|  |  |
| --- | --- |
| **Categorical Variables** | 28 (56% of total) |
| One category contains at least 90% of observations | 10 (28.0%) |
| No category contains at least 90% of observations | 18 (72.0%) |
| **Continuous or Ordinal Variables** | 22 (44% of total) |
| Containing larger than 90% zero-values | 3 (13.6%) |
| Containing fewer than 90% zero-values | 19 (86.4%) |

**Table 1. Summary of variables in Iowa Housing data set**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **K-fold cross-validation error (K=10)** | | **Bootstrap re-sampling error (B=10)** | |
| **Prediction Method** | Root mean square error (RMSE) | Mean absolute error (MAE) | Root mean square error (RMSE) | Mean absolute error (MAE) |
| Decision tree (CART) | 49322.68 | 33722.26 | 45215.75 | 30833.37 |
| Random forest | 39995.81 | 25200.23 | 33240.84 | 21016.23 |

**Table 3. Summary of variables in Iowa Housing data set**