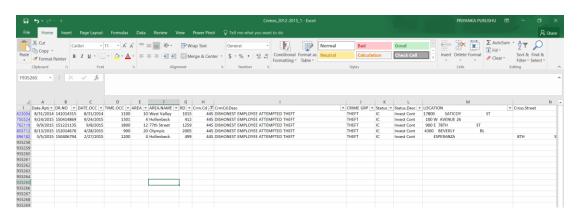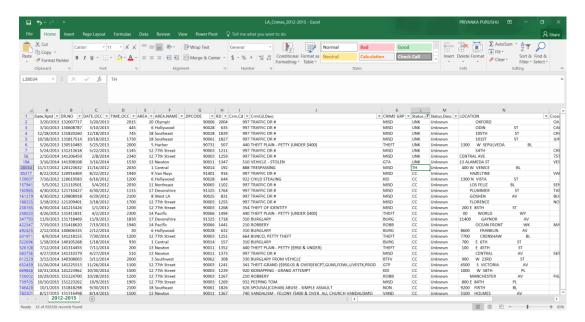# DATA CLEANING

The dataset required very little cleaning to be done. The corrections made on the existing dataset are as follows:
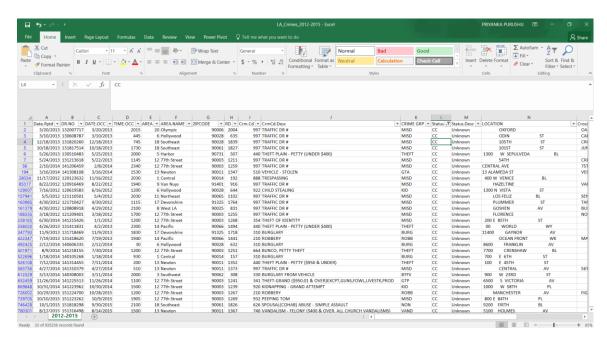
1. Missing Value – Crime description missing

    After cleaning:

2. Mismatch in the status value. UNK is the description for status type CC. TH is unknown.

    Hence status UNK and TH has been replaced with CC because of the same description.

After cleaning:

3. Addition of a new columns named Crime group and zip code. This was done to group the various crime codes which would help in portraying a better data visualization.