

# Assignment 4 - Real Estate

Bharath Dasari

21/09/2021

Data Cleansing

```
library(readxl)
REAL_ESTATE_SALES_DATA <- read_excel("REAL ESTATE SALES DATA.xlsx")
View(REAL_ESTATE_SALES_DATA)
real.estate<- as.data.frame(REAL_ESTATE_SALES_DATA)
str(real.estate)
```

```
## 'data.frame':    522 obs. of  13 variables:
## $ Id           : num  1 2 3 4 5 6 7 8 9 10 ...
## $ Sales price: num  360000 340000 250000 205500 275500 ...
## $ Square feet: num  3032 2058 1780 1638 2196 ...
## $ Bedrooms    : num  4 4 4 4 4 4 3 2 3 3 ...
## $ Bathrooms   : num  4 2 3 2 3 3 2 1 2 3 ...
## $ AC          : num  1 1 1 1 1 1 1 1 1 0 ...
## $ Garage      : num  2 2 2 2 2 5 2 1 2 1 ...
## $ Pool        : num  0 0 0 0 0 1 0 0 0 0 ...
## $ Year        : num  1972 1976 1980 1963 1968 ...
## $ Quality     : num  2 2 2 2 2 2 2 2 3 3 ...
## $ Style       : num  1 1 1 1 7 1 7 1 1 1 ...
## $ lot size    : num  22221 22912 21345 17342 21786 ...
## $ Highway     : num  0 0 0 0 0 0 0 0 0 0 ...
```

```
real.estate$Pool<- as.factor(real.estate$Pool)
real.estate$AC<- as.factor(real.estate$AC)
real.estate$Highway<- as.factor(real.estate$Highway)
real.estate$Quality<-as.factor(real.estate$Quality)
str(real.estate)
```

```
## 'data.frame':    522 obs. of  13 variables:
## $ Id           : num  1 2 3 4 5 6 7 8 9 10 ...
## $ Sales price: num  360000 340000 250000 205500 275500 ...
## $ Square feet: num  3032 2058 1780 1638 2196 ...
## $ Bedrooms    : num  4 4 4 4 4 4 3 2 3 3 ...
## $ Bathrooms   : num  4 2 3 2 3 3 2 1 2 3 ...
## $ AC          : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 1 ...
## $ Garage      : num  2 2 2 2 2 5 2 1 2 1 ...
## $ Pool        : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 1 1 1 ...
## $ Year        : num  1972 1976 1980 1963 1968 ...
## $ Quality     : Factor w/ 3 levels "1","2","3": 2 2 2 2 2 2 2 2 3 3 ...
## $ Style       : num  1 1 1 1 7 1 7 1 1 1 ...
## $ lot size    : num  22221 22912 21345 17342 21786 ...
## $ Highway     : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
```

```
View(real.estate)
```

# Cumulative Logit model

```
library(ordinal)
attach(real.estate)
model.fit1<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                Garage+Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
summary(model.fit1)
```

```
## formula:
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC + Garage +
Pool + Year + `lot size` + Highway
##
## link threshold nobs logLik AIC niter max.grad cond.H
## logit flexible 522 -189.55 403.09 9(0) 5.29e-08 2.8e+15
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## `Sales price` -1.966e-05  3.040e-06  -6.467 1.00e-10 ***
## `Square feet` -1.431e-03  4.477e-04  -3.196  0.00139 **
## Bedrooms      5.250e-01  1.672e-01   3.140  0.00169 **
## Bathrooms     -9.014e-01  2.217e-01  -4.066  4.78e-05 ***
## AC1           -1.179e+00  4.129e-01  -2.855  0.00431 **
## Garage        -5.412e-01  3.114e-01  -1.738  0.08226 .
## Pool1         3.977e-01  5.303e-01   0.750  0.45331
## Year          -5.854e-02  1.162e-02  -5.038  4.70e-07 ***
## `lot size`    -6.260e-07  1.267e-05  -0.049  0.96058
## Highway1     -4.336e-01  7.607e-01  -0.570  0.56866
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Threshold coefficients:
##      Estimate Std. Error z value
## 1|2  -133.04      23.15  -5.747
## 2|3  -123.44      22.81  -5.411
```

Important Predictors:

- 1.Sales price
- 2.Square feet
- 3.Bedrooms
- 4.Bathrooms
- 5.AC
- 6.Year

Interpretation of Coefficients:

1. For every unit increase in Sales price the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.999980340193257 controlling for other variables.
2. For every unit increase in square feet the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.998570023392284 controlling for other variables.
3. For every unit increase in number of Bedrooms the odds of quality of construction being on higher quality changes multiplicatively by a factor of 1.69045884837909 controlling for other variables.
4. For every unit increase in number of Bathrooms the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.406000860469356 controlling for other variables.
5. For every availability of AC, the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.307586171030328 controlling for other variables.
6. For every unit increase in the no of cars in Garage the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.582049373882101 controlling for other variables.
7. For every availability of Pool, the odds of quality of construction being on higher quality changes multiplicatively by a factor of 1.48839744368959 controlling for other variables.
8. For every increase in Year the odds of quality of construction being on higher quality changes multiplicatively by a factor of 0.943140514024583 controlling for other variables.

```

ctable <- coef(summary(model.fit1))
p <- round(pnorm(abs(ctable[, "z value"])), lower.tail = FALSE) * 2, digits = 3)
ctable <- cbind(ctable, "p-value" = p)
ci <- confint.default(model.fit1)
ctable <- cbind(ctable, round(exp(cbind(OR = coef(model.fit1), ci)), digits = 3))
ctable

```

```

##              Estimate Std. Error   z value    Pr(>|z|) p-value    OR
## 1|2          -1.330367e+02 2.314907e+01 -5.74695608 9.086432e-09  0.000 0.000
## 2|3          -1.234352e+02 2.281373e+01 -5.41056594 6.282588e-08  0.000 0.000
## `Sales price` -1.966124e-05 3.040232e-06 -6.46701773 9.995593e-11  0.000 1.000
## `Square feet` -1.431018e-03 4.477076e-04 -3.19632152 1.391919e-03  0.001 0.999
## Bedrooms      5.249620e-01 1.671771e-01  3.14015544 1.688582e-03  0.002 1.690
## Bathrooms     -9.014348e-01 2.216941e-01 -4.06612062 4.780220e-05  0.000 0.406
## AC1            -1.178663e+00 4.128826e-01 -2.85471595 4.307534e-03  0.004 0.308
## Garage         -5.411935e-01 3.114352e-01 -1.73774022 8.225660e-02  0.082 0.582
## Pool1          3.976836e-01 5.303143e-01  0.74990169 4.533139e-01  0.453 1.488
## Year           -5.853971e-02 1.161903e-02 -5.03826306 4.697755e-07  0.000 0.943
## `lot size`     -6.260411e-07 1.266525e-05 -0.04942983 9.605768e-01  0.961 1.000
## Highway1       -4.336341e-01 7.607233e-01 -0.57002869 5.686582e-01  0.569 0.648
##              2.5 % 97.5 %
## 1|2              0.000 0.000
## 2|3              0.000 0.000
## `Sales price`    1.000 1.000
## `Square feet`    0.998 0.999
## Bedrooms         1.218 2.346
## Bathrooms        0.263 0.627
## AC1              0.137 0.691
## Garage           0.316 1.072
## Pool1            0.526 4.208
## Year             0.922 0.965
## `lot size`       1.000 1.000
## Highway1         0.146 2.879

```

## 1.Sales Price

```
model.fit2<- clm(as.factor(Quality) ~ `Square feet`+Bedrooms+Bathrooms+AC+
                Garage+Pool+Year+`lot size`+Highway,data=real.estate)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
##   - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit2)
```

```
## Likelihood ratio tests of cumulative link models:
##
##           formula:
## model.fit2 as.factor(Quality) ~ `Square feet` + Bedrooms + Bathrooms + AC + Garage + Pool
+ Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##           link: threshold:
## model.fit2 logit flexible
## model.fit1 logit flexible
##
##           no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit2      11 464.59 -221.29
## model.fit1      12 403.09 -189.55  63.495  1  1.607e-15 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Sales Price is not a significant variable

Ha: Sales Price is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Sales Price is a significant variable.

## 2.Square feet

```
model.fit3<- clm(as.factor(Quality) ~ `Sales price`+Bedrooms+Bathrooms+AC+
                Garage+Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
##   - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit3)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit3 as.factor(Quality) ~ `Sales price` + Bedrooms + Bathrooms + AC + Garage + Pool
+ Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit3 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC   logLik LR.stat df Pr(>Chisq)
## model.fit3      11 412.42 -195.21
## model.fit1      12 403.09 -189.55  11.327  1   0.000764 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Square feet is not a significant variable

Ha: Square feet is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Square feet is a significant variable.

### 3.Bedrooms

```
model.fit4<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bathrooms+AC+
Garage+Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit4)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit4 as.factor(Quality) ~ `Sales price` + `Square feet` + Bathrooms + AC + Garage +
Pool + Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit4 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC   logLik LR.stat df Pr(>Chisq)
## model.fit4      11 411.23 -194.61
## model.fit1      12 403.09 -189.55  10.135  1   0.001455 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Bedrooms is not a significant variable

Ha: Bedrooms is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Bedrooms is a significant variable.

#### 4. Bathrooms

```
model.fit5<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+AC+
Garage+Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit5)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit5 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + AC + Garage + P
ool + Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit5 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC   logLik LR.stat df Pr(>Chisq)
## model.fit5      11 419.13 -198.56
## model.fit1      12 403.09 -189.55  18.038  1  2.165e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Bathrooms is not a significant variable

Ha: Bathrooms is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Bathrooms is a significant variable.

## 5.AC

```
model.fit6<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+
Garage+Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit6)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit6 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + Gar
age + Pool + Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit6 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit6      11 410.02 -194.01
## model.fit1      12 403.09 -189.55   8.928  1   0.002808 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: AC is not a significant variable

Ha: AC is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Bathrooms is a significant variable.

## 6. Garage

```
model.fit7<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
Pool+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit7)
```



```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit7 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Pool + Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit7 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit7      11 404.34 -191.17
## model.fit1      12 403.09 -189.55  3.2482  1    0.0715 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Garage is not a significant variable

Ha: Garage is a significant variable

Since p-value is greater than  $\alpha = 0.01$ , we fail to reject Ho.

Therefore, Garage is a not significant variable.

## 7.Pool

```
model.fit8<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
Garage+Year+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit8)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit8 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Year + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit8 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit8      11 401.65 -189.82
## model.fit1      12 403.09 -189.55  0.5584  1    0.4549
```

Ho: Pool is not a significant variable

Ha: Pool is a significant variable

Since p-value is greater than  $\alpha = 0.01$ , we fail to reject Ho.

Therefore, Pool is a not significant variable.

8.Year

```
model.fit9<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                Garage+Pool+`lot size`+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit9)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit9 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + `lot size` + Highway
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit9 logit flexible
## model.fit1 logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit9      11 429.59 -203.80
## model.fit1      12 403.09 -189.55  28.499  1 9.373e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Ho: Year is not a significant variable

Ha: Year is a significant variable

Since p-value is less than  $\alpha = 0.01$ , we reject Ho.

Therefore, Year is a significant variable.

9.lot size

```
model.fit10<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                 Garage+Pool+Year+Highway)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit10)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit10 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + Highway
## model.fit1  as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit10 logit flexible
## model.fit1  logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit10      11 401.09 -189.55
## model.fit1       12 403.09 -189.55  0.0024  1    0.9606
```

Ho: lot size is not a significant variable

Ha: lot size is a significant variable

Since p-value is greater than  $\alpha = 0.01$ , we fail to reject Ho.

Therefore, lot size is a not significant variable.

10.Highway

```
model.fit11<- clm(as.factor(Quality) ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
Garage+Pool+Year+`lot size`)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
anova(model.fit1,model.fit11)
```

```
## Likelihood ratio tests of cumulative link models:
##
##          formula:
## model.fit11 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size`
## model.fit1  as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Garage + Pool + Year + `lot size` + Highway
##          link: threshold:
## model.fit11 logit flexible
## model.fit1  logit flexible
##
##          no.par    AIC  logLik LR.stat df Pr(>Chisq)
## model.fit11      11 401.42 -189.71
## model.fit1       12 403.09 -189.55  0.3274  1    0.5672
```

Ho: Highway is not a significant variable

Ha: Highway is a significant variable

Since p-value is greater than  $\alpha = 0.01$ , we fail to reject Ho.

Therefore, Highway is a not significant variable.

Predictors to be retained are:

Important Predictors:

- 1.Sales price
- 2.Square feet
- 3.Bedrooms
- 4.Bathrooms
- 5.AC
- 6.Year

Backward elimination method:

```
library(bootStepAIC)
```

```
## Loading required package: MASS
```

```
stepAIC(model.fit1,direction = "backward")
```

```
## Start:  AIC=403.09
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms +
##   Bathrooms + AC + Garage + Pool + Year + `lot size` + Highway
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
##           Df    AIC
## - `lot size`    1 401.09
## - Highway      1 401.42
## - Pool         1 401.65
## <none>         403.09
## - Garage       1 404.34
## - AC           1 410.02
## - Bedrooms     1 411.23
## - `Square feet` 1 412.42
## - Bathrooms    1 419.13
## - Year         1 429.59
## - `Sales price` 1 464.59
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
##
## Step: AIC=401.09
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms +
## Bathrooms + AC + Garage + Pool + Year + Highway
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
##           Df    AIC
## - Highway    1 399.43
## - Pool        1 399.66
## <none>        401.09
## - Garage      1 402.34
## - AC          1 408.19
## - Bedrooms    1 409.23
## - `Square feet` 1 410.56
## - Bathrooms   1 417.14
## - Year        1 429.15
## - `Sales price` 1 466.77
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
##
## Step: AIC=399.43
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms +
## Bathrooms + AC + Garage + Pool + Year
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```



```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
##           Df    AIC
## - Pool           1 398.03
## <none>           399.43
## - Garage         1 400.78
## - AC             1 406.29
## - Bedrooms       1 407.53
## - `Square feet`  1 408.78
## - Bathrooms      1 415.42
## - Year           1 428.09
## - `Sales price`  1 464.91
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
##
## Step:  AIC=398.03
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms +
##   Bathrooms + AC + Garage + Year
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.
##   Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue
## - Rescale variables?
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
##           Df      AIC  
## <none>      398.03  
## - Garage      1 399.28  
## - AC          1 404.62  
## - Bedrooms    1 406.08  
## - `Square feet` 1 407.22  
## - Bathrooms   1 413.67  
## - Year        1 427.37  
## - `Sales price` 1 462.98
```

```
## formula:  
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC + Garage +  
Year  
##  
## link threshold nobs logLik AIC niter max.grad cond.H  
## logit flexible 522 -190.01 398.03 9(0) 5.76e-08 4.1e+15  
##  
## Coefficients:  
## `Sales price` `Square feet` Bedrooms Bathrooms AC1  
## -1.957e-05 -1.409e-03 5.222e-01 -8.872e-01 -1.135e+00  
## Garage Year  
## -5.429e-01 -5.934e-02  
##  
## Threshold coefficients:  
## 1|2 2|3  
## -134.5 -124.9
```

```
mod.reduced<-clm(as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms +  
AC + Garage + Year, data=real.estate)
```

```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
summary(mod.reduced)
```

```
## formula:  
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC + Garage +  
Year  
## data: real.estate  
##  
## link threshold nobs logLik AIC niter max.grad cond.H  
## logit flexible 522 -190.01 398.03 9(0) 5.76e-08 4.1e+15  
##  
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## `Sales price` -1.957e-05 2.951e-06 -6.632 3.32e-11 ***  
## `Square feet` -1.409e-03 4.453e-04 -3.163 0.00156 **  
## Bedrooms 5.222e-01 1.670e-01 3.127 0.00176 **  
## Bathrooms -8.872e-01 2.201e-01 -4.031 5.56e-05 ***  
## AC1 -1.135e+00 4.037e-01 -2.811 0.00493 **  
## Garage -5.429e-01 3.117e-01 -1.741 0.08161 .  
## Year -5.934e-02 1.147e-02 -5.174 2.29e-07 ***  
## ---  
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Threshold coefficients:  
## Estimate Std. Error z value  
## 1|2 -134.46 22.84 -5.886  
## 2|3 -124.88 22.49 -5.552
```

```
anova(model.fit1 , mod.reduced , type = 1)
```

```
## Likelihood ratio tests of cumulative link models:  
##  
## formula:  
## mod.reduced as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC  
+ Garage + Year  
## model.fit1 as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC  
+ Garage + Pool + Year + `lot size` + Highway  
## link: threshold:  
## mod.reduced logit flexible  
## model.fit1 logit flexible  
##  
## no.par AIC logLik LR.stat df Pr(>Chisq)  
## mod.reduced 9 398.03 -190.01  
## model.fit1 12 403.09 -189.55 0.9376 3 0.8164
```

Variables to be retained:

Sales price

Square feet

Bedrooms

Bathrooms

AC

Year

## Binary Logistic Regression Model

```
estate.data<- as.data.frame(REAL_ESTATE_SALES_DATA)
estate.data$Pool<- as.factor(estate.data$Pool)
estate.data$AC<- as.factor(estate.data$AC)
estate.data$Highway<- as.factor(estate.data$Highway)
library(car)
```

```
## Loading required package: carData
```

```
estate.data$Quality_High <- recode(estate.data$Quality,"1=1;c(2,3)=0")
estate.data$Quality_Medium <- recode(estate.data$Quality,"2=1;c(1,3)=0")
estate.data$Quality_Low <- recode(estate.data$Quality,"3=1;c(1,2)=0")
str(estate.data)
```

```
## 'data.frame':    522 obs. of  16 variables:
## $ Id             : num  1 2 3 4 5 6 7 8 9 10 ...
## $ Sales price    : num  360000 340000 250000 205500 275500 ...
## $ Square feet    : num  3032 2058 1780 1638 2196 ...
## $ Bedrooms       : num  4 4 4 4 4 4 3 2 3 3 ...
## $ Bathrooms      : num  4 2 3 2 3 3 2 1 2 3 ...
## $ AC             : Factor w/ 2 levels "0","1": 2 2 2 2 2 2 2 2 2 1 ...
## $ Garage         : num  2 2 2 2 2 5 2 1 2 1 ...
## $ Pool           : Factor w/ 2 levels "0","1": 1 1 1 1 1 2 1 1 1 1 ...
## $ Year           : num  1972 1976 1980 1963 1968 ...
## $ Quality        : num  2 2 2 2 2 2 2 3 3 ...
## $ Style          : num  1 1 1 1 7 1 7 1 1 1 ...
## $ lot size       : num  22221 22912 21345 17342 21786 ...
## $ Highway        : Factor w/ 2 levels "0","1": 1 1 1 1 1 1 1 1 1 1 ...
## $ Quality_High   : num  0 0 0 0 0 0 0 0 0 0 ...
## $ Quality_Medium : num  1 1 1 1 1 1 1 0 0 ...
## $ Quality_Low    : num  0 0 0 0 0 0 0 0 1 1 ...
```

```
View(estate.data)
```

```
attach(estate.data)
```

```
## The following objects are masked from real.estate:
##
## AC, Bathrooms, Bedrooms, Garage, Highway, Id, lot size, Pool,
## Quality, Sales price, Square feet, Style, Year
```

```
model.bin1<- glm(Quality_High ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                Garage+Pool+Year+`lot size`+Highway,family = binomial(link = "logit"))
summary(model.bin1)
```

```
##
## Call:
## glm(formula = Quality_High ~ `Sales price` + `Square feet` +
##      Bedrooms + Bathrooms + AC + Garage + Pool + Year + `lot size` +
##      Highway, family = binomial(link = "logit"))
##
## Deviance Residuals:
##      Min        1Q    Median        3Q        Max
## -3.3751  -0.1051  -0.0403   0.0000   2.6958
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -4.519e+01  1.439e+03  -0.031   0.9750
## `Sales price`  2.034e-05  3.516e-06   5.787 7.18e-09 ***
## `Square feet`  9.282e-04  6.107e-04   1.520   0.1285
## Bedrooms      -5.666e-01  2.934e-01  -1.931   0.0535 .
## Bathrooms      4.025e-01  3.330e-01   1.209   0.2267
## AC1            1.674e+01  1.439e+03   0.012   0.9907
## Garage         8.570e-01  4.970e-01   1.725   0.0846 .
## Pool1         -3.868e-01  7.882e-01  -0.491   0.6236
## Year           7.897e-03  2.213e-02   0.357   0.7212
## `lot size`    -2.354e-05  2.265e-05  -1.039   0.2986
## Highway1       1.454e+00  1.705e+00   0.853   0.3939
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 403.92  on 521  degrees of freedom
## Residual deviance: 110.24  on 511  degrees of freedom
## AIC: 132.24
##
## Number of Fisher Scoring iterations: 19
```

```
model.bin2<- glm(Quality_Medium ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                Garage+Pool+Year+`lot size`+Highway,family = binomial(link = "logit"))
summary(model.bin2)
```

```
##
## Call:
## glm(formula = Quality_Medium ~ `Sales price` + `Square feet` +
##       Bedrooms + Bathrooms + AC + Garage + Pool + Year + `lot size` +
##       Highway, family = binomial(link = "logit"))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -3.2048  -0.8443   0.4748   0.8148   2.3498
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  -7.616e+01  1.691e+01  -4.504 6.68e-06 ***
## `Sales price` -1.356e-05  1.917e-06  -7.074 1.50e-12 ***
## `Square feet`  5.231e-04  3.251e-04   1.609 0.107601
## Bedrooms      8.308e-02  1.345e-01   0.618 0.536746
## Bathrooms     1.082e+00  1.909e-01   5.669 1.43e-08 ***
## AC1           1.551e+00  3.327e-01   4.662 3.13e-06 ***
## Garage        -7.346e-02  2.044e-01  -0.359 0.719253
## Pool1         6.034e-02  4.125e-01   0.146 0.883718
## Year          3.748e-02  8.680e-03   4.318 1.57e-05 ***
## `lot size`    4.095e-05  1.079e-05   3.794 0.000148 ***
## Highway1     -3.463e-01  6.844e-01  -0.506 0.612876
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 717.19  on 521  degrees of freedom
## Residual deviance: 549.18  on 511  degrees of freedom
## AIC: 571.18
##
## Number of Fisher Scoring iterations: 4
```

```
model.bin3<- glm(Quality_Low ~ `Sales price`+`Square feet`+Bedrooms+Bathrooms+AC+
                  Garage+Pool+Year+`lot size`+Highway,family = binomial(link = "logit"))
summary(model.bin3)
```

```
##
## Call:
## glm(formula = Quality_Low ~ `Sales price` + `Square feet` + Bedrooms +
##     Bathrooms + AC + Garage + Pool + Year + `lot size` + Highway,
##     family = binomial(link = "logit"))
##
## Deviance Residuals:
##      Min       1Q   Median       3Q      Max
## -2.9434  -0.2669  -0.0159   0.1997   3.5518
##
## Coefficients:
##              Estimate Std. Error z value Pr(>|z|)
## (Intercept)  1.701e+02  3.034e+01   5.606 2.07e-08 ***
## `Sales price` -1.318e-05  5.882e-06  -2.240 0.025110 *
## `Square feet` -2.710e-03  7.777e-04  -3.485 0.000493 ***
## Bedrooms      5.834e-01  2.298e-01   2.539 0.011115 *
## Bathrooms     -1.423e+00  3.297e-01  -4.316 1.59e-05 ***
## AC1           -1.013e+00  4.649e-01  -2.179 0.029368 *
## Garage        -4.229e-01  4.468e-01  -0.946 0.343926
## Pool1         9.414e-01  7.739e-01   1.216 0.223829
## Year          -8.150e-02  1.533e-02  -5.315 1.07e-07 ***
## `lot size`    -2.802e-06  1.636e-05  -0.171 0.864028
## Highway1      1.133e-01  9.014e-01   0.126 0.899990
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##      Null deviance: 649.79  on 521  degrees of freedom
## Residual deviance: 246.09  on 511  degrees of freedom
## AIC: 268.09
##
## Number of Fisher Scoring iterations: 8
```

## Comparison of slope

Variable Cumulative Binary:High Quality Binary:Medium Quality Binary: Low Quality

Sales price -0.000020 0.00002 -0.00001 -0.000013

Square feet -0.001431 0.00093 0.00052 -0.002710

Bedrooms 0.525000 -0.56660 0.08308 0.583400

Bathrooms -0.901400 0.40250 1.08200 -1.423000

AC1 -1.179000 16.74000 1.55100 -1.013000

Garage -0.541200 0.85700 -0.07346 -0.422900

Pool1 0.397700 -0.38680 0.06034 0.941400

Year -0.058540 0.00790 0.03748 -0.081500

lot size -0.000001 -0.00002 0.00004 -0.000003

Highway1 -0.433600 1.45400 -0.34630 0.113300

## Final Model

```
mod.final<-clm(as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC
+ Year + Year, data=real.estate)
```



```
## Warning: Using formula(x) is deprecated when x is a character vector of length > 1.  
## Consider formula(paste(x, collapse = " ")) instead.
```

```
## Warning: (2) Model is nearly unidentifiable: very large eigenvalue  
## - Rescale variables?  
## In addition: Absolute and relative convergence criteria were met
```

```
summary(mod.final)
```

```
## formula:  
## as.factor(Quality) ~ `Sales price` + `Square feet` + Bedrooms + Bathrooms + AC + Year + Year  
## data:    real.estate  
##  
## link threshold nobs logLik AIC niter max.grad cond.H  
## logit flexible 522 -191.64 399.28 9(0) 9.28e-08 1.6e+16  
##  
## Coefficients:  
## Estimate Std. Error z value Pr(>|z|)  
## `Sales price` -2.036e-05 2.909e-06 -7.000 2.56e-12 ***  
## `Square feet` -1.491e-03 4.433e-04 -3.364 0.000767 ***  
## Bedrooms 5.242e-01 1.671e-01 3.137 0.001707 **  
## Bathrooms -8.833e-01 2.180e-01 -4.051 5.09e-05 ***  
## AC1 -1.220e+00 3.989e-01 -3.059 0.002219 **  
## Year -6.309e-02 1.121e-02 -5.629 1.82e-08 ***  
## ---  
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
##  
## Threshold coefficients:  
## Estimate Std. Error z value  
## 1|2 -141.19 22.42 -6.297  
## 2|3 -131.54 22.06 -5.964
```