GVCdtLAB
an initiative by CIRANO

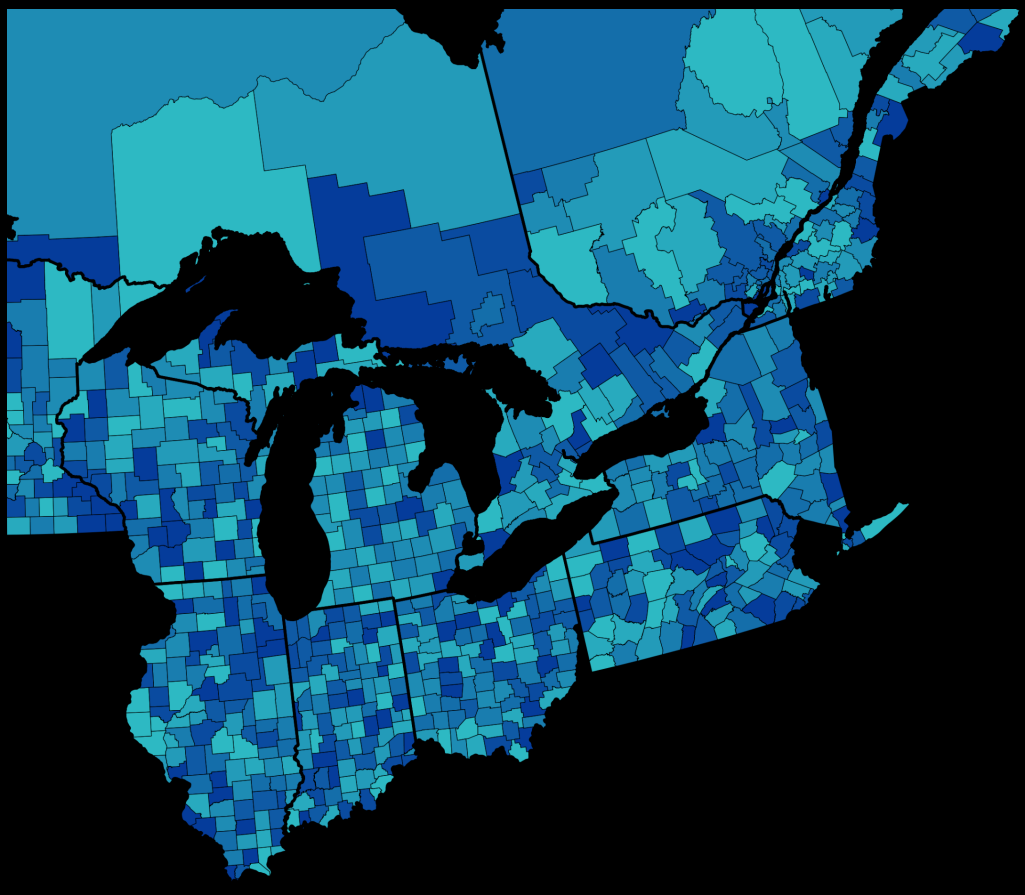
CIRANO
Knowledge into action

EXECUTIVE SUMMARY

March 2025

Unveiling Quebec's economic clusters: Harnessing machine learning and real-time data from the registre des entreprises

Lucien Chaffa and Thierry Warin



Abstract

This study explores the dynamics of economic clusters in Québec by applying machine learning algorithms to near real-time, firm-level data from the Registre des Entreprises du Québec (REQ). Through a novel integration of advanced data science techniques and traditional economic analysis, the research identifies and characterizes economic clusters, offering insights into their formation, evolution, and regional economic impact. Employing algorithms such as K-Means and Hierarchical Clustering, the analysis leverages the granularity of REQ data to examine cluster attributes, including industry composition, spatial distribution, and temporal trajectories. The findings demonstrate the efficacy of machine learning in revealing intricate patterns and inter-industry linkages, shedding light on emerging clusters and shifts in economic activities. This study provides valuable guidance for policymakers by highlighting opportunities for targeted economic development, strategic resource allocation, and enhanced regional competitiveness.

Introduction and Contribution

This research contributes to the economic literature by:

- Introducing an ML-driven approach to analyze economic clusters, moving beyond traditional static methods.
- Leveraging real-time data to capture evolving industry dynamics, allowing for more responsive policy interventions.
- Demonstrating how unsupervised learning techniques (e.g., K-Means, Hierarchical Clustering) can reveal hidden patterns in industrial growth trajectories and firm interactions.
- Providing empirical insights into the geographic and temporal distribution of clusters, highlighting strategic growth opportunities.

Methodology

The study applies machine learning algorithms to classify and track economic clusters based on industry growth patterns. The methodology involves:

- **Data Collection**: Utilizing high-frequency REQ data to examine firm characteristics, locations, and industry evolution.
- **Clustering Algorithms**: Implementing unsupervised ML techniques, including K-Means, Hierarchical Clustering, and the Louvain community detection algorithm, to identify clusters based on firm growth trajectories.
- **Validation**: Employing a rigorous validation framework that assesses the coherence and distinctiveness of the identified clusters.
- **Spatial & Temporal Analysis**: Mapping cluster locations across Québec and analyzing industry growth trends over time.

Key Findings

1. **Economic Clusters Exhibit Distinct Growth Trajectories**: Machine learning effectively identifies industry clusters based on firm growth rates, revealing intricate inter-industry linkages.
2. **Geographic Concentration of Clusters**: Clusters are highly concentrated in major urban areas like Montreal, indicating strong regional agglomeration effects.
3. **Temporal Trends Highlight Economic Shifts**: The number of firms within clusters fluctuates over time, reflecting macroeconomic conditions, policy impacts, and sectoral transformations.
4. **Industry Interdependence**: Strong correlations between industries suggest that economic shocks or policy changes affecting one sector can have spillover effects on related industries.

Policy Implications

The study's insights offer several policy recommendations:

- **Targeted Economic Support**: Policymakers can use ML-driven cluster identification to allocate resources efficiently, supporting high-growth sectors and emerging industries.
- **Regional Development Strategies**: Understanding the geographic distribution of clusters enables the design of policies that strengthen regional industrial ecosystems.
- **Resilience and Competitiveness**: Identifying inter-industry dependencies allows for proactive interventions to mitigate economic shocks and enhance cluster resilience.
- **Innovation and Entrepreneurship Promotion**: Encouraging collaboration within clusters can foster innovation, knowledge spillovers, and new business formation.

Conclusion

This research underscores the transformative potential of machine learning in economic analysis, providing a real-time, data-driven approach to understanding industrial clusters. By integrating advanced computational techniques with economic theory, this study offers a novel framework for policymakers and researchers to analyze, monitor, and support regional economic development in Québec. Moving forward, further refinement of ML techniques and expansion of real-time data integration can enhance the precision and applicability of economic cluster analysis.

GVC*dt*LAB
an initiative by CIRANO

Suite 1400
1130 Sherbrooke West
Montréal, QC, Canada H3A 2M8
www.gvcdtlab.com

Learn more about
our project