Advanced Marketing Analytics

Luca Sangiovanni, Tommaso Premoli & Jan Philip Richter

TOPICS

Data Driven Attribution Model	DDA vs Last Click Model, Comparison & Considerations
Market Basket Analysis	Apriori Algorithm, Frequent Itemsets & Association Rules
CLTV	Current and future CLTV, RFM

DATA DRIVEN ATTRIBUTION MODEL

Data-Driven Attribution Model (DDA) …

- This model is used to identify the amount of credit each of the different marketing channels should receive for having generated a conversion
- DDA aims to provide a more accurate representation of the impact of each marketing channel.
- An attribution window of conversions is calculated. We will use a time period of 30 days. So, only visits made by the customer 30 days before the conversion will be taken into account.

Dataset	
fullVisitorId	Visitor ID
date	Date of the visit
visitStartTime	Timestamp of the visit
channelGrouping	Channel grouping associated with the session
utm_source	Traffic source
utm_medium	Traffic medium
transactionRevenue	Revenue generated from the transaction
transactionId	Unique identifier for each transaction

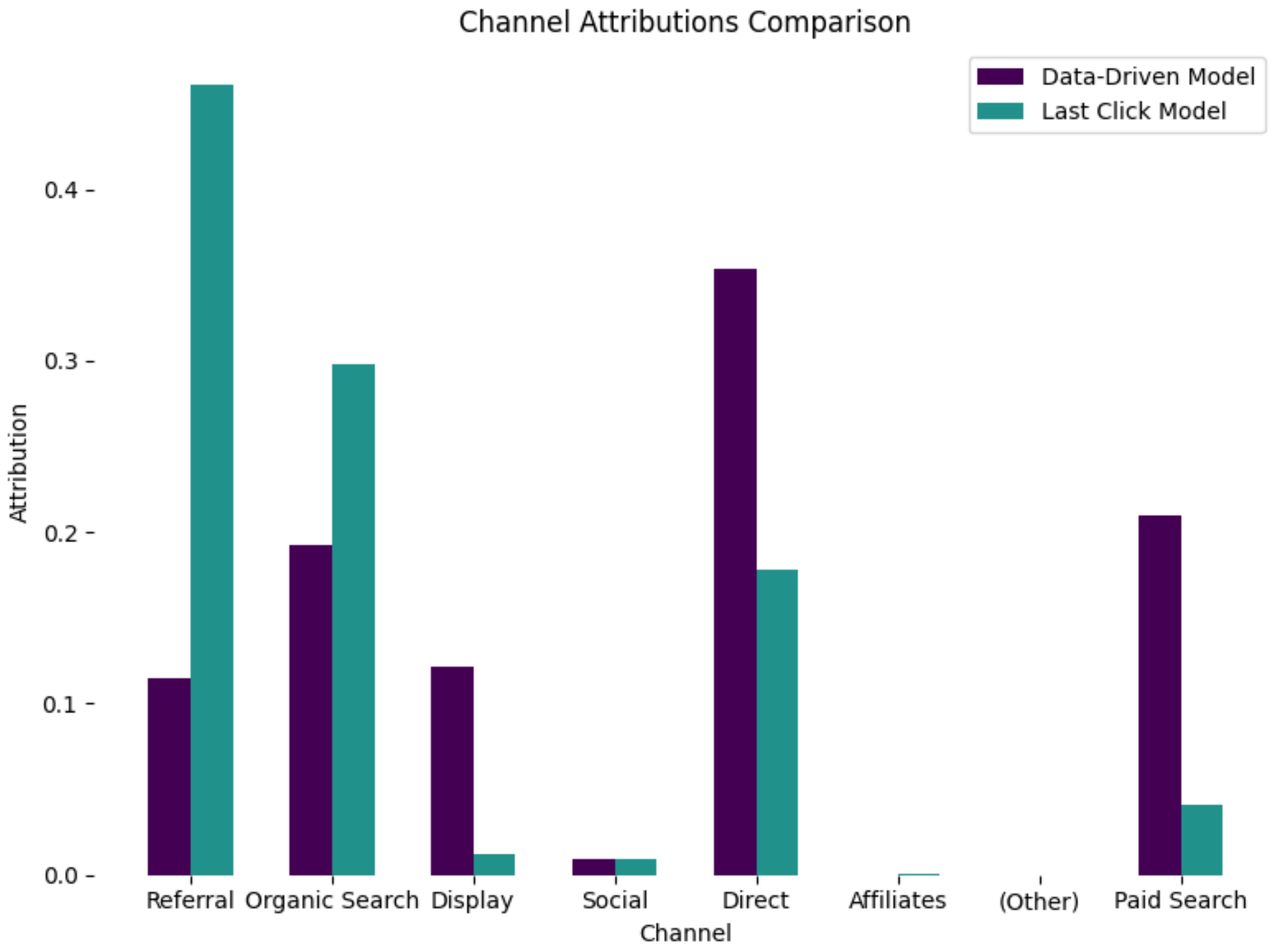
… vs Last Click Model

- The Last Click Model is used to identify the last customer marketing channel in the buyer's journey before conversion
- This model gives 100% credit to the last ‘click’ made by the customer, that is the last marketing channel used

Our <u>objectives</u> are:

- To build a data-driven attribution model for assigning credits to channels related to total conversions
- To compare the weights of our model with the Last Click attribution model

Comparison of the models & Considerations



- The DDA allocations are more equally distributed among the different channels according to their actual contribution to conversion
- Some channels in which more is invested according to the Last Click model are taken into account less for the DDA model such as referrals and organic search
- Even less relevant channels receive a share of the attribution if they have contributed in some way to customer acquisition like display and paid search

MARKET BASKET ANALYSIS

APRIORI ALGORITHM

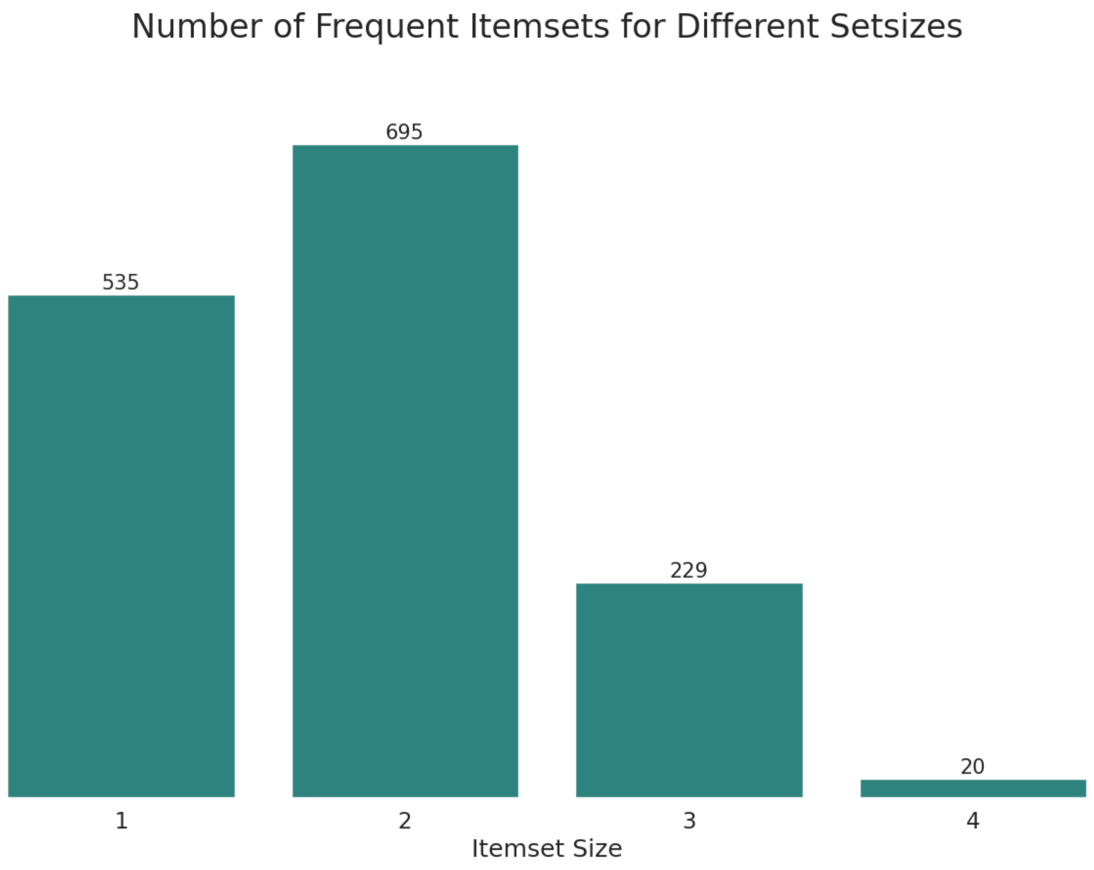
- The apriori algorithm finds all frequent itemsets
- An itemset is frequent, if its support lies above a predetermined threshold
- We consider an itemset frequent if its support is greater than 0.01
- For the analysis we only consider the top 1,000 most sold products

Dataset	
Product_SKU	Product ID
Business_Unit	Top category of the product
Product_Category	Second Level category of the product
Product_SubCategory	Third level category of the product
Ticket_Number	Basket ID
Customer_Loyalty	Boolean indicator if customer is subscribed to loyalty programme
Customer_CardNumber	Customer ID
Store	Store ID
Purchase_Date	Date of the purchase
Price	Price of the product
Quantity	Quantity purchase

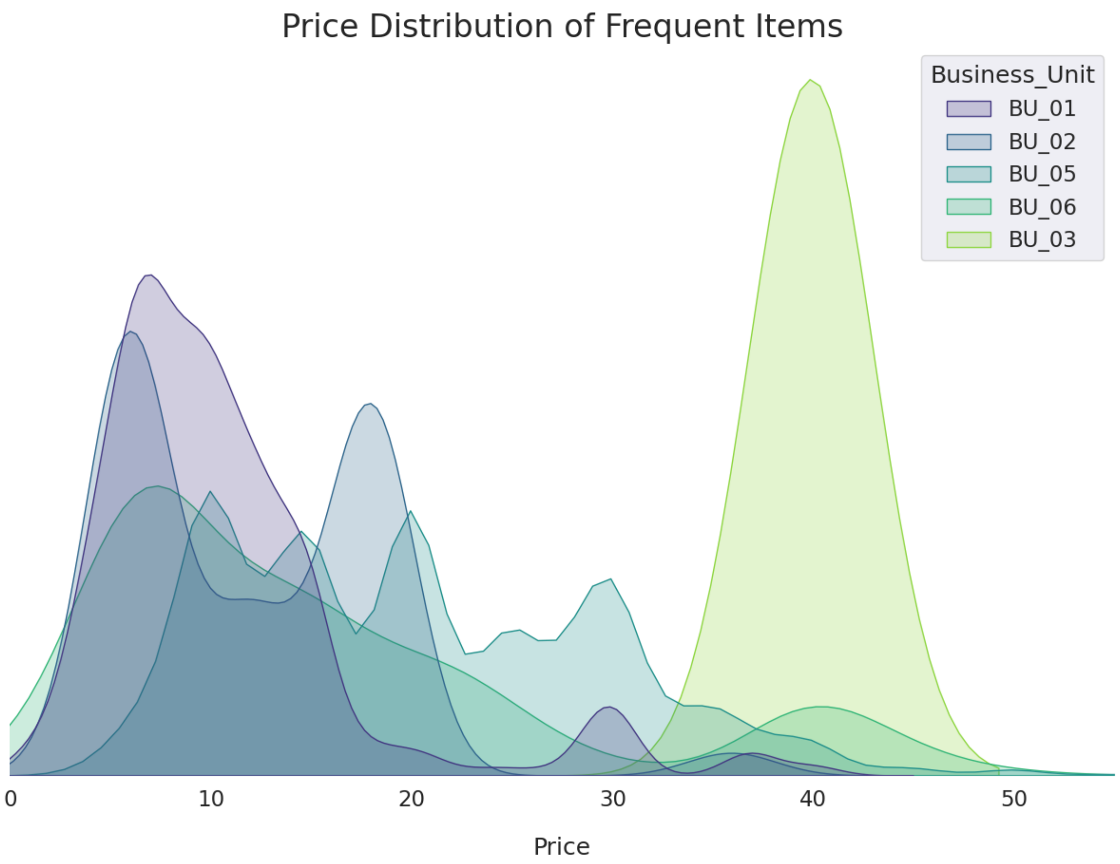
FREQUENT ITEMSETS

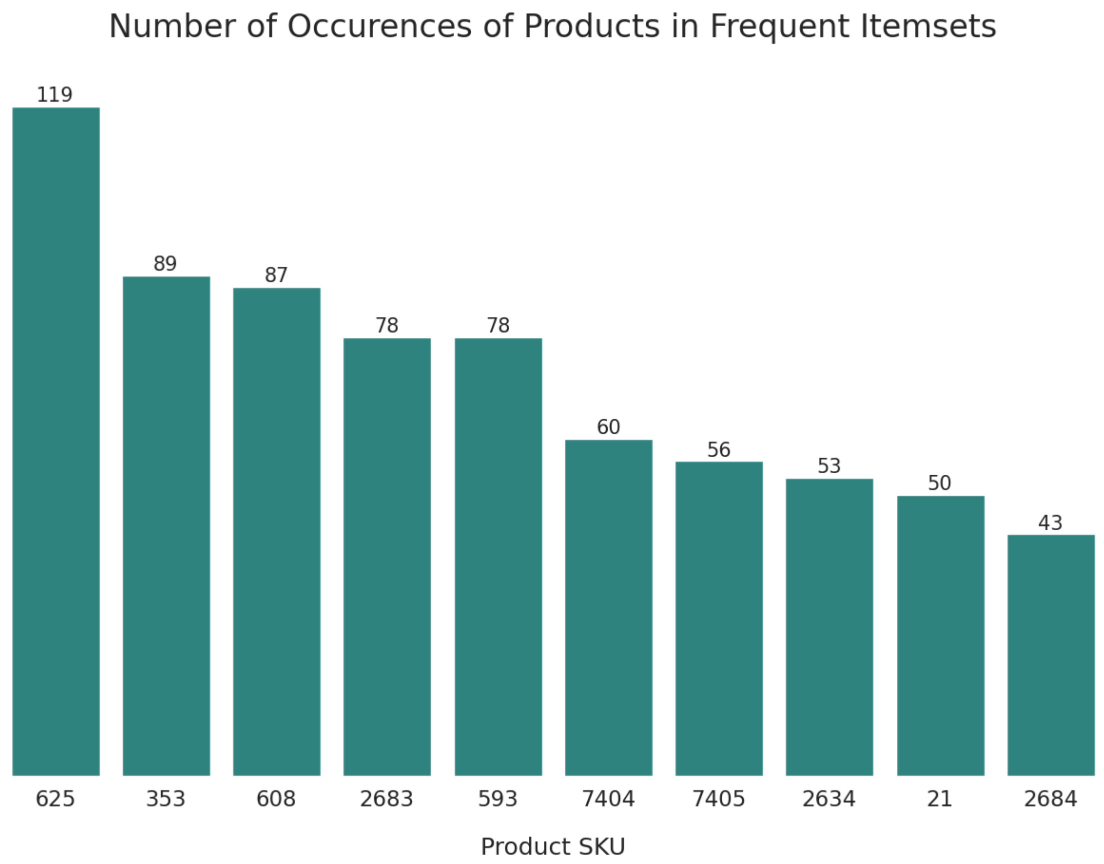
Top 5 frequent itemsets for each setsize:

1. {625} - **Support: 0.1158**
2. {9789, 9788} - **Support: 0.0767**
3. {625, 608, 593} - **Support: 0.0263**
4. {625, 608, 593, 353} - **Support: 0.0151**



FREQUENT ITEMSETS





ASSOCIATION RULES

Top 5 highest lift association rules:

1. {9335, 9788} => {9789, 9334} **Confidence: 100% | Lift: 74.46**
2. {9789, 9334} => {9335, 9788} **Confidence: 95.83% | Lift: 74.46**
3. {9789, 9335} => {9788, 9334} **Confidence: 100% | Lift: 74.46**
4. {9334, 9788} => {9789, 9335} **Confidence: 95.83% | Lift: 74.46**
5. {1747} => {1074} **Confidence: 80.77% | Lift: 48.11**

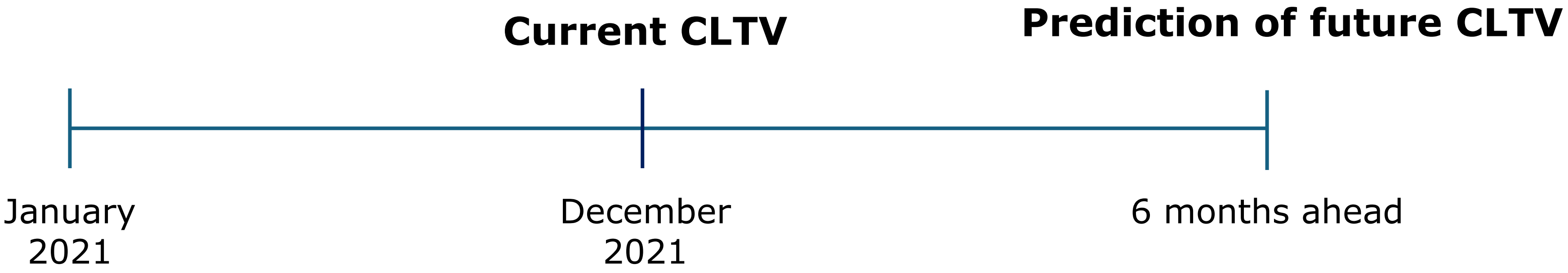
CUSTOMER LIFETIME VALUE

WHAT IS THE CLTV?

The **Customer Lifetime Value (CLTV)** is a metric that allows a business to see how much revenue a customer might bring in over time.

We used two different approaches in its calculation:

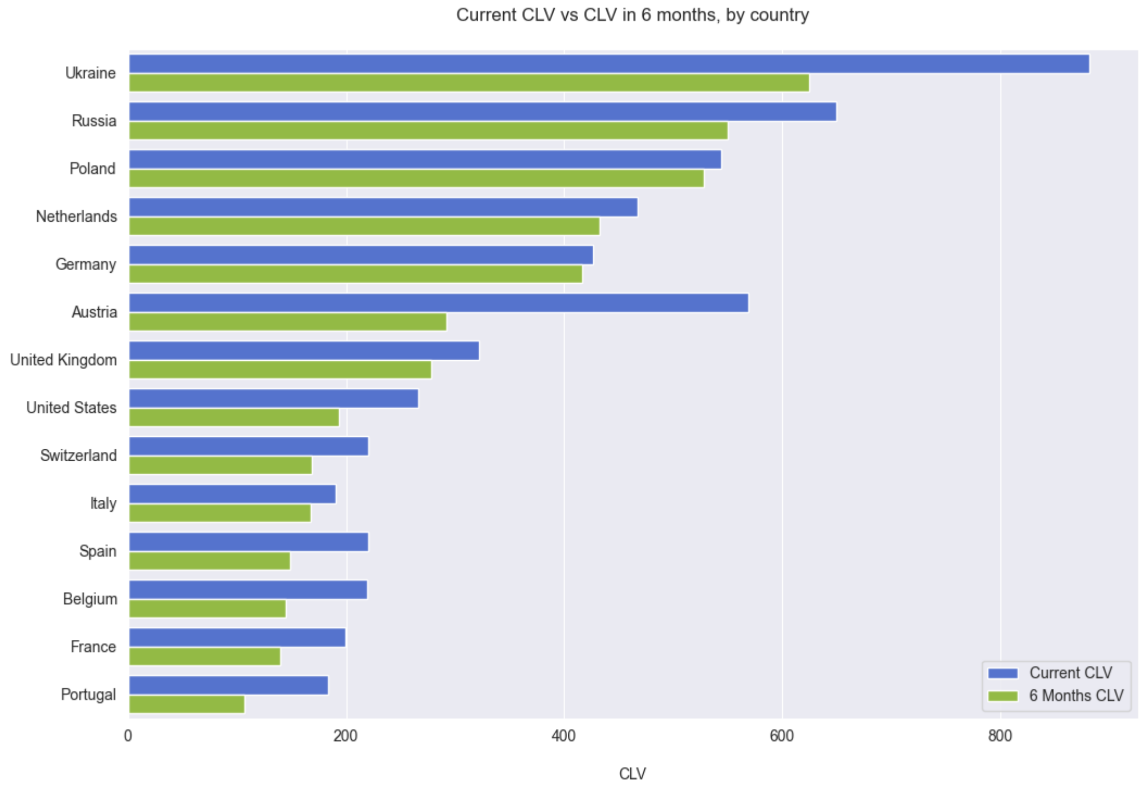
- *Current (or historical) CLTV.*
- *Predictive CLTV*.



CURRENT VS FUTURE CLTV

When calculating the present CLTV, for each customer we take into account his average purchase value, his average frequency rate and his average lifespan.

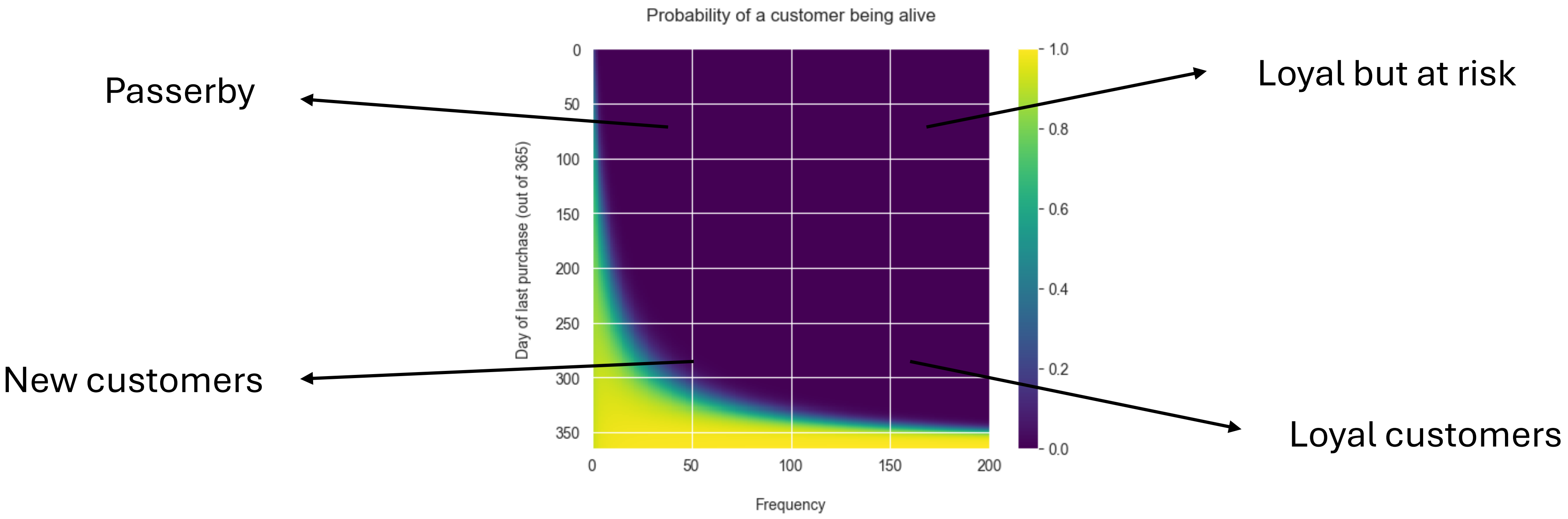
For the future values, a combination of the *BG/NBD model* and the *Gamma-Gamma model* has been used.



ADDING THE RFM

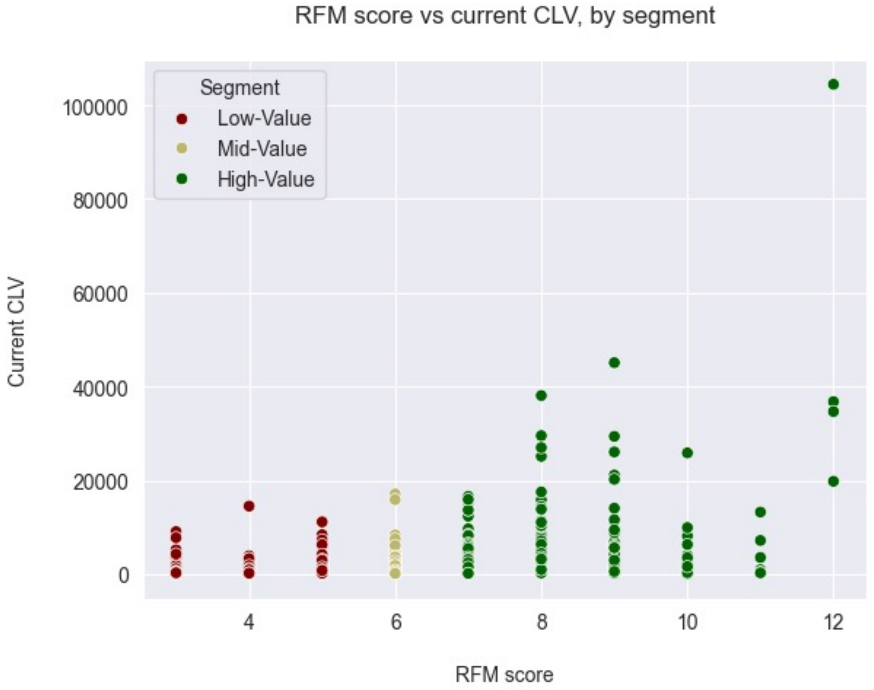
When predicting the CLTV of each customer, it may be useful to use the **RFM model**, which gives us more insights on each customer by taking into account:

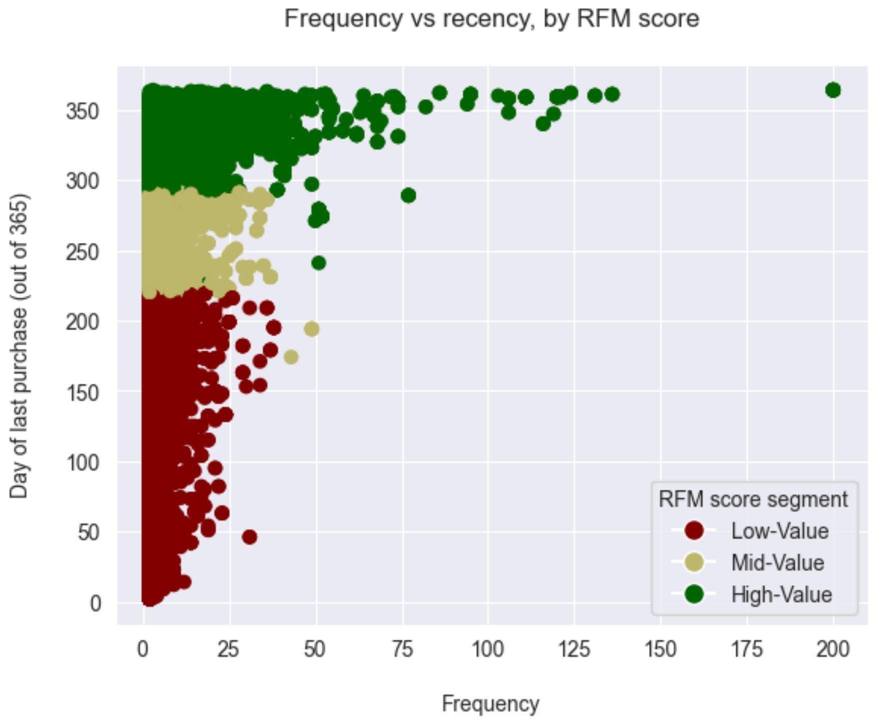
- *Recency (R)*: how much time has passed since the last purchase.
- *Frequency (F)*: how often a purchase is made in a given time period.
- *Monetary value (M)*: how much has been spent, in total, in a given time period.



CLTV VS RFM

By looking at the values of the RFM, we can compute, for each customer, a RFM score. Then, from that, we can group the customers into clusters (Low-Value, Mid-Value, High-Value).
Of course, customers that have a higher value will also have a higher CLV.





THANK YOU FOR YOUR ATTENTION!