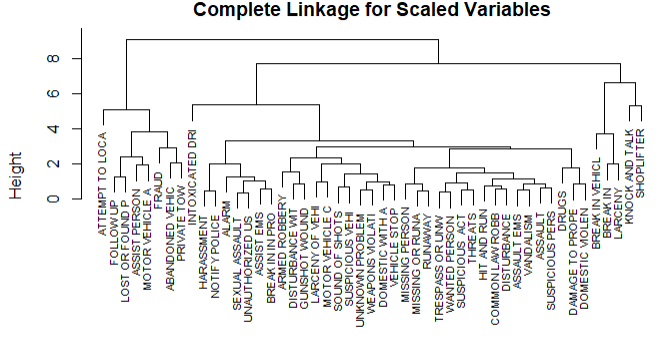
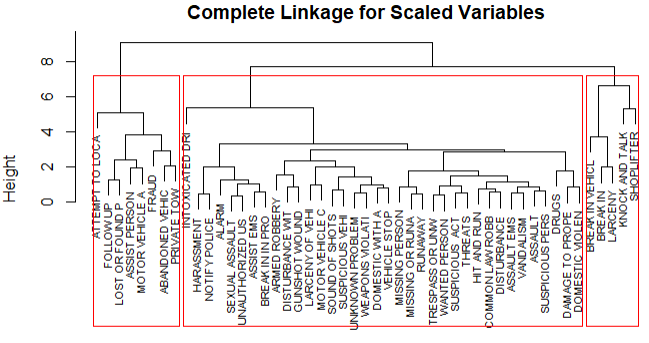
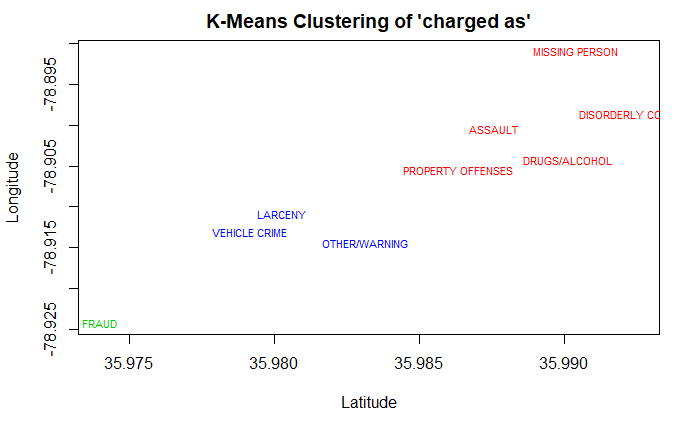
reportedas:

* Hierarchical Clustering
  + Grouped data by “reportedas”, and then calculated summary stats:
    - Counts
    - Latitude mean and sd
    - Longitude mean and sd
    - Hour mean and sd
  + For sake of dendrogram size, only used counts > 150
  + Used complete linkage because distribution is spherical, also used scaled data because variables are measured under different scales (counts vs. latitude/longitude vs hour)
  + 
  + 
  + Findings:
    - Makes sense that home break-ins, vehicle break-ins, larceny, and shoplifting are all clustered together
    - Some other pairs that make sense:
      * Abandoned vehicle/private towing
      * Missing person/runaway
    - Some interesting pairs:
      * Damage to property/domestic violence
      * Sound of shots/suspicious vehicles
* K-means clustering
  + Basically, was almost the exact same dendrogram as hierarchical clustering
  + 98% (48/49) of the crimes were classified into the same cluster in K-means as hierarchical
  + The only difference was knock and talk (residential search request)
    - Make note of this on dendrogram: knock and talk group 3 → 2

\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

Charged as

* K-Means Clustering
  + Grouped data by “charged as” and then calculated summary stats:
    - Mean latitude
    - Mean longitude
  + k=3
  + 
  + F
* Hierarchical clustering
  + Same graph as K-means; 100% similarity
  + Complete linkage of scaled data
  + 