

# Web Marketing

Giovanni Cornacchia  
Alberto Gadda  
Paolo Guerini Rocco

830631  
824029  
826236



Anno accademico: 2021-2022

# Business Questions

## RFM

L'obiettivo della segmentazione deterministica (RFM) è quello di classificare la customer base in diversi segmenti sulla base della distribuzione di tre variabili chiave:

- **Recency**
- **Frequency**
- **Monetary**

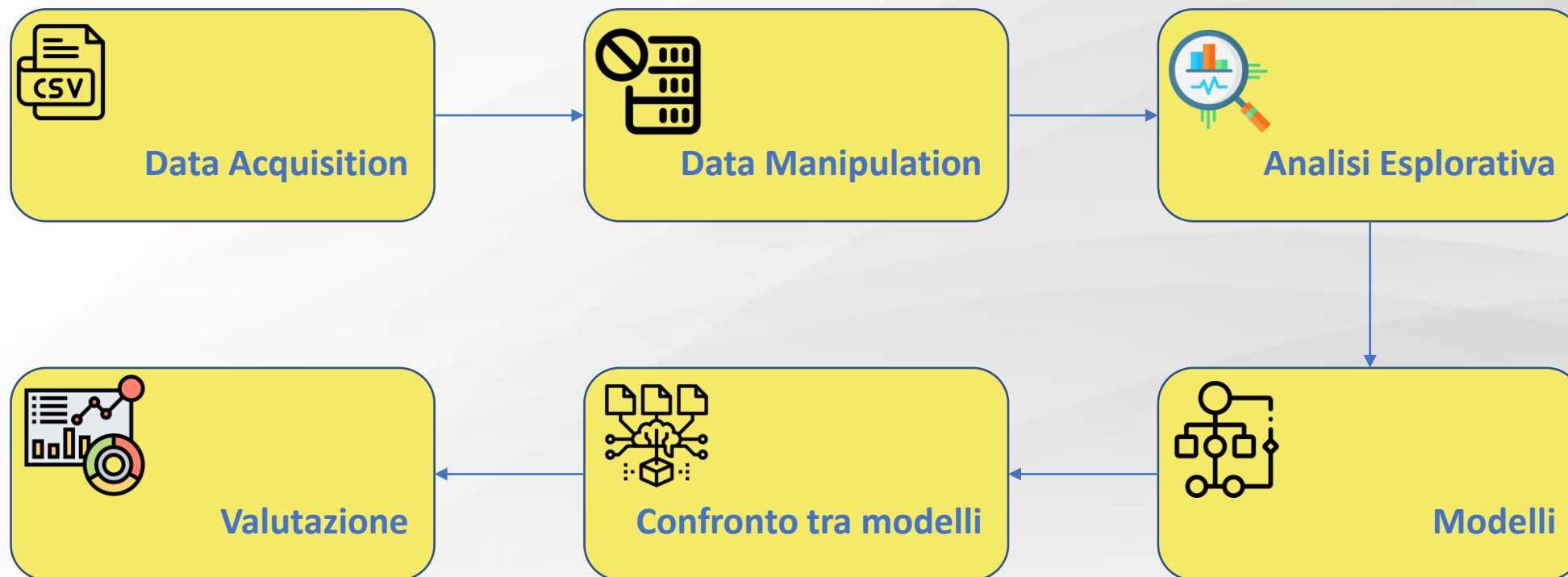
## MBA

L'analisi delle associazioni tra prodotti ha l'obiettivo di individuare regole di associazione tra item che appaiono congiuntamente in un evento, identificando pattern di prodotti/categorie che vengono acquistati con maggiore probabilità

## CHURN

L'obiettivo del modello di churn consiste nell'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

# Workflow



# Datasets



## FIDELITY

Cliente, Fedeltà, Negozio, Account Principale, Tipo Fedeltà, Status, Data Attivazione



## ACCOUNT

Cliente, Email, Telefono, Indirizzo, Tipo Account, Professione



## ADDRESS

Indirizzo, CAP, Provincia, Regione



## PRIVACY

Cliente, Privacy, Profiling, Direct Marketing



## CAMPAIGN

Campagna, Tipo Campagna



## EVENT

Evento, Cliente, Campagna, Consegna, Tipo Evento, Data



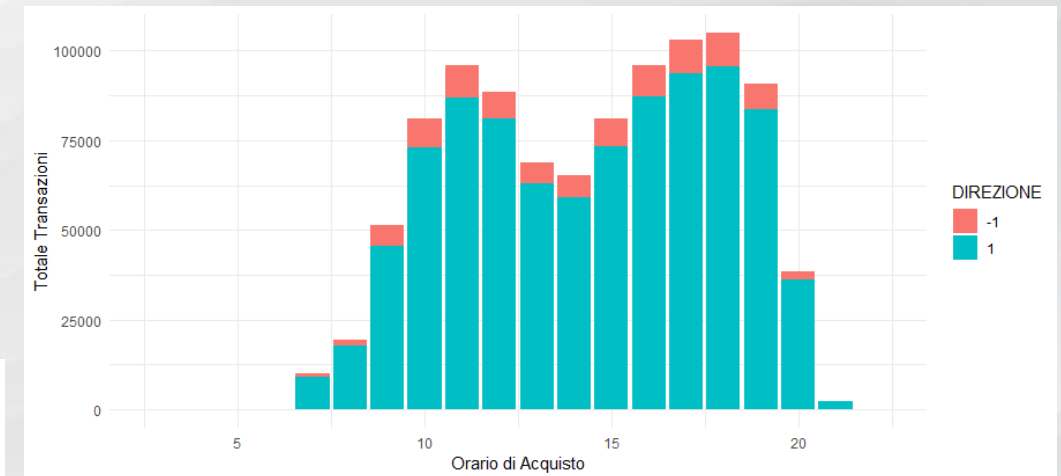
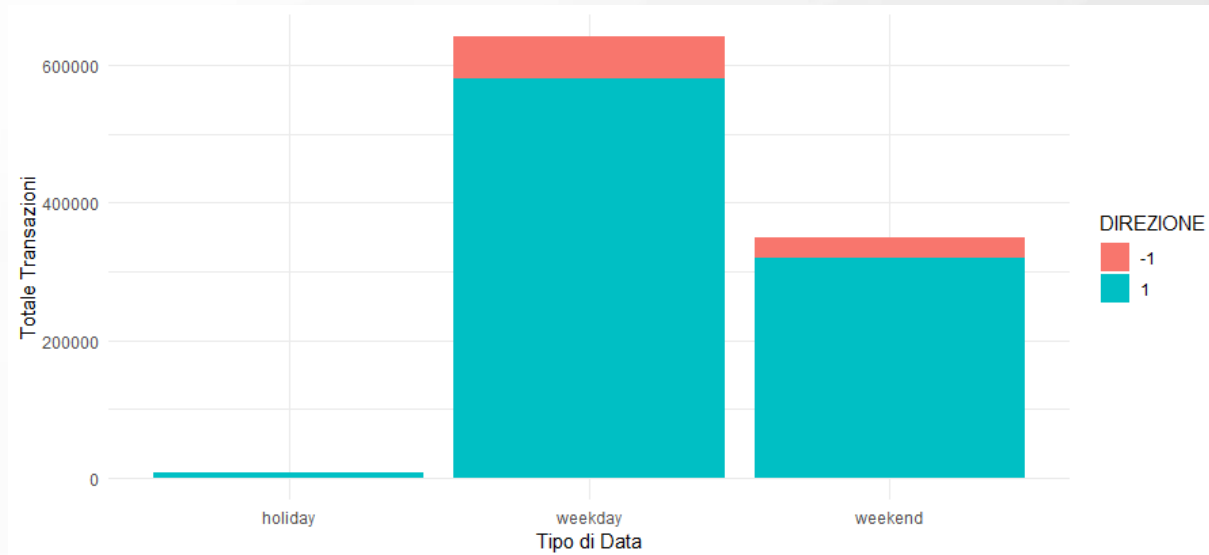
## TICKETS

Scontrino, Cliente, Negozio, Articolo, Reparto, Direzione, Importo, Sconto, Data

# Analisi Esplorativa

L'analisi esplorativa è una fase fondamentale che deve essere svolta per prendere confidenza con i dati.

Guardando la distribuzione degli acquisti nei diversi orari (a destra) si nota che le fasce orarie preferite sono prima di pranzo (10-12) e nel tardo pomeriggio (16-18). Non si nota una significativa variazione del numero di resi nelle varie fasce.



Emerge (nel grafico a sinistra) che nei giorni feriali si concentra il maggior numero di acquisti, dimezzandosi nel weekend per poi pressoché eliminarsi nei giorni di vacanza.

# RFM

Il modello RFM serve per segmentare la clientela sfruttando diverse sue caratteristiche. In questo modo è possibile per l'azienda intraprendere azioni adatte e personalizzate per ogni tipologia di cliente. Le metriche utilizzate ai fini di questa analisi sono:

- **Recency:** quanto tempo fa il cliente ha effettuato l'ultimo acquisto?
- **Frequency:** quanti acquisti ha effettuato il cliente nel periodo di riferimento?
- **Monetary Value:** quanto ha speso il cliente nel periodo di riferimento?

Questa suddivisione è stata effettuata sulla base dello storico delle transazioni dei clienti, usando le soglie numeriche e dei terzili di seguito riportate:

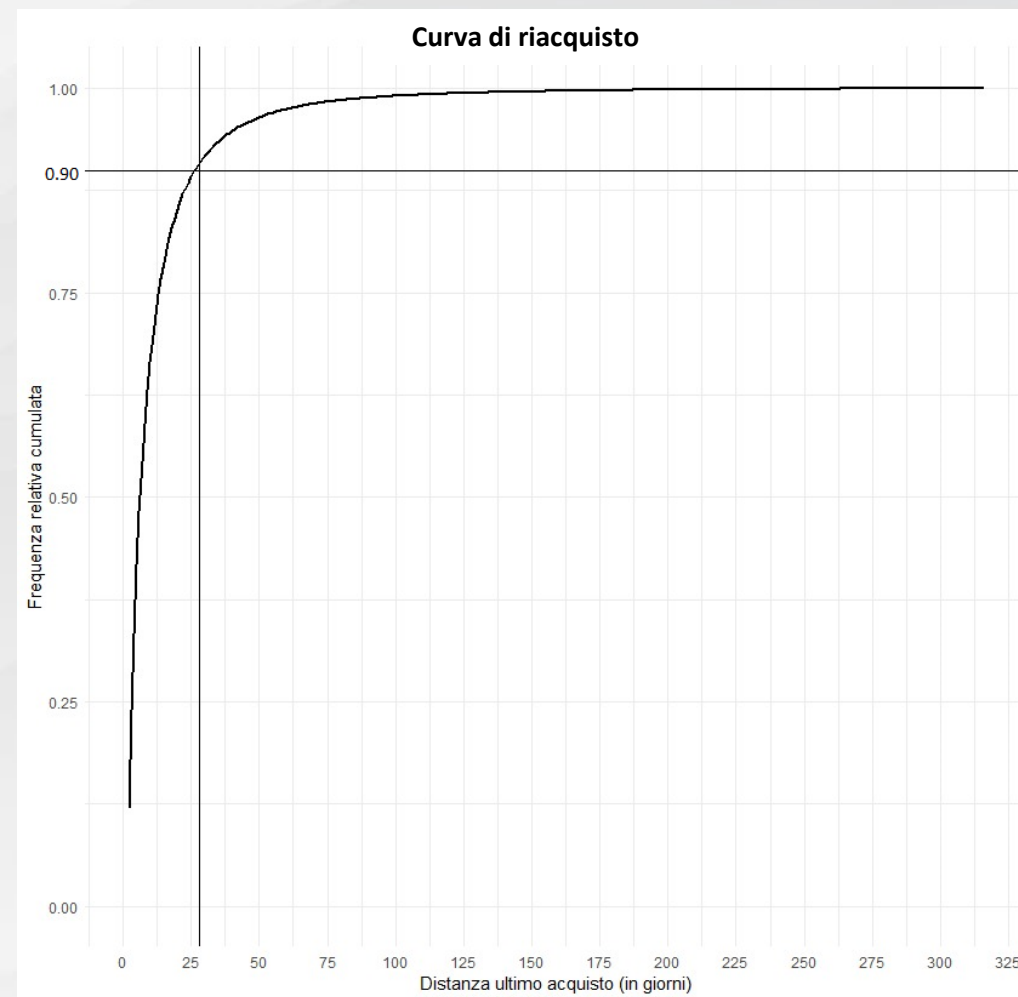
	Recency	Frequency	Monetary
LOW	<33%	1 purchase	<33%
MEDIUM	$33\% < x < 66\%$	2 purchase	$33\% < x < 66\%$
HIGH	>66%	3+ purchase	>66%

# RFM: reference period

Il modello RFM richiede di selezionare un periodo di riferimento su cui basare l'analisi.

Comunemente si utilizza la **curva di riacquisto** per identificare il numero di giorni entro cui il 90% dei clienti acquista nuovamente un prodotto. In questo caso questa soglia corrisponde a un periodo di 4 settimane.

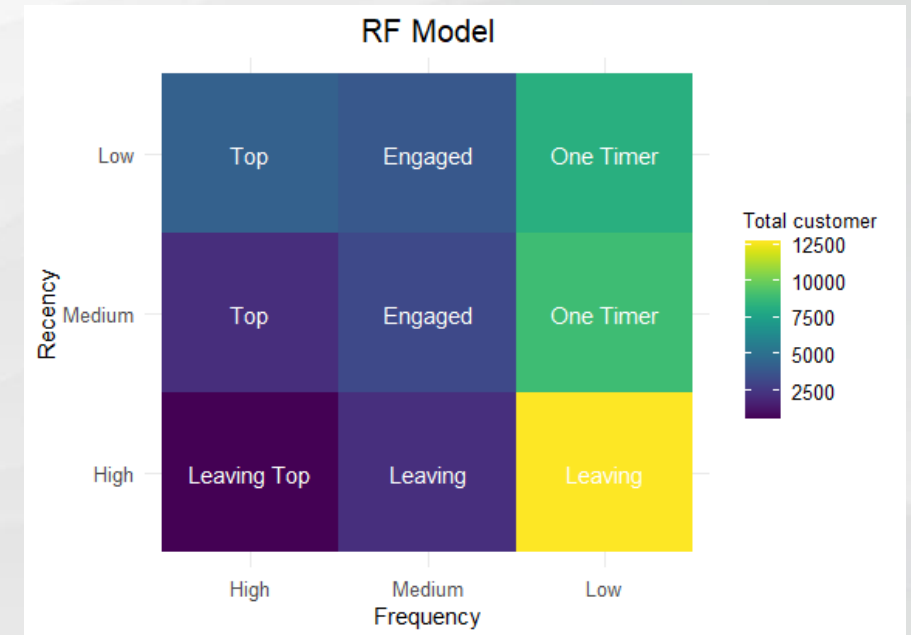
Per semplicità nei calcoli il periodo di riferimento viene esteso a un mese.



# RFM: Risultati

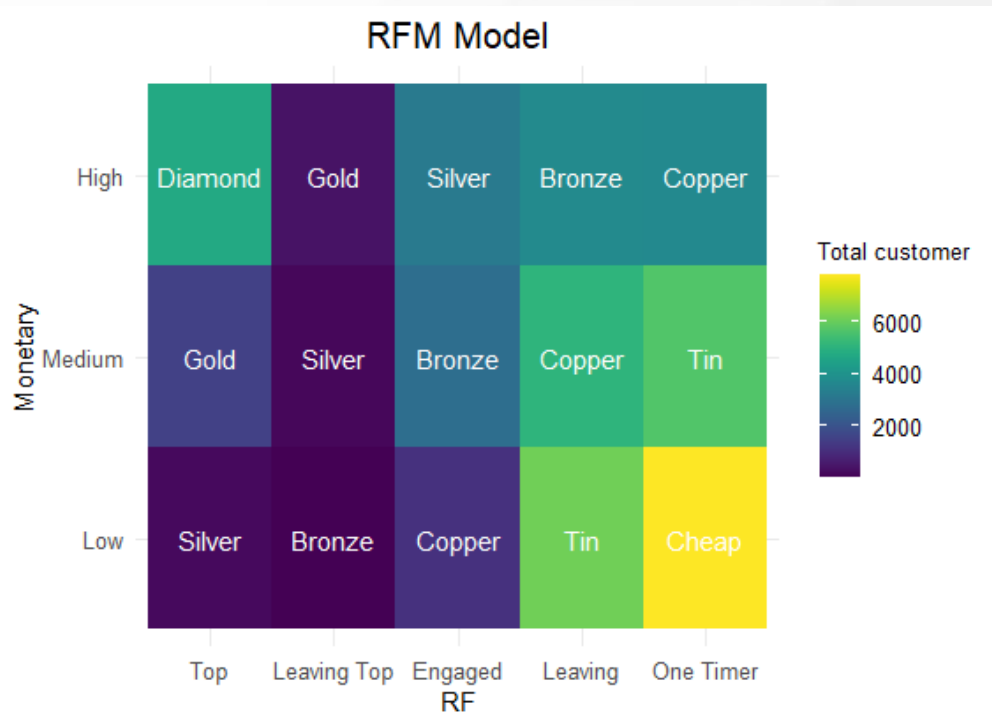
La suddivisione effettuata sulla base della frequency e della recency porta a segmentare la clientela come nell'immagine a destra.

La categoria più presente è quella dei clienti «leaving» mentre quella meno popolosa è quella dei clienti «leaving top».



Incrociando questa segmentazione con la misura Monetary, si ottiene una nuova suddivisione della clientela, come da immagine a sinistra.

Il segmento più numeroso è quello composto dai clienti «cheap».





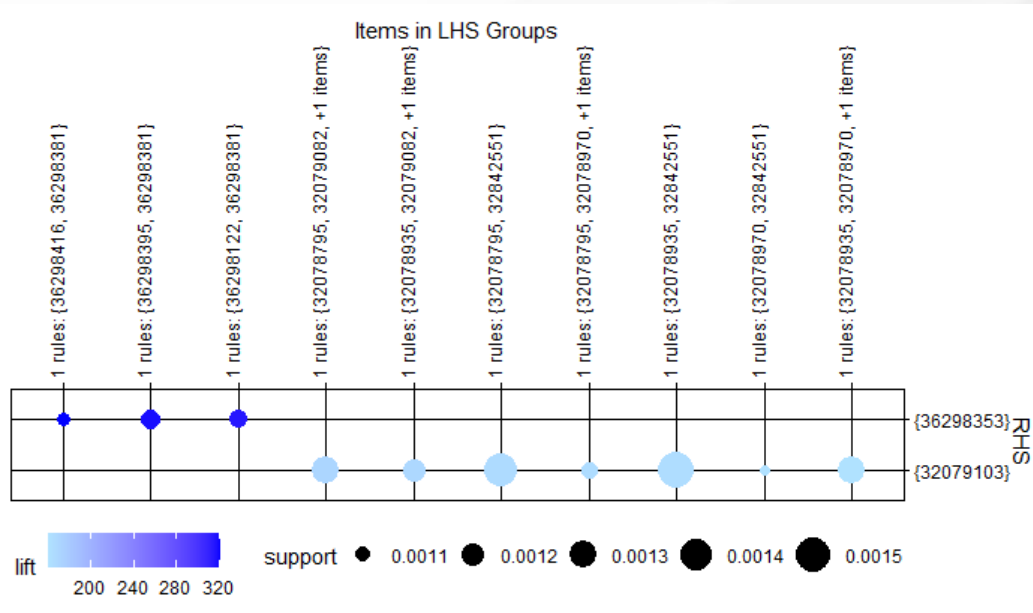
# MBA

La **Market Basket Analysis** serve ad individuare regole di associazione tra i prodotti che vengono acquistati in maniera congiunta. Tramite questa tecnica è possibile identificare prodotti che influenzano il cliente all'acquisto di altri prodotti (Cross-selling).

Per svolgere questa analisi si sono andate a considerare le regole con supporto superiore a 0.001 e confidenza superiore a 0.8, in modo tale da prendere in considerazione solo quelle più rilevanti.

Nel grafico si mostrano le 10 regole associative più significative, evidenziando per ognuna di esse la confidenza il lift e il supporto.

Si può notare che le regole associative in alto a sinistra hanno un lift doppio rispetto a quelle in basso a destra.



# CHURN

Il modello di churn ha come obiettivo di prevedere la probabilità di abbandono di un cliente. In questo modo è possibile avere un'idea più chiara della situazione e intraprendere azioni preventive nei confronti dei clienti a rischio di abbandono.

E' stato identificato un periodo di lookback di 9 mesi e un periodo di holdout di 3 mesi, in modo tale da essere sicuri che il cliente abbia perso fiducia nell'azienda qualora non effettuasse neanche un acquisto da 3 periodi di riferimento consecutivi.

Modelli Utilizzati:

- Regression Trees
- Random Forest
- Logistic Regression
- Lasso Regression

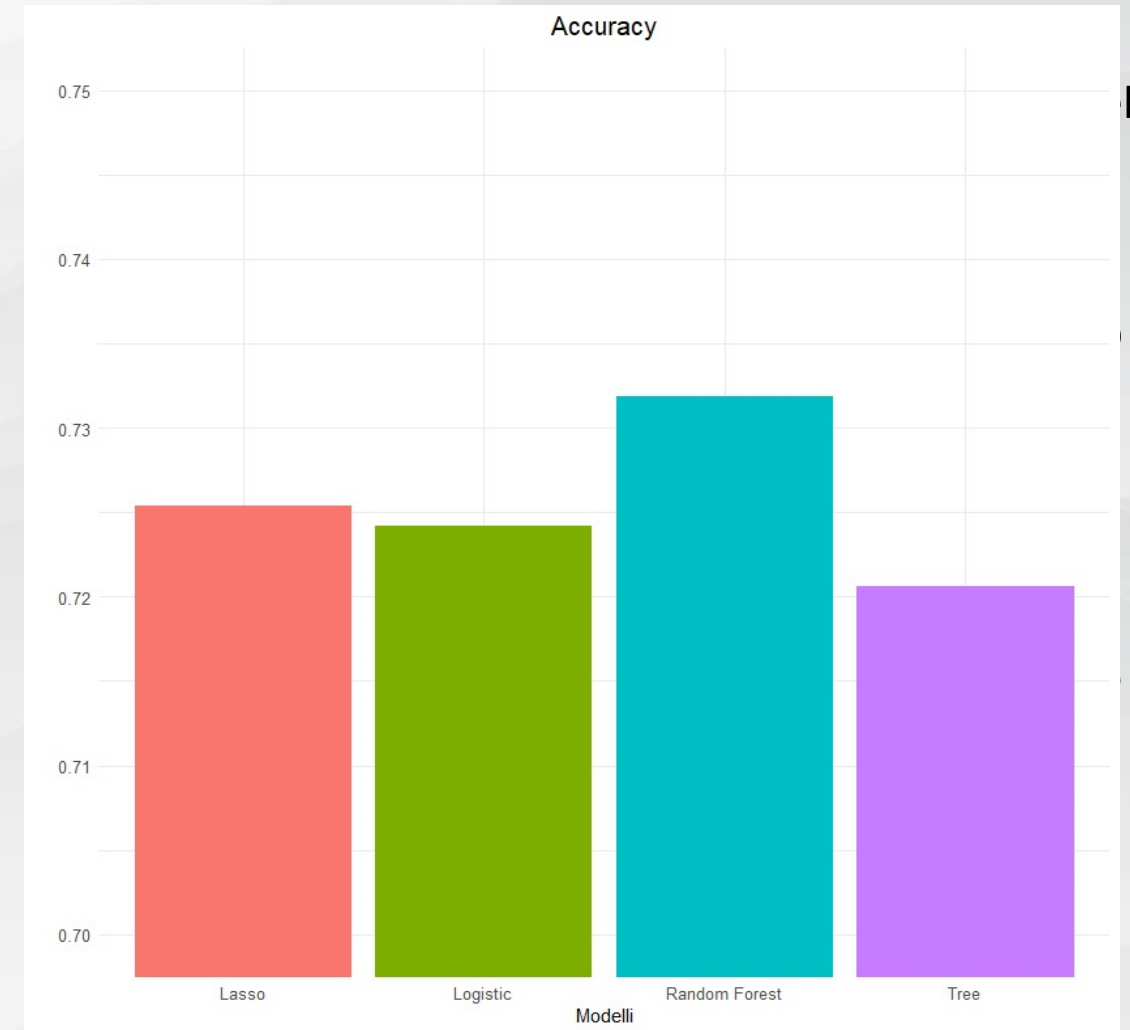


# CHURN

Il modello più performante in termini di accuracy è una **Random Forest** con un numero di alberi pari a 300 e le seguenti variabili esplicative:

- Recency
- Spesa
- Last\_cod\_fid
- Typ\_cli\_account
- Flag\_privacy\_1
- Flag\_direct\_mkt
- Refund\_ratio

Il modello ottiene una accuracy pari a 73.19%, maggiore di 0.65% rispetto al secondo modello più performante che è una regressione logistica.

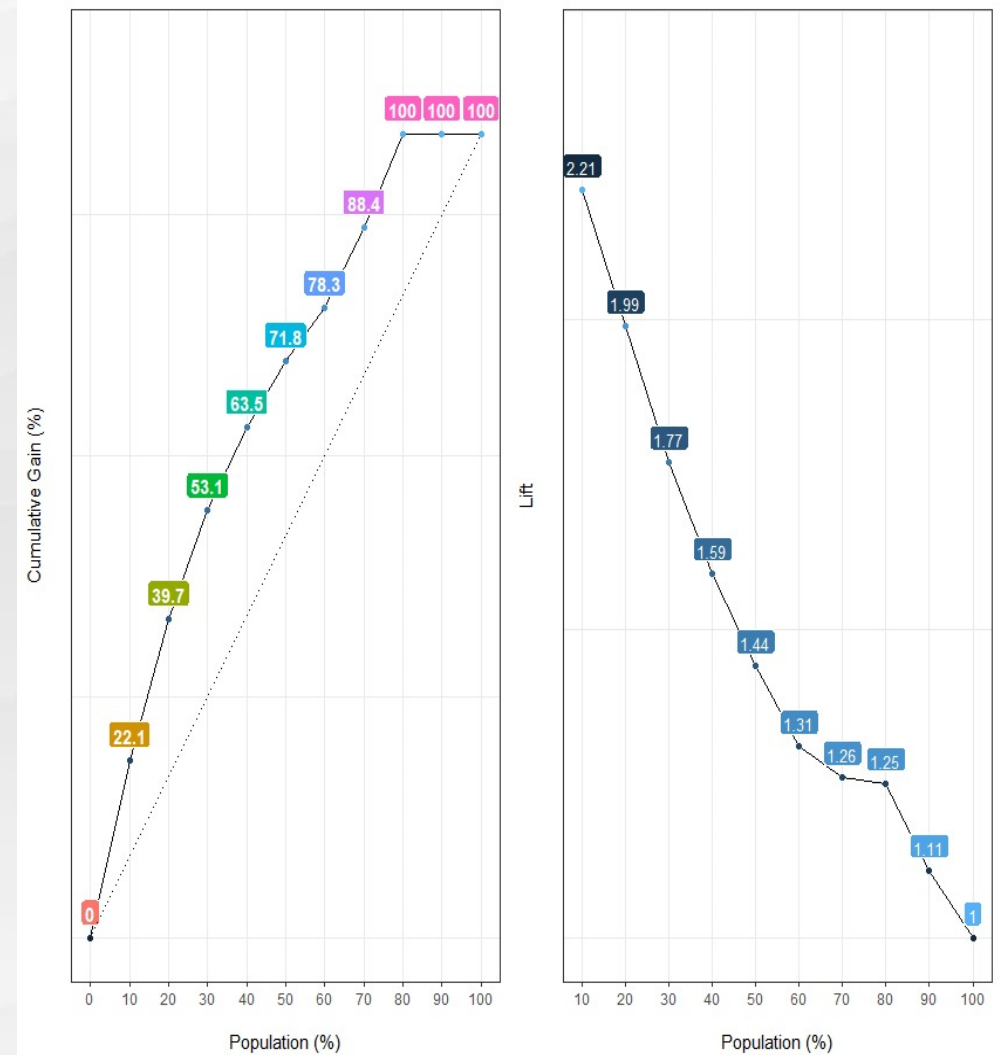


# CHURN

La curva delle cumulative gains mostra che selezionando il 50% dei potenziali churner ritenuti più probabili dal modello, si sarà in grado di intervenire sul 71.8% dei churner totali.

La Curva di lift mostra che selezionando tale soglia di clienti il modello è capace di identificare i churner 1.44 volte meglio di una scelta casuale.

La soglia del 50% funge da trade-off tra il costo dell'intervento sui churner prima che lascino la compagnia e la sua scalabilità per raggiungere il maggior numero di clienti.



# Tipi di azioni Data-Driven

Il lavoro descritto ha come obiettivo quello di poter prendere decisioni guidate dai dati. Le azioni che possono essere prese utilizzando i modelli sviluppati possono essere:

**RFM**



- Creazione di campagne di **email marketing** personalizzate in base ai diversi clienti
- Creazione di **offerte e promozioni** rivolte ai clienti di maggior valore (es: clienti Diamond)

**MBA**



- Ottimizzare la **disposizione dei prodotti** sugli scaffali
- Effettuare **promozioni** i abbinando prodotti in modo tale da favorire il cross-selling

**CHURN**



- **Prevenire l'abbandono del cliente** intraprendendo campagne finalizzate a riconquistare la sua **fiducia**

**Grazie per l'attenzione !**