

## Ćwiczenia 10

1. Dla zbioru `mtcars` przeprowadź hierarchiczną oraz niehierarchiczną analizę skupień. Ile skupień wydaje się najbardziej sensownym wyborem? Czy wybór miary niepodobieństwa oraz metody wiązania skupień ma istotny wpływ na liczbę skupień? Czy indeks CH oraz współczynnik zarysu potwierdzają wybraną optymalną liczbę skupień?
2. (S) Zbiór danych `votes.repub` z pakietu `cluster` zawiera informacje na temat procenta głosów oddanych na republikanów w wyborach prezydenckich w USA w latach 1856-1976. Korzystając z odległości taksówkowej jako miary niepodobieństwa oraz pełnej metody wiązania skupień wykonaj na tym zbiorze hierarchiczną analizę skupień. Narysuj dendrogram. Jaka liczba skupień wydaje się najbardziej sensowna? Następnie w układzie dwóch pierwszych składowych głównych zwizualizuj wyniki (punkty odpowiadające różnym stanom mają mieć ten sam kolor jeżeli należą do tego samego skupienia). Spróbuj do wykresu dodać nazwy stanów.