



Trabalho 11

Nome da Atividade: K-means

Nome e Matrícula: Lucas Lima do Nascimento - 12111ECP024

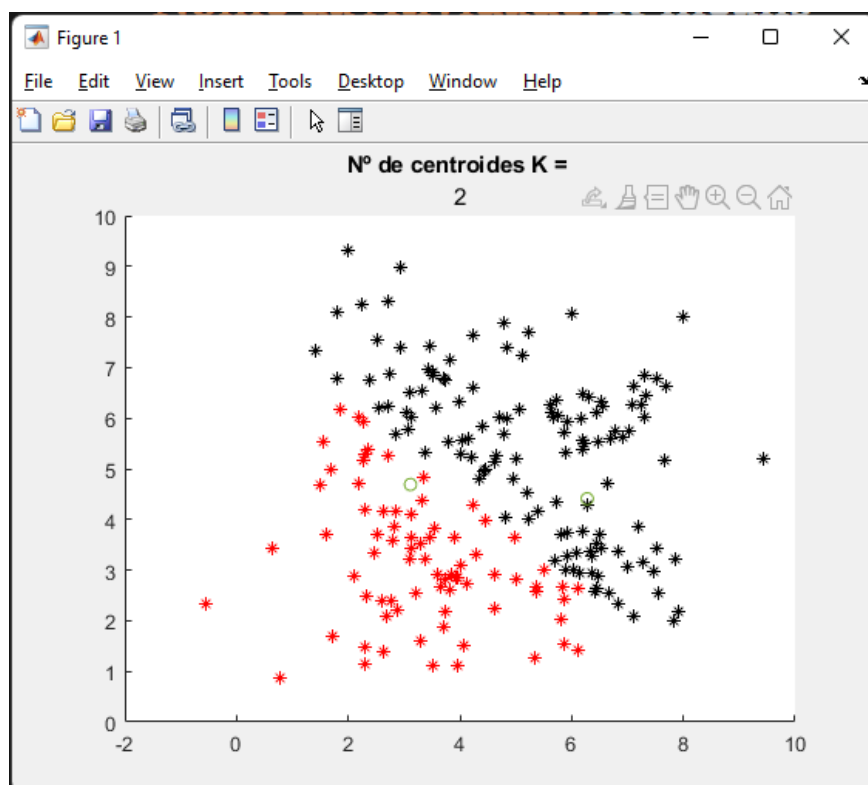
Introdução:

Nesse trabalho, fiz a implementação do algoritmo K-means para clusterização de dados.

200 dados foram fornecidos, tanto para X quanto para Y.

Resultados:

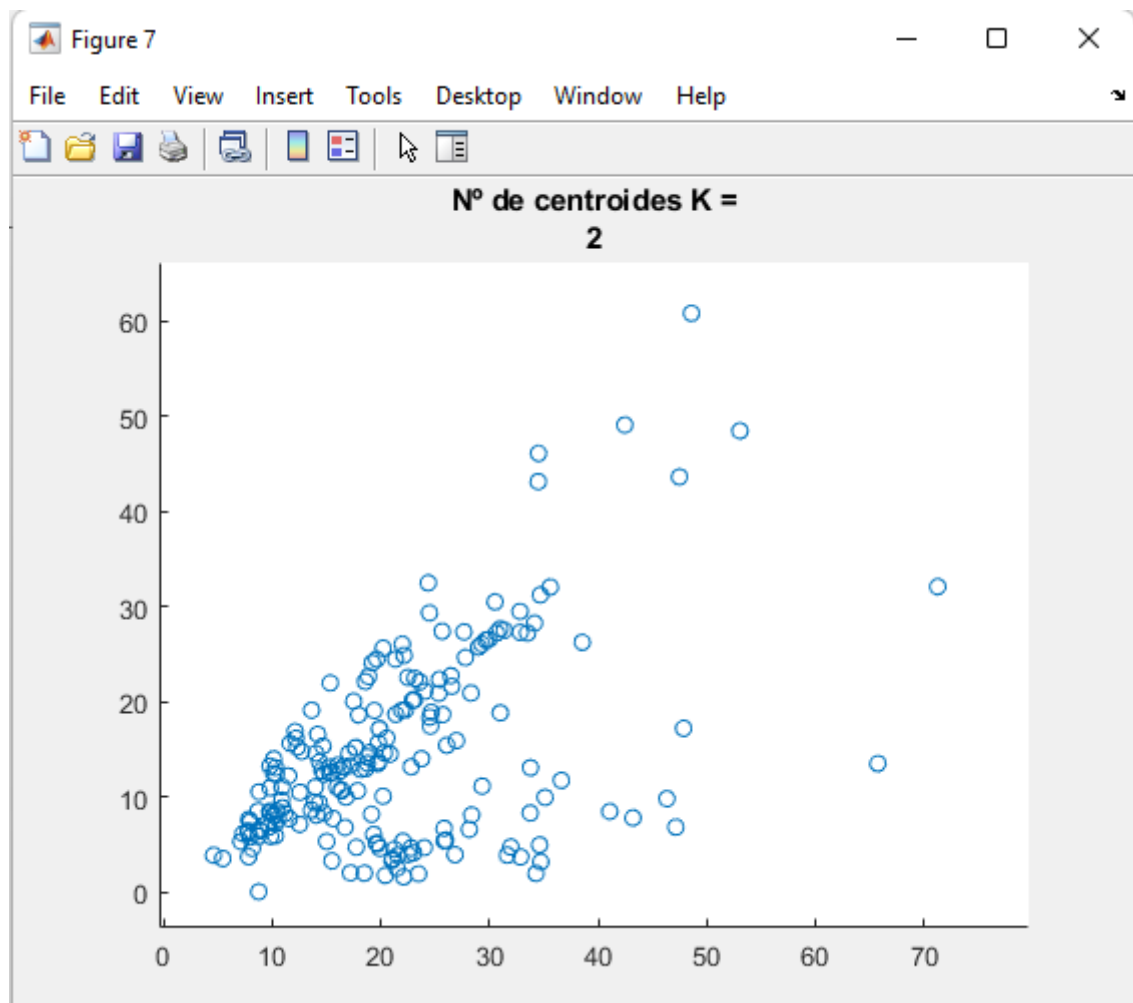
Para $K = 2$ centróides, obtive esse gráfico:



```
ans =  
  
2.3712
```

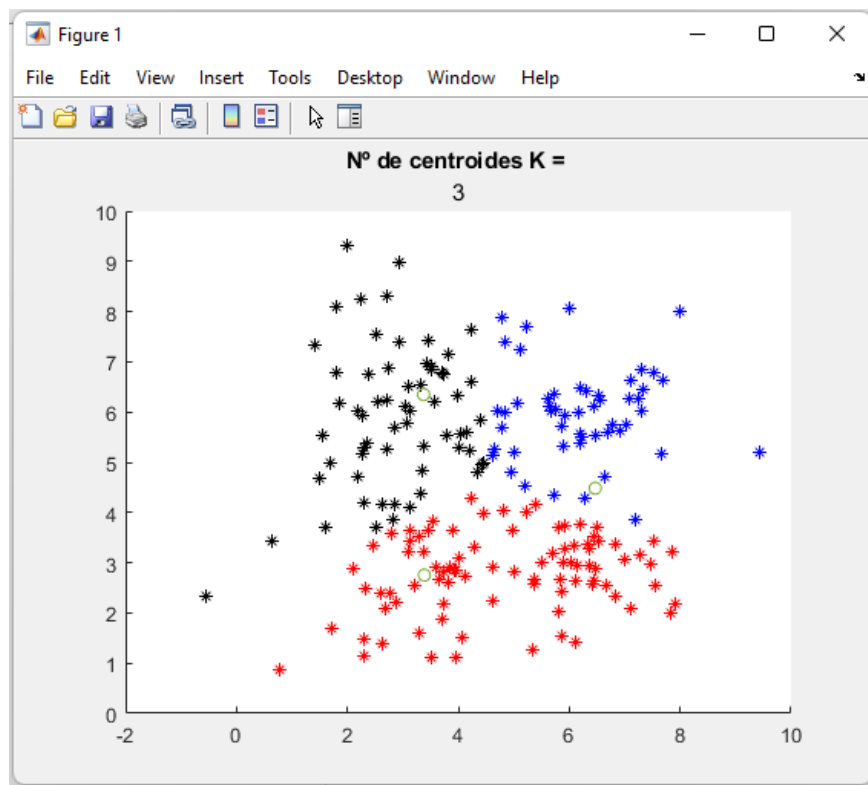
Obtive como média de distância entre os centróides e os pontos 2.3712.

Erro quadrático:



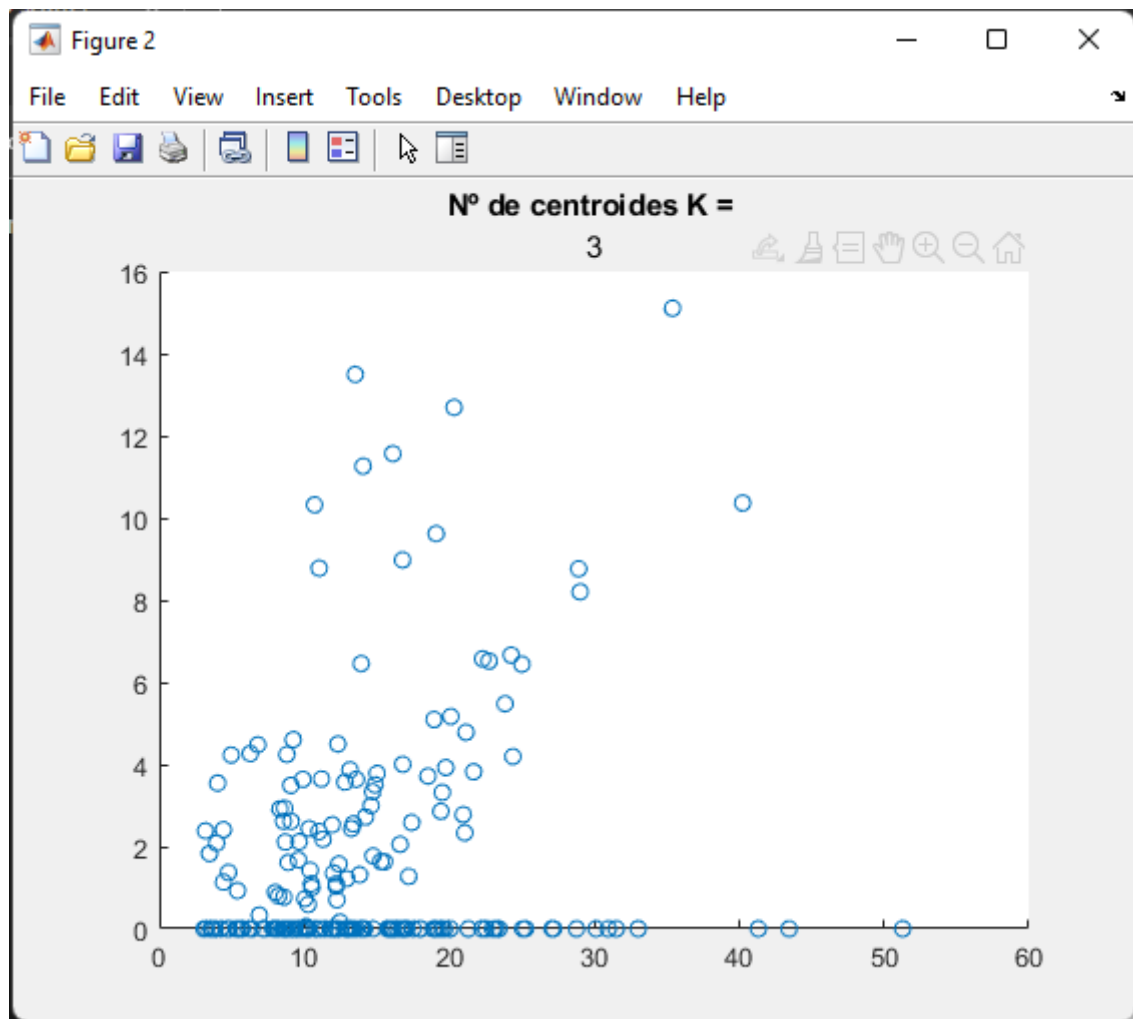
Para K = 3 centróides, obtive esse gráfico:

```
ans =  
  
2.8955
```



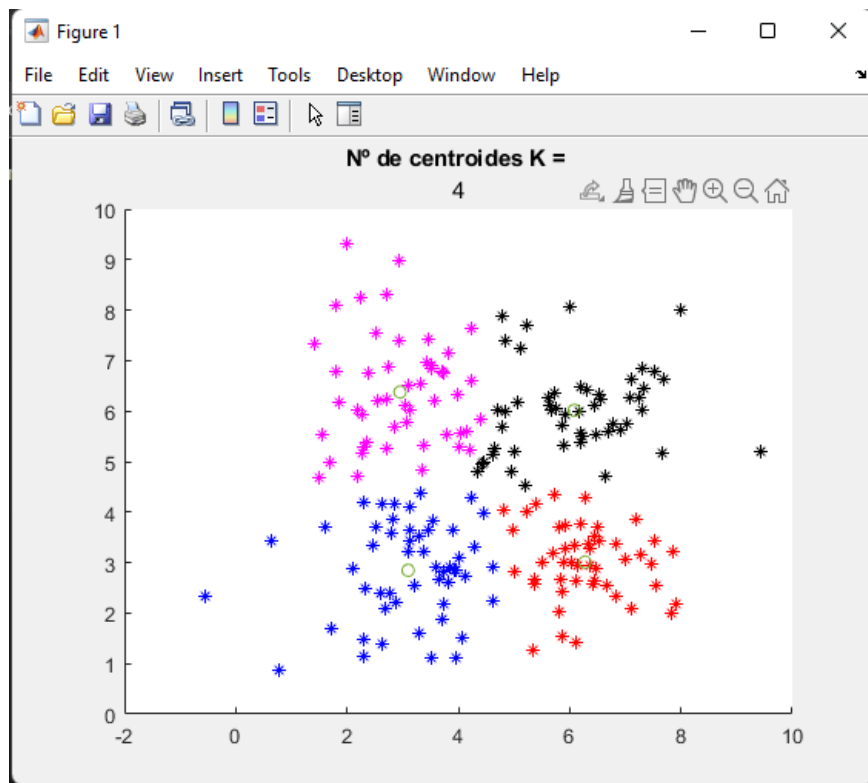
Obtive como média de distância entre os centróides e os pontos 2.8955.

Erro quadrático:



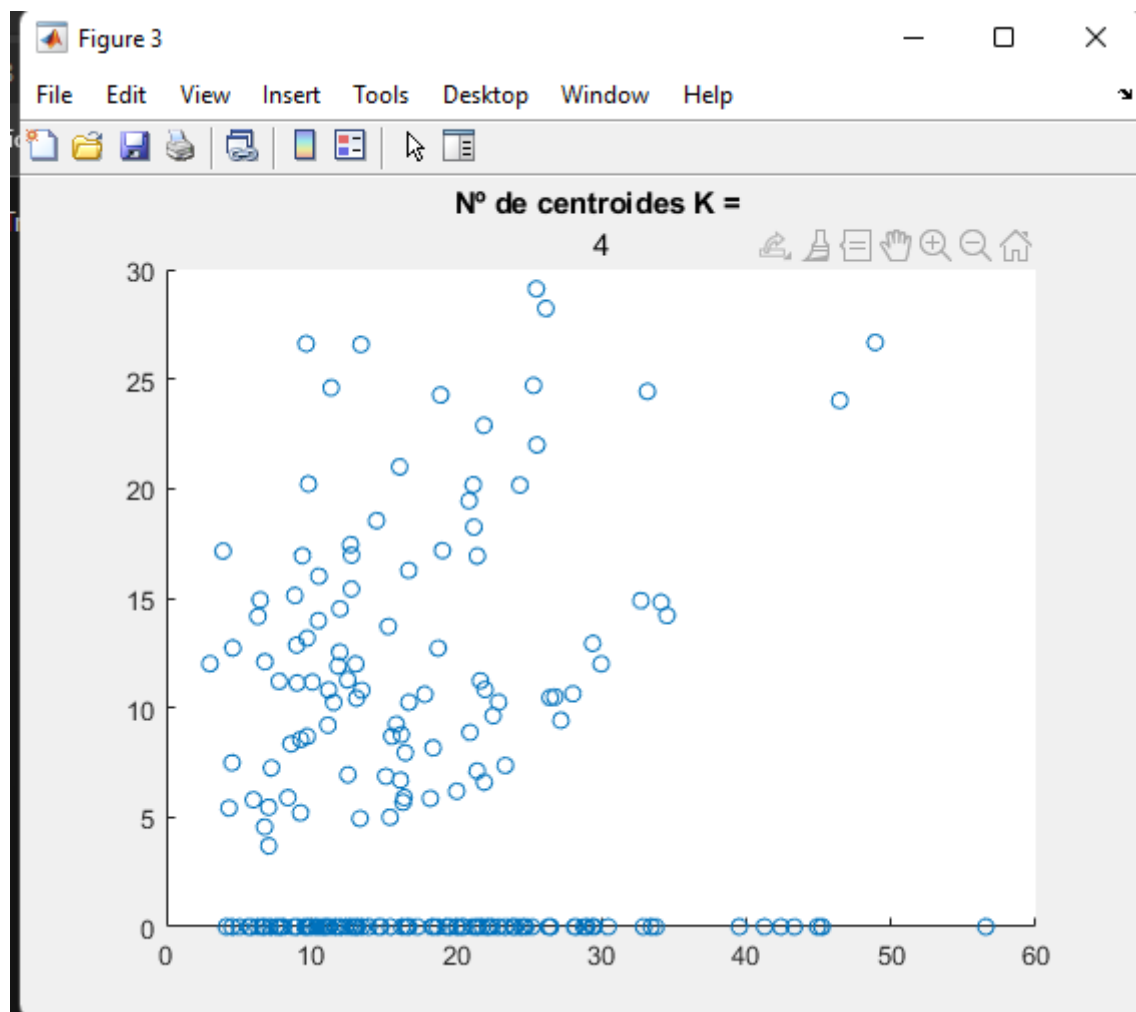
Para $K = 4$ centróides, obtive esse gráfico:

```
ans =  
2.8959
```



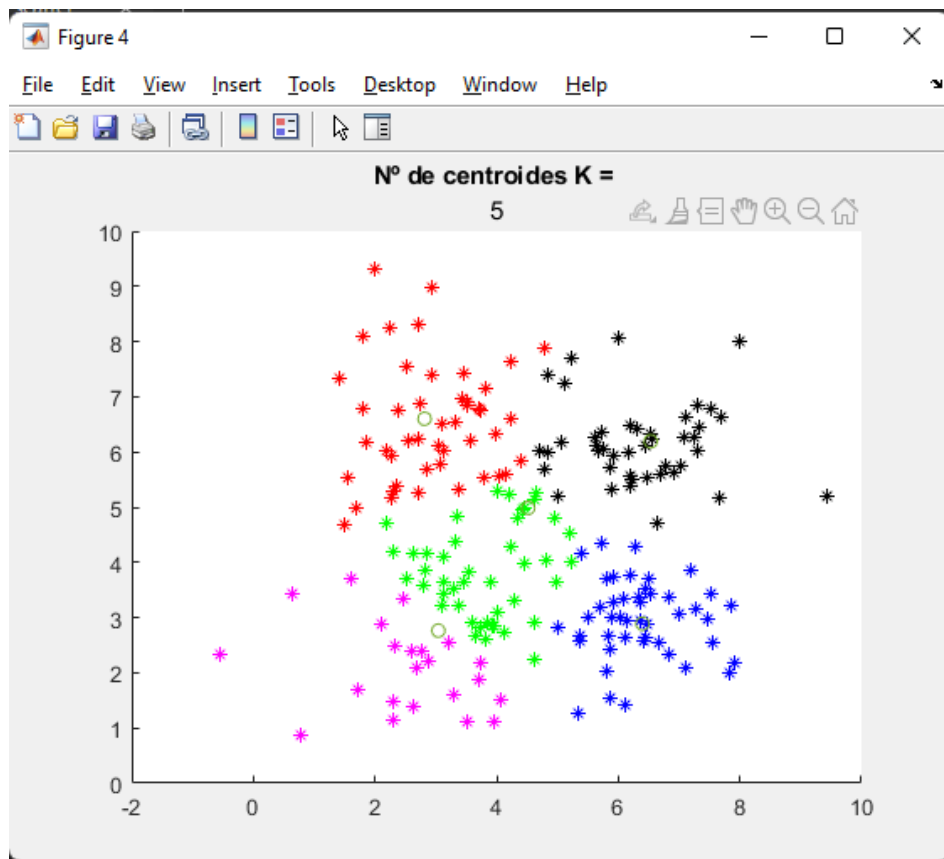
Obtive como média de distância entre os centróides e os pontos 2.8959.

Erro quadrático:



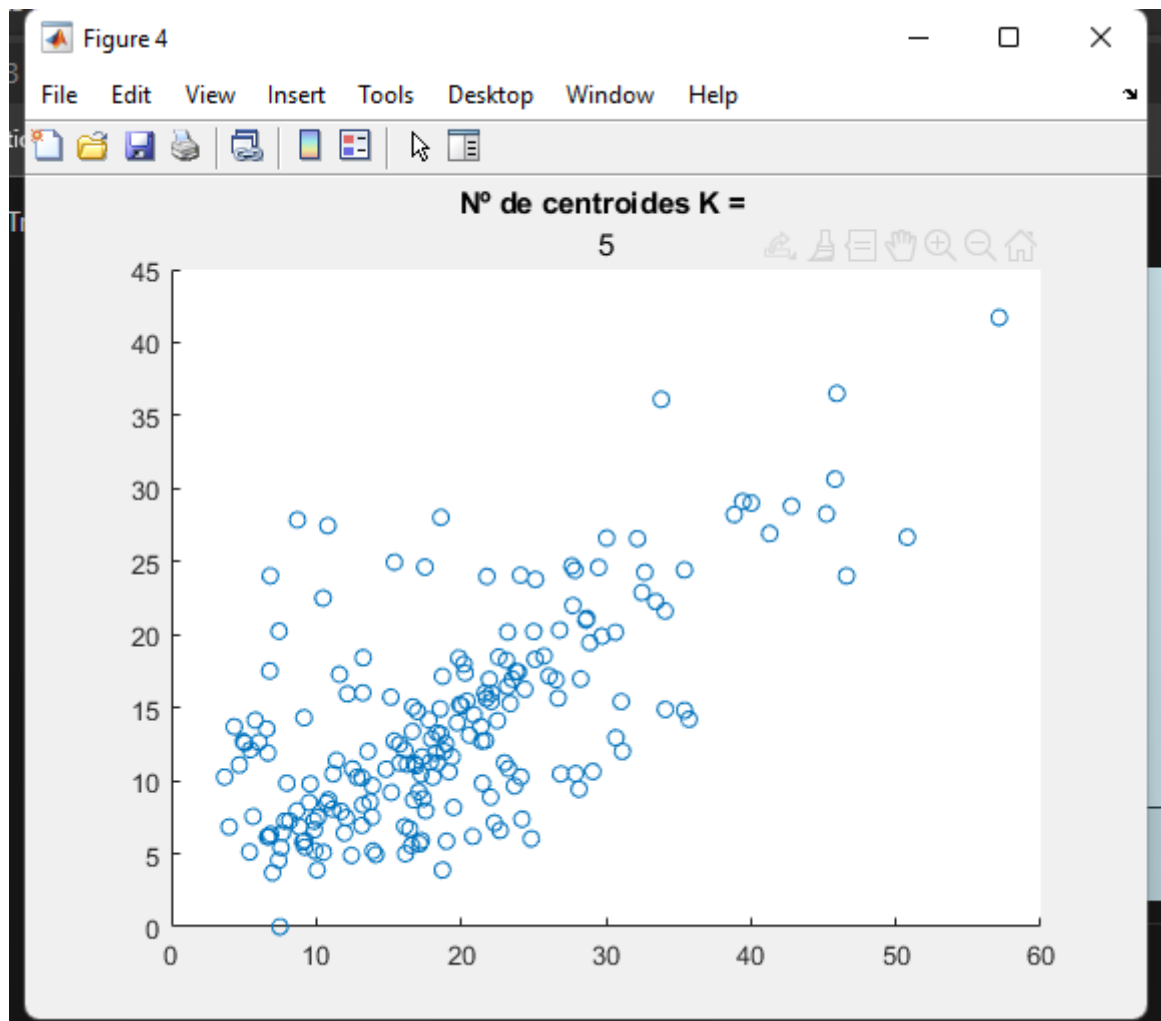
Para $K = 5$ centróides, obtive esse gráfico:

```
ans =  
3.1518
```



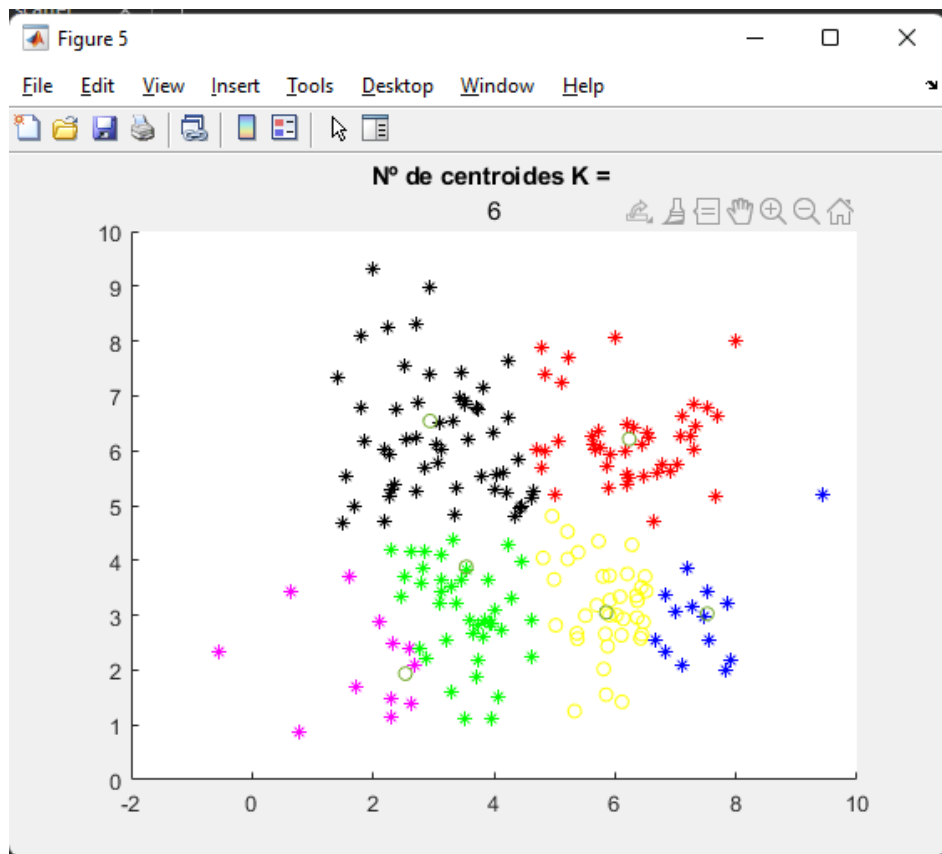
Obtive como média de distância entre os centróides e os pontos 3.1518.

Erro quadrático:



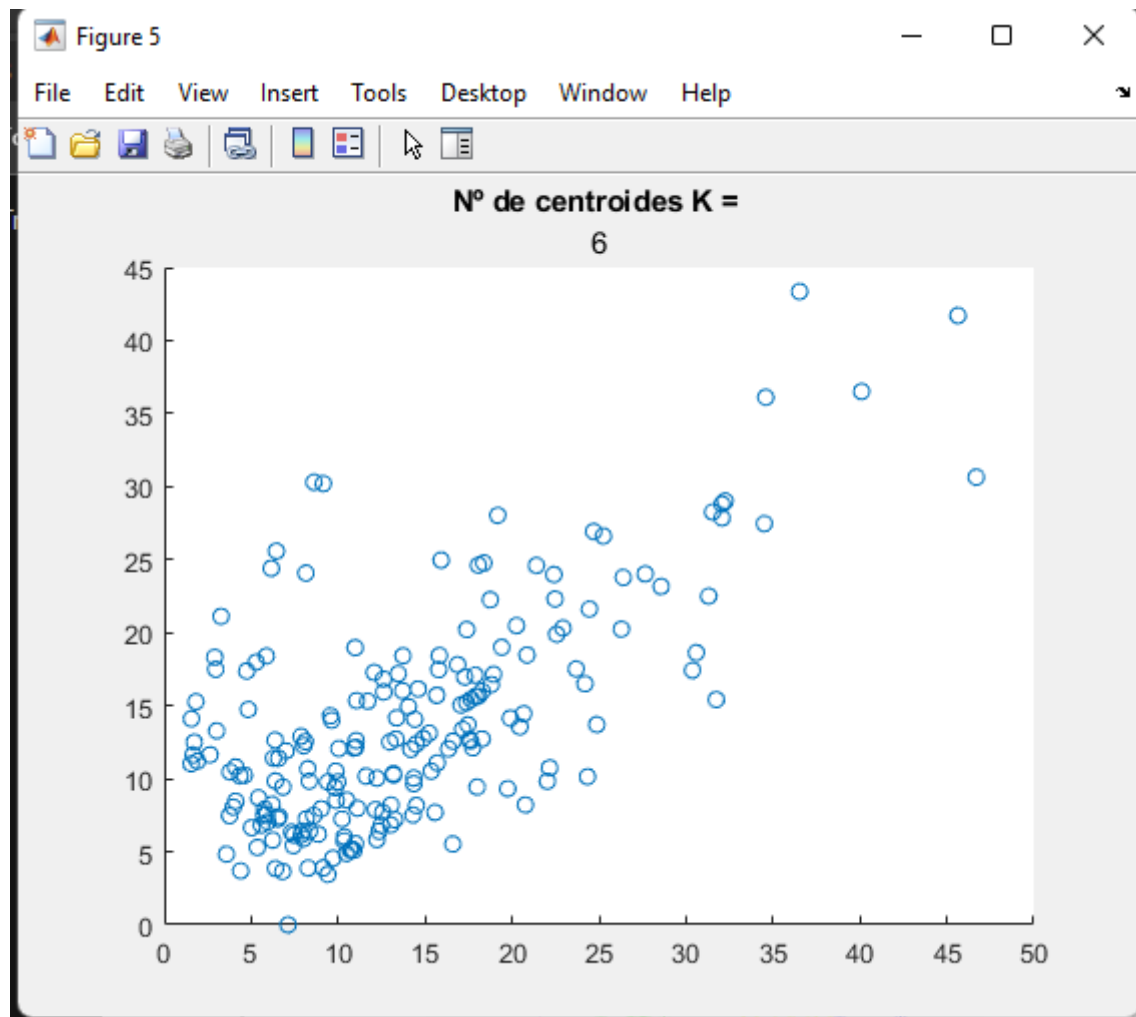
Para $K = 6$ centróides, obtive esse gráfico:

```
ans =  
3.6362
```



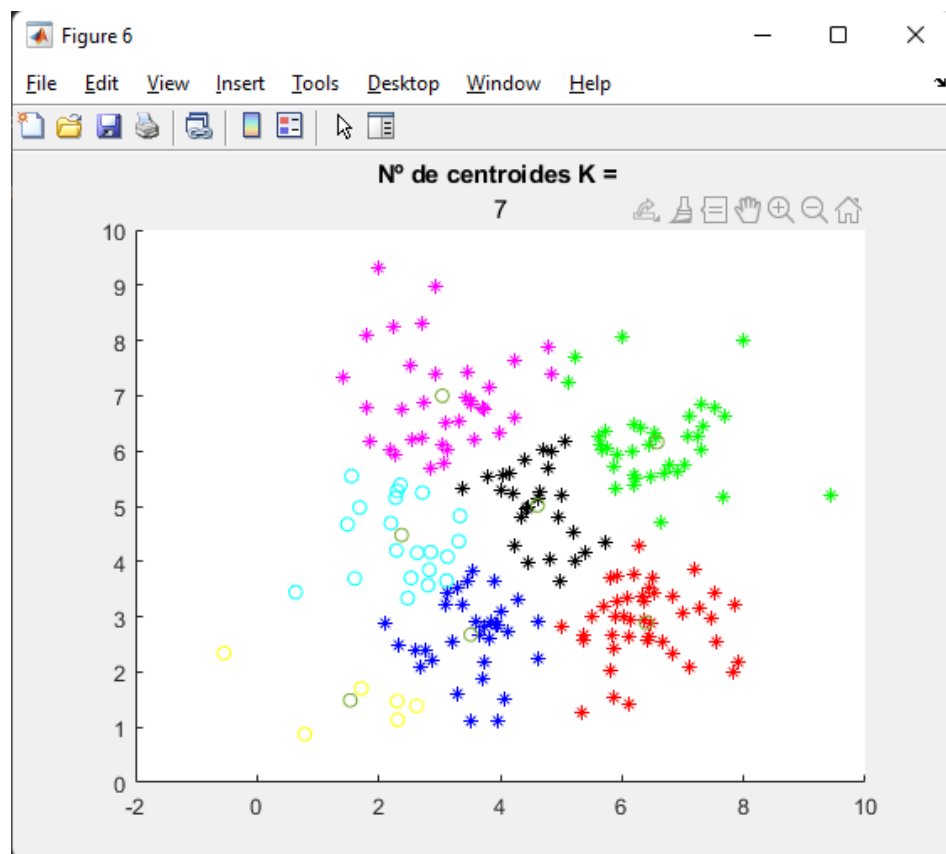
Obtive como média de distância entre os centróides e os pontos 3.6362.

Erro quadrático:



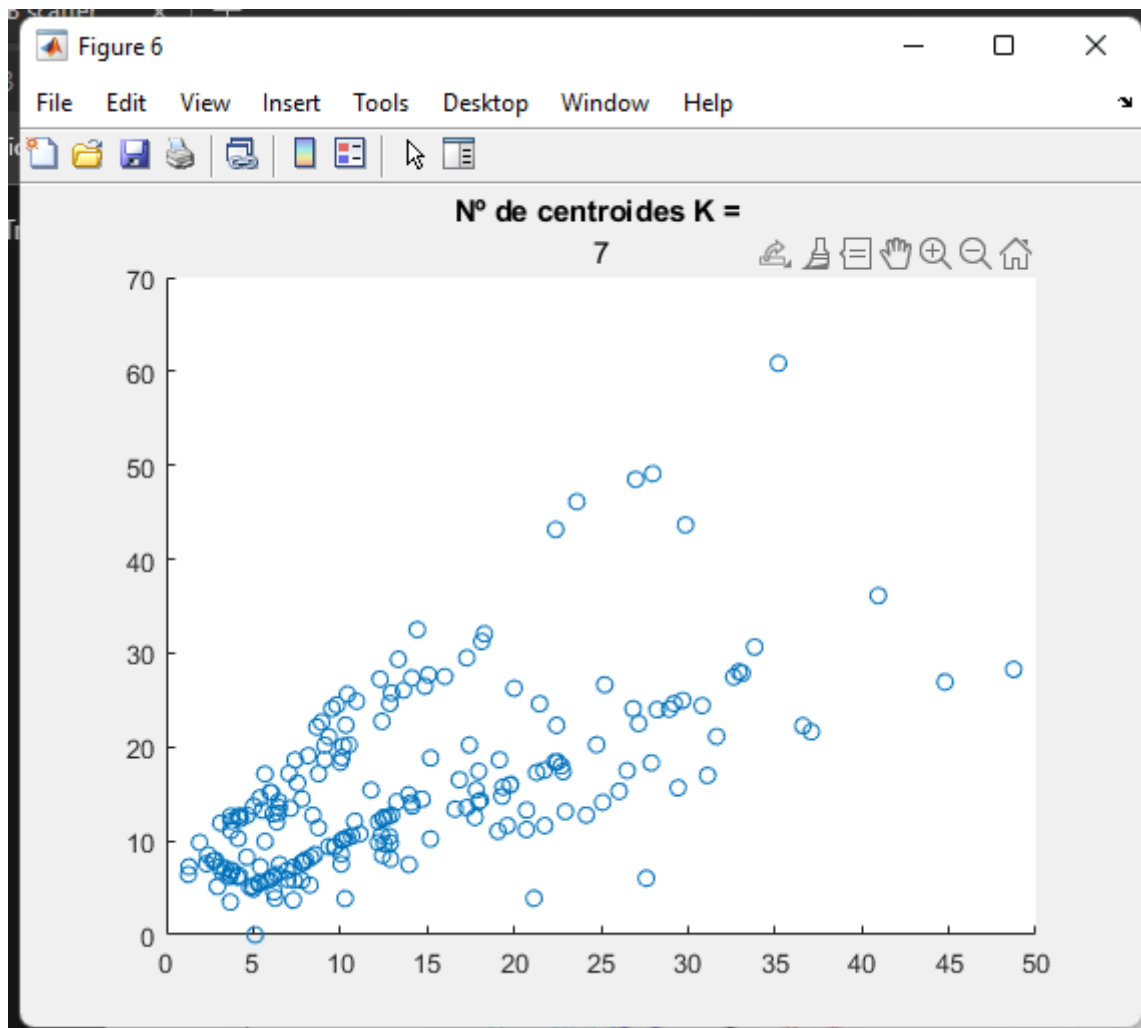
Para $K = 7$ centróides, obtive esse gráfico:

```
ans =  
3.9275
```



Obtive como média de distância entre os centróides e os pontos 3.9275.

Erro quadrático:



Para o conjunto de K testado (de 2 a 7), a melhor clusterização encontrada (com menor média de distância entre o centróide e os pontos) foi a de K = 2.

Código:

Em anexo ao envio.