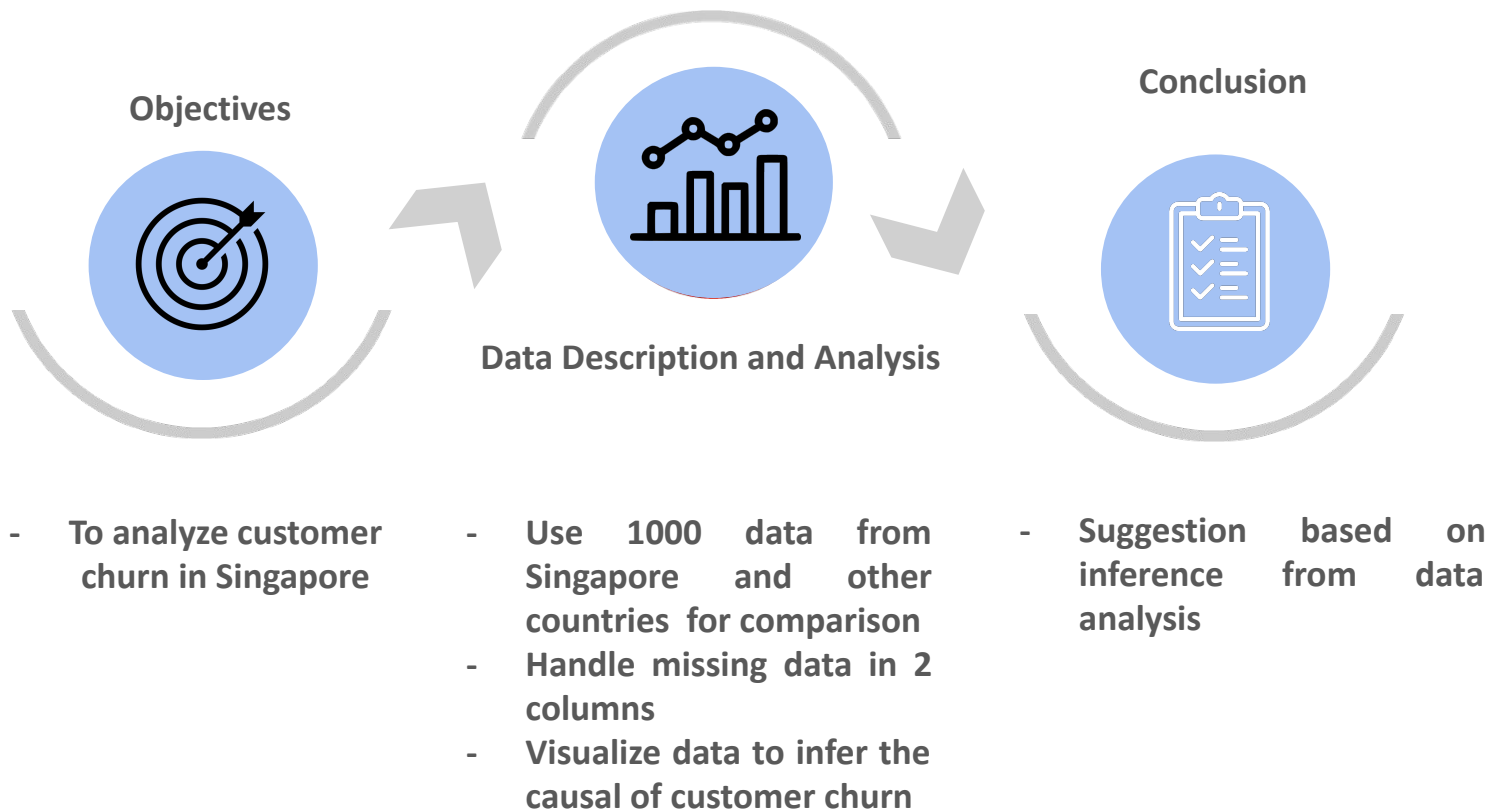


# AVAIL EDA Analysis

Iqbal Ridalta Putra





# OBJECTIVES

## Problems



Avail is a news streaming platform supporting local language



Deep learning implementation to create mouth movement for the news



In Singapore, the 1st 4 months establish rapid growth before stagnated

## Goals



Evaluate performance in Singapore, analyze churn and decide business strategy for the future recommendation

# DATA ANALYSIS - MISSING DATA

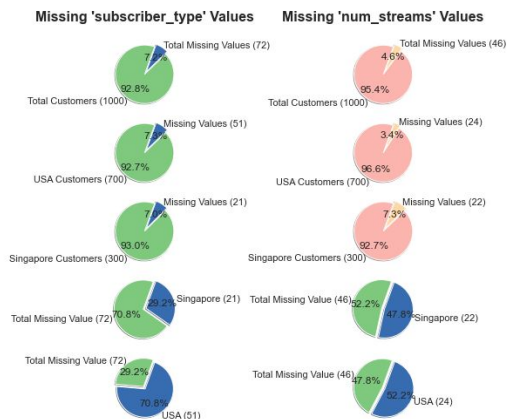
	customer_id	country_name	age	customer_name	is_subscriber	subscriber_type	num_streams
0	1	united_states	21.0	Kasen Todd	True	aavail_premium	23.0
1	2	singapore	31.0	Ensley Garza	False	NaN	12.0
2	3	united_states	22.0	Lillian Carey	False	aavail_premium	22.0
3	4	united_states	21.0	Beau Christensen	True	aavail_basic	19.0
4	5	singapore	22.0	Ernesto Gibson	True	aavail_premium	23.0

## Data Columns

```
customer_id      0
country_name     0
age              0
customer_name    0
is_subscriber    0
subscriber_type  72
num_streams      46
dtype: int64
```

## Missing Value

- subscriber\_type
- num\_streams



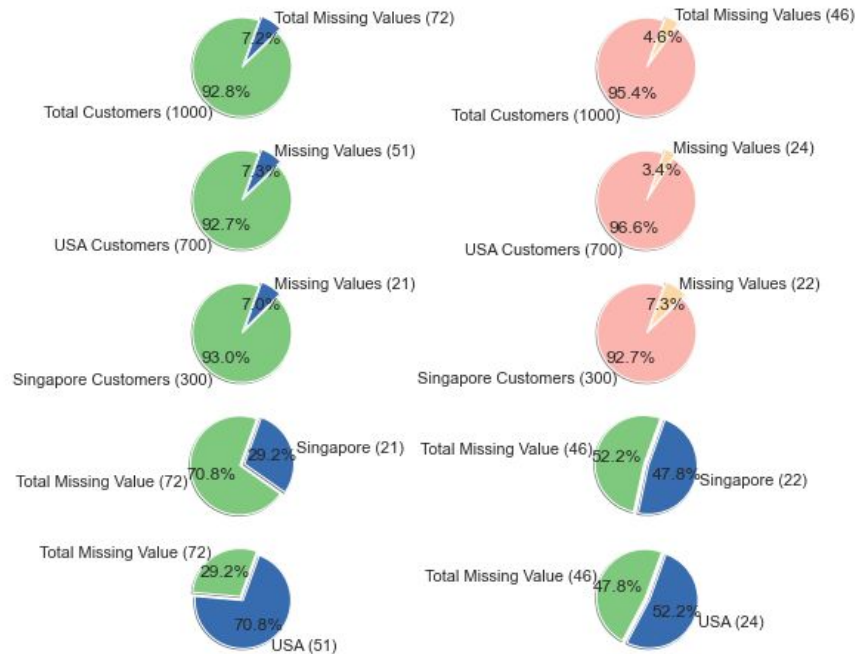
## Strategy in Handling Missing Value

- Assumption to apply mode and mean method to the missing value is that the data is centralized so we try to predict the missing data is close to the centre of available data
- Since subscriber\_type is a categorical data, simply it can be imputed using mode method
- since number\_streamers is a numerical data, simply it can be imputed using mean method

# DATA ANALYSIS - MISSING VALUE

Missing 'subscriber\_type' Values

Missing 'num\_streams' Values



## Missing Value Handling

1. In subscriber\_type columns, there are 72 missing values which are 7.2% of the total data. USA customer missing value at 7.3%, relatively higher in percentage than Singapore missing value at 7%.
2. In number\_subscribers columns, there are 46 missing values which are 4.6% of the total data. USA customer missing value at 3.4%, relatively higher in percentage than Singapore missing value at 7.3%.

# DATA ANALYSIS - CHURN INVESTIGATION



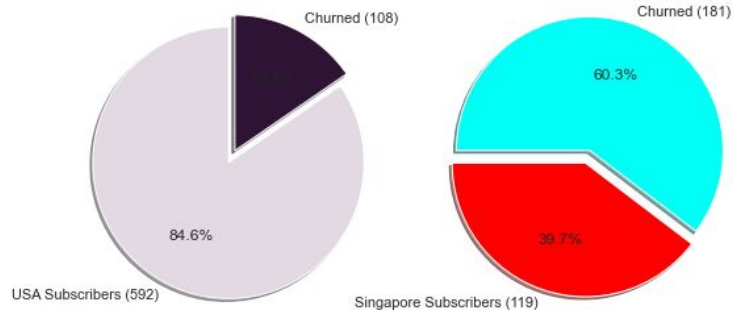
## Correlation

1. Weak correlation between age and churn
2. weak and inversely proportionate relationship between number of streams and churn
3. negative correlation between age and number of streams

# DATA ANALYSIS - CHURN INVESTIGATION

USA Market

Singapore Market

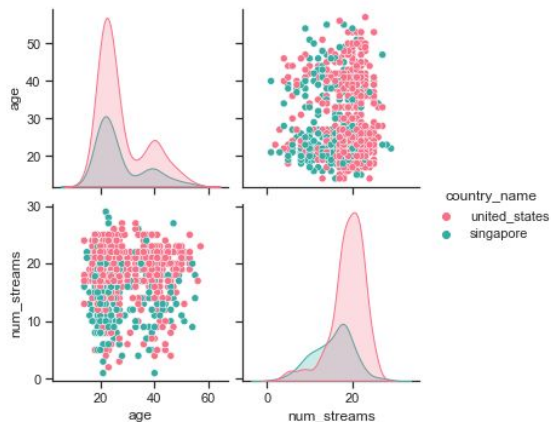


## Churn

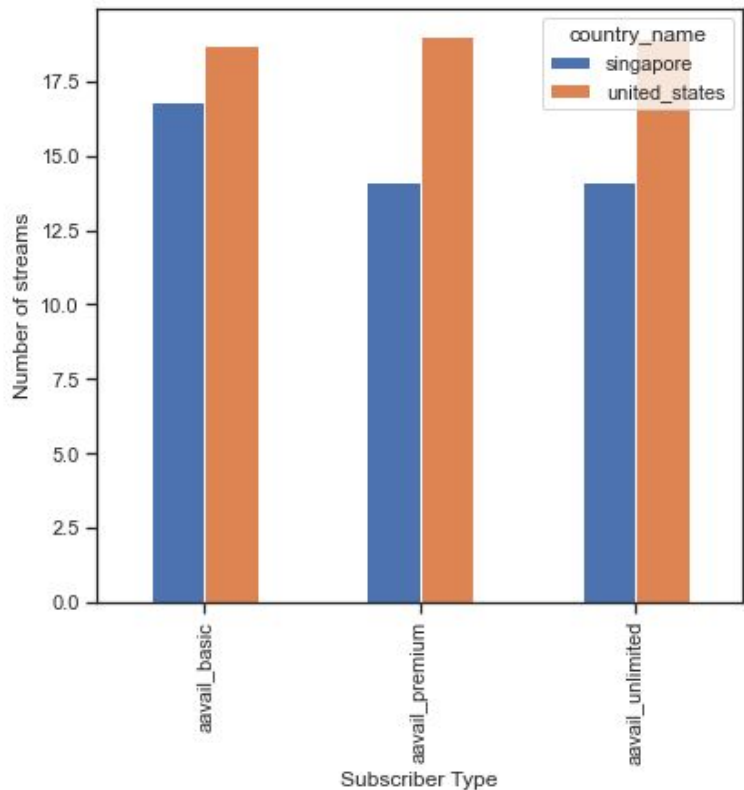
1. Total churn rate is 24%
2. Singapore reaches 49% (quite higher than the rest) and USA only 15%

## Correlation

1. Churn and age has no significant correlation
2. There is a slight increase in churn when number of streamer reach certain value (~17-18)



# DATA ANALYSIS - CHURN INVESTIGATION



## Churn

1. Overall, USA perform higher in all categories of subscriber type than Singapore
2. Based on the data, more popular subscriber type in Singapore is basic.



# CONCLUSION

1. By referring to age, the correlation is higher at 20-40 years. Avail should focus on marketing that targets this segment: social media
2. Investigate further data to understand the high churn rate in Singapore. It can be on pricing, or loyalty program
3. Focus more on the basic subscriber type, and reduce either premium or unlimited type (or merge them) to make segmentation clearer

A horizontal bar at the top of the slide, composed of a small blue rectangle on the left and a long grey rectangle extending to the right.

**THANK YOU**

A thin red horizontal line positioned below the "THANK YOU" text.