Chapter 4

# Try It Yourself: Determine Outliers Using the Quartile and Mean and Standard Deviation Methods

## Excel Instructions

1. Open videogames.csv in Excel.
2. Verify that global\_sales is in Column J. Use this variable to calculate Mean and Standard Deviation.
   1. Find the mean.
      1. Type into Cell R2 “Mean:”
      2. Type into Cell S2:

=AVERAGE(J:J)

* 1. Find the standard deviation.
     1. Type into Cell R3: “Standard Deviation:”
     2. Type into Cell S3:

=STDEV.S(J:J)

1. Use global\_sales to calculate quartiles and IQR:
   1. Find Q1 (25th Percentile)
      1. Type into Cell R4 “Q1 (25th Percentile)”
      2. Type into Cell S4:

=QUARTILE.INC(J:J, 1)

* 1. Find Q3 (75th Percentile)
     1. Type into Cell R5 “Q3 (75th Percentile)”
     2. Type into Cell S5

=QUARTILE.INC(J:J, 3)

* 1. Find IQR
     1. Type into Cell R6 “IQR (Q3 - Q1)”
     2. Type into Cell S6:

=S5 − S4

1. Calculate the Outlier Thresholds
   1. Quartile Method for Outliers
      1. Mild Low: Q1 – 1.5 \* IQR
         1. Type into Cell R7 “Quartile Mild Low”
         2. Type into Cell S7:

=S4 − 1.5\*S6

* + 1. Mild High: Q3 + 1.5 \* IQR
       1. Type into Cell R8 “Quartile Mild High”
       2. Type into Cell S8:

=S5 + 1.5\*S6

* + 1. Regular Low: Q1 – 3 \* IQR
       1. Type into Cell R9 “Quartile Regular Low”
       2. Type into Cell S9:

=S4 − 3\*S6

* + 1. Regular High: Q3 + 3 \* IQR
       1. Type into Cell R10 “Quartile Regular High”
       2. Type into Cell S10:

=S5 + 3\*S6

* 1. Mean/SD Method for Outliers
     1. Mild Low: Mean – 2 \* SD
        1. Type into Cell R11 “Mean/SD Mild Low”
        2. Type into Cell S11:

=S2 − 2\*S3

* + 1. Mild High: Mean + 2 \* SD
       1. Type into Cell R12 “Mean/SD Mild High”
       2. Type into Cell S12:

=S2 + 2\*S3

* + 1. Regular Low: Mean – 3 \* SD
       1. Type into Cell R13 “Mean/SD Regular Low”
       2. Type into Cell S13:

=S2 − 3\*S3

* + 1. Regular High: Mean + 3 \* SD
       1. Type into Cell R14 “Mean/SD Regular High”
       2. Type into Cell S14:

=S2 + 3\*S3

1. Count the Number of Outliers   
   Use the following formulas to count outliers of each high threshold.
   1. Calculate the Quartile Mild Outlier High
      1. Type into Cell R15   
         “Outlier Count for Quartile Mild High”
      2. Type into Cell S15:

=COUNTIF(J:J, ">" &S8)

* 1. Calculate the Quartile Regular Outlier High
     1. Type into Cell R16   
        “Outlier Count for Quartile Regular High”
     2. Type into Cell S16:

=COUNTIF(J:J, ">" &S10)

* 1. Calculate the Mean/SD Mild Outlier High
     1. Type into Cell R17 “Outlier Count for Mean/SD Mild High”
     2. Type into Cell S17:

=COUNTIF(J:J, ">" &S12)

* 1. Calculate the Mean/SD Regular Outlier High
     1. Type into Cell R18 “Outlier Count for Mean/SD Regular High”
     2. Type into Cell S18:

=COUNTIF(J:J, ">" &S14)