# Lawrence Nagbe

## BAN 502 Project1

### February 26, 2023

Loading all libraries #

library(tidymodels)

## Warning: package 'tidymodels' was built under R version 4.2.2

## ── Attaching packages ────────────────────────────────────── tidymodels 1.0.0 ──

## ✔ broom 1.0.3 ✔ recipes 1.0.5  
## ✔ dials 1.1.0 ✔ rsample 1.1.1  
## ✔ dplyr 1.1.0 ✔ tibble 3.1.8  
## ✔ ggplot2 3.4.1 ✔ tidyr 1.3.0  
## ✔ infer 1.0.4 ✔ tune 1.0.1  
## ✔ modeldata 1.1.0 ✔ workflows 1.1.3  
## ✔ parsnip 1.0.4 ✔ workflowsets 1.0.0  
## ✔ purrr 1.0.1 ✔ yardstick 1.1.0

## Warning: package 'broom' was built under R version 4.2.2

## Warning: package 'dials' was built under R version 4.2.2

## Warning: package 'dplyr' was built under R version 4.2.2

## Warning: package 'ggplot2' was built under R version 4.2.2

## Warning: package 'infer' was built under R version 4.2.2

## Warning: package 'parsnip' was built under R version 4.2.2

## Warning: package 'purrr' was built under R version 4.2.2

## Warning: package 'recipes' was built under R version 4.2.2

## Warning: package 'rsample' was built under R version 4.2.2

## Warning: package 'tidyr' was built under R version 4.2.2

## Warning: package 'tune' was built under R version 4.2.2

## Warning: package 'workflows' was built under R version 4.2.2

## Warning: package 'workflowsets' was built under R version 4.2.2

## Warning: package 'yardstick' was built under R version 4.2.2

## ── Conflicts ───────────────────────────────────────── tidymodels\_conflicts() ──  
## ✖ purrr::discard() masks scales::discard()  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ recipes::step() masks stats::step()  
## • Dig deeper into tidy modeling with R at https://www.tmwr.org

library(tidyverse)

## Warning: package 'tidyverse' was built under R version 4.2.2

## Warning: package 'readr' was built under R version 4.2.2

## Warning: package 'stringr' was built under R version 4.2.2

## Warning: package 'forcats' was built under R version 4.2.2

## Warning: package 'lubridate' was built under R version 4.2.2

## ── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0 ──  
## ✔ forcats 1.0.0 ✔ readr 2.1.4  
## ✔ lubridate 1.9.2 ✔ stringr 1.5.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## ✖ readr::col\_factor() masks scales::col\_factor()  
## ✖ purrr::discard() masks scales::discard()  
## ✖ dplyr::filter() masks stats::filter()  
## ✖ stringr::fixed() masks recipes::fixed()  
## ✖ dplyr::lag() masks stats::lag()  
## ✖ readr::spec() masks yardstick::spec()  
## ℹ Use the ]8;;http://conflicted.r-lib.org/conflicted package]8;; to force all conflicts to become errors

library(lmtest)

## Warning: package 'lmtest' was built under R version 4.2.2

## Loading required package: zoo

## Warning: package 'zoo' was built under R version 4.2.2

##   
## Attaching package: 'zoo'  
##   
## The following objects are masked from 'package:base':  
##   
## as.Date, as.Date.numeric

library(rpart)

## Warning: package 'rpart' was built under R version 4.2.2

##   
## Attaching package: 'rpart'  
##   
## The following object is masked from 'package:dials':  
##   
## prune

library(rpart.plot)

## Warning: package 'rpart.plot' was built under R version 4.2.2

library(rattle)

## Warning: package 'rattle' was built under R version 4.2.2

## Loading required package: bitops  
## Rattle: A free graphical interface for data science with R.  
## Version 5.5.1 Copyright (c) 2006-2021 Togaware Pty Ltd.  
## Type 'rattle()' to shake, rattle, and roll your data.

library(RColorBrewer)  
library(caret)

## Warning: package 'caret' was built under R version 4.2.2

## Loading required package: lattice  
##   
## Attaching package: 'caret'  
##   
## The following objects are masked from 'package:yardstick':  
##   
## precision, recall, sensitivity, specificity  
##   
## The following object is masked from 'package:purrr':  
##   
## lift

library(plyr)

## Warning: package 'plyr' was built under R version 4.2.2

## ------------------------------------------------------------------------------  
## You have loaded plyr after dplyr - this is likely to cause problems.  
## If you need functions from both plyr and dplyr, please load plyr first, then dplyr:  
## library(plyr); library(dplyr)  
## ------------------------------------------------------------------------------  
##   
## Attaching package: 'plyr'  
##   
## The following object is masked from 'package:purrr':  
##   
## compact  
##   
## The following objects are masked from 'package:dplyr':  
##   
## arrange, count, desc, failwith, id, mutate, rename, summarise,  
## summarize

library(mice)

## Warning: package 'mice' was built under R version 4.2.2

##   
## Attaching package: 'mice'  
##   
## The following object is masked from 'package:stats':  
##   
## filter  
##   
## The following objects are masked from 'package:base':  
##   
## cbind, rbind

library(VIM)

## Warning: package 'VIM' was built under R version 4.2.2

## Loading required package: colorspace

## Warning: package 'colorspace' was built under R version 4.2.2

## Loading required package: grid  
## VIM is ready to use.  
##   
## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues  
##   
## Attaching package: 'VIM'  
##   
## The following object is masked from 'package:rattle':  
##   
## wine  
##   
## The following object is masked from 'package:recipes':  
##   
## prepare  
##   
## The following object is masked from 'package:datasets':  
##   
## sleep

library(ggcorrplot)

## Warning: package 'ggcorrplot' was built under R version 4.2.2

library(MASS)

## Warning: package 'MASS' was built under R version 4.2.2

##   
## Attaching package: 'MASS'  
##   
## The following object is masked from 'package:dplyr':  
##   
## select

library(car)

## Warning: package 'car' was built under R version 4.2.2

## Loading required package: carData

## Warning: package 'carData' was built under R version 4.2.2

##   
## Attaching package: 'car'  
##   
## The following object is masked from 'package:purrr':  
##   
## some  
##   
## The following object is masked from 'package:dplyr':  
##   
## recode

library(skimr)

## Warning: package 'skimr' was built under R version 4.2.2

Loading dataset #

ames\_student\_1 <- read\_csv("ames\_student-1.csv",   
 col\_types = cols(Lot\_Frontage = col\_number(),   
 Lot\_Area = col\_number(), Year\_Built = col\_number(),   
 Year\_Remod\_Add = col\_number(), Mas\_Vnr\_Area = col\_number(),   
 BsmtFin\_SF\_1 = col\_number(), BsmtFin\_SF\_2 = col\_number(),   
 Bsmt\_Unf\_SF = col\_number(), Total\_Bsmt\_SF = col\_number(),   
 First\_Flr\_SF = col\_number(), Second\_Flr\_SF = col\_number(),   
 Low\_Qual\_Fin\_SF = col\_number(), Gr\_Liv\_Area = col\_number(),   
 Bsmt\_Full\_Bath = col\_number(), Bsmt\_Half\_Bath = col\_number(),   
 Full\_Bath = col\_number(), Half\_Bath = col\_number()))

Reviewing the structure and summary

str(ames\_student\_1)

## spc\_tbl\_ [2,053 × 81] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ MS\_SubClass : chr [1:2053] "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" ...  
## $ MS\_Zoning : chr [1:2053] "Residential\_Low\_Density" "Residential\_High\_Density" "Residential\_Low\_Density" "Residential\_Low\_Density" ...  
## $ Lot\_Frontage : num [1:2053] 141 80 81 93 74 78 43 39 0 85 ...  
## $ Lot\_Area : num [1:2053] 31770 11622 14267 11160 13830 ...  
## $ Street : chr [1:2053] "Pave" "Pave" "Pave" "Pave" ...  
## $ Alley : chr [1:2053] "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" ...  
## $ Lot\_Shape : chr [1:2053] "Slightly\_Irregular" "Regular" "Slightly\_Irregular" "Regular" ...  
## $ Land\_Contour : chr [1:2053] "Lvl" "Lvl" "Lvl" "Lvl" ...  
## $ Utilities : chr [1:2053] "AllPub" "AllPub" "AllPub" "AllPub" ...  
## $ Lot\_Config : chr [1:2053] "Corner" "Inside" "Corner" "Corner" ...  
## $ Land\_Slope : chr [1:2053] "Gtl" "Gtl" "Gtl" "Gtl" ...  
## $ Neighborhood : chr [1:2053] "North\_Ames" "North\_Ames" "North\_Ames" "North\_Ames" ...  
## $ Condition\_1 : chr [1:2053] "Norm" "Feedr" "Norm" "Norm" ...  
## $ Condition\_2 : chr [1:2053] "Norm" "Norm" "Norm" "Norm" ...  
## $ Bldg\_Type : chr [1:2053] "OneFam" "OneFam" "OneFam" "OneFam" ...  
## $ House\_Style : chr [1:2053] "One\_Story" "One\_Story" "One\_Story" "One\_Story" ...  
## $ Overall\_Qual : chr [1:2053] "Above\_Average" "Average" "Above\_Average" "Good" ...  
## $ Overall\_Cond : chr [1:2053] "Average" "Above\_Average" "Above\_Average" "Average" ...  
## $ Year\_Built : num [1:2053] 1960 1961 1958 1968 1997 ...  
## $ Year\_Remod\_Add : num [1:2053] 1960 1961 1958 1968 1998 ...  
## $ Roof\_Style : chr [1:2053] "Hip" "Gable" "Hip" "Hip" ...  
## $ Roof\_Matl : chr [1:2053] "CompShg" "CompShg" "CompShg" "CompShg" ...  
## $ Exterior\_1st : chr [1:2053] "BrkFace" "VinylSd" "Wd Sdng" "BrkFace" ...  
## $ Exterior\_2nd : chr [1:2053] "Plywood" "VinylSd" "Wd Sdng" "BrkFace" ...  
## $ Mas\_Vnr\_Type : chr [1:2053] "Stone" "None" "BrkFace" "None" ...  
## $ Mas\_Vnr\_Area : num [1:2053] 112 0 108 0 0 20 0 0 0 0 ...  
## $ Exter\_Qual : chr [1:2053] "Typical" "Typical" "Typical" "Good" ...  
## $ Exter\_Cond : chr [1:2053] "Typical" "Typical" "Typical" "Typical" ...  
## $ Foundation : chr [1:2053] "CBlock" "CBlock" "CBlock" "CBlock" ...  
## $ Bsmt\_Qual : chr [1:2053] "Typical" "Typical" "Typical" "Typical" ...  
## $ Bsmt\_Cond : chr [1:2053] "Good" "Typical" "Typical" "Typical" ...  
## $ Bsmt\_Exposure : chr [1:2053] "Gd" "No" "No" "No" ...  
## $ BsmtFin\_Type\_1 : chr [1:2053] "BLQ" "Rec" "ALQ" "ALQ" ...  
## $ BsmtFin\_SF\_1 : num [1:2053] 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_Type\_2 : chr [1:2053] "Unf" "LwQ" "Unf" "Unf" ...  
## $ BsmtFin\_SF\_2 : num [1:2053] 0 144 0 0 0 0 0 0 0 0 ...  
## $ Bsmt\_Unf\_SF : num [1:2053] 441 270 406 1045 137 ...  
## $ Total\_Bsmt\_SF : num [1:2053] 1080 882 1329 2110 928 ...  
## $ Heating : chr [1:2053] "GasA" "GasA" "GasA" "GasA" ...  
## $ Heating\_QC : chr [1:2053] "Fair" "Typical" "Typical" "Excellent" ...  
## $ Central\_Air : chr [1:2053] "Y" "Y" "Y" "Y" ...  
## $ Electrical : chr [1:2053] "SBrkr" "SBrkr" "SBrkr" "SBrkr" ...  
## $ First\_Flr\_SF : num [1:2053] 1656 896 1329 2110 928 ...  
## $ Second\_Flr\_SF : num [1:2053] 0 0 0 0 701 678 0 0 0 0 ...  
## $ Low\_Qual\_Fin\_SF : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Gr\_Liv\_Area : num [1:2053] 1656 896 1329 2110 1629 ...  
## $ Bsmt\_Full\_Bath : num [1:2053] 1 0 0 1 0 0 0 1 1 1 ...  
## $ Bsmt\_Half\_Bath : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Full\_Bath : num [1:2053] 1 1 1 2 2 2 2 2 2 1 ...  
## $ Half\_Bath : num [1:2053] 0 0 1 1 1 1 0 0 0 1 ...  
## $ Bedroom\_AbvGr : num [1:2053] 3 2 3 3 3 3 2 2 3 2 ...  
## $ Kitchen\_AbvGr : num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Kitchen\_Qual : chr [1:2053] "Typical" "Typical" "Good" "Excellent" ...  
## $ TotRms\_AbvGrd : num [1:2053] 7 5 6 8 6 7 5 5 6 5 ...  
## $ Functional : chr [1:2053] "Typ" "Typ" "Typ" "Typ" ...  
## $ Fireplaces : num [1:2053] 2 0 0 2 1 1 0 1 0 1 ...  
## $ Fireplace\_Qu : chr [1:2053] "Good" "No\_Fireplace" "No\_Fireplace" "Typical" ...  
## $ Garage\_Type : chr [1:2053] "Attchd" "Attchd" "Attchd" "Attchd" ...  
## $ Garage\_Finish : chr [1:2053] "Fin" "Unf" "Unf" "Fin" ...  
## $ Garage\_Cars : num [1:2053] 2 1 1 2 2 2 2 2 2 2 ...  
## $ Garage\_Area : num [1:2053] 528 730 312 522 482 470 506 608 420 506 ...  
## $ Garage\_Qual : chr [1:2053] "Typical" "Typical" "Typical" "Typical" ...  
## $ Garage\_Cond : chr [1:2053] "Typical" "Typical" "Typical" "Typical" ...  
## $ Paved\_Drive : chr [1:2053] "Partial\_Pavement" "Paved" "Paved" "Paved" ...  
## $ Wood\_Deck\_SF : num [1:2053] 210 140 393 0 212 360 0 237 483 192 ...  
## $ Open\_Porch\_SF : num [1:2053] 62 0 36 0 34 36 82 152 21 0 ...  
## $ Enclosed\_Porch : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Three\_season\_porch: num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Screen\_Porch : num [1:2053] 0 120 0 0 0 0 144 0 0 0 ...  
## $ Pool\_Area : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Pool\_QC : chr [1:2053] "No\_Pool" "No\_Pool" "No\_Pool" "No\_Pool" ...  
## $ Fence : chr [1:2053] "No\_Fence" "Minimum\_Privacy" "No\_Fence" "No\_Fence" ...  
## $ Misc\_Feature : chr [1:2053] "None" "None" "Gar2" "None" ...  
## $ Misc\_Val : num [1:2053] 0 0 12500 0 0 0 0 0 500 0 ...  
## $ Mo\_Sold : num [1:2053] 5 6 6 4 3 6 1 3 3 2 ...  
## $ Year\_Sold : num [1:2053] 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...  
## $ Sale\_Type : chr [1:2053] "WD" "WD" "WD" "WD" ...  
## $ Sale\_Condition : chr [1:2053] "Normal" "Normal" "Normal" "Normal" ...  
## $ Longitude : num [1:2053] -93.6 -93.6 -93.6 -93.6 -93.6 ...  
## $ Latitude : num [1:2053] 42.1 42.1 42.1 42.1 42.1 ...  
## $ Above\_Median : chr [1:2053] "Yes" "No" "Yes" "Yes" ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. MS\_SubClass = col\_character(),  
## .. MS\_Zoning = col\_character(),  
## .. Lot\_Frontage = col\_number(),  
## .. Lot\_Area = col\_number(),  
## .. Street = col\_character(),  
## .. Alley = col\_character(),  
## .. Lot\_Shape = col\_character(),  
## .. Land\_Contour = col\_character(),  
## .. Utilities = col\_character(),  
## .. Lot\_Config = col\_character(),  
## .. Land\_Slope = col\_character(),  
## .. Neighborhood = col\_character(),  
## .. Condition\_1 = col\_character(),  
## .. Condition\_2 = col\_character(),  
## .. Bldg\_Type = col\_character(),  
## .. House\_Style = col\_character(),  
## .. Overall\_Qual = col\_character(),  
## .. Overall\_Cond = col\_character(),  
## .. Year\_Built = col\_number(),  
## .. Year\_Remod\_Add = col\_number(),  
## .. Roof\_Style = col\_character(),  
## .. Roof\_Matl = col\_character(),  
## .. Exterior\_1st = col\_character(),  
## .. Exterior\_2nd = col\_character(),  
## .. Mas\_Vnr\_Type = col\_character(),  
## .. Mas\_Vnr\_Area = col\_number(),  
## .. Exter\_Qual = col\_character(),  
## .. Exter\_Cond = col\_character(),  
## .. Foundation = col\_character(),  
## .. Bsmt\_Qual = col\_character(),  
## .. Bsmt\_Cond = col\_character(),  
## .. Bsmt\_Exposure = col\_character(),  
## .. BsmtFin\_Type\_1 = col\_character(),  
## .. BsmtFin\_SF\_1 = col\_number(),  
## .. BsmtFin\_Type\_2 = col\_character(),  
## .. BsmtFin\_SF\_2 = col\_number(),  
## .. Bsmt\_Unf\_SF = col\_number(),  
## .. Total\_Bsmt\_SF = col\_number(),  
## .. Heating = col\_character(),  
## .. Heating\_QC = col\_character(),  
## .. Central\_Air = col\_character(),  
## .. Electrical = col\_character(),  
## .. First\_Flr\_SF = col\_number(),  
## .. Second\_Flr\_SF = col\_number(),  
## .. Low\_Qual\_Fin\_SF = col\_number(),  
## .. Gr\_Liv\_Area = col\_number(),  
## .. Bsmt\_Full\_Bath = col\_number(),  
## .. Bsmt\_Half\_Bath = col\_number(),  
## .. Full\_Bath = col\_number(),  
## .. Half\_Bath = col\_number(),  
## .. Bedroom\_AbvGr = col\_double(),  
## .. Kitchen\_AbvGr = col\_double(),  
## .. Kitchen\_Qual = col\_character(),  
## .. TotRms\_AbvGrd = col\_double(),  
## .. Functional = col\_character(),  
## .. Fireplaces = col\_double(),  
## .. Fireplace\_Qu = col\_character(),  
## .. Garage\_Type = col\_character(),  
## .. Garage\_Finish = col\_character(),  
## .. Garage\_Cars = col\_double(),  
## .. Garage\_Area = col\_double(),  
## .. Garage\_Qual = col\_character(),  
## .. Garage\_Cond = col\_character(),  
## .. Paved\_Drive = col\_character(),  
## .. Wood\_Deck\_SF = col\_double(),  
## .. Open\_Porch\_SF = col\_double(),  
## .. Enclosed\_Porch = col\_double(),  
## .. Three\_season\_porch = col\_double(),  
## .. Screen\_Porch = col\_double(),  
## .. Pool\_Area = col\_double(),  
## .. Pool\_QC = col\_character(),  
## .. Fence = col\_character(),  
## .. Misc\_Feature = col\_character(),  
## .. Misc\_Val = col\_double(),  
## .. Mo\_Sold = col\_double(),  
## .. Year\_Sold = col\_double(),  
## .. Sale\_Type = col\_character(),  
## .. Sale\_Condition = col\_character(),  
## .. Longitude = col\_double(),  
## .. Latitude = col\_double(),  
## .. Above\_Median = col\_character()  
## .. )  
## - attr(\*, "problems")=<externalptr>

summary(ames\_student\_1)

## MS\_SubClass MS\_Zoning Lot\_Frontage Lot\_Area   
## Length:2053 Length:2053 Min. : 0.00 Min. : 1300   
## Class :character Class :character 1st Qu.: 43.00 1st Qu.: 7500   
## Mode :character Mode :character Median : 62.00 Median : 9548   
## Mean : 57.38 Mean : 10258   
## 3rd Qu.: 78.00 3rd Qu.: 11600   
## Max. :313.00 Max. :215245   
## Street Alley Lot\_Shape Land\_Contour   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## Utilities Lot\_Config Land\_Slope Neighborhood   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## Condition\_1 Condition\_2 Bldg\_Type House\_Style   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## Overall\_Qual Overall\_Cond Year\_Built Year\_Remod\_Add  
## Length:2053 Length:2053 Min. :1875 Min. :1950   
## Class :character Class :character 1st Qu.:1953 1st Qu.:1965   
## Mode :character Mode :character Median :1972 Median :1993   
## Mean :1971 Mean :1984   
## 3rd Qu.:2000 3rd Qu.:2004   
## Max. :2010 Max. :2010   
## Roof\_Style Roof\_Matl Exterior\_1st Exterior\_2nd   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## Mas\_Vnr\_Type Mas\_Vnr\_Area Exter\_Qual Exter\_Cond   
## Length:2053 Min. : 0.0 Length:2053 Length:2053   
## Class :character 1st Qu.: 0.0 Class :character Class :character   
## Mode :character Median : 0.0 Mode :character Mode :character   
## Mean : 103.8   
## 3rd Qu.: 164.0   
## Max. :1600.0   
## Foundation Bsmt\_Qual Bsmt\_Cond Bsmt\_Exposure   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
## BsmtFin\_Type\_1 BsmtFin\_SF\_1 BsmtFin\_Type\_2 BsmtFin\_SF\_2   
## Length:2053 Min. :1.00 Length:2053 Min. : 0.00   
## Class :character 1st Qu.:3.00 Class :character 1st Qu.: 0.00   
## Mode :character Median :3.00 Mode :character Median : 0.00   
## Mean :4.21 Mean : 52.57   
## 3rd Qu.:7.00 3rd Qu.: 0.00   
## Max. :7.00 Max. :1526.00   
## Bsmt\_Unf\_SF Total\_Bsmt\_SF Heating Heating\_QC   
## Min. : 0.0 Min. : 0 Length:2053 Length:2053   
## 1st Qu.: 226.0 1st Qu.: 793 Class :character Class :character   
## Median : 460.0 Median : 988 Mode :character Mode :character   
## Mean : 561.2 Mean :1055   
## 3rd Qu.: 801.0 3rd Qu.:1304   
## Max. :2336.0 Max. :5095   
## Central\_Air Electrical First\_Flr\_SF Second\_Flr\_SF   
## Length:2053 Length:2053 Min. : 432 Min. : 0.0   
## Class :character Class :character 1st Qu.: 882 1st Qu.: 0.0   
## Mode :character Mode :character Median :1088 Median : 0.0   
## Mean :1168 Mean : 326.1   
## 3rd Qu.:1402 3rd Qu.: 701.0   
## Max. :5095 Max. :1862.0   
## Low\_Qual\_Fin\_SF Gr\_Liv\_Area Bsmt\_Full\_Bath Bsmt\_Half\_Bath   
## Min. : 0.000 Min. : 480 Min. :0.0000 Min. :0.00000   
## 1st Qu.: 0.000 1st Qu.:1137 1st Qu.:0.0000 1st Qu.:0.00000   
## Median : 0.000 Median :1447 Median :0.0000 Median :0.00000   
## Mean : 4.973 Mean :1499 Mean :0.4301 Mean :0.05796   
## 3rd Qu.: 0.000 3rd Qu.:1737 3rd Qu.:1.0000 3rd Qu.:0.00000   
## Max. :1064.000 Max. :5095 Max. :3.0000 Max. :2.00000   
## Full\_Bath Half\_Bath Bedroom\_AbvGr Kitchen\_AbvGr   
## Min. :0.000 Min. :0.0000 Min. :0.000 Min. :1.000   
## 1st Qu.:1.000 1st Qu.:0.0000 1st Qu.:2.000 1st Qu.:1.000   
## Median :2.000 Median :0.0000 Median :3.000 Median :1.000   
## Mean :1.564 Mean :0.3751 Mean :2.855 Mean :1.047   
## 3rd Qu.:2.000 3rd Qu.:1.0000 3rd Qu.:3.000 3rd Qu.:1.000   
## Max. :4.000 Max. :2.0000 Max. :6.000 Max. :3.000   
## Kitchen\_Qual TotRms\_AbvGrd Functional Fireplaces   
## Length:2053 Min. : 3.000 Length:2053 Min. :0.000   
## Class :character 1st Qu.: 5.000 Class :character 1st Qu.:0.000   
## Mode :character Median : 6.000 Mode :character Median :1.000   
## Mean : 6.442 Mean :0.603   
## 3rd Qu.: 7.000 3rd Qu.:1.000   
## Max. :15.000 Max. :4.000   
## Fireplace\_Qu Garage\_Type Garage\_Finish Garage\_Cars   
## Length:2053 Length:2053 Length:2053 Min. :0.000   
## Class :character Class :character Class :character 1st Qu.:1.000   
## Mode :character Mode :character Mode :character Median :2.000   
## Mean :1.774   
## 3rd Qu.:2.000   
## Max. :5.000   
## Garage\_Area Garage\_Qual Garage\_Cond Paved\_Drive   
## Min. : 0 Length:2053 Length:2053 Length:2053   
## 1st Qu.: 320 Class :character Class :character Class :character   
## Median : 478 Mode :character Mode :character Mode :character   
## Mean : 472   
## 3rd Qu.: 576   
## Max. :1488   
## Wood\_Deck\_SF Open\_Porch\_SF Enclosed\_Porch Three\_season\_porch  
## Min. : 0.00 Min. : 0.00 Min. : 0.00 Min. : 0.000   
## 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 0.00 1st Qu.: 0.000   
## Median : 0.00 Median : 27.00 Median : 0.00 Median : 0.000   
## Mean : 93.52 Mean : 48.17 Mean : 23.02 Mean : 2.799   
## 3rd Qu.: 168.00 3rd Qu.: 72.00 3rd Qu.: 0.00 3rd Qu.: 0.000   
## Max. :1424.00 Max. :742.00 Max. :584.00 Max. :407.000   
## Screen\_Porch Pool\_Area Pool\_QC Fence   
## Min. : 0.00 Min. : 0.000 Length:2053 Length:2053   
## 1st Qu.: 0.00 1st Qu.: 0.000 Class :character Class :character   
## Median : 0.00 Median : 0.000 Mode :character Mode :character   
## Mean : 16.68 Mean : 1.339   
## 3rd Qu.: 0.00 3rd Qu.: 0.000   
## Max. :576.00 Max. :800.000   
## Misc\_Feature Misc\_Val Mo\_Sold Year\_Sold   
## Length:2053 Min. : 0.00 Min. : 1.000 Min. :2006   
## Class :character 1st Qu.: 0.00 1st Qu.: 4.000 1st Qu.:2007   
## Mode :character Median : 0.00 Median : 6.000 Median :2008   
## Mean : 60.12 Mean : 6.189 Mean :2008   
## 3rd Qu.: 0.00 3rd Qu.: 8.000 3rd Qu.:2009   
## Max. :17000.00 Max. :12.000 Max. :2010   
## Sale\_Type Sale\_Condition Longitude Latitude   
## Length:2053 Length:2053 Min. :-93.69 Min. :41.99   
## Class :character Class :character 1st Qu.:-93.66 1st Qu.:42.02   
## Mode :character Mode :character Median :-93.64 Median :42.03   
## Mean :-93.64 Mean :42.03   
## 3rd Qu.:-93.62 3rd Qu.:42.05   
## Max. :-93.58 Max. :42.06   
## Above\_Median   
## Length:2053   
## Class :character   
## Mode :character   
##   
##   
##

Converting Character to Factor

#ames\_student\_1 = ames\_student\_1 %>% mutate(MS\_SubClass = as.factor(MS\_SubClass))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(MS\_Zoning = as.factor(MS\_Zoning))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Street = as.factor(Street))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Alley = as.factor(Alley))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Lot\_Shape = as.factor(Lot\_Shape))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Land\_Contour = as.factor(Land\_Contour))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Utilities = as.factor(Utilities))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Lot\_Config = as.factor(Lot\_Config))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Land\_Slope = as.factor(Land\_Slope))  
#ames\_student\_1 = ames\_student\_1 %>% mutate(Neighborhood = as.factor(Neighborhood))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Condition\_1 = as.factor(Condition\_1))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Condition\_2 = as.factor(Condition\_2))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Bldg\_Type = as.factor(Bldg\_Type))  
ames\_student\_1 = ames\_student\_1 %>% mutate(House\_Style = as.factor(House\_Style))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Overall\_Qual = as.factor(Overall\_Qual))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Overall\_Cond = as.factor(Overall\_Cond))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Roof\_Style = as.factor(Roof\_Style))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Roof\_Matl = as.factor(Roof\_Matl))   
ames\_student\_1 = ames\_student\_1 %>% mutate(Exterior\_1st = as.factor(Exterior\_1st))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Exterior\_2nd = as.factor(Exterior\_2nd))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Mas\_Vnr\_Type = as.factor(Mas\_Vnr\_Type))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Exter\_Qual = as.factor(Exter\_Qual))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Exter\_Cond = as.factor(Exter\_Cond))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Foundation = as.factor(Foundation))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Bsmt\_Qual = as.factor(Bsmt\_Qual))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Bsmt\_Cond = as.factor(Bsmt\_Cond))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Bsmt\_Exposure = as.factor(Bsmt\_Exposure))  
ames\_student\_1 = ames\_student\_1 %>% mutate(BsmtFin\_Type\_1 = as.factor(BsmtFin\_Type\_1))  
ames\_student\_1 = ames\_student\_1 %>% mutate(BsmtFin\_Type\_2 = as.factor(BsmtFin\_Type\_2))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Central\_Air = as.factor(Central\_Air))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Electrica = as.factor(Electrical))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Kitchen\_Qual = as.factor(Kitchen\_Qual))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Functional = as.factor(Functional))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Fireplace\_Qu = as.factor(Fireplace\_Qu))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Garage\_Finish = as.factor(Garage\_Finish))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Garage\_Cond = as.factor(Garage\_Cond))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Garage\_Qual = as.factor(Garage\_Qual))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Paved\_Drive = as.factor(Paved\_Drive))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Pool\_QC = as.factor(Pool\_QC))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Fence = as.factor(Fence))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Misc\_Feature = as.factor(Misc\_Feature))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Sale\_Type = as.factor(Sale\_Type))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Sale\_Condition = as.factor(Sale\_Condition))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Above\_Median = as.factor(Above\_Median))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Heating\_QC = as.factor(Heating\_QC))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Heating = as.factor(Heating))  
ames\_student\_1 = ames\_student\_1 %>% mutate(Electrical = as.factor(Electrical))

Output of the dataset into factor

str(ames\_student\_1)

## spc\_tbl\_ [2,053 × 82] (S3: spec\_tbl\_df/tbl\_df/tbl/data.frame)  
## $ MS\_SubClass : chr [1:2053] "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" ...  
## $ MS\_Zoning : chr [1:2053] "Residential\_Low\_Density" "Residential\_High\_Density" "Residential\_Low\_Density" "Residential\_Low\_Density" ...  
## $ Lot\_Frontage : num [1:2053] 141 80 81 93 74 78 43 39 0 85 ...  
## $ Lot\_Area : num [1:2053] 31770 11622 14267 11160 13830 ...  
## $ Street : chr [1:2053] "Pave" "Pave" "Pave" "Pave" ...  
## $ Alley : chr [1:2053] "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" ...  
## $ Lot\_Shape : chr [1:2053] "Slightly\_Irregular" "Regular" "Slightly\_Irregular" "Regular" ...  
## $ Land\_Contour : chr [1:2053] "Lvl" "Lvl" "Lvl" "Lvl" ...  
## $ Utilities : chr [1:2053] "AllPub" "AllPub" "AllPub" "AllPub" ...  
## $ Lot\_Config : chr [1:2053] "Corner" "Inside" "Corner" "Corner" ...  
## $ Land\_Slope : chr [1:2053] "Gtl" "Gtl" "Gtl" "Gtl" ...  
## $ Neighborhood : chr [1:2053] "North\_Ames" "North\_Ames" "North\_Ames" "North\_Ames" ...  
## $ Condition\_1 : Factor w/ 9 levels "Artery","Feedr",..: 3 2 3 3 3 3 3 3 3 3 ...  
## $ Condition\_2 : Factor w/ 8 levels "Artery","Feedr",..: 3 3 3 3 3 3 3 3 3 3 ...  
## $ Bldg\_Type : Factor w/ 5 levels "Duplex","OneFam",..: 2 2 2 2 2 2 4 4 2 2 ...  
## $ House\_Style : Factor w/ 8 levels "One\_and\_Half\_Fin",..: 3 3 3 3 8 8 3 3 3 3 ...  
## $ Overall\_Qual : Factor w/ 10 levels "Above\_Average",..: 1 2 1 6 2 1 9 9 1 6 ...  
## $ Overall\_Cond : Factor w/ 9 levels "Above\_Average",..: 2 1 1 2 2 1 2 2 6 2 ...  
## $ Year\_Built : num [1:2053] 1960 1961 1958 1968 1997 ...  
## $ Year\_Remod\_Add : num [1:2053] 1960 1961 1958 1968 1998 ...  
## $ Roof\_Style : Factor w/ 6 levels "Flat","Gable",..: 4 2 4 4 2 2 2 2 2 2 ...  
## $ Roof\_Matl : Factor w/ 6 levels "CompShg","Metal",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Exterior\_1st : Factor w/ 16 levels "AsbShng","AsphShn",..: 4 14 15 4 14 14 7 6 7 7 ...  
## $ Exterior\_2nd : Factor w/ 17 levels "AsbShng","AsphShn",..: 11 15 16 4 15 15 7 6 7 7 ...  
## $ Mas\_Vnr\_Type : Factor w/ 5 levels "BrkCmn","BrkFace",..: 5 4 2 4 4 2 4 4 4 4 ...  
## $ Mas\_Vnr\_Area : num [1:2053] 112 0 108 0 0 20 0 0 0 0 ...  
## $ Exter\_Qual : Factor w/ 4 levels "Excellent","Fair",..: 4 4 4 3 4 4 3 3 4 4 ...  
## $ Exter\_Cond : Factor w/ 5 levels "Excellent","Fair",..: 5 5 5 5 5 5 5 5 3 5 ...  
## $ Foundation : Factor w/ 6 levels "BrkTil","CBlock",..: 2 2 2 2 3 3 3 3 3 3 ...  
## $ Bsmt\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 3 6 3 3 3 3 ...  
## $ Bsmt\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 3 6 6 6 6 6 6 6 6 6 ...  
## $ Bsmt\_Exposure : Factor w/ 5 levels "Av","Gd","Mn",..: 2 4 4 4 4 4 4 4 4 2 ...  
## $ BsmtFin\_Type\_1 : Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_SF\_1 : num [1:2053] 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_Type\_2 : Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 7 4 7 7 7 7 7 7 7 7 ...  
## $ BsmtFin\_SF\_2 : num [1:2053] 0 144 0 0 0 0 0 0 0 0 ...  
## $ Bsmt\_Unf\_SF : num [1:2053] 441 270 406 1045 137 ...  
## $ Total\_Bsmt\_SF : num [1:2053] 1080 882 1329 2110 928 ...  
## $ Heating : Factor w/ 6 levels "Floor","GasA",..: 2 2 2 2 2 2 2 2 2 2 ...  
## $ Heating\_QC : Factor w/ 5 levels "Excellent","Fair",..: 2 5 5 1 3 1 1 1 1 3 ...  
## $ Central\_Air : Factor w/ 2 levels "N","Y": 2 2 2 2 2 2 2 2 2 2 ...  
## $ Electrical : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ First\_Flr\_SF : num [1:2053] 1656 896 1329 2110 928 ...  
## $ Second\_Flr\_SF : num [1:2053] 0 0 0 0 701 678 0 0 0 0 ...  
## $ Low\_Qual\_Fin\_SF : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Gr\_Liv\_Area : num [1:2053] 1656 896 1329 2110 1629 ...  
## $ Bsmt\_Full\_Bath : num [1:2053] 1 0 0 1 0 0 0 1 1 1 ...  
## $ Bsmt\_Half\_Bath : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Full\_Bath : num [1:2053] 1 1 1 2 2 2 2 2 2 1 ...  
## $ Half\_Bath : num [1:2053] 0 0 1 1 1 1 0 0 0 1 ...  
## $ Bedroom\_AbvGr : num [1:2053] 3 2 3 3 3 3 2 2 3 2 ...  
## $ Kitchen\_AbvGr : num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Kitchen\_Qual : Factor w/ 5 levels "Excellent","Fair",..: 5 5 3 1 5 3 3 3 5 3 ...  
## $ TotRms\_AbvGrd : num [1:2053] 7 5 6 8 6 7 5 5 6 5 ...  
## $ Functional : Factor w/ 8 levels "Maj1","Maj2",..: 8 8 8 8 8 8 8 8 8 8 ...  
## $ Fireplaces : num [1:2053] 2 0 0 2 1 1 0 1 0 1 ...  
## $ Fireplace\_Qu : Factor w/ 6 levels "Excellent","Fair",..: 3 4 4 6 6 3 4 6 4 5 ...  
## $ Garage\_Type : chr [1:2053] "Attchd" "Attchd" "Attchd" "Attchd" ...  
## $ Garage\_Finish : Factor w/ 4 levels "Fin","No\_Garage",..: 1 4 4 1 1 1 3 3 1 4 ...  
## $ Garage\_Cars : num [1:2053] 2 1 1 2 2 2 2 2 2 2 ...  
## $ Garage\_Area : num [1:2053] 528 730 312 522 482 470 506 608 420 506 ...  
## $ Garage\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Garage\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Paved\_Drive : Factor w/ 3 levels "Dirt\_Gravel",..: 2 3 3 3 3 3 3 3 3 3 ...  
## $ Wood\_Deck\_SF : num [1:2053] 210 140 393 0 212 360 0 237 483 192 ...  
## $ Open\_Porch\_SF : num [1:2053] 62 0 36 0 34 36 82 152 21 0 ...  
## $ Enclosed\_Porch : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Three\_season\_porch: num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Screen\_Porch : num [1:2053] 0 120 0 0 0 0 144 0 0 0 ...  
## $ Pool\_Area : num [1:2053] 0 0 0 0 0 0 0 0 0 0 ...  
## $ Pool\_QC : Factor w/ 5 levels "Excellent","Fair",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ Fence : Factor w/ 5 levels "Good\_Privacy",..: 5 3 5 5 3 5 5 5 1 5 ...  
## $ Misc\_Feature : Factor w/ 5 levels "Elev","Gar2",..: 3 3 2 3 3 3 3 3 5 3 ...  
## $ Misc\_Val : num [1:2053] 0 0 12500 0 0 0 0 0 500 0 ...  
## $ Mo\_Sold : num [1:2053] 5 6 6 4 3 6 1 3 3 2 ...  
## $ Year\_Sold : num [1:2053] 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...  
## $ Sale\_Type : Factor w/ 10 levels "COD","Con","ConLD",..: 10 10 10 10 10 10 10 10 10 10 ...  
## $ Sale\_Condition : Factor w/ 6 levels "Abnorml","AdjLand",..: 5 5 5 5 5 5 5 5 5 5 ...  
## $ Longitude : num [1:2053] -93.6 -93.6 -93.6 -93.6 -93.6 ...  
## $ Latitude : num [1:2053] 42.1 42.1 42.1 42.1 42.1 ...  
## $ Above\_Median : Factor w/ 2 levels "No","Yes": 2 1 2 2 2 2 2 2 2 2 ...  
## $ Electrica : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...  
## - attr(\*, "spec")=  
## .. cols(  
## .. MS\_SubClass = col\_character(),  
## .. MS\_Zoning = col\_character(),  
## .. Lot\_Frontage = col\_number(),  
## .. Lot\_Area = col\_number(),  
## .. Street = col\_character(),  
## .. Alley = col\_character(),  
## .. Lot\_Shape = col\_character(),  
## .. Land\_Contour = col\_character(),  
## .. Utilities = col\_character(),  
## .. Lot\_Config = col\_character(),  
## .. Land\_Slope = col\_character(),  
## .. Neighborhood = col\_character(),  
## .. Condition\_1 = col\_character(),  
## .. Condition\_2 = col\_character(),  
## .. Bldg\_Type = col\_character(),  
## .. House\_Style = col\_character(),  
## .. Overall\_Qual = col\_character(),  
## .. Overall\_Cond = col\_character(),  
## .. Year\_Built = col\_number(),  
## .. Year\_Remod\_Add = col\_number(),  
## .. Roof\_Style = col\_character(),  
## .. Roof\_Matl = col\_character(),  
## .. Exterior\_1st = col\_character(),  
## .. Exterior\_2nd = col\_character(),  
## .. Mas\_Vnr\_Type = col\_character(),  
## .. Mas\_Vnr\_Area = col\_number(),  
## .. Exter\_Qual = col\_character(),  
## .. Exter\_Cond = col\_character(),  
## .. Foundation = col\_character(),  
## .. Bsmt\_Qual = col\_character(),  
## .. Bsmt\_Cond = col\_character(),  
## .. Bsmt\_Exposure = col\_character(),  
## .. BsmtFin\_Type\_1 = col\_character(),  
## .. BsmtFin\_SF\_1 = col\_number(),  
## .. BsmtFin\_Type\_2 = col\_character(),  
## .. BsmtFin\_SF\_2 = col\_number(),  
## .. Bsmt\_Unf\_SF = col\_number(),  
## .. Total\_Bsmt\_SF = col\_number(),  
## .. Heating = col\_character(),  
## .. Heating\_QC = col\_character(),  
## .. Central\_Air = col\_character(),  
## .. Electrical = col\_character(),  
## .. First\_Flr\_SF = col\_number(),  
## .. Second\_Flr\_SF = col\_number(),  
## .. Low\_Qual\_Fin\_SF = col\_number(),  
## .. Gr\_Liv\_Area = col\_number(),  
## .. Bsmt\_Full\_Bath = col\_number(),  
## .. Bsmt\_Half\_Bath = col\_number(),  
## .. Full\_Bath = col\_number(),  
## .. Half\_Bath = col\_number(),  
## .. Bedroom\_AbvGr = col\_double(),  
## .. Kitchen\_AbvGr = col\_double(),  
## .. Kitchen\_Qual = col\_character(),  
## .. TotRms\_AbvGrd = col\_double(),  
## .. Functional = col\_character(),  
## .. Fireplaces = col\_double(),  
## .. Fireplace\_Qu = col\_character(),  
## .. Garage\_Type = col\_character(),  
## .. Garage\_Finish = col\_character(),  
## .. Garage\_Cars = col\_double(),  
## .. Garage\_Area = col\_double(),  
## .. Garage\_Qual = col\_character(),  
## .. Garage\_Cond = col\_character(),  
## .. Paved\_Drive = col\_character(),  
## .. Wood\_Deck\_SF = col\_double(),  
## .. Open\_Porch\_SF = col\_double(),  
## .. Enclosed\_Porch = col\_double(),  
## .. Three\_season\_porch = col\_double(),  
## .. Screen\_Porch = col\_double(),  
## .. Pool\_Area = col\_double(),  
## .. Pool\_QC = col\_character(),  
## .. Fence = col\_character(),  
## .. Misc\_Feature = col\_character(),  
## .. Misc\_Val = col\_double(),  
## .. Mo\_Sold = col\_double(),  
## .. Year\_Sold = col\_double(),  
## .. Sale\_Type = col\_character(),  
## .. Sale\_Condition = col\_character(),  
## .. Longitude = col\_double(),  
## .. Latitude = col\_double(),  
## .. Above\_Median = col\_character()  
## .. )  
## - attr(\*, "problems")=<externalptr>

summary(ames\_student\_1)

## MS\_SubClass MS\_Zoning Lot\_Frontage Lot\_Area   
## Length:2053 Length:2053 Min. : 0.00 Min. : 1300   
## Class :character Class :character 1st Qu.: 43.00 1st Qu.: 7500   
## Mode :character Mode :character Median : 62.00 Median : 9548   
## Mean : 57.38 Mean : 10258   
## 3rd Qu.: 78.00 3rd Qu.: 11600   
## Max. :313.00 Max. :215245   
##   
## Street Alley Lot\_Shape Land\_Contour   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Utilities Lot\_Config Land\_Slope Neighborhood   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Condition\_1 Condition\_2 Bldg\_Type House\_Style   
## Norm :1771 Norm :2027 Duplex : 76 One\_Story :1052   
## Feedr : 113 Feedr : 12 OneFam :1706 Two\_Story : 590   
## Artery : 67 Artery : 4 Twnhs : 67 One\_and\_Half\_Fin: 225   
## RRAn : 35 PosA : 4 TwnhsE : 157 SLvl : 90   
## PosN : 24 PosN : 3 TwoFmCon: 47 SFoyer : 56   
## RRAe : 19 RRAe : 1 Two\_and\_Half\_Unf: 19   
## (Other): 24 (Other): 2 (Other) : 21   
## Overall\_Qual Overall\_Cond Year\_Built Year\_Remod\_Add  
## Average :587 Average :1143 Min. :1875 Min. :1950   
## Above\_Average:518 Above\_Average: 376 1st Qu.:1953 1st Qu.:1965   
## Good :411 Good : 286 Median :1972 Median :1993   
## Very\_Good :237 Very\_Good : 98 Mean :1971 Mean :1984   
## Below\_Average:169 Below\_Average: 73 3rd Qu.:2000 3rd Qu.:2004   
## Excellent : 70 Fair : 35 Max. :2010 Max. :2010   
## (Other) : 61 (Other) : 42   
## Roof\_Style Roof\_Matl Exterior\_1st Exterior\_2nd Mas\_Vnr\_Type   
## Flat : 14 CompShg:2023 VinylSd:705 VinylSd:699 BrkCmn : 17   
## Gable :1607 Metal : 1 MetalSd:319 MetalSd:317 BrkFace: 638   
## Gambrel: 14 Roll : 1 Wd Sdng:313 Wd Sdng:302 CBlock : 1   
## Hip : 404 Tar&Grv: 17 HdBoard:303 HdBoard:277 None :1231   
## Mansard: 9 WdShake: 8 Plywood:151 Plywood:190 Stone : 166   
## Shed : 5 WdShngl: 3 CemntBd: 90 CmentBd: 90   
## (Other):172 (Other):178   
## Mas\_Vnr\_Area Exter\_Qual Exter\_Cond Foundation   
## Min. : 0.0 Excellent: 78 Excellent: 9 BrkTil:216   
## 1st Qu.: 0.0 Fair : 21 Fair : 43 CBlock:880   
## Median : 0.0 Good : 682 Good : 213 PConc :911   
## Mean : 103.8 Typical :1272 Poor : 1 Slab : 36   
## 3rd Qu.: 164.0 Typical :1787 Stone : 6   
## Max. :1600.0 Wood : 4   
##   
## Bsmt\_Qual Bsmt\_Cond Bsmt\_Exposure BsmtFin\_Type\_1  
## Excellent :178 Excellent : 3 Av : 284 ALQ :298   
## Fair : 57 Fair : 76 Gd : 199 BLQ :196   
## Good :849 Good : 80 Mn : 179 GLQ :578   
## No\_Basement: 57 No\_Basement: 57 No :1331 LwQ :106   
## Poor : 1 Poor : 4 No\_Basement: 60 No\_Basement: 57   
## Typical :911 Typical :1833 Rec :216   
## Unf :602   
## BsmtFin\_SF\_1 BsmtFin\_Type\_2 BsmtFin\_SF\_2 Bsmt\_Unf\_SF   
## Min. :1.00 ALQ : 42 Min. : 0.00 Min. : 0.0   
## 1st Qu.:3.00 BLQ : 47 1st Qu.: 0.00 1st Qu.: 226.0   
## Median :3.00 GLQ : 23 Median : 0.00 Median : 460.0   
## Mean :4.21 LwQ : 64 Mean : 52.57 Mean : 561.2   
## 3rd Qu.:7.00 No\_Basement: 58 3rd Qu.: 0.00 3rd Qu.: 801.0   
## Max. :7.00 Rec : 79 Max. :1526.00 Max. :2336.0   
## Unf :1740   
## Total\_Bsmt\_SF Heating Heating\_QC Central\_Air Electrical   
## Min. : 0 Floor: 1 Excellent:1040 N: 137 FuseA : 126   
## 1st Qu.: 793 GasA :2019 Fair : 61 Y:1916 FuseF : 33   
## Median : 988 GasW : 21 Good : 333 FuseP : 6   
## Mean :1055 Grav : 6 Poor : 1 SBrkr :1887   
## 3rd Qu.:1304 OthW : 1 Typical : 618 Unknown: 1   
## Max. :5095 Wall : 5   
##   
## First\_Flr\_SF Second\_Flr\_SF Low\_Qual\_Fin\_SF Gr\_Liv\_Area   
## Min. : 432 Min. : 0.0 Min. : 0.000 Min. : 480   
## 1st Qu.: 882 1st Qu.: 0.0 1st Qu.: 0.000 1st Qu.:1137   
## Median :1088 Median : 0.0 Median : 0.000 Median :1447   
## Mean :1168 Mean : 326.1 Mean : 4.973 Mean :1499   
## 3rd Qu.:1402 3rd Qu.: 701.0 3rd Qu.: 0.000 3rd Qu.:1737   
## Max. :5095 Max. :1862.0 Max. :1064.000 Max. :5095   
##   
## Bsmt\_Full\_Bath Bsmt\_Half\_Bath Full\_Bath Half\_Bath   
## Min. :0.0000 Min. :0.00000 Min. :0.000 Min. :0.0000   
## 1st Qu.:0.0000 1st Qu.:0.00000 1st Qu.:1.000 1st Qu.:0.0000   
## Median :0.0000 Median :0.00000 Median :2.000 Median :0.0000   
## Mean :0.4301 Mean :0.05796 Mean :1.564 Mean :0.3751   
## 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:2.000 3rd Qu.:1.0000   
## Max. :3.0000 Max. :2.00000 Max. :4.000 Max. :2.0000   
##   
## Bedroom\_AbvGr Kitchen\_AbvGr Kitchen\_Qual TotRms\_AbvGrd   
## Min. :0.000 Min. :1.000 Excellent: 142 Min. : 3.000   
## 1st Qu.:2.000 1st Qu.:1.000 Fair : 50 1st Qu.: 5.000   
## Median :3.000 Median :1.000 Good : 790 Median : 6.000   
## Mean :2.855 Mean :1.047 Poor : 1 Mean : 6.442   
## 3rd Qu.:3.000 3rd Qu.:1.000 Typical :1070 3rd Qu.: 7.000   
## Max. :6.000 Max. :3.000 Max. :15.000   
##   
## Functional Fireplaces Fireplace\_Qu Garage\_Type   
## Typ :1896 Min. :0.000 Excellent : 21 Length:2053   
## Min2 : 54 1st Qu.:0.000 Fair : 56 Class :character   
## Min1 : 51 Median :1.000 Good :538 Mode :character   
## Mod : 27 Mean :0.603 No\_Fireplace:993   
## Maj1 : 15 3rd Qu.:1.000 Poor : 36   
## Maj2 : 6 Max. :4.000 Typical :409   
## (Other): 4   
## Garage\_Finish Garage\_Cars Garage\_Area Garage\_Qual   
## Fin :509 Min. :0.000 Min. : 0 Excellent: 2   
## No\_Garage:109 1st Qu.:1.000 1st Qu.: 320 Fair : 85   
## RFn :563 Median :2.000 Median : 478 Good : 16   
## Unf :872 Mean :1.774 Mean : 472 No\_Garage: 109   
## 3rd Qu.:2.000 3rd Qu.: 576 Poor : 2   
## Max. :5.000 Max. :1488 Typical :1839   
##   
## Garage\_Cond Paved\_Drive Wood\_Deck\_SF Open\_Porch\_SF   
## Excellent: 1 Dirt\_Gravel : 163 Min. : 0.00 Min. : 0.00   
## Fair : 53 Partial\_Pavement: 42 1st Qu.: 0.00 1st Qu.: 0.00   
## Good : 10 Paved :1848 Median : 0.00 Median : 27.00   
## No\_Garage: 109 Mean : 93.52 Mean : 48.17   
## Poor : 8 3rd Qu.: 168.00 3rd Qu.: 72.00   
## Typical :1872 Max. :1424.00 Max. :742.00   
##   
## Enclosed\_Porch Three\_season\_porch Screen\_Porch Pool\_Area   
## Min. : 0.00 Min. : 0.000 Min. : 0.00 Min. : 0.000   
## 1st Qu.: 0.00 1st Qu.: 0.000 1st Qu.: 0.00 1st Qu.: 0.000   
## Median : 0.00 Median : 0.000 Median : 0.00 Median : 0.000   
## Mean : 23.02 Mean : 2.799 Mean : 16.68 Mean : 1.339   
## 3rd Qu.: 0.00 3rd Qu.: 0.000 3rd Qu.: 0.00 3rd Qu.: 0.000   
## Max. :584.00 Max. :407.000 Max. :576.00 Max. :800.000   
##   
## Pool\_QC Fence Misc\_Feature Misc\_Val   
## Excellent: 2 Good\_Privacy : 81 Elev: 1 Min. : 0.00   
## Fair : 1 Good\_Wood : 77 Gar2: 5 1st Qu.: 0.00   
## Good : 1 Minimum\_Privacy : 225 None:1978 Median : 0.00   
## No\_Pool :2047 Minimum\_Wood\_Wire: 9 Othr: 3 Mean : 60.12   
## Typical : 2 No\_Fence :1661 Shed: 66 3rd Qu.: 0.00   
## Max. :17000.00   
##   
## Mo\_Sold Year\_Sold Sale\_Type Sale\_Condition Longitude   
## Min. : 1.000 Min. :2006 WD :1789 Abnorml: 121 Min. :-93.69   
## 1st Qu.: 4.000 1st Qu.:2007 New : 163 AdjLand: 5 1st Qu.:-93.66   
## Median : 6.000 Median :2008 COD : 54 Alloca : 16 Median :-93.64   
## Mean : 6.189 Mean :2008 ConLD : 16 Family : 30 Mean :-93.64   
## 3rd Qu.: 8.000 3rd Qu.:2009 ConLI : 8 Normal :1712 3rd Qu.:-93.62   
## Max. :12.000 Max. :2010 CWD : 8 Partial: 169 Max. :-93.58   
## (Other): 15   
## Latitude Above\_Median Electrica   
## Min. :41.99 No :1010 FuseA : 126   
## 1st Qu.:42.02 Yes:1043 FuseF : 33   
## Median :42.03 FuseP : 6   
## Mean :42.03 SBrkr :1887   
## 3rd Qu.:42.05 Unknown: 1   
## Max. :42.06   
##

Assigning NA’s to all variables with zeros

ames\_student\_1[ames\_student\_1 == 0] <- NA  
ames\_student\_1

## # A tibble: 2,053 × 82  
## MS\_Sub…¹ MS\_Zo…² Lot\_F…³ Lot\_A…⁴ Street Alley Lot\_S…⁵ Land\_…⁶ Utili…⁷ Lot\_C…⁸  
## <chr> <chr> <dbl> <dbl> <chr> <chr> <chr> <chr> <chr> <chr>   
## 1 One\_Sto… Reside… 141 31770 Pave No\_A… Slight… Lvl AllPub Corner   
## 2 One\_Sto… Reside… 80 11622 Pave No\_A… Regular Lvl AllPub Inside   
## 3 One\_Sto… Reside… 81 14267 Pave No\_A… Slight… Lvl AllPub Corner   
## 4 One\_Sto… Reside… 93 11160 Pave No\_A… Regular Lvl AllPub Corner   
## 5 Two\_Sto… Reside… 74 13830 Pave No\_A… Slight… Lvl AllPub Inside   
## 6 Two\_Sto… Reside… 78 9978 Pave No\_A… Slight… Lvl AllPub Inside   
## 7 One\_Sto… Reside… 43 5005 Pave No\_A… Slight… HLS AllPub Inside   
## 8 One\_Sto… Reside… 39 5389 Pave No\_A… Slight… Lvl AllPub Inside   
## 9 One\_Sto… Reside… NA 7980 Pave No\_A… Slight… Lvl AllPub Inside   
## 10 One\_Sto… Reside… 85 10176 Pave No\_A… Regular Lvl AllPub Inside   
## # … with 2,043 more rows, 72 more variables: Land\_Slope <chr>,  
## # Neighborhood <chr>, Condition\_1 <fct>, Condition\_2 <fct>, Bldg\_Type <fct>,  
## # House\_Style <fct>, Overall\_Qual <fct>, Overall\_Cond <fct>,  
## # Year\_Built <dbl>, Year\_Remod\_Add <dbl>, Roof\_Style <fct>, Roof\_Matl <fct>,  
## # Exterior\_1st <fct>, Exterior\_2nd <fct>, Mas\_Vnr\_Type <fct>,  
## # Mas\_Vnr\_Area <dbl>, Exter\_Qual <fct>, Exter\_Cond <fct>, Foundation <fct>,  
## # Bsmt\_Qual <fct>, Bsmt\_Cond <fct>, Bsmt\_Exposure <fct>, …

Using Select Function to delete variables with “0”

ames <- subset(ames\_student\_1, select = - c(Three\_season\_porch, Pool\_Area, Low\_Qual\_Fin\_SF, Bsmt\_Half\_Bath))  
str(ames)

## tibble [2,053 × 78] (S3: tbl\_df/tbl/data.frame)  
## $ MS\_SubClass : chr [1:2053] "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" ...  
## $ MS\_Zoning : chr [1:2053] "Residential\_Low\_Density" "Residential\_High\_Density" "Residential\_Low\_Density" "Residential\_Low\_Density" ...  
## $ Lot\_Frontage : num [1:2053] 141 80 81 93 74 78 43 39 NA 85 ...  
## $ Lot\_Area : num [1:2053] 31770 11622 14267 11160 13830 ...  
## $ Street : chr [1:2053] "Pave" "Pave" "Pave" "Pave" ...  
## $ Alley : chr [1:2053] "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" ...  
## $ Lot\_Shape : chr [1:2053] "Slightly\_Irregular" "Regular" "Slightly\_Irregular" "Regular" ...  
## $ Land\_Contour : chr [1:2053] "Lvl" "Lvl" "Lvl" "Lvl" ...  
## $ Utilities : chr [1:2053] "AllPub" "AllPub" "AllPub" "AllPub" ...  
## $ Lot\_Config : chr [1:2053] "Corner" "Inside" "Corner" "Corner" ...  
## $ Land\_Slope : chr [1:2053] "Gtl" "Gtl" "Gtl" "Gtl" ...  
## $ Neighborhood : chr [1:2053] "North\_Ames" "North\_Ames" "North\_Ames" "North\_Ames" ...  
## $ Condition\_1 : Factor w/ 9 levels "Artery","Feedr",..: 3 2 3 3 3 3 3 3 3 3 ...  
## $ Condition\_2 : Factor w/ 8 levels "Artery","Feedr",..: 3 3 3 3 3 3 3 3 3 3 ...  
## $ Bldg\_Type : Factor w/ 5 levels "Duplex","OneFam",..: 2 2 2 2 2 2 4 4 2 2 ...  
## $ House\_Style : Factor w/ 8 levels "One\_and\_Half\_Fin",..: 3 3 3 3 8 8 3 3 3 3 ...  
## $ Overall\_Qual : Factor w/ 10 levels "Above\_Average",..: 1 2 1 6 2 1 9 9 1 6 ...  
## $ Overall\_Cond : Factor w/ 9 levels "Above\_Average",..: 2 1 1 2 2 1 2 2 6 2 ...  
## $ Year\_Built : num [1:2053] 1960 1961 1958 1968 1997 ...  
## $ Year\_Remod\_Add: num [1:2053] 1960 1961 1958 1968 1998 ...  
## $ Roof\_Style : Factor w/ 6 levels "Flat","Gable",..: 4 2 4 4 2 2 2 2 2 2 ...  
## $ Roof\_Matl : Factor w/ 6 levels "CompShg","Metal",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Exterior\_1st : Factor w/ 16 levels "AsbShng","AsphShn",..: 4 14 15 4 14 14 7 6 7 7 ...  
## $ Exterior\_2nd : Factor w/ 17 levels "AsbShng","AsphShn",..: 11 15 16 4 15 15 7 6 7 7 ...  
## $ Mas\_Vnr\_Type : Factor w/ 5 levels "BrkCmn","BrkFace",..: 5 4 2 4 4 2 4 4 4 4 ...  
## $ Mas\_Vnr\_Area : num [1:2053] 112 NA 108 NA NA 20 NA NA NA NA ...  
## $ Exter\_Qual : Factor w/ 4 levels "Excellent","Fair",..: 4 4 4 3 4 4 3 3 4 4 ...  
## $ Exter\_Cond : Factor w/ 5 levels "Excellent","Fair",..: 5 5 5 5 5 5 5 5 3 5 ...  
## $ Foundation : Factor w/ 6 levels "BrkTil","CBlock",..: 2 2 2 2 3 3 3 3 3 3 ...  
## $ Bsmt\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 3 6 3 3 3 3 ...  
## $ Bsmt\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 3 6 6 6 6 6 6 6 6 6 ...  
## $ Bsmt\_Exposure : Factor w/ 5 levels "Av","Gd","Mn",..: 2 4 4 4 4 4 4 4 4 2 ...  
## $ BsmtFin\_Type\_1: Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_SF\_1 : num [1:2053] 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_Type\_2: Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 7 4 7 7 7 7 7 7 7 7 ...  
## $ BsmtFin\_SF\_2 : num [1:2053] NA 144 NA NA NA NA NA NA NA NA ...  
## $ Bsmt\_Unf\_SF : num [1:2053] 441 270 406 1045 137 ...  
## $ Total\_Bsmt\_SF : num [1:2053] 1080 882 1329 2110 928 ...  
## $ Heating : Factor w/ 6 levels "Floor","GasA",..: 2 2 2 2 2 2 2 2 2 2 ...  
## $ Heating\_QC : Factor w/ 5 levels "Excellent","Fair",..: 2 5 5 1 3 1 1 1 1 3 ...  
## $ Central\_Air : Factor w/ 2 levels "N","Y": 2 2 2 2 2 2 2 2 2 2 ...  
## $ Electrical : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ First\_Flr\_SF : num [1:2053] 1656 896 1329 2110 928 ...  
## $ Second\_Flr\_SF : num [1:2053] NA NA NA NA 701 678 NA NA NA NA ...  
## $ Gr\_Liv\_Area : num [1:2053] 1656 896 1329 2110 1629 ...  
## $ Bsmt\_Full\_Bath: num [1:2053] 1 NA NA 1 NA NA NA 1 1 1 ...  
## $ Full\_Bath : num [1:2053] 1 1 1 2 2 2 2 2 2 1 ...  
## $ Half\_Bath : num [1:2053] NA NA 1 1 1 1 NA NA NA 1 ...  
## $ Bedroom\_AbvGr : num [1:2053] 3 2 3 3 3 3 2 2 3 2 ...  
## $ Kitchen\_AbvGr : num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Kitchen\_Qual : Factor w/ 5 levels "Excellent","Fair",..: 5 5 3 1 5 3 3 3 5 3 ...  
## $ TotRms\_AbvGrd : num [1:2053] 7 5 6 8 6 7 5 5 6 5 ...  
## $ Functional : Factor w/ 8 levels "Maj1","Maj2",..: 8 8 8 8 8 8 8 8 8 8 ...  
## $ Fireplaces : num [1:2053] 2 NA NA 2 1 1 NA 1 NA 1 ...  
## $ Fireplace\_Qu : Factor w/ 6 levels "Excellent","Fair",..: 3 4 4 6 6 3 4 6 4 5 ...  
## $ Garage\_Type : chr [1:2053] "Attchd" "Attchd" "Attchd" "Attchd" ...  
## $ Garage\_Finish : Factor w/ 4 levels "Fin","No\_Garage",..: 1 4 4 1 1 1 3 3 1 4 ...  
## $ Garage\_Cars : num [1:2053] 2 1 1 2 2 2 2 2 2 2 ...  
## $ Garage\_Area : num [1:2053] 528 730 312 522 482 470 506 608 420 506 ...  
## $ Garage\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Garage\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Paved\_Drive : Factor w/ 3 levels "Dirt\_Gravel",..: 2 3 3 3 3 3 3 3 3 3 ...  
## $ Wood\_Deck\_SF : num [1:2053] 210 140 393 NA 212 360 NA 237 483 192 ...  
## $ Open\_Porch\_SF : num [1:2053] 62 NA 36 NA 34 36 82 152 21 NA ...  
## $ Enclosed\_Porch: num [1:2053] NA NA NA NA NA NA NA NA NA NA ...  
## $ Screen\_Porch : num [1:2053] NA 120 NA NA NA NA 144 NA NA NA ...  
## $ Pool\_QC : Factor w/ 5 levels "Excellent","Fair",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ Fence : Factor w/ 5 levels "Good\_Privacy",..: 5 3 5 5 3 5 5 5 1 5 ...  
## $ Misc\_Feature : Factor w/ 5 levels "Elev","Gar2",..: 3 3 2 3 3 3 3 3 5 3 ...  
## $ Misc\_Val : num [1:2053] NA NA 12500 NA NA NA NA NA 500 NA ...  
## $ Mo\_Sold : num [1:2053] 5 6 6 4 3 6 1 3 3 2 ...  
## $ Year\_Sold : num [1:2053] 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...  
## $ Sale\_Type : Factor w/ 10 levels "COD","Con","ConLD",..: 10 10 10 10 10 10 10 10 10 10 ...  
## $ Sale\_Condition: Factor w/ 6 levels "Abnorml","AdjLand",..: 5 5 5 5 5 5 5 5 5 5 ...  
## $ Longitude : num [1:2053] -93.6 -93.6 -93.6 -93.6 -93.6 ...  
## $ Latitude : num [1:2053] 42.1 42.1 42.1 42.1 42.1 ...  
## $ Above\_Median : Factor w/ 2 levels "No","Yes": 2 1 2 2 2 2 2 2 2 2 ...  
## $ Electrica : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...

summary(ames)

## MS\_SubClass MS\_Zoning Lot\_Frontage Lot\_Area   
## Length:2053 Length:2053 Min. : 21.00 Min. : 1300   
## Class :character Class :character 1st Qu.: 59.00 1st Qu.: 7500   
## Mode :character Mode :character Median : 68.00 Median : 9548   
## Mean : 69.13 Mean : 10258   
## 3rd Qu.: 80.00 3rd Qu.: 11600   
## Max. :313.00 Max. :215245   
## NA's :349   
## Street Alley Lot\_Shape Land\_Contour   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Utilities Lot\_Config Land\_Slope Neighborhood   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Condition\_1 Condition\_2 Bldg\_Type House\_Style   
## Norm :1771 Norm :2027 Duplex : 76 One\_Story :1052   
## Feedr : 113 Feedr : 12 OneFam :1706 Two\_Story : 590   
## Artery : 67 Artery : 4 Twnhs : 67 One\_and\_Half\_Fin: 225   
## RRAn : 35 PosA : 4 TwnhsE : 157 SLvl : 90   
## PosN : 24 PosN : 3 TwoFmCon: 47 SFoyer : 56   
## RRAe : 19 RRAe : 1 Two\_and\_Half\_Unf: 19   
## (Other): 24 (Other): 2 (Other) : 21   
## Overall\_Qual Overall\_Cond Year\_Built Year\_Remod\_Add  
## Average :587 Average :1143 Min. :1875 Min. :1950   
## Above\_Average:518 Above\_Average: 376 1st Qu.:1953 1st Qu.:1965   
## Good :411 Good : 286 Median :1972 Median :1993   
## Very\_Good :237 Very\_Good : 98 Mean :1971 Mean :1984   
## Below\_Average:169 Below\_Average: 73 3rd Qu.:2000 3rd Qu.:2004   
## Excellent : 70 Fair : 35 Max. :2010 Max. :2010   
## (Other) : 61 (Other) : 42   
## Roof\_Style Roof\_Matl Exterior\_1st Exterior\_2nd Mas\_Vnr\_Type   
## Flat : 14 CompShg:2023 VinylSd:705 VinylSd:699 BrkCmn : 17   
## Gable :1607 Metal : 1 MetalSd:319 MetalSd:317 BrkFace: 638   
## Gambrel: 14 Roll : 1 Wd Sdng:313 Wd Sdng:302 CBlock : 1   
## Hip : 404 Tar&Grv: 17 HdBoard:303 HdBoard:277 None :1231   
## Mansard: 9 WdShake: 8 Plywood:151 Plywood:190 Stone : 166   
## Shed : 5 WdShngl: 3 CemntBd: 90 CmentBd: 90   
## (Other):172 (Other):178   
## Mas\_Vnr\_Area Exter\_Qual Exter\_Cond Foundation   
## Min. : 1.0 Excellent: 78 Excellent: 9 BrkTil:216   
## 1st Qu.: 120.0 Fair : 21 Fair : 43 CBlock:880   
## Median : 200.0 Good : 682 Good : 213 PConc :911   
## Mean : 257.9 Typical :1272 Poor : 1 Slab : 36   
## 3rd Qu.: 334.5 Typical :1787 Stone : 6   
## Max. :1600.0 Wood : 4   
## NA's :1227   
## Bsmt\_Qual Bsmt\_Cond Bsmt\_Exposure BsmtFin\_Type\_1  
## Excellent :178 Excellent : 3 Av : 284 ALQ :298   
## Fair : 57 Fair : 76 Gd : 199 BLQ :196   
## Good :849 Good : 80 Mn : 179 GLQ :578   
## No\_Basement: 57 No\_Basement: 57 No :1331 LwQ :106   
## Poor : 1 Poor : 4 No\_Basement: 60 No\_Basement: 57   
## Typical :911 Typical :1833 Rec :216   
## Unf :602   
## BsmtFin\_SF\_1 BsmtFin\_Type\_2 BsmtFin\_SF\_2 Bsmt\_Unf\_SF   
## Min. :1.00 ALQ : 42 Min. : 28.0 Min. : 14.0   
## 1st Qu.:3.00 BLQ : 47 1st Qu.: 178.5 1st Qu.: 285.0   
## Median :3.00 GLQ : 23 Median : 360.0 Median : 522.0   
## Mean :4.21 LwQ : 64 Mean : 423.2 Mean : 613.2   
## 3rd Qu.:7.00 No\_Basement: 58 3rd Qu.: 603.0 3rd Qu.: 840.0   
## Max. :7.00 Rec : 79 Max. :1526.0 Max. :2336.0   
## Unf :1740 NA's :1798 NA's :174   
## Total\_Bsmt\_SF Heating Heating\_QC Central\_Air Electrical   
## Min. : 105.0 Floor: 1 Excellent:1040 N: 137 FuseA : 126   
## 1st Qu.: 814.8 GasA :2019 Fair : 61 Y:1916 FuseF : 33   
## Median :1001.0 GasW : 21 Good : 333 FuseP : 6   
## Mean :1084.7 Grav : 6 Poor : 1 SBrkr :1887   
## 3rd Qu.:1313.2 OthW : 1 Typical : 618 Unknown: 1   
## Max. :5095.0 Wall : 5   
## NA's :57   
## First\_Flr\_SF Second\_Flr\_SF Gr\_Liv\_Area Bsmt\_Full\_Bath Full\_Bath   
## Min. : 432 Min. : 110.0 Min. : 480 Min. :1.000 Min. :1.000   
## 1st Qu.: 882 1st Qu.: 595.0 1st Qu.:1137 1st Qu.:1.000 1st Qu.:1.000   
## Median :1088 Median : 748.0 Median :1447 Median :1.000 Median :2.000   
## Mean :1168 Mean : 777.5 Mean :1499 Mean :1.036 Mean :1.572   
## 3rd Qu.:1402 3rd Qu.: 901.0 3rd Qu.:1737 3rd Qu.:1.000 3rd Qu.:2.000   
## Max. :5095 Max. :1862.0 Max. :5095 Max. :3.000 Max. :4.000   
## NA's :1192 NA's :1201 NA's :10   
## Half\_Bath Bedroom\_AbvGr Kitchen\_AbvGr Kitchen\_Qual   
## Min. :1.000 Min. :1.000 Min. :1.000 Excellent: 142   
## 1st Qu.:1.000 1st Qu.:2.000 1st Qu.:1.000 Fair : 50   
## Median :1.000 Median :3.000 Median :1.000 Good : 790   
## Mean :1.023 Mean :2.865 Mean :1.047 Poor : 1   
## 3rd Qu.:1.000 3rd Qu.:3.000 3rd Qu.:1.000 Typical :1070   
## Max. :2.000 Max. :6.000 Max. :3.000   
## NA's :1300 NA's :7   
## TotRms\_AbvGrd Functional Fireplaces Fireplace\_Qu  
## Min. : 3.000 Typ :1896 Min. :1.000 Excellent : 21   
## 1st Qu.: 5.000 Min2 : 54 1st Qu.:1.000 Fair : 56   
## Median : 6.000 Min1 : 51 Median :1.000 Good :538   
## Mean : 6.442 Mod : 27 Mean :1.168 No\_Fireplace:993   
## 3rd Qu.: 7.000 Maj1 : 15 3rd Qu.:1.000 Poor : 36   
## Max. :15.000 Maj2 : 6 Max. :4.000 Typical :409   
## (Other): 4 NA's :993   
## Garage\_Type Garage\_Finish Garage\_Cars Garage\_Area   
## Length:2053 Fin :509 Min. :1.000 Min. : 160.0   
## Class :character No\_Garage:109 1st Qu.:1.000 1st Qu.: 368.0   
## Mode :character RFn :563 Median :2.000 Median : 484.0   
## Unf :872 Mean :1.872 Mean : 498.2   
## 3rd Qu.:2.000 3rd Qu.: 576.0   
## Max. :5.000 Max. :1488.0   
## NA's :108 NA's :108   
## Garage\_Qual Garage\_Cond Paved\_Drive Wood\_Deck\_SF   
## Excellent: 2 Excellent: 1 Dirt\_Gravel : 163 Min. : 12.0   
## Fair : 85 Fair : 53 Partial\_Pavement: 42 1st Qu.: 125.0   
## Good : 16 Good : 10 Paved :1848 Median : 171.0   
## No\_Garage: 109 No\_Garage: 109 Mean : 198.1   
## Poor : 2 Poor : 8 3rd Qu.: 240.0   
## Typical :1839 Typical :1872 Max. :1424.0   
## NA's :1084   
## Open\_Porch\_SF Enclosed\_Porch Screen\_Porch Pool\_QC   
## Min. : 4.00 Min. : 16.0 Min. : 40.0 Excellent: 2   
## 1st Qu.: 38.25 1st Qu.: 87.0 1st Qu.:141.0 Fair : 1   
## Median : 64.00 Median :134.0 Median :174.0 Good : 1   
## Mean : 86.29 Mean :140.2 Mean :185.1 No\_Pool :2047   
## 3rd Qu.:108.00 3rd Qu.:184.0 3rd Qu.:216.0 Typical : 2   
## Max. :742.00 Max. :584.0 Max. :576.0   
## NA's :907 NA's :1716 NA's :1868   
## Fence Misc\_Feature Misc\_Val Mo\_Sold   
## Good\_Privacy : 81 Elev: 1 Min. : 80 Min. : 1.000   
## Good\_Wood : 77 Gar2: 5 1st Qu.: 450 1st Qu.: 4.000   
## Minimum\_Privacy : 225 None:1978 Median : 600 Median : 6.000   
## Minimum\_Wood\_Wire: 9 Othr: 3 Mean : 1668 Mean : 6.189   
## No\_Fence :1661 Shed: 66 3rd Qu.: 1275 3rd Qu.: 8.000   
## Max. :17000 Max. :12.000   
## NA's :1979   
## Year\_Sold Sale\_Type Sale\_Condition Longitude Latitude   
## Min. :2006 WD :1789 Abnorml: 121 Min. :-93.69 Min. :41.99   
## 1st Qu.:2007 New : 163 AdjLand: 5 1st Qu.:-93.66 1st Qu.:42.02   
## Median :2008 COD : 54 Alloca : 16 Median :-93.64 Median :42.03   
## Mean :2008 ConLD : 16 Family : 30 Mean :-93.64 Mean :42.03   
## 3rd Qu.:2009 ConLI : 8 Normal :1712 3rd Qu.:-93.62 3rd Qu.:42.05   
## Max. :2010 CWD : 8 Partial: 169 Max. :-93.58 Max. :42.06   
## (Other): 15   
## Above\_Median Electrica   
## No :1010 FuseA : 126   
## Yes:1043 FuseF : 33   
## FuseP : 6   
## SBrkr :1887   
## Unknown: 1   
##   
##

Changing NAs in to median values final cleaning of missingness

ames$Bsmt\_Full\_Bath[is.na(ames$Bsmt\_Full\_Bath)] <- median(ames$Bsmt\_Full\_Bath, na.rm = TRUE)  
ames$Enclosed\_Porch[is.na(ames$Enclosed\_Porch)] <- median(ames$Enclosed\_Porch, na.rm = TRUE)  
ames$Wood\_Deck\_SF[is.na(ames$Wood\_Deck\_SF)] <- median(ames$Wood\_Deck\_SF, na.rm = TRUE)  
ames$Open\_Porch\_SF[is.na(ames$Open\_Porch\_SF)] <- median(ames$Open\_Porch\_SF, na.rm = TRUE)  
ames$Misc\_Val[is.na(ames$Misc\_Val)] <- median(ames$Misc\_Val, na.rm = TRUE)  
ames$Screen\_Porch[is.na(ames$Screen\_Porch)] <- median(ames$Screen\_Porch, na.rm = TRUE)  
ames$Fireplaces[is.na(ames$Fireplaces)] <- median(ames$Fireplaces, na.rm = TRUE)  
ames$Half\_Bath[is.na(ames$Half\_Bath)] <- median(ames$Half\_Bath, na.rm = TRUE)  
ames$BsmtFin\_SF\_2[is.na(ames$BsmtFin\_SF\_2)] <- median(ames$BsmtFin\_SF\_2, na.rm = TRUE)  
ames$Second\_Flr\_SF[is.na(ames$Second\_Flr\_SF)] <- median(ames$Second\_Flr\_SF, na.rm = TRUE)  
ames$Mas\_Vnr\_Area[is.na(ames$Mas\_Vnr\_Area)] <- median(ames$Mas\_Vnr\_Area, na.rm = TRUE)  
  
str(ames)

## tibble [2,053 × 78] (S3: tbl\_df/tbl/data.frame)  
## $ MS\_SubClass : chr [1:2053] "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" "One\_Story\_1946\_and\_Newer\_All\_Styles" ...  
## $ MS\_Zoning : chr [1:2053] "Residential\_Low\_Density" "Residential\_High\_Density" "Residential\_Low\_Density" "Residential\_Low\_Density" ...  
## $ Lot\_Frontage : num [1:2053] 141 80 81 93 74 78 43 39 NA 85 ...  
## $ Lot\_Area : num [1:2053] 31770 11622 14267 11160 13830 ...  
## $ Street : chr [1:2053] "Pave" "Pave" "Pave" "Pave" ...  
## $ Alley : chr [1:2053] "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" "No\_Alley\_Access" ...  
## $ Lot\_Shape : chr [1:2053] "Slightly\_Irregular" "Regular" "Slightly\_Irregular" "Regular" ...  
## $ Land\_Contour : chr [1:2053] "Lvl" "Lvl" "Lvl" "Lvl" ...  
## $ Utilities : chr [1:2053] "AllPub" "AllPub" "AllPub" "AllPub" ...  
## $ Lot\_Config : chr [1:2053] "Corner" "Inside" "Corner" "Corner" ...  
## $ Land\_Slope : chr [1:2053] "Gtl" "Gtl" "Gtl" "Gtl" ...  
## $ Neighborhood : chr [1:2053] "North\_Ames" "North\_Ames" "North\_Ames" "North\_Ames" ...  
## $ Condition\_1 : Factor w/ 9 levels "Artery","Feedr",..: 3 2 3 3 3 3 3 3 3 3 ...  
## $ Condition\_2 : Factor w/ 8 levels "Artery","Feedr",..: 3 3 3 3 3 3 3 3 3 3 ...  
## $ Bldg\_Type : Factor w/ 5 levels "Duplex","OneFam",..: 2 2 2 2 2 2 4 4 2 2 ...  
## $ House\_Style : Factor w/ 8 levels "One\_and\_Half\_Fin",..: 3 3 3 3 8 8 3 3 3 3 ...  
## $ Overall\_Qual : Factor w/ 10 levels "Above\_Average",..: 1 2 1 6 2 1 9 9 1 6 ...  
## $ Overall\_Cond : Factor w/ 9 levels "Above\_Average",..: 2 1 1 2 2 1 2 2 6 2 ...  
## $ Year\_Built : num [1:2053] 1960 1961 1958 1968 1997 ...  
## $ Year\_Remod\_Add: num [1:2053] 1960 1961 1958 1968 1998 ...  
## $ Roof\_Style : Factor w/ 6 levels "Flat","Gable",..: 4 2 4 4 2 2 2 2 2 2 ...  
## $ Roof\_Matl : Factor w/ 6 levels "CompShg","Metal",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ Exterior\_1st : Factor w/ 16 levels "AsbShng","AsphShn",..: 4 14 15 4 14 14 7 6 7 7 ...  
## $ Exterior\_2nd : Factor w/ 17 levels "AsbShng","AsphShn",..: 11 15 16 4 15 15 7 6 7 7 ...  
## $ Mas\_Vnr\_Type : Factor w/ 5 levels "BrkCmn","BrkFace",..: 5 4 2 4 4 2 4 4 4 4 ...  
## $ Mas\_Vnr\_Area : num [1:2053] 112 200 108 200 200 20 200 200 200 200 ...  
## $ Exter\_Qual : Factor w/ 4 levels "Excellent","Fair",..: 4 4 4 3 4 4 3 3 4 4 ...  
## $ Exter\_Cond : Factor w/ 5 levels "Excellent","Fair",..: 5 5 5 5 5 5 5 5 3 5 ...  
## $ Foundation : Factor w/ 6 levels "BrkTil","CBlock",..: 2 2 2 2 3 3 3 3 3 3 ...  
## $ Bsmt\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 3 6 3 3 3 3 ...  
## $ Bsmt\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 3 6 6 6 6 6 6 6 6 6 ...  
## $ Bsmt\_Exposure : Factor w/ 5 levels "Av","Gd","Mn",..: 2 4 4 4 4 4 4 4 4 2 ...  
## $ BsmtFin\_Type\_1: Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_SF\_1 : num [1:2053] 2 6 1 1 3 3 1 3 1 3 ...  
## $ BsmtFin\_Type\_2: Factor w/ 7 levels "ALQ","BLQ","GLQ",..: 7 4 7 7 7 7 7 7 7 7 ...  
## $ BsmtFin\_SF\_2 : num [1:2053] 360 144 360 360 360 360 360 360 360 360 ...  
## $ Bsmt\_Unf\_SF : num [1:2053] 441 270 406 1045 137 ...  
## $ Total\_Bsmt\_SF : num [1:2053] 1080 882 1329 2110 928 ...  
## $ Heating : Factor w/ 6 levels "Floor","GasA",..: 2 2 2 2 2 2 2 2 2 2 ...  
## $ Heating\_QC : Factor w/ 5 levels "Excellent","Fair",..: 2 5 5 1 3 1 1 1 1 3 ...  
## $ Central\_Air : Factor w/ 2 levels "N","Y": 2 2 2 2 2 2 2 2 2 2 ...  
## $ Electrical : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ First\_Flr\_SF : num [1:2053] 1656 896 1329 2110 928 ...  
## $ Second\_Flr\_SF : num [1:2053] 748 748 748 748 701 678 748 748 748 748 ...  
## $ Gr\_Liv\_Area : num [1:2053] 1656 896 1329 2110 1629 ...  
## $ Bsmt\_Full\_Bath: num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Full\_Bath : num [1:2053] 1 1 1 2 2 2 2 2 2 1 ...  
## $ Half\_Bath : num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Bedroom\_AbvGr : num [1:2053] 3 2 3 3 3 3 2 2 3 2 ...  
## $ Kitchen\_AbvGr : num [1:2053] 1 1 1 1 1 1 1 1 1 1 ...  
## $ Kitchen\_Qual : Factor w/ 5 levels "Excellent","Fair",..: 5 5 3 1 5 3 3 3 5 3 ...  
## $ TotRms\_AbvGrd : num [1:2053] 7 5 6 8 6 7 5 5 6 5 ...  
## $ Functional : Factor w/ 8 levels "Maj1","Maj2",..: 8 8 8 8 8 8 8 8 8 8 ...  
## $ Fireplaces : num [1:2053] 2 1 1 2 1 1 1 1 1 1 ...  
## $ Fireplace\_Qu : Factor w/ 6 levels "Excellent","Fair",..: 3 4 4 6 6 3 4 6 4 5 ...  
## $ Garage\_Type : chr [1:2053] "Attchd" "Attchd" "Attchd" "Attchd" ...  
## $ Garage\_Finish : Factor w/ 4 levels "Fin","No\_Garage",..: 1 4 4 1 1 1 3 3 1 4 ...  
## $ Garage\_Cars : num [1:2053] 2 1 1 2 2 2 2 2 2 2 ...  
## $ Garage\_Area : num [1:2053] 528 730 312 522 482 470 506 608 420 506 ...  
## $ Garage\_Qual : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Garage\_Cond : Factor w/ 6 levels "Excellent","Fair",..: 6 6 6 6 6 6 6 6 6 6 ...  
## $ Paved\_Drive : Factor w/ 3 levels "Dirt\_Gravel",..: 2 3 3 3 3 3 3 3 3 3 ...  
## $ Wood\_Deck\_SF : num [1:2053] 210 140 393 171 212 360 171 237 483 192 ...  
## $ Open\_Porch\_SF : num [1:2053] 62 64 36 64 34 36 82 152 21 64 ...  
## $ Enclosed\_Porch: num [1:2053] 134 134 134 134 134 134 134 134 134 134 ...  
## $ Screen\_Porch : num [1:2053] 174 120 174 174 174 174 144 174 174 174 ...  
## $ Pool\_QC : Factor w/ 5 levels "Excellent","Fair",..: 4 4 4 4 4 4 4 4 4 4 ...  
## $ Fence : Factor w/ 5 levels "Good\_Privacy",..: 5 3 5 5 3 5 5 5 1 5 ...  
## $ Misc\_Feature : Factor w/ 5 levels "Elev","Gar2",..: 3 3 2 3 3 3 3 3 5 3 ...  
## $ Misc\_Val : num [1:2053] 600 600 12500 600 600 600 600 600 500 600 ...  
## $ Mo\_Sold : num [1:2053] 5 6 6 4 3 6 1 3 3 2 ...  
## $ Year\_Sold : num [1:2053] 2010 2010 2010 2010 2010 2010 2010 2010 2010 2010 ...  
## $ Sale\_Type : Factor w/ 10 levels "COD","Con","ConLD",..: 10 10 10 10 10 10 10 10 10 10 ...  
## $ Sale\_Condition: Factor w/ 6 levels "Abnorml","AdjLand",..: 5 5 5 5 5 5 5 5 5 5 ...  
## $ Longitude : num [1:2053] -93.6 -93.6 -93.6 -93.6 -93.6 ...  
## $ Latitude : num [1:2053] 42.1 42.1 42.1 42.1 42.1 ...  
## $ Above\_Median : Factor w/ 2 levels "No","Yes": 2 1 2 2 2 2 2 2 2 2 ...  
## $ Electrica : Factor w/ 5 levels "FuseA","FuseF",..: 4 4 4 4 4 4 4 4 4 4 ...

summary(ames)

## MS\_SubClass MS\_Zoning Lot\_Frontage Lot\_Area   
## Length:2053 Length:2053 Min. : 21.00 Min. : 1300   
## Class :character Class :character 1st Qu.: 59.00 1st Qu.: 7500   
## Mode :character Mode :character Median : 68.00 Median : 9548   
## Mean : 69.13 Mean : 10258   
## 3rd Qu.: 80.00 3rd Qu.: 11600   
## Max. :313.00 Max. :215245   
## NA's :349   
## Street Alley Lot\_Shape Land\_Contour   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Utilities Lot\_Config Land\_Slope Neighborhood   
## Length:2053 Length:2053 Length:2053 Length:2053   
## Class :character Class :character Class :character Class :character   
## Mode :character Mode :character Mode :character Mode :character   
##   
##   
##   
##   
## Condition\_1 Condition\_2 Bldg\_Type House\_Style   
## Norm :1771 Norm :2027 Duplex : 76 One\_Story :1052   
## Feedr : 113 Feedr : 12 OneFam :1706 Two\_Story : 590   
## Artery : 67 Artery : 4 Twnhs : 67 One\_and\_Half\_Fin: 225   
## RRAn : 35 PosA : 4 TwnhsE : 157 SLvl : 90   
## PosN : 24 PosN : 3 TwoFmCon: 47 SFoyer : 56   
## RRAe : 19 RRAe : 1 Two\_and\_Half\_Unf: 19   
## (Other): 24 (Other): 2 (Other) : 21   
## Overall\_Qual Overall\_Cond Year\_Built Year\_Remod\_Add  
## Average :587 Average :1143 Min. :1875 Min. :1950   
## Above\_Average:518 Above\_Average: 376 1st Qu.:1953 1st Qu.:1965   
## Good :411 Good : 286 Median :1972 Median :1993   
## Very\_Good :237 Very\_Good : 98 Mean :1971 Mean :1984   
## Below\_Average:169 Below\_Average: 73 3rd Qu.:2000 3rd Qu.:2004   
## Excellent : 70 Fair : 35 Max. :2010 Max. :2010   
## (Other) : 61 (Other) : 42   
## Roof\_Style Roof\_Matl Exterior\_1st Exterior\_2nd Mas\_Vnr\_Type   
## Flat : 14 CompShg:2023 VinylSd:705 VinylSd:699 BrkCmn : 17   
## Gable :1607 Metal : 1 MetalSd:319 MetalSd:317 BrkFace: 638   
## Gambrel: 14 Roll : 1 Wd Sdng:313 Wd Sdng:302 CBlock : 1   
## Hip : 404 Tar&Grv: 17 HdBoard:303 HdBoard:277 None :1231   
## Mansard: 9 WdShake: 8 Plywood:151 Plywood:190 Stone : 166   
## Shed : 5 WdShngl: 3 CemntBd: 90 CmentBd: 90   
## (Other):172 (Other):178   
## Mas\_Vnr\_Area Exter\_Qual Exter\_Cond Foundation   
## Min. : 1.0 Excellent: 78 Excellent: 9 BrkTil:216   
## 1st Qu.: 200.0 Fair : 21 Fair : 43 CBlock:880   
## Median : 200.0 Good : 682 Good : 213 PConc :911   
## Mean : 223.3 Typical :1272 Poor : 1 Slab : 36   
## 3rd Qu.: 200.0 Typical :1787 Stone : 6   
## Max. :1600.0 Wood : 4   
##   
## Bsmt\_Qual Bsmt\_Cond Bsmt\_Exposure BsmtFin\_Type\_1  
## Excellent :178 Excellent : 3 Av : 284 ALQ :298   
## Fair : 57 Fair : 76 Gd : 199 BLQ :196   
## Good :849 Good : 80 Mn : 179 GLQ :578   
## No\_Basement: 57 No\_Basement: 57 No :1331 LwQ :106   
## Poor : 1 Poor : 4 No\_Basement: 60 No\_Basement: 57   
## Typical :911 Typical :1833 Rec :216   
## Unf :602   
## BsmtFin\_SF\_1 BsmtFin\_Type\_2 BsmtFin\_SF\_2 Bsmt\_Unf\_SF   
## Min. :1.00 ALQ : 42 Min. : 28.0 Min. : 14.0   
## 1st Qu.:3.00 BLQ : 47 1st Qu.: 360.0 1st Qu.: 285.0   
## Median :3.00 GLQ : 23 Median : 360.0 Median : 522.0   
## Mean :4.21 LwQ : 64 Mean : 367.9 Mean : 613.2   
## 3rd Qu.:7.00 No\_Basement: 58 3rd Qu.: 360.0 3rd Qu.: 840.0   
## Max. :7.00 Rec : 79 Max. :1526.0 Max. :2336.0   
## Unf :1740 NA's :174   
## Total\_Bsmt\_SF Heating Heating\_QC Central\_Air Electrical   
## Min. : 105.0 Floor: 1 Excellent:1040 N: 137 FuseA : 126   
## 1st Qu.: 814.8 GasA :2019 Fair : 61 Y:1916 FuseF : 33   
## Median :1001.0 GasW : 21 Good : 333 FuseP : 6   
## Mean :1084.7 Grav : 6 Poor : 1 SBrkr :1887   
## 3rd Qu.:1313.2 OthW : 1 Typical : 618 Unknown: 1   
## Max. :5095.0 Wall : 5   
## NA's :57   
## First\_Flr\_SF Second\_Flr\_SF Gr\_Liv\_Area Bsmt\_Full\_Bath Full\_Bath   
## Min. : 432 Min. : 110.0 Min. : 480 Min. :1.000 Min. :1.000   
## 1st Qu.: 882 1st Qu.: 748.0 1st Qu.:1137 1st Qu.:1.000 1st Qu.:1.000   
## Median :1088 Median : 748.0 Median :1447 Median :1.000 Median :2.000   
## Mean :1168 Mean : 760.4 Mean :1499 Mean :1.015 Mean :1.572   
## 3rd Qu.:1402 3rd Qu.: 748.0 3rd Qu.:1737 3rd Qu.:1.000 3rd Qu.:2.000   
## Max. :5095 Max. :1862.0 Max. :5095 Max. :3.000 Max. :4.000   
## NA's :10   
## Half\_Bath Bedroom\_AbvGr Kitchen\_AbvGr Kitchen\_Qual   
## Min. :1.000 Min. :1.000 Min. :1.000 Excellent: 142   
## 1st Qu.:1.000 1st Qu.:2.000 1st Qu.:1.000 Fair : 50   
## Median :1.000 Median :3.000 Median :1.000 Good : 790   
## Mean :1.008 Mean :2.865 Mean :1.047 Poor : 1   
## 3rd Qu.:1.000 3rd Qu.:3.000 3rd Qu.:1.000 Typical :1070   
## Max. :2.000 Max. :6.000 Max. :3.000   
## NA's :7   
## TotRms\_AbvGrd Functional Fireplaces Fireplace\_Qu  
## Min. : 3.000 Typ :1896 Min. :1.000 Excellent : 21   
## 1st Qu.: 5.000 Min2 : 54 1st Qu.:1.000 Fair : 56   
## Median : 6.000 Min1 : 51 Median :1.000 Good :538   
## Mean : 6.442 Mod : 27 Mean :1.087 No\_Fireplace:993   
## 3rd Qu.: 7.000 Maj1 : 15 3rd Qu.:1.000 Poor : 36   
## Max. :15.000 Maj2 : 6 Max. :4.000 Typical :409   
## (Other): 4   
## Garage\_Type Garage\_Finish Garage\_Cars Garage\_Area   
## Length:2053 Fin :509 Min. :1.000 Min. : 160.0   
## Class :character No\_Garage:109 1st Qu.:1.000 1st Qu.: 368.0   
## Mode :character RFn :563 Median :2.000 Median : 484.0   
## Unf :872 Mean :1.872 Mean : 498.2   
## 3rd Qu.:2.000 3rd Qu.: 576.0   
## Max. :5.000 Max. :1488.0   
## NA's :108 NA's :108   
## Garage\_Qual Garage\_Cond Paved\_Drive Wood\_Deck\_SF   
## Excellent: 2 Excellent: 1 Dirt\_Gravel : 163 Min. : 12.0   
## Fair : 85 Fair : 53 Partial\_Pavement: 42 1st Qu.: 171.0   
## Good : 16 Good : 10 Paved :1848 Median : 171.0   
## No\_Garage: 109 No\_Garage: 109 Mean : 183.8   
## Poor : 2 Poor : 8 3rd Qu.: 171.0   
## Typical :1839 Typical :1872 Max. :1424.0   
##   
## Open\_Porch\_SF Enclosed\_Porch Screen\_Porch Pool\_QC   
## Min. : 4.00 Min. : 16 Min. : 40 Excellent: 2   
## 1st Qu.: 57.00 1st Qu.:134 1st Qu.:174 Fair : 1   
## Median : 64.00 Median :134 Median :174 Good : 1   
## Mean : 76.44 Mean :135 Mean :175 No\_Pool :2047   
## 3rd Qu.: 72.00 3rd Qu.:134 3rd Qu.:174 Typical : 2   
## Max. :742.00 Max. :584 Max. :576   
##   
## Fence Misc\_Feature Misc\_Val Mo\_Sold   
## Good\_Privacy : 81 Elev: 1 Min. : 80.0 Min. : 1.000   
## Good\_Wood : 77 Gar2: 5 1st Qu.: 600.0 1st Qu.: 4.000   
## Minimum\_Privacy : 225 None:1978 Median : 600.0 Median : 6.000   
## Minimum\_Wood\_Wire: 9 Othr: 3 Mean : 638.5 Mean : 6.189   
## No\_Fence :1661 Shed: 66 3rd Qu.: 600.0 3rd Qu.: 8.000   
## Max. :17000.0 Max. :12.000   
##   
## Year\_Sold Sale\_Type Sale\_Condition Longitude Latitude   
## Min. :2006 WD :1789 Abnorml: 121 Min. :-93.69 Min. :41.99   
## 1st Qu.:2007 New : 163 AdjLand: 5 1st Qu.:-93.66 1st Qu.:42.02   
## Median :2008 COD : 54 Alloca : 16 Median :-93.64 Median :42.03   
## Mean :2008 ConLD : 16 Family : 30 Mean :-93.64 Mean :42.03   
## 3rd Qu.:2009 ConLI : 8 Normal :1712 3rd Qu.:-93.62 3rd Qu.:42.05   
## Max. :2010 CWD : 8 Partial: 169 Max. :-93.58 Max. :42.06   
## (Other): 15   
## Above\_Median Electrica   
## No :1010 FuseA : 126   
## Yes:1043 FuseF : 33   
## FuseP : 6   
## SBrkr :1887   
## Unknown: 1   
##   
##

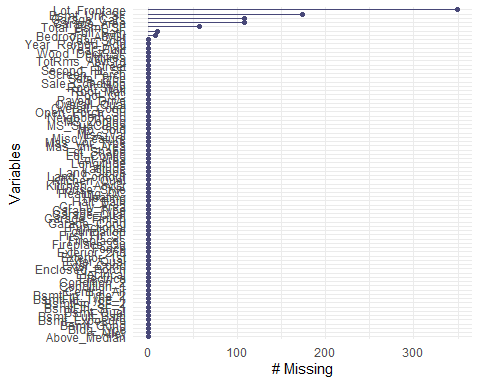
Using visualization to depicts missingness (naniar package)

library(naniar)

##   
## Attaching package: 'naniar'

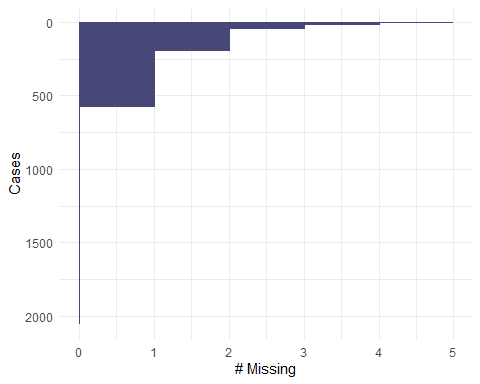
## The following object is masked from 'package:skimr':  
##   
## n\_complete

gg\_miss\_var(ames)

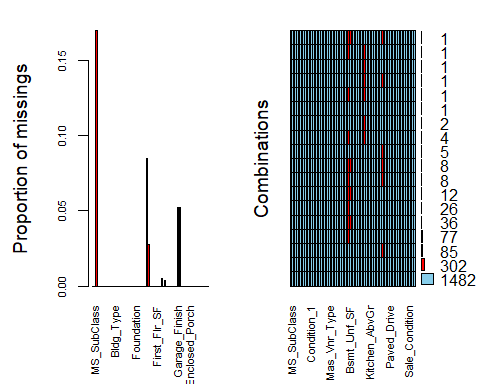


Using visualization to explore the depth of the data.

gg\_miss\_case(ames)

 # The VIM package is used in the visualization below (“aggr)

vim\_plot = aggr(ames, numbers = TRUE, prop = c(TRUE, FALSE), cex.axis=.7)



Visualization using (ggcorr) function to create a correlation matrix, using 15 variables

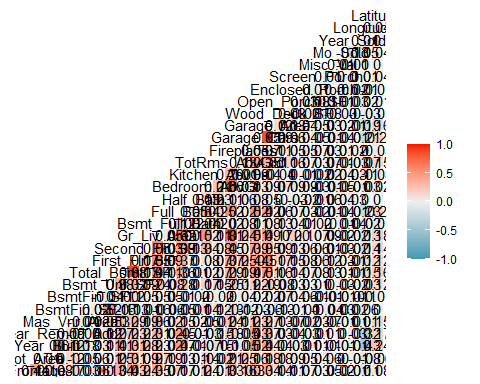
library(GGally)

## Warning: package 'GGally' was built under R version 4.2.2

## Registered S3 method overwritten by 'GGally':  
## method from   
## +.gg ggplot2

ggcorr(ames, label = "TRUE", label\_round = 2)

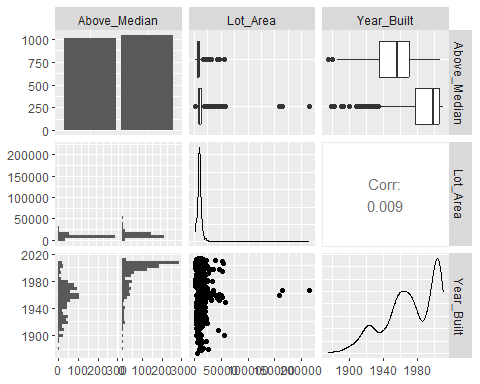
## Warning in ggcorr(ames, label = "TRUE", label\_round = 2): data in column(s)  
## 'MS\_SubClass', 'MS\_Zoning', 'Street', 'Alley', 'Lot\_Shape', 'Land\_Contour',  
## 'Utilities', 'Lot\_Config', 'Land\_Slope', 'Neighborhood', 'Condition\_1',  
## 'Condition\_2', 'Bldg\_Type', 'House\_Style', 'Overall\_Qual', 'Overall\_Cond',  
## 'Roof\_Style', 'Roof\_Matl', 'Exterior\_1st', 'Exterior\_2nd', 'Mas\_Vnr\_Type',  
## 'Exter\_Qual', 'Exter\_Cond', 'Foundation', 'Bsmt\_Qual', 'Bsmt\_Cond',  
## 'Bsmt\_Exposure', 'BsmtFin\_Type\_1', 'BsmtFin\_Type\_2', 'Heating', 'Heating\_QC',  
## 'Central\_Air', 'Electrical', 'Kitchen\_Qual', 'Functional', 'Fireplace\_Qu',  
## 'Garage\_Type', 'Garage\_Finish', 'Garage\_Qual', 'Garage\_Cond', 'Paved\_Drive',  
## 'Pool\_QC', 'Fence', 'Misc\_Feature', 'Sale\_Type', 'Sale\_Condition',  
## 'Above\_Median', 'Electrica' are not numeric and were ignored



Using ggpairs for selected variable visualizations (Above\_Median, Lot\_Area, Year\_ Built) Assessing the correlations between the response variable and two other variables.

ggpairs(ames, columns = c("Above\_Median", "Lot\_Area", "Year\_Built"))

## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.  
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.



Finally, looking at how some Lot\_Area is larger in size and are Above\_Median

ggplot(ames, aes(x = Lot\_Area, y = Above\_Median)) + geom\_boxplot() + geom\_jitter(alpha = 0.15) + theme() +  
 theme(axis.text.x = element\_text(angle = 90))

