**How to Databases**

**Wikidata –**

Wikidata is a large database that stores their data in a Resource Description Framework (RDF) format. This format creates a network of all its datapoints, where each individual point has a unique identifier. Points are connected by their *relationships* to other points, making it easy to find datapoints that share a particular feature or belong to a certain category. That is the basis of how RDF databases are structured and what allows users to explore the relationship between datapoints.

**Making Sense of Databases SPARQL –** **Sparql Protocol And RDF Query Language**

To find relationships between datapoints in online databases such as Wikidata, we need to write a program, called a query, at will allow users to access the data they need. SPARQL is a query language specifically used for Wikidata as well as other compatible bases. As an example, we can use SPAQRL to specify what datapoint and feature we want to begin with and request the query to find all other datapoints that share that feature. The same goes for categories, were we can request to be shown all datapoints that belong to a given category.

A summary of how our SPAQRL query retrieves data is as follows:

Any object that has Health specialty = oncology 🡪 of those objects objects that have a genetic association 🡪 determine which cancers we want to focus on

**Visual Representation (d3.js)**