PREDICT MEAT PRODUCT CONSUMPTION

PROYECTO FINAL
VARINI-BERRONDO

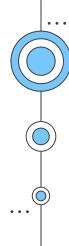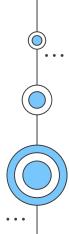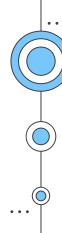# Table of Contents
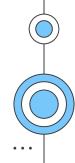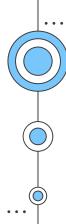
01
SUPERSTORE

Food retail store!

The Superstore has gathered all the information during last year's campaign, such as age, education, marital status, income, and purchase history, etc.
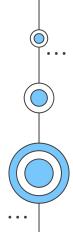
Objective: increase meat sales and decipher consumer behavior in the face of an exciting marketing campaign. How will we achieve it? Through the magic of data.
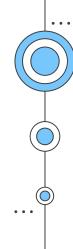
Context: Take an analytical approach that will examine all the factors that influence our meat eaters. We will explore educational levels, ages, genders, household types.
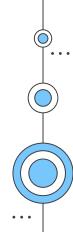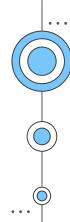
COLUMN	DESCRIPTION	DATATYPE
ID	Unique ID of each customer	Integer
Year_Birth	Age of the customer	Integer
Complain	1 if the customer complained in the last 2 years	String
Dt_Customer	date of customer's enrollment with the company	DateTime
Education	customer's level of education	String
Marital	customer's marital status	String
Kidhome	number of small children in customer's household	Integer
Teenhome	number of teenagers in customer's household	Integer
Income	customer's yearly household income	Decimal
MntFishProducts	the amount spent on fish products in the last 2 years	Integer
MntMeatProducts	the amount spent on meat products in the last 2 years	Integer
MntFruits	the amount spent on fruits products in the last 2 years	Integer
MntSweetProducts	amount spent on sweet products in the last 2 years	Integer
MntWines	the amount spent on wine products in the last 2 years	Integer
MntGoldProds	the amount spent on gold products in the last 2 years	Integer
NumDealsPurchases	number of purchases made with discount	Integer
NumCatalogPurchases	number of purchases made using catalog (buying goods to be shipped through the mail)	Integer
NumStorePurchases	number of purchases made directly in stores	Integer
NumWebPurchases	number of purchases made through the company's website	Integer
NumWebVisitsMonth	number of visits to company's website in the last month	Integer
Recency	number of days since the last purchase	Integer
Response	1 if customer accepted the offer in the last campaign, 0 otherwise	Integer
Complain	1 if the customer complained in the last 2 years	Integer

02

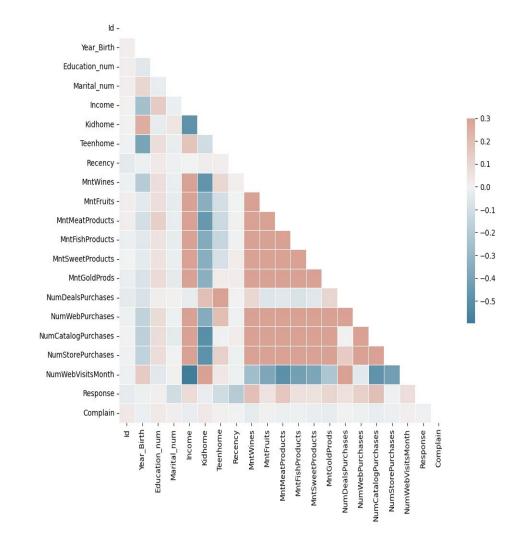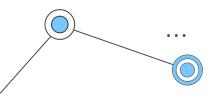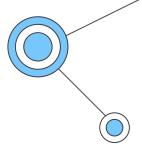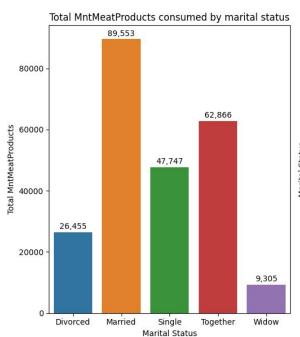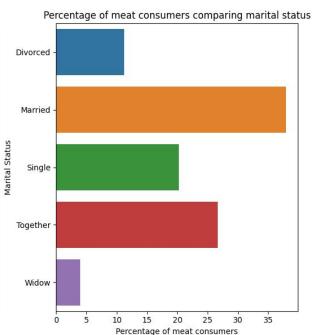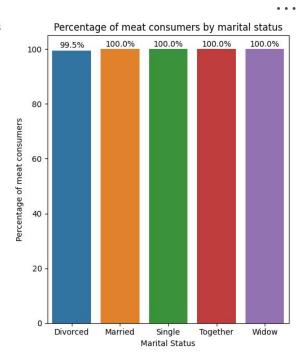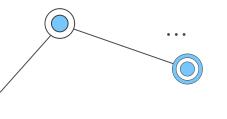## EXPLORATORY DATA ANALYSIS

Correlation analysis

1 **What are the marital states that consume more meat and its percentage?**

Total MntMeatProducts consumed by marital status

89,553

80000

62,866

60000

47,747

Total MntMeatProducts

40000

26,455

20000

9,305

0
Divorced   Married   Single   Together   Widow
Marital Status

Percentage of meat consumers comparing marital status

Divorced

Married

Marital Status

Single

Together

Widow

0   5   10   15   20   25   30   35
Percentage of meat consumers

Percentage of meat consumers by marital status

100   99.5%   100.0%   100.0%   100.0%   100.0%

80

60

Percentage of meat consumers

40

20

0
Divorced   Married   Single   Together   Widow
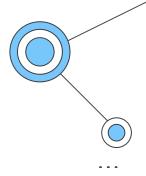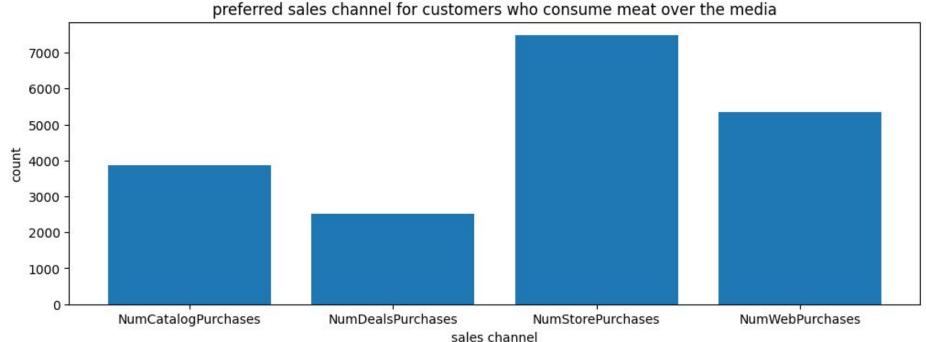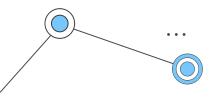Marital Status

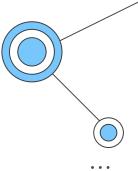**2** What is the preferred sales channel for customers who consume meat and are over the mean value?

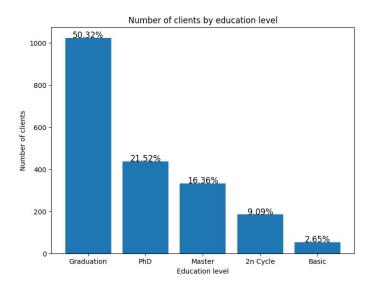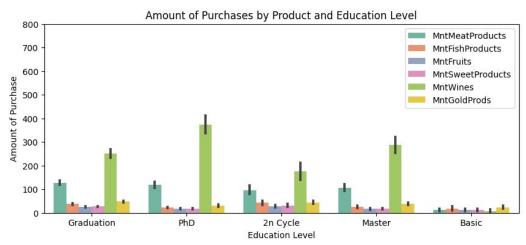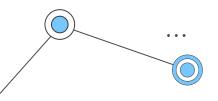preferred sales channel for customers who consume meat over the media

3 What kind of education do customers who consume meat have?

Number of clients by education level

Amount of Purchases by Product and Education Level

Legend:
- MntMeatProducts
- MntFishProducts
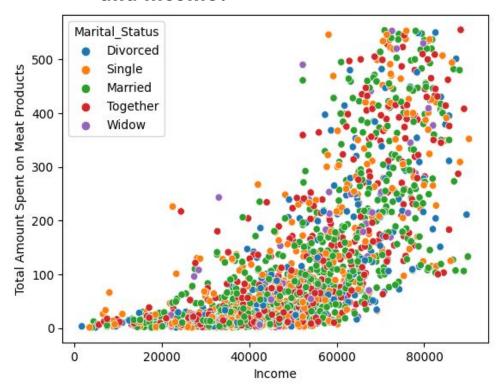- MntFruits
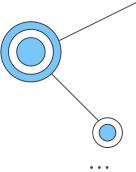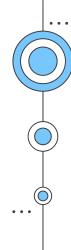- MntSweetProducts
- MntWines
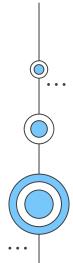- MntGoldProds

4 What relation is between meat consumption and income?
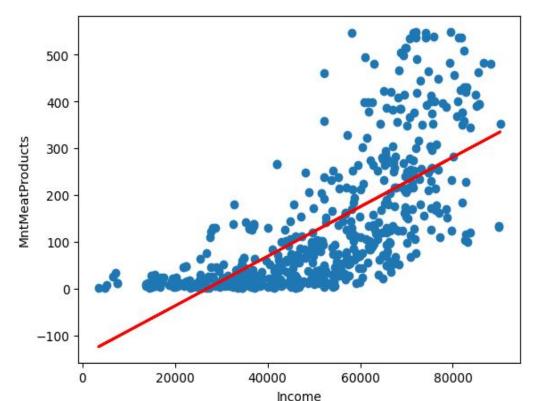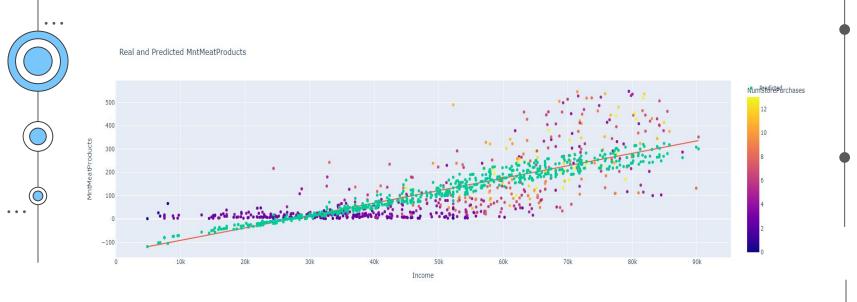
03

MODELING

Trying to predict
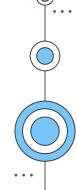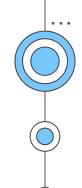
LINEAL REGRESSION

We are going to work with a linear regression model because it establishes a linear relationship between the dependent variable (meat consumption) and the independent variables (characteristics or predictor variables).

If meat consumption is expected to increase or decrease proportionally to certain predictor variables, the linear regression model is appropriate.
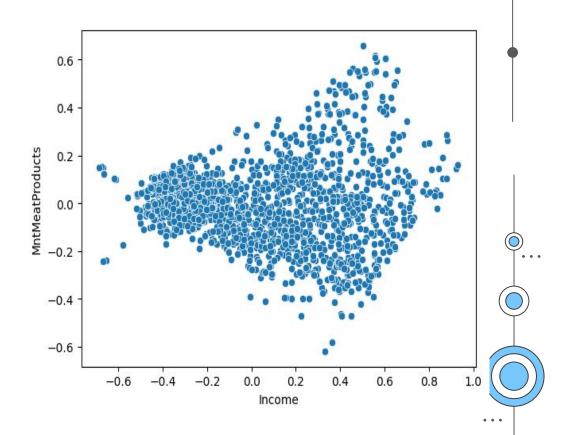
Real and Predicted MntMeatProducts

The second model, which incorporates both income and the number of store purchases as predictors, demonstrates superior performance in predicting meat product consumption compared to the first model that solely relied on income.
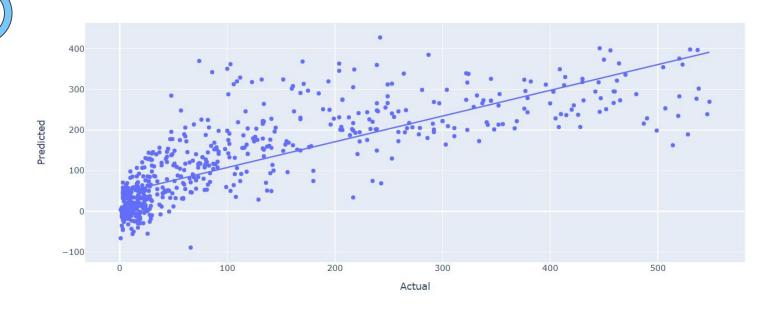
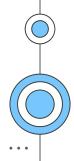# PCA (Principal Component Analysis)

The PCA results provide insights into the variance explained by each component and the cumulative variance explained by the selected number of components.

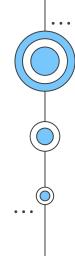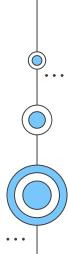Actual vs Predicted MntMeatProducts

Based on these results, we can conclude that the linear regression model, trained on the principal components obtained from PCA, demonstrates a reasonable level of predictive performance.

04

CONCLUSIONS

Our insights

01 Higher incomes and who made more purchases consume more meat products.

02 Higher education levels showed greater preference for meat products.

03 Correlations between meat and income, number of purchases, and educational level

04 Linear regression models had moderate predictive power

Conclusion

In conclusion, understanding the relationships between marital status, sales channels, education, and income provides valuable insights for marketing strategies aimed at increasing meat sales.

Targeting married individuals, leveraging effective sales channels, and considering income levels can help optimize marketing campaigns and product offerings to cater to the preferences and behaviors of the target audience.

Thanks!