Prepare Data

Frank Neugebauer

9/10/2019

Preparing Agency Data

The agency data has been prepared and profiled. With this assumption, the data itself is mostly categorical and still contains at least one feature that is not useful (i.e., `master_company`). Each remaining feature must be carefully considered for deletion, transformation, or simply to remain as-is.

The original data set had proprietary information that had to be scrubbed. Here is a summary of that effort:

- The column names were all changed to remove upper case and spaces
- The `account_name` values were all replaced with randomly generated names
- The `master_company` names were all changed to random names
- The `effective_date` values were all altered, but kept in the same quarter
- The `annual_premium` and `written_premium` values were all altered
- The `branch_name` values were all altered to random names
- The `assigned_agent` values were all altered to random names