Analisi del comportamento del consumatore

Dataset: Mall_Customers

Source: kaggle.com

Descrizione del dataset

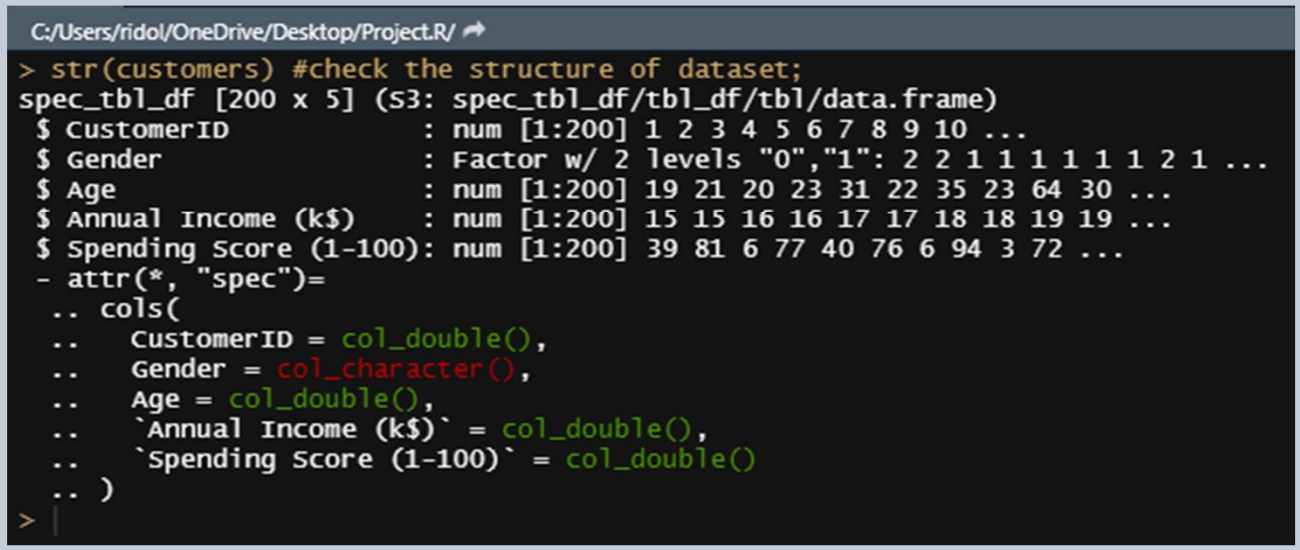
Customer_ID: codice identificativo del cliente del centro commerciale

Gender: genere del cliente

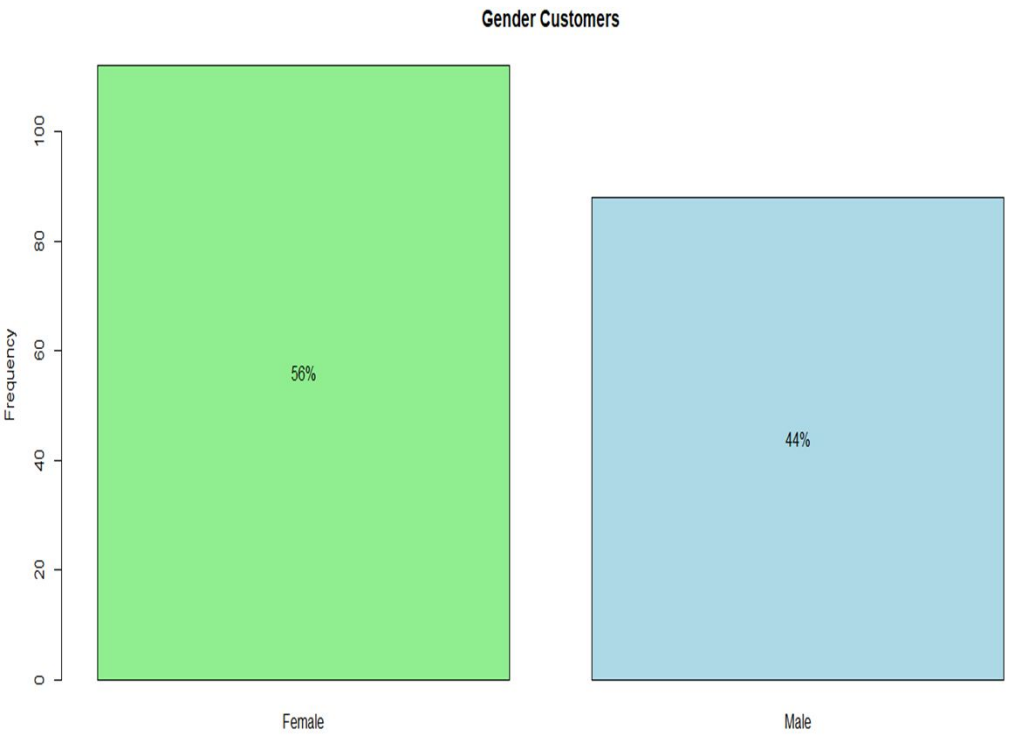
Age: Età del cliente

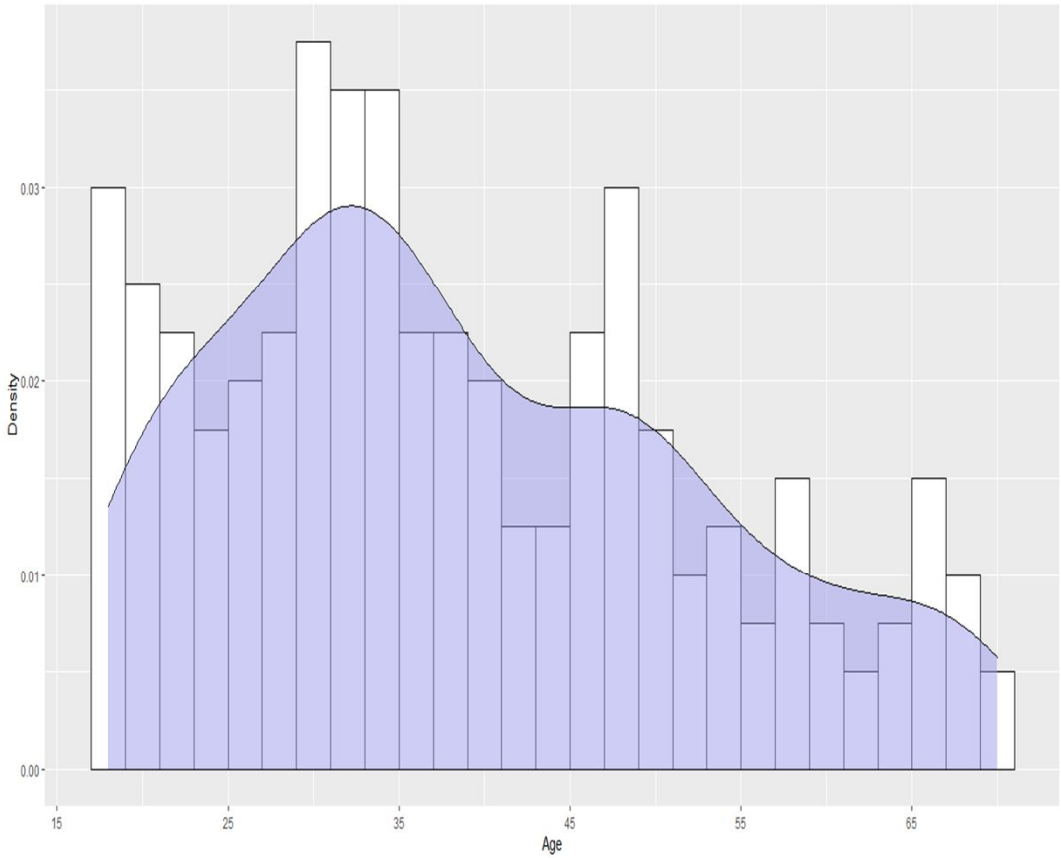
Annual_Income: reddito annuale percepito dal cliente in migliaia di dollari $

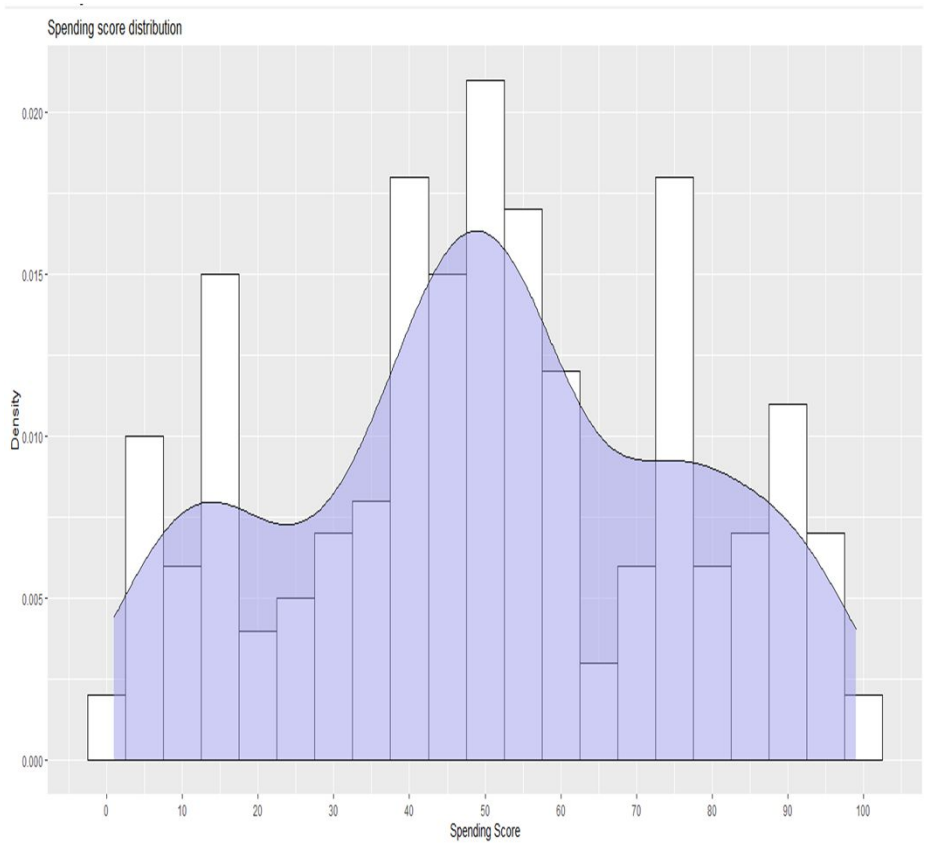
Spending_Score: punteggio attribuito ad ogni cliente in base ai suoi comportamenti di acquisto(0-100)

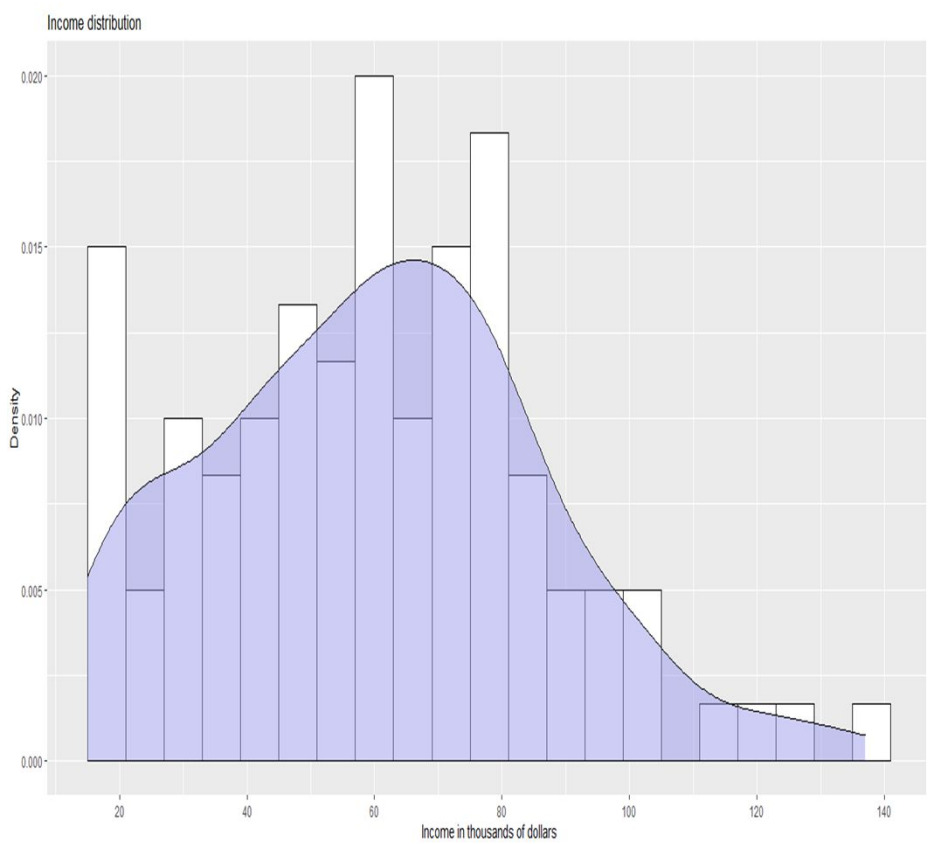


Analisi esplorativa






Spending score distribution
Density
0.020
0.015
0.010
0.005
0.000
0
10
20
30
40
50
60
70
80
90
100
Spending Score


Income distribution
Density
0.020
0.015
0.010
0.005
0.000
20
40
60
80
100
120
140
Income in thousands of dollars

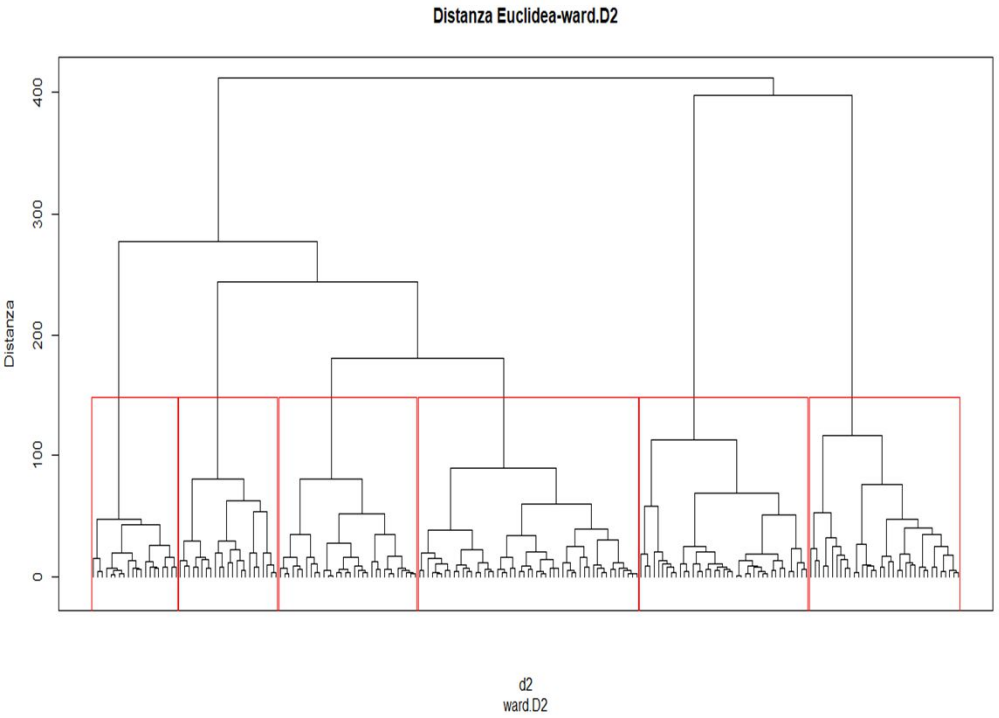
Cluster analysis

- Obiettivo: raggruppare i clienti per individuare quali tra essi possono essere potenzialmente influenzabili da politiche di marketing

- Variabili considerate:

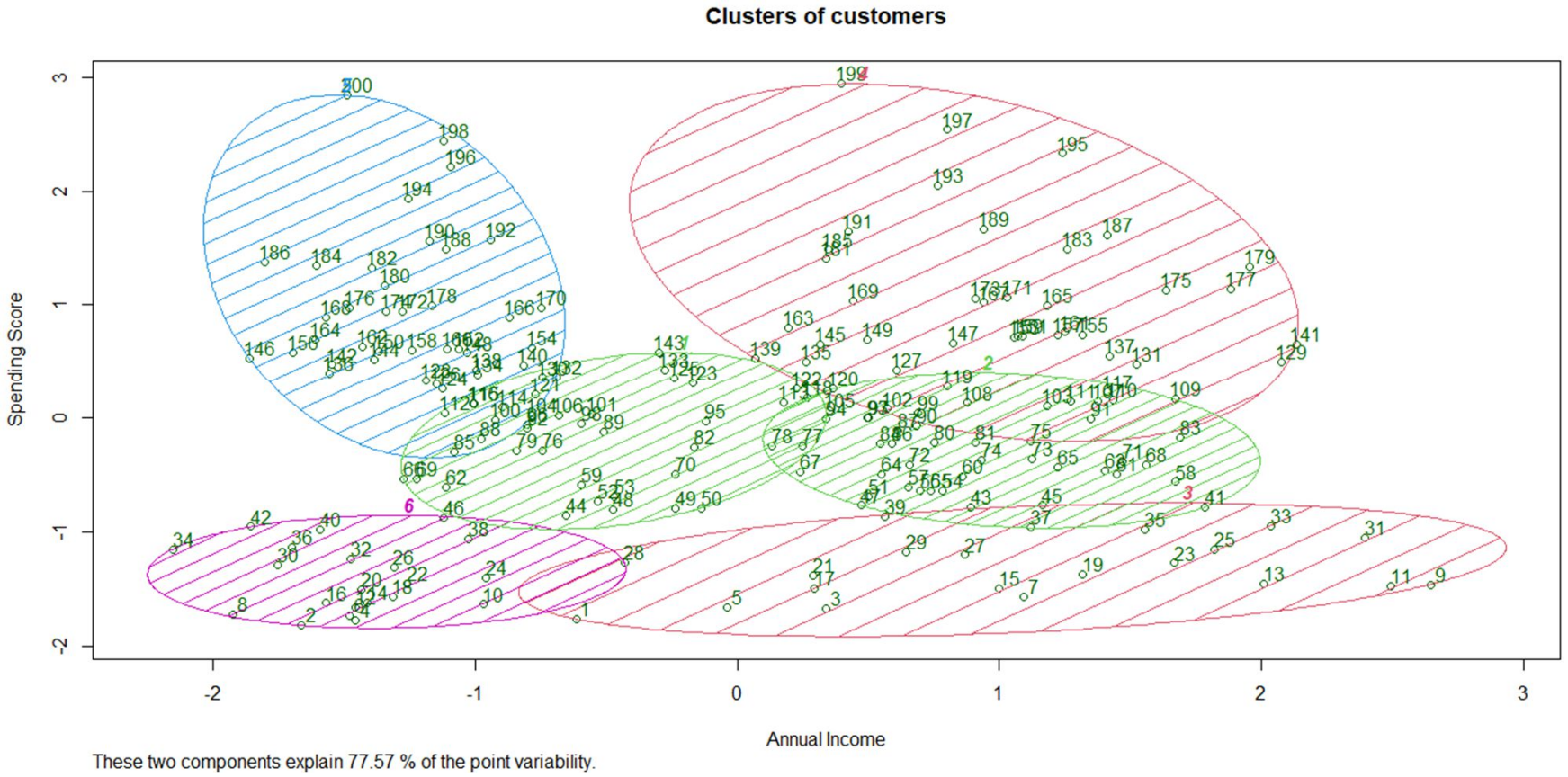
(clustering gerarchico)	(k-means algorithm)
Age	*Age*
Annual_Income	*Annual_Income*
Spending_Score	*Spending_Score*

Risultati Cluster Analysis

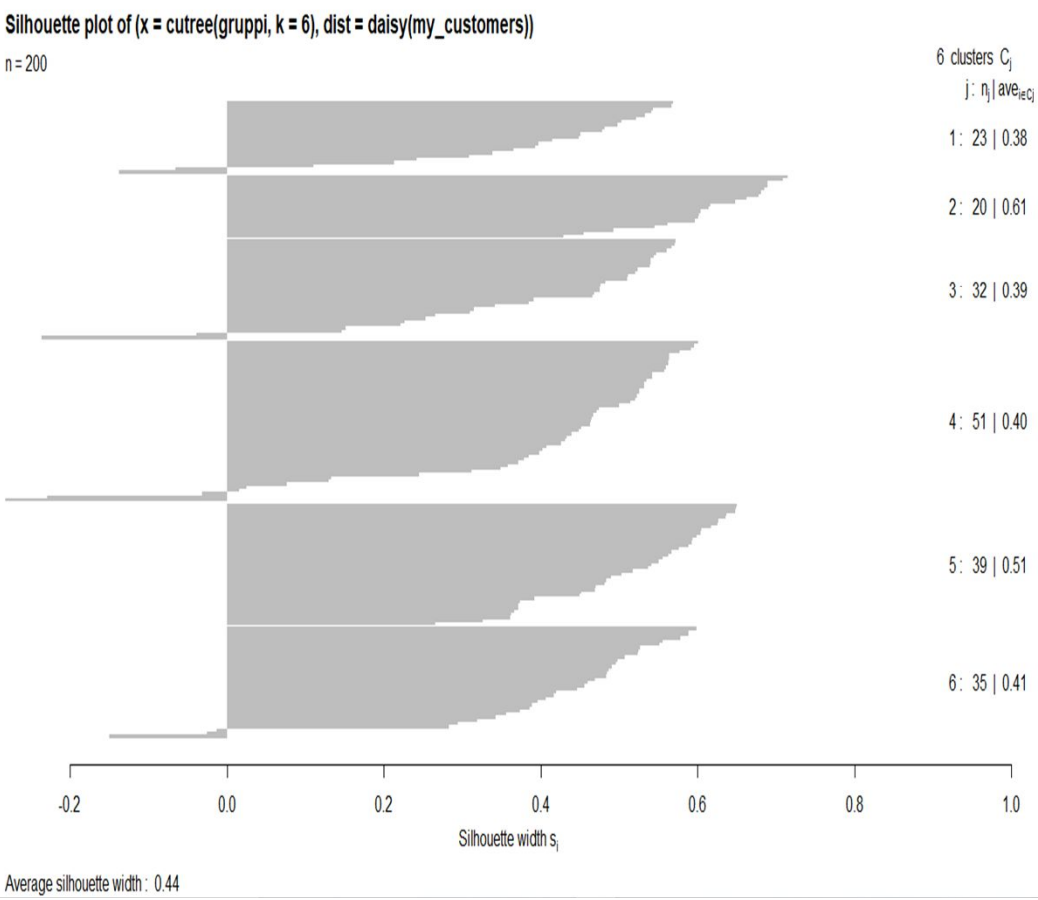


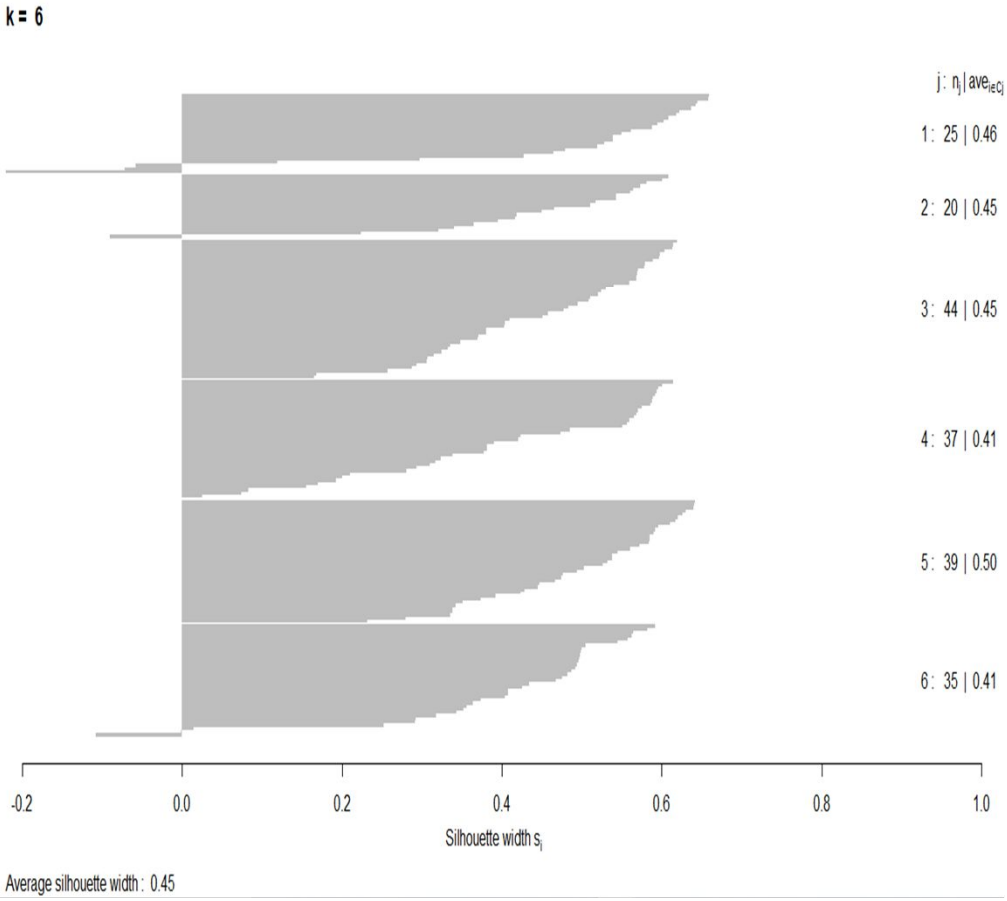
	classi	count	mean_Age	mean_AnnualIncome	mean_SpendingScore
1	1	23	45.21739	26.30435	20.91304
2	2	20	24.85000	24.95000	81.00000
3	3	32	24.53125	54.18750	50.25000
4	4	51	53.21569	55.11765	49.47059
5	5	39	32.69231	86.53846	82.12821
6	6	35	41.68571	88.22857	17.28571

‘K-Means algorithm’

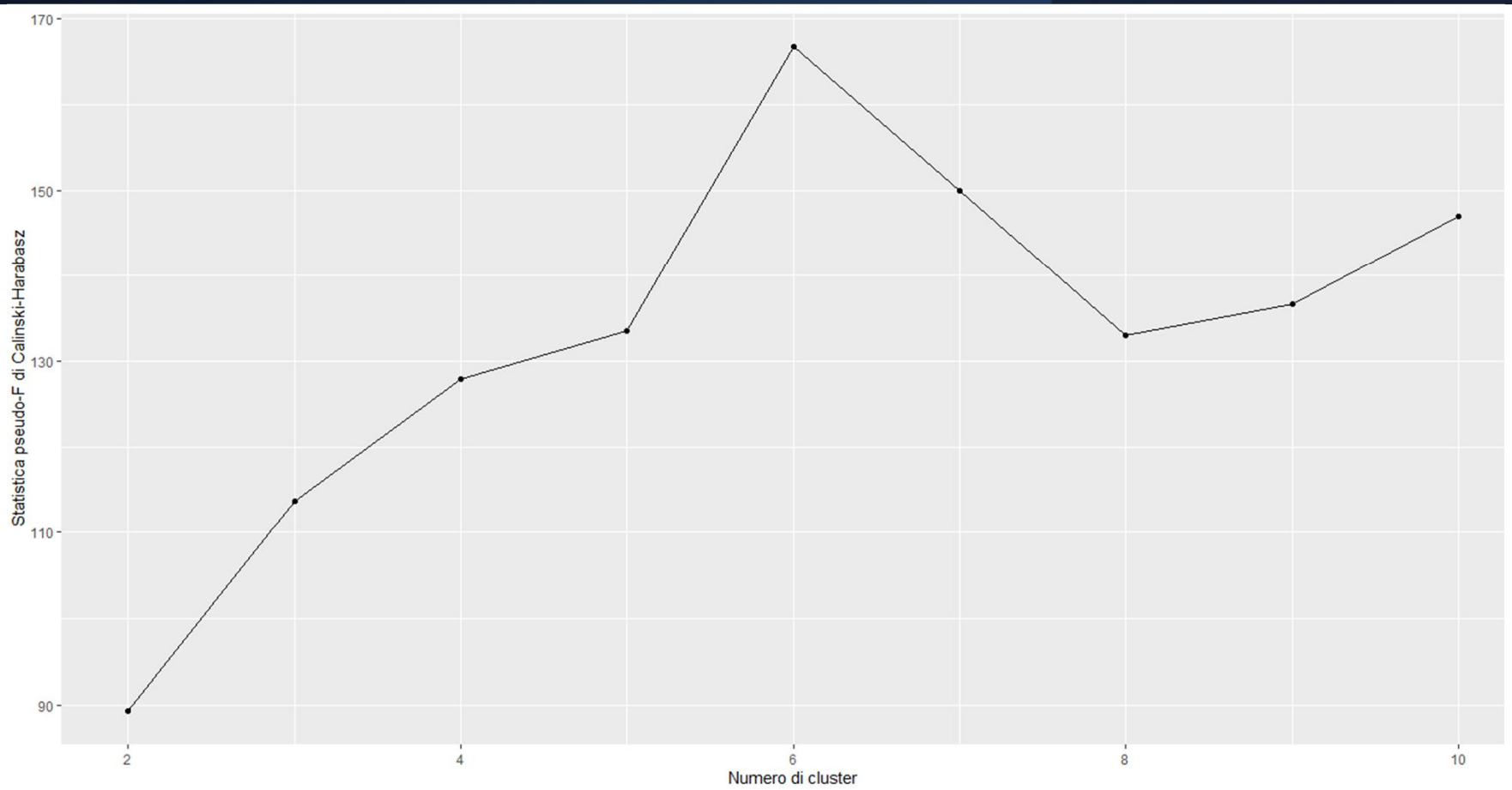


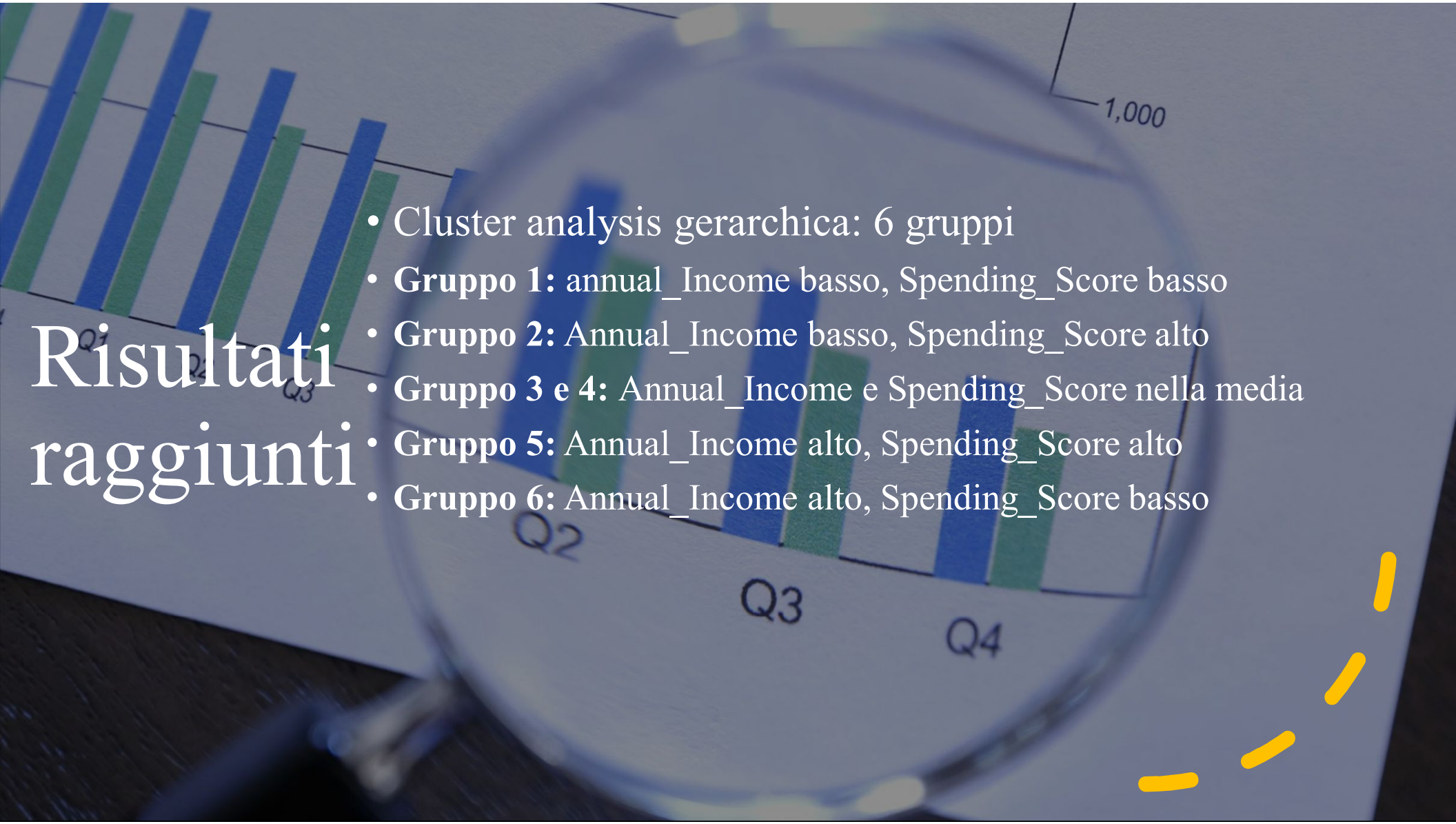
Silhouette_score: gerarchico e k-means





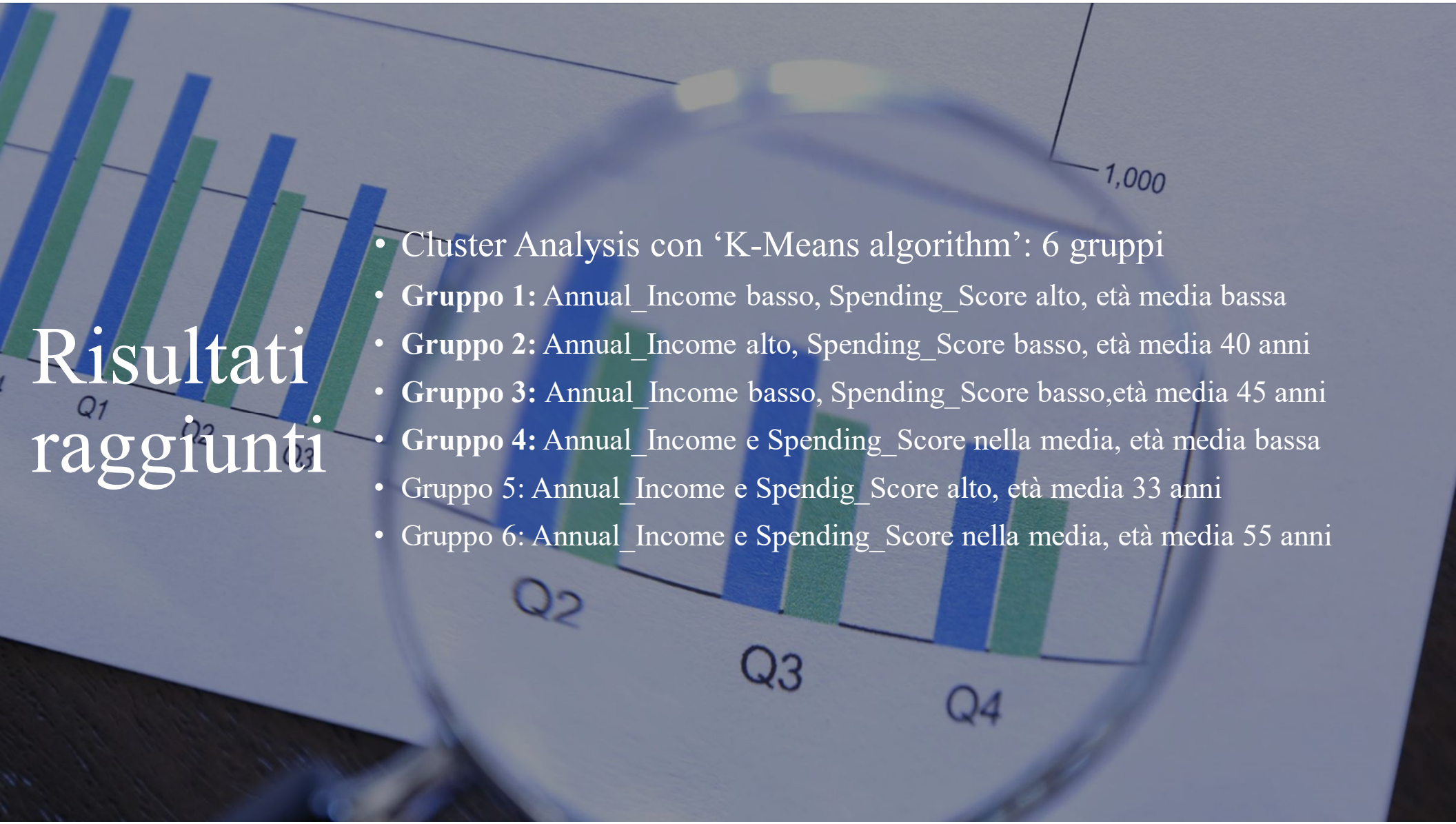
Pseudo F-index





Risultati raggiunti

- Cluster analysis gerarchica: 6 gruppi
- **Gruppo 1:** annual_Income basso, Spending_Score basso
- **Gruppo 2:** Annual_Income basso, Spending_Score alto
- **Gruppo 3 e 4:** Annual_Income e Spending_Score nella media
- **Gruppo 5:** Annual_Income alto, Spending_Score alto
- **Gruppo 6:** Annual_Income alto, Spending_Score basso



Risultati raggiunti

- Cluster Analysis con ‘K-Means algorithm’: 6 gruppi
- **Gruppo 1:** Annual_Income basso, Spending_Score alto, età media bassa
- **Gruppo 2:** Annual_Income alto, Spending_Score basso, età media 40 anni
- **Gruppo 3:** Annual_Income basso, Spending_Score basso,età media 45 anni
- **Gruppo 4:** Annual_Income e Spending_Score nella media, età media bassa
- Gruppo 5: Annual_Income e Spendig_Score alto, età media 33 anni
- Gruppo 6: Annual_Income e Spending_Score nella media, età media 55 anni

Grazie per l'attenzione

- Lavoro realizzato da *Ridolfi Lorenzo*