

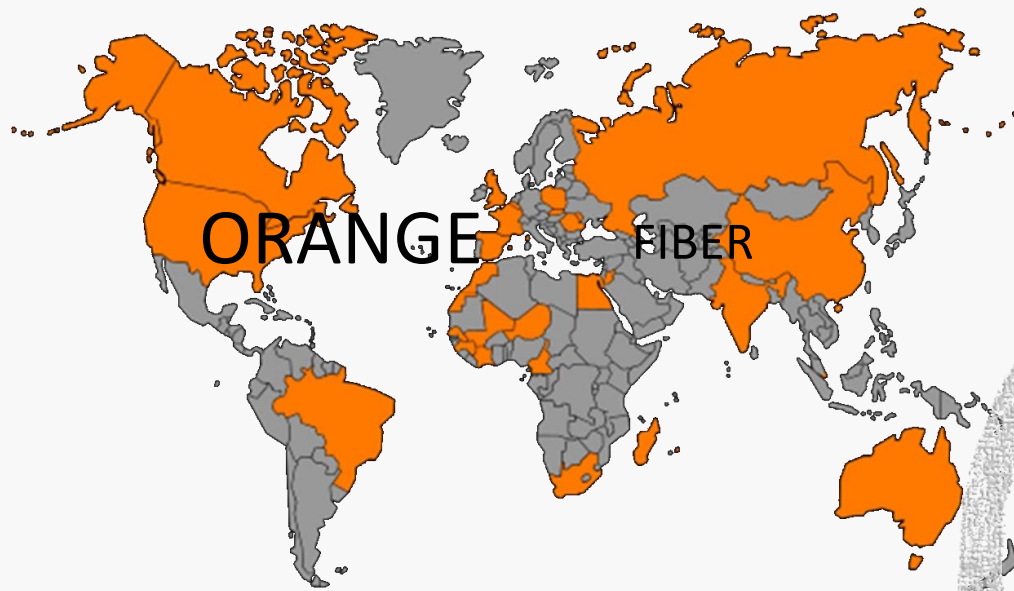


orangeTM



ORANGE FIBER

orangeTM



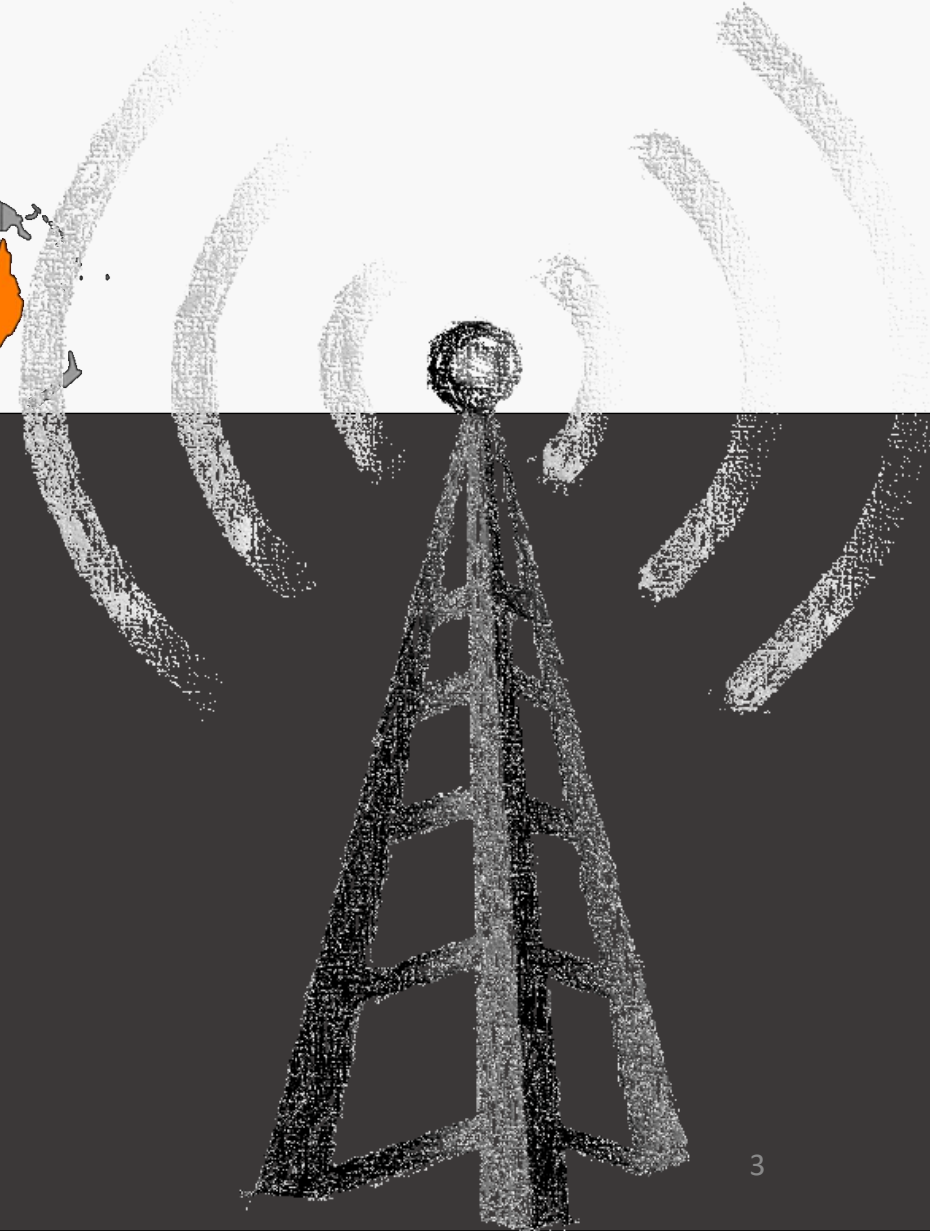
CODING ACADEMY BY ORANGE

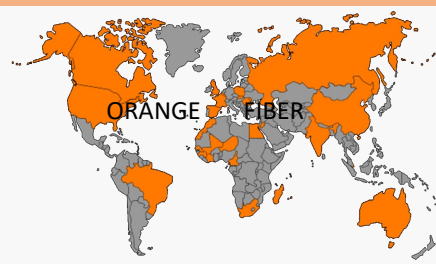
TASK 1

RAZAN SALMAN

NEDAL ALTITI

LAITH RASHEED





O U T L I N E S

PROBLEM DEFINITION

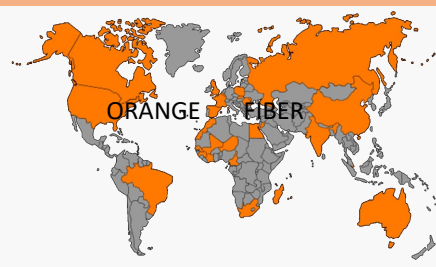
PROBLEM STRUCTURE

FEATURES

HYPOTHESIS

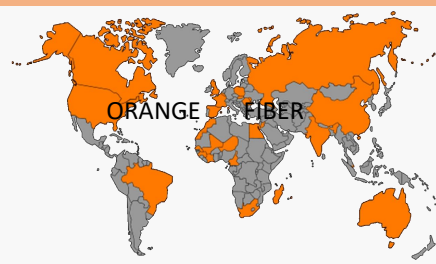
RECOMMENDATION





PROBLEM DEFINITION



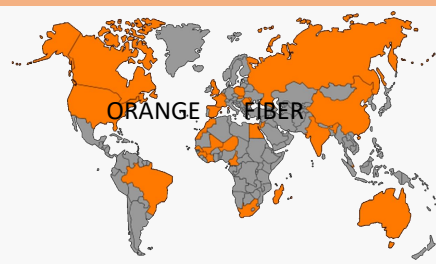


PROBLEM DEFINITION

INITIAL SITUATION

Dataset that includes details about fiber customers, demographics, subscription technical information, and contract type





PROBLEM DEFINITION

PROJECT SCOPE/ USE CASE

1. OUR OBJECTIVE
2. BUISSNIS OBJECTIVE

OUR OBJECTIVE

identify customers who are at risk of churning and check their loyalty and satisfaction

BUISSNES OBJECTIVE

REDUCE THE CHURN RATE AND
IMPROVING CUSTOMER RETENTION



What is the churn rate?

“Churn rate is a metric used to measure the rate at which customers or subscribers discontinue using a company's products or services over a given period of time”



What is the churn rate?

“Churn rate is a metric used to measure the rate at which customers or subscribers discontinue using a company's products or services over a given period of time”

REASONS

- 1- CUSTOMER DIES OR GOES OUT OF BUSINESS.
- 2- JUST DO NOT ACHIEVE THEIR DESIRED OUTCOME.



What is the churn rate?

“Churn rate is a metric used to measure the rate at which customers or subscribers discontinue using a company's products or services over a given period of time”

CHURN RATE CALCULATION

CHURN RATE CAPTURES THE NUMBER OF PEOPLE WE RETAIN AT THE END OF A TIME PERIOD.

CHURN RATE =

CUSTOMERS AT START OF USAGE INTERVAL – CUSTOMERS AT END OF USAGE INTERVAL

CUSTOMERS AT START OF USAGE INTERVAL

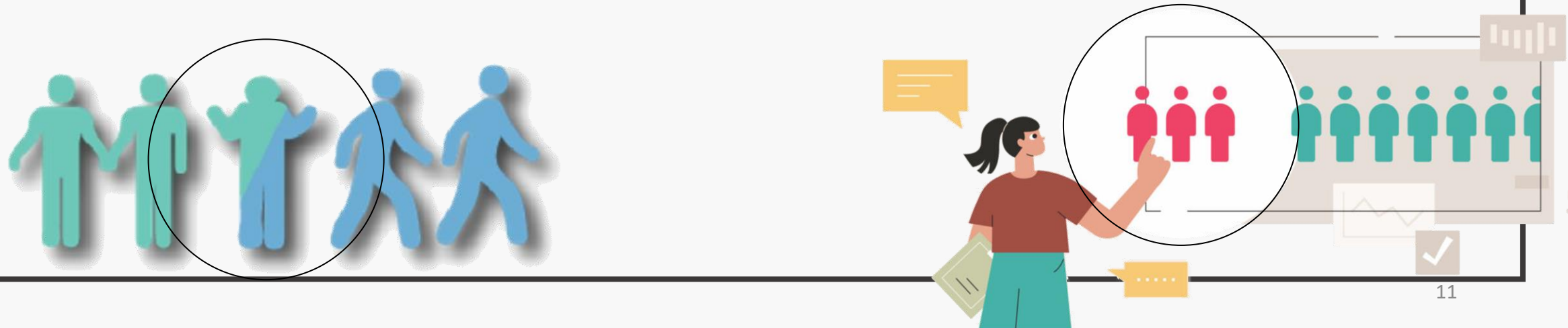


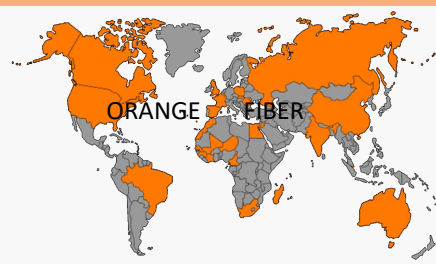
What is the churn rate?

“Churn rate is a metric used to measure the rate at which customers or subscribers discontinue using a company's products or services over a given period of time”

SOLUTIONS

identify customers who are at risk of churning and take steps to retain them.



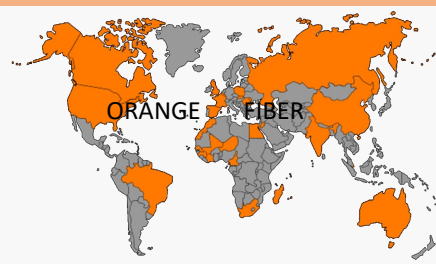


PROBLEM DEFINITION

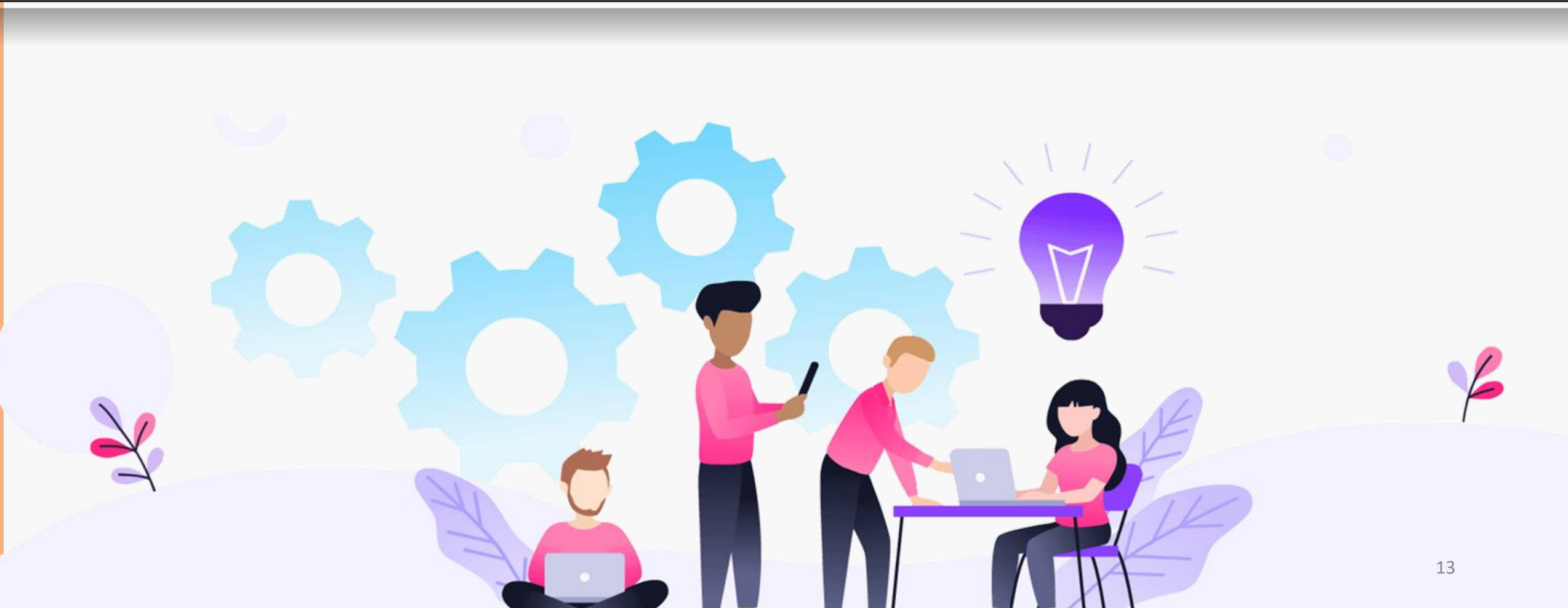
SUCCESS CRITERIA

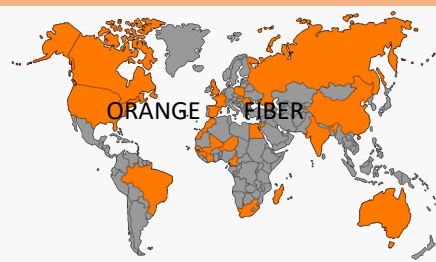
in general, an **annual** churn rate of 5% is seen as a reasonable benchmark





PROBLEM STRUCTURE





PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

FRAME THE PROBLEM

GET THE DATA

EDA

MISINTERPRETATION
SOME
FEATURES

MISSING VALUES

OUTLIERS AND
VARIATION

IMBALANCE
DATA

CATEGORICAL
DATA

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

FRAME THE PROBLEM

GET THE DATA

EDA

MISINTERPRETATION SOME
FEATURES

1. SEARCH
2. ASK EXPERTS
3. CASE STUDIES

MISSING VALUES

1. Structured
2. MAR
3. MNAR

OUTLIERS AND
VARIATION

1. Transformation
2. Modeling
- 3- Normalization

IMBALANCE DATA

1. Non representative
2. skewness

CATEGORICAL
DATA

1. Encoding to convert them to numerical values

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

MISSING VALUES

| | columns | number of missing values |
|----|--------------------------------|--------------------------|
| 0 | OF_PREV_SPEED | 74714 |
| 1 | GB_TOTAL_CONSUMPTION_Month3 | 4242 |
| 2 | Disconnection_TOTAL_MAX_day | 625 |
| 3 | Disconnection_TOTAL_MIN_day | 625 |
| 4 | Disconnection_TOTAL_SUM_Month | 625 |
| 5 | Disconnection_TOTAL_MEAN_Month | 625 |
| 6 | GB_TOTAL_CONSUMPTION_Month2 | 367 |
| 7 | GOVERNORATE | 323 |
| 8 | GB_TOTAL_CONSUMPTION_Month1 | 155 |
| 9 | CUSTOMER_GENDER | 54 |
| 10 | LAST_LINK_QUALITY | 28 |
| 11 | LAST_LINK_STATUS | 28 |
| 12 | LAST_POWER_VALIDATION | 28 |
| 13 | LAST_LINK_PRIORITY | 28 |

We have 82,467 missing values in total

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

MISSING VALUES

Methods to handle the missing values

Numerical Variable

- 1- delete the missing values **x**
- 2- using the mean **x**
- 3- using the median, better for outlier and skewness
- 4- iterative imputation

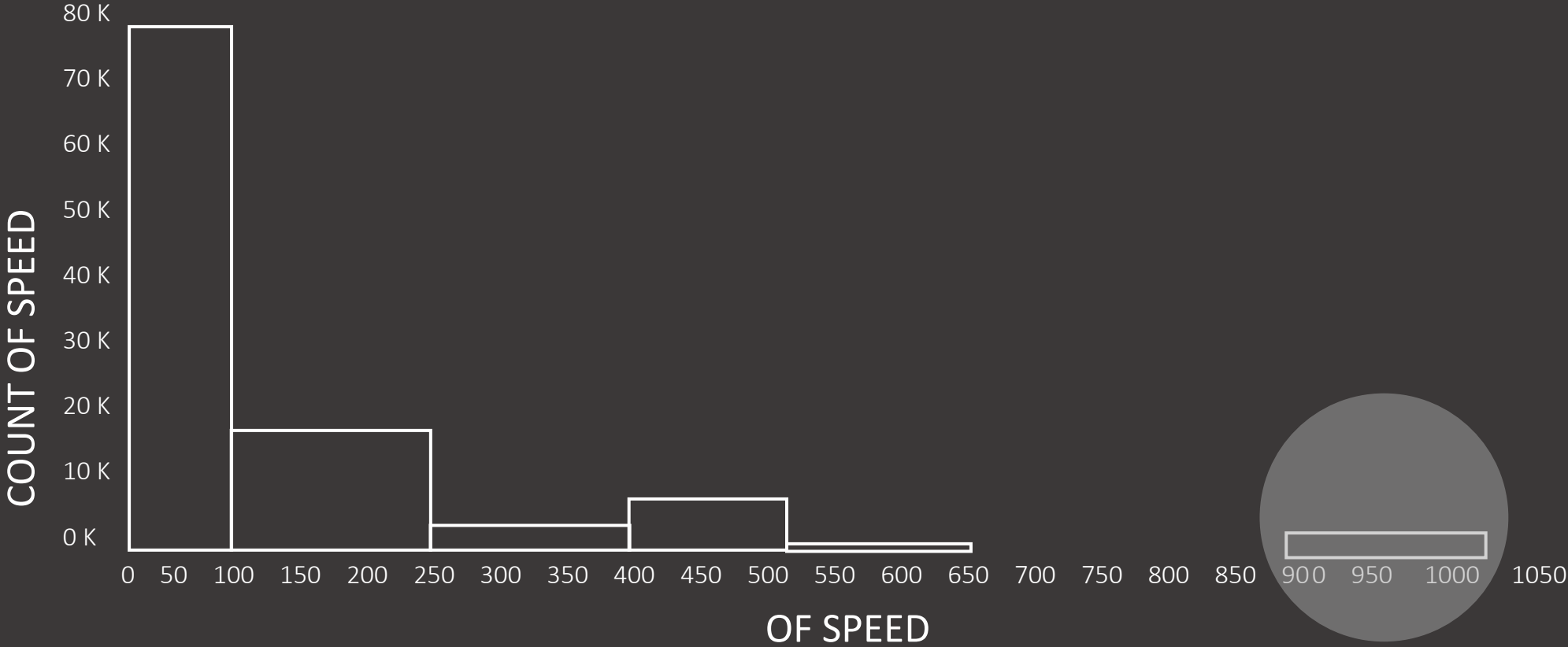
Categorical Variable

- 1- delete the missing values
- 2- impute the missing values with most frequent variable
- 3- iterative imputation

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

Outliers

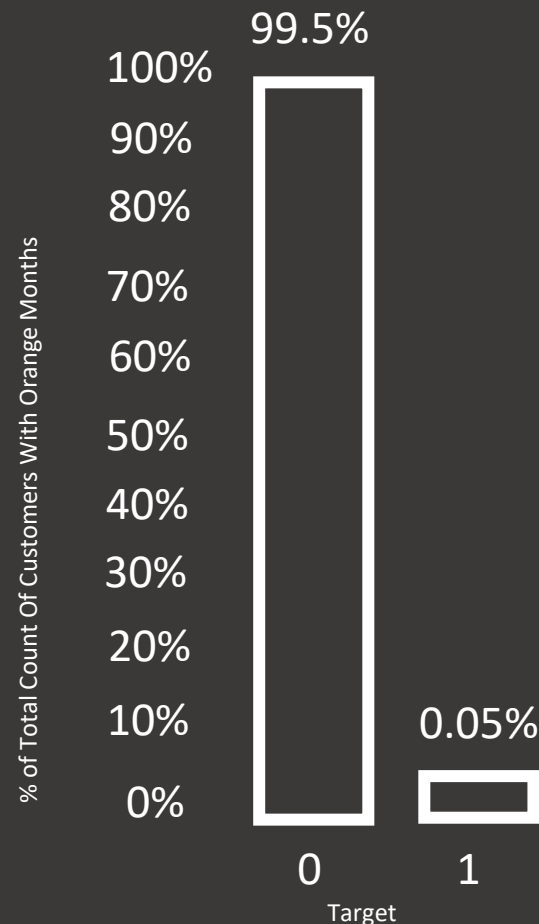


PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

IMBALANCE DATA

CHURNED vs UNCHARNED CUSTOMERS



Target 0 : people who Stay with Orange

Target 1 : People who left Orange

Methods to handle the imbalance data

1- Combine oversampling and undersampling and keep in mind to take the sample that is representative to the population

2- Oversampling synthetic Data

3- Use algorithms with weighted average

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

Categorical data

Methods to handle categorical data

- 1- OneHotEncoding, e.g “Gender, Migration Flag”.
- 2- LabelEncoding, e.g “age group”.

OneHotEncoding

| TARGET | F | M |
|--------|---|---|
| 0 | 0 | 1 |
| 1 | 1 | 0 |

LabelEncoding

| AGE GROUP | CATEGORICAL | TARGET |
|-----------|-------------|--------|
| CHILD | 1 | 1 |
| TENEEAGER | 2 | 1 |
| ADULT | 3 | 0 |
| SENIOR | 4 | 0 |

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

FRAME THE PROBLEM

GET THE DATA

EDA

MISINTERPRETATION
SOME
FEATURES

1. SEARCH
2. ASK EXPERTS
3. CASE STUDIES

MISSING VALUES

1. Structured
2. MAR
3. MNAR

OUTLIERS AND
VARIATION

1. Transformation
2. Modeling
- 3- Normalization

IMBALANCE DATA

1. Non representative
2. skewness

CATEGORICAL
DATA

1. Encoding to convert them to numerical values

PROBLEM STRUCTURE

“WHAT COULD BE THE KEY ELEMENT OF THE PROBLEM?”

FRAME THE PROBLEM

GET THE DATA

EDA

MISINTERPRETAT
ION SOME
FEATURES

MISSING VALUES

OUTLIERS AND
VARIATION

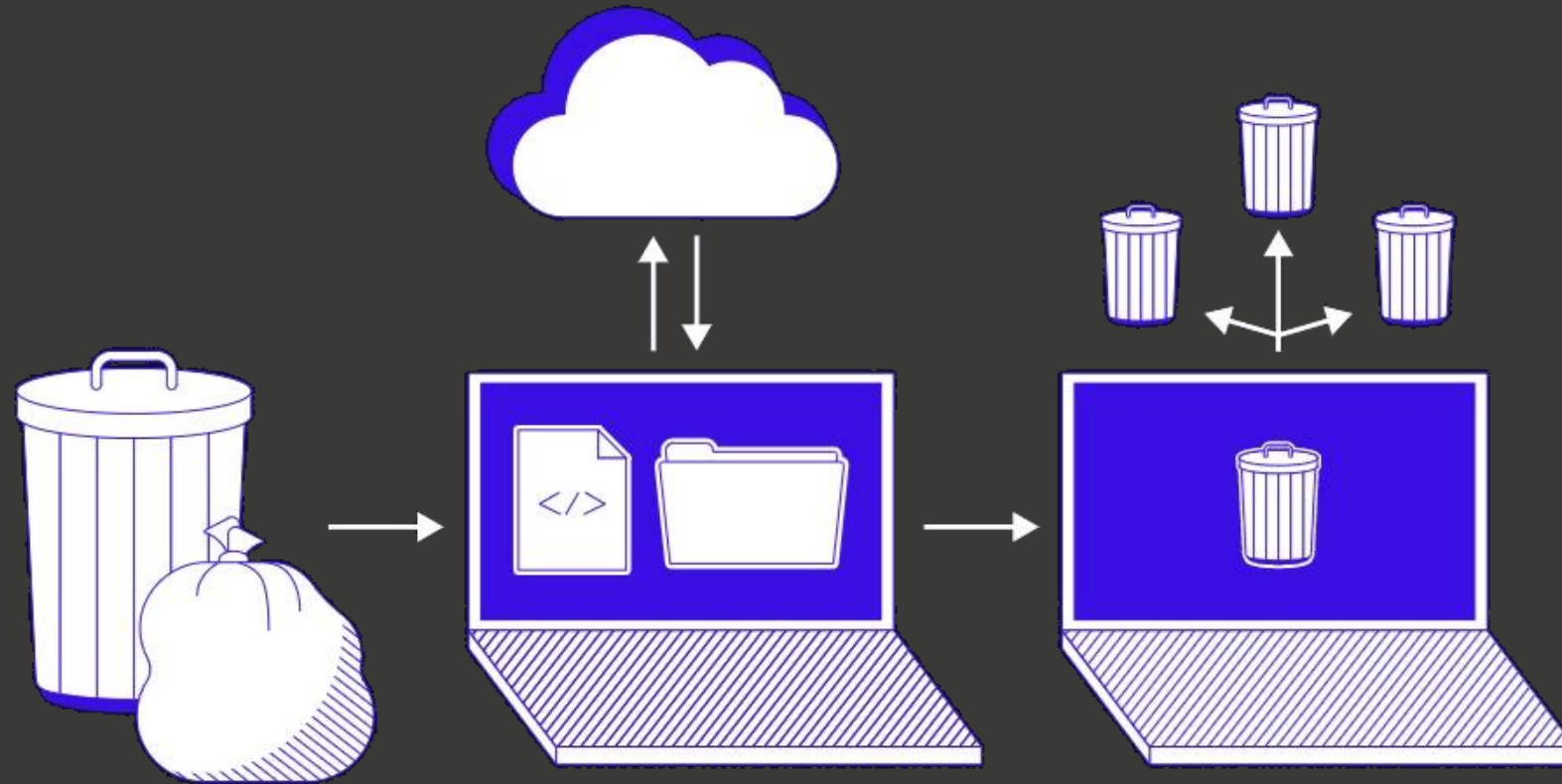
IMBALANCE DATA

CATEGOR-ICAL
DATA

DATA GAB

DATA GAB

GARBAGE IN, GARBAGE OUT



Bad Data

Great Model

Bad Prediction

FRAME THE PROBLEM

GET THE DATA

EDA

MISINTERPRETAT
ION SOME
FEATURES

MISSING VALUES

OUTLIERS AND
VARIATION

IMBALANCE DATA

CATEGOR-ICAL
DATA

DATA GAB

FRAME THE PROBLEM

GET THE DATA

EDA

DATA GAB

VISUALIZATION

EXPLORE DIFFERENT MODELS

FINE TUNNING THE MODEL

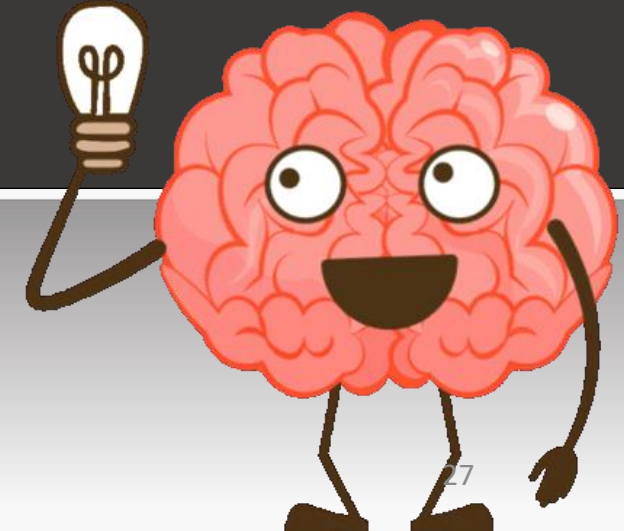
EVALUATION METRICS

COMBINING PRECISION AND RECALL, F1 SCORE, AREA UNDER CURVE ROC

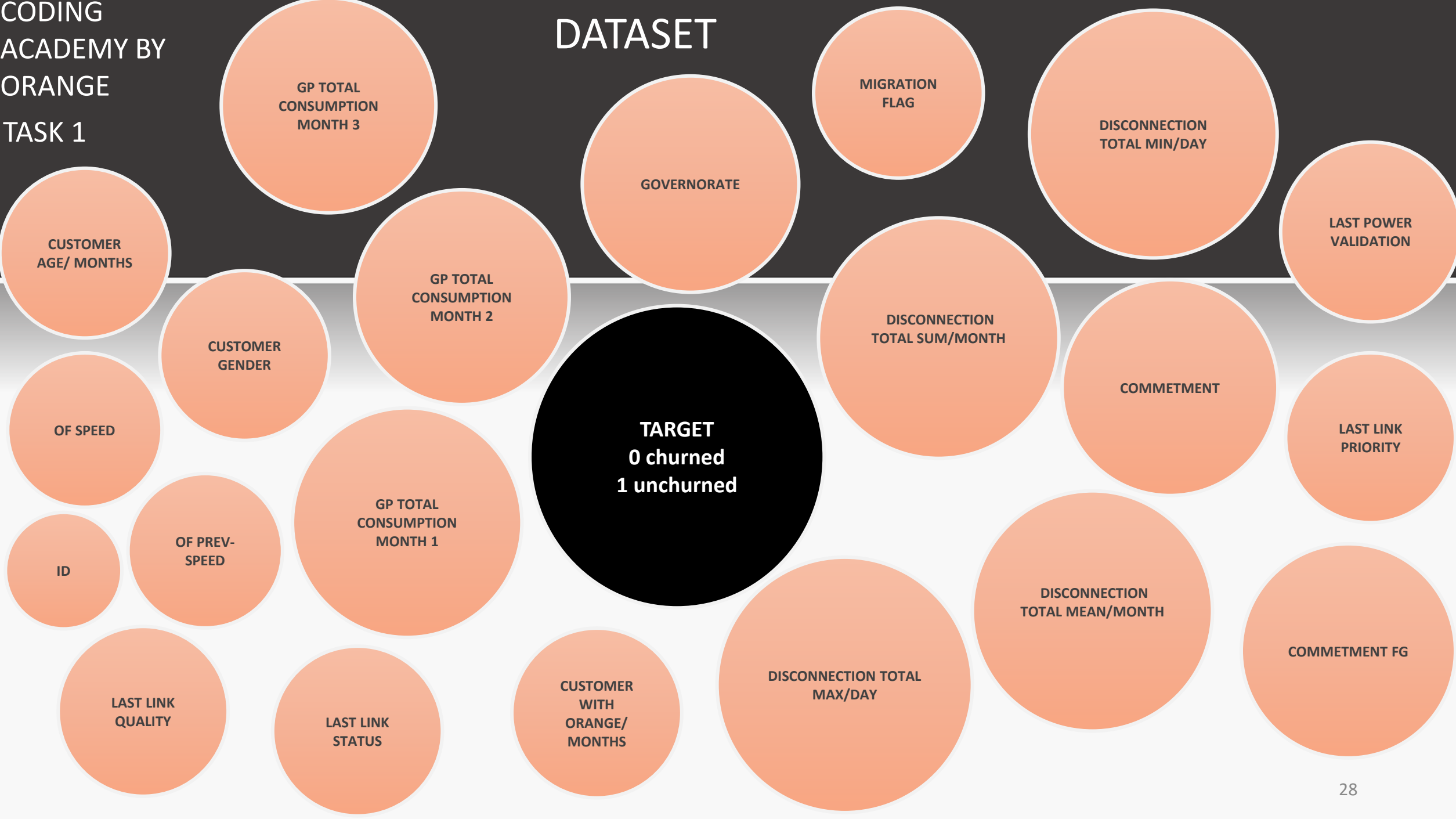
PRESENT SOLUTION

An isometric illustration within a light blue oval frame depicting a cloud deployment process. It features several server racks, a large laptop, and three stylized human figures in purple shirts and yellow pants. One figure stands on a server rack, another stands next to a server, and a third is positioned near the laptop. A large blue cloud with a white arrow points downwards towards the servers. Various colorful geometric shapes like cubes, spheres, and a donut chart are scattered around the central elements. The text "DEPLOY THE MODEL" is overlaid in the center in a large, bold, black font.

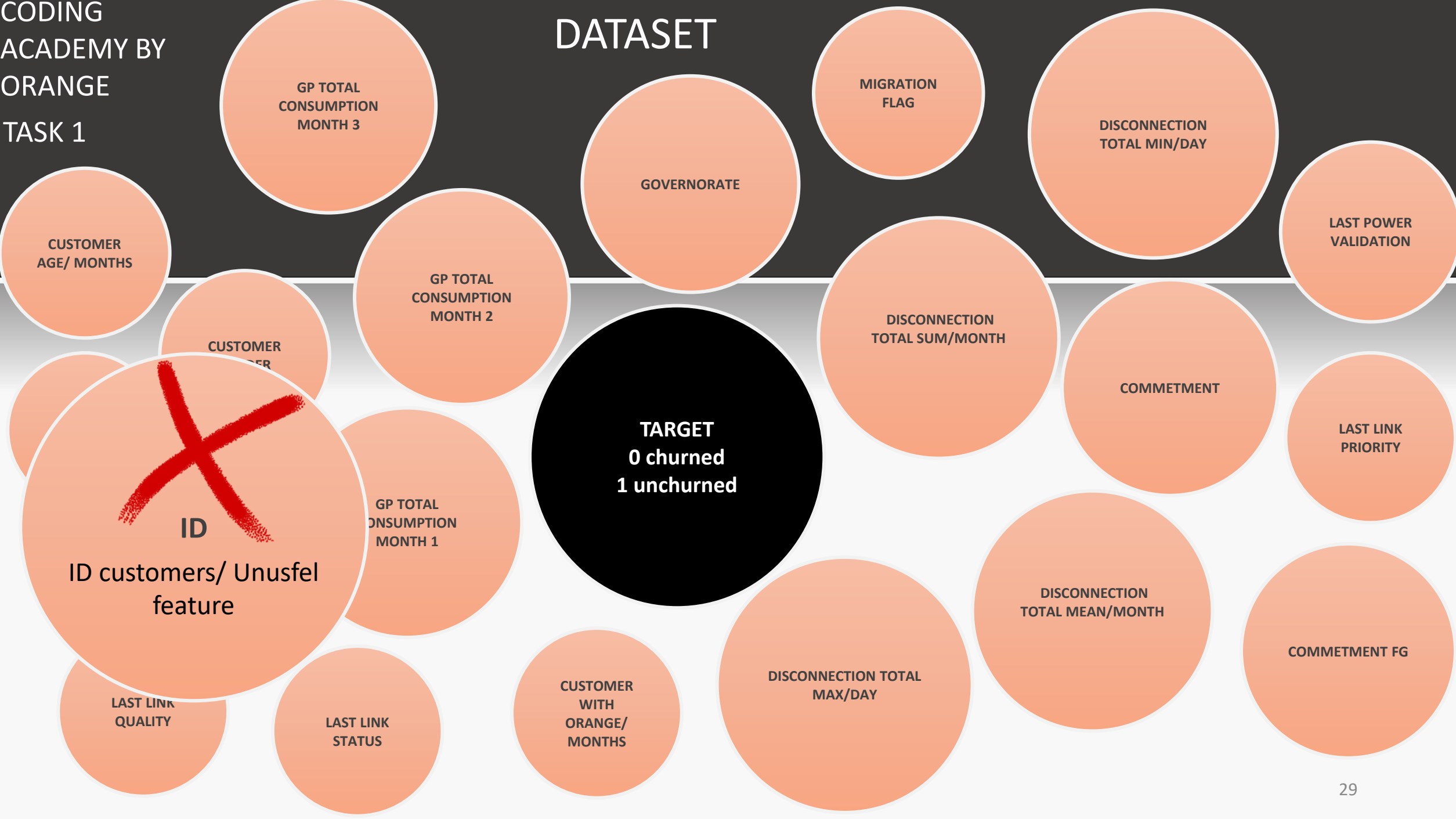
DEPLOY THE MODEL

[illegible]

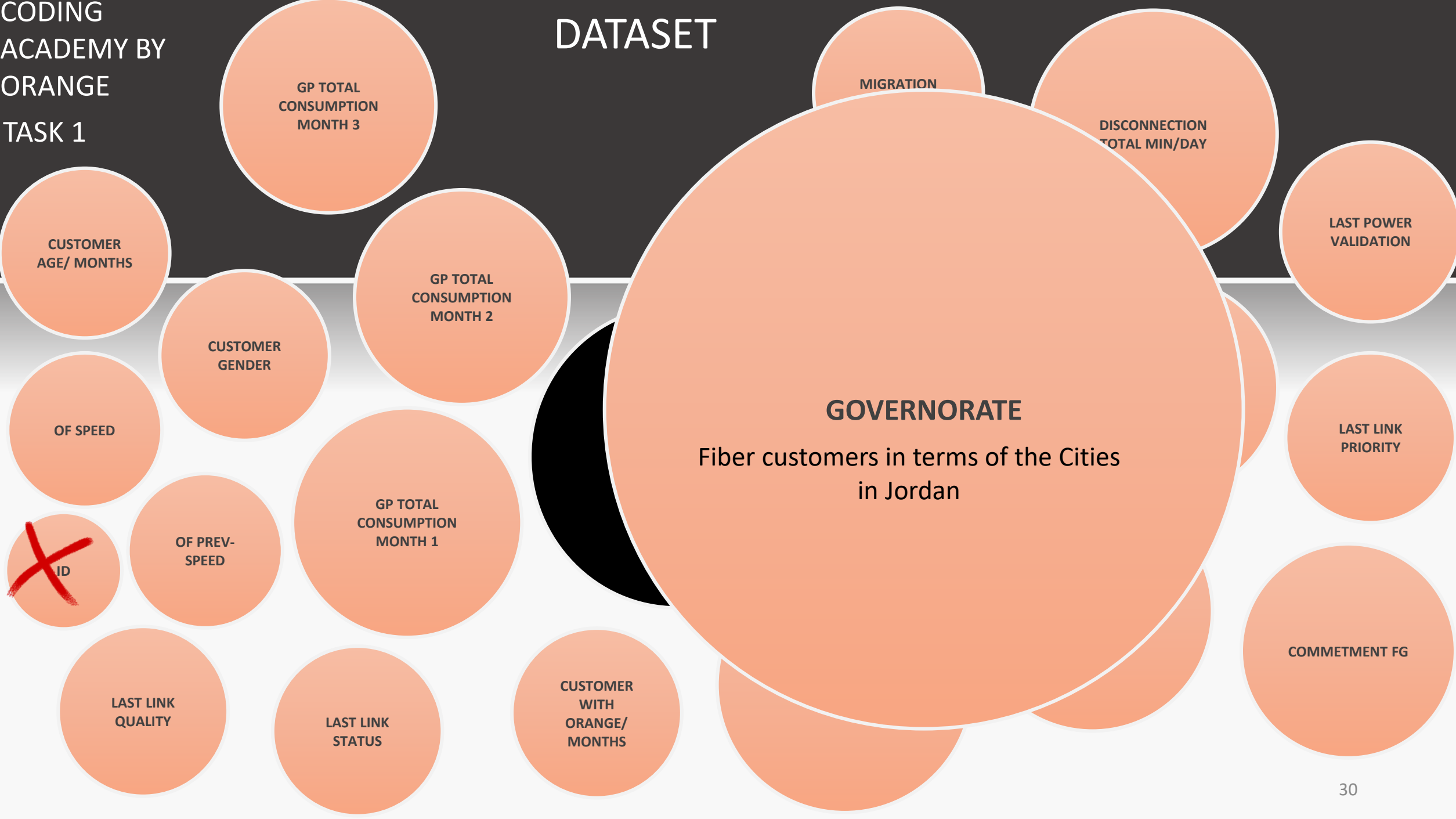
DATASET



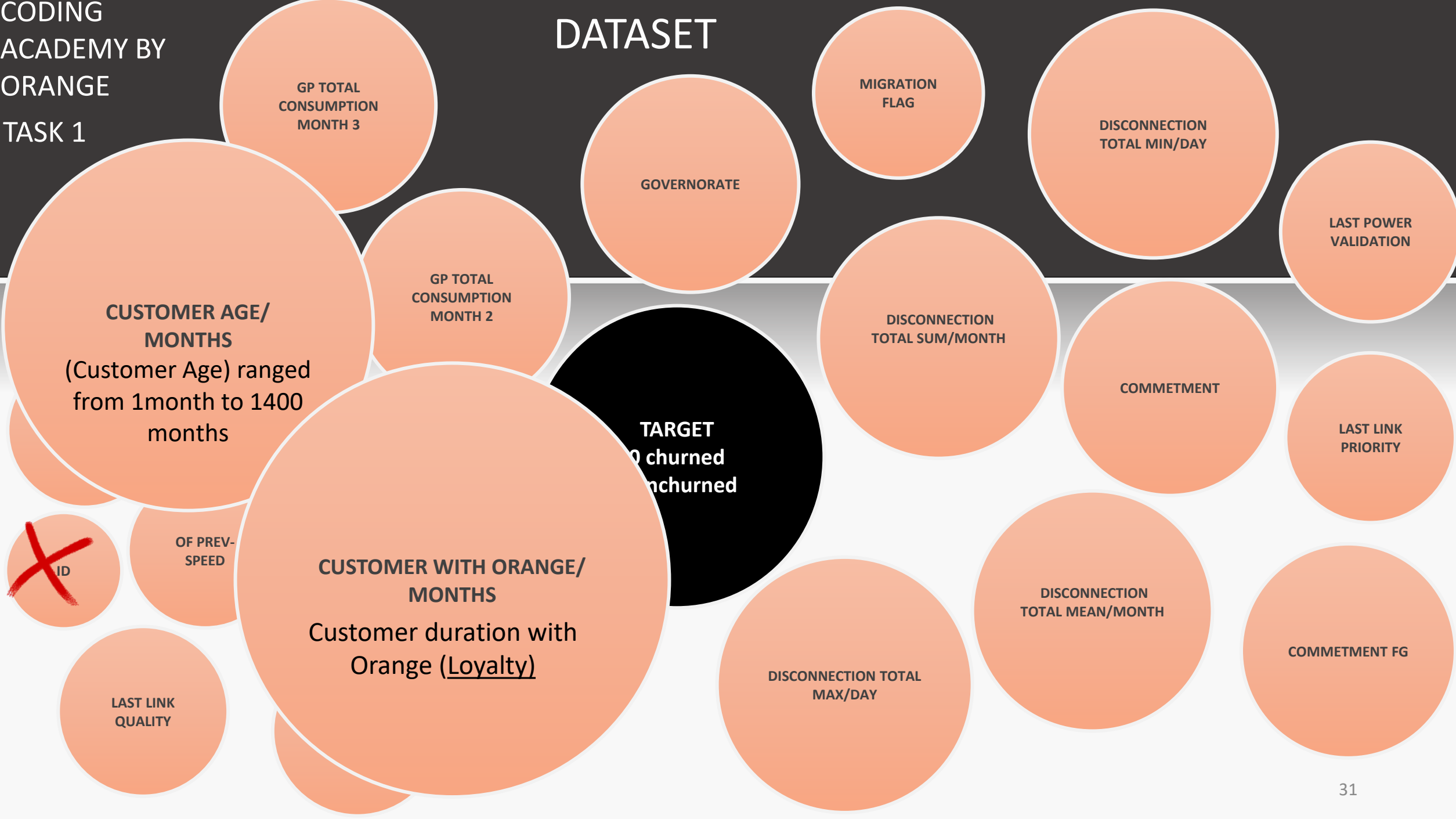
DATASET



DATASET



DATASET



DATASET

**CUSTOMER AGE/
MONTHS**
(Customer Age) ranged
from 1month to 1400
months

GP TOTAL
CONSUMPTION
MONTH 3

OF PREV-
SPEED

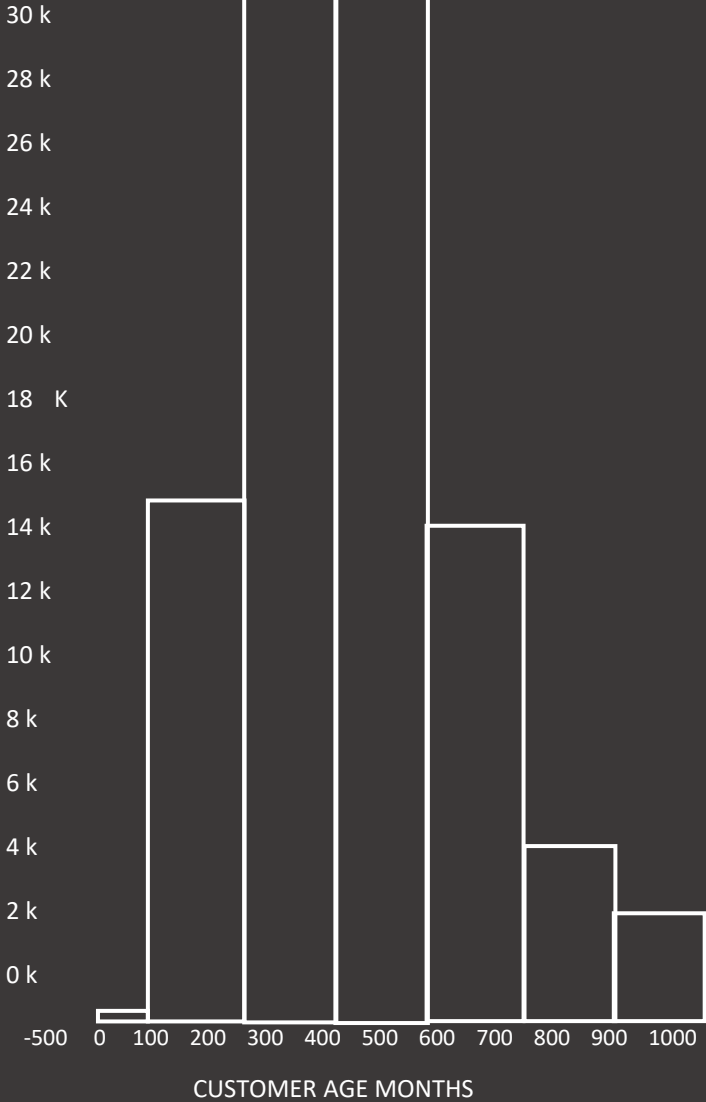
GP TO
CONSUM
MON

~~ID~~

LAST LINK
QUALITY

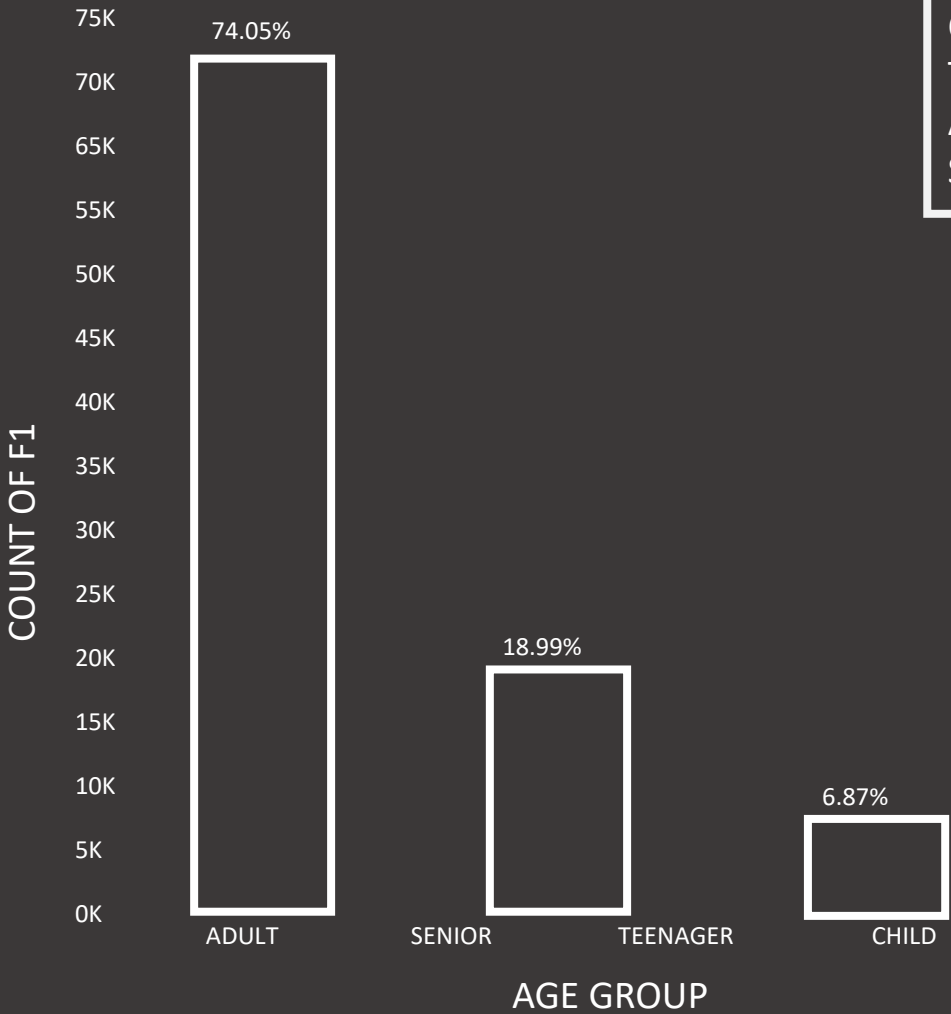
LAST LINK
STATUS

COUNT OF CUSTOMERS AGE MONTHS



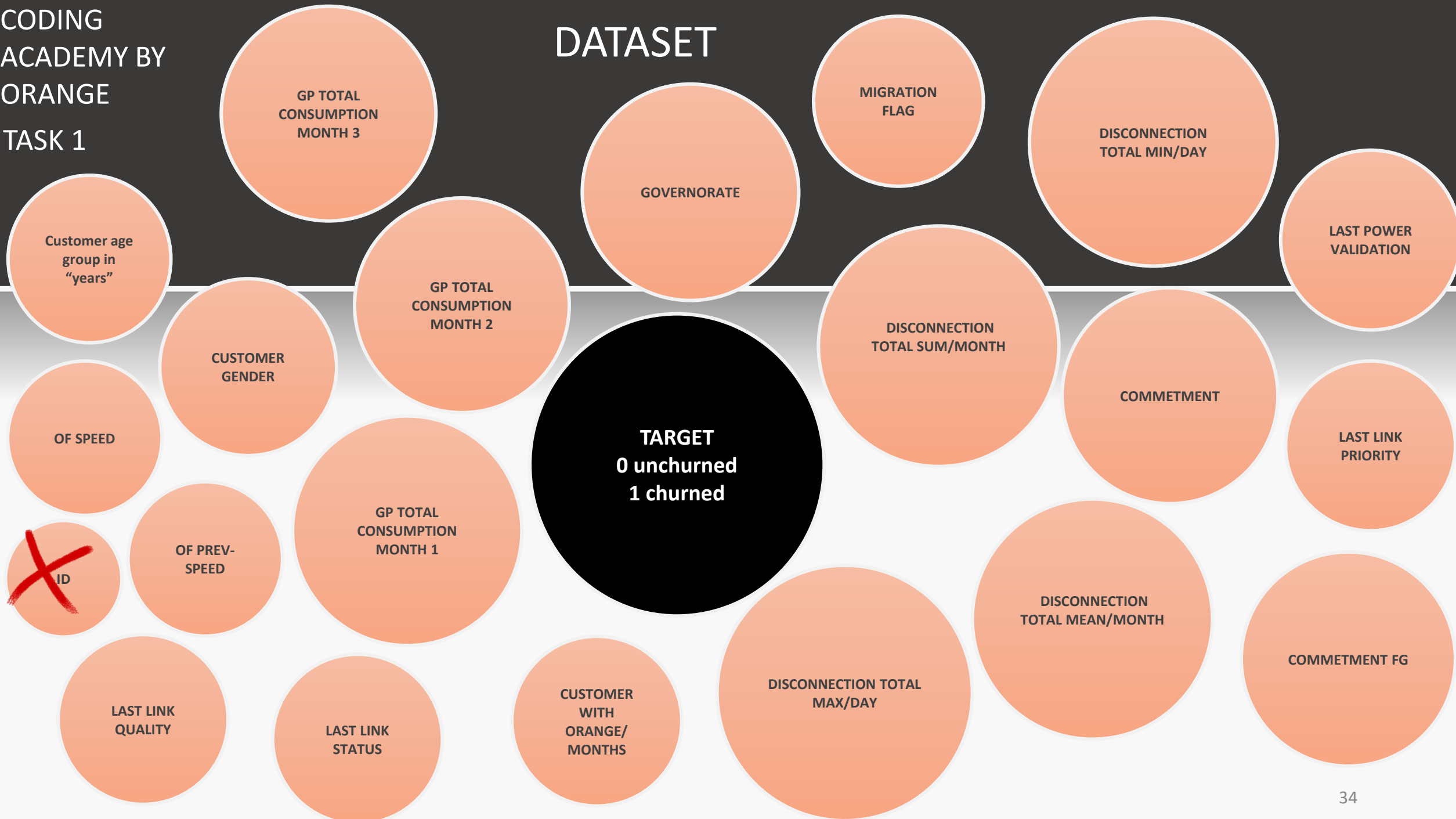
DATASET

Convert it to years and Categorize “ Customer Age “ to Age group.

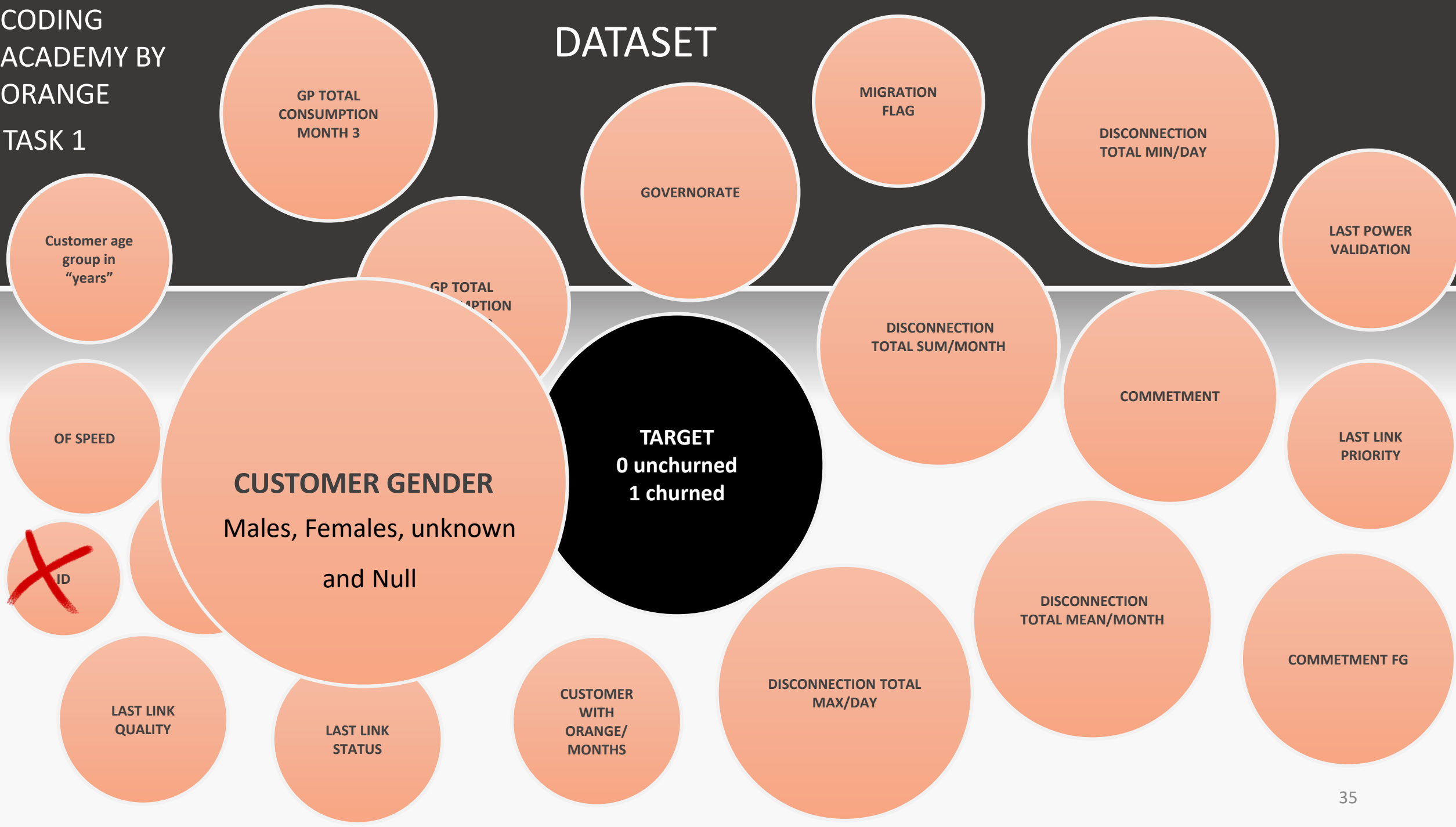


| | |
|----------|----------|
| Child | Under 18 |
| Teenager | 19-25 |
| Adult | 25-60 |
| Seniors | Above 60 |

DATASET

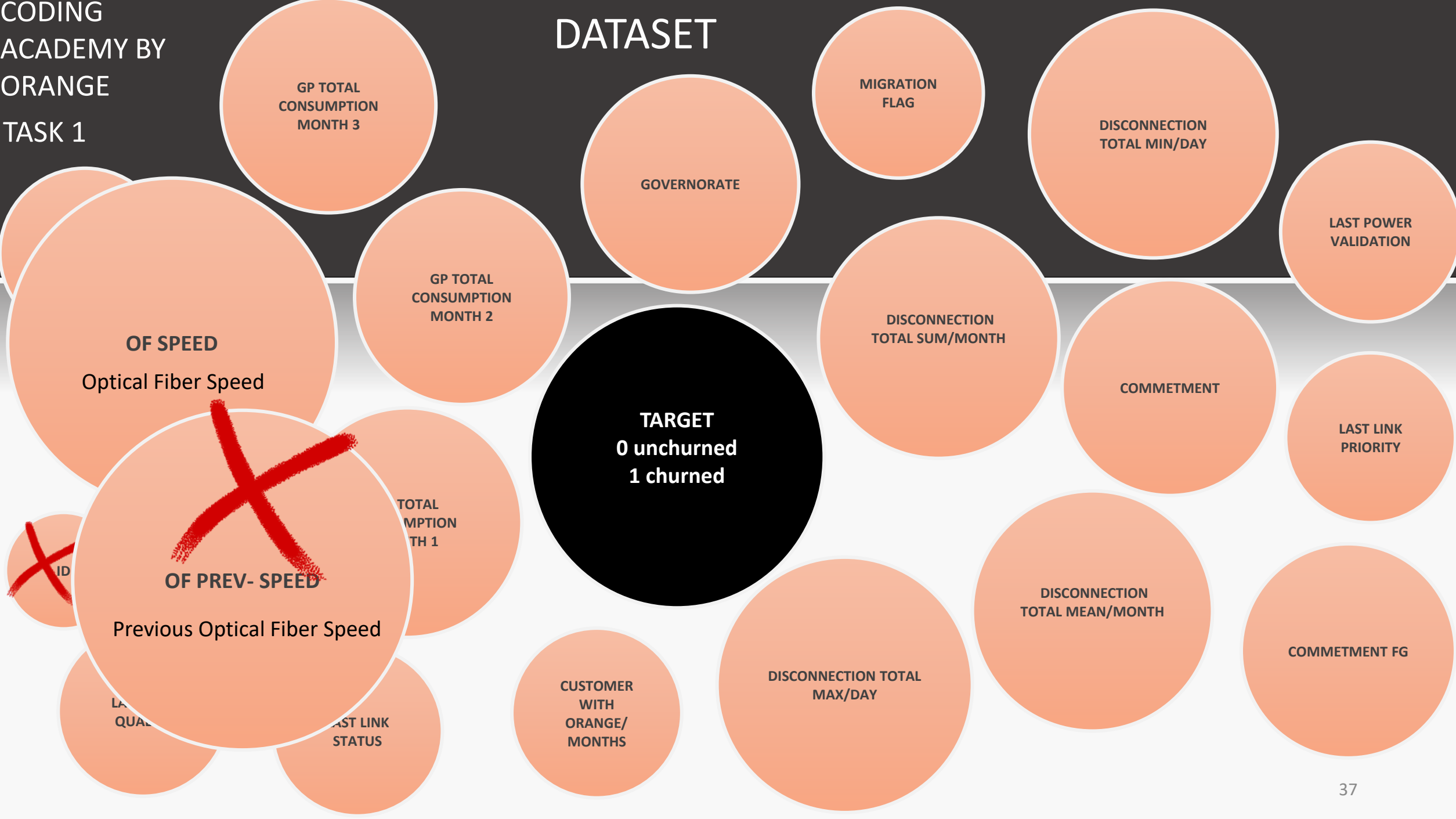


DATASET

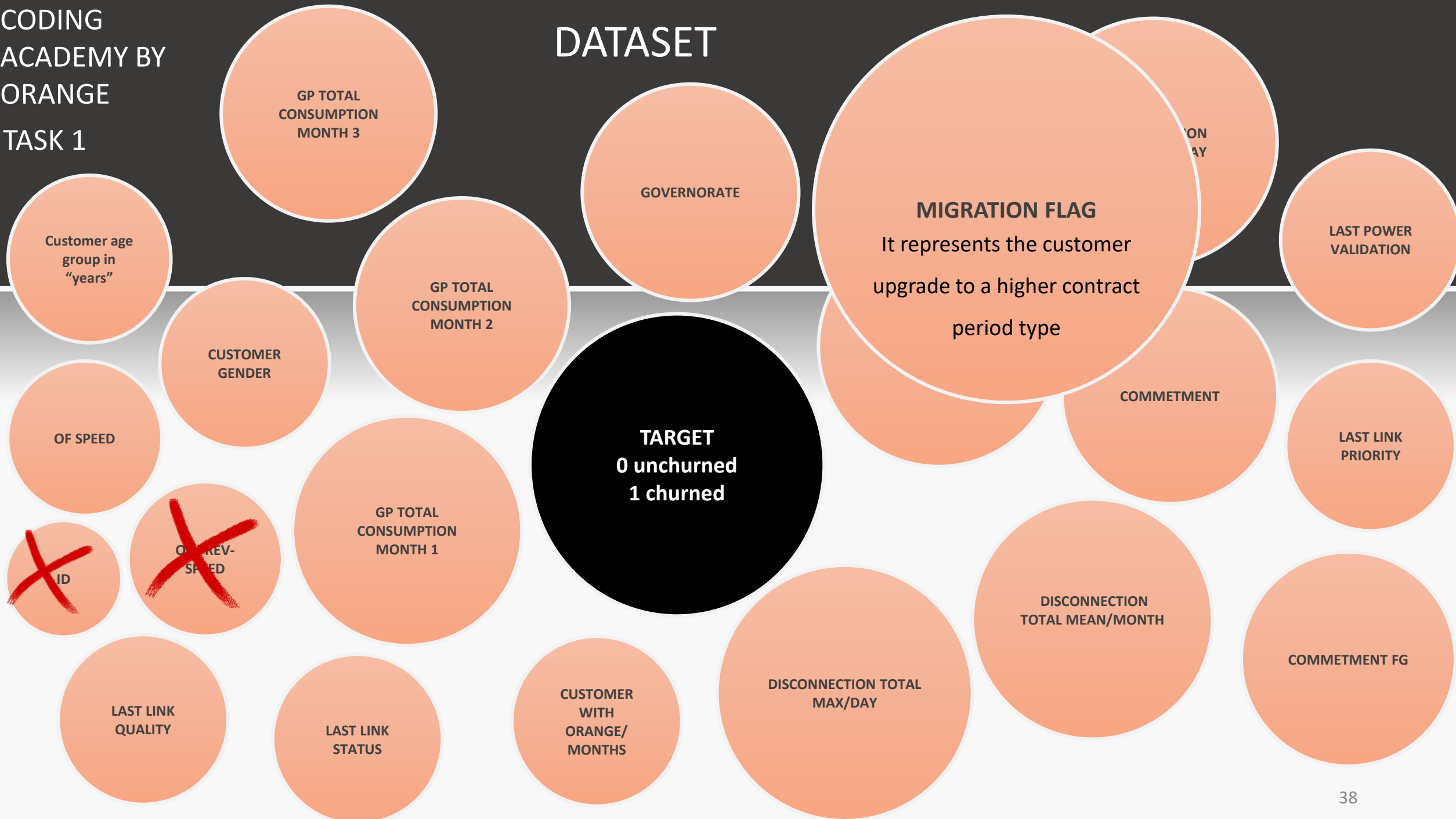




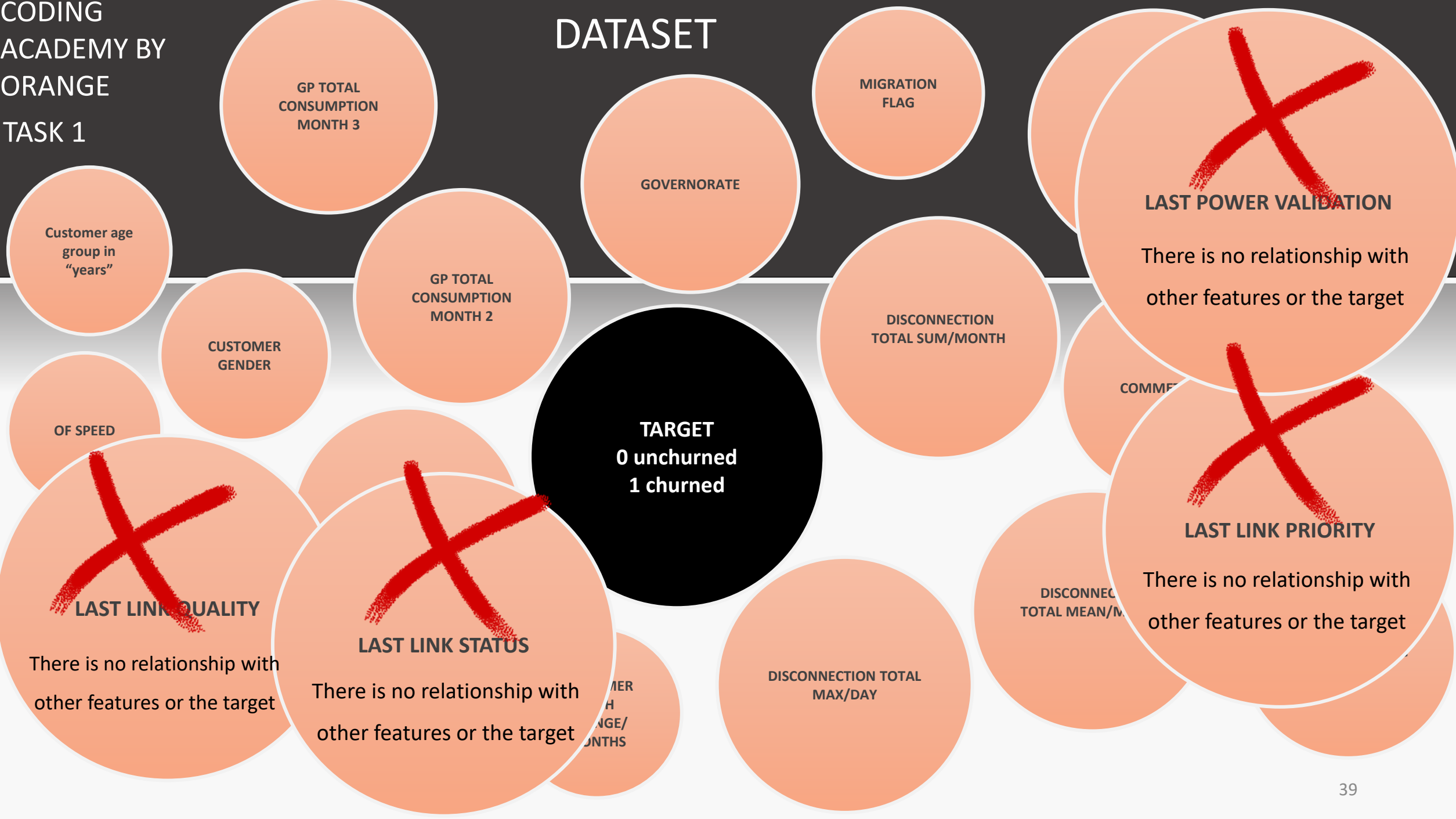
DATASET



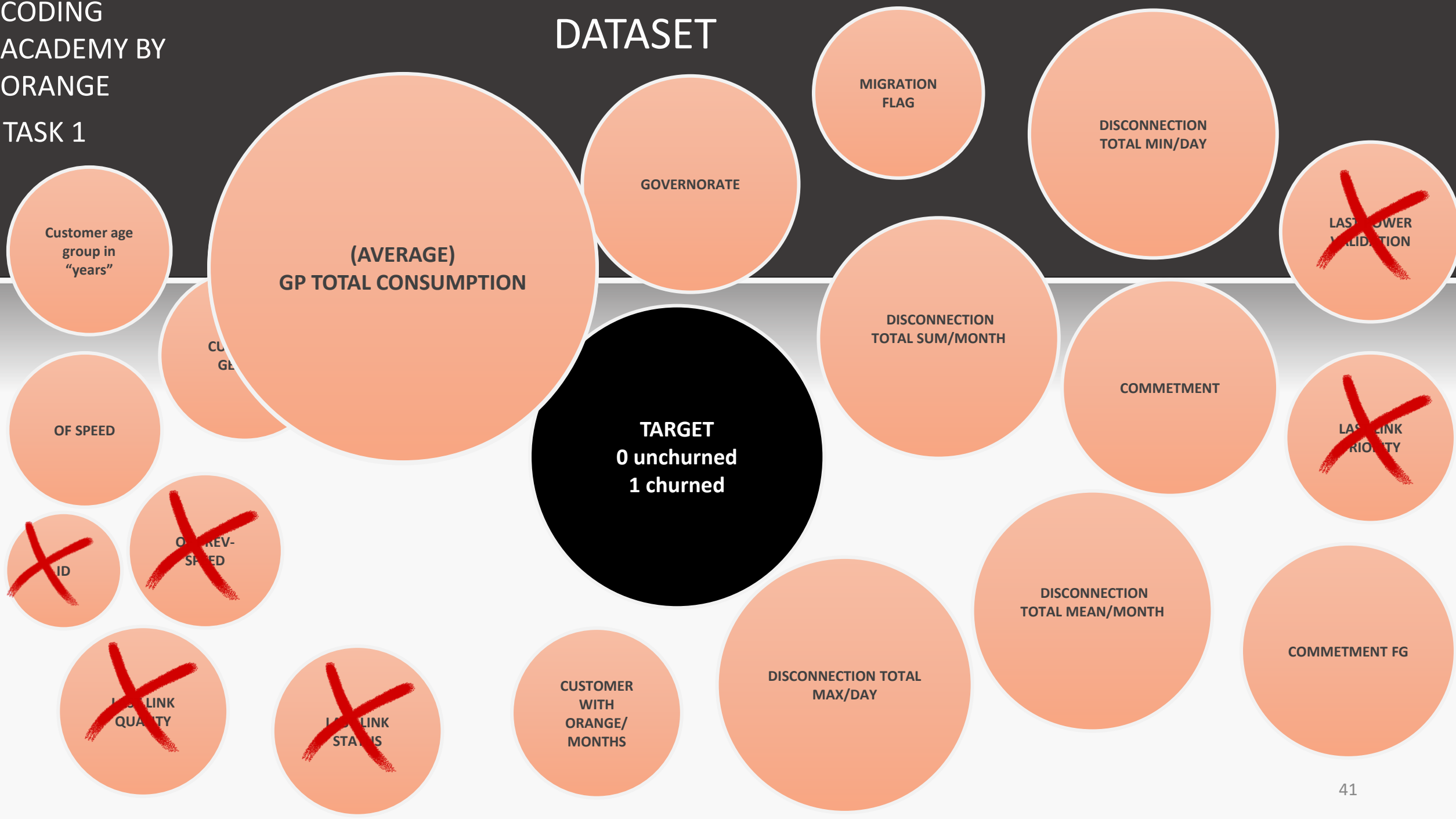
DATASET

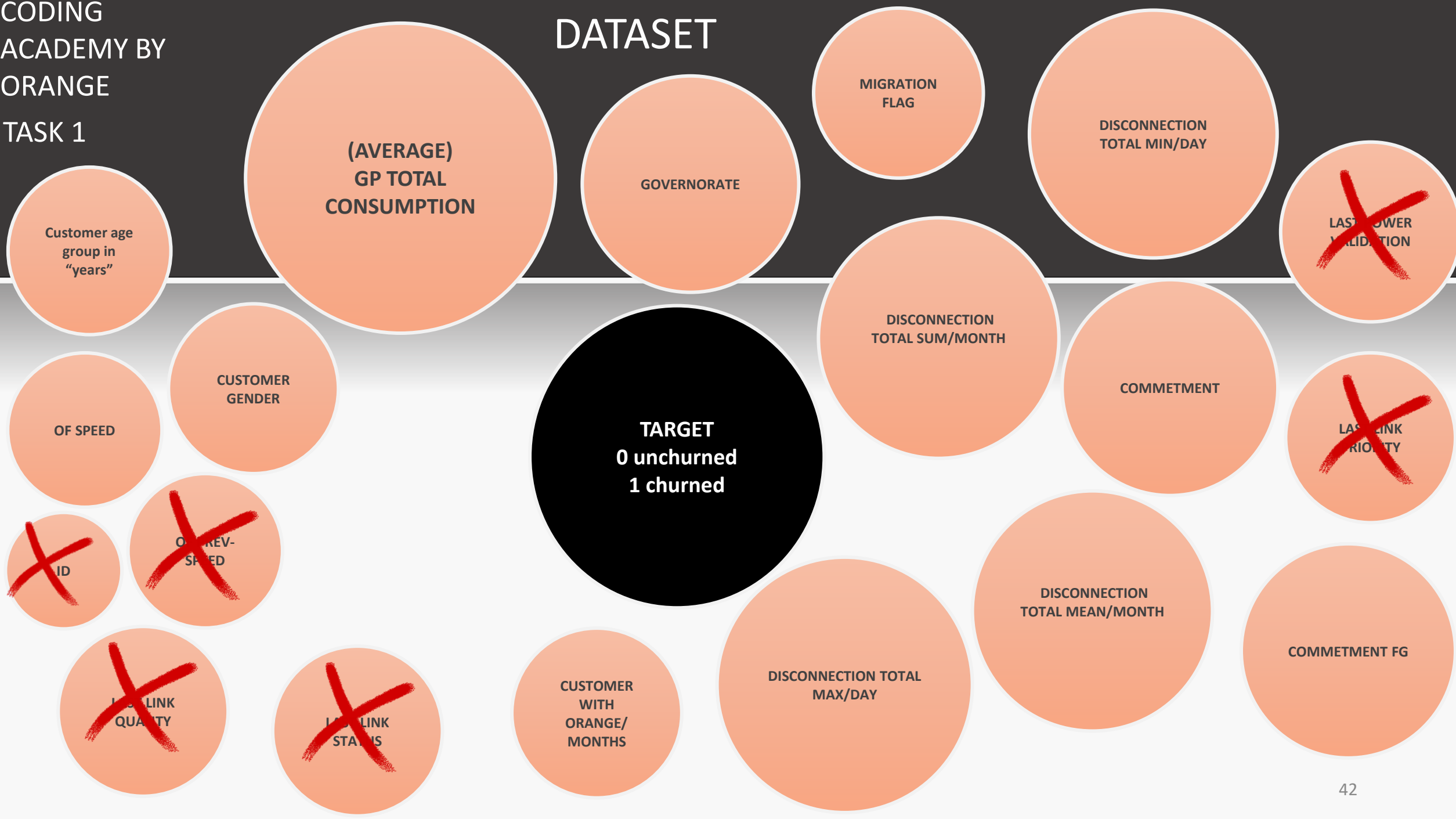


DATASET

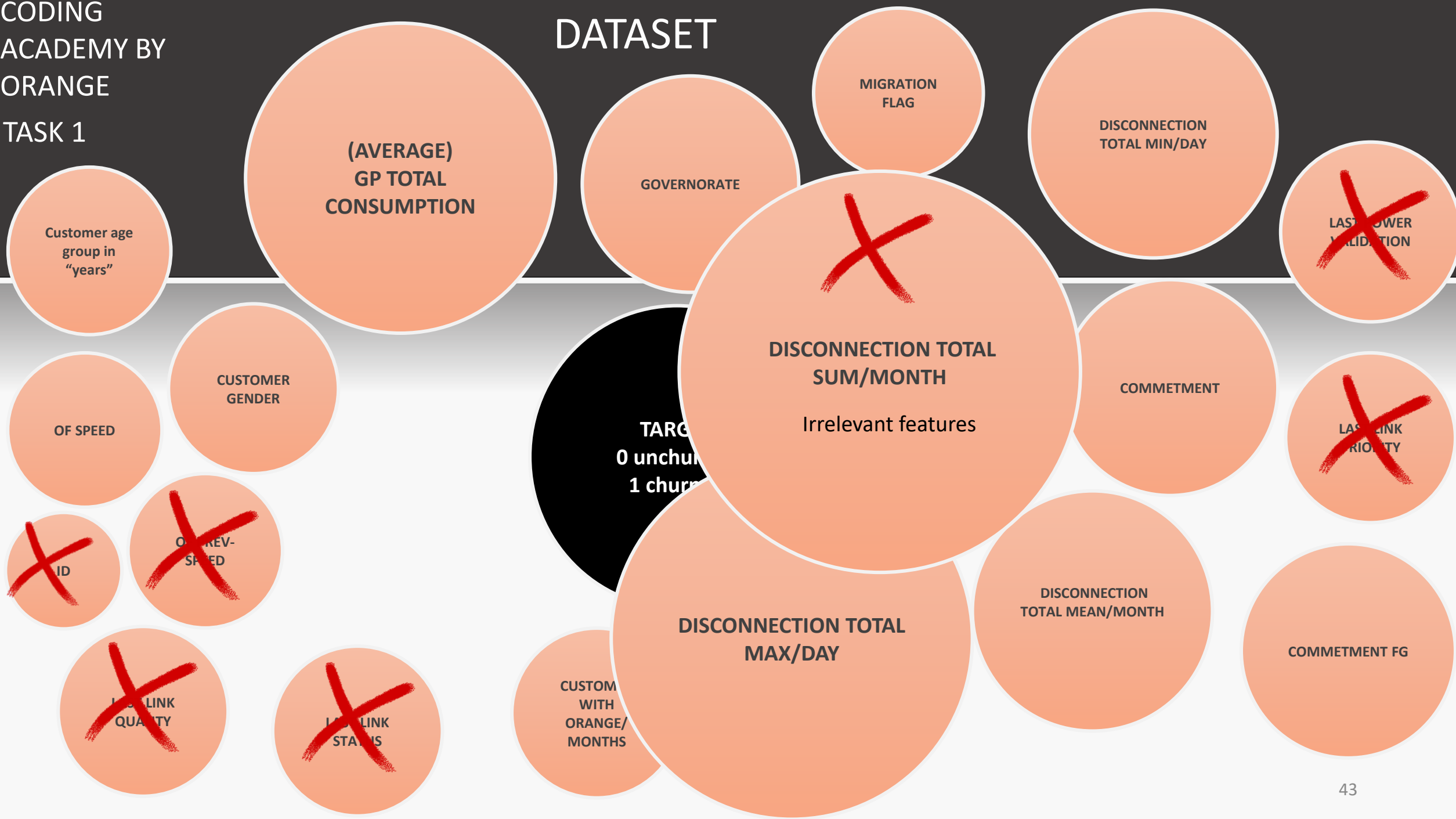


DATASET

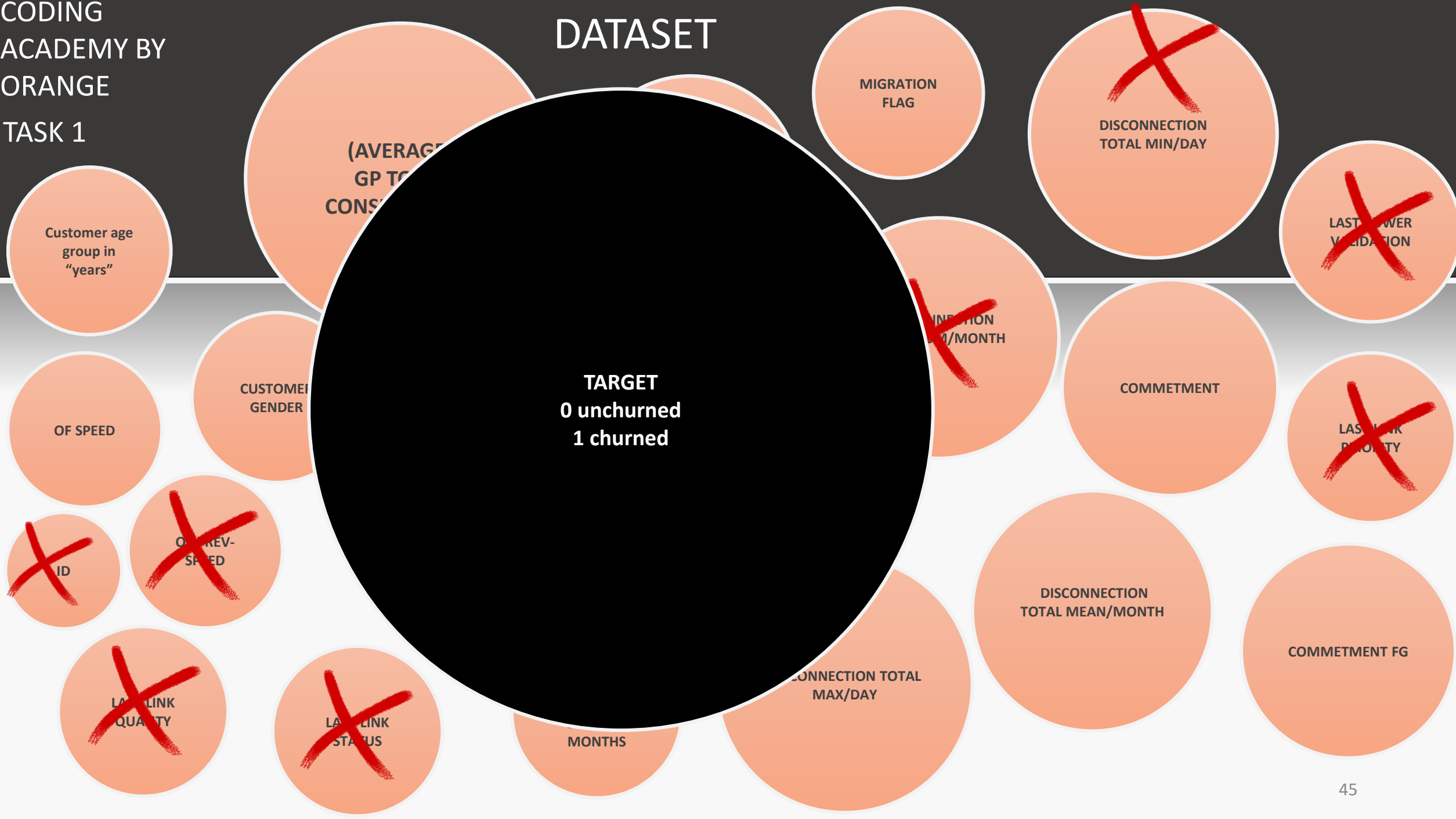




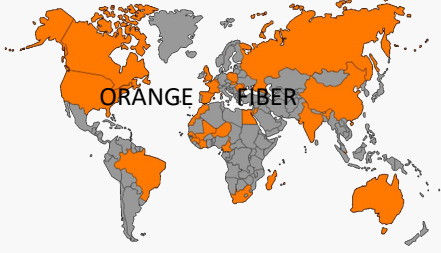
DATASET



DATASET

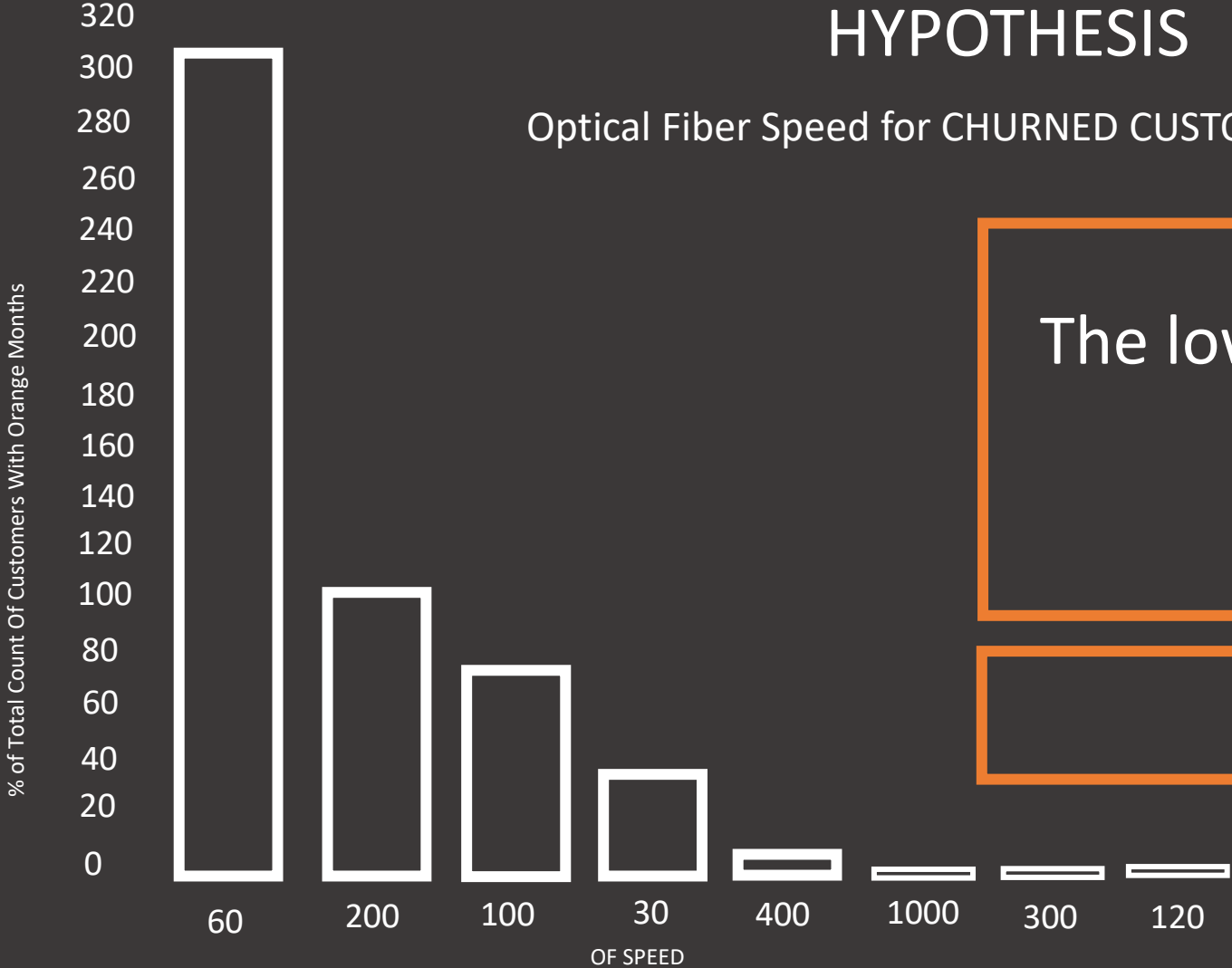


[illegible]



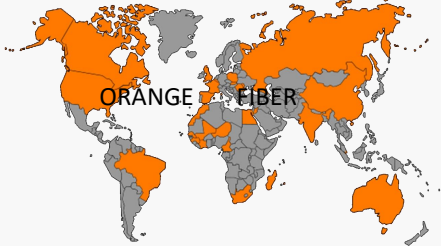
HYPOTHESIS

Optical Fiber Speed for CHURNED CUSTOMERS

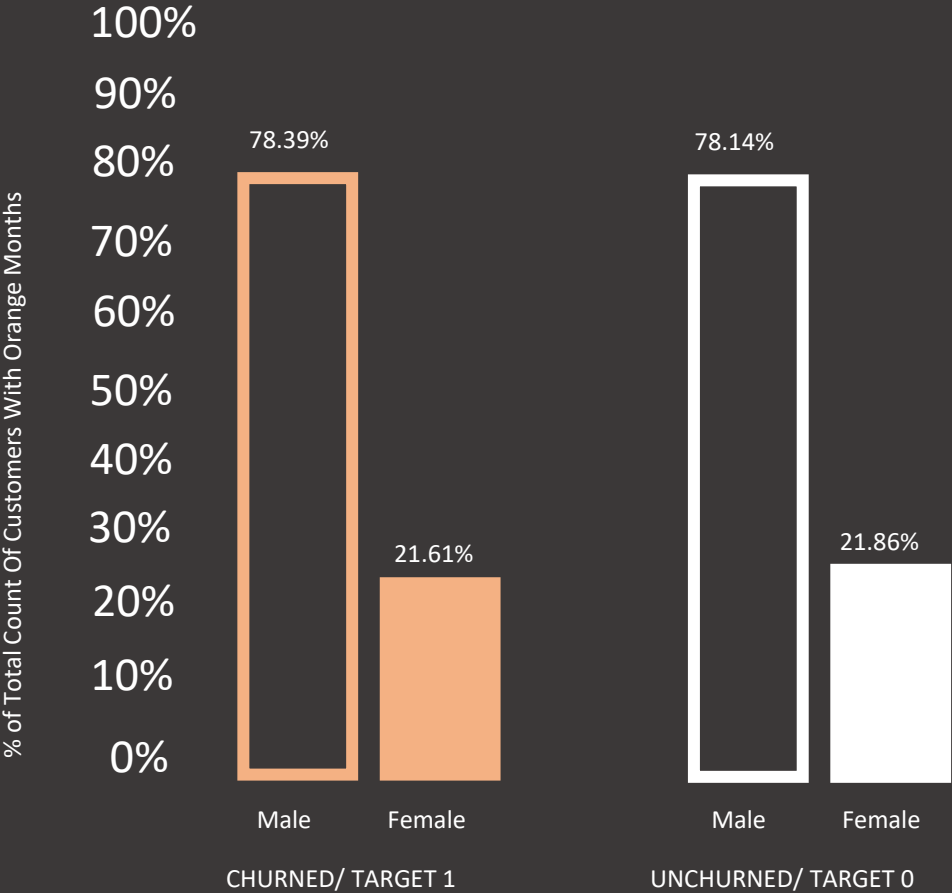


The lower the OF-SPEED, the higher a customer being churned

CORRECT

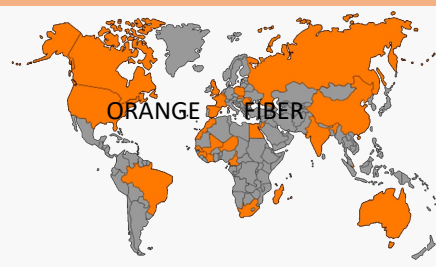


HYPOTHESIS



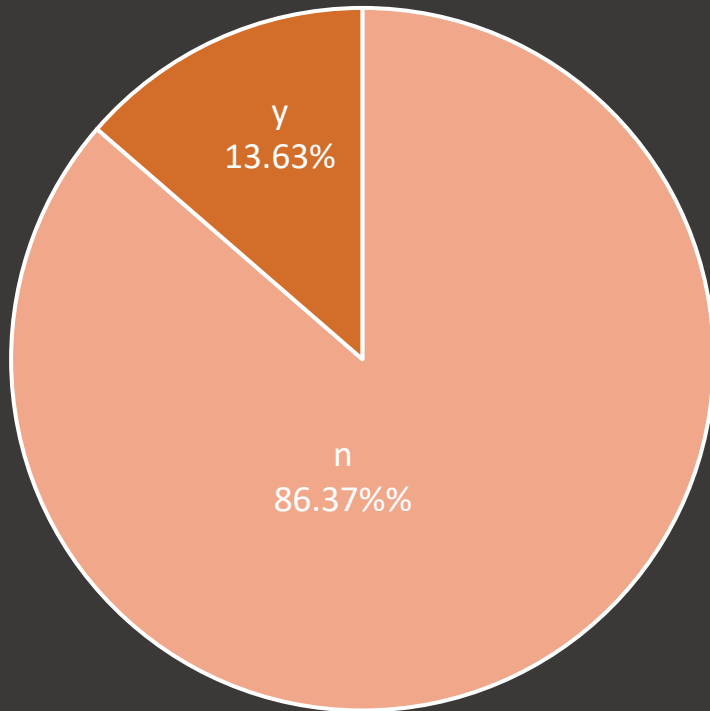
The probability of a male customer to be churned is higher than the probability of a female customer to be churned

WRONG



HYPOTHESIS

Churned Customers Who Either Upgraded Their Contract Or Not

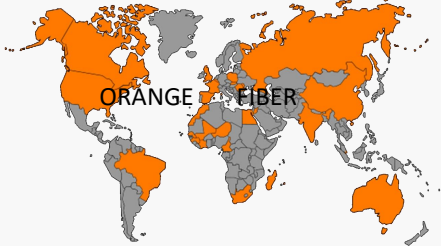


n: did not upgrade
y: upgraded

Customers who don't perform migration flag (n), have a higher probability to being churned

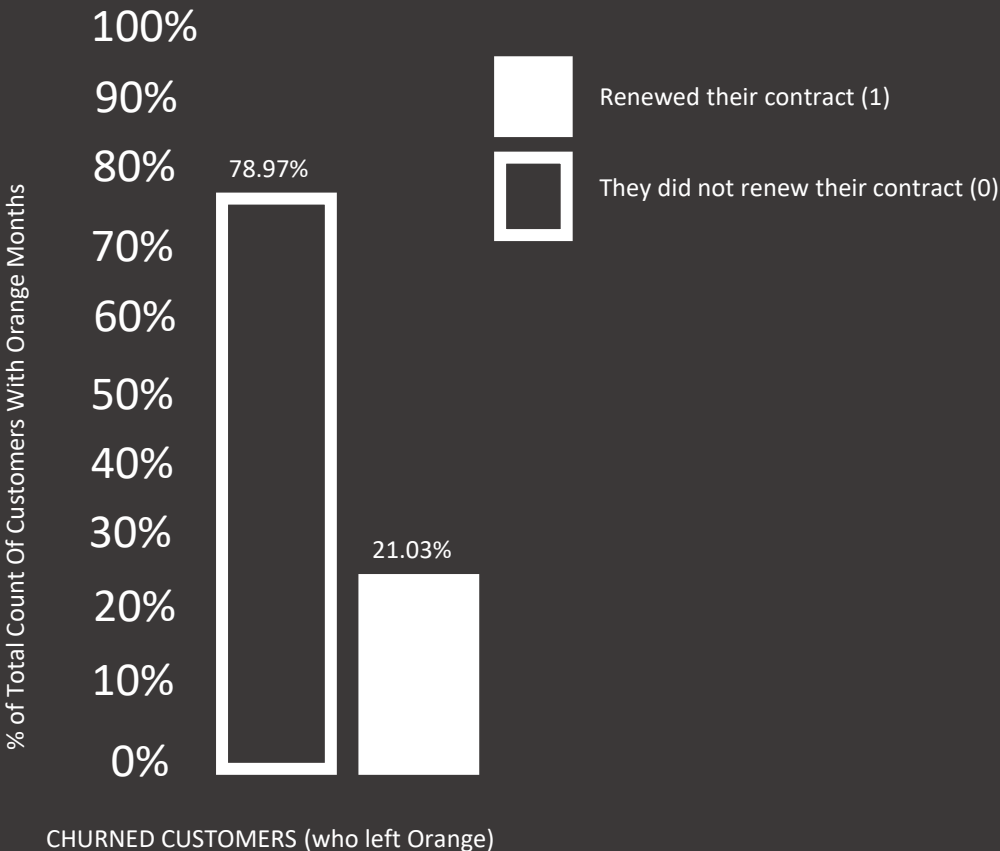
Migration Flag: The customers who upgraded their contract

CORRECT



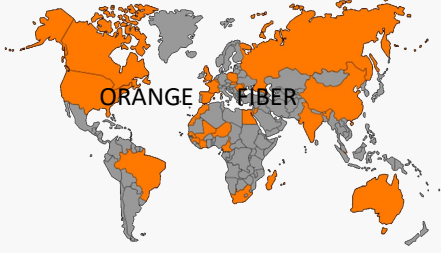
HYPOTHESIS

Churned Customers Who Either Renewed Their Contracts Or Not



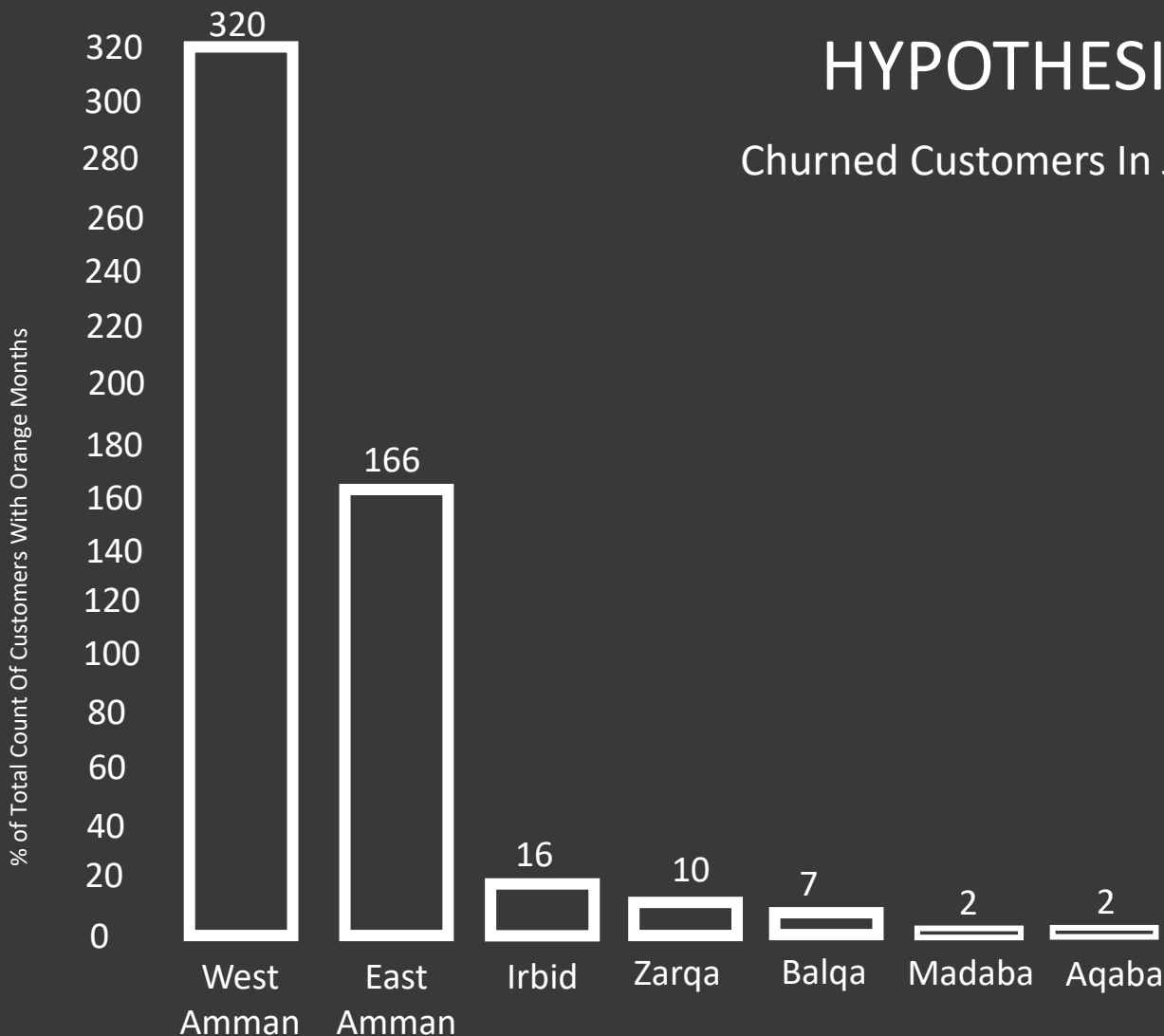
Customers who don't commit with their contract period are most likely to being churned more than the customers who commit with their contract period

CORRECT



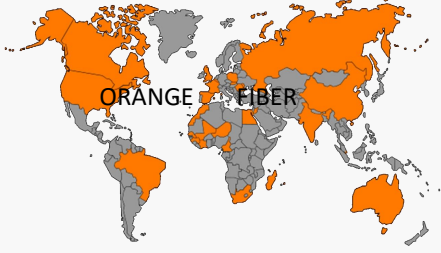
HYPOTHESIS

Churned Customers In Jordan



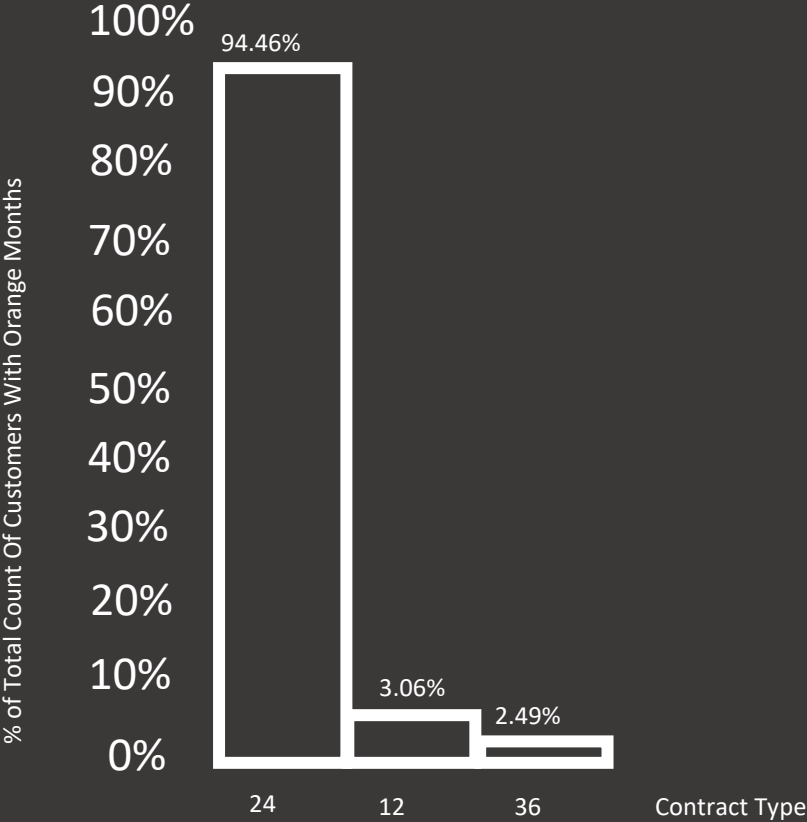
Higher churned customers
are from West Amman

WE CAN NOT PROVE IT



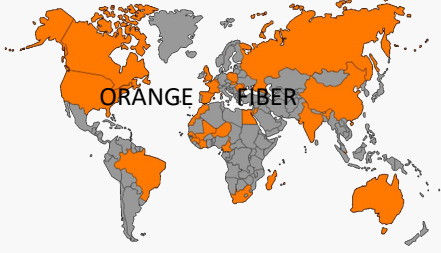
HYPOTHESIS

Churned Customers vs contract type



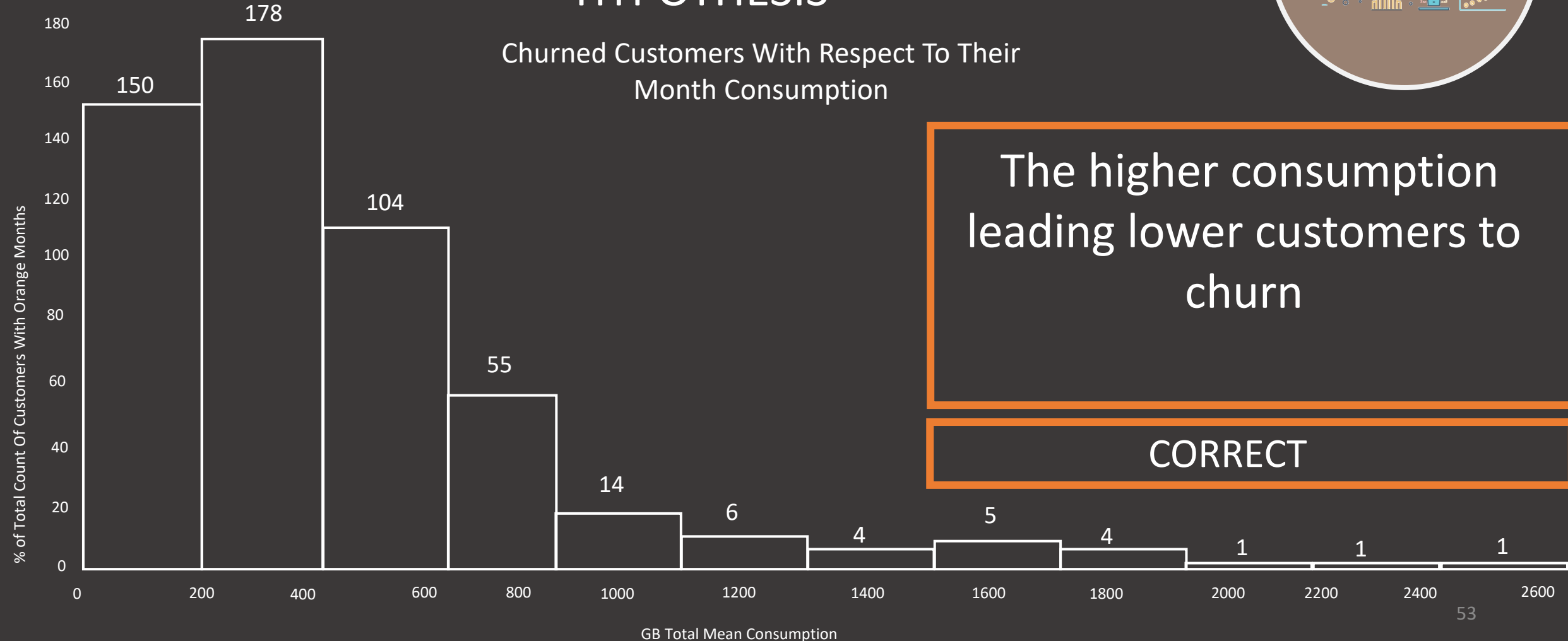
Customers with 24 months contract type are having higher probability to churn

CORRECT



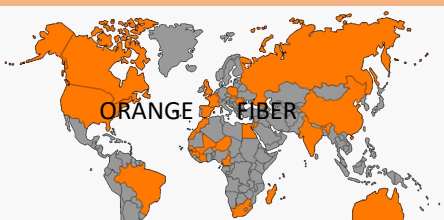
HYPOTHESIS

Churned Customers With Respect To Their
Month Consumption



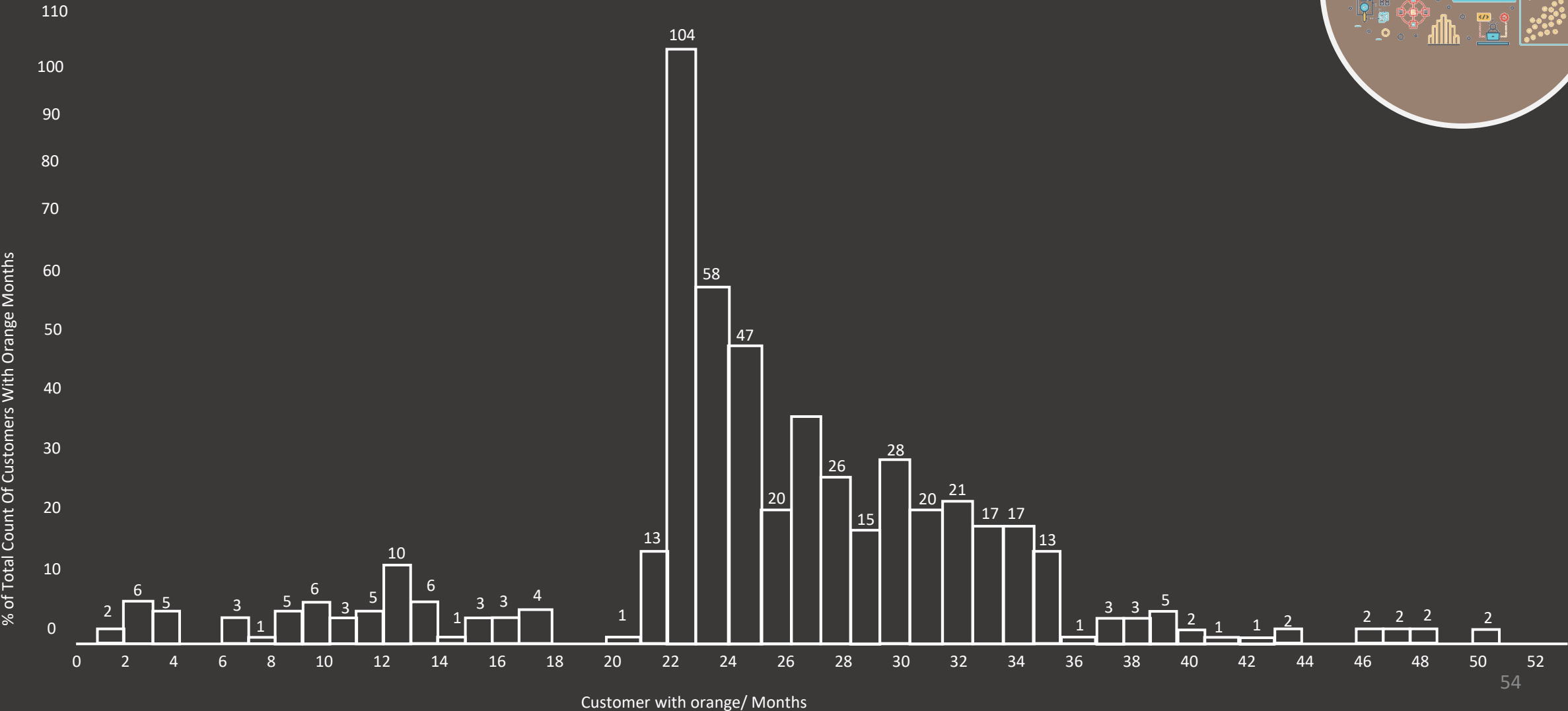
The higher consumption
leading lower customers to
churn

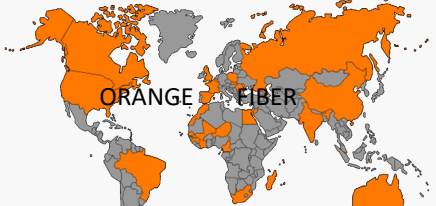
CORRECT



TASK 1

HYPOTHESIS

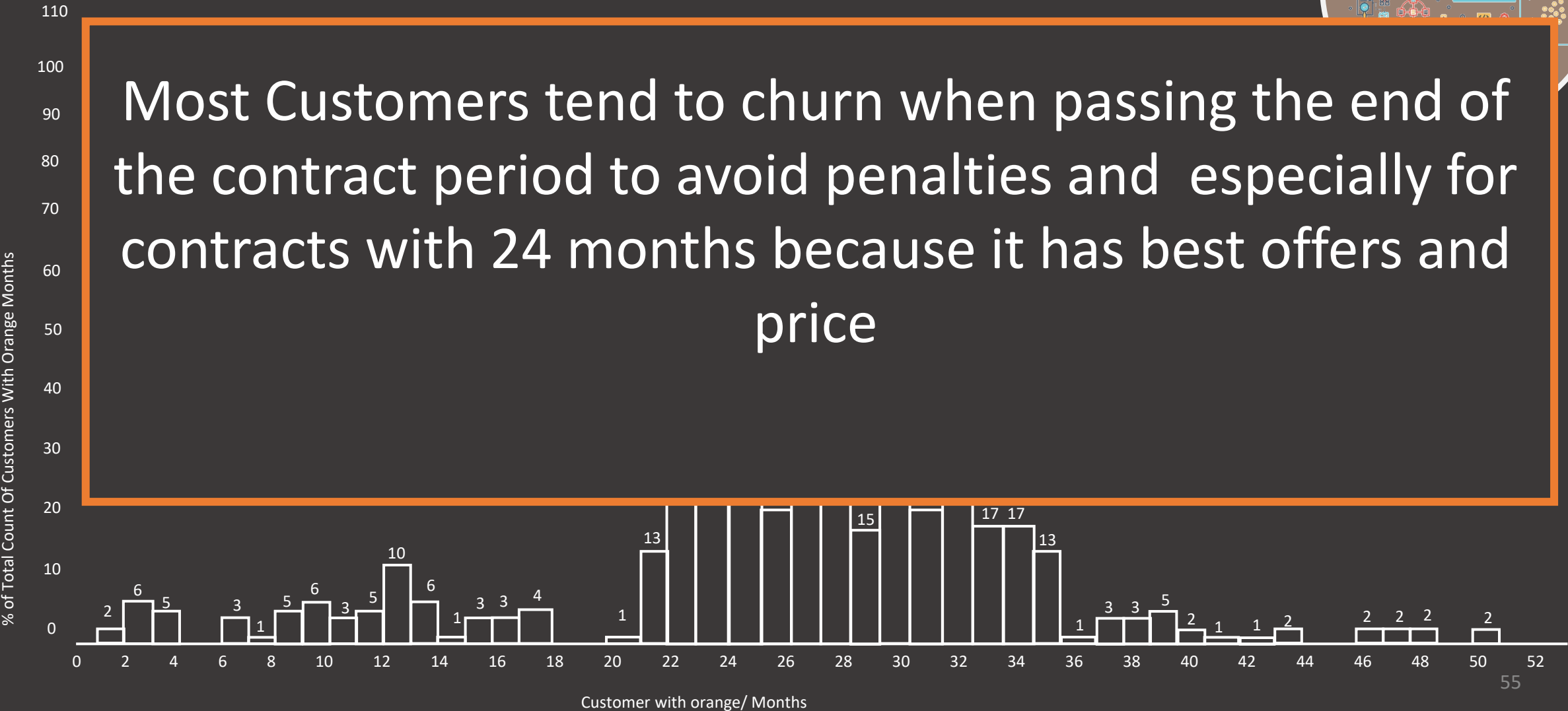


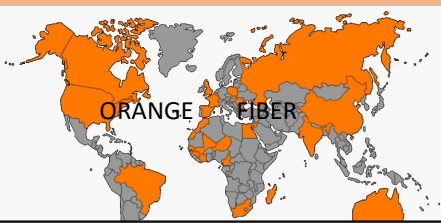


TASK 1

HYPOTHESIS

Most Customers tend to churn when passing the end of the contract period to avoid penalties and especially for contracts with 24 months because it has best offers and price



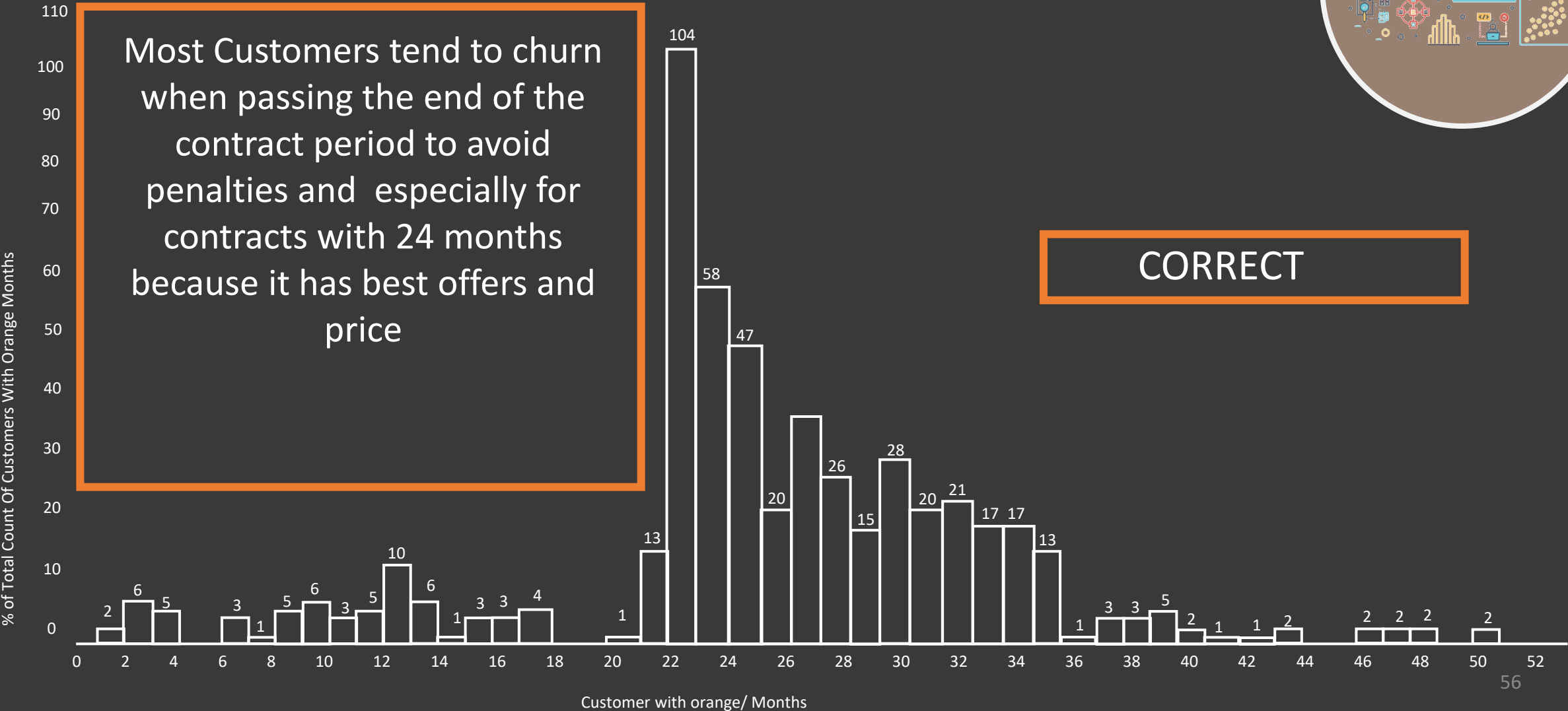


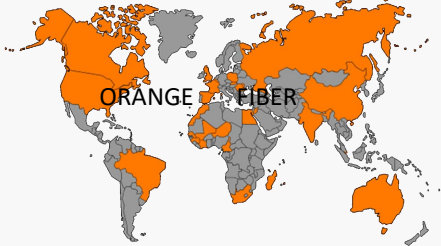
TASK 1

HYPOTHESIS

Most Customers tend to churn when passing the end of the contract period to avoid penalties and especially for contracts with 24 months because it has best offers and price

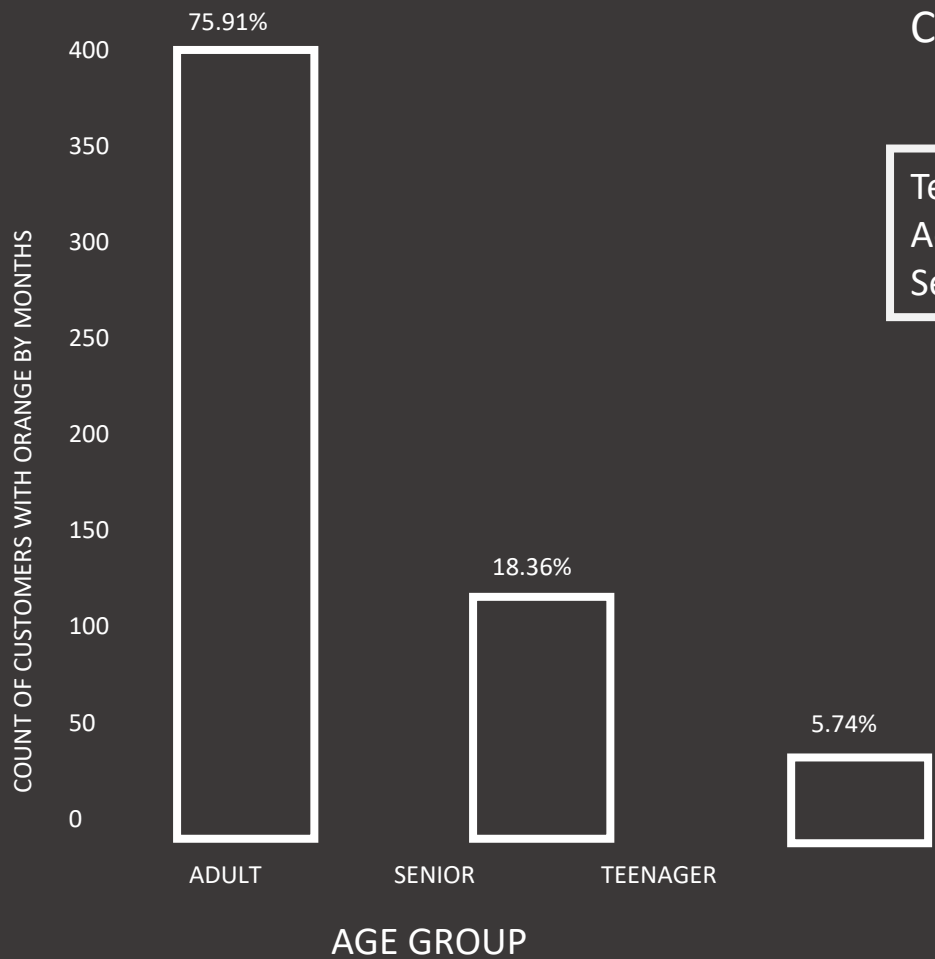
CORRECT





HYPOTHESIS

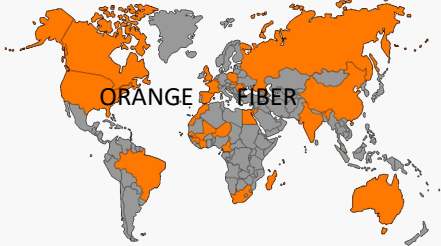
Churned Customers age group



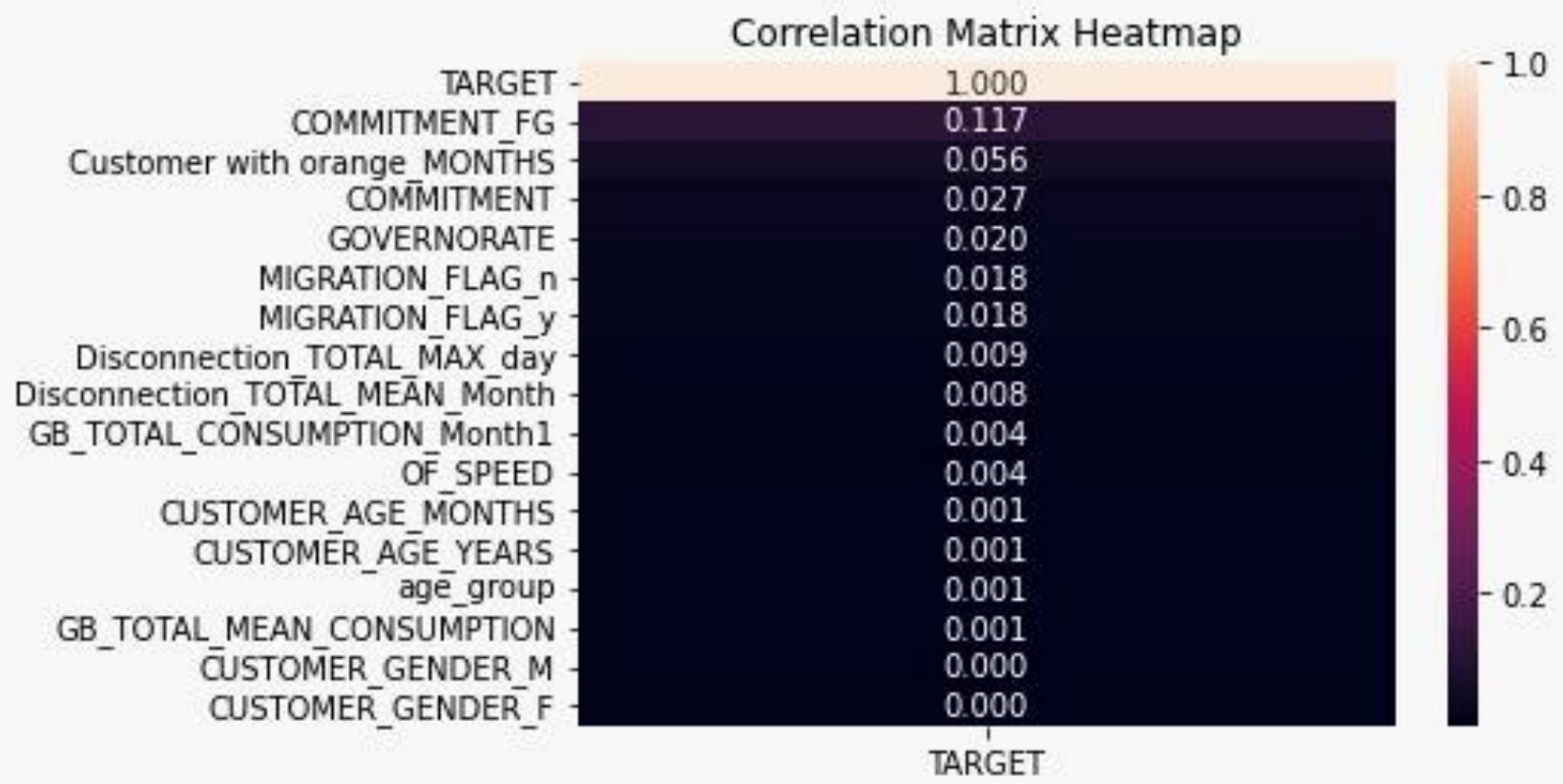
| | |
|----------|----------|
| Teenager | 19-25 |
| Adult | 25-60 |
| Seniors | Above 60 |

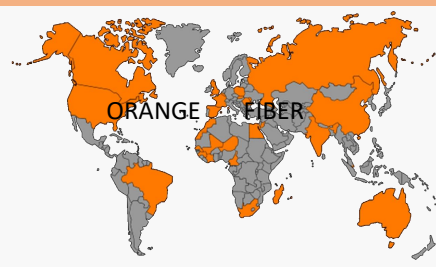
Adults are the “age group”
that churned the most

CORRECT



CORELATION MATRIX HEATMAP





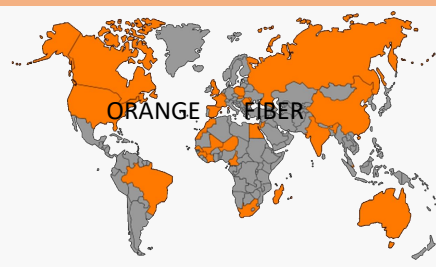
Future work

1- track down who participated in the data collection and who made the data.

2-collecting more features for our customers

- time-based features to capture the patterns and trends in customer behavior over time.
- education level, and occupation could provide insights into which customer groups are more likely to churn.
- customer service interactions, such as the number of complaints or inquiries, can provide insights into customer satisfaction with the service and identify areas for improvement.
- whether a customer has a spouse or not

3- Incorporate unstructured data such as social media or customer service transcripts that can provide insights into customer behavior and attitudes.



RECOMMENDATION

Ensuring the Supply Chain & Technical Departments are providing a high of-speed all the time, and continuous check for any technical issue could lead to disconnection, because these two are the main factors could lead a customer to churn

Offering attractive deals to customers before the end of their contract is a proactive way to prevent churn

it's important to note that churn rate is not the only metric that companies should consider when evaluating customer satisfaction and loyalty. Other factors such as customer engagement, customer lifetime value, and customer referral rates should also be taken into account

accurate measurement of churn rate requires careful consideration of the time interval and definition of churn, as well as the data quality and business context.