



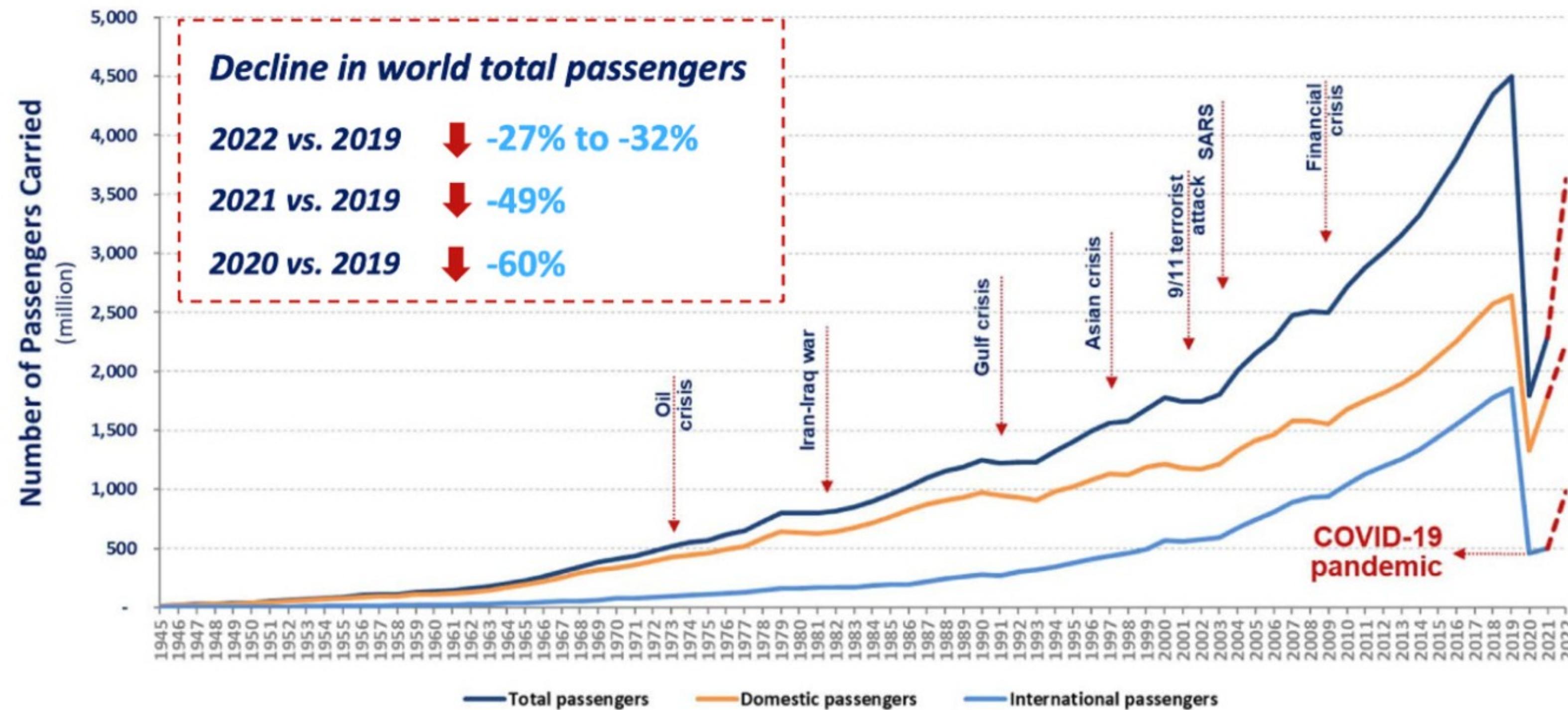
By Justin Jimenez

A large commercial airplane is shown from a low angle, flying towards the viewer against a backdrop of a vibrant orange and yellow sunset. The sky is filled with wispy clouds. The airplane's wings and engines are clearly visible.

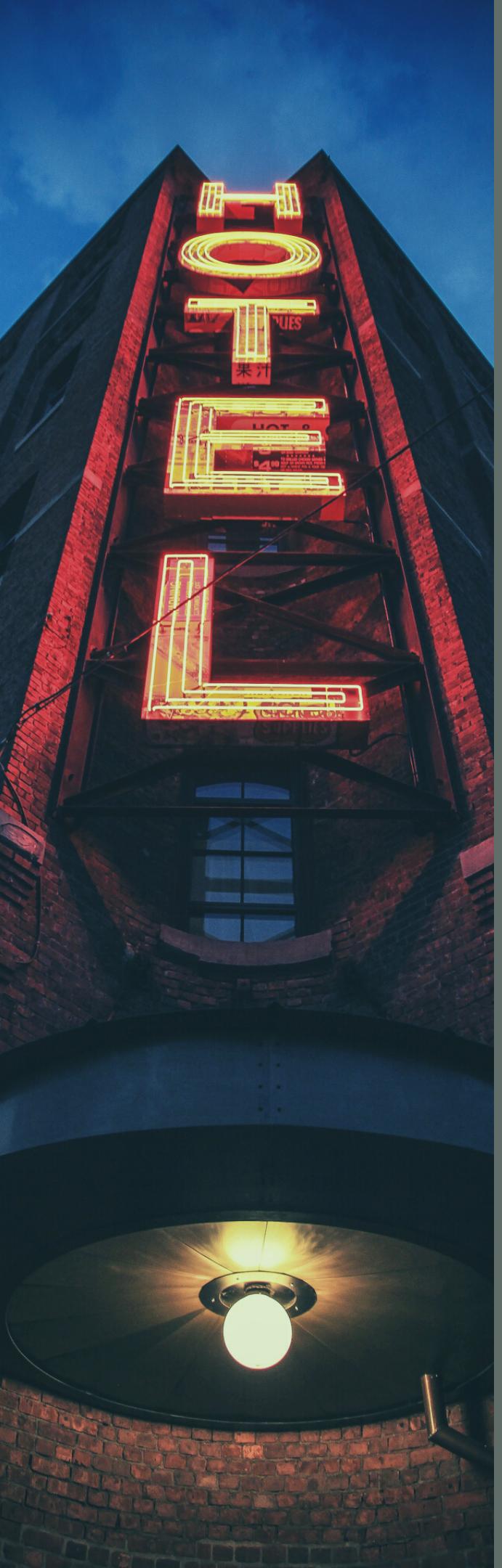
**"GLOBAL TOURISM
SUFFERED ITS WORST
YEAR ON RECORD IN
2020, WITH
INTERNATIONAL
ARRIVALS DROPPING BY
74%."**

- WORLD TOURISM ORGANIZATION

World passenger traffic evolution 1945 – 2022



Source: ICAO Air Transport Reporting Form A and A-S



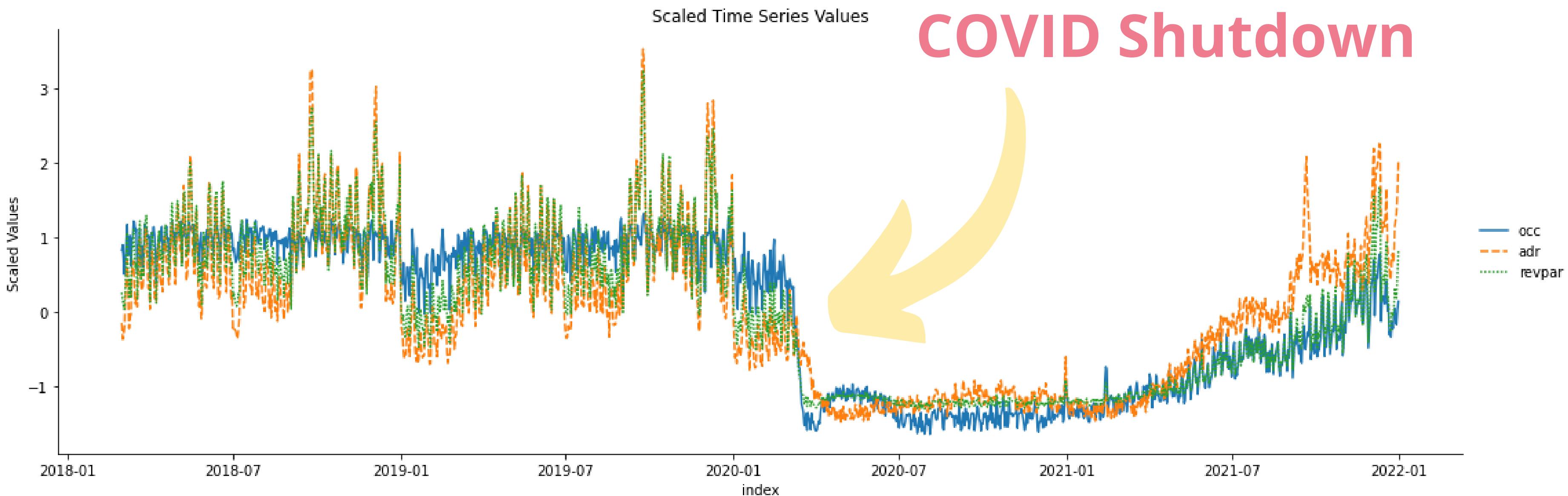
PROBLEM SCOPE DEFINITION

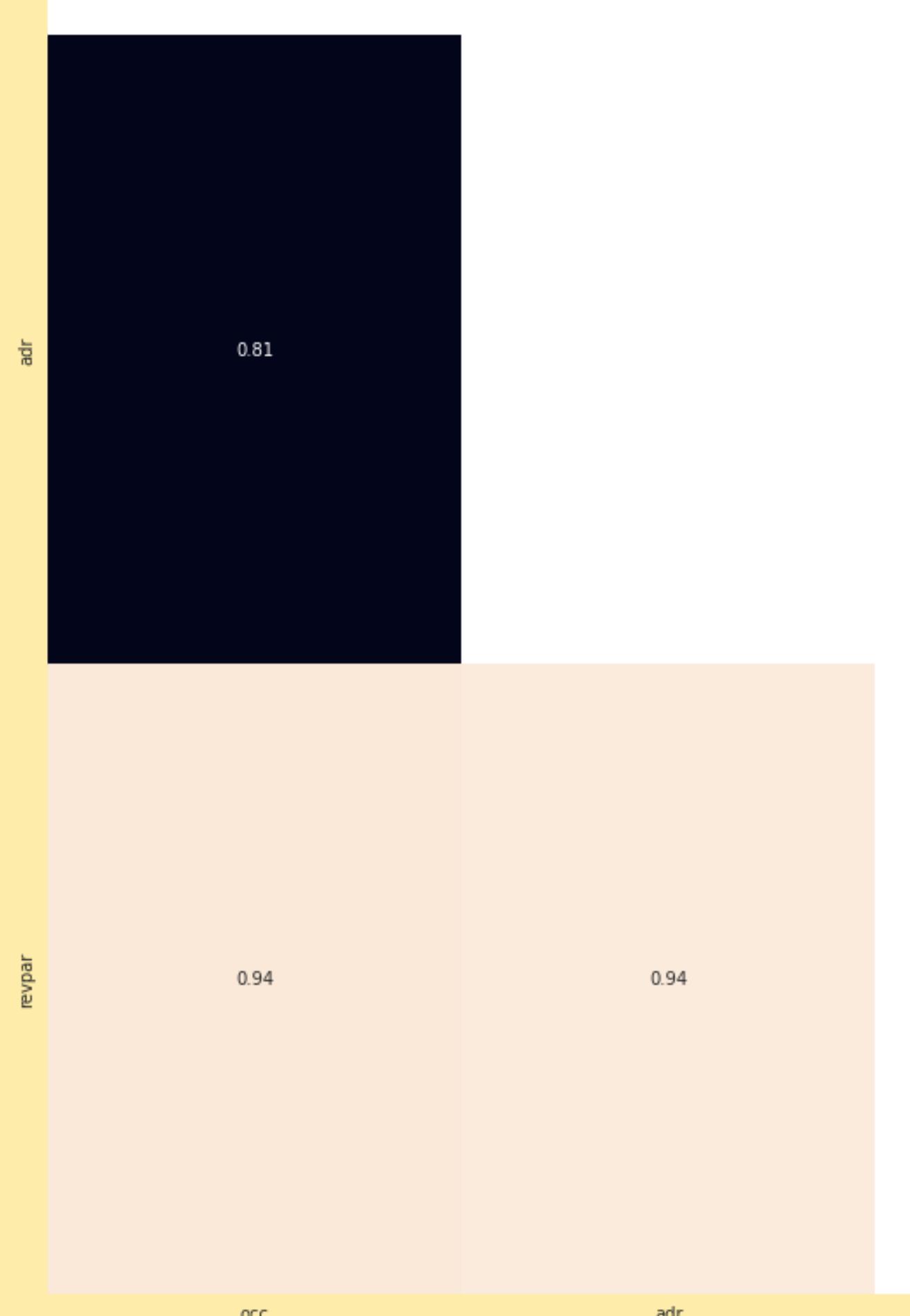
- 1 VISUALIZE THE DATA
- 2 IDENTIFY SEASONAL TRENDS
- 3 EXPLORE RELATIONSHIPS
- 4 DEVELOP A FORECASTING MODEL
- 5 RECOMMEND PLAN FOR IMPROVEMENT



- 1400 days of market data
- Starting in 03/2018, ending in 12/2021
- Including market occupation, average rate of rooms sold, and revenue per available room

Scaled Feature Values

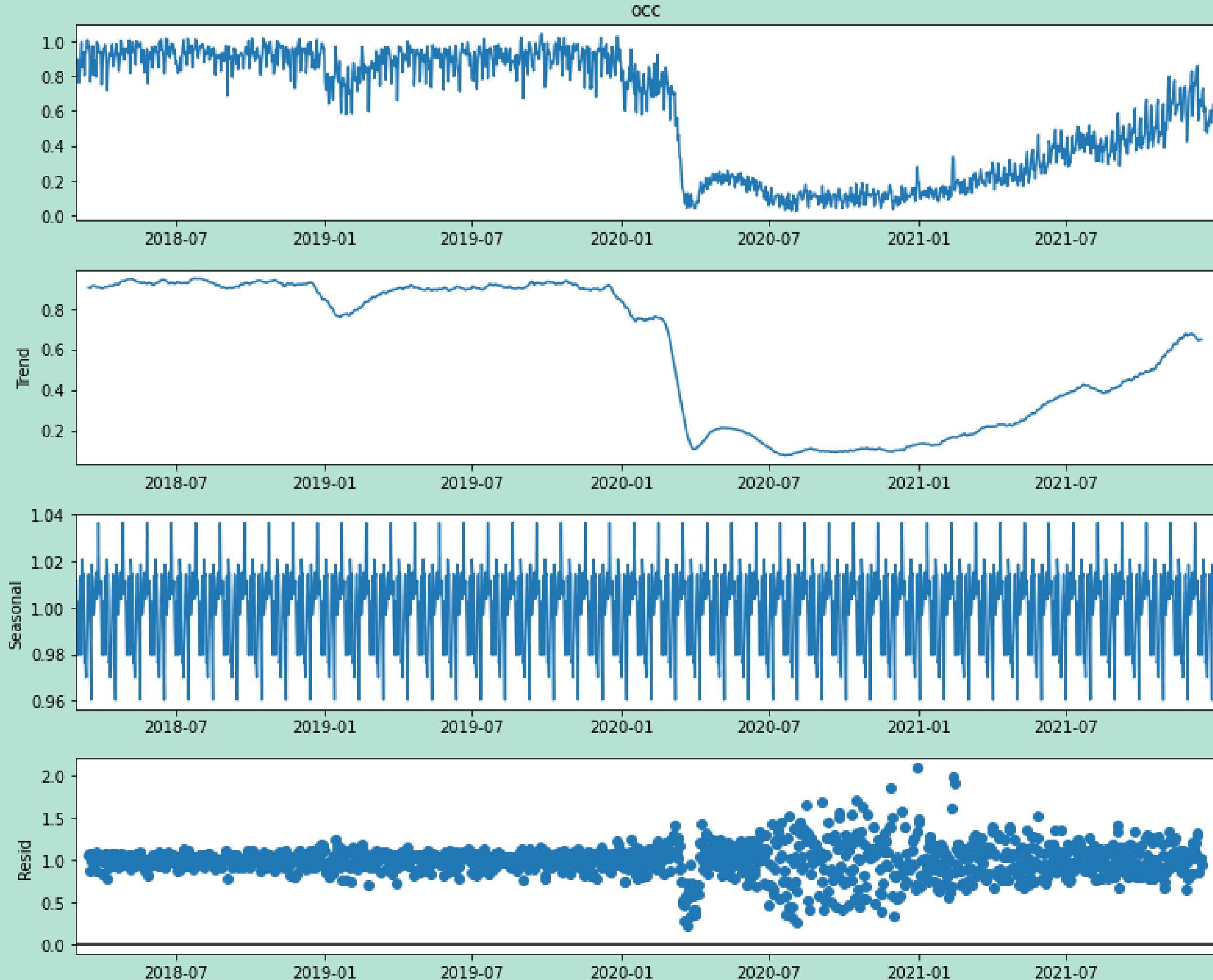




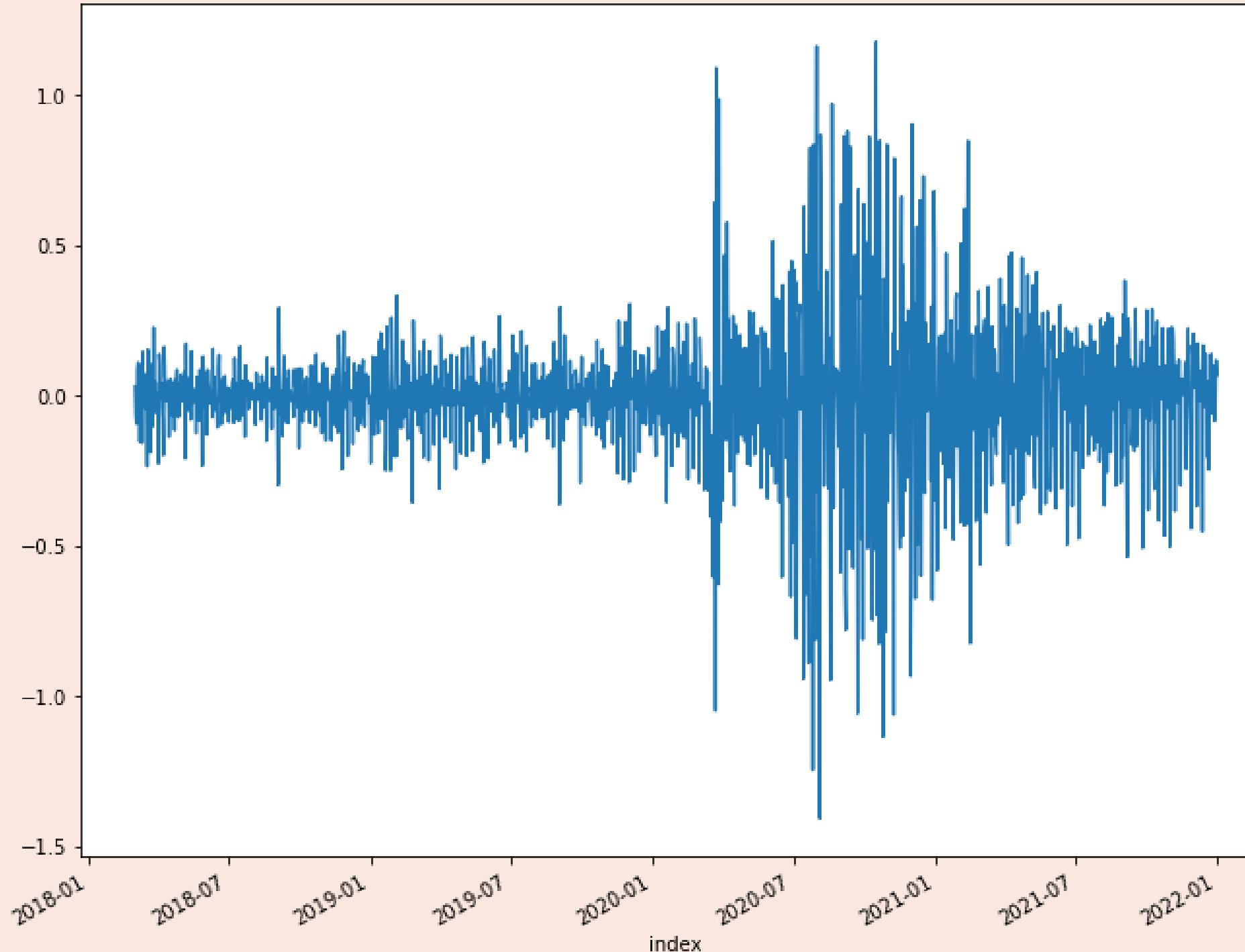
- **OCC** strongly correlated with **RevPAR**
- **OCC** inversely correlated with **ADR**
- **ADR** and **RevPAR** correlated with each other

Market Occupation Decomposition

- Seasonal peaks in summer and winter
- Strong stability pre-pandemic
- Low occupation and variance early pandemic
- Trend increasing towards pre-pandemic levels with increased variance

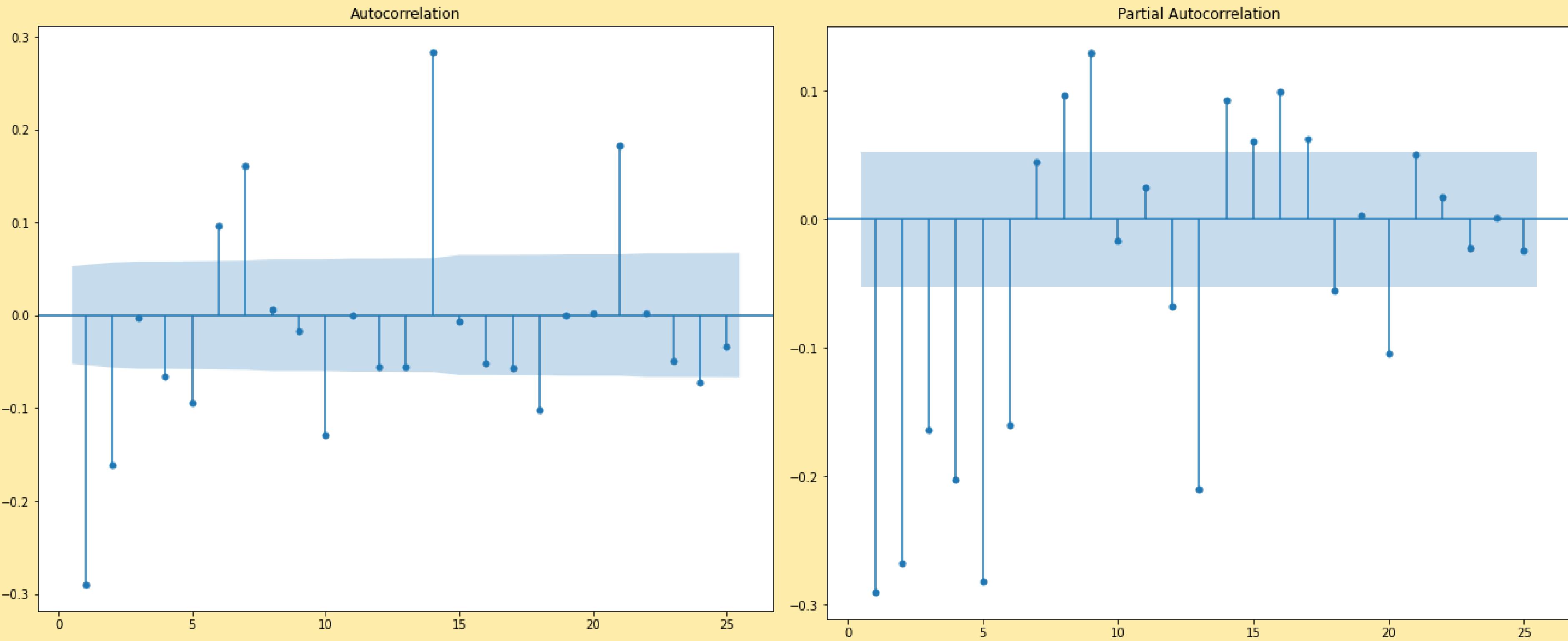


Stationarity



- Original data failed KPSS and ADF stationarity tests
- Use of log transform and first difference creates validated stationarity

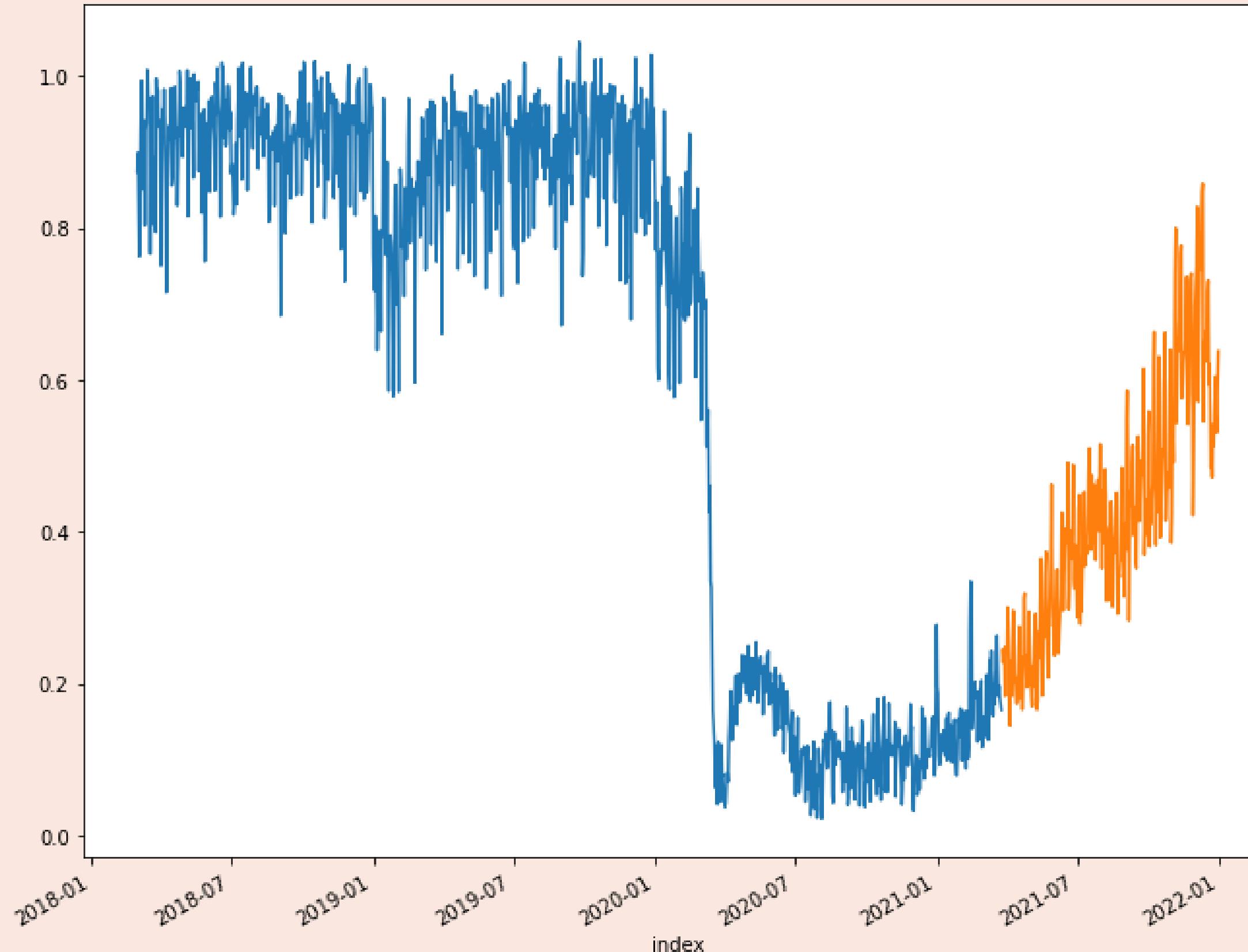
Autocorrelation



ARIMA Forecasting Model

- Baseline forecasting model
- 80/20 Train-Test Split
- Trained on log transformed data
- Order values selected using grid search process
- Evaluated on MSE
- Selected model of order (2,1,2), MSE = 0.029

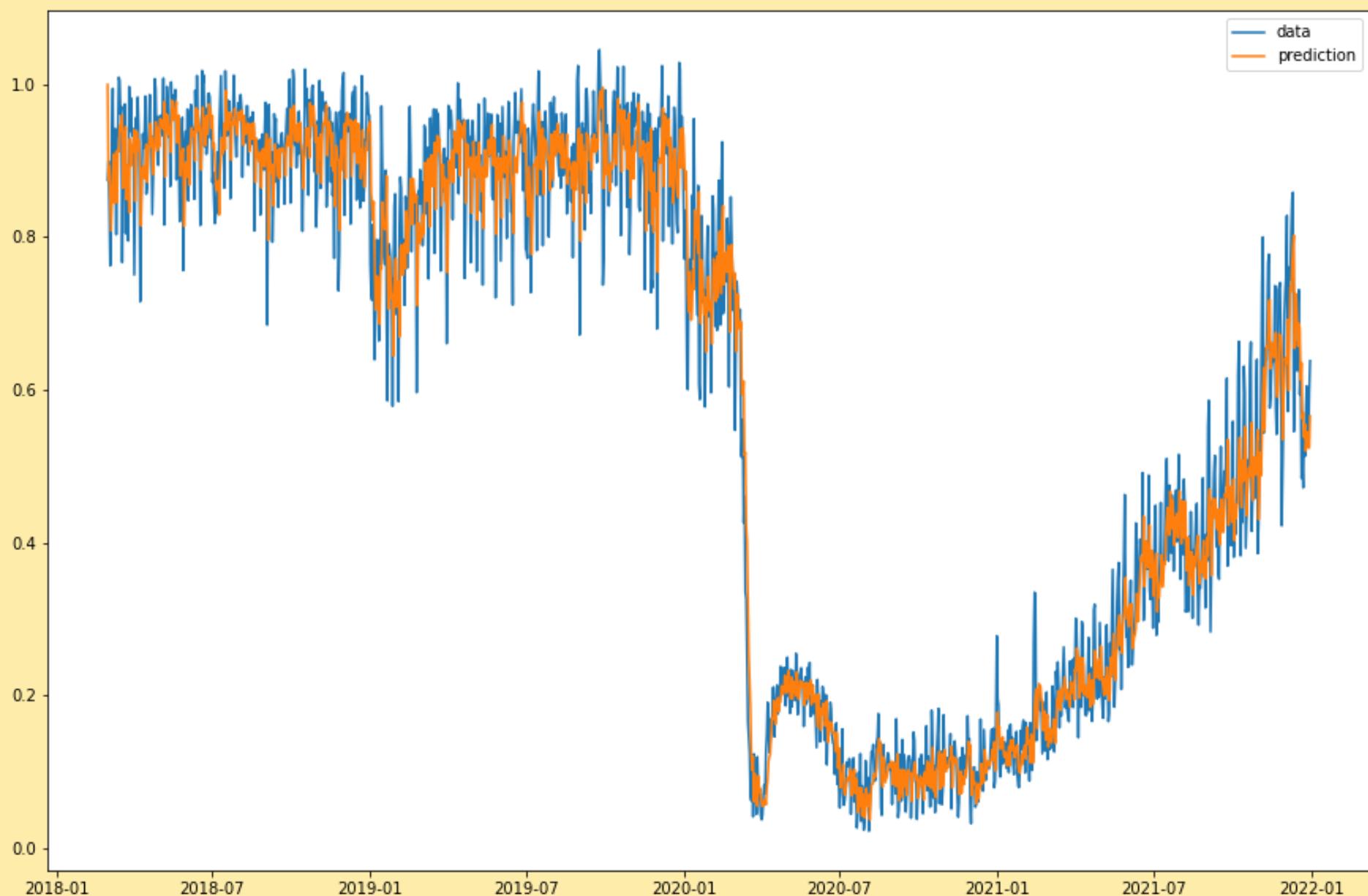
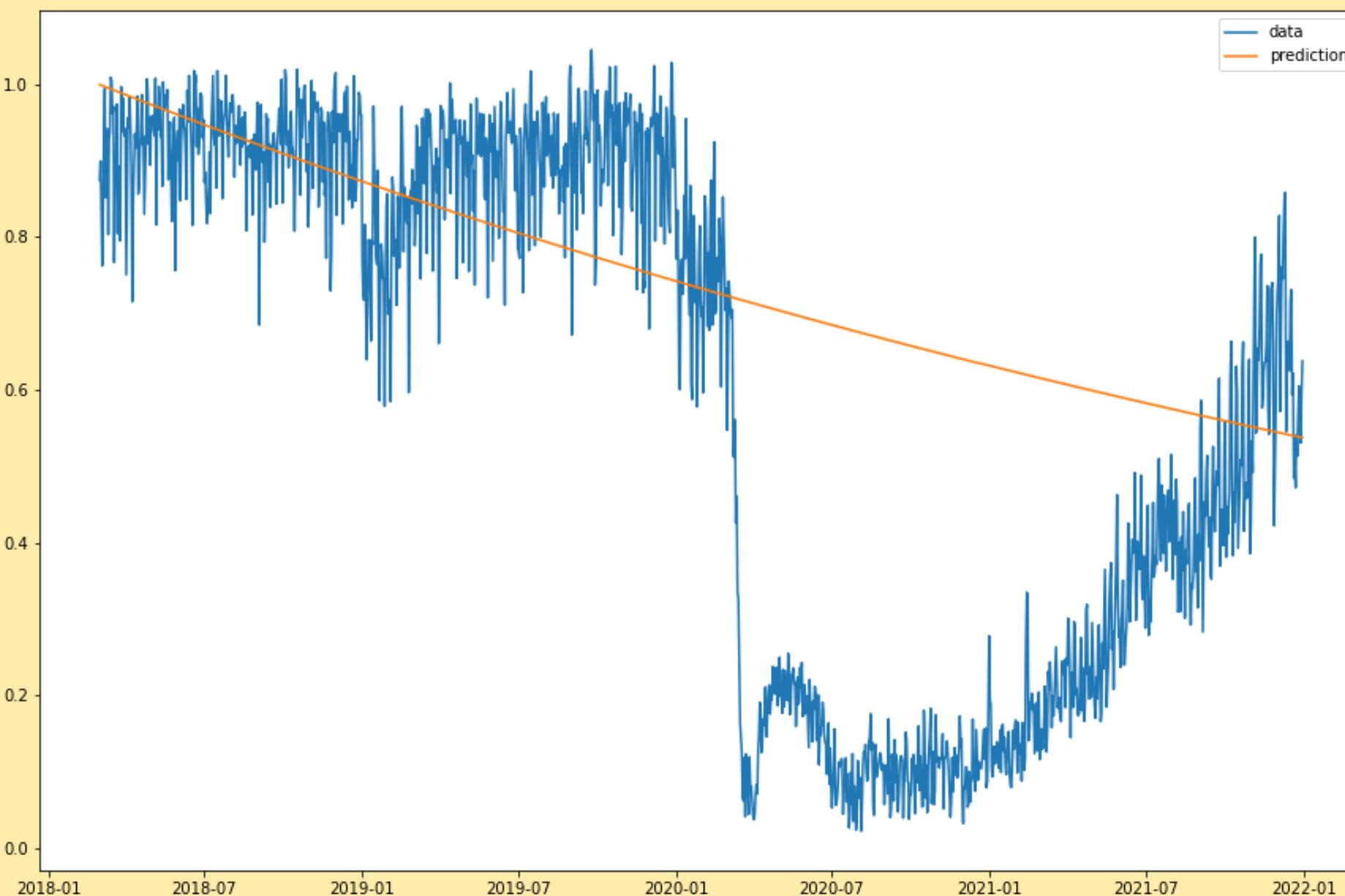
Train-Test Split



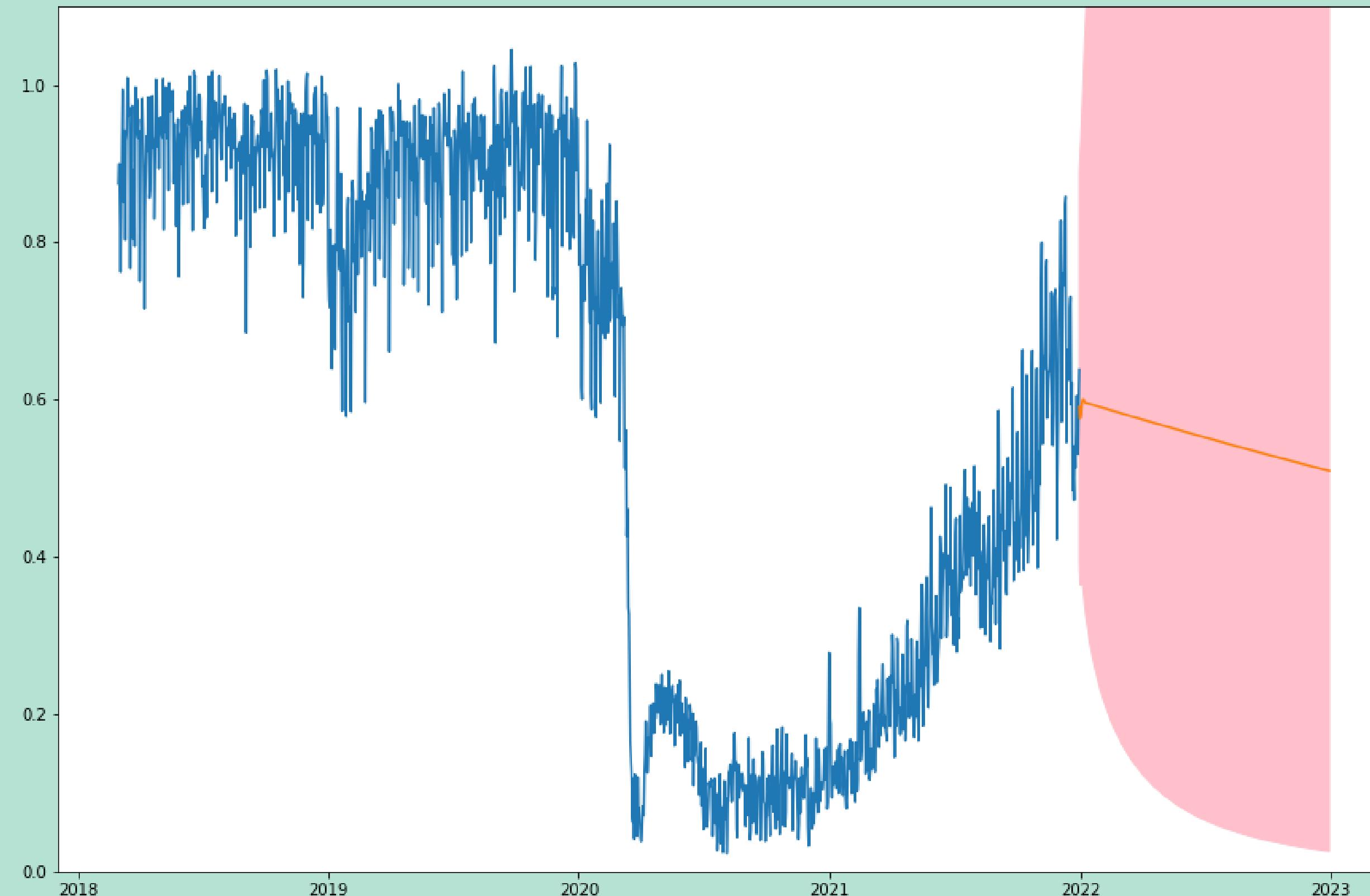
Predictions

Dynamic

Non-Dynamic

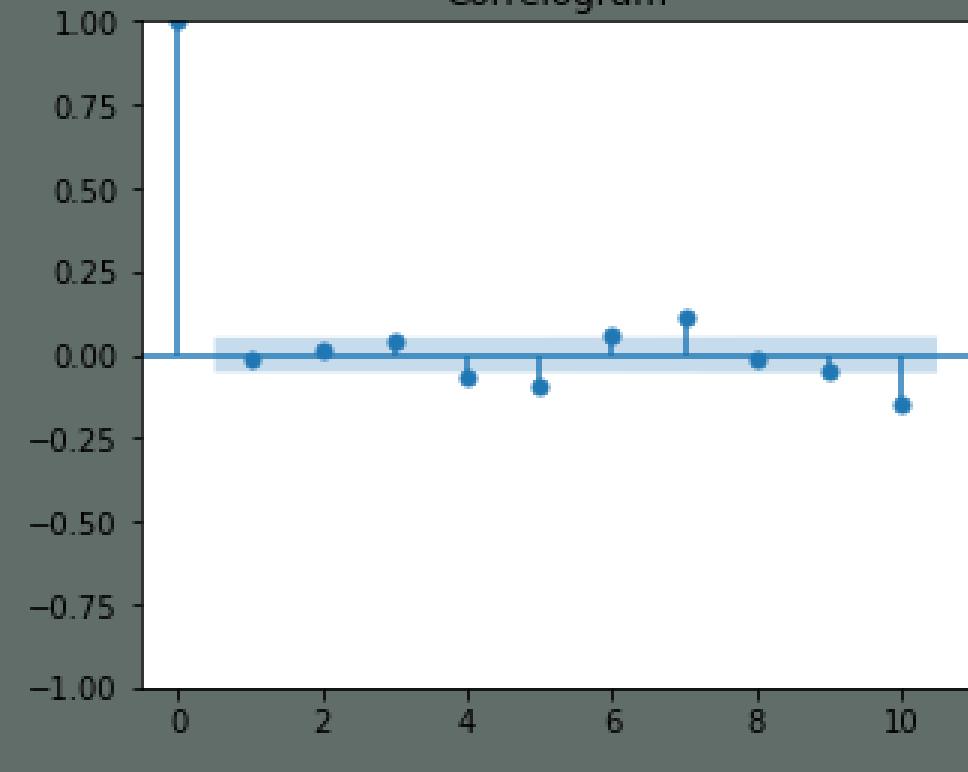
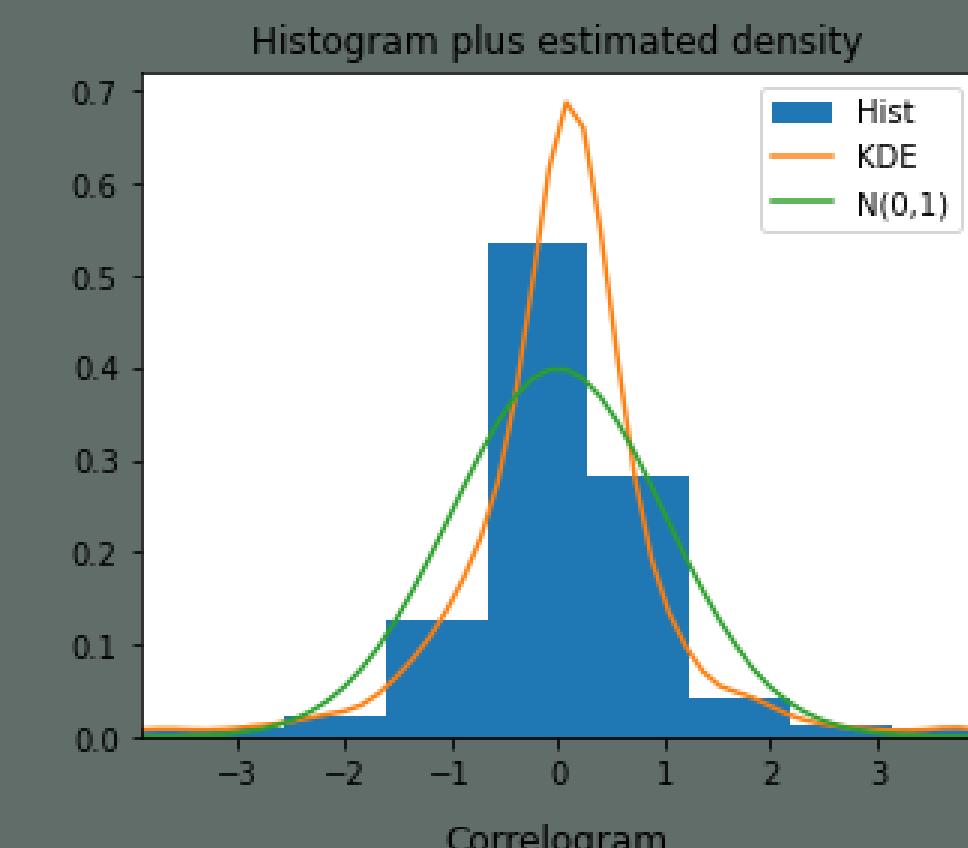
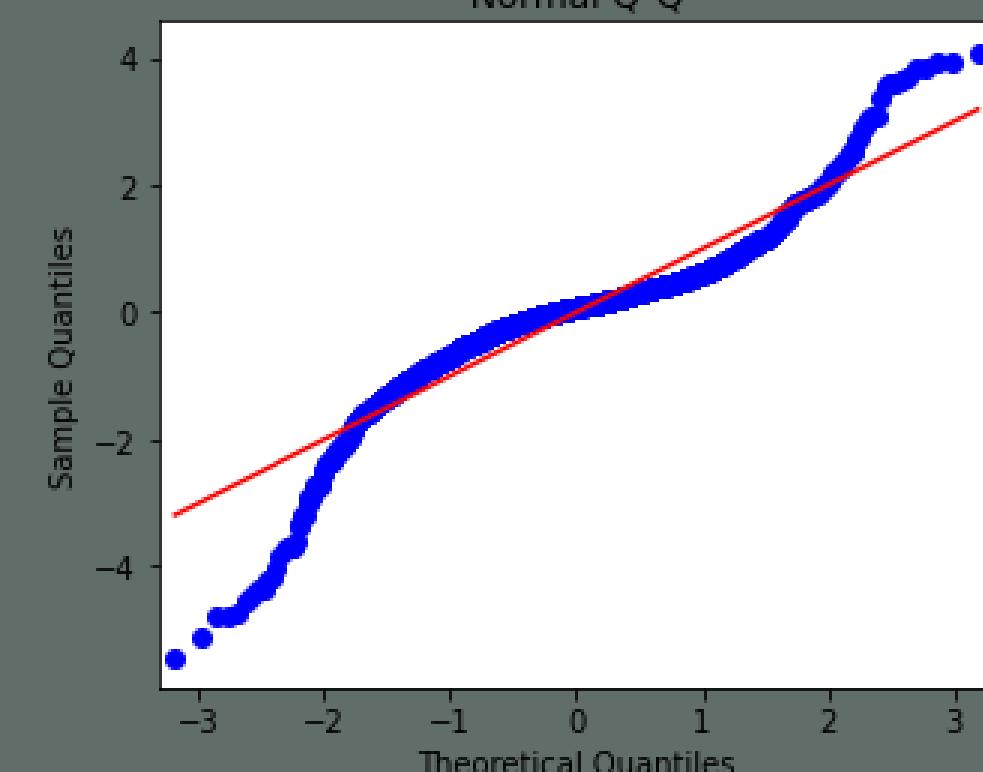
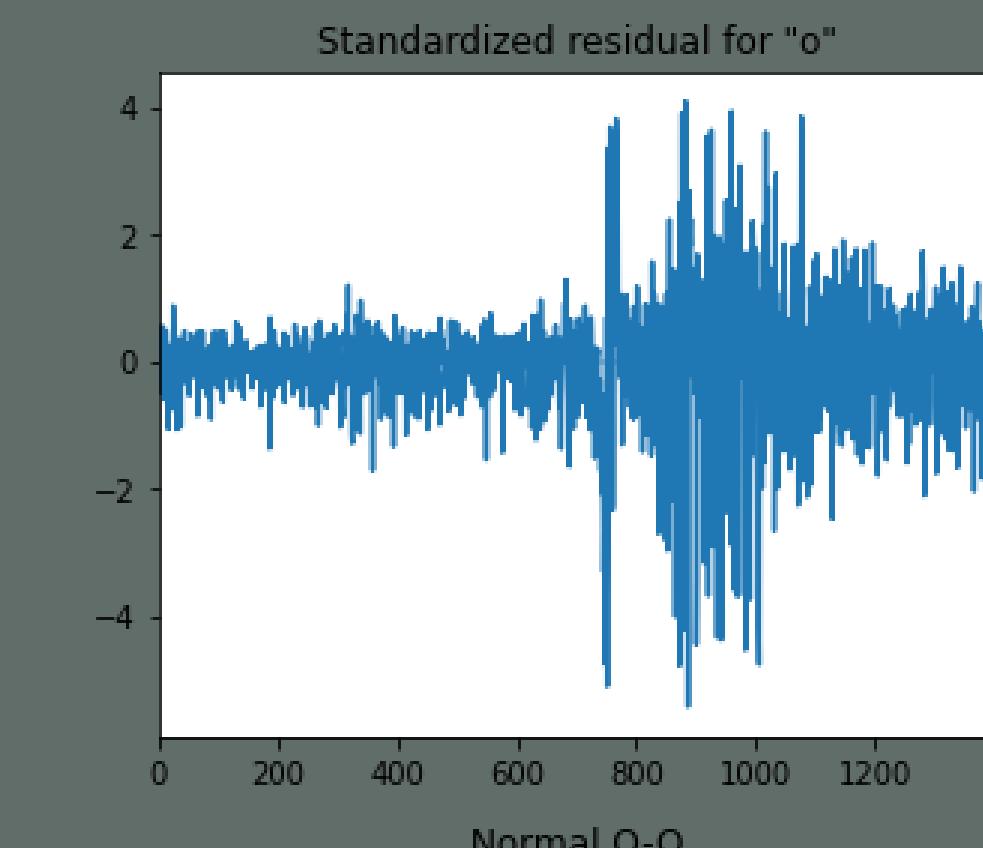


Forecasting



Diagnostics

Dep. Variable:	occ	No. Observations:	1401			
Model:	SARIMAX(2, 1, 2)	Log Likelihood	209.630			
Date:	Fri, 18 Feb 2022	AIC	-407.261			
Time:	16:31:15	BIC	-375.795			
Sample:	0	HQIC	-395.498			
	- 1401					
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
intercept	-0.0002	0.001	-0.194	0.847	-0.002	0.002
ar.L1	0.9257	0.053	17.628	0.000	0.823	1.029
ar.L2	-0.3977	0.023	-17.525	0.000	-0.442	-0.353
ma.L1	-1.4352	0.052	-27.348	0.000	-1.538	-1.332
ma.L2	0.6145	0.040	15.472	0.000	0.537	0.692
sigma2	0.0436	0.001	49.726	0.000	0.042	0.045
Ljung-Box (L1) (Q):	0.37	Jarque-Bera (JB):	2037.32			
Prob(Q):	0.54	Prob(JB):	0.00			
Heteroskedasticity (H):	9.45	Skew:	-0.75			
Prob(H) (two-sided):	0.00	Kurtosis:	8.72			





IMPROVEMENT PLAN

5 Steps for Improvement

Incorporate Seasonal
and Exogenous factors

Develop RNN model for
comparison

Prepare the selected
model for deployment

Develop GARCH
model for comparison

Evaluate and select
the best forecasting
model

