Clients Clustering

Autor: Alexander Tamayo

Index

Introduction

Car Doctor is an auto repair shop located in Guadalajara, Jalisco that has several clients and they want to know what segments of clients they have.

Data Understanding

Car Doctor provided us with a csv file with the following useful data:

- idCliente
- Nombre Cliente
- Venta

Exploratory Data Analysis (EDA)

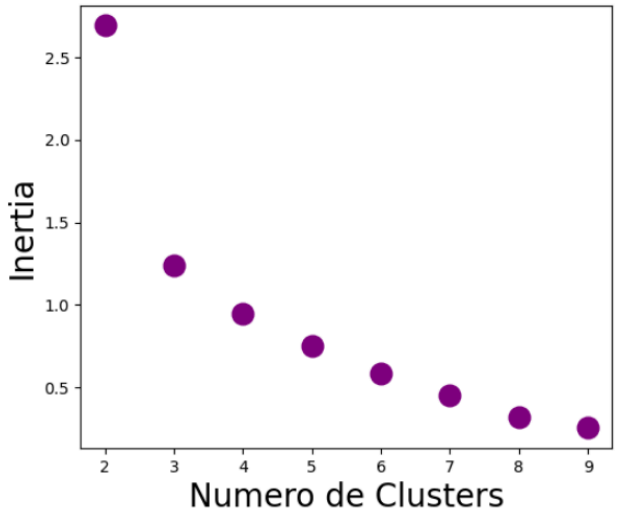
After some exploration, we were able to group the number of sells and total sales per client.

	idCliente	NumeroVentas	TotalVentas
0	7	1	550.0
1	24	8	20927.0
2	29	5	11900.0
3	30	3	15901.0
4	31	2	12360.0

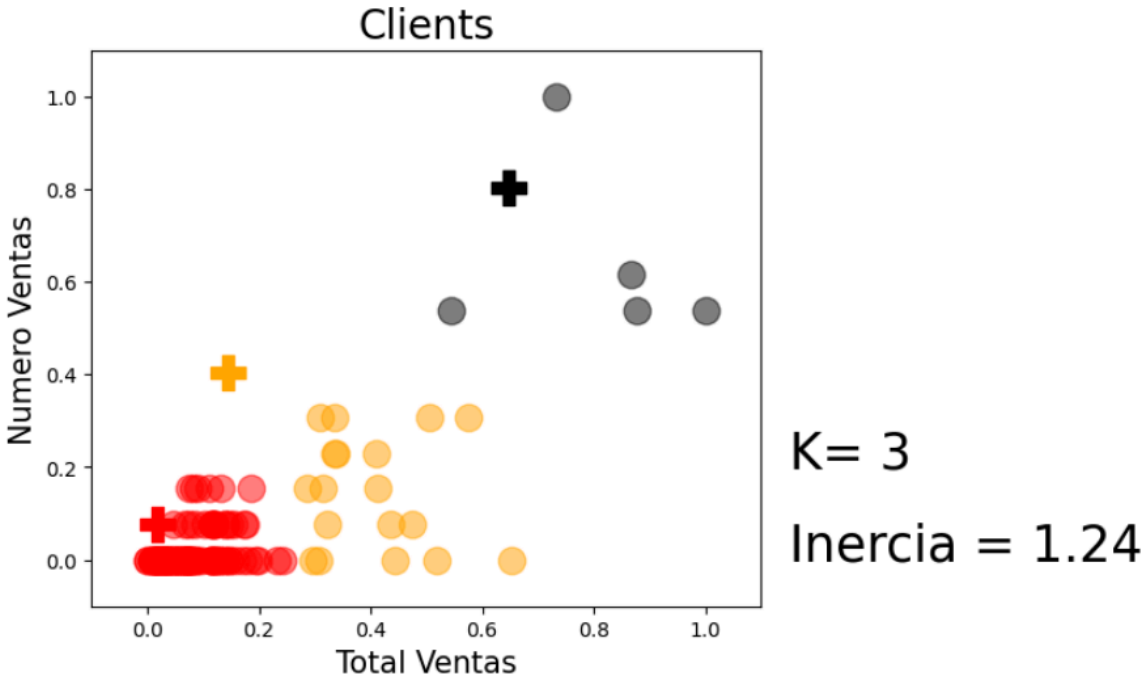
The model k means was used to cluster every client, based on 3 main clusters.

	NumeroVentas	TotalVentas	Cluster
0	0.000000	0.014285	0
1	0.538462	0.543516	2
2	0.307692	0.309067	1
3	0.153846	0.412981	1
4	0.076923	0.321014	1

Results and Interpretation



Using the elbow method we could be able to understand that using 3 clusters was the best approach for this model.



3 clusters were identified:

- Cluster 1: This cluster of clients doesn´t assist to the business often and doesn't spend so much money in it´s service, usually only visiting for one unique cheap service.
- Cluster 2: These are clients that have purchased more than one service and assist to the business regularly.
- Cluster 3: These clients assist to the business very often and have purchased a big amount of money, usually they are companies.

Conclusion

Car Doctor encounters three main clusters of clients, based on this info they could create strategies personally designed for each cluster.