

# Project Proposal for Data Management

Valerio Massimo Camaiani 1935012

Francesco Bianchi 1942637

## Project type: NoSql usage and analysis

### Project description

The project involves the usage of a graph database to model a supply chain, entities like retail shops, manufactures and suppliers will be defined and their interaction will be studied based on the need to deliver products made of certain components.

The main areas of interest for this project are:

- Analyzing suppliers performances
- Identification of bottlenecks in the execution of a given scenario
- Best fallback procedure in case of random failure of one of the components
- Benchmarking of the above scenarios
- \*Depending on whether the above analysis dont end up in particularly interesting results:  
Performance analysis with a SQL database.

### Dataset description

Taking inspiration from a couple of public dataset available we generated a synthetic dataset with the following characteristics. (This is a first draft, future iterations might add or modify some of the attributes to allow more interesting analysis)

### Nodes

#### 1. Product

- a. ID (Integer): Unique identifier for each product.
- b. Sell\_Price (Float): The selling price of the product.

#### 2. Component

- a. ID (Integer): Unique identifier for each component.

#### 3. Manufacturer

- a. ID (Integer): Unique identifier for each manufacturer.

#### 4. Supplier

- a. ID (Integer): Unique identifier for each supplier.

- b. Component\_ID (Integer): ID of the component they supply. This creates a direct link to the Component node.
- c. Price (Float): Price at which the supplier provides the component.
- d. Time (Integer): Time taken to deliver the component.
- e. Max\_Cap (Integer): Maximum capacity of components the supplier can deliver in a specific time frame.

## 5. Retail

- a. ID (Integer): Unique identifier for each retail entity.

## Relationships

1. **MAKES** (Manufacturer to Product)
  - a. N\_Stock (Integer): Number of units of the product currently in stock at the manufacturer.
  - b. Cost (Float): Manufacturing cost per unit of the product.
  - c. Time (Integer): Time taken to manufacture the product.
  - d. Max\_Cap (Integer): Maximum production capacity of the manufacturer for the product.
2. **SUPPLIES** (Supplier to Manufacturer)
  - a. Component\_ID (Integer): ID of the component being supplied, linking back to the Component node.
  - b. Max\_Cap (Integer): Number of units supplied per delivery.
3. **COMPOSES** (Component to Product)
  - a. No attributes, simply indicates which components are part of a product.
4. **OFFERS** (Retail to Product)
  - a. No attributes, indicating which products are offered by a retail entity.
5. **Ships** (Either from Supplier to Manufacturer or from Manufacturer to Retail)
  - a. N\_Items (Integer): Number of items per container.
  - b. Time (Integer): Time taken for the shipment from source to destination.
  - c. Cost (Float): Cost of the shipment.