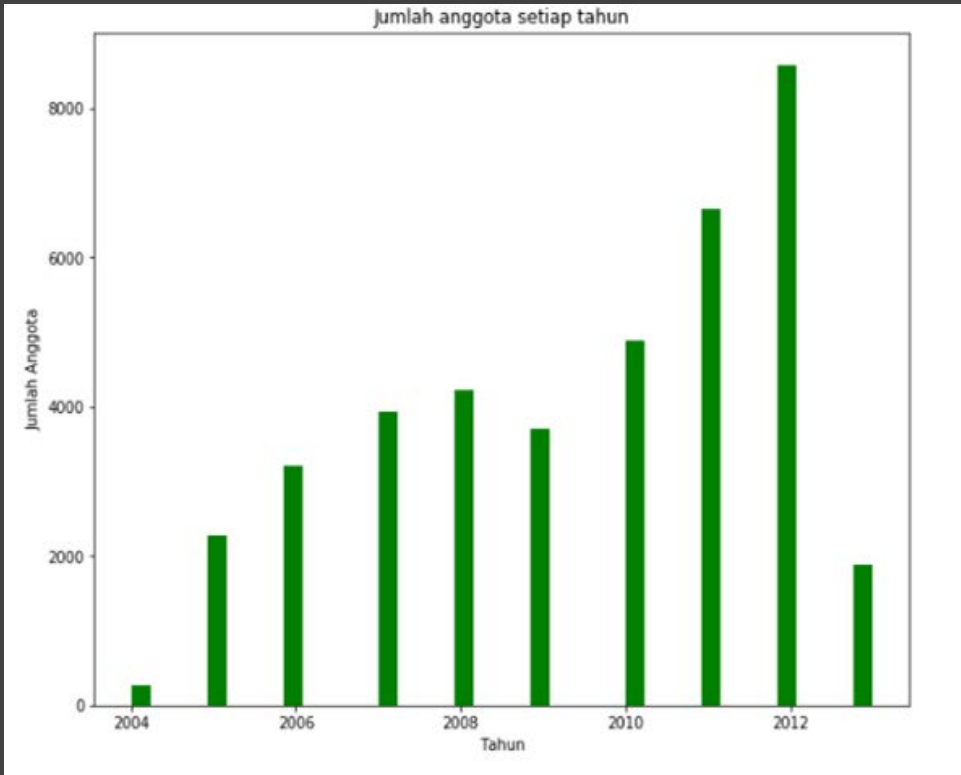
BP Sum

Seg Km Sum

Points Sum

# EXPLORATORY DATA

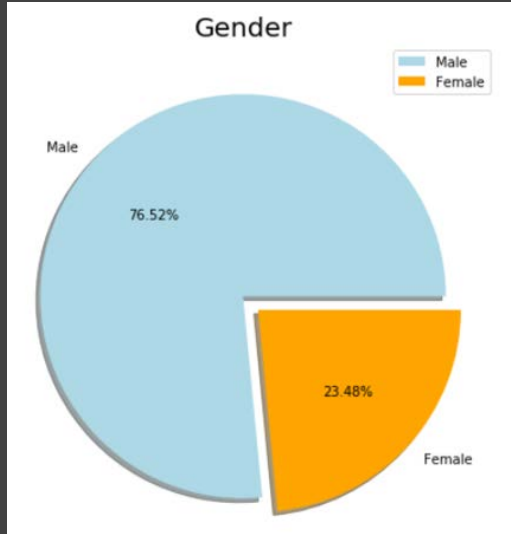
## 1. Berapa perolehan customer setiap tahunnya?



Insight :

Dalam setiap tahun ada penambahan perolehan customer

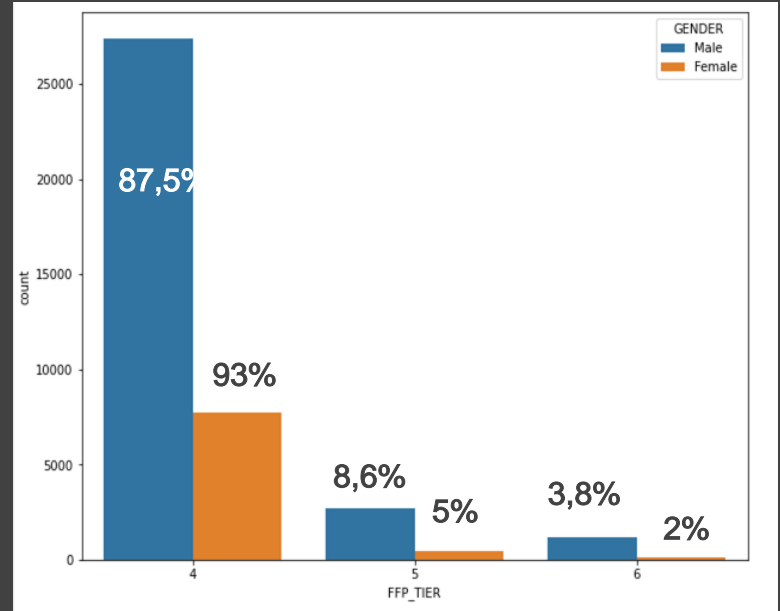
## 2. Gender apa yang paling banyak?



### Insight :

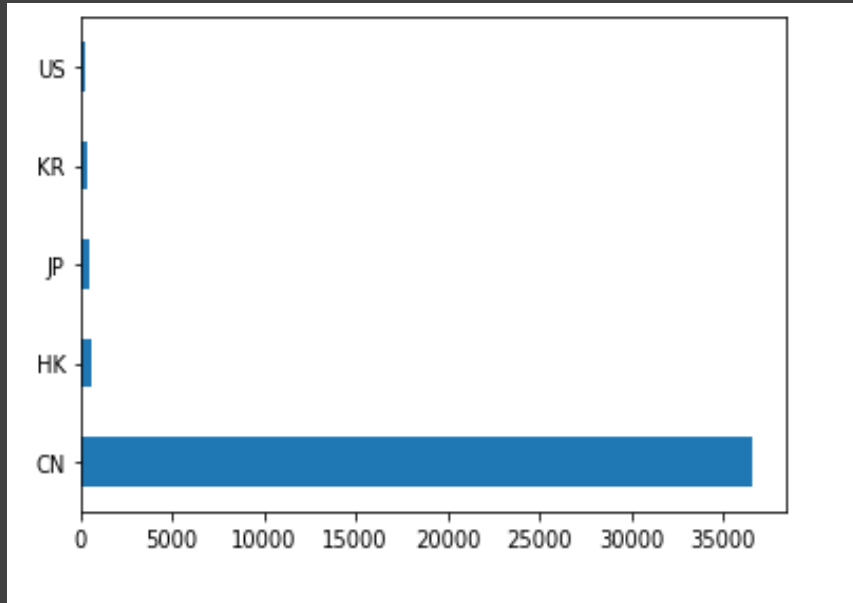
Gender yang paling banyak secara global tiap tier adalah male (laki laki).

## 3. Gender yang paling banyak setiap tier?



## EXPLORATORY DATA

4. Country mana yang paling banyak menjadi customer?

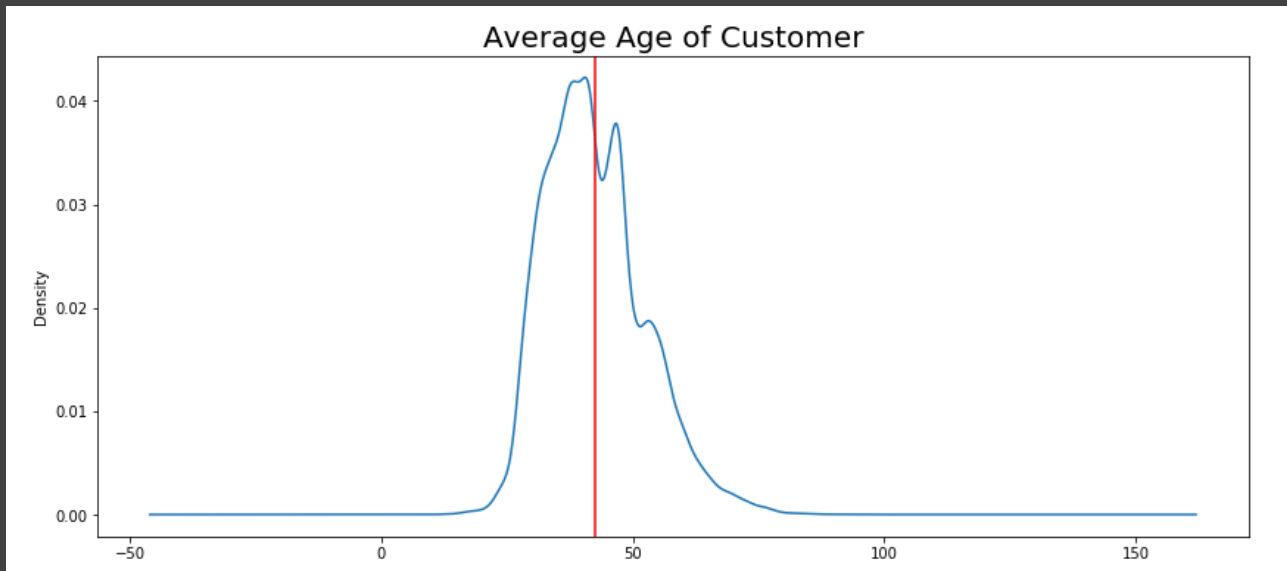


Insight :

Country customer yang paling banyak berada di CN (China).

# EXPLORATORY DATA

## 5. Berapa rata - rata umur customer ?



Insight :

Rata rata umur customer adalah 42 tahun.

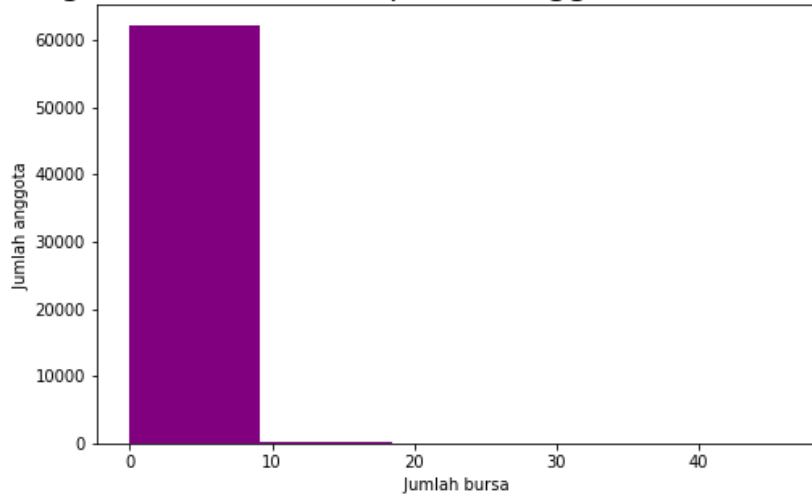
```
flight['AGE'].mean()
```

```
42.484967655981634
```



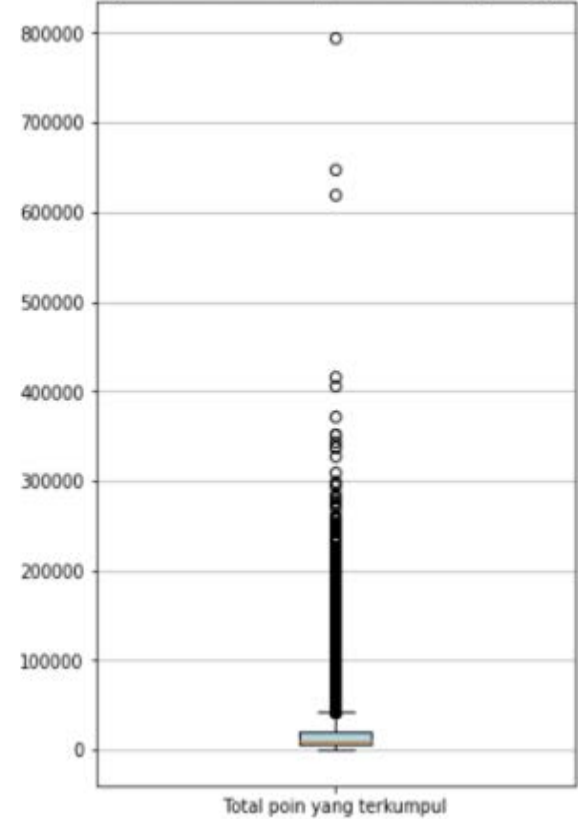
# ANALISIS DISTRIBUSI

Histogram distribusi berapa kali anggota menukarkan poin



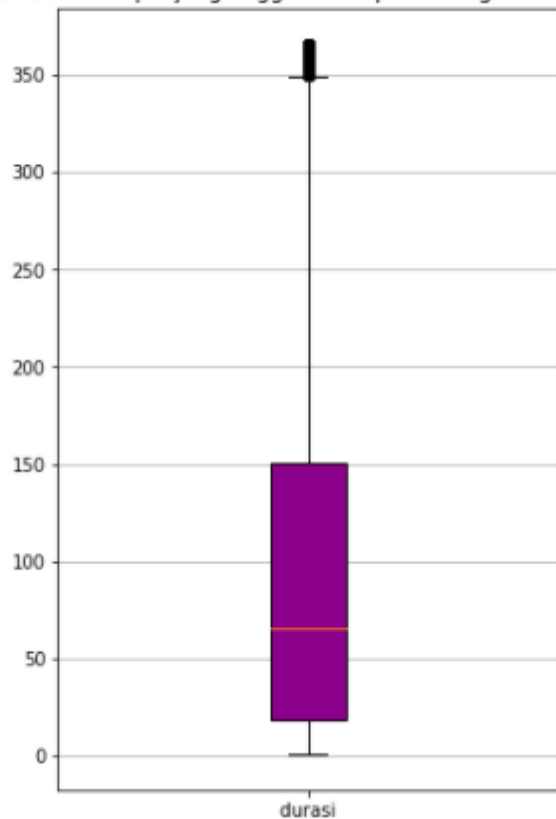
EXCHANGE\_COUNT

Bagan kotak dari total poin kumulatif pelanggan



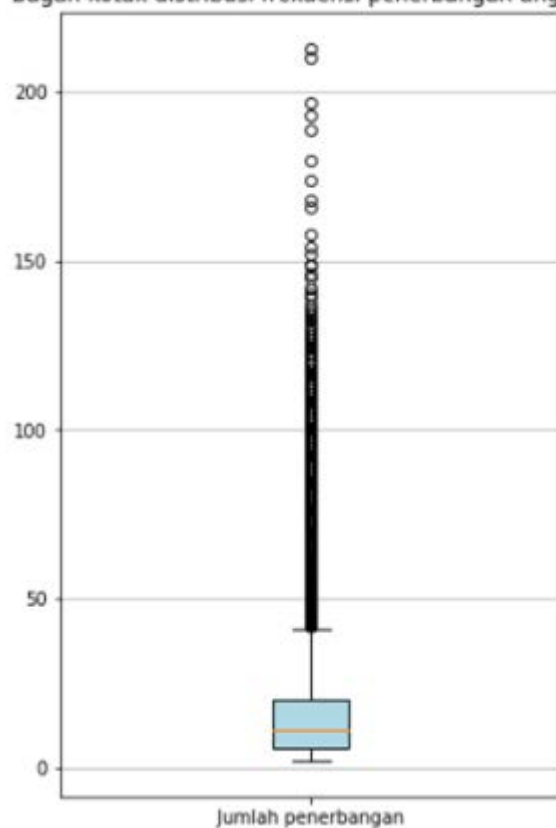
POINT\_SUM

Diagram kotak distribusi panjang anggota dari penerbangan terakhir hingga akhir

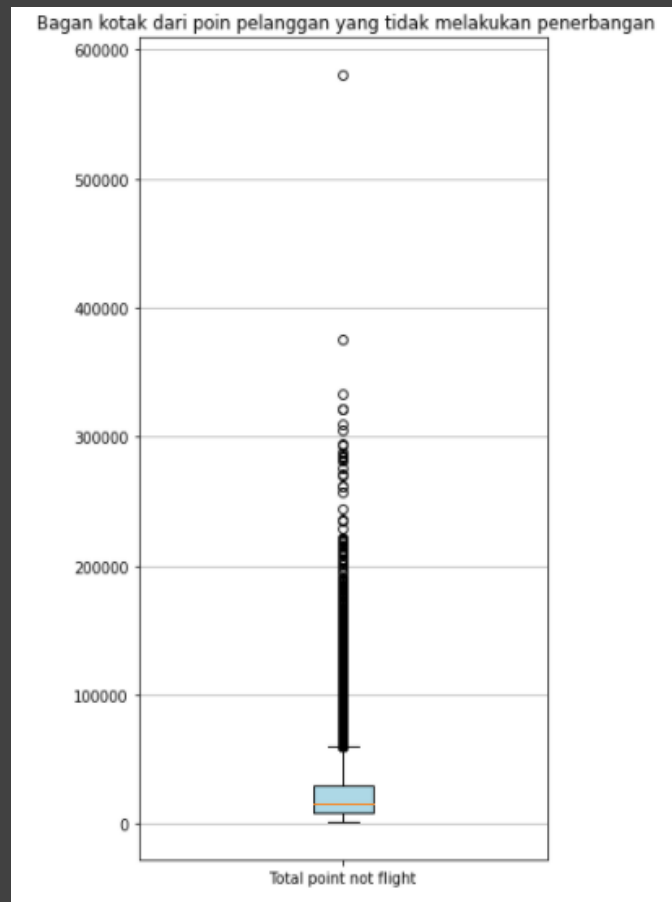


LOAD\_DATA

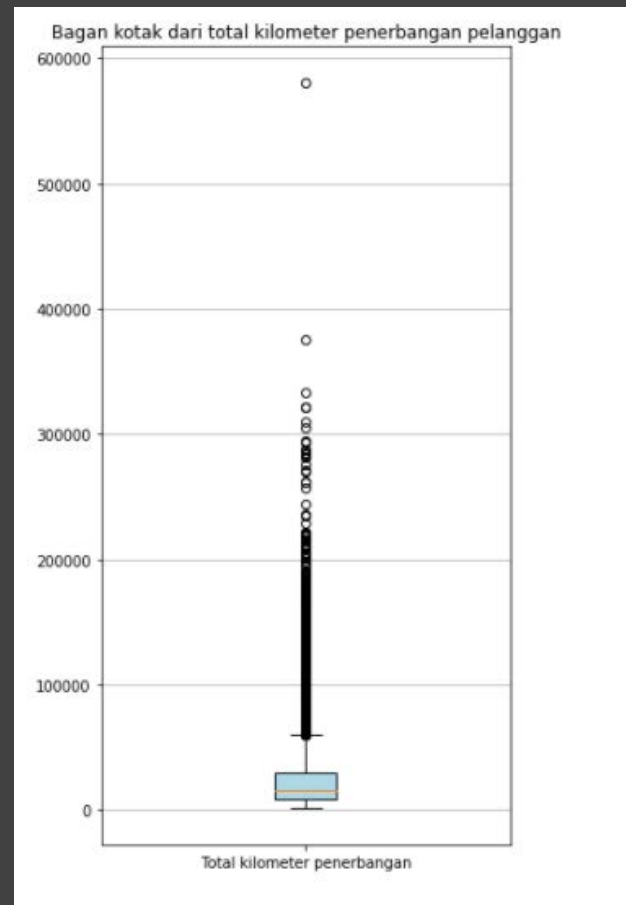
Bagan kotak distribusi frekuensi penerbangan anggota



FLIGHT\_COUNT



Point\_NotFlight



SEG\_KM\_SUM

# FEATURE ENGINEERING

## MEMBER\_TIME

FFP\_DATE – LOAD\_DATE

### MEMBER TIME

LAMANYA WAKTU MENJADI  
ANGGOTA

## FLIGHT\_TIME

LAST FLIGHT\_DATE –  
FIRST\_FLIGHT\_DATE

### FLIGHT TIME

WAKTU TERBANG

## COUNTRY

WORK COUNTRY

### COUNTRY

COUNTRY CHINA ATAU SELAIN  
CHINA(OTHER)

## TIME INTERVAL\_DIFF

MAX\_INTERVAL –  
AVG\_INTERVAL

### INTERVAL DIFF

SELISIH INTERVAL

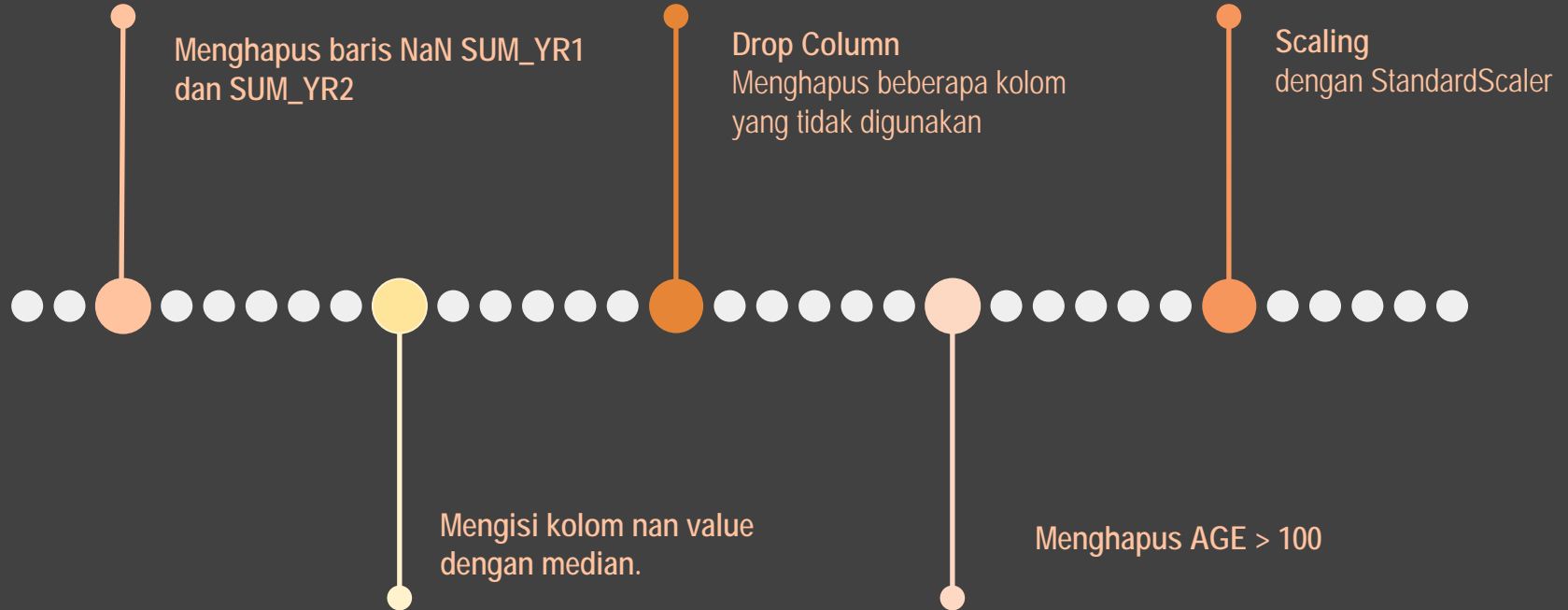
## YEAR

FFP\_DATE

### YEAR

EXTRACT TAHUN DARI FFP DATE

# FEATURE ENGINEERING

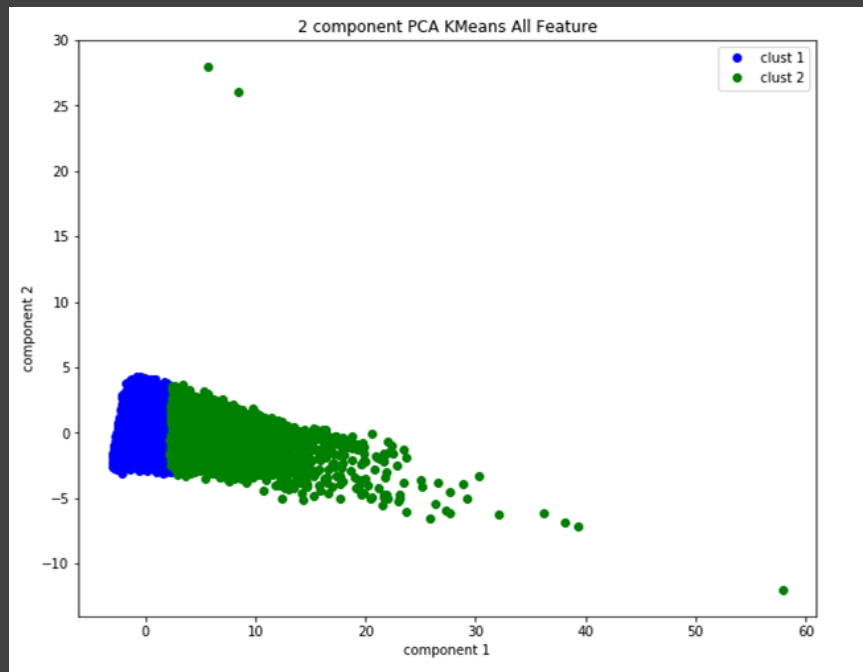


# 05

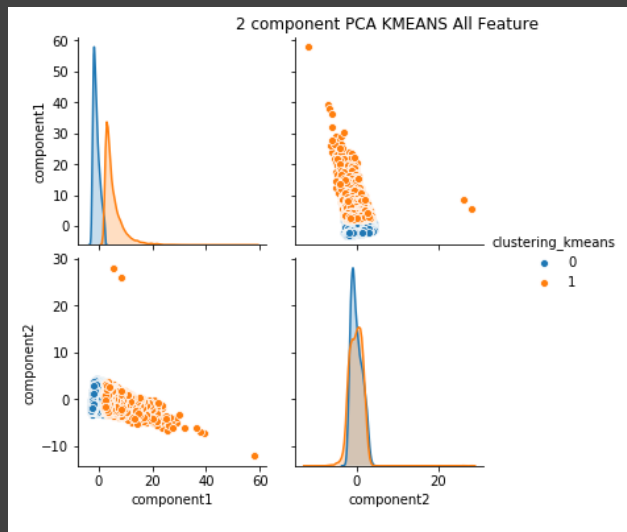
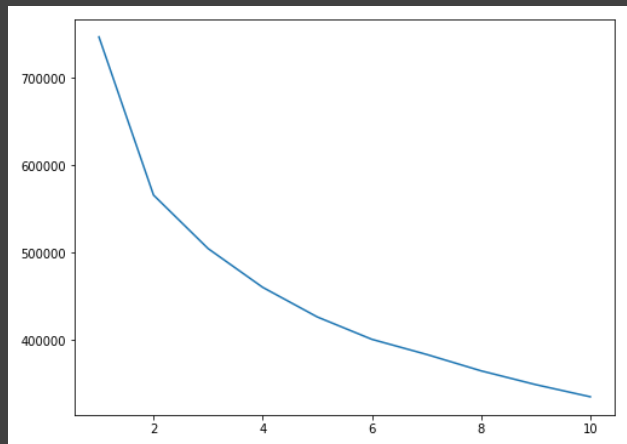
## Modeling



# KMEANS

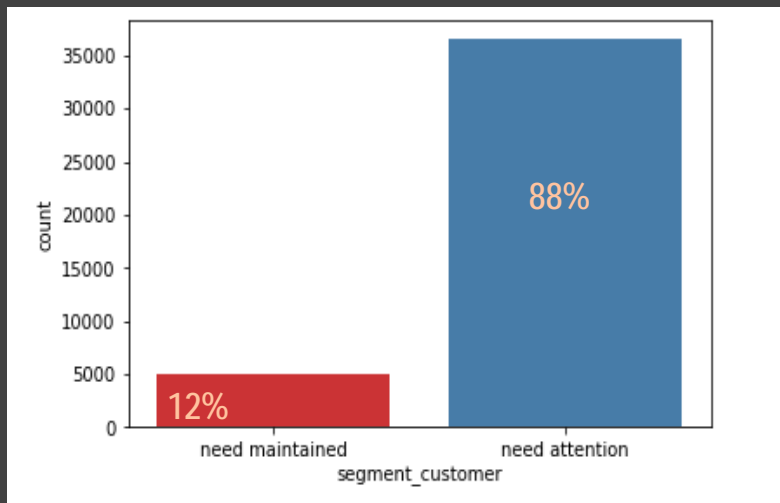


silhoutte score = 0.42234042104550923



# KARAKTERISTIK KMEANS

	kmeans_clustering	FFP_TIER	FLIGHT_COUNT	AVG_FARE_PER_KM	TIME_INTERVAL_DIFF	Points_Sum	Point_NotFlight	FLIGHT_TIME	Total Score
0	0	-0.236048	-0.268773	-0.273165	0.086899	-0.250570	-0.083748	-0.085846	25.945635
1	1	1.715660	1.953516	1.985435	-0.631608	1.821211	0.608704	0.623952	33.686915



**CLUSTER 1 = NEED MAINTAINED**  
**CLUSTER 0 = NEED ATTENTION**

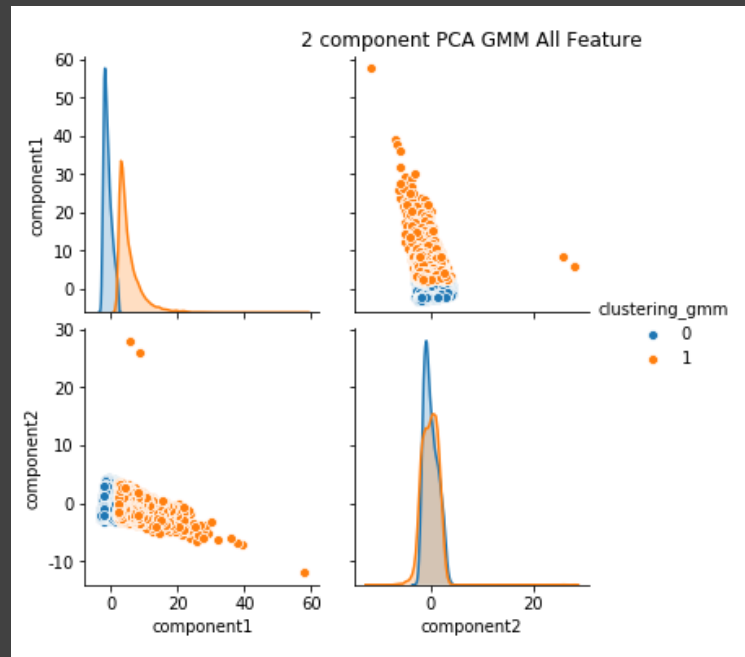
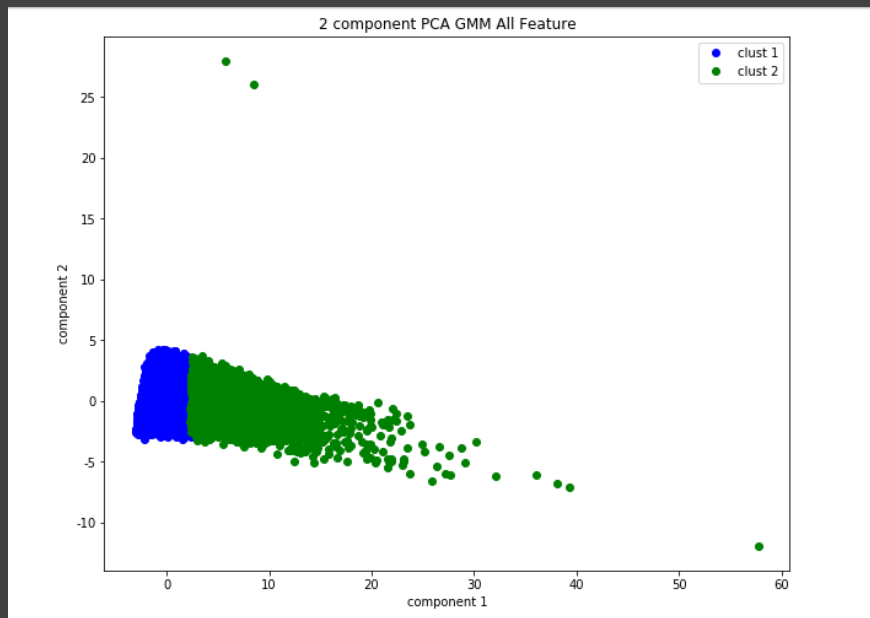
## INSIGHT

Jumlah Cluster **need maintained** sedikit

```
need attention    0.879056
need maintained   0.120944
```



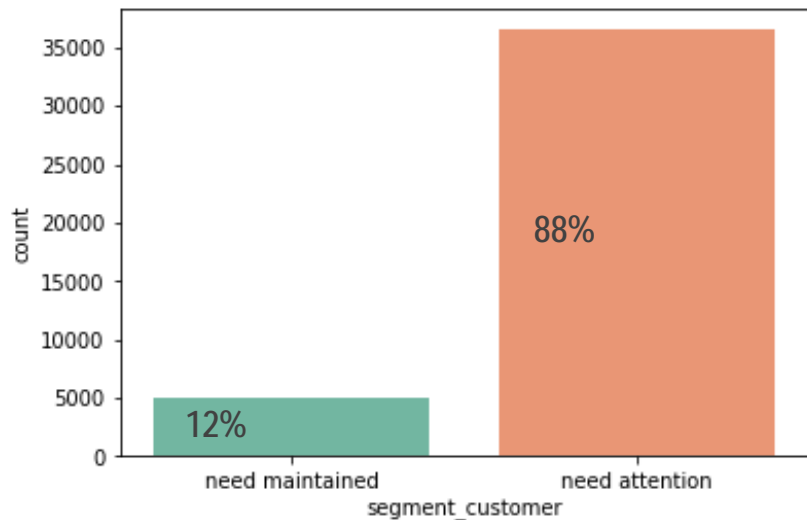
# GAUSSIAN MIXTURE MODEL (GMM)



silhoutte score = 0.43206970801171274

# KARAKTERISTIK GMM

	gmm_clust	FFP_TIER	FLIGHT_COUNT	AVG_FARE_PER_KM	TIME_INTERVAL_DIFF	Points_Sum	Point_NotFlight	FLIGHT_TIME	Total Score
0	0	-0.236048	-0.268773	-0.273165	0.086899	-0.250570	-0.083748	-0.085846	25.945635
1	1	1.715660	1.953516	1.985435	-0.631608	1.821211	0.608704	0.623952	33.686915



**CLUSTER 1 = NEED MAINTAINED**  
**CLUSTER 0 = NEED ATTENTION**

## INSIGHT

Jumlah Cluster **need maintained** sedikit

need attention	0.879056
need maintained	0.120944

## 06. Insights



# INSIGHTS

Penting untuk memahami informasi tentang pelanggan



Pertahankan interaksi kepada pelanggan



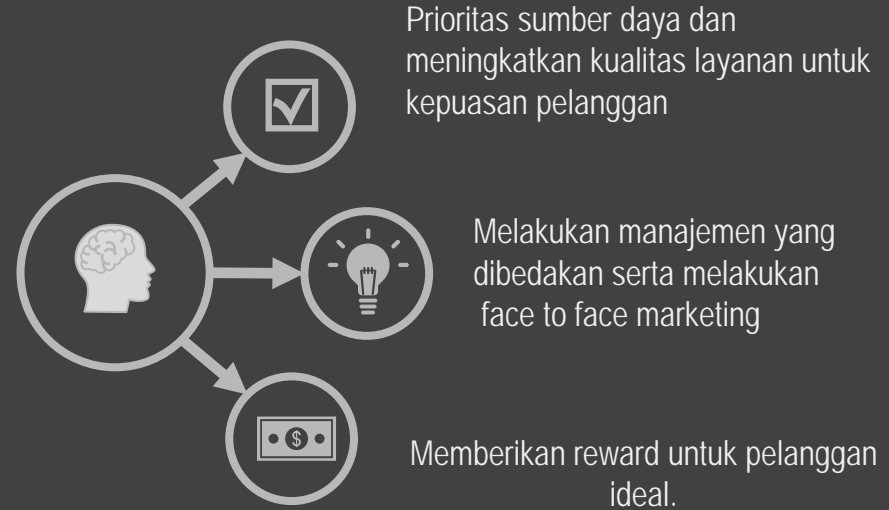
Mengadopsi metode pemasaran tertentu agar dapat memperpanjang siklus hidup pelanggan



## Need Attention

Kelompok yang harus diperhatikan, karena memiliki score analisis yang rendah. Namun sayangnya kelompok ini sangat banyak.

88%



## Need Maintained

Kelompok yang harus dipertahankan, karena memiliki score analisis yang tinggi. Namun sayangnya kelompok ini hanya sedikit

12%

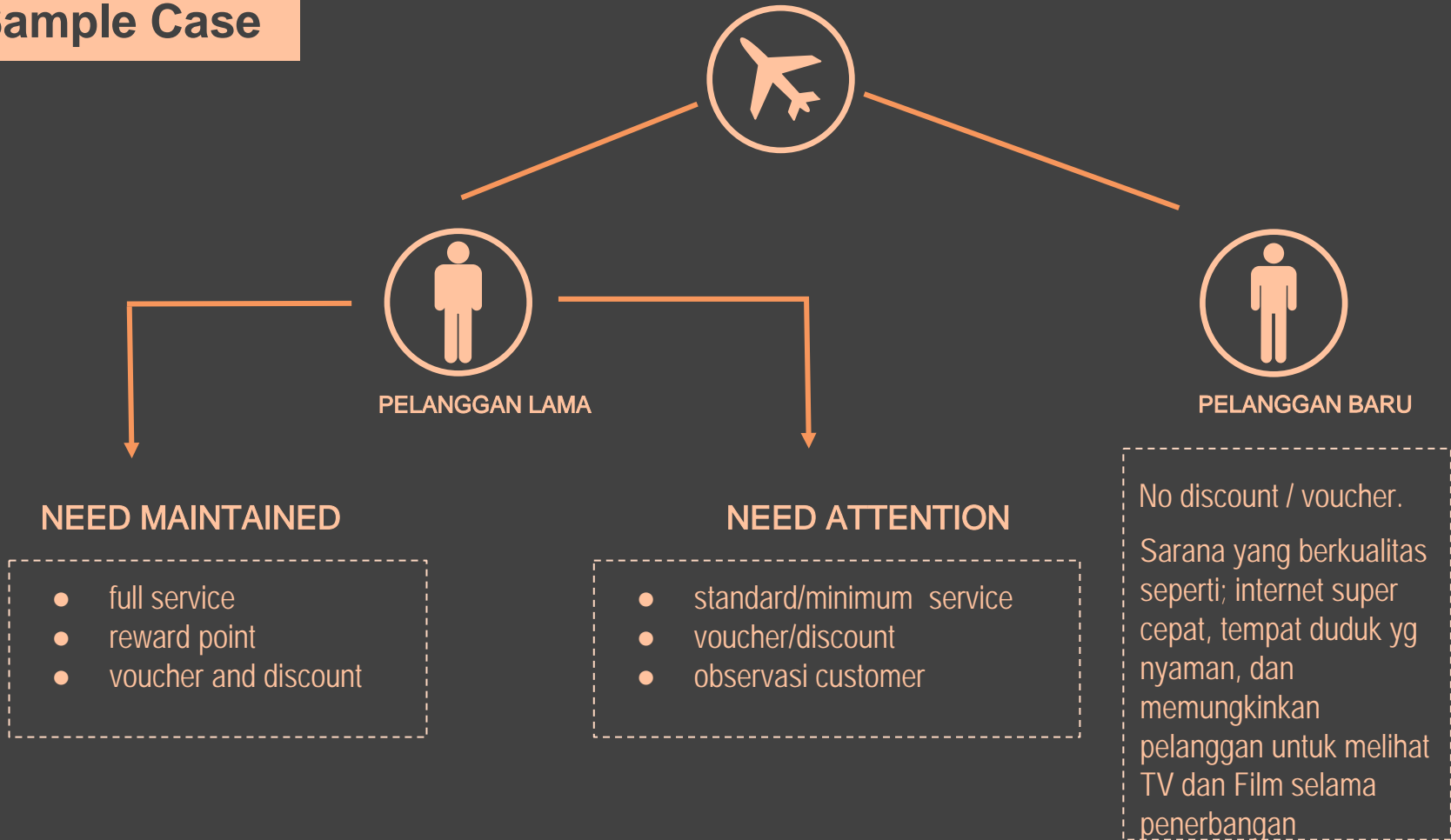
## 07. Recommendation



## How to Increase and Reduce Customer Degradation?

- Melakukan survey terhadap pelanggan terkait kualitas layanan yg dirasakan dan persepsi tentang kepuasan layanan terhadap maskapai (Heskett *et al.*, 1997)
- Be creative and unique
- Memberikan pengalaman terbang terbaik dan reward point agar pelanggan tetap setia dan loyal kepada maskapai.
- Menggunakan sosial media untuk mempromosikan dan menjangkau pelanggan.
- Menjangkau pelanggan secara offline dan online dengan cara membangun citra maskapai menjadi perusahaan terbaik dan terpercaya dengan memberi dampak langsung seperti ikut berkontribusi dalam bantuan sosial.
- Fokus ke pelanggan lama daripada pelanggan baru
- Menggunakan power influencer

## Sample Case



## REFERENSI

Alisanta. 2014. Mengenal Frequent Flyer Program. Javamilk. Diakses dari

<https://javamilk.com/2014/05/26/mengenal-frequent-flyer-program/>

Chauhan, A. 2019. 5 Innovative Marketing Strategies Used By Airlines. hotelspeak. Diakses dari

<https://www.hotelspeak.com/2019/07/5-innovative-marketing-strategies-used-by-airlines/>

<https://www.bp.com/>

Dolnicar, S., Grabler, K., Grün, B., and Kulnig, A. 2011. Key drivers of airline loyalty. *Tourism Management* (32). 1020

Heskett, J., Sasser, W., and Schlesinger, L. 1997. *The service profit chain. How leading companies link profit and growth to loyalty, satisfaction, and value*. New York: The Free Press.





THANK YOU!

# Appendix

# LRPMC



**LENGTH**

Long time being  
customer

FFP\_DATE\_LOAD\_TIME



**RECENCY**

Last time the  
customer took  
aircraft

LAST\_TO\_END



**FREQUENCY**

Flight Frequency

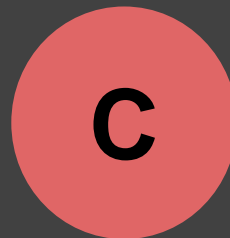
FLIGHT\_COUNT



**MONETARY**

Cumulation of  
mileage

SEG\_KM\_SUM

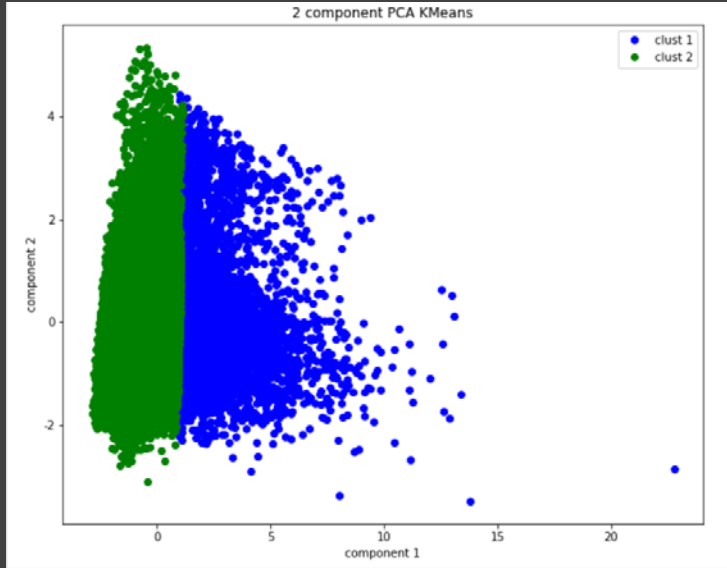


**COEFFICIENT**

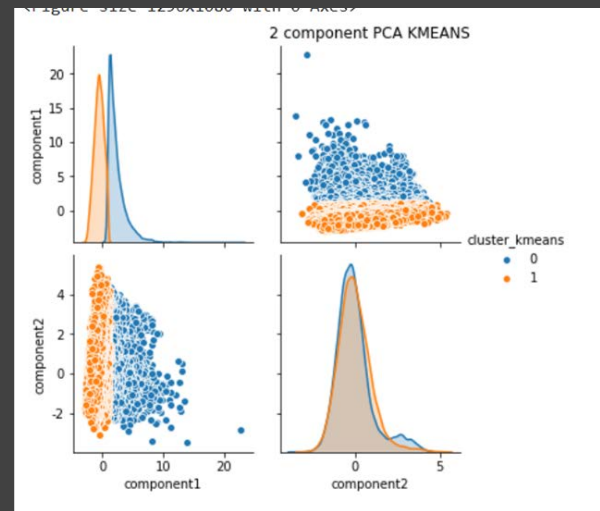
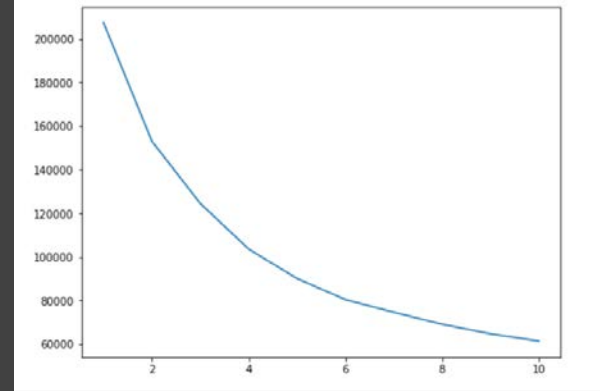
Coefficient  
Discount

AVG\_DISCOUNT

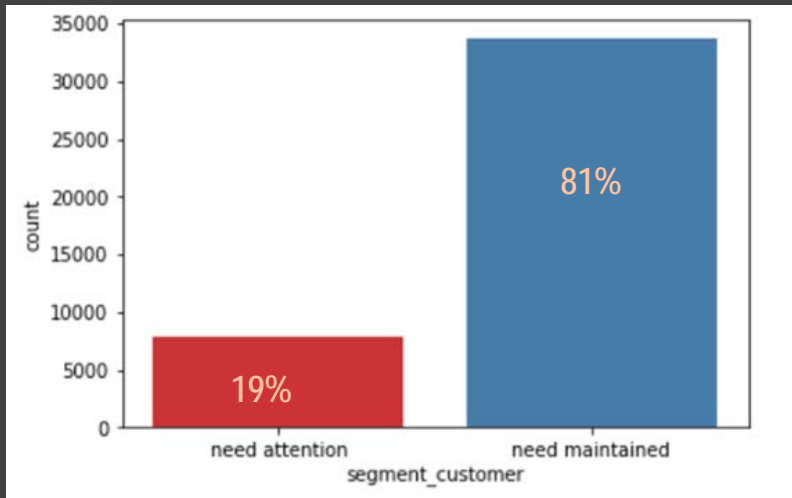
# KMEANS LRFMC



silhoutte score = 0.34577582301247156



# KARAKTERISTIK



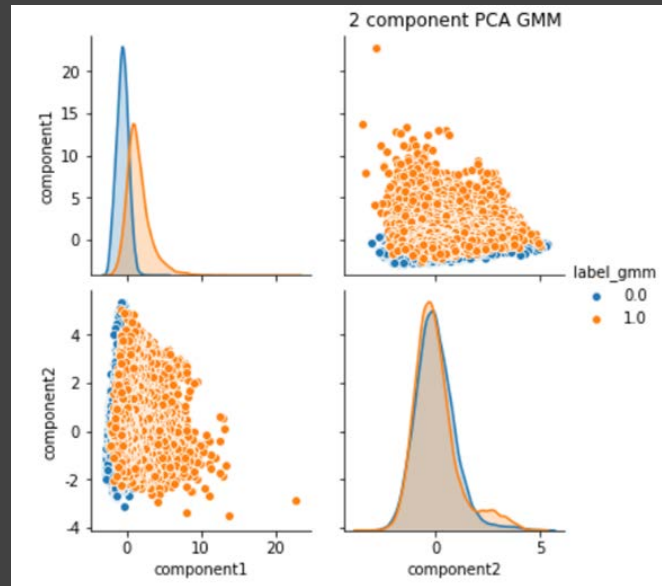
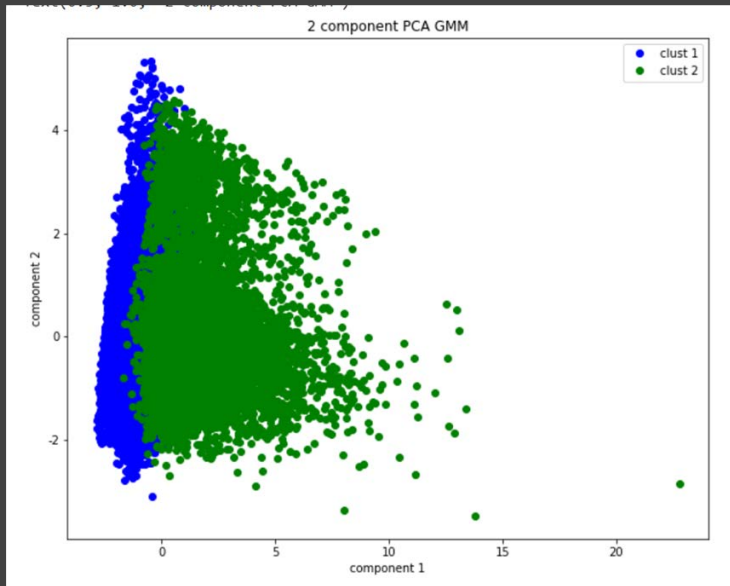
## INSIGHT

Jumlah Cluster **need maintained** sedikit

```
need maintained    0.810454
need attention     0.189546
```

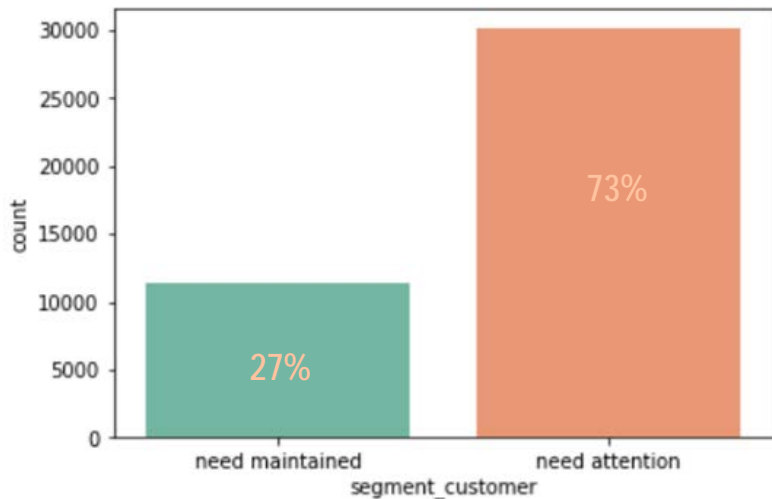
	kmeans_clust	Length	Recency	Frequency	Monetary	Coefficient	L_Quartile	R_Quartile	F_Quartile	M_Quartile	C_Quartile	Total Score
0	0	0.522546	-0.747316	1.553658	1.494433	0.365165	3.044605	3.478714	3.920829	3.934299	2.840640	17.219088
1	1	-0.122211	0.174779	-0.363364	-0.349512	-0.085403	2.372080	2.283719	2.056114	2.164329	2.420317	11.296558

# GMM LRFMC



silhoutte score = 0.34577582301247156

# KARAKTERISTIK



## INSIGHT

Jumlah Cluster **need maintained** sedikit

need attention	0.72516
need maintained	0.27484

	gmm_clust	Length	Recency	Frequency	Monetary	Coefficient	L_Quartile	R_Quartile	F_Quartile	M_Quartile	C_Quartile	Total Score
0	0	-0.08887	0.235494	-0.414932	-0.419592	-0.138502	2.405580	2.224747	1.952500	2.030161	2.390301	11.003288
1	1	0.23448	-0.621345	1.094788	1.107082	0.365435	2.747502	3.263453	3.615513	3.739001	2.789395	16.154864