**Assessing Google Trends Data**

**Executive summary**

The information on economic indicators is crucial for policymaking and taking decision at right time but this information is usually available with a lag. So, the need for alternative data sources is of growing importance for both supplementing Statistics Canada’s data holdings and for nowcasting economic activity. The goal of this project is to develop a methodology to predict macroeconomic factors such as Gross Domestic Product (GDP), retail trade sales and retail e-commerce sales in real time by using the real time data source, Google Trends. Google Trends provides daily, weekly, and monthly reports on the volume of Google queries related to different industries which can help to understand the business cycles and provide signals about multiple aspects of the economy that can further be used to estimate the economic factors in real time. The nowcasting of economic indicators will provide more timely information for policymaking.

1. **Introduction**

Macroeconomic factors are the key drivers of economy and timely information on macroeconomic factors helps in good policymaking. However, this information is available with a lag, for instance, the data for the present month’s GDP is generally published in the coming month/quarter which causes delay in decision-making. To overcome this issue of delayed information gave rise to nowcasting approach and this approach is fetching the interest of economists and researchers nowadays as this modern approach provides the information on economic indicators in real-time. Traditional macroeconomic indicators have some lag, and to fill this gap of information, Google Trends have been widely used as it may help in predicting the present [1]. The volume of queries on different industries may be correlated with the current level of economic activities in respective industry and may help to predict the subsequent data release [1].

Many researchers have used Google trends for nowcasting the economic activity. Google Trends provide information of business cycles and economic activities in economy and the salient features of these business cycles can be captured with few unknown factors using dynamic factor analysis models [2]. These models are applicable to high-dimensional data and can reduce the dimensionality of economic systems. DFM became the mainstream tool for nowcasting GDP growth over the time. Later on, new techniques emerged, and researchers have started to use machine learning algorithms for nowcasting economic factors. Woloszko [3] proposed a weekly tracker to estimate GDP in 46 Organisation for Economic Co-operation and Development (OECD) countries and G20 countries (excluding European Union). The proposed OECD tracker is based on a machine learning algorithm that estimates the relationship between Google Trends variables and GDP growth.

Dauphin et al. [4] have also used Google Trends data to estimate GDP growth, they provide comparative analysis of different nowcasting approaches such as Auto-Regressive models, Dynamic Factor Model (DFM) and some machine learning algorithms like Regularized Regression models, Random Forest, Support Vector Machine (SVM) and Neural Networks, and state that there is no one-size fits all model as different models are suitable for different datasets. Richardson et al. [5] used machine learning algorithms to nowcast GDP growth in New Zealand and their results show that machine learning algorithms boosted trees, SVM and neural networks outperformed the traditional autoregressive models for their study. The aforementioned studies indicate that traditional econometrics models and machine learning models both can be used for the nowcasting economic factors and success, and accuracy of the model may vary for different datasets. Therefore, a comparative study between traditional and modern machine learning algorithms may be more appropriate to fit a model on data in hand.

**References**

[1] H. Choi, H. Varian, Predicting the present with Google Trends, *Economic record*, *88 (2012)*, 2-9.

[2] Stock, J.H. and Watson, M.W., 2016. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. In *Handbook of macroeconomics* (Vol. 2, pp. 415-525). Elsevier.

[3] Woloszko, N. (2020). Tracking activity in real time with Google Trends, OECD Economics Department Working Papers, No. 1634, OECD Publishing, Paris.

[4] Dauphin, M.J.F., Dybczak, M.K., Maneely, M., Sanjani, M.T., Suphaphiphat, M.N., Wang, Y. and Zhang, H., 2022. *Nowcasting GDP-A Scalable Approach Using DFM, Machine Learning and Novel Data, Applied to European Economies*. International Monetary Fund.

[5] Richardson, A., van Florenstein Mulder, T. and Vehbi, T., 2021. Nowcasting GDP using machine-learning algorithms: A real-time assessment. *International Journal of Forecasting*, *37*(2), pp.941-948.

**INTRODUCTION**

**AIMS AND OBJECTIVES:**

**DATA SET:**

**DELIVERABLES AND SCHEDULE/ TIMELINE:**