Customer Segmentation : Clustering and Analysis

Table of contents :

- 1.Introduction to project
- 2.Data Analysis-Descriptive,Exploratory
- 3.Data Analysis - Clustering
- 4.Conclusions

1.Introduction:

It’s a small customer segmentation project based on information from a mall.

Data set used in this project contains **200 unique customers** and their **4 demographic characteristics**:

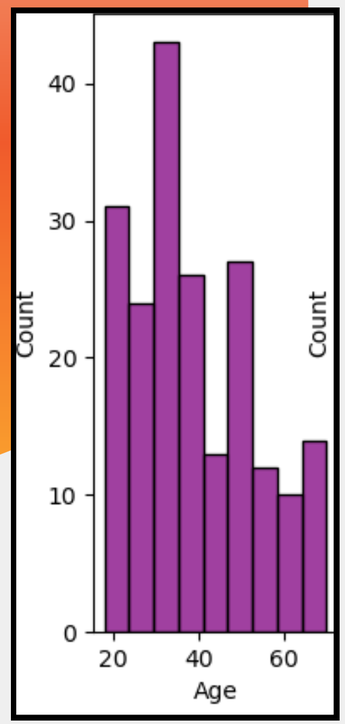
- age, gender, annual income, spending score

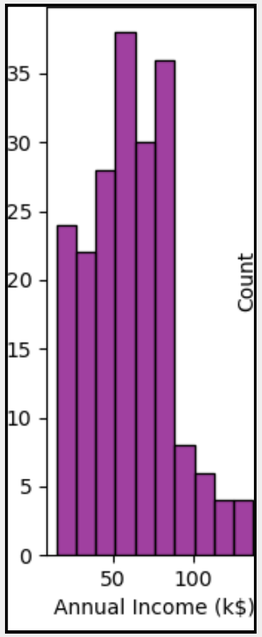
The purpose of this analysis is to collect informations about customers and **improve marketing strategies and decisions** regarding the examined mall.

2.Data Analysis - Descriptive,Exploratory

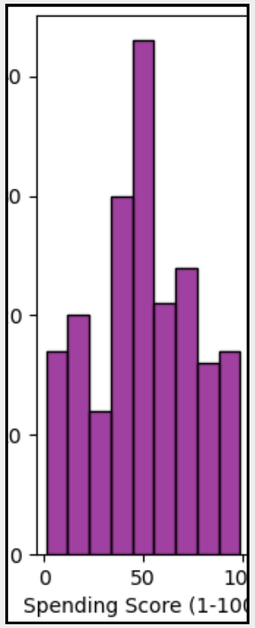
From 200 customers collected
4 features:

Age:
mean is 39 years



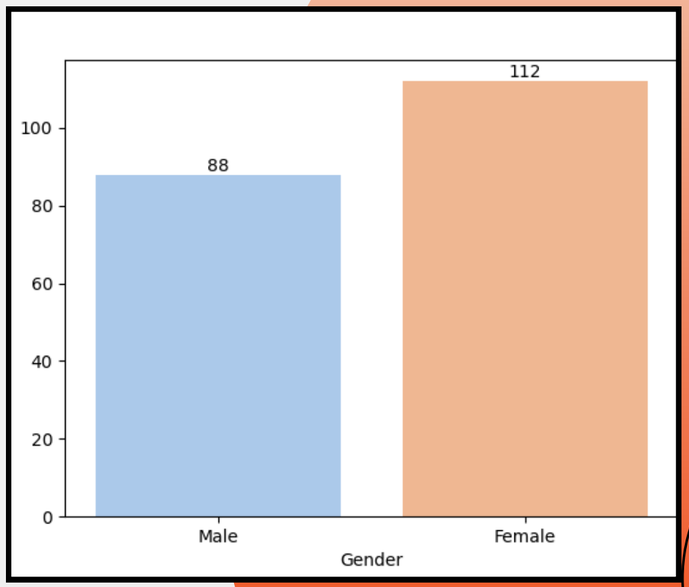


Annual Income(k$):
mean is 60k$



Spend Score(1-100):
mean is 50

Gender:



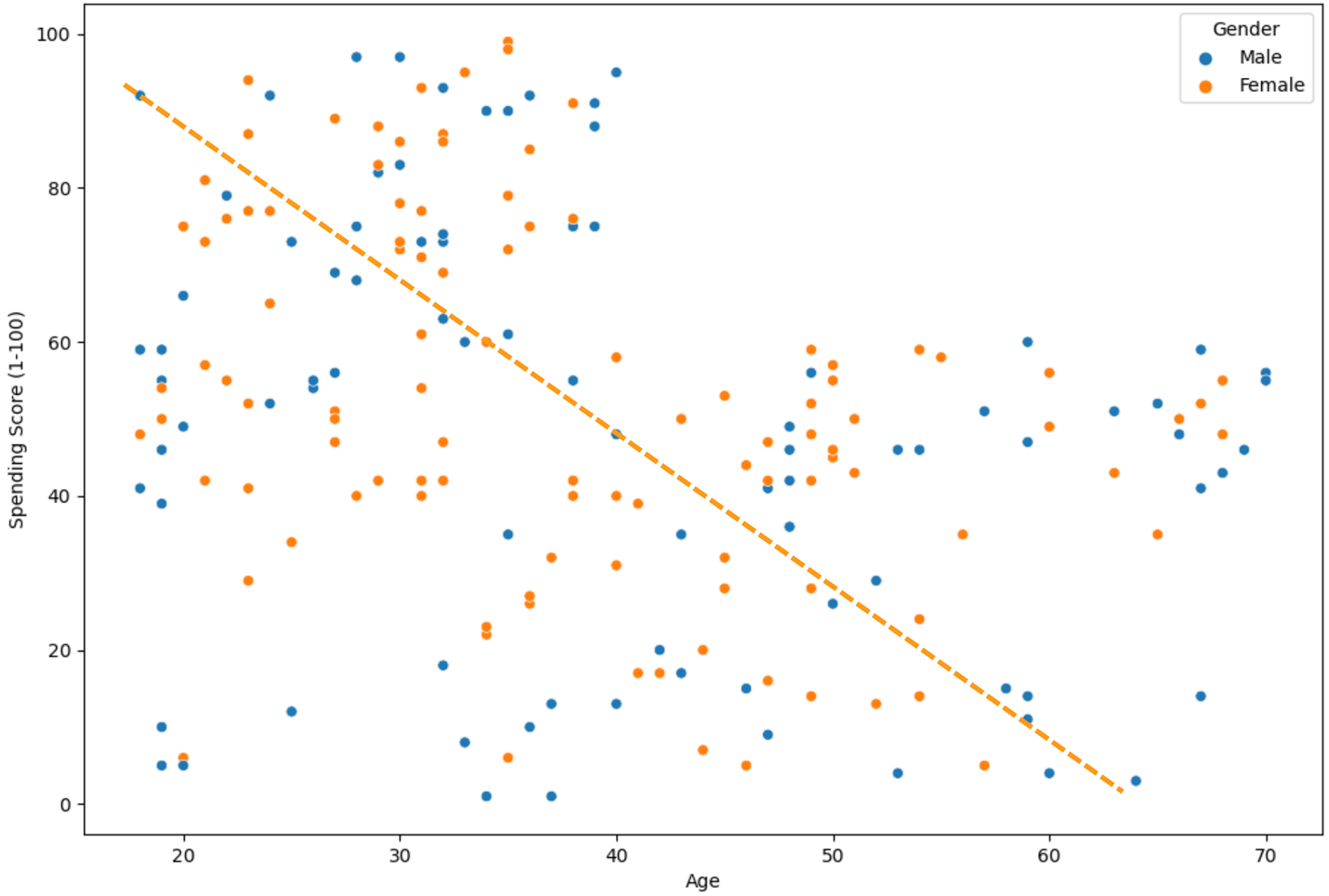
Correlations:

Using Spearman's correlation and heatmap to visualize it, a negative correlation appears between a feature "Age" and "Spending Score" and it is statistically significant.

Age-Spending Score Correlation:

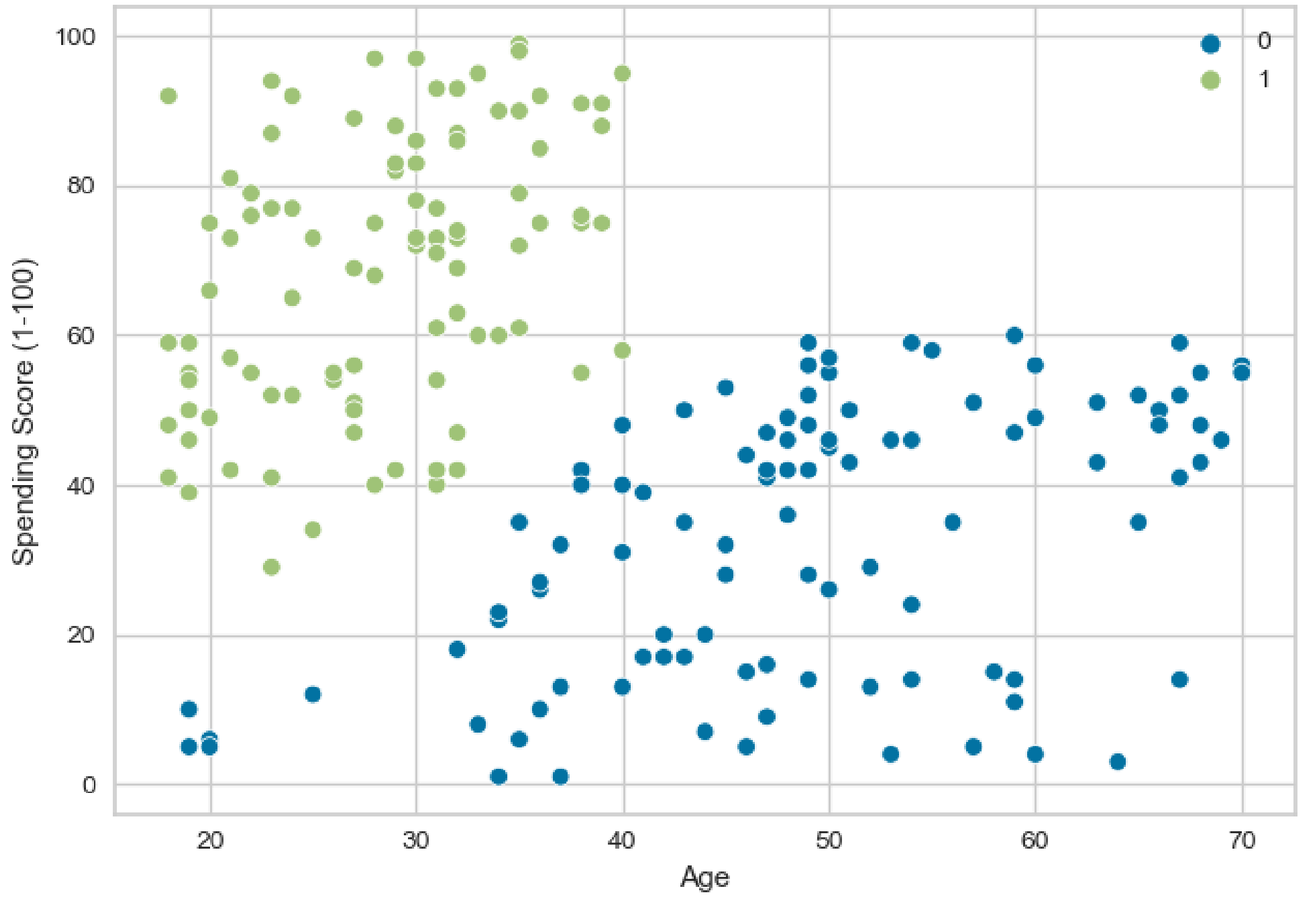
Conclusions:

- As people get older, their spending score decrease
- the correlation looks the same for males and females
- people over age around 40 have a spending score below 60.



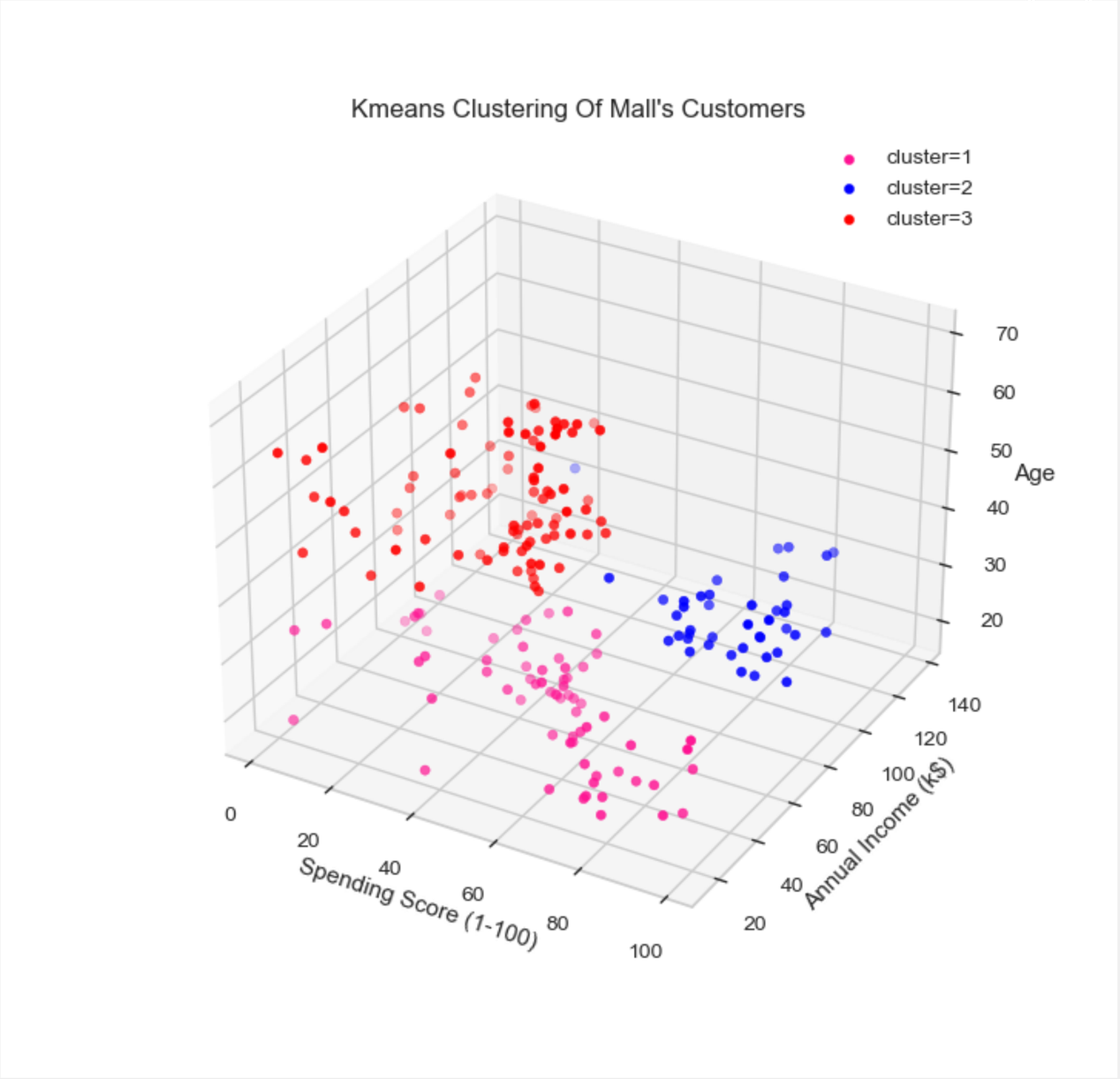
3.Data Analysis-Clustering:

Clusters of subgroups were made with **K-Means algorithm** that parted people into 2 groups **based on their age and spending score**.
Now it is easier to see a certain group of people and their age that has a **spending score higer** then the avarege.
People **over age 40 have a spending score less then 60**



3D Clustering:

Does the amount of the annual income influence spending score? Made clusters shows subgroups that customers can make.They are based on their **age**, **spending score and annual income.**
Taking into consideration those clusters, customers with age below 40 may have a high spending score despite a low annual income or a high annual income.



4.Conclusions:

Most of customers are around age 18-40. People over age 40 are much less numerous.

Most of customers have an avarege or low annual income but it doesn’t influence their spending score. People over age 40 do not have a spending score higher then 60

Based on customer's age, spending score and annual income it is possible to differ 3 groups that can help to better understand customer's needs

End

Thank you