



Identify Potential Product Savings by an Innovative Network Graph Clustering Method

AUTHORS:

- Ge Zhuo Ran, Financial Analyst
- Wang Min Hao, Assistant Director, Finance



INTRODUCTION

A project was set up to analyze product prices, whether there are different prices sourced among institutions for a similar item.

From the purchase orders (PO) data collected, we notice there is a high percentage of the medical supplies item not having a standardized material coding and description.

It is challenging to cross-compare PO data among the institutions on the same medical supply item due to the inconsistent data.

OBJECTIVE

Leveraging on PO data, group similar/identical medical supplies purchased across institutions and compared the prices to identify potential savings.

METHODOLOGY

We explored using the Innovative Network Graph Clustering Method. The PO data is innovatively represented in a large network graph. Each unique PO line forms a node. And the connectivity between nodes (edge) is evaluated by both description and price similarity.

Similar/identical items can be clustered together by locating a densely connected subgraph from this method.

RESULTS

The network graph clustering method is evaluated to be >80% accurate.

We have completed the analysis of PO data from 2017-18 for 26 PHIs. Using the network graph clustering method, we identified potential annual savings of **\$3.1 Million**.

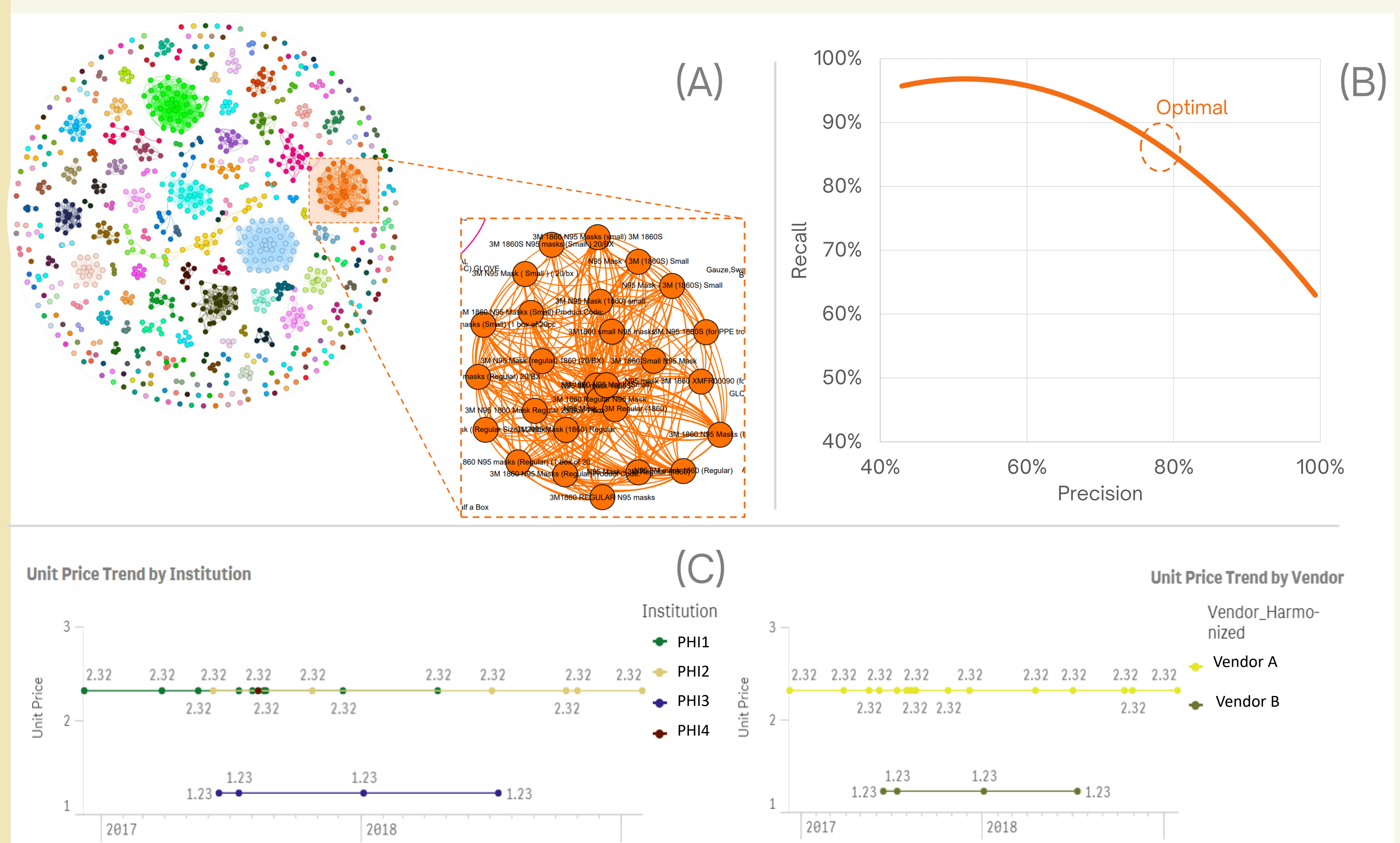


ANALYSIS

(A) Visual illustration of the network graph, showing various short texts of 3M 1860 N95 mask. The Network Graph Clustering Method group similar short texts and form a densely connected subgraph.

(B) The Network Graph Clustering Method achieved >80% recall and precision level on optimal parameters (Evaluated on 1,000 labelled data).

(C) Illustrated potential savings on item (IV Adm Set): PHI 3 purchased the item from Vendor B with a lower price than other institutions.



CONCLUSION

Traditionally, analyzing PO data across institutions is challenging, if not impossible, due to the lack of standard coding and description. The network graph clustering method successfully tackled this difficulty and achieved good results. Subsequently, the technique has also been applied in other applications, such as spend and saving analysis and Covid-19 incremental cost analysis.