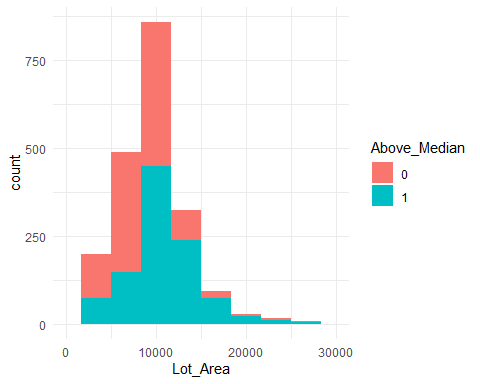
# BAN 502 - Course Project Phase 1

## Khayrayyah Haamid-Day

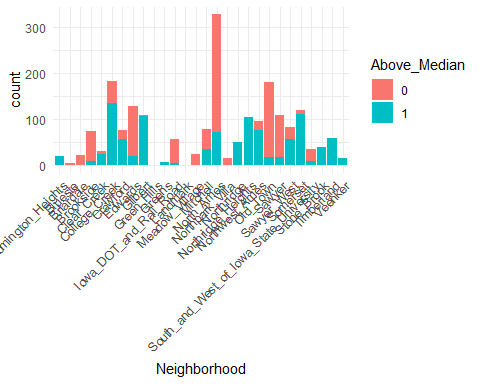
ames\_student = ames\_student %>% mutate\_if(is.character, as.factor) %>%  
 mutate(Above\_Median = fct\_recode(Above\_Median, "0" = "No", "1" = "Yes"))  
#summary(ames\_student)

I picked variables that I believe, after researching this, have the most influence on the pricing of a home. If I am buying a home, I would have a list of must haves, such as only wanting to live in a particular neighborhood with a large yard. I would also be open to paying extra for a finished basement and 2 full baths even if it was not on my list, especially if it is high quality. There were a few other variables that I added, such as month sold, after testing them out on esquisse while there a couple of variables I initially picked decided not to include in the PowerPoint as I do not believe there is much of a relationship.

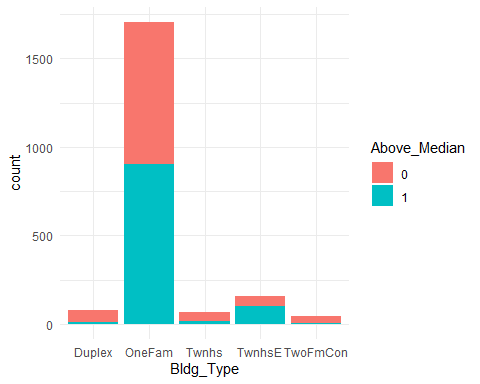
#esquisser()  
  
ggplot(ames\_student) +  
 aes(x = Lot\_Area, fill = Above\_Median) +  
 geom\_histogram(bins = 10L) +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 xlim(0, 30000)



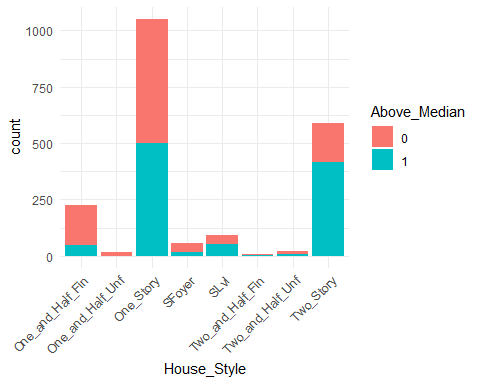
ggplot(ames\_student) +  
 aes(x = Neighborhood, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))



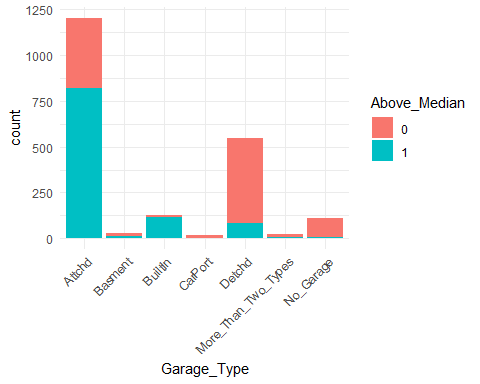
ggplot(ames\_student) +  
 aes(x = Bldg\_Type, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



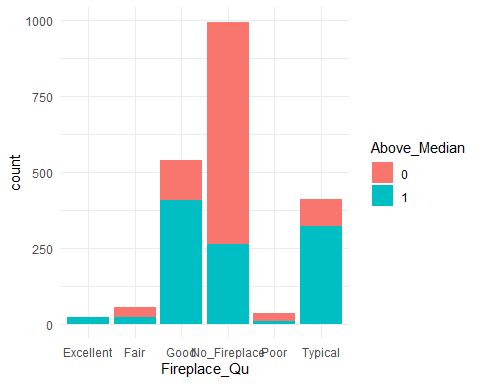
ggplot(ames\_student) +  
 aes(x = House\_Style, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))



ggplot(ames\_student) +  
 aes(x = Garage\_Type, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))

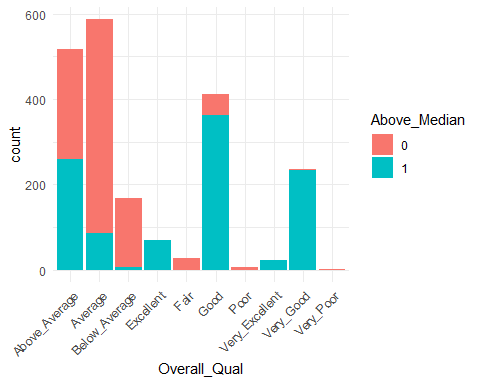


ggplot(ames\_student) +  
 aes(x = Fireplace\_Qu, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()

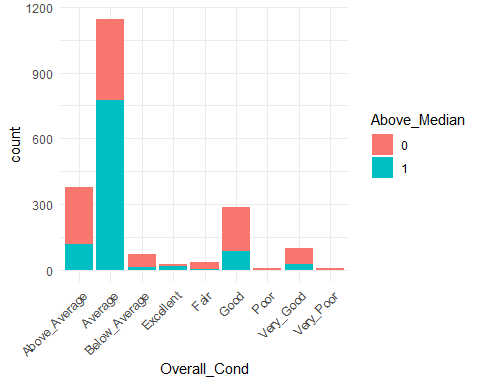


Lot Area - Is skewed to the left, though there isn’t much data for larger lot sizes, appears that the more sqfeet increases median price Neighborhood - Neighborhoods appear to impact if a house will be Above Median Bldg Type - Does not appear to be much of a relationship House Style - Stronger relationship with two-story homes being above median Garage Type - Attached and built-in garages could be above median in comparison to no garage and detached Fireplace Quality - Looks normally distributed; having no fireplace decrease the house price while a good and typical quality increases price above median

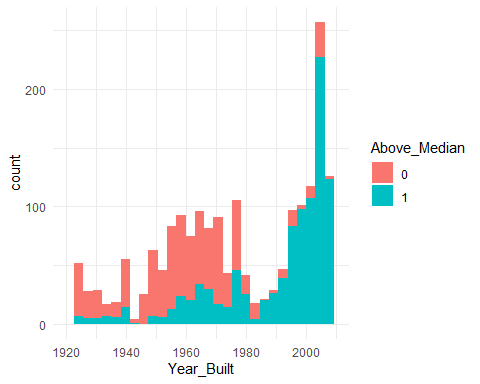
ggplot(ames\_student) +  
 aes(x = Overall\_Qual, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))



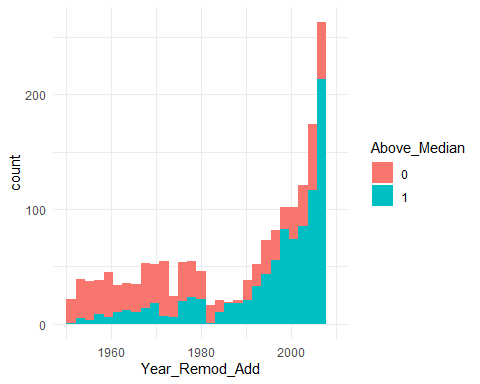
ggplot(ames\_student) +  
 aes(x = Overall\_Cond, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 theme(axis.text.x = element\_text(angle = 45, vjust = 1, hjust = 1))



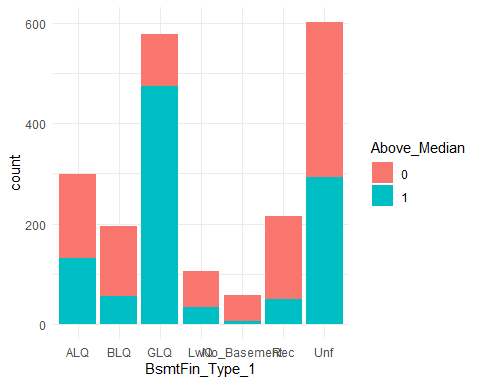
ggplot(ames\_student) +  
 aes(x = Year\_Built, fill = Above\_Median) +  
 geom\_histogram(bins = 30L) +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 xlim(1920, 2010)



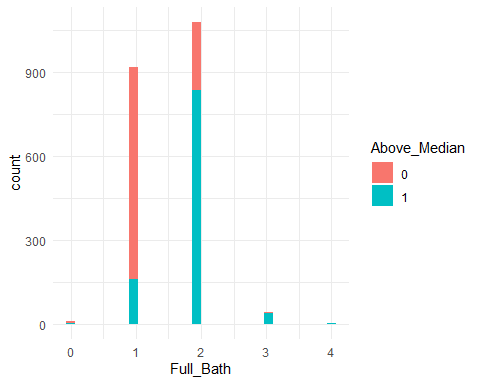
ggplot(ames\_student) +  
 aes(x = Year\_Remod\_Add, fill = Above\_Median) +  
 geom\_histogram(bins = 30L) +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal() +  
 xlim(1950, 2010)



ggplot(ames\_student) +  
 aes(x = BsmtFin\_Type\_1, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()

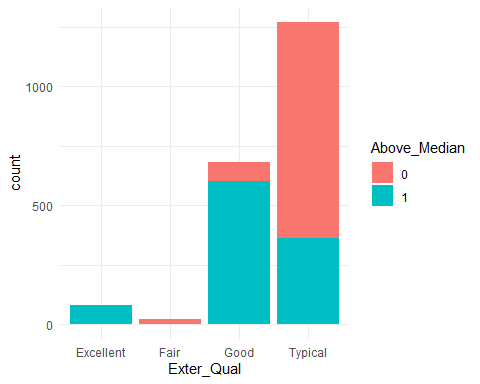


ggplot(ames\_student) +  
 aes(x = Full\_Bath, fill = Above\_Median) +  
 geom\_histogram(bins = 30L) +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()

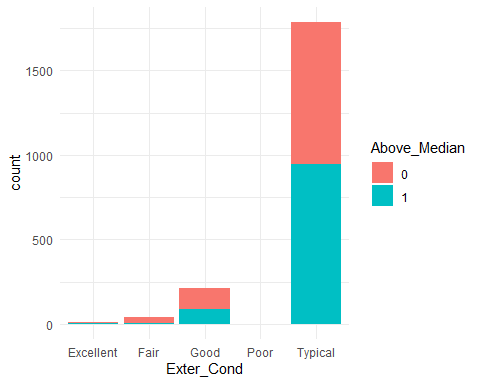


Overall quality - overall quality of material used in home affects price of house being above/below median Overall condition - overall condition of the home may not have enough data; an average home is above median but it is odd that a above average house does not increase above median Year built - Skewed to the right; houses built after 1980/1990 have a above median price Year Remodeled - Skewed to the right; similar to year built, after 1980/1990 homes have a above median price Basement Finished - the rating of a finished basement at good living quarters have above median prices Full bath - there may not be many homes that have 2+ full baths; having 2 full baths drastically increases the price

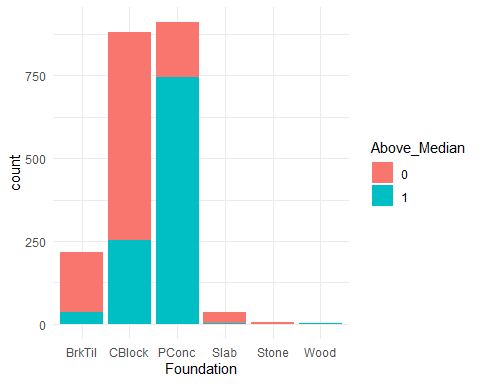
ggplot(ames\_student) +  
 aes(x = Exter\_Qual, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



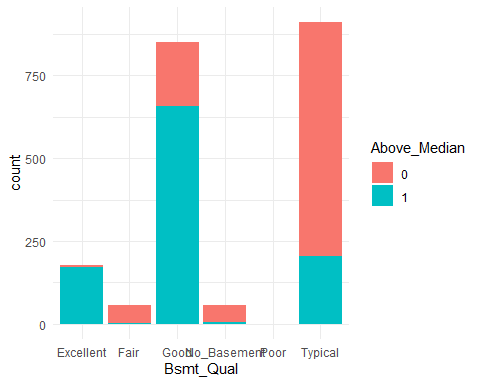
ggplot(ames\_student) +  
 aes(x = Exter\_Cond, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



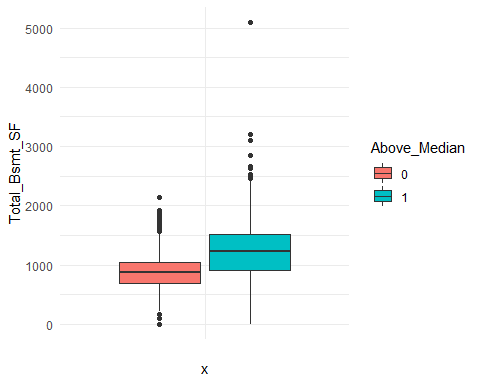
ggplot(ames\_student) +  
 aes(x = Foundation, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



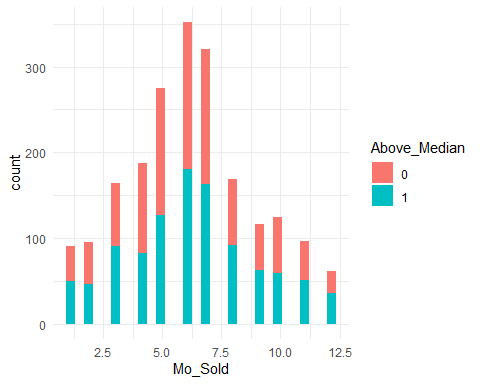
ggplot(ames\_student) +  
 aes(x = Bsmt\_Qual, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



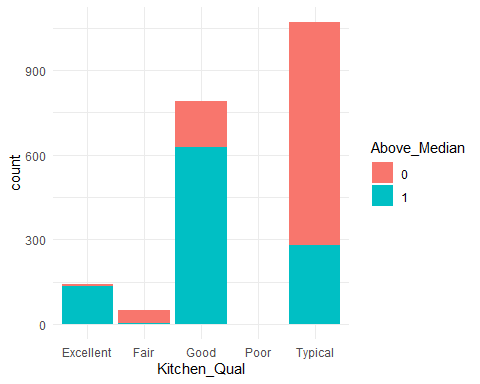
ggplot(ames\_student) +  
 aes(x = "", y = Total\_Bsmt\_SF, fill = Above\_Median) +  
 geom\_boxplot() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



ggplot(ames\_student) +  
 aes(x = Mo\_Sold, fill = Above\_Median) +  
 geom\_histogram(bins = 30L) +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



ggplot(ames\_student) +  
 aes(x = Kitchen\_Qual, fill = Above\_Median) +  
 geom\_bar() +  
 scale\_fill\_hue(direction = 1) +  
 theme\_minimal()



Exterior quality - skewed to right; good and excellent exterior quality of material increases the price above median Exterior cond - condition of exterior may not have a lot data, does not appear to have a strong relationship Foundation - skewed to left; only having a poured concrete increased above median which cinder block, brick and tile, and slab decrease Basement Qual - basement quality of good and excellent increase price above median Total basement sf - more square feet in the basement appears to impact above median Month sold - normally distributed; there is a relationship between month sold and above median Kitchen quality - kitchen quality being good and excellent increase price above median