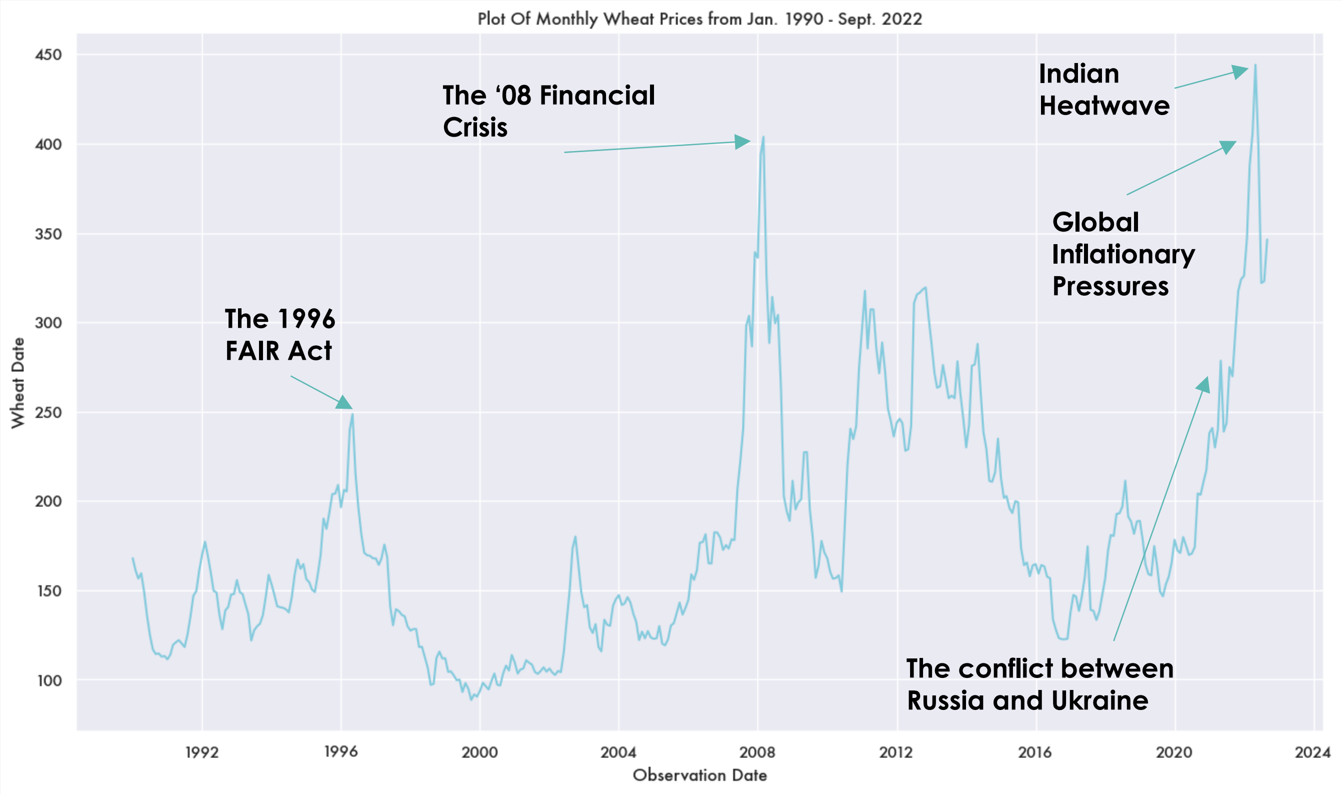
TIME SERIES ANALYSIS OF WHEAT PRICES

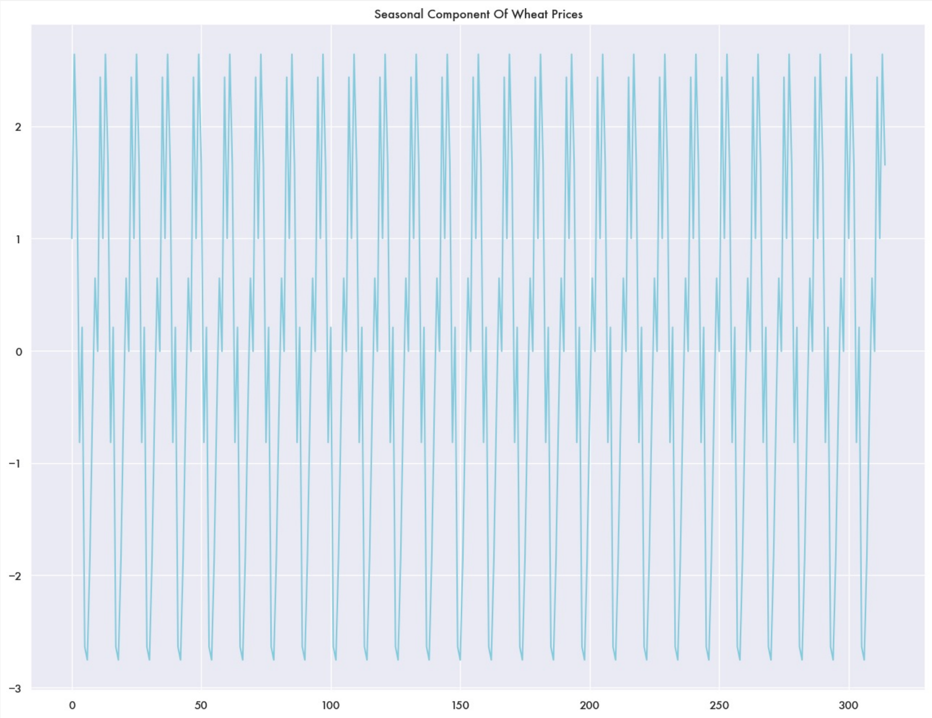
Dhruv Akshay Pandit 58802
Gabriel Dias Pereira 58454
Maria Louro Pereira 53554
Johana Pertoldová 58966

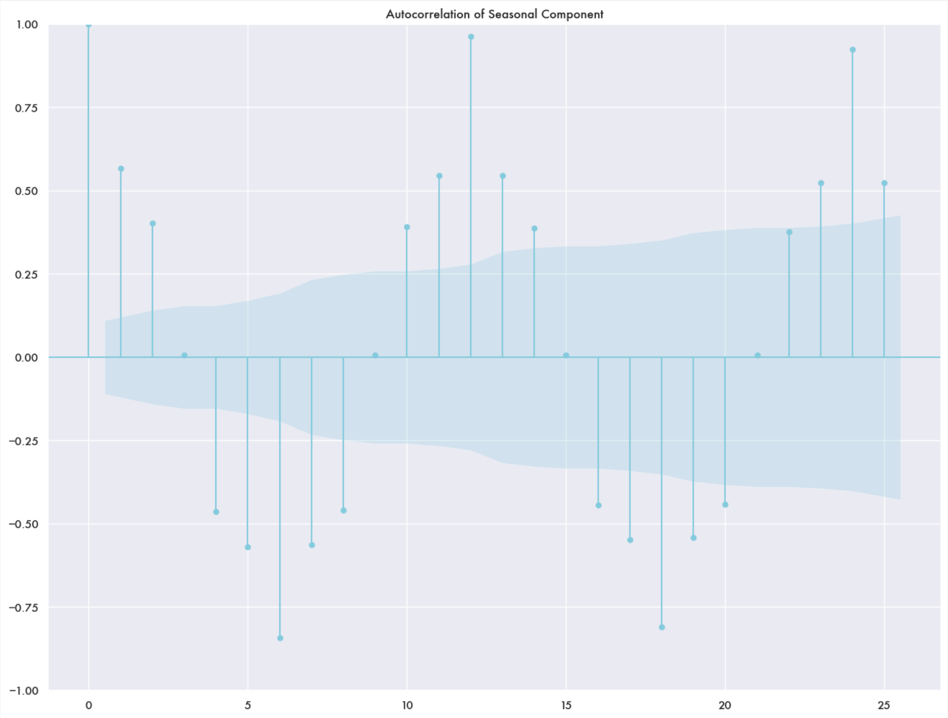
WHEAT PRICES DATASET AND OUTLIERS

- Wheat is a fundamental commodity and the most widely grown crop overall
- Frequency: Monthly
- Data: from 1990 to 2022, (393 observations)
- Prices are period averages in nominal U.S. dollars
- Large volume of export
- Units: U.S. Dollars per Metric Ton, Not Seasonally Adjusted



SEASONALITY OF DATA



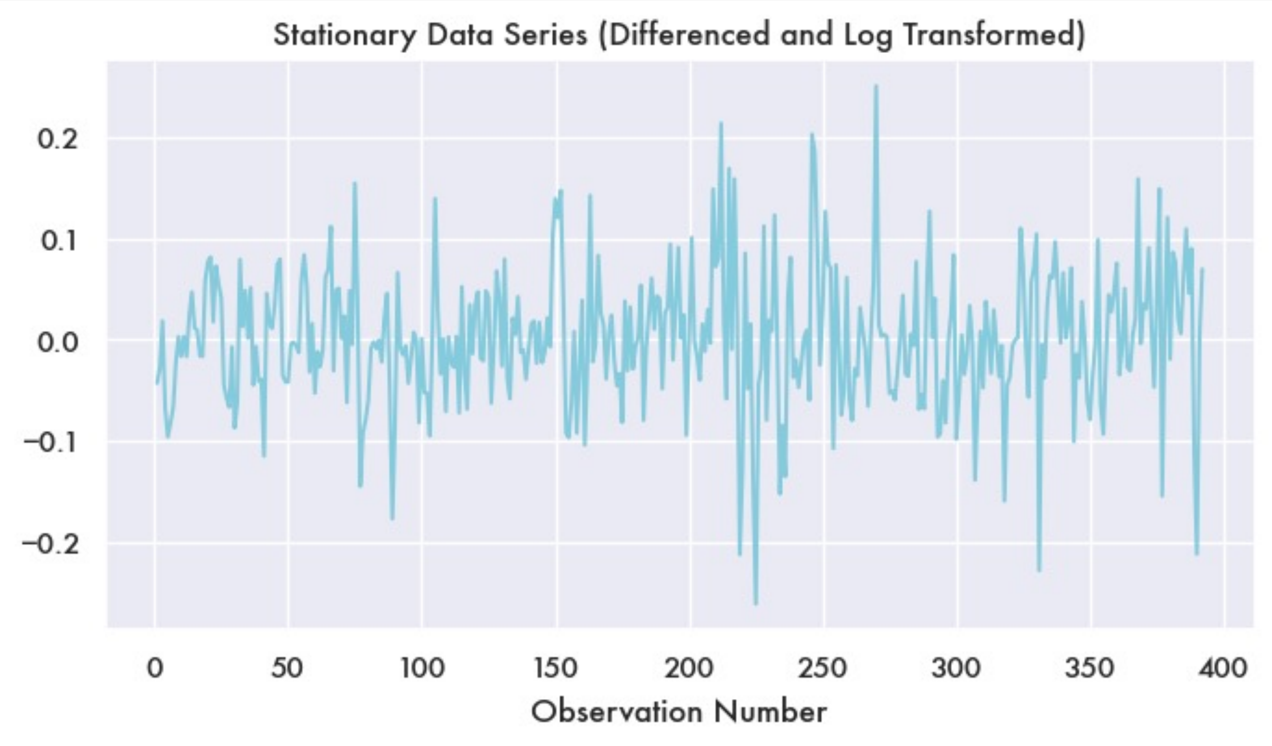


- Possible seasonal component present
- Wheat prices could be related to the harvest season

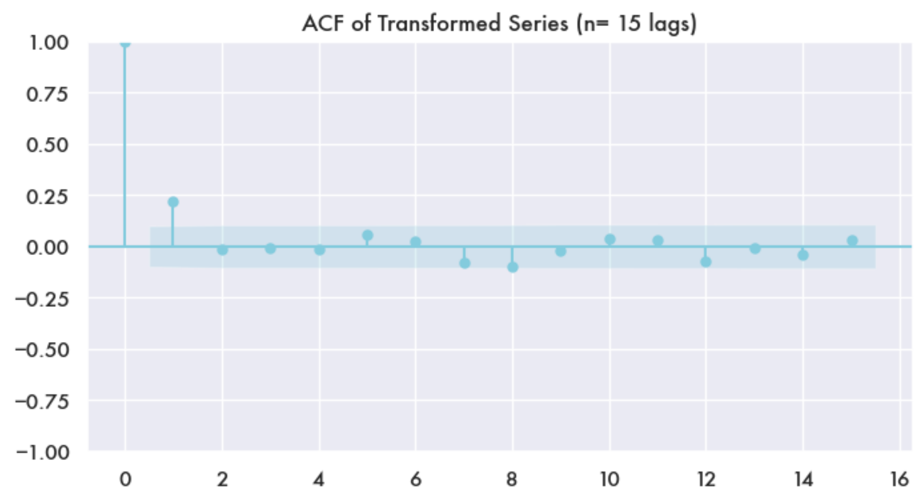
- ACF of seasonal component repeats after n lags
- Plot shows us that seasonality of data is 12 (n = 12)
- This makes sense since our dataset is monthly

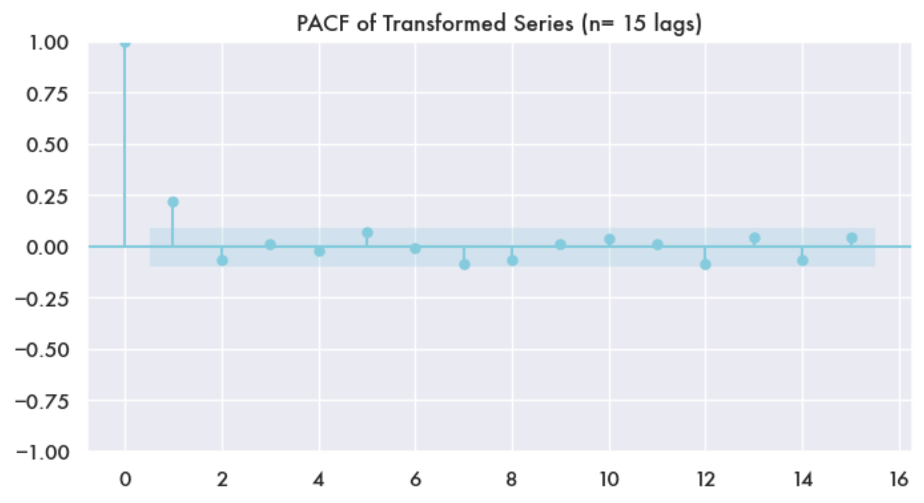
STATIONARITY OF DATA

- Define function within python to check stationarity
- Intitial data is non-stationary:
  - ADF test statistic: -2.271
  - P-value: 0.182
- Stabilise variance with log transformation and take first difference of obtained series
- Transformed data is stationary:
  - ADF test statistic:-15.727063808321123
  - P-value: 1.3029254061006328e-28



ACF AND PACF OF TRANSFORMED DATA





- ACF of transformed data cuts off after lag 1
- Process could be MA (1), with seasonality of 12

- PACF of transformed data cuts off after lag 1
- Process could be AR (1), with seasonality of 12

OUR MODEL SELECTION

- Import auto_arima from pmdarima to select optimal SARIMAX model
- Define seasonal component to be 12 (m = 12)
- Results show that a SARIMAX model with order = (1, 1, 1) and seasonal order = (1, 0 , 1, 12) with AIC of 3217.431 is the best model for the given data
- Model is fitted using the above parameters, and trained on the training set, defined for 80% of the observations

```
from pmdarima import auto_arima
auto_arima(df_wheat["Wheat_Price"],  m = 12, Trace = True, seasonal=True).summary()
```

[134]

SARIMAX Results

Dep. Variable:	y	No. Observations:	393
Model:	SARIMAX(1, 1, 1)x(1, 0, 1, 12)	Log Likelihood	-1603.716
Date:	Fri, 16 Dec 2022	AIC	3217.431
Time:	15:34:49	BIC	3237.287
Sample:	0	HQIC	3225.301
	- 393		
Covariance Type:	opg		

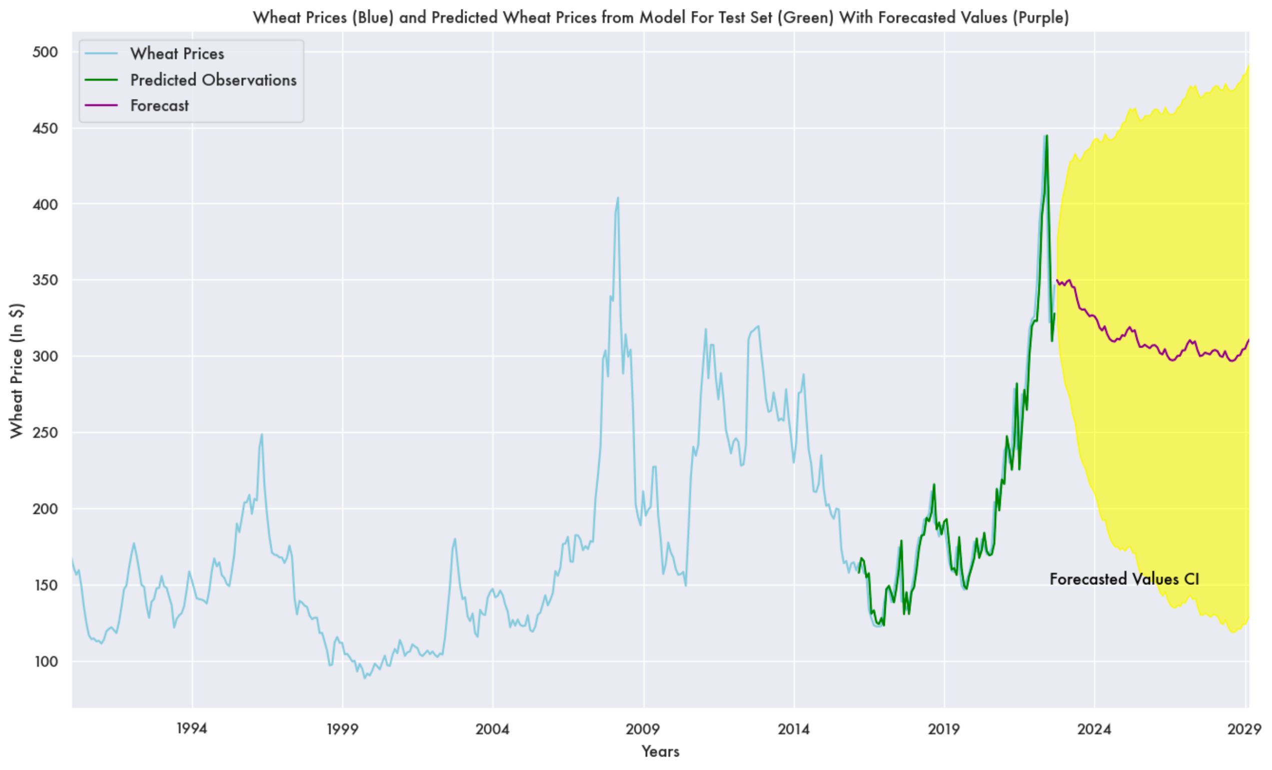
	coef	std err	z	P>|z|	[0.025	0.975]
ar.L1	-0.3610	0.118	-3.067	0.002	-0.592	-0.130
ma.L1	0.5956	0.096	6.215	0.000	0.408	0.783
ar.S.L12	0.6547	0.226	2.902	0.004	0.212	1.097
ma.S.L12	-0.7589	0.202	-3.759	0.000	-1.155	-0.363
sigma2	209.0846	8.134	25.705	0.000	193.142	225.027

Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):	470.75
Prob(Q):	0.96	Prob(JB):	0.00
Heteroskedasticity (H):	3.98	Skew:	0.05
Prob(H) (two-sided):	0.00	Kurtosis:	8.37

Warnings:
[1] Covariance matrix calculated using the outer product of gradients (complex-step).

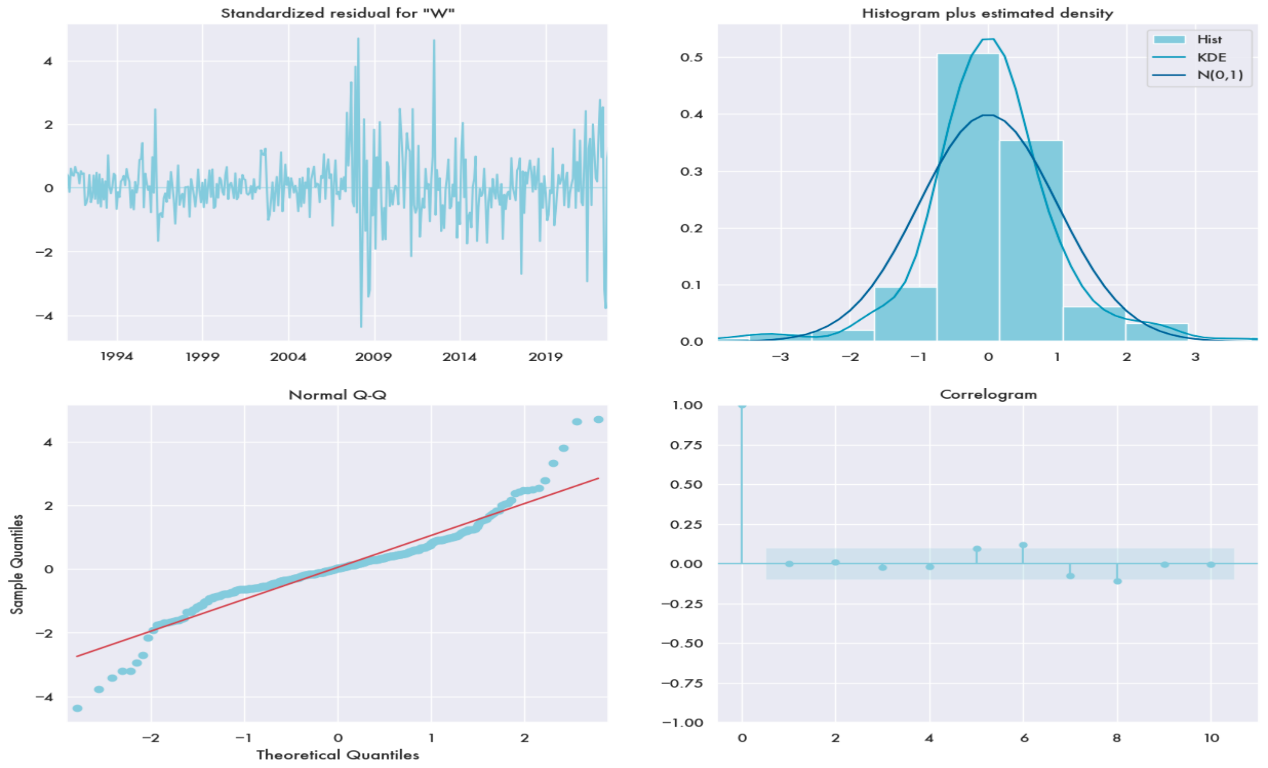
FORECASTING WHEAT PRICES USING THE MODEL

- Prediction shows a sufficient fit for our predicted values in green
- Our forecasted values for the wheat prices until 2029 exhibit a downward trend, with local seasonal peaks
- Confidence interval for our forecasted values is given in yellow
- RMSE obtained for the model is: 17.61



MODEL DIAGNOSTICS

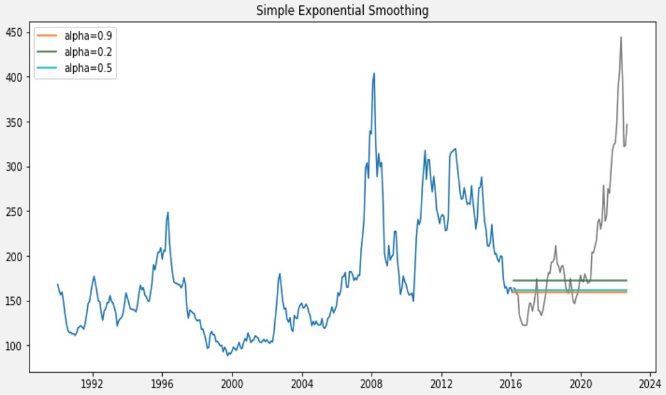
- Our fitted model diagnostics are shown on the right
- The quantile-quantile (QQ) plot on the bottom left shows us that our data is not normal distributed
- The correlogram on the bottom right shows there is no autocorrelation in the residuals
  - Uncorrelated Residuals with mean close to zero

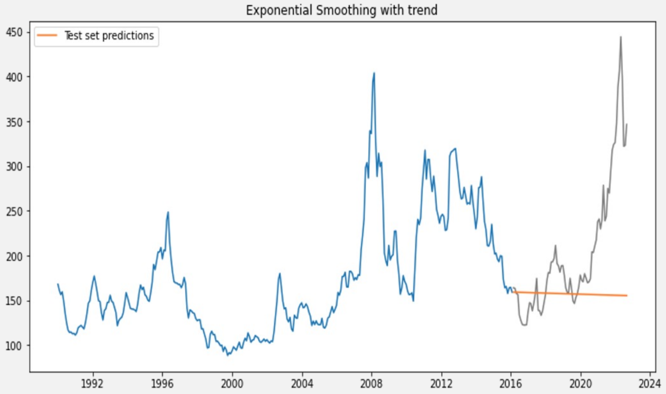


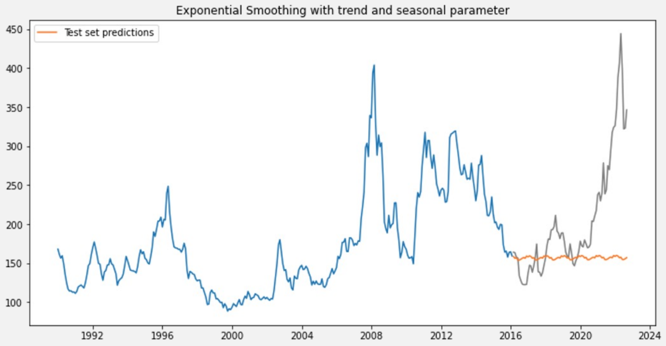
EXPONENTIAL SMOOTHING METHODS

- We estimated three different exponential methods: simple, with trend, and with trend and seasonal components
- We used the 'statsmodels' python package
- We trained and analyzed the models with the same time intervals for training and test set as in SARIMA (so, both results are comparable)

Exponential method	RMSE
Simple exponential	80.32 (alpha=0.2)
W/ trend (Holt)	87.91
W/ trend and seasonal (Winters)	87.13







FORECASTING (EXPONENTIAL SMOOTHING)

- We decided to use the Holt-Winters method to forecast future values
- We chose the best seasonal method with the train-test approach, then, trained the model again to predict the next 24 months
- The best fitted model was using additive seasonal method
- The 24-month forecast shows an upward trend (in opposite to the downward trend of the SARIMAX output), with seasonal cycles

- Forecast next values with Holt Winters method

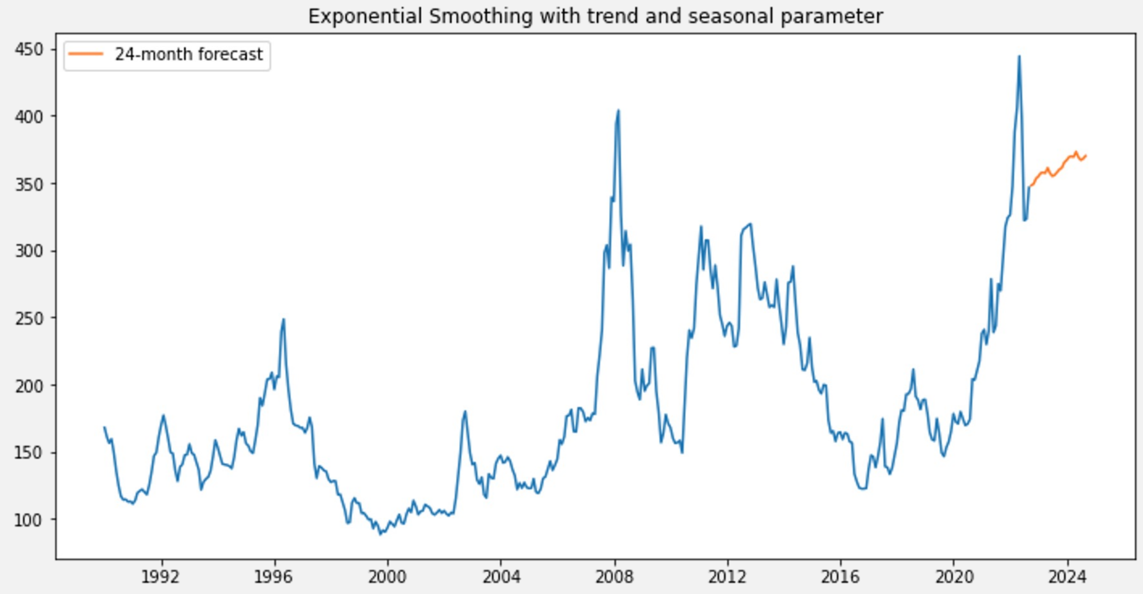
```python
# Holt Winters
holt = ExponentialSmoothing(
    df_wheat["wheat_price"],
    trend='add',
    seasonal="add",
    seasonal_periods=12
).fit()
holt_forecasts = holt.forecast(24).rename("Holt Winters Forecast")

# Plotting the forecasted values
fig, ax = plt.subplots(figsize=(12, 6))
ax.plot(df_wheat["observation_date"], df_wheat["wheat_price"])
ax.plot(holt_forecasts.index, holt_forecasts, label="24-month forecast", color='#ff7823')

plt.title("Exponential Smoothing with trend and seasonal parameter")

plt.legend();
```

✓ 0.1s



MODEL COMPARISONS

SARIMAX

- The SARIMA model showed better fitting results out-of-sample in comparison to the exponential smoothing methods
- Showed root mean squared error of 17.61
- Predicted values closely follow observations from test dataset
- Forecasted values exhibit a downward trend

EXPONENTIAL SMOOTHING

- Exponential smoothing model demonstrated poorer fitting
- Lowest root mean squared error of 80.32, while chosen mothed showed RMSE of 87.91
- Predicted values differ from observations
- Forecasted values exhibit an upward trend

FUTURE CONSIDERATIONS

- Addition of exogenous variables to the dataset to train the SARIMAX model could improve the foracasted values
  - Possible variables could be oil prices, inflation data and the USD value
- Expand dataset from monthly to daily
- Investigate other methods of forecasting data, perhaps implementing machine learning

THANK YOU

Questions?

Dhruv, Gabriel, Maria, and Johana

REFERENCES

- Crain, Susan J. and Lee, Jae Ha, Volatility in Wheat Spot and Futures Markets, 1950-1993: Government Farm Programs, Seasonality, and Causality. J. OF FINANCE, Vol. 51 No.1, March 1996, Available at SSRN: https://ssrn.com/abstract=7123
- Ghoshray, A. (2010) A reexamination of trends in primary commodity prices, Journal of Development Economics. North-Holland. Available at: https://www.sciencedirect.com/science/article/pii/S0304387810000386
- International Monetary Fund, Global price of Wheat, retrieved from FRED, Federal Reserve Bank of St. Louis; https://fred.stlouisfed.org/series/PWHEAMTUSDM, November 22, 2022.
- Jian Yang , Michael S. Haigh & David J. Leatham (2001) Agricultural liberalization policy and commodity price volatility: a GARCH application, Applied Economics Letters, 8:9, 593-598, DOI: 10.1080/13504850010018734
- Prem, K. (2018) Price elasticity of supply of paddy, wheat, cotton and mustard. Economic & Political Weekly. Available at: https://www.researchgate.net/figure/Price-Elasticity-of-Supply-of-Paddy-Wheat-Cotton-and-Mustard_tbl2_329586223
- (2022) Wheat crisis: Wheat prices increased in international market, Ground Report. Available at: https://groundreport.in/wheat-crisis-wheat-prices-increased-in-international-market/ (Accessed: November 23, 2022).