Beer!

Jasmine Dumas

December 12, 2015

1. **Data source**:

beer = read.table("http://www.craftbeeranalytics.com/uploads/3/3/8/9/3389428/ratebeer\_beerjobber.txt", header=TRUE, sep="")  
save(beer, file="beer.RData")  
head(beer)

## name brewer  
## 1 Abbey Monks\u0092 Ale Abbey Beverage Company  
## 2 Abbey Monks\u0092 Tripel Abbey Beverage Company  
## 3 Abbey Monks\u0092 Wit Abbey Beverage Company  
## 4 Alameda Barn Owl Imperial Brown Ale Alameda Brewing Co.  
## 5 Alameda Black Bear XX Stout Alameda Brewing Co.  
## 6 Alameda El Torero Organic IPA Alameda Brewing Co.  
## style abv ratings score.overall score.by.style  
## 1 Belgian Ale 5.2 96 50 49  
## 2 Abbey Tripel 8.0 3 NA NA  
## 3 Belgian White (Witbier) 5.1 46 23 19  
## 4 Brown Ale 7.9 13 74 81  
## 5 Foreign Stout 6.8 172 94 76  
## 6 India Pale Ale (IPA) 7.2 56 74 43

colnames(beer) # we'd be predicting column score.overall using abv, ratings, & score.by.style

## [1] "name" "brewer" "style" "abv"   
## [5] "ratings" "score.overall" "score.by.style"

1. **Data Clean**:

beer <- na.omit(beer)

1. **Research Question**: What contributes to beer rating scores overall, abv, style, brewer?
2. **Dependent Variable**: score.overall
3. **Independent Variable**: score.by.style, name, brewer, style, abv, ratings
4. **Quantitative vs. Qualitative**: The quantitative variables are score.overall, abv, ratings,score.by.style. The qualitative variable is name and brewer.
5. **Correlation and Scatter plots**: We can try a linear model initially.

# no multicollinearity between continuous variables - want to be low!  
cor(beer$abv, beer$ratings)

## [1] 0.1430207

cor(beer$abv, beer$score.by.style)

## [1] 0.107154

cor(beer$ratings, beer$score.by.style)

## [1] 0.2549337

# want correlation between response and covariates - want to be high!  
cor(beer$score.overall, beer$abv)

## [1] 0.4758994

cor(beer$score.overall, beer$ratings)

## [1] 0.3086524

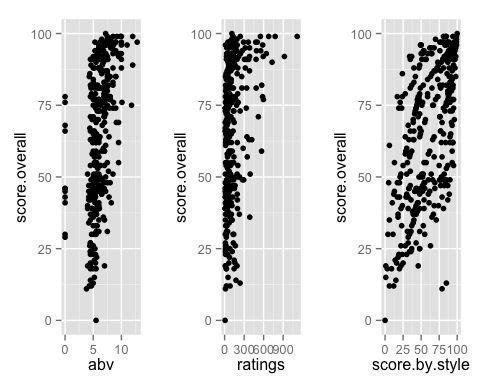
cor(beer$score.overall, beer$score.by.style)

## [1] 0.6066654

# what do these variables all look like against the response y variable  
library(ggplot2)  
a=ggplot(beer, aes(y = score.overall, x = abv)) + geom\_point() # log?  
b=ggplot(beer, aes(y = score.overall, x = ratings)) + geom\_point() # log?  
c=ggplot(beer, aes(y = score.overall, x = score.by.style)) + geom\_point() # weird-linear?  
# these graphs look weird maybe try inverse or logrithm scale and re-graph?  
source('~/Desktop/depaul/CSC423/multiplot.R', echo=TRUE)

##   
## > multiplot <- function(..., plotlist = NULL, file,   
## + cols = 1, layout = NULL) {  
## + library(grid)  
## + plots <- c(list(...), plotlist)  
## + .... [TRUNCATED]

multiplot(a, b, c, cols=3)



1. **Building the model(s)**: The initial model has a high F-Statistic, significant p-value, and High R-squared/adj R-squared value which could be improved upon my removing the variable of brewer, which has many not significant p-values.

After running the second model each of the co-variates deems to be significant (<0.05) except the abv value, the R-squared and Adjusted R-squared are lower than before but are still high (*at 95%, and 94% respectively which means that beer.model1 can account for 95%, and 94% of the error present in the data)*, and the F-Statistic is High along with the significant p-value.

The next 2 models seeks to fit the parsimonious attempts of simplifying the model while producing the aforementioned high F-Statistic, significant p-value, and High R-squared/adj R-squared values.

The automated step-wise regression algorithms verifies that the Lowest AIC score is achieved with the two co-variates of score.by.style + style

# full model  
beer.model = lm(score.overall ~ abv + ratings + score.by.style + brewer + style, data=beer)   
summary(beer.model) # High F-statistic, High R-squared/adj R-squared value

##   
## Call:  
## lm(formula = score.overall ~ abv + ratings + score.by.style +   
## brewer + style, data = beer)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -14.521 -3.278 0.000 2.934 17.814   
##   
## Coefficients: (1 not defined because of singularities)  
## Estimate Std. Error  
## (Intercept) 22.501763 7.196998  
## abv 0.263789 0.290101  
## ratings -0.006820 0.003505  
## score.by.style 0.625776 0.020156  
## brewerAlameda Brewing Co. 8.482498 5.824043  
## brewerAle Industries 14.771222 8.486819  
## brewerAmerican Brewing Company 4.308142 6.311022  
## brewerAppalachian Brewing Company (PA) 4.040653 5.326286  
## brewerBent River Brewing Company 4.465501 6.837067  
## brewerBison Brewing Company 5.134515 5.644076  
## brewerBlue Mountain Brewery 10.151851 5.712417  
## brewerBlue Point Brewing 6.329580 5.346322  
## brewerBoxcar Brewing Company -8.735900 6.102582  
## brewerBrewer\u0092s Alley Restaurant & Brewery 3.994800 6.018330  
## brewerCarolina Brewing Company 7.068853 5.809236  
## brewerCavalry Brewing Company 1.324203 6.143387  
## brewerColorado Cider Co. -15.454786 8.226849  
## brewerCrow Peak Brewing Company 8.320945 6.328713  
## brewerDevils Backbone Brewing Company 2.002875 6.822908  
## brewerDevils Canyon Brewing Co. 6.195032 6.018215  
## brewerDiamond Knot Brewery 4.125524 5.563029  
## brewerDicks Brewing Company 8.349254 5.329566  
## brewerDry Dock Brewing Co. 6.071808 6.301619  
## brewerDuClaw Brewing Company 8.805548 5.320990  
## brewerEddyline Restaurant and Brewery 0.820398 5.904812  
## brewerElevation Beer Company 5.520867 7.812773  
## brewerElysian Brewing Co. 12.114578 5.349251  
## brewerErie Brewing Company 4.367627 5.464159  
## brewerEvil Genius Beer Company 3.534196 6.159517  
## brewerFigure Eight Brewing LLC -0.780034 6.135084  
## brewerFlat 12 Bierwerks 6.381920 5.777496  
## brewerFort Collins Brewery 8.728682 5.833715  
## brewerFrench Broad Brewing Co. 7.122094 5.698862  
## brewerGreat River Brewery 5.852322 6.018870  
## brewerHales Ales 8.562541 5.508580  
## brewerHarlem Brewing Company 7.623857 8.339696  
## brewerHeretic Brewing Company 1.336523 6.022102  
## brewerHigh Noon Saloon & Brewery 0.716413 6.205506  
## brewerHighland Brewing Company 9.195669 5.493809  
## brewerHilliard\u0092s Brewing Company 3.531114 6.644103  
## brewerHopworks Urban Brewery 6.424324 5.926130  
## brewerLa Cumbre Brewing Company 1.425306 8.050169  
## brewerLaurelwood Public House & Brewery 7.330016 5.565252  
## brewerLazy Magnolia Brewing Company 6.078554 5.771092  
## brewerLeft Coast/Oggis Pizza and Brewing Company 4.811738 5.759288  
## brewerMad River Brewery 8.454134 5.747438  
## brewerMadison River Brewing Company -0.790687 5.926726  
## brewerMarble Brewery 3.837506 5.889432  
## brewerMillstream Brewing Company 5.724432 5.387335  
## brewerMission Brewery 3.262050 5.672320  
## brewerMoylans Brewery & Restaurant 3.006386 5.501880  
## brewerNarragansett Brewing Company 4.497825 5.877550  
## brewerOdin Brewing Company 7.793030 6.533353  
## brewerPale Horse Brewing Company 7.573883 6.809429  
## brewerPhiladelphia Brewing Company 2.756365 5.307971  
## brewerPort City Brewing Company 6.800746 5.659262  
## brewerRahr & Sons Brewing 1.833127 6.063768  
## brewerRed Lodge Ales 3.982634 5.630086  
## brewerRenegade Brewing Company 9.805514 7.900531  
## brewerRio Grande Brewing Co. (Sierra Blanca Brewing) 0.620635 6.290024  
## brewerRivertown Brewing 3.868337 5.248642  
## brewerSanta Fe Brewing Company 6.466106 5.494908  
## brewerSierra Blanca Brewing Co 1.735987 5.583230  
## brewerSix Rivers Brewery 4.102379 5.682184  
## brewerStraight To Ale 4.443409 7.990520  
## brewerTerrapin Beer Company 7.921199 6.177003  
## brewerThunderhead Brewing Company 1.514937 7.111648  
## brewerTwisted Pine Brewing -1.555676 5.716633  
## brewerUpslope Brewing Company 3.407582 5.773095  
## brewerVoodoo Brewing Co. 5.737933 5.629712  
## styleAbbey Tripel 8.394655 6.664550  
## styleAltbier -5.774166 5.952843  
## styleAmber Ale -13.519711 5.245309  
## styleAmerican Pale Ale 1.523029 4.973159  
## styleAmerican Strong Ale 20.705923 5.119791  
## styleBarley Wine 23.826306 5.757386  
## styleBelgian Ale -4.334474 6.496316  
## styleBelgian Strong Ale 9.287018 6.963854  
## styleBelgian White (Witbier) -11.970404 5.331904  
## styleBerliner Weisse 11.259340 8.379767  
## styleBière de Garde -5.592234 8.527838  
## styleBlack IPA 25.933340 6.700280  
## styleBohemian Pilsener -18.669362 6.173070  
## styleBrown Ale -6.178933 5.130451  
## styleCalifornia Common -9.758405 8.679384  
## styleCider NA NA  
## styleClassic German Pilsener -27.403523 6.924676  
## styleCream Ale -32.354621 7.060512  
## styleDoppelbock 2.103145 8.026878  
## styleDortmunder/Helles -29.623592 5.927405  
## styleDry Stout 5.844091 6.018426  
## styleDunkel -27.797805 8.041290  
## styleDunkler Bock -8.962482 6.651520  
## styleEnglish Pale Ale -18.588133 6.924326  
## styleEnglish Strong Ale 0.528681 8.002945  
## styleForeign Stout 9.520270 6.846091  
## styleFruit Beer -23.369246 5.322322  
## styleGerman Hefeweizen -15.992143 5.994191  
## styleGolden Ale/Blond Ale -21.119189 5.683677  
## styleHeller Bock -17.400725 6.755413  
## styleImperial Stout 30.568414 5.197961  
## styleImperial/Double IPA 25.948729 5.129070  
## styleImperial/Strong Porter 39.322899 8.006166  
## styleIndia Pale Ale (IPA) 12.359871 4.955677  
## styleIrish Ale -17.057579 6.139811  
## styleKölsch -27.196792 5.625288  
## styleLambic - Fruit 24.580720 7.996046  
## styleLambic - Gueuze 42.863038 7.979405  
## styleLambic - Unblended 16.125752 8.047691  
## styleOktoberfest/Märzen -25.274374 5.323168  
## styleOld Ale 10.673579 8.282627  
## stylePale Lager -66.617989 7.324140  
## stylePilsener -31.122886 5.756353  
## stylePorter 8.664853 5.079093  
## stylePremium Bitter/ESB -12.001113 6.248699  
## stylePremium Lager -42.426247 6.880556  
## styleSaison 19.746398 7.043779  
## styleSchwarzbier -2.496530 7.001026  
## styleScotch Ale 17.486330 5.369023  
## styleScottish Ale -4.995499 5.937094  
## styleSmoked 8.741628 8.315996  
## styleSour Ale/Wild Ale 28.056240 6.586082  
## styleSpecialty Grain -5.659704 5.720840  
## styleSpice/Herb/Vegetable -11.890902 5.044040  
## styleStout 8.036409 5.598755  
## styleStrong Pale Lager/Imperial Pils -8.678294 8.540139  
## styleSweet Stout 10.136642 5.274300  
## styleVienna -30.454358 5.707029  
## styleWeizen Bock 13.821124 7.953598  
## styleWheat Ale -27.305449 5.589263  
## t value Pr(>|t|)   
## (Intercept) 3.127 0.001989 \*\*   
## abv 0.909 0.364112   
## ratings -1.946 0.052843 .   
## score.by.style 31.047 < 2e-16 \*\*\*  
## brewerAlameda Brewing Co. 1.456 0.146583   
## brewerAle Industries 1.740 0.083065 .   
## brewerAmerican Brewing Company 0.683 0.495500   
## brewerAppalachian Brewing Company (PA) 0.759 0.448828   
## brewerBent River Brewing Company 0.653 0.514302   
## brewerBison Brewing Company 0.910 0.363892   
## brewerBlue Mountain Brewery 1.777 0.076820 .   
## brewerBlue Point Brewing 1.184 0.237629   
## brewerBoxcar Brewing Company -1.432 0.153596   
## brewerBrewer\u0092s Alley Restaurant & Brewery 0.664 0.507479   
## brewerCarolina Brewing Company 1.217 0.224874   
## brewerCavalry Brewing Company 0.216 0.829524   
## brewerColorado Cider Co. -1.879 0.061524 .   
## brewerCrow Peak Brewing Company 1.315 0.189845   
## brewerDevils Backbone Brewing Company 0.294 0.769356   
## brewerDevils Canyon Brewing Co. 1.029 0.304346   
## brewerDiamond Knot Brewery 0.742 0.459063   
## brewerDicks Brewing Company 1.567 0.118539   
## brewerDry Dock Brewing Co. 0.964 0.336259   
## brewerDuClaw Brewing Company 1.655 0.099269 .   
## brewerEddyline Restaurant and Brewery 0.139 0.889617   
## brewerElevation Beer Company 0.707 0.480478   
## brewerElysian Brewing Co. 2.265 0.024430 \*   
## brewerErie Brewing Company 0.799 0.424900   
## brewerEvil Genius Beer Company 0.574 0.566660   
## brewerFigure Eight Brewing LLC -0.127 0.898934   
## brewerFlat 12 Bierwerks 1.105 0.270441   
## brewerFort Collins Brewery 1.496 0.135914   
## brewerFrench Broad Brewing Co. 1.250 0.212622   
## brewerGreat River Brewery 0.972 0.331874   
## brewerHales Ales 1.554 0.121418   
## brewerHarlem Brewing Company 0.914 0.361555   
## brewerHeretic Brewing Company 0.222 0.824554   
## brewerHigh Noon Saloon & Brewery 0.115 0.908187   
## brewerHighland Brewing Company 1.674 0.095479 .   
## brewerHilliard\u0092s Brewing Company 0.531 0.595592   
## brewerHopworks Urban Brewery 1.084 0.279432   
## brewerLa Cumbre Brewing Company 0.177 0.859617   
## brewerLaurelwood Public House & Brewery 1.317 0.189070   
## brewerLazy Magnolia Brewing Company 1.053 0.293283   
## brewerLeft Coast/Oggis Pizza and Brewing Company 0.835 0.404289   
## brewerMad River Brewery 1.471 0.142628   
## brewerMadison River Brewing Company -0.133 0.893982   
## brewerMarble Brewery 0.652 0.515293   
## brewerMillstream Brewing Company 1.063 0.289053   
## brewerMission Brewery 0.575 0.565779   
## brewerMoylans Brewery & Restaurant 0.546 0.585283   
## brewerNarragansett Brewing Company 0.765 0.444878   
## brewerOdin Brewing Company 1.193 0.234133   
## brewerPale Horse Brewing Company 1.112 0.267147   
## brewerPhiladelphia Brewing Company 0.519 0.604043   
## brewerPort City Brewing Company 1.202 0.230673   
## brewerRahr & Sons Brewing 0.302 0.762681   
## brewerRed Lodge Ales 0.707 0.480020   
## brewerRenegade Brewing Company 1.241 0.215783   
## brewerRio Grande Brewing Co. (Sierra Blanca Brewing) 0.099 0.921484   
## brewerRivertown Brewing 0.737 0.461838   
## brewerSanta Fe Brewing Company 1.177 0.240473   
## brewerSierra Blanca Brewing Co 0.311 0.756127   
## brewerSix Rivers Brewery 0.722 0.471020   
## brewerStraight To Ale 0.556 0.578675   
## brewerTerrapin Beer Company 1.282 0.200961   
## brewerThunderhead Brewing Company 0.213 0.831492   
## brewerTwisted Pine Brewing -0.272 0.785757   
## brewerUpslope Brewing Company 0.590 0.555582   
## brewerVoodoo Brewing Co. 1.019 0.309132   
## styleAbbey Tripel 1.260 0.209048   
## styleAltbier -0.970 0.333039   
## styleAmber Ale -2.577 0.010555 \*   
## styleAmerican Pale Ale 0.306 0.759682   
## styleAmerican Strong Ale 4.044 7.09e-05 \*\*\*  
## styleBarley Wine 4.138 4.86e-05 \*\*\*  
## styleBelgian Ale -0.667 0.505278   
## styleBelgian Strong Ale 1.334 0.183609   
## styleBelgian White (Witbier) -2.245 0.025684 \*   
## styleBerliner Weisse 1.344 0.180346   
## styleBière de Garde -0.656 0.512611   
## styleBlack IPA 3.870 0.000140 \*\*\*  
## styleBohemian Pilsener -3.024 0.002765 \*\*   
## styleBrown Ale -1.204 0.229645   
## styleCalifornia Common -1.124 0.262010   
## styleCider NA NA   
## styleClassic German Pilsener -3.957 0.000100 \*\*\*  
## styleCream Ale -4.582 7.42e-06 \*\*\*  
## styleDoppelbock 0.262 0.793538   
## styleDortmunder/Helles -4.998 1.12e-06 \*\*\*  
## styleDry Stout 0.971 0.332518   
## styleDunkel -3.457 0.000647 \*\*\*  
## styleDunkler Bock -1.347 0.179122   
## styleEnglish Pale Ale -2.684 0.007775 \*\*   
## styleEnglish Strong Ale 0.066 0.947385   
## styleForeign Stout 1.391 0.165641   
## styleFruit Beer -4.391 1.70e-05 \*\*\*  
## styleGerman Hefeweizen -2.668 0.008156 \*\*   
## styleGolden Ale/Blond Ale -3.716 0.000253 \*\*\*  
## styleHeller Bock -2.576 0.010605 \*   
## styleImperial Stout 5.881 1.37e-08 \*\*\*  
## styleImperial/Double IPA 5.059 8.42e-07 \*\*\*  
## styleImperial/Strong Porter 4.912 1.68e-06 \*\*\*  
## styleIndia Pale Ale (IPA) 2.494 0.013308 \*   
## styleIrish Ale -2.778 0.005902 \*\*   
## styleKölsch -4.835 2.39e-06 \*\*\*  
## styleLambic - Fruit 3.074 0.002358 \*\*   
## styleLambic - Gueuze 5.372 1.85e-07 \*\*\*  
## styleLambic - Unblended 2.004 0.046228 \*   
## styleOktoberfest/Märzen -4.748 3.55e-06 \*\*\*  
## styleOld Ale 1.289 0.198764   
## stylePale Lager -9.096 < 2e-16 \*\*\*  
## stylePilsener -5.407 1.56e-07 \*\*\*  
## stylePorter 1.706 0.089316 .   
## stylePremium Bitter/ESB -1.921 0.055980 .   
## stylePremium Lager -6.166 2.98e-09 \*\*\*  
## styleSaison 2.803 0.005474 \*\*   
## styleSchwarzbier -0.357 0.721711   
## styleScotch Ale 3.257 0.001290 \*\*   
## styleScottish Ale -0.841 0.400966   
## styleSmoked 1.051 0.294241   
## styleSour Ale/Wild Ale 4.260 2.95e-05 \*\*\*  
## styleSpecialty Grain -0.989 0.323515   
## styleSpice/Herb/Vegetable -2.357 0.019213 \*   
## styleStout 1.435 0.152488   
## styleStrong Pale Lager/Imperial Pils -1.016 0.310577   
## styleSweet Stout 1.922 0.055813 .   
## styleVienna -5.336 2.21e-07 \*\*\*  
## styleWeizen Bock 1.738 0.083554 .   
## styleWheat Ale -4.885 1.90e-06 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.041 on 238 degrees of freedom  
## Multiple R-squared: 0.9596, Adjusted R-squared: 0.9378   
## F-statistic: 44.12 on 128 and 238 DF, p-value: < 2.2e-16

# remove brewer  
beer.model1 = lm(score.overall ~ abv + ratings + score.by.style + style, data=beer)   
summary(beer.model1) # High F-statistic, High R-squared/adj R-squared value

##   
## Call:  
## lm(formula = score.overall ~ abv + ratings + score.by.style +   
## style, data = beer)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.8047 -3.3136 -0.0985 4.0117 24.4522   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 27.654443 4.744288 5.829  
## abv 0.304527 0.256139 1.189  
## ratings -0.004995 0.002472 -2.021  
## score.by.style 0.673950 0.014636 46.048  
## styleAbbey Tripel 7.973348 6.196424 1.287  
## styleAltbier -10.549431 5.384420 -1.959  
## styleAmber Ale -16.845722 4.727586 -3.563  
## styleAmerican Pale Ale -1.346310 4.561642 -0.295  
## styleAmerican Strong Ale 18.631642 4.713666 3.953  
## styleBarley Wine 20.008081 5.253523 3.809  
## styleBelgian Ale -7.452130 5.368179 -1.388  
## styleBelgian Strong Ale 8.150218 6.190808 1.317  
## styleBelgian White (Witbier) -16.029521 4.811131 -3.332  
## styleBerliner Weisse 5.490370 7.584265 0.724  
## styleBière de Garde -5.281861 7.579265 -0.697  
## styleBlack IPA 23.652989 6.169936 3.834  
## styleBohemian Pilsener -21.987587 5.678361 -3.872  
## styleBrown Ale -9.601081 4.661549 -2.060  
## styleCalifornia Common -14.647200 7.603968 -1.926  
## styleCider -24.810933 6.186889 -4.010  
## styleClassic German Pilsener -32.646385 6.205354 -5.261  
## styleCream Ale -34.141323 6.201576 -5.505  
## styleDoppelbock -0.109162 7.576884 -0.014  
## styleDortmunder/Helles -34.685173 5.387520 -6.438  
## styleDry Stout 1.186526 5.398656 0.220  
## styleDunkel -33.491694 7.585561 -4.415  
## styleDunkler Bock -13.071079 6.179022 -2.115  
## styleEnglish Pale Ale -21.472439 6.199392 -3.464  
## styleEnglish Strong Ale -5.070295 7.568346 -0.670  
## styleForeign Stout 9.500725 6.215598 1.529  
## styleFruit Beer -28.136610 4.843906 -5.809  
## styleGerman Hefeweizen -18.845472 5.394671 -3.493  
## styleGolden Ale/Blond Ale -23.435233 4.996461 -4.690  
## styleHeller Bock -21.496045 6.170220 -3.484  
## styleImperial Stout 27.395638 4.769934 5.743  
## styleImperial/Double IPA 22.309836 4.640998 4.807  
## styleImperial/Strong Porter 36.797804 7.578858 4.855  
## styleIndia Pale Ale (IPA) 9.404556 4.466409 2.106  
## styleIrish Ale -21.513881 5.645621 -3.811  
## styleKölsch -30.398383 5.124676 -5.932  
## styleLambic - Fruit 19.697348 7.563649 2.604  
## styleLambic - Gueuze 39.627127 7.563494 5.239  
## styleLambic - Unblended 10.344677 7.575060 1.366  
## styleOktoberfest/Märzen -28.262789 4.850876 -5.826  
## styleOld Ale 10.241173 7.559409 1.355  
## stylePale Lager -71.631418 6.239188 -11.481  
## stylePilsener -35.391828 5.221257 -6.778  
## stylePorter 5.945551 4.613370 1.289  
## stylePremium Bitter/ESB -12.702579 5.659969 -2.244  
## stylePremium Lager -45.394492 6.224201 -7.293  
## styleSaison 17.101729 6.231680 2.744  
## styleSchwarzbier -6.164050 6.233466 -0.989  
## styleScotch Ale 14.152053 4.889827 2.894  
## styleScottish Ale -7.799332 5.383122 -1.449  
## styleSmoked 9.569238 7.637045 1.253  
## styleSour Ale/Wild Ale 29.295431 6.189355 4.733  
## styleSpecialty Grain -8.728866 5.059111 -1.725  
## styleSpice/Herb/Vegetable -14.954355 4.545662 -3.290  
## styleStout 4.622577 5.019361 0.921  
## styleStrong Pale Lager/Imperial Pils -8.516360 7.577386 -1.124  
## styleSweet Stout 6.927213 4.764939 1.454  
## styleVienna -35.963362 5.089455 -7.066  
## styleWeizen Bock 10.023126 7.556420 1.326  
## styleWheat Ale -31.984750 4.906146 -6.519  
## Pr(>|t|)   
## (Intercept) 1.43e-08 \*\*\*  
## abv 0.235405   
## ratings 0.044171 \*   
## score.by.style < 2e-16 \*\*\*  
## styleAbbey Tripel 0.199158   
## styleAltbier 0.051000 .   
## styleAmber Ale 0.000425 \*\*\*  
## styleAmerican Pale Ale 0.768091   
## styleAmerican Strong Ale 9.62e-05 \*\*\*  
## styleBarley Wine 0.000169 \*\*\*  
## styleBelgian Ale 0.166095   
## styleBelgian Strong Ale 0.189000   
## styleBelgian White (Witbier) 0.000970 \*\*\*  
## styleBerliner Weisse 0.469676   
## styleBière de Garde 0.486410   
## styleBlack IPA 0.000154 \*\*\*  
## styleBohemian Pilsener 0.000132 \*\*\*  
## styleBrown Ale 0.040288 \*   
## styleCalifornia Common 0.055008 .   
## styleCider 7.65e-05 \*\*\*  
## styleClassic German Pilsener 2.71e-07 \*\*\*  
## styleCream Ale 7.87e-08 \*\*\*  
## styleDoppelbock 0.988515   
## styleDortmunder/Helles 4.73e-10 \*\*\*  
## styleDry Stout 0.826189   
## styleDunkel 1.41e-05 \*\*\*  
## styleDunkler Bock 0.035212 \*   
## styleEnglish Pale Ale 0.000610 \*\*\*  
## styleEnglish Strong Ale 0.503410   
## styleForeign Stout 0.127424   
## styleFruit Beer 1.59e-08 \*\*\*  
## styleGerman Hefeweizen 0.000548 \*\*\*  
## styleGolden Ale/Blond Ale 4.13e-06 \*\*\*  
## styleHeller Bock 0.000567 \*\*\*  
## styleImperial Stout 2.26e-08 \*\*\*  
## styleImperial/Double IPA 2.41e-06 \*\*\*  
## styleImperial/Strong Porter 1.93e-06 \*\*\*  
## styleIndia Pale Ale (IPA) 0.036060 \*   
## styleIrish Ale 0.000168 \*\*\*  
## styleKölsch 8.17e-09 \*\*\*  
## styleLambic - Fruit 0.009662 \*\*   
## styleLambic - Gueuze 3.02e-07 \*\*\*  
## styleLambic - Unblended 0.173070   
## styleOktoberfest/Märzen 1.45e-08 \*\*\*  
## styleOld Ale 0.176504   
## stylePale Lager < 2e-16 \*\*\*  
## stylePilsener 6.35e-11 \*\*\*  
## stylePorter 0.198463   
## stylePremium Bitter/ESB 0.025536 \*   
## stylePremium Lager 2.66e-12 \*\*\*  
## styleSaison 0.006425 \*\*   
## styleSchwarzbier 0.323519   
## styleScotch Ale 0.004077 \*\*   
## styleScottish Ale 0.148414   
## styleSmoked 0.211171   
## styleSour Ale/Wild Ale 3.40e-06 \*\*\*  
## styleSpecialty Grain 0.085479 .   
## styleSpice/Herb/Vegetable 0.001121 \*\*   
## styleStout 0.357809   
## styleStrong Pale Lager/Imperial Pils 0.261937   
## styleSweet Stout 0.147040   
## styleVienna 1.10e-11 \*\*\*  
## styleWeizen Bock 0.185693   
## styleWheat Ale 2.95e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.169 on 303 degrees of freedom  
## Multiple R-squared: 0.9463, Adjusted R-squared: 0.9351   
## F-statistic: 84.76 on 63 and 303 DF, p-value: < 2.2e-16

# remove abv  
beer.model2 = lm(score.overall ~ ratings + score.by.style + style, data=beer)   
summary(beer.model2) # better significant results

##   
## Call:  
## lm(formula = score.overall ~ ratings + score.by.style + style,   
## data = beer)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.3808 -3.2848 -0.0807 4.0663 24.6375   
##   
## Coefficients:  
## Estimate Std. Error t value  
## (Intercept) 29.681391 4.430388 6.700  
## ratings -0.005001 0.002473 -2.022  
## score.by.style 0.678039 0.014236 47.629  
## styleAbbey Tripel 8.624215 6.176389 1.396  
## styleAltbier -11.313214 5.349594 -2.115  
## styleAmber Ale -17.537186 4.694865 -3.735  
## styleAmerican Pale Ale -1.967154 4.534735 -0.434  
## styleAmerican Strong Ale 18.768365 4.715467 3.980  
## styleBarley Wine 21.061701 5.181758 4.065  
## styleBelgian Ale -7.945603 5.355748 -1.484  
## styleBelgian Strong Ale 8.748327 6.174530 1.417  
## styleBelgian White (Witbier) -16.688578 4.782339 -3.490  
## styleBerliner Weisse 4.782279 7.565984 0.632  
## styleBière de Garde -4.931818 7.578693 -0.651  
## styleBlack IPA 23.599198 6.173965 3.822  
## styleBohemian Pilsener -22.846206 5.636078 -4.054  
## styleBrown Ale -10.257994 4.631835 -2.215  
## styleCalifornia Common -15.094401 7.599822 -1.986  
## styleCider -25.052514 6.187755 -4.049  
## styleClassic German Pilsener -33.421290 6.175225 -5.412  
## styleCream Ale -34.882617 6.174348 -5.650  
## styleDoppelbock 0.157881 7.578703 0.021  
## styleDortmunder/Helles -35.496444 5.347766 -6.638  
## styleDry Stout 0.284205 5.348677 0.053  
## styleDunkel -34.099650 7.573451 -4.503  
## styleDunkler Bock -13.417532 6.176344 -2.172  
## styleEnglish Pale Ale -22.098152 6.181212 -3.575  
## styleEnglish Strong Ale -5.099650 7.573451 -0.673  
## styleForeign Stout 9.295430 6.217423 1.495  
## styleFruit Beer -28.556624 4.834290 -5.907  
## styleGerman Hefeweizen -19.649088 5.355797 -3.669  
## styleGolden Ale/Blond Ale -24.224671 4.955511 -4.888  
## styleHeller Bock -21.492260 6.174415 -3.481  
## styleImperial Stout 27.942321 4.750948 5.881  
## styleImperial/Double IPA 22.664817 4.634532 4.890  
## styleImperial/Strong Porter 37.076984 7.580370 4.891  
## styleIndia Pale Ale (IPA) 9.081346 4.461159 2.036  
## styleIrish Ale -21.940861 5.638018 -3.892  
## styleKölsch -31.485967 5.045799 -6.240  
## styleLambic - Fruit 19.372000 7.563837 2.561  
## styleLambic - Gueuze 39.410312 7.566436 5.209  
## styleLambic - Unblended 9.893424 7.570689 1.307  
## styleOktoberfest/Märzen -28.854806 4.828532 -5.976  
## styleOld Ale 10.120935 7.563872 1.338  
## stylePale Lager -72.652988 6.183939 -11.749  
## stylePilsener -36.182547 5.182246 -6.982  
## stylePorter 5.407518 4.594242 1.177  
## stylePremium Bitter/ESB -13.284952 5.642566 -2.354  
## stylePremium Lager -46.311763 6.180398 -7.493  
## styleSaison 16.139968 6.183153 2.610  
## styleSchwarzbier -7.044730 6.193505 -1.137  
## styleScotch Ale 14.349278 4.890335 2.934  
## styleScottish Ale -8.491839 5.355155 -1.586  
## styleSmoked 8.755311 7.611471 1.150  
## styleSour Ale/Wild Ale 29.638636 6.186823 4.791  
## styleSpecialty Grain -9.255538 5.043106 -1.835  
## styleSpice/Herb/Vegetable -15.569277 4.519212 -3.445  
## styleStout 4.084051 5.002280 0.816  
## styleStrong Pale Lager/Imperial Pils -8.608968 7.582137 -1.135  
## styleSweet Stout 6.461126 4.752014 1.360  
## styleVienna -36.690968 5.055961 -7.257  
## styleWeizen Bock 9.992779 7.561514 1.322  
## styleWheat Ale -32.616823 4.880573 -6.683  
## Pr(>|t|)   
## (Intercept) 1.01e-10 \*\*\*  
## ratings 0.044076 \*   
## score.by.style < 2e-16 \*\*\*  
## styleAbbey Tripel 0.163637   
## styleAltbier 0.035262 \*   
## styleAmber Ale 0.000224 \*\*\*  
## styleAmerican Pale Ale 0.664744   
## styleAmerican Strong Ale 8.62e-05 \*\*\*  
## styleBarley Wine 6.13e-05 \*\*\*  
## styleBelgian Ale 0.138960   
## styleBelgian Strong Ale 0.157553   
## styleBelgian White (Witbier) 0.000555 \*\*\*  
## styleBerliner Weisse 0.527812   
## styleBière de Garde 0.515701   
## styleBlack IPA 0.000160 \*\*\*  
## styleBohemian Pilsener 6.41e-05 \*\*\*  
## styleBrown Ale 0.027524 \*   
## styleCalifornia Common 0.047914 \*   
## styleCider 6.54e-05 \*\*\*  
## styleClassic German Pilsener 1.26e-07 \*\*\*  
## styleCream Ale 3.70e-08 \*\*\*  
## styleDoppelbock 0.983393   
## styleDortmunder/Helles 1.46e-10 \*\*\*  
## styleDry Stout 0.957659   
## styleDunkel 9.58e-06 \*\*\*  
## styleDunkler Bock 0.030597 \*   
## styleEnglish Pale Ale 0.000407 \*\*\*  
## styleEnglish Strong Ale 0.501231   
## styleForeign Stout 0.135936   
## styleFruit Beer 9.32e-09 \*\*\*  
## styleGerman Hefeweizen 0.000288 \*\*\*  
## styleGolden Ale/Blond Ale 1.65e-06 \*\*\*  
## styleHeller Bock 0.000573 \*\*\*  
## styleImperial Stout 1.07e-08 \*\*\*  
## styleImperial/Double IPA 1.63e-06 \*\*\*  
## styleImperial/Strong Porter 1.63e-06 \*\*\*  
## styleIndia Pale Ale (IPA) 0.042653 \*   
## styleIrish Ale 0.000122 \*\*\*  
## styleKölsch 1.47e-09 \*\*\*  
## styleLambic - Fruit 0.010916 \*   
## styleLambic - Gueuze 3.52e-07 \*\*\*  
## styleLambic - Unblended 0.192266   
## styleOktoberfest/Märzen 6.40e-09 \*\*\*  
## styleOld Ale 0.181876   
## stylePale Lager < 2e-16 \*\*\*  
## stylePilsener 1.83e-11 \*\*\*  
## stylePorter 0.240108   
## stylePremium Bitter/ESB 0.019188 \*   
## stylePremium Lager 7.37e-13 \*\*\*  
## styleSaison 0.009494 \*\*   
## styleSchwarzbier 0.256251   
## styleScotch Ale 0.003599 \*\*   
## styleScottish Ale 0.113840   
## styleSmoked 0.250933   
## styleSour Ale/Wild Ale 2.60e-06 \*\*\*  
## styleSpecialty Grain 0.067440 .   
## styleSpice/Herb/Vegetable 0.000651 \*\*\*  
## styleStout 0.414889   
## styleStrong Pale Lager/Imperial Pils 0.257091   
## styleSweet Stout 0.174945   
## styleVienna 3.32e-12 \*\*\*  
## styleWeizen Bock 0.187318   
## styleWheat Ale 1.12e-10 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.173 on 304 degrees of freedom  
## Multiple R-squared: 0.9461, Adjusted R-squared: 0.9351   
## F-statistic: 85.99 on 62 and 304 DF, p-value: < 2.2e-16

# remove ratings  
beer.model3 = lm(score.overall ~ score.by.style + style, data=beer)   
summary(beer.model3) # better significant results

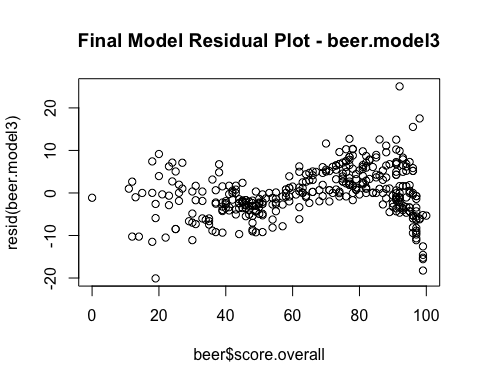
##   
## Call:  
## lm(formula = score.overall ~ score.by.style + style, data = beer)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -20.0895 -3.3172 -0.0302 3.9047 25.0566   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)  
## (Intercept) 29.86026 4.45187 6.707 9.62e-11  
## score.by.style 0.66918 0.01361 49.154 < 2e-16  
## styleAbbey Tripel 8.31193 6.20563 1.339 0.181432  
## styleAltbier -11.11292 5.37568 -2.067 0.039553  
## styleAmber Ale -17.89978 4.71512 -3.796 0.000177  
## styleAmerican Pale Ale -2.37735 4.55307 -0.522 0.601950  
## styleAmerican Strong Ale 18.68479 4.73909 3.943 1.00e-04  
## styleBarley Wine 20.24728 5.19216 3.900 0.000119  
## styleBelgian Ale -7.72206 5.38164 -1.435 0.152343  
## styleBelgian Strong Ale 8.68807 6.20563 1.400 0.162521  
## styleBelgian White (Witbier) -16.71209 4.80647 -3.477 0.000581  
## styleBerliner Weisse 4.71828 7.60412 0.620 0.535398  
## styleBière de Garde -4.41768 7.61267 -0.580 0.562136  
## styleBlack IPA 23.68051 6.20500 3.816 0.000164  
## styleBohemian Pilsener -23.00378 5.66399 -4.061 6.21e-05  
## styleBrown Ale -10.37557 4.65485 -2.229 0.026544  
## styleCalifornia Common -15.52944 7.63513 -2.034 0.042822  
## styleCider -24.56421 6.21426 -3.953 9.60e-05  
## styleClassic German Pilsener -33.18429 6.20528 -5.348 1.75e-07  
## styleCream Ale -34.67296 6.20465 -5.588 5.09e-08  
## styleDoppelbock -0.24395 7.61435 -0.032 0.974463  
## styleDortmunder/Helles -35.74811 5.37331 -6.653 1.33e-10  
## styleDry Stout 0.06193 5.37455 0.012 0.990814  
## styleDunkel -33.74094 7.60960 -4.434 1.29e-05  
## styleDunkler Bock -13.36858 6.20748 -2.154 0.032053  
## styleEnglish Pale Ale -22.27794 6.21178 -3.586 0.000390  
## styleEnglish Strong Ale -4.74094 7.60960 -0.623 0.533736  
## styleForeign Stout 8.08609 6.21983 1.300 0.194565  
## styleFruit Beer -28.49690 4.85861 -5.865 1.17e-08  
## styleGerman Hefeweizen -19.38747 5.38127 -3.603 0.000367  
## styleGolden Ale/Blond Ale -24.53744 4.97810 -4.929 1.36e-06  
## styleHeller Bock -21.34970 6.20518 -3.441 0.000661  
## styleImperial Stout 27.49389 4.76973 5.764 2.01e-08  
## styleImperial/Double IPA 21.80296 4.63818 4.701 3.93e-06  
## styleImperial/Strong Porter 36.76360 7.61705 4.826 2.20e-06  
## styleIndia Pale Ale (IPA) 8.56004 4.47619 1.912 0.056768  
## styleIrish Ale -22.13671 5.66565 -3.907 0.000115  
## styleKölsch -31.37488 5.07097 -6.187 1.97e-09  
## styleLambic - Fruit 19.63519 7.60090 2.583 0.010252  
## styleLambic - Gueuze 39.38747 7.60463 5.179 4.05e-07  
## styleLambic - Unblended 9.92824 7.60889 1.305 0.192938  
## styleOktoberfest/Märzen -28.73070 4.85252 -5.921 8.62e-09  
## styleOld Ale 10.29682 7.60156 1.355 0.176558  
## stylePale Lager -72.73339 6.21503 -11.703 < 2e-16  
## stylePilsener -35.73339 5.20362 -6.867 3.68e-11  
## stylePorter 5.21017 4.61639 1.129 0.259944  
## stylePremium Bitter/ESB -13.27468 5.67105 -2.341 0.019888  
## stylePremium Lager -46.04910 6.21023 -7.415 1.21e-12  
## styleSaison 16.06421 6.21426 2.585 0.010200  
## styleSchwarzbier -8.02266 6.20576 -1.293 0.197067  
## styleScotch Ale 13.77266 4.90666 2.807 0.005324  
## styleScottish Ale -9.06004 5.37478 -1.686 0.092884  
## styleSmoked 7.04154 7.60231 0.926 0.355055  
## styleSour Ale/Wild Ale 29.57554 6.21798 4.756 3.04e-06  
## styleSpecialty Grain -9.41491 5.06795 -1.858 0.064170  
## styleSpice/Herb/Vegetable -15.63326 4.54192 -3.442 0.000658  
## styleStout 2.96147 4.99647 0.593 0.553812  
## styleStrong Pale Lager/Imperial Pils -8.09442 7.61612 -1.063 0.288713  
## styleSweet Stout 6.31934 4.77549 1.323 0.186731  
## styleVienna -37.01159 5.07899 -7.287 2.73e-12  
## styleWeizen Bock 10.01888 7.59968 1.318 0.188382  
## styleWheat Ale -32.66352 4.90516 -6.659 1.28e-10  
##   
## (Intercept) \*\*\*  
## score.by.style \*\*\*  
## styleAbbey Tripel   
## styleAltbier \*   
## styleAmber Ale \*\*\*  
## styleAmerican Pale Ale   
## styleAmerican Strong Ale \*\*\*  
## styleBarley Wine \*\*\*  
## styleBelgian Ale   
## styleBelgian Strong Ale   
## styleBelgian White (Witbier) \*\*\*  
## styleBerliner Weisse   
## styleBière de Garde   
## styleBlack IPA \*\*\*  
## styleBohemian Pilsener \*\*\*  
## styleBrown Ale \*   
## styleCalifornia Common \*   
## styleCider \*\*\*  
## styleClassic German Pilsener \*\*\*  
## styleCream Ale \*\*\*  
## styleDoppelbock   
## styleDortmunder/Helles \*\*\*  
## styleDry Stout   
## styleDunkel \*\*\*  
## styleDunkler Bock \*   
## styleEnglish Pale Ale \*\*\*  
## styleEnglish Strong Ale   
## styleForeign Stout   
## styleFruit Beer \*\*\*  
## styleGerman Hefeweizen \*\*\*  
## styleGolden Ale/Blond Ale \*\*\*  
## styleHeller Bock \*\*\*  
## styleImperial Stout \*\*\*  
## styleImperial/Double IPA \*\*\*  
## styleImperial/Strong Porter \*\*\*  
## styleIndia Pale Ale (IPA) .   
## styleIrish Ale \*\*\*  
## styleKölsch \*\*\*  
## styleLambic - Fruit \*   
## styleLambic - Gueuze \*\*\*  
## styleLambic - Unblended   
## styleOktoberfest/Märzen \*\*\*  
## styleOld Ale   
## stylePale Lager \*\*\*  
## stylePilsener \*\*\*  
## stylePorter   
## stylePremium Bitter/ESB \*   
## stylePremium Lager \*\*\*  
## styleSaison \*   
## styleSchwarzbier   
## styleScotch Ale \*\*   
## styleScottish Ale .   
## styleSmoked   
## styleSour Ale/Wild Ale \*\*\*  
## styleSpecialty Grain .   
## styleSpice/Herb/Vegetable \*\*\*  
## styleStout   
## styleStrong Pale Lager/Imperial Pils   
## styleSweet Stout   
## styleVienna \*\*\*  
## styleWeizen Bock   
## styleWheat Ale \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 6.205 on 305 degrees of freedom  
## Multiple R-squared: 0.9453, Adjusted R-squared: 0.9344   
## F-statistic: 86.46 on 61 and 305 DF, p-value: < 2.2e-16

# automated selection algorithm using stepwise regression for the final   
full.beer = step(beer.model3, direction="both")

## Start: AIC=1395.84  
## score.overall ~ score.by.style + style  
##   
## Df Sum of Sq RSS AIC  
## <none> 11741 1395.8  
## - style 60 123983 135725 2174.1  
## - score.by.style 1 93013 104754 2197.0

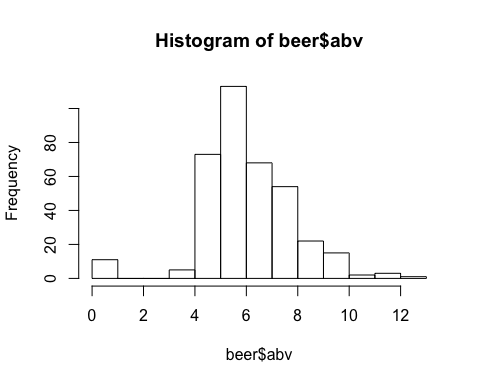
1. **Checking of assumptions:** A noticeable sign of slight multi-colinearity can be observed in checking the correlation between cor(beer$ratings, beer$score.by.style)
2. **Checking for interaction terms:** Without expansive domain experience in the craft beer industry and with the few amounts of co-variates I would not test or see the rationale behind interaction terms for this model.
3. **Checking for higher order models:** As noticed in the scatter plots, there are some intriguing distributions of the co-variates against the response variable, y. Slight curves could signal a second order model or a different transformation.
4. **Examination of residuals:** The residual plot takes a look at the observed values versus the predicted values from the final model. (Residual = Observed – Predicted; or e = y - y-hat). There doesn't seem to be an extreme obvious trend in the dispersion of the plot.

plot(beer$score.overall, resid(beer.model3), main = "Final Model Residual Plot - beer.model3")

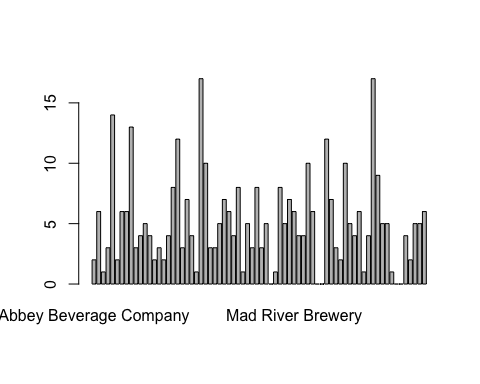


1. **Transformations if appropriate:** Check for distributions of co-variates in histograms/plots.

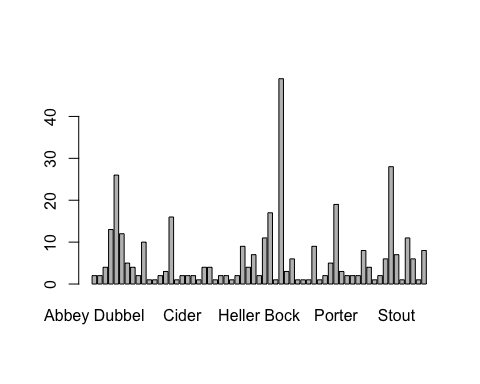
hist(beer$abv)



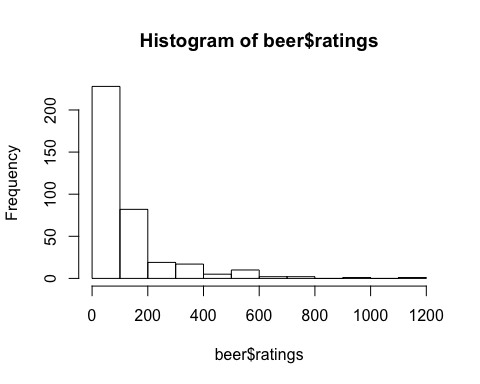
plot(beer$brewer)



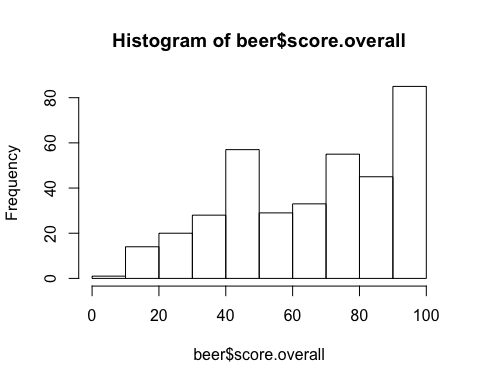
plot(beer$style)



hist(beer$ratings)



hist(beer$score.overall)



hist(beer$score.by.style)

