Variables

ORI: Originating Agency Identifier
AgencyName: Agency name assigned by State Police
m1: crime count total from January1 to January31
m2: crime count total from February1 to February28
m3: crime count total from March1 to March31
m4: crime count total from April1 to April30
m5: crime count total from May1 to May31
m6: crime count total from June1 to June30
m7: crime count total from July1 to July31
m8: crime count total from August1 to August31
m9: crime count total from September1 to September30
m10: crime count total from October1 to October31
m11: crime count total from November1 to November30
m12: crime count total from December1 to December31

Steps for using the “Data Quality” macro-enabled Excel workbook.

Step1: Enter data

Data variable needed: ORI (column A), Agency Name (column B), and monthly (January – December) crime count data (columns C through N).

Step1a: Obtain the aggregate monthly IBR property crime data by agency and enter or paste data in the “p” worksheet of the “Data Quality” workbook.

Step1b: Obtain the aggregate monthly IBR violent crime data by agency and enter or paste data in the “v” worksheet of the “Data Quality” workbook.

Step1c: Obtain the aggregate monthly IBR non-index crime data by agency and enter or paste data in the “z” worksheet of the “Data Quality” workbook.

Step2: Zero Classifier

Step2a: Click on the grey "Click Here to CLASSIFY ZEROs" button (available on any of the “p”, “v”, or “z” worksheets) to run the macro built to identify agencies with zero reports suspected as missing data.

Step2b: Classify the highlighted zeros (highlighted in grey) as missing values (the “0” will need to be manually replaced with “.”) or true zero (leave the data as is) according to the displayed guidelines (guidelines are also available in the report). You will need to click “OK” on the displayed guidelines to proceed.

Step3: Outlier Detection

Step3a: Click on the grey "Click Here to generate NEW DATA to check for outliers" button on the “ZeroCheck” worksheet to run the macro that will generate separate worksheets for property and
violent crime data. The macro will create four new worksheets, “OutlierCheckV”, “OutlierCheckP”, “AgencyPlotsV”, and “AgencyPlotsP”.

Step3b: Click on the grey “Click Here to Check For OUTLIERS” button (available on “OutlierCheckV” and “OutlierCheckP” worksheets) to run the macro built to identify agencies with suspected irregular data.

Step 3c: Data suspected as irregular will be highlighted according to the thresholds set for the Yi and Rr outlier detection methods (located at the top of the worksheet). The default threshold/comparison values are set to Yi= 4 (and 0.25) and Rr=2, but can be manually adjusted.

Step 3d: Classify the highlighted data (highlighted in red or orange) as an outlier (the “data” will need to be manually replaced with “.”) according to the data pattern. You will need to click “OK” on the displayed prompt to proceed.

Step 3e: Data visualization may assist with determining whether data is irregular. To visualize data as a histogram, dot plot, and line chart, click on the “Click Here to Go To PLOTS” (available on “OutlierCheckV” and “OutlierCheckP” worksheets) to run the macro built to plot data.

**Step4: Generate Clean Data**

Step4a: Once outliers are identified and the data is replaced with “.”, click on either the “Click Here to Generate Clean Data to use for Analysis” button (available on “OutlierCheckV” and “OutlierCheckP” worksheets) to run the macro built to produce a clean data set to be used for analysis. The new data set will consist of the ORI in column A, Agency Name in column B, and monthly (January – December) crime count data in columns C through N.

Step4b: Copy and Paste data as needed.