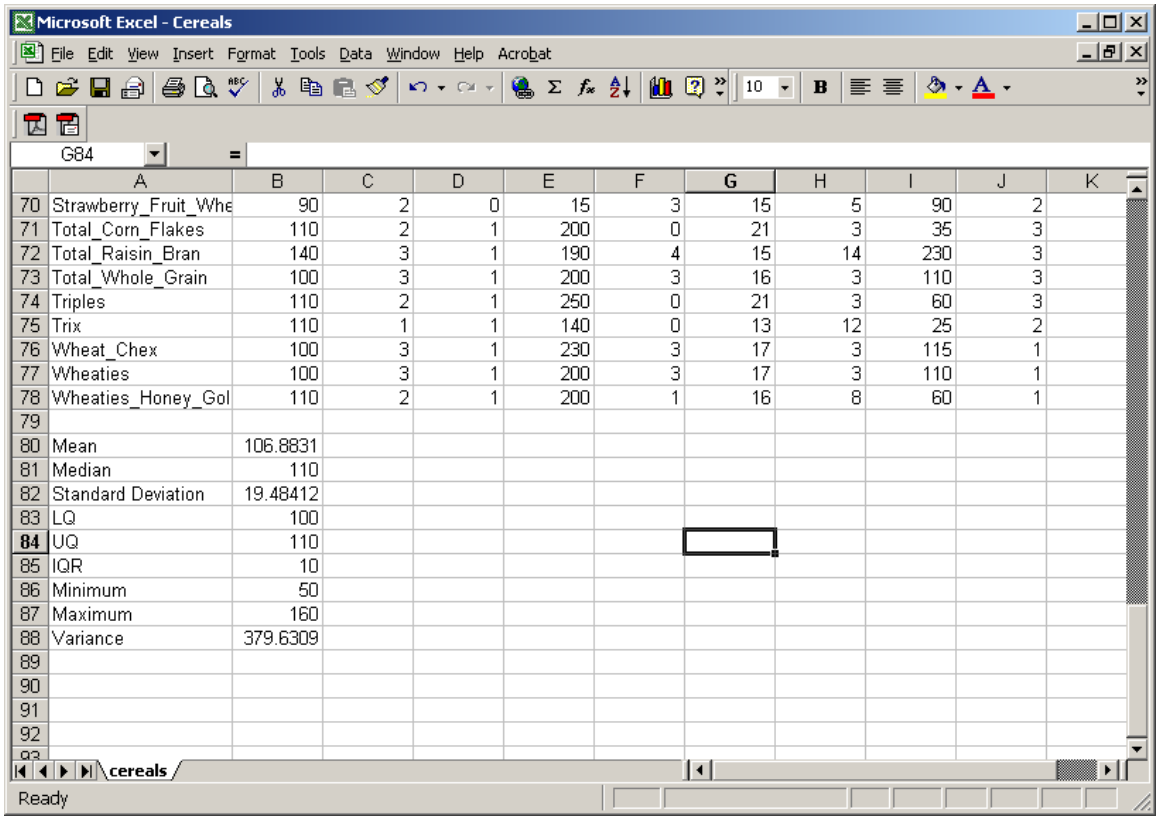
MATH 124 – Computing summary statistics

bolstad_math124@bmbolstad.com
http://math124sfsu.bmbolstad.com

The purpose of this document is to guide you through the steps needed to compute summary statistics of variables in Excel. Again this document makes use of the cereals data file. You should of course modify these instructions to deal with data for the assignment.

1. First load the spreadsheet data you have already stored.
2. First we will calculate the mean calories. Go to the cell A80 and type the word "Mean". Now move to cell B80 and type "=average(B2:B78)". Press enter the mean value should now appear.
3. Now we calculate the median calories. Go to the cell A81 and type the word "Median". Next move to the cell B81 and type "=median(B2:B78)". Press enter and the median value should appear.
4. Now the Standard deviation. Go the cell A82 and type the words "Standard deviation". Next move to the cell B82 and type "=stdev(B2:B78)"
5. The lower quartile. Go to the cell A83 and type the word "LQ". Next move to the cell B83 and type "=quartile(B2:B78,1)
6. The Upper quartile. Go to the cell A84 and type the word "UQ". Next move to the cell B84 and type "=quartile(B2:B78,3)
7. The IQR. Go to the cell A85 and type the word IQR. Next move to the cell B85 and type "=B84-B83"
8. The minimum. Go to the cell A86 and type the word "Minimum". Next move to the cell B86 and type "=min(B2:B78)"
9. The maximum. Go to the cell A87 and type the word "Maximum". Next move to the cell B87 and type "=max(B2:B78)"
10. The variance. Go to the cell A88 and type the word "Variance". Next move to the cell B88 and type "=var(B2:B78)"
11. Your spreadsheet window should look like the following:



	A	B	C	D	E	F	G	H	I	J	K
70	Strawberry_Fruit_Whe	90	2	0	15	3	15	5	90	2	
71	Total_Corn_Flakes	110	2	1	200	0	21	3	35	3	
72	Total_Raisin_Bran	140	3	1	190	4	15	14	230	3	
73	Total_Whole_Grain	100	3	1	200	3	16	3	110	3	
74	Triples	110	2	1	250	0	21	3	60	3	
75	Trix	110	1	1	140	0	13	12	25	2	
76	Wheat_Chex	100	3	1	230	3	17	3	115	1	
77	Wheaties	100	3	1	200	3	17	3	110	1	
78	Wheaties_Honey_Gol	110	2	1	200	1	16	8	60	1	
79											
80	Mean	106.8831									
81	Median	110									
82	Standard Deviation	19.48412									
83	LQ	100									
84	UQ	110									
85	IQR	10									
86	Minimum	50									
87	Maximum	160									
88	Variance	379.6309									
89											
90											
91											
92											

If you want to calculate the summary statistics of the sodium values you should repeat the steps above but use “E2:E78” in place of “B2:B78”. And you could do a similar thing to get the summary values of any of the other variables.

Note that you should probably remove any of the missing data values (the ones with value of –1) by locating those cells and deleting that observation (ie make that square blank).