

# Analyze Customer Segmentation for Marketing Campaigns

## Introduction

In this project, **we want to identify customer segments for a mail-order sales company in Germany** using unsupervised learning techniques. By comparing clusters of demographic data and its customer data, we find out the characteristics of its customers. Since the segmentations include the gender, personality, and financial situations of customers, **the company can target potential audiences in their marketing campaigns**, from content creation to marketing channels.

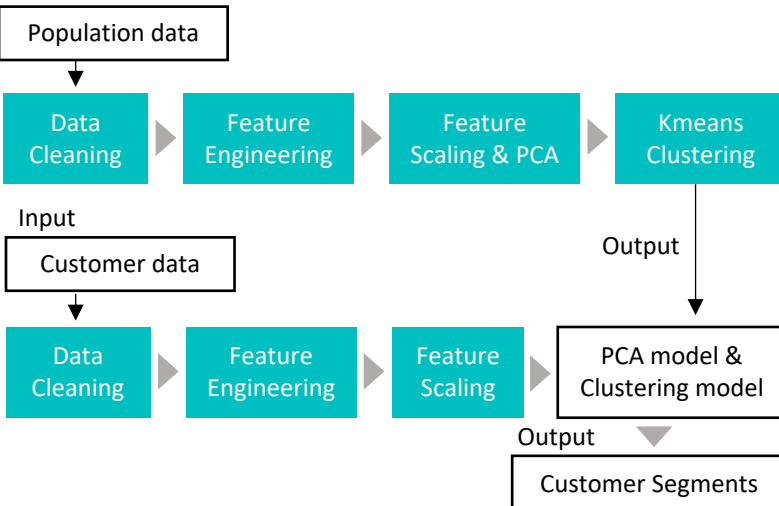
## Data

1. Population data: demographics data for the German population
2. Customer Data: demographics data for customers of the mail-order company

The variables include gender, buying channels, personality, life stage, household type, income, social status, health, neighborhood features, etc.

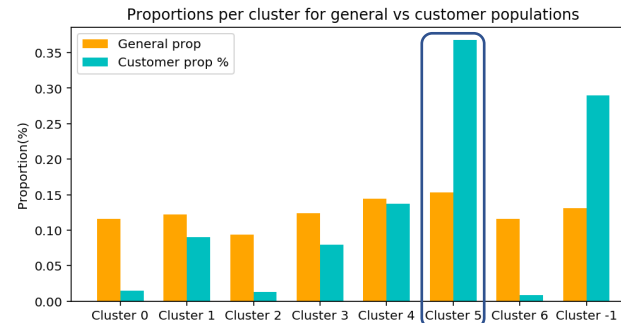
## Methodology

### Input



Note: The PCA model condenses 68 variables into 25 components.

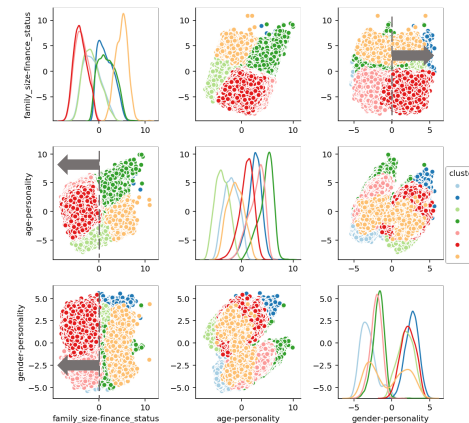
## Identify customer segments



Note: Cluster -1 is missing data for both; therefore, we don't consider them.

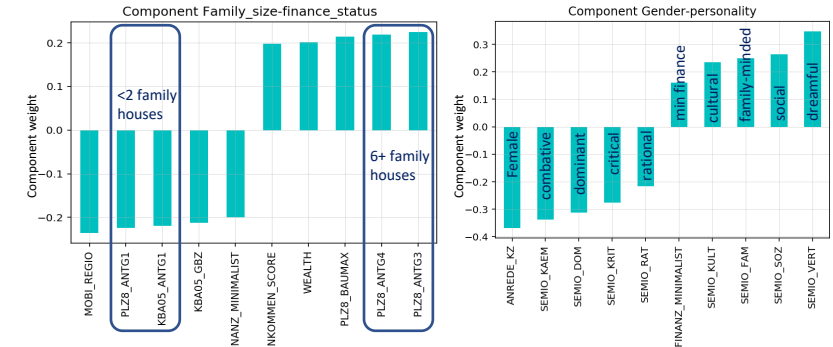
Using the same clustering model, since the proportion of customer population in Cluster 5 is higher than the proportion in the general population, we hypothesize that people with Cluster 5 characteristics are more likely to be our customers.

## Investigate the characteristics of the segment – Cluster 5



Cluster 5 (red) seems to be strongly associated with negative values of the **family\_size-finance\_status** component and positive values of the **gender-personality** component.

## Analyze the target customer profile through component analysis



Note: Wealth - 1: Wealthy Households, 5: Poorer Households.

Note: The personality scale is highest affinity = 1, lowest affinity = 5.

We analyze the significant variables positively and negatively associated with the **family\_size-finance\_status** component and **gender-personality** component.

- **family\_size-finance\_status** component: Variables with positive values describe people living in a high-density areas with lower income. In contrast, strong **negative values indicate low-density areas with more financial affluence**.
- **gender-personality** component: **The positive values describe males that are less dreamful, less social, less family-oriented, and conservative return shopping type.**

## Conclusion: Customer profiles – 35% of customers



Affluent households in low density area (1-2 family houses in the PLZ8 region)



Practical males who don't social much