* Hedonic Pricing Model?
* Automated Valuation Model (DECISION TREE!!) - Visual Studio
  + Fitting issues?
* Determine key/important attribute(s) for each table
* Crime data (x/ycoord) joined with Tax Parcel (x/ycoord)
  + Where did tax parcel come from?
    - <https://gisdata-piercecowa.opendata.arcgis.com/datasets/tax-parcels/data>
  + Link with a radius
  + Home data???
* Time constraint (past 12 months)

-----

* Crime in pierce county ranked scaled from 1 to 100 (low to high)
* Forecasting tool
* Appraisal and sale; crime and police
  + Police station dataset: <https://gisdata-piercecowa.opendata.arcgis.com/datasets/police-stations/data>
* GIS tools – APIs (lat and long) convert JSON format

-----

Data Preprocessing

* Classification or numerical prediction
  + Where is our problem statement leading to?
  + Crime classification: low, med, high
  + Crime prediction: forecast which areas will have high or lower crime rates?
* Evaluation
  + Classification: confusion matrix (many ratios: accuracy, precision, recall, etc.)
  + Prediction: MSE
* Clustering as a preprocessing tool

Choose and algorithm -> feed in clusters

Association rules

Classify level of crime for location -> decision tree?

Decision Tree -> random forest, Neural Network, Bayesian, Clustering (segmentation) [3]

At least 3 models

10-20 models...

Correlation analysis

(correlation matrix, chi-squared)

US Census Bureau – income data??

* Average income for a given city??

Review Bayesian classification and decision tree

Review SQL Server and Visual Studio