**Table 12.** Derived attributes were from feature engineering process

|  |  |  |
| --- | --- | --- |
| Name | Type | Description |
| Missing | Numeric | Sum of missing values in all the columns for a given row |
| Calendar Age | Numeric | Derived value SaleYear – Production Year. Filled missing values with -1 |
| Model\_popularity\_count | Numeric | Count of number times the model is sold in a state |
| Base\_model\_popularity\_count | Numeric | Count of number times the model is sold in a state |
| Popularity\_ratio | Numeric | Ratio of model popularity count and base model popularity count |
| State\_sale\_count | Numeric | Count of sales in a given state |
| State\_exp | Numeric | Rank of states by expensiveness of saleprice |
| Month | Categorical | month + 12\*(Sale year -1989 ) |
| Cluster | Categorical (values 0-59) | Derived value [month + 12\*(Saleyear-1989)] % 4 |
| Historical\_mean\_price\_1 | Numeric | Group by modelID & cluster and take mean of the saleprice for each group |
| Historical\_mean\_count\_1 | Numeric | Same as above instead of mean of Saleprice count of SalesID is used for each model cluster grouping |
| Historical\_machine\_price | Numeric | Same as historical model price instead of modelID – MachineID is used |
| Missing\_Blade\_Type | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Coupler | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Drive\_System | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Forks | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Grouser\_Type | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Pattern\_Changer | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_ProductSize | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Ride\_Control | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Ripper | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Stick\_Length | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Thumb | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Tire\_Size | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Track\_Type | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Transmission | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Travel\_Controls | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_Undercarriage\_Pad\_Width | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_UsageBand | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_fiModelDescriptor | Categorical (values 0-1) | Whether the column value is present or not |
| Missing\_fiSecondaryDesc | Categorical (values 0-1) | Whether the column value is present or not |
| Usage | Numeric | Operating Hours divided by (100 + age) column values |