



Data Science & Data Analytics Portfolio

Amodya Satriatama Subagio

2 0 2 4



1

Hello,

I'm Amodya Satriatama Subagio

2

3

4

5

6

7

8

9

Highly adaptable and rapid-learning analyst with 2 years of experience in harnessing the power of data to uncover valuable insights that drive strategic decision-making with solid understanding of various analytical tools and programming languages, including SQL, Python, Looker Studio, Tableau, Google Workspace, and other tools.



<https://github.com/amodyas/data-science-portfolio>



[Amodya Subagio](#)



amossubagio@gmail.com



+628111475017



Technical Skills

1

2

3

4

5

6

7

8

9

Language

Python, SQL

Libraries & Tools

Pandas, NumPy, Scikit-Learn, Matplotlib, Seaborn, Jupyter Notebook, Looker Studio, Tableau

Databases

MySQL

Others

Git, Adobe Photoshop, Google Workspace, Microsoft Office



1

2

3

4

5

6

7

8

9

Education

Institut Teknologi Sepuluh Nopember (ITS) - Bachelor of Engineering

Engineering Physics (Teknik Fisika), Faculty of Industrial and Systems Engineering

Certifications

- Data Science Bootcamp, Rakamin Academy (2024)

Certification ID 2304DS 3822024

- Phyton Programming Essentials, Rakamin Academy (2024)

Credential ID 23208_DATA1SCIENCEIII2023-EG

- SQL - MySQL for Data Analytics and Business Intelligence, Udemy (2021)

Credential ID UC-354a73c9-2f78-4c90-861b-729943a24519

1

2

3

4

5

6

7

8

9

Project Portfolios

1

2

3

4

5

6

7

8

9

Project 1

Banking Marketing Targets

Conducted in order to obtain Data Science Bootcamp certification from Rakamin Academy, this project titled "Bank UNO Marketing Targets" focuses on **analyzing customer profiles and predicting their likelihood to subscribe to term deposits** at UNO Bank following a marketing campaign.

Utilizing tools such as **Pandas, NumPy, Matplotlib, Seaborn, and Scikit-Learn**, the team conducted **exploratory data analysis, data preprocessing, and data modeling** with algorithms like **Logistic Regression, Decision Tree, and XGBoost**.

Key findings include the **identification of significant predictors for subscription likelihood**, ultimately aiming to increase term deposit subscriptions. The model recommended targeting customers predicted to respond positively, **achieving a conversion rate increase of above 15%**.

This project further awarded as the runner up for best final project from the certification program.

[Click here to see project on GitHub](#)

Problem Statement

Jumlah data nasabah yang membuka deposito berjangka pada UNO Bank dari total 45.211 nasabah yang ada, **hanya sekitar 5.289 nasabah atau 11,7 % nasabah saja yang membuka deposito berjangka.**

Permintaan dari manajemen UNO Bank itu terdapat peningkatan, **minimal sebesar 15%** dari total nasabah UNO bank yang membuka deposito berjangka.

Diantaranya terdapat 4.369 nasabah atau 9.66% nasabah yang dihubungi melalui telepon cellular dan 390 nasabah atau 0.86% nasabah yang dihubungi melalui telepon rumah.



Project 1

Business Insights

1

2

3

4

5

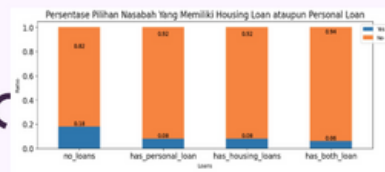
6

7

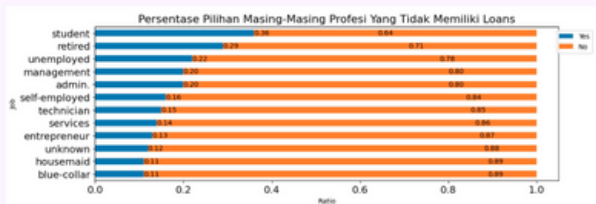
8

9

Business Insight



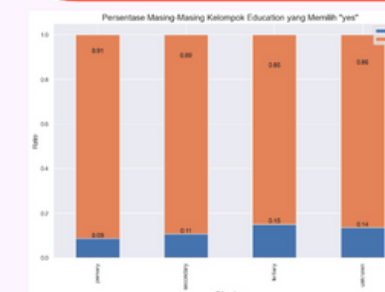
- Nasabah pembuka deposito berjangka umumnya **tidak memiliki kredit perumahan atau pinjaman pribadi**.
- Profesi dominan di antara nasabah ini termasuk **mahasiswa, pensiunan, pengangguran, staf administrasi, dan manajemen**.
- Disarankan untuk merancang strategi khusus untuk menarik pekerja kasar dan pembantu rumah tangga, yang menunjukkan tingkat partisipasi rendah.



Business Insight

- Uji hipotesis menunjukkan semakin banyak kampanye yang dilakukan, semakin banyak nasabah yang menolak untuk membuka deposito berjangka.
- Direkomendasikan untuk mengoptimalkan frekuensi kampanye agar tidak berdampak negatif terhadap keputusan nasabah dan dianggap sebagai spam.

Business Insight



- Persentase kelompok dengan pendidikan **tertiary** (lulusan S1 atau lebih) merupakan yang **terbanyak** mendaftar deposito berjangka, sekitar 15% dari total pendaftar dalam kelompoknya, diikuti oleh kelompok dengan pendidikan tidak diketahui, kemudian secondary (SMP dan SMA), dan terakhir primary (SD ke bawah).
- Saat hanya mempertimbangkan persentase dari **semua pelanggan yang memilih 'yes'**, kelompok pendidikan **secondary** paling banyak, diikuti oleh kelompok pendidikan tertiary, lalu primary, dan kelompok dengan pendidikan tidak diketahui.
- Berdasarkan insight ini, direkomendasikan untuk menargetkan kampanye pada kelompok **pendidikan tertiary dan secondary** untuk nasabah baru.

Business Insight



- Bisa dilihat melalui visualisasi diatas bahwa tipe contact **cellular** merupakan tipe contact yang paling banyak menghasilkan customer untuk mendaftar deposito berjangka.
- Prioritaskan kontak melalui **cellular** dalam kampanye deposito berjangka, mengingat efektivitasnya yang tinggi.
- Pertimbangkan email dan WhatsApp sebagai langkah awal kampanye sebelum panggilan langsung, untuk meningkatkan peluang pendaftaran.

Business Insight

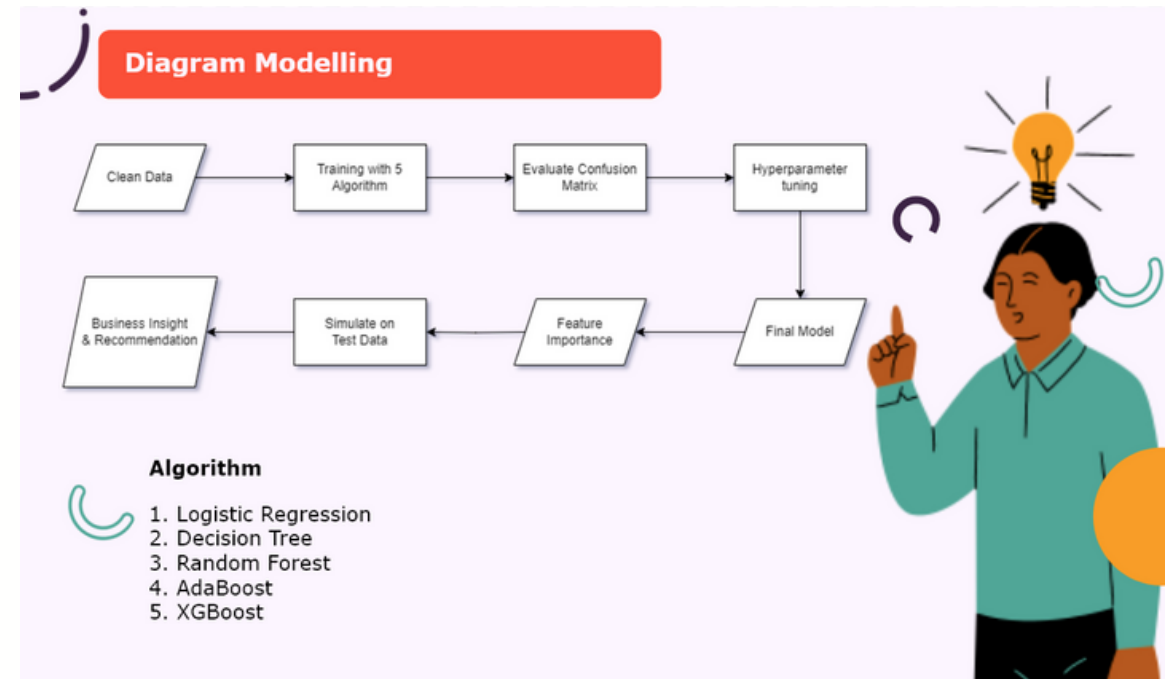


- Single** memiliki persentase tingkat pendaftaran deposito berjangka **terbesar**, 15%, diikuti oleh yang **bercerai**, dan paling rendah di antara yang **menikah** (10%).
- Berdasarkan **semua pelanggan yang memilih 'yes'**, kelompok yang **sudah menikah** mendominasi dalam memilih 'yes', kemudian diikuti oleh kelompok **single**, dan yang **bercerai** adalah yang paling sedikit.
- Untuk **single**, tawarkan konsep tabung untuk tujuan besar seperti pernikahan dengan bunga deposito yang menarik. Sampaikan pesan bahwa berinvestasi sekarang dapat membantu mencapai impian mereka di masa depan.
- Untuk yang **menikah/cerai**, fokuskan pada perlindungan finansial keluarga dan masa depan anak-anak. Gunakan narasi seperti "Investasi deposito Anda hari ini menjamin pendidikan anak dan pensiun yang nyaman," menekankan pada keamanan dan perencanaan finansial untuk keluarga.

Business Insights Summary

- Term deposit customers often **lack housing and personal loans**.
- Key customer groups include **students, retirees, and managerial staff**.
- Over-campaigning** leads to resistance; optimal frequency is crucial.
- Tertiary and secondary** educated individuals are **prime targets**.
- Cellular contact outperforms** other communication methods in effectiveness.
- The majority of term deposit customers consist of individuals who are **married**, however **only 10%** of married customers have a term deposit; tailored messaging emphasizing **future planning and financial security is advised**.

Data Modelling

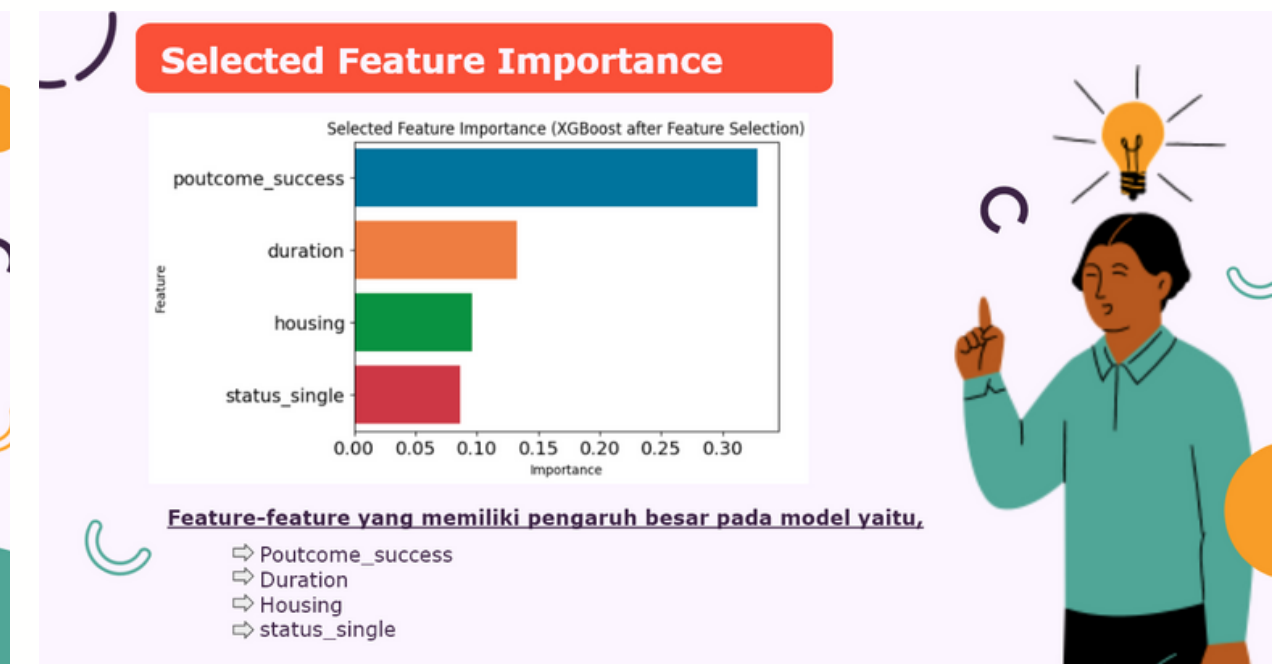


Evaluasi Confusion Matrix

Hyperparameter Tuning						
Algorithm	AUC-Score Train	AUC-Score Test	Recall	Precision	F1	Accuracy
Logistic Regression	0.88	0.89	0.65	0.78	0.68	0.90
Decision Tree	0.77	0.78	0.65	0.77	0.69	0.90
Random Forest	0.89	0.92	0.58	0.82	0.61	0.90
AdaBoost	0.90	0.90	0.66	0.77	0.69	0.90
XGBoost	0.90	0.92	0.68	0.77	0.72	0.90

Final Model XGBoost

Final Model XGBoost					
AUC-Score Train	AUC-Score Test	Recall	Precision	F1	Accuracy
0.88	0.86	0.66	0.78	0.69	0.90



The team utilized various algorithms like **Logistic Regression, Decision Trees, Random Forest, AdaBoost, and XGBoost** for modeling. The outcome was evaluated using metrics like **AUC-Score, Recall, Precision, F1-Score, and Accuracy**. **XGBoost emerged as the final model** due to its balanced performance across these metrics.

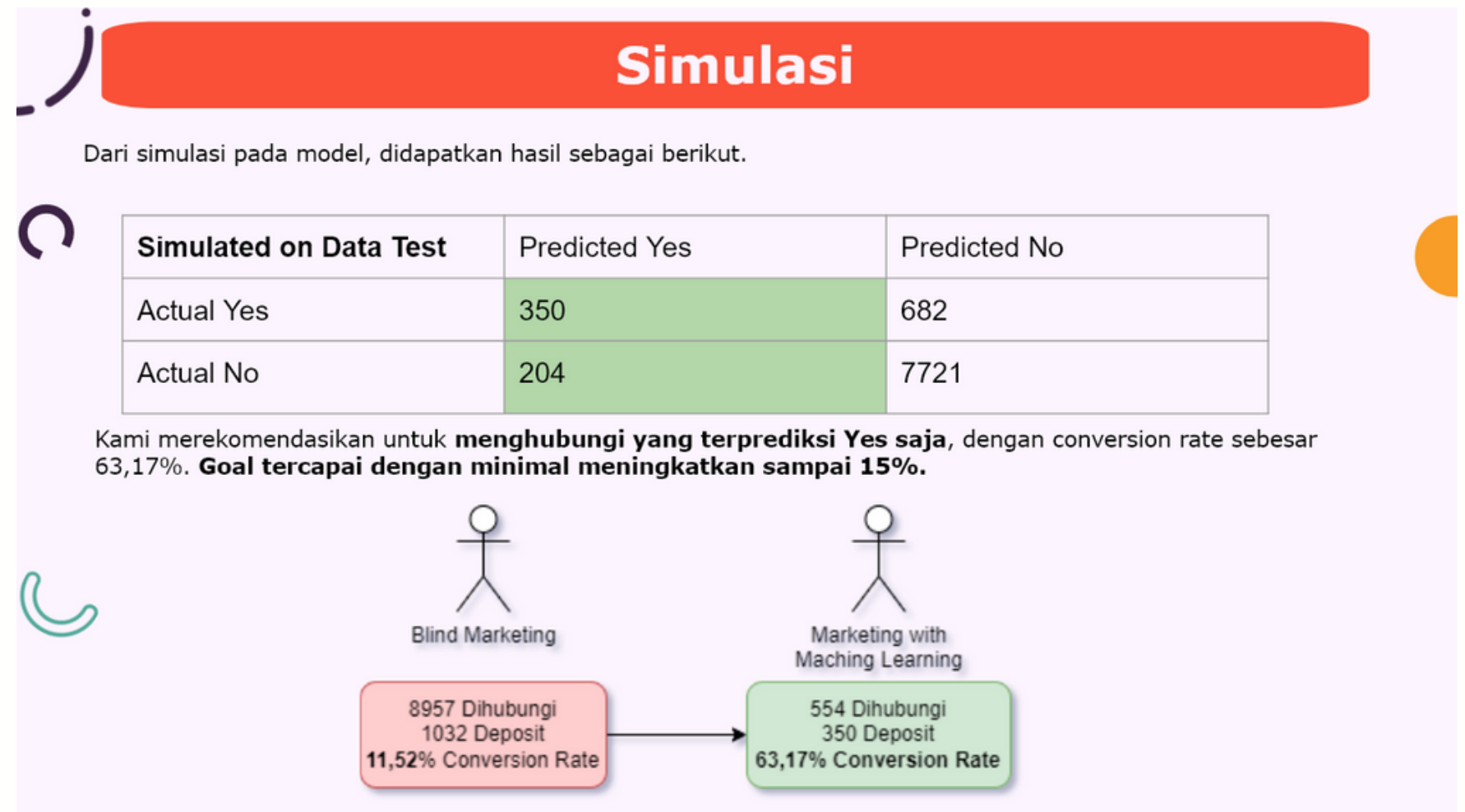
The model recommended targeting customers predicted to respond positively, achieving a **conversion rate increase of over 15%**.

Project 1

Outcome & Recommendations

Business Insights Summary

- **Focus on Non-Loan Holders:** Target individuals without home loans with tailored offers.
- **Optimize Campaigns:** Ensure a minimum of 4-minute calls and adjust frequency to avoid spam perceptions.
- **Emphasize Family Planning:** Use narratives showcasing term deposits' benefits for family financial planning.
- **Education Targeting:** Aim at tertiary and secondary educated groups with relevant offers.
- **Tailor Messages:** Craft promotional messages to fit demographic specifics, focusing on long-term goals for singles and financial security for families.



Simulation

The simulation of the UNO Bank marketing campaign on test data showed that focusing on customers predicted as 'Yes' could lead to a substantial increase in conversion rates, potentially achieving a 6317% increase and greatly exceeding the initial goal of a 15% improvement

Project 2

Airline Customer Value Analysis

The second project, titled "Airline Customer Value Analysis," delves into **unsupervised learning to cluster airline customers based on their value.**

Utilizing tools such as **Python, Pandas, Scikit-Learn, and Matplotlib**, the team conducted **exploratory data analysis, feature engineering, and K-means clustering.**

The project **successfully identified four distinct customer clusters**, each with unique characteristics regarding flight habits, financial engagement, and loyalty program involvement. **Key findings include tailored business recommendations for each cluster to enhance customer engagement and loyalty.**

[Click here to see project on GitHub](#)

Deskripsi Dataset

Code	Description
MEMBER_NO-b	: ID Member
FFP_DATE	: Frequent Flyer Program Join Date
FIRST_FLIGHT_DATE	: Tanggal Penerbangan pertama
GENDER	: Jenis Kelamin
FFP_TIER	: Tier dari Frequent Flyer Program
WORK_CITY	: Kota Asal
WORK_PROVINCE	: Provinsi Asal
WORK_COUNTRY	: Negara Asal
AGE	: Umur Customer
LOAD_TIME	: Tanggal data diambil
FLIGHT_COUNT	: Jumlah penerbangan Customer
BP_SUM	: Rencana Perjalanan
SUM_YR_1	: Fare Revenue
SUM_YR_2	: Votes Prices
SEG_KM_SUM	: Total jarak(km) penerbangan yg sudah dilakukan
LAST_FLIGHT_DATE	: Tanggal penerbangan terakhir
LAST_TO_END	: Jarak waktu penerbangan terakhir ke pesanan penerbangan paling akhir
AVG_INTERVAL	: Rata-rata jarak waktu
MAX_INTERVAL	: Maksimal jarak waktu
EXCHANGE_COUNT	: Jumlah penukaran
avg_discount	: Rata rata discount yang didapat customer
Points_Sum	: Jumlah poin yang didapat customer
Point_NotFlight	: point yang tidak digunakan oleh members

Project 2

Exploratory Data Analysis

1

2

3

4

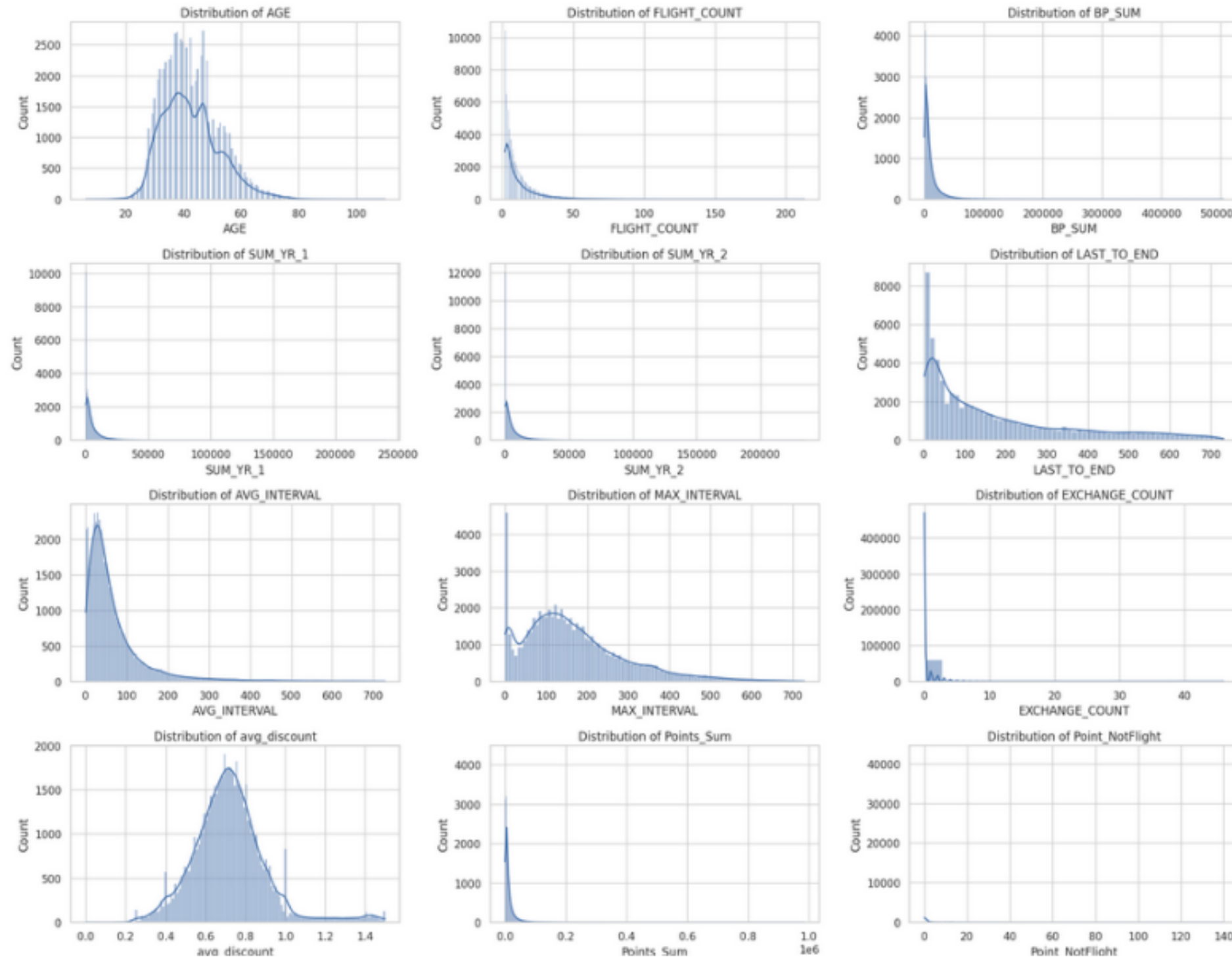
5

6

7

8

9



Statistical Analysis of Numeric Columns:

- **AGE:** The majority of customers are between 35 and 48 years old.
- **FLIGHT_COUNT:** Shows a wide range, indicating a mix of frequent and occasional flyers.
- **SEG_KM_SUM (Total Flight Distance):** Indicates a right-skewed distribution, suggesting most customers fly short distances.
- **AVG_INTERVAL (Average Time Between Flights):** Right-skewed, many customers do not fly frequently.
- **avg_discount:** Mostly clustered around 0.6 to 0.8, forming a normal distribution.
- **Points_Sum:** Highly varied, with most customers having fewer points.

Project 2

Exploratory Data Analysis

1

2

3

4

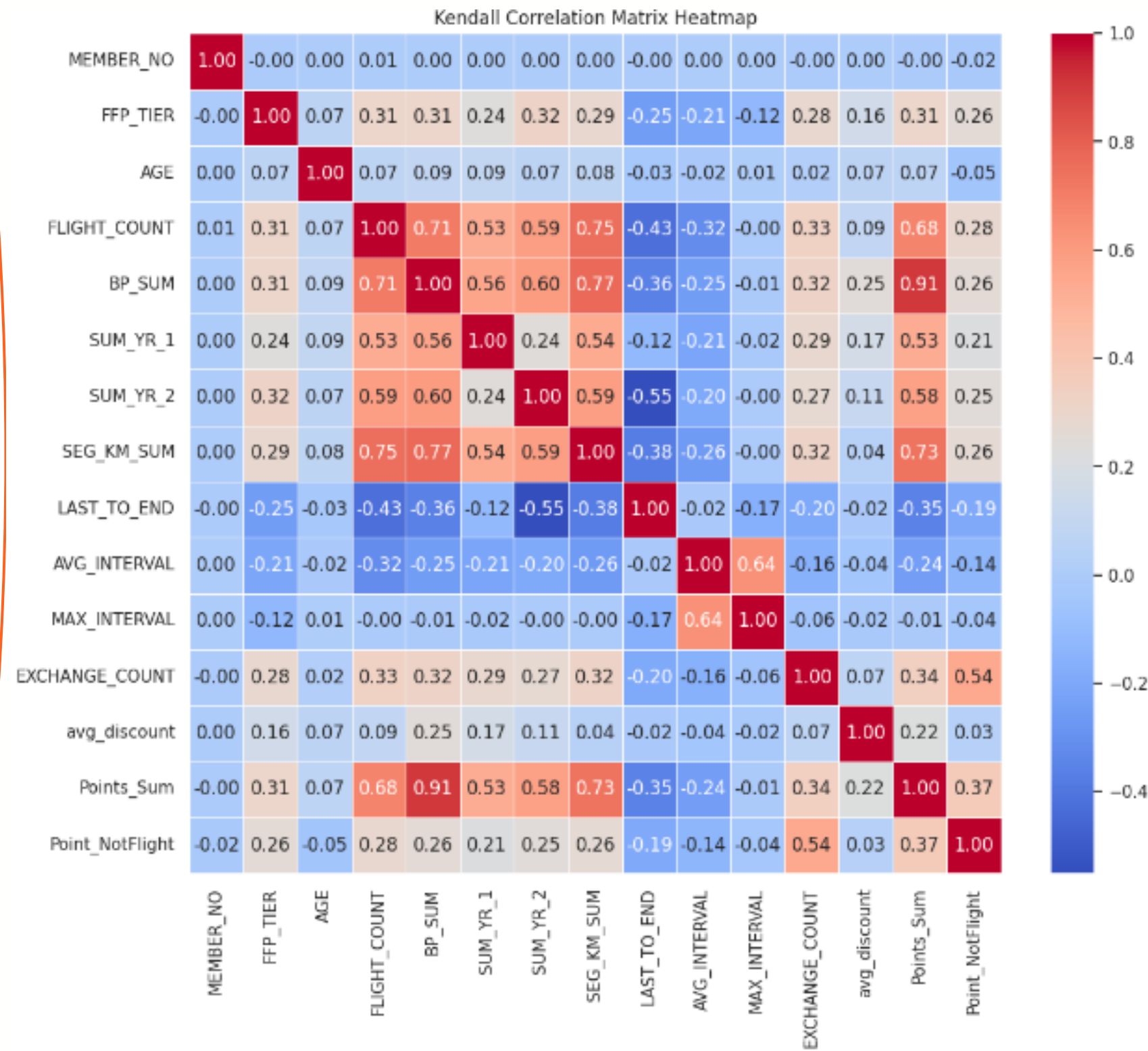
5

6

7

8

9



FFP_TIER shows a moderate positive correlation with **FLIGHT_COUNT**, **BP_SUM**, **SUM_YR_1**, **SUM_YR_2**, **SEG_KM_SUM**, and **Points_Sum**, indicating that higher tiers in the frequent flyer program are associated with more flights, further distances, higher fare earnings, and more accumulated points.

FLIGHT_COUNT is strongly correlated with **BP_SUM**, **SUM_YR_1**, **SUM_YR_2**, **SEG_KM_SUM**, and **Points_Sum**, suggesting frequent flyers tend to earn more, accumulate more points, and travel longer distances.

Project 2

K-Means Clustering

1

2

3

4

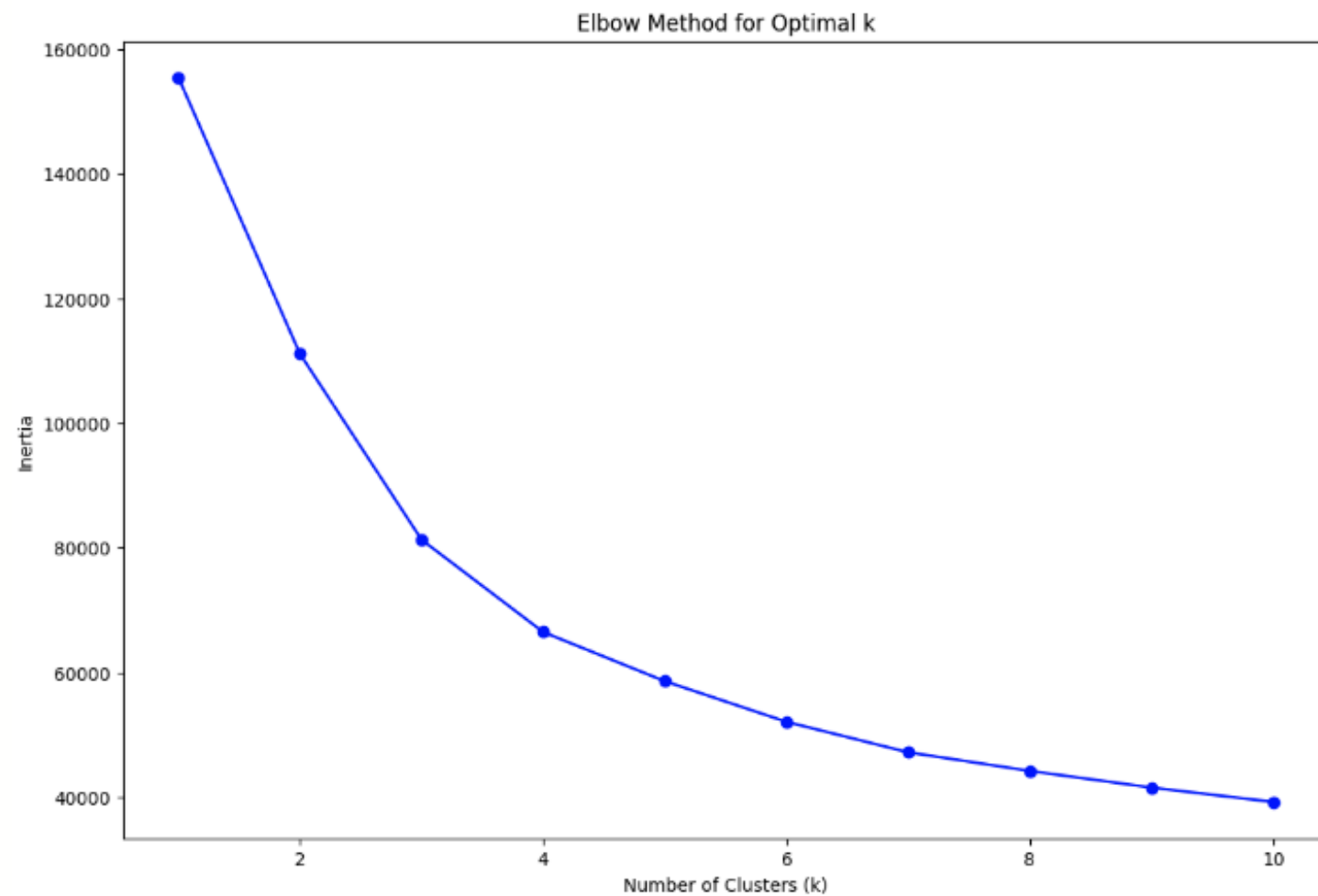
5

6

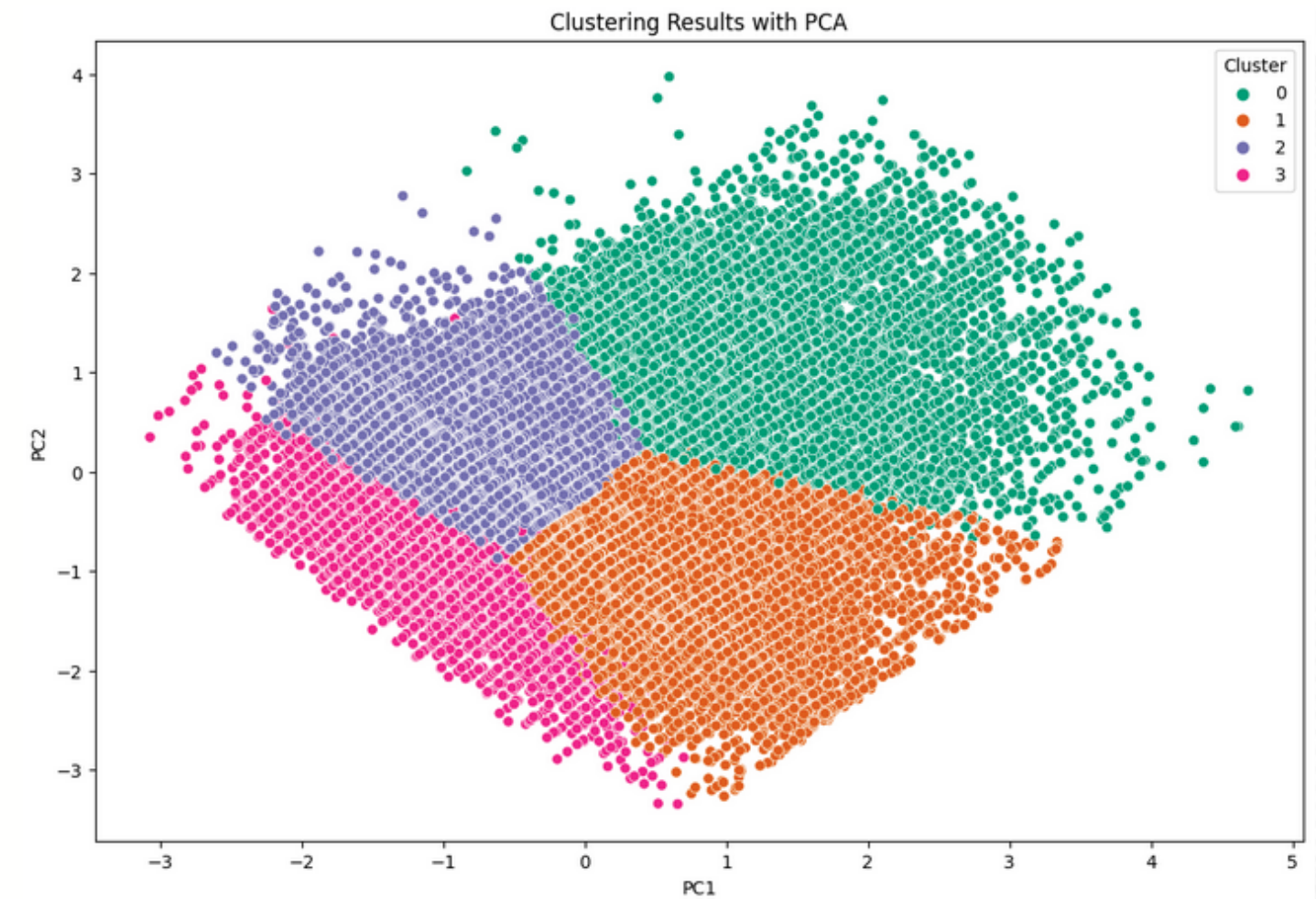
7

8

9



By using the **Elbow Method**, the number of clusters that can be used is 4 clusters.



Based on visualization using PCA method, the cluster has been divided into 4 clusters (labeled 0-3), which obtained **0.31 of Silhouette Score**.

Project 2

Business Insight & Recommendation

1

2

3

4

5

6

7

8

9

Business Insight

There are four distinct customer clusters identified within an airline's customer base.

- **Cluster 0: Established Regulars**
Middle-aged, frequent long-distance flyers with high fare revenue and moderate loyalty engagement.
- **Cluster 1: Older Infrequent Flyers**
Older age group, less frequent flyers with lower fare spending and minimal loyalty engagement.
- **Cluster 2: Young Frequent Flyers**
Younger demographic, regular short-distance flyers with moderate fare revenue and varied loyalty engagement.
- **Cluster 3: Occasional Flyers**
Young to middle-aged, least frequent and shortest distance flyers with lowest fare revenue and minimal loyalty engagement.

Business Recommendation

Focusing on tailored strategies to enhance customer engagement and spending.

- **Cluster 0: Established Regulars**
Enhance rewards, use personalized marketing to maintain their high engagement.
- **Cluster 1: Older Infrequent Flyers**
Increase loyalty program awareness for older, less frequent flyers, and implement retargeting campaigns.
- **Cluster 2: Young Frequent Flyers**
Offer youth-oriented programs for young frequent flyers, including deals and social media engagements.
- **Cluster 3: Occasional Flyers**
Develop incentives for more frequent flying, communicate the benefits of the loyalty program.

Project 3

Telco Customers Churn Analysis

	customerID	gender	SeniorCitizen	Partner	Dependents	tenure	PhoneService	MultipleLines	InternetService	OnlineSecurity	...
0	7590-VHVEG	Female	0	Yes	No	1	No	No phone service	DSL	No	...
1	5575-GNVDE	Male	0	No	No	34	Yes	No	DSL	Yes	...
2	3668-QPYBK	Male	0	No	No	2	Yes	No	DSL	Yes	...
3	7795-CFOCW	Male	0	No	No	45	No	No phone service	DSL	Yes	...
4	9237-HQITU	Female	0	No	No	2	Yes	No	Fiber optic	No	...

5 rows × 21 columns

Understanding the dataset:

Presented a set of comprehensive data on customer interactions and demographics. Each row represents a unique customer; columns detail various attributes such as service usage, charges, and personal demographics.

Purpose: To derive actionable insights for minimizing churn and enhancing customer loyalty.

[Click here to see project on GitHub](#)

Project 3 - Telco Customers Churn Analysis

1

2

3

4

5

6

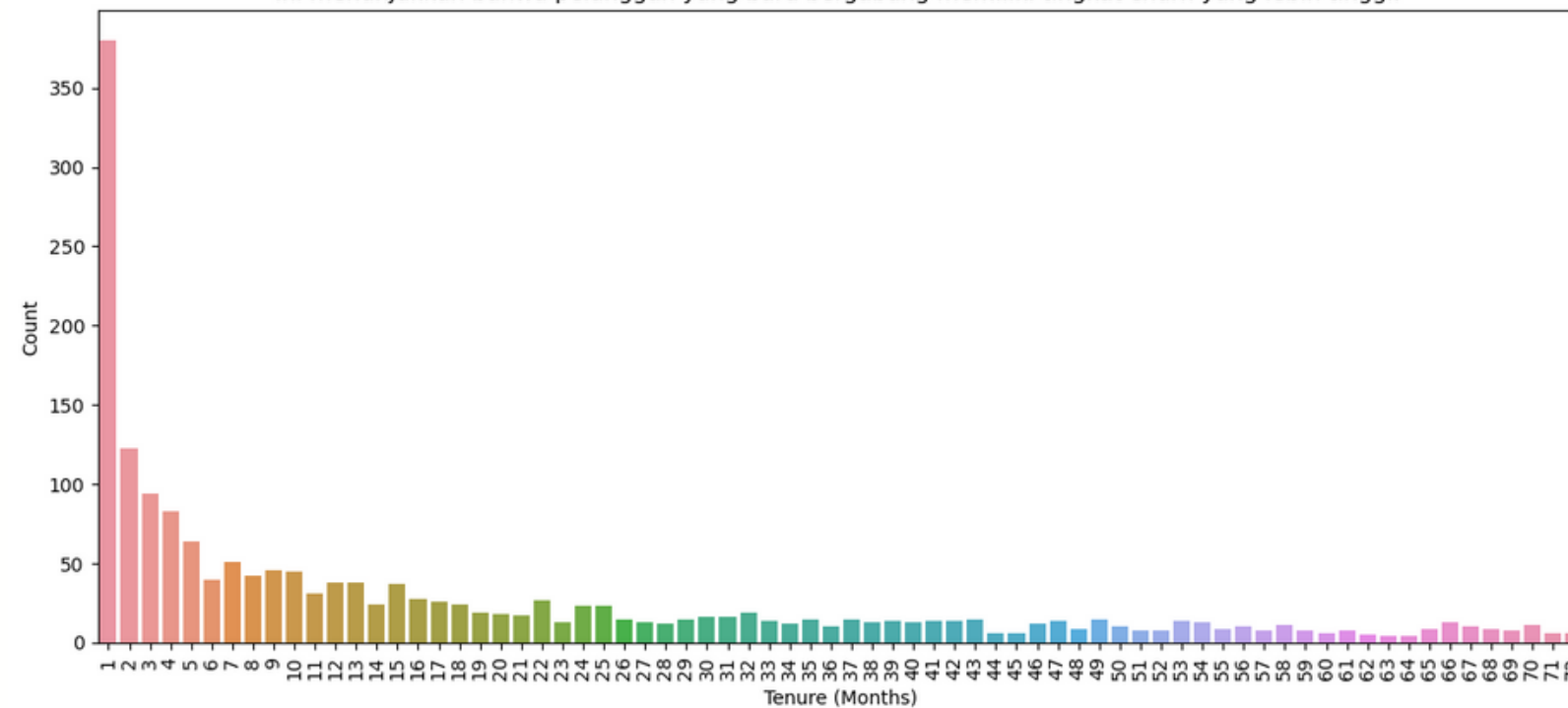
7

8

9

Churn vs. Tenure

Dari Plot 1, dapat dilihat bahwa banyak pelanggan yang melakukan churn pada saat tenure sekitar 1-2 bulan. Ini menunjukkan bahwa pelanggan yang baru bergabung memiliki tingkat churn yang lebih tinggi.



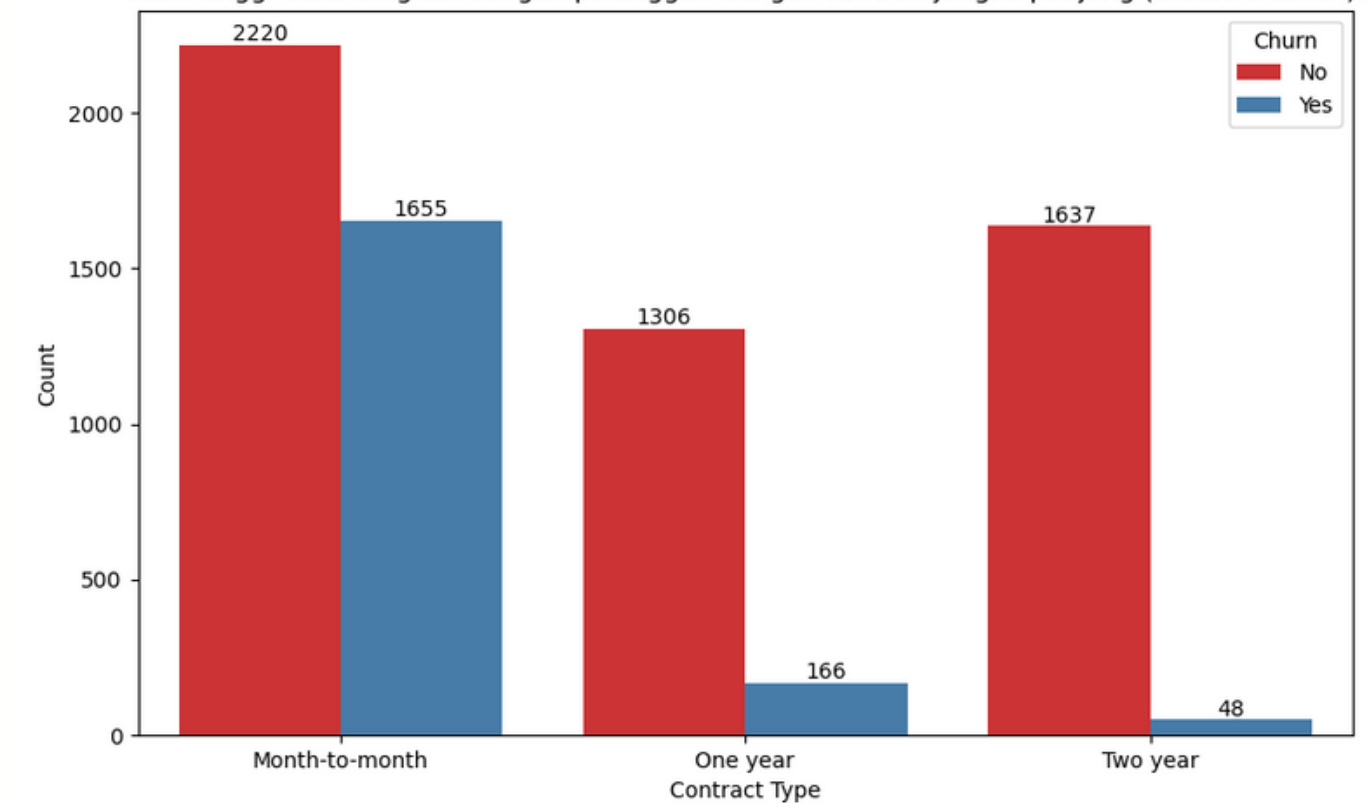
Churn vs. Tenure

Insight: Customers are more likely to churn within the first 1-2 months of service.

Action: Implement a targeted onboarding process with follow-up check-ins and special offers to new customers to increase early engagement and reduce early churn.

Churn vs. Contract Type

Dari Plot 2, dapat dilihat bahwa pelanggan dengan kontrak bulanan memiliki tingkat churn yang jauh lebih tinggi dibandingkan dengan pelanggan dengan kontrak jangka panjang (1 atau 2 tahun).



Churn vs. Contract Type

Insight: Customers with month-to-month contracts are churning at a higher rate compared to those with one or two-year contracts.

Action: Encourage customers to move to longer-term contracts through loyalty discounts and highlight the benefits of commitment over a month-to-month plan.

Project 3 - Telco Customers Churn Analysis

1

2

3

4

5

6

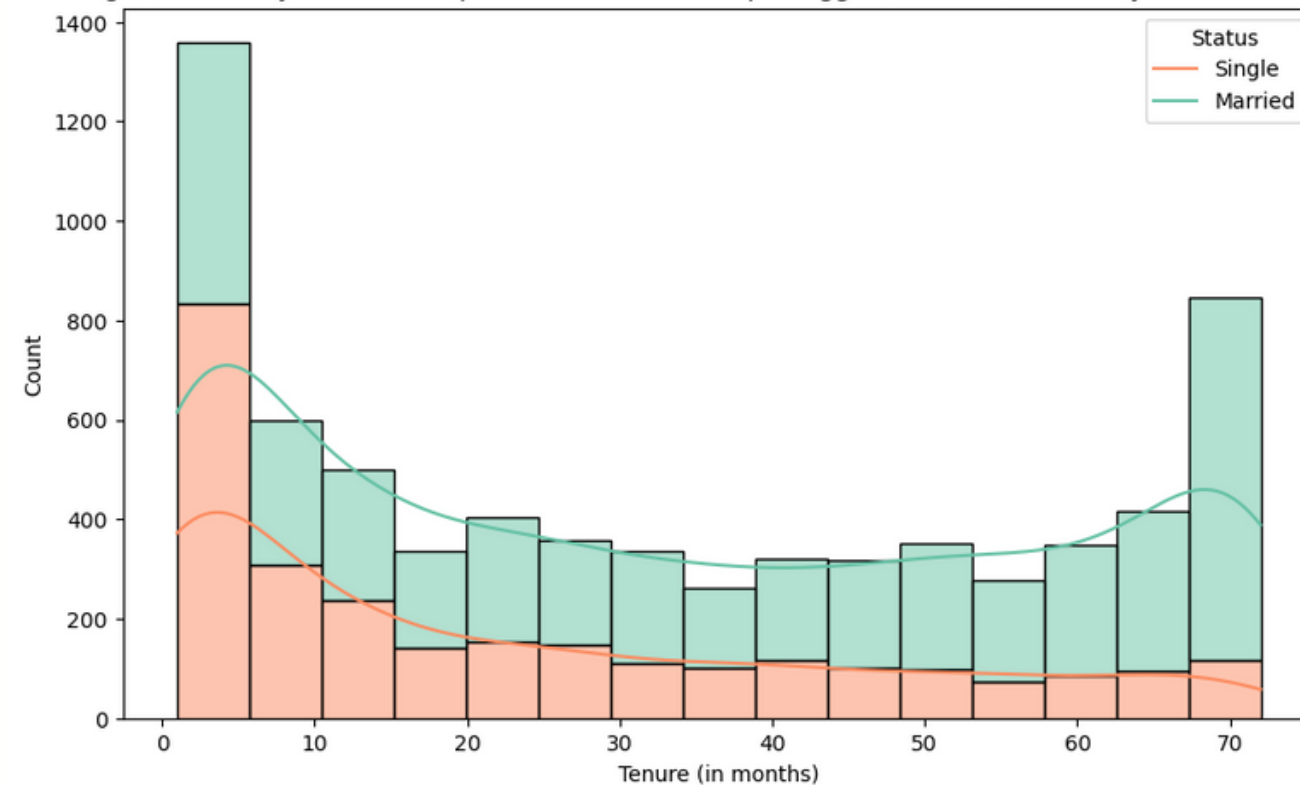
7

8

9

Tenure vs. Status

Dari Plot 3, ditunjukkan persebaran status pelanggan (Single vs. Married) berdasarkan tenure. Pelanggan 'Single' lebih banyak terutama pada tenure awal dan pelanggan 'Married' lebih banyak di tenure lama.



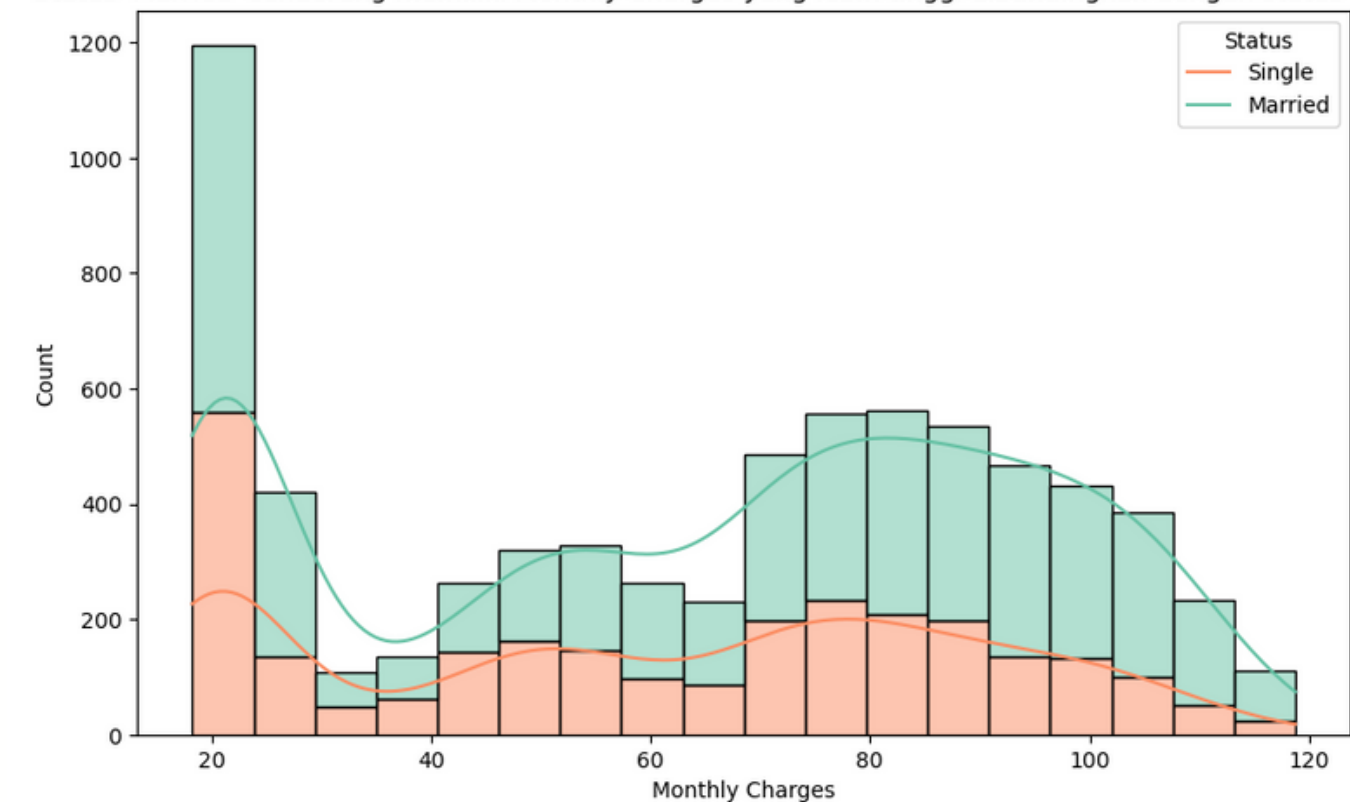
Tenure vs. Status (Single vs. Married)

Insight: Single customers show higher churn rates in early tenure, while married customers show more stability over time.

Action: Tailor marketing and retention efforts to address the specific needs and preferences of single customers, particularly in the initial stages of their tenure.

Monthly Charges vs. Status

Dari Plot 4, dapat dilihat adanya perbedaan dalam distribusi Monthly Charges berdasarkan status pelanggan. Status 'Married' cenderung memiliki Monthly Charges yang lebih tinggi dibandingkan dengan status 'Single'.



Monthly Charges vs. Status

Insight: Married customers tend to have higher monthly charges than single customers.

Action: Review pricing structures and consider family or couple plans that provide better value, potentially reducing churn among customers with higher monthly charges.

Project 3 - Telco Customers Churn Analysis

1

2

3

4

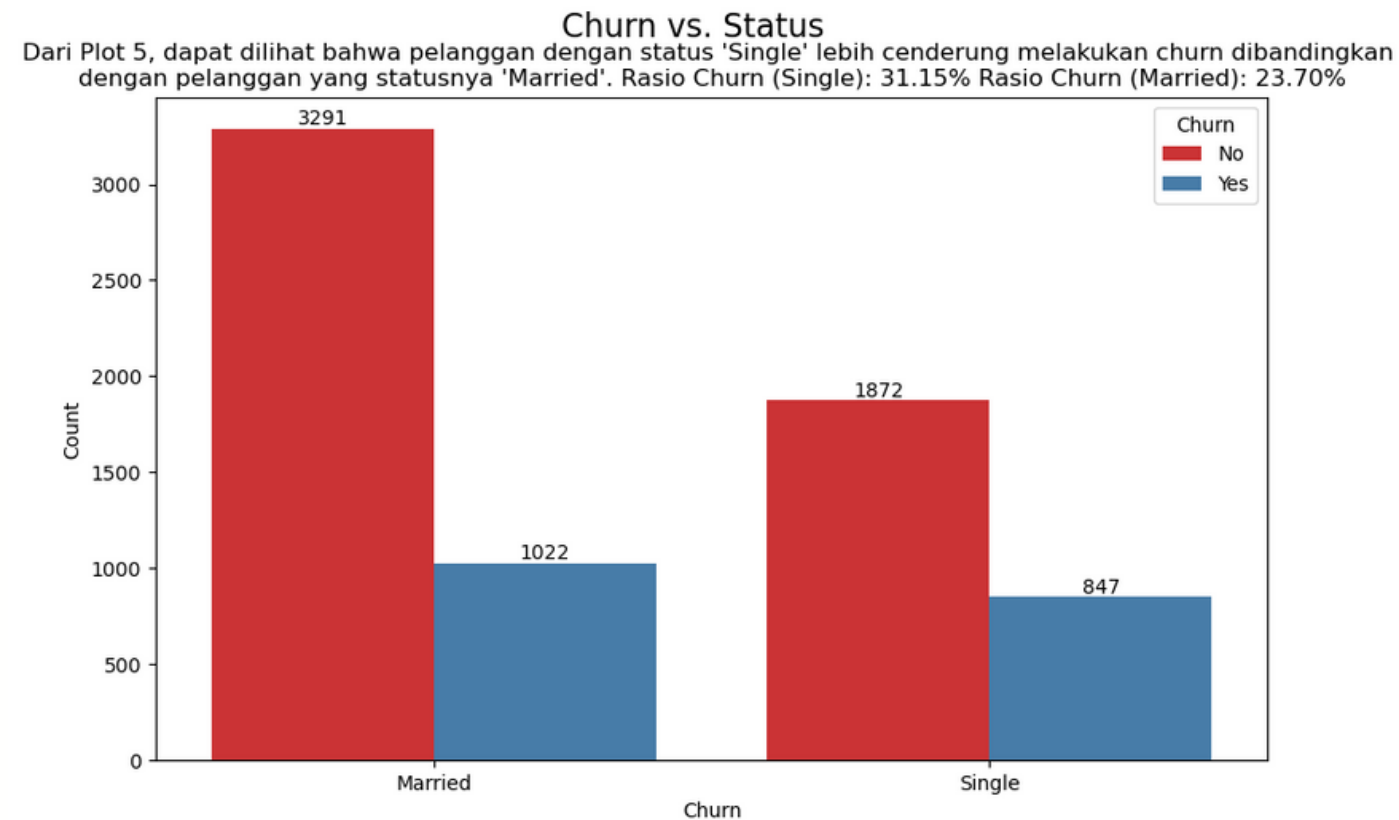
5

6

7

8

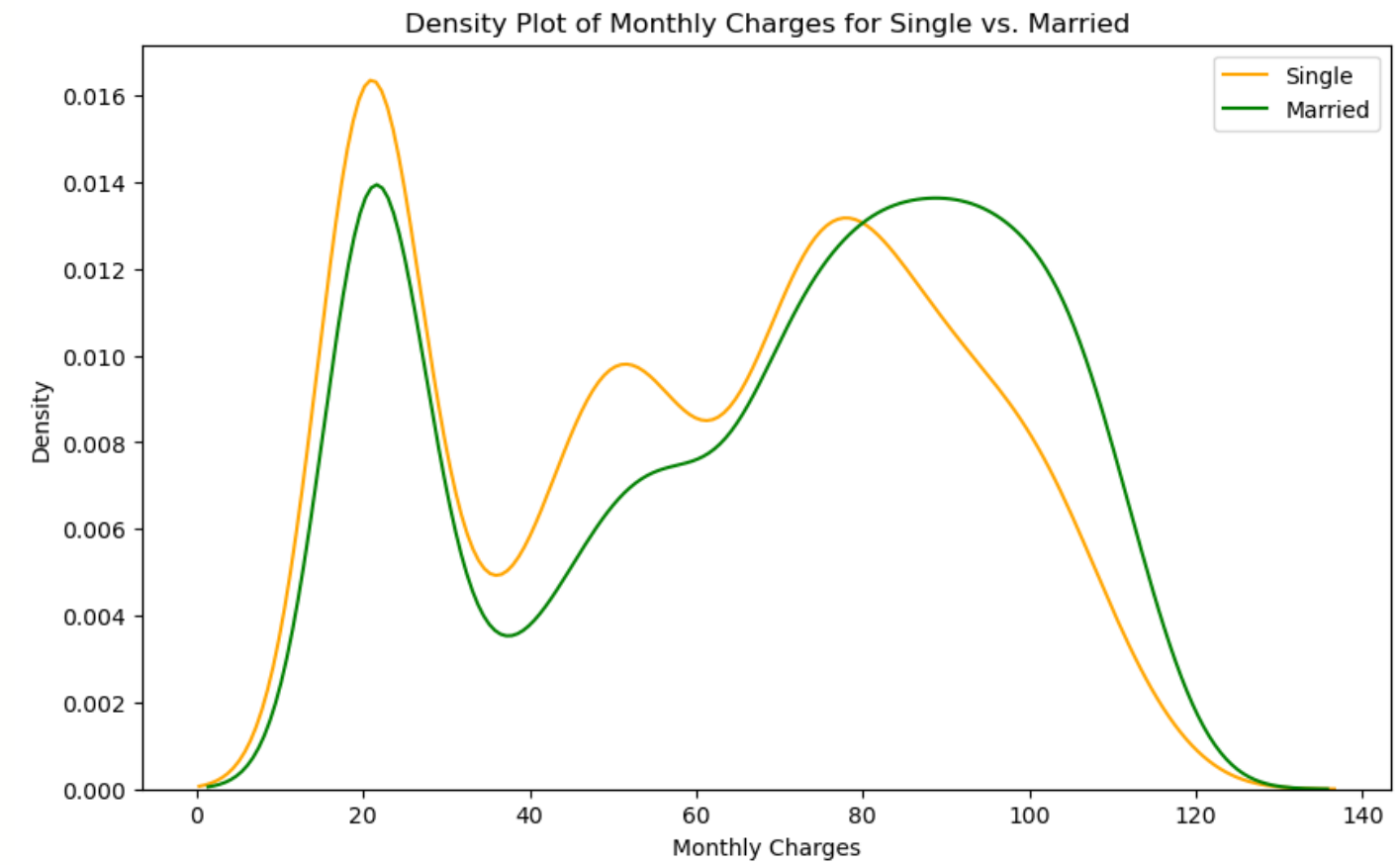
9



Churn vs. Status

Insight: Single customers have a higher churn rate compared to married customers.

Action: Develop targeted retention strategies for single customers, such as community building or social features that may increase their engagement and loyalty.

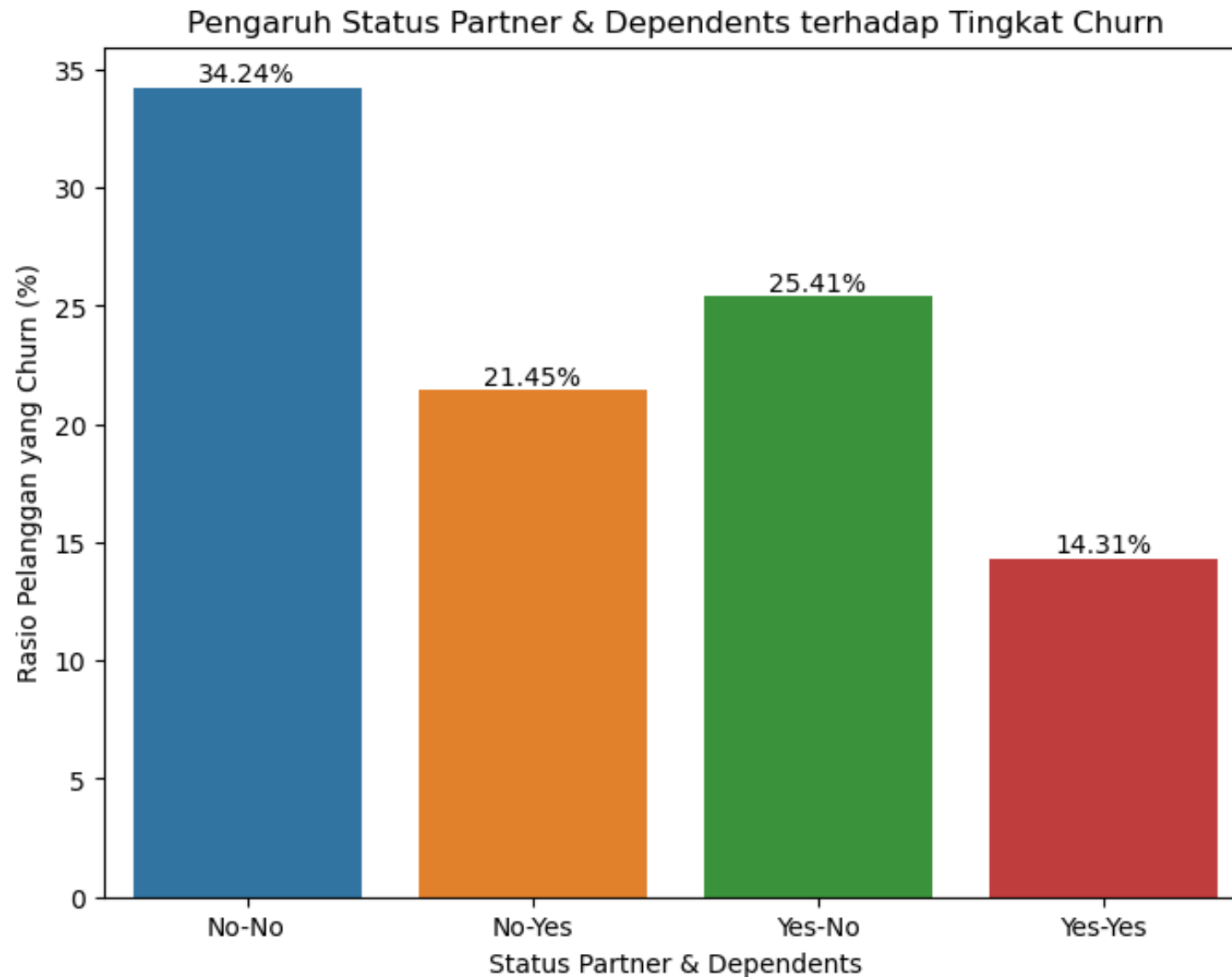


Density Plot of Monthly Charges for Single vs. Married

Insight: There are distinct peaks in monthly charges for single and married customers, indicating different usage patterns or plan preferences.

Action: Offer personalized plan recommendations based on usage patterns to ensure customers feel they are getting the best value for their money.

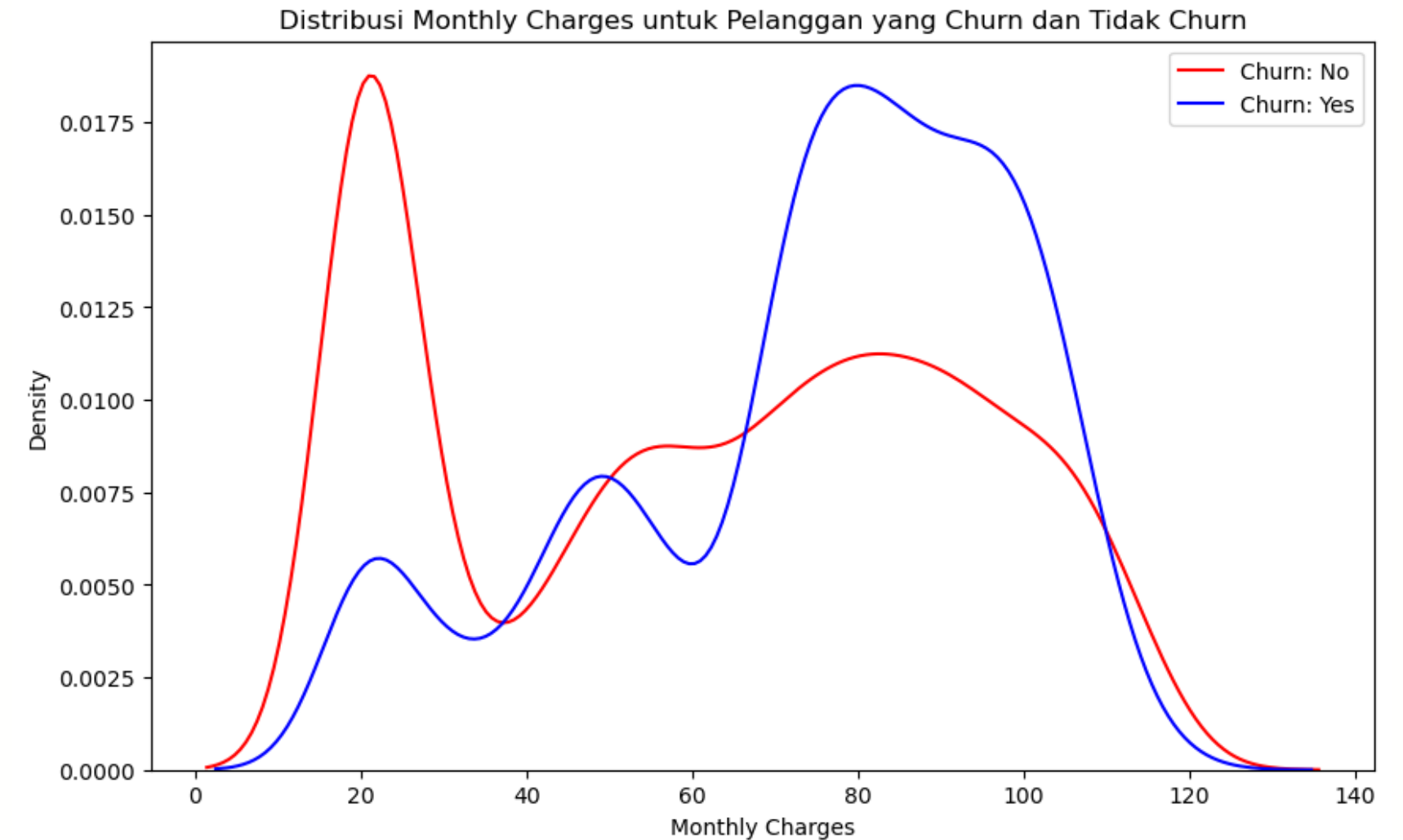
Project 3 - Telco Customers Churn Analysis



Influence of Partner & Dependents Status on Churn Rate

Insight: Customers without partners and dependents have the highest churn rate.

Action: Create targeted support and service bundles for individuals without partners or dependents to better meet their unique needs and reduce churn.



Distribution of Monthly Charges for Churning and Non-Churning Customers

Insight: Churning customers tend to have varying monthly charges, with some peaks possibly indicating price sensitivity.

Action: Investigate price points causing churn and consider offering tiered services or discounts to retain price-sensitive customers.

Project 3

Tanamduit Dashboard - SQL & Looker Studio



```
select m.customer_id,cust.Customer_name,
order_id, m.product_code, product_name,trx_date,trx_units,nav_value,trx_amount,order_status,
(case trx_type
when 1 then 'Subscription'
when 2 then 'Redemption'
when 3 then 'Switching'
when 5 then 'Dividen'
else '-' end) as Trx_type
from smc_order_mutualfund m
inner join
(select customer_id,concat(firstname," ",lastname) as Customer_name,'Individual' as Customer_Type,sid
from smc_customer_individual_profiles
union
select customer_id,corporate_name as Customer_name,'Corporate' as Customer_Type,sid
from smc_customer_corporate_profiles) as Cust on m.customer_id = cust.customer_id
where order_status = 6
```

```
select m.customer_id,cust.Customer_name,
order_id, m.product_code, product_name,trx_datetime,trx_amount,order_status,
(case trx_type
when 1 then 'Subscription'
when 2 then 'Redemption'
else '-' end) as Trx_type
from smc_order_bond m
inner join
(select customer_id,concat(firstname," ",lastname) as Customer_name,'Individual' as Customer_Type,sid
from smc_customer_individual_profiles
where open_sec_account = 4) as Cust on m.customer_id = cust.customer_id
where order_status = 4
```

```
select order_ref,product_code,product_name,
m.user_id,concat(firstname," ",lastname) as Customer_name,'Individual' as Customer_Type,
trx_datetime,total_premi,order_status
from smc_order_insurance m
inner join smc_customer_individual_profiles i on i.user_id = m.user_id
where order_status = 2
```

Description: The project focuses on developing an interactive dashboard to monitor and analyze customer financial activities and product engagements with Tanamduit, a financial services provider.

Tools & Technologies: Utilizing **SQL** and **Looker Studio** connected to MySQL databases, the project aimed to provide real-time insights into customer demographics, transaction behaviors, and investment patterns.

Key Findings or Results: Key outcomes included enhanced data-driven decision-making for targeted marketing strategies and product development, leading to improved customer engagement and satisfaction.



Thank you!

Reach me here

Phone : +628111475017

Email : amossubagio@gmail.com

Linkedin : [Amodya Subagio](#)