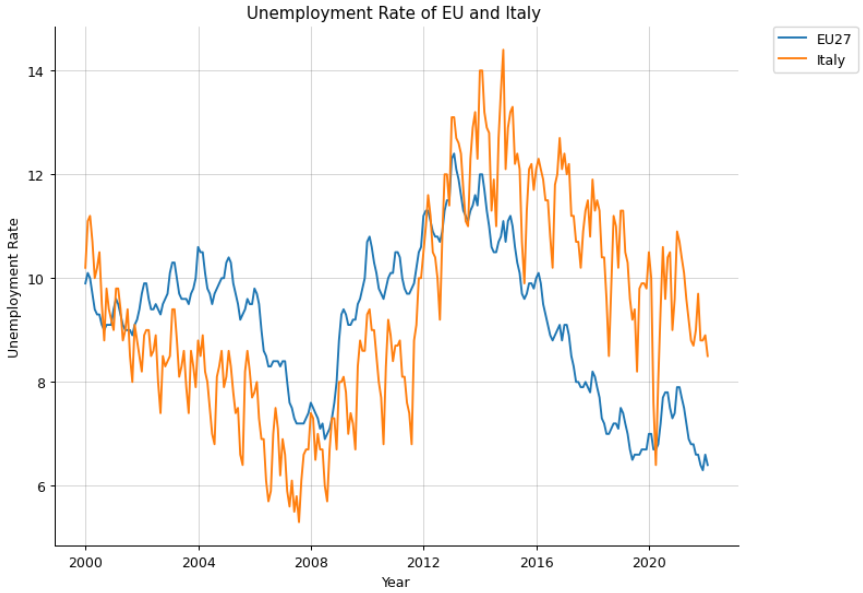
UNEMPLOYMENT RATE BETWEEN EU AND ITALY

Forecasting and projecting the rate of two geographies in order to comprehend Italy's economic status in comparison to Europe and its rank, since Italy is currently one of the nations with the largest debt in the EU. Specifically, to examine if there are any significant effects on the unemployment rate beginning in 2020 as a result of Brexit, the United Kingdom no longer being a member of the EU, and the onset of the pandemic.

Methodology: Using Python to analyze data, ARIMA is used to anticipate future time steps to make predictions and compare the results.

Results:



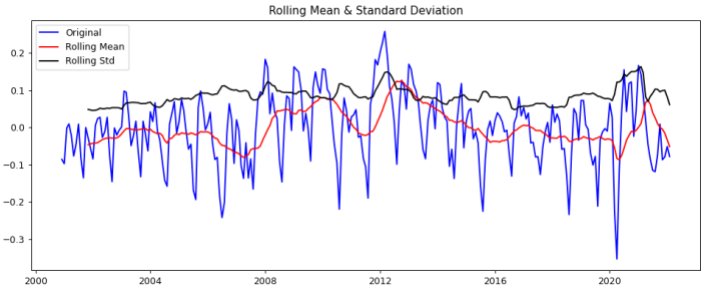
Unemployment rates of Italy and Europe (with 27 countries)
From 2000 to Jan 2022

According to the graph, Italy's unemployment rate in 2012 was much higher than that of the European Union (just 27 nations excluding the United Kingdom). Surprisingly, when Covid launched in 2020, Italy's rate fell far lower than usual while it should have been increasing cause of the situation and stay-at-home policy.

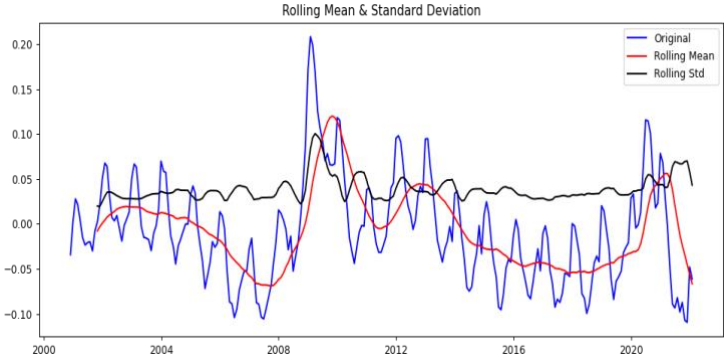
The stationarity of the data series is first seen using ACF and PACF plots, and then tested using the Statistics Dickey-Fuller test statistic. It has been discovered that Europe's and Italy's unemployment rates are nonstationary.

However, after computing the starting difference of the data series' logarithmic values, the same types of plots and statistics show that the data is stationary. Taking the log of the dependent variable also ensures that the rolling mean and standard deviation (as seen in the graphs below) are almost horizontal. The p-value is less than 0.05, and the ADF Statistic is close to critical. As a result, the time series has reached its stationary point.

ADF Statistic for IT: -2.864195
p-value: 0.049673
Critial Values:
   1%, -3.457664132155201
Critial Values:
   5%, -2.8735585105960224
Critial Values:
   10%, -2.5731749894132916



Italy's rolling mean and standard deviation



ADF Statistic for EU: -3.681559
p-value: 0.004380
Critial Values:
   1%, -3.458128284586202
Critial Values:
   5%, -2.873761835239286
Critial Values:
   10%, -2.5732834559706235

Europe's rolling mean and standard deviation

The best ARIMA model was chosen based on criteria such as BIC and RMSE. The model with the lowest criterion values is deemed the best model. It fits an ARIMA(3,1,3) model. This sets the lag value to 3 for autoregression, uses a difference order of 1 to make the time series stationary, and uses a moving average model of 3. This analysis predicts and reports values from 2000 to Jan 2022.

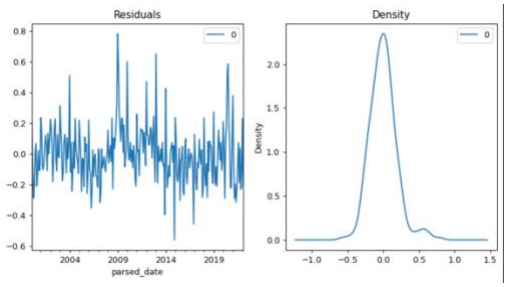
ARIMA Model Results

Dep. Variable:	D.Value	No. Observations:	265
Model:	ARIMA(3, 1, 3)	Log Likelihood	69.756
Method:	mle	S.D. of innovations	0.183
Date:	Fri, 15 Apr 2022	AIC	-123.511
Time:	08:55:03	BIC	-94.873
Sample:	02-01-2000	HQIC	-112.005
	- 02-01-2022		

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.0128	0.010	-1.334	0.183	-0.032	0.006
ar.L1.D.Value	0.7317	0.010	75.604	0.000	0.713	0.751
ar.L2.D.Value	0.7322	0.014	52.814	0.000	0.705	0.759
ar.L3.D.Value	-0.9865	0.009	-105.310	0.000	-1.007	-0.970
ma.L1.D.Value	-0.7866	0.030	-26.578	0.000	-0.845	-0.729
ma.L2.D.Value	-0.7534	0.030	-25.152	0.000	-0.812	-0.695
ma.L3.D.Value	0.9877	0.031	31.464	0.000	0.926	1.049

Roots

	Real	Imaginary	Modulus	Frequency
AR.1	-1.0030	-0.0000j	1.0030	-0.5000
AR.2	0.8718	-0.4986j	1.0043	-0.0827
AR.3	0.8718	+0.4986j	1.0043	0.0827
MA.1	-1.0122	-0.0000j	1.0122	-0.5000
MA.2	0.8875	-0.4611j	1.0001	-0.0763
MA.3	0.8875	+0.4611j	1.0001	0.0763



Europe's ARIMA

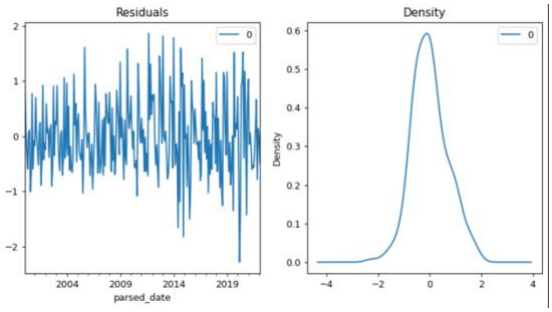
ARIMA Model Results

Dep. Variable:	D.Value	No. Observations:	265
Model:	ARIMA(3, 1, 0)	Log Likelihood	-279.808
Method:	mle	S.D. of innovations	0.695
Date:	Fri, 15 Apr 2022	AIC	569.617
Time:	08:54:49	BIC	587.516
Sample:	02-01-2000	HQIC	576.808
	- 02-01-2022		

	coef	std err	z	P>|z|	[0.025	0.975]
const	-0.0070	0.030	-0.233	0.816	-0.066	0.052
ar.L1.D.Value	-0.0734	0.061	-1.202	0.230	-0.193	0.046
ar.L2.D.Value	-0.2227	0.060	-3.741	0.000	-0.339	-0.106
ar.L3.D.Value	-0.1219	0.061	-2.004	0.046	-0.241	-0.003

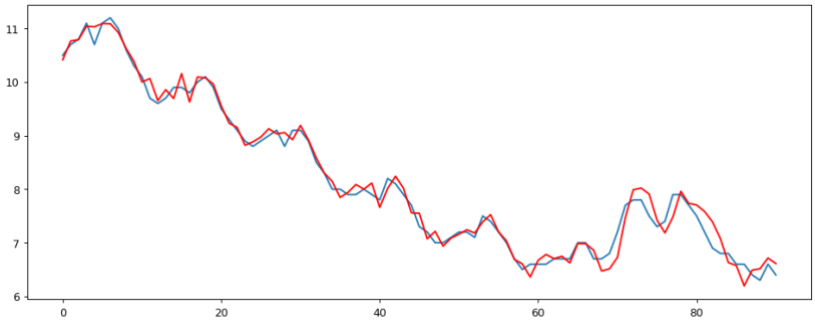
Roots

	Real	Imaginary	Modulus	Frequency
AR.1	0.4449	-1.6796j	1.7375	-0.2088
AR.2	0.4449	+1.6796j	1.7375	0.2088
AR.3	-2.7163	-0.0000j	2.7163	-0.5000

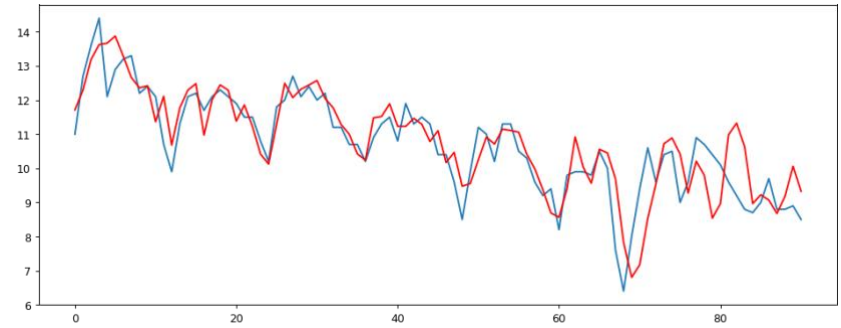


Italy's ARIMA

Next, for predicting, we split the training dataset into train and test sets, use the train set to fit the model, and generate a prediction for each element on the test set. see how the model compares to the original time series. Then, calculate a final root mean squared error score (RMSE) for the predictions, providing a point of comparison for other ARIMA configurations. And using model of forecast residual error, would be useful to make better predictions. With a good estimate of forecast error at a time step, we can make better predictions. For example, we can add the expected forecast error to a prediction to correct it and in turn improve the skill of the model. Finally, the expected values for the test dataset are plotted (blue) compared to the predicted (red).



Prediction of residual error time series of Europe
RMSE = 0.179



Prediction of residual error time series of Italy
RMSE = 0.795

As can be seen from the graphs, the actual expected value and the predicted value are almost same, except for the last part of Italy (from 80), which is somewhat different. The quantitative goal variable is the unemployment rate, and this approach may be used to get expectations prior to the release of official statistics, making them especially valuable for monitoring economic events and short-term forecasting. In conclusion, a high unemployment rate might also be troublesome if it is accompanied with low inflation, earnings, economic growth, and policies. And, since the Covid began, as well as the Brexit, it appears that neither has had a significant impact on the average unemployment rate of 27 European nations in general, and specifically Italy. Furthermore, in forecast, Italy even has a far higher unemployment rate than the rate of Europe.