What a smoothie!

Teachers Teaching with Technology™
Professional Development from Texas Instruments

Teacher Notes and Answers

7 8 9 10 **11** **12**

TI-Nspire CAS | How-to | Teacher | 60 min

Introduction

In this task, you will construct a template for smoothing a time series, and representing it visually. It takes a little while to construct, but if saved it can be used to help with smoothing series investigations, regardless of the size of the data set. The following smoothing methods will be included in the template:

Mean based	**Median based**
• *3-point moving mean*	• *3-point moving median*
• *5-point moving mean*	• *5-point moving median*
• *4-point moving mean (centred)*	• *4-point moving mean (centred)*

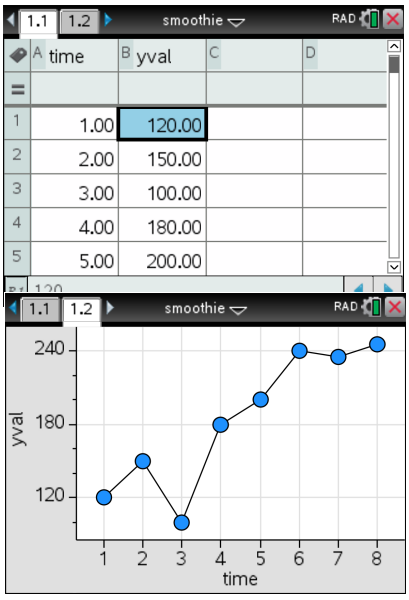
Construct a spreadsheet template for smoothed variables

The first job is to construct a spreadsheet template that allows the user to enter time series data, and have the smoothed variable values calculated, observable and graphable. To check the success of our template, we will first enter and use the following time series data:

time	1	2	3	4	5	6	7	8
yval	120	150	100	180	200	240	235	245

To enter the data on a spreadsheet and create a time series plot using the TI-Nspire CAS:

- Press [on] **> New Document** and select **Add Lists & Spreadsheet**.
- Press [on] **> Settings > Document Settings** to adjust settings so that **Display Digits = Fix 2** and **Calculation Mode = Approximate**.
- Click **OK** to save the settings changes, and then select **Current** to return to the spreadsheet.
- Locate the cursor at the top of column A and type the heading "**time**"
- Locate the cursor at the top of column B and type the heading "**yval**"
- Enter the data in the first two columns (should appear as shown right)
- Press [ctrl] [doc] and select **Add Data & Statistics**.
- Press [tab] to select **time** as *explanatory* variable
- Press [tab] to select **yval** as *response* variable
- Press [menu] **> Plot Type >X-Y Line Plot** (result shown right)

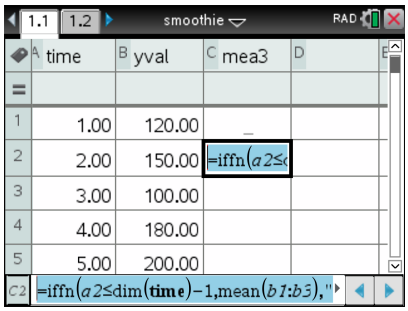


Author: D. Tynan

Constructing formulae for 3-point and 5-point moving mean values

It would be useful if the spreadsheet was able to work correctly when a time series with a different number of values is used. To do this, we will use a formula which 'checks' how many values there are in the current series, and displays all relevant smoothed values. For example, consider the formula below, which is to be stored in cell C2.

=iffn(a2<=dim(time)-1,mean(b1:b3),"")



This command can be interpreted in the following manner:

"If the contents of cell A2 is a number less than or equal to 7 (i.e. one less than the number of values in the time series), then display the mean of cells B1, B2 and B3, otherwise display nothing"

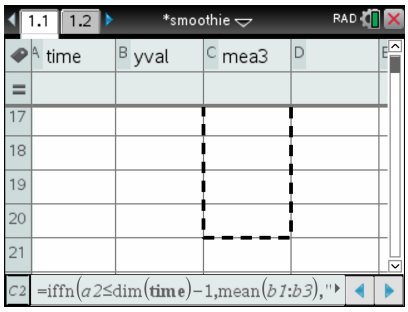
Question 1.

Look carefully at the formula above typed into cell C2. Why is it in C2 and not C1?

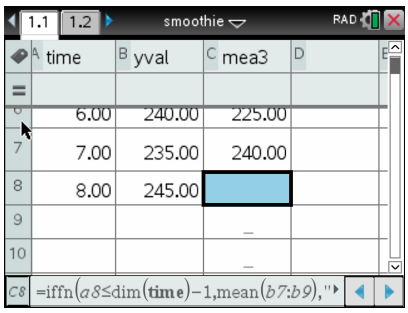
The first calculated value of a 3-point moving mean series is centred on the second time period, not the first time period.

The formula entered in cell C2 can be filled down as far as needed so that it finds the mean of the correct 3 nearby values of the time series for that time value. For our purposes, we will fill down so that the template works for a time series with up to 20 time series values.

To do this using the TI-Nspire CAS:



- Locate the cursor at the top of column C and type the heading "**mea3**"
- Locate the cursor in cell C2 and type in the formula ***=iffn(a2<=dim(time)-1,mean(b1:b3),"")***
  [Note: the "quotes" character can be entered by pressing ctrl ×]
- Press enter to calculate the value obtained from this formula (the value displayed should be 123.33)
- Locate the cursor again in cell C2
- To fill the formula down in column C, press menu **> Data > Fill**
- Press ▼ until you reach cell C20, and then press enter
  (the formula will now have been filled down to C20)



Note that the cells from C8 to C20 are blank, as the values in the corresponding cells in column A are not less than or equal to 7.

Question 2.

Write a similar formula to be placed in cell D3 that will return the first 5-point moving mean smoothed value. [Hint: remember it should only return a smoothed value from the 3rd value to the 3rd last value in the series].

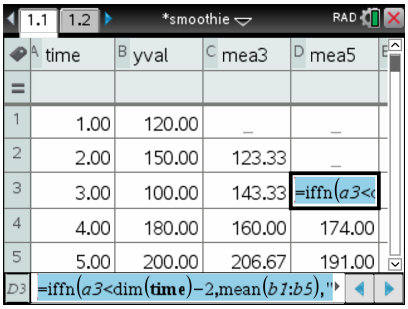
The following formula will work:

=iffn(a3<=dim(time)-2,mean(b1:b5),"")

Author: D. Tynan

Once you are happy that your formula answer to Question 2 is correct, go to the top of column D and type the column heading "**mea5**", and then enter your formula in cell D3.
Then fill it down to cell D20.

The screen should resemble the one shown right.



Constructing formulae for 3-point and 5-point moving median values

The process is fairly similar for 3-pont and 5-point moving medians. However, we will need to use the inbuilt median command.

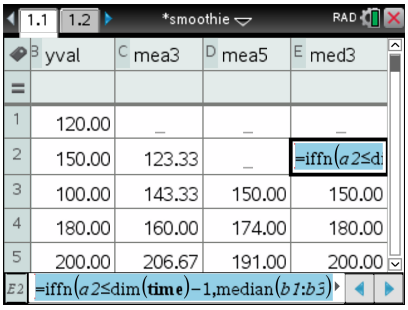
In column E, type a column heading "med3" and then construct a formula in E2 that will return the first 3-point moving median smoothed value. Then fill the formula down to cell E20.

Question 3.

What formula did you use in E2? What value is obtained from this formula?

The correct formula is
=iffn(a2<=dim(time)-1,median(b1:b3),"")

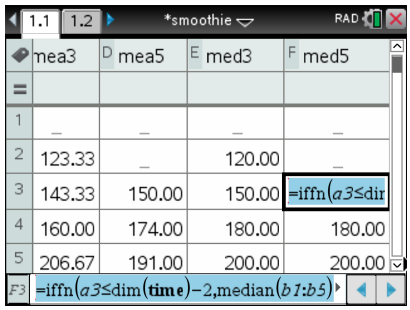
The value obtained should be 120.00



Question 4.

In column F, type a column heading "med5" and then construct a formula in F3 that will return the first 5-point moving median smoothed value. Then fill the formula down to cell F20.

=iffn(a3<=dim(time)-2,median(b1:b5),"")



Here is the spreadsheet after the extra formulae have been added. Press ctrl S to save your work.

	A time	B yval	C mea3	D mea5	E med3	F med5
1	1.00	120.00	_	_	_	_
2	2.00	150.00	123.33	_	120.00	_
3	3.00	100.00	143.33	150.00	150.00	150.00
4	4.00	180.00	160.00	174.00	180.00	180.00
5	5.00	200.00	206.67	191.00	200.00	200.00
6	6.00	240.00	225.00	220.00	235.00	235.00
7	7.00	235.00	240.00		240.00	
8	8.00	245.00				
9			_	_	_	_
10			_	_	_	_

Author: D. Tynan

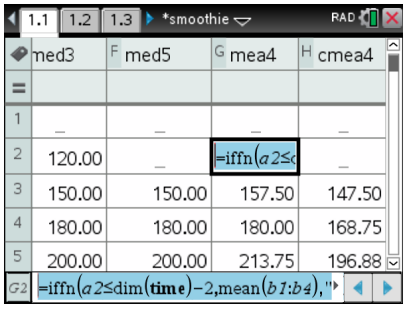
Constructing formulae for 4-point and centred 4-point moving mean values

First consider a 4-point smoothed value, which is inconveniently centred halfway between the 2nd and 3rd time period (i.e. *time* = 2.5). Recall that to deal with this, adjacent 4-point moving mean values can be centred on an actual time period via a further 2-point moving mean smoothing process. For example, this extra smoothing would centre 4-point moving mean values centred on *time* = 2.5 and 3.5 so that it is centred on *time* = 3.

In terms of the spreadsheet, we will construct a formula for the 4-point moving mean values which is not correctly centred (for example, it's harder to introduce an extra line for time = 2.5 and 3.5 value!). However the formula will be correctly aligned again for the centred 4-point moving mean values.

To do this using the TI-Nspire CAS :

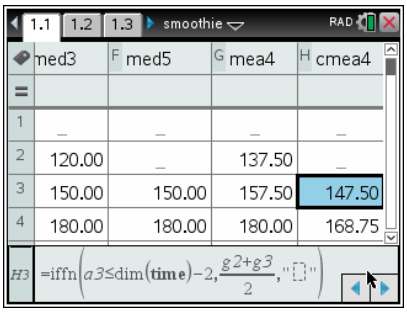
- Locate the cursor at the top of column G and type a column heading "**mea4**"
- Locate the cursor in the cell **G2** and type in the formula ***=iffn(a2<=dim(time)-2,mean(b1:b4),"")***
- Press enter to calculate the value obtained from this formula (the value displayed should be 137.50)
- Fill the formula down to G20



Note that these 4-point moving mean values are **NOT** correctly aligned with the time periods. The values will be correct, but the time period would need to be reduced by 0.5 to align them correctly. This issue is resolved for the centred 4-point moving mean smoothed values in the next step.

To create the centred 4-point smoothed moving mean values using the TI-Nspire CAS :

- Locate the cursor at the top of column H and type a column heading "**cmea4**"
- Locate the cursor in the cell **H3** and type in the formula ***=iffn(a3<=dim(time)-2,(g2+g3)/2,"")***
- Press enter to calculate the value obtained from this formula (the value displayed should be 147.50)
- Fill the formula down to **H20**



Constructing formulae for 4-point and centred 4-point moving median values

Question 5.

Repeat the above process to calculate both the 4-point moving *median* and centred 4-point moving *median* values.

4-point moving median smoothed values (start in I2):

=iffn(a2<=dim(time)-2,median(b1:b4),"")

Centred 4-point moving median smoothed values (in J3):

=iffn(a3<=dim(time)-2,(i2+i3)/2,"")

Author: D. Tynan

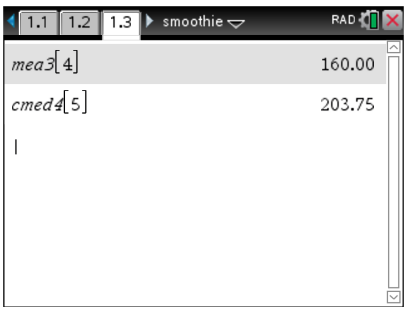
Using the "smoothie" template

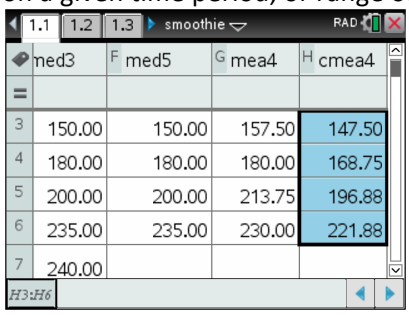
If all formulas are correct, the following results should be visible for the dataset provided. Check that your template has generated the same smoothed values. Then save the file again (ctrl S)

	A time	B yval	C mea3	D mea5	E med3	F med5	G mea4	H cmea4	I med4	J cmed4
1	1.00	120.00								
2	2.00	150.00	123.33		120.00		137.50		135.00	
3	3.00	100.00	143.33	150.00	150.00	150.00	157.50	147.50	165.00	150.00
4	4.00	180.00	160.00	174.00	180.00	180.00	180.00	168.75	190.00	177.50
5	5.00	200.00	206.67	191.00	200.00	200.00	213.75	196.88	217.50	203.75
6	6.00	240.00	225.00	220.00	235.00	235.00	230.00	221.88	237.50	227.50
7	7.00	235.00	240.00		240.00					
8	8.00	245.00								

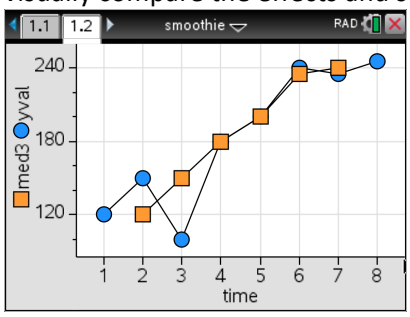
The template can be used to:

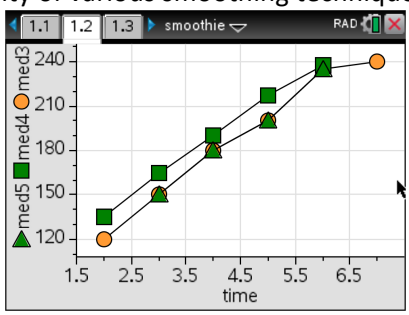
- find the correct smoothed values based on a given time period, or range of time periods



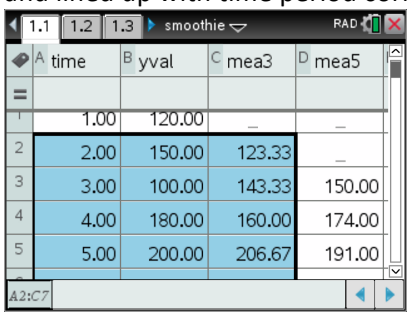


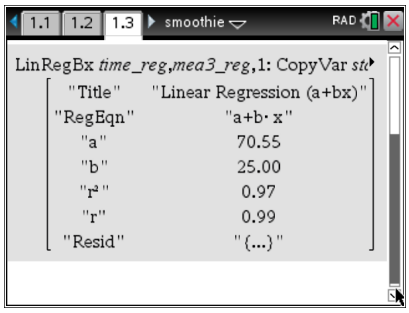
- visually compare the effects and suitability of various smoothing techniques





- create smoothed bivariate data sets upon which regression methods may be applied (would need to be copied and pasted to new lists to ensure that the two variables had same number of values, and lined up with time period correctly





VIP Note!: To use the template with a new time series, select the **time** column and press ▲ until the entire column is selected. Then press menu **> Data > Clear Data** to clear the data (this does not delete the variable name, which is important as these names are used in the formulas in other columns).
Repeat for the **yval** column, and then enter the data for the new time series.

Author: D. Tynan

Teacher notes

- This is not so much a task as a tool – constructed by students so that they have a clearer understanding of how it works, and its limitations (and how to get around them hopefully!). It is best constructed as a whole class activity.
- The smoothing methods are chosen based on the common ones used in Year 12 Further Maths.
- Although useful for SAC and exam purposes, it is still important that students be able to look at a time series provided and quickly determine a smoothed value for a particular time period
- Formulas that include conditions can look a bit difficult, but are included so that the smoothed values will still be calculated for time series with up to 20 time periods listed (i.e. $n = 20$). If there are more time periods listed ($n > 20$), then the formulas could just be filled down further.
- Remind students about the importance of clearing that data for **time** and **yval** first (MENU-3-4) before entering a new time series. This ensures that the two variable names are retained, and the formulas entered in other columns still work correctly. Might be useful to try a different time series to check they are clear about this. Here's one!

Year	Time	No. of employees	3 point moving average?
1994	1	800	
1995	2	875	
1996	3	910	
1997	4	1005	
1998	5	1150	
1999	6	1100	
2000	7	1103	
2001	8	1105	
2002	9	1155	
2003	10	1200	

Should get following results:

	A time	B yval	C mea3	D mea5	E med3	F med5	G mea4	H cmea4	I med4	J cmed4
=										
1	1.00	800.00	_	_	_	_	_	_	_	_
2	2.00	875.00	861.67	_	875.00	_	897.50	_	892.50	_
3	3.00	910.00	930.00	948.00	910.00	910.00	985.00	941.25	957.50	925.00
4	4.00	1005.00	1021.67	1008.00	1005.00	1005.00	1041.25	1013.13	1052.50	1005.00
5	5.00	1150.00	1085.00	1053.60	1100.00	1100.00	1089.50	1065.38	1101.50	1077.00
6	6.00	1100.00	1117.67	1092.60	1103.00	1103.00	1114.50	1102.00	1104.00	1102.75
7	7.00	1103.00	1102.67	1122.60	1103.00	1105.00	1115.75	1115.13	1104.00	1104.00
8	8.00	1105.00	1121.00	1132.60	1105.00	1105.00	1140.75	1128.25	1130.00	1117.00
9	9.00	1155.00	1153.33		1155.00					
10	10.00	1200.00								

Author: D. Tynan