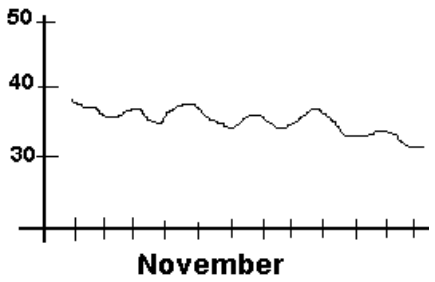
Normal Temperatures using Cubic Splines - Prof. Richard B. Goldstein

PROBLEM: How are "normal" high, low, and mean temperatures calculated for each city in the United States?

SOLUTION:

(1) The U.S. Weather Bureau does not take the literal average (arithmetic mean) on each particular day! Why not? Because it would be an erratic graph (as shown on the right) rapidly changing day-by-day up and down.



(2) The procedure used involves several steps.

(A) Collect data over at least 30 years.

(B) For each month take the mean (high, low, and daily mean temperature). For example, suppose we have 35 years of data for Providence (actually the state airport in Warwick). We would have 35 X 30 = 1,050 days in November from which to find the monthly average high by summing all 1,050 values and dividing by 1,050. The table given in 1982 for Providence is shown below:

Monthly Means for Providence, RI

	Low	Mean	High
January	20.6	28.4	36.2
February	21.2	29.4	37.6
March	29.0	36.9	44.7
April	37.8	47.3	56.7
May	46.9	56.9	66.8
June	56.5	66.4	76.3
July	63.0	72.1	81.1
August	61.0	70.4	79.8
September	53.6	63.4	73.1
October	43.4	53.7	63.9
November	34.6	43.3	52.0
December	23.4	31.5	39.6

(C) Next, the data is "fit" and the daily values are calculated using the fit. But how is this done?

(3) One can fit the 12 numbers in any column by using a FOURIER SERIES because this is cyclical data. Although this gives a fairly good fit, it was not chosen by the U.S. Weather Bureau. Perhaps, the reason was that on would need to assume all months are equal in length - say 30, days.

(4) The choice made was to use a CUBIC SPLINE fit.

(A) Attempt #1 Consider the 12 points (low temperatures):

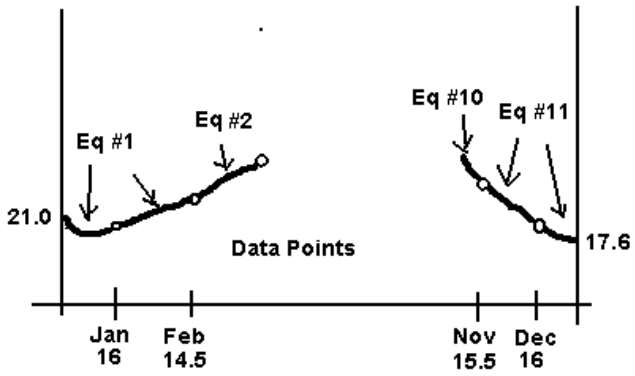
	x(i)	y(i)	
(1 + 31)/2 =	16.0	20.6	Here we can use
(32 + 59)/2 =	45.5	21.2	unequally spaced points.
	75.0	29.0	The midpoint day number
	105.5	37.8	for each month is x(i)
	136.0	46.9	
	166.5	56.5	Feb 1 = day #32
	197.5	63.0	Feb 28 = day #59
	228.0	61.0	
	258.5	53.5	etc.
	289.0	43.4	
	319.5	34.6	
(335 + 365)/2 =	350.0	23.4	

With <u>12</u> points one find <u>11</u> cubic equations.

<u>For</u>	<u>Use</u>	<u>As an</u>
Jan 1 to 15	eq #1	extrapolation
Jan 16 to Feb 14	eq #1	interpolation
Feb 15 to Mar 15	eq #2	interpolation
...	...	...
Nov 16 to Dec 15	eq #11	interpolation
Dec 16 to Dec 31	eq #11	extrapolation

This creates an unexpected problem:

THE ENDS DON'T MEET!



For example, by reading from the fit we find:

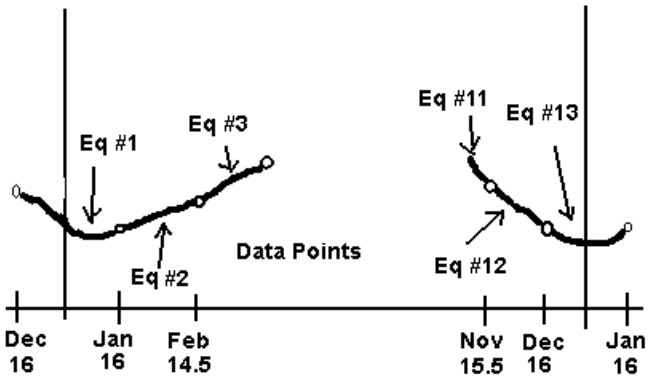
Early January			Late December		
Day #1	Jan 1	21.0	Day #359	Dec 25	19.9
Day #2	Jan 2	21.0	Day #360:	Dec 26	19.5
Day #3	Jan 3	21.0	Day #361	Dec 27	19.1
Day #4	Jan 4	21.0	Day #362	Dec 28	18.7
Day #5	Jan 5	21.0	Day #363	Dec 29	18.4
Day #6	Jan 6	21.0	Day #364	Dec 30	18.0
Day #7	Jan 7	20.9	Day #365	Dec 31	17.6

This does not lead to a very smooth transition from Dec 31 (17,6) to Jan 1 (21.0). How can we make the curves meet?

(B) Attempt #2 - Consider adding one extra point at each end - that is, using14 points. That would look like:

x(i)	y(i)		
- 15.0	23.4	Z	prior December
16.0	20.6		
45.5	21.2		
75.0	29.0		
...	...		
289.0	43.4		
319.5	34.6		
350.0	23.4		
381.0	20.6	Z	following January

The period between Dec 16 and Jan 16 now has **two** cubic curves representing it. Take the average of these results (cubic eq #1 and #13).



	Fit from cubic					Fit from cubic		
Jan	Eq #1	Eq #13	rounded mean		Dec	Eq #1	Eq #13	rounded mean
1	21.8	21.0	21.4	node<	16	23.4	23.4	23.4
2	21.7	20.9	21.4		17	23.3	23.2	23.3
3	21.6	20.9	21.3		18	23.2	22.9	23.1
4	21.5	20.9	21.2		19	23.1	22.7	22.9
5	21.4	20.8	21.1		20	23.0	22.5	22.8
6	21.4	20.8	21.1		21	22.9	22.3	22.6
7	21.3	20.7	21.0		22	22.8	22.1	22.5
8	21.2	20.7	21.0		23	22.7	22.0	22.4
9	21.1	20.7	20.9		24	22.6	21.8	22.2
10	21.0	20.7	20.9		25	22.5	21.7	22.1
11	21.0	20.7	20.9		26	22.4	21.6	22.0
12	20.9	20.6	20.8		27	22.3	21.5	21.9
13	20.8	20.6	20.7		28	22.2	21.3	21.8
14	20.7	20.6	20.7		29	22.1	21.3	21.7
15	20.6	20.6	20.7		30	22.0	21.2	21.6
16	20.6	20.6	20.6	=node	31	21.9	21.1	21.5

Now the ends meet! Dec 31 has a mean of 21.5 and Jan 1 has 21.4. What have we sacrificed? Some smoothness. The first and second derivatives don't match exactly on Jan 16 or Dec 16, but because we know have a continuous curve this would hardly be noticed.

Finally, what do you do about February 29? Use the same values as February 28.

The final result:

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
1	21.4	20.12	24.5	33.7	42.3	52.0	60.6	62.9	57.6	48.3	39.0	28.5
2	21.4	0.1	24.8	34.0	42.6	52.3	60.8	62.9	57.3	48.0	38.7	28.1
3	21.3	20.2	25.0	34.3	42.9	52.6	61.0	62.8	57.1	47.6	38.5	27.7
4	21.2	20.2	25.3	34.5	43.2	53.0	61.2	62.7	56.8	47.3	38.2	27.3
5	21.1	20.2	25.6	34.8	43.5	53.3	61.4	62.6	56.6	46.9	37.9	27.0
6	21.1	20.3	25.9	35.1	43.8	53.6	61.6	62.5	56.3	46.6	37.6	26.6
7	21.0	20.4	26.2	35.4	44.1	53.9	61.8	62.3	56.0	46.3	37.3	26.2
8	21.0	20.4	26.5	35.7	44.4	54.2	62.0	62.2	55.7	45.9	37.0	25.9
9	20.9	20.5	26.8	36.0	44.7	54.5	62.1	62.1	55.5	45.6	36.7	25.5
10	20.9	20.6	27.1	36.2	45.0	54.8	62.3	61.9	55.2	45.3	36.4	25.2
11	20.9	20.7	27.5	36.5	45.3	55.1	62.4	61.8	54.9	44.9	36.1	24.8
12	20.8	20.9	27.8	36.8	45.7	55.5	62.5	61.7	54.6	44.6	35.8	24.5
13	20.7	21.0	28.1	37.1	46.0	55.8	62.7	61.5	54.3	44.3	35.5	24.2
14	20.7	21.1	28.4	37.4	46.3	56.1	62.8	61.3	54.0	44.0	35.1	23.9
15	20.7	21.3	28.7	37.7	46.6	56.4	62.9	61.2	53.7	43.7	34.8	23.7
16	20.6	21.4	29.0	37.9	46.9	56.6	63.0	61.0	53.3	43.4	34.4	23.4
17	20.5	21.6	29.3	38.2	47.2	56.9	63.0	60.8	53.0	43.1	34.1	23.3
18	20.5	21.8	29.6	38.5	47.5	57.2	63.1	60.6	52.7	42.8	33.7	23.1
19	20.4	22.0	29.9	38.8	47.8	57.5	63.2	60.5	52.4	42.5	33.3	22.9
20	20.4	22.2	30.2	39.1	48.2	57.8	63.2	60.3	52.0	42.2	32.9	22.8
21	20.3	22.4	30.5	39.4	48.5	58.1	63.2	60.1	51.7	42.0	32.5	22.6
22	20.3	22.7	30.8	39.7	48.8	58.4	63.3	59.9	51.4	41.7	32.1	22.5
23	20.2	22.9	31.1	40.0	49.1	58.6	63.3	59.7	51.0	41.4	31.7	22.4
24	20.2	23.1	31.4	40.2	49.4	58.9	63.3	59.5	50.7	41.2	31.3	22.2
25	20.2	23.4	31.7	40.5	49.8	59.2	63.3	59.2	50.4	40.9	30.9	22.1
26	20.1	23.7	32.0	40.8	50.1	59.4	63.2	59.0	50.0	40.6	30.5	22.0
27	20.1	23.9	32.3	41.1	50.4	59.7	63.2	58.8	49.7	40.4	30.1	21.9
28	20.1	24.2	32.5	41.4	50.7	59.9	63.2	58.6	49.3	40.1	29.7	21.8
29	20.1		32.8	41.7	51.1	60.1	63.1	58.3	49.0	39.8	29.3	21.7
30	20.1		33.1	42.0	51.4	60.4	63.1	58.1	48.6	39.6	28.9	21.6
31	20.1		33.4		51.7		63.0	57.8		39.3		21.5