

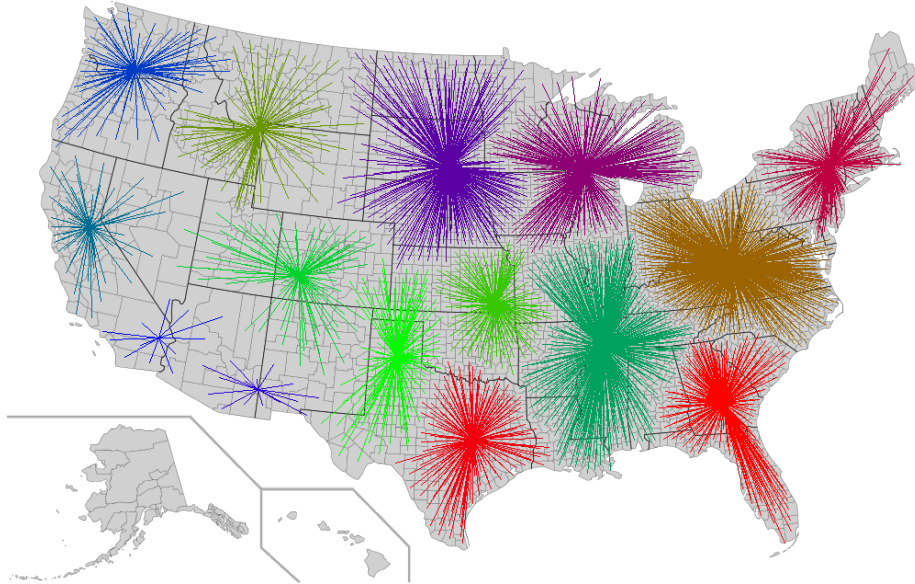
Lab. 4 - Clustering di dati medici

Ballarin Simone, Gobbo Alessio, Rossi Daniel

June 2019

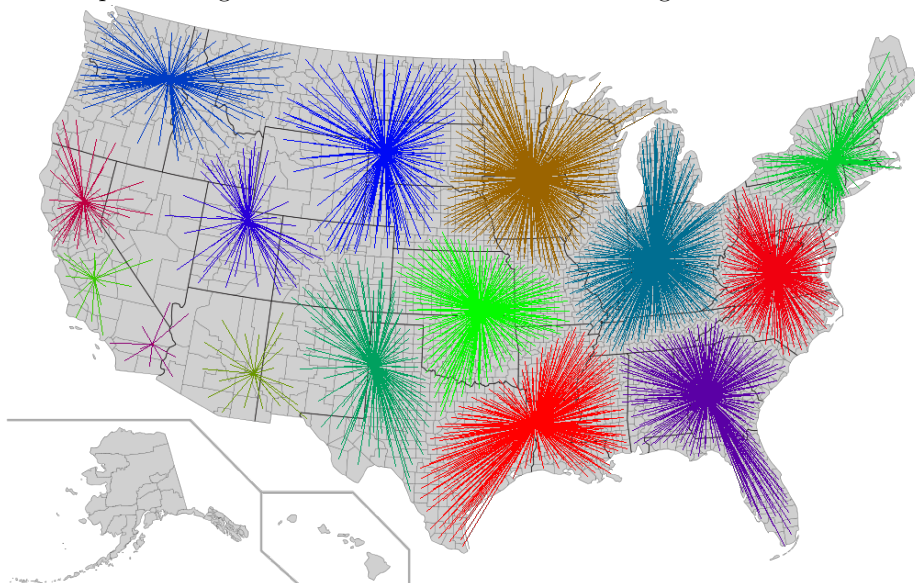
Domanda 1

Viene riportato il grafico ottenuto dall'esecuzione dell'algoritmo Kmeans.



Domanda 2

Viene riportato il grafico ottenuto dall'esecuzione dell'algoritmo Gerarchico.



Domanda 3

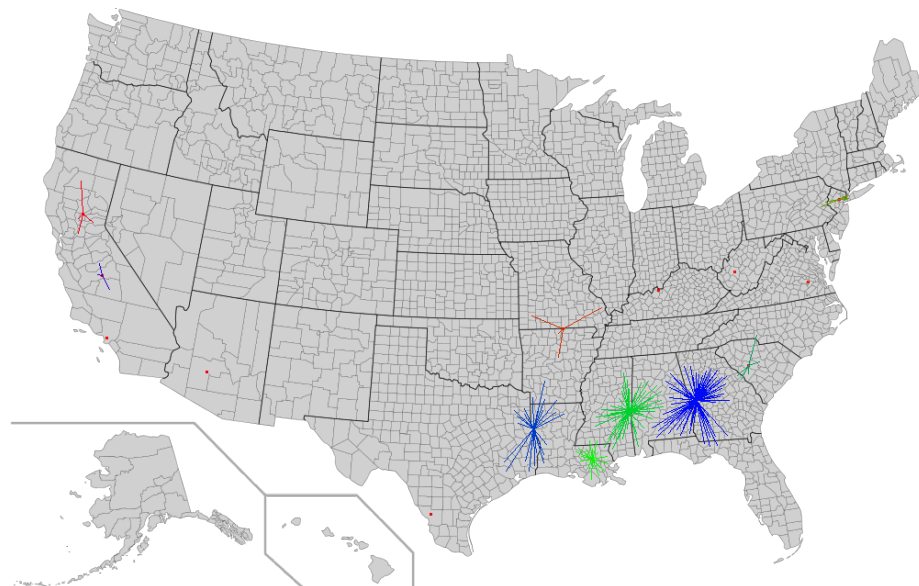
Dati q il numero di iterazioni di *kmeans*, k il numero di clustering e n la cardinalità dell'insieme dei punti, possiamo quantificare le seguenti complessità dei due algoritmi di clustering:

- La complessità del clustering gerarchico è $O(n h(n))$, dove $h(n)$ è la complessità della ricerca dei punti più vicini. Questa ricerca è stata effettuata con la funzione *FastestClosestPair*, la quale ha complessità $O(n \log(n))$; Quindi complessivamente il clustering gerarchico ha complessità $O(n^2 \log(n))$.
- La complessità di *kmeans* invece è di $O(q n k)$.

Ipotizzando quindi q e k fattori molto piccoli, quindi trascurabili come segue $q \approx 1$ e $k \approx 1$, giungiamo alla conclusione che la complessità del clustering gerarchico rimanga inalterata, mentre *kmeans* giunga a $O(n)$.

Da quando appena esposto risulta evidente come *kmeans* sia in termini asintotici, e considerando q e k insignificanti, significativamente più efficienti.

Domanda 4



Domanda 5

