

## What are the axes of the 'chemical structure' space plot?

Written by Ed Champness

Tuesday, 29 September 2009 19:55 - Last Updated Tuesday, 16 February 2016 13:45

---

It is not possible to quantify the axes of the structural chemical space plot as they do not correspond to a particular property or descriptor. The structural chemical space plot is a two-dimensional approximation of a multi-dimensional space, where similarity between each molecule in the data set represents a single dimension in the multi-dimensional space. Using either 'Visual Clustering' (tSNE - t-Distributed Stochastic Neighbour Embedding) or Principle Component Analysis (PCA), this is reduced to a two-dimensional space that maximizes the visible variation. As a result, the two dimensions are effectively a function of the molecules' similarities and do not have an explicit meaning. The dimensions are best considered as distance measures such that the closer molecules are in the plot, the greater their similarity. Bear in mind that the more molecules there are in the plot, the more approximate the visible distance will be for any two individual points, although the overall trend covers the diversity of the entire set.