# Final Project - BAN 502

## Underwood, Katie

library(tidyverse)

## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --

## v ggplot2 3.3.2 v purrr 0.3.4  
## v tibble 3.0.4 v dplyr 1.0.2  
## v tidyr 1.1.2 v stringr 1.4.0  
## v readr 1.4.0 v forcats 0.5.0

## -- Conflicts ------------------------------------------ tidyverse\_conflicts() --  
## x dplyr::filter() masks stats::filter()  
## x dplyr::lag() masks stats::lag()

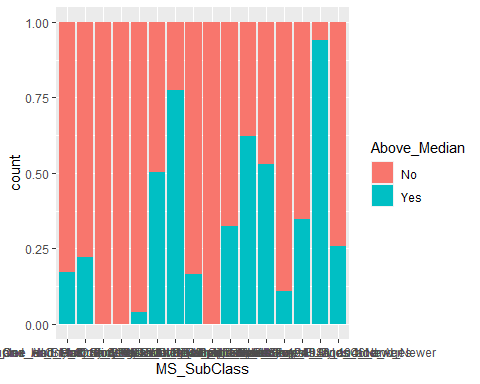
ames\_student <- read\_csv("ames\_student.csv")

## Warning: Missing column names filled in: 'X1' [1]

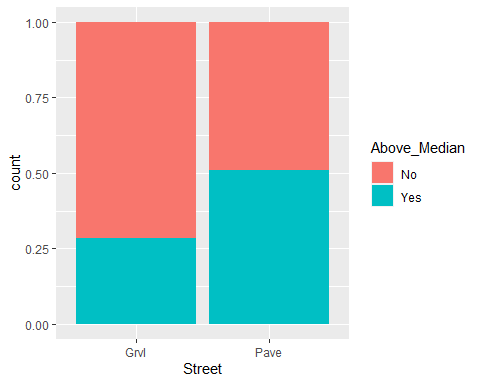
##   
## -- Column specification --------------------------------------------------------  
## cols(  
## .default = col\_character(),  
## X1 = col\_double(),  
## Lot\_Frontage = col\_double(),  
## Lot\_Area = col\_double(),  
## Year\_Built = col\_double(),  
## Year\_Remod\_Add = col\_double(),  
## Mas\_Vnr\_Area = col\_double(),  
## BsmtFin\_SF\_1 = col\_double(),  
## BsmtFin\_SF\_2 = col\_double(),  
## Bsmt\_Unf\_SF = col\_double(),  
## Total\_Bsmt\_SF = col\_double(),  
## First\_Flr\_SF = col\_double(),  
## Second\_Flr\_SF = col\_double(),  
## Low\_Qual\_Fin\_SF = col\_double(),  
## Gr\_Liv\_Area = col\_double(),  
## Bsmt\_Full\_Bath = col\_double(),  
## Bsmt\_Half\_Bath = col\_double(),  
## Full\_Bath = col\_double(),  
## Half\_Bath = col\_double(),  
## Bedroom\_AbvGr = col\_double(),  
## Kitchen\_AbvGr = col\_double()  
## # ... with 15 more columns  
## )  
## i Use `spec()` for the full column specifications.

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

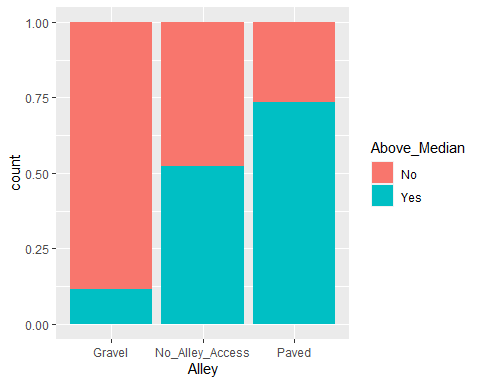
ggplot(ames\_student\_clean, aes(x = MS\_SubClass, fill = Above\_Median)) + geom\_bar(position = "fill")



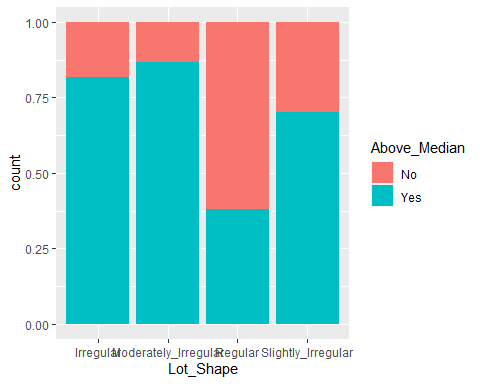
ggplot(ames\_student\_clean, aes(x = Street, fill = Above\_Median)) + geom\_bar(position = "fill")



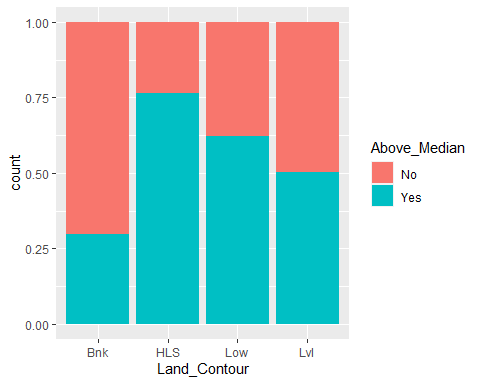
ggplot(ames\_student\_clean, aes(x = Alley, fill = Above\_Median)) + geom\_bar(position = "fill")



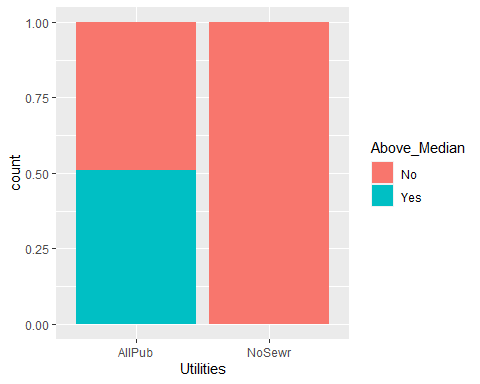
ggplot(ames\_student\_clean, aes(x = Lot\_Shape, fill = Above\_Median)) + geom\_bar(position = "fill")



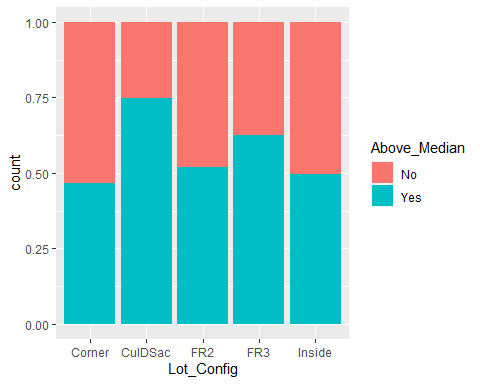
ggplot(ames\_student\_clean, aes(x = Land\_Contour, fill = Above\_Median)) + geom\_bar(position = "fill")



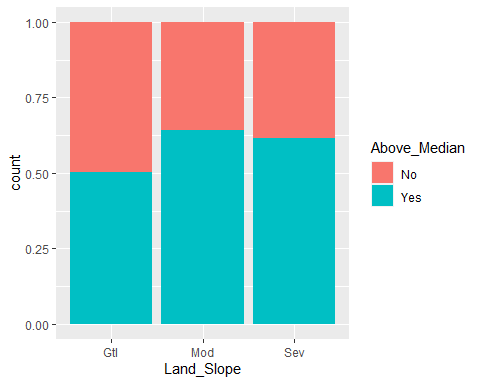
ggplot(ames\_student\_clean, aes(x = Utilities, fill = Above\_Median)) + geom\_bar(position = "fill")



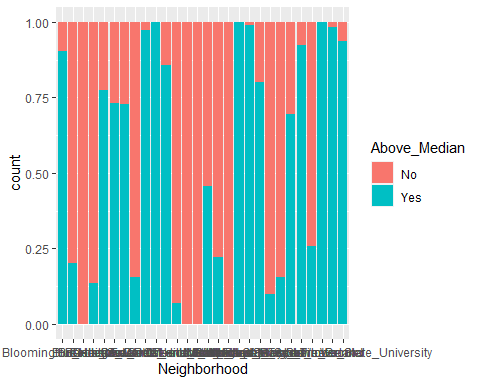
ggplot(ames\_student\_clean, aes(x = Lot\_Config, fill = Above\_Median)) + geom\_bar(position = "fill")



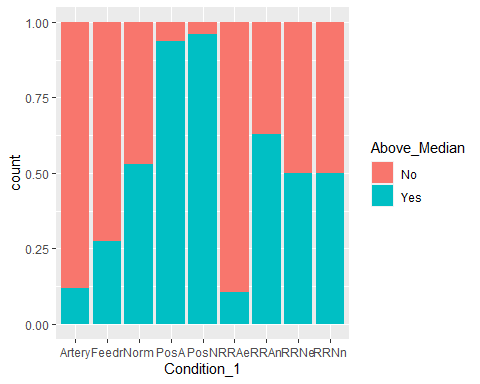
ggplot(ames\_student\_clean, aes(x = Land\_Slope, fill = Above\_Median)) + geom\_bar(position = "fill")



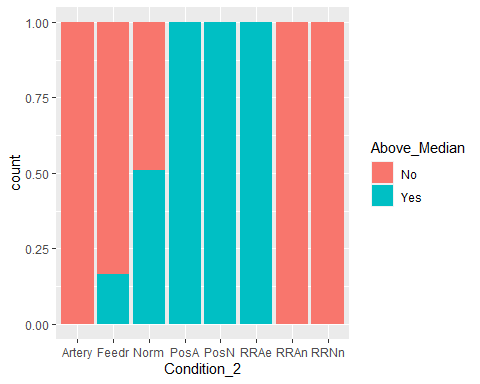
ggplot(ames\_student\_clean, aes(x = Neighborhood, fill = Above\_Median)) + geom\_bar(position = "fill")



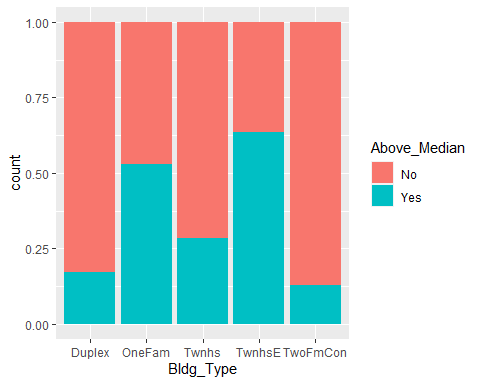
ggplot(ames\_student\_clean, aes(x = Condition\_1, fill = Above\_Median)) + geom\_bar(position = "fill")



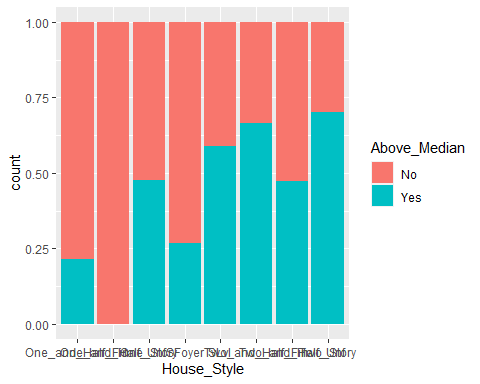
ggplot(ames\_student\_clean, aes(x = Condition\_2, fill = Above\_Median)) + geom\_bar(position = "fill")



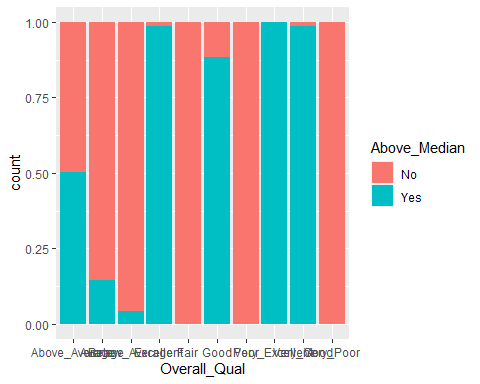
ggplot(ames\_student\_clean, aes(x = Bldg\_Type, fill = Above\_Median)) + geom\_bar(position = "fill")



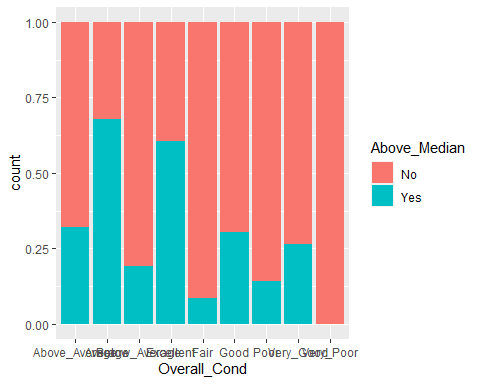
ggplot(ames\_student\_clean, aes(x = House\_Style, fill = Above\_Median)) + geom\_bar(position = "fill")



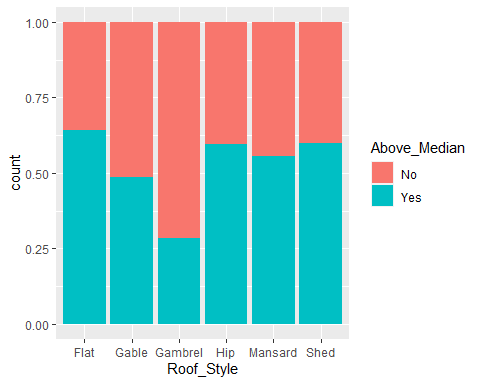
ggplot(ames\_student\_clean, aes(x = Overall\_Qual, fill = Above\_Median)) + geom\_bar(position = "fill")



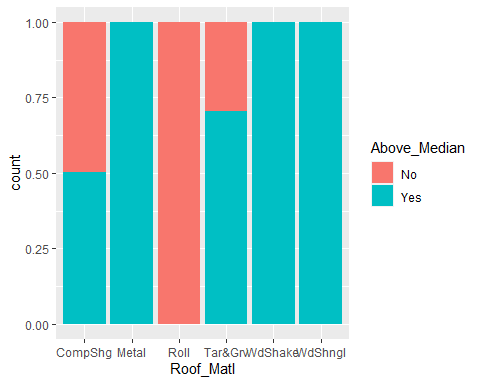
ggplot(ames\_student\_clean, aes(x = Overall\_Cond, fill = Above\_Median)) + geom\_bar(position = "fill")



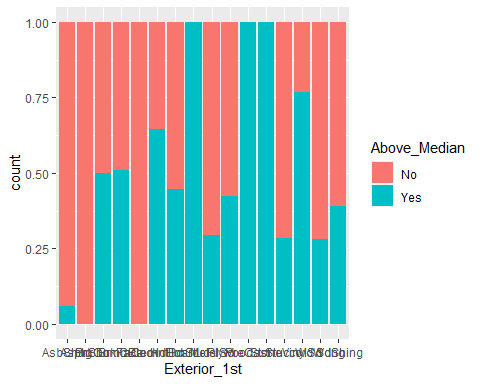
ggplot(ames\_student\_clean, aes(x = Roof\_Style, fill = Above\_Median)) + geom\_bar(position = "fill")



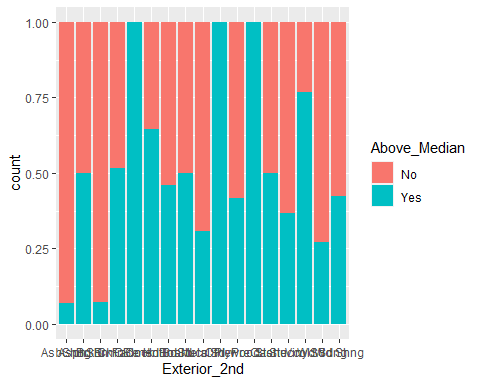
ggplot(ames\_student\_clean, aes(x = Roof\_Matl, fill = Above\_Median)) + geom\_bar(position = "fill")



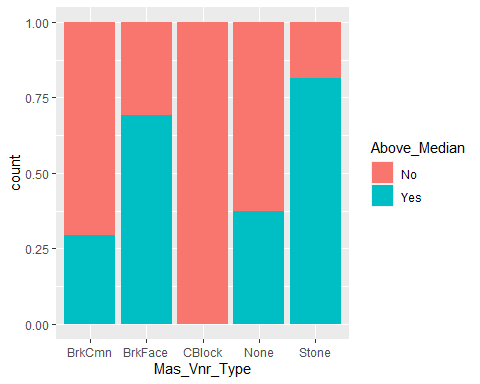
ggplot(ames\_student\_clean, aes(x = Exterior\_1st, fill = Above\_Median)) + geom\_bar(position = "fill")



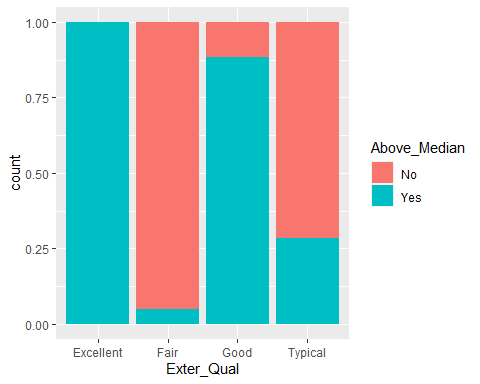
ggplot(ames\_student\_clean, aes(x = Exterior\_2nd, fill = Above\_Median)) + geom\_bar(position = "fill")



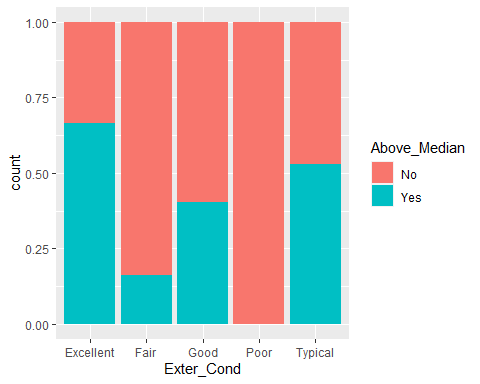
ggplot(ames\_student\_clean, aes(x = Mas\_Vnr\_Type, fill = Above\_Median)) + geom\_bar(position = "fill")



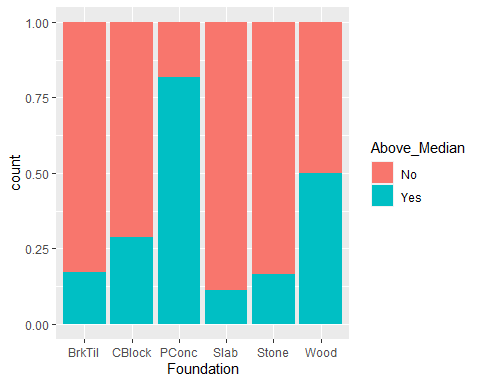
ggplot(ames\_student\_clean, aes(x = Exter\_Qual, fill = Above\_Median)) + geom\_bar(position = "fill")



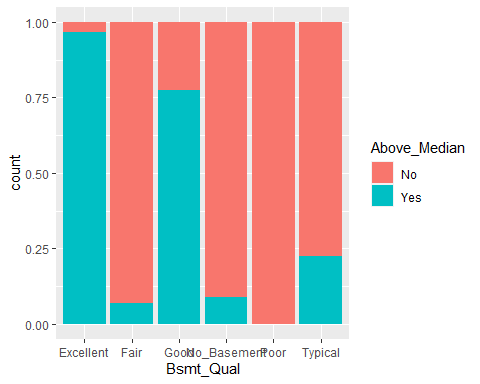
ggplot(ames\_student\_clean, aes(x = Exter\_Cond, fill = Above\_Median)) + geom\_bar(position = "fill")



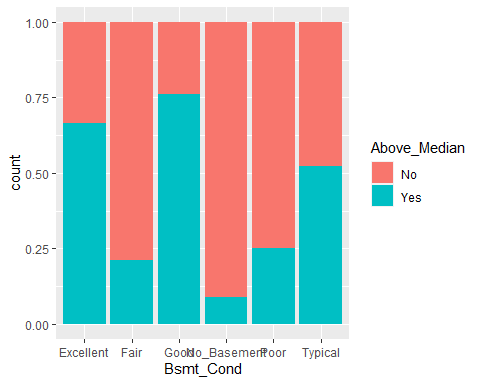
ggplot(ames\_student\_clean, aes(x = Foundation, fill = Above\_Median)) + geom\_bar(position = "fill")



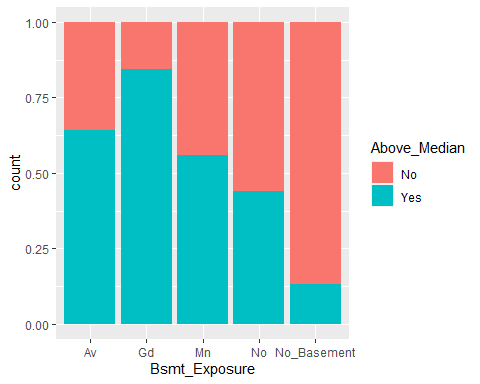
ggplot(ames\_student\_clean, aes(x = Bsmt\_Qual, fill = Above\_Median)) + geom\_bar(position = "fill")



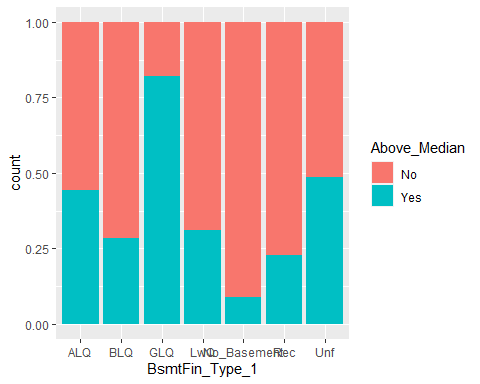
ggplot(ames\_student\_clean, aes(x = Bsmt\_Cond, fill = Above\_Median)) + geom\_bar(position = "fill")



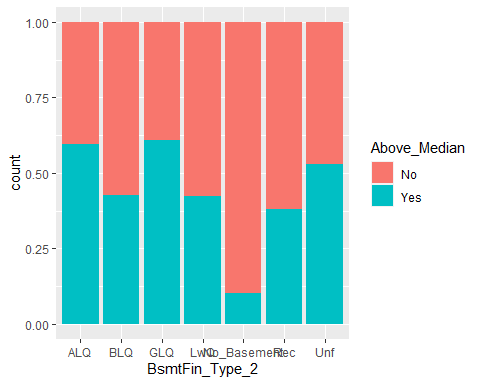
ggplot(ames\_student\_clean, aes(x = Bsmt\_Exposure, fill = Above\_Median)) + geom\_bar(position = "fill")



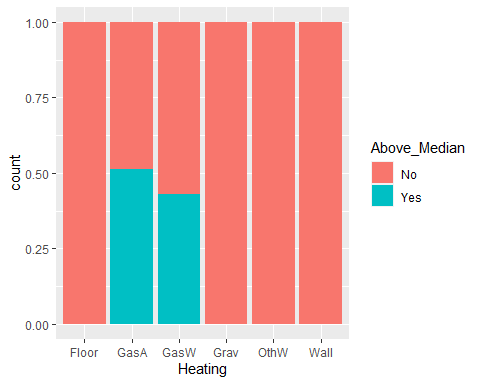
ggplot(ames\_student\_clean, aes(x = BsmtFin\_Type\_1, fill = Above\_Median)) + geom\_bar(position = "fill")



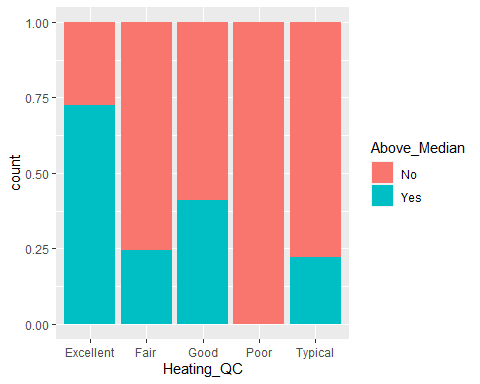
ggplot(ames\_student\_clean, aes(x = BsmtFin\_Type\_2, fill = Above\_Median)) + geom\_bar(position = "fill")



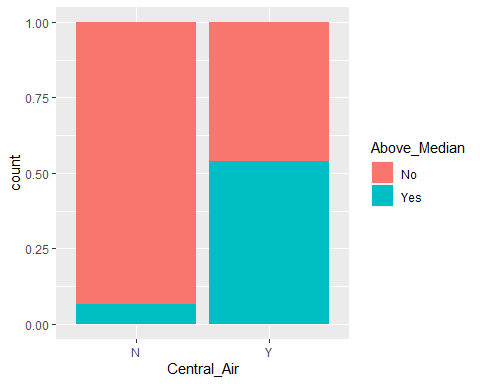
ggplot(ames\_student\_clean, aes(x = Heating, fill = Above\_Median)) + geom\_bar(position = "fill")



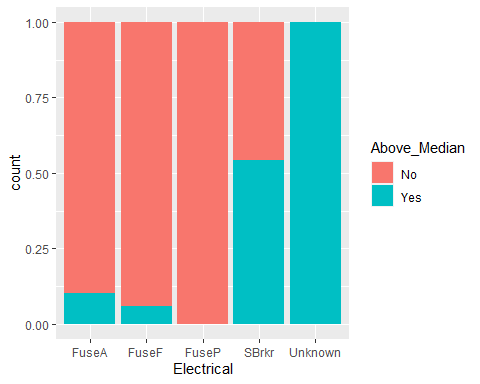
ggplot(ames\_student\_clean, aes(x = Heating\_QC, fill = Above\_Median)) + geom\_bar(position = "fill")



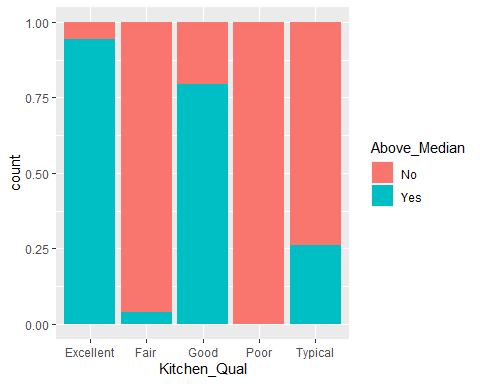
ggplot(ames\_student\_clean, aes(x = Central\_Air, fill = Above\_Median)) + geom\_bar(position = "fill")



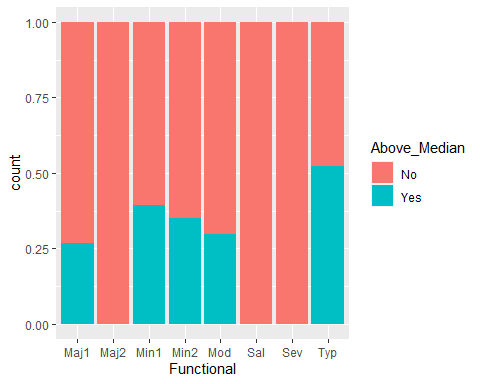
ggplot(ames\_student\_clean, aes(x = Electrical, fill = Above\_Median)) + geom\_bar(position = "fill")



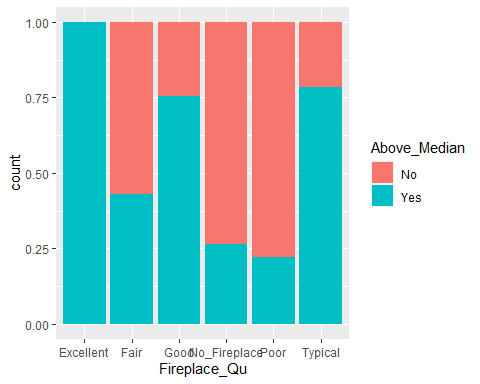
ggplot(ames\_student\_clean, aes(x = Kitchen\_Qual, fill = Above\_Median)) + geom\_bar(position = "fill")



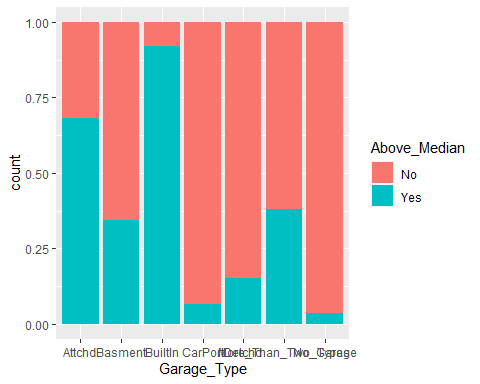
ggplot(ames\_student\_clean, aes(x = Functional, fill = Above\_Median)) + geom\_bar(position = "fill")



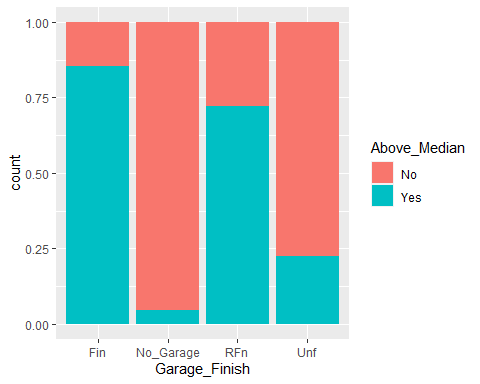
ggplot(ames\_student\_clean, aes(x = Fireplace\_Qu, fill = Above\_Median)) + geom\_bar(position = "fill")



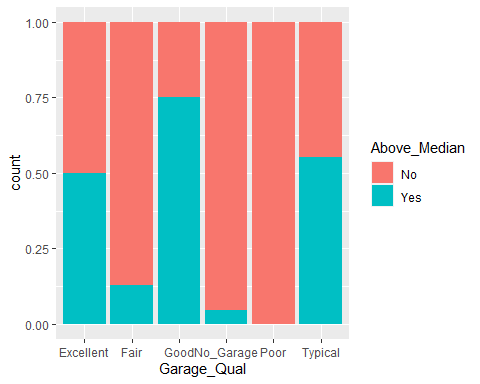
ggplot(ames\_student\_clean, aes(x = Garage\_Type, fill = Above\_Median)) + geom\_bar(position = "fill")



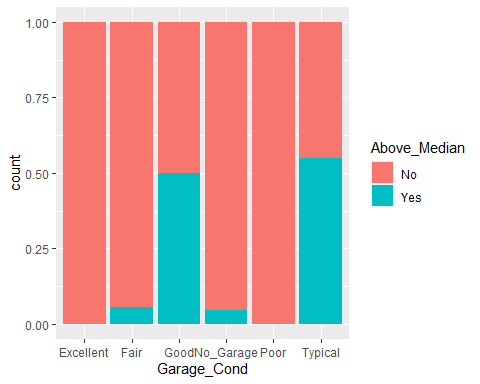
ggplot(ames\_student\_clean, aes(x = Garage\_Finish, fill = Above\_Median)) + geom\_bar(position = "fill")



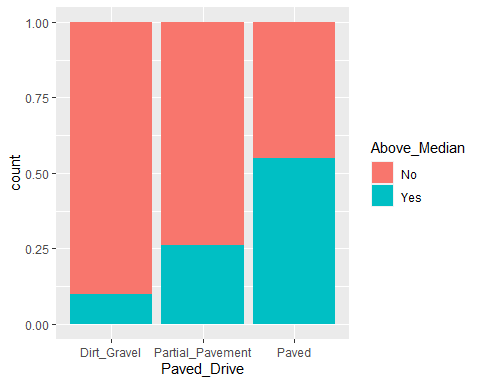
ggplot(ames\_student\_clean, aes(x = Garage\_Qual, fill = Above\_Median)) + geom\_bar(position = "fill")



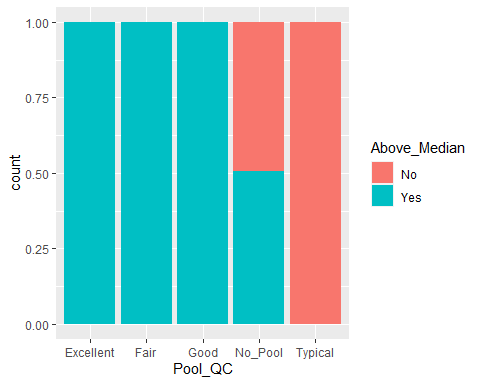
ggplot(ames\_student\_clean, aes(x = Garage\_Cond, fill = Above\_Median)) + geom\_bar(position = "fill")



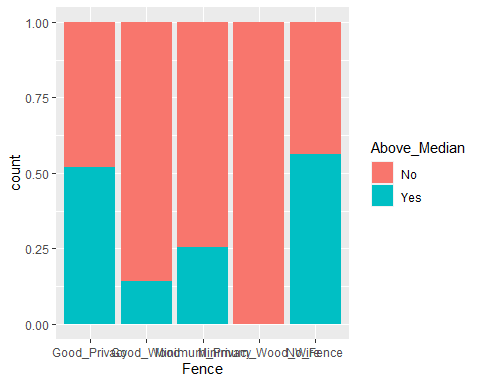
ggplot(ames\_student\_clean, aes(x = Paved\_Drive, fill = Above\_Median)) + geom\_bar(position = "fill")



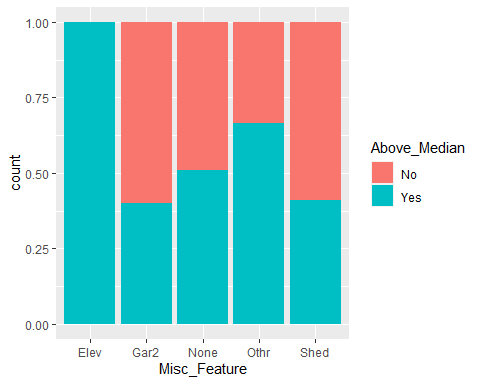
ggplot(ames\_student\_clean, aes(x = Pool\_QC, fill = Above\_Median)) + geom\_bar(position = "fill")



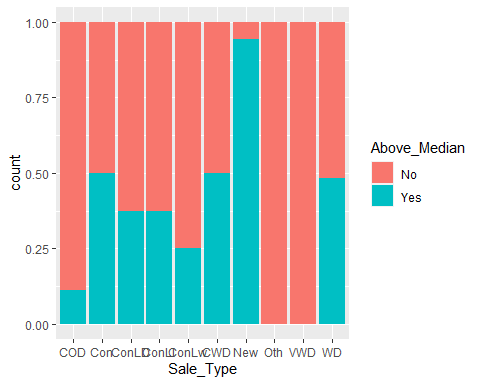
ggplot(ames\_student\_clean, aes(x = Fence, fill = Above\_Median)) + geom\_bar(position = "fill")



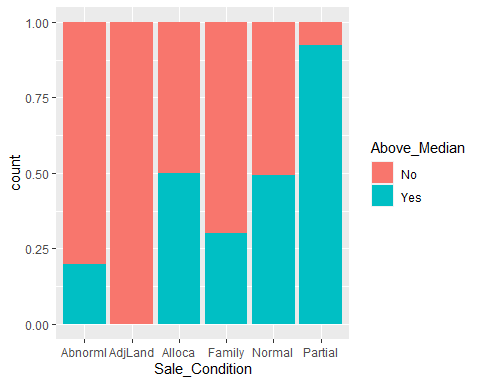
ggplot(ames\_student\_clean, aes(x = Misc\_Feature, fill = Above\_Median)) + geom\_bar(position = "fill")



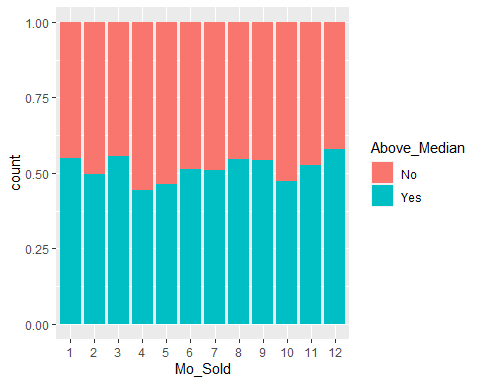
ggplot(ames\_student\_clean, aes(x = Sale\_Type, fill = Above\_Median)) + geom\_bar(position = "fill")



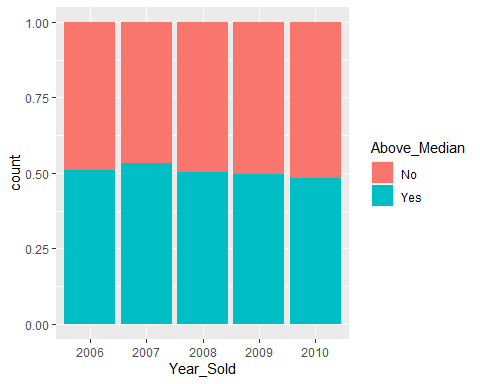
ggplot(ames\_student\_clean, aes(x = Sale\_Condition, fill = Above\_Median)) + geom\_bar(position = "fill")



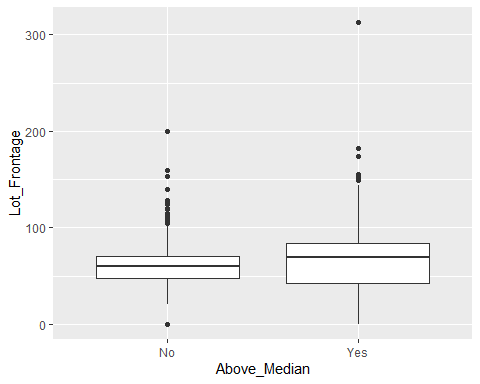
ggplot(ames\_student\_clean, aes(x = Mo\_Sold, fill = Above\_Median)) + geom\_bar(position = "fill")



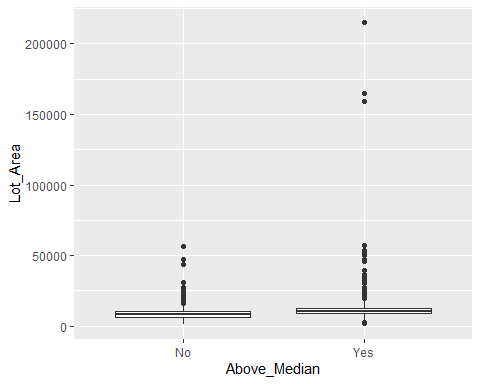
ggplot(ames\_student\_clean, aes(x = Year\_Sold, fill = Above\_Median)) + geom\_bar(position = "fill")



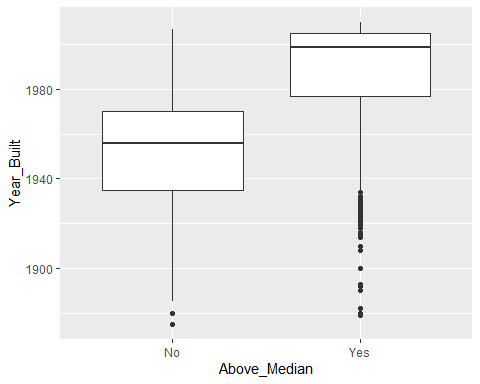
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Lot\_Frontage)) + geom\_boxplot()



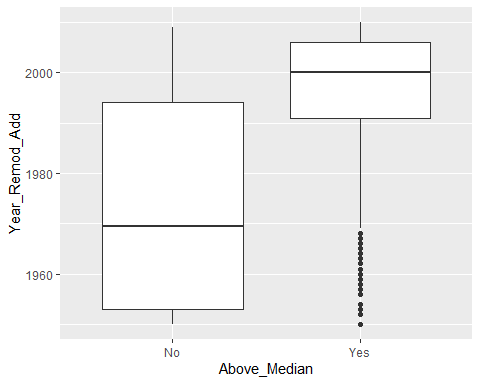
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Lot\_Area)) + geom\_boxplot()



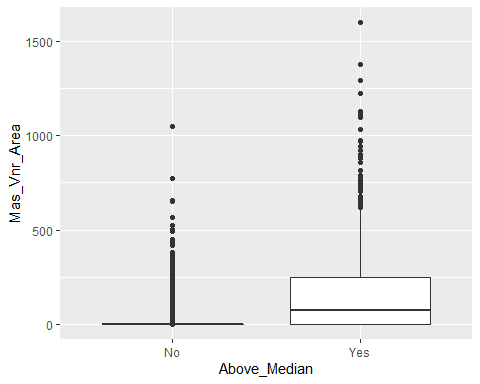
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Year\_Built)) + geom\_boxplot()



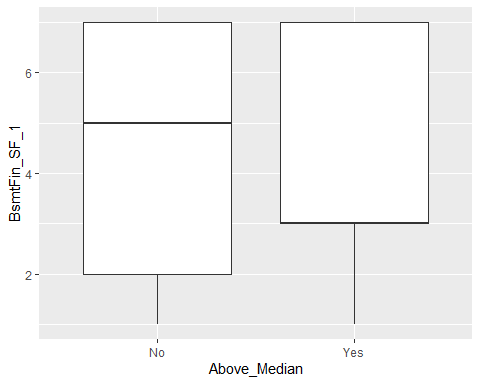
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Year\_Remod\_Add)) + geom\_boxplot()



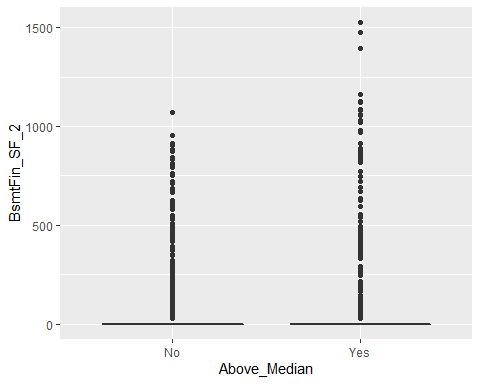
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Mas\_Vnr\_Area)) + geom\_boxplot()



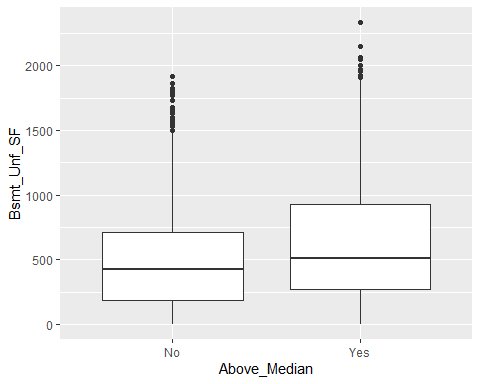
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=BsmtFin\_SF\_1)) + geom\_boxplot()



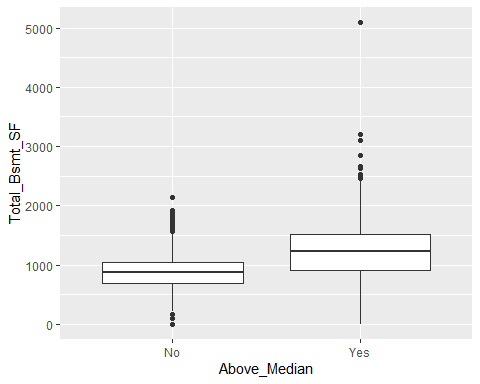
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=BsmtFin\_SF\_2)) + geom\_boxplot()



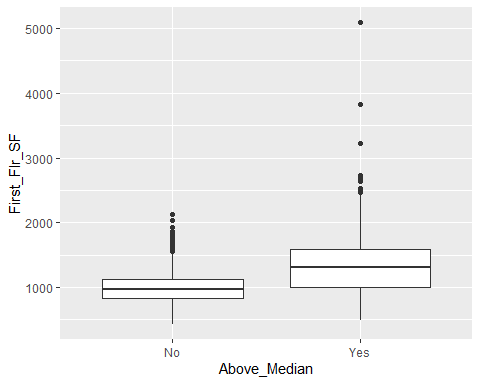
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Bsmt\_Unf\_SF)) + geom\_boxplot()



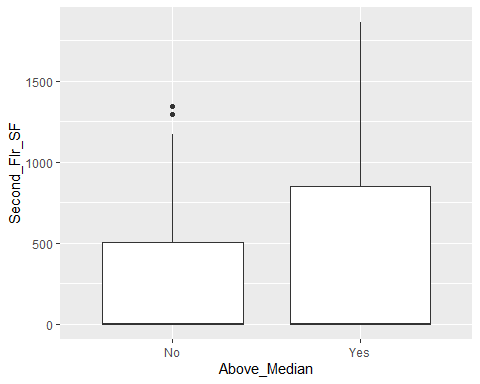
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Total\_Bsmt\_SF)) + geom\_boxplot()



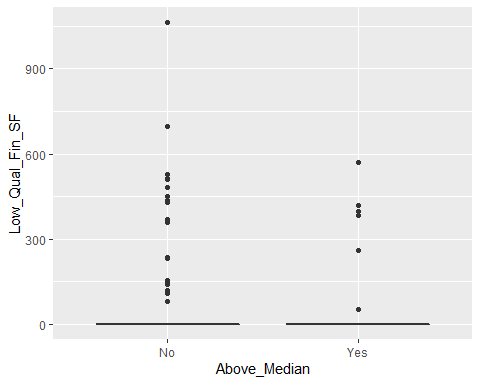
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=First\_Flr\_SF)) + geom\_boxplot()



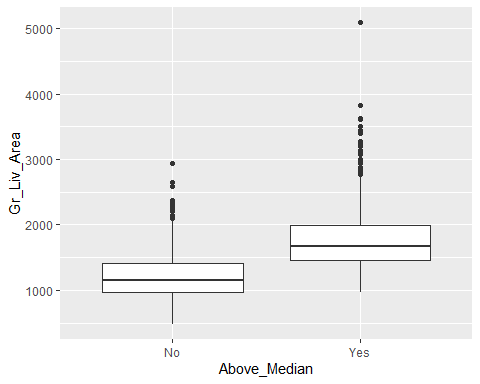
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Second\_Flr\_SF)) + geom\_boxplot()



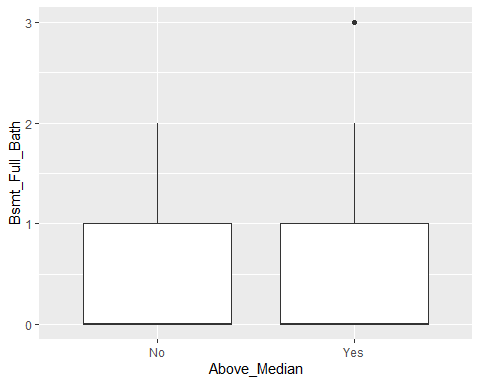
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Low\_Qual\_Fin\_SF)) + geom\_boxplot()



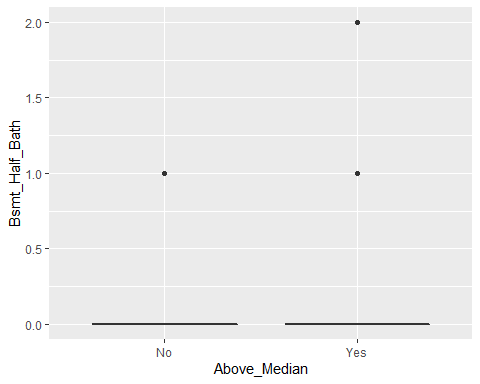
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Gr\_Liv\_Area)) + geom\_boxplot()



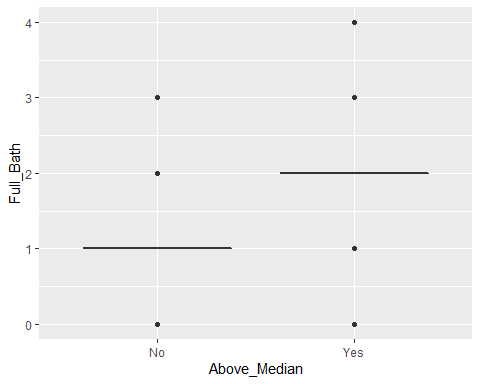
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Bsmt\_Full\_Bath)) + geom\_boxplot()



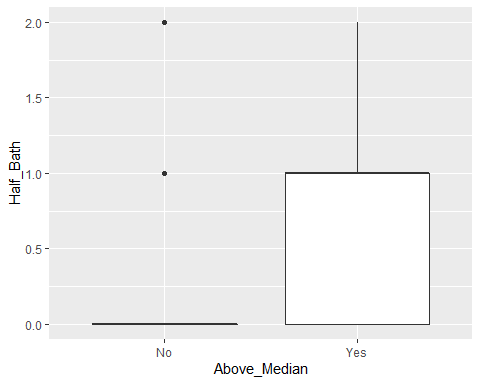
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Bsmt\_Half\_Bath)) + geom\_boxplot()



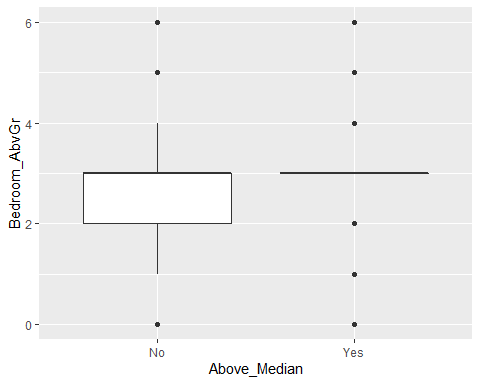
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Full\_Bath)) + geom\_boxplot()



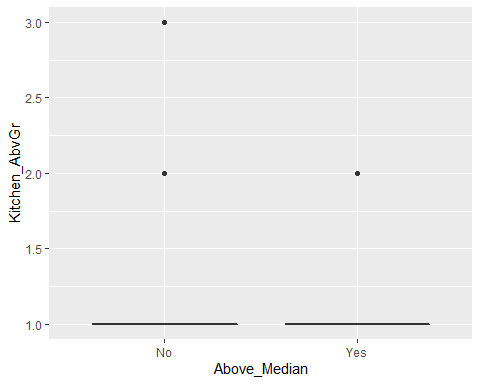
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Half\_Bath)) + geom\_boxplot()



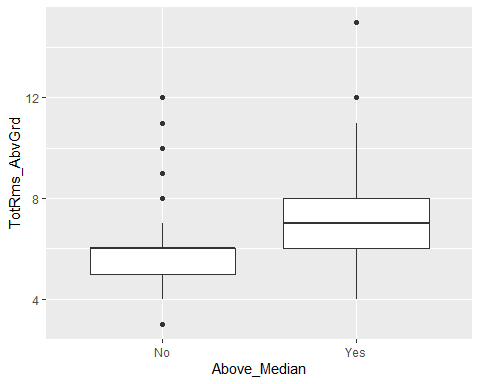
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Bedroom\_AbvGr)) + geom\_boxplot()



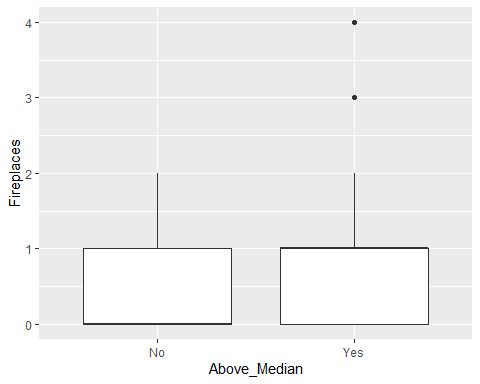
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Kitchen\_AbvGr)) + geom\_boxplot()



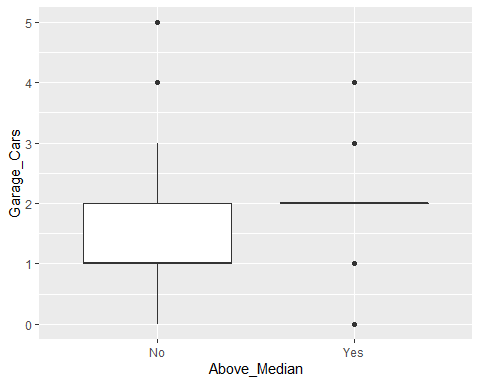
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=TotRms\_AbvGrd)) + geom\_boxplot()



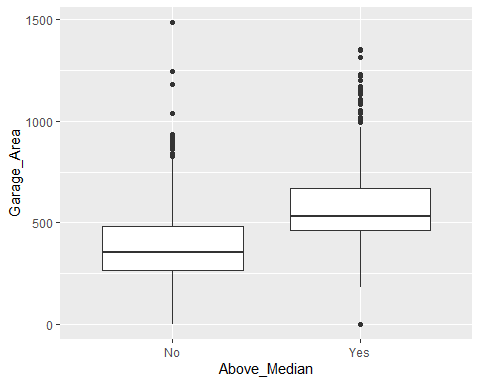
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Fireplaces)) + geom\_boxplot()



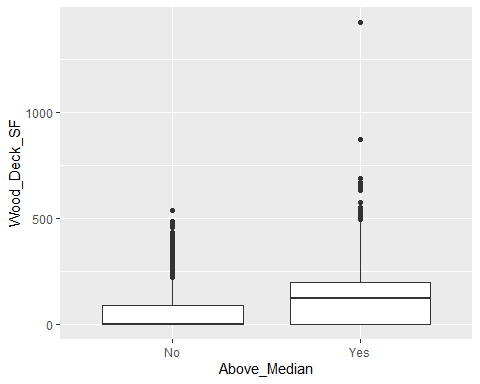
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Garage\_Cars)) + geom\_boxplot()



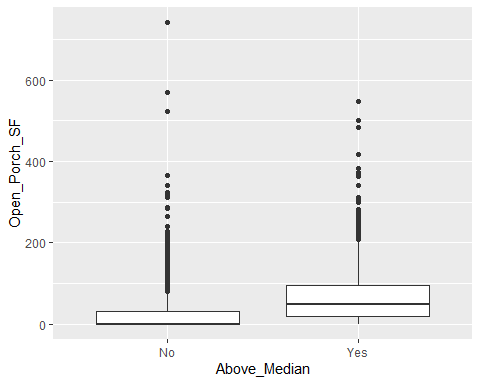
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Garage\_Area)) + geom\_boxplot()



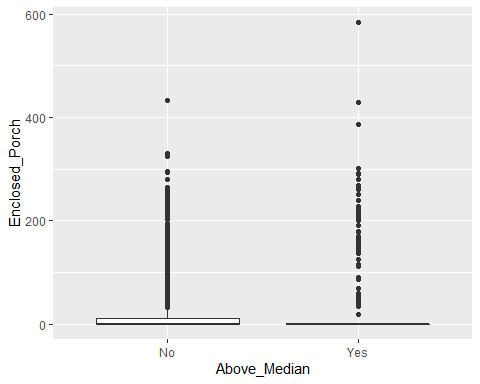
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Wood\_Deck\_SF)) + geom\_boxplot()



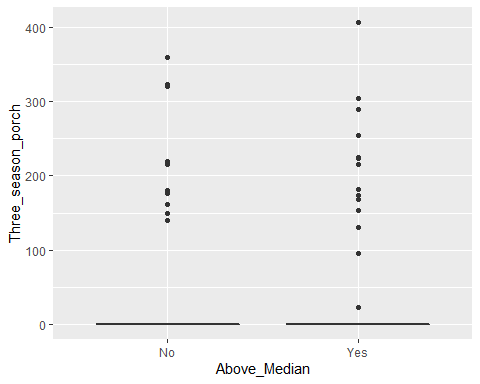
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Open\_Porch\_SF)) + geom\_boxplot()



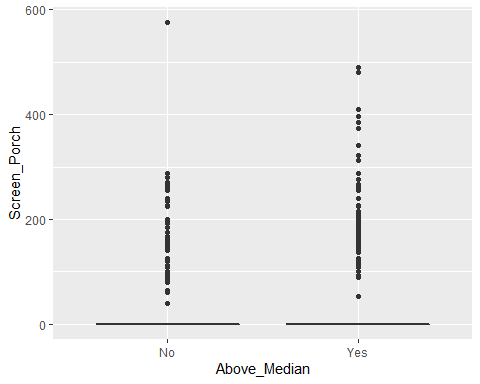
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Enclosed\_Porch)) + geom\_boxplot()



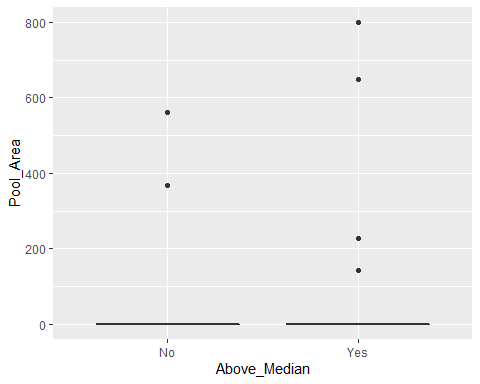
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Three\_season\_porch)) + geom\_boxplot()



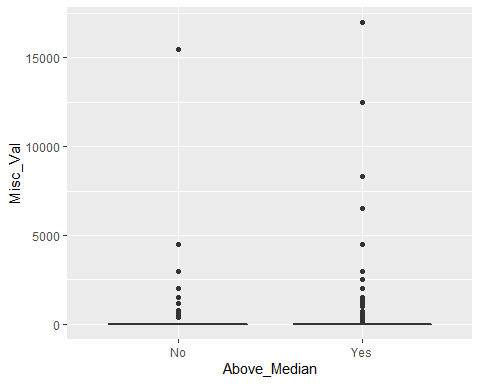
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Screen\_Porch)) + geom\_boxplot()



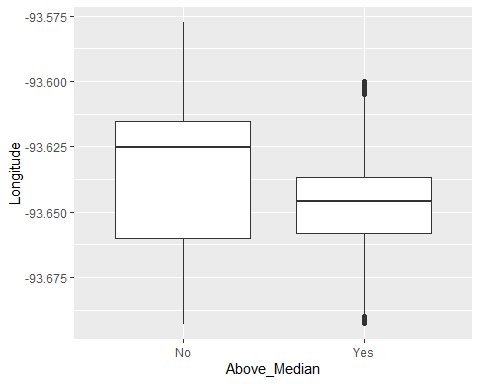
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Pool\_Area)) + geom\_boxplot()



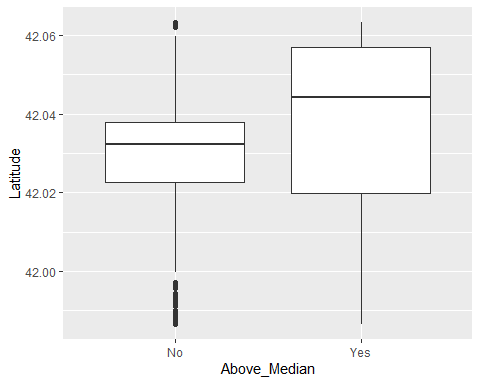
ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Misc\_Val)) + geom\_boxplot()



ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Longitude)) + geom\_boxplot()



ggplot(ames\_student\_clean, aes(x = Above\_Median, y=Latitude)) + geom\_boxplot()



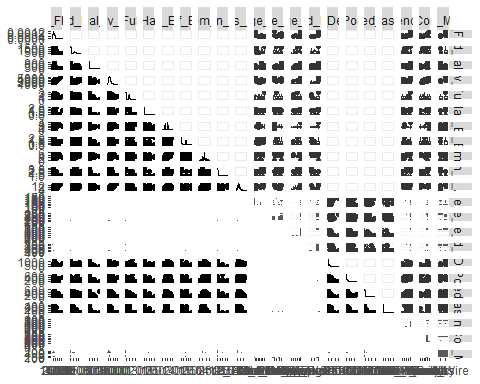
library(GGally)

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

ggpairs(ames\_student\_clean, columns = c(44:53,55,59,60,64:69,73,79,82))

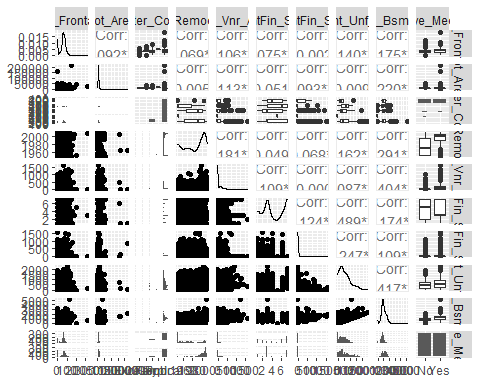
## `stat\_bin()` using `bins = 30`. Pick better value with `binwidth`.

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



ggpairs(ames\_student\_clean, columns = c(4,5,29,21,27,35,37,38,39,82))

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



library(ranger)  
library(vip)

##   
## Attaching package: 'vip'

## The following object is masked from 'package:utils':  
##   
## vi

library(tidymodels)

## -- Attaching packages -------------------------------------- tidymodels 0.1.2 --

## v broom 0.7.2 v recipes 0.1.15  
## v dials 0.0.9 v rsample 0.0.8   
## v infer 0.5.4 v tune 0.1.2   
## v modeldata 0.1.0 v workflows 0.2.1   
## v parsnip 0.1.5 v yardstick 0.0.7

## -- Conflicts ----------------------------------------- tidymodels\_conflicts() --  
## x scales::discard() masks purrr::discard()  
## x dplyr::filter() masks stats::filter()  
## x recipes::fixed() masks stringr::fixed()  
## x dplyr::lag() masks stats::lag()  
## x yardstick::spec() masks readr::spec()  
## x recipes::step() masks stats::step()

rf\_folds = vfold\_cv(ames\_student\_clean, v =5)  
  
ames\_recipe = recipe(Above\_Median ~., ames\_student\_clean) %>%  
 step\_dummy(all\_nominal(), -all\_outcomes())  
  
rf\_model = rand\_forest(mtry = tune(), min\_n = tune(), trees = 100) %>%  
 set\_engine("ranger", importance = "permutation") %>%  
 set\_mode("classification")  
  
ames\_wflow =  
 workflow() %>%  
 add\_model(rf\_model) %>%  
 add\_recipe(ames\_recipe)  
  
rf\_grid = grid\_regular(  
 mtry(range = c(2, 8)),  
 min\_n(range = c(5, 20)), levels = 10  
)  
  
rf\_res\_tuned = tune\_grid(ames\_wflow,   
 resamples = rf\_folds,  
 grid = rf\_grid)

##   
## Attaching package: 'rlang'

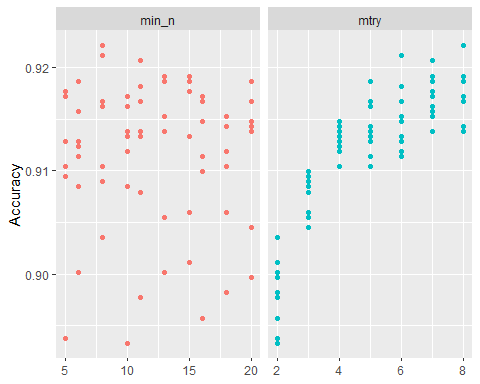
## The following objects are masked from 'package:purrr':  
##   
## %@%, as\_function, flatten, flatten\_chr, flatten\_dbl, flatten\_int,  
## flatten\_lgl, flatten\_raw, invoke, list\_along, modify, prepend,  
## splice

##   
## Attaching package: 'vctrs'

## The following object is masked from 'package:dplyr':  
##   
## data\_frame

## The following object is masked from 'package:tibble':  
##   
## data\_frame

rf\_res\_tuned %>%  
 collect\_metrics() %>%  
 filter(.metric == "accuracy") %>%  
 select(mean, min\_n, mtry) %>%  
 pivot\_longer(min\_n:mtry,  
 values\_to = "value",  
 names\_to = "parameter"  
 ) %>%  
 ggplot(aes(value, mean, color = parameter)) +  
 geom\_point(show.legend = FALSE) +  
 facet\_wrap(~parameter, scales = "free\_x") +  
 labs(x = NULL, y = "Accuracy")



best\_rf = select\_best(rf\_res\_tuned, "accuracy")  
  
final\_rf = finalize\_workflow(  
 ames\_wflow,  
 best\_rf  
)  
  
final\_rf\_fit = fit(final\_rf, ames\_student\_clean)  
  
final\_rf\_fit %>% pull\_workflow\_fit() %>% vip(geom = "point")

