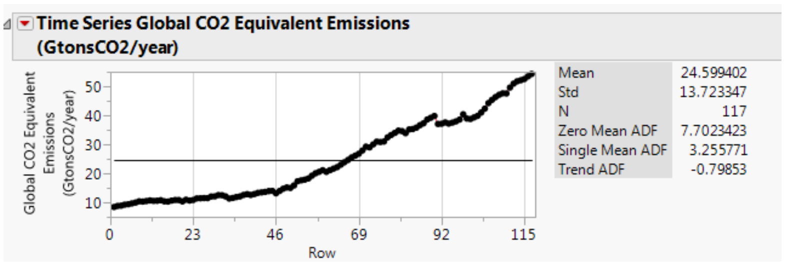
Vitid Nakareseisoon(ID: 9623-5872-43)

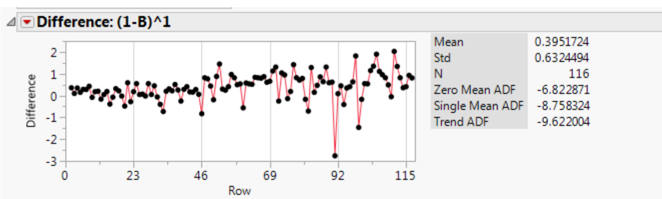
1. (*Dataset: Climate.xlsx*) There has been a lot of press recently about *climate change* and CO_2 emissions[1]. The data series for this problem contains data on *Global CO_2 Equivalent Emissions* (Giga-tons of CO_2 per year) from 1900 – 2016. The 2015 Paris Agreement pledge was to lower emissions by at least 26 percent below 2005 levels by 2025.

a. Create a solid time series model for this data and include your final output here. Briefly list the steps you took to come to this model.

ANS:



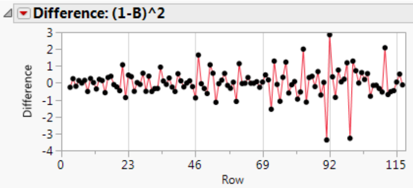
I plot the time series for CO2, all Dicky-Fuller T-statistics indicate that the series is not stationary.



Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	0.1887	4.2397	0.0395*	1	0.1887
2	0.0554	4.6081	0.0999	2	0.0205
3	0.1557	7.5449	0.0564	3	0.1469
4	0.0822	8.3709	0.0789	4	0.0281
5	0.0794	9.1493	0.1033	5	0.0559
6	0.0962	10.2998	0.1126	6	0.0529
7	0.1289	12.3863	0.0885	7	0.0927
8	0.1584	15.5664	0.0490*	8	0.1082
9	0.2209	21.8071	0.0095*	9	0.1664
10	0.2054	27.2545	0.0024*	10	0.1258
11	-0.0453	27.5222	0.0038*	11	-0.1500
12	-0.0540	27.9066	0.0057*	12	-0.1081
13	0.0002	27.9066	0.0093*	13	-0.0672
14	-0.0639	28.4541	0.0124*	14	-0.0983
15	0.1015	29.8490	0.0125*	15	0.1081
16	-0.0013	29.8492	0.0188*	16	-0.0784
17	0.0046	29.8521	0.0274*	17	-0.0174
18	0.0718	30.5723	0.0322*	18	0.0124
19	0.1591	34.1435	0.0177*	19	0.1684
20	-0.1414	36.9959	0.0117*	20	-0.1701
21	-0.0475	37.3209	0.0155*	21	0.0882
22	-0.0111	37.3389	0.0217*	22	-0.0320
23	-0.0404	37.5795	0.0282*	23	-0.0005
24	0.0216	37.6490	0.0377*	24	0.0188
25	-0.0716	38.4208	0.0421*	25	-0.1479

I take 1 difference and noted that all Dicky-Fuller T-statistics indicate a stationary so I(1) is considered. ACF graph indicates that MA() term can go upward to MA(10) and PACF graph indicates that AR() term can go upward to AR(1).

Difference: (1-B)^2



Mean	0.003913
Std	0.8081602
N	115
Zero Mean ADF	-16.63303
Single Mean ADF	-16.56075
Trend ADF	-16.48699

Lag	AutoCorr	Ljung-Box Q	p-Value	Lag	Partial
0	1.0000	.	.	0	1.0000
1	-0.4196	20.7824	<.0001*	1	-0.4196
2	-0.1414	23.1616	<.0001*	2	-0.3853
3	0.1075	24.5510	<.0001*	3	-0.1943
4	-0.0462	24.8096	<.0001*	4	-0.1890
5	-0.0150	24.8370	0.0001*	5	-0.1652
6	-0.0137	24.8603	0.0004*	6	-0.1924
7	0.0132	24.8822	0.0008*	7	-0.1777
8	-0.0232	24.9501	0.0016*	8	-0.2203
9	0.0463	25.2223	0.0027*	9	-0.1696
10	0.1443	27.8910	0.0019*	10	0.1037
11	-0.1503	30.8127	0.0012*	11	0.0557
12	-0.0433	31.0575	0.0019*	12	-0.0001
13	0.0758	31.8146	0.0026*	13	0.0273

I also try I(2) but its PACF graph indicates over differencing so I narrow down the model to I(1).

I run ARIMA model group based on the narrowed down parameters and obtain the following top candidate models:

Model	DF	Variance	AIC	SBC	RSquare	-2LogLH	Weights
ARIMA(1, 1, 1)	113	0.3839529	221.33543	229.59620	0.998	215.33543	0.265969
ARI(1, 1)	114	0.3924584	222.71454	228.22172	0.998	218.71454	0.133463
IMA(1, 1)	114	0.3928017	222.81599	228.32317	0.998	218.81599	0.126862
ARIMA(1, 1, 2)	112	0.3858666	222.87169	233.88605	0.998	214.87169	0.123378
ARIMA(1, 1, 3)	111	0.3849018	223.59703	237.36498	0.998	213.59703	0.085848
IMA(1, 3)	112	0.3927687	224.79868	235.81304	0.998	216.79868	0.047076
IMA(1, 2)	113	0.396271	224.81362	233.07439	0.998	218.81362	0.046725
I(1)	115	0.4034704	224.90176	227.65535	0.998	222.90176	0.044711
ARIMA(1, 1, 4)	110	0.3883917	225.59033	242.11188	0.998	213.59033	0.031688
IMA(1, 4)	111	0.3952822	226.50739	240.27534	0.998	216.50739	0.020033
ARIMA(1, 1, 5)	109	0.3887639	226.73073	246.00586	0.998	212.73073	0.017917
IMA(1, 10)	105	0.3742986	227.10164	257.39113	0.998	205.10164	0.014884
ARIMA(1, 1, 6)	108	0.3895234	227.92161	249.95033	0.998	211.92161	0.009878

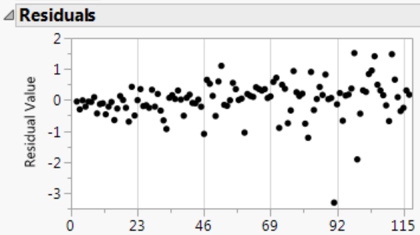
The top 3 models – ARIMA(1,1,1), ARIMA(1,1,0), and ARIMA(0,1,1) have comparable prediction performance and satisfactory residual plots but all parameters of ARIMA(1,1,1) are significant as 1% level of significant(the other two models do not) so I decide ARIMA(1,1,1) is the best model.

Parameter Estimates

Term	Lag	Estimate	Std Error	t Ratio	Prob>|t|	Constant Estimate
AR1	1	0.93545055	0.0717773	13.03	<.0001*	0.02642825
MA1	1	0.84026426	0.1032684	8.14	<.0001*	
Intercept	0	0.40942636	0.1312976	3.12	0.0023*	

Forecast

Residuals



b. Calculate a prediction for 2025 if the emissions trend continues the same way.

ANS: 59.56 Giga-tons of CO2 per year.

c. What is a 95% interval for your above prediction? Do you think there is a chance to meet the Paris Agreement pledge if steps aren't taken?

ANS: Interval: 64.46 - 54.67 Giga-tons of CO_2 per year. At 95% interval, there is no chance to meet the goal, as the lower band of the prediction is still much higher than the CO_2 level in 2005(45.35 Giga-tons of CO_2 per year).

2. (*Dataset: Unemployment.xlsx*) The following variables represent the Help Wanted Index and Unemployment Rate in the USA form the years 1969 through 2000.

HWI: National help wanted index

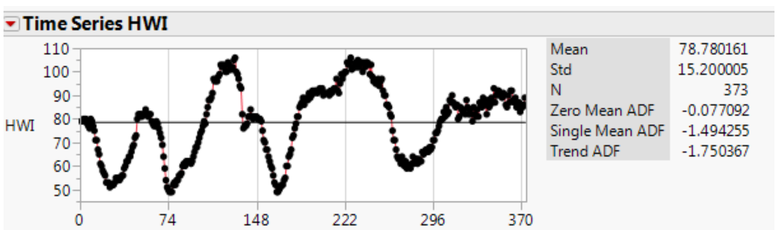
UNRATE: Unemployment rate %

a. Based on *just* the top Time Series plot in JMP, do either of these series look stationary?

ANS: No. The values of both series seem to come back to the mean but it does not look like they have a constant variance.

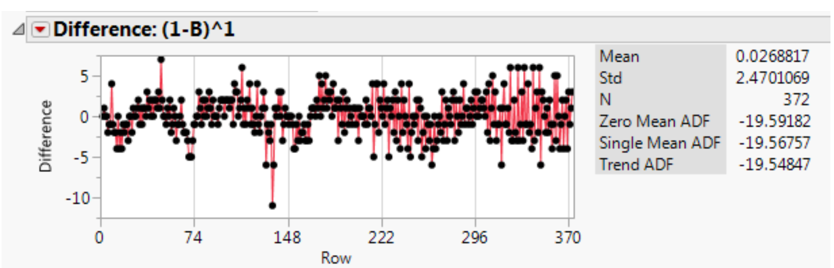
b. Use our standard statistical procedures to determine if the *HWI* variable is stationary.
i. Include here your basic output results (no graphs are necessary here) and your decision.
ii. If your series is NOT stationary, what do you need to do to fix it? Prove that your modification creates stationarity.

ANS:



HW1 is not stationary as all of its Dicky-Fuller T-statistics do not pass the 1% significant threshold.

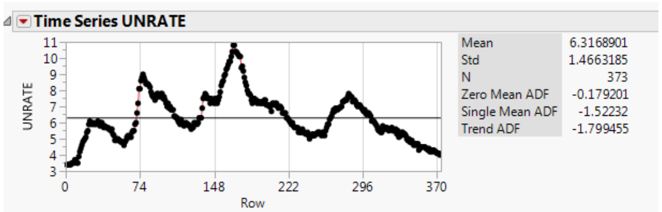
It can be made stationary by taking 1 difference, I(1). All Dicky-Fuller T-statistics pass 1% significant threshold and indicate stationary.



c. Use our standard statistical procedures to determine if the *UNRATE* variable is stationary.

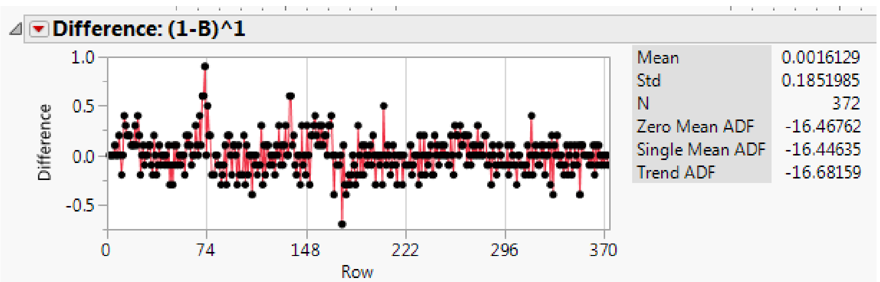
i. Include here your basic output results (no graphs are necessary here) and your decision.
ii. If your series is NOT stationary, what do you need to do to fix it? Prove that your modification creates stationarity.

ANS:



UNRATE is not stationary as all of its Dicky-Fuller T-statistics do not pass the 1% significant threshold.

It can be made stationary by taking 1 difference, I(1). All Dicky-Fuller T-statistics pass 1% significant threshold and indicate stationary.



d. Now create a simple regression model using *UNRATE* to predict *HWI* without any modifications, but be sure to include a constant.
i. What is your regression equation?
ii. Test this model for the possibility of *cointegration* and report your statistical results. What do you determine?

ANS:

Parameter Estimates

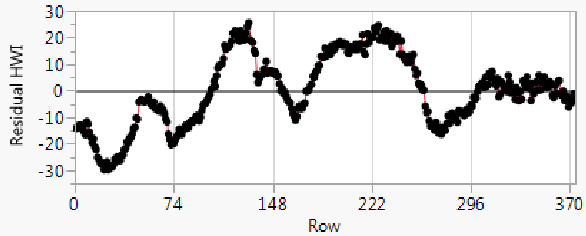
Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	109.82819	3.072768	35.74	<.0001*
UNRATE	-4.915082	0.473838	-10.37	<.0001*

The equation is:

HW1 = 109.83 - 4.92*UNRATE

I investigate the residual's time series of the model:

Time Series Residual HWI



Mean	5.143e-15
Std	13.382758
N	373
Zero Mean ADF	-1.87458
Single Mean ADF	-1.872323
Trend ADF	-1.779794

Dicky-Fuller T-statistics indicate a non-stationary so there may not be a cointegration between both series.

3. (*Dataset: PlantExpenditures.xlsx*) The goal of this question is to use a ***distributed lag***, or ***dynamic regression*** model to determine the parameters of the long and short-run impact of manufacturing sales (***Sales***, in billions of seasonally adjusted dollars) on the investment in equipment (***Expend***, also in billions of seasonally adjusted dollars).

a. Use the *Koyck Approach* to estimate the parameters you need. What is your output for this model?

ANS:

I run the regression model Expend ~ lag(Expend,1) + Sales, and obtain the following result:

Observations (or Sum Wgts) 21

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Ratio
Model	2	52944.232	26472.1	690.0561
Error	18	690.521	38.4	Prob > F
C. Total	20	53634.753		<.0001*

Parameter Estimates

Term	Estimate	Std Error	t Ratio	Prob>|t|
Intercept	-15.10403	4.72945	-3.19	0.0050*
lag(Expend)	0.2716761	0.114858	2.37	0.0294*
Sales	0.6292734	0.097819	6.43	<.0001*

The estimated lambda for Koyck is 0.27. The estimated alpha is -15.1/(1-0.27) = -20.68

Both median lag and mean lag indicate that incorporating only 1 lag period is enough. The estimated model is:

$$Expend = -20.68 + 0.63 * Sales + (0.63 * 0.27) * lag(Sales, 1)$$

b. What is the short-run impact of a one-unit (one billion dollar) increase in *Sales*?

ANS: From the equation above one unit increased in Sales resulted in 0.63 unit increased in Expend.

c. What is your long-run demand function for expenditures?

ANS:

The estimated total impact is 0.63/(1-0.27) = 0.86, so the long-run function is:

$$Expend = -20.68 + 0.86 * Sales$$

d. How many lags of *Sales* would you suggest including? (Explain why)

ANS:

From a., the estimated lambda is 0.27

The number of lag used can be calculated from Median Lag and Mean Lag

For Median Lag, the estimated Median Lag = $-\frac{\ln(2)}{\ln(0.27)}$ = 0.53

For Mean Lag, the estimated Mean Lag = $\frac{0.27}{1-0.27}$ = 0.37

Both approaches indicate that all changes can be accounted by incorporating just one lag period, so, only 1 lag of Sales is sufficient.