Balance the data set

Use ROSE library

Run two regression with and without subtypes

Median of household??

Make number of houses binary, 1 house or more than 1

Reduce subtypes by looking at the main categories grouped in it

Take median of income by converting it into numeric?