HW 1 – Business analytics.

Due Week 2 before class (See Canvas). Late submissions receive half credit.

To solve part b) below, you will need to do the following:

a. Use the Generalized Analytics Procedure (GAP) to set up your problem as follows:
    i. Define your model in words
        1. Identify the objective function in words
        2. Identify the random variables in words
        3. Identify the decision variables in words
        4. Identify the constraints in words
    ii. Formulate your model mathematically
        1. Define the decision variables
        2. Define the random variables
        3. Define the objective function in terms of decision variables and random variables
        4. Define the constraints in terms of the decision variables. Please include any non-negativity constraints in your formulation
b. Solve the problem in Excel
c. Answer the questions stated in the problem (in words).

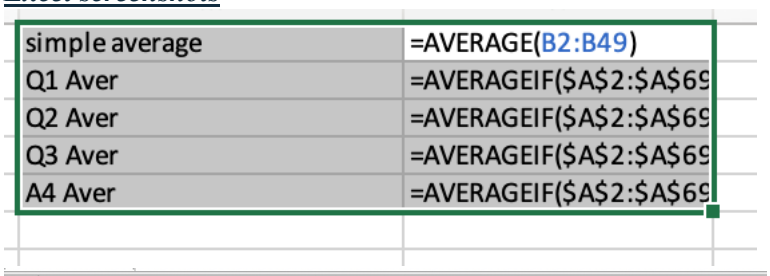
Please submit only one file in PDF format with your write-up. Do not submit your Excel file. Your writeup must include the screenshots from your Excel Spreadsheets. If you make any additional assumptions, state them clearly.

Australian Tabaco Production

a) Tab "training data" of HW1 spreadsheet.xlsx spreadsheet contains 1990-2001 quarterly Australian Tobacco Production (in metric tons). Use the following three forecasting methods (see Lecture 1 slides for details) to forecast the production in the 10 quarters starting with Q1 2002, using the 1990-2001 data as your training data:

- Simple Average
- Seasonal Naïve
- Seasonal Average

Excel screenshots

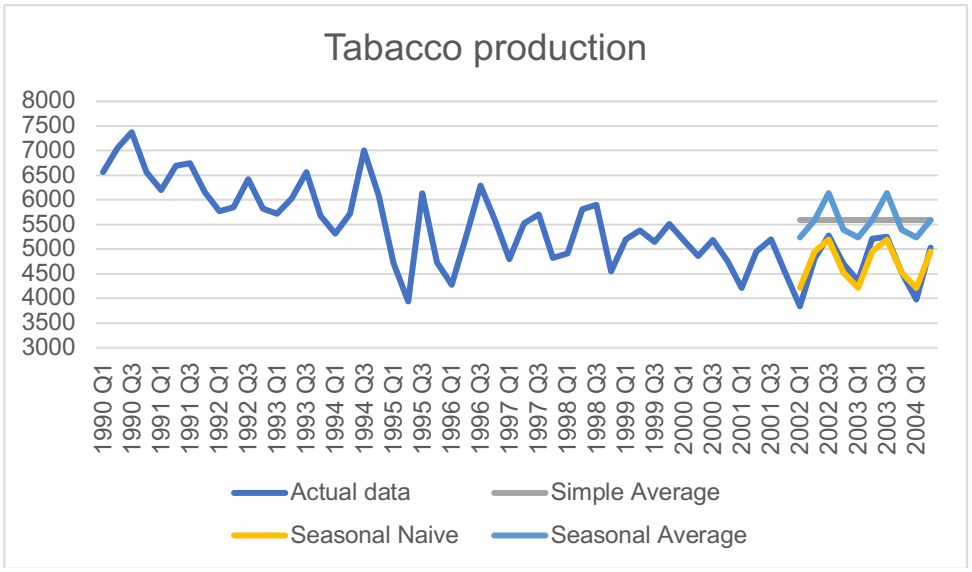


simple average	=AVERAGE(B2:B49)
Q1 Aver	=AVERAGEIF($A$2:$A$69
Q2 Aver	=AVERAGEIF($A$2:$A$69
Q3 Aver	=AVERAGEIF($A$2:$A$69
A4 Aver	=AVERAGEIF($A$2:$A$69

	A	B	C	D	E	F
1	Quarter	Actual (Metric tons)	Prediction	Simple Average	Seasonal Naïve	Seasonal Average
50	2002 Q1			5592.15	4217	5239.17
51	2002 Q2			5592.15	4959	5590
52	2002 Q3			5592.15	5196	6139.42
53	2002 Q4			5592.15	4522	5400
54	2003 Q1			5592.15	4217	5239.17
55	2003 Q2			5592.15	4959	5590
56	2003 Q3			5592.15	5196	6139.42
57	2003 Q4			5592.15	4522	5400
58	2004 Q1			5592.15	4217	5239.166667
59	2004 Q2			5592.15	4959	5590

Create a line graph in Excel showing the predictions from the three forecasts. In addition to the forecasts, include the actual data (from 1990 to 2004) in your graph. Which method do you expect to perform best/worst?

Answer



Seasonal naïve uses the last observation from 2001. Simple average uses average of all previous observations. Seasonal average uses quarterly averages. I expect the simple average method performs worst since it uses all the data without considering seasonal factors. So using the last data to predict for tobacco production isn't very practical. While the seasonal average is best since it calculates the seasonal situation and predicts using the all the data in the past.

b) You have been tasked to select the best of the three methods to forecast the production in the next 10 quarters (Q1 2002 – Q2 2004). Which method do you prefer? Answer this question comparing the forecasts from each method against the actual (realized) production quantities found in Q1 2002 – Q2 2004 displayed in tab "full data". Follow the procedure on the previous page (in red), and an evaluation method of your choice (MAD or MSE).

Excel screenshots

		Forecasts		
	Truth	Simple Average	Seasonal Naïve	Seasonal Average
2002 Q1	3843	5592.14583333333	4217	5239.16666666667
2002 Q2	4806	5592.14583333333	4959	5590
2002 Q3	5280	5592.14583333333	5196	6139.41666666667
2002 Q4	4709	5592.14583333333	4522	5400
2003 Q1	4362	5592.14583333333	4217	5239.16666666667
2003 Q2	5210	5592.14583333333	4959	5590
2003 Q3	5258	5592.14583333333	5196	6139.41666666667
2003 Q4	4526	5592.14583333333	4522	5400
2004 Q1	3974	5592.14583333333	4217	5239.16666666667
2004 Q2	5027	5592.14583333333	4959	5590

	Deviations		
	Simple Average	Seasonal Naïve	Seasonal Average
	=C3-B3	=D3-B3	=E3-B3
	=C4-B4	=D4-B4	=E4-B4
	=C5-B5	=D5-B5	=E5-B5
	=C6-B6	=D6-B6	=E6-B6
	=C7-B7	=D7-B7	=E7-B7
	=C8-B8	=D8-B8	=E8-B8
	=C9-B9	=D9-B9	=E9-B9
	=C10-B10	=D10-B10	=E10-B10
	=C11-B11	=D11-B11	=E11-B11
	=C12-B12	=D12-B12	=E12-B12

	Absolute Deviations		
	Simple Average	Seasonal Naïve	Seasonal Average
	=ABS(C3-B3)	=ABS(D3-B3)	=ABS(E3-B3)
	=ABS(C4-B4)	=ABS(D4-B4)	=ABS(E4-B4)
	=ABS(C5-B5)	=ABS(D5-B5)	=ABS(E5-B5)
	=ABS(C6-B6)	=ABS(D6-B6)	=ABS(E6-B6)
	=ABS(C7-B7)	=ABS(D7-B7)	=ABS(E7-B7)
	=ABS(C8-B8)	=ABS(D8-B8)	=ABS(E8-B8)
	=ABS(C9-B9)	=ABS(D9-B9)	=ABS(E9-B9)
	=ABS(C10-B10)	=ABS(D10-B10)	=ABS(E10-B10)
	=ABS(C11-B11)	=ABS(D11-B11)	=ABS(E11-B11)
	=ABS(C12-B12)	=ABS(D12-B12)	=ABS(E12-B12)
MAD	=AVERAGE(B19:B28)	=AVERAGE(C19:C28)	=AVERAGE(D19:D28)

	Squared Deviations		
	Simple Average	Seasonal Naïve	Seasonal Average
	=$H3^2	=$I3^2	=$J3^2
	=$H4^2	=$I4^2	=$J4^2
	=$H5^2	=$I5^2	=$J5^2
	=$H6^2	=$I6^2	=$J6^2
	=$H7^2	=$I7^2	=$J7^2
	=$H8^2	=$I8^2	=$J8^2
	=$H9^2	=$I9^2	=$J9^2
	=$H10^2	=$I10^2	=$J10^2
	=$H11^2	=$I11^2	=$J11^2
	=$H12^2	=$I12^2	=$J12^2
MSE	=AVERAGE(H19:H28)	=AVERAGE(I19:I28)	=AVERAGE(J19:J28)

	Absolute Deviations		
	Simple Aver:	Seasonal Naïv	Seasonal Average
	1749.15	374.00	1396.17
	786.15	153.00	784.00
	312.15	84.00	859.42
	883.15	187.00	691.00
	1230.15	145.00	877.17
	382.15	251.00	380.00
	334.15	62.00	881.42
	1066.15	4.00	874.00
	1618.15	243.00	1265.17
	565.15	68.00	563.00
MAD	892.65	157.10	857.13

	Squared Deviations		
	Simple Average	Seasonal Naïv	Seasonal Average
	3059511.15	139876.00	1949281.36
	618025.27	23409.00	614656.00
	97435.02	7056.00	738597.01
	779946.56	34969.00	477481.00
	1513258.77	21025.00	769421.36
	146035.44	63001.00	144400.00
	111653.44	3844.00	776895.34
	1136666.94	16.00	763876.00
	2618395.94	59049.00	1600646.69
	319389.81	4624.00	316969.00
MSE	1040031.83	35686.90	815222.38

Answer

Define Model in Words

O: Objective: what are we trying to achieve?

Predict tobacco production (Q1 2002 – Q2 2004).

R: Random Variables: what are the unknowns/risks?

Tobacco production (Q1 2002 – Q2 2004).

D: Decision Variables: What will you calculate/solve for?

Predictions for tobacco production (Q1 2002 – Q2 2004)

C: Constraints: What are the values that decision variables can(not) take?

Prediction numbers must be non-negative

I prefer the seasonal naive method. Compared the MAD or MSE of all three methods, we can find that the seasonal naive is in an acceptable range (the criteria here is using the 50 percentiles to calculate all the MAD and MSE relatively).

c) The Australian government is planning to collect a tax on tobacco production. Specifically, the tax will be $100 per metric ton in 2002, and $200 per metric ton in 2003-2004. Based on your best prediction, what will be the total tax revenue? Use the "SUMPRODUCT" formula to obtain your result.

Note: Your submission should include a screenshot showing how you used the "SUMPRODUCT" formula (You can display all formulas in an Excel spreadsheet by going to Formulas tab/Show formulas).

Excel screenshots

Seasonal Naïve	Tax per ton(AUD)	Annual Tax(Million AUD)
4217	100	=E5*D5
4959	100	=E6*D6
5196	100	=E7*D7
4522	100	=E8*D8
4217	200	=E9*D9
4959	200	=E10*D10
5196	200	=E11*D11
4522	200	=E12*D12
4217	200	=E13*D13
4959	200	=E14*D14
	Total:	=SUM(F5:F14)

Total (Sumproduct):	=SUMPRODUCT(D5:D14,E5:E14)

Seasonal Naïve	Tax per ton(AUD)	Annual Tax(Million AUD)
4217	100	421,700.00
4959	100	495,900.00
5196	100	519,600.00
4522	100	452,200.00
4217	200	843,400.00
4959	200	991,800.00
5196	200	1,039,200.00
4522	200	904,400.00
4217	200	843,400.00
4959	200	991,800.00
	Total:	7,503,400.00

Total (Sumproduct):	7503400

Answer

The total tax revenue will be AUD 7,503,400.