

# Time Series Analysis for Chicago's Trade Data

MSCA 31006 IP03 Time Series Analysis and Forecasting

2023, Spring

Group 5

Mia Song / Minh Vo / Rolamjaya Hotmartua / Soohyun Iris Lee / Xiao Pang

# Table of Contents

1

**Goal of Analysis**

2

**Data Description**

3

**Experimental Analysis**

4

**Model Exploration**

5

**Model Selection**

6

**Conclusion and  
Future Work**

# Goal of Analysis

The goal of prediction for Chicago's import data is to forecast future import trends accurately, enabling businesses / policymakers to make informed decisions, optimize operations, & drive economic growth.



**Accurately predict monthly import trends in Chicago**



**Pre-Covid vs. Post-Covid :**  
Does the model predict Pre-covid better?



**Pre-Covid vs. Post-Covid :**  
Which model performs better for each period?

*Compared ETS, SARIMA,  
Regression w/ ARMA errors, VAR Models*

# Business Opportunity

**Accurate import forecasting for Chicago informs** economic planning, trade analysis, infrastructure planning, market opportunities, policy-making, and risk management. It enables informed decisions, capitalizing on opportunities, and fostering sustainable economic growth.



## Economic Planning

Businesses/city officials can plan and allocate resources effectively, and attract investments

## Logistics Planning

Logistics capabilities to support the efficient movement of goods

## Policy & Regulatory

Develop & implement trade policies, regulations, and incentives that support economic growth, job creation, and sustainable development

# Data Description

1. **Dataset Size:** 254 monthly observations across 9 variables
2. **Target Variable:** 'Import' (in \$billion)
3. **Time Span:** January 2002 to February 2023
  - Pre-Covid: 2002~2018 -> forecast 2019
  - Post-Covid: 2002~2021 -> forecast 2022
4. **Predictor Variables:** Supplemented by the following economic indices from the same time period, i.e.
  - Export (in \$billion)
  - CPI (Consumer Price Index)
  - PPI (Producer Price Index)
  - Bond (10-year Government Bond Yield)
  - Sentiment (Consumer Sentiment Indicator from the University of Michigan)
  - USDX (US Dollar Index)
  - Uncertainty (Economic Policy Uncertainty Index for the United States)
5. **Source of Data:** United States Census Bureau (Trade Data) and The Fed St. Louis (Other Economic Data)

# Experimental Analysis: Stationarity

## 1. What Is Stationarity?

- Stationarity in time series refers to a property where a series' statistical properties (the mean, variance, and autocorrelation structure) remain constant over time. (= consistent patterns)

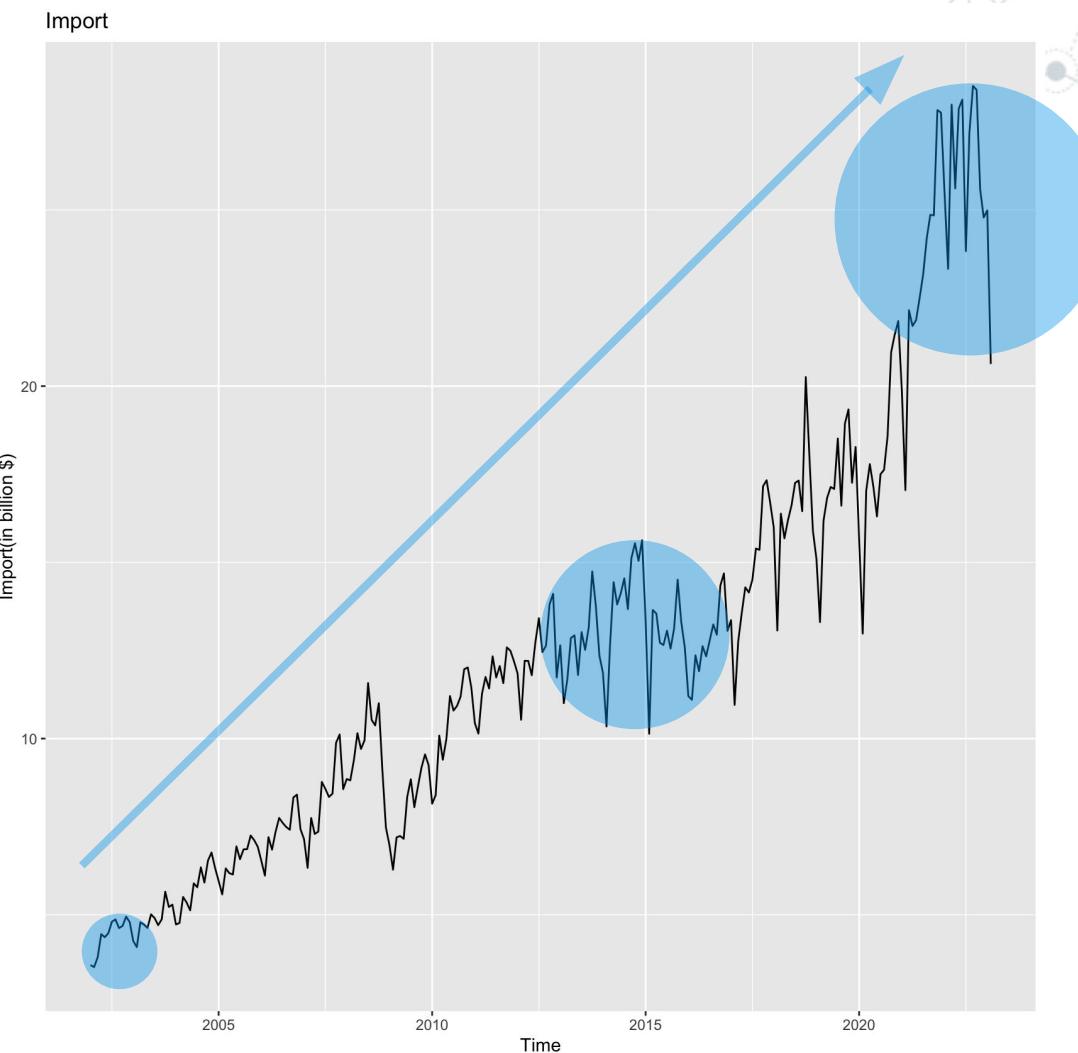
## 2. Why Is Stationarity Important?

- Easy to model
- Accurate forecasting / Avoids misleading results → Meaningful analysis
- Validates assumptions of many time series models such as ARIMA

# Experimental Analysis: Stationarity

Does the Original Data Satisfy the Stationarity Assumption? **No.**

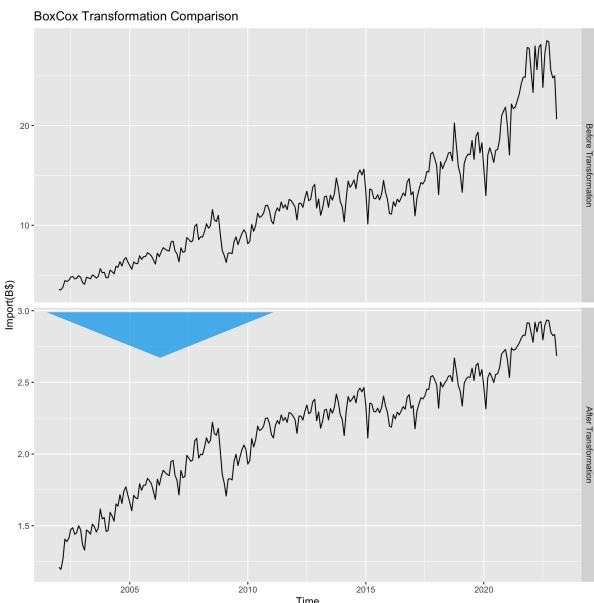
- Upon visual inspection, it is clear that the original data does not satisfy the stationarity assumption due to an observable **1) increasing trend** and **2) inconsistent variance**.
- The KPSS that assesses the stationarity of a time series confirms the non-stationarity of the data due to ***the p-value of 0.01(<0.05)***.



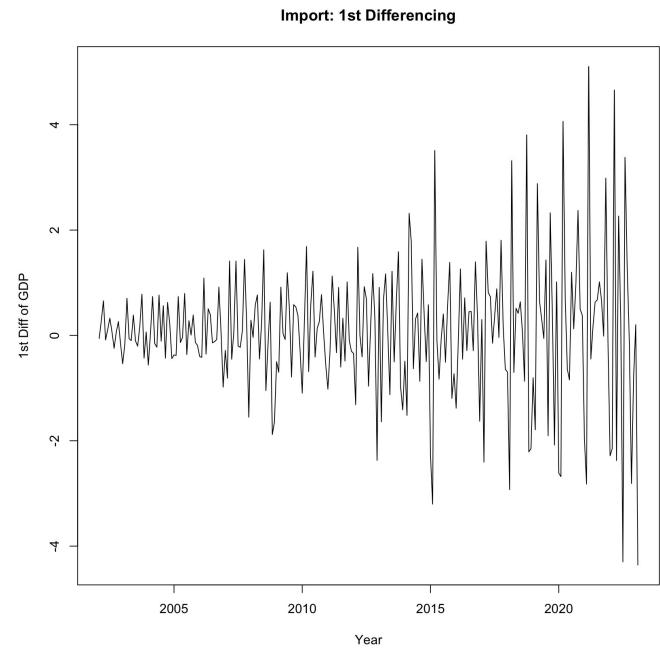
# Experimental Analysis: Stationarity

- We'll address the non-stationarity in our data through two key steps:
  1. **Box-Cox Transformation** to stabilize any varying variance.
  2. **1st order differencing** to mitigate upward trends.

Box-Cox Transformation



1st Order Differencing

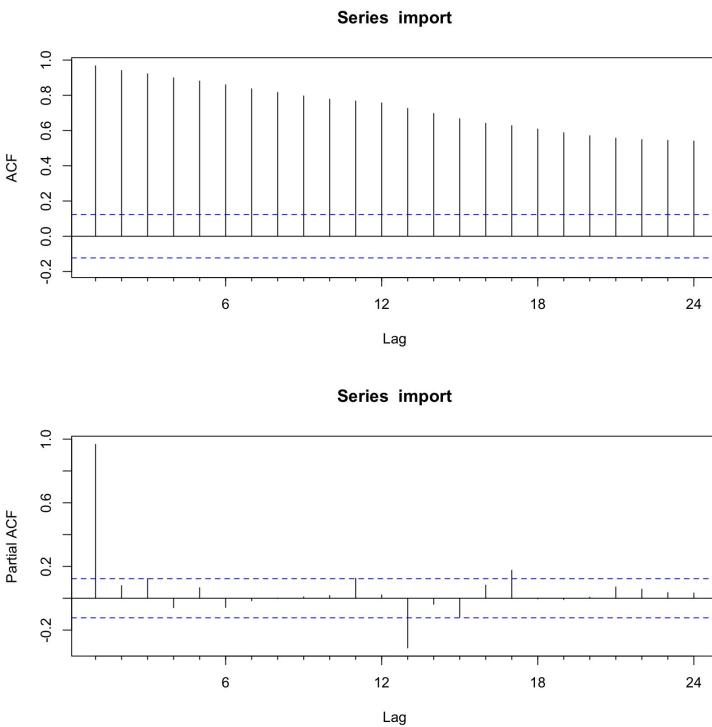


Applying the Box-Cox transformation (with  $\lambda=-0.081$ ) stabilized our time series. However, the first-order differencing applied to the original data was insufficient to fully eliminate the trend. → **Combine two methods to achieve stationarity.**

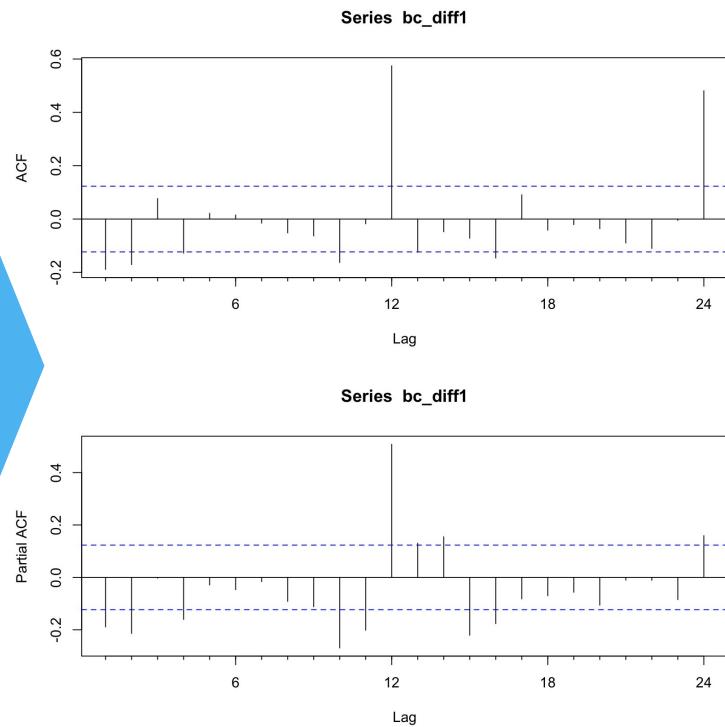
# Experimental Analysis: Stationarity

## ACF and PACF

Original Data

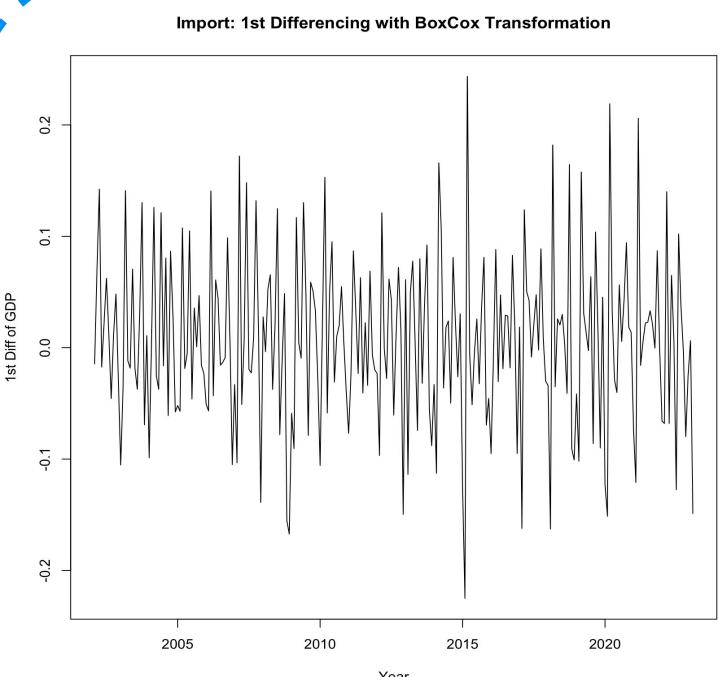


BoxCox Trans. + 1<sup>st</sup> order Differencing



Data after

Box-Cox Trans + 1<sup>st</sup> order Differencing

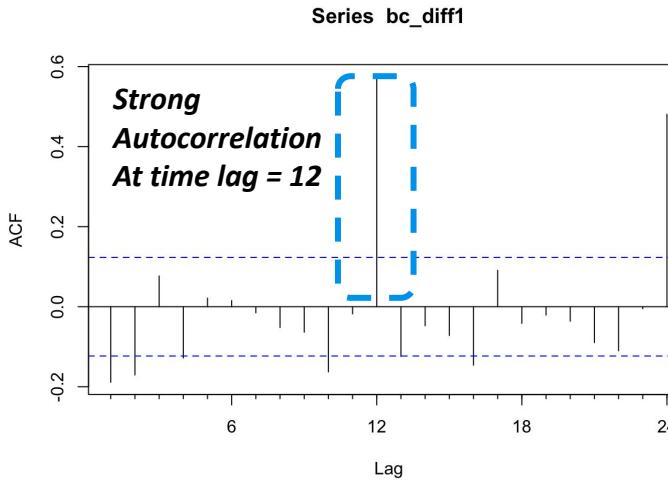


This data will form the foundation  
for our upcoming analysis.

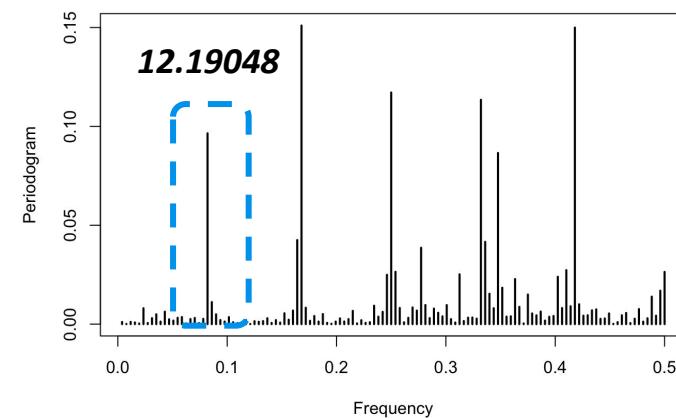
# Experimental Analysis: Seasonality

We have confirmed the presence of seasonality with a lag of 12 in our time series data.

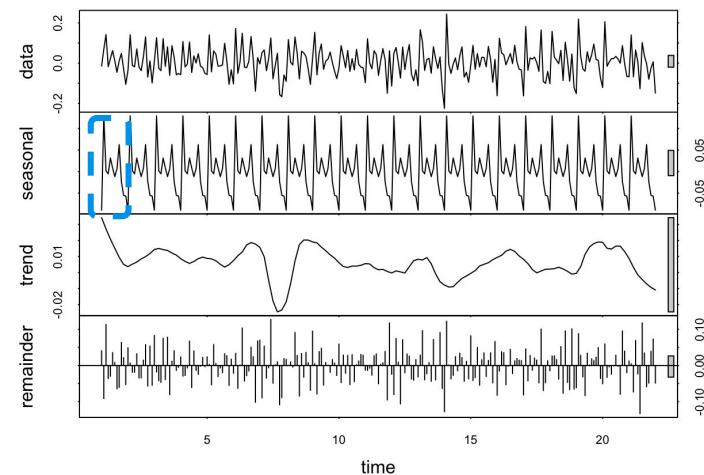
## ACF Plot



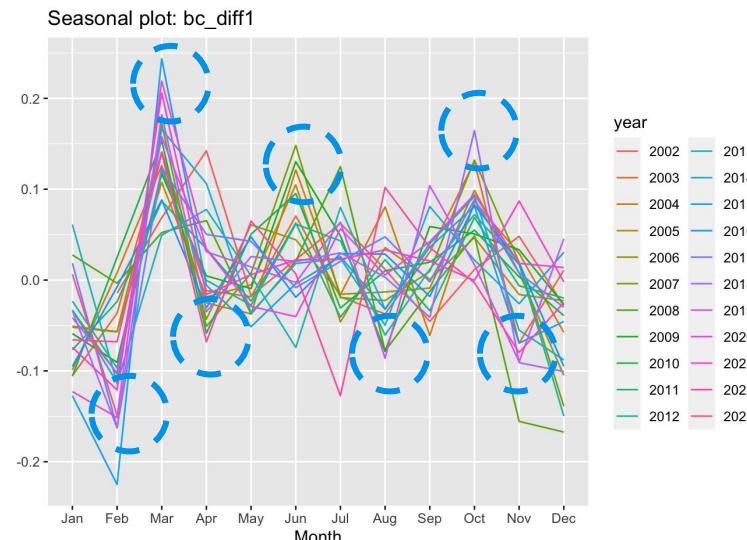
## Periodogram



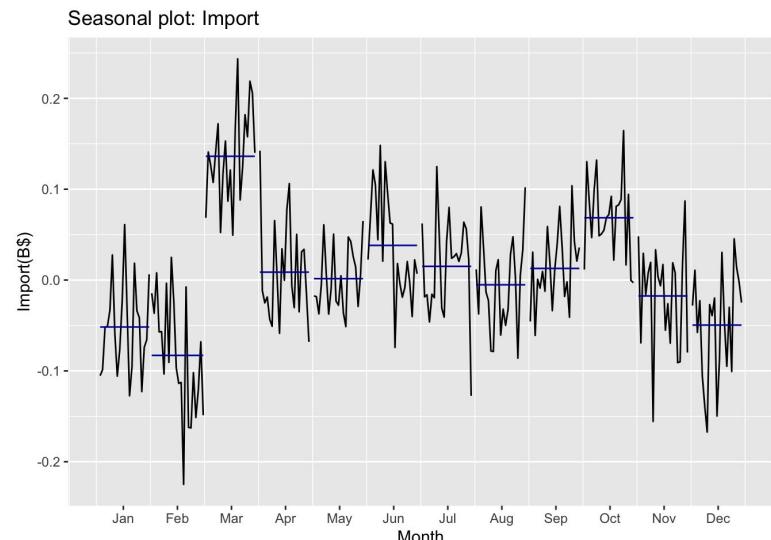
## STL



## Seasonal Plot



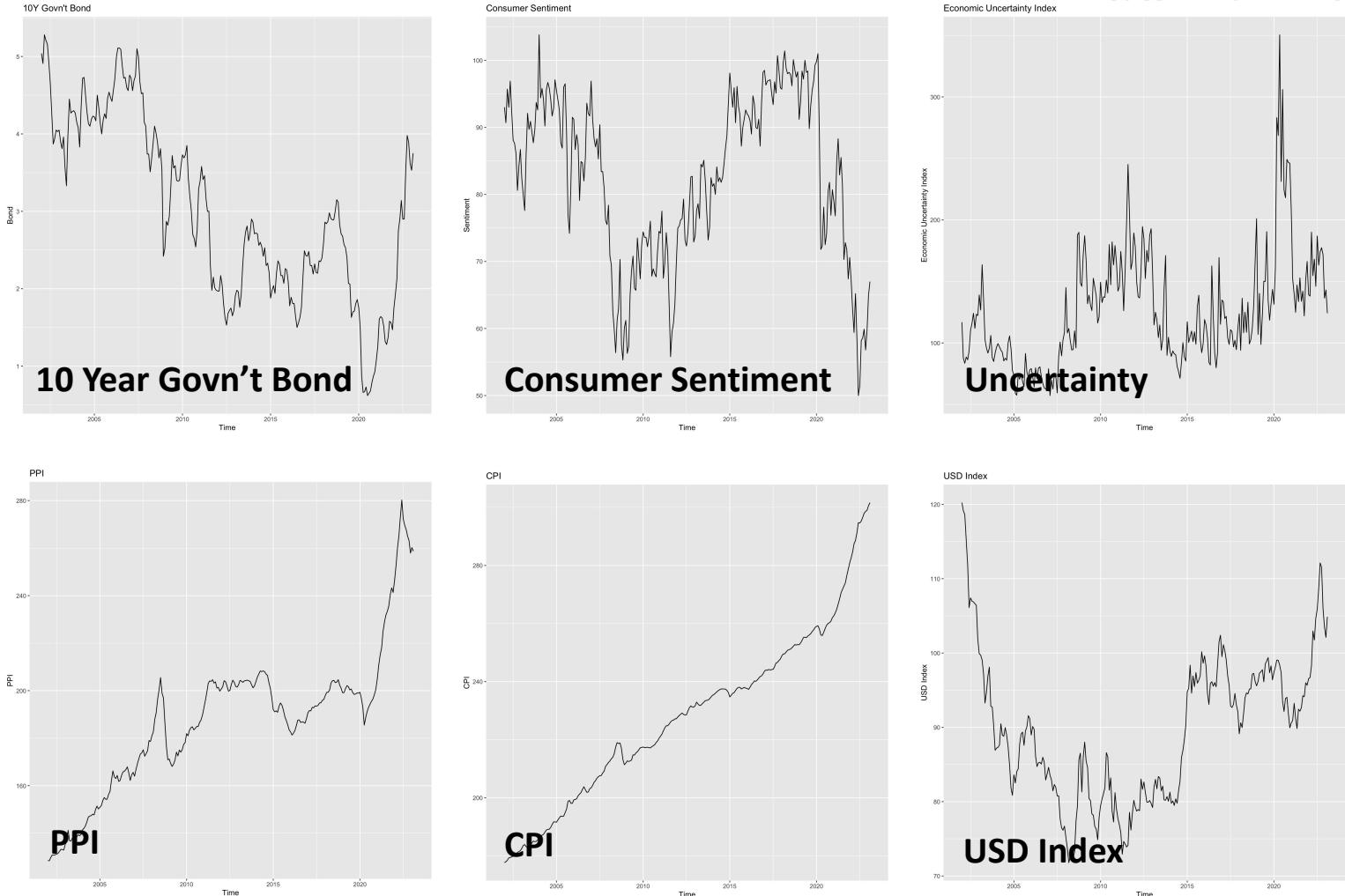
## Subseries Plot (Seasonal Plot)



Seasonal Patterns throughout a year  
1) high in Mar, Jun, Oct 2) low in Feb, Apr, Aug, and Nov

# Experimental Analysis: Correlation

- In our model exploration, we'll be employing advanced techniques, such as Regression with ARIMA error modeling and VAR and VARIMA for handling multivariate time series.
- This necessitates examining time series plots for each variable and assessing their correlations.

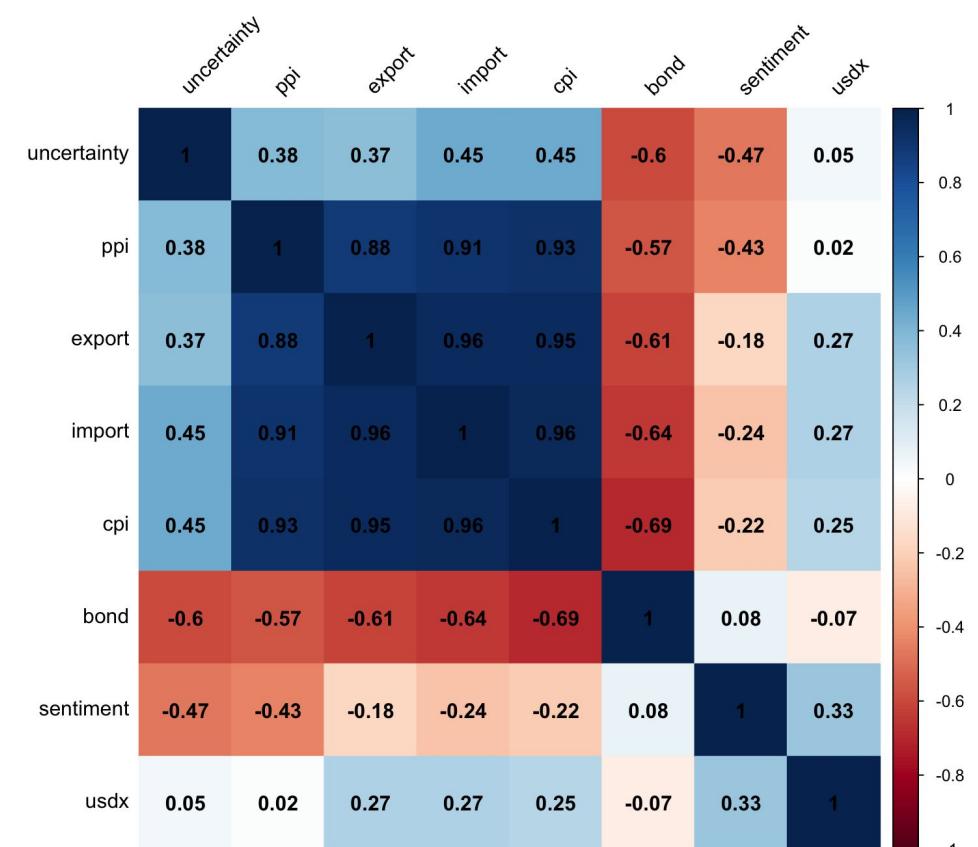


# Experimental Analysis: Correlation

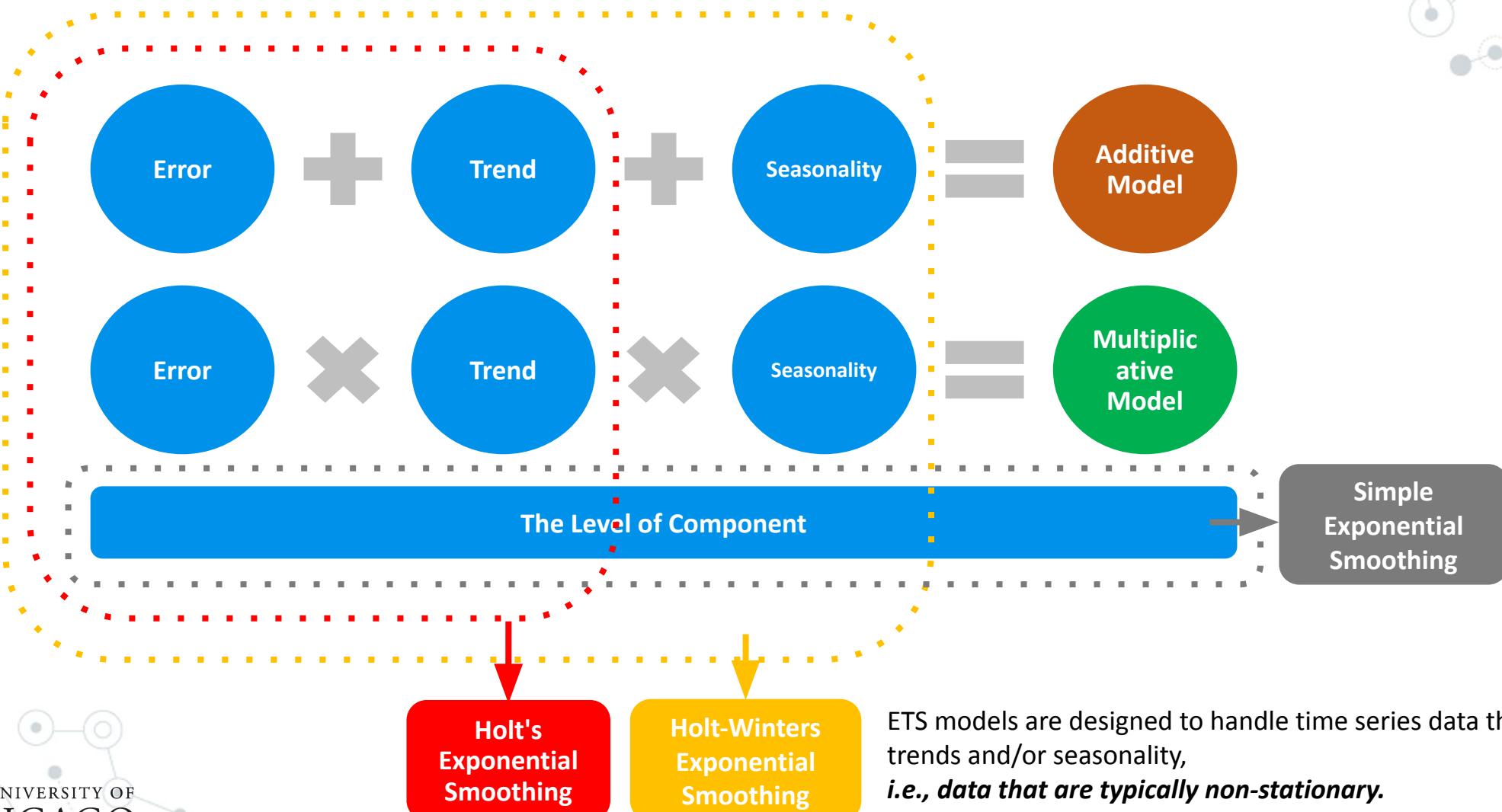
## The Definition of Variables

Variables	Definition
Import & Export	<ul style="list-style-type: none"> <li>The value of goods and services that the US buys from/sells to other countries</li> </ul>
10-Year Govn't Bond	<ul style="list-style-type: none"> <li>A debt security issued by the US government in 10 years</li> </ul>
Consumer Sentiment	<ul style="list-style-type: none"> <li>How people feel about the economy and their financial situation</li> <li>An indicator of consumer spending, essential for economic growth</li> </ul>
Economic Uncertainty	<ul style="list-style-type: none"> <li>The unpredictability of the future state of the economy</li> </ul>
PPI (Producer Price Index)	<ul style="list-style-type: none"> <li>The average change over time in selling prices received by domestic producers for their output</li> <li>A leading indicator of consumer price inflation</li> </ul>
CPI (Consumer Price Index)	<ul style="list-style-type: none"> <li>The average change over time in prices paid by consumers for a market basket of consumer goods and services, a measure of inflation</li> </ul>
USD Index	<ul style="list-style-type: none"> <li>The value of the US dollar against a basket of foreign currencies</li> </ul>

## Correlation among Variables



# Model Exploration(1) ETS Overview



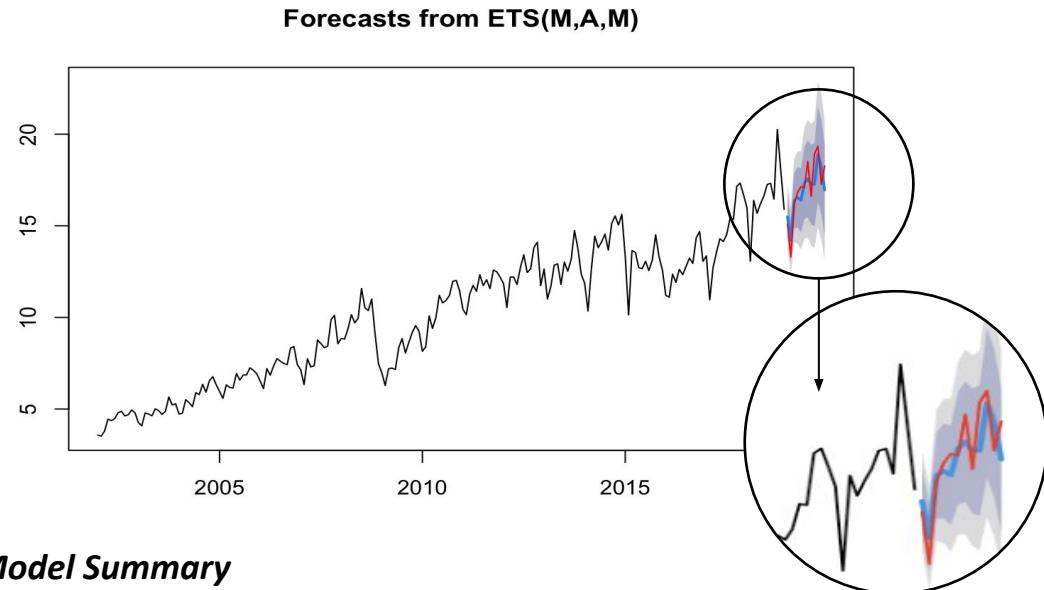
# Model Exploration(1) ETS Before COVID19

TS Models	E/T/S	( $\alpha$ , $\beta$ , $\gamma$ )	RMSE (train)	AICc	RMSE (test)	Ljung-Box (P-Value)
Seasonal Naïve	N/A	-	1.527044	-	1.248044	2.2e-16
Simple Exponential Smoothing	(A, N,N)	-	0.9509481	-	1.643461	2.2e-16
Holt's Exponential Smoothing	(A, A,N)	-	0.6488222	-	0.8469332	0.001823
	(A, A <sub>d'</sub> ,N)	-	0.6597823	-	1.105793	0.01986
	(A, M,N)	-	0.5818108	-	0.8660047	0.01154
	(A, M <sub>d'</sub> , N)	-	0.5807925	-	0.9971667	0.0577
Holt-Winters Exponential Smoothing	(M,A,M)	(0.5353, 1e-04, 2e-04)	0.5906786	851.2826	0.8525379	0.01196
	(A,A,A)	(0.4797, 1e-04, 0.2997)	0.6488222	945.6878	0.8469332	0.001823
	(M,A,A)	(0.5215, 1e-04, 0.206)	0.6566577	884.2588	0.8176148	0.02815
	(M,M,M)	(0.554, 1e-04, 1e-04)	0.5900927	850.9813	0.8680253	0.01866

The Holt-Winters Exponential Smoothing model with (M,A,M) parameters appears to be the best option.

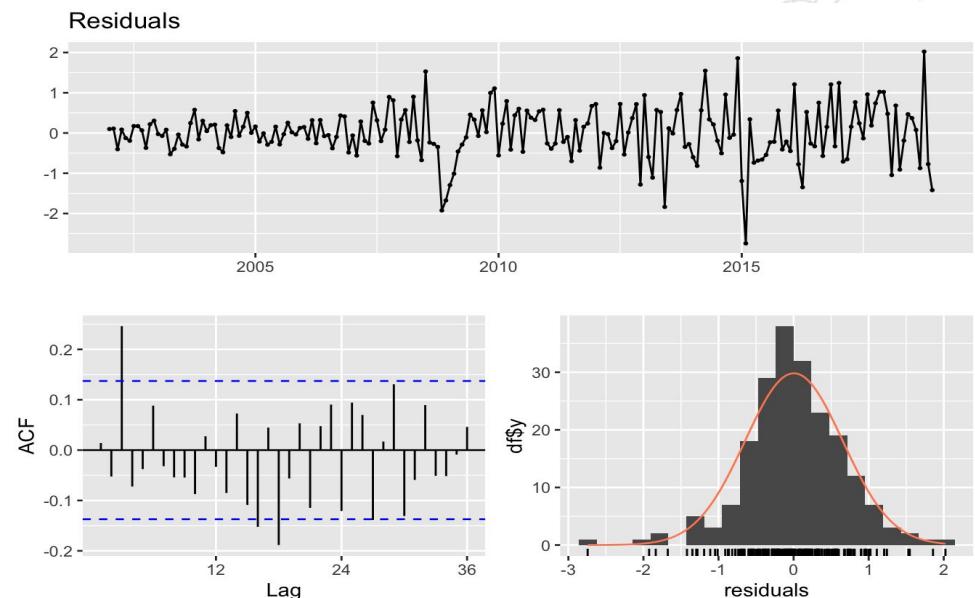
It offers a balanced performance with fairly low training and test RMSE values (0.5906786 and 0.8525379, respectively), a moderate AICc score (851.2826), and a comparatively high Ljung-Box P-value (0.01196), suggesting that its residuals are closer to white noise compared to other models.

# Model Exploration(1) ETS Pre COVID-19



## Model Summary

1. ETS(M,A,M): Error(Multiplicative), Trend(Additive), Seasonality(Multiplicative)
2. AICc: 851.2826
3. RMSE: 0.5906786(training), 0.8525379(test)
4. Smoothing Parameters ( $\alpha$ ,  $\beta$ ,  $\gamma$ ): (0.5353, 1e-04, 2e-04)



## Check Residuals

- Ljung-Box test
- p-value: 0.001823 (Total lags used: 24)
- There's autocorrelation in the residuals of a time series model.

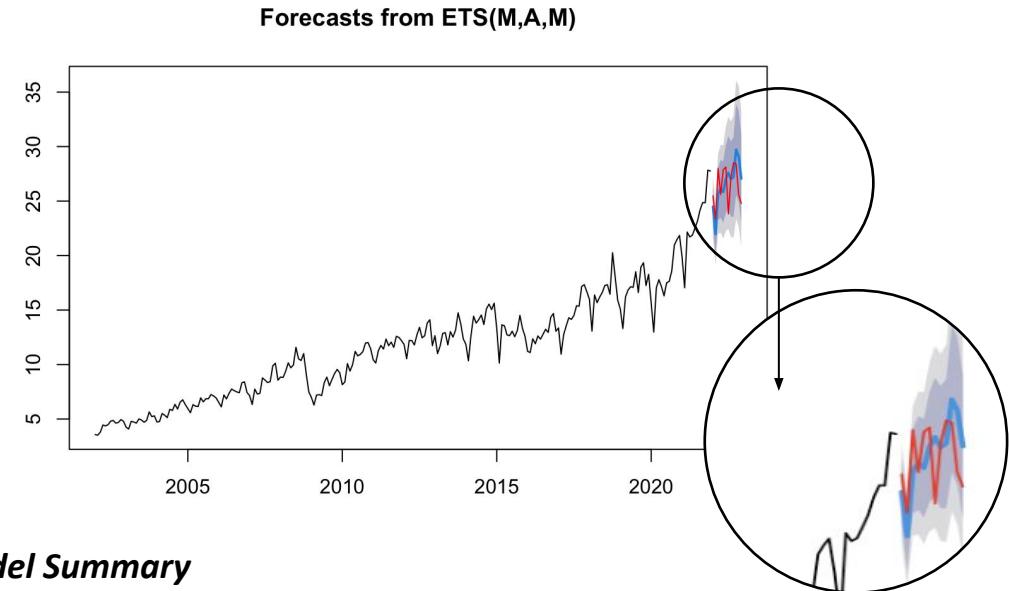
Based on the result we see from the Ljung-Box test, this outcome implies that the model fails to sufficiently capture the data's temporal dependencies and inherent patterns.

# Model Exploration(1) ETS After COVID19

TS Models	E/T/S	( $\alpha$ , $\beta$ , $\gamma$ )	RMSE(train)	AICc	RMSE(test)	Ljung-Box (P-Value)
Seasonal Naïve	N/A	-	1.926457	-	4.445163	2.2e-16
Simple Exponential Smoothing	(A, N,N)	-	1.118179	-	2.103524	2.2e-16
Holt's Exponential Smoothing	(A, A,N)	-	0.7700349	-	3.432513	0.01716
	(A, A <sub>d</sub> ,N)	-	0.7972801	-	3.829269	0.0001161
	(A, M,N)	-	0.7014787	-	3.174066	5.467e-05
	(A, M <sub>d</sub> , N)	-	0.6975741	-	3.633945	0.0006899
Holt-Winters Exponential Smoothing	(M,A,M)	(0.5311, 1e-04, 1e-04)	0.7294272	1097.287	2.019115	0.004212
	(A,A,A)	(0.4867, 0.0276, 0.346)	0.7700349	1226.677	3.432513	0.01716
	(M,A,A)	(0.5565, 1e-04, 0.2409)	0.773105	1129.110	1.907392	0.06159
	(M,M,M)	(0.5646, 1e-04, 9e-04)	0.7196395	1097.301	2.680248	0.003535

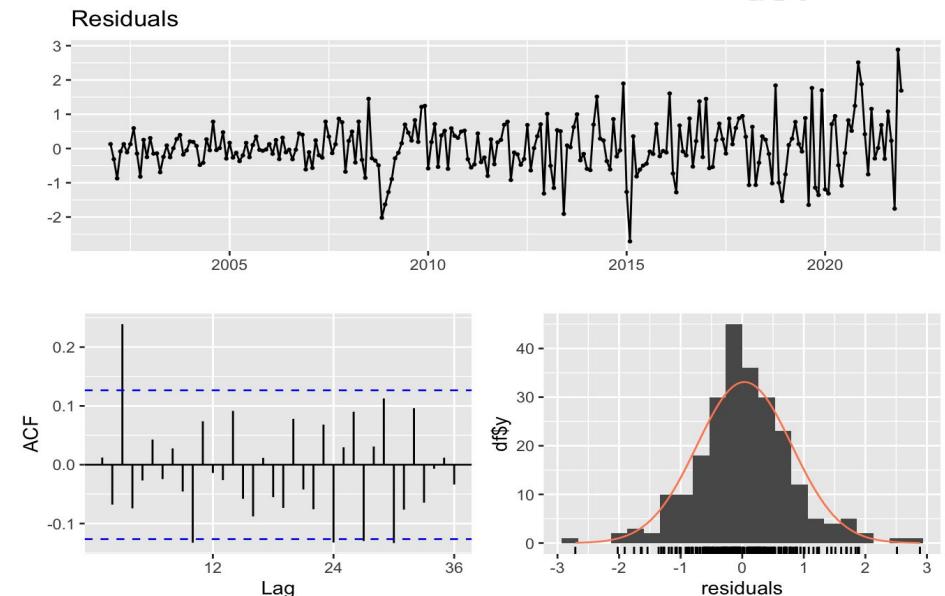
The ETS(M,A,M) model stands out due to its consistently strong performance, as indicated by its favorable AICc and RMSE values for both the training and test sets.

# Model Exploration(1) ETS Post COVID-19



## Model Summary

1. ETS(M,A,M): Error(Multiplicative), Trend(Additive), Seasonality(Multiplicative)
2. AICc: **1097.287**
3. RMSE: **0.7294272 (train); 2.019115 (test)**
4. Smoothing Parameters ( $\alpha$ ,  $\beta$ ,  $\gamma$ ): (0.5311, 1e-04, 1e-04)



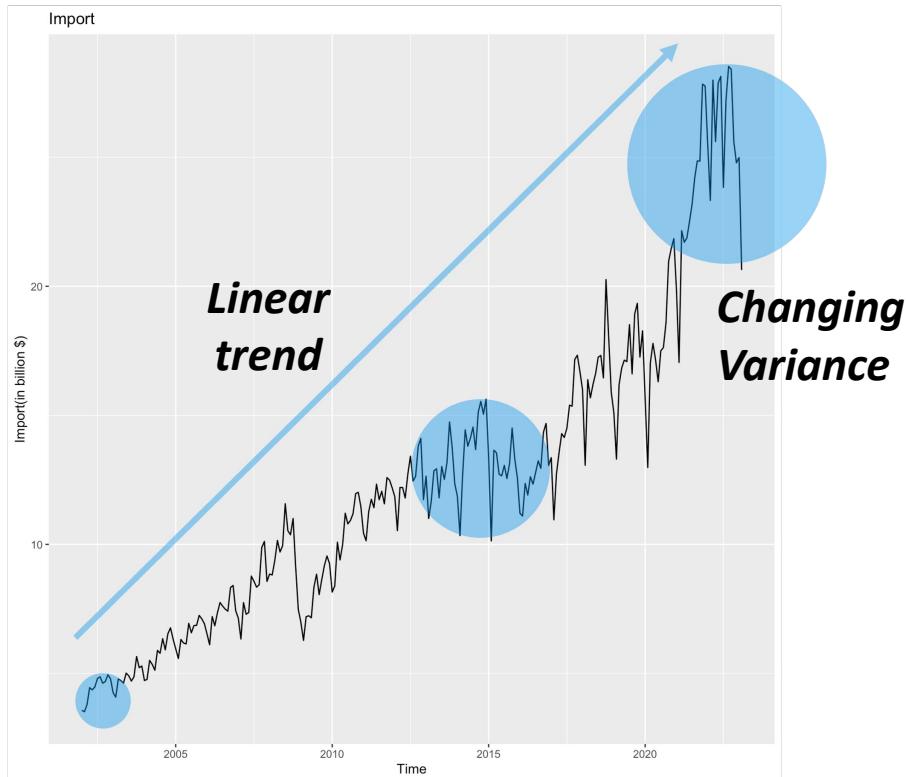
## Check Residuals

- Ljung-Box test
- p-value: 0.004212 (Total lags used: 24)
- There's autocorrelation in the residuals of a time series model.

In both predictions before and after the COVID-19 period, the ETS(M,A,M) model consistently outperformed other models. However, we observed a decline in the model's performance when incorporating the COVID-19 pandemic period for prediction.

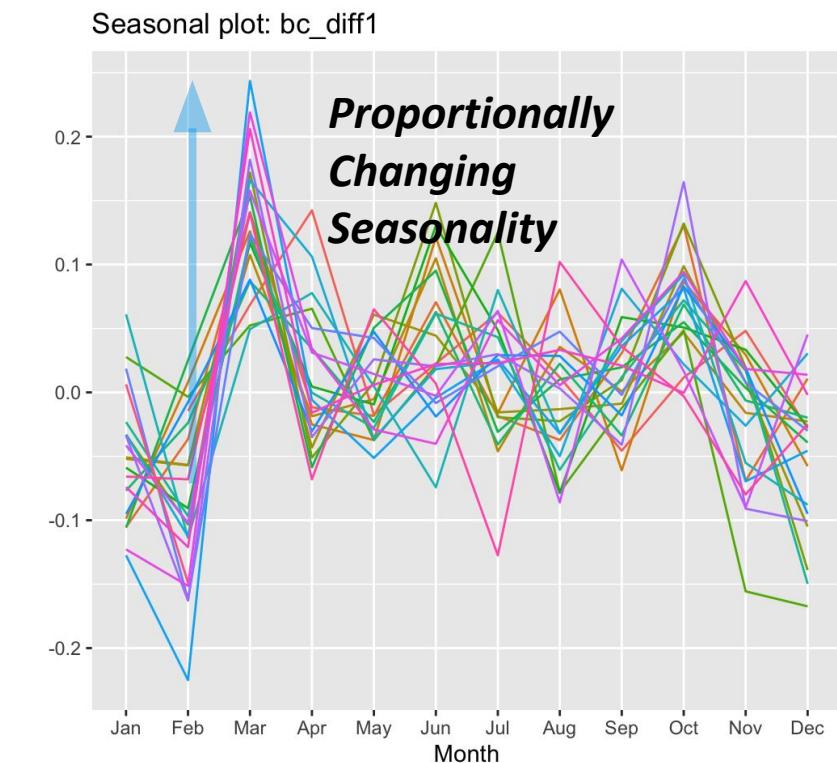
# Model Exploration

## (1) ETS Summary



### Model Selection

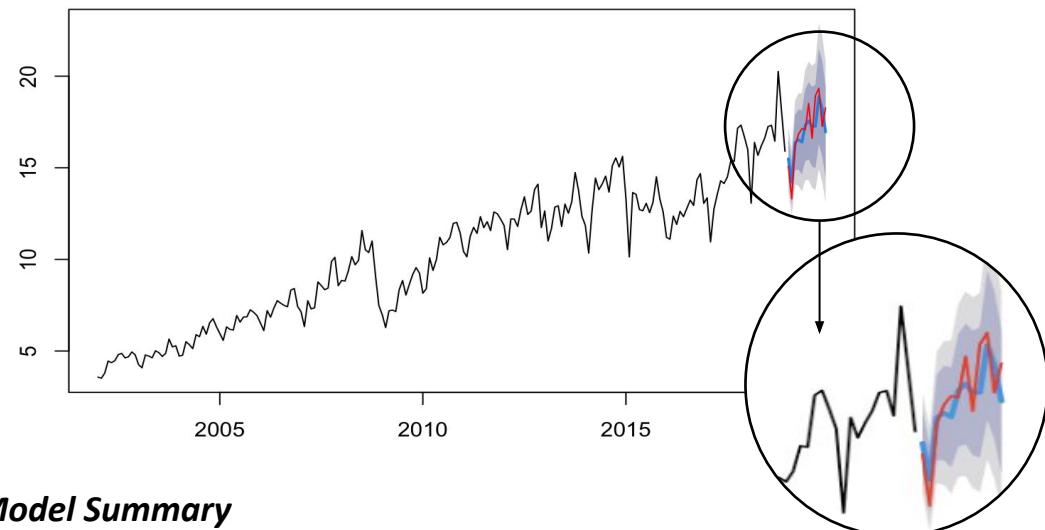
- **The  $ETS(M,A,M)$  model** is chosen as the best model due to its ability to capture the changing variance over time (Multiplicative Error - M), linear trend (Additive Trend - A), and proportionally changing seasonality pattern (Multiplicative Seasonality - M).



# Model Exploration(1) ETS Summary

Pre Covid, ETS(M,A,M)

Forecasts from ETS(M,A,M)

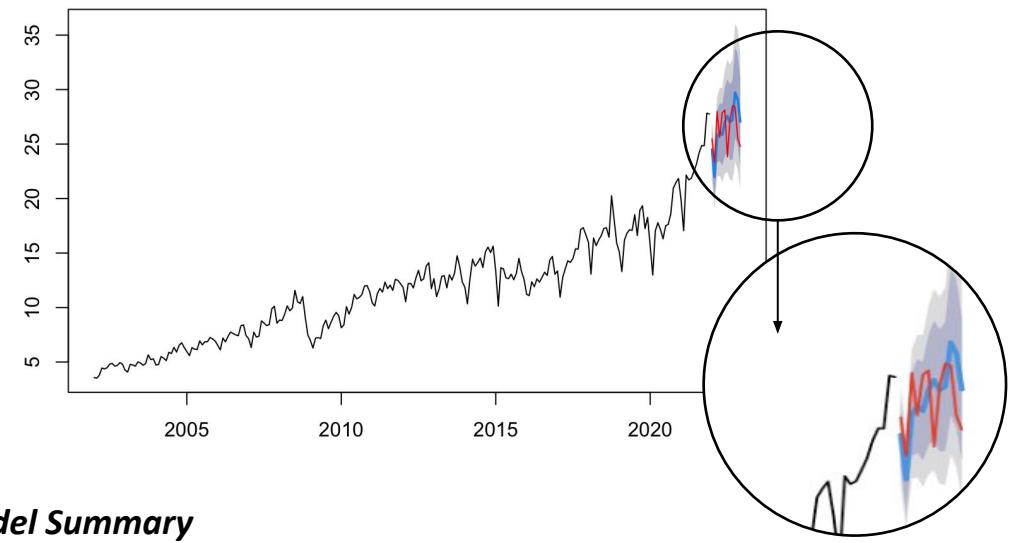


## Model Summary

1. ETS(M,A,M): Error(Multiplicative), Trend(Additive), Seasonality(Multiplicative)
2. AICc: 851.2826
3. RMSE: 0.5906786(train), 0.8525379(test)
4. Smoothing Parameters ( $\alpha, \beta, \gamma$ ): (0.5353, 1e-04, 2e-04)

Post Covid, ETS(M,A,M)

Forecasts from ETS(M,A,M)

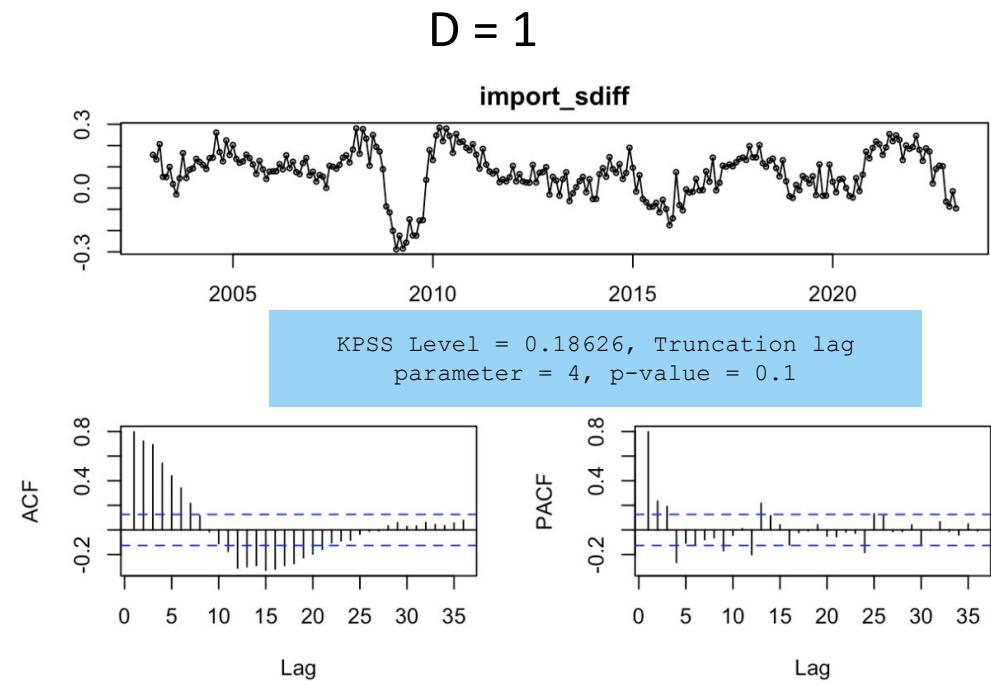
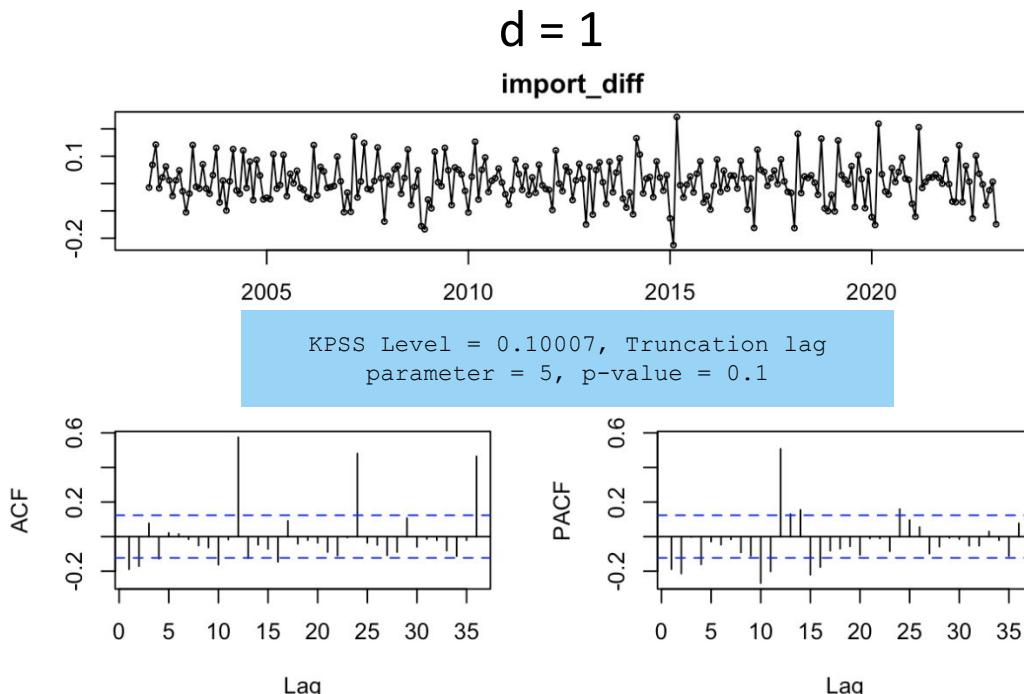


## Model Summary

1. ETS(M,A,M): Error(Multiplicative), Trend(Additive), Seasonality(Multiplicative)
2. AICc: **1097.287**
3. RMSE: **0.7294272 (train); 2.019115 (test)**
4. Smoothing Parameters ( $\alpha, \beta, \gamma$ ): (0.5311, 1e-04, 1e-04)

# Model Exploration (2) SARIMA

SARIMA Assumption - Time series must be stationary.



Exploration Method: `Auto.Arima()` and explore other orders with different differencing.

# Model Exploration (2) - SARIMA Pre COVID

No	Model		AICc	RMSE (train)	Ljung-Box (P-Value)	RMSE (test)
	Order (p,d,q)(P,D,Q)[N]	drift				
1	(4,0,0)(1,1,2)[12]	Yes	-470.5	0.5971973	0.4979	0.9223954
2	(4,0,0)(1,1,2)[12]	No	-467.13	0.6043885	0.3191	0.9219941
3	(4,0,0)(0,1,1)[12]	Yes	-474.7	0.5988067	0.6329	0.9224962
4	(4,0,0)(0,1,1)[12]	No	-471.38	0.6054706	0.4527	0.9207406
5	(3,0,1)(0,1,1)[12]	Yes	-471.64	0.5991217	0.199	0.9189453
6	(3,0,1)(0,1,1)[12]	No	-469.23	0.6053335	0.1559	0.8641353
7	(2,0,3)(0,1,1)[12]	Yes	-470.98	0.6009981	0.2485	0.9144917
8	(2,0,3)(0,1,1)[12]	No	-467.07	0.6086834	0.142	0.9842110
9	(4,1,0)(1,0,2)[12]	Yes	N/A	N/A	N/A	N/A
10	(4,1,0)(0,0,1)[12]	Yes	-402.05	0.8027907	4.839e-08	1.1848420
11	(3,1,1)(0,0,1)[12]		-402.36	0.8038033	6.836e-08	1.5335435
12	(2,1,0)(1,0,0)[12]	Yes	-440.15	0.7246152	0.0006764	1.0411485

← from Auto.Arima()

## Guide for exploration:

1. tsdisplay()
2. Extended Autocorrelation Function [eacf()]

AR/MA	0 1 2 3 4 5 6 7 8 9 10 11 12 13
0	x x x x x x x x x x x x x x
1	x x o x o o o o o o x o o
2	x x o x o o o o o o x o o
3	o x o x o o o o o o x x o
4	o x x x o o o o o o x x o
5	x o x x o o o o o o x o x
6	x o o x o o o o o o x o x
7	x o o x o o o o o o x o x

3. Significance of parameter
4. CheckResiduals()

# Model Exploration (2) - SARIMA for Post COVID

No	Model		AICc	RMSE (train)	Ljung-Box (P-Value)	RMSE (test)
	Order (p,d,q)(P,D,Q)[N]	drift				
1	(4,0,0)(1,1,2)[12]	Yes	-701.41	0.706554	0.5409	2.365628
2	(4,0,0)(1,1,2)[12]	No	-697.64	0.7151124	0.3938	2.8653706
3	(4,0,0)(0,1,1)[12]	Yes	-705.38	0.7090993	0.6448	2.3751877
4	(4,0,0)(0,1,1)[12]	No	-701.72	0.7170508	0.5179	2.8429648
5	(3,0,2)(0,1,1)[12]	Yes	-701.83	0.7106919	0.4118	2.4440392
6	(3,0,2)(0,1,1)[12]	No	-698.7	0.718369	0.3459	2.797201
7	(2,0,3)(1,1,2)[12]	Yes	-698.84	0.7060893	0.2978	2.3099154
8	(2,0,3)(1,1,2)[12]	No	-694.07	0.7166455	0.1551	2.9972462
9	(4,1,0)(1,0,2)[12]	Yes	N/A	N/A	N/A	N/A
10	(4,1,0)(0,0,1)[12]	Yes	-613.83	0.9223215	1.108e-10	2.6477560
11	(2,1,3)(1,0,2)[12]	Yes	-723.26	0.7174144	0.2935	3.1392563
12	(2,1,0)(1,0,0)[12]	Yes	-664.68	0.819143	0.0005101	4.117066

← from Auto.Arima()

## Guide for exploration:

1. tsdisplay()
  2. Extended Autocorrelation
  3. Significance of parameter
  4. CheckResiduals()
- |   |   |   |   |   |   |   |   |   |   |    |    |    |    |
|---|---|---|---|---|---|---|---|---|---|----|----|----|----|
| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 |
| 0 | x | x | x | x | x | x | x | x | x | x  | x  | x  | x  |
| 1 | x | x | o | x | o | o | o | o | o | x  | o  | o  | o  |
| 2 | x | x | o | x | o | o | o | o | o | x  | x  | o  | o  |
| 3 | o | x | o | x | o | o | o | o | o | x  | x  | o  | o  |
| 4 | o | x | x | o | o | o | o | o | o | x  | x  | o  | o  |
| 5 | x | o | o | x | x | o | o | o | o | x  | o  | o  | o  |
| 6 | x | o | o | x | o | o | o | o | o | x  | x  | o  | o  |
| 7 | x | x | o | x | o | o | o | o | o | x  | o  | o  | o  |

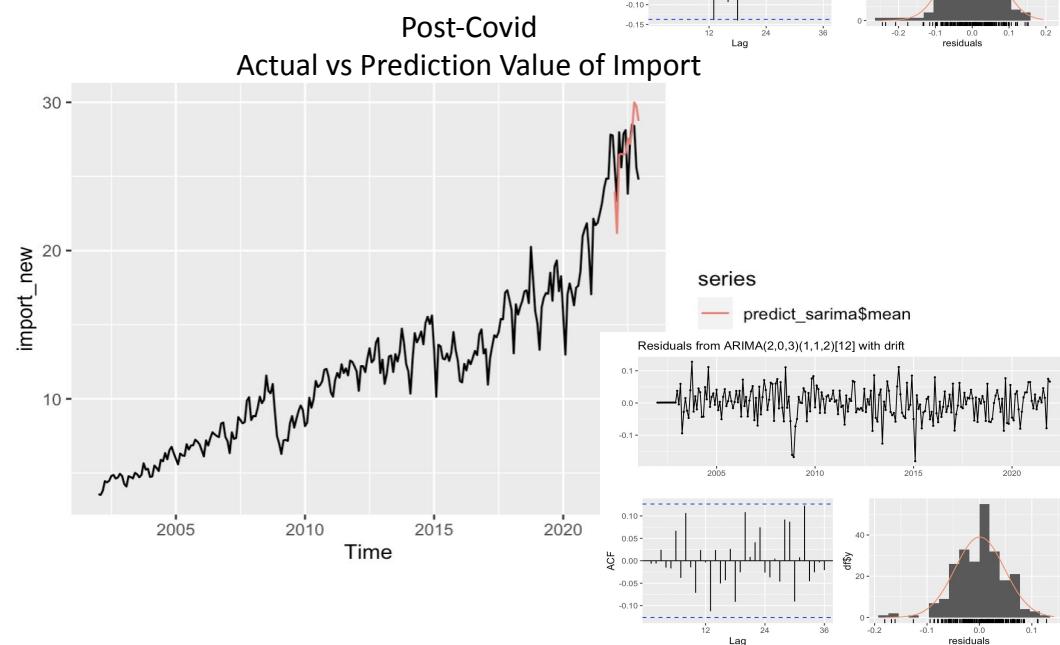
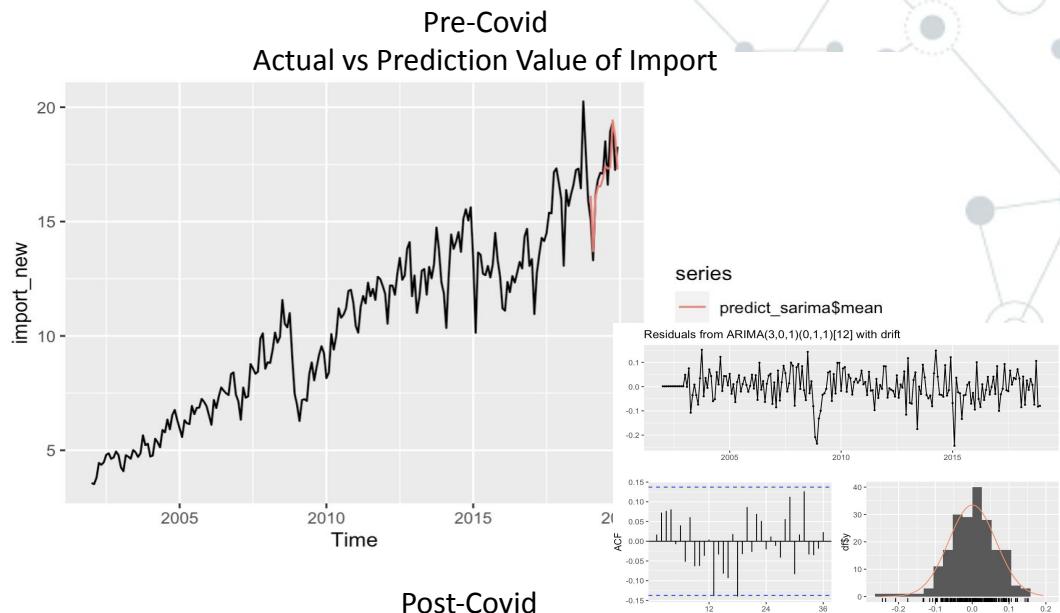
# Model Exploration (2) - Best sARIMA Model

Period	Best sARIMA Model		RMSE (test)
	Order (p,d,q)(P,D,Q)[N]	drift	
Pre-COVID	(3,0,1)(0,1,1)[12]	Yes	0.8641353
Post-COVID	(2,0,3)(1,1,2)[12]	No	2.3099154

Period	Base Model	RMSE (test)
Pre-COVID	Seasonal Naive Forecast	1.248044
Post-COVID	Naive Forecast	2.235439

## Notes:

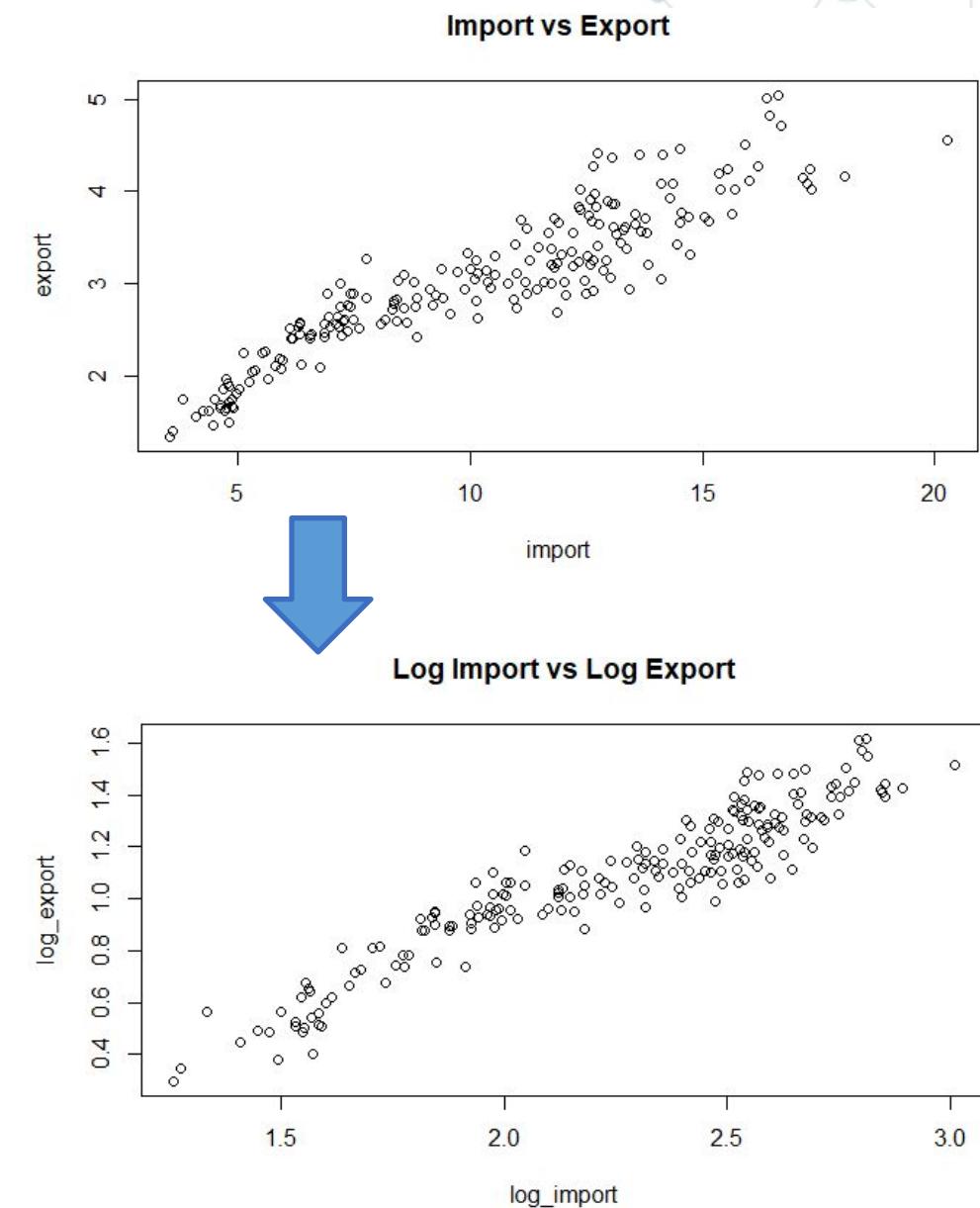
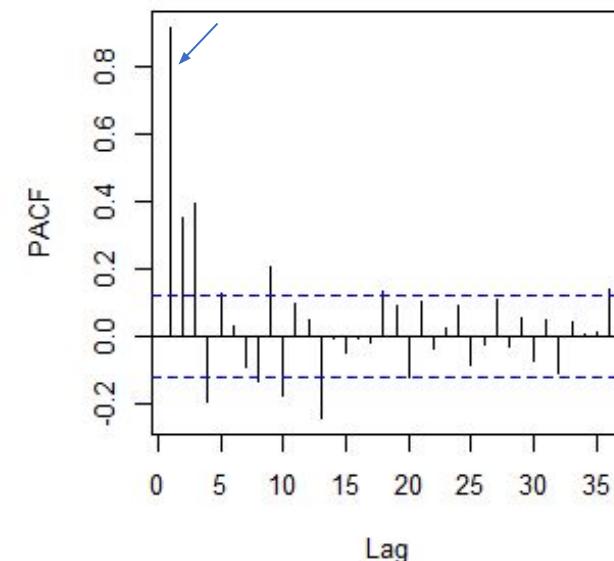
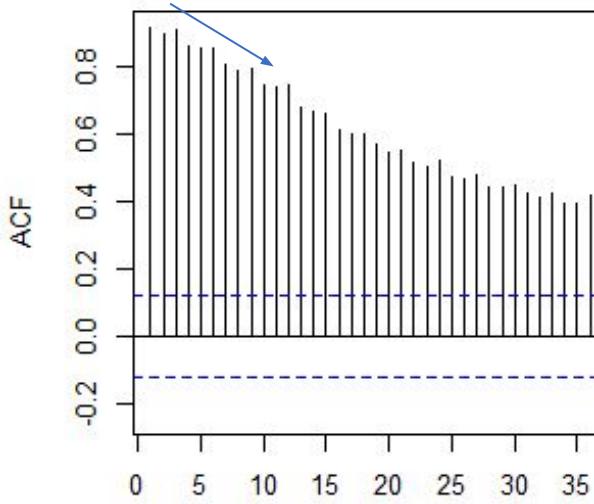
- For Pre-COVID, Arima forecast is better than the base model.
- For Post-COVID, Arima forecast is worse than the base model.
- COVID era changes the data generating process that could not be accounted by Arima Model.



# Model Exploration (3): Regression with ARIMA Errors

## Export as the explanatory variable

- Choosing Export as the covariate for the model
- Strong correlation with Import (95.65%)



# Model Exploration (3) – Regression with ARIMA error

## Pre COVID (2019)

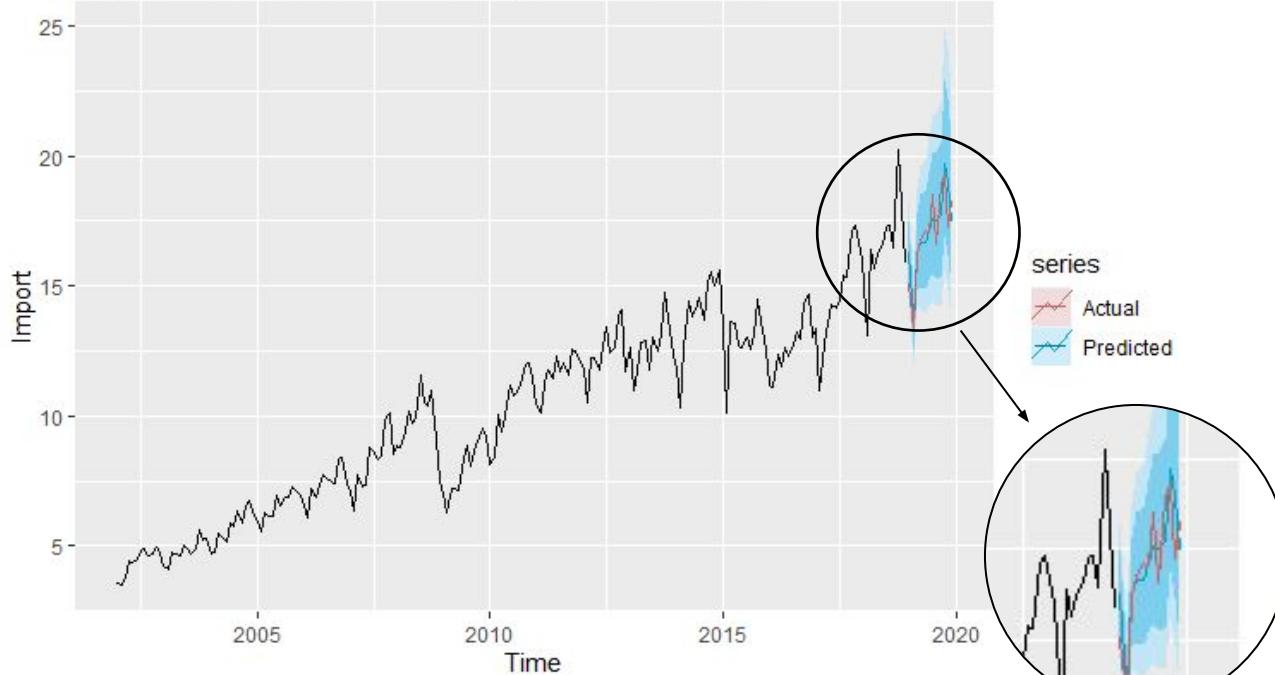
No	Model	AICc	RMSE (train)	P-Value (Ljung-Box)
	Order (p,d,q)(P,D,Q)[N]			
from Auto.Arima()	1 (1,0,3)(1,1,2)[12]	-523.6	0.05479793	0.1094
	2 (1,0,3)(1,1,1)[12]	-520.98	0.05567779	0.05388
	3 (1,0,3)(0,1,1)[12]	-522.96	0.05574829	0.05427
	4 (1,1,3)(0,1,1)[12]	-522.42	0.05547876	0.1040
	5 (1,0,4)(1,1,2)[12]	-519.15	0.05539141	0.08216
	6 (1,0,4)(1,1,1)[12]	-521.08	0.05544364	0.1108
	7 (2,1,0)(1,1,1)[12]	-521.39	0.05571836	0.01013
	8 (2,1,1)(0,1,1)[12]	<b>-525.53</b>	0.05535418	0.1381
	9 (2,1,2)(1,1,1)[12]	-522.36	0.05523922	0.1837
	10 (2,1,2)(1,1,2)[12]	-520.24	0.05522163	0.1297



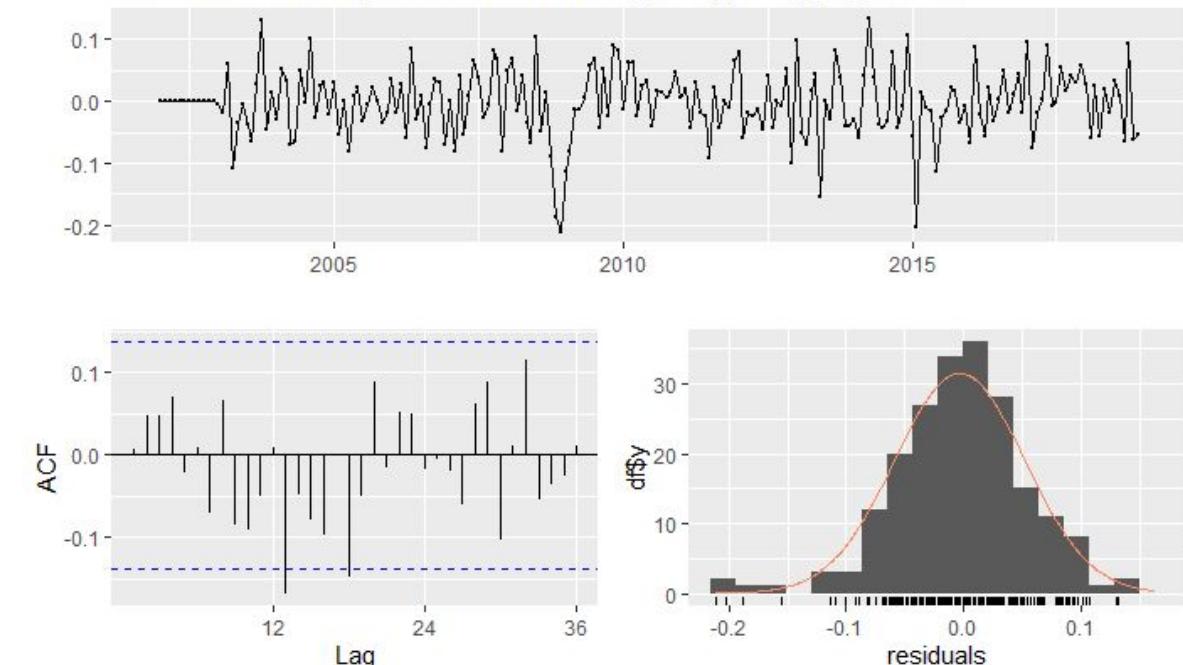
# Model Exploration (3) –

## Best Regression with ARIMA error Model Pre COVID

Import Forecast in 2019 with Export as Covariate



Residuals from Regression with ARIMA(2,1,1)(0,1,1)[12] errors



### Model Summary

Regression with ARIMA(2,1,1)(0,1,1)[12] errors

AICc: -525.53

RMSE (train): 0.0554

RMSE (test): 0.8446

### Check Residuals

Ljung-Box test:

p-value = 0.1381, Total lags used: 24

Autocorrelations still exist

# Model Exploration (3) – Regression with ARIMA error

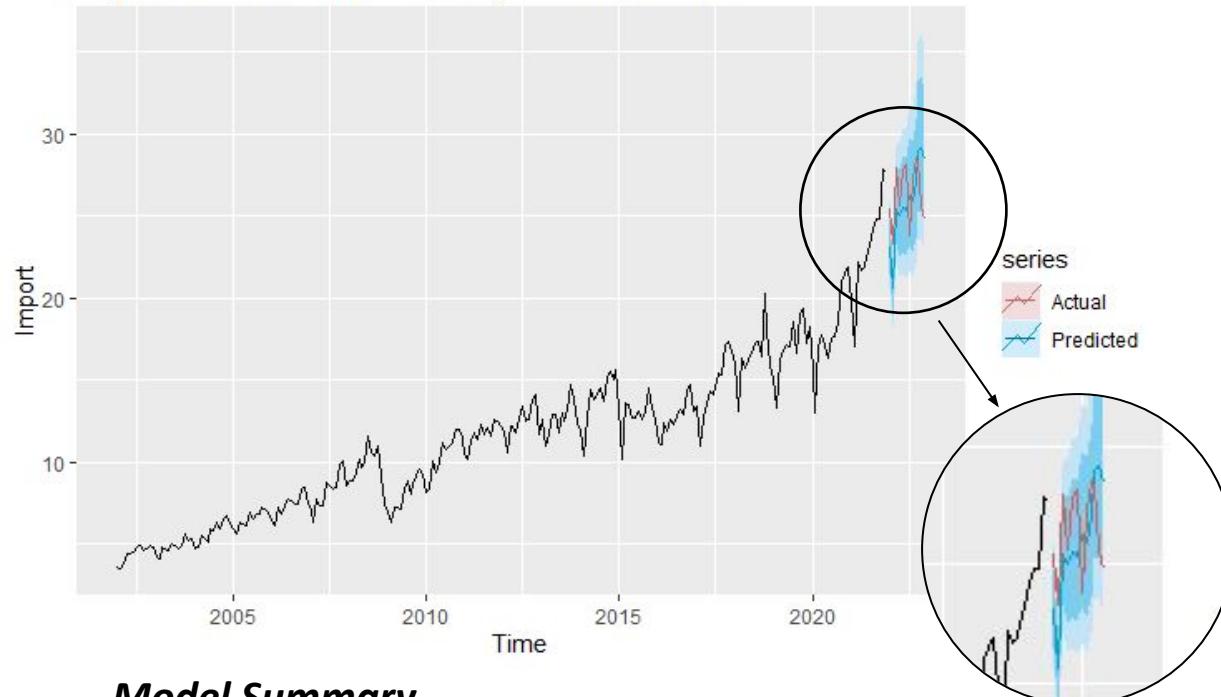
## Post COVID (2022)

No	Model	AICc	RMSE (train)	P-Value (Ljung-Box)
	Order (p,d,q)(P,D,Q)[N]			
from Auto.Arima()	1	<b>(4,0,0)(1,1,2)[12]</b>	<b>-633.24</b>	0.05469901
	2	(2,0,2)(0,1,1)[12]	-614.46	5.092e-05
	3	(3,0,1)(0,1,1)[12]	-631.27	0.1713
	4	(3,0,1)(1,1,1)[12]	-629.13	0.1362
	5	(3,1,0)(0,1,1)[12]	-632.84	0.4102
	6	(3,1,1)(0,1,1)[12]	-630.85	0.3841
	7	(4,0,0)(0,1,1)[12]	-632.58	0.3754
	8	(4,1,1)(1,1,1)[12]	-628.49	0.232
	9	(5,0,0)(0,1,1)[12]	-630.49	0.3311
	10	(5,1,1)(0,1,1)[12]	-627	0.2825

Best model

# Model Exploration (3) – Best Regression with ARIMA errors Model Post COVID

Import Forecast in 2022 with Export as Covariate



## Model Summary

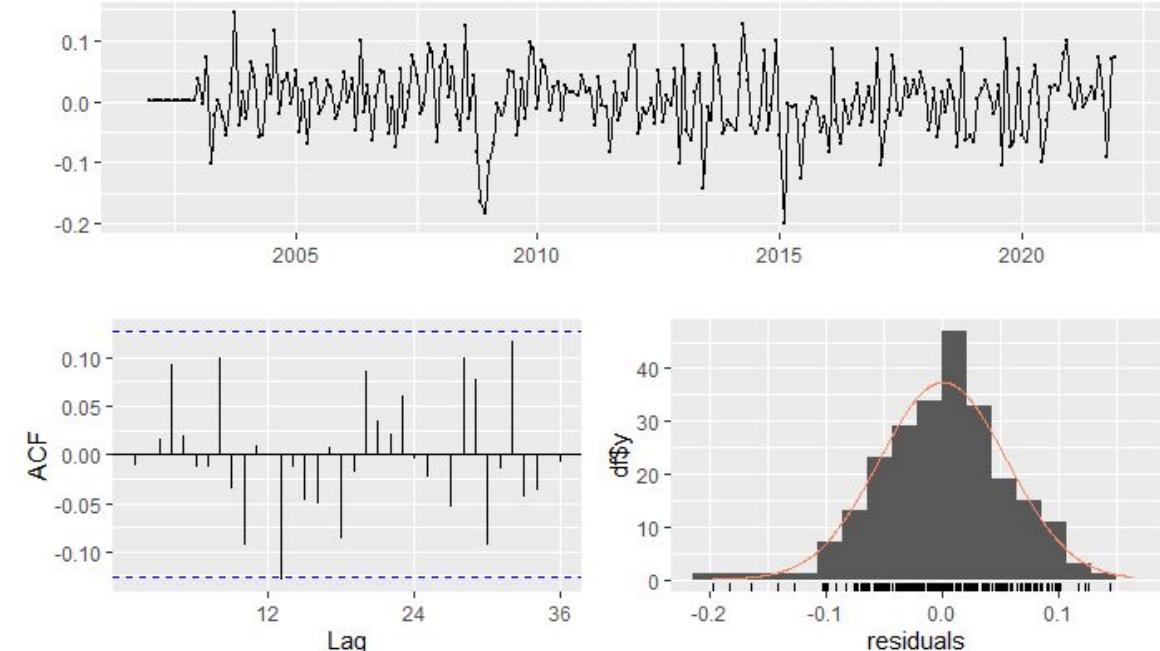
Regression with ARIMA(4,0,0)(1,1,2)[12] errors

AICc: -633.24

RMSE (train): 0.0547

RMSE (test): 2.3918

Residuals from Regression with ARIMA(4,0,0)(1,1,2)[12] errors

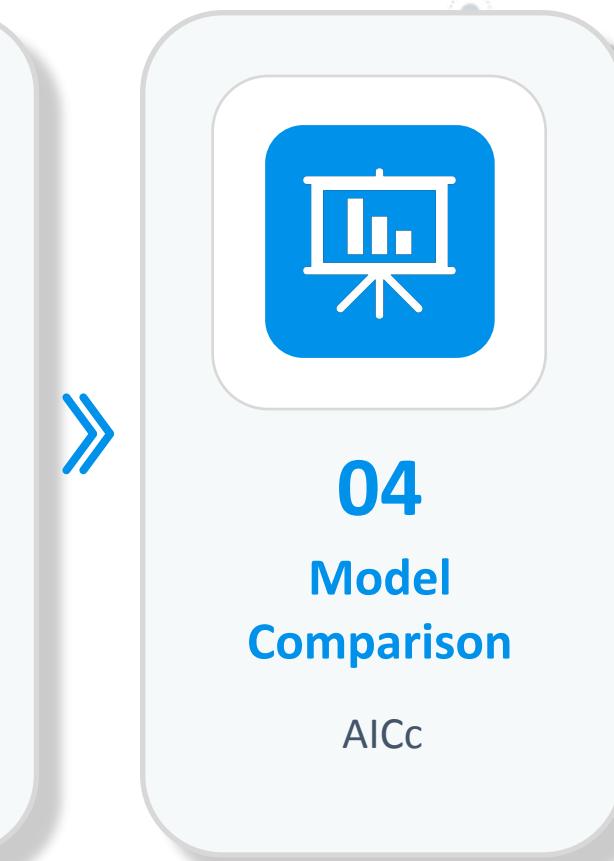
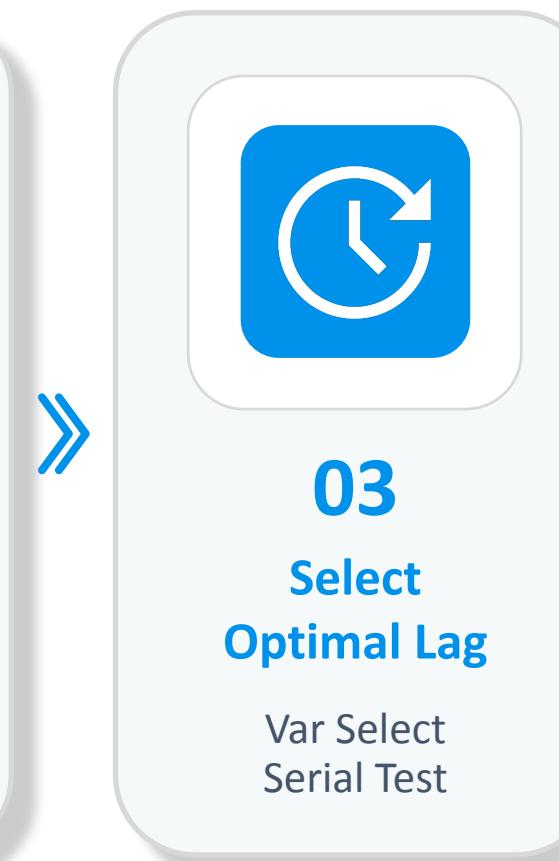
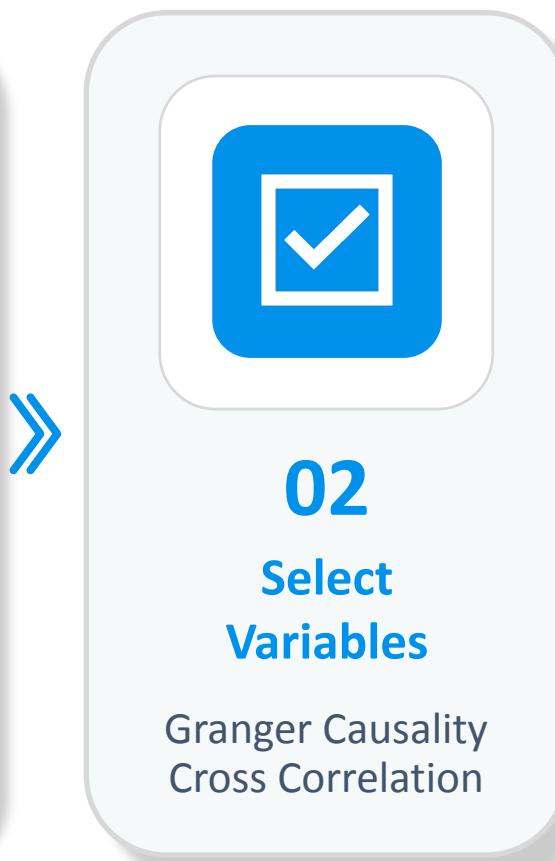
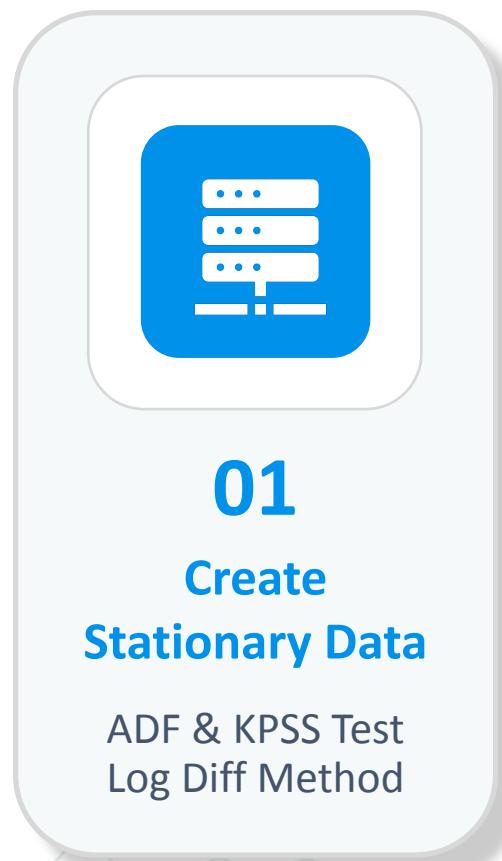


## Check Residuals

Ljung-Box test:

p-value = 0.3682, Total lags used: 24

# Model Exploration (4) Building VAR Model



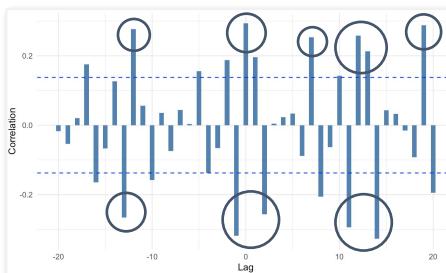
# Model Exploration (4) VAR Model > Select Variables

Granger  
Causality

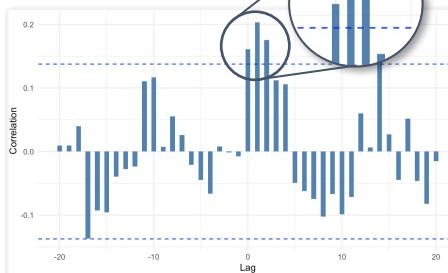
Pre Covid

Variables	P value	Significant
EXPORT	0.00012	Caused
CPI	0.00045	Caused
PPI	0.00064	Caused
...	...	...
UNCERTAINTY	0.39899	Not Caused

Cross  
Correlation



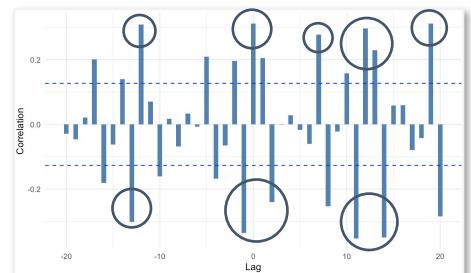
< Import vs. Export >



< Import vs. PPI >

Post Covid

Variables	P value	Significant
EXPORT	0.00000	Caused
CPI	0.00052	Caused
PPI	0.00141	Caused
...	...	...
UNCERTAINTY	0.27819	Not Caused



< Import vs. Export >



< Import vs. PPI >



# Model Exploration (4) VAR Model > Select Optimal Lags

VAR  
Select

Serial Test

Pre Covid

AIC (n)	HQ (n)	SC (n)	FPE (n)
2	2	2	2

Lags	P-value	Autocorrelation
2	0.0969	Not Autocorrelated
3	0.1252	Not Autocorrelated
4	0.19486	Not Autocorrelated
5	0.22673	Not Autocorrelated

Post Covid

AIC (n)	HQ (n)	SC (n)	FPE (n)
2	2	2	2

Lags	P-value	Autocorrelation
2	0.03088	Autocorrelated
3	0.0337	Autocorrelated
4	0.02565	Autocorrelated
5	0.12602	Not Autocorrelated
6	0.07209	Not Autocorrelated
7	0.0404	Autocorrelated

# Model Exploration (4) VAR Model Comparison

## Pre Covid

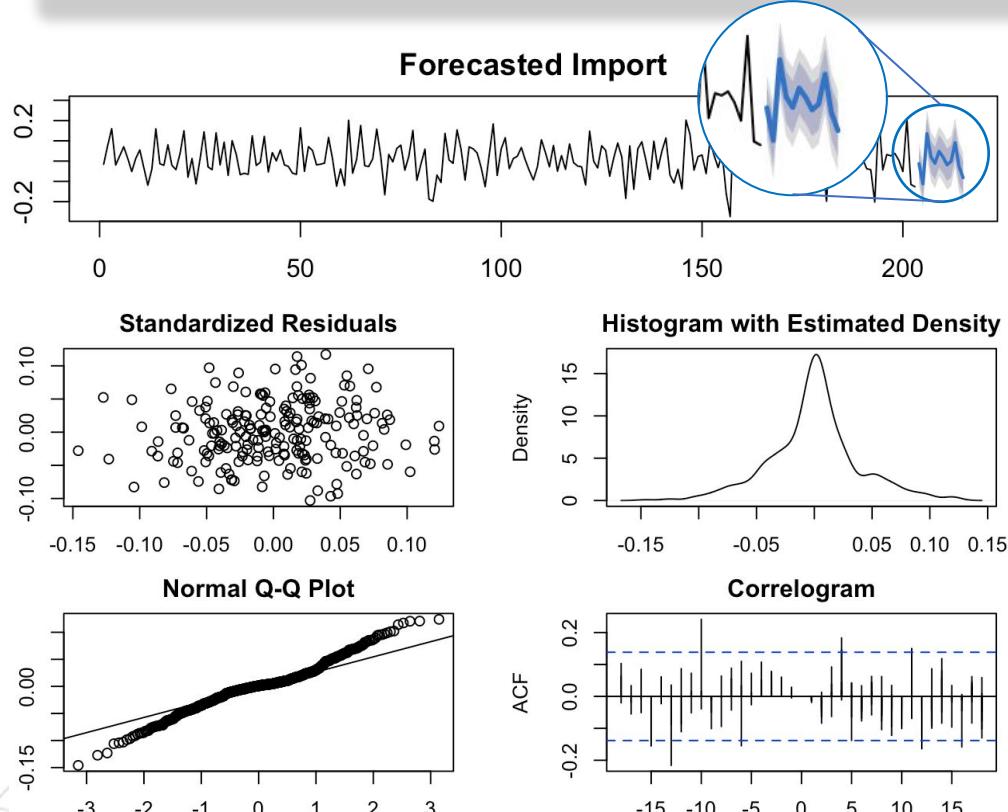
Lags	MAE	MSE	RMSE	AICc	MAPE
2	0.7923706	0.9031720	0.9503536	-2487.399	0.04560345
3	0.7850147	0.8518980	0.9229832	-2469.384	0.04571006
4	0.7846202	0.7942794	0.8912235	-2545.401	0.04620134
5	0.8424218	0.9754128	0.9876299	-2433.698	0.04900531

## Post Covid

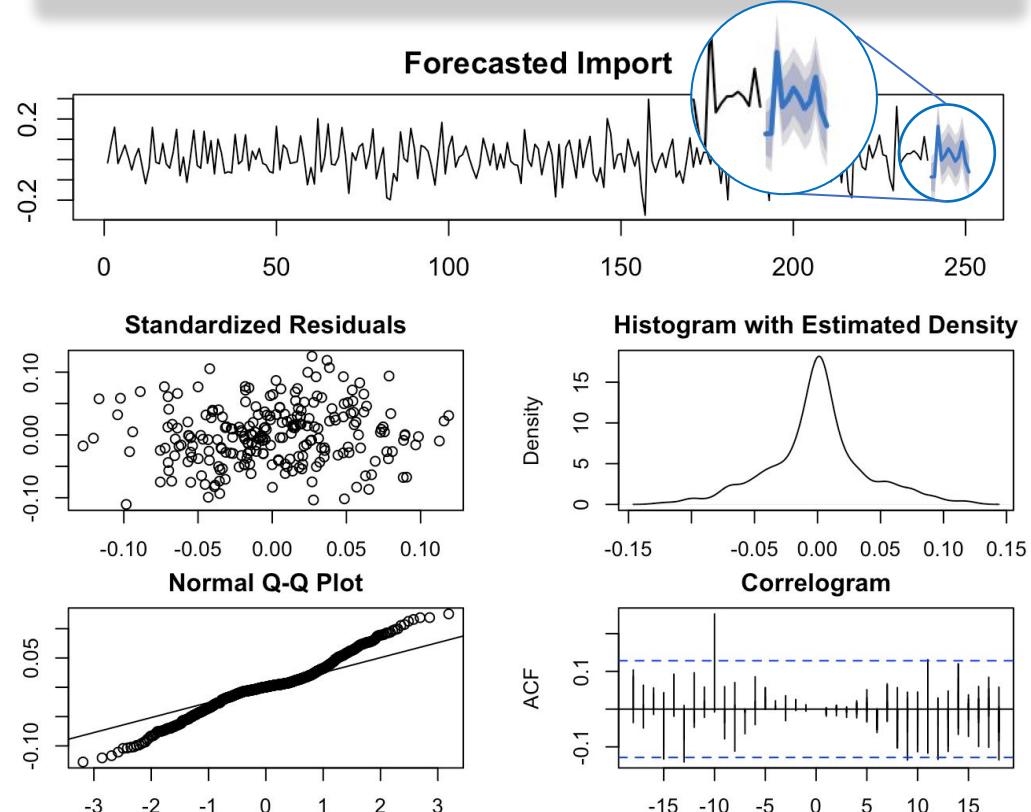
Lags	MAE	MSE	RMSE	AICc	MAPE
5	2.823877	14.67617	3.830949	-2858.115	0.1091495
6	2.517569	11.59846	3.405651	-2833.328	0.0974212

# Model Exploration (4) VAR Model Comparison

Pre Covid, VAR(2)



Post Covid, VAR(5)



# Model Selection

## Best Model:

- Pre COVID forecast (2019): Regression with ARIMA Errors (2,1,1)(0,1,1)[12]
- Post COVID forecast (2022): ETS (M,A,M).

### Pre COVID (2019)

Models	RMSE
ETS	0.8525
SARIMA	0.8641
Regression with ARIMA errors	0.8446
VAR	0.9504

### Post COVID (2022)

Models	RMSE
ETS	2.0191
SARIMA	2.3099
Regression with ARIMA errors	2.3918
VAR	3.8309

# Conclusion and Future Work

**Conclusion:** Performance of models **reduce** after COVID-19 period.

- Increased Volatility: Higher volatility and uncertainty introduced in 2020, making time series models less accurate.
- Model Overfitting: Potential overfitting to the data from 2002-2018 leads to poor generalization when encountering the unique patterns of the 2019-2021 period.
- Seasonality Impact: Increased importance of seasonality during and after the pandemic due to lockdowns, reopening, and changing consumer behavior.

**Future Work:**

- Examine additional potential explanatory variables.
- Try forecasting using other models (E.g. TBATS, Dynamic Harmonic Regression, ARCH/GARCH, etc).
- Build the model and forecast based on the industry.



THANK YOU

# Roles and Responsibilities

Topic #	Contents	Point Person
1	Goal of Analysis	Iris / Xiao
2	Data Description	Jaya
3	EDA	Mia / Xiao
4	Model Exploration (4+)	Exponential Smoothing (Mia) SARIMA (Jaya) Regression with ARIMA errors (Minh) VAR and VARIMA (Iris, Xiao)
5	Model Evaluation & Selection	All members
6	Conclusion and Future work	Minh