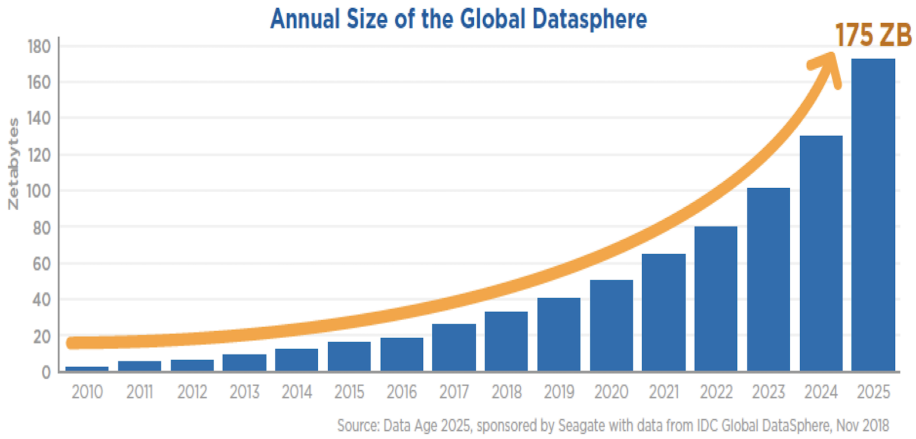
Alberto Andrés Valdés González.
Degree: Mathematical Engineer.
Work position: Data Scientist.
Mail: anvaldes@uc.cl/alberto.valdes.gonzalez.96@gmail.com
Location: Santiago, Chile.

Topological Data Analysis (TDA)

Topological Data Analysis or TDA is an up-and-coming approach to data analysis that focuses on looking at the shape of data.

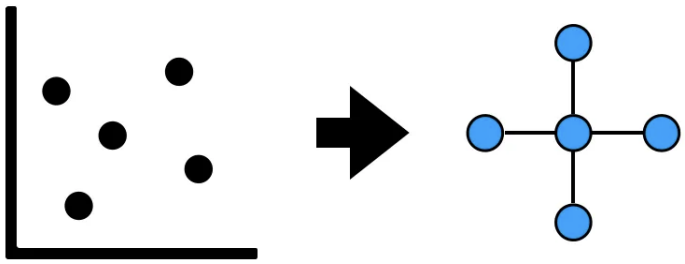
Introduction

Data volumes across differents domains seem to be increasing at an accelerating rate and in many cases is a good thing. For example it served as the fuel for so many technological innovations in the past decade like recommendation systems and NLP processing. But there's a problem: "More data, more problems".



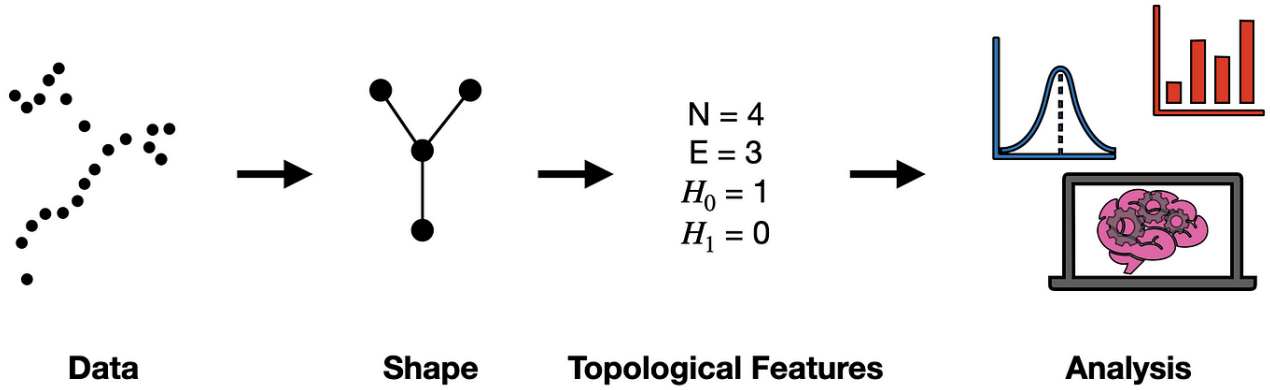
The data from the real world is often **noisy** and have **high dimensional**.

"Data has shape and shape matters."



The basic idea of TDA is to take data and translate it into some kind of shape. The reason is that the undelying shape of our data **is often much more robust to noise and perturbations** and also the TDA algorithm are readily **applicable to high dimensional data**.

TDA Pipeline



Techniques

1. Mapper Algorithm: With this algorithm we can take any generic data and translate it into a graph.

2. Persistent Homology: Persistent homology (PH) is a mathematical tool in computational topology that measures the topological features of data that persist across multiple scales with applications ranging from biological networks to social networks.