

ETL Project Proposal

Brian Klovert, Yamini Sasidhar, Youqing Ye, Fardi Yueng

For this ETL project, we plan to extract, transform, and load datasets for use in the analysis of housing market trends in New Brunswick. The three datasets we plan to acquire are based on the following variables: property, local education, and crime rates. We plan on using datasets from Kaggle, Data World, and potentially a personal outside resource. We decided to go with the previously mentioned variables as they are generally the most important factors when deciding on which real estate to purchase. The questions we have proposed are as follows:

- Do New Brunswick housing prices go down in parallel with the high crime rates?
- When thinking about local education, do higher quality school systems increase housing prices?
- Is it accurate to say when school systems are strong, crime rates are typically lower?
 - Is that the case for New Brunswick?
- Does the price of a home reduce if the property per square foot is lower?
 - Is that relevant?

Due to the time constraints of this project, we've decided to only look at New Brunswick, but if given more time we would analyze how the variables affect housing prices across different states in the country. We found real estate interesting as the current housing market is in good shape. We want to test if the variables that we've selected play a part in the price being relatively higher or lower.