Question 1.

Assumption that we're using a .05 significance level when answering the questions below.

There are a number of ways to test for outliers using the grubbs.test function. The answer to the outlier question is dependent on the testing method used.

I've configured the grubbs.test to check if lowest and highest values are two outliers on opposite tails of the sample and to treat it as a two-sided test:

grubbs.test(uscrime_df\$Crime, type = 11, opposite = FALSE, two.sided = TRUE) This results in a p-value < 2.2e-16 which allows us to reject the null hypothesis and accept the alternative hypothesis that 342 (lowest-crime) and 1993 (highest-crime) are outliers.

However, if you test for a high/low outlier individually you receive different results. Checking if the highest value is an outlier with the following function:

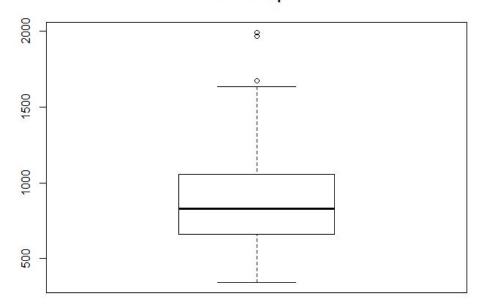
grubbs.test(uscrime_df\$Crime, type = 10, opposite = FALSE, two.sided = FALSE)
The result is a p-value of 0.07887 which doesn't allow us to reject the null hypothesis.

Checking if the lowest value is an outlier with the following function:

grubbs.test(uscrime_df\$Crime, type = 10, opposite = TRUE, two.sided = FALSE)
The result is a p-value of 1, which doesn't allow us to reject the null hypothesis.

Also, if you use a box plot to plot crime, 342 doesn't appear as an outlier

US Crime Rate per 100k



Question 2.

A change detection algorithm would be useful in tracking refrigerator temperature. We'd want to track to make sure that the temperature stays between 34 and 40 degrees. This would ensure that refrigerated food maintains a safe temperature. I'd place a low threshold (T) value with a low C because I'd want the algorithm to be more sensitive and allow me to respond quickly if temperatures exceeded the safety threshold.

Question 3.1

Because we're interested in when summer ends in Atlanta (not when it experiences a heat wave), I've used one-sided CUSUM focused on a decrease from the mean.

Equation:
$$S_t = max\{0, S_{t-1} + (x_t - \mu - C)\}$$

With the assumption that μ is 83° and a low C of 1 (keeping the CUSUM model more sensitive to variation), I found that a threshold (T) level of 50 produced a fairly consistent "summer ending date" across the years while also minimizing the number of "false alarms" where the threshold was briefly exceeded.

Fig. 1: Table of "Unofficial" End of Summer Dates - when T is exceeded and stays exceeded

"Unoffocial" End of Summer								
Avg.	6-Oct							
1996	1-Oct							
1997	27-Sep							
1998	9-Oct							
1999	28-Sep							
2000	18-Sep							
2001	28-Sep							
2002	29-Sep							
2003	29-Sep							
2004	21-Sep							
2005	12-Oct							
2006	30-Sep							
2007	14-Oct							
2008	8-Oct							
2009	30-Sep							
2010	4-Oct							
2011	3-Oct							
2012	8-Oct							
2013	29-Sep							
2014	4-Oct							
2015	26-Sep							

Fig. 2: Plot of St values over the years:



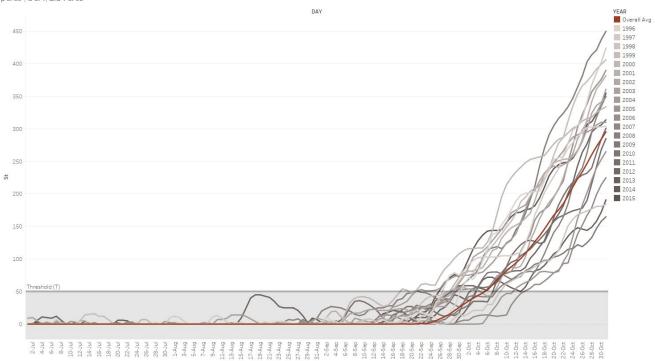


Fig. 3: Chart of St values across the years (red text exceeds threshold)

CUSUM - Daily Temperature µ of 83°, C of 1, and T of 50

AY	Overall	1996	1997	1998	1999	2000	2001	2002	2003	2004	YEAR 2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015
-Jul																					
-Jul																			6.0		
Jul Jul																					
Jul																					
Jul Jul																					
Jul											0.0										
D-Jul L-Jul											3.0										
2-Jul																					
3-Jul 4-Jul					14.0 15.0														4.0		
5-Jul																					
5-Jul 7-Jul																					
8-Jul 9-Jul																					
9-Jul																					
1-Jul 2-Jul																				5.0	
2-Jul 3-Jul																					
4-Jul 5-Jul																					
5-Jul																					
7-Jul B-Jul																					
i-Jul																					
D-Jul L-Jul											4.0										
Aug																					
Aug																					
Aug Aug																					
Aug Aug																					
Aug																					
Aug Aug																					
-Aug																					
-Aug																					
-Aug -Aug										9.0											
-Aug										14.0											
-Aug -Aug										14.0									9.0		
7-Aug B-Aug																			40.0		
-Aug										4.0 0.0									45.0 45.0		
-Aug																			43.0		
-Aug -Aug																			41.0 35.0		
B-Aug			4.0											4.0							
I-Aug -Aug			4.0																		
-Aug																					
7-Aug 8-Aug														8.0 4.0							
-Aug																					
l-Aug								9.0													
Sep																					
Sep Sep						4.0									24.0						
Sep		4.0																			
Sep Sep																					
Sep																					
Sep Sep						40.0															
0-Sep						42.0											34.0				
1-Sep 2-Sep						40.0 36.0				3.0											4.0
3-Sep		4.0																			
4-Sep 5-Sep												14.0					8.0			4:0 6.0	
6-Sep																	14.0				
7-Sep B-Sep						45.0 54.0															
9-5ep						52.0							14.0		47.0			4.0			14.0
D-Sep 1-Sep	0.0					47.0 52.0				44.0 51.0					48.0 51.0			7.0			
2-Sep						61.0			24.0	53.0					49.0						
-Sep -Sep		16.0	14.0		39.0 43.0	62.0				51.0 51.0			7.0		48.0 45.0		24.0				
-Sep	4.7				44.0	60.0				52.0					40.0					14.0	42.0
Sep Sep			42.0 60.0		44.0	74.0		34.0 41.0		55.0 65.0							14.0		40.0		53.0 64.0
-Sep			74.0		50.0	92.0	50.0	44.0	36.0	69.0				34.0					48.0	34.0	71.0
-Sep -Sep			79.0 75.0		55.0 66.0	101.0	61.0 72.0	53.0 56.0	52.0 63.0	73.0 75.0		47.0 54.0	1.0		48.0 56.0				54.0 59.0	45.0 43.0	76.0
Oct		61.0	82.0		75.0	113.0	79.0	56.0	73.0	75.0		53.0		34.0	62.0		34.0		59.0		84.0
Oct Oct		71.0	91.0 98.0		82.0	116.0 116.0	81.0 82.0	54.0 52.0	87.0 103.0	75.0 77.0		54.0 54.0	6.0	42.0	69.0 75.0		48.0		59.0 59.0		100.0
Oct	44.1	81.0	102.0		91.0	117.0	84.0	52.0	108.0	78.0		52.0		43.0	83.0	58.0	57.0		56.0	57.0	128.0
Oct Oct	47.9 53.5	97.0 115.0	103.0		100.0	117.0	87.0 99.0	47.0 43.0	112.0	80.0 87.0		48.0 54.0	14.0	42.0	103.0	70.0	56.0 57.0		54.0 52.0	68.0 72.0	137.0
let	59.9	137.0	103.0		120.0	142.0	113.0	45.0	128.0	94.0		64.0		42.0	117.0	80.0	60.0	48.0	60.0	72.0	144.0
Oct Oct	67.1 74.9	141.0	103.0	43.0 53.0	129.0	169.0 196.0	116.0 132.0	56.0 72.0	137.0	103.0		74.0		52.0 59.0	119.0	80:0 77.0	64.0 74.0	67.0 79.0	70.0 76.0	68.0	144.0
Oct	81.7	163.0	105.0	62.0	149.0	214.0	141.0	84.0	155.0	125.0	47.0	79.0	14.0	64.0	124.0	73.0	88.0	88.0	78.0	60.0	156.0
Oct Oct	87.9 94.1	176.0 189.0	105.0	66.0 70.0	158.0 170.0	225.0	148.0 152.0	88.0 86.0	171.0 175.0	130.0	47.0 55.0	81.0 92.0		68.0 73.0	129.0 145.0	69.0 71.0	105.0	95.0 98.0	81.0 82.0	56.0 53.0	165.0
-Oct	100.6	198.0	110.0	74.0	188.0	241.0	156.0	89.0	179.0	157.0	60.0	112.0	47.0	78.0	154.0	73.0	122.0	105.0	82.0		172.0
Oct Oct	108.4 117.5	201.0	124.0	81.0 84.0	195.0	248.0 253.0	163.0	103.0	183.0 196.0	176.0 196.0	64.0	125.0 137.0	51.0 55.0	80.0	170.0	82.0 86.0	127.0	110.0	87.0 101.0	57.0 70.0	176.0
-Oct	126.9	204.0	168.0	88.0	209.0	255.0	190.0	144.0	206.0	207.0	73.0	160.0	61.0	80.0	212.0	92.0	127.0	123.0	109.0	82.0	182.0
-Oct	135.8 145.9	204.0	184.0	93.0	211.0	257.0	212.0	162.0	220.0	214.0 223.0	80.0	171.0 176.0	61.0	93.0	243.0	94.0 98.0	124.0	130.0	119.0	90.0	194.0
Oct	156.8	239.0	215.0	97.0	238.0	268.0	241.0	187.0	239.0	237.0	80.0	182.0	72.0	125.0	291.0	98.0	141.0	147.0	147.0	99.0	228.0
Oct	167.6 175.5	253.0 256.0	227.0	104.0	260.0 278.0	277.0 284.0	248.0	196.0 207.0	243.0	248.0 257.0	79.0 81.0	195.0 208.0	79.0 83.0	141.0 151.0	305.0 316.0	103.0	164.0 183.0	158.0 164.0	159.0 169.0	108.0	239.0
-Oct -Oct	185.1	257.0	259.0	132.0	287.0	287.0	253.0	225.0	245.0	266.0	96.0	220.0	93.0	165.0	324.0	109.0	197,0	167.0	182.0	125.0	248.0
-Oct	197.5	270.0	278.0	151.0	312.0	294.0	254.0	248.0	249.0	278.0	108.0	249.0	94.0	185.0	334.0	115.0	209.0	171.0	201.0	135.0	249.0
-Oct	210.6	279.0 288.0	298.0 305.0	161.0 168.0	335.0 353.0	301.0	257,0	262.0	258.0 267.0	287.0	134.0 162.0	275.0 302.0	117.0	213.0	347.0 364.0	116.0	218.0	174.0 176.0	217.0	143.0	255.0
-Oct	233.9	295.0	316.0	171.0	366.0	312.0	284.0	298.0	281.0	294.0	183.0	322.0	152.0	240.0	381.0	119.0	229.0	178.0	264.0	146.0	281.0
7-0ct 3-0ct	247.0 260.4	302.0	341.0	174.0 177.0	373.0	316.0 318.0	315.0 342.0	311.0	299.0 324.0	295.0	202.0	338.0 357.0	167.0 179.0	263.0	403.0	125.0 133.0	232.0	190.0 216.0	277.0	144.0	307.0
	273.0	303.0	386.0	181.0	392.0	325.0	361.0	325.0	336.0 341.0	306.0 310.0	240.0	367.0	199.0	318.0	421.0	147.0	263.0	242.0	302.0	158.0	323.0
0-Oct	284.0	303.0	402.0	181.0	399.0	330.0	371.0	339.0				376.0	214.0	335.0		158.0	284.0	268.0	306.0	172.0	335.0

Question 3.2

Using the tables and charts from my answer to 3.1, I'd make the judgement that Atlanta's summer climate hasn't gotten warmer. The summer end dates appear to fall within normal variation.

Analysis was done using Alteryx and Tableau

