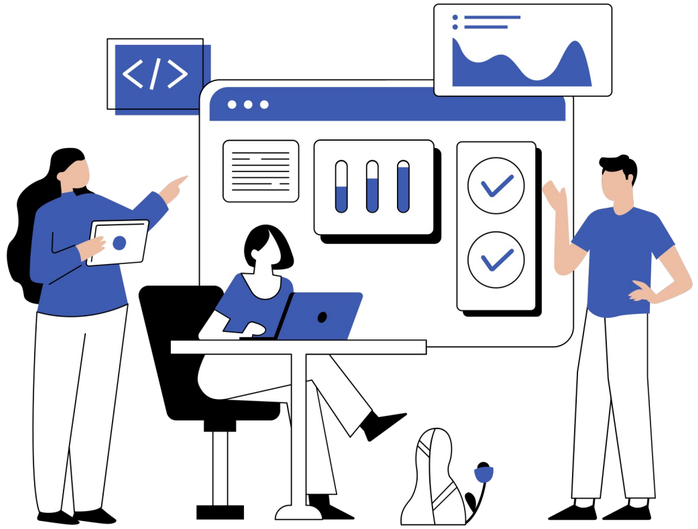
KATION

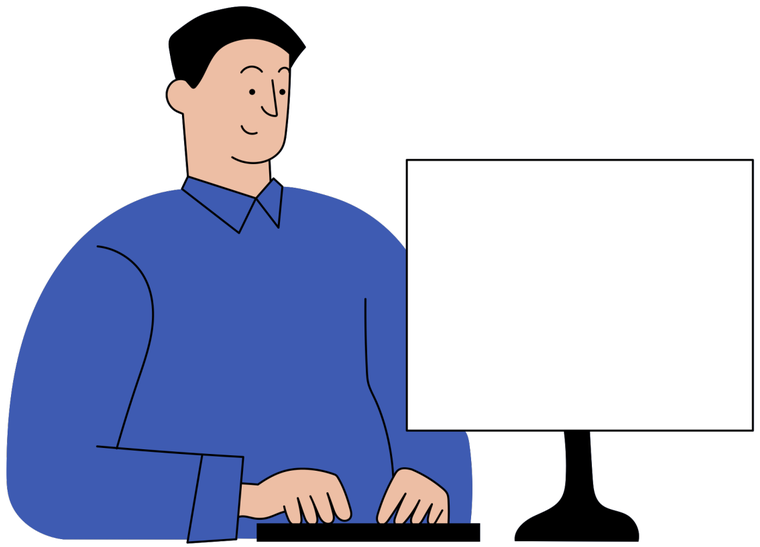
WQD 7005 DATA MINING



Evaluating Telco Campaign Performance And Predicting Campaign Offer Takers

KHAR SHIN YIN (S2145827) LEE SHIN EE (S2132380) OOI SHI YUAN (17088226) SIM LIN ZHENG (S2102170) YU YUEN HERN (S2121801)

Content



- Introduction
- Analysis Goal & Objective
- Dataset Description
- SEMMA: Sample
- SEMMA: Explore
- Conclusion

Introduction

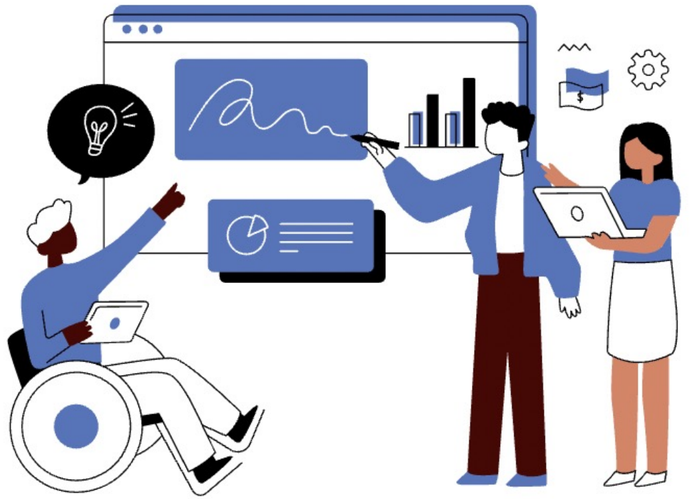
Platform migration in Kation

- An initiative called ‘Right Planning’ is conducted
  - to migrate old rate plans to new rate plans
- Business goal: To standardize the rate plans & enhance customers’ experience and users’ experience.
- Decided to launch a pilot campaign.

Analysis Goal & Objective

To assess pilot campaign performance before implementing to the entire customer base.

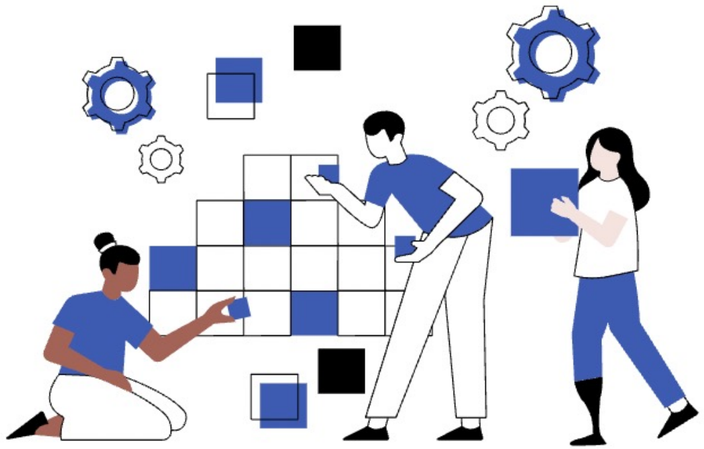
- To evaluate the effectiveness of ‘Right Planning’ pilot campaign.
- To determine campaign takers’ profile.
- To predict campaign takers based on usage and revenue behavior.



Dataset Description

Campaign data in customer level

- Demographic
- Usage before and after campaign
- Revenue before and after campaign
- Status before and after campaign



Demographic & Status

ID
TENURE
AGE
GENDER
NATIONALITY
STATE
STATUS_BEFORE
STATUS_AFTER
OFFER_TAKER
OFFER_TAKE_UP_DT
ACTIVITY_DAYS_AFTER
ACTVIITY_STATUS_AFTER

Usage

DATA_USG_BEFORE
DATA_USG_AFTER
DATA_PURC_BEFORE
DATA_PURC_AFTER
RLD_IND_BEFORE
RLD_IND_AFTER
VOICE_USG_BEFORE
VOICE_USG_AFTER

Revenue

ARPU_BEFORE
ARPU_AFTER
CPA_RVN_BEFORE
CPA_RVN_AFTER
DATA_CHRG_BEFORE
DATA_CHRG_AFTER
RLD_AMT_BEFORE
RLD_AMT_AFTER

Sample

7272 records, 27 features

Metadata Completed.

Library: DATASET
Data Source: Dataset
Role: Raw

Basic

Role	Level	Count
Input	Interval	7
Input	Nominal	20

Metadata Completed.

Library: DATASET
Data Source: Dataset
Role: Raw

Advance

Role	Level	Count
Input	Binary	5
Input	Interval	7
Input	Nominal	5
Rejected	Nominal	8
Rejected	Unary	2

Sample

Reclassification

Name	Role	Level
ACTIVITY_STATUS_AFTER	INPUT	NOMINAL
ACTVIITY_DAYS_AFTER	INPUT	INTERVAL
AGE	INPUT	INTERVAL
ARPU_AFTER	REJECTED	NOMINAL
ARPU_BEFORE	REJECTED	NOMINAL
CPA_RVN_AFTER	REJECTED	NOMINAL
CPA_RVN_BEFORE	REJECTED	NOMINAL
DATA_CHRG_AFTER	REJECTED	NOMINAL
DATA_CHRG_BEFORE	REJECTED	NOMINAL
DATA_PURC_AFTER	INPUT	BINARY
DATA_PURC_BEFORE	INPUT	BINARY
DATA_USG_AFTER	INPUT	INTERVAL
DATA_USG_BEFORE	INPUT	INTERVAL
GENDER	INPUT	NOMINAL
NATIONALITY	REJECTED	UNARY
OFFER_TAKER	INPUT	NOMINAL
OFFER_TAKE_UP_DT	INPUT	BINARY
RLD_AMT_AFTER	REJECTED	NOMINAL
RLD_AMT_BEFORE	REJECTED	NOMINAL
RLD_IND_AFTER	INPUT	BINARY
RLD_IND_BEFORE	INPUT	BINARY
STATE	INPUT	NOMINAL
STATUS_AFTER	INPUT	NOMINAL
STATUS_BEFORE	REJECTED	UNARY
TENURE	INPUT	INTERVAL
VOICE_USG_AFTER	INPUT	INTERVAL
VOICE_USG_BEFORE	INPUT	INTERVAL

Advance

Name	Role	Level
ACTIVITY_STATUS_AFTER	INPUT	NOMINAL
ACTVIITY_DAYS_AFTER	INPUT	INTERVAL
AGE	INPUT	INTERVAL
ARPU_AFTER	INPUT	NOMINAL
ARPU_BEFORE	INPUT	NOMINAL
CPA_RVN_AFTER	INPUT	NOMINAL
CPA_RVN_BEFORE	INPUT	NOMINAL
DATA_CHRG_AFTER	INPUT	NOMINAL
DATA_CHRG_BEFORE	INPUT	NOMINAL
DATA_PURC_AFTER	INPUT	BINARY
DATA_PURC_BEFORE	INPUT	BINARY
DATA_USG_AFTER	INPUT	INTERVAL
DATA_USG_BEFORE	INPUT	INTERVAL
GENDER	INPUT	NOMINAL
NATIONALITY	INPUT	NOMINAL
OFFER_TAKER	TARGET	BINARY
OFFER_TAKE_UP_DT	INPUT	NOMINAL
RLD_AMT_AFTER	INPUT	NOMINAL
RLD_AMT_BEFORE	INPUT	NOMINAL
RLD_IND_AFTER	INPUT	BINARY
RLD_IND_BEFORE	INPUT	BINARY
STATE	INPUT	NOMINAL
STATUS_AFTER	INPUT	NOMINAL
STATUS_BEFORE	INPUT	NOMINAL
TENURE	INPUT	INTERVAL
VOICE_USG_AFTER	INPUT	INTERVAL
VOICE_USG_BEFORE	INPUT	INTERVAL

Manual

Explore

Univariate Analysis

STATE
GENDER

Inconsistent

AGE

Intentional

ARPU_BEFORE
ARPU_AFTER
CPA_RVN_BEFORE
CPA_RVN_AFTER
DATA_CHRG_BEFORE
DATA_CHRG_AFTER
RLD_AMT_BEFORE
RLD_AMT_AFTER

Noisy & Incomplete

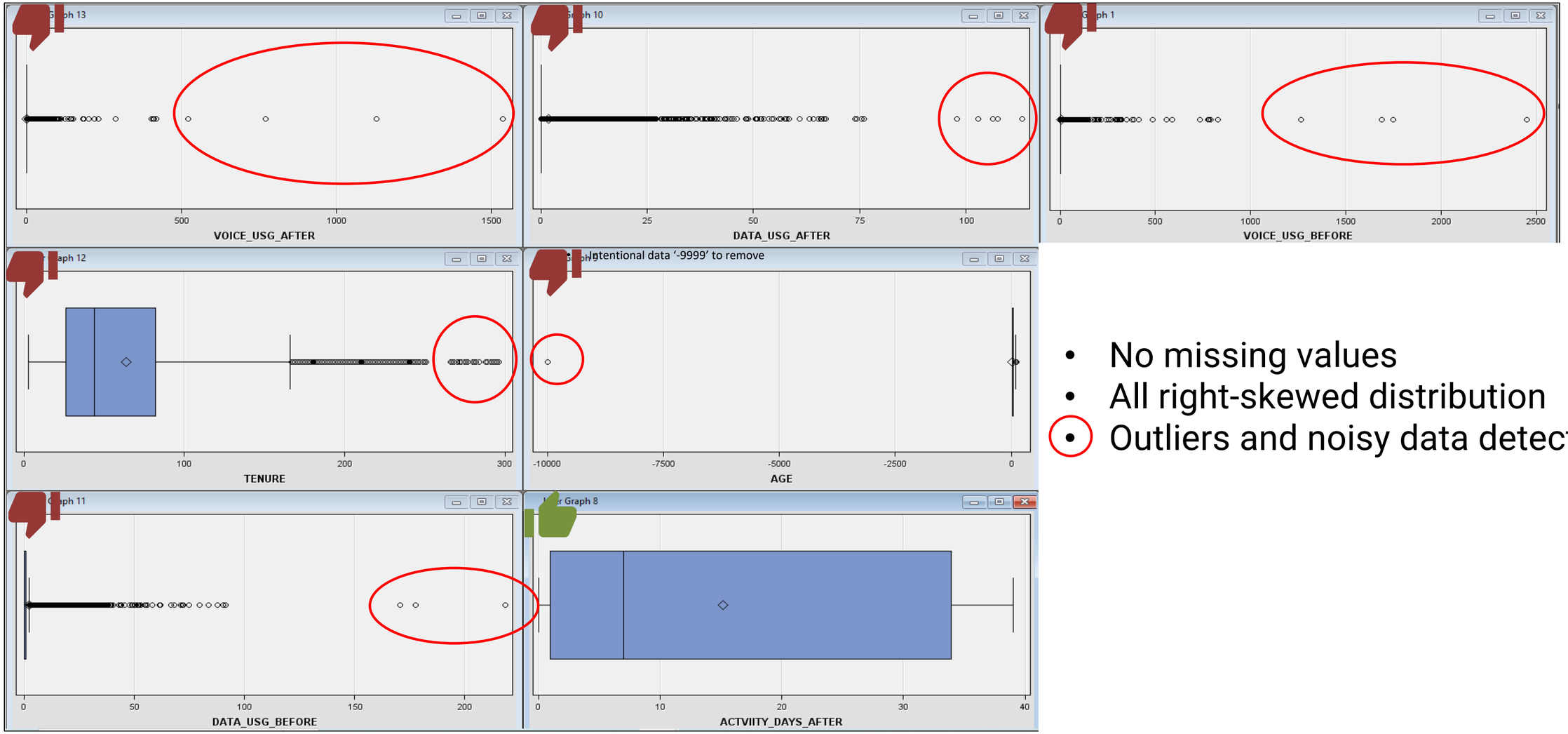
OFFER_TAKE_UP_DT

Incomplete

- There are 4 issues found: Intentional, Noisy, Incomplete & Inconsistent data
- Perform data cleaning during Modify stage.

Explore

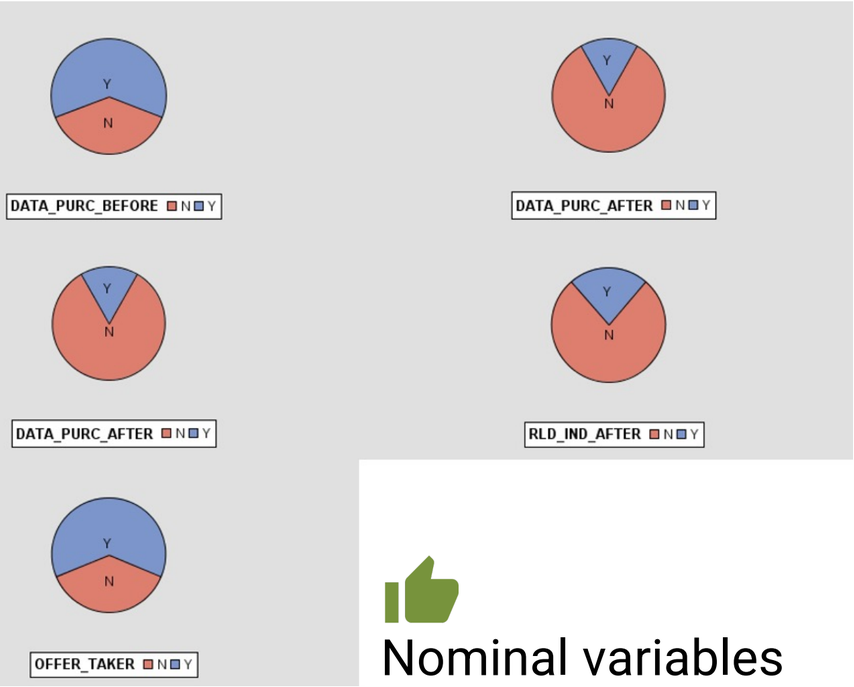
Univariate Analysis | Boxplot



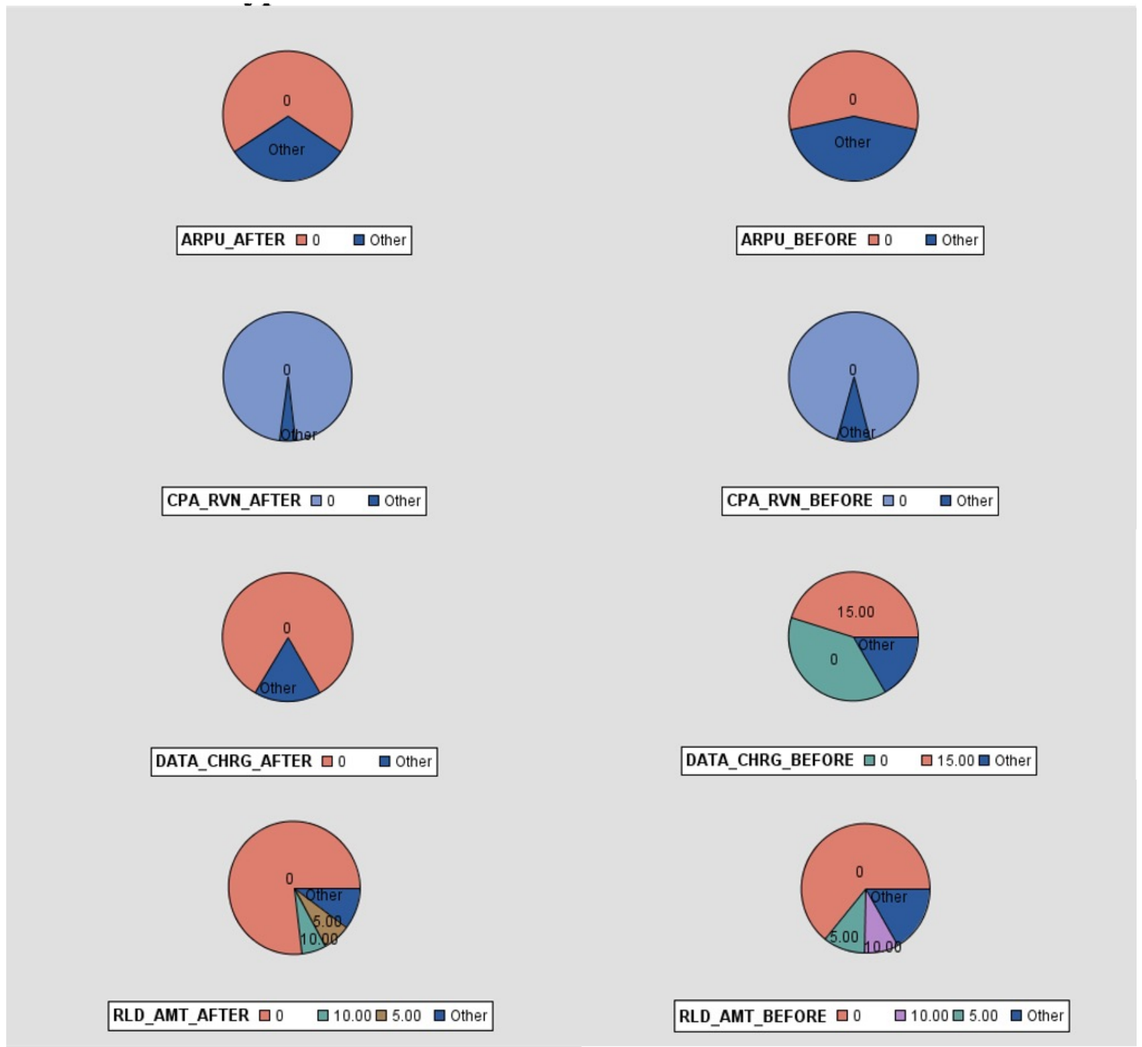
- No missing values
- All right-skewed distribution
- Outliers and noisy data detected

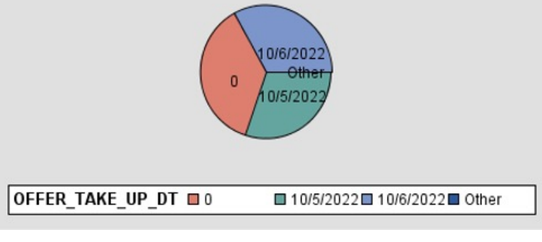
Explore

Univariate Analysis | Pie chart



Nominal variables

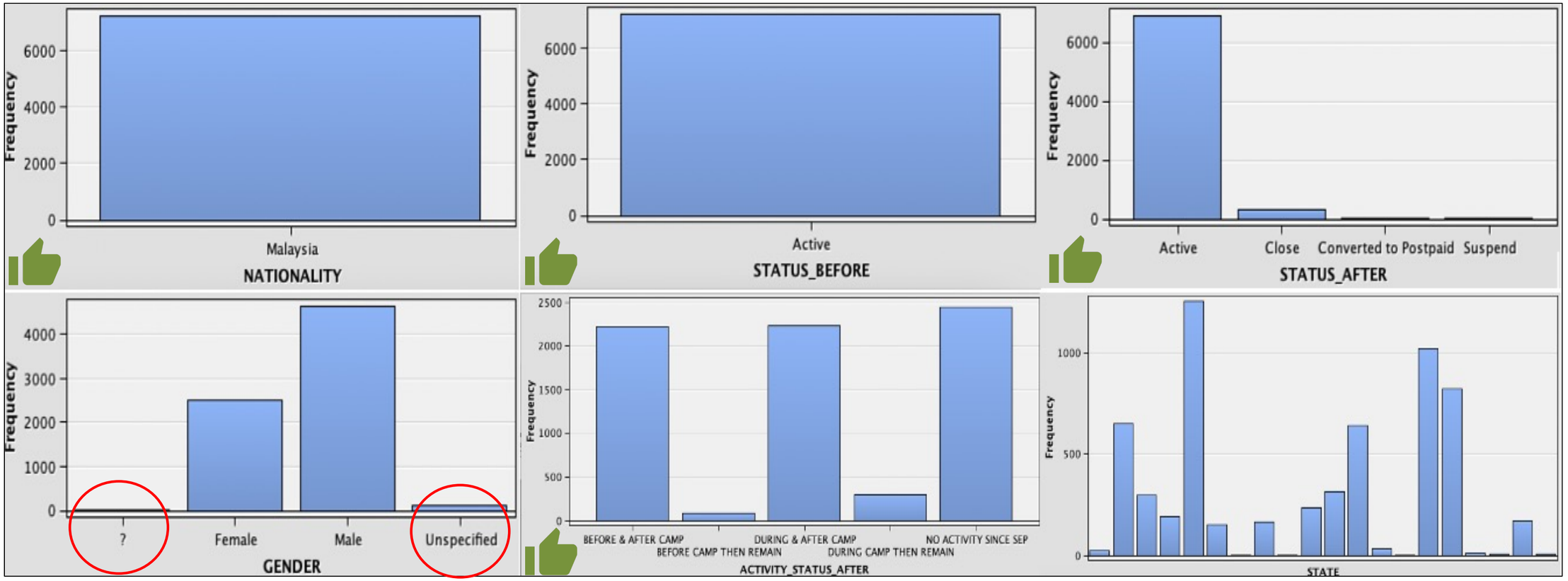




Many missing values
Wrongly tagged as nominal

Explore

Univariate Analysis | Bar Chart

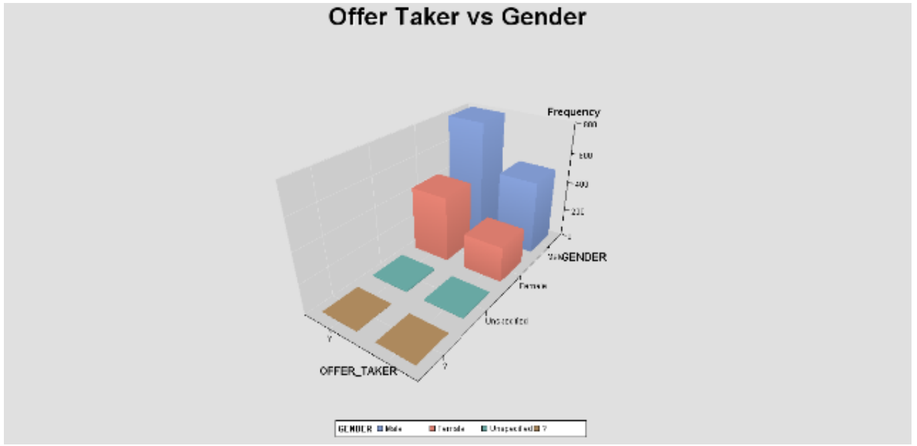


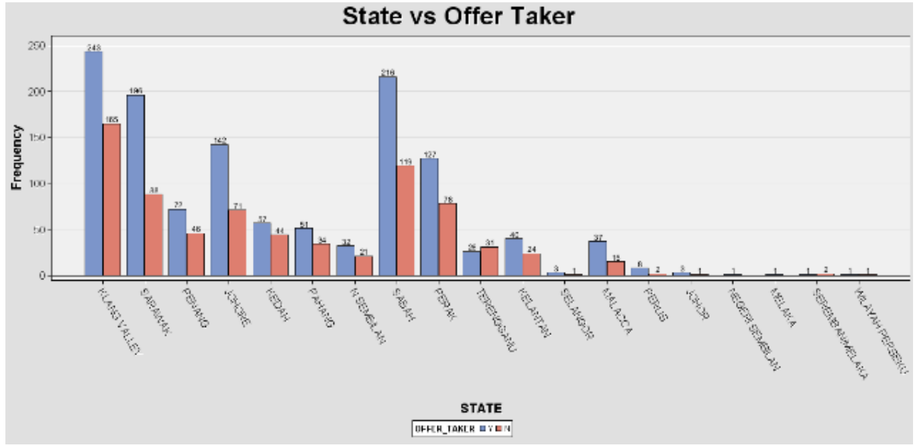
Inconsistent value:
'?' and 'unspecified' is same

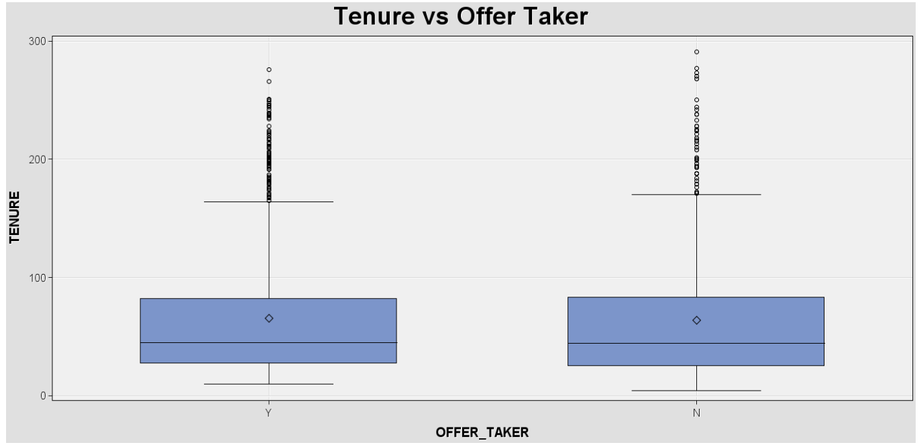
Inconsistent value:
more than 13 states

Explore

Bivariate Analysis



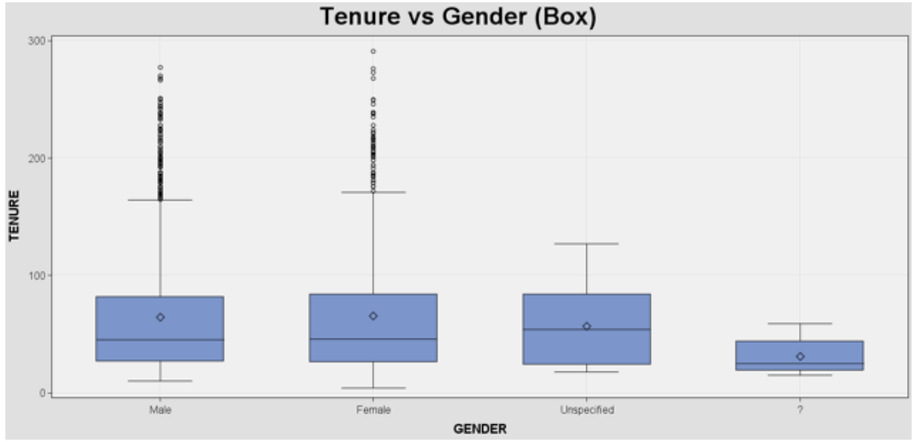


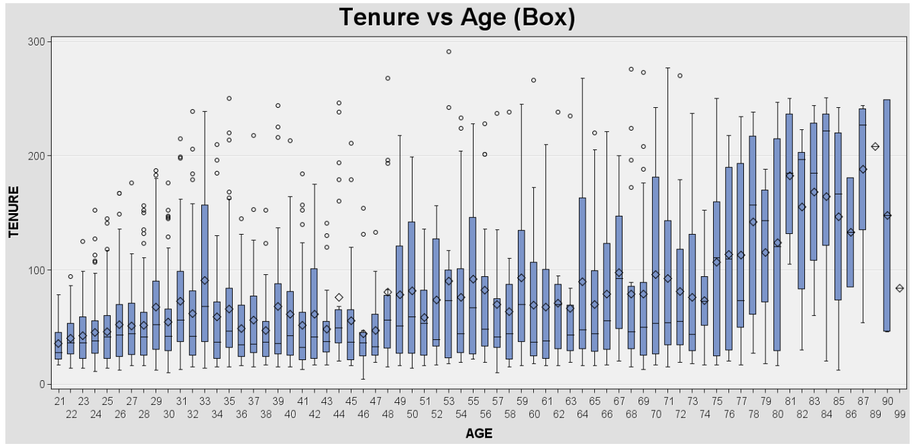


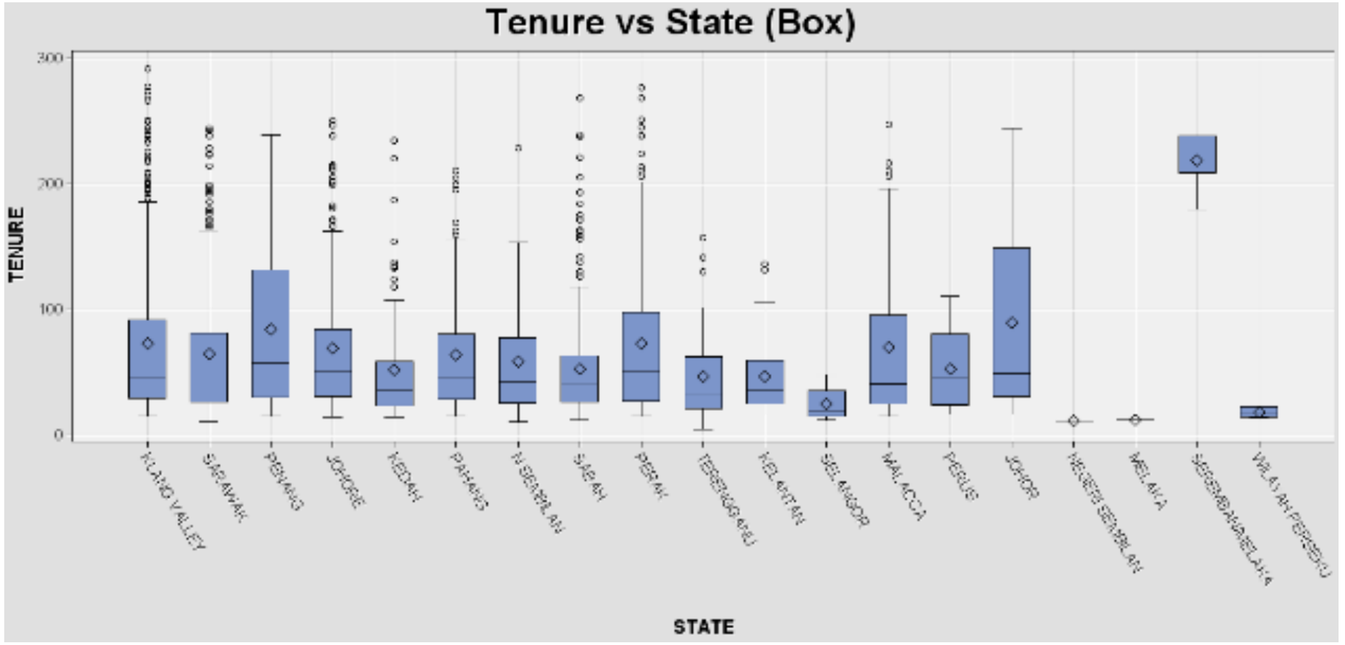
- Males likely to join campaign, most participants from Klang Valley, median age of 34 for both, median tenure of ~45

Explore

Bivariate Analysis



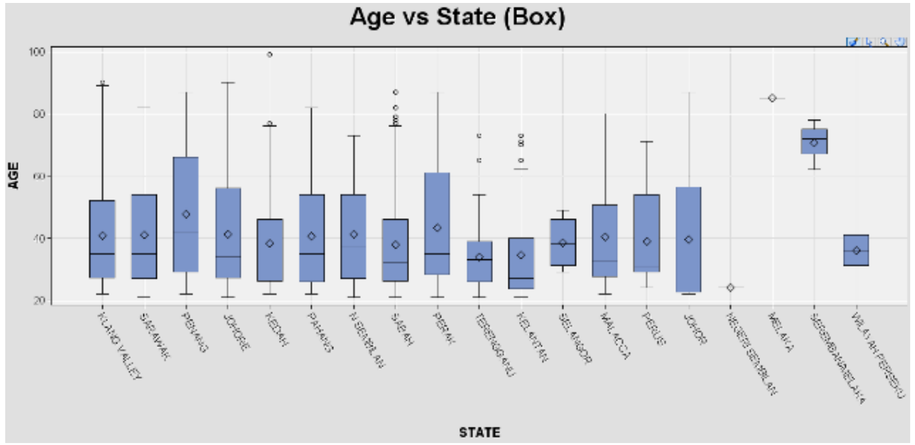


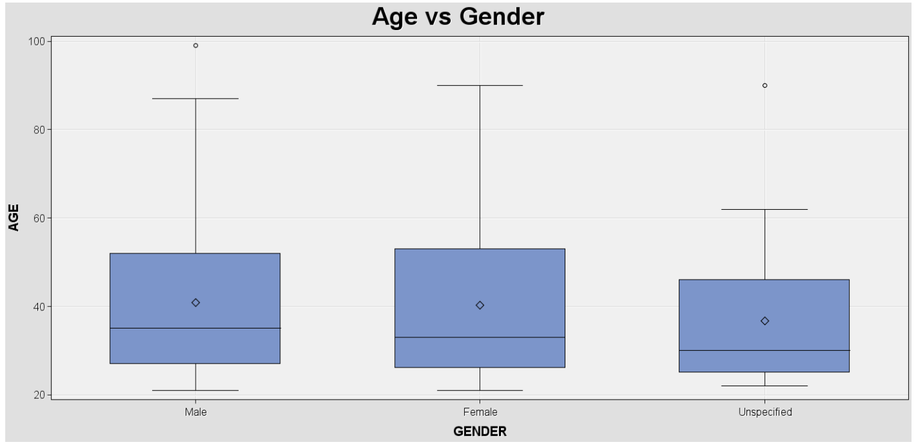


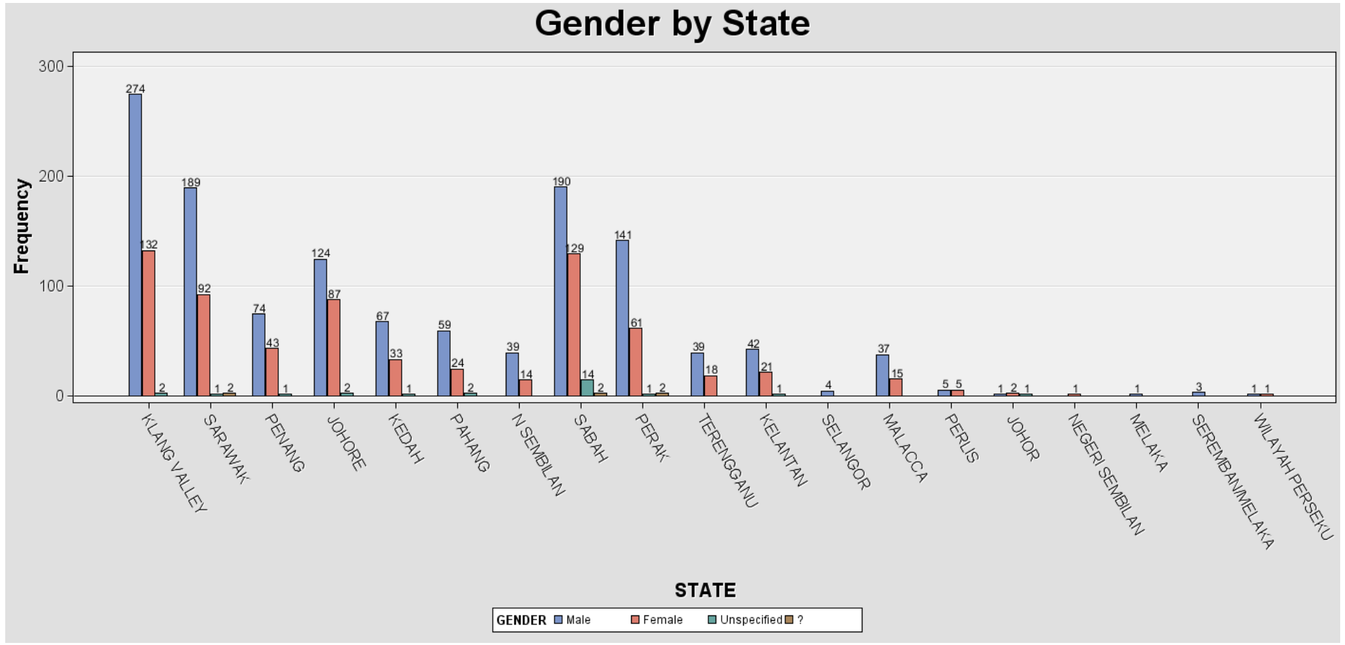
- Similar tenure distribution for both male & female)
- Younger age = shorter tenure
- Varying tenure gap in all states – surprisingly short tenures in Kedah

Explore

Bivariate Analysis



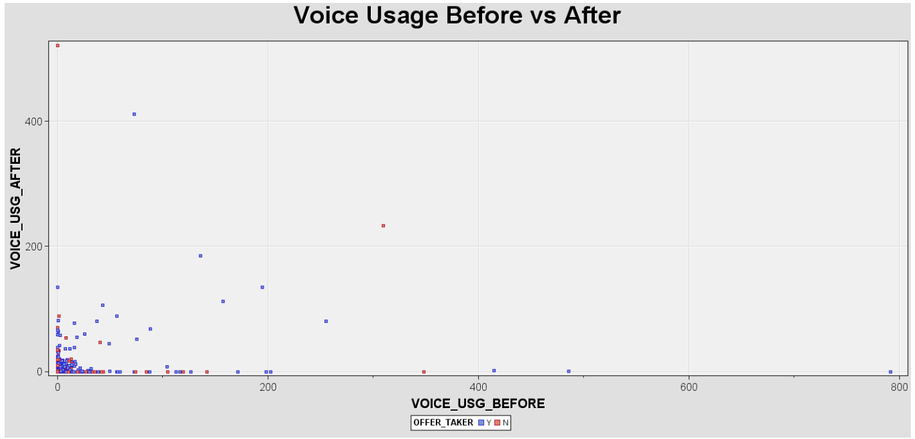


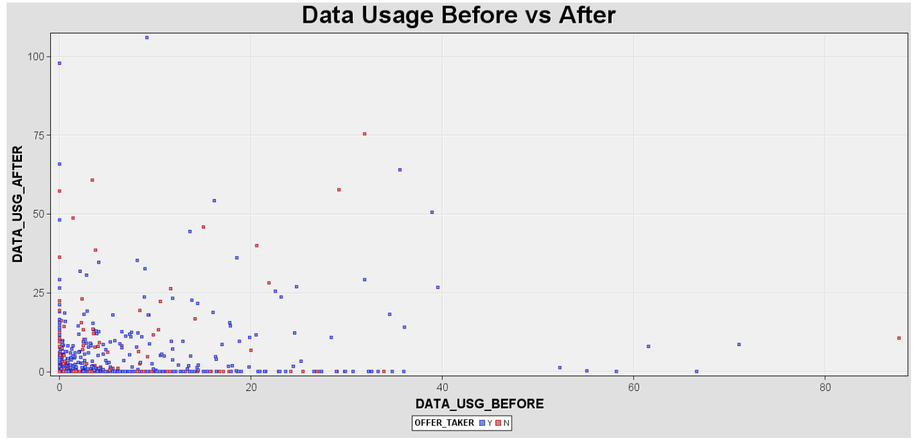


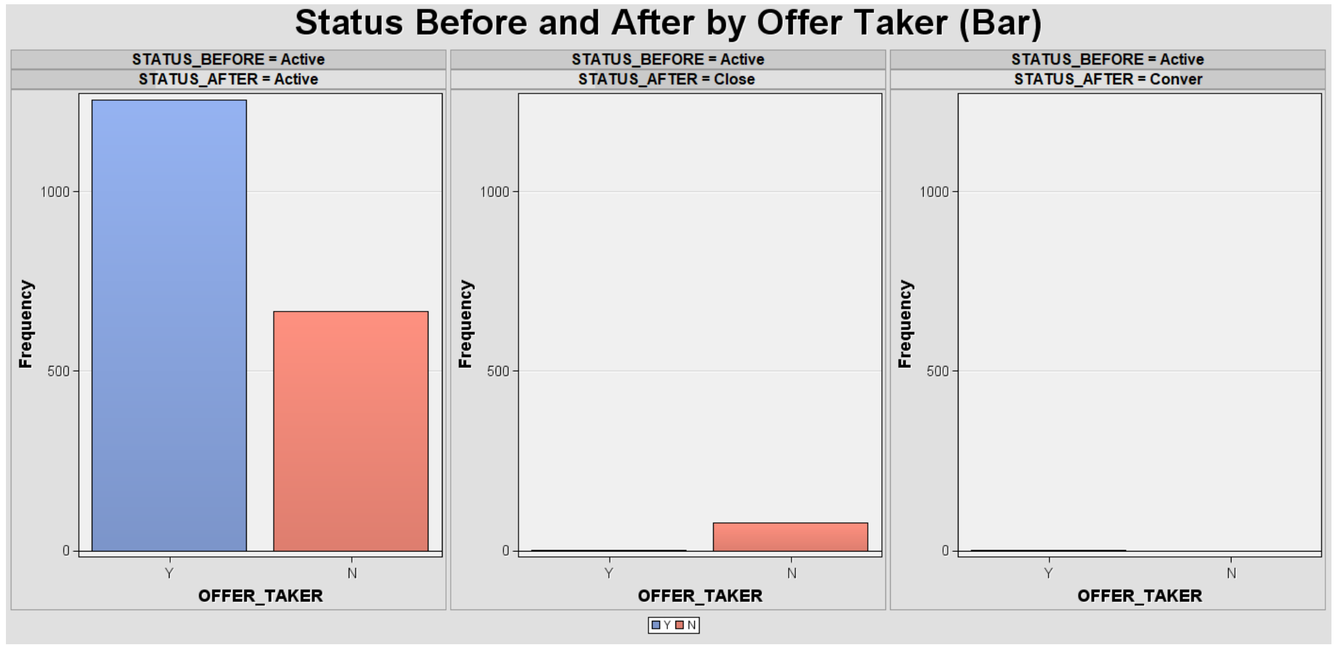
- Exceptionally young group of participants in Terengganu and Kelantan
- Similar median age for male and female (~34)
- Sabah has the most customers with undisclosed gender

Explore

Multivariate Analysis



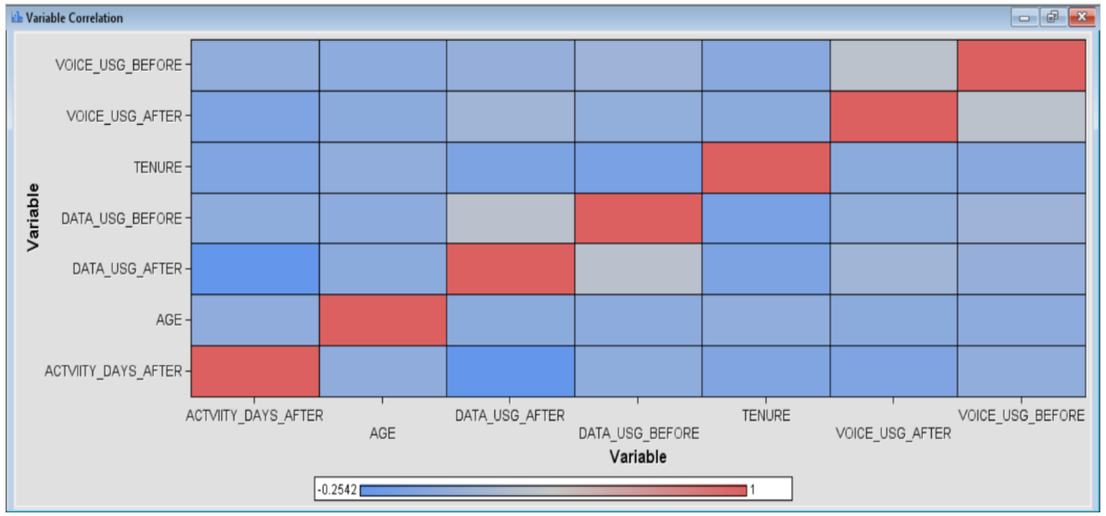


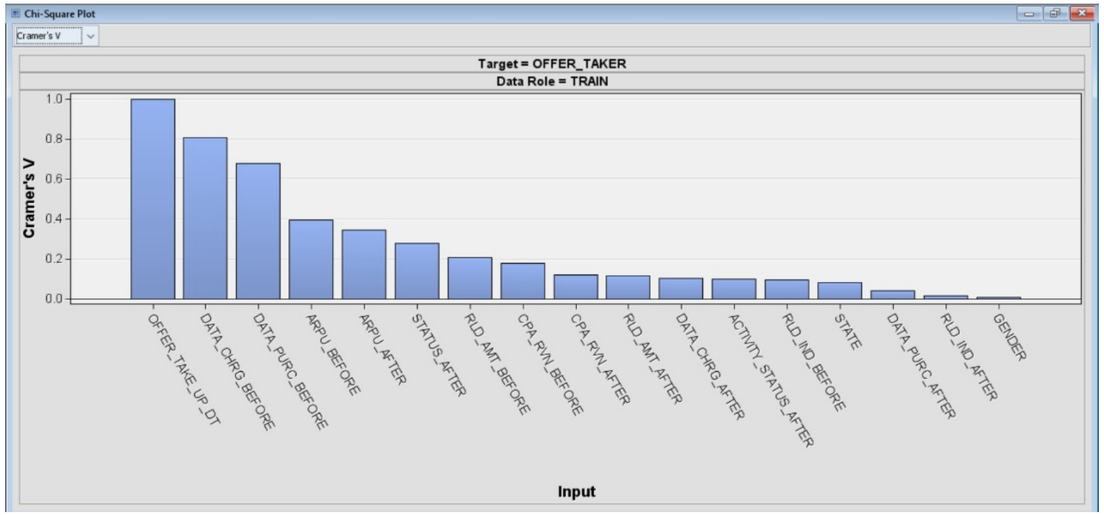


- No significant changes on the data and voice usage before and after campaign
- Takers more likely to remain active while non-takers more likely to churn

Explore

Correlation Analysis





For interval variables:

- No correlation value > 0.9
- No variables were removed

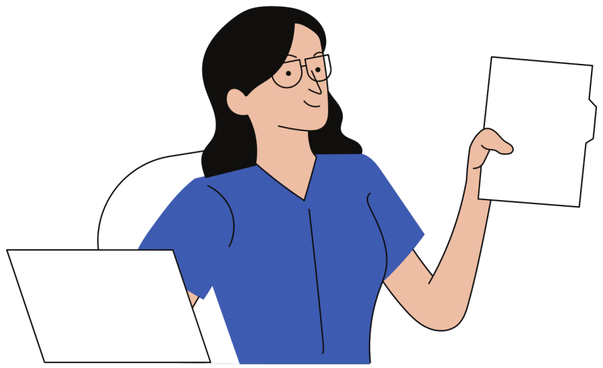
For nominal variables:

- OFFER_TAKE_UP_DT showed a prefect association

Conclusion

The dataset needs to undergo pre-processing as there are incomplete, noisy, inconsistent and intentional data.

Based on exploration, overview and hypothesis are generated and will further justify in the process of Modify, Model and Assess.



01

To assess the effectiveness of “Right Planning” pilot campaign

- 4573 takers (63% opt in rate)
- 2921 takers remain active after campaign (64% active rate)
- Moderately success

02

To identify campaign takers’ profile

- Target group is sample of Kation’s prepaid Malaysian
- Higher opt in rate among male
- Age group between 22 - 36 years old with tenure > 1 year
- Mostly from Klang Valley, then Sabah and Sarawak

That's a wrap!

Thank you.

10th Dec 2022