

Data Science Lab

Case Study

Customer Segmentation

Beatrice Fumagalli – 784549

Matteo Porcino – 748876

Pierluigi Tagliabue – 835211

Contesto

MIFID II – Markets in financial instruments directive (2014/65/EU)

Scopo: Aumento della tutela per gli investitori

- Maggior numero di informazioni per il cliente
- Nuove imposizioni per imprese e intermediari
- Maggiore chiarezza sui costi dei prodotti finanziari
- Migliore adeguatezza del rischio

In breve: maggiore trasparenza verso gli investitori e prodotti finanziari conformi ai propri bisogni e caratteristiche

Customer Segmentation

La segmentazione del cliente è il processo di suddivisione della clientela in gruppi di individui che sono simili in modi specifici rilevanti per il marketing

- Creare e comunicare messaggi di marketing per target specifici
- Selezionare il canale di comunicazione migliore per il segmento
- Identificare modi per migliorare i propri prodotti
- Instaurare una migliore customer relationship

Variabili
Geografiche

Variabili
Demografiche

Variabili
Comportamentali

Variabili
Psicografiche

L'algoritmo di clustering utilizzato è K-Means

Per la scelta del numero di Cluster sono stati utilizzati:

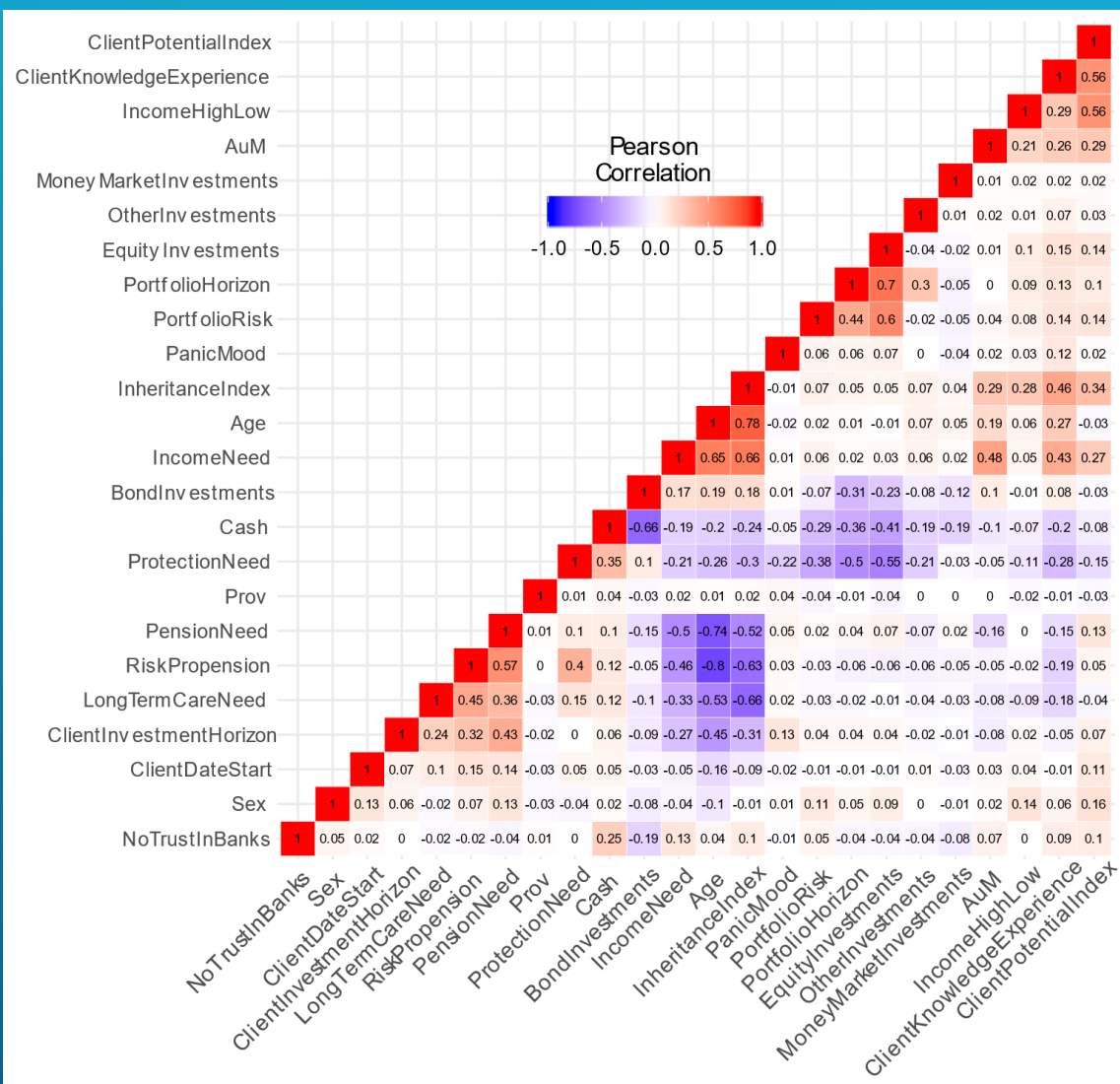
- L'indice di Silhouette
- Il metodo dell'Elbow
- Il decremento della varianza intra-cluster

Dataset

- 25 variabili
- 5000 osservazioni iniziali

```
'data.frame': 5000 obs. of 25 variables:
 $ ClientID      : Factor w/ 5000 levels "Client1","Client10",...: 1 1112 2223 3334 4445 4557 4668 4779 4890 2 ...
 $ RiskPropension : num  0.598 0.501 0.554 0.441 0.541 ...
 $ PortfolioRisk  : num  0.0878 0.0623 0 0.0323 0.0379 ...
 $ ClientInvestmentHorizon : num  20.78 15.23 9.68 9.56 17.94 ...
 $ PortfolioHorizon : num  3.7 2.73 0.5 2.57 2.49 ...
 $ ClientKnowledgeExperience: num  0.286 0.714 0.524 0.619 0.429 ...
 $ ClientPotentialIndex : num  0.1485 0.1528 0.2436 0.0896 0.1626 ...
 $ IncomeHighLow : int  0 0 0 0 0 0 0 0 0 0 ...
 $ Sex            : int  1 1 0 0 0 1 1 1 1 0 ...
 $ AuM            : int  13719 45973 10233 25144 222892 34746 313942 124883 75549 25948 ...
 $ Age            : int  22 61 24 46 57 50 53 80 91 52 ...
 $ IncomeNeed     : num  0.0792 0.2816 0.124 0.2096 0.3201 ...
 $ LongTermCareNeed : num  0.798 0.175 0.883 0.548 0.908 ...
 $ ProtectionNeed  : num  0.204 0.154 0.77 0.226 0.704 ...
 $ PensionNeed     : num  0.634 0.629 0.664 0.218 0.506 ...
 $ InheritanceIndex : num  0.1148 0.5054 0.0699 0.4157 0.3351 ...
 $ PanicMood      : int  0 0 0 0 0 0 0 0 0 0 ...
 $ ClientDateStart : int  2015 2016 2015 2015 2015 2015 2014 2015 2016 2015 ...
 $ NoTrustInBanks  : int  0 0 0 0 0 0 0 0 1 0 ...
 $ Prov           : Factor w/ 102 levels "AG","AL","AN",...: 60 76 17 74 51 89 100 NA 85 15 ...
 $ BondInvestments : num  0.573 0.534 0 0.643 0.171 ...
 $ EquityInvestments : num  0.427 0.447 0 0.122 0.27 ...
 $ MoneyMarketInvestments : num  0 0.0124 0 0.0407 0.5332 ...
 $ OtherInvestments : num  0 0 0 0.1943 0.0045 ...
 $ Cash            : num  0 0.007 1 0.0002 0.0215 ...
```

Matrice di correlazione



Variabili correlate positivamente:

- Age
- Inheritance Index
- Income Need

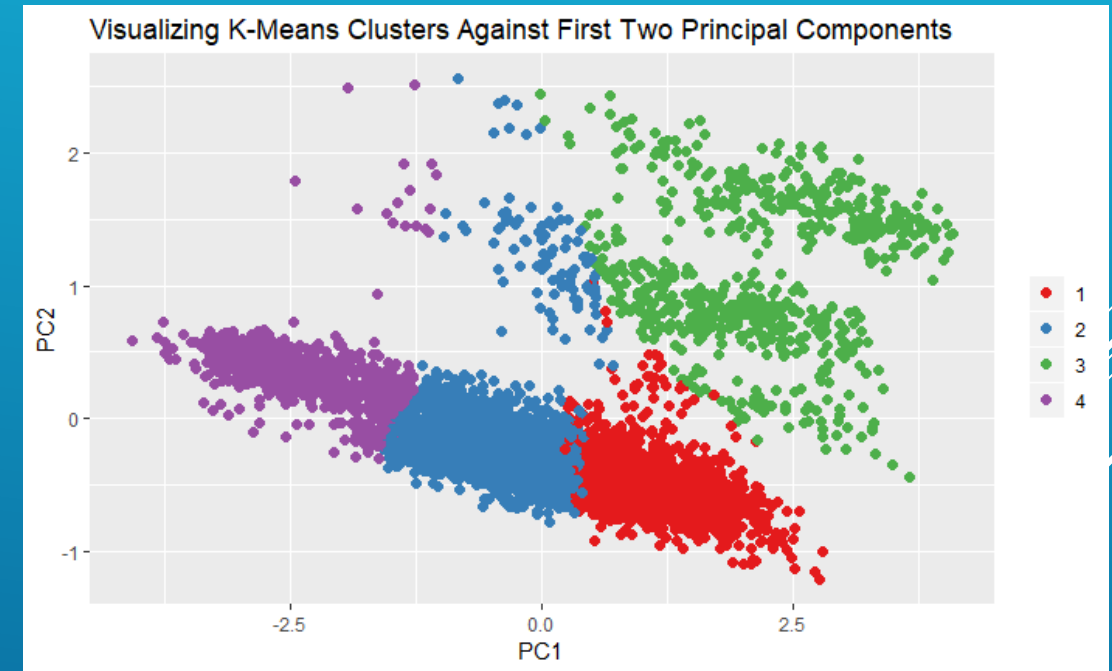
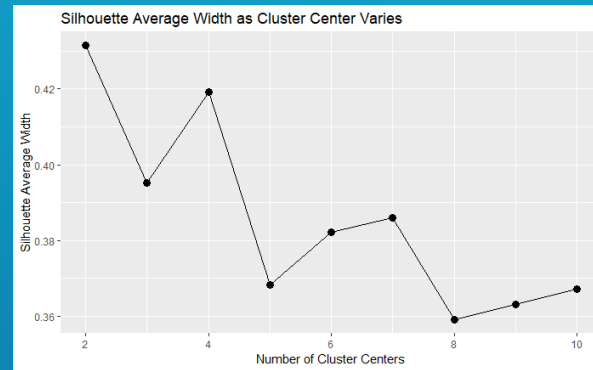
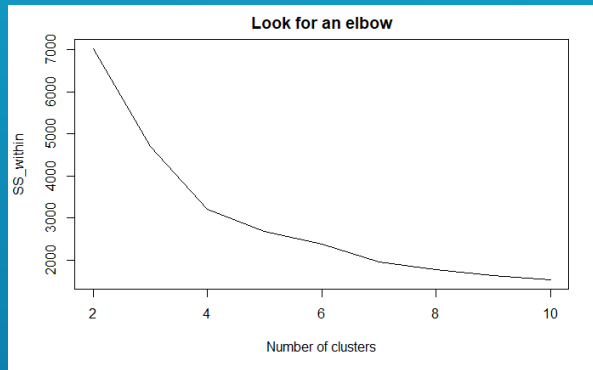
Variabili correlate negativamente:

- Age
- Risk Propension
- Pension Need

Clustering 1 – correlazione positiva

Variabili correlate positivamente - Age, IncomeNeed, InheritanceIndex

Il numero di Cluster suggerito è 4 - Dimensioni (639, 887, 1442, 1909) – bontà della clusterizzazione: 78%



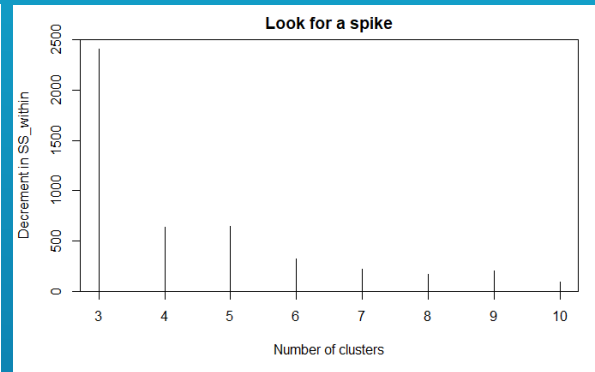
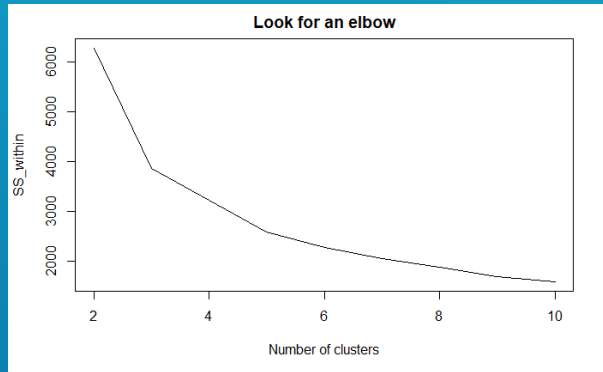
- 1) clienti over 60, medio/alto bisogno di reddito e alto bisogno di ottimizzazione successoria/fiscale
- 2) clienti tra 40 e 60, bisogno medio di reddito e di ottimizzazione successoria/fiscale
- 3) clienti over 60, elevato bisogno di reddito e alto bisogno di ottimizzazione successoria/fiscale
- 4) clienti under 40, basso bisogno di reddito e di ottimizzazione successoria/fiscale

	x1	x2	x3	x4
Age	0.9570606	-0.30550501	0.7106575	-1.410352
IncomeNeed	0.2002391	-0.38080293	2.0971939	-1.016797
InheritanceIndex	0.7069353	-0.08554861	0.9630185	-1.658915

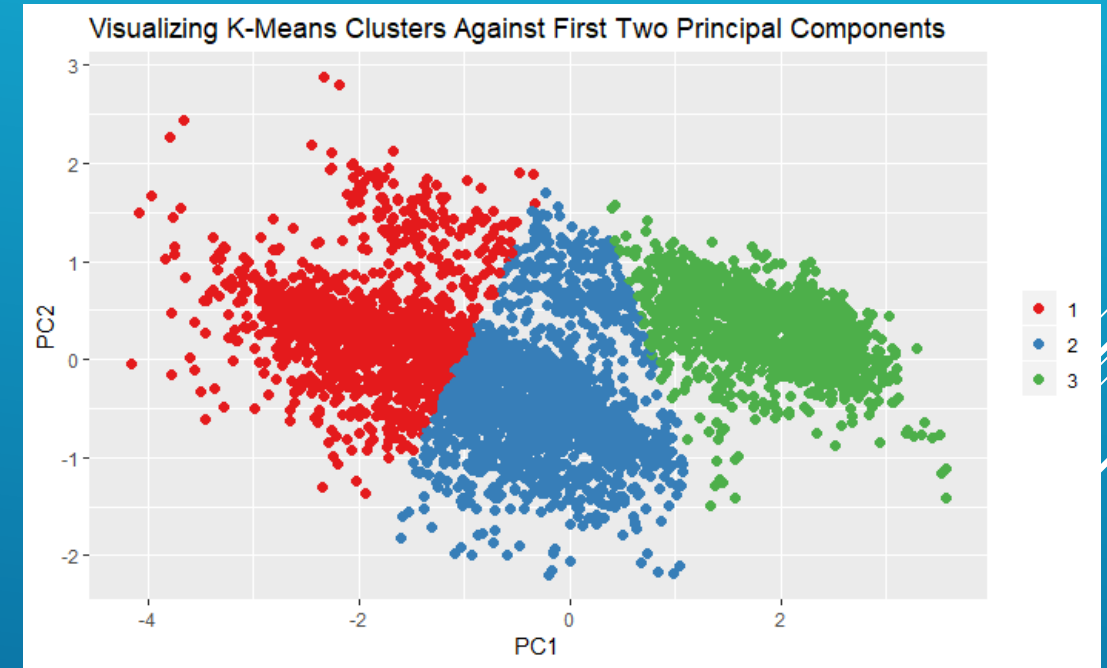
Clustering 2 – correlazione negativa

Variabili correlate negativamente - Age, RiskPropension, PensionNeed

Il numero di Cluster suggerito è 3 - Dimensioni (2045, 1548, 1284) – bontà della clusterizzazione: 73,6%



- 1) clienti under 40 con maggior propensione al rischio e maggior bisogno di previdenza integrativa
- 2) clienti tra 40 e 60 con media propensione al rischio e bisogno medio di previdenza integrativa
- 3) clienti over 60 bassa propensione al rischio e il minor bisogno di previdenza integrativa



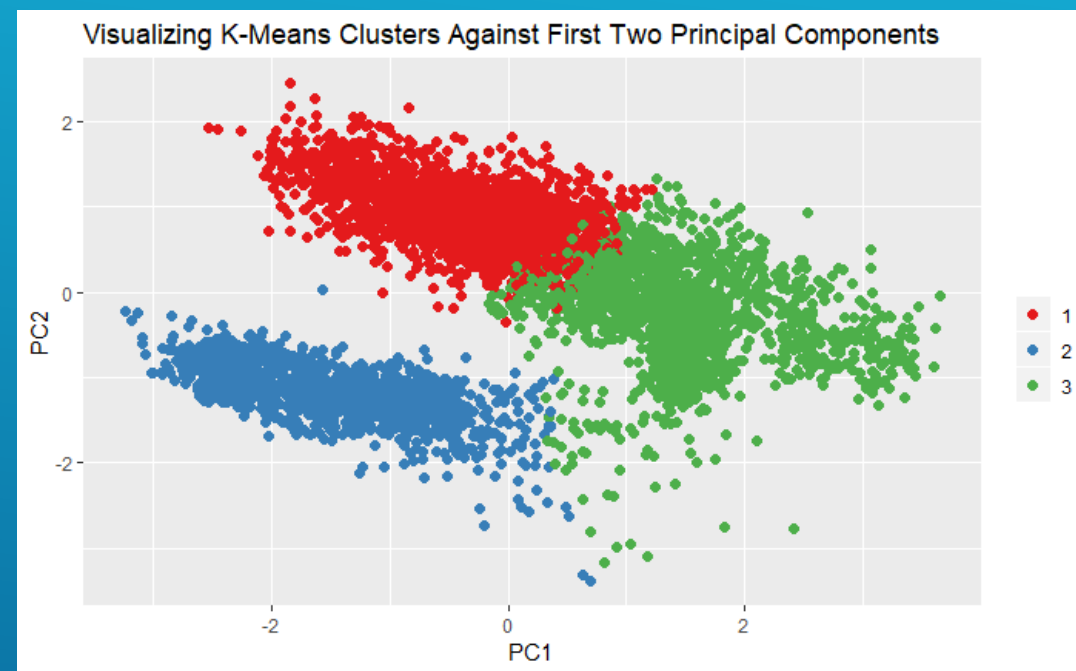
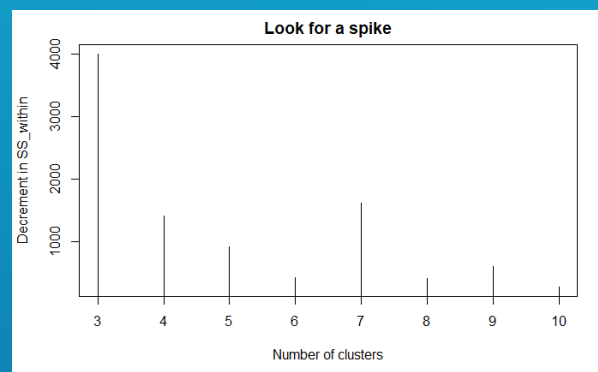
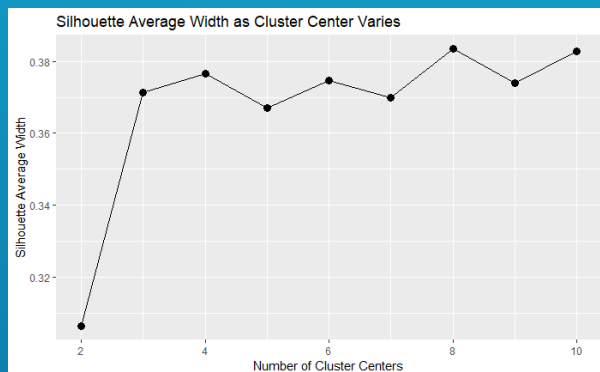
	x1	x2	x3
Age	-1.1953738	-0.08426688	1.1028332
RiskPropension	1.2772103	-0.12544640	-0.8936693
PensionNeed	0.7850503	0.45864324	-1.2570607

Clustering 3 – bisogni dei clienti – IncomeNeed, LongTermCareNeed, ProtectionNeed, PensionNeed

Bisogni personali della clientela:

reddito, cure a lungo termine da anziano, protezione del capitale o della persona e previdenza integrativa

Il numero di Cluster suggerito è 3 - Dimensioni (1957, 1176, 1744) – bontà della clusterizzazione: 53,1%



Need\Age	20 - 45	45 - 60	Over 60
Income	Basso	Medio	Alto
LongTermCare	Alto	Medio	Basso
Protection	Alto	Basso	Medio
Pension	Alto	Alto	Basso

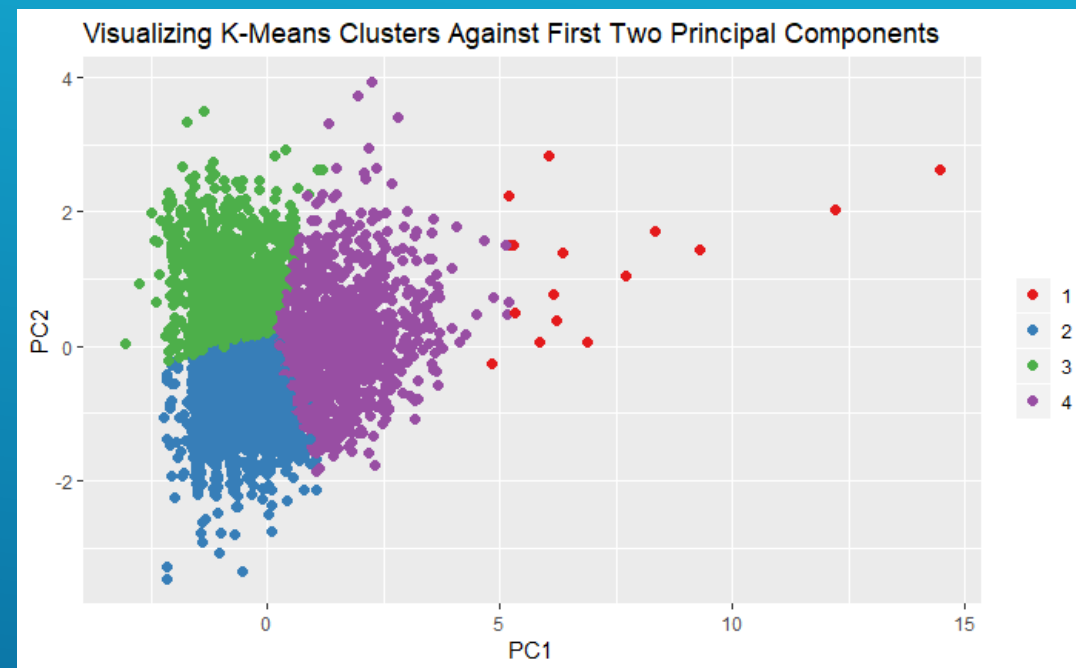
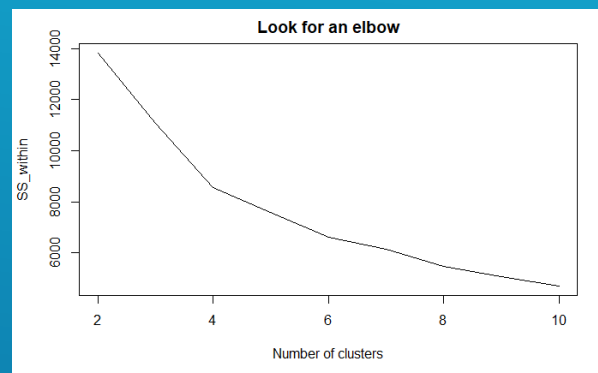
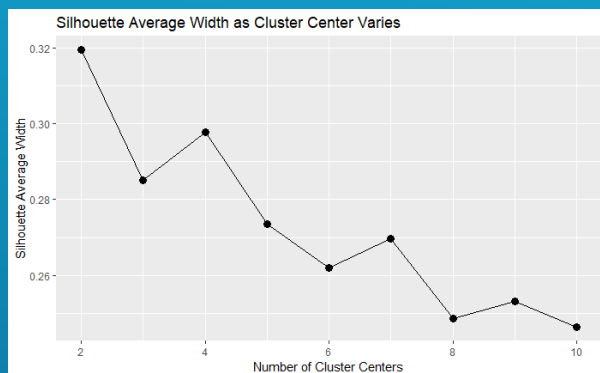
	X1	X2	X3
IncomeNeed	-0.3140271	-0.6853588	0.8145258
LongTermCareNeed	0.1725429	0.5408026	-0.5582858
ProtectionNeed	-0.7567218	1.5712513	-0.2103709
PensionNeed	0.6220751	0.6275600	-1.1212222

Clustering 4 – profilo finanziario – Risk Propension, ClientKnowledge, ClientPotentialIndex, AuM

Profilo finanziario del cliente:

Propensione al rischio, livello di cultura finanziaria del cliente, potenziale di crescita del cliente, denaro investito

Il numero di Cluster suggerito è 4 - Dimensioni (21, 1763, 1575, 1518) – bontà della clusterizzazione: 56,2%



	cluster 1	cluster 2	cluster 3	cluster 4
RiskPropension	Medio	Basso	Alto	Medio
ClientKnowledge	Alto	Medio	Basso	Alto
ClientPotential	Alto	Basso	Medio	Alto
AuM	Molto alto	Basso	Medio/basso	Alto

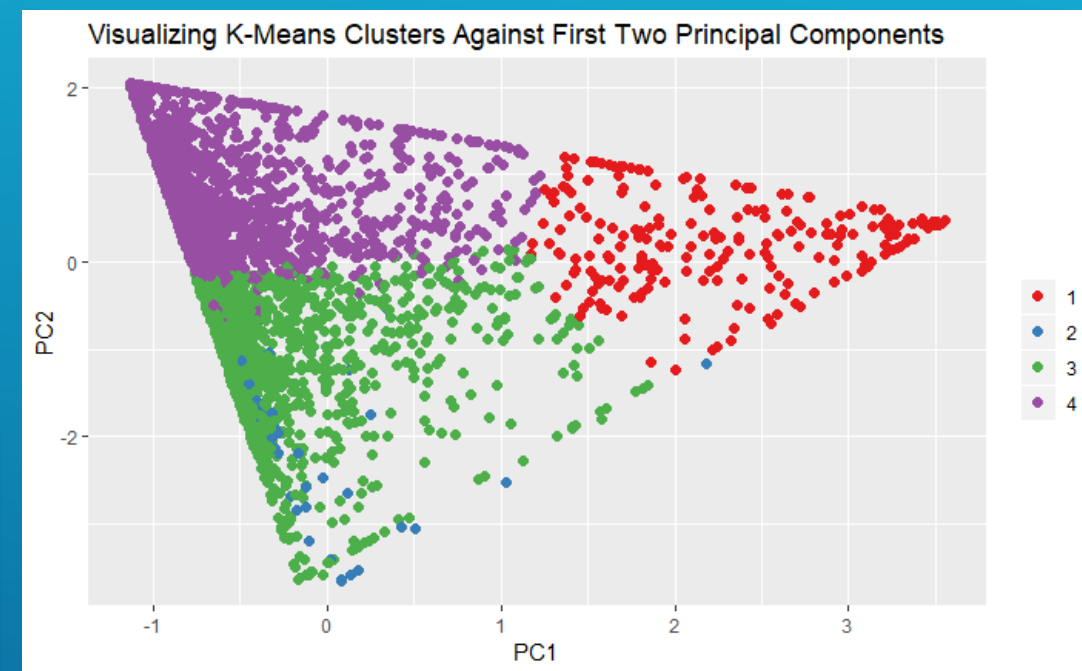
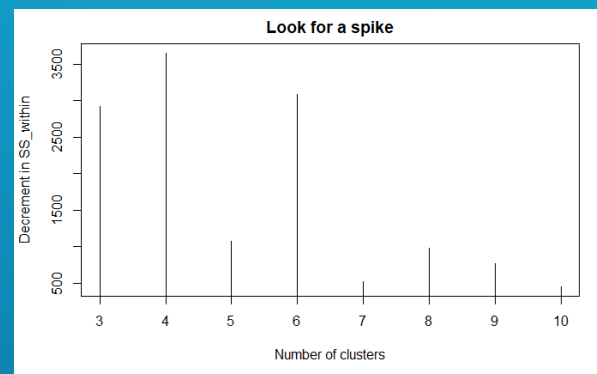
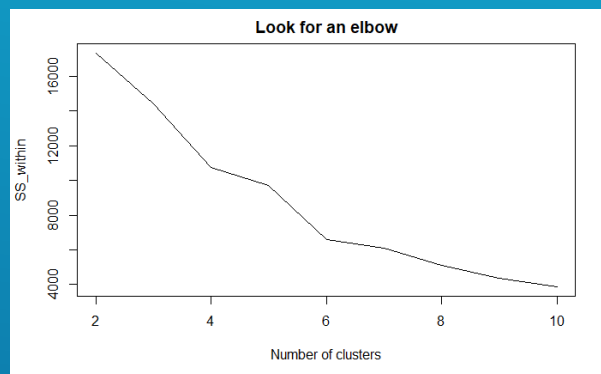
	x1	x2	x3	x4
RiskPropension	-0.07015624	-0.7639277	1.0270131	-0.1773833
ClientKnowledgeExperience	1.37083714	-0.2763690	-0.6950087	1.0231157
ClientPotentialIndex	1.80544032	-0.6475522	-0.3284056	1.0678255
AuM	11.36599826	-0.1778697	-0.2341128	0.2922436

Clustering 5 – prodotti finanziari– Bond, Equity, MoneyMarket, Other, Cash

Prodotti finanziari offerti:

Investimenti in azioni, obbligazioni, strumenti monetari, altro, di cassa

Il numero di Cluster suggerito è 4 - Dimensioni (653, 73, 1776, 2375) – bontà della clusterizzazione: 57,1%



	Portafoglio 1	Portafoglio 2	Portafoglio 3	Portafoglio 4
Azioni	25 %	15 %	20 %	45 %
Obbligazioni	15 %	5 %	50 %	25 %
Strumenti monetari	5 %	70 %	0 %	5 %
altro	5 %	5 %	15 %	5 %
cassa	50 %	5 %	15 %	20 %

	x1	x2	x3	x4
BondInvestments	-1.5698826	-1.1624937	-0.43146000	0.79000771
EquityInvestments	-0.9741624	-0.5724334	0.97796956	-0.44587715
MoneyMarketInvestments	-0.4197332	6.4220174	0.00898437	-0.08870642
OtherInvestments	-0.4242789	-0.3415262	0.28216182	-0.08384584
Cash	2.3477196	-0.4649089	-0.36901974	-0.35526042

AdviseOnly™

Assegnazione portafogli – clustering 4 e 5

Clustering 4 – profilo finanziario

Cluster 1: propensione al rischio media, AuM elevato e alto potenziale di crescita

Cluster 2: bassa propensione al rischio e poco capitale investito, basso potenziale di crescita

Cluster 3: clienti avventati con poco capitale investito e basso potenziale di crescita

Cluster 4: ottima cultura finanziaria, grande quantità di denaro investito, ottimo potenziale di crescita

Clustering 5 – prodotti finanziari

Portafoglio 1: il denaro investito in prodotti finanziari e monetari è la metà di quello finalizzato all'investimento ed è investito in strumenti finanziari molto rischiosi

Portafoglio 2: maggior parte della liquidità è investita in strumenti monetari (breve termine, basso rischio)

Portafoglio 3: portafoglio più rischioso con grossa parte di liquidità investita in strumenti finanziari ad alto rischio

Portafoglio 4: portafoglio rischioso con grande quantità di denaro investito in strumenti finanziari e piccola parte in strumenti monetari

Grazie per l'attenzione
