Project 2 : Explartion of Red Wine

Isabel María Villalba Jiménez

6 de diciembre de 2016

```
##  [1] "X"                    "fixed.acidity"        "volatile.acidity"
##  [4] "citric.acid"          "residual.sugar"       "chlorides"
##  [7] "free.sulfur.dioxide"  "total.sulfur.dioxide" "density"
## [10] "pH"                   "sulphates"            "alcohol"
## [13] "quality"
```

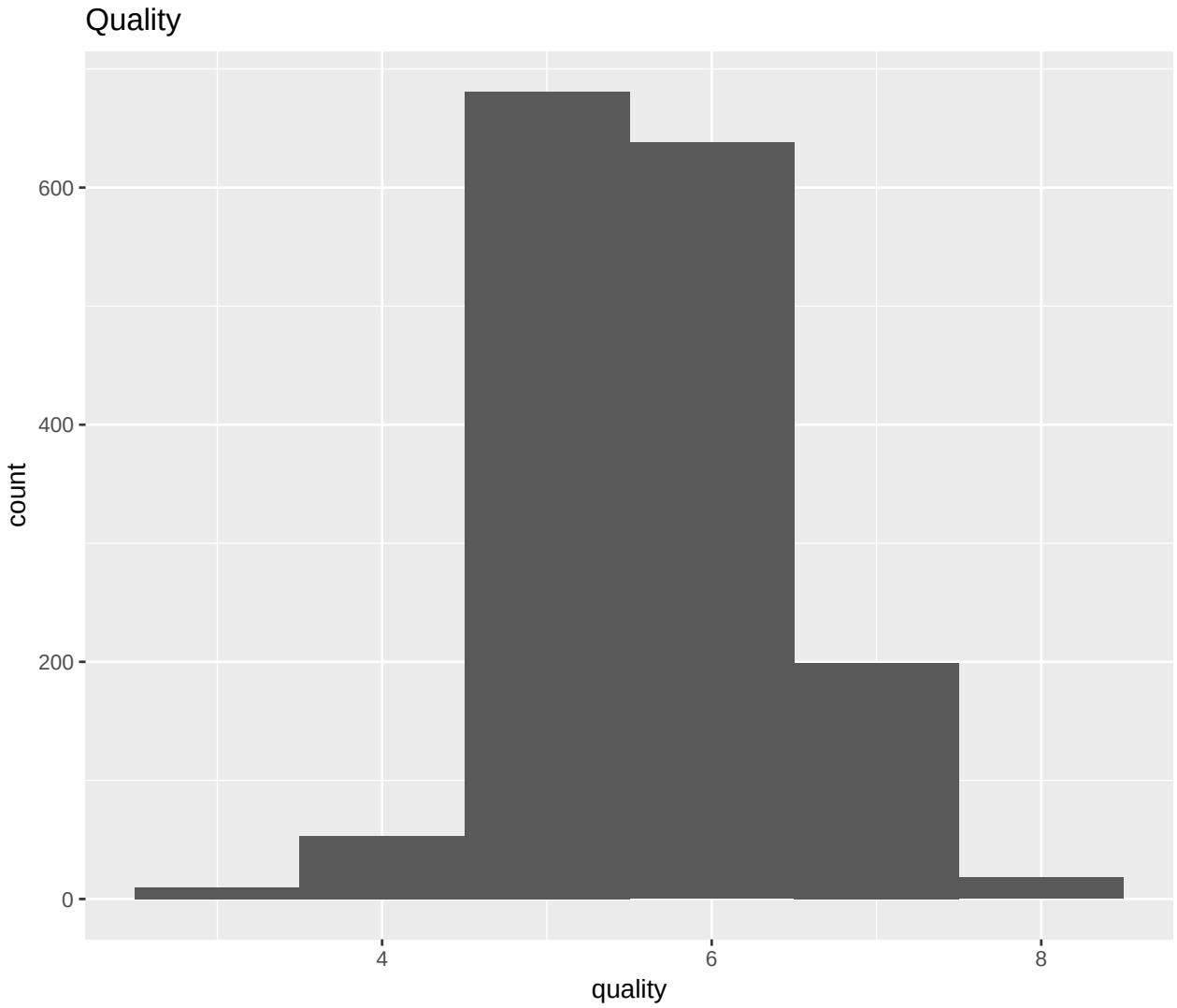
```
##        X          fixed.acidity   volatile.acidity  citric.acid
##  Min.   :   1.0   Min.   : 4.60   Min.   :0.1200   Min.   :0.000
##  1st Qu.: 400.5   1st Qu.: 7.10   1st Qu.:0.3900   1st Qu.:0.090
##  Median : 800.0   Median : 7.90   Median :0.5200   Median :0.260
##  Mean   : 800.0   Mean   : 8.32   Mean   :0.5278   Mean   :0.271
##  3rd Qu.:1199.5   3rd Qu.: 9.20   3rd Qu.:0.6400   3rd Qu.:0.420
##  Max.   :1599.0   Max.   :15.90   Max.   :1.5800   Max.   :1.000
##  residual.sugar     chlorides       free.sulfur.dioxide
##  Min.   : 0.900   Min.   :0.01200   Min.   : 1.00
##  1st Qu.: 1.900   1st Qu.:0.07000   1st Qu.: 7.00
##  Median : 2.200   Median :0.07900   Median :14.00
##  Mean   : 2.539   Mean   :0.08747   Mean   :15.87
##  3rd Qu.: 2.600   3rd Qu.:0.09000   3rd Qu.:21.00
##  Max.   :15.500   Max.   :0.61100   Max.   :72.00
##  total.sulfur.dioxide    density             pH          sulphates
##  Min.   :  6.00       Min.   :0.9901   Min.   :2.740   Min.   :0.3300
##  1st Qu.: 22.00       1st Qu.:0.9956   1st Qu.:3.210   1st Qu.:0.5500
##  Median : 38.00       Median :0.9968   Median :3.310   Median :0.6200
##  Mean   : 46.47       Mean   :0.9967   Mean   :3.311   Mean   :0.6581
##  3rd Qu.: 62.00       3rd Qu.:0.9978   3rd Qu.:3.400   3rd Qu.:0.7300
##  Max.   :289.00       Max.   :1.0037   Max.   :4.010   Max.   :2.0000
##     alcohol         quality
##  Min.   : 8.40   Min.   :3.000
##  1st Qu.: 9.50   1st Qu.:5.000
##  Median :10.20   Median :6.000
##  Mean   :10.42   Mean   :5.636
##  3rd Qu.:11.10   3rd Qu.:6.000
##  Max.   :14.90   Max.   :8.000
```

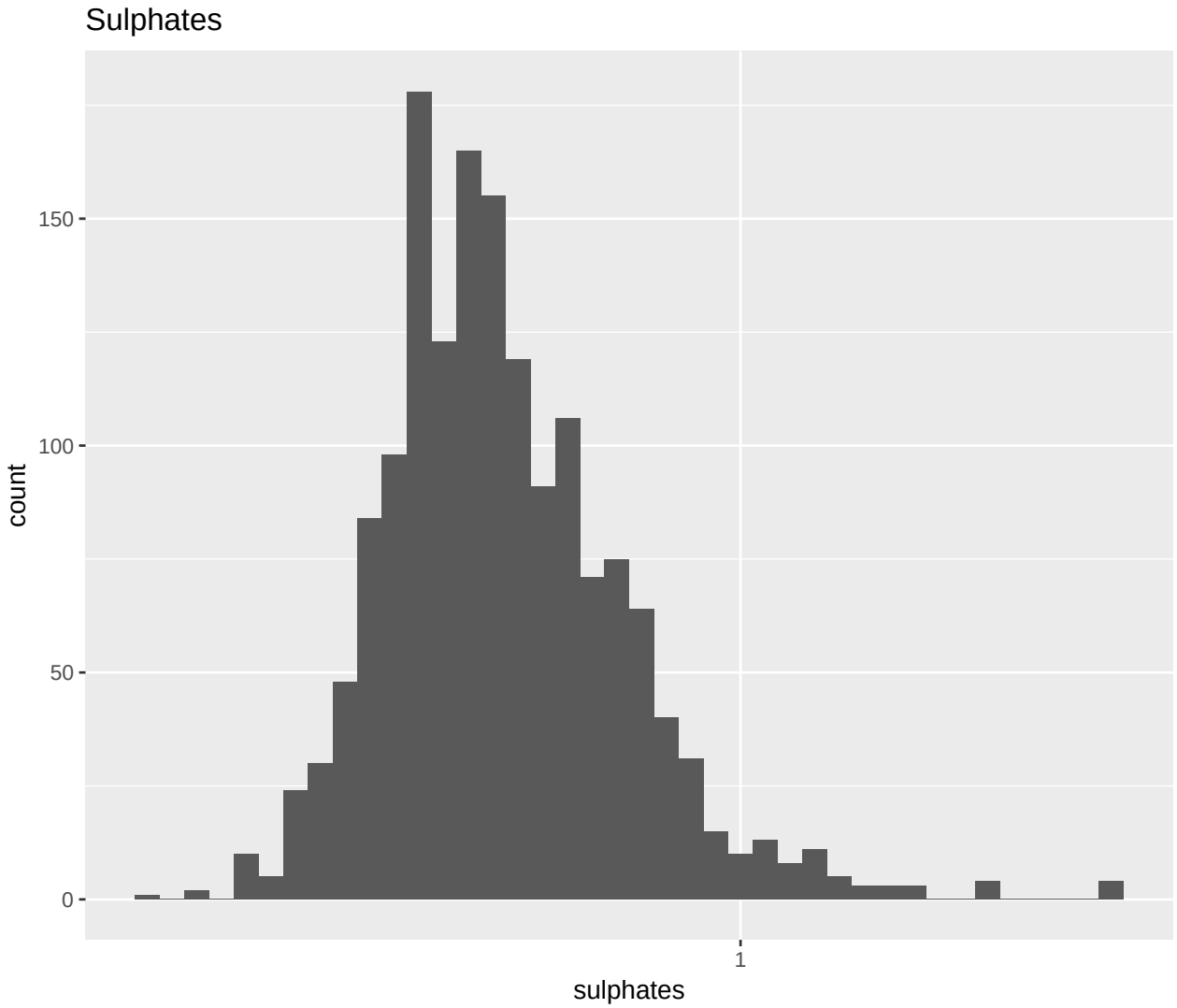
```
## 'data.frame':    1599 obs. of  15 variables:
##  $ X                   : int  1 2 3 4 5 6 7 8 9 10 ...
##  $ fixed.acidity       : num  7.4 7.8 7.8 11.2 7.4 7.4 7.9 7.3 7.8 7.5 ...
##  $ volatile.acidity    : num  0.7 0.88 0.76 0.28 0.7 0.66 0.6 0.65 0.58 0.5 ...
##  $ citric.acid         : num  0 0 0.04 0.56 0 0 0.06 0 0.02 0.36 ...
##  $ residual.sugar      : num  1.9 2.6 2.3 1.9 1.9 1.8 1.6 1.2 2 6.1 ...
##  $ chlorides           : num  0.076 0.098 0.092 0.075 0.076 0.075 0.069 0.065 0.073 0.071 ...
##  $ free.sulfur.dioxide : num  11 25 15 17 11 13 15 15 9 17 ...
##  $ total.sulfur.dioxide: num  34 67 54 60 34 40 59 21 18 102 ...
##  $ density             : num  0.998 0.997 0.997 0.998 0.998 ...
##  $ pH                  : num  3.51 3.2 3.26 3.16 3.51 3.51 3.3 3.39 3.36 3.35 ...
##  $ sulphates           : num  0.56 0.68 0.65 0.58 0.56 0.56 0.46 0.47 0.57 0.8 ...
##  $ alcohol             : num  9.4 9.8 9.8 9.8 9.4 9.4 9.4 10 9.5 10.5 ...
```

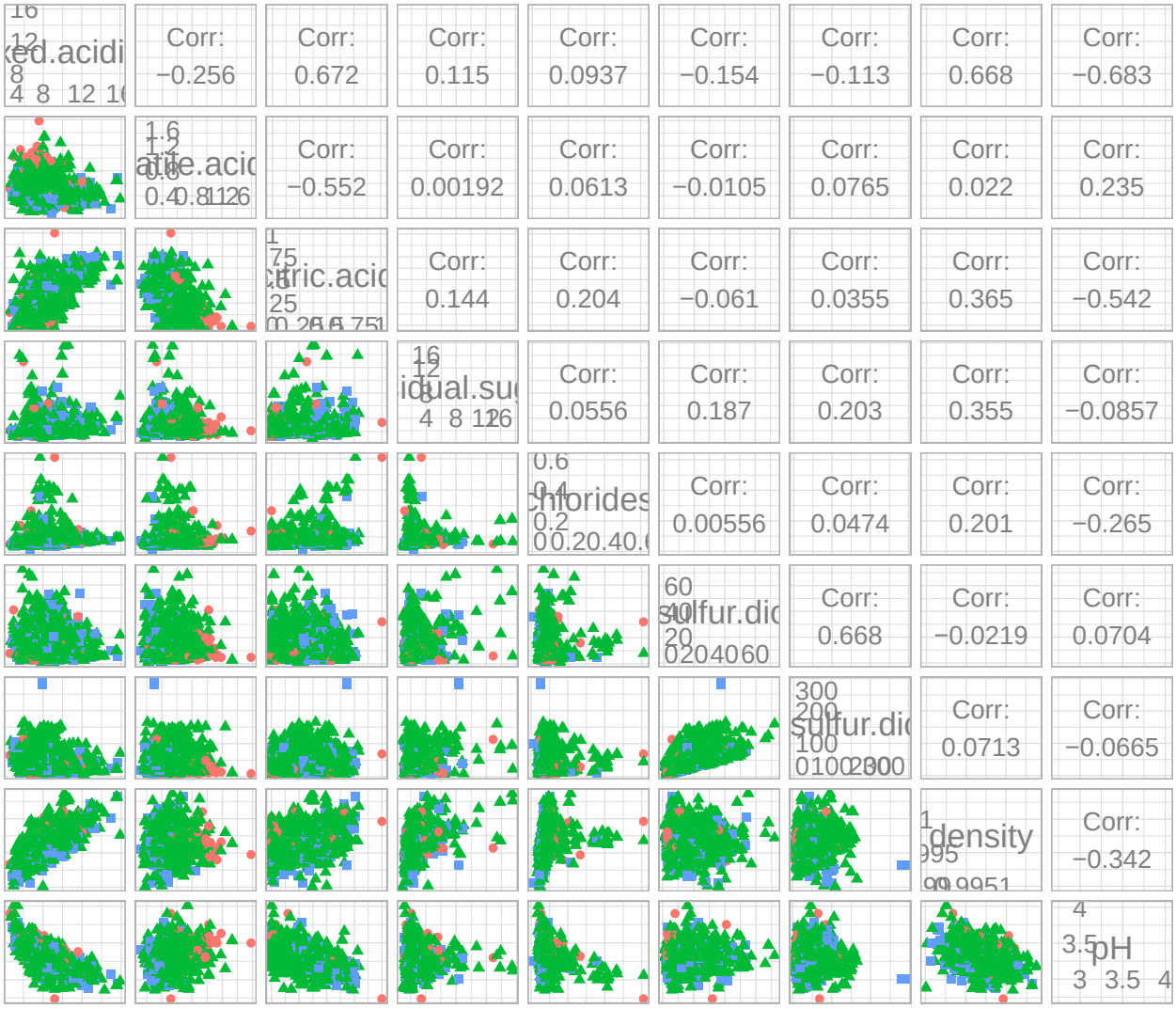
```
##  $ quality             : int  5 5 5 6 5 5 5 7 7 5 ...
##  $ quality.factor      : Factor w/ 6 levels "3","4","5","6",..: 3 3 3 4 3 3 3 5 5 3 ...
##  $ quality.cat         : Factor w/ 3 levels "bad","medium",..: 2 2 2 2 2 2 2 3 3 2 ...
```

```
## [1] medium good   bad
## Levels: bad medium good
```

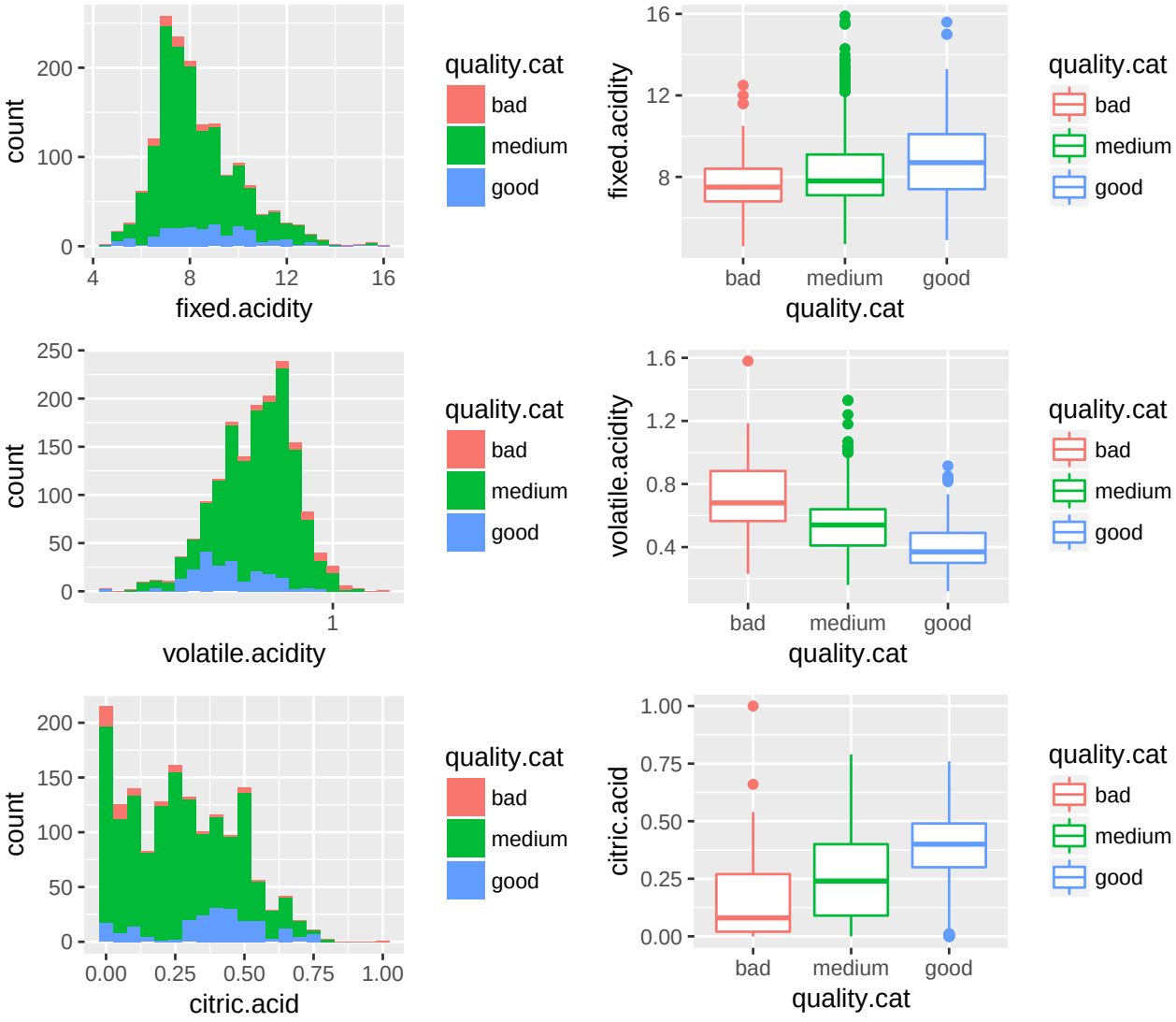
Univariate Plots Section



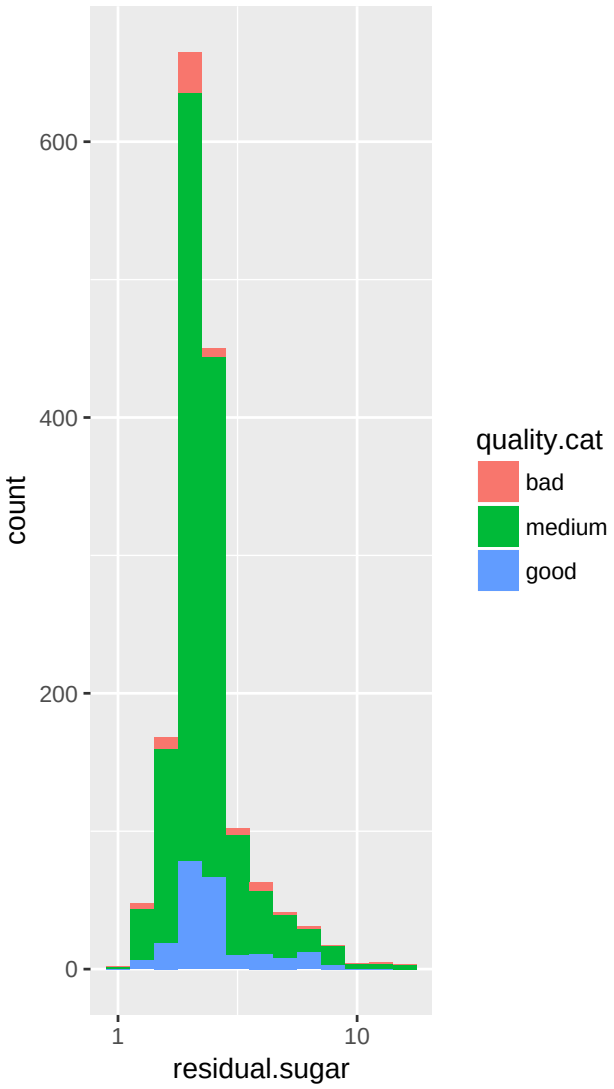

Sulphates
150
100
50
0
count
1
sulphates

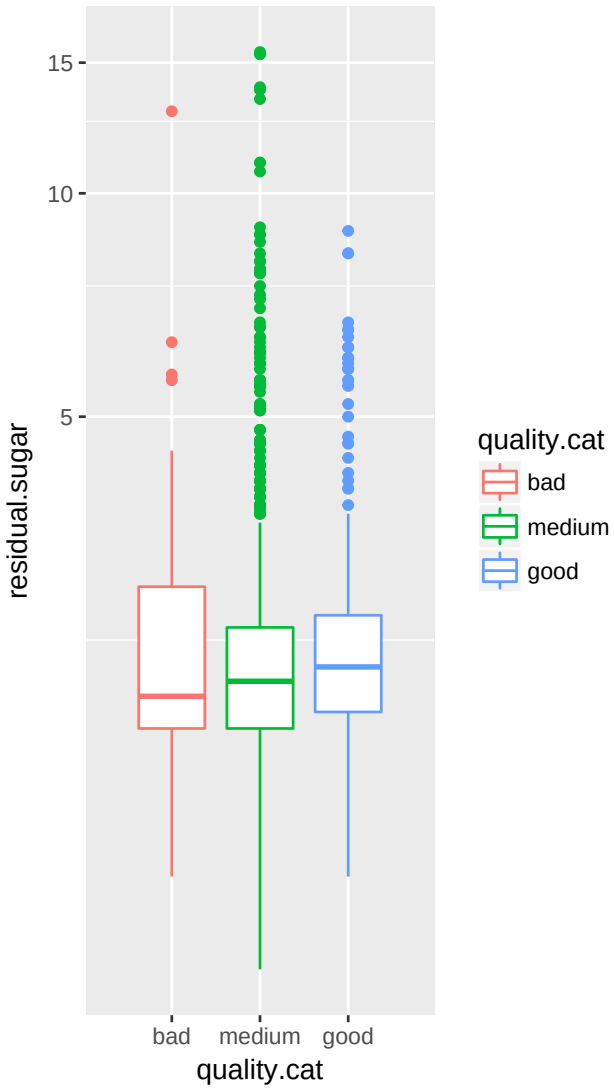

Corr: −0.256
Corr: 0.672
Corr: 0.115
Corr: 0.0937
Corr: −0.154
Corr: −0.113
Corr: 0.668
Corr: −0.683
Corr: −0.552
Corr: 0.00192
Corr: 0.0613
Corr: −0.0105
Corr: 0.0765
Corr: 0.022
Corr: 0.235
Corr: 0.144
Corr: 0.204
Corr: −0.061
Corr: 0.0355
Corr: 0.365
Corr: −0.542
Corr: 0.0556
Corr: 0.187
Corr: 0.203
Corr: 0.355
Corr: −0.0857
Corr: 0.00556
Corr: 0.0474
Corr: 0.201
Corr: −0.265
Corr: 0.668
Corr: −0.0219
Corr: 0.0704
Corr: 0.0713
Corr: −0.0665
Corr: −0.342
density
pH

Fixed and volatile acidity

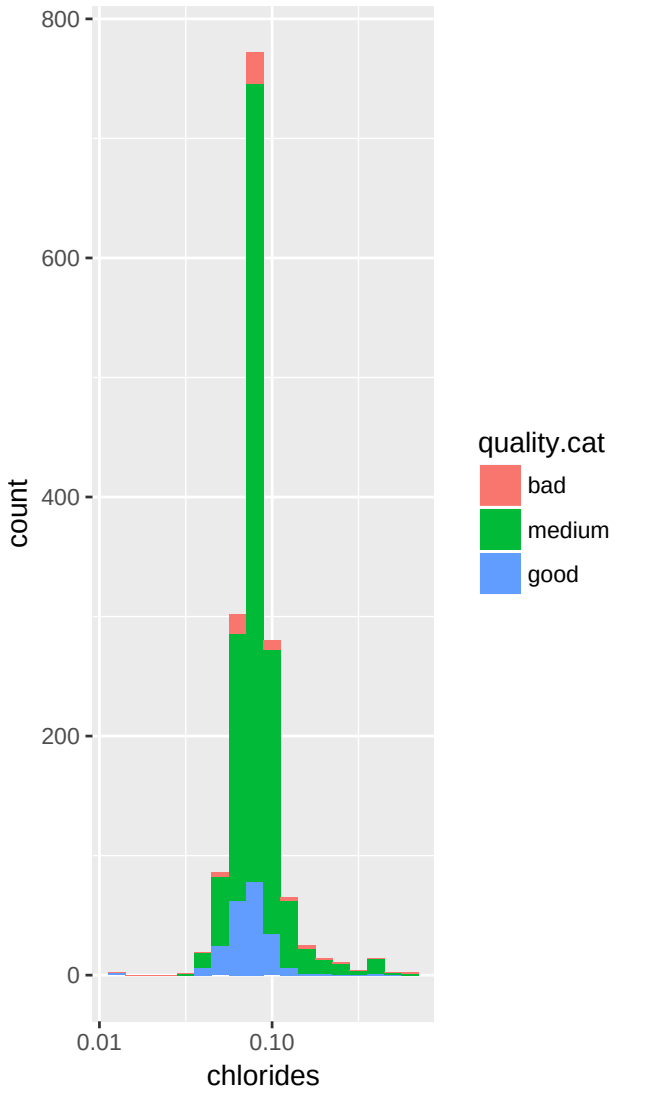


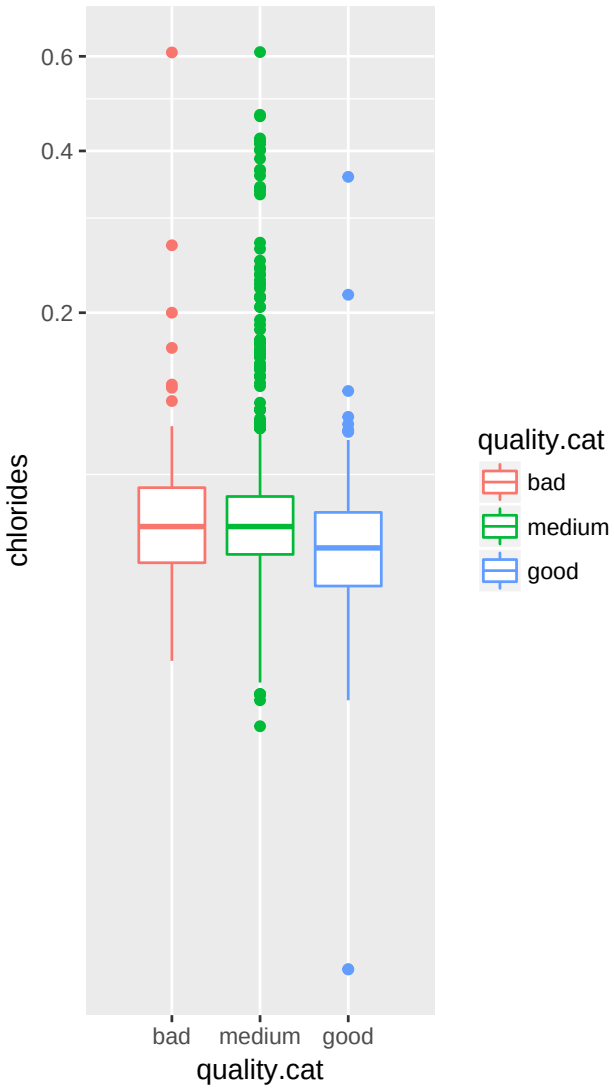

```
## NULL
```


600
400
200
0
count
1
10
residual.sugar
quality.cat
bad
medium
good

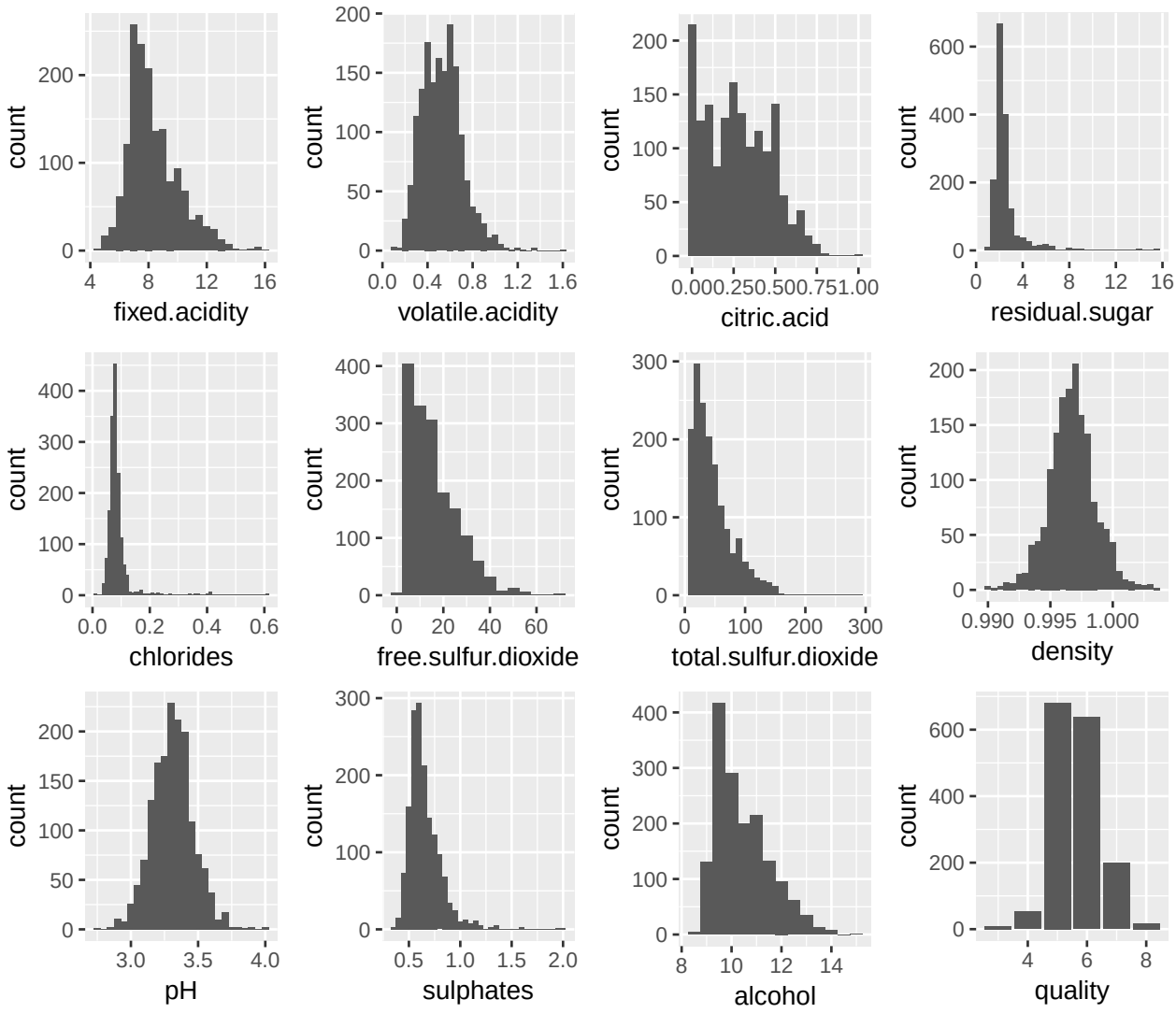

15
10
5
residual.sugar
bad
medium
good
quality.cat
quality.cat
bad
medium
good

```
## NULL
```


800
600
400
200
0
count
0.01
0.10
chlorides
quality.cat
bad
medium
good


0.6
0.4
0.2
chlorides
bad
medium
good
quality.cat
quality.cat
bad
medium
good

```
## NULL
```


count
fixed.acidity
volatile.acidity
citric.acid
residual.sugar
chlorides
free.sulfur.dioxide
total.sulfur.dioxide
density
pH
sulphates
alcohol
quality

```
## NULL
```

Univariate Analysis

What is the structure of your dataset?

What is/are the main feature(s) of interest in your dataset?

What other features in the dataset do you think will help support your investigation into your feature(s) of interest?

Did you create any new variables from existing variables in the dataset?

Of the features you investigated, were there any unusual distributions? Did you perform any operations on the data to tidy, adjust, or change the form of the data? If so, why did you do this?

Bivariate Plots Section

Bivariate Analysis

Talk about some of the relationships you observed in this part of the investigation. How did the feature(s) of interest vary with other features in the dataset?

Did you observe any interesting relationships between the other features (not the main feature(s) of interest)?

What was the strongest relationship you found?

Multivariate Plots Section

Multivariate Analysis

Talk about some of the relationships you observed in this part of the investigation. Were there features that strengthened each other in terms of looking at your feature(s) of interest?

Were there any interesting or surprising interactions between features?

OPTIONAL: Did you create any models with your dataset? Discuss the strengths and limitations of your model.

Final Plots and Summary

Plot One

Description One

Plot Two

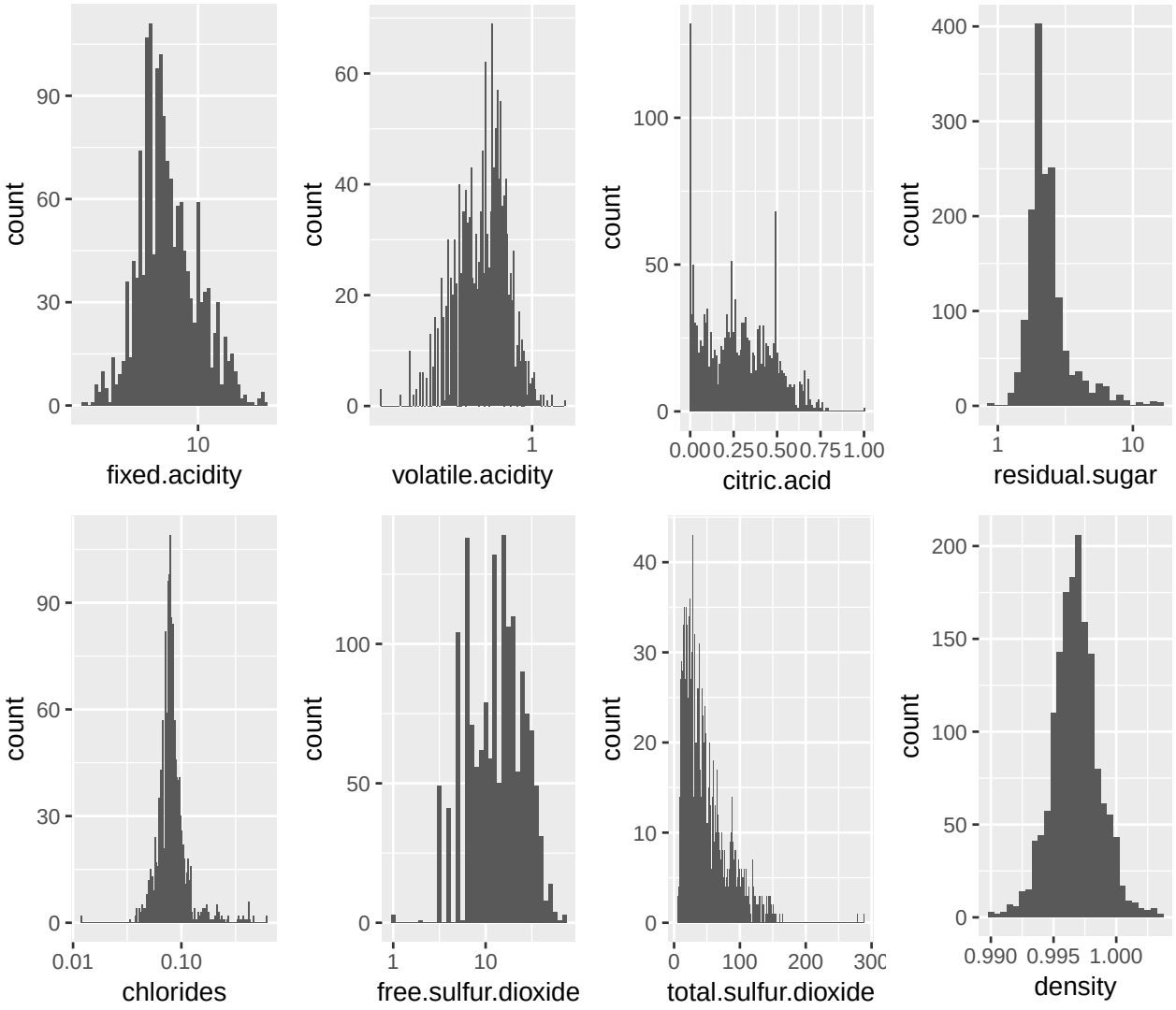
Description Two

Plot Three

Description Three

Reflection

Others




```
## NULL
```