

Input: a data set containing MAF or GISTIC scores (e.g., as obtained from cBio portal) (Cerami et al. (2012); Beroukhim et al. (2007)).

Result: a data structure containing Boolean flags for “events,” relative frequencies and other metadata.

- 1 From the data set (depending on the data format) derive a Boolean matrix M , where each entry $\langle i, j \rangle$ is **true** if event i is “present” in sample/patient j ;
- 2 **forall** *events* e **do**
 - 3 Compute the *frequency* of the event e in the data set and save it in a map F ;
 - 4 Compute the *joint probability* of co-occurrence of pair of events in the data set and save it in a map C ;
- 5 **end**
- 6 **return** *A data structure comprising the Boolean matrix M , the maps F and C and other metadata.*