Zachary Tankersley

Southern New Hampshire University

DAT-375-T2962

December 4, 2022

**Module Six Assignment**

As part of an exploratory analysis using periodic sampling, I will attempt to answer the following questions:

* The top three crimes from both the sample and larger data set
* Similarities and differences between both analyses
* Describing the purpose of taking samples in analysis projects

1. The top three crime types for the sample data are: Burglary (8 instances), Theft -All Other (6 instances), and Burglary -Residence (5 instances).
2. The top three crime types for the full data set are Burglary (93 instances), Theft -All Other (52 instances), and Criminal Mischief/Vandalism (46 instances).

The two analyses are very close in the first and second place categories while only differing between the third-place categories wherein the sample data indicates that residential burglary is more common, and in the main data criminal mischief/vandalism is more common. For these particular analyses I would go so far as to say that the sample data is not quite as accurate, criminal mischief is likely to be more common in almost any city nationwide than is residential burglary, but not by an impossible margin.

The purpose of sampling is to limit the amount of data that you need to analyze, increasing the speed at which an analysis can be performed, while limiting the resources needed to do so. Depending on the type of analysis, in this case randomized (periodically), sampling can also eliminate issues with the data in regard to things like over representation.