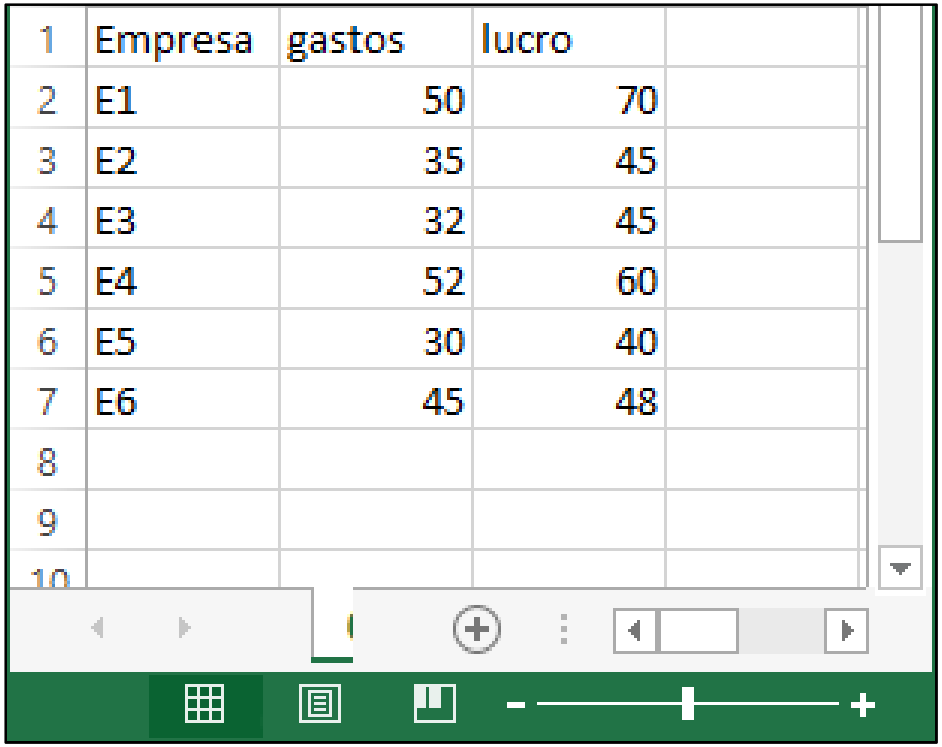
Estatística Aplicada à Administração II

Tópico: Análise de Clusters no R

Base de dados

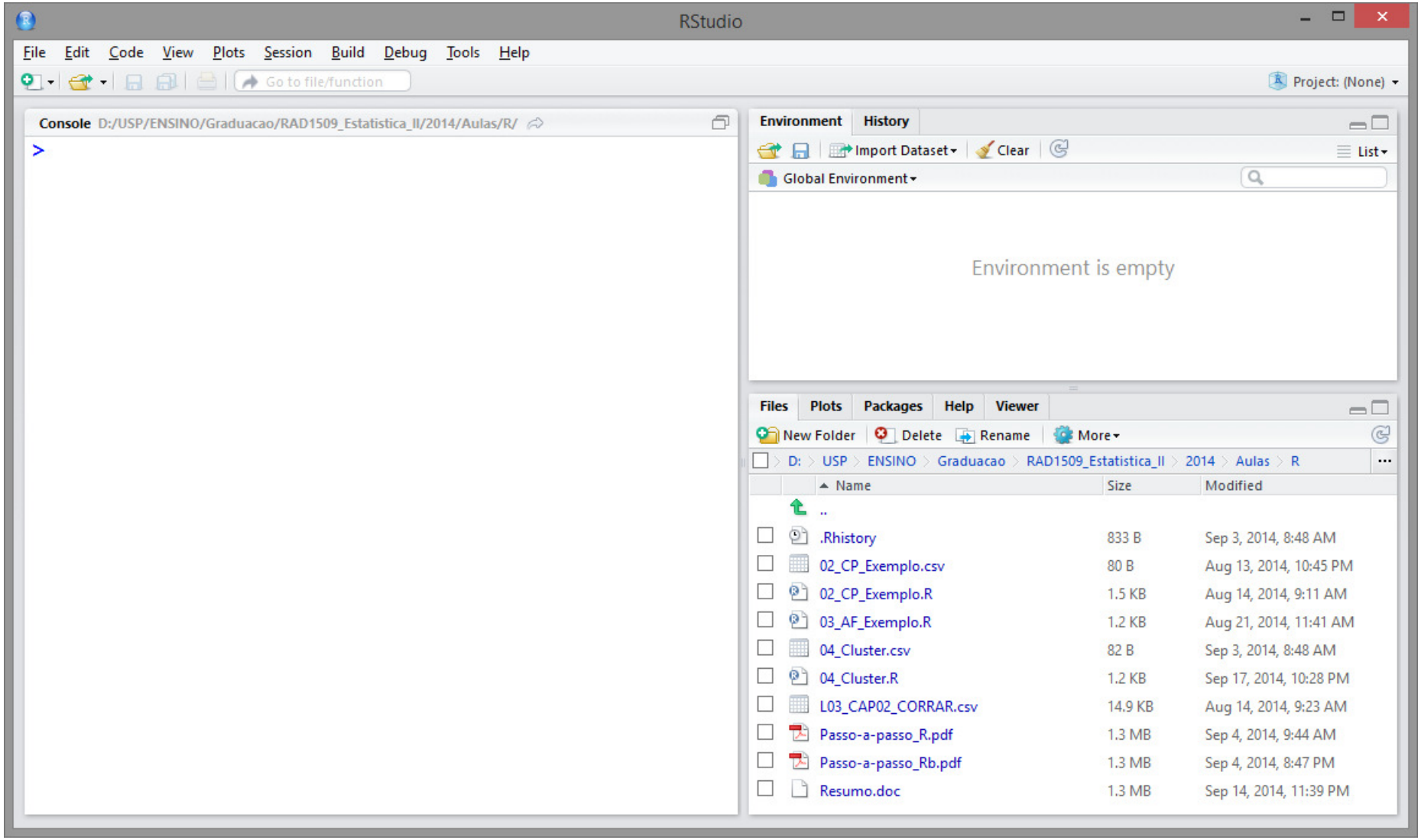
04_Cluster.csv



	Empresa	gastos	lucro
2	E1	50	70
3	E2	35	45
4	E3	32	45
5	E4	52	60
6	E5	30	40
7	E6	45	48
8			
9			
10			

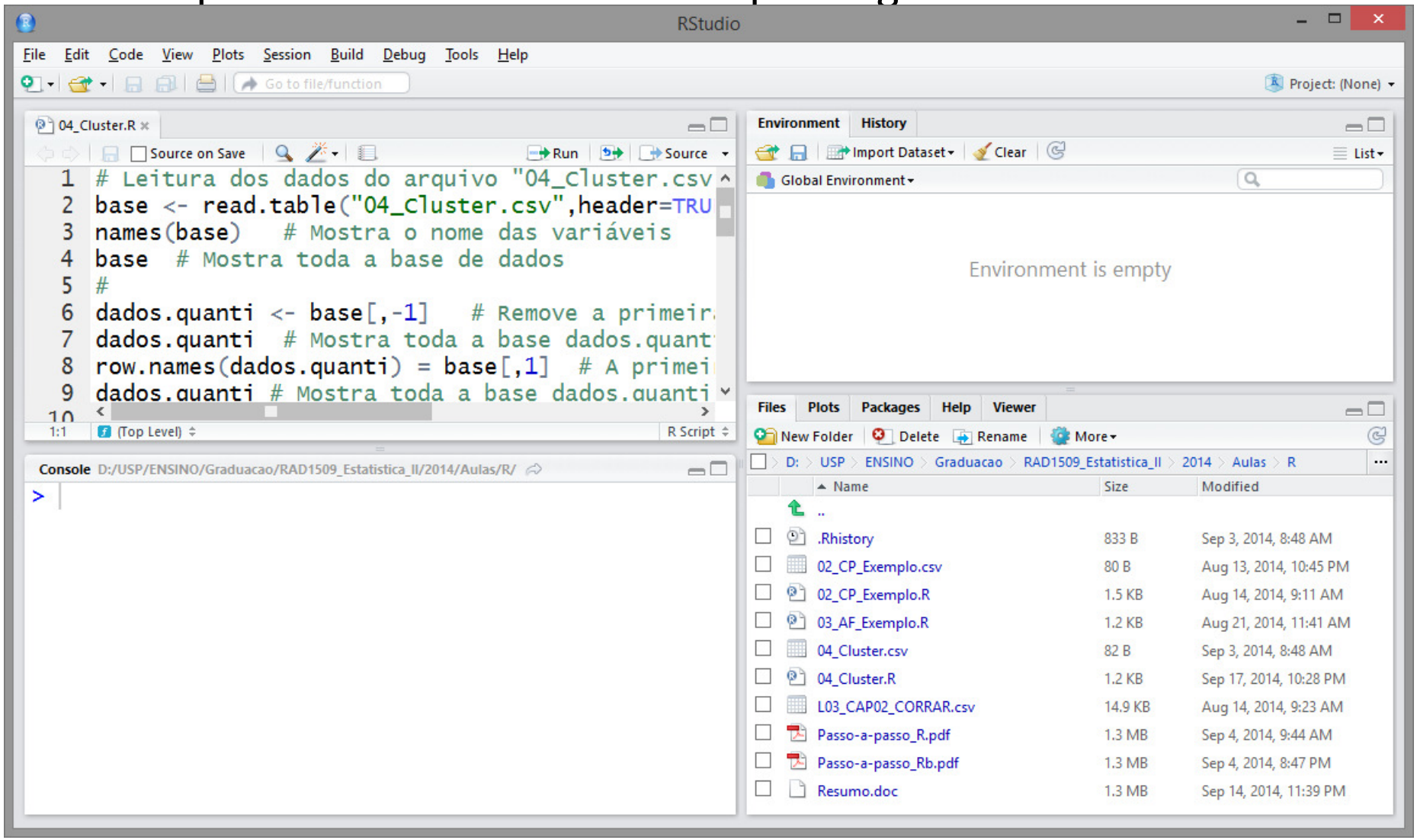
Análise de Agrupamentos no R

Considere o diretório de trabalho no qual esteja o arquivo **.csv**



Análise de Agrupamentos no R

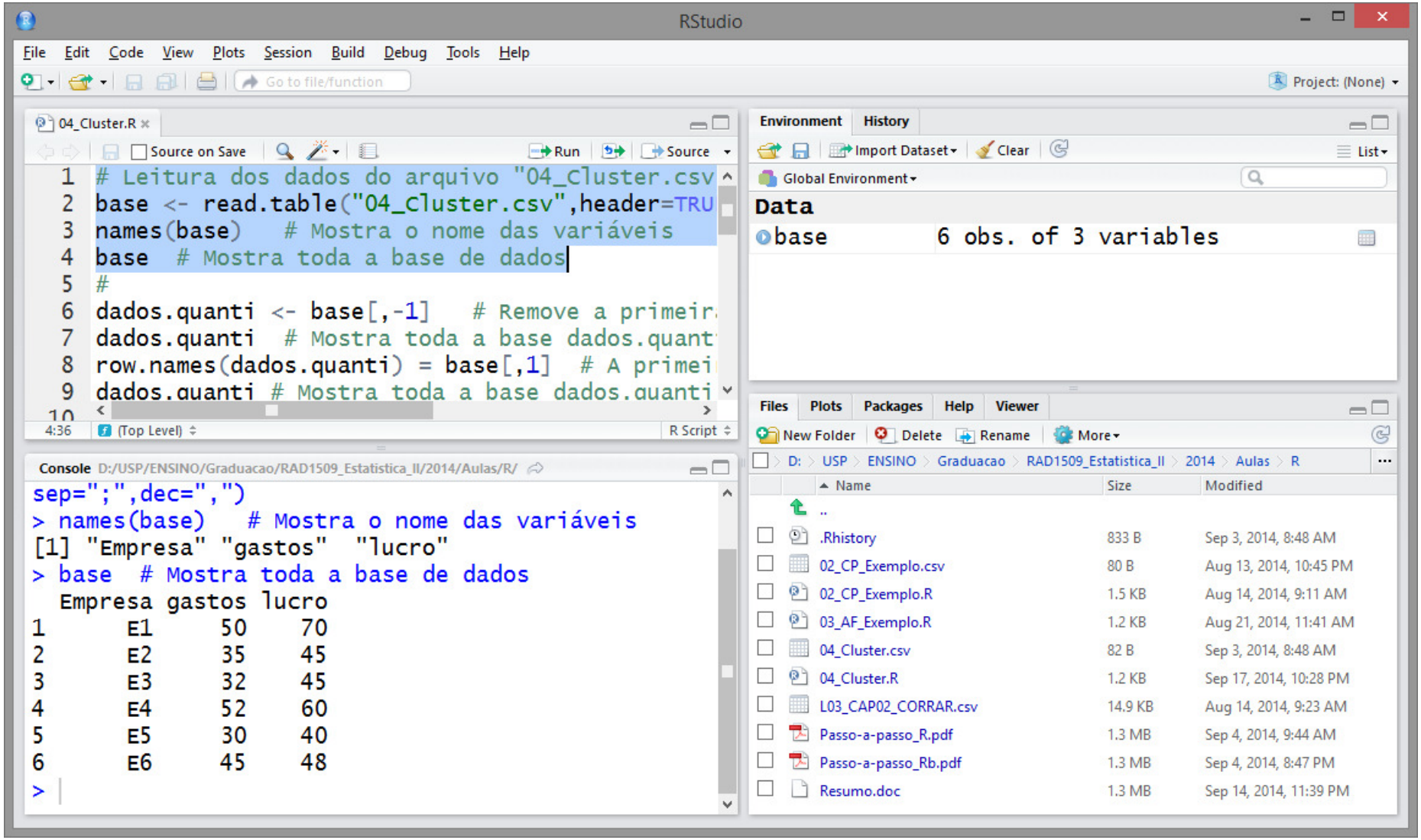
Clique no arquivo ".R" para abrir o script, ou então comece escrever um novo script: File → New File → R Script. Digite os comandos



Análise de Agrupamentos no R

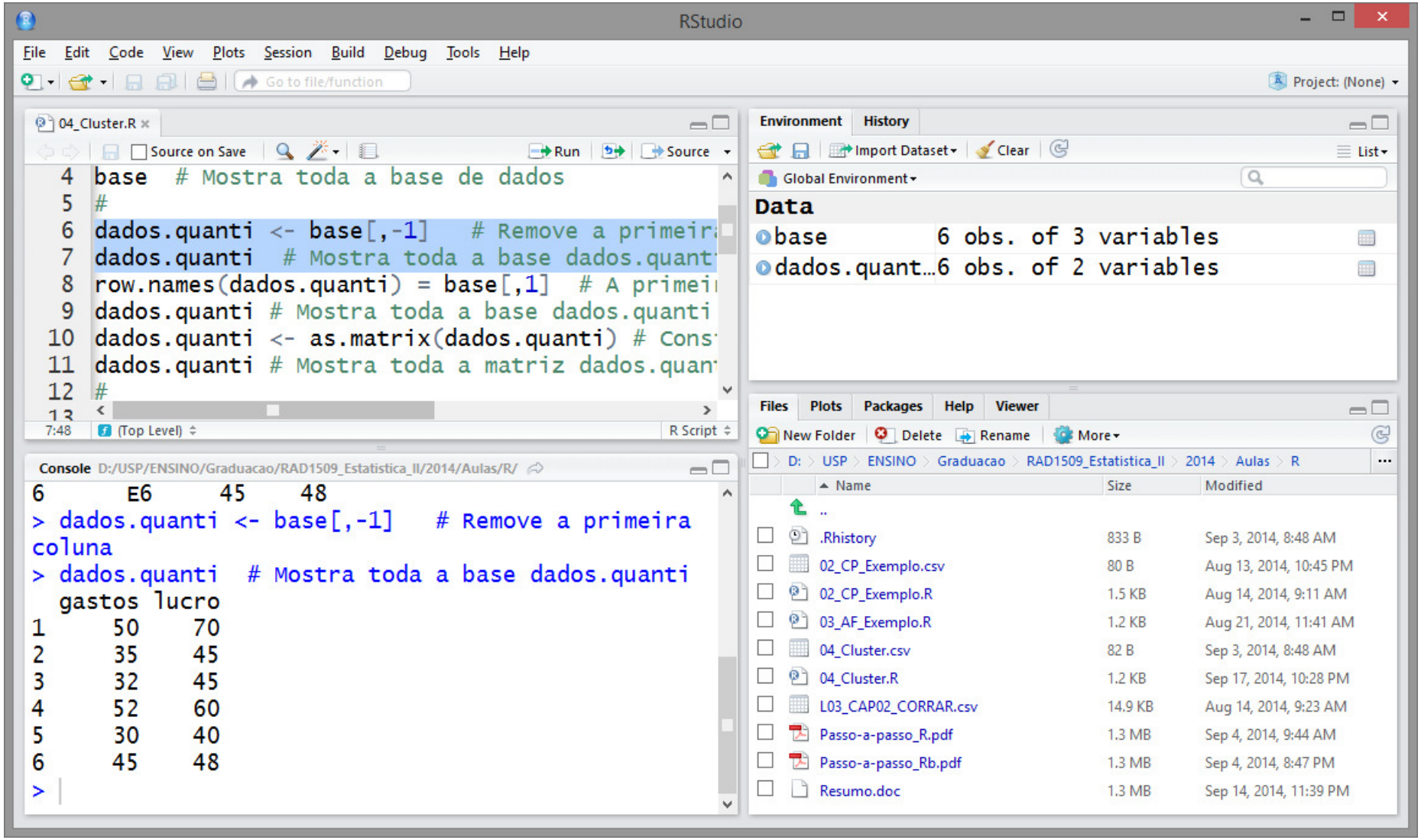
Faça a leitura dos dados do arquivo .csv para o data.frame **base**.

Oberve o conteúdo do data.frame base.



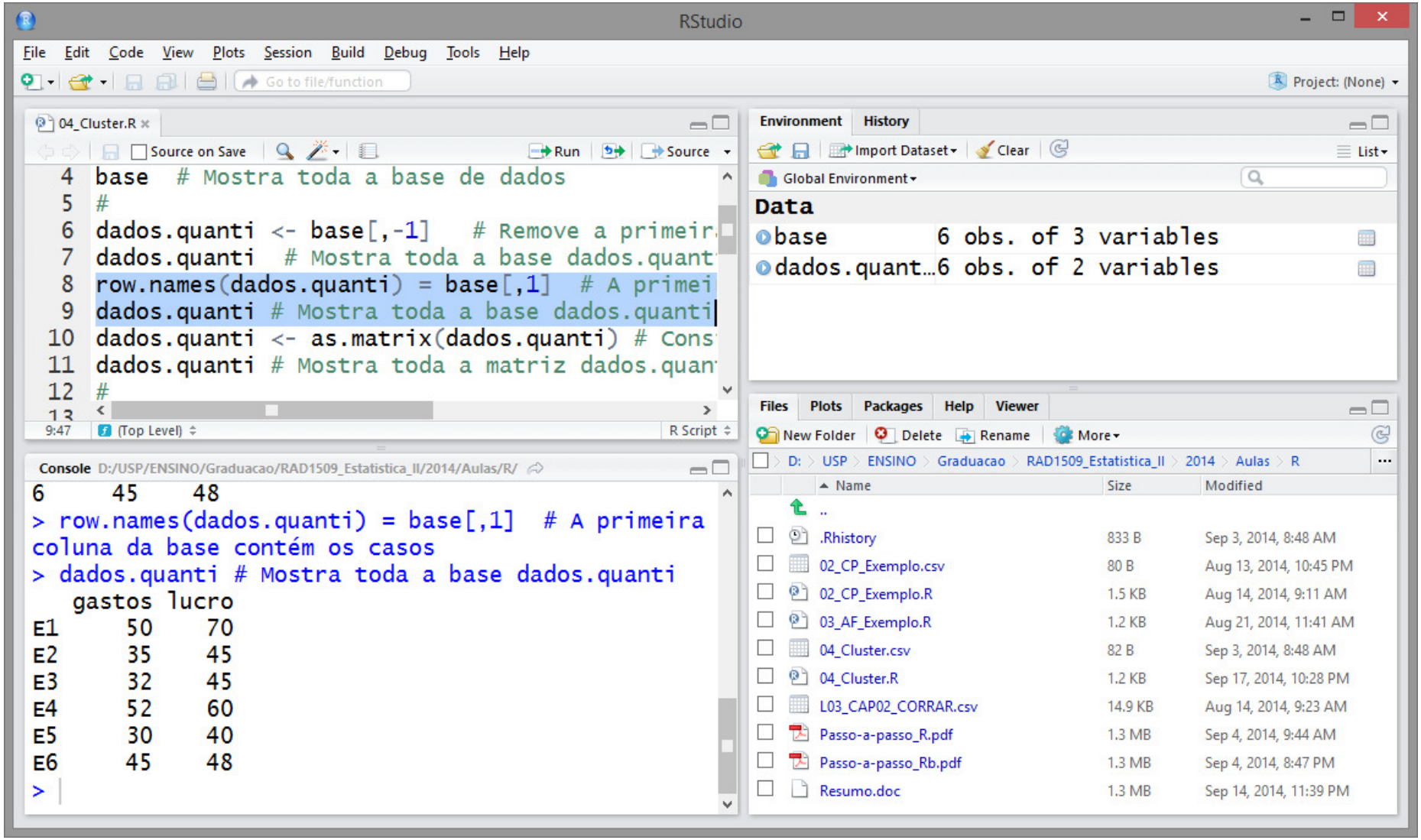
Análise de Agrupamentos no R

Remova a primeira coluna para ficar com dados quantitativos



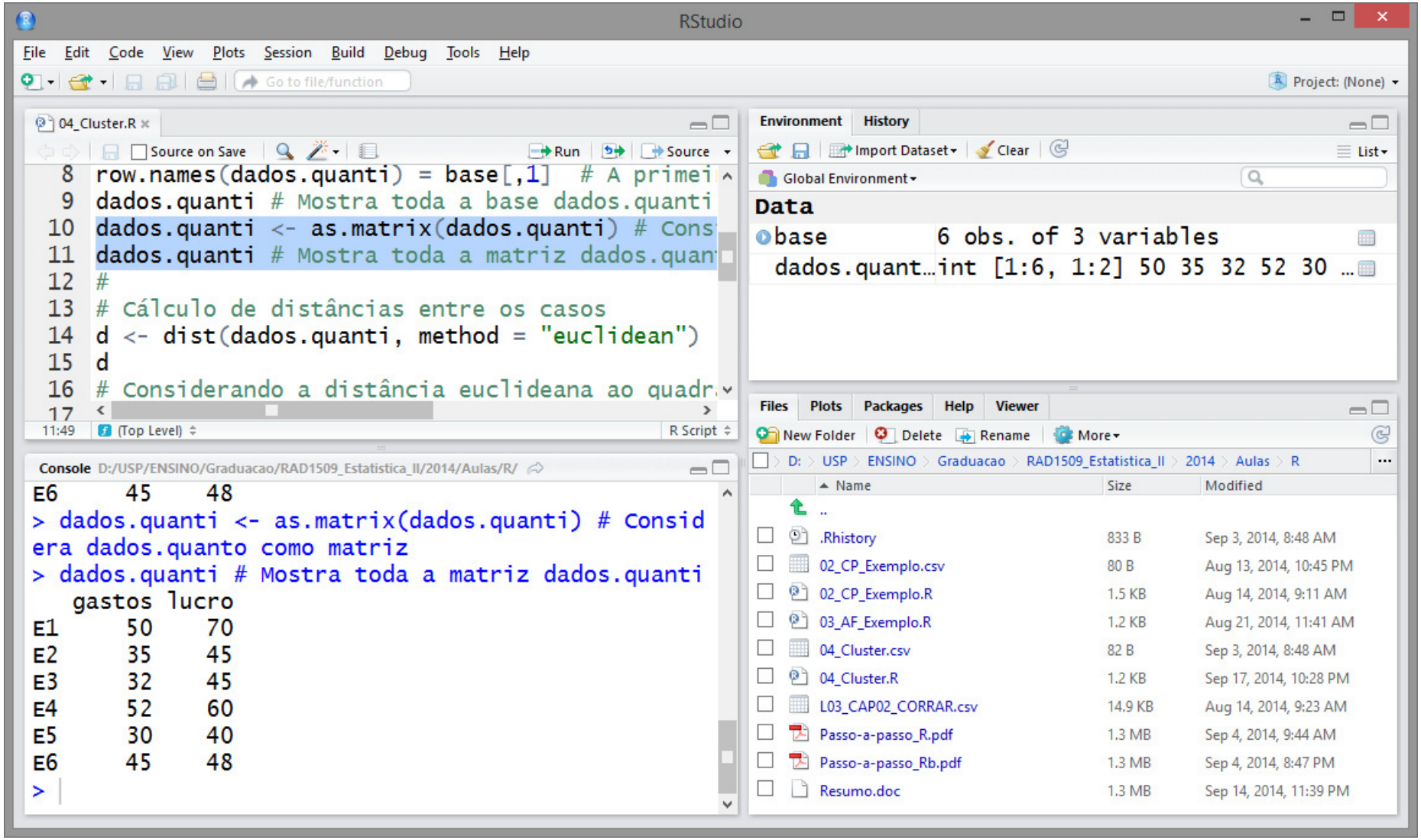
Análise de Agrupamentos no R

Coloque nomes nas linhas de acordo com os casos (1ª coluna da base)



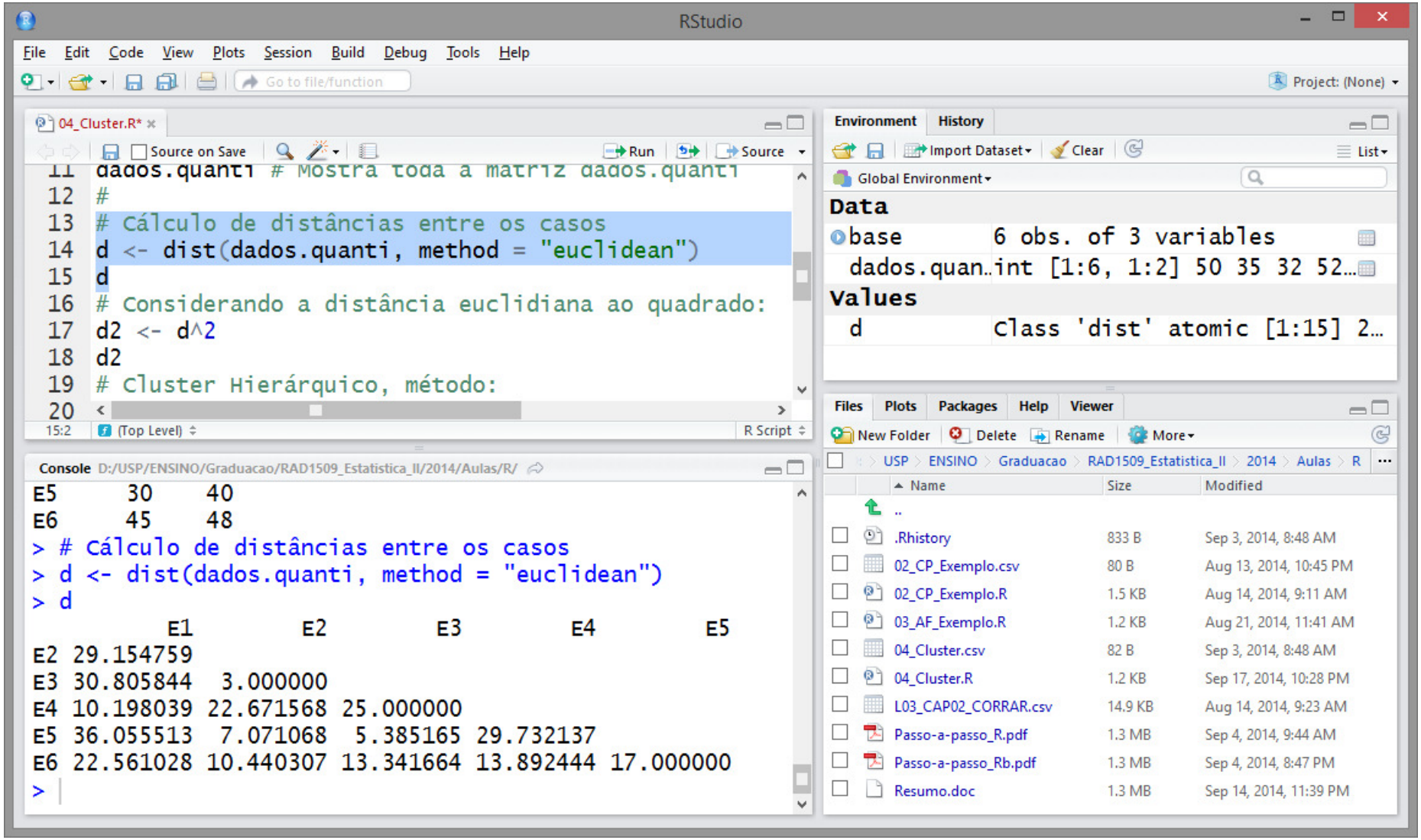
Análise de Agrupamentos no R

Mude a base dados.quanti para uma forma de matriz para poder obter a distância entre os casos.



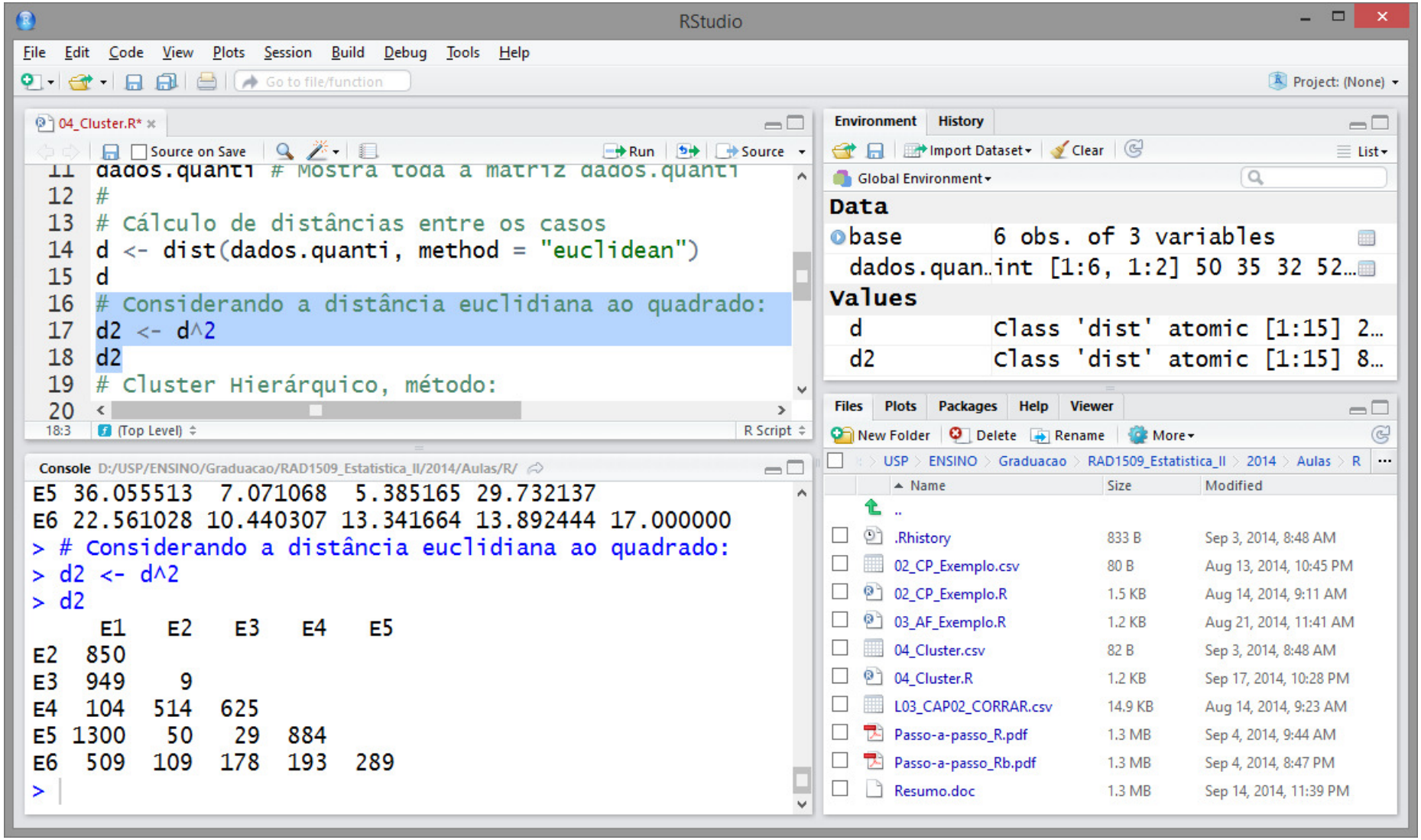
Análise de Agrupamentos no R

Calcule a matriz de distâncias entre os casos E1, E2, ...



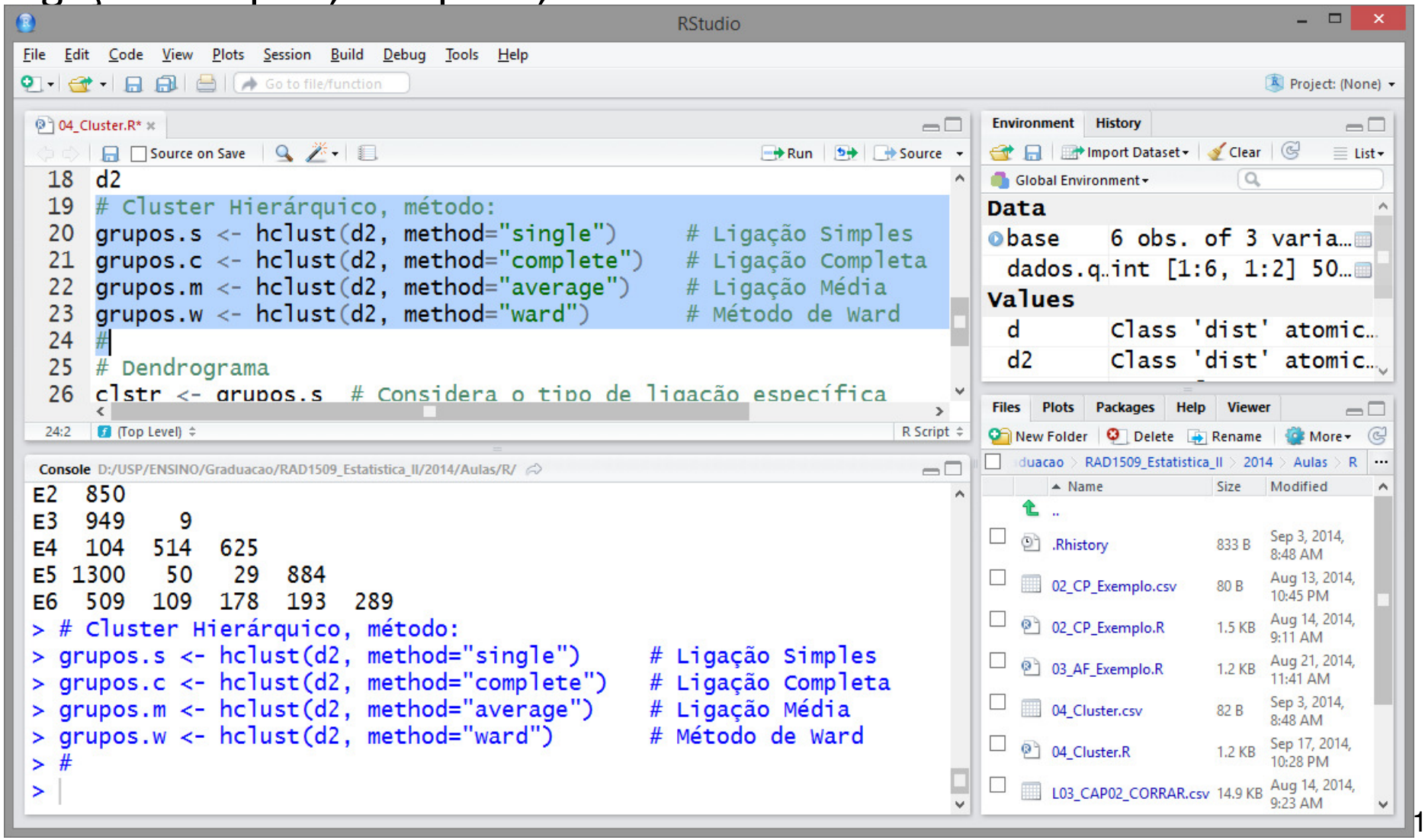
Análise de Agrupamentos no R

Calcule a distância euclidiana ao quadrado.



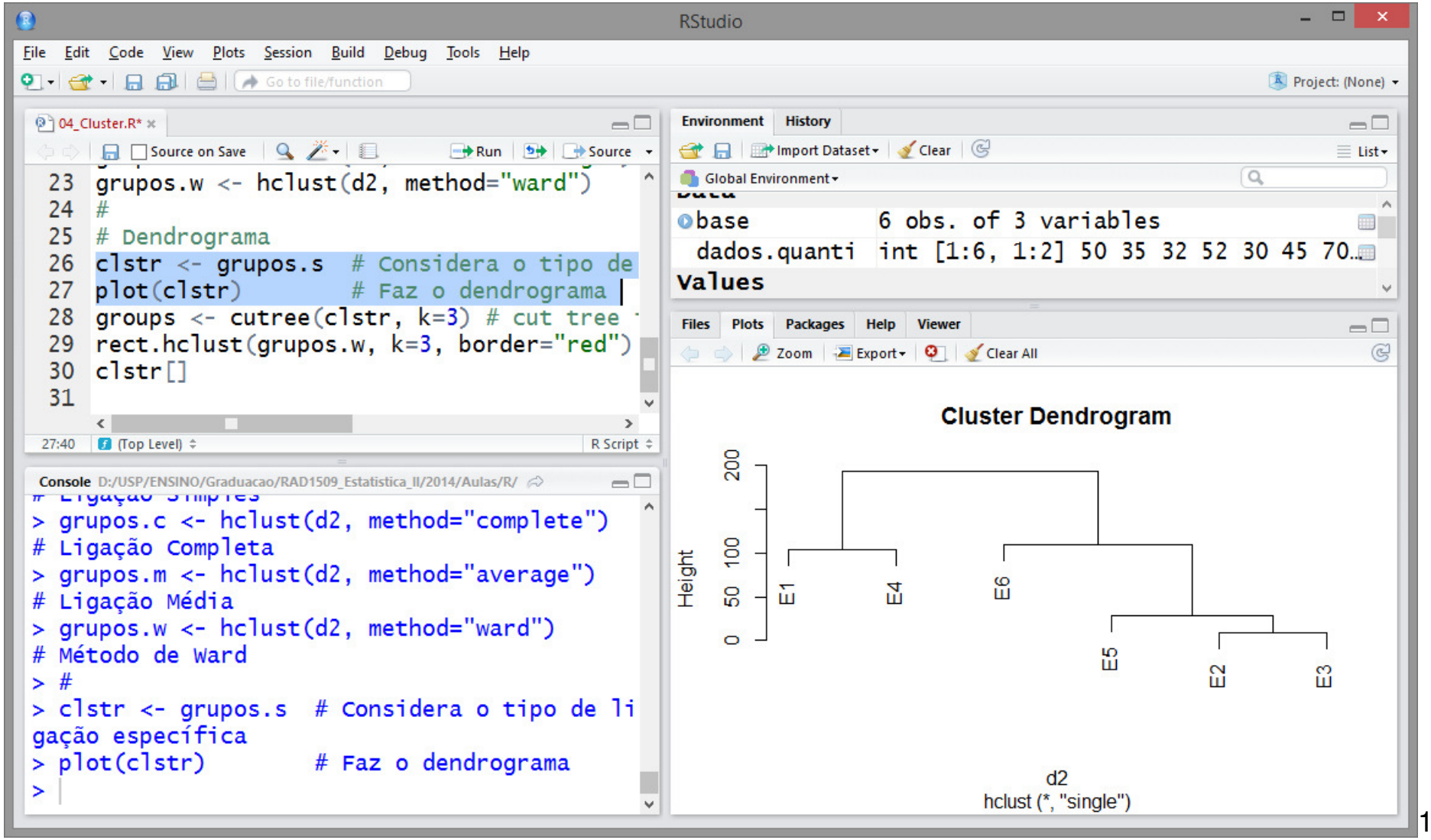
Análise de Agrupamentos no R

Obtenção de Clusters pelo método hierárquico com vários métodos de ligação: simples, completa, média e Ward.



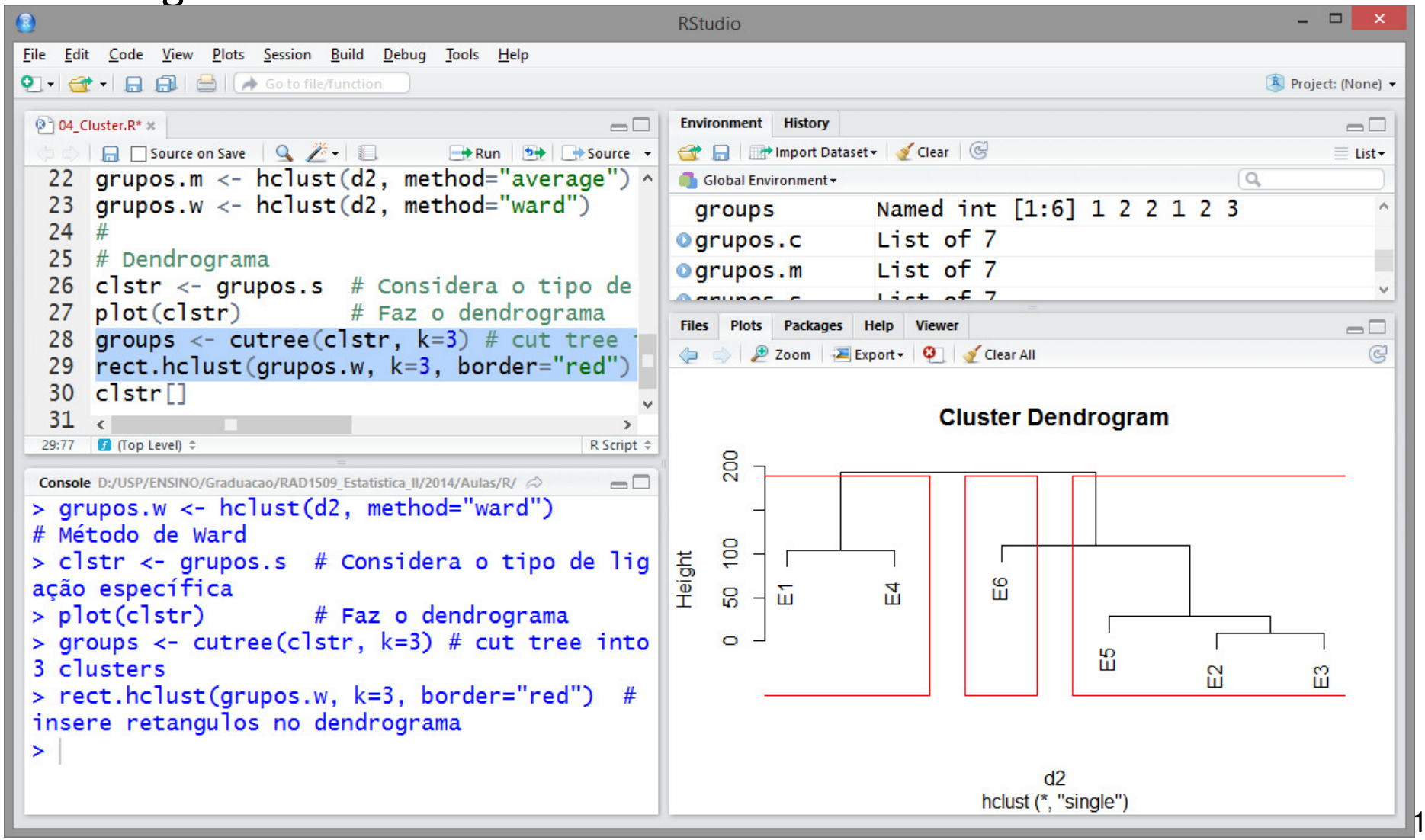
Análise de Agrupamentos no R

Selecione um dos modelos feitos anteriormente e faça o dendrograma.



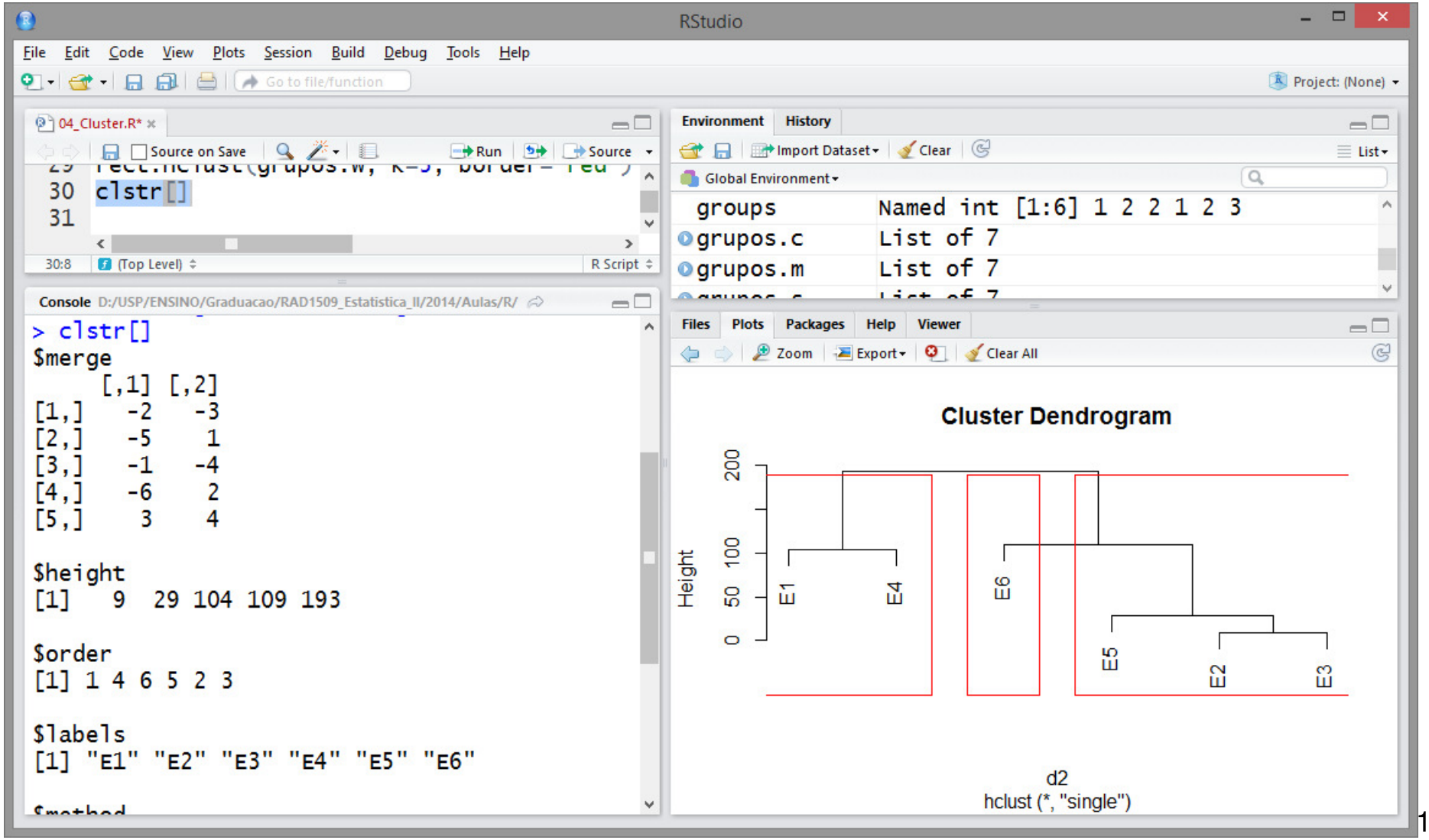
Análise de Agrupamentos no R

Selecione um número de grupos adequado para ser observado no dendrograma.



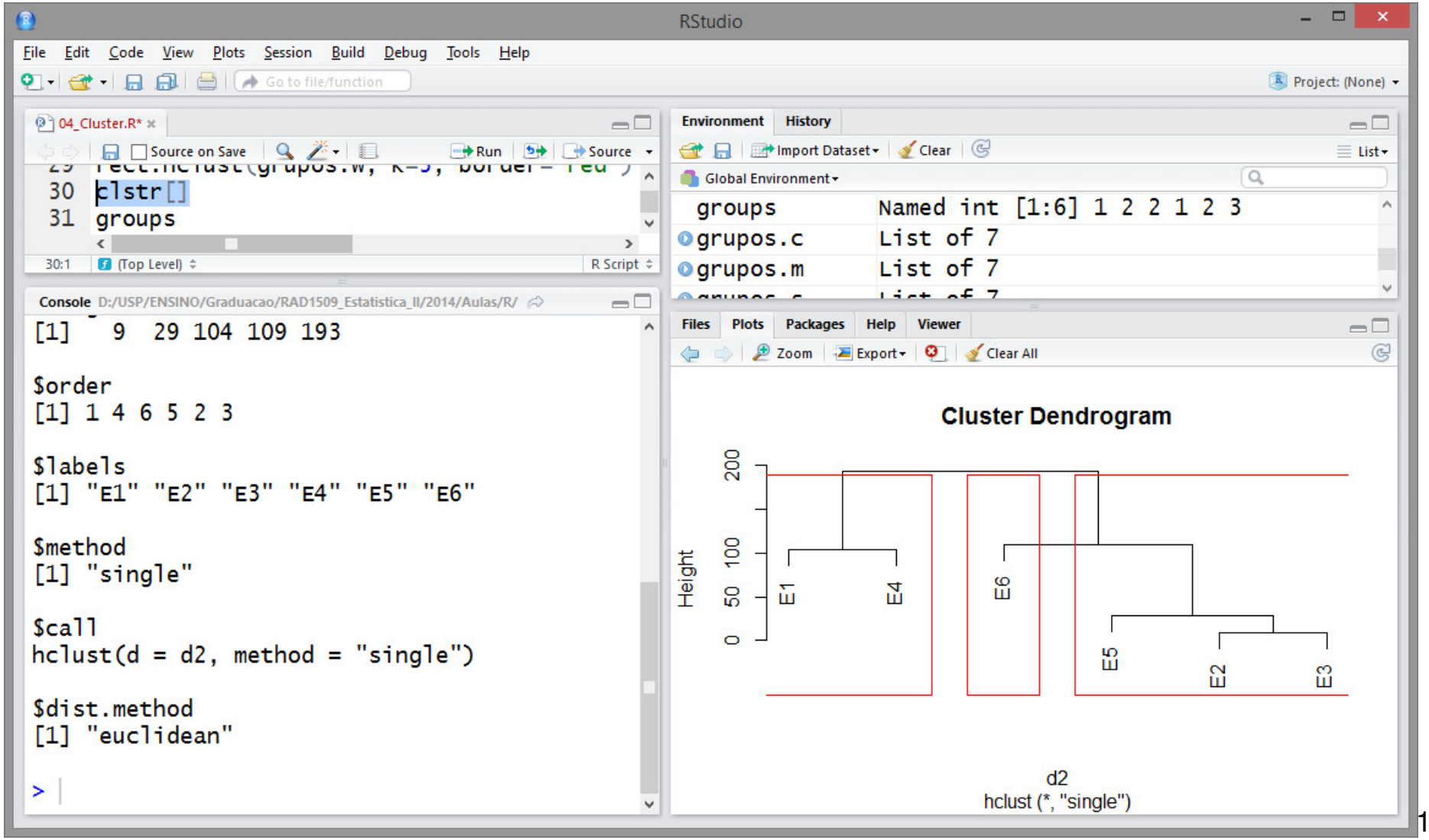
Análise de Agrupamentos no R

Verifique a análise completa.



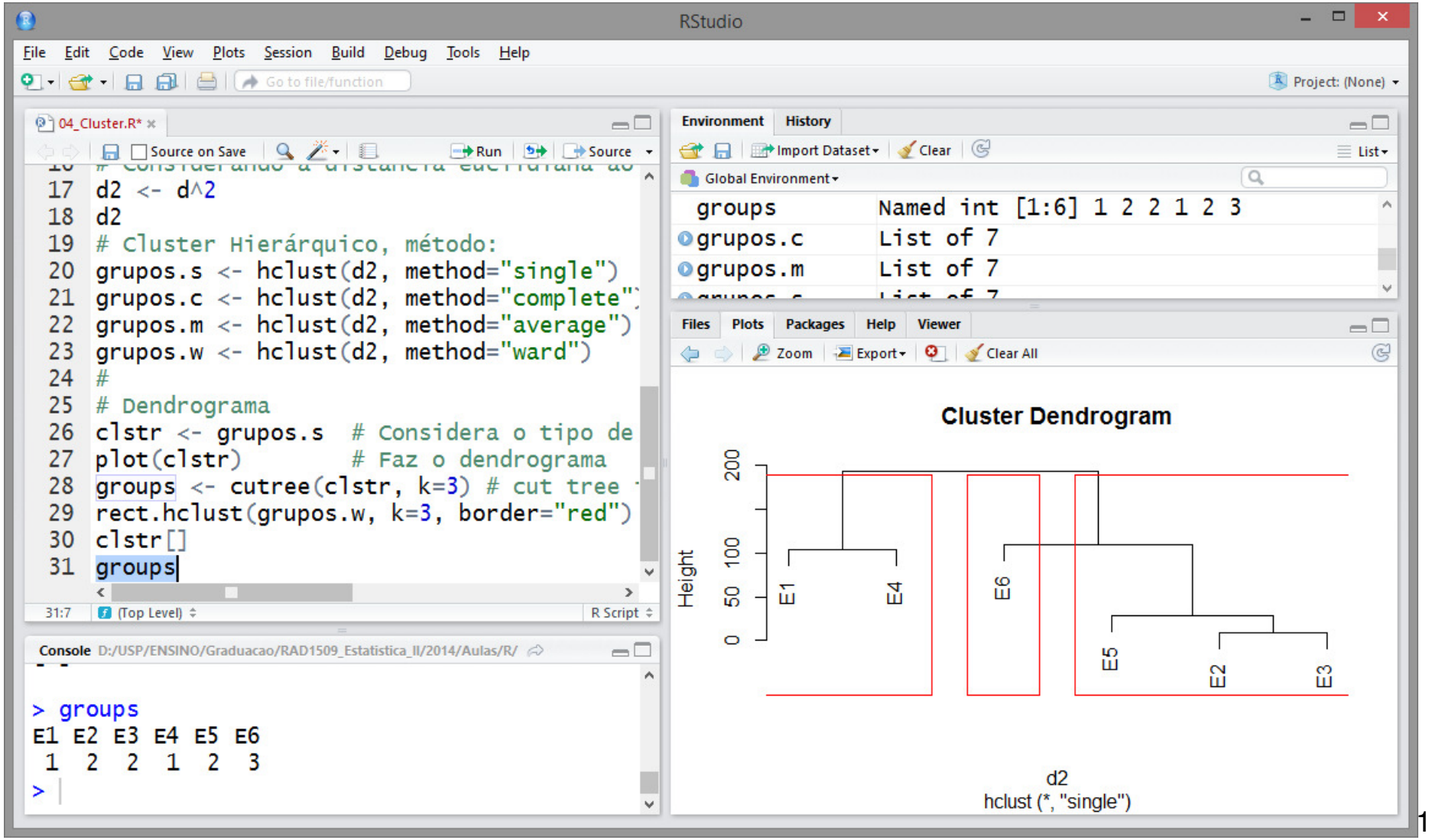
Análise de Agrupamentos no R

Verifique a análise completa.

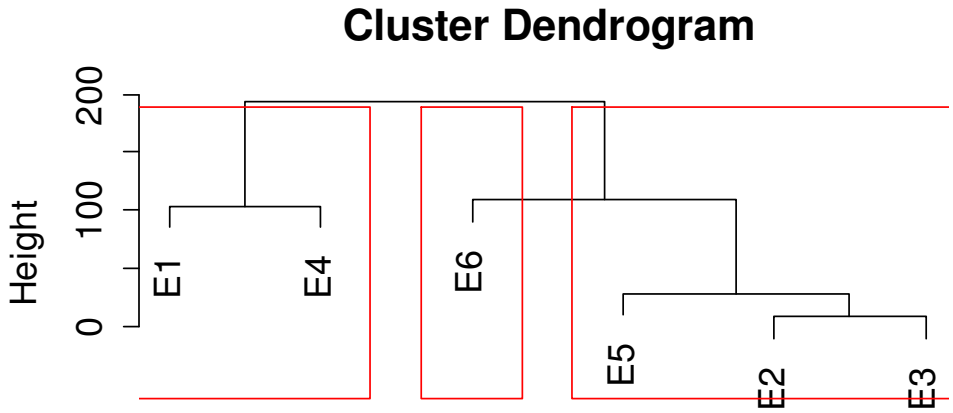


Análise de Agrupamentos no R

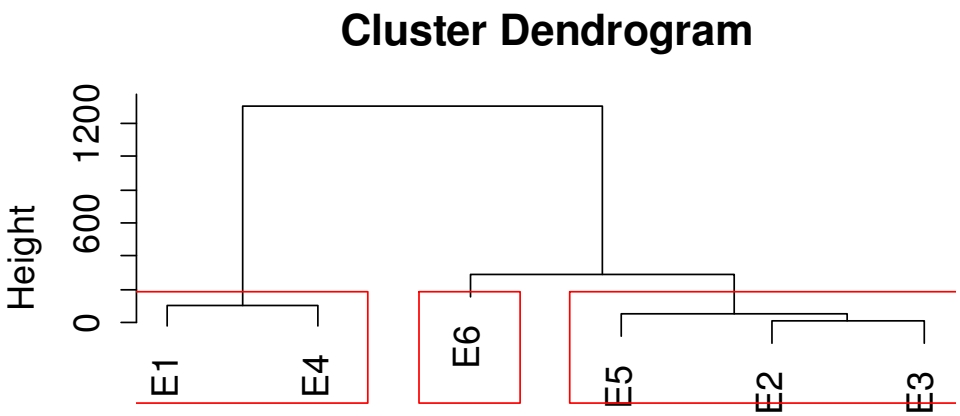
Obtenha a lista dos membros de cada agrupamento



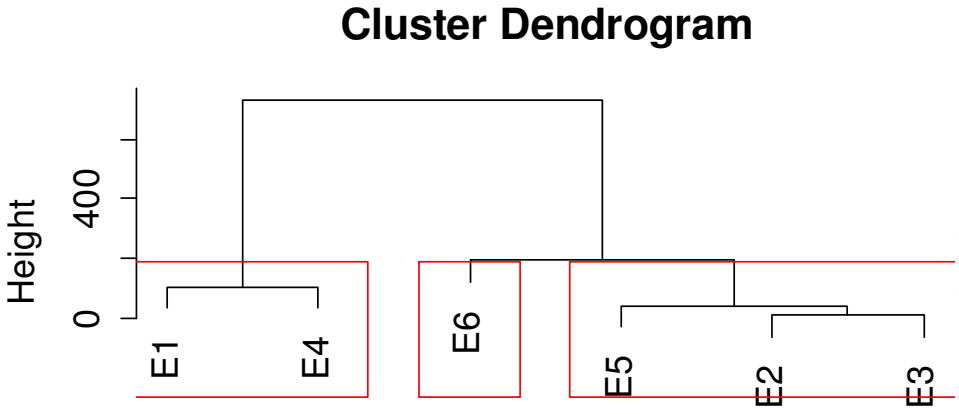
Análise de Agrupamentos no R



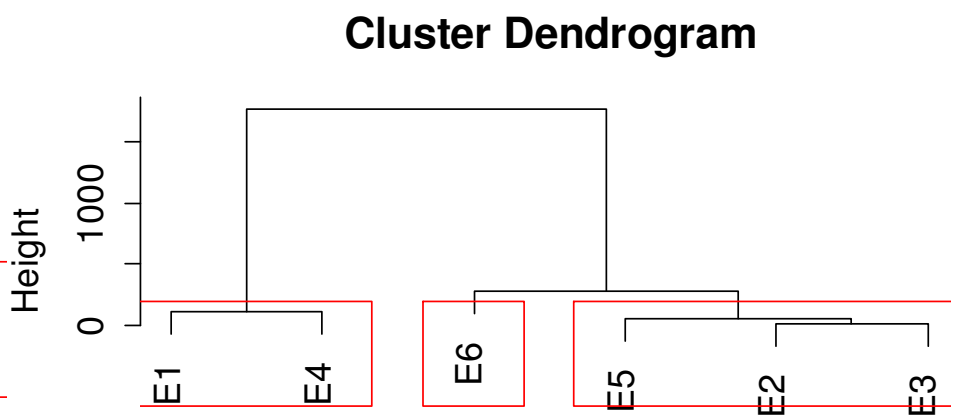
d2
hclust (*, "single")



d2
hclust (*, "complete")



d2
hclust (*, "average")



d2
hclust (*, "ward")