



How can ACME enter the Spanish health market?



Gianni Miller | Maria Gonzalez | Yannick Van Dam | Göktuğ Aşcı | Zohak Mirza | Joaquin Calderon | Gerardo Gandara
Prof. Jorge Centeno



Table of Contents

01 Introduction

02 Assumptions & Limitations

03 Product Information

04 Traces Analysis

05 Data Quality Analysis

06 Product Health Information

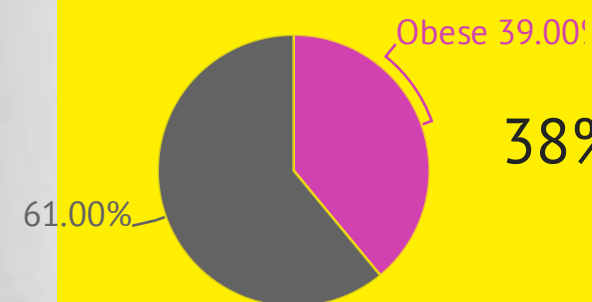
07 Go to Market Strategy

INTRODUCTION

Group E

01

Supermarket chain in the United States of America



38% of the world's population is obese

03

Variety of products both healthy and unhealthy

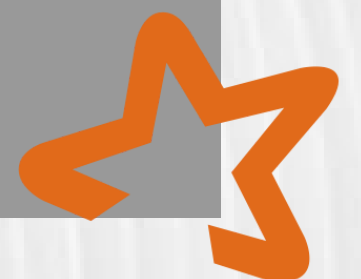
04

Healthy food market has grown to 259 billion a year



05

Number of Centenarians living in Spain 4.3 times higher than the world average



Assumptions & Limitations

Group E

ASSUMPTIONS

- 1. NHS data on health representative for a daily healthy intake
- 2. Food is representative for a healthy life
- 3. Spain is health oriented → good market (mediterranean diet)
- 4. Health education in Europe still in the making
- 5. Chosen variables of interest = representative



LIMITATIONS

- 1. Relatively small data set with many null values
- 2. Messy data set
- 3. Focus on small fraction of attributes



Product Information - Product Statistics

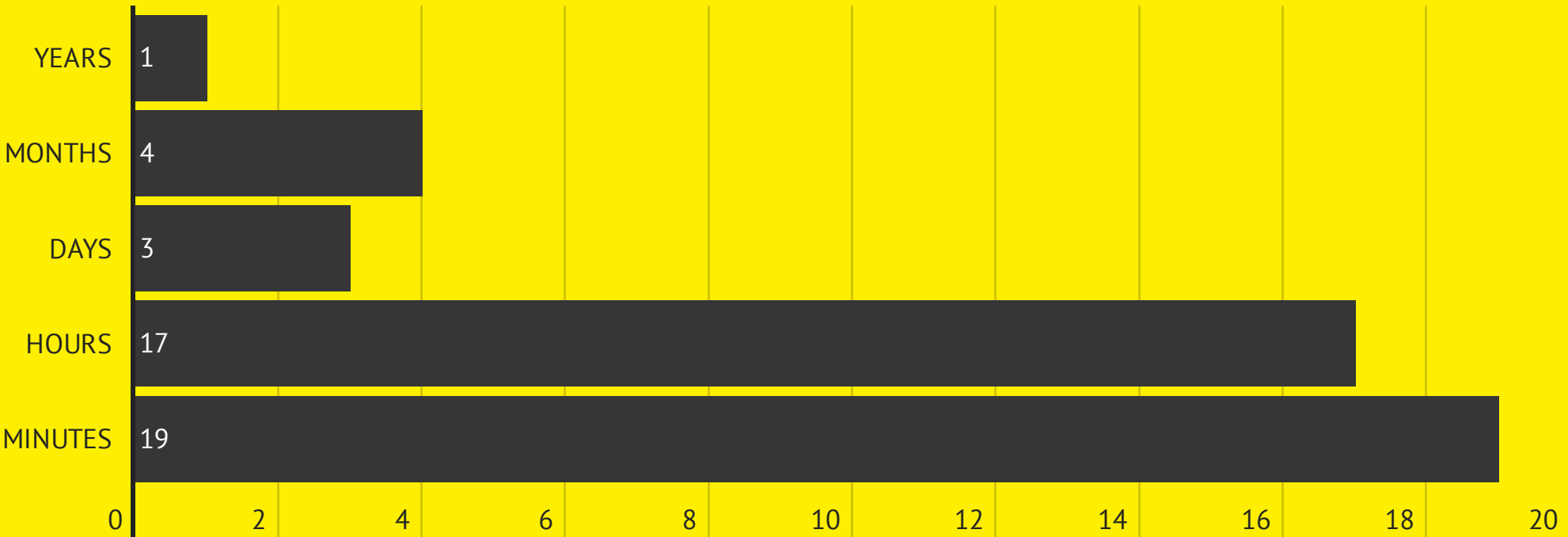
Group E

3
days

9
years











Average shelf life:








Product Information - Product Categories

Group E

Top 5 selling categories

		Main Category	Count
		Cheeses	1219
		Biscuits	1085
		Extra Virgin Olive Oils	1084
		Prepared Meats	884
		Serrano Ham	740

Products by category

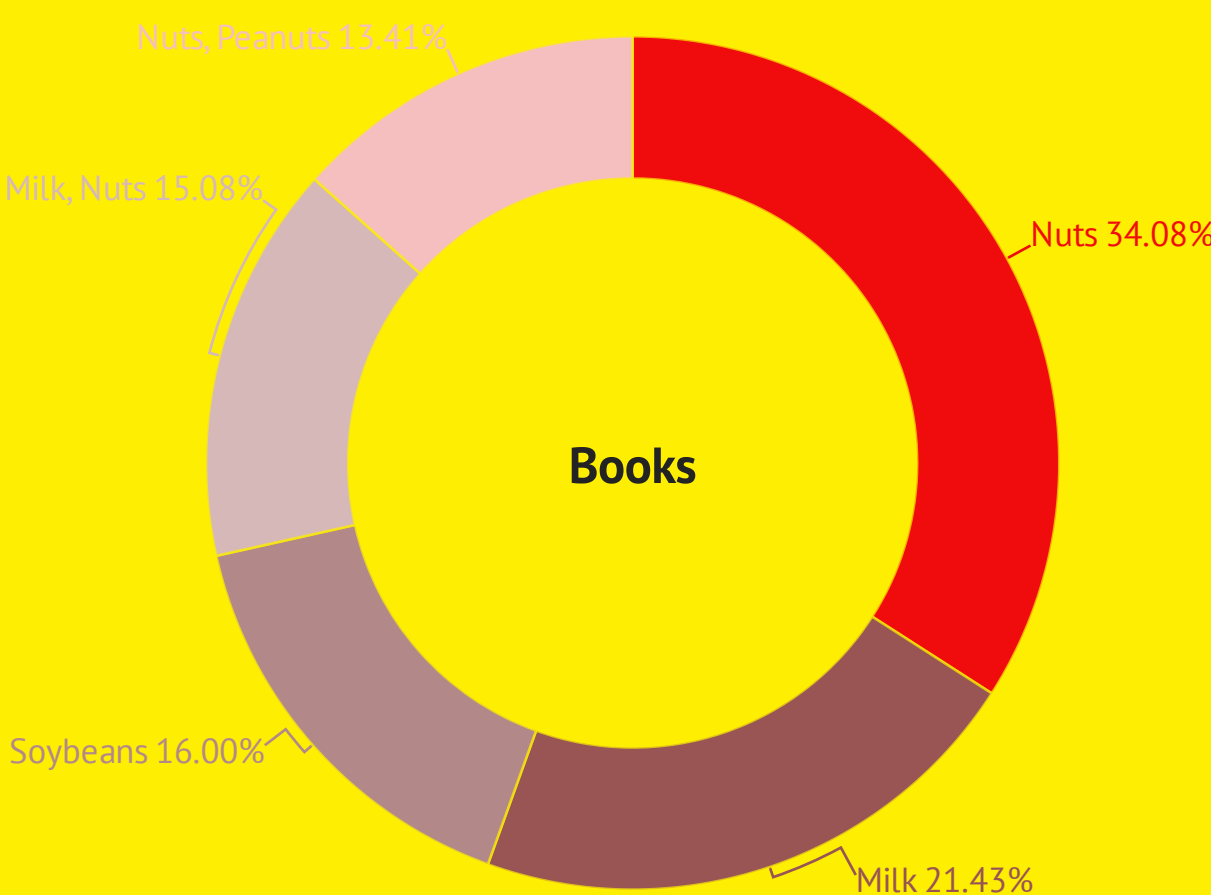
		Products Name	Main Category
		galego	11% white wine
		curcuma rama	3
		miel acacia	acacia honeys
		acacia con trozos de panal	acacia honeys
		panal miel de acacia	acacia honeys








Traces Analysis

Group E

Top 5 product trace information



Product trace examples

	Product Traces	Product Name
	almendra, avellana, nuez-pecan, pistacho	ser chocolate
	balance-diet-and-a-healthy-lifestyle-is-recommended	out unsweetened milk
	caffeina	infuso erbe al limone gr 50
	cebada, contiene-derivados-de-soja-puede-contener	caldo de verduras
	celery	crema de pollo con verduras

96% do not have complete traces info



Data Quality Analysis

Group E

% of products without complete analysis

27%

% of products without additives info

95%

Number of products with complete info

149625

% of products without traces info

96%



Product Health Information - NHS Health Category Classification

Group E



Total Fat

High: 17.5g fat per 100g

Low: 3g fat per 100g



Saturated Fat

High: 5g saturated fat per 100g

Low: 1.5g saturated fat per 100g



Sugars

High: 22.5g sugar per 100g

Low: 5g fat per 100g



Salt

High: 1.5g salt per 100g

Low: 0.3g salt per 100g



Product Health Information - Data Profiling

Group E

```
interestColumnsDF = (  
  df  
  .where(col("fat_100g").isNotNull())  
  .where(col("saturated-fat_100g").isNotNull())  
  .where(col("sugars_100g").isNotNull())  
  .where(col("salt_100g").isNotNull())  
  .withColumn("total_fat_level_100g",  
    when(col("fat_100g")<=3, "low")  
    .when((col("fat_100g")>3) & (col("fat_100g")<=17.5), "normal")  
    .otherwise("high")  
  )  
  .withColumn("saturated_fat_level_100g",  
    when(col("saturated-fat_100g")<=1.5, "low")  
    .when((col("saturated-fat_100g")>1.5) & (col("saturated-fat_100g")<=5.0), "normal")  
    .otherwise("high")  
  )  
  .withColumn("salt_level_100g",  
    when(col("salt_100g")<=0.3, "low")  
    .when((col("salt_100g")>0.3) & (col("salt_100g")<=1.5), "normal")  
    .otherwise("high")  
  )  
  .withColumn("sugar_level_100g",  
    when(col("sugars_100g")<=5, "low")  
    .when((col("sugars_100g")>5) & (col("sugars_100g")<=22.5), "normal")  
    .otherwise("high")  
  )  
  .cache() # optimization to make the processing faster
```



Product Health Information - Health Category Statistics

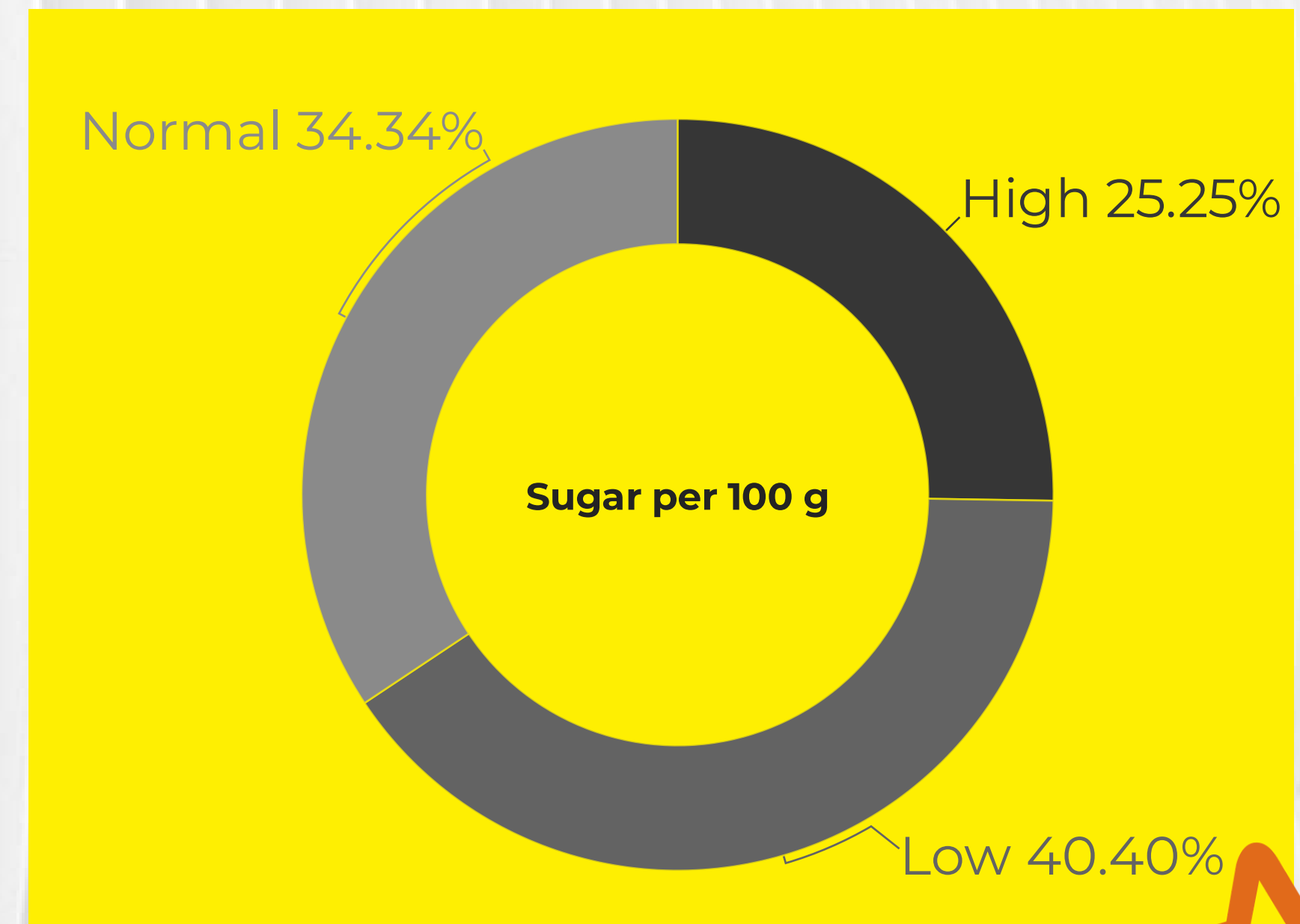
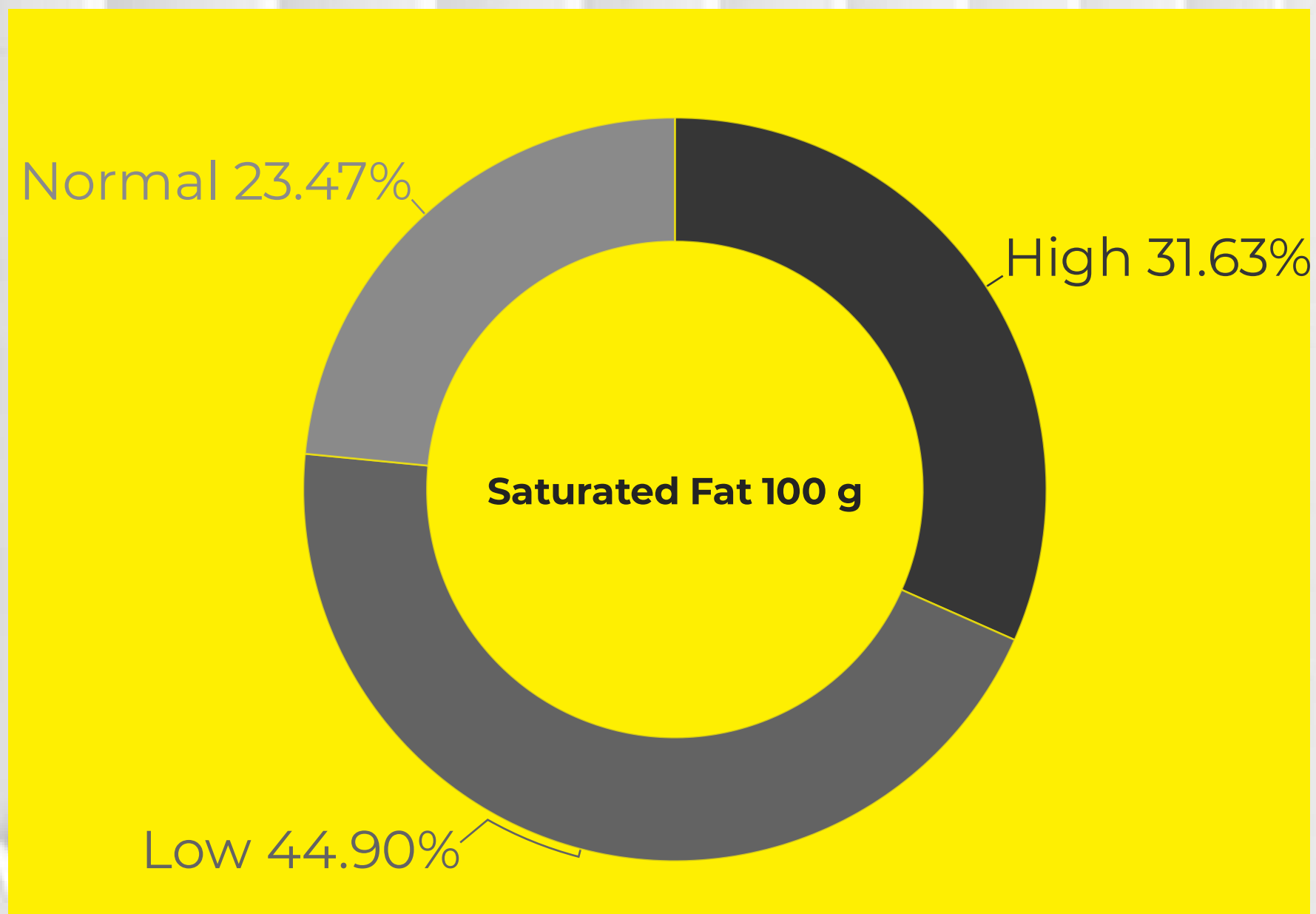
Group E

Saturated Fat 100 g

Sugar per 100 g

Sugar per 100 g

Total Fat per 100 g





Product Health Information

2179

Number of Healthy Products :

Go To Market Strategy

Group E

MEASUREMENTS

- QUESTION:

How tall are you & how much do you weigh?

- GOAL:

Health information

ACTIVITY

- QUESTION:

How many times a week do you do what sports?

- GOAL:

Required kcal & protein intake

TARGET

- QUESTION:

What is your goal?

GOAL:

Target

SOLUTION:

ML + RE algorithms to build perfect weekly baskets
(privacy through Blockchain, anonymising accounts)



Go To Market Strategy

Group E

BASKETS

Composition healthy products

FEEDBACK

What did you like last week?
Cheat day this week?

SUGGESTIONS

“Check out this new basket?”
Select order of last week



APPLICATION



Go To Market Strategy - Marketing

Group E

