

CH 3: DATA DESCRIPTION

A number of measures of a data set such as mean, variance, standard deviation, etc, are computed using the Minitab statistical function Descriptive Statistics. For this discussion we will use again the data from example 2 of section 2, chapter 2: E-C02-S02-02 - Temperatures.

- Click on **Stat** → **Basic Statistics** → **Display Descriptive Statistics**
- Click inside the box for **Variables** and then double click the variable **Temperatures** showing on the left of the window.
- Click on **Statistics**
- The following window will appear

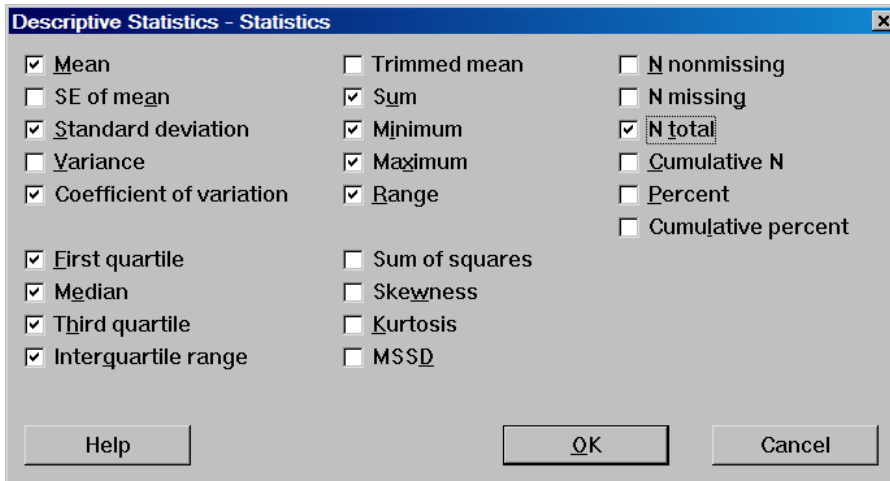


Figure 3.1

- Check the boxes as shown and click OK and OK
- The output will be placed in the Session Window as shown below.

Descriptive Statistics: TEMPERATURES								
Variable	Total Count	Mean	StDev	Variance	CoefVar	Sum	Minimum	Q1
TEMPERATURES	50	114.10	6.23	38.79	5.46	5705.00	100.00	110.00
Variable	Median	Q3	Maximum	Range	IQR			
TEMPERATURES	114.00	118.00	134.00	34.00	8.0			

In the following sections when we come upon a measure that is calculated with this Minitab function, we will simply reference this function without any further elaboration.

3.1 Measures of Central Tendency

The four major measures of Central Tendency are the mean, mode, median, and midrange. They are calculated as follows:

- Mean: Descriptive Statistics
- Median: Descriptive Statistics
- Midrange: Use Calculator: $(\text{Max}(C1) + \text{Min}(C1)) / 2$
- Mode: No easy way. One way is as follows: Tally the column. (in our case C1). This will produce two columns: one with the distinct values in the data set, the other with the frequency of their occurrences. If there are too many rows of data, sort the two columns on the frequency column to see what value has the highest occurrence.

3.2 Measures of Variation

- Sample Variation: Descriptive Statistics
- Sample Standard Deviation: Descriptive Statistics
- Population Variation: Use Calculator and the definition of Variation
- Population Standard Deviation: Use Calculator; square root of Population Variation
- Coefficient of Variation: Descriptive Statistics
- Range Rule of Thumb: Use Calculator: $\text{Range} / 4$

3.3 Measures of Position

- Standard Scores: **Calc** → **Standardize**; The default option for the Standard Scores formula ("Subtract mean and divide by std. Dev") is the one we want. The output column contains the standard scores of each element of the original column.
- Quartiles: Descriptive Statistics
- Interquartile Range (IQR): Descriptive Statistics

3.4 Exploratory Data Analysis

- Five-Number Summary Descriptive Statistics
- Outliers Boxplot (see below)
- Boxplot **Graph** → **Boxplot** and choose **Simple**
Place "Temperatures" in **Graph Variables**
Click on **Labels** and enter a title for the chart .
Click on **Data View** and check the **Interquartile range box** and **Outlier symbols** check boxes. Click OK twice.

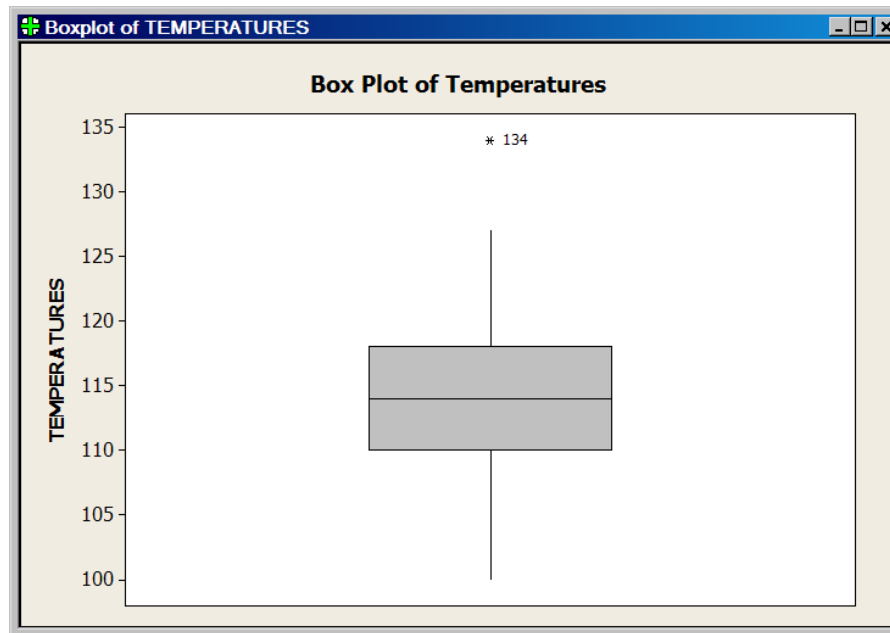


Figure 3.2

Observe that the value 134 was identified as an outlier. You may want to verify this by performing the Outlier calculation described in the book, page 149.