Stats 101A - Final Project Training Data Set

Hao Ma

Winter 2019

Final version of my training data set:

```
dim(train)
```

```
## [1] 2500    85
```

```
colnames(train)
```

```
##  [1] "Ob"           "MSSubClass"   "MSZoning"     "LotFrontage"
##  [5] "LotArea"      "Street"       "Alley"        "LotShape"
##  [9] "LandContour"  "Utilities"    "LotConfig"    "LandSlope"
## [13] "Neighborhood" "Condition1"   "Condition2"   "BldgType"
## [17] "HouseStyle"   "OverallQual"  "OverallCond"  "YearBuilt"
## [21] "YearRemodAdd" "RoofStyle"    "RoofMatl"     "Exterior1st"
## [25] "Exterior2nd"  "MasVnrType"   "MasVnrArea"   "ExterQual"
## [29] "ExterCond"    "Foundation"   "BsmtQual"     "BsmtCond"
## [33] "BsmtExposure" "BsmtFinType1" "BsmtFinSF1"   "BsmtFinType2"
## [37] "BsmtFinSF2"   "BsmtUnfSF"    "TotalBsmtSF"  "Heating"
## [41] "HeatingQC"    "CentralAir"   "Electrical"   "X1stFlrSF"
## [45] "X2ndFlrSF"    "LowQualFinSF" "GrLivArea"    "BsmtFullBath"
## [49] "BsmtHalfBath" "FullBath"     "HalfBath"     "BedroomAbvGr"
## [53] "KitchenAbvGr" "KitchenQual"  "TotRmsAbvGrd" "Functional"
## [57] "Fireplaces"   "FireplaceQu"  "GarageType"   "GarageYrBlt"
## [61] "GarageFinish" "GarageCars"   "GarageArea"   "GarageQual"
## [65] "GarageCond"   "PavedDrive"   "WoodDeckSF"   "OpenPorchSF"
## [69] "EnclosedPorch" "X3SsnPorch"  "ScreenPorch"  "PoolArea"
## [73] "PoolQC"       "Fence"        "MiscFeature"  "MiscVal"
## [77] "MoSold"       "YrSold"       "SaleType"     "SaleCondition"
## [81] "SalePrice"    "Age1"         "Age2"         "Ngroup"
## [85] "MasVnr"
```

```
head(train)
```

```
##   Ob MSSubClass MSZoning LotFrontage LotArea Street Alley LotShape
## 1  1         50       RL          82   12375   Pave  <NA>      Reg
## 2  2         30       RM          60   10800   Pave  Grvl      Reg
## 3  3         45       RM          58    6380   Pave  <NA>      Reg
## 4  4         20       RL          60    7200   Pave  <NA>      Reg
## 5  5         60       FV         100   13162   Pave  <NA>      Reg
## 6  6         80       RL          61    9734   Pave  <NA>      IR1
##   LandContour Utilities LotConfig LandSlope Neighborhood Condition1
## 1         Lvl    AllPub    Inside       Gtl       Sawyer      Feedr
## 2         Lvl    AllPub    Inside       Gtl      OldTown       Norm
## 3         Lvl    AllPub    Inside       Gtl      BrkSide       Norm
## 4         Lvl    AllPub    Inside       Gtl        NAmes       Norm
## 5         Lvl    AllPub    Corner       Gtl      Somerst      Feedr
## 6         Lvl    AllPub    Inside       Gtl      Gilbert       RRAn
##   Condition2 BldgType HouseStyle OverallQual OverallCond YearBuilt
## 1       Norm     1Fam     1.5Fin           5           5      1951
```

```
## 2      Norm    1Fam    1Story          4          7         1885
## 3      Norm    1Fam    1.5Unf          5          6         1922
## 4      Norm    1Fam    1Story          5          8         1950
## 5      Norm    1Fam    2Story          9          5         2006
## 6      Norm    1Fam     SLvl           7          5         2004
##    YearRemodAdd RoofStyle RoofMatl Exterior1st Exterior2nd MasVnrType
## 1          1951     Gable  CompShg      HdBoard      HdBoard      Stone
## 2          1995   Mansard  CompShg      VinylSd      VinylSd       None
## 3          1950     Gable  CompShg      MetalSd      MetalSd       None
## 4          2002     Gable  CompShg      VinylSd      VinylSd       None
## 5          2006     Gable  CompShg      VinylSd      VinylSd       None
## 6          2004     Gable  CompShg      VinylSd      VinylSd       None
##    MasVnrArea ExterQual ExterCond Foundation BsmtQual BsmtCond BsmtExposure
## 1          41        TA        Fa     CBlock       TA       TA           No
## 2           0        TA        TA     BrkTil       Fa       TA           No
## 3           0        TA        TA     BrkTil       TA       Fa           No
## 4           0        TA        TA     CBlock       TA       TA           No
## 5           0        Gd        TA      PConc       Ex       TA           No
## 6           0        Gd        TA      PConc       Gd       TA           Mn
##    BsmtFinType1 BsmtFinSF1 BsmtFinType2 BsmtFinSF2 BsmtUnfSF TotalBsmtSF
## 1           BLQ        329          Unf          0       477         806
## 2           Unf          0          Unf          0       641         641
## 3           Unf          0          Unf          0       993         993
## 4           ALQ        398          BLQ        149       317         864
## 5           GLQ       1836          Unf          0       200        2036
## 6           GLQ        241          Rec        113        30         384
##    Heating HeatingQC CentralAir Electrical X1stFlrSF X2ndFlrSF LowQualFinSF
## 1     GasA        TA          Y      SBrkr      1081       341            0
## 2     GasA        Gd          Y      SBrkr      1047         0            0
## 3     GasA        TA          Y      FuseA      1048         0            0
## 4     GasA        Gd          Y      SBrkr       864         0            0
## 5     GasA        Ex          Y      SBrkr      2036       604            0
## 6     GasA        Ex          Y      SBrkr       744       630            0
##    GrLivArea BsmtFullBath BsmtHalfBath FullBath HalfBath BedroomAbvGr
## 1       1422            1            0        1        0            3
## 2       1047            0            0        1        0            2
## 3       1048            0            0        1        0            2
## 4        864            1            0        1        0            3
## 5       2640            1            0        3        1            3
## 6       1374            0            0        2        1            3
##    KitchenAbvGr KitchenQual TotRmsAbvGrd Functional Fireplaces FireplaceQu
## 1             1          TA            7        Typ          1          TA
## 2             1          TA            6        Typ          0        <NA>
## 3             1          TA            5        Typ          1          Gd
## 4             1          Gd            5        Typ          0        <NA>
## 5             1          Ex           11        Typ          1          Gd
## 6             1          Gd            7        Typ          0        <NA>
##    GarageType GarageYrBlt GarageFinish GarageCars GarageArea GarageQual
## 1      Detchd        1951          Unf          1        288         TA
## 2      Detchd        1954          Unf          1        273         Fa
## 3      Detchd        1922          Unf          1        280         TA
## 4      Detchd        1980          RFn          2        720         TA
## 5      Attchd        2006          RFn          3        792         TA
## 6     BuiltIn        2004          Fin          2        400         TA
```

```
##    GarageCond PavedDrive WoodDeckSF OpenPorchSF EnclosedPorch X3SsnPorch
## 1         TA          Y          0           0             0          0
## 2         Fa          N          0           0             0          0
## 3         TA          Y          0           0           116          0
## 4         TA          Y        194           0             0          0
## 5         TA          Y          0         265             0          0
## 6         TA          Y          0           0             0          0
##    ScreenPorch PoolArea PoolQC Fence MiscFeature MiscVal MoSold YrSold
## 1           0        0   <NA>  GdWo        <NA>       0      6   2006
## 2           0        0   <NA>  <NA>        Shed     450      8   2007
## 3           0        0   <NA>  <NA>        <NA>       0      8   2006
## 4           0        0   <NA>  <NA>        <NA>       0      7   2007
## 5           0        0   <NA>  <NA>        <NA>       0     11   2006
## 6           0        0   <NA>  <NA>        <NA>       0      5   2009
##    SaleType SaleCondition SalePrice Age1 Age2  Ngroup        MasVnr
## 1        WD        Normal  131878.1   68   68 Group 3 Small MasVnr
## 2        WD        Normal  104069.3  134   24 Group 3    No MasVnr
## 3        WD        Normal  116843.7   97   69 Group 3    No MasVnr
## 4        WD        Normal  132932.2   69   17 Group 1    No MasVnr
## 5       New       Partial  429077.4   13   13 Group 1    No MasVnr
## 6        WD        Normal  174735.6   15   15 Group 1    No MasVnr
```