




```
import pandas as pd




tabela = pd.read_csv("mall_customers.csv")
display(tabela)
```

	CustomerID	Gender	Age	Annual Income (k\$)	Spending Score (1-100)	
0	1	Male	19	15	39	
1	2	Male	21	15	81	
2	3	Female	20	16	6	
3	4	Female	23	16	77	
4	5	Female	31	17	40	
...	
195	196	Female	35	120	79	
196	197	Female	45	126	28	
197	198	Male	32	126	74	
198	199	Male	32	137	18	
199	200	Male	30	137	83	

200 rows × 5 columns

Próximas etapas: [Gerar código com tabela](#) [New interactive sheet](#)

```
colunas_cluster = ["Annual Income (k$)", "Spending Score (1-100)"]
tabela_cluster = tabela[colunas_cluster].copy()
display(tabela_cluster)
```

	Annual Income (k\$)	Spending Score (1-100)	
0	15	39	
1	15	81	
2	16	6	
3	16	77	
4	17	40	
...	
195	120	79	
196	126	28	
197	126	74	
198	137	18	
199	137	83	




200 rows × 2 columns

Próximas etapas: [Gerar código com tabela_cluster](#) [New interactive sheet](#)

```
from sklearn.preprocessing import StandardScaler
from sklearn.cluster import KMeans

normalizador = StandardScaler()

tabela_cluster[colunas_cluster] = normalizador.fit_transform(tabela_cluster[colunas_cluster])
display(tabela_cluster)
```




	Annual Income (k\$)	Spending Score (1-100)	
0	-1.738999	-0.434801	
1	-1.738999	1.195704	
2	-1.700830	-1.715913	
3	-1.700830	1.040418	
4	-1.662660	-0.395980	
...	
195	2.268791	1.118061	
196	2.497807	-0.861839	
197	2.497807	0.923953	
198	2.917671	-1.250054	
199	2.917671	1.273347	

200 rows × 2 columns

Próximas etapas:

[Gerar código com tabela_cluster](#)[New interactive sheet](#)

```
modelo_kmeans = KMeans(n_clusters=5)
modelo_kmeans.fit(tabela_cluster)
tabela_cluster["Cluster"] = modelo_kmeans.labels_
display(tabela_cluster)
```

	Annual Income (k\$)	Spending Score (1-100)	Cluster	
0	-1.738999	-0.434801	3	
1	-1.738999	1.195704	4	
2	-1.700830	-1.715913	3	
3	-1.700830	1.040418	4	
4	-1.662660	-0.395980	3	
...	
195	2.268791	1.118061	2	
196	2.497807	-0.861839	0	
197	2.497807	0.923953	2	
198	2.917671	-1.250054	0	
199	2.917671	1.273347	2	

200 rows × 3 columns

Próximas etapas:

[Gerar código com tabela_cluster](#)[New interactive sheet](#)

```
import plotly.express as px

grafico = px.scatter(tabela_cluster, x="Annual Income (k$)", y="Spending Score (1-100)", color="Cluster")
grafico.show()
```

