Course 3 Advanced Machine Learning
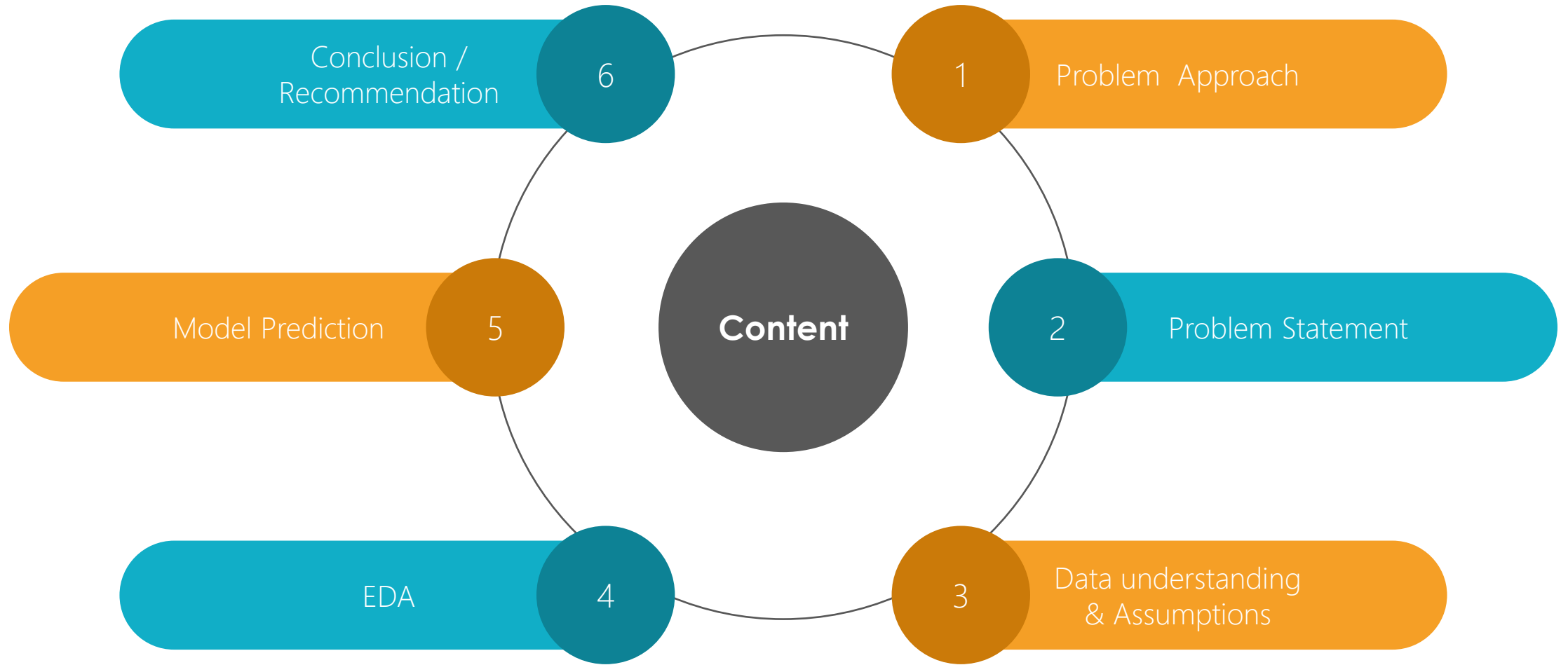
Telecom Churn Case Study
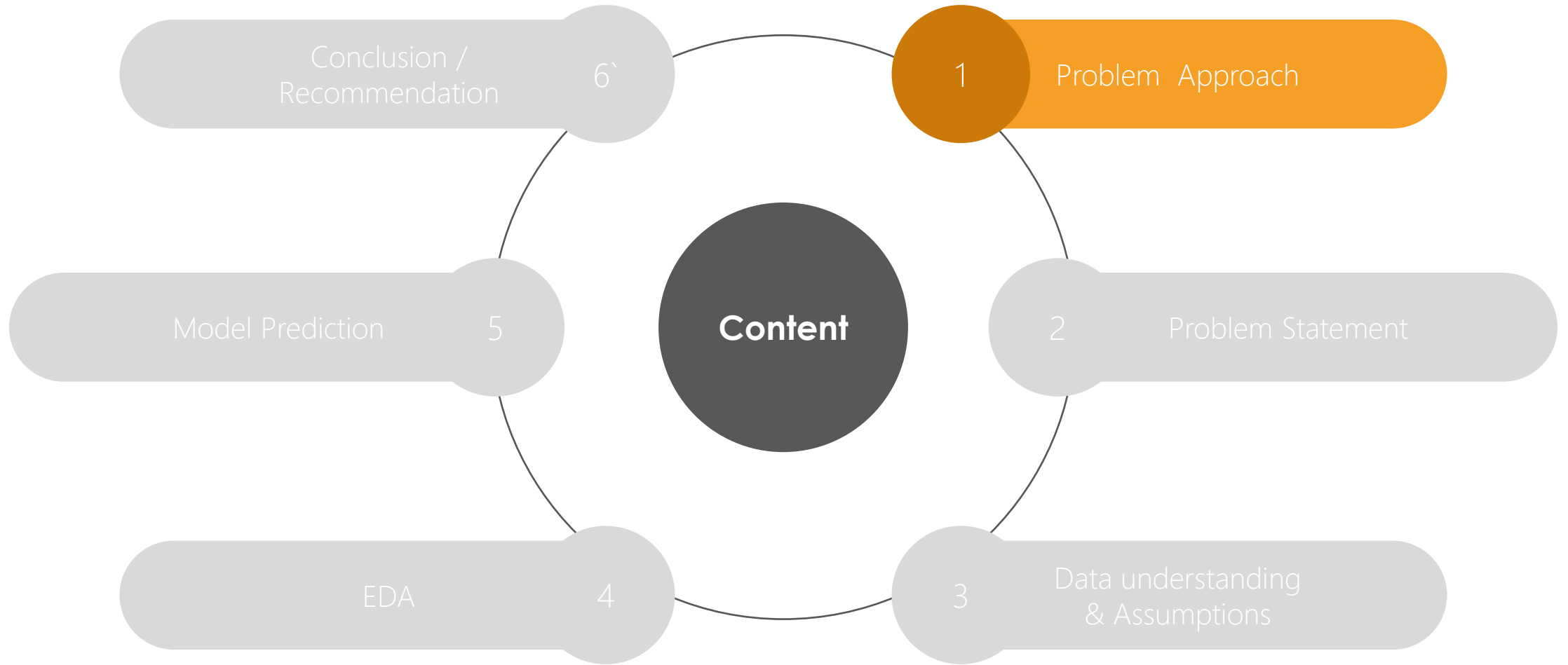
Team:
Akshatha Shetty (akshathashetty20@gmail.com)
UpGrad & IIITB | Data Science Program - October 2022 (DSC 49)
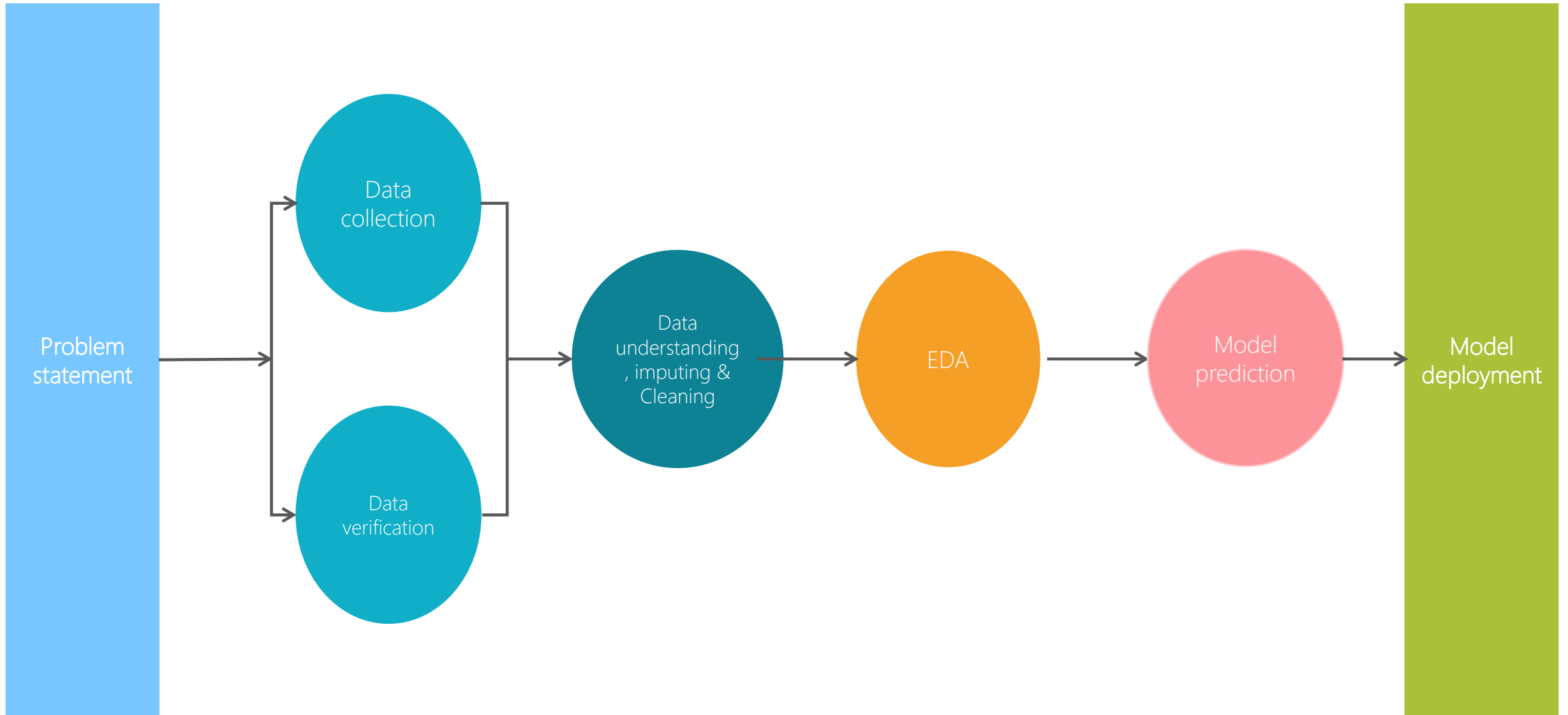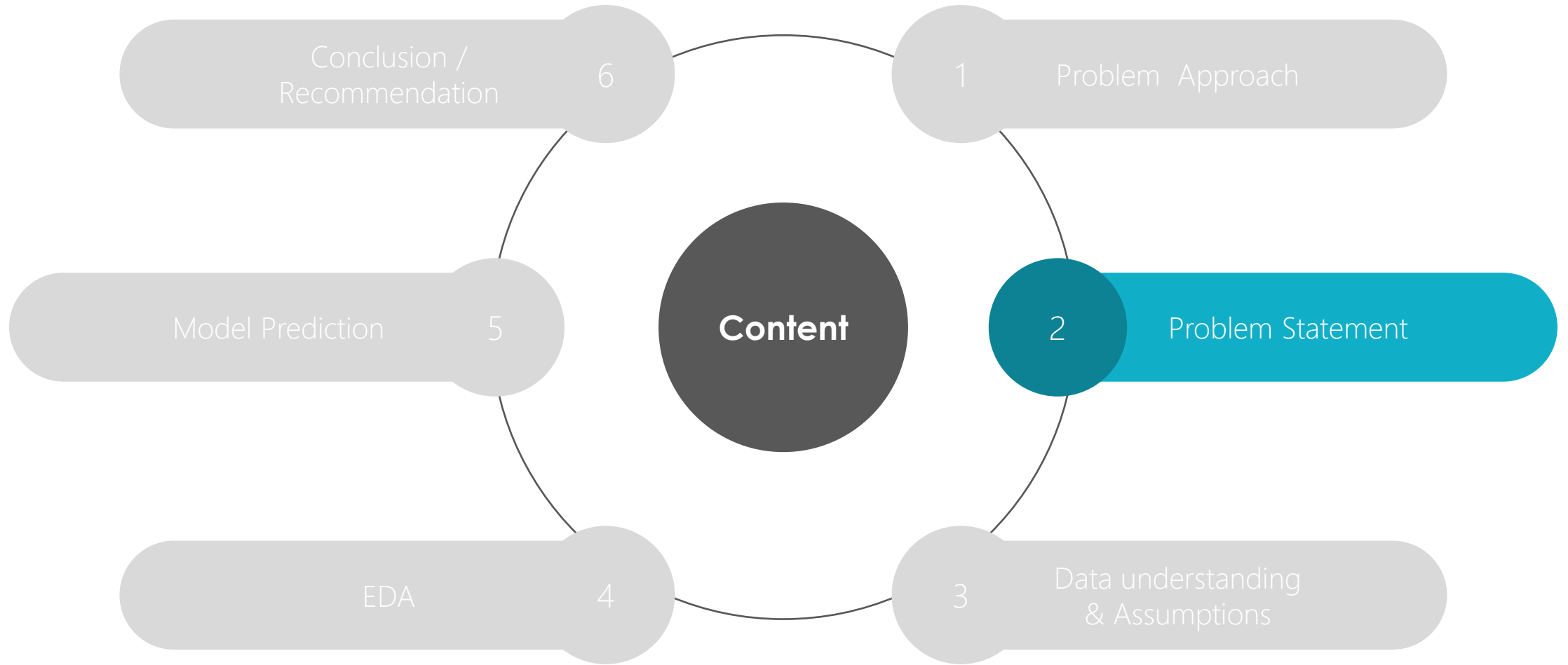
Content

Content

6 Conclusion / Recommendation

1 Problem Approach

5 Model Prediction

2 Problem Statement

4 EDA

3 Data understanding & Assumptions

Content

Conclusion / Recommendation 6`

6`

1 Problem Approach

Model Prediction 5

Content

2 Problem Statement

EDA 4

3 Data understanding & Assumptions

Problem Approach

Problem statement

Data collection

Data verification

Data understanding, imputing & Cleaning

EDA

Model prediction

Model deployment

Content

Conclusion / Recommendation 6

1 Problem Approach

Model Prediction 5

Content

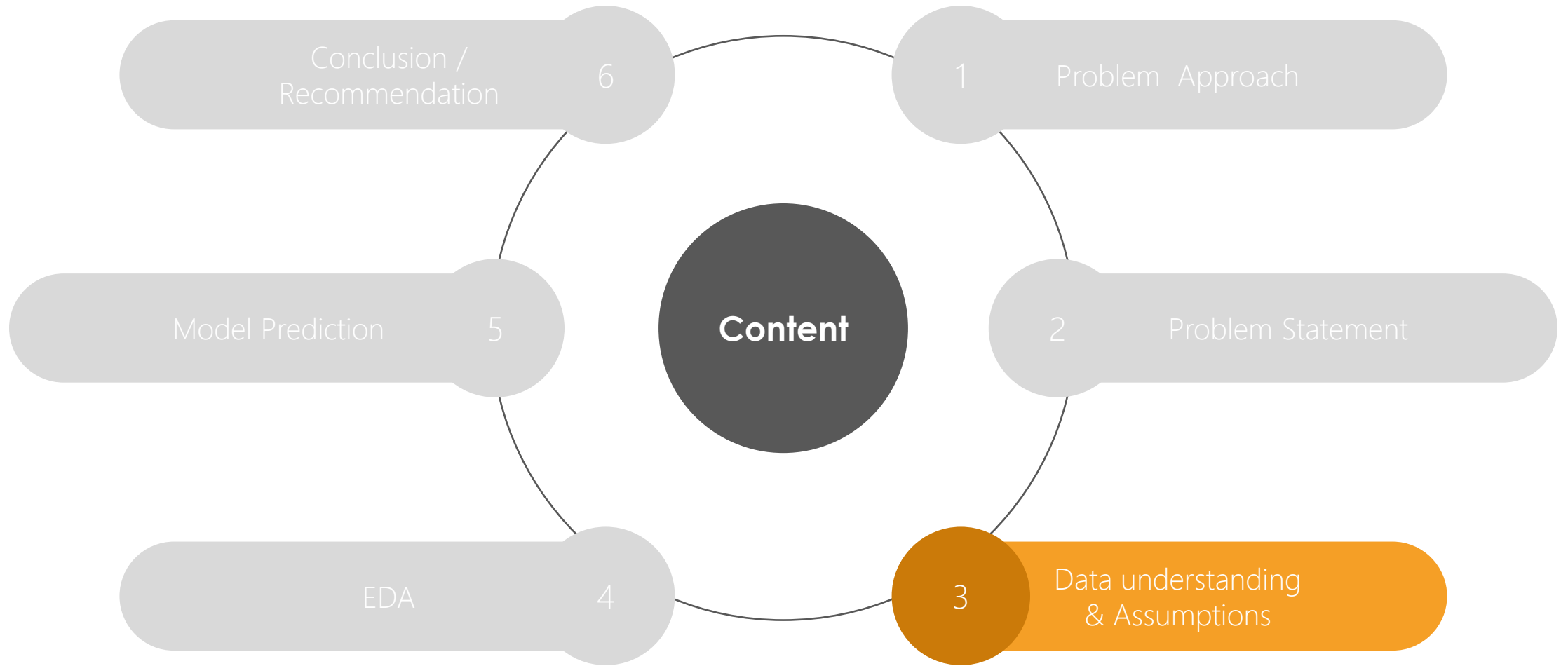2 Problem Statement

EDA 4

3 Data understanding & Assumptions

Problem Statement

➢ In the telecom industry, customers are able to choose from multiple service providers and actively switch from one operator to another. In this highly competitive market, the telecommunications industry experiences an average of 15-25% annual churn rate. Given the fact that it costs 5-10 times more to acquire a new customer than to retain an existing one, customer retention has now become even more important than customer acquisition.

➢ For many incumbent operators, retaining high profitable customers is the number one business goal. To reduce customer churn, telecom companies need to predict which customers are at high risk of churn.

➢ In this project, you will analyse customer-level data of a leading telecom firm, build predictive models to identify customers at high risk of churn and identify the main indicators of churn.
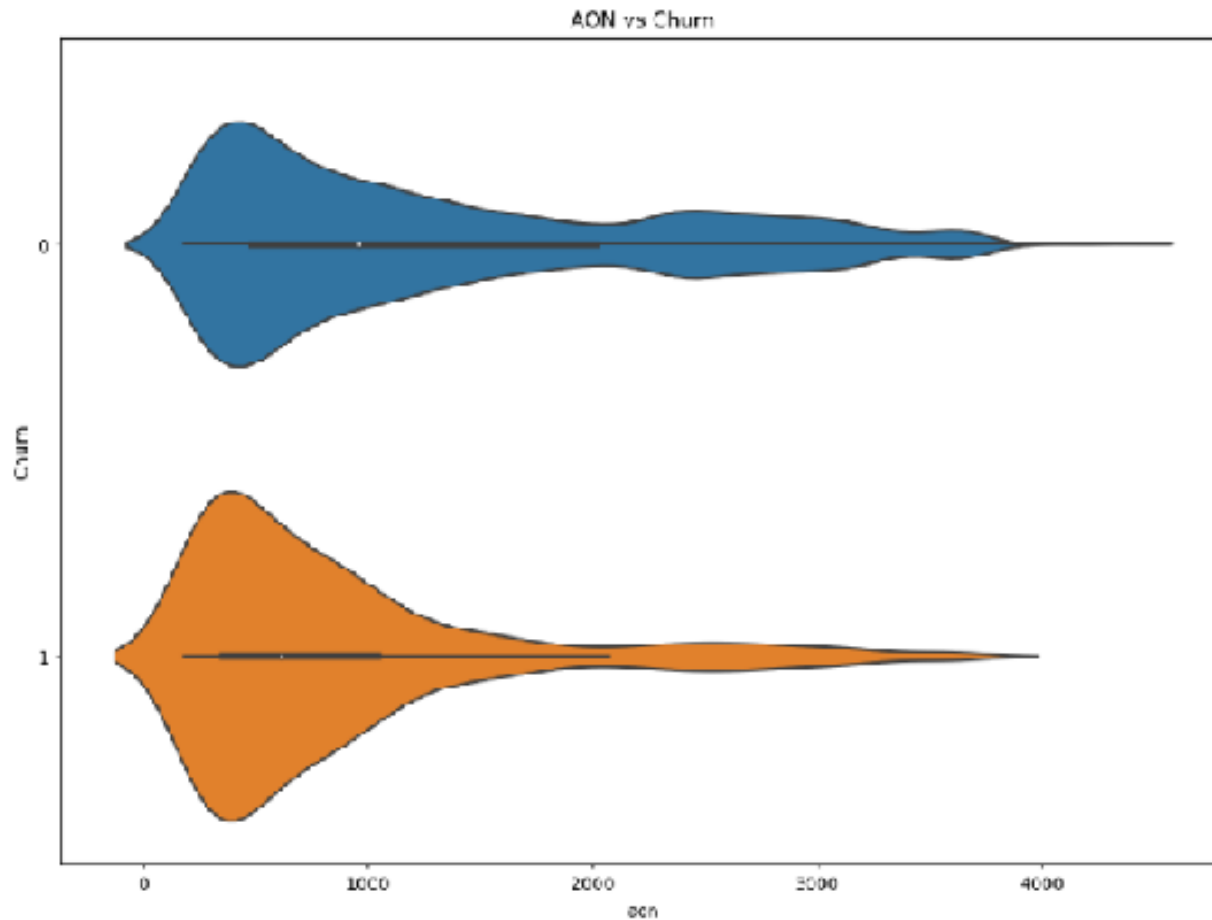
Content

- 6 Conclusion / Recommendation
- 1 Problem Approach
- Content
- 2 Problem Statement
- 5 Model Prediction
- 4 EDA
- 3 Data understanding & Assumptions

Data Understanding

➢ Total outgoing mou for each customer with missing *_mou data is 0.0
➢ Total incoming mou for each customer with missing *_mou data is 0.0
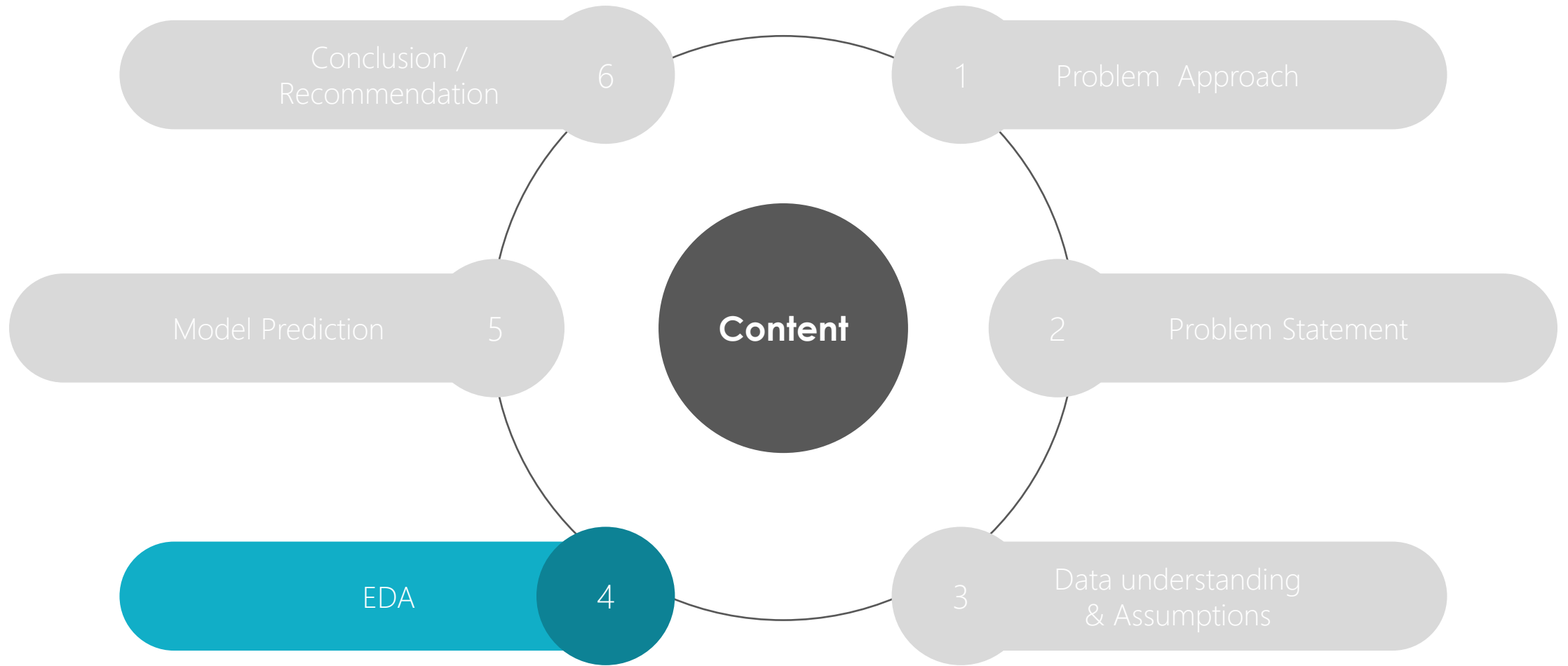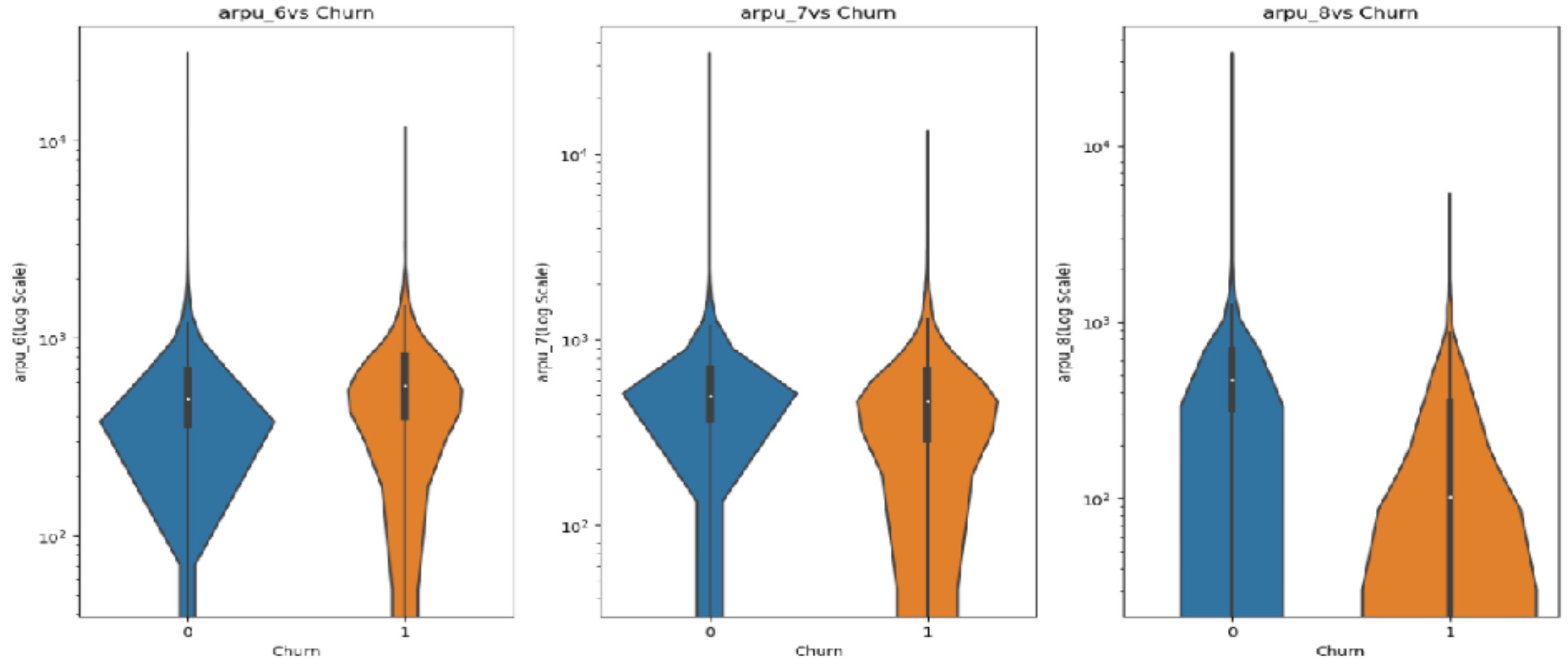➢ Imputing missing values using mode.

Data Assumptions

- The Customers with lesser 'AON' are more likely to churn when compared to customers with higher 'AON'.
- Most customers prefer the plans of '0' category.

Content

Conclusion / Recommendation 6

1 Problem Approach

Model Prediction 5

Content

2 Problem Statement

EDA 4

3 Data understanding & Assumptions

Exploratory Data Analysis

- ➤ Revenue generated by the customers who are about to churn is very unstable
- ➤ Customers whose arpu decreases in 7th month are more likely to churn when compared to ones with increase in arpu.

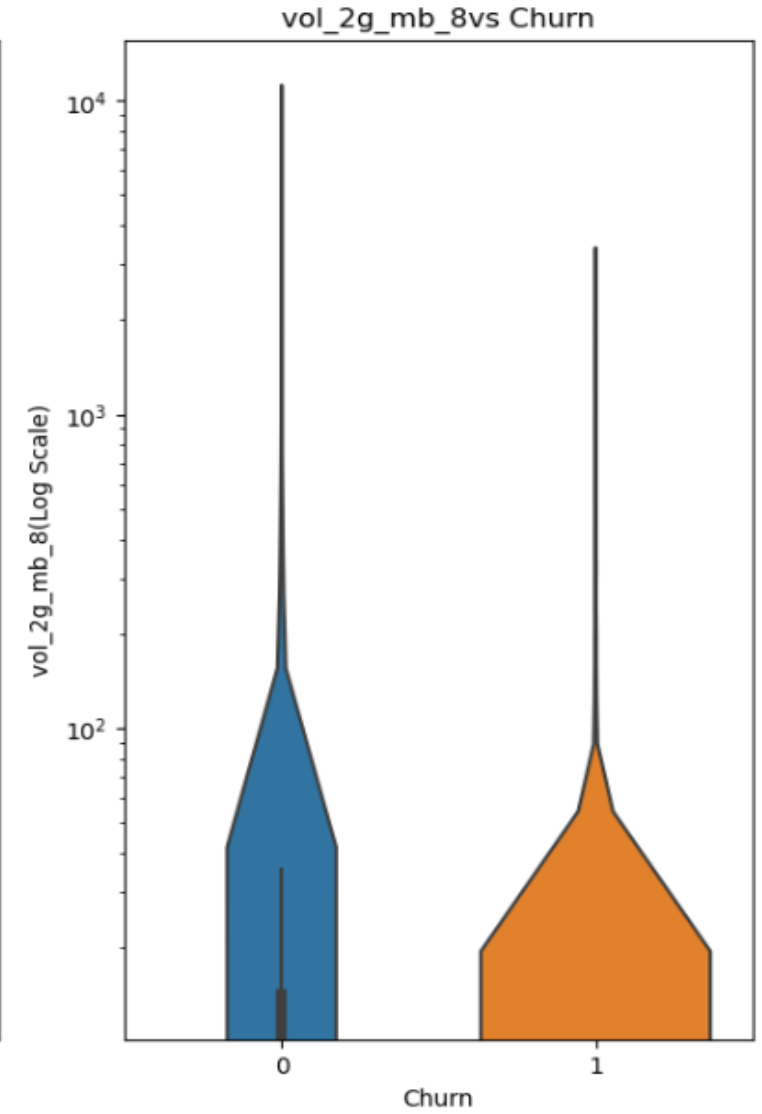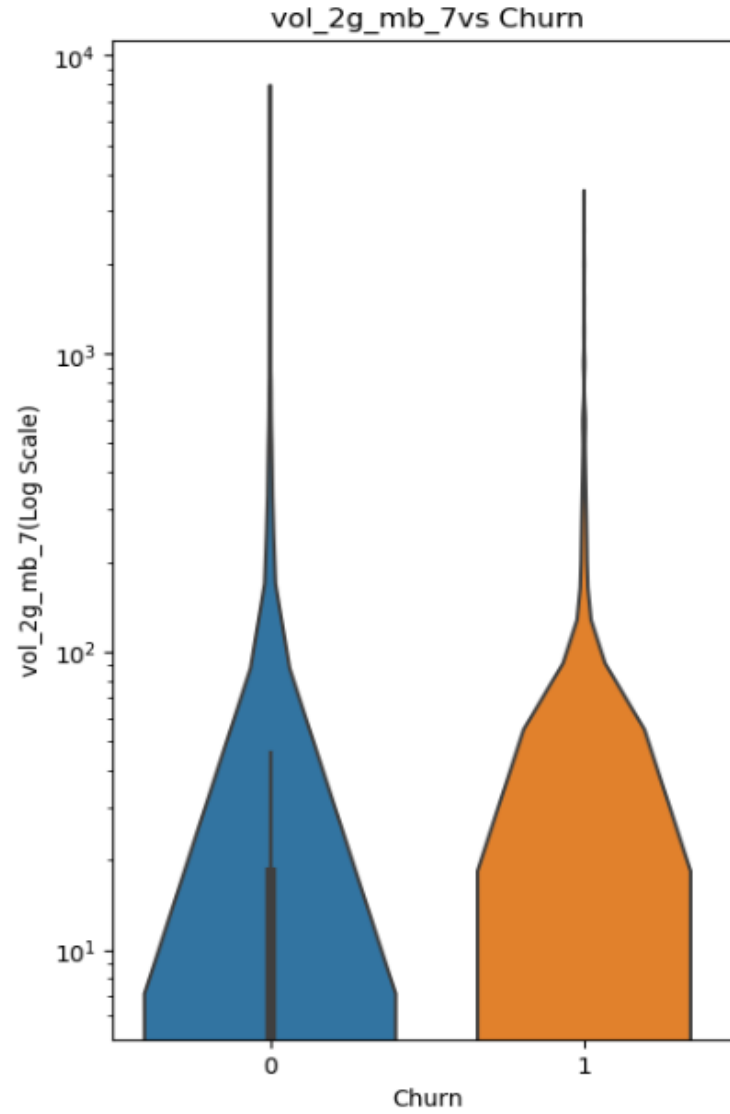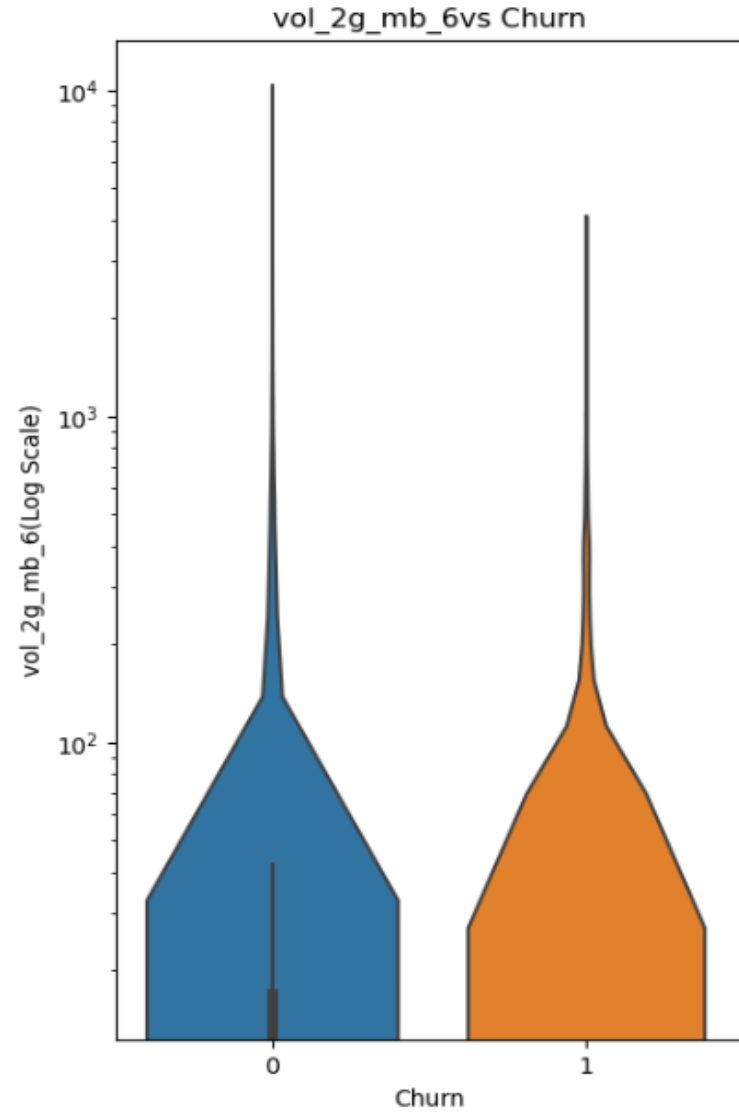# Exploratory Data Analysis
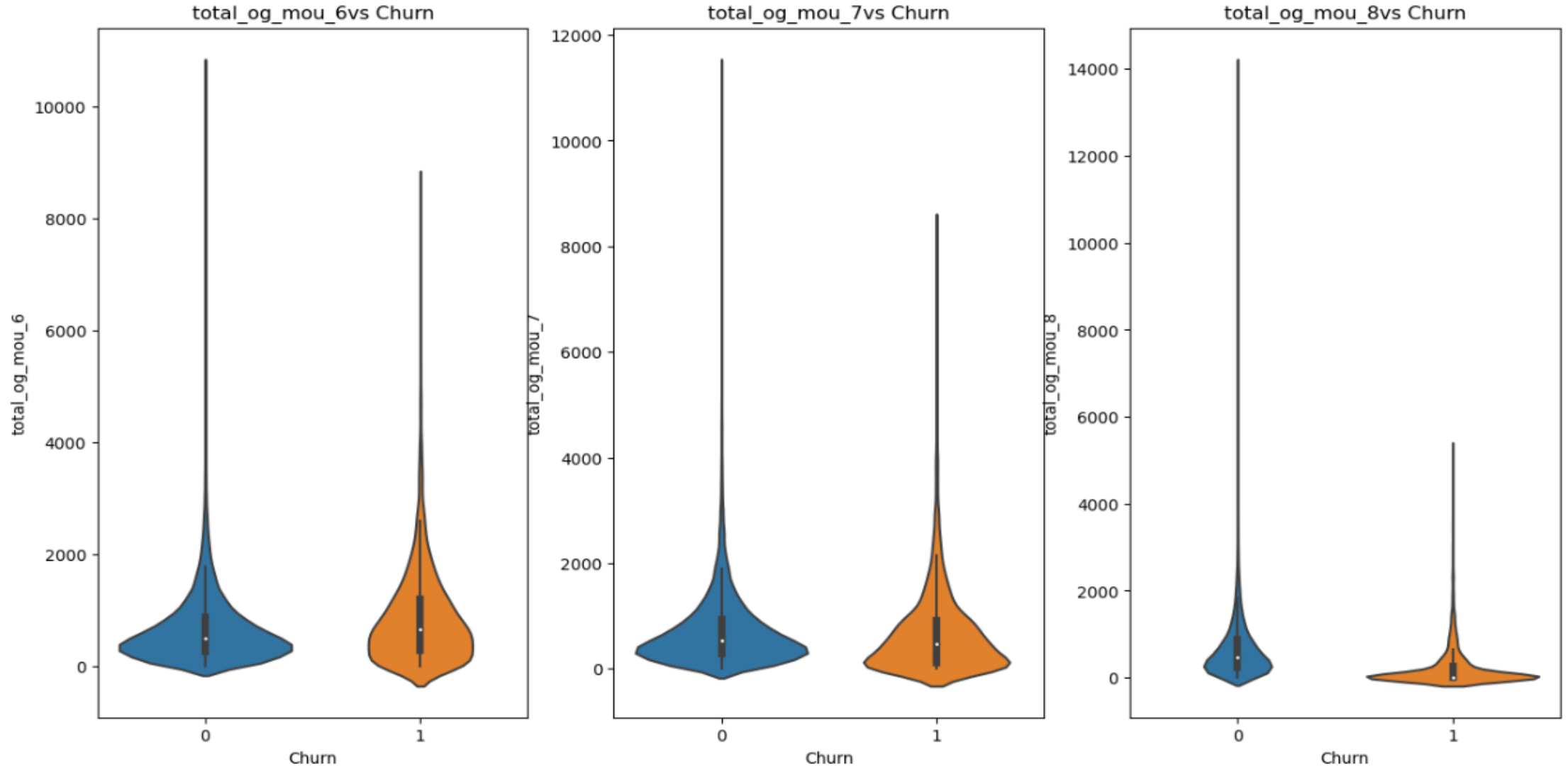
> ➤ Customer whose out going minutes of usage is less in 8th month compared to 6th and 7th month are more likely to churn
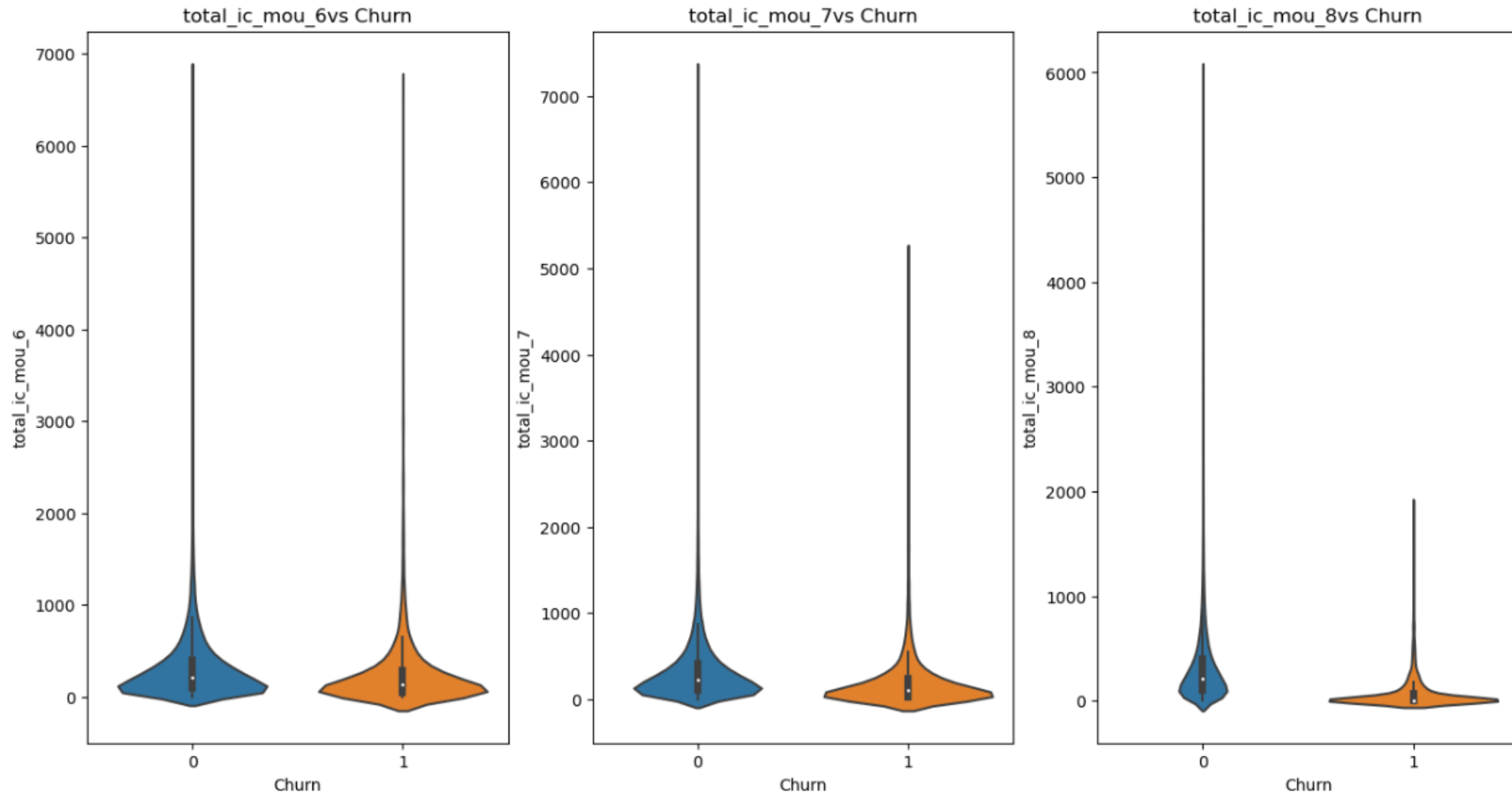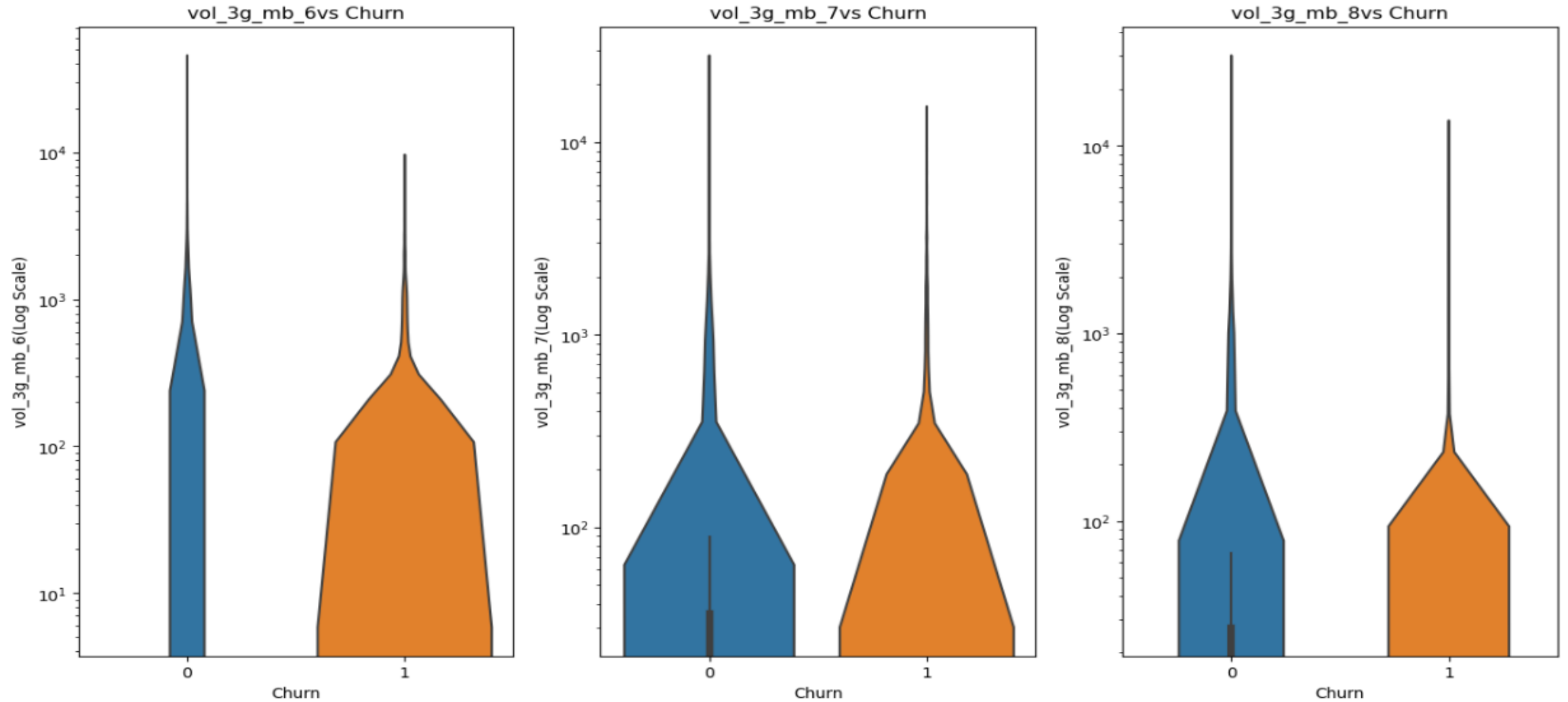
Analysis

➢ Customer whose incoming minutes of usage is less in 8th month compared to 6th and 7th month are more likely to churn
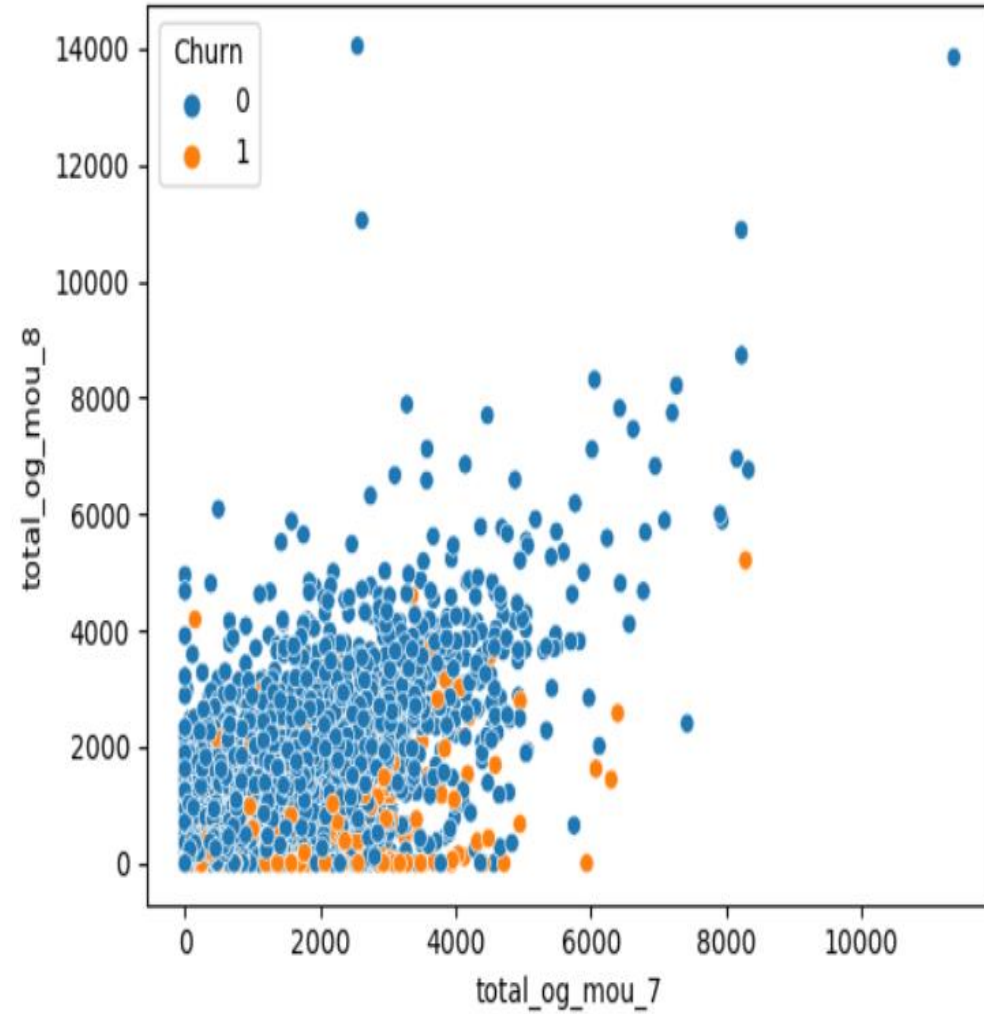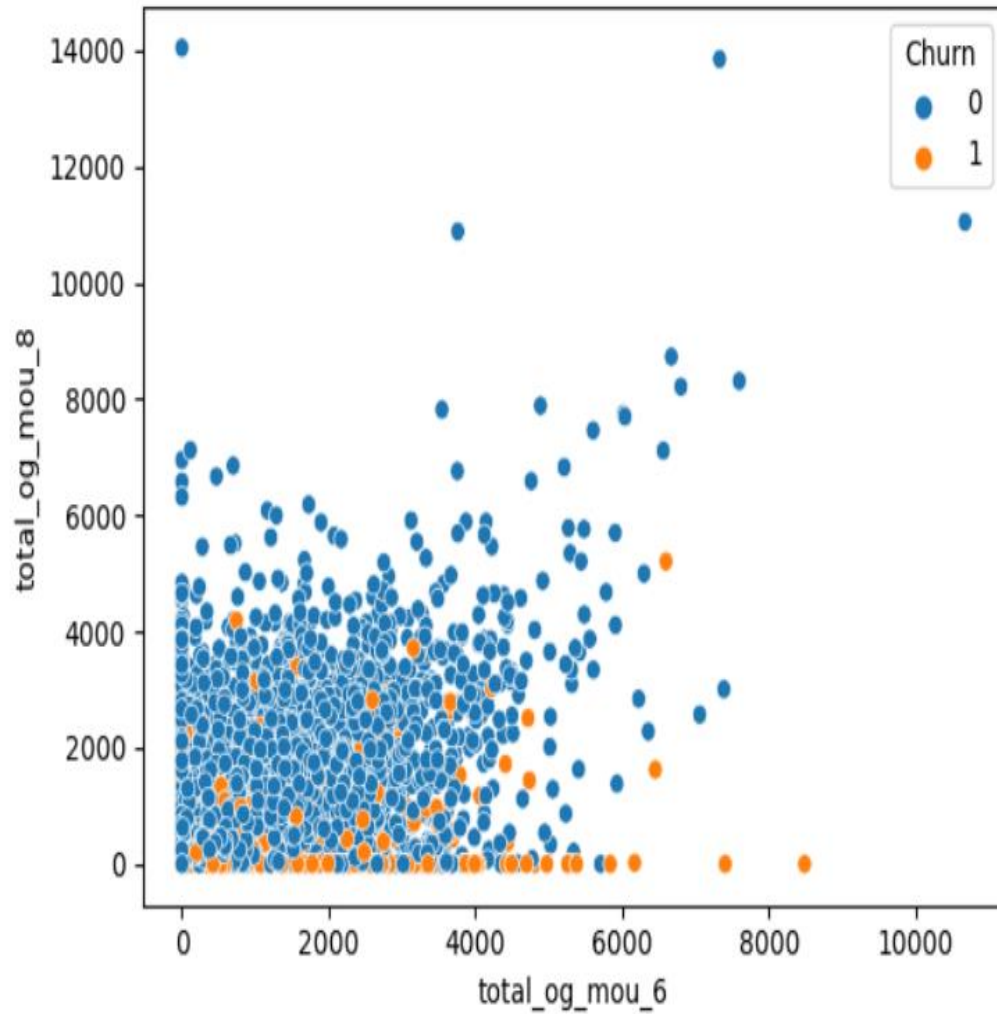
Analysis

➢ Customers with reduced usage of 2g data in the month of Aug compared to june and july are more likely to churn.
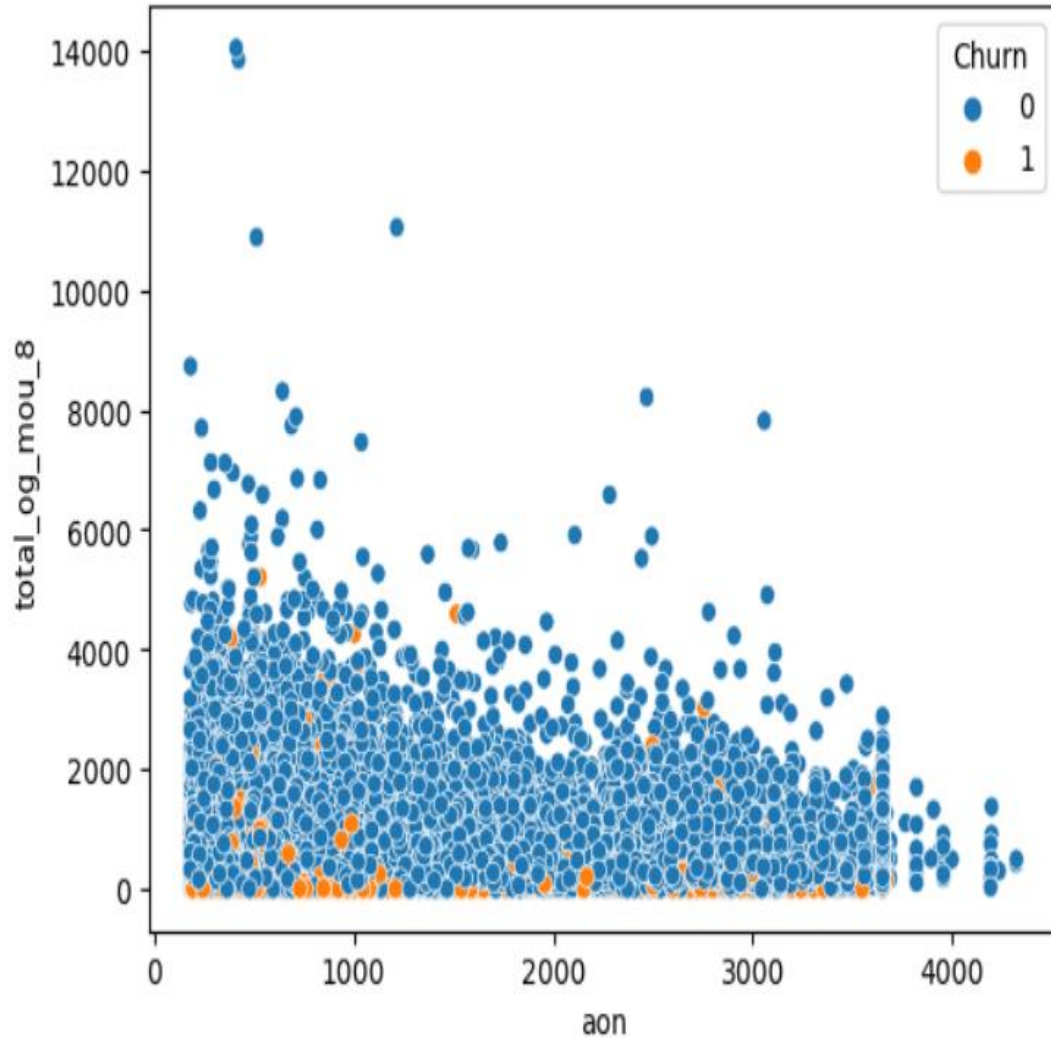
Analysis

Customers with fall in usage of 3g data in the Aug month compared to june and july usage is more likely to churn
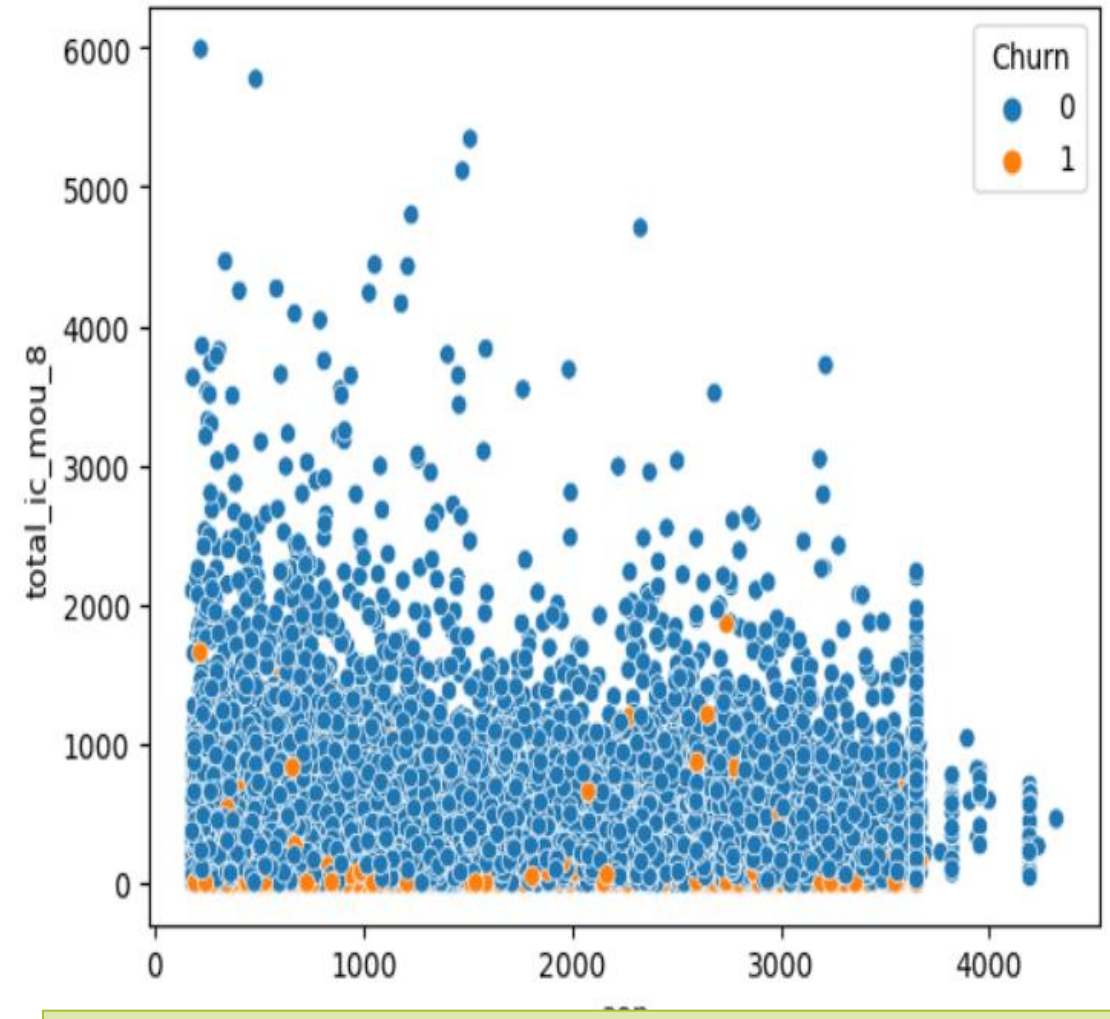
Correlation Analysis

Customers with lower out going minutes of user are more likely to churn than high og_mou

Correlation Analysis
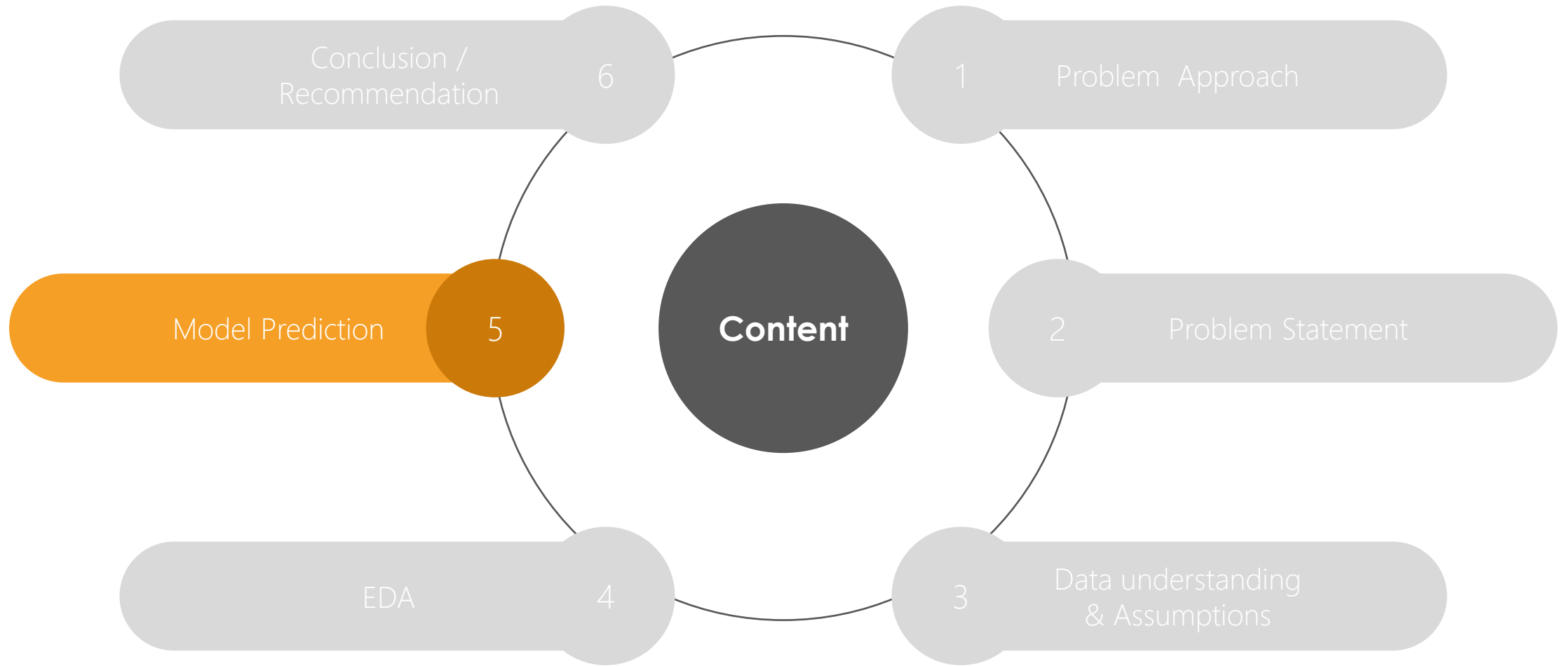

Here, customers with lesser total_og_mou_8 and aon are more likely to churn

Here, customers with lesser total_ic_mou_8 are more likely to churn irrespective of aon. And customer <2000 total_ic_mou_8 are more likely to churn

Content

Conclusion / Recommendation 6

1 Problem Approach

Model Prediction 5

Content

2 Problem Statement

EDA 4

3 Data understanding & Assumptions

Model Prediction

Train performance:
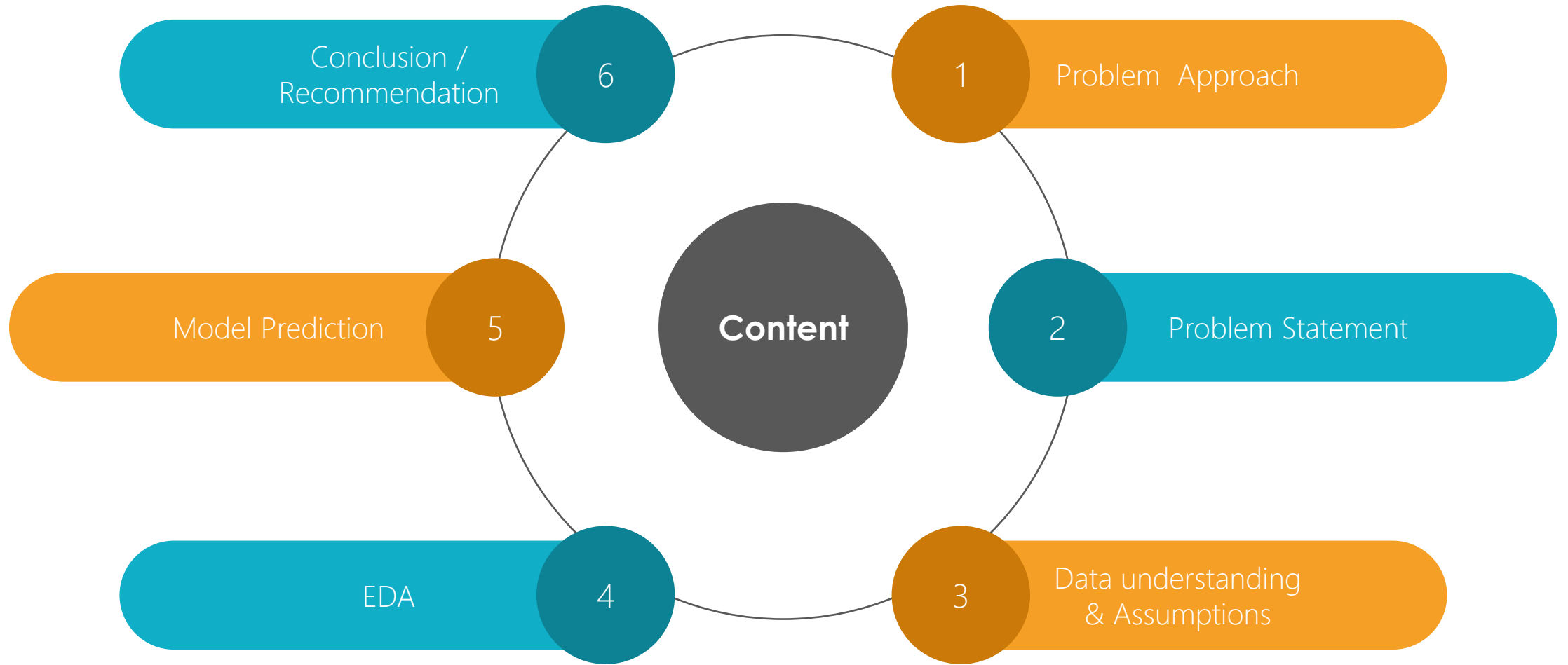- Accuracy: 63%
- Sensitivity: 91%
- Specificity: 60%

Test data:
- Accuracy: 18%
- Sensitivity: 97%
- Specificity: 10%
- Precision: 9%
- F1-Score : 16%

PCA + Logistic Regression Model: It has an ROC score of 0.87, sensitivity of 97%

Content

Conclusion / Recommendation

6

Problem Approach

1

Content

Model Prediction

5

Problem Statement

2

EDA

4

Data understanding & Assumptions

3

Conclusion / Recommendation

Following are the strongest indicators of Churn:

➢ Customers who churn are more likely to be users of 'monthly 2g package-0 in action period when all other factors are held constant. This is the strongest indicator of shurn.
➢ Customers who churn show lower number of recharges done in action period , When all other factors are held constant. This is the stronges indicator of churn
➢ Customers who churn show lower average monthly std incomingcalls compared to others.
➢ Customers who churn show lower average monthly isd outgoing compared to others.

Not Likely leads for X Education:

➢ Concentrate on users who users of less than 1.3 standard deviation of 'monthly 2g 'package. They are most likely to churn

➢ Concentrate on users who recharge less number of times.

➢ Models with high sensitivity are best for predicting churn . Use the PCA + Logistic Regression model to predict churn. It has an ROC score of 0.87, sensitivity of 97%

Thank You

24Slides