

This Dataset named House Sales in King County contains all the deal price for houses in King county. It includes houses sold between May 2014 and May 2015. The dataset contains twenty variables along with tens of thousands of records for each. Approximately sixty percent of the variables are quantifiable. Factors such as square feet in bedroom, latitude, and longitude can be considered as continuous, while years, number of floors and grade for condition are discrete variables. The rest of them are mostly binary data, such as waterfront, renovated and backyard.

Overall, the dataset is suitable for evaluating a multivariate linear regression. The response variable is the price of the house in King County, while factors of interest are location (defined by latitude and longitude), square feet for different kinds of rooms, number of rooms, and grade of condition. Logistic regression can be conducted on it as well, for the richness of binary variables. For this scenario, the response variable can be whether the house is waterfront or has a backyard, while independent factors can be any of the variables.