Beer!

Jasmine Dumas

October 28, 2015

```r
#maybe beer data
beer = read.table("http://www.craftbeeranalytics.com/uploads/3/3/8/9/3389428/ratebeer_beerjobber.txt", l
beer <- na.omit(beer)
head(beer)
```

```
##                                 name                   brewer
## 1              Abbey Monks\u0092 Ale 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.
## 8        Alameda Klickitat Pale Ale    Alameda Brewing Co.
##                       style abv ratings score.overall score.by.style
## 1               Belgian Ale 5.2      96            50             49
## 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
## 8         American Pale Ale 5.3      61            75             67
```

```r
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"
```

```r
beer.model1 = lm(score.overall ~ abv + ratings + score.by.style + brewer, data=beer) # low F-statistic
summary(beer.model1)
```

```
##
## Call:
## lm(formula = score.overall ~ abv + ratings + score.by.style +
##     brewer, data = beer)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -35.503  -9.078   0.009   8.185  47.517
##
## Coefficients:
##                                   Estimate Std. Error
## (Intercept)                      -2.063739  11.029285
## abv                               4.823654   0.493029
## ratings                           0.023574   0.007496
## score.by.style                    0.354358   0.039296
## brewerAlameda Brewing Co.        25.613985  12.401813
## brewerAle Industries             32.602104  18.610577
```

```
## brewerAmerican Brewing Company                              20.271655  13.860327
## brewerAppalachian Brewing Company (PA)                       10.276268  11.405152
## brewerBent River Brewing Company                             11.383453  15.072532
## brewerBison Brewing Company                                  12.800075  12.373496
## brewerBlue Mountain Brewery                                  19.636117  12.508307
## brewerBlue Point Brewing                                      7.485508  11.577373
## brewerBoxcar Brewing Company                                 -5.566236  13.782675
## brewerBrewer\u0092s Alley Restaurant & Brewery               13.449466  13.080319
## brewerCarolina Brewing Company                               15.440743  12.635875
## brewerCavalry Brewing Company                                11.959239  13.083659
## brewerColorado Cider Co.                                     -0.899429  15.215825
## brewerCrow Peak Brewing Company                              10.648352  13.756216
## brewerDevils Backbone Brewing Company                        14.095972  15.196162
## brewerDevils Canyon Brewing Co.                              24.397450  13.061585
## brewerDiamond Knot Brewery                                   10.476951  12.035378
## brewerDicks Brewing Company                                  21.423391  11.558622
## brewerDry Dock Brewing Co.                                   30.107586  13.919866
## brewerDuClaw Brewing Company                                 26.200251  12.154315
## brewerEddyline Restaurant and Brewery                         4.863433  13.053689
## brewerElevation Beer Company                                 14.679968  18.611396
## brewerElysian Brewing Co.                                    23.019999  11.443879
## brewerErie Brewing Company                                    1.928521  11.699782
## brewerEvil Genius Beer Company                                3.008693  13.749845
## brewerFigure Eight Brewing LLC                               16.909931  13.786675
## brewerFlat 12 Bierwerks                                      26.258765  12.717157
## brewerFort Collins Brewery                                   14.333465  12.321932
## brewerFrench Broad Brewing Co.                               19.131124  12.401225
## brewerGreat River Brewery                                    18.152002  13.121507
## brewerHales Ales                                             24.300190  11.934062
## brewerHarlem Brewing Company                                -11.186019  18.451692
## brewerHeretic Brewing Company                                29.900419  12.723570
## brewerHigh Noon Saloon & Brewery                             -8.498791  13.750064
## brewerHighland Brewing Company                               17.871757  12.148384
## brewerHilliard\u0092s Brewing Company                        32.949598  13.850728
## brewerHopworks Urban Brewery                                 29.851638  12.748937
## brewerLa Cumbre Brewing Company                              29.257328  18.649242
## brewerLaurelwood Public House & Brewery                      24.713625  12.066129
## brewerLazy Magnolia Brewing Company                          21.514891  12.689644
## brewerLeft Coast/Oggis Pizza and Brewing Company             17.303589  12.236725
## brewerMad River Brewery                                      14.400913  12.641947
## brewerMadison River Brewing Company                           4.165780  13.044918
## brewerMarble Brewery                                         24.099125  13.281163
## brewerMillstream Brewing Company                              6.198793  11.769147
## brewerMission Brewery                                        13.796479  12.397699
## brewerMoylans Brewery & Restaurant                           12.515142  11.923679
## brewerNarragansett Brewing Company                           -9.282352  12.113993
## brewerOdin Brewing Company                                   13.222595  13.872543
## brewerPale Horse Brewing Company                             25.267753  15.066741
## brewerPhiladelphia Brewing Company                           10.665080  11.679919
## brewerPort City Brewing Company                              17.846144  12.681720
## brewerRahr & Sons Brewing                                    -4.724961  13.049993
## brewerRed Lodge Ales                                          3.224594  12.318306
## brewerRenegade Brewing Company                               25.180121  18.469441
## brewerRio Grande Brewing Co. (Sierra Blanca Brewing)         -5.764029  13.061171
```

```
## brewerRivertown Brewing                            13.683274  11.294708
## brewerSanta Fe Brewing Company                      17.363445  11.884309
## brewerSierra Blanca Brewing Co                       5.587135  12.603067
## brewerSix Rivers Brewery                            13.696665  12.609297
## brewerStraight To Ale                               28.777048  18.566031
## brewerTerrapin Beer Company                         13.612763  13.373094
## brewerThunderhead Brewing Company                  -13.102400  15.129508
## brewerTwisted Pine Brewing                           0.356491  12.612691
## brewerUpslope Brewing Company                         0.845362  12.634198
## brewerVoodoo Brewing Co.                            13.677199  12.431032
##                                                     t value Pr(>|t|)
## (Intercept)                                          -0.187  0.85170
## abv                                                   9.784  < 2e-16 ***
## ratings                                               3.145  0.00183 **
## score.by.style                                        9.018  < 2e-16 ***
## brewerAlameda Brewing Co.                             2.065  0.03976 *
## brewerAle Industries                                  1.752  0.08084 .
## brewerAmerican Brewing Company                        1.463  0.14464
## brewerAppalachian Brewing Company (PA)                0.901  0.36831
## brewerBent River Brewing Company                      0.755  0.45070
## brewerBison Brewing Company                           1.034  0.30176
## brewerBlue Mountain Brewery                           1.570  0.11752
## brewerBlue Point Brewing                              0.647  0.51841
## brewerBoxcar Brewing Company                         -0.404  0.68661
## brewerBrewer\u0092s Alley Restaurant & Brewery       1.028  0.30468
## brewerCarolina Brewing Company                        1.222  0.22269
## brewerCavalry Brewing Company                         0.914  0.36143
## brewerColorado Cider Co.                             -0.059  0.95290
## brewerCrow Peak Brewing Company                       0.774  0.43950
## brewerDevils Backbone Brewing Company                 0.928  0.35437
## brewerDevils Canyon Brewing Co.                       1.868  0.06276 .
## brewerDiamond Knot Brewery                            0.871  0.38472
## brewerDicks Brewing Company                           1.853  0.06481 .
## brewerDry Dock Brewing Co.                            2.163  0.03134 *
## brewerDuClaw Brewing Company                          2.156  0.03192 *
## brewerEddyline Restaurant and Brewery                 0.373  0.70973
## brewerElevation Beer Company                          0.789  0.43088
## brewerElysian Brewing Co.                             2.012  0.04517 *
## brewerErie Brewing Company                            0.165  0.86919
## brewerEvil Genius Beer Company                        0.219  0.82694
## brewerFigure Eight Brewing LLC                        1.227  0.22097
## brewerFlat 12 Bierwerks                               2.065  0.03981 *
## brewerFort Collins Brewery                            1.163  0.24566
## brewerFrench Broad Brewing Co.                        1.543  0.12397
## brewerGreat River Brewery                             1.383  0.16759
## brewerHales Ales                                      2.036  0.04262 *
## brewerHarlem Brewing Company                         -0.606  0.54482
## brewerHeretic Brewing Company                         2.350  0.01943 *
## brewerHigh Noon Saloon & Brewery                     -0.618  0.53699
## brewerHighland Brewing Company                        1.471  0.14232
## brewerHilliard\u0092s Brewing Company                 2.379  0.01800 *
## brewerHopworks Urban Brewery                          2.342  0.01987 *
## brewerLa Cumbre Brewing Company                       1.569  0.11775
## brewerLaurelwood Public House & Brewery               2.048  0.04142 *
```

```
## brewerLazy Magnolia Brewing Company                          1.695  0.09103 .
## brewerLeft Coast/Oggis Pizza and Brewing Company             1.414  0.15839
## brewerMad River Brewery                                      1.139  0.25556
## brewerMadison River Brewing Company                          0.319  0.74969
## brewerMarble Brewery                                         1.815  0.07060 .
## brewerMillstream Brewing Company                             0.527  0.59880
## brewerMission Brewery                                        1.113  0.26668
## brewerMoylans Brewery & Restaurant                           1.050  0.29475
## brewerNarragansett Brewing Company                          -0.766  0.44414
## brewerOdin Brewing Company                                   0.953  0.34129
## brewerPale Horse Brewing Company                             1.677  0.09458 .
## brewerPhiladelphia Brewing Company                           0.913  0.36192
## brewerPort City Brewing Company                              1.407  0.16040
## brewerRahr & Sons Brewing                                   -0.362  0.71756
## brewerRed Lodge Ales                                         0.262  0.79368
## brewerRenegade Brewing Company                               1.363  0.17381
## brewerRio Grande Brewing Co. (Sierra Blanca Brewing)  -0.441  0.65931
## brewerRivertown Brewing                                      1.211  0.22668
## brewerSanta Fe Brewing Company                               1.461  0.14506
## brewerSierra Blanca Brewing Co                               0.443  0.65786
## brewerSix Rivers Brewery                                     1.086  0.27826
## brewerStraight To Ale                                        1.550  0.12221
## brewerTerrapin Beer Company                                  1.018  0.30954
## brewerThunderhead Brewing Company                           -0.866  0.38718
## brewerTwisted Pine Brewing                                   0.028  0.97747
## brewerUpslope Brewing Company                                0.067  0.94670
## brewerVoodoo Brewing Co.                                     1.100  0.27212
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 15.05 on 297 degrees of freedom
## Multiple R-squared:  0.6866, Adjusted R-squared:  0.6138
## F-statistic: 9.429 on 69 and 297 DF,  p-value: < 2.2e-16
```

```r
# re-run model
beer.model2 = lm(score.overall ~ abv + ratings + score.by.style, data=beer) # better significant result
summary(beer.model2)
```

```
##
## Call:
## lm(formula = score.overall ~ abv + ratings + score.by.style,
##     data = beer)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -60.612 -12.262   1.678  12.149  51.591
##
## Coefficients:
##                 Estimate Std. Error t value Pr(>|t|)
## (Intercept)     2.257859   3.355362   0.673  0.50143
## abv             5.211170   0.461658  11.288  < 2e-16 ***
## ratings         0.018356   0.005852   3.137  0.00185 **
## score.by.style 0.481523   0.032870  14.649  < 2e-16 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 16.3 on 363 degrees of freedom
## Multiple R-squared:  0.551,  Adjusted R-squared:  0.5473
## F-statistic: 148.5 on 3 and 363 DF,  p-value: < 2.2e-16
```

```
# another run with style as a dummy variable
beer.model3 = lm(score.overall ~ abv + ratings + score.by.style + style, data=beer)
summary(beer.model3)
```

```
##
## 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
```

```r
# stepwise regression for the final beer.model
full.beer = step(beer.model3, direction="both")
```

```
## Start:  AIC=1393.23
## score.overall ~ abv + ratings + score.by.style + style
```

```
##
##                 Df Sum of Sq    RSS    AIC
## - abv            1        54  11586 1392.9
## <none>                        11532 1393.2
## - ratings        1       155  11687 1396.2
## - style         60     84898  96430 2052.6
## - score.by.style 1     80699  92231 2154.3
##
## Step:  AIC=1392.94
## score.overall ~ ratings + score.by.style + style
##
##                 Df Sum of Sq    RSS    AIC
## <none>                        11586 1392.9
## + abv            1        54  11532 1393.2
## - ratings        1       156  11741 1395.8
## - style         60    118692 130278 2161.1
## - score.by.style 1     86453  98039 2174.7
```

```r
# no multicollinearity between variables
cor(beer$abv, beer$ratings)
```

```
## [1] 0.1430207
```

```r
cor(beer$abv, beer$score.by.style)
```
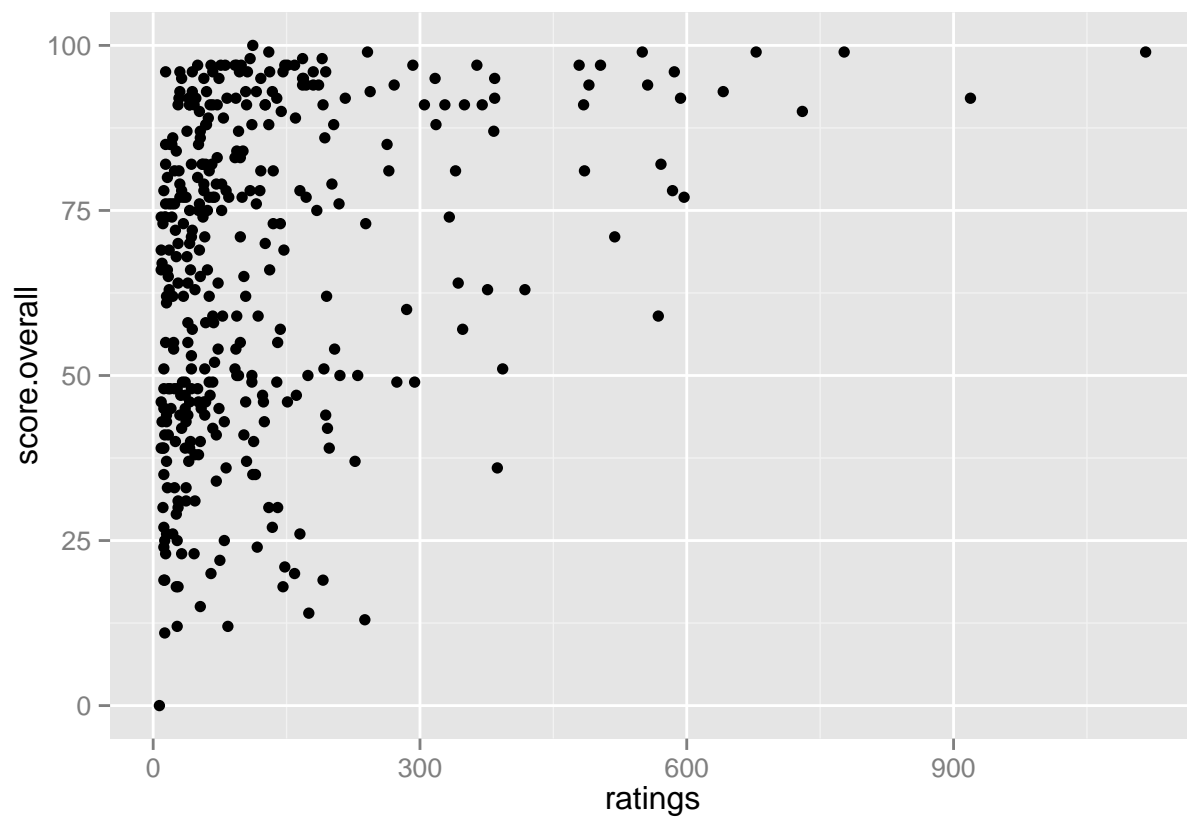
```
## [1] 0.107154
```

```r
cor(beer$ratings, beer$score.by.style)
```
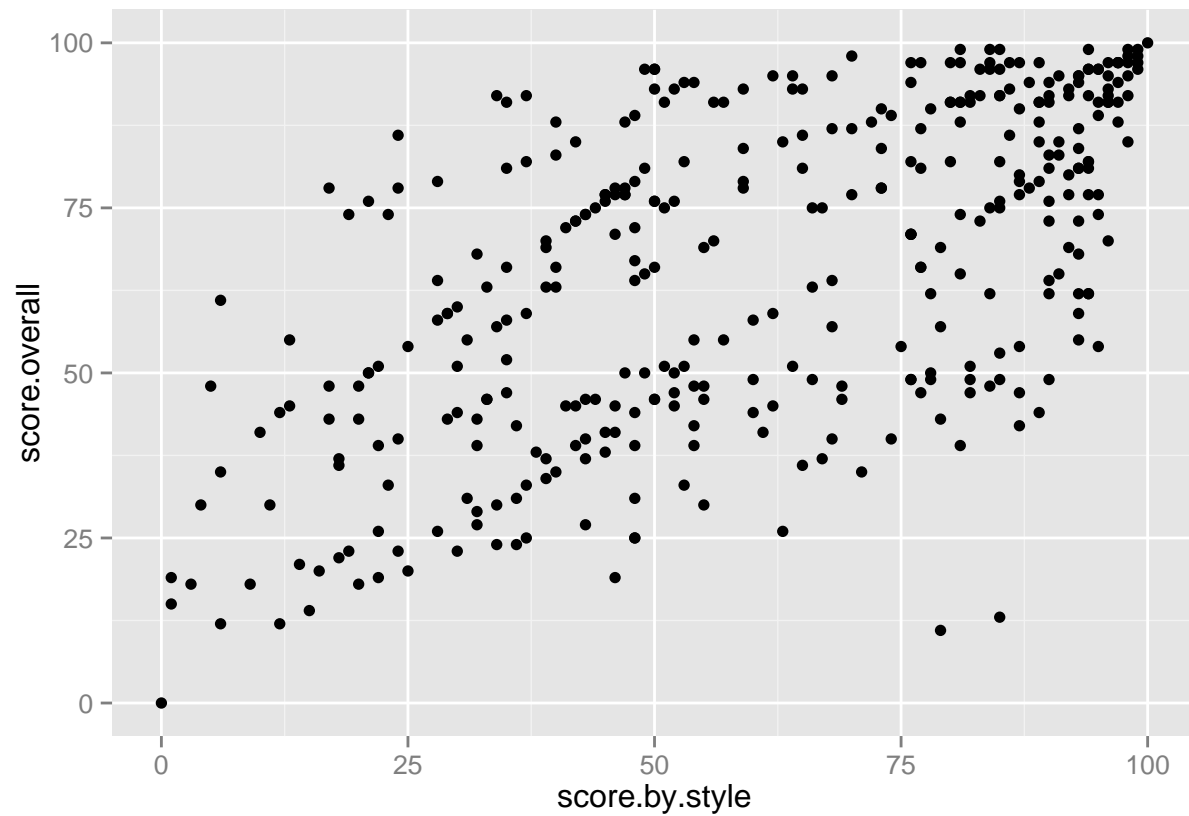
```
## [1] 0.2549337
```

```r
# what do these variables all look like against the response y varaible
library(ggplot2)
ggplot(beer, aes(y = score.overall, x = abv)) + geom_point() # log?
```
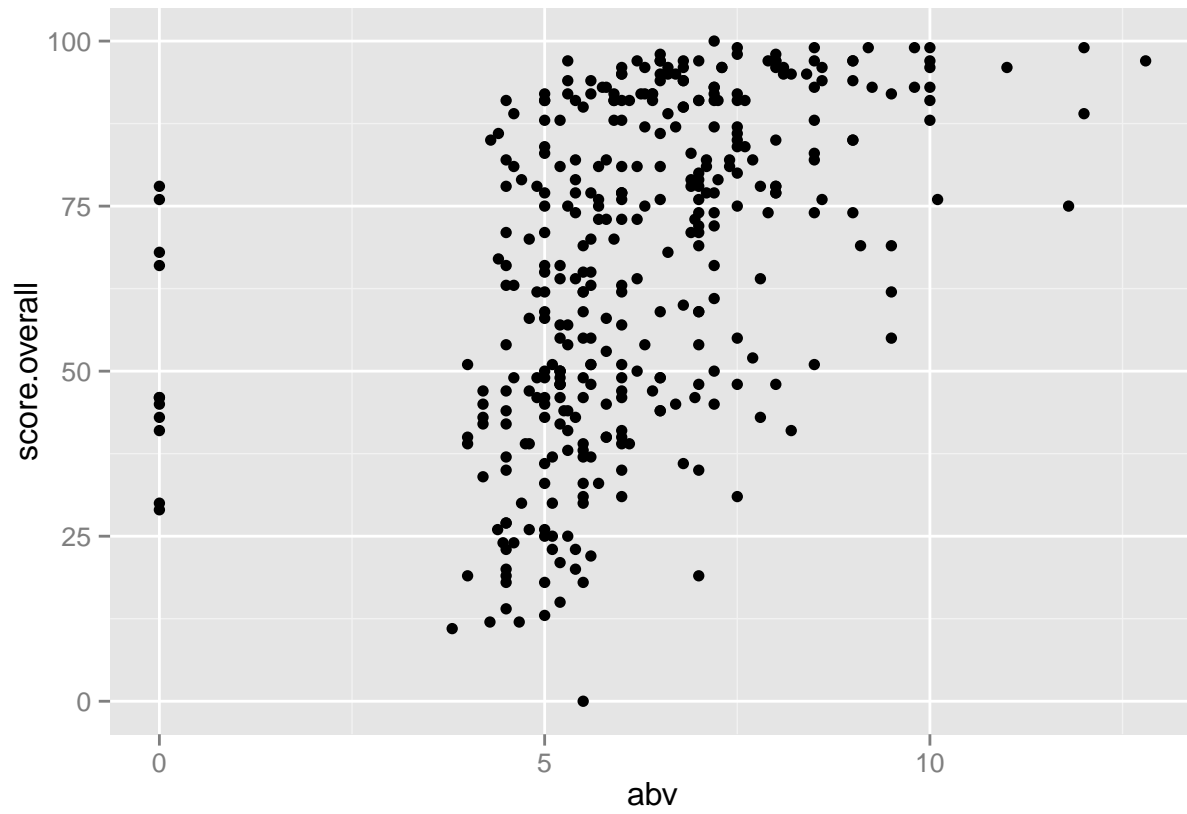
```
ggplot(beer, aes(y = score.overall, x = ratings)) + geom_point() # log?
```

```r
ggplot(beer, aes(y = score.overall, x = score.by.style)) + geom_point() # weird-linear?
```



```r
# these graphs look weird maybe try inverse or logrithm scale and re-graph?
```