The Club Q8 Challenge

QTEM Data Challenge Fall 2021

FJS TEAM

15.01.2022

OUR TEAM

Fabian Pfeffer

Technical University of Munich

Sara Luppi

Tilburg University

Joao Machado

University of Porto

TABLE OF CONTENTS

Introduction

01

This study aims at improving Q8’s loyalty program in Italy

- By recognizing patterns through classification and regression techniques
- By using the results of the analyses to support managerial decisions concerning customer satisfaction and retention

RESEARCH QUESTIONS

The overall objective is to increase the number of active members, both by attracting new members and by keeping current members.

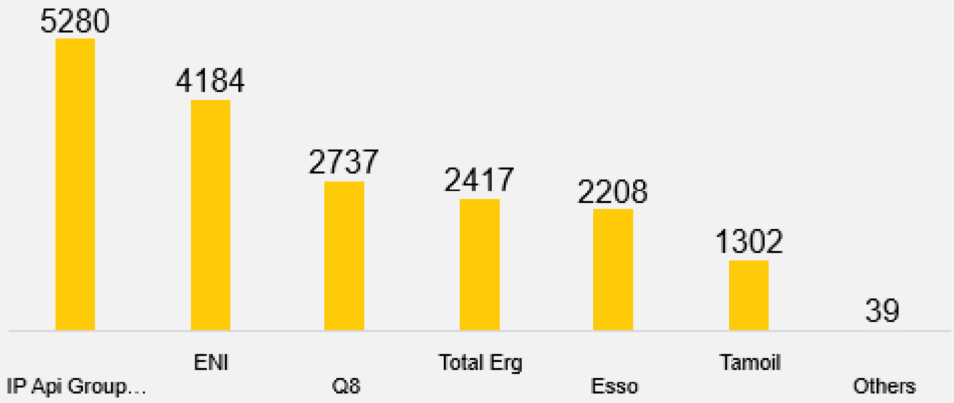
How can Q8 segment the members of the loyalty program?

Which customers are more likely to churn?

Business Background

02

Number of Petrol Stations in Italy in 2021



Competitors use loyalty programs that offer discounts and prizes too.
E.g., ENI has You&Eni and Total Erg has Box Più.

LITERATURE

CUSTOMER REFUELING BEHAVIOR & LOYALTY PROGRAMS

What do customers value about gas stations’ loyalty programs (in the Russian market)?
“To have a loyal customer, gas stations [‘ loyalty programs] have to have **quality** products, namely, fuel and store products. Second, the **price** should be competitive. Third, cleanliness at gas stations.“

“The study demonstrated that the **access** to station from both directions, road barricades in direction of station, to be located on a **local or state** road, and the **speed limit** on the front road have been the major factors for the gas station site selection.“

“Loyalty was significantly affected by consumer **satisfaction**, **involvement**, **frequency** of gasoline purchases, information route and customer's occupation. “

DATA DESCRIPTION

Program Members

[redacted]

Master data

Contains columns such as:

- *Member ID,*
- *Date of birth,*
- *Amount of points,*
- *etc.*

Claimed Prizes Transactions

Transactional data

Contains columns such as:

- *Date of transaction,*
- *Price category,*
- *Cost of price in points,*
- *etc.*

[redacted]

Refueling Transactions

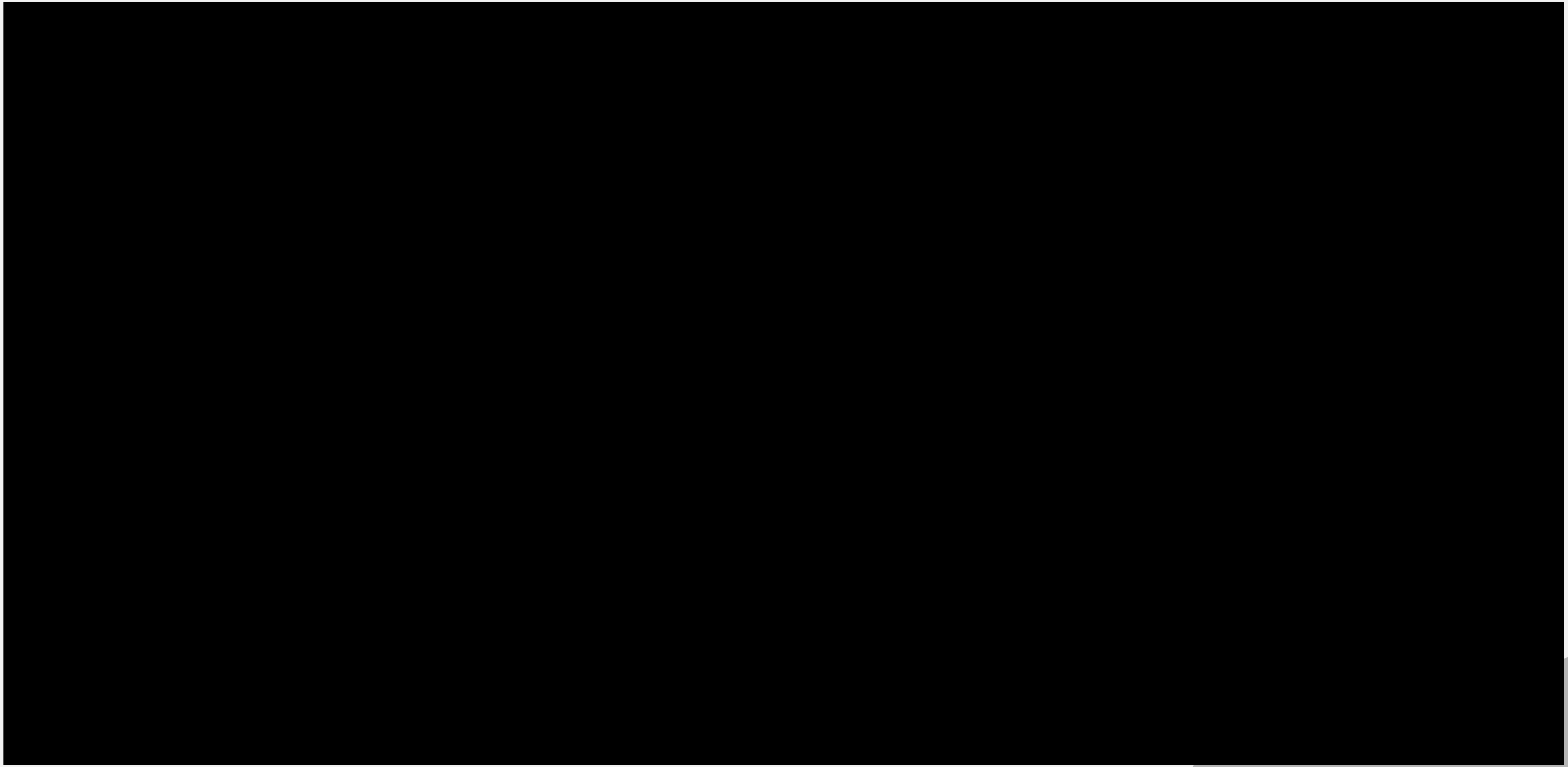
[redacted]

Transactional data

Contains columns such as:

- *Date of refueling transaction,*
- *Type of product,*
- *Received points,*
- *etc.*

DATA DESCRIPTION



DATA DESCRIPTION

CLAIMED PRIZES

500 k claimed prizes

230 k members that reedemed

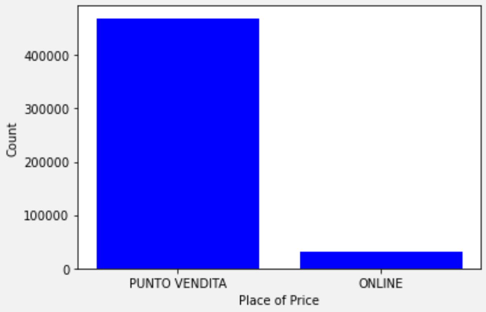
Redemption per prize

DESCRIZIONE	count
BUONO SCONTO CARB 4E 430 PUNTI	88.550
BUONO SCONTO CARB 4E 860 PUNTI	70.302
BUONO SCONTO CARB 4E 1290 PUNTI	66.117
BUONO SCONTO CARB 4E 2580 PUNTI	47.553
CONSEGNA PASTA GAROFALO	30.722
PRENOTAZIONE BICCHIERE GUZZINI 4PZ	15.623
CONSEGNA CODICE GIFT AMILON 5?	14.970
PRENOTAZIONE MASCHERINA TOMBOLINI	8.342
PRENOTAZIONE PHILIPS RASOIO	6.823
PRENOTAZIONE MISUR PRESSIONE	5.309
PRENOTAZIONE KIT 38 UTENSILI	5.232
PRENOTAZIONE C. LETTO HENNE'	5.092
CONSEGNA CODICE GIFT AMILON 20?	5.054
PRENOTAZIONE BICCHIERE GUZZINI	4.923
PRENOTAZIONE SET SPUGNE C.I.	4.797

Redemption per group

RAGGRUPPAMENTO_MERCEOLOGICO	count
BUONO SCONTO CARBURANTE	272.660
PER LA TUA CASA	68.450
PER TE	30.569
ELETTRONICA	24.924
TAVOLA & CUCINA	20.407
CASA	18.529
ELETTRODOMESTICI	17.489
PER IL TUO BENESSERE	14.858
TEMPO LIBERO	9.569
MODA	7.237
SPORT	3.438
PER IL TUO AMBIENTE	2.849
PER IL TUO BAMBINO	2.837
PER IL TUO VIAGGIO	2.171
BAMBINI	1.749
AMICI A 4 ZAMPE	1.554
GUARDA & PRENOTA	1.366
BELLEZZA & BENESSERE	698
PER IL TUO CUCCIOLO	60
ONLUS & PROGETTI	35

Place of redemption



DATA DESCRIPTION

REFUELINGS

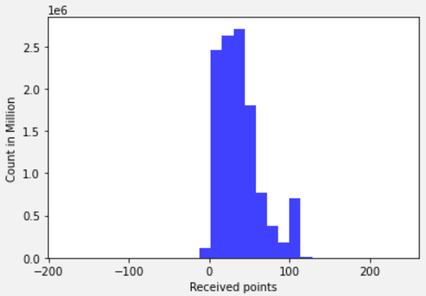
11,8 mln refueling transactions

1 mln members

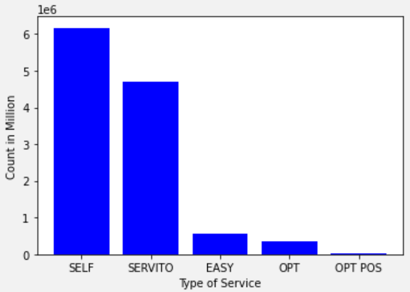
Transactions per product

PRODOTTO	count
CARICO GASOLIO	5.835.363
CARICO BENZINA	4.403.314
CARICO GPL	641.881
CARICO HIPERFORM DIESEL	551.675
CARICO HIPERFORM 100 OTTANI	228.215
CARICO GASOLIO PESANTE	77.482
CARICO METANO	42.962
ANNULLO CARICO GASOLIO	3.084
ANNULLO CARICO BENZINA	2.454
ANNULLO CARICO HIPERFORM DIESEL	625
ANNULLO CARICO GASOLIO PESANTE	267
ANNULLO CARICO HIPERFORM 100 OTTANI	232
ANNULLO CARICO GPL	197
CARICO METANO PESANTE	107
ANNULLO CARICO METANO	4

Points received



Type of service



	LITRES
Min	1
Max	1200
Mean	40.56

Data Analysis

03

FINAL DATASET

870 k rows x 28 variables

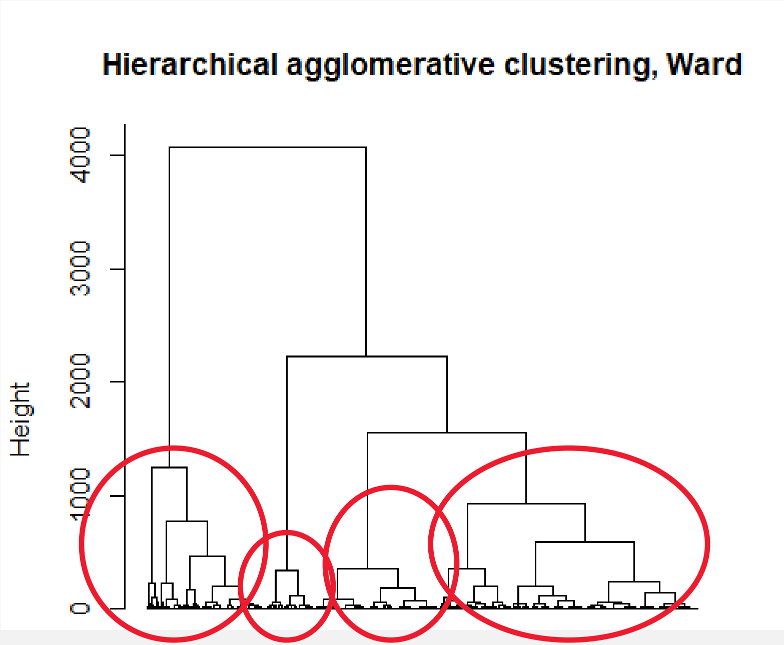
Joined dataset of the previous ones that has been cleaned

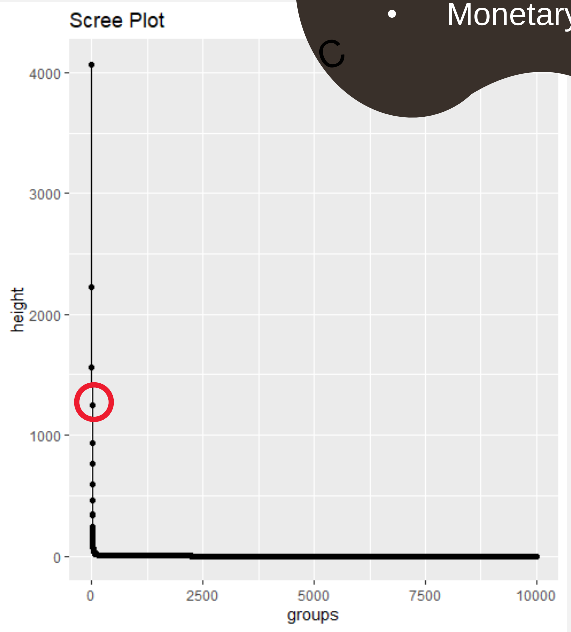
Grouped by COD_PAN_DA_POS:

- Each row is a specific customer
- Variables are summarized

CLUSTER ANALYSIS

HIERARCHICAL AGGLOMERATIVE CLUSTERING





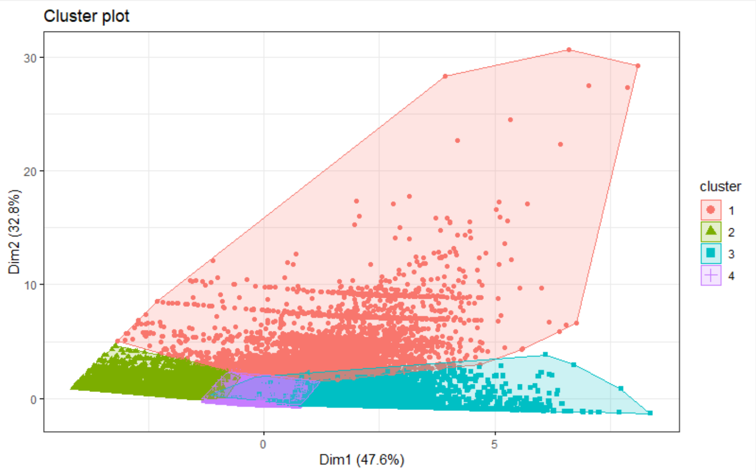
CLUSTERING VARIABLES

- Recency (R)
- Frequency (F)
- Monetary value (M)

4 OPTIMAL CLUSTERS

CLUSTER ANALYSIS

K-MEANS (WITH 4 CLUSTER CENTERS)



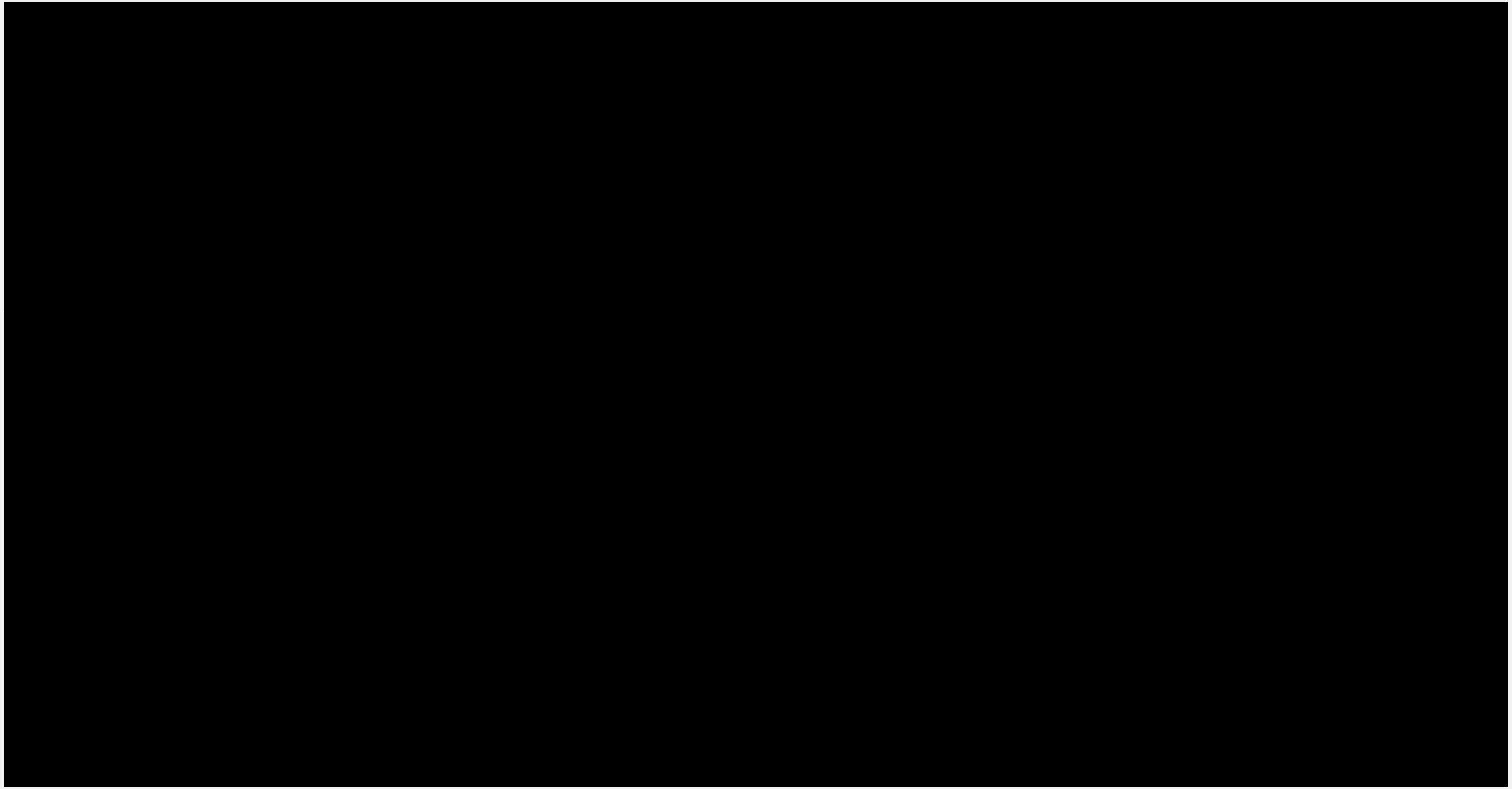
Number of people in each cluster

1	2	3	4
118075	38242	6851	22053

With a dimension reduction technique, we can plot the clusters on 2 dimensions

	Dim1	Dim2	Dim3
[redacted]	-0.337	0.092	-0.937
	0.925	-0.150	-0.347
	0.173	0.984	0.035

VARIABLES DIFFERENCES BY CLUSTER



CLUSTERS OVERVIEW

1) The Normal

- Middle ground
- Does not score high or low on most dimensions

2) The Active

- Drives a lot
- Interacts frequently
- Refuels frequently
- Can afford a lot of prizes

4) The Dissatisfied

- Has churned or is likely to
- Rarely refuels

3) The Money

- Spends significant amounts of money and points on prizes
- Has been in the program for long

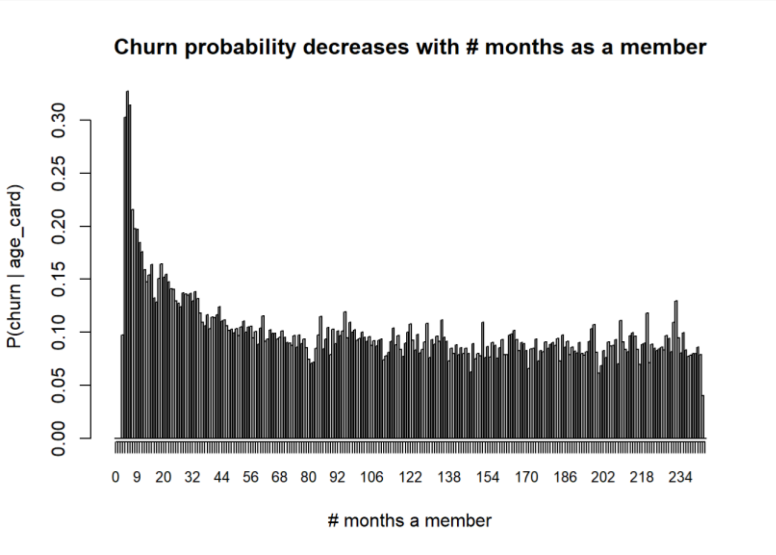
CHURN MANAGEMENT

Predicting if a customer is going to become inactive

churn
0
1

Has made transactions in the past 3 months

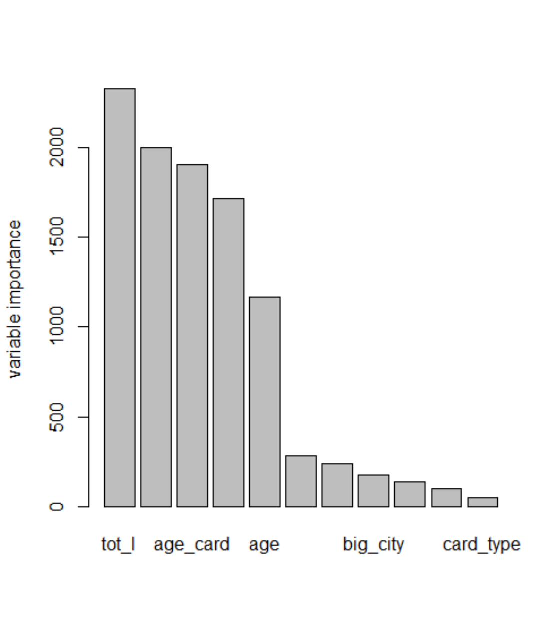
Has NOT made transactions in the past 3 months



Logistic Regression

Balanced Accuracy = 0.51

Random Forests




```
model1.rf <- ranger(churn ~ sex+age+zone+big_city+
                        age_card+point_balance+card_type+
                        redeemed+selling_mode+tot_l+tot_points,
```

Balanced Accuracy = 0.67

Suggestions

04

CLUSTER 2

The **most active group of program members**. They have by far the highest number of refuelings in the observed period (39,6 on average). In addition, they also claimed the most prices (3,43 on average). Very few customers in this cluster are considered to churn and they **tend to be younger** than the customers in the other clusters.

Investigate sources of customer satisfaction

Frequent refuelings and claims of prizes. This cluster **has a high retention rate and a high customer satisfaction.**

To profit from this, we suggest to **investigate the reasons behind it**. This could be done by supplying surveys to this segment. These surveys could ask program members which features of the program are most valuable to them or why they chose Q8 over competing loyalty programs.

Never change a winning team - **Do not spend additional budget on this segment**

They are already very active. It is essential that these customers are retained and are not deterred from continuing to use the loyalty program. However, it **may not be worth to spend more resources than necessary** since they are already using the program actively.

We therefore propose to **not implement major changes** and **to invest the budget in other clusters**. Since members of this segment are already satisfied, they **will likely not appreciate major changes to the program**. Prizes or prize categories that are popular within this cluster should be maintained.

CLUSTER 3

Spends most points and money on prizes. Although the cluster is comparably small, its members are **willing to spend more money** on the loyalty program and therefore may have a higher profitability than customers who only use points to buy prices.

Tailor custom advertisement to specific preferences of the cluster

They save up their points and invest them into prizes that are comparably expensive. On average, they spend roughly 50 € for their prizes in addition to the required points.

Members of this segment can be **provided with custom advertisemen**t that **shows them the prizes in the higher cost segment.**

Since the cluster seems to save up their points for this type of rewards, they can be **motivated to use the program more actively and continuously** if they are provided with information of the prizes they are interested in.

CLUSTER 4

51% inactive members. The similarity of the behavioral variables in this cluster implies, that the still-active members may also be dissatisfied with the program.

To improve the loyalty program, it is important to increase the customer satisfaction and **win these members back.**

Increase number of points

This cluster has a **low mean number of points** compared to the other clusters. This prevents them from claiming prizes that require a higher number of points.

By conducting a **one-time giveaway of a fixed amount of points**, these customers are reminded of the program and can claim a price that they were previously not able to afford.

Give discounts for popular prices

Cluster 1, 2, and 3 show a lower churn rate (close to 0) than cluster 4 (51%). However, Cluster 4 and Cluster 1 have claimed a similar number of prizes on average (1.84 vs 1.66).

By **offering discounts on** the quantity of points or the monetary contribution for **prizes that are popular** with cluster 1, 2, and 3, members of cluster 4 can be enabled to afford prices that the other, more satisfied customers have chosen.

Popular prizes that can be used in this approach are e.g., **"Per la tua casa" or "Per te"**

LIMITATIONS

- Grouping the observations and summarizing the variables causes some loss of information

- HAC cannot be performed on a large dataset, so we used a random sample from the data.

Still, the K-means results have been influenced by that.

- The churn variable is unbalanced, some models do not yield good predictions because of this

THANKS!

Appendix

REFERENCES

Nechipurenko, D. (2019). Value of customer loyalty programs for consumers.

Semih, T., & Seyhan, S. (2011). A multi-criteria factor evaluation model for gas station site selection. evaluation, 2(1), 12-21.

Lee, A., Huh, E. J., & Jeon, H. R. (2012).
A Study on the Effects of Consumer Satisfaction on Loyalty According to Involvement Focused on the Gas Station Service. Korean Journal of Human Ecology, 21(2), 241-256.

Grouped Dataset

Demographical variables

[redacted]

Membership variables

[redacted]

Prize redemption variables

[redacted]

Refueling variables

[redacted]

Cluster	Age	Points	Age of Card	Max points for price	Contribution in €	Nr. of Prices	Total litres	Total points	Nr. of refuelings	Last refueling in days	Churn indicator
1	54,4	1521	2379	1186	1,94	1,84	553	547	13,3	15,8	0
2	50,4	2837	2048	1464	2,84	3,43	1709	1676	39,6	8,03	0,00143
3	51,3	3782	2710	4906	50,7	2,67	1300	1290	25	15,7	0,0211
4	52,5	844	2115	1031	1,66	1,66	196	193	4,7	99,1	0,51