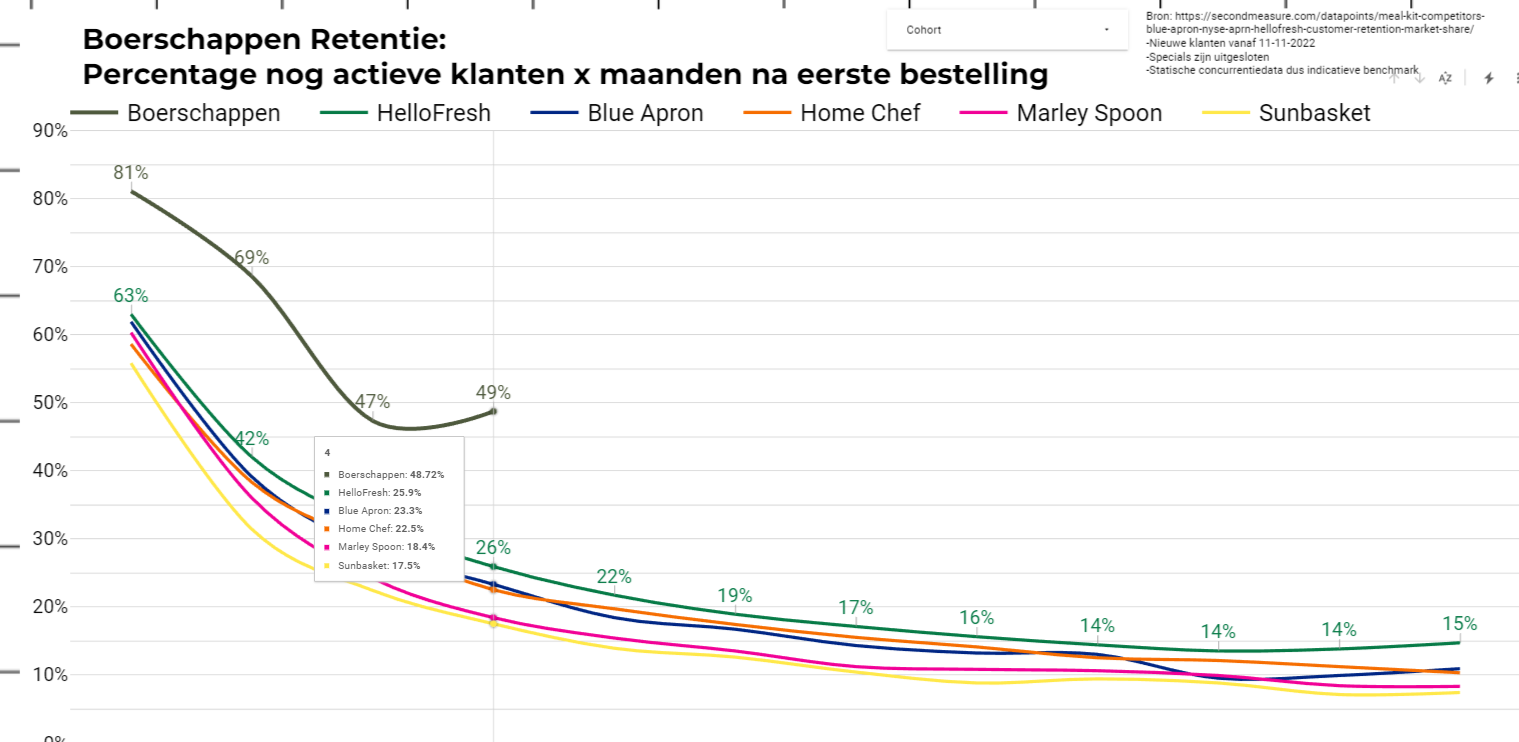
## Situation

I have a bigquery db with ecommerce information. It is a food box subscription service. So people subscribe to a foodbox. and get the food delivered

based on order information I want to predict when a customer will leave

Currently they have the following insight with looker.

Here they show the retention rate for boerschappen



Goal: As a business manager I want to see the cancels of consumers in the future. Per month, per week en drilled down to which consumer individually. When I know I’ll be able to communicate with them to avoid cancellation

## 

## Project

This project has two phases

1. analyze data, create the model, test the model and see the results locally with the CSV
2. deploy the model to google, link it to the customer bigquery db and make available the looker dashboard

Next phase

* Predict the other status changes
* Add additional consumer meta data and see if there are any correlations

## File

I have created a CSV to local testing purposes it can be downloaded here

<https://drive.google.com/drive/folders/1cIzlvGR3JkgPoGoVZJwXL_RjDpxzf0tL?usp=sharing>

In the excel sheet the file is explained

<https://docs.google.com/spreadsheets/d/1SpVsy-jo37t_gg3cBSJQOsBAGM_ZwymCF1h24wiW_E4/edit?usp=share_link>

Every event the consumer does an event is stored. There are several events

These are the most important ones

What can we predict as to when a customer

3 cancels the subscription

4 change the subscription

7 pauses

6 reactivates

It concerns when someone 3 cancel a subscription at this time

## Project Detailed Steps

1. analyze data, create the model, test the model and see the results
   1. 1: My target variable will be =the time a customer cancels their subscription.
   2. 2: potential predictors that may be useful in predicting the target variable. This could include factors such as the customer's order history, subscription details, payment information, and any other relevant information that is available in your database.
   3. 3:Data preparation
   4. 4: Split your data into training and test sets
   5. 5: Select an ML algorithm
   6. 6: Train your model
   7. 7: Evaluate your model
   8. 8: Ready the model and save it
2. deploy the model to google, link it to the customer bigquery db and make available the looker dashboard

## Output Deliverable

* The output of an ML model would typically be a predicted value for the time when a customer is likely to cancel their subscription. This prediction can be represented as a numerical value or a probability score, depending on the specific algorithm used to build the model.
* Looker dashboard.
  + Like the dashboard above in a graph per month: expectation of cancels
  + Drill down per week
  + Dril down which consumer might cancel: table

## Possible other phase

In addition to the predicted churn time, the ML model may also provide insights into the most important features that contribute to customer churn. This can help businesses to identify the key factors that drive customer churn and take proactive measures to retain customers before they cancel their subscription.

But For this i will charge Extra