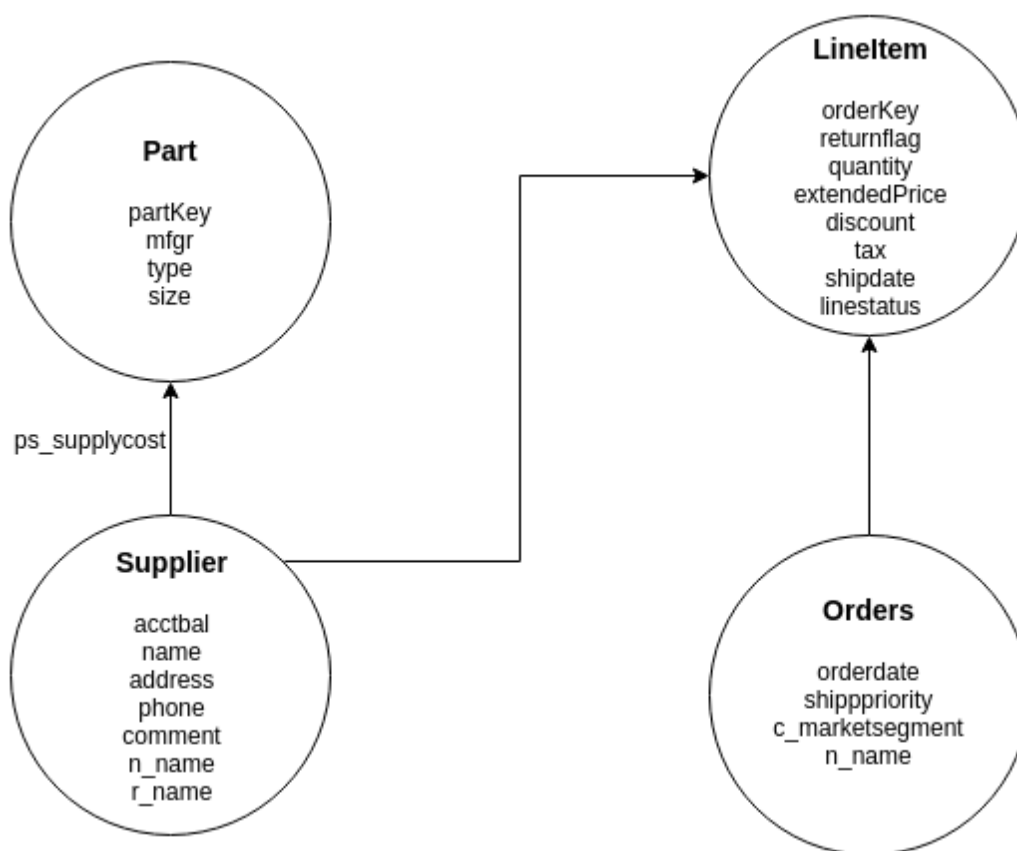


# CBDE

## Laboratori de Neo4J

### MODEL



Model 1: Graf que representa el model de la base de dades

## EXPLICACIÓ DEL MODEL

Hem decidit analitzar les quatre consultes per separat i a posteriori unir els models resultants i fusionar-los en un sol. El graf final és el que es pot observar al *Model 1*.

Hem decidit crear el node *Order* perquè es utilitza en gran part de les consultes donades. Els atributs són tots aquells que s'utilitzen en algun moment de l'execució de les consultes (*orderdate* i *shippriority*). I a més a més hem decidit afegir l'atribut *marketsegment* de *Customer* i *name* de *Nation* perquè, donat l'esquema relacionat, hem entès que per cada *Order* sempre tindrem un *Customer* i una *Nation*. En conseqüència, podem integrar ambdós atributs dins el node d'*Orders*.

Llavors, navegant pel graf arribem al node *LineItem*. Hem unit amb una aresta de *Order* a *LineItem* donat que per cada order hi haurà molt lineitems. Com a atributs, hem afegit els que són utilitzats algun cop en les consultes dins l'entitat relacional donada de *Lineitem*. A diferència d'abans, no afegim cap atribut d'alguna entitat extra perquè no és necessari.

Per altra banda tenim els dos nodes *Supplier* i *Part*. En l'esquema relacionat existeix una altra entitat anomenada *PartSupplier* que en aquest model està representat com l'aresta que uneix dos nodes. Hem decidit afegir-li l'atribut *supplycost*, ja que és l'únic utilitzat per les quatre consultes. Aquesta relació evita replicar dades en el node *Lineitem* i també evita crear una aresta de *Part* a *Lineitem*.

Com a atributs del node *Supplier*, hem decidit agafar els que les consultes utilitzen. Per part del node *Supplier*, hem realitzat el mateix i afegint també el *name* de *Nation* i *Region* perquè ambdós s'utilitzen i la cardinalitat de *Supplier* a *Nation* és 1 a 1 i de *Nation* a *Region* també d'1 a 1, en conseqüència, es poden afegir al node de *Supplier*.

## EXPLICACIÓ DELS ÍNDEXS

Referent als índexs d'aquest model, després d'estudiar bé la seva eficàcia i validesa, hem decidit instaurar-ne dos:

- *LineItem(shipdate)*: Hem decidit indexar la base de dades per aquest atribut perquè tant en la primera com en la tercera consulta de les quatre s'utilitza com a restricció del *where*. Això aportarà una molt millor performance a les consultes.
- *Orders(orderdate)*: Després d'una anàlisi consulta per consulta, hem arribat a la conclusió que seria el millor atribut per aplicar-li un índex en el nostre model de base de dades. Gran raó són les condicions de les consultes 3 i 4, on utilitzen aquest atribut com filtratge de la query.