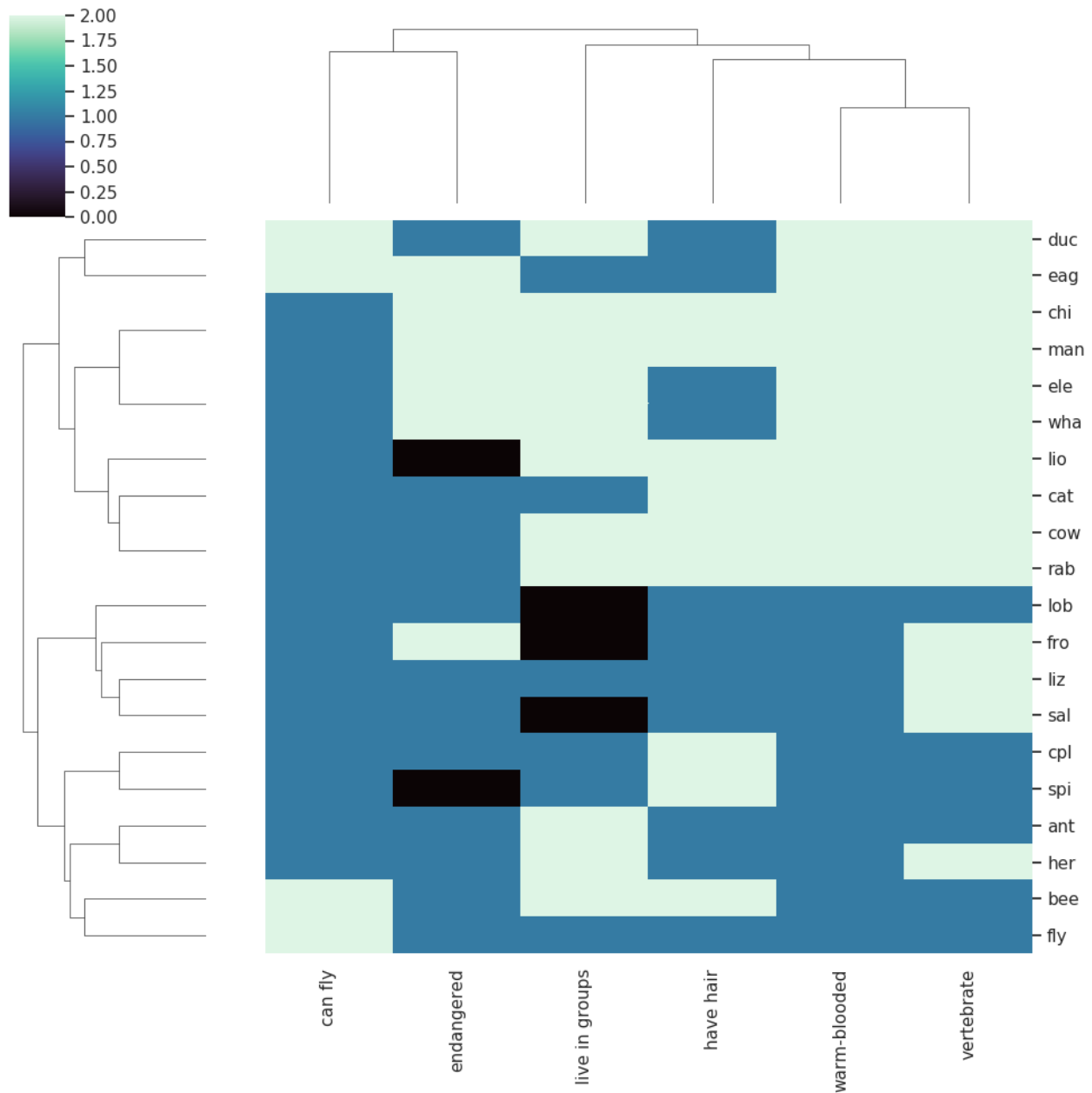


Zadanie 7

Bartosz Kozłowski

Korzystając z zbioru danych animals.csv oraz biblioteki seaborn i pandas narysowano heatmap'e z klastrowaniem. Wartości NaN w zbiorze animals uzupełniono wartościami 0.



Obliczono także z wykorzystaniem biblioteki sklearn skuteczność metody k-means dla zbioru iris. Obliczenia skuteczności wykonano krosvalidację dla kombinacji 2 zmiennych, 3 zmiennych i dla wszystkich zmiennych. W przypadku k-means obliczona skuteczność oznacza jak daleko znajdują się testowane punkty od centrów skupień, gdzie otrzymane wartości są ujemne, a najlepszym wynikiem jest wartość najbliższa 0. Najlepsze wyniki są osiągane dla parametrów (1,3), (1,2,3) oraz (0,1,2,3). Z tych wyników można przypuszczać, że największy wpływ mają parametry '1' i '3'.

Parametry: 0, 1
-15.8798784463
Parametry: 0, 2
-29.4701868495
Parametry: 0, 3
-17.3864460424
Parametry: 1, 2
-22.8247218861
Parametry: 1, 3
-9.0340792391
Parametry: 2, 3
-24.948737804
Parametry: 0, 1, 2
-14.9297627389
Parametry: 0, 1, 3
-10.8290527777
Parametry: 0, 2, 3
-14.0459335765
Parametry: 1, 2, 3
-10.3974123445
Parametry: 0, 1, 2, 3
-8.32302135353