

# Capstone Session - 1

## Marketing Campaign Segmentation

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Agenda

- 1) Marketing Campaign Segmentation Problem Statement
- 2) Solution Steps Walkthrough
- 3) General Best Practices
- 4) Q&A

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Marketing Campaign Segmentation Dataset

- The dataset consists of 2240 rows and 27 columns.
- Out of 27 columns, 24 columns are numeric and 3 columns are of object data type.
- There are a few missing values in the Income variable.
- The ID column is an identifier which is unique for each customer.

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Problem Definition and Objective

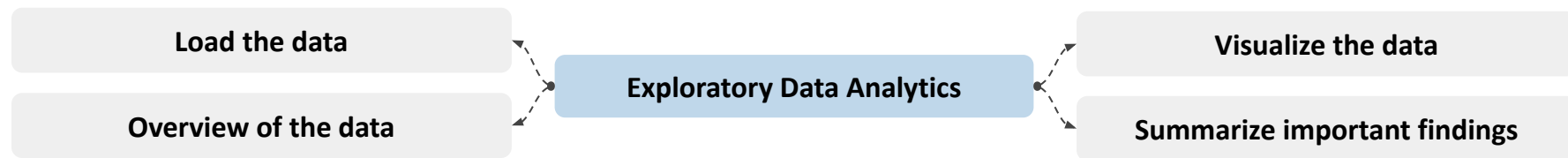
- Customer segmentation is the process of dividing a dataset of customers into groups of similar customers based on certain common characteristics, usually for the purpose of understanding the population dataset in a better fashion.
- Understanding customer behavior and characteristics is usually a critical part of the marketing operations of any business or organization, with direct consequences on sales & marketing strategy.
- Customer segmentation is often viewed as a means to achieve a better return on investment from marketing efforts, and make organizations more efficient in terms of utilizing their money, time, and other critical resources in custom marketing strategies for different groups of customers based on their unique needs and motivations.
- Using Unsupervised Learning ideas such as Dimensionality Reduction and Clustering, the objective is to come up with the best possible customer segments using the given customer dataset.

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Solution Step 1



Example of questions that can be answered by EDA:

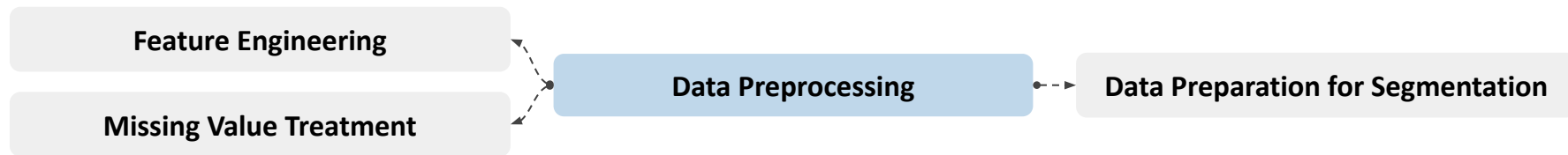
1. Is there any missing values in the dataset?
2. Explore the summary statistics of the dataset?
3. Are there any extreme values in the dataset?
4. Is there skewed distribution in the dataset?
5. Why should we drop the identifier column?

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

## Solution Step 2



Example of questions that can be answered by Data Preprocessing:

1. Why do we need to do feature engineering?
2. Should the missing values need to be imputed or dropped?
3. Why should we drop irrelevant columns before performing segmentation?
4. Is there a need to scale the data?

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

## Solution Step 3

Applying T-SNE and PCA on the data

Model Building

Cluster Profiling

Example of questions that can be answered by Model Building:

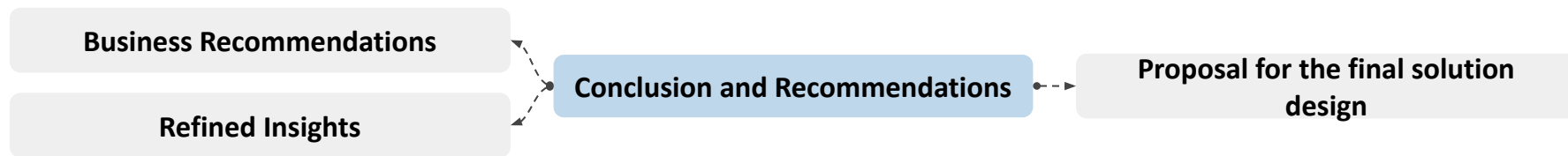
1. What are the different clustering methods that can be applied to solve this problem?
2. Why should we apply clustering algorithm on the PCA data?
3. Which method can be used to determine the optimal number of clusters?
4. Why do we need to profile the clusters?

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

## Solution Step 4



Example of questions that can be answered by Conclusions and Recommendations:

1. What are the refined insights from EDA and model building?
2. What observations and insights can be drawn from the measure of success?
3. How do different techniques perform? Which one is performing relatively better? Is there scope to improve the performance further?
4. What is proposal for final solution design? Why is this the best solution to adopt? What are expected benefits and costs (assume numbers) of this solution design?

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.



# General Best Practices

Some of the best practices for submission:

- **Address all key questions in the rubric:** Make sure to read the rubric carefully and understand all the requirements. Address all the key questions asked in the rubric in your submission.
- **Provide observations and insights:** Provide observation and insight for every important output, such as cluster profiles, summary statistics, missing values detection and treatment. This will help to make your work more understandable and actionable.
- **Explain your design steps:** Explain the steps you took to design your solution approach. This will help the reader to understand the overview of your solution approach and how you arrived at your final model.
- **Comparing the model performance:** To compare the model performance before and after applying dimensionality reduction techniques will help in understanding the importance of dimensionality reduction techniques and choosing the best model.

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.

# Q&A

This file is meant for personal use by jacesca@gmail.com only.

Sharing or publishing the contents in part or full is liable for legal action.

Proprietary content. © Great Learning. All Rights Reserved. Unauthorized use or distribution prohibited.