Aggregate Data

* Backwards selection (variable selection)
* Filter out types of crime (crime category)
* Filter out city/cities (Puyallup/South Hill)
* Filter out ZIP code

Build out models

* Decision tree
* Neural network
* Logit (RStudio)

City selected:

* Puyallup/South Hill
  + ZIP code: 98371, 98372, 98373, 98374, 98375
    - Provided by <https://gisdata-piercecowa.opendata.arcgis.com/datasets/zip-codes/data?orderBy=FULLNAME&page=5>

Distance of Crime: <2 miles from parcel -> 0.25 mi -> 1 mi

Categorizing Crime (refer to SQL code):

* Personal
* Property
* Drug
* Homicide
* Other

Create dummy variables from crime category so individual crim category is a variable?

* Help determine the number of type of crime that occurred at a given parcel nuber