

Presented By:

Aishwarya Sharma
Harpreet Kaur
Jagdeep Brar

IMAGE: <https://wallpapersafari.com/w/swTjG4>



Nowcasting Macroeconomic Indicators using Google Trends



THE
UNIVERSITY OF
BRITISH
COLUMBIA

Date: 31st May, 2022



OUTLINE



- 1 INTRODUCTION TO TEAM AND CLIENT
- 2 OVERVIEW OF PROJECT
- 3 DATA SET
- 4 METHODOLOGY
- 5 OBTAINED RESULTS
- 6 ROADBLOCKS
- 7 LEARNINGS AND FUTURE PLANS



Introduction To Team





AISHWARYA SHARMA

- Bachelors in Computer Science
- Former Software Developer in Hewlett Packard Enterprise (HPE)
- Passion for data and loves to play guitar in free time

A dark blue background featuring a complex network graph. The graph consists of numerous small, semi-transparent colored dots (yellow, orange, red, green) connected by thin white lines forming a web-like structure. This pattern repeats across the slide, creating a sense of depth and connectivity.

JAGDEEP BRAR

- Ph.D. in Applied Mathematics
- Worked on Credit Risk Management Projects as a Postdoc
- Likes model fitting and cooking!!



HARPREET KAUR

- Gold Medalist in Applied Mathematics
- Former Software engineer
- Love hiking, making and tinkering

Introduction To Clients

Statistics
Canada



A large, abstract graphic occupies the left side of the slide, featuring a complex network of interconnected nodes (represented by small circles) and connecting lines (represented by thin blue lines). The background is a dark blue gradient, and the nodes are colored in various shades of blue, white, yellow, and orange, creating a sense of depth and connectivity.

NICK NEWSTEAD

- Economist, Sociologist
- Data Exploration and Integration
- Strong Data Science Background



MARINA SMAILES

- Senior Analyst
- Data Exploration and Integration Lab
- Microdata, Demographic Analysis

OVERVIEW OF PROJECT

- ▶ Economic indicators are crucial indicators for a country's policy and decision making but information contains a lag.
- ▶ Aim to develop a model for Statistics Canada to nowcast macroeconomic factors in real time using Google search trends.

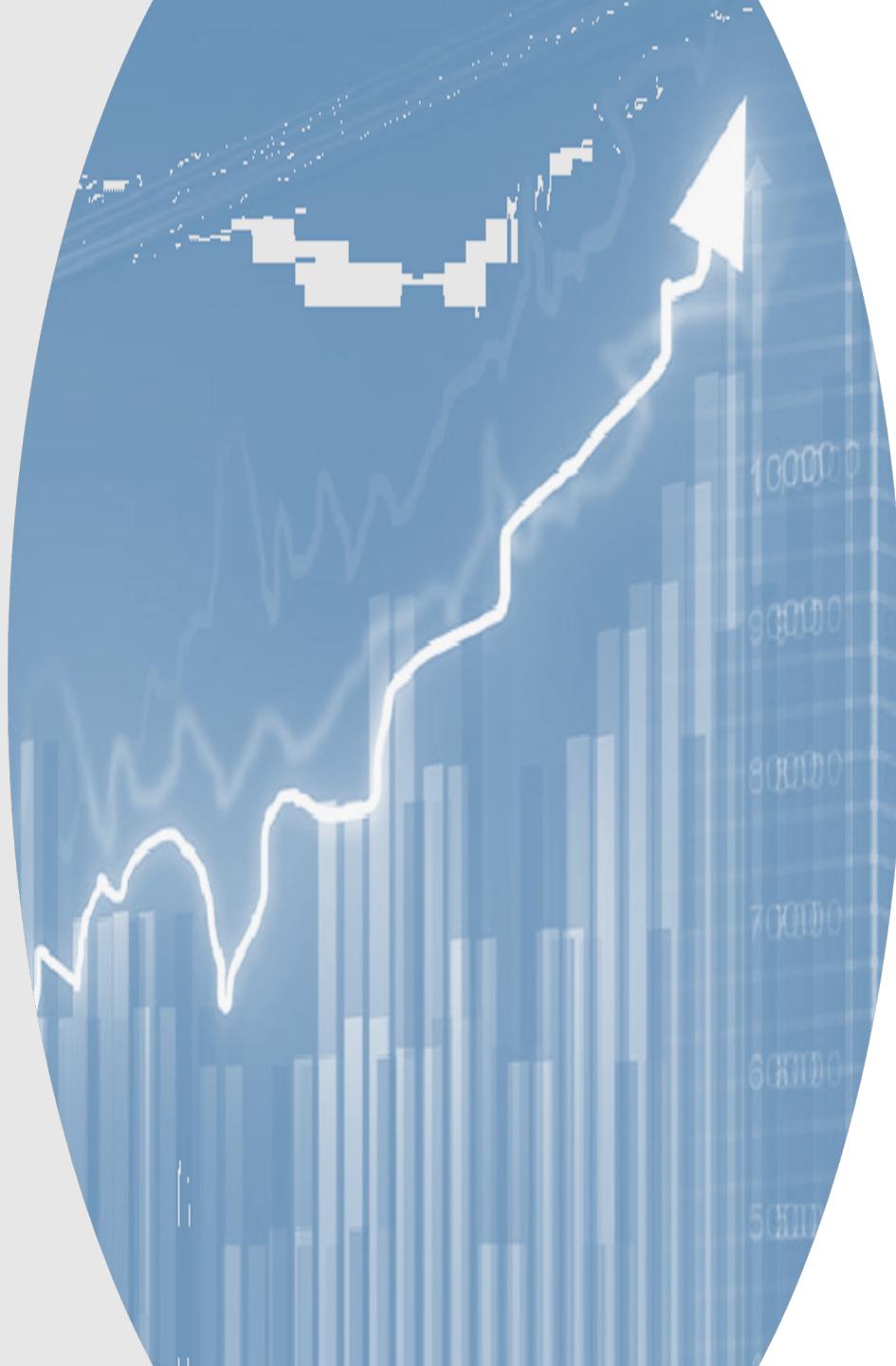
Economic indicators: Retail trade sales, E-commerce sales and Gross Domestic Product

- ▶ Use related queries, topics, keywords of different industrial categories to capture business cycles for economic factors.

RESEARCH QUESTIONS

- ▶ Nowcast quarterly Gross Domestic Product (GDP)
- ▶ Nowcast monthly retail trade sales
- ▶ Nowcast monthly retail E-commerce sales

Nowcast: “predicting the present”



DATASET

- ▶ Statistics Canada's historical time series data for the indicators.
- ▶ Google Trends extract in real time for different macroeconomic indicators as per the related queries, topics, and keywords.
- ▶ Google Trends data available from the year 2004 onwards.



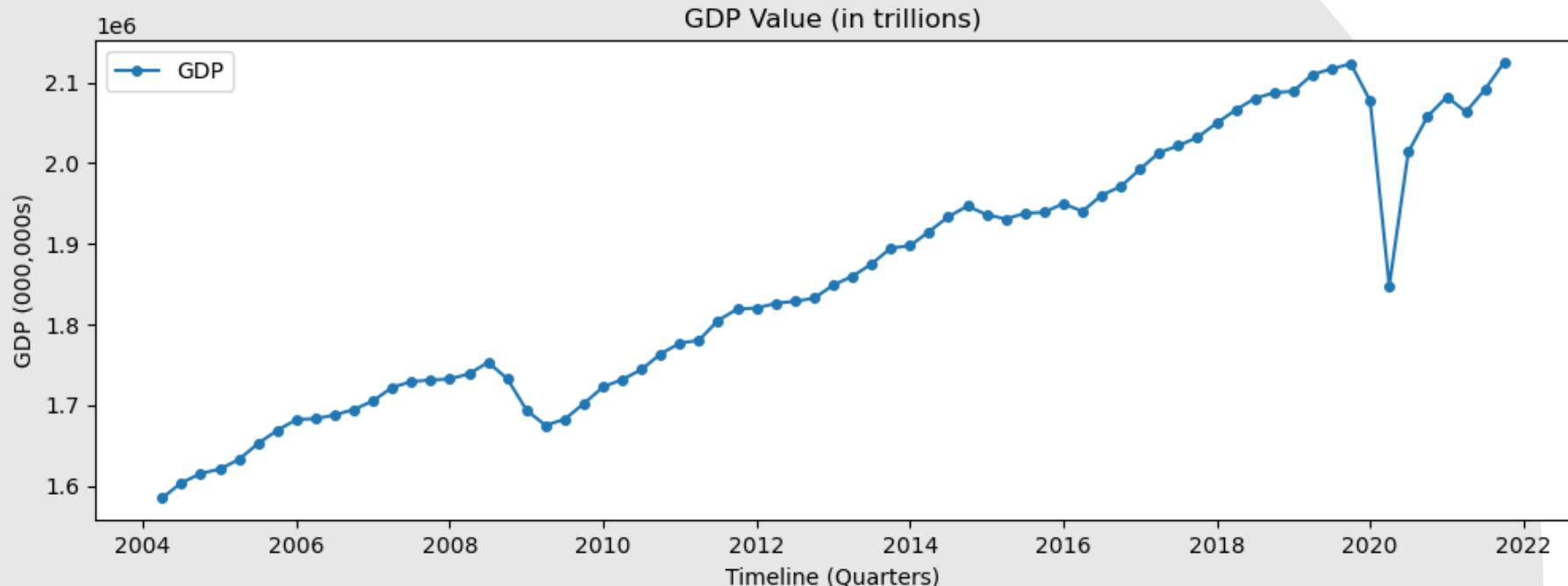
Statistics Canada Data

Economic Indicator	Timeline	Frequency
Retail trade sales	1991-2022	Monthly
E-commerce sales	2016-2022	Monthly
GDP	1997-2021	Quarterly

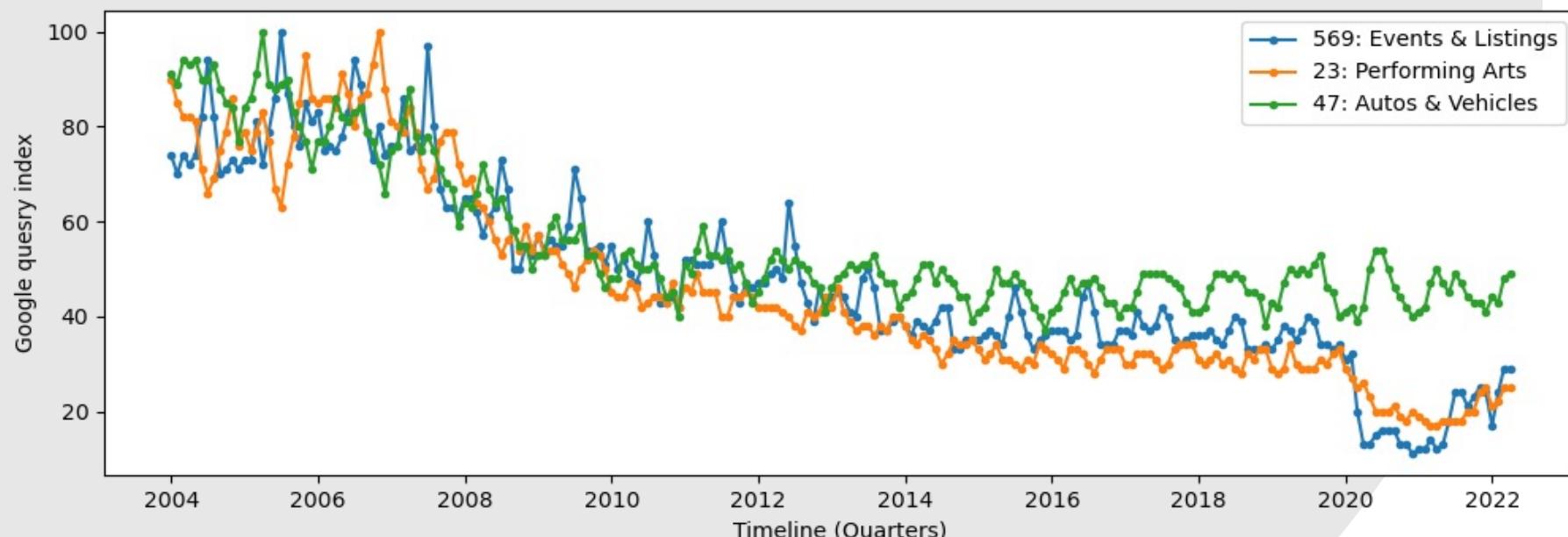
Selected Data for Analysis

Economic Indicator	Timeline	Frequency	Num. of observations	Num. of predictors (trends)
Retail trade sales	2004-2022	Monthly	217	396
E-commerce sales	2016-2022	Monthly	74	31
GDP	2004-2022	Quarterly	72	446

GDP time series



Google Trends time series





METHODOLOGY



STATIONARY
TIME SERIES

MODELLING

ECONOMETRIC
MODELS

MACHINE
LEARNING
MODELS

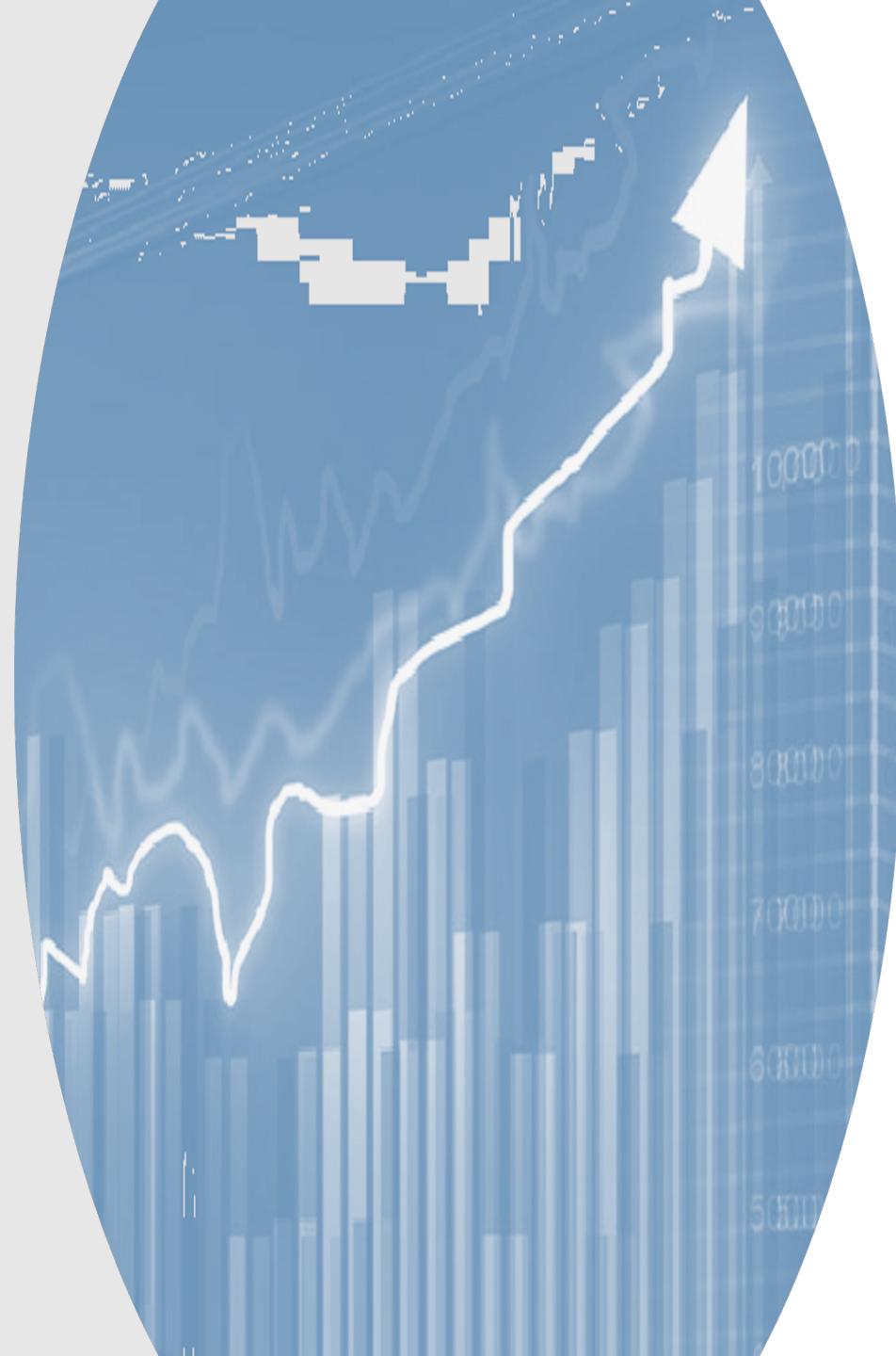


Time Series Stationary

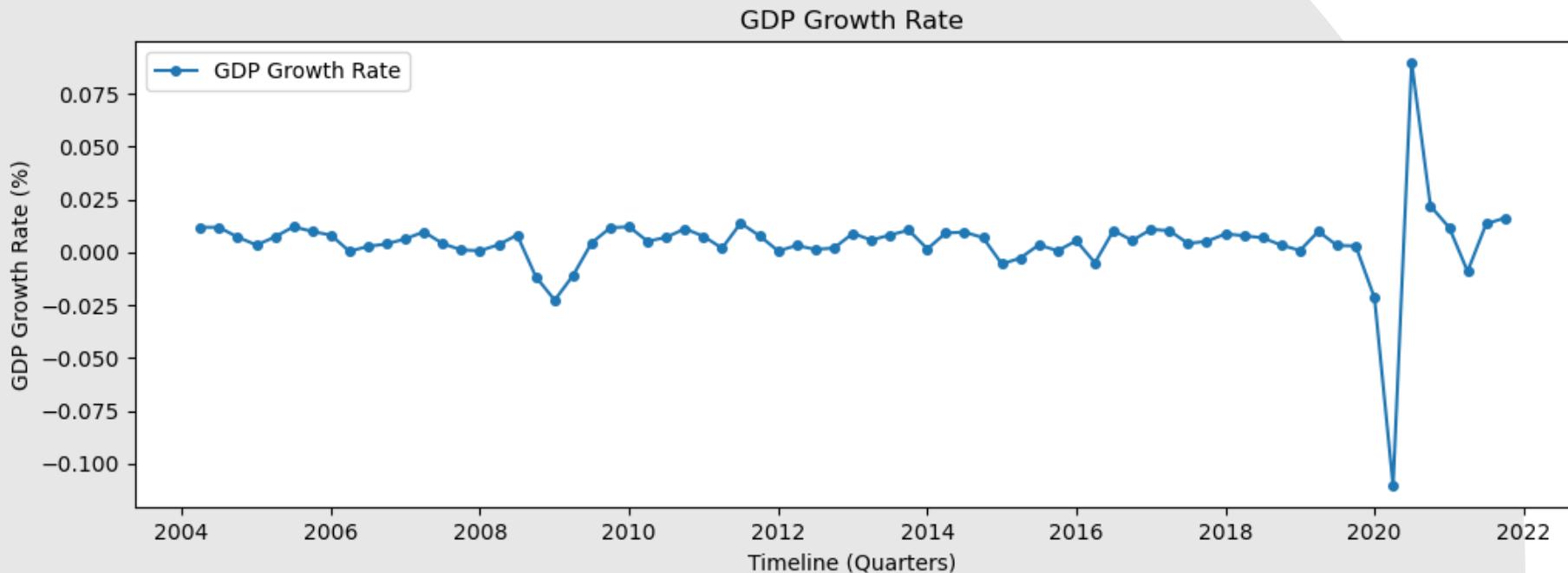
Normalized Data

Removed Trend

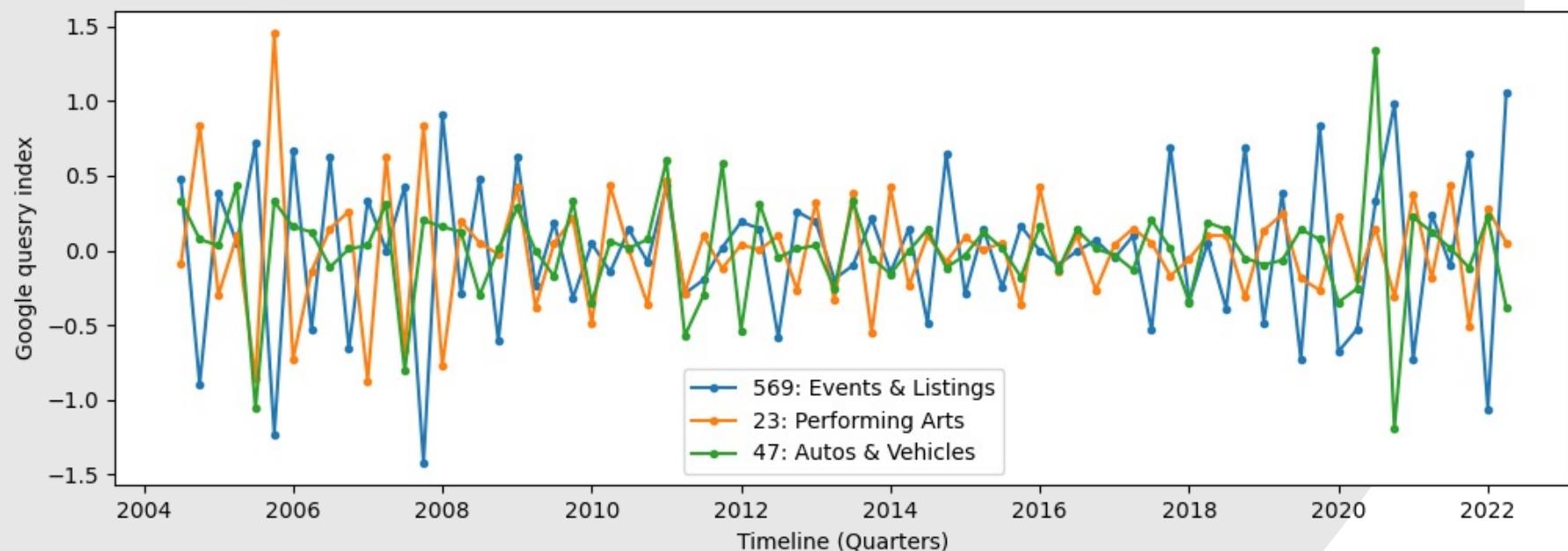
Removed Seasonality



Stationary time
series (GDP
Growth rate)



Stationary
Google Trends
time series



ECONOMETRIC MODELS

1

Dynamic Factor Model (DFM)

- ▶ Best suited model for time series when predictors are more than samples
- ▶ Reduces dimensionality of time series data
- ▶ Resultant factors work as efficient predictors for ARIMA model



②

Autoregressive Integrated Moving Average (ARIMA)

- ▶ Widely used model for time series data
- ▶ Lag selection for AR and MA order by using PACF and ACF plots
- ▶ Used factors of Google Trends data as predictors
- ▶ Implemented one-step ahead rolling prediction



MACHINE LEARNING MODELS

1

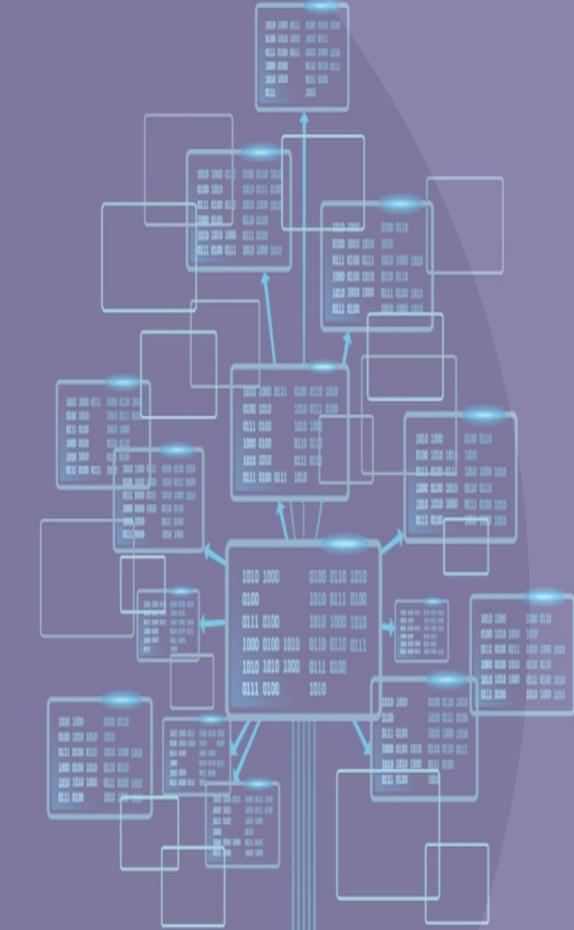
LASSO

- ▶ Fitted model on growth rate for three economic indicators
- ▶ Rolling predictions to predict growth rate
- ▶ Calculated original values by using predicted growth rate
- ▶ Used cross validation to tune parameter



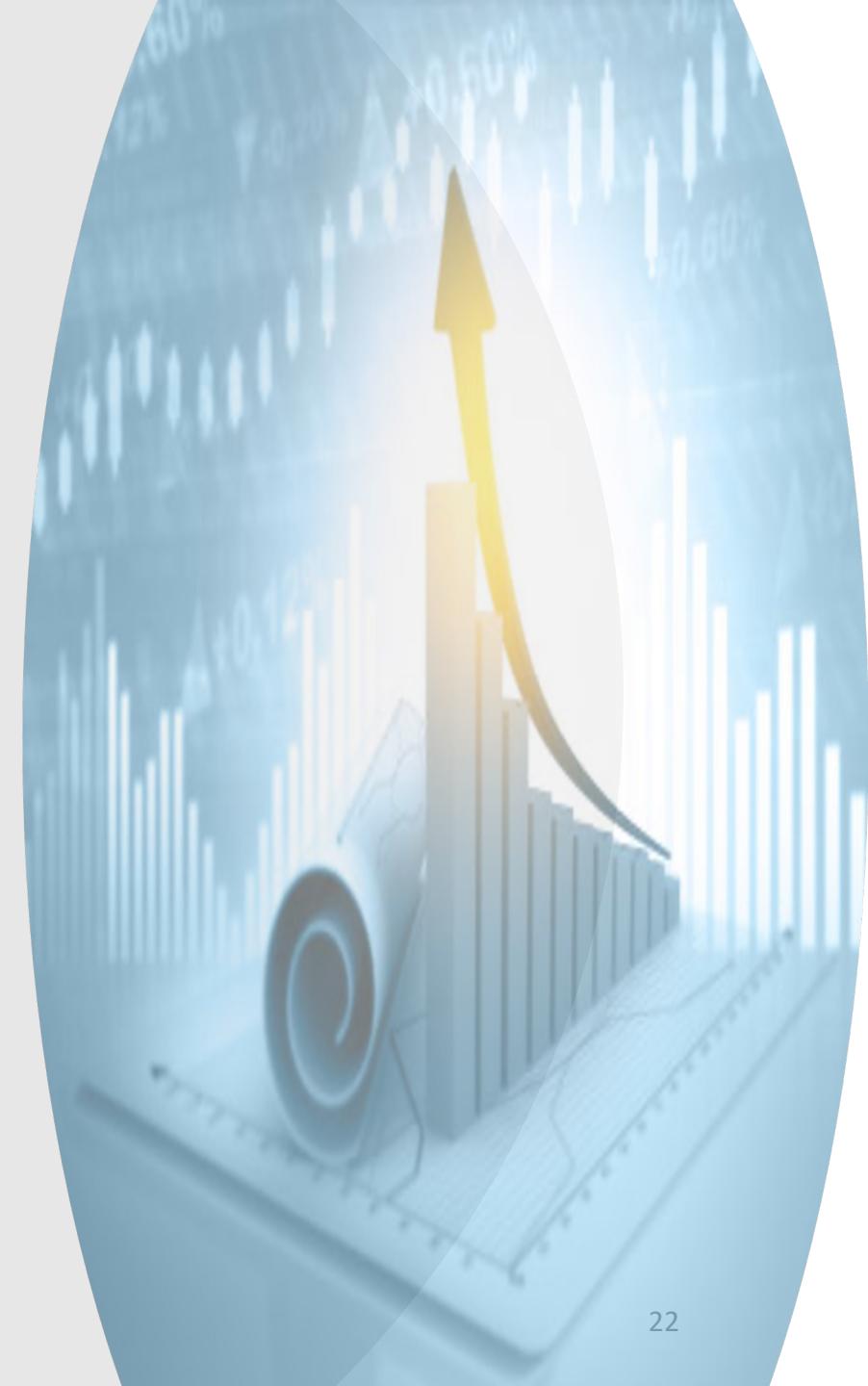
Random Forest

- ▶ Applied PCA on Google Trends data to get components
- ▶ Fitted model on components to predict growth rate
- ▶ Calculated original values using predicted growth rate
- ▶ Cross validation to get appropriate number of trees



XG Boost

- ▶ Applied PCA on Google Trends to get components
- ▶ Fitted model on components to predict growth rate for three economic indicators
- ▶ Calculated original values by using predicted growth rate
- ▶ Cross validation to tune number of trees and penalty parameter (*in progress*)



OBTAINED RESULTS

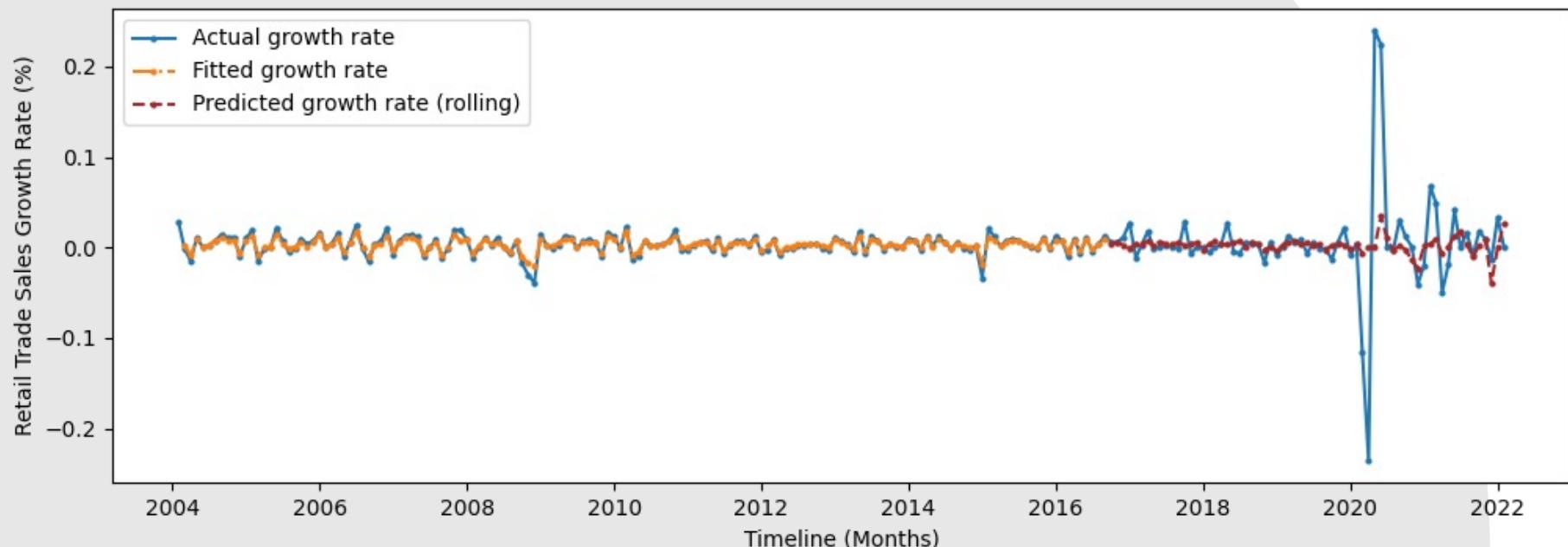


Retail Trade Sales

- ▶ **Response:** Monthly retail trade sales growth rate
- ▶ **Predictors:** Categories & Keywords

Method	Prediction Error (RMSE)	Parameter Tuning
DFM + ARIMA	2,828,358	Number of factors
LASSO	2,379,342	Penalty parameter
PCA + Random Forest	2,281,435	Number of trees
PCA + XGBoost	3,410,734	<i>In progress</i>

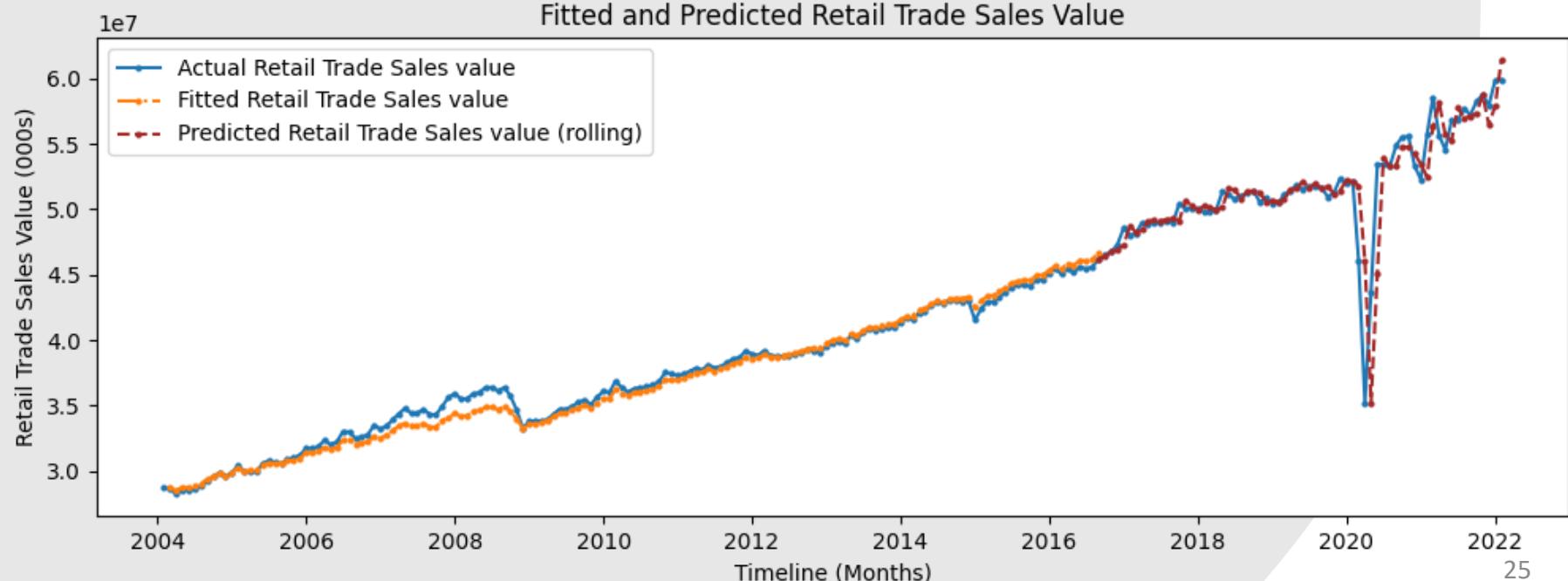
Fitted and Predicted Retail Trade Sales Growth Rate



Retail Trade Sales
Growth Rate
Prediction

Retail Trade Sales
Value Prediction

Fitted and Predicted Retail Trade Sales Value

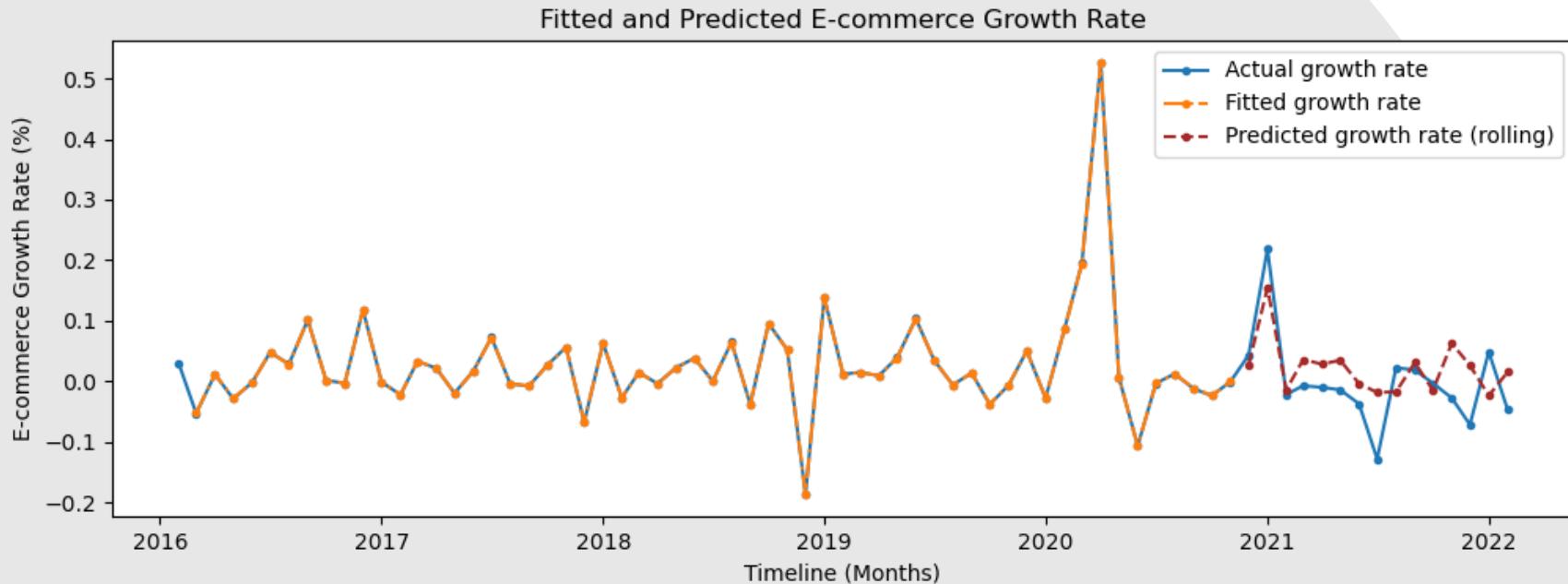


E-Commerce Retail Trade Sales

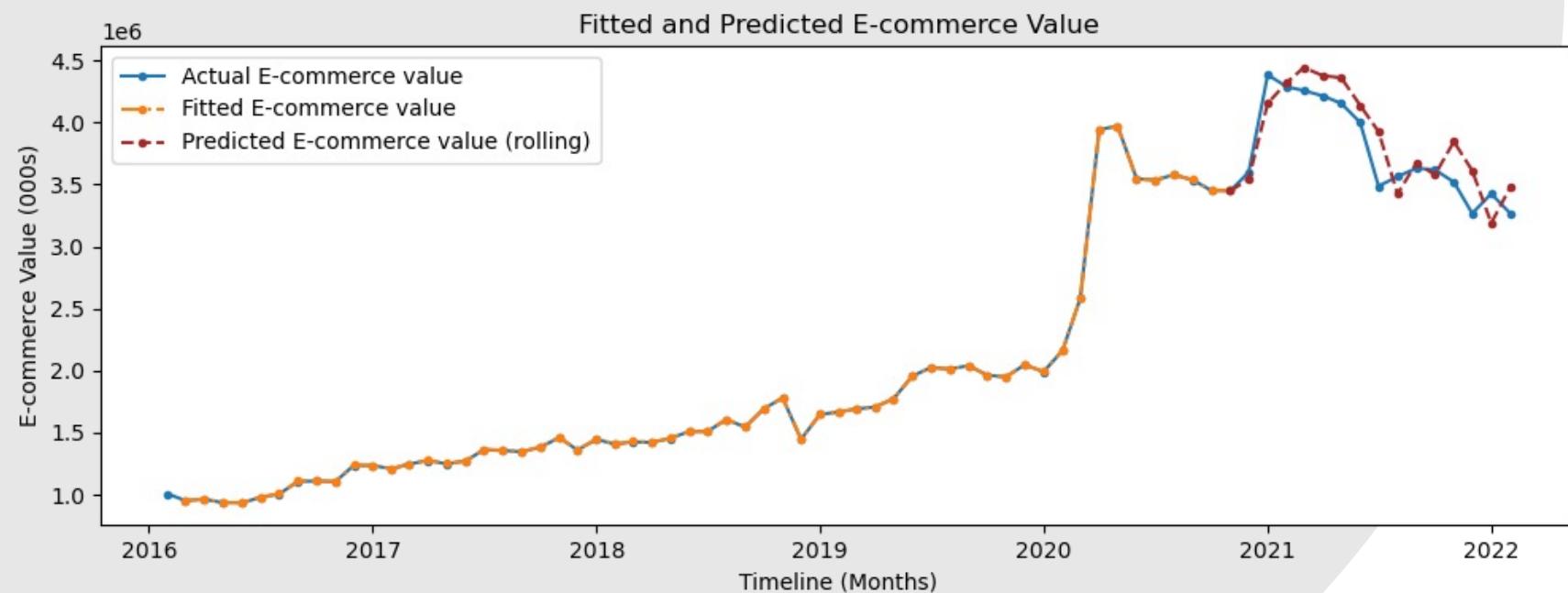
- ▶ **Response:** Monthly e-commerce retail trade sales growth rate
- ▶ **Predictors:** Category & Keywords

Method	Prediction Error (RMSE)	Parameter Tuning
ARIMA	390,077	--
LASSO	246,766	Penalty parameter
Random Forest	260,128	Number of trees
XGBoost	212,289	<i>In progress</i>

E-commerce Growth Rate Prediction



E-commerce Value Prediction

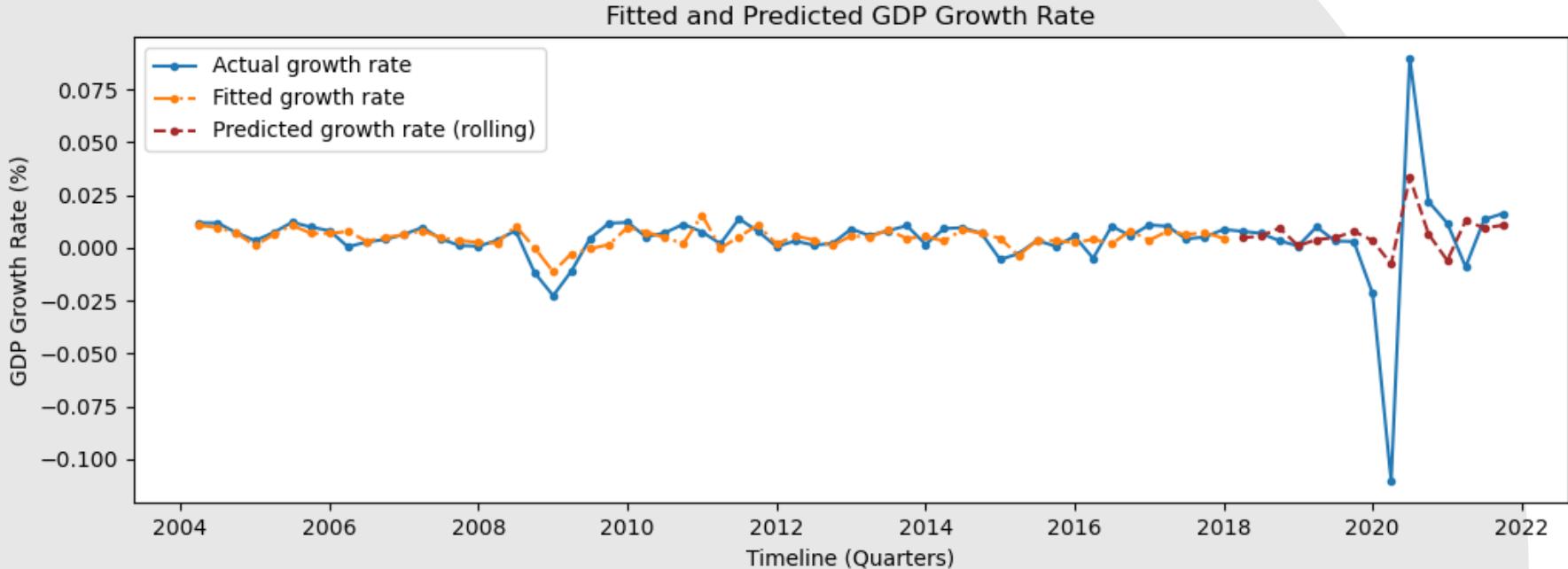


Gross Domestic Product (GDP)

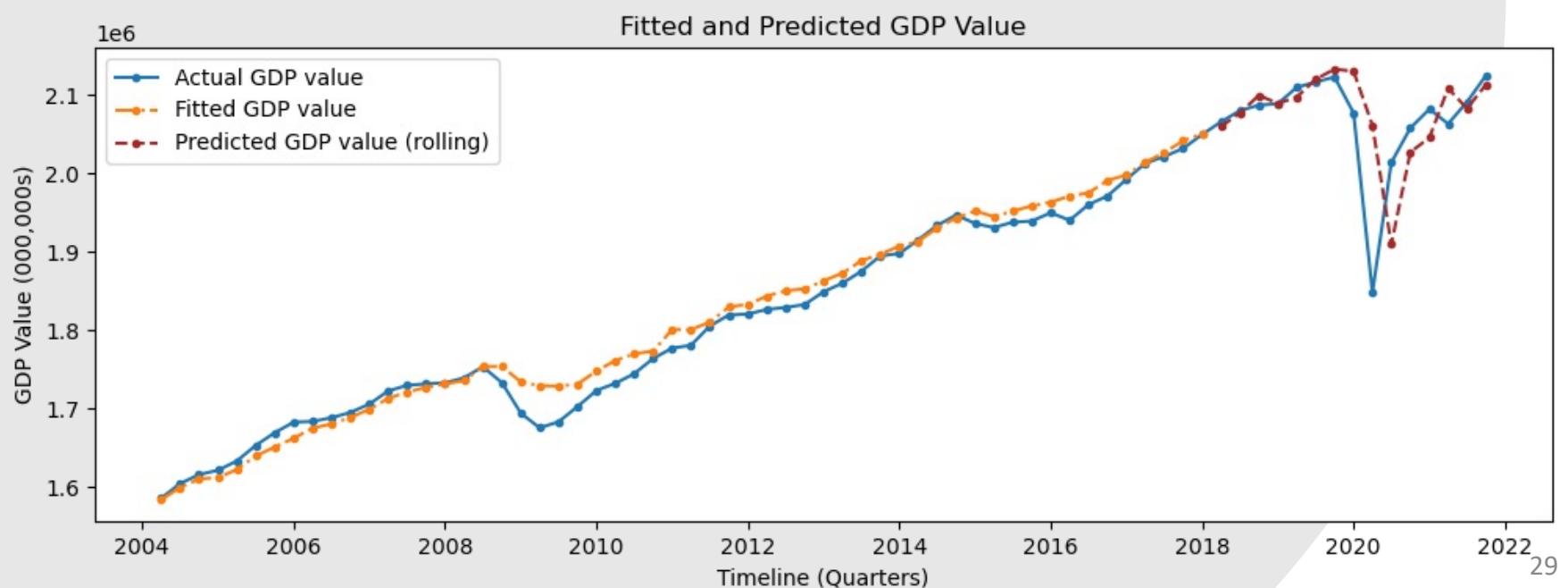
- ▶ **Response:** Quarterly GDP growth rate
- ▶ **Predictors:** Categories & Keywords

Method	Prediction Error (RMSE)	Parameter Tuning
DFM + ARIMA	65,511	Number of factors
LASSO	84,146	Penalty parameter
PCA + Random Forest	78,651	Number of trees
PCA + XGBoost	83,641	<i>In progress</i>

GDP
Growth Rate
Prediction

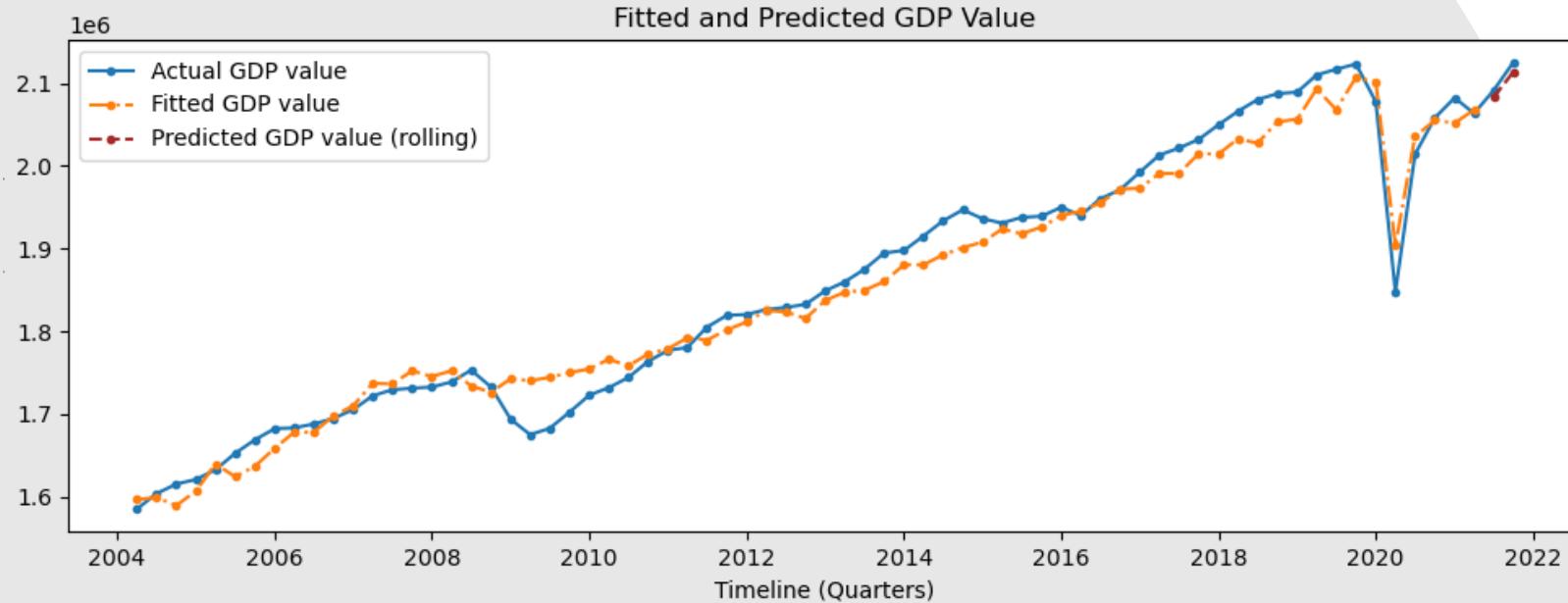


GDP Value
Prediction

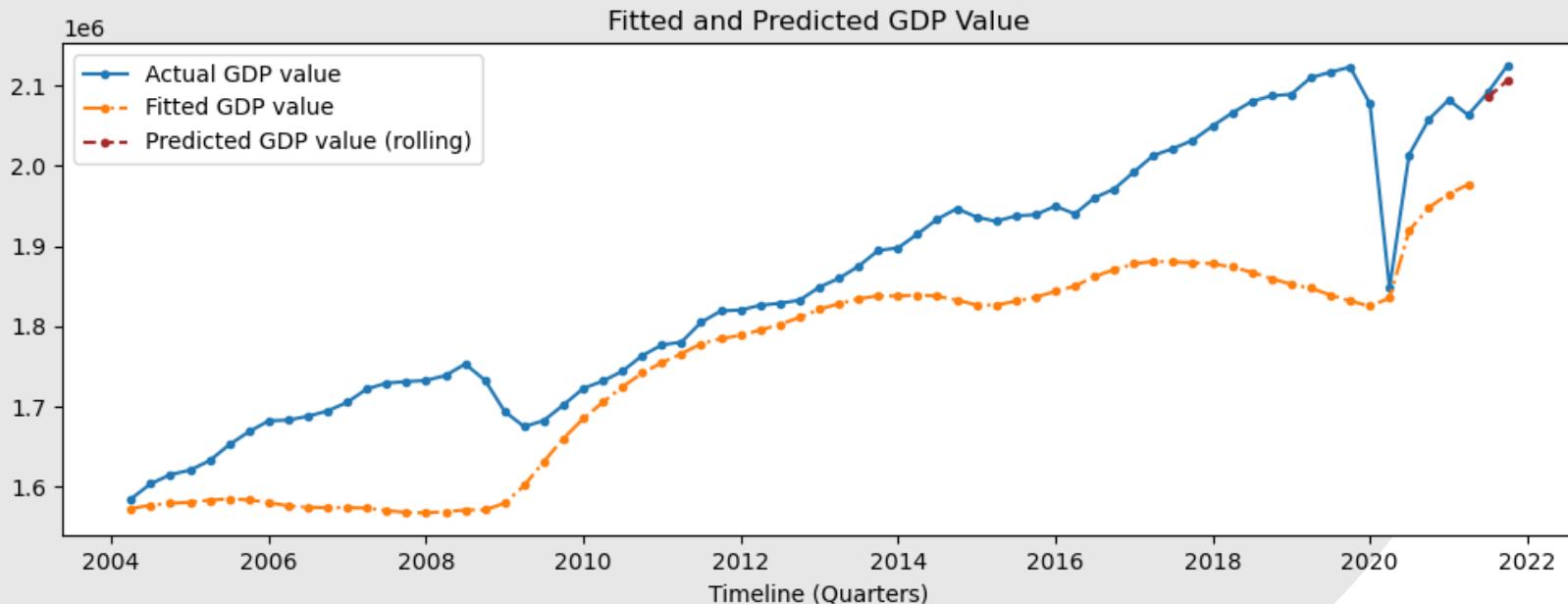


Does Google Trends really help?

Model fit using
Google Trends



Model fit without
using Google Trends

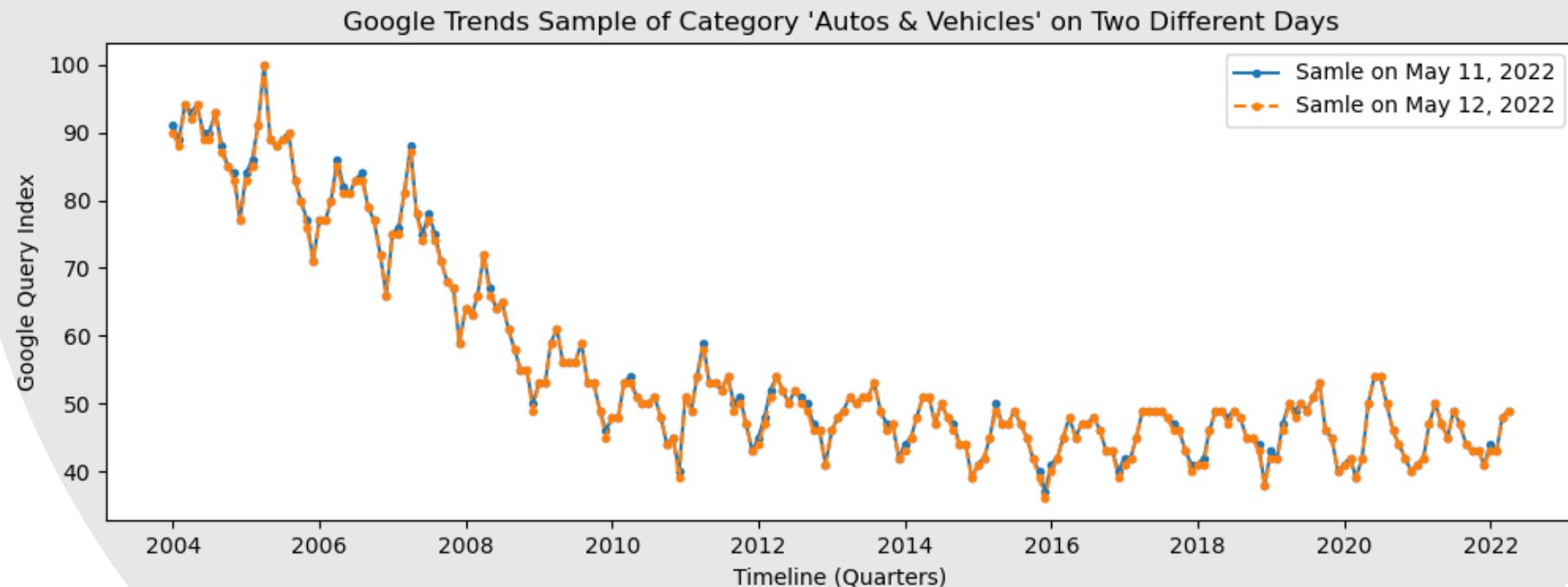


Roadblock

► API Calls on different days to extract trends data

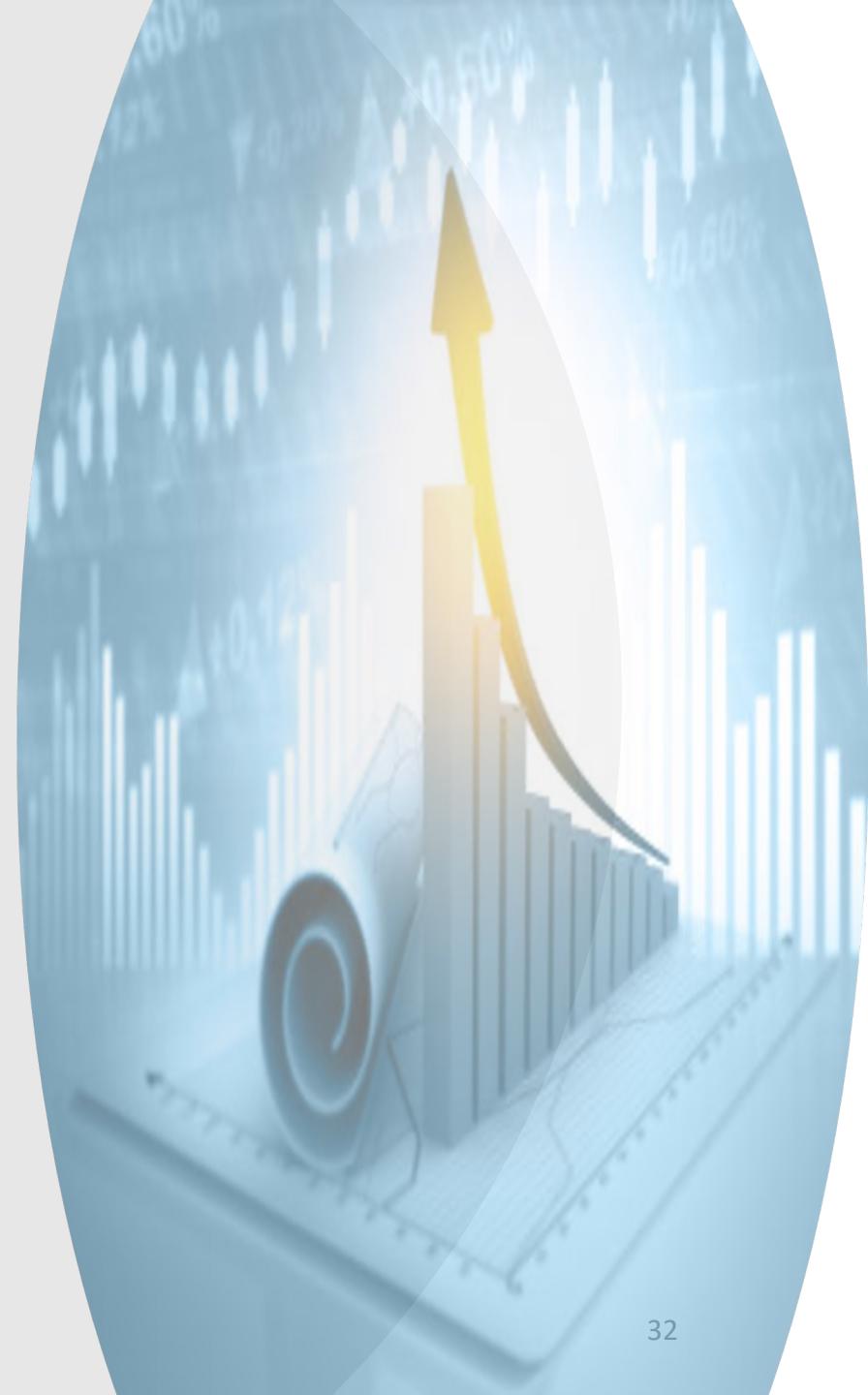
Issue: Many API calls, Require time

Solution: Move forward with one sample



Learnings

- ▶ Use of Google Trends/pytrends for nowcasting
- ▶ DFM for dimensionality reduction of time series data
- ▶ Machine learning models for time series analysis



Future Plan

- ▶ Cross Validation of XGBoost
- ▶ Bootstrap to get prediction interval
- ▶ Visualizations and dashboard
- ▶ A detailed report



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.
6. Logo source: [UBC Logo](#), [Statistics Canada](#), [Google Trends](#)



THANK YOU !!

QUESTIONS