

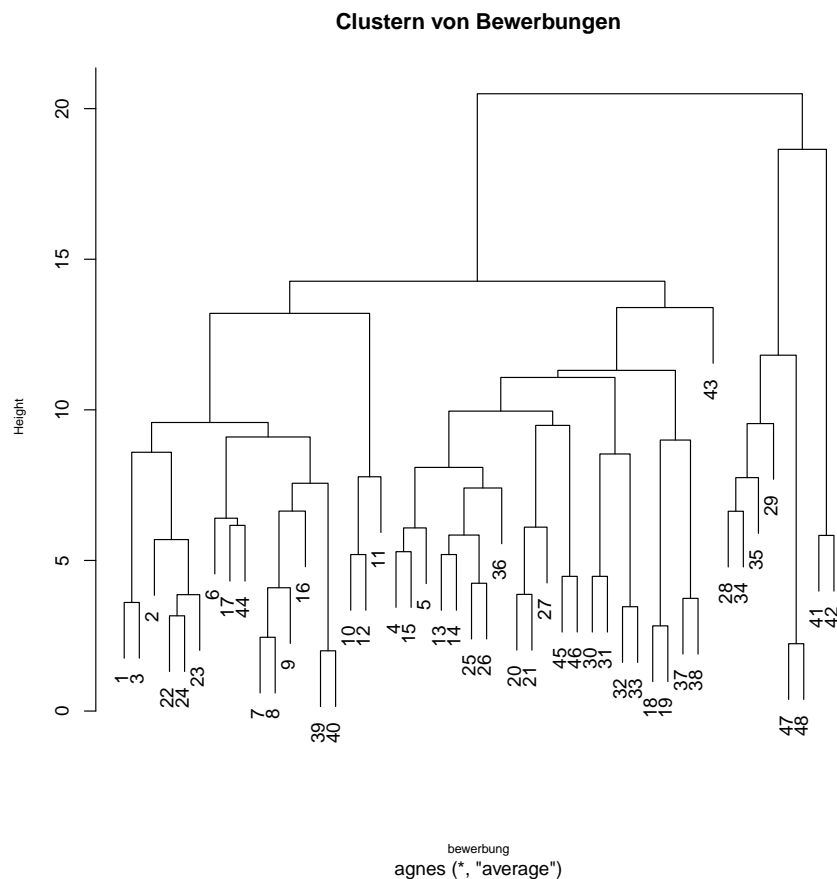
## Lösung zu Kapitel 12: Beispiel 3

Die Daten beziehen sich auf 48 Bewerbungen um eine Position in einem Unternehmen. Diese Bewerbungen wurden anhand der 15 Variablen im Datenfile `bewerbung.csv` bewertet (je höher der Wert, desto stärker ist die Eigenschaft ausgeprägt).

- Gibt es Gruppen ähnlicher Bewerbungen?

R

```
> library("cluster")
> bewerbung <- read.csv2("bewerbung.csv", header = TRUE)
> attach(bewerbung)
> clust_bewerbung <- agnes(bewerbung)
> pltree(clust_bewerbung, main = "Clustern von Bewerbungen")
```



**Abbildung 1:** Dendrogramm: Bewerbungen

Aus dem Ergebnis muss eine Aufteilung in nicht zu viele aber doch homogene Cluster ausgewählt werden. Zieht man im Dendrogramm (► Abbildung 1) in der Höhe von ca. 12 die Reißleine, verbleiben Cluster mit den Beobachtungen (die Cluster im Dendrogramm von rechts her interpretierend):

- 41, 42

- 28, 34, 35, 29, 47, 48
- 43
- eine Klasse mit Beobachtungen 4 .. 38
- eine Klasse mit Beobachtungen 1 .. 11

R

```
> detach(bewerbung)
```