THE UNIVERSITY OF TEXAS AT AUSTIN
McCombs School of Business

STA 372.5 **Spring 2018**

HOMEWORK #6 – Due Wednesday, March 28

<u>Questions 1 and 2 need to be turned in. Questions 3, 4 and 5 do not need to be turned in.</u>

1. Suppose the following exponential smoothing model is fit to the data:

$$Y_t = L_{t-1} + T_{t-1} + \varepsilon_t \qquad \varepsilon_t \text{ iid } N(0, \sigma^2)$$

$$L_{t-1} = \alpha Y_{t-1} + (1 - \alpha)(L_{t-2} + T_{t-2})$$

$$T_{t-1} = \beta(L_{t-1} - L_{t-2}) + (1 - \beta)T_{t-2}$$

where $\alpha = 0.3$, $\beta = 0.4$, $\sigma = 1$, $L_0 = 30$, and $T_0 = 0.5$. Suppose the $t = 50^{\text{th}}$ observation is $Y_{50} = 40.0$, $L_{49} = 41.0$ and $T_{49} = 1.0$.

(a) What is the distribution of Y_{51} given information through time period $t = 50$? What is the forecast of Y_{51} given information through $t = 50$? What is the probability the actual value of Y_{51} that occurs is within one unit of the forecasted value?

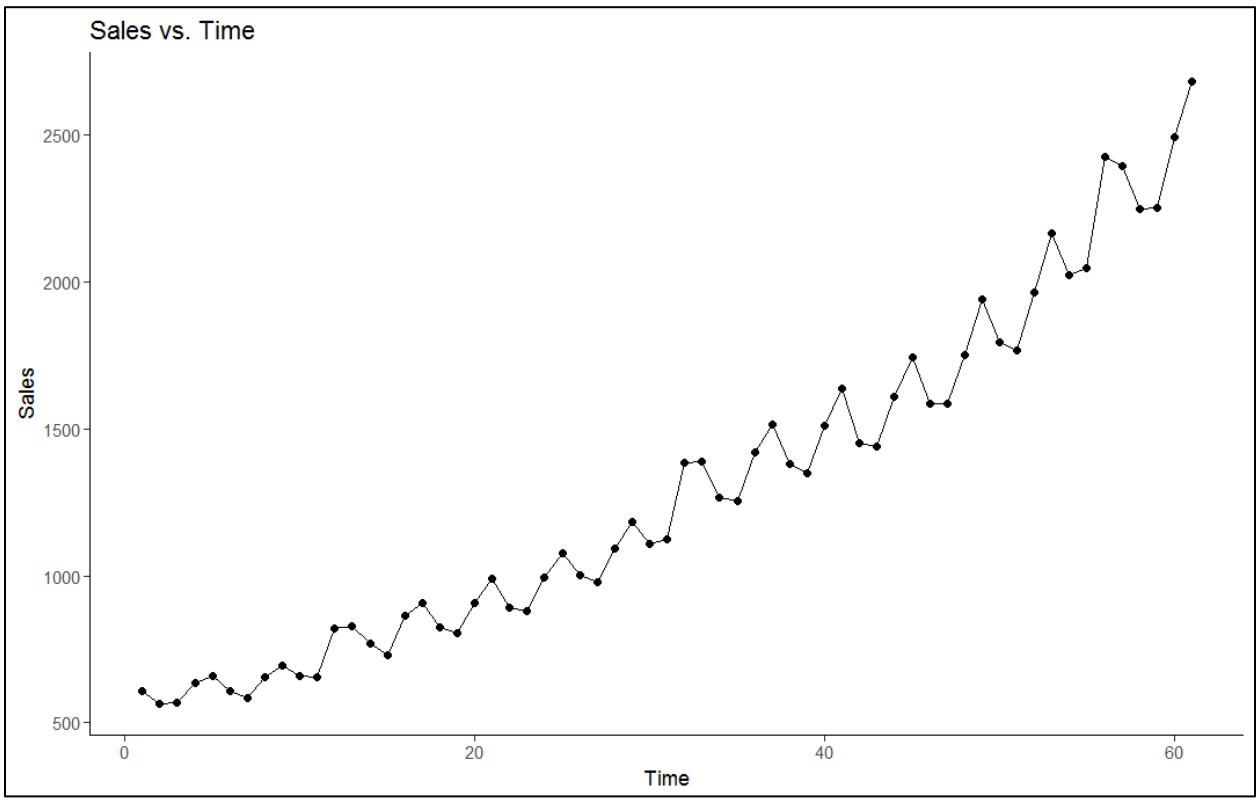
(b) What is the distribution of Y_{52} given information through time period $t = 50$? What is the forecast of Y_{52} given information through $t = 50$? What is the probability the actual value of Y_{52} that occurs is within one unit of the forecasted value?

2. Southeastern Associates Inc. (SAI) provides consulting activities to supermarket chains and to large, individually owned grocery stores located primarily in the southeastern part of the United States. As part of its own planning activities, SAI regularly forecasts sales of its target group of supermarket firms. One member of the target group is Publix Super Markets Inc., which operates about 525 supermarkets, mostly in Florida and Georgia. Publix sells groceries and dairy, produce, deli, bakery, meat, seafood, housewares, and health and beauty care items. The firm also makes dairy, bakery, and deli products.

   The data for this case is in the file STA372_Homework6_Question2.dat on the *Data sets* page of the Canvas class website. This file contains Publix's sales for 61 quarters. The file contains *Time* in the first column, the variable *Quarter* in the second column, and Publix's sales in the third column. You will need to seasonally adjust Publix's sales using a multiplicative model (or equivalently, using an additive model for the log of sales).

   You have been asked to forecast Publix's total sales for four quarters beyond the last quarter for which data are available. This means you are asked to forecast their sales in quarters 62-65.

   A time series plot of Publix's sales data is shown below.



The Marketing Manager of SAI, Mr. Tom Janke, also wants to have a deeper understanding of the seasonal character of Publix's sales. Such knowledge can be helpful, he has often said, in managing inventory and storage policies, part-time employment offers, and other activities that are influenced by seasonal variations. He also desires to have a better understanding of the trend in sales data; he has often commented on his interest in seeing trends - in addition to seasonal movements - in sales data.

You should use R to answer parts (a) - (f).

(a) Compute and plot the seasonal indices for log(*Sales*). Is the pattern of seasonality in Publix's sales fairly consistent through time? Why or why not? What quarter within the year are Publix's sales typically the highest?

(b) Compute and plot the seasonally adjusted sales values, denoted A_t. Does the seasonal adjustment do a good job of seasonally adjusting sales? Why or why not?

(c) Is there a consistent trend in seasonally adjusted sales or does it vary through time?

(d) Use Holt's exponential smoothing model to model and forecast seasonally adjusted sales (A_t).

Estimate α, β, L_0 and T_0, and compute the in-sample forecasts and also the four out-of-sample forecasts for time periods 62-65.

Plot the seasonally adjusted sales and their in-sample forecasts on the same graph. Do the in-sample forecasts do a good job tracking the seasonally adjusted sales?

How good are the in-sample forecasts in terms of their RMSE?

(e) Plot the errors from Holt's model and also compute their autocorrelation function. Do they appear to be independent? Why is it important to have independent errors?

(f) Using the results from part (d), provide forecasts of Publix's sales (not seasonally adjusted sales) in time periods 61-64. Given the RMSE of the in-sample forecasts and the plot of the errors from Holt's model for seasonally adjusted sales, how confident are you that the forecasts for periods 61-64 are likely to be accurate?

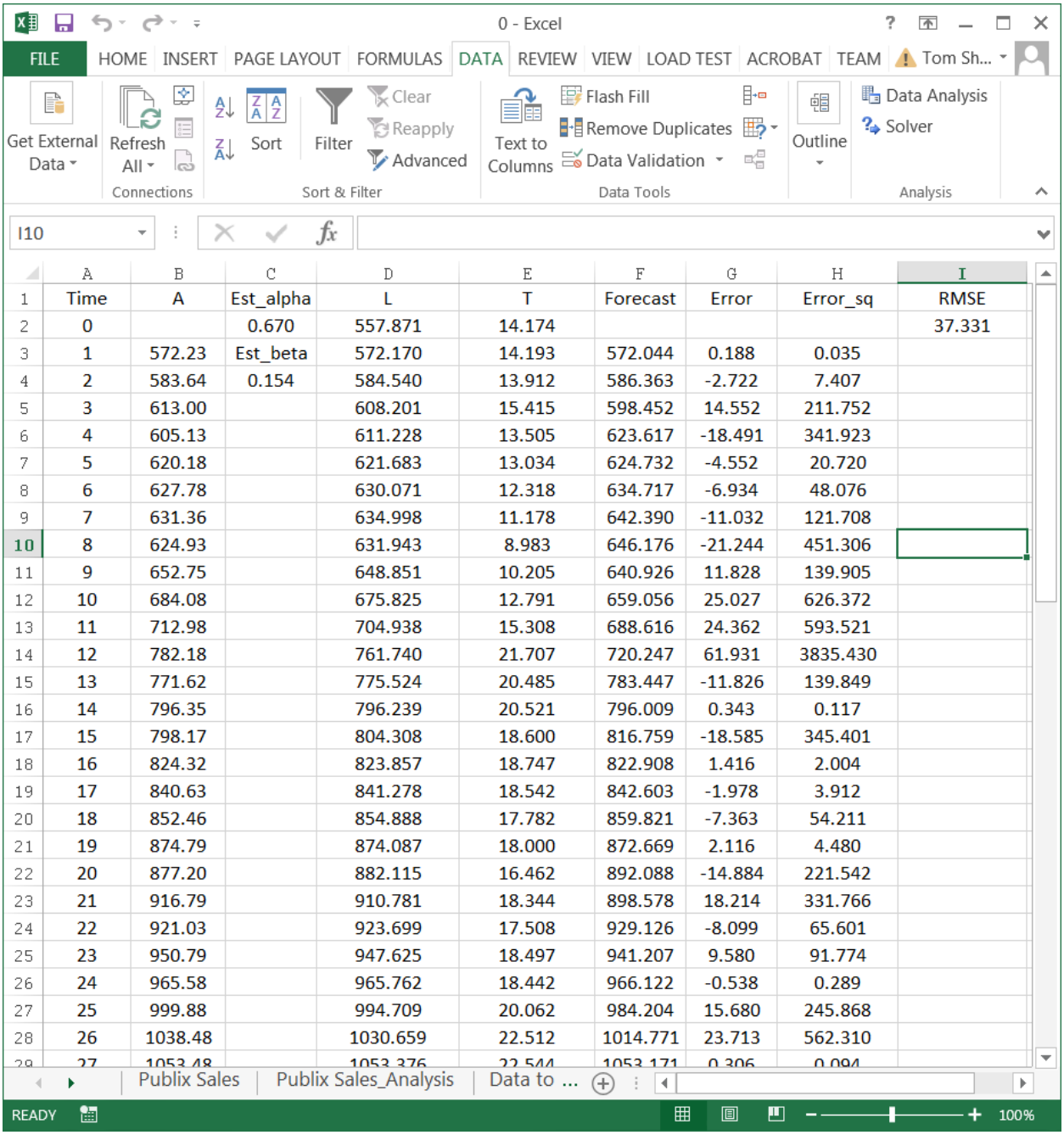
The following questions do not need to be turned in. However, I strongly recommend that you do them. I will distribute the answers to these problems with the answers to problems 1 and 2 that you do need to turn in.

3. Publix's seasonally adjusted sales data (A_t) are in the Excel spreadsheet STA372_Homework6_Question3.xlsx on the *Data sets* page of the Canvas class website.

   Set up a spreadsheet to compute the in-sample forecasts of Publix's sales using Holt's model in problem 2(d). Use Solver to estimate α, β, L_0 and T_0. The estimates of these values should be approximately the same as those obtained using the *holt* command in R in problem 2 (subject to differences due to different optimization routines being used in R and Solver - the estimate of beta will be somewhat different).

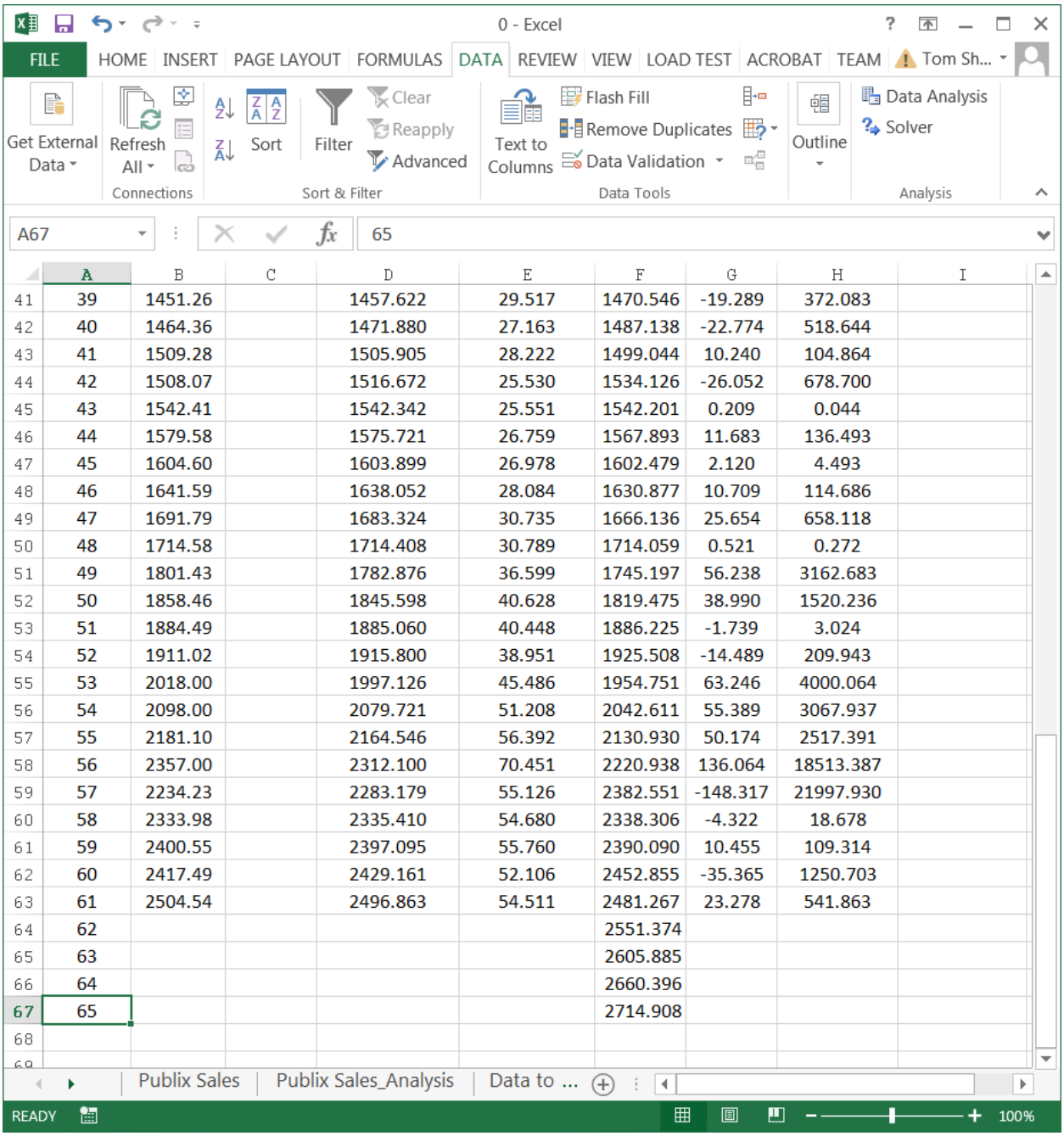
   Screenshots of the Excel implementation of Solver for Holt's model are given on the next two pages so you can check your estimates.

Screenshot of Excel spreadsheet implementing Holt's method



	A	B	C	D	E	F	G	H	I
1	Time	A	Est_alpha	L	T	Forecast	Error	Error_sq	RMSE
2	0		0.670	557.871	14.174				37.331
3	1	572.23	Est_beta	572.170	14.193	572.044	0.188	0.035	
4	2	583.64	0.154	584.540	13.912	586.363	-2.722	7.407	
5	3	613.00		608.201	15.415	598.452	14.552	211.752	
6	4	605.13		611.228	13.505	623.617	-18.491	341.923	
7	5	620.18		621.683	13.034	624.732	-4.552	20.720	
8	6	627.78		630.071	12.318	634.717	-6.934	48.076	
9	7	631.36		634.998	11.178	642.390	-11.032	121.708	
10	8	624.93		631.943	8.983	646.176	-21.244	451.306	
11	9	652.75		648.851	10.205	640.926	11.828	139.905	
12	10	684.08		675.825	12.791	659.056	25.027	626.372	
13	11	712.98		704.938	15.308	688.616	24.362	593.521	
14	12	782.18		761.740	21.707	720.247	61.931	3835.430	
15	13	771.62		775.524	20.485	783.447	-11.826	139.849	
16	14	796.35		796.239	20.521	796.009	0.343	0.117	
17	15	798.17		804.308	18.600	816.759	-18.585	345.401	
18	16	824.32		823.857	18.747	822.908	1.416	2.004	
19	17	840.63		841.278	18.542	842.603	-1.978	3.912	
20	18	852.46		854.888	17.782	859.821	-7.363	54.211	
21	19	874.79		874.087	18.000	872.669	2.116	4.480	
22	20	877.20		882.115	16.462	892.088	-14.884	221.542	
23	21	916.79		910.781	18.344	898.578	18.214	331.766	
24	22	921.03		923.699	17.508	929.126	-8.099	65.601	
25	23	950.79		947.625	18.497	941.207	9.580	91.774	
26	24	965.58		965.762	18.442	966.122	-0.538	0.289	
27	25	999.88		994.709	20.062	984.204	15.680	245.868	
28	26	1038.48		1030.659	22.512	1014.771	23.713	562.310	

Screenshot of Excel spreadsheet implementing Holt's method (continued)



	A	B	C	D	E	F	G	H	I
41	39	1451.26		1457.622	29.517	1470.546	-19.289	372.083	
42	40	1464.36		1471.880	27.163	1487.138	-22.774	518.644	
43	41	1509.28		1505.905	28.222	1499.044	10.240	104.864	
44	42	1508.07		1516.672	25.530	1534.126	-26.052	678.700	
45	43	1542.41		1542.342	25.551	1542.201	0.209	0.044	
46	44	1579.58		1575.721	26.759	1567.893	11.683	136.493	
47	45	1604.60		1603.899	26.978	1602.479	2.120	4.493	
48	46	1641.59		1638.052	28.084	1630.877	10.709	114.686	
49	47	1691.79		1683.324	30.735	1666.136	25.654	658.118	
50	48	1714.58		1714.408	30.789	1714.059	0.521	0.272	
51	49	1801.43		1782.876	36.599	1745.197	56.238	3162.683	
52	50	1858.46		1845.598	40.628	1819.475	38.990	1520.236	
53	51	1884.49		1885.060	40.448	1886.225	-1.739	3.024	
54	52	1911.02		1915.800	38.951	1925.508	-14.489	209.943	
55	53	2018.00		1997.126	45.486	1954.751	63.246	4000.064	
56	54	2098.00		2079.721	51.208	2042.611	55.389	3067.937	
57	55	2181.10		2164.546	56.392	2130.930	50.174	2517.391	
58	56	2357.00		2312.100	70.451	2220.938	136.064	18513.387	
59	57	2234.23		2283.179	55.126	2382.551	-148.317	21997.930	
60	58	2333.98		2335.410	54.680	2338.306	-4.322	18.678	
61	59	2400.55		2397.095	55.760	2390.090	10.455	109.314	
62	60	2417.49		2429.161	52.106	2452.855	-35.365	1250.703	
63	61	2504.54		2496.863	54.511	2481.267	23.278	541.863	
64	62					2551.374			
65	63					2605.885			
66	64					2660.396			
67	65					2714.908			
68									

4. If there is an increasing trend in the data but you used a simple exponential smoothing model (instead of Holt's exponential smoothing model that is specifically designed to model trends), would $\alpha = 0.9$ or $\alpha = 0.1$ give a smaller RMSE? Why?

5. True or False: If a random walk model is used to model the pattern in a data set of 30 observations, the best forecast of the 31st observation will be the sample mean of the previous 30 observations. Briefly explain why you answered true or false.