

**Northwind Traders** ®

# Enhancing Product Recommendations and Warehouse Efficiency through Data-Driven Analytics

DATASCI 205 Project 3

Courtney Chen, Hong Hu, Kevin Pradjinata

Confidential

Copyright ©



# Neo4j's ability to detect relationships in the data enables diverse use cases, including warehouse and cross-selling opportunities.

## Current situation

Suboptimal warehouse storage design

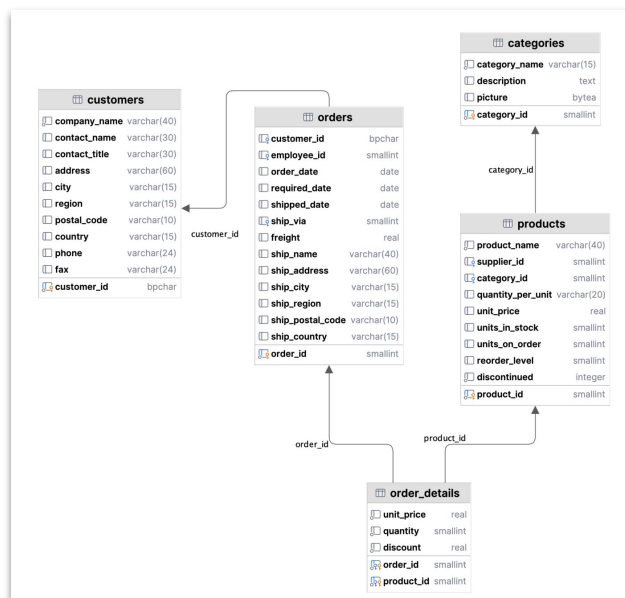
Purely transactional checkout process

Limited integration of analytics into decision-making

Confidential

Copyright ©

## Leveraging the Northwind database



- The Northwind database is a complete sales database containing **14 tables**.
- We will focus on: **customers, orders, categories, products, and order details**.
- Using **Neo4j**, we can harness relationships to inform warehouse placement and tailor recommendations.

# Using Neo4j's triangle count, frequently co-purchased products can be recommended to customers to increase cart value.

## Application of Neo4j's triangle count

- Neo4j triangle count computes the number of co-purchase "triangles" each product appears in.
- A higher triangle count means a product is widely purchased with other products.
- Triangle count can be used to provide tailored recommendations at checkout to customers.


## Most co-purchased products

	product_id	product_name	triangle_count
0	60	Camembert Pierrot	1074
1	41	Jack's New England Clam Chowder	1050
2	2	Chang	1039
3	77	Original Frankfurter grüne Soße	1019
4	13	Konbu	1001

## Recommender system

 Your cart:

Item	Price
Aniseed Syrup	\$9.50
Cart Total	\$9.50

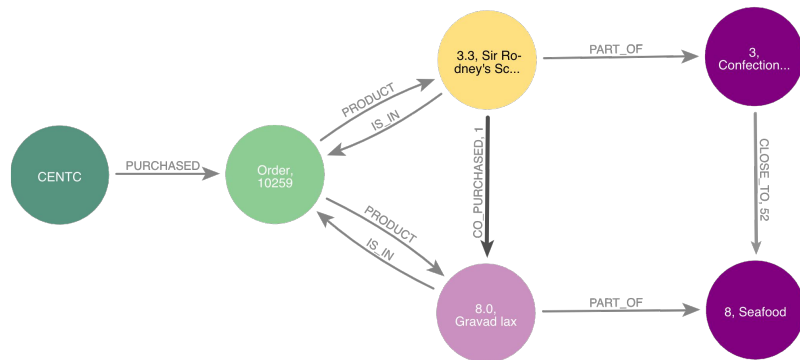
 Based on your cart, you might also like:  
60: Camembert Pierrot (Triangle Count: 1074)  
41: Jack's New England Clam Chowder (Triangle Count: 1050)  
77: Original Frankfurter grüne Soße (Triangle Count: 1019)

# To optimize warehouse placement, we identified product categories that benefit from being stored in close proximity.

## Description of the CLOSE\_TO product link

- Frequently co-purchased products are connected by a weighted CO\_PURCHASED relationship, capturing the strength of their association.
- A weighted CLOSE\_TO relationship between categories guides how closely products should be stored in the warehouse.
- Grouping related products minimizes travel distance, accelerating pick times and boosting efficiency.

## Sample snapshot of product relationships



# The hybrid community ID combines Louvain clustering with product-level requirements to improve warehouse organization.

## Updates to Louvain's community ID

- The metric assigned by the Louvain algorithm does not account for specific storage needs.
  - o Certain product categories (e.g., dairy, seafood) require freezer storage.
  - o Others (e.g., grains/cereals, condiments) need dry storage.
  - o Some (e.g., meat/poultry, produce) require refrigerator storage.
- Updating the default metric to incorporate product-specific storage needs ensures food safety while preserving meaningful product groupings.

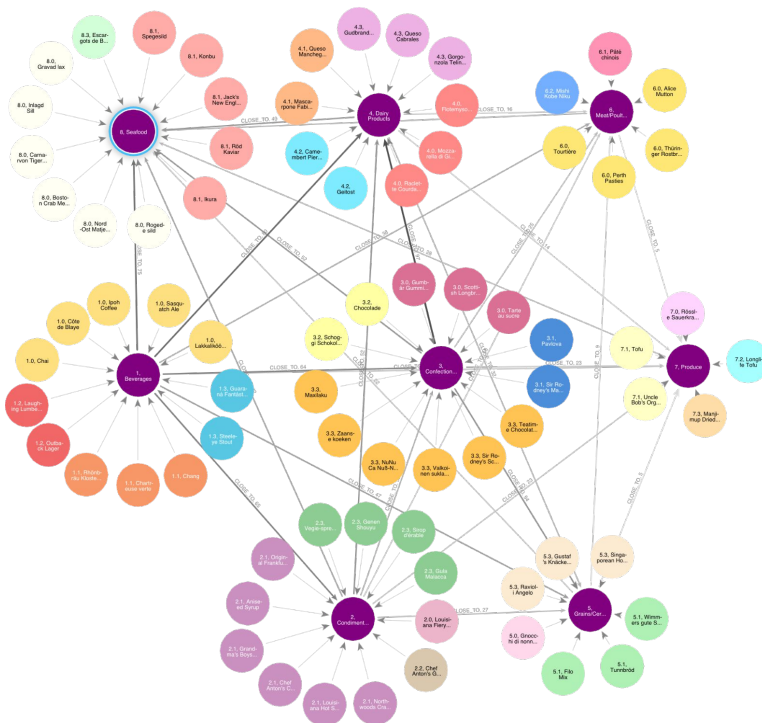
## Composition of the hybrid community ID



# The final warehouse blueprint optimizes placement based on product category, Louvain communities, & domain knowledge.

## The placement logic is based on 3 principles:

- **Category Principle:** Group products of the same category in one section. Stronger 'CLOSE\_TO' relationships mean their sections should be closer (e.g., Seafood next to Beverages).
- **Closeness Principle:** Products in the same Louvain community should be grouped tightly within their category and positioned near related products across categories. (e.g. Boston Crab Meat [8.0] close to Chai [1.0]).
- **Common Sense Principle:** Place categories requiring similar storage equipment close to each other (e.g., Condiments beside Grains/Cereals).



## Summary



Tapping into historical sales data from the existing Northwind Traders database enables greater efficiency in both inventory management and user experience.



Neo4j identifies consumer preferences at both individual and group levels, enabling customized experiences while driving net sales and increasing mean cart value.



Discovering product communities based on co-purchases, storage needs, and product communities offers opportunities to reduce warehouse labor costs and travel times.



Next steps: 3D digital twin in Simio that models batch picking time through a warehouse, with and without strategic storage, to evaluate differences in performance.

# Thank you! Questions?