

Web Marketing and Communication Management

Progetto a cura di:

- Emanuela Elli (892901)
- Federica Madon (825628)
- Tommaso Strada (829351)



Business questions



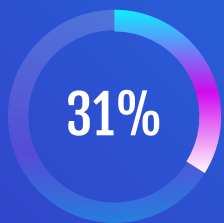
1) *Valorizzazione* della relazione con il *cliente*:

- Come rendere la relazione più redditizia?
- Quali sono le caratteristiche della client base e come può essere divisa?
- Su quali tipi di clienti bisogna investire?
- Chi sono i potenziali churners e perché lo sono?

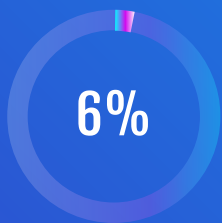
2) *Valorizzazione* dei *prodotti*:

- Quali sono i prodotti più acquistati?
- Su quale combinazione di prodotti conviene investire?

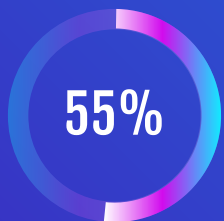
CUSTOMER EXPLORATION



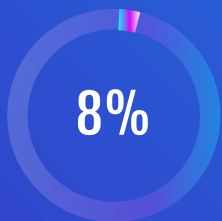
PREMIUM



PREMIUM_BIZ

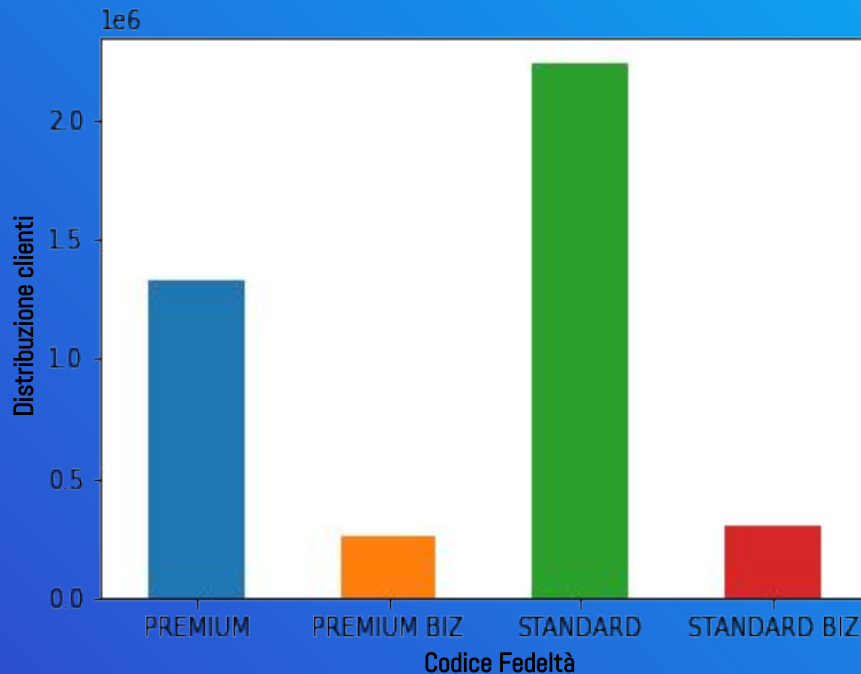


STANDARD



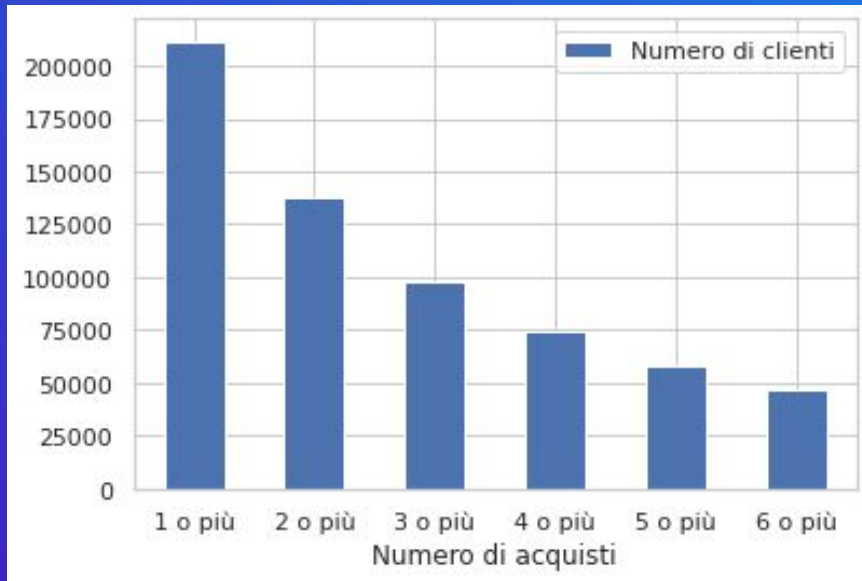
STANDARD_BIZ

Customers fidelity class count



CUSTOMER EXPLORATION

Customer distribution for number of purchases



I 3 clienti che registrano il numero maggiore di acquisti sono:

376925
177 acquisti

117212
155 acquisti

248975
154 acquisti

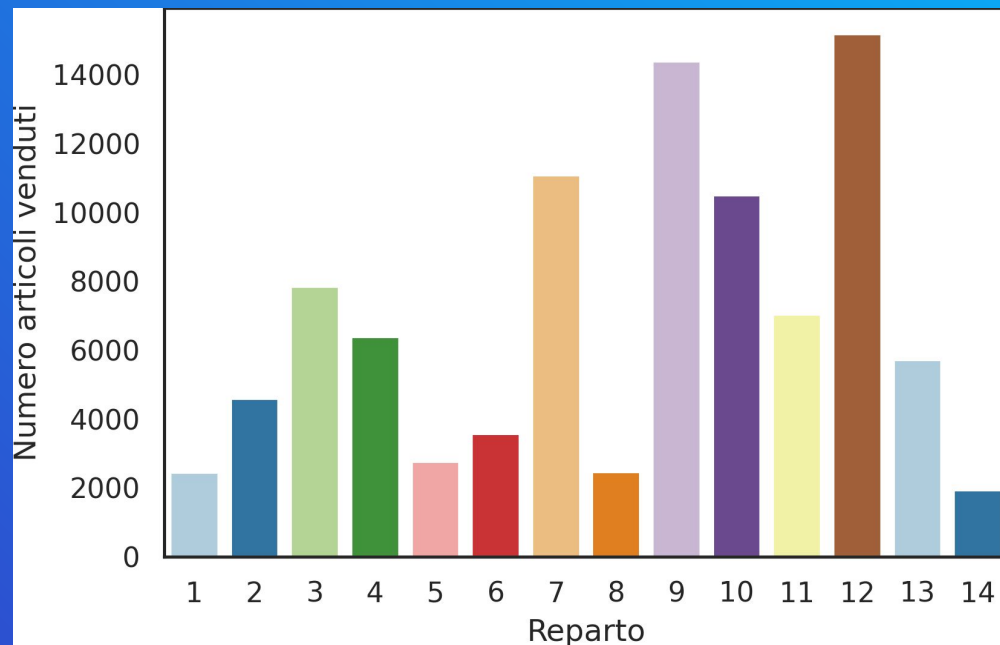
PRODUCTS EXPLORATION

In seguito alla data preparation ci sono:

- 902994 scontrini diversi
- 95826 articoli differenti venduti distribuiti per 14 reparti

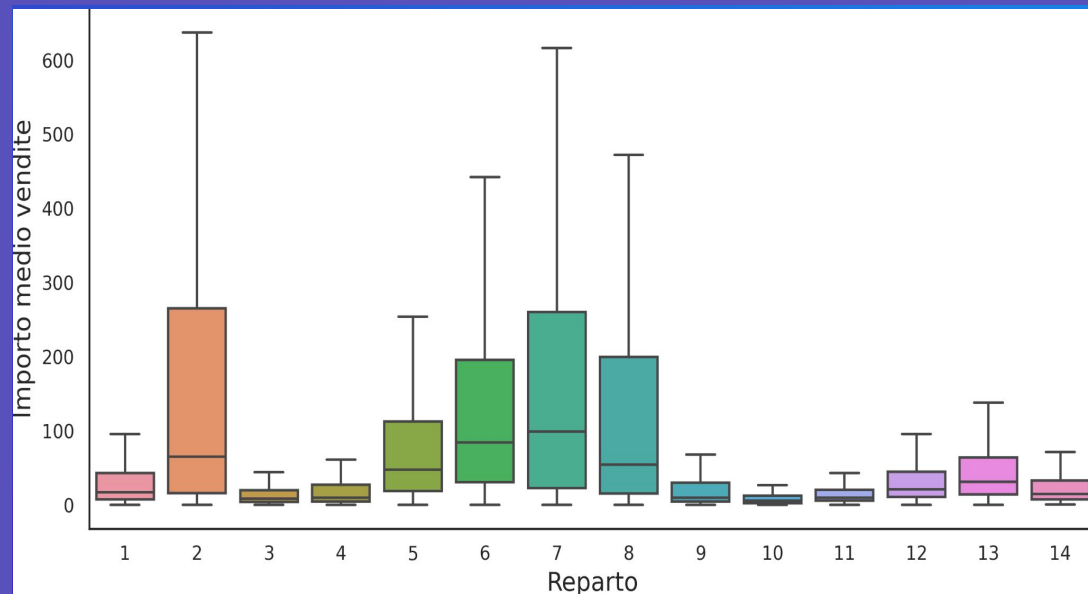
Gli articoli acquistati maggiormente appartengono al reparto 12

Product distribution for departments



PRODUCTS EXPLORATION

Average sales amount for departments



Gli articoli venduti nel reparto 7 sono quelli con mediana più elevata.

DATA DRIVEN-STRATEGY

Business questions



Models



Actions data driven



Data cleaning



Evaluation

DATA PREPARATION

Duplicati

Vengono rimosse le **righe** che fanno riferimento allo **stesso articolo** ma con **importi diversi**.

Resi

Eliminazione **righe** che fanno riferimento ad un **reso** e non ad un acquisto.

Importo netto

Creazione variabile **IMPORTO_NETTO** ed eliminazione delle righe con **valore negativo**.

Formattazione

Formattazione variabili nel **formato corrispondente** (es: DATA = DATETIME).

MODELS

Segmentare i propri clienti per valore in modo tale da pianificare differenti strategie di marketing e comunicazione



Valorizzazione del cliente

Prevenire l'abbandono dei clienti aumentando fidelizzazione e retention

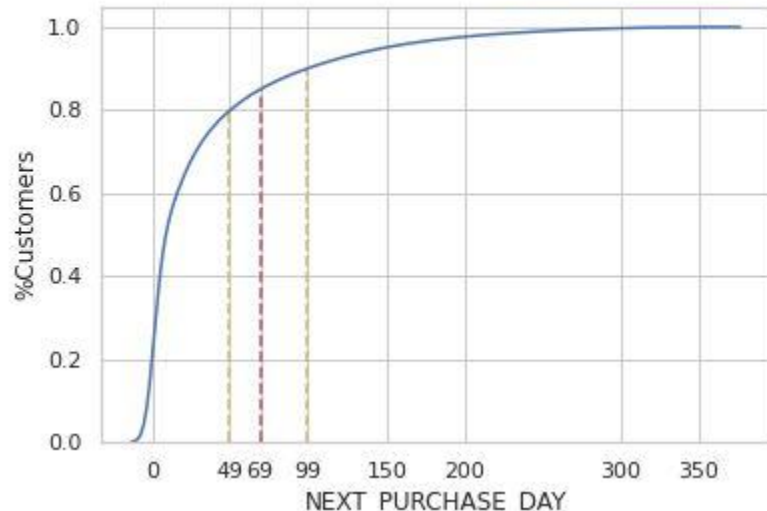
Studio **affinità** dei **prodotti** per aumentare le vendite

Valorizzazione del prodotto

RFM PREPARATION

Next Purchase Day Curve

85% degli utenti riacquista entro 69 giorni



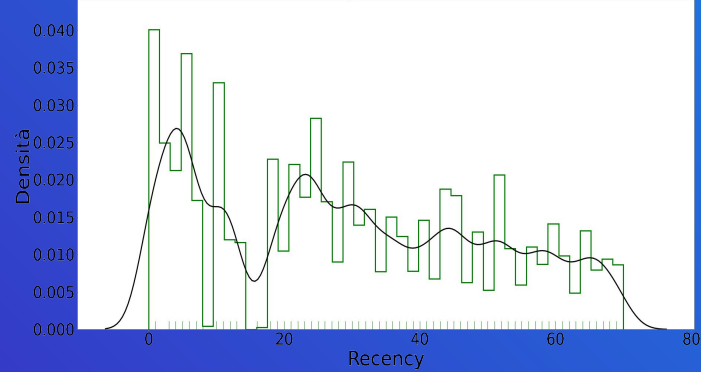
Preprocessing

- ❑ Divisione customer base in attiva e inattiva. Il cut point è stato fissato al 2019-02-19, cioè 69 giorni prima dell'ultima data disponibile.
- ❑ Costruzione metriche: Recency, Frequency, Monetary Value
- ❑ Costruzione RF Matrix e RFM Matrix.

RFM

Distributions

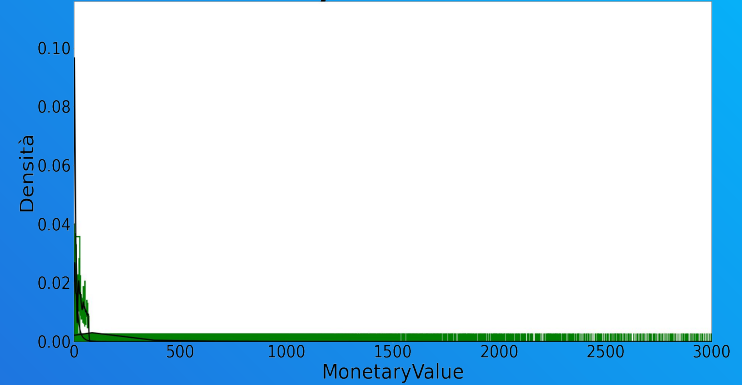
Recency Distribution



Tempo trascorso dall'ultimo ordine o dall'ultima interazione con il prodotto



Monetary Value Distribution

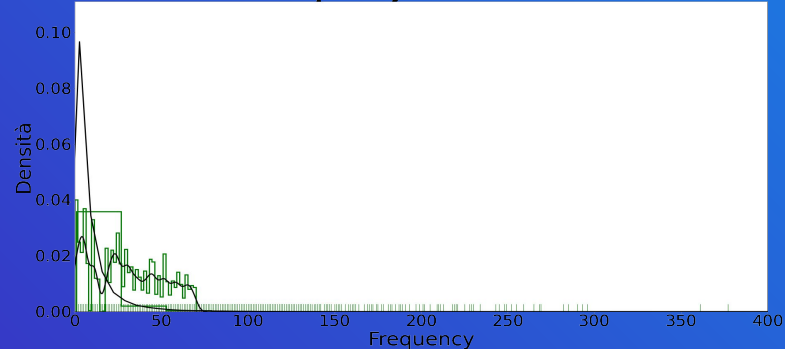


Valore totale o medio delle transazioni



Numero totale di transazioni

Frequency Distribution



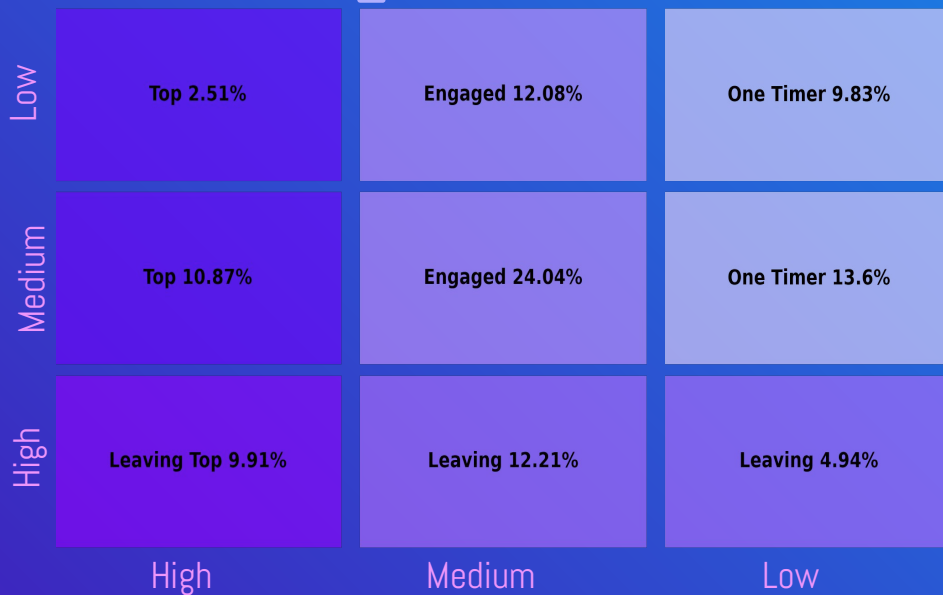
RFM

RF Matrix

Engaged(36.12%):Clienti che hanno acquistato recentemente o abbastanza recentemente e con media frequenza

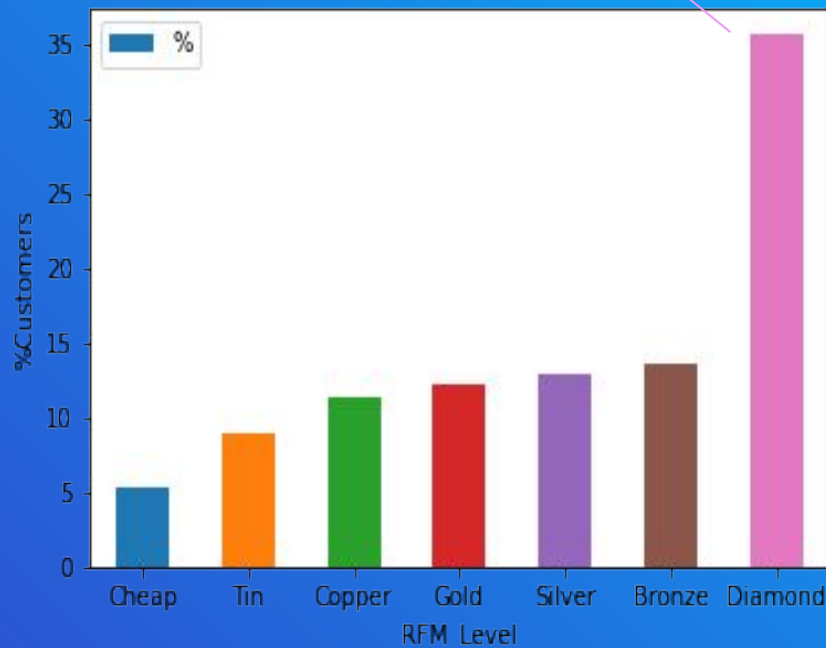


FREQUENCY



RF Matrix

Diamond(35.56%):Clienti con alti valori di Monetary e Top RF

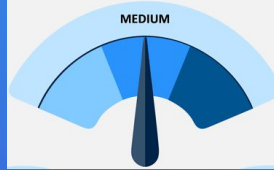


RFM INSIGHTS



HIGHER VALUE CUSTOMERS

Trattenere questi tipi di clienti, proponendo prodotti con maggior valore.



MEDIUM VALUE CUSTOMERS

Raccomandare altri prodotti e proporre programmi di fedeltà.



LOWER VALUE CUSTOMERS

Riacquistare questo tipo di clienti e aumentare la retention

CHURN PROPENSITY

»» Riprendendo il concetto di cut point definito alla slide 10, dividiamo in:

- Lookback period
1/05/2018 - 19/02/2019
- Holdout period
19/02/2019 - 20/04/2019

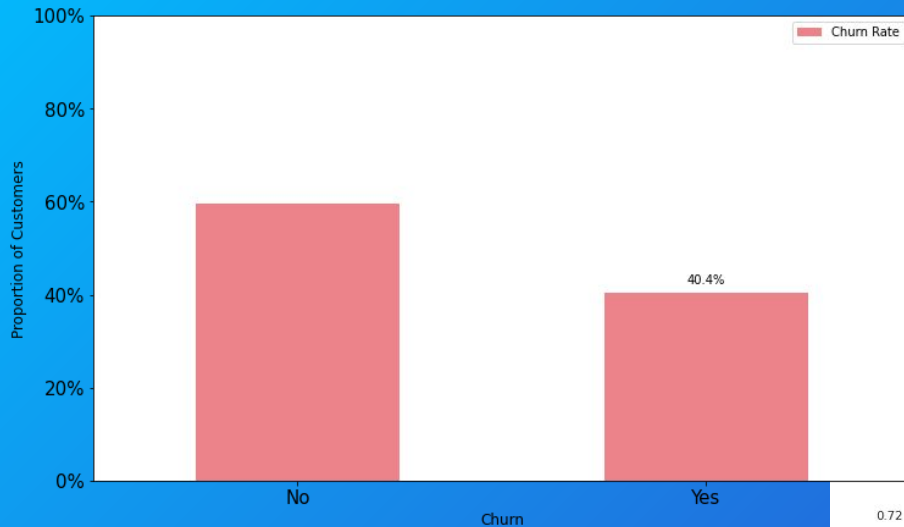
»» Features considerate

- Type fidelity program
- Main account
- Active account
- First and last date activation
- Reference store
- Number of subscriptions for each client
- email account provider
- phone number added
- type client account
- client job

»» Algoritmi utilizzati

- Logistic regression
- SVC
- Kernel SVM
- KNN
- Gaussian NB
- Decision Tree Classifier
- Random Forest

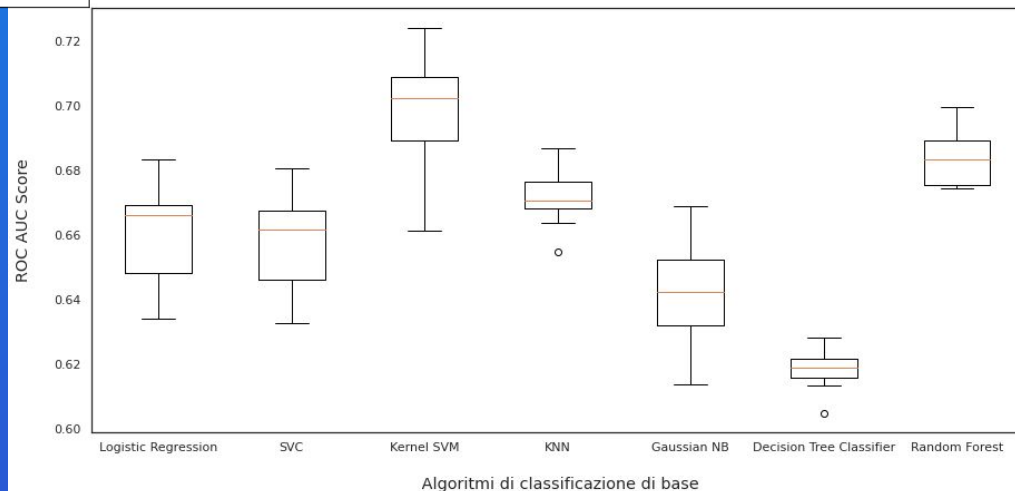
Overall Churn Rate



Definizione dei churn a partire
dalla data definita alla slide

Comparazione delle curve ROC degli
algoritmi utilizzati.

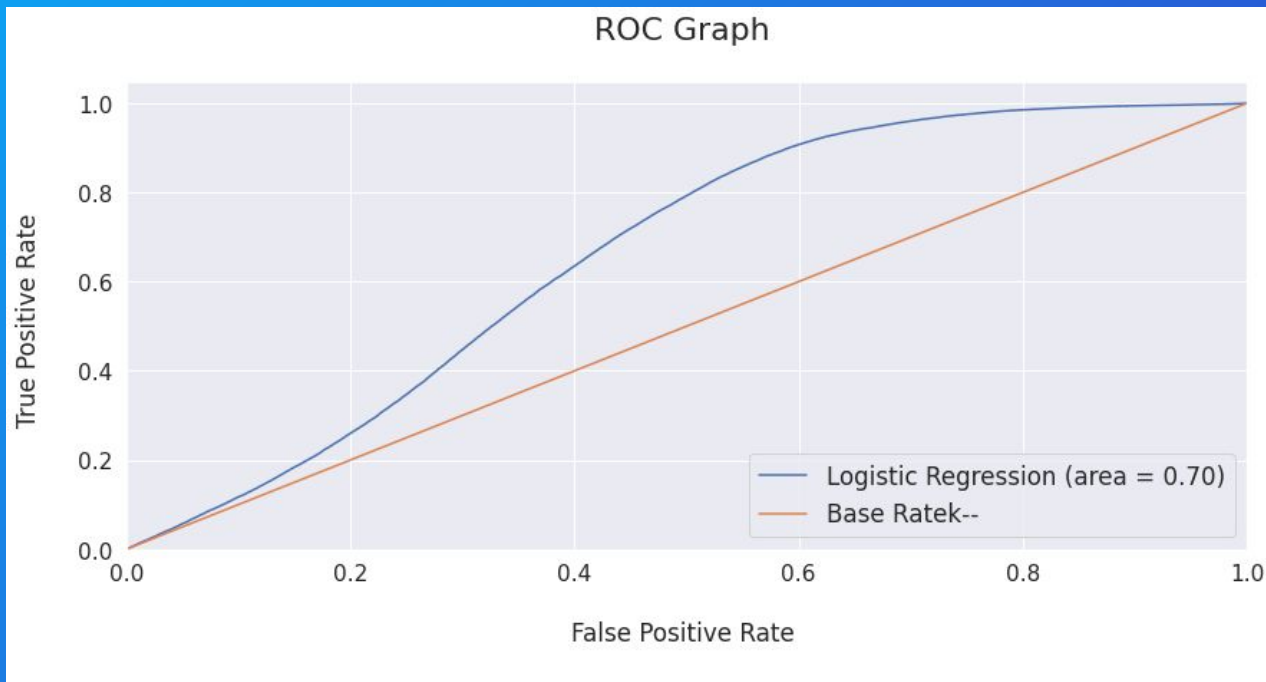
ROC AUC Comparison



CHURN PROPENSITY

La **regressione logistica** è un modello ottimale siccome ha la combinazione relativamente più alta di *precision*, *recall* e *F1 score*.

Si valuta il modello usando il grafico della curva ROC che mostra la capacità di un modello di distinguere tra le classi in base al punteggio medio AUC.



La **linea arancione** rappresenta la curva ROC di un classificatore casuale (un buon classificatore cerca di rimanere il più lontano possibile da quella linea).

Come mostrato in figura, il modello di regressione logistica perfezionato ha mostrato un punteggio AUC più elevato rispetto al modello casuale.

MARKET BASKET ANALYSIS

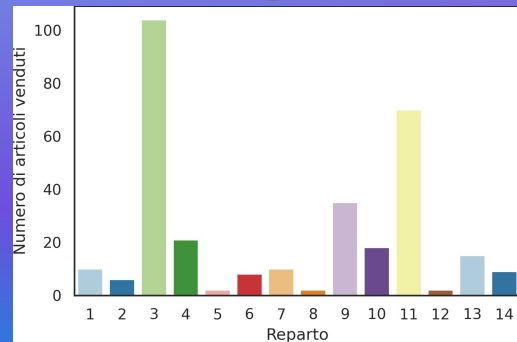


Overview of high
support articles

188938 scontrini
diversi - 62582
articoli differenti

Selezionati gli articoli
con un valore di
supporto maggiore di
0.1%

Gli articoli ad alto
supporto sono 312



Periodo di riferimento dal 19/02/2019 per coerenza con gli altri modelli.

MARKET BASKET ANALYSIS

390 itemsets interessanti con supporto > 0.1% tutti appartenenti al reparto 3

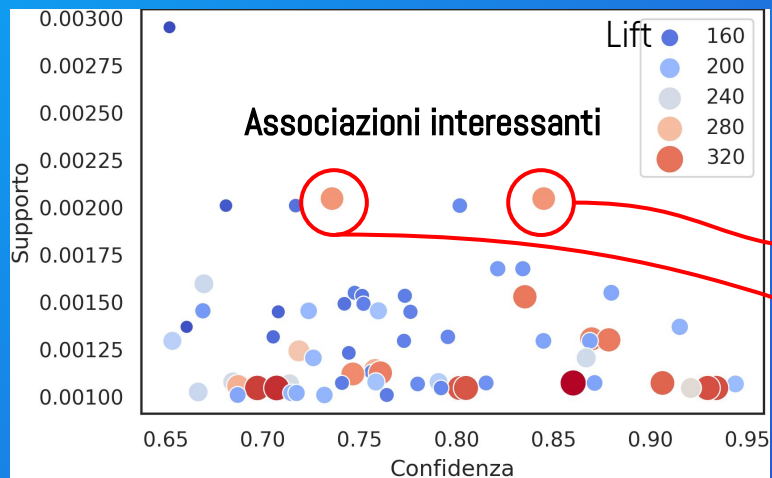
Apriori Algorithm

Association Rules

Pruning delle relazioni

Association rules

66 associazioni rilevanti con confidenza > 0.65 e lift > 1



36298353 → 36298381

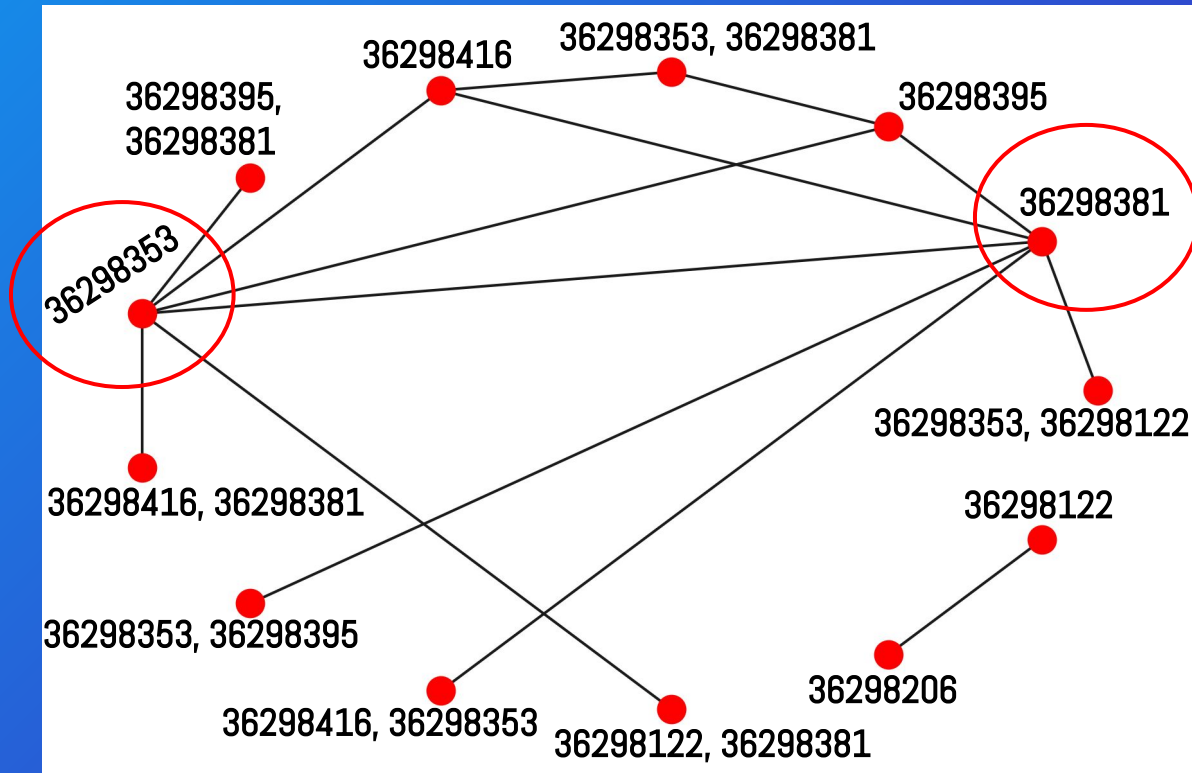
36298381 → 36298353

MARKET BASKET ANALYSIS



- Gli articoli **36298353** e **36298381** sono i prodotti più “driven”
- L'articolo **36298353** è quello più acquistato tra quelli con supporto maggiore
- L'associazione (**36298353**, **36298122**) → 36298381 risulta quella con lift maggiore (> 354)
- I clienti che seguono la regola con lift maggiore acquistano, in media, i tre prodotti ogni **17 giorni**

Top 15 association rules



AZIONI DATA DRIVEN



Move Silver and Bronze customers to Gold level

Dopo il livello Diamond, la maggior parte della customer base si divide tra il livello Silver e Bronze. Risulta quindi opportuno ingaggiare questi tipi di clienti aumentando la loyalty e sviluppando azioni di up/cross selling (ad esempio attraverso la creazione di bundle promozionali di prodotti sia nello store fisico che online). Alcune possibili strategie: campagne email personalizzate, sconti differenziati in base alle abitudini d'acquisto, costruire una relazione con il cliente a partire dal processo di on-boarding, offrire ricompense.



CRM

Gestione della relazione col cliente in modo tale da incrementare la loyalty, advocacy e opportunità di vendita.



Layout optimization

Possibilità di ottimizzare la collocazione dei prodotti (es. shelf) collocando in modo strategico prodotti trainanti e trainati.



Previsione dei potenziali churn e conseguenti azioni di caring

Mantenere i potenziali churner è meno costoso dell'acquisizione nuovi clienti per cui è bene attuare azioni di caring no per perderli. Esempi di azioni mirate: affrontare offerte speciali, razionalizzazione e diversificazione di la strategia di contatto marketing (tramite call center, e-mail, ecc.) in base al rischio di abbandono.



Integrazione degli insight all'interno della marketing automation

Per arricchire la contact strategy attraverso il "next best product to buy" in base alla relazione tra item trainanti e trainati.



Promozioni

Le promozioni possono essere definite calcolando in media ogni quanto vengono acquistate determinate associazioni di prodotti.



**Grazie per
l'attenzione!**