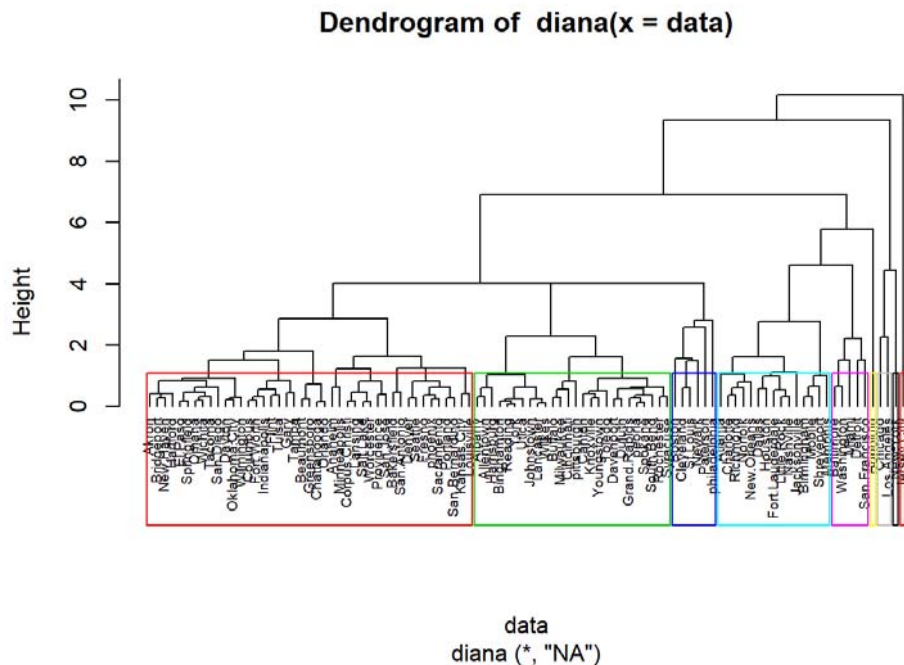


También puede visualizar los clústeres dentro del dendrograma colocando bordes como se muestra a continuación

```
pltree(hc4, hang=-1, cex = 0.6)
rect.hclust(hc4, k = 9, border = 2:10)
```



## dendextend

Puede hacer muchas otras manipulaciones en dendrogramas con el paquete `dendextend`, como - cambiar el color de las etiquetas, cambiar el tamaño del texto de las etiquetas, cambiar el tipo de línea de ramas, sus colores, etc. También puede cambiar el texto de la etiqueta así como también ordenarlo. Puede ver las muchas opciones en la viñeta en línea del paquete.

Otra función importante que ofrece el paquete `dendextend` es el `tanglegram`. Se usa para comparar dos dendrogramas (con el mismo conjunto de etiquetas), uno frente al otro y con sus etiquetas conectadas por líneas.

Como ejemplo, comparemos los métodos de vinculación único y completo para la función de Agnes.

```
library("dendextend")
```

```
hc_single <- agnes(data, method = "single") hc_complete <- agnes(data, method = "complete")
```

converting to dendrogram objects as `dendextend` works with dendrogram objects

```
hc_single <- as.dendrogram(hc_single) hc_complete <- as.dendrogram(hc_complete)
```

```
tanglegram(hc_single, hc_complete)
```

Esto es útil al comparar dos métodos. Como se ve en la figura, uno puede relacionarse con la metodología utilizada para construir los conglomerados al observar esta comparación. Puede encontrar más funcionalidades del paquete `dendextend` aquí .

## Notas finales

Espero que comprendas mejor los algoritmos de agrupamiento que con los que comenzaste. Discutimos sobre las técnicas de agrupamiento divisional y de aglomeración y cuatro métodos de vinculación, a saber, el método individual, completo, promedio y de Ward. A continuación, implementamos las técnicas discutidas en R usando un conjunto de datos numéricos. Tenga en cuenta que no teníamos ninguna variable categórica en el conjunto de datos que utilizamos. Debe tratar las variables categóricas para incorporarlas a un algoritmo de agrupamiento. Por último, discutimos un par de diagramas para visualizar los grupos / grupos formados. Tenga en cuenta que se ha asumido que el valor de 'k' (número de clústeres) es conocido. Sin embargo, este no es siempre el caso. Hay una serie de heurísticas y reglas generales para elegir el número de clústeres. Una heurística dada funcionará mejor en algunos conjuntos de datos que otros. Lo mejor es aprovechar el conocimiento del dominio para ayudar a establecer la cantidad de clústeres, si es posible. De lo contrario,

