

LUCA LAZZATI

MATRICOLA: 850334



# DIGITAL MARKETING ANALYSIS

Digital Marketing Project

A.A. 2019/2020

# Strategy: a model for data-driven actions

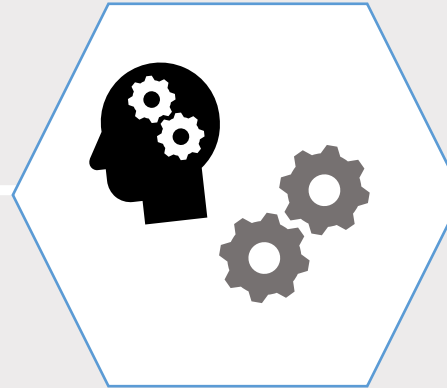
## BUSINESS QUESTIONS



## PULIZIA DEI DATI



## MODELLAZIONE



## VALUTAZIONE



# Modelli Sviluppati



**RFM Model:** Consente di scattare un'istantanea del valore attuale della customer base, identificando segmenti di client ai quali proporre azioni di marketing ad hoc e valutare l'eventuale migrazione dei client tra i diversi segmenti per implementare azioni correttive



**RFM Clustering:** Consente di identificare cluster di client con comportamenti distintivi e non identificabili a priori, derivandone insight per sviluppare azioni specifiche di marketing



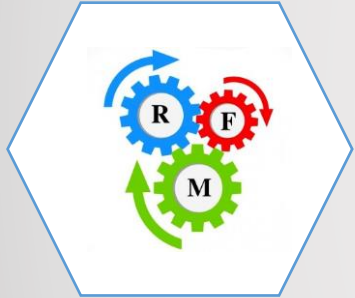
**Propensity to churn:** Consente di assegnare a ciascun cliente la sua probabilità di abbandono, in modo da implementare specifiche azioni di marketing correttive finalizzate a trattenere i clienti a più alto valore

# Business Questions



## RFM

Quanti sono i clienti che hanno **acquistato** poco/tanto **nell'ultimo periodo**? Quanti sono i clienti che **spendono tanto**? Quanti invece sono clienti «**occasional**i»?



## Clustering

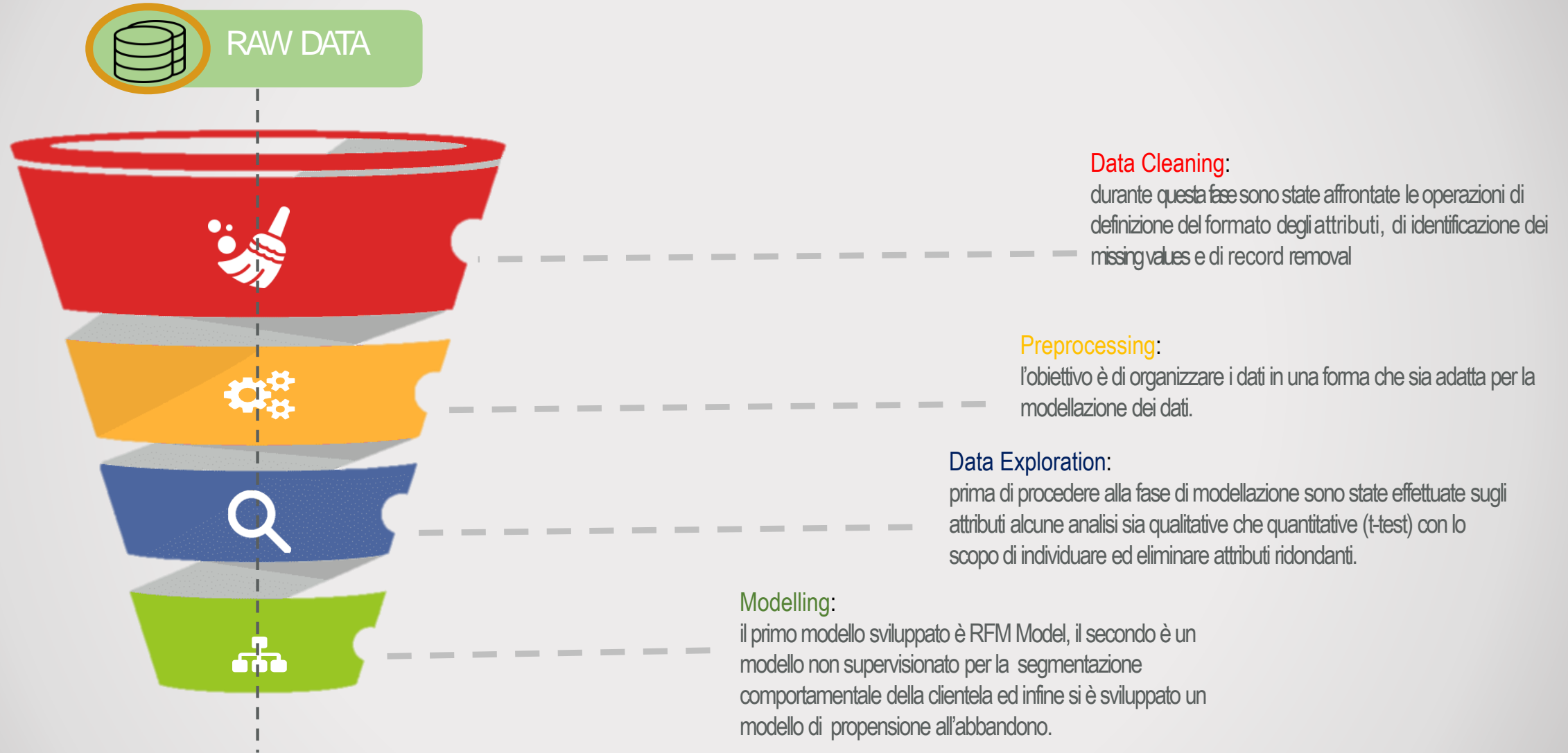
Quanti sono i clienti che fanno parte della **Fascia Alta**? Quanti della **fascia Media**? Quanti della **fascia Bassa**? Quanti sono i clienti **più importanti**?



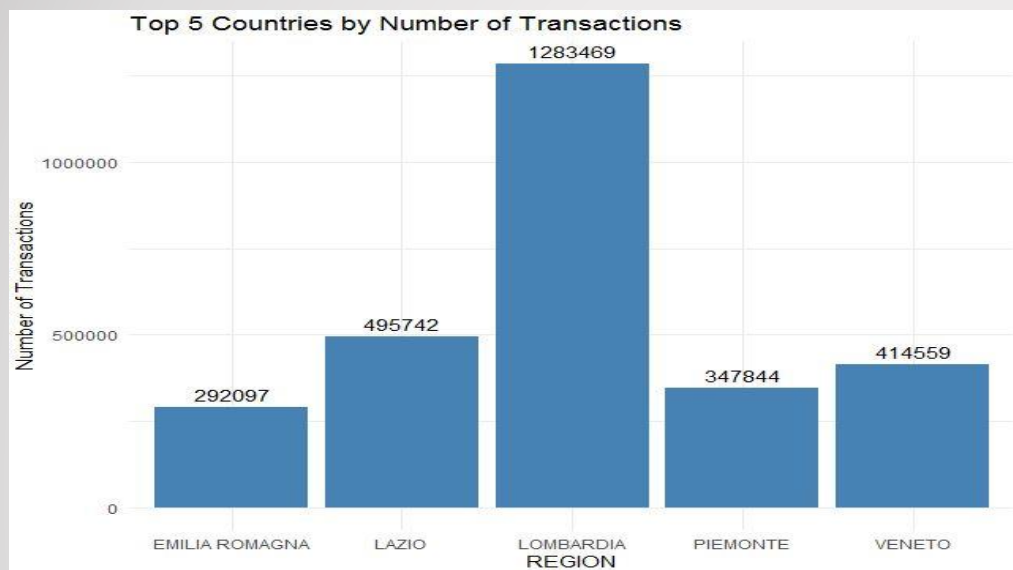
## Churn

Qual è, per ciascun cliente, **la probabilità di abbandono**? Quali **azioni** di marketing correttive, finalizzate a trattenere i clienti a **più alto valore**, si possono implementare?

# Analysis workflow: a data-driven communication strategy



# Data Exploration

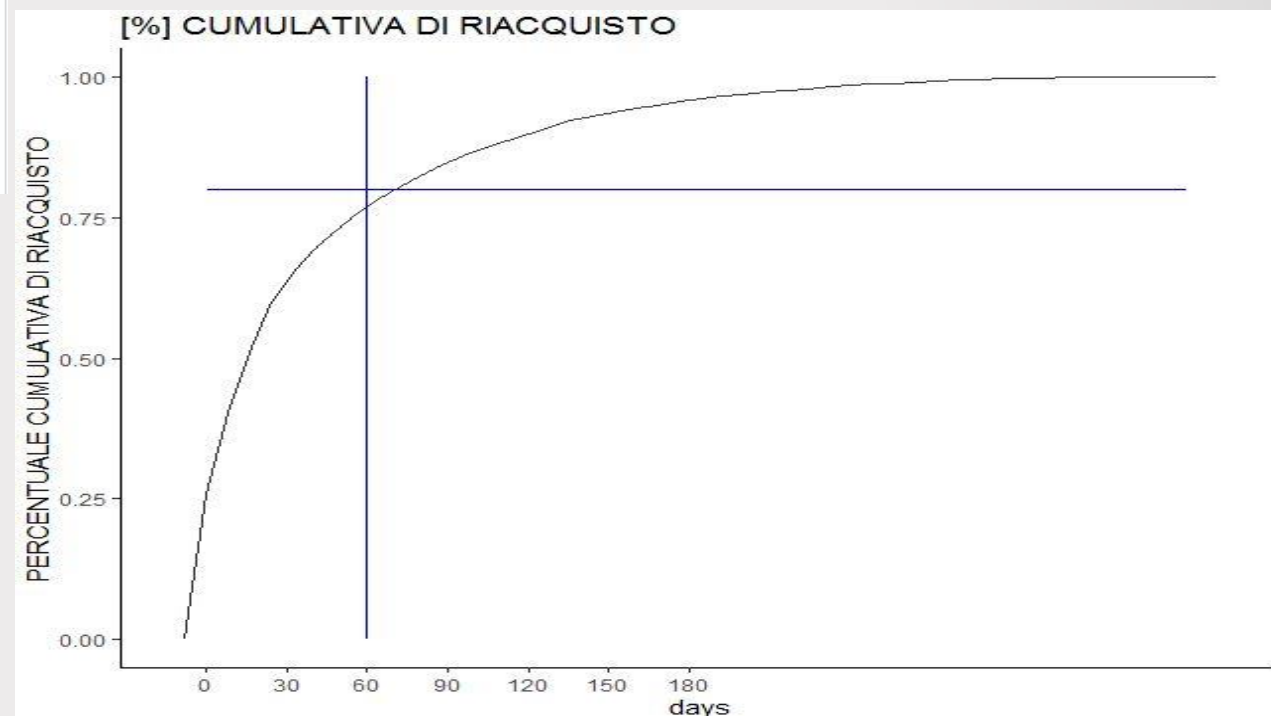


Per il calcolo della soglia dipende dalla scala temporale dell'attività degli Utenti e questa informazione può essere intuita **dall'analisi dei giorni per il prossimo acquisto.**

80% dei clienti **riacquista** dopo 60 giorni.

Viene perciò individuata come data soglia il 28/02/2019 (ultima data di rivelazione il 30/04/2019)

Dato che la Lombardia ha conseguito la maggior parte di di transazioni in un anno, ho voluto creare i miei modelli RFM, Clustering e Churn filtrando per la **Lombardia**



# RFM Model

Analisi relativa al periodo: 28/02/2019 - 30/04/2019    N Record = 35734

## Variabili Esplicative:

- ✓ ID\_CLI
- ✓ ID\_SCONTRINO
- ✓ DATA
- ✓ QUANTITÀ TOTALE
- ✓ IMPORTO TOTALE
- ✓ RECENCY
- ✓ FREQUENCY
- ✓ MONETARY

## Creazione Datamart:

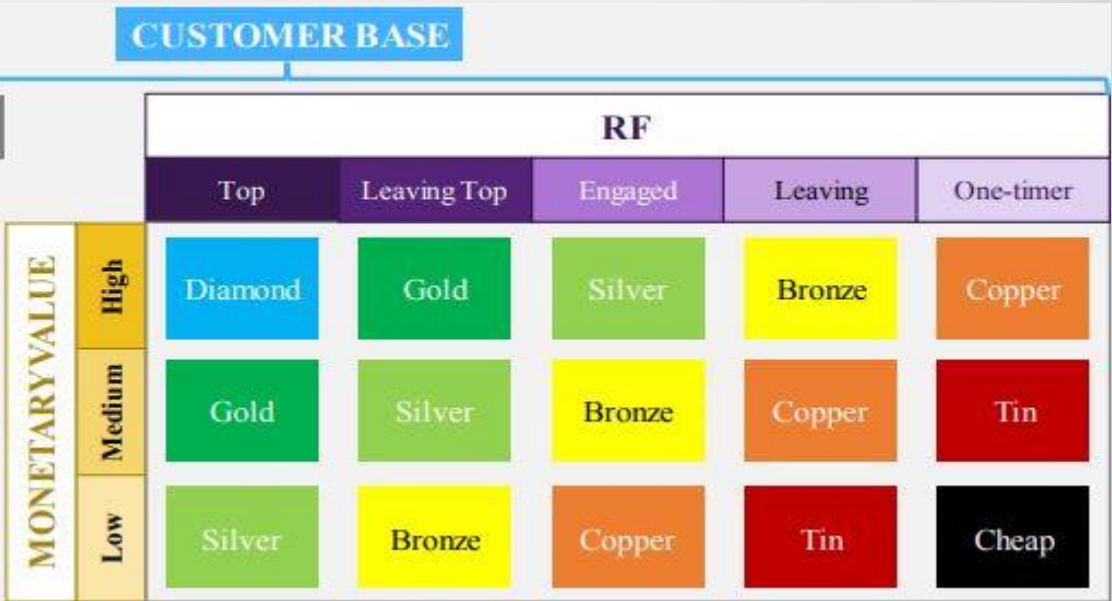
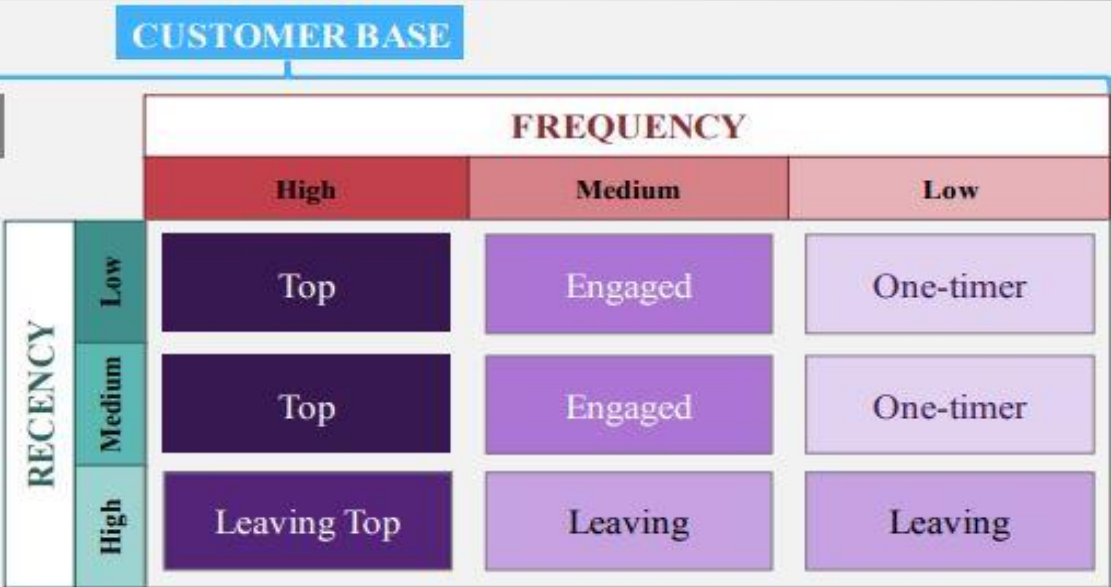
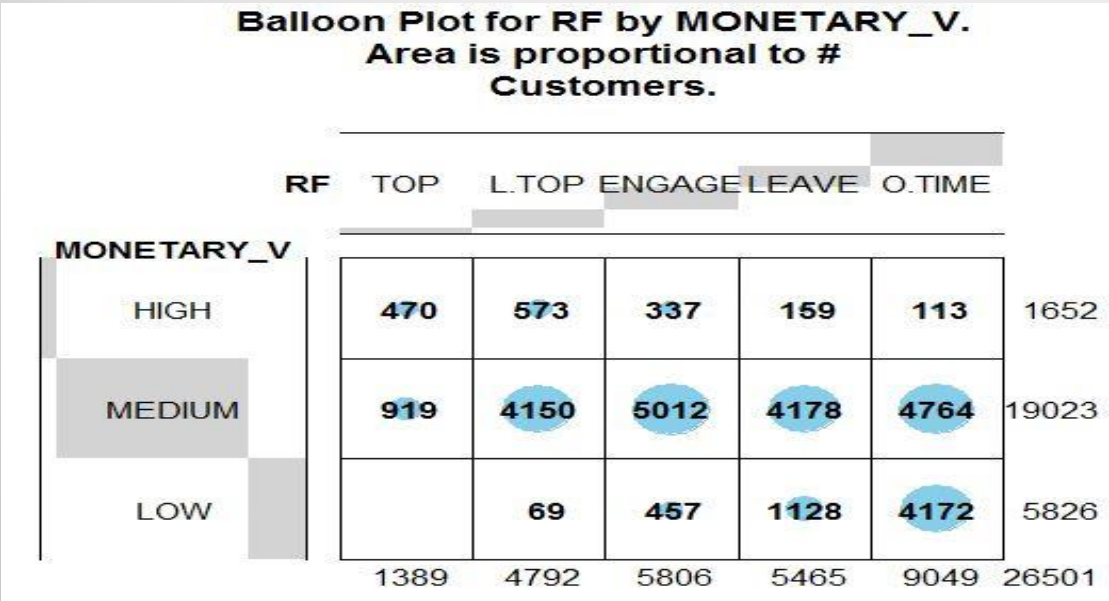
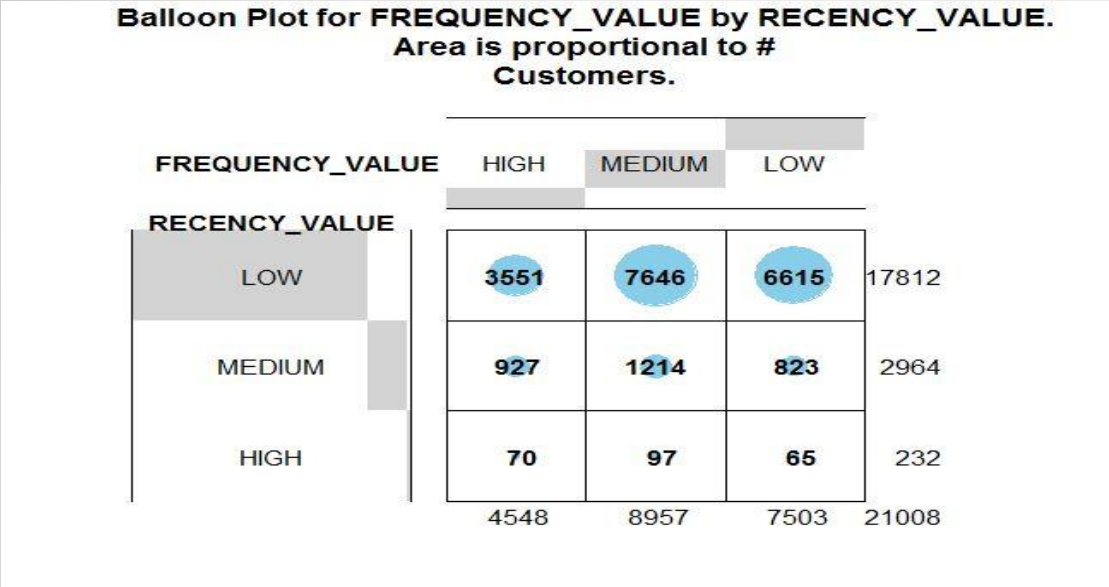
- ✓ Controllo e **selezione** delle potenziali variabili da inserire del modello
- ✓ Definizione di una **finestra temporale** di due mesi
- ✓ Creazione del **datamart** di analisi

## Metodologia:

- ✓ Creazione dei **Segmenti** per identificazione degli utenti (attraverso i **Quartili**) e Visualizzazioni
- ✓ Rimozione **Utenti Inattivi**, prendo in considerazione il comportamento degli **Utenti Attivi**
- ✓ Implementazione di una **RFM-Matrix**



# RF/RFM-Matrix VALUTAZIONI





# Clustering RFM

Analisi relativa al periodo: 28/02/2019 - 30/04/2019    N Record = 35734

## Variabili Esplicative:

- ✓ NUMERO DI UTENTI
- ✓ RECENCY MEDIA
- ✓ FREQUENCY MEDIA
- ✓ MONETARY MEDIA
- ✓ ENTRATE TOTALI

## Creazione Datamart:

- ✓ Controllo e **selezione** delle potenziali variabili da inserire del modello
- ✓ Creazione del **datamart** di analisi

## Metodologia:

- ✓ Creazione di 3 cluster per individuare le **Fascia Bassa, Fascia Media e Fascia Alta**
- ✓ Implementazione dell'algoritmo di Clustering per **Segmentazione Comportamentale**

# Clustering RFM VALUTAZIONI

## Clienti Occasionali

Numero di Utenti:  
13464 su 35734

Numero Medio Giorni  
di Inattività= 221

Numero Medio di  
Acquisti = 2 all'anno

Scontrino Medio = 18  
Euro all'anno

## Clienti Abitudinali

Numero di Utenti:  
19321 su 35734

Numero Medio Giorni di  
Inattività= 69

Numero Medio di  
Acquisti = 3 all'anno

Scontrino Medio = 164  
Euro all'anno

## Clienti Top

Numero di Utenti:  
2949 su 35734

Numero Medio Giorni di  
Inattività= 63

Numero Medio di  
Acquisti = 12  
all'anno

Scontrino Medio = 3442  
Euro all'anno

# Churn Model

Analisi relativa al periodo: 28/02/2019 - 30/04/2019    N Record = 51550

## Variabili Esplicative:

- ✓ NUMERO SCONTRINI
- ✓ SCONTRINO MEDIO
- ✓ SCONTO COMPLESSIVO PER CLIENTE
- ✓ NUMERO MEDIO DI ARTICOLI PER SCONTRINO
- ✓ CODICE FIDELTY
- ✓ ID\_NEG

## Creazione Datamart:

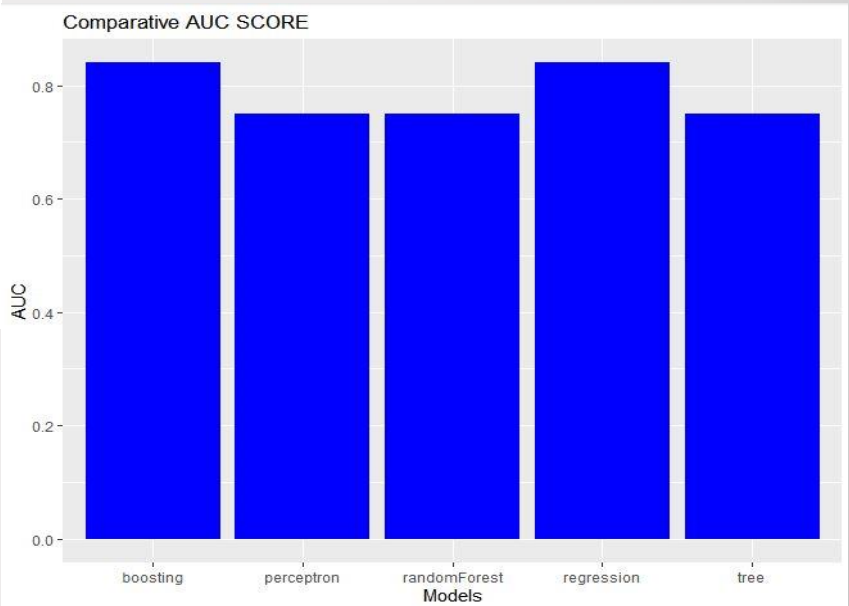
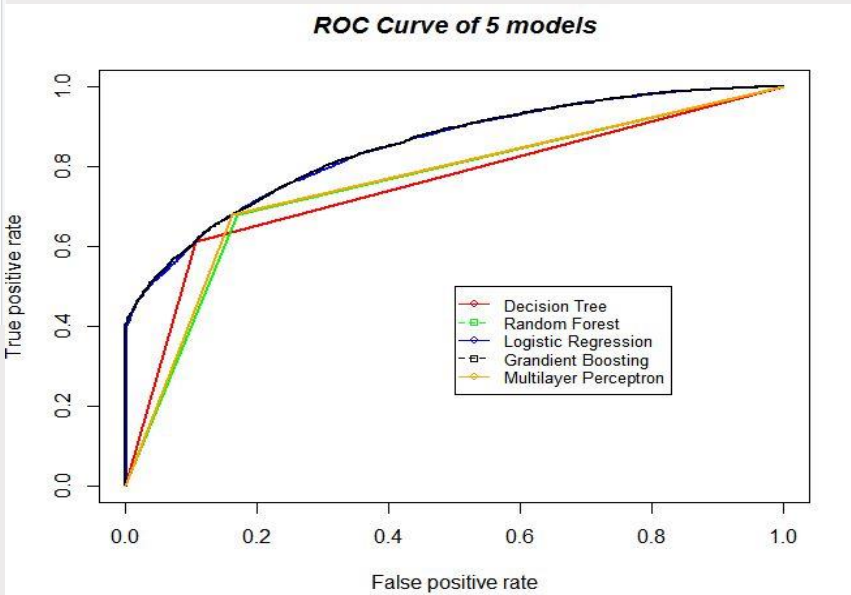
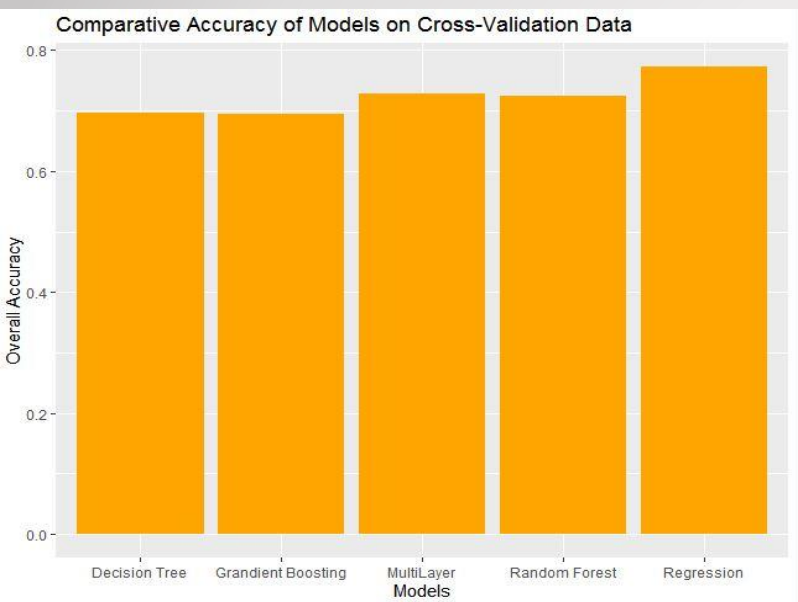
- ✓ Controllo e **selezione** delle potenziali variabili da inserire del modello
- ✓ Definizione **Soglia** temporale di riacquisto in 60 giorni
- ✓ Esclusione dei clienti che hanno effettuato il primo acquisto **dopo** il 28/02
- ✓ **Standardizzazione** delle variabili numeriche e creazione dataset

## Metodologia:

- ✓ Partizione dei dati in **train e test**
- ✓ Implementazione di diversi tipi di **modelli** quali Decision Tree, Random Forest, Logistic Regression, Gradient Boosting e MultuLayer Perceptron

**VARIABLE CHURN:** assume **valore 0** se il consumatore non riacquista nella finestra temporale definita, assume **valore 1** se il cliente verrà classificato come churner

# Churn Model VALUTAZIONI



# Azioni Data Driven



## **CRM:**

Gestione virtuosa della Customer Base al fine di incrementare le opportunità di vendita e ridurre i costi



## **Segmentazione:**

Conoscenza del target di riferimento che viene clusterizzato per definire messaggi e promozioni più efficaci ed efficienti



## **Concept Marketing:**

Cambio del programma di fidelizzazione per i potenziali churner al fine di comprenderne i bisogni e migliorarne la soddisfazione

GRAZIE PER L'ATTENZIONE