Question 1

a) The plot of the coffee price index can be seen in Figure 1 in the appendix. There is an obvious upward trend, but seasonality cannot be observed from the time series plot.

b) In Figure 2, AC is high at all lags, which is a strong sign of a significant trend in the data. In PACF plot, there is a spike with high coefficient at lag 1 and a spike with relatively less and negative coefficient at lag 2. By only looking at PACF plot, I would say that the data is coming from AR(2) model. In Figure 3, it can be inferred that the detrended data is coming from either AR(2) or MA(2) distribution.

c) MAE of AR(2) model with detrending: 2.70268
MAPE of AR(2) model with detrending: 0.03528
RMSE of AR(2) model with detrending: 3.94290

d) MAE of MA(2) model with detrending: 2.71354
MAPE of MA(2) model with detrending: 0.03550
RMSE of MA(2) model with detrending: 3.93859

From Figure 4 and 6, it can be seen all coefficients except intercepts are significant for both models. MA(2) model has slightly less AIC and BIC values. It also has less RMSE than AR(2). Therefore, MA(2) looks like a better model for this data.

e) MAE of AR(2) model with detrending on test set: 12.27281
MAPE of AR(2) model with detrending on test set: 0.12987
RMSE of AR(2) model with detrending on test set: 14.13357

MAE of MA(2) model with detrending on test set: 12.14096
MAPE of MA(2) model with detrending on test set: 0.12811
RMSE of MA(2) model with detrending on test set: 13.97662

I split the data with 75-25 split for training and test sets. Test set errors of the models trained with training sets are given above.

f)

Method	Spec.	RMSE (Train)	MAPE (Test)
Benchmark (HW1)	-	4.14	0.036
ARIMA(2,1,0)	$\phi_1 = 0.2079$ $\phi_2 = 0.1099$	3.943	0.130
ARIMA(0,1,2)	$\theta_1 = 0.2083$ $\theta_2 = 0.1543$	3.939	0.128

Question 2

a) Time series plot can be seen in Figure 8.

b) Results of SARIMA model can be seen in Figure 9. Since we only detrend and deseasonalized, there is only intercept, which has a p-value of 0.963. Therefore, only detrending and deseasonalizing is definitely not enough.

c) Corresponding plots can be seen in Figure 10, 11 and 12.

d) I used AR(2) with detrending & deseasonalizing and AR(2)+MA(3) with detrending & deseasonalizing. The corresponding SARIMA results can be seen in Figure 13 and 14.

MAE of AR(2) model with detrending and deseasonalizing: 14.05637
MAPE of AR(2) model with detrending and deseasonalizing: 0.03323
RMSE of AR(2) model with detrending and deseasonalizing: 18.36874

MAE of AR(2)&MA(3) model with detrending and deseasonalizing: 13.18680
MAPE of AR(2)&MA(3) model with detrending and deseasonalizing: 0.03104
RMSE of AR(2)&MA(3) model with detrending and deseasonalizing: 16.57181

e) I used AR(2) & MA(3) model since it has a lower RMSE value.

One-quarter ahead forecast for quarter 155: 403.25083
95 percent confidence interval for the forecast: (370.771, 435.731)

f) MAE of AR(2) model with detrending and deseasonalizing on test set: 13.72438
MAPE of AR(2) model with detrending and deseasonalizing on test set: 0.03371
RMSE of AR(2) model with detrending and deseasonalizing on test set: 18.33355

MAE of AR(2)&MA(3) model with detrending and deseasonalizing on test set: 12.62688
MAPE of AR(2)&MA(3) model with detrending and deseasonalizing on test set: 0.03106
RMSE of AR(2)&MA(3) model with detrending and deseasonalizing on test set: 16.85520

I split the data with 75-25 split for training and test sets. Test set errors of the models trained with training sets are given above.

g)

Method	**Spec.**	**RMSE (Train)**	**MAPE (Test)**
Benchmark	-	-	-
SARIMA(2,1,0)(0,1,0,4)	ϕ_1 = -0.9534 ϕ_2 = -0.5409	18.369	0.034
SARIMA(2,1,3)(0,1,0,4)	ϕ_1 = -0.4025 ϕ_2 = -0.6638 θ_1 = -0.7007 θ_2 = 0.8461 θ_3 = -0.8434	16.571	0.031

Question 3

a) OLS results of dummies can be seen in Figure 15. R^2 is 0.21 and MSE is 7476.83910.

b) OLS results of dummies and trend term can be seen in Figure 16. R^2 is 0.782 and MSE is 2062.08204. R^2 has improved significantly and MSE decreased a lot. Therefore, the model with trend term explains the variance of the data better which means the data has trend.

c) **For the OLS without trend:**
One-quarter ahead forecast for quarter 155: 377.52105
95 percent confidence interval for the forecast: (208.045, 546.997)

For the OLS with trend:

One-quarter ahead forecast for quarter 155: 506.64019
95 percent confidence interval for the forecast: (417.638, 595.642)

d)

Method	**Coefficients**	**R^2**	**MSE**
Model without trend	M1 = -67.33 M2 = -115.38 M3 = -102.25	0.21	7476.839
Model with trend	t = 1.6554 M1 = -65.67 M2 = -115.38 M3 = -100.59	0.782	2062.082

Appendix

Figure 1: Time series plot of coffee price index

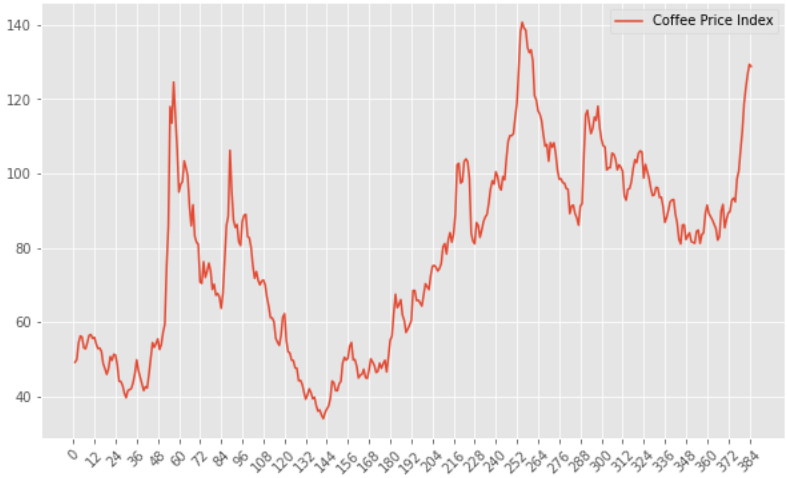

Figure 2: ACF and PACF plots of coffee price index

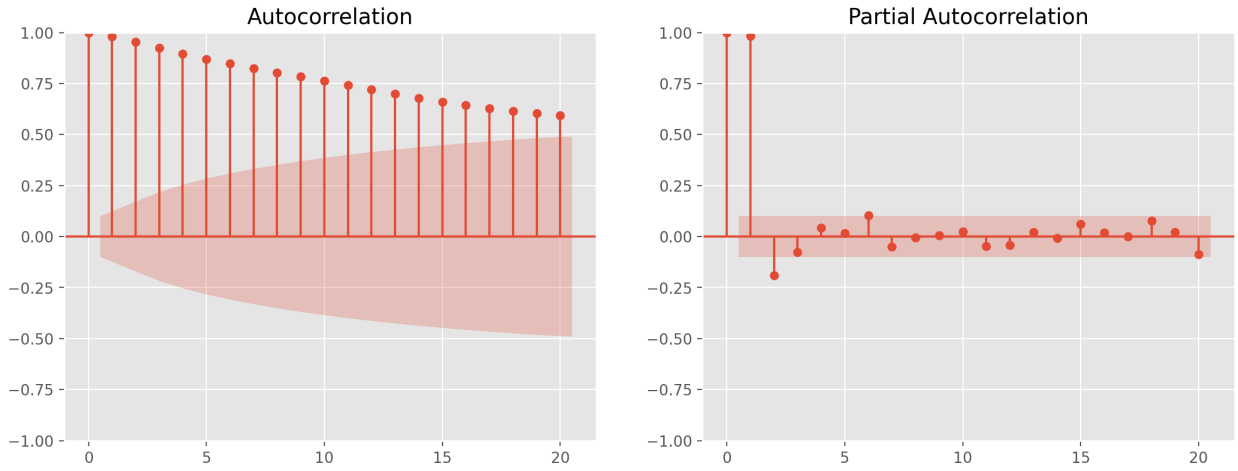

Figure 3: ACF and PACF plots of detrended data

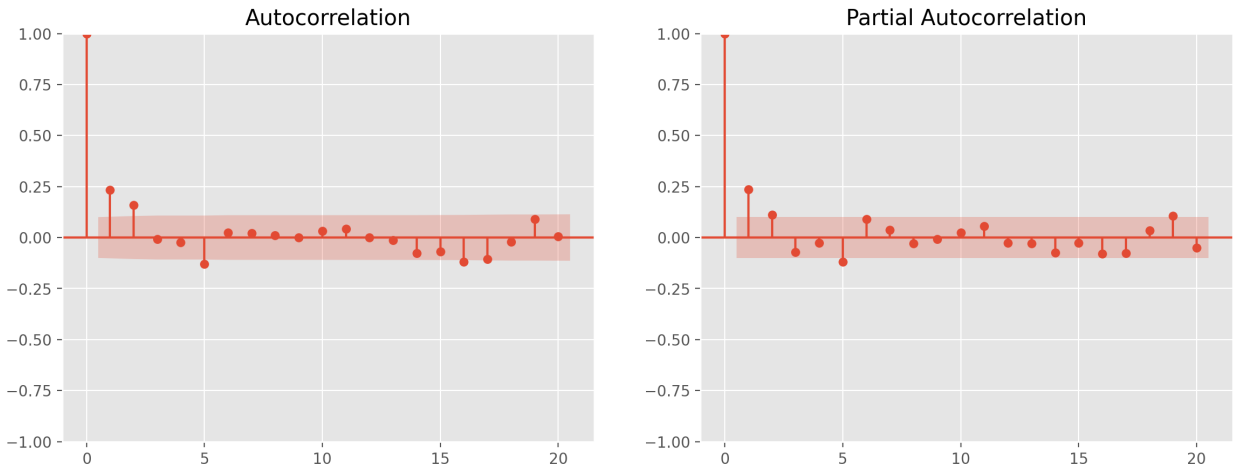

Figure 4: AR(2) results of detrended data

SARIMAX Results

Dep. Variable:	price	**No. Observations:**	385
Model:	SARIMAX(2, 1, 0)	**Log Likelihood**	-1071.722
Date:	Mon, 04 Apr 2022	**AIC**	2151.443
Time:	21:33:35	**BIC**	2167.246
Sample:	0	**HQIC**	2157.711
	- 385		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	0.1432	0.233	0.615	0.539	-0.313	0.600
ar.L1	0.2079	0.028	7.529	0.000	0.154	0.262
ar.L2	0.1099	0.035	3.109	0.002	0.041	0.179
sigma2	15.5459	0.630	24.686	0.000	14.312	16.780

Ljung-Box (L1) (Q):	0.03	**Jarque-Bera (JB):**	1124.85
Prob(Q):	0.87	**Prob(JB):**	0.00
Heteroskedasticity (H):	0.41	**Skew:**	0.93
Prob(H) (two-sided):	0.00	**Kurtosis:**	11.17

Figure 5: AR(2) model vs. real data

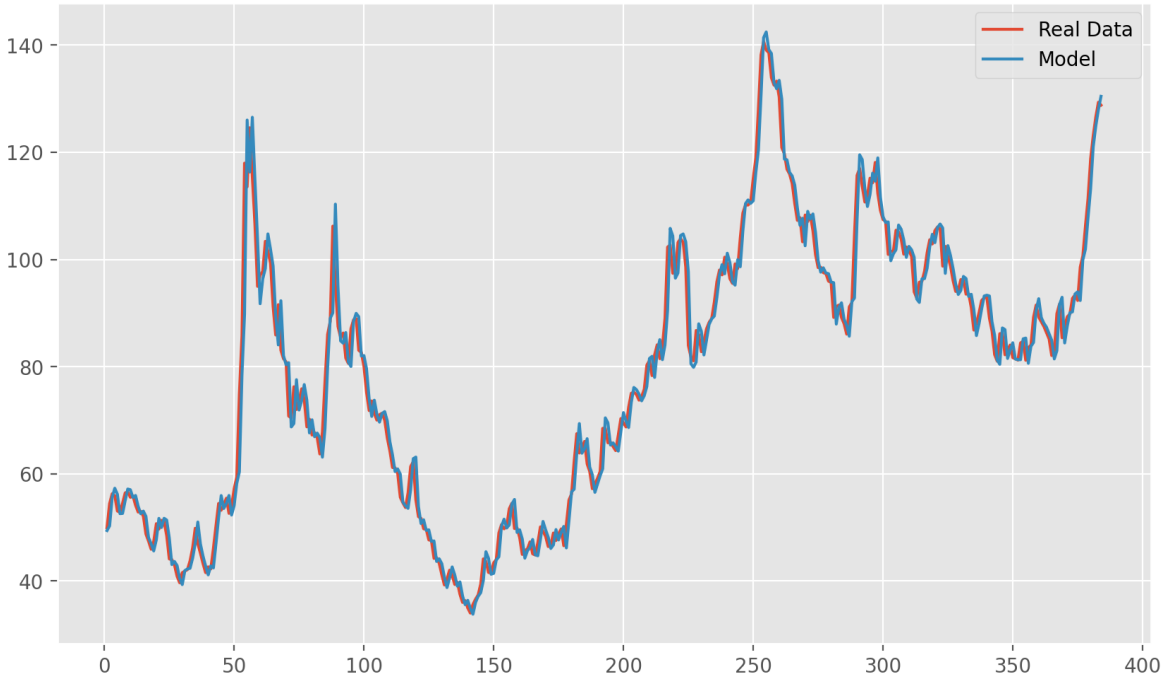

Figure 6: MA(2) results of detrended data

SARIMAX Results

Dep. Variable:	price	**No. Observations:**	385
Model:	SARIMAX(0, 1, 2)	**Log Likelihood**	-1071.301
Date:	Mon, 04 Apr 2022	**AIC**	2150.603
Time:	21:33:35	**BIC**	2166.405
Sample:	0	**HQIC**	2156.871
	- 385		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	0.2076	0.306	0.679	0.497	-0.392	0.807
ma.L1	0.2083	0.030	6.949	0.000	0.150	0.267
ma.L2	0.1549	0.037	4.139	0.000	0.082	0.228
sigma2	15.5119	0.663	23.405	0.000	14.213	16.811

Ljung-Box (L1) (Q):	0.01	**Jarque-Bera (JB):**	1148.32
Prob(Q):	0.92	**Prob(JB):**	0.00
Heteroskedasticity (H):	0.42	**Skew:**	1.01
Prob(H) (two-sided):	0.00	**Kurtosis:**	11.23

Figure 7: Real data vs. MA(2) model

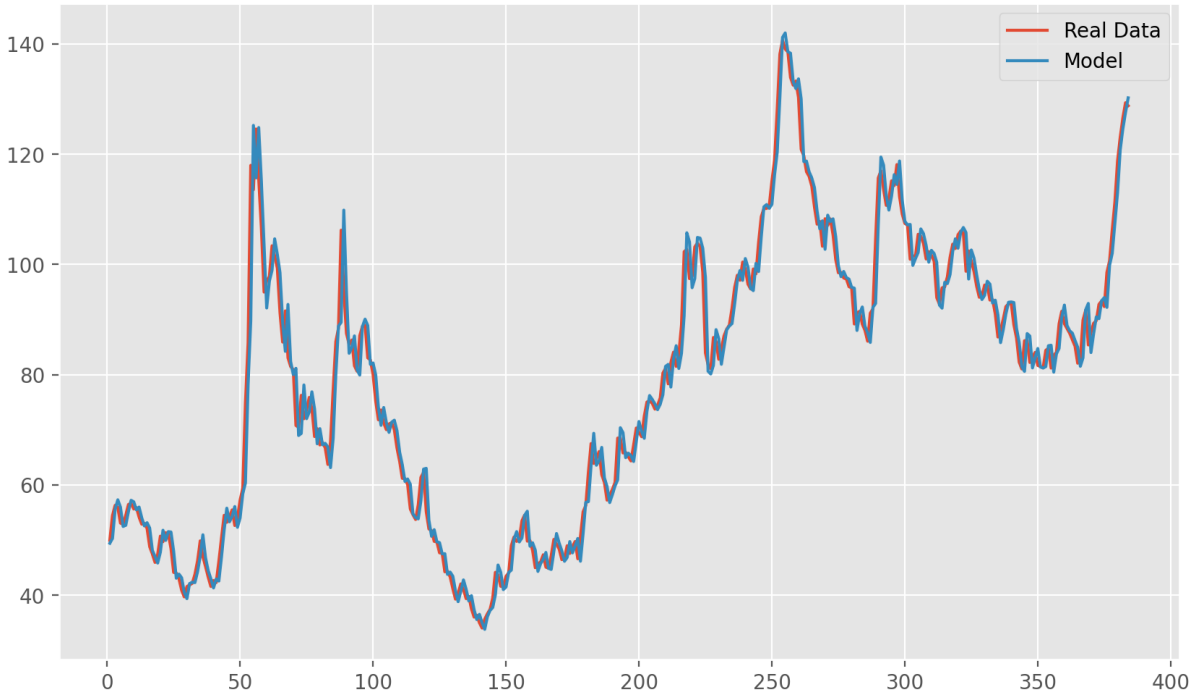

Figure 8: Time series plot of Australian beer production

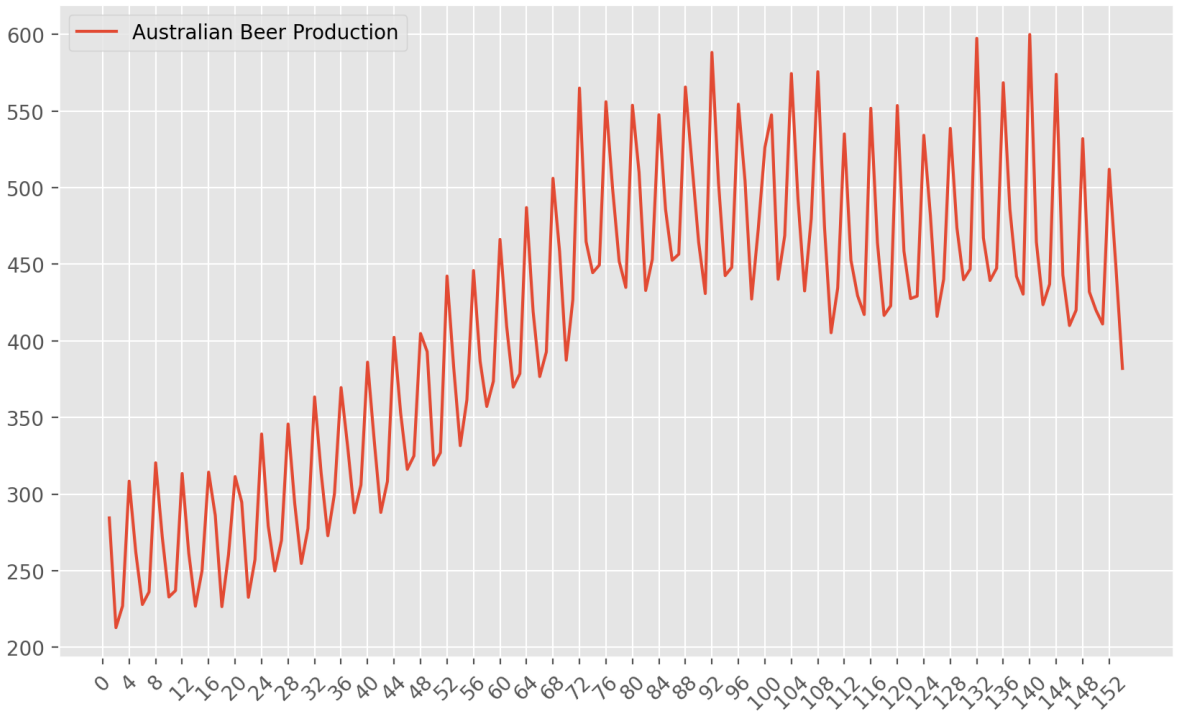

Figure 9: SARIMA results of detrended and deseasonalized AusBeer data

SARIMAX Results

Dep. Variable:	prod	**No. Observations:**	154
Model:	SARIMAX(0, 1, 0)x(0, 1, 0, 4)	**Log Likelihood**	-705.976
Date:	Mon, 04 Apr 2022	**AIC**	1415.953
Time:	21:33:36	**BIC**	1421.961
Sample:	0	**HQIC**	1418.394
	- 154		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	-0.1047	2.283	-0.046	0.963	-4.579	4.370
sigma2	763.8084	78.855	9.686	0.000	609.255	918.361

Ljung-Box (L1) (Q):	55.25	**Jarque-Bera (JB):**	2.99
Prob(Q):	0.00	**Prob(JB):**	0.22
Heteroskedasticity (H):	3.04	**Skew:**	-0.20
Prob(H) (two-sided):	0.00	**Kurtosis:**	3.56

Figure 10: ACF and PACF plots of detrended AusBeer data

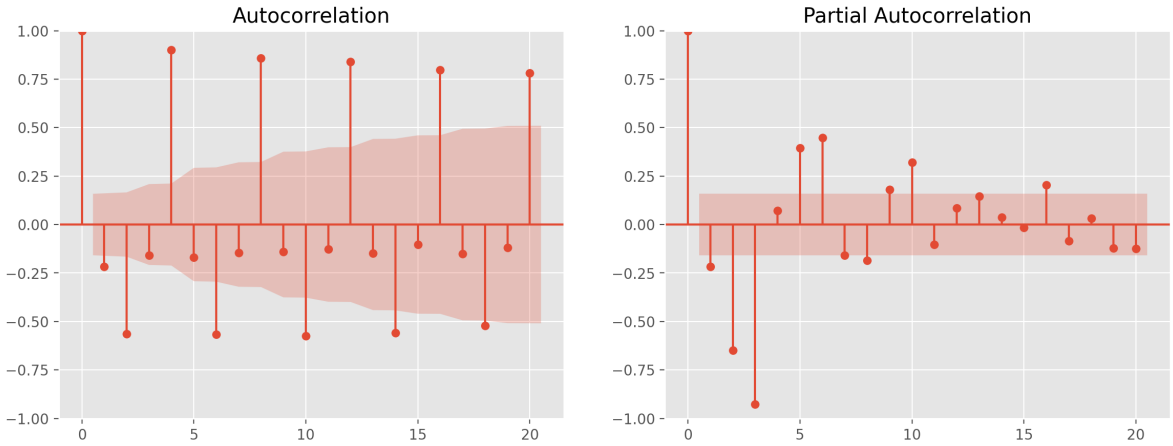

Figure 11: ACF and PACF plots of deseasonalized AusBeer data

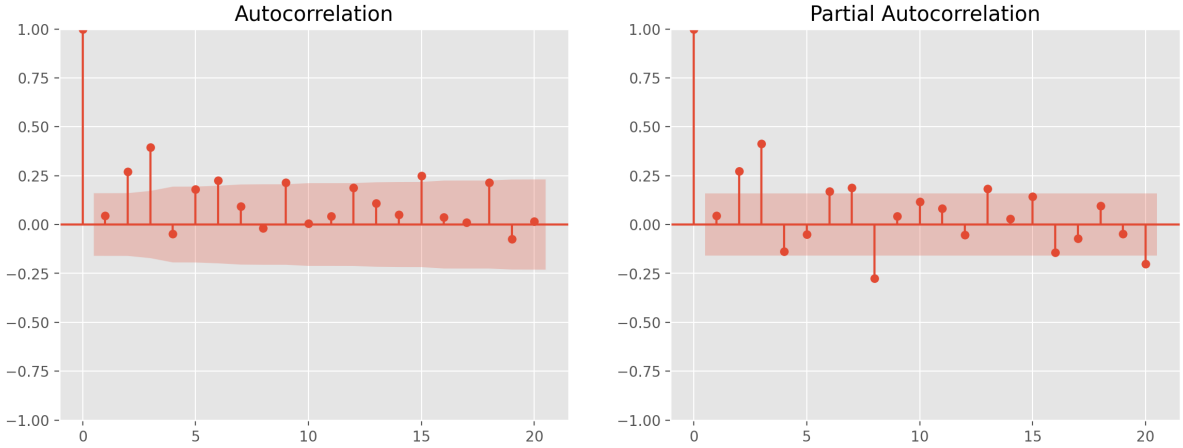

Figure 12: ACF and PACF plots of both detrended and deseasonalized AusBeer data

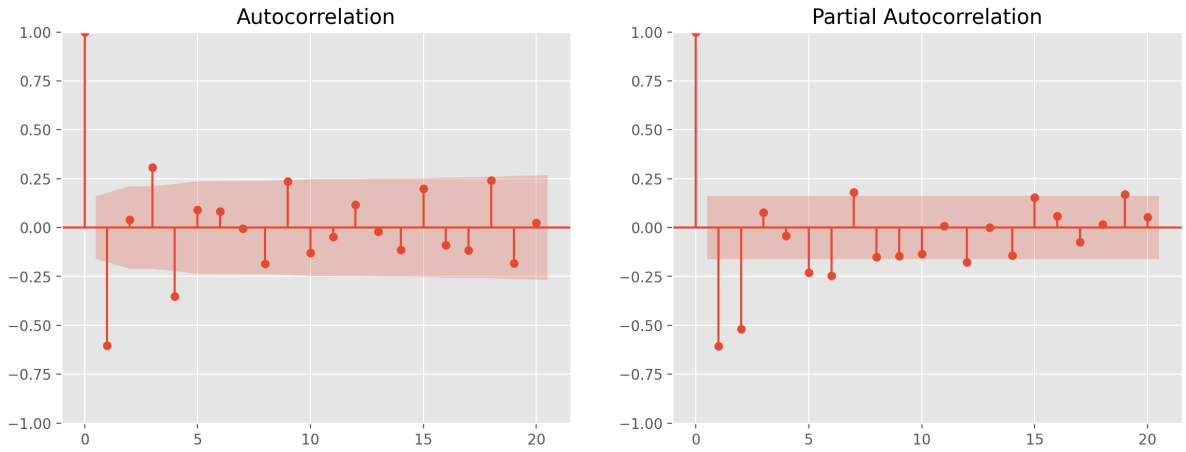

Figure 13: AR(2) results of AusBeer data

SARIMAX Results

Dep. Variable:	prod	**No. Observations:**	154
Model:	SARIMAX(2, 1, 0)x(0, 1, 0, 4)	**Log Likelihood**	-645.877
Date:	Mon, 04 Apr 2022	**AIC**	1299.754
Time:	21:33:37	**BIC**	1311.770
Sample:	0	**HQIC**	1304.636
	- 154		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	-0.2012	1.577	-0.128	0.898	-3.292	2.889
ar.L1	-0.9534	0.062	-15.260	0.000	-1.076	-0.831
ar.L2	-0.5409	0.073	-7.414	0.000	-0.684	-0.398
sigma2	338.2869	31.493	10.742	0.000	276.562	400.012

Ljung-Box (L1) (Q):	0.18	**Jarque-Bera (JB):**	11.78
Prob(Q):	0.67	**Prob(JB):**	0.00
Heteroskedasticity (H):	3.65	**Skew:**	0.17
Prob(H) (two-sided):	0.00	**Kurtosis:**	4.34

Figure 14: AR(2) & MA(3) results of AusBeer data

SARIMAX Results

Dep. Variable:	prod	**No. Observations:**	154
Model:	SARIMAX(2, 1, 3)x(0, 1, [], 4)	**Log Likelihood**	-631.745
Date:	Mon, 04 Apr 2022	**AIC**	1277.490
Time:	21:33:37	**BIC**	1298.518
Sample:	0	**HQIC**	1286.033
	- 154		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
intercept	-0.1479	0.418	-0.354	0.723	-0.966	0.671
ar.L1	-0.4025	0.059	-6.774	0.000	-0.519	-0.286
ar.L2	-0.6638	0.063	-10.567	0.000	-0.787	-0.541
ma.L1	-0.7007	0.053	-13.198	0.000	-0.805	-0.597
ma.L2	0.8461	0.045	18.956	0.000	0.759	0.934
ma.L3	-0.8434	0.058	-14.610	0.000	-0.957	-0.730
sigma2	269.2314	31.183	8.634	0.000	208.114	330.349

Ljung-Box (L1) (Q):	1.06	**Jarque-Bera (JB):**	0.75
Prob(Q):	0.30	**Prob(JB):**	0.69
Heteroskedasticity (H):	3.41	**Skew:**	-0.06
Prob(H) (two-sided):	0.00	**Kurtosis:**	3.33

Figure 15: OLS results of AusBeer data with dummies

OLS Regression Results

Dep. Variable:	prod	**R-squared:**	0.210
Model:	OLS	**Adj. R-squared:**	0.194
Method:	Least Squares	**F-statistic:**	13.29
Date:	Mon, 04 Apr 2022	**Prob (F-statistic):**	9.73e-08
Time:	21:33:37	**Log-Likelihood:**	-905.32
No. Observations:	154	**AIC:**	1819.
Df Residuals:	150	**BIC:**	1831.
Df Model:	3		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	479.7684	14.213	33.756	0.000	451.685	507.852
M1	-67.3300	19.971	-3.371	0.001	-106.790	-27.870
M2	-115.3812	19.971	-5.778	0.000	-154.842	-75.921
M3	-102.2474	20.100	-5.087	0.000	-141.963	-62.532

Omnibus:	54.951	**Durbin-Watson:**	0.079
Prob(Omnibus):	0.000	**Jarque-Bera (JB):**	16.787
Skew:	-0.574	**Prob(JB):**	0.000226
Kurtosis:	1.860	**Cond. No.**	4.82

Figure 16: OLS results of AusBeer data with dummies and trend variable

OLS Regression Results

Dep. Variable:	prod	**R-squared:**	0.782
Model:	OLS	**Adj. R-squared:**	0.776
Method:	Least Squares	**F-statistic:**	133.7
Date:	Mon, 04 Apr 2022	**Prob (F-statistic):**	2.97e-48
Time:	21:33:37	**Log-Likelihood:**	-806.14
No. Observations:	154	**AIC:**	1622.
Df Residuals:	149	**BIC:**	1637.
Df Model:	4		
Covariance Type:	nonrobust		

	coef	std err	t	P>|t|	[0.025	0.975]
Intercept	350.6493	9.935	35.296	0.000	331.018	370.280
t	1.6554	0.084	19.780	0.000	1.490	1.821
M1	-65.6746	10.523	-6.241	0.000	-86.469	-44.880
M2	-115.3812	10.523	-10.965	0.000	-136.175	-94.588
M3	-100.5920	10.591	-9.497	0.000	-121.521	-79.663

Omnibus:	3.876	**Durbin-Watson:**	0.285
Prob(Omnibus):	0.144	**Jarque-Bera (JB):**	2.965
Skew:	0.208	**Prob(JB):**	0.227
Kurtosis:	2.462	**Cond. No.**	400.

Figure 17: Comparison of best OLS and ARIMA models with real data