

### PROVA TEORICO\_PRATICA 1 - - Estratta

Dato un *dataset* composto da 100 linee cellulari tumorali, ognuna delle quali associata all'espressione di una coorte di 10000 geni, elaborare lo pseudo-codice di un algoritmo in grado di dividere i campioni in gruppi il più possibile omogenei rispetto all'espressione dei loro geni.

### PROVA TEORICO-PRATICA 2 -Non estratta

Elaborare lo pseudo-codice per estrarre tutti i *Repetitive Elements* a partire da un *file* in formato **fastq**, tenendo presente i limiti dei *software* di allineamento di *read* su genoma murino.

### PROVA TEORICO-PRATICA 3 -Non estratta

Dato un *dataset* di espressione genica associata a diverse linee cellulari, elaborare lo pseudo-codice per effettuare la normalizzazione dei dati in modo che siano confrontabili, considerando anche il caso in cui i campioni siano pochi (e.g.  $< 5$ ) .