project 2

Group 11

```
#packages library
#|echo: false
library(tidyverse)
library(moderndive)
library(gapminder)
library(sjPlot)
library(stats)
library(jtools)
```

1 Introduction

(include research question)

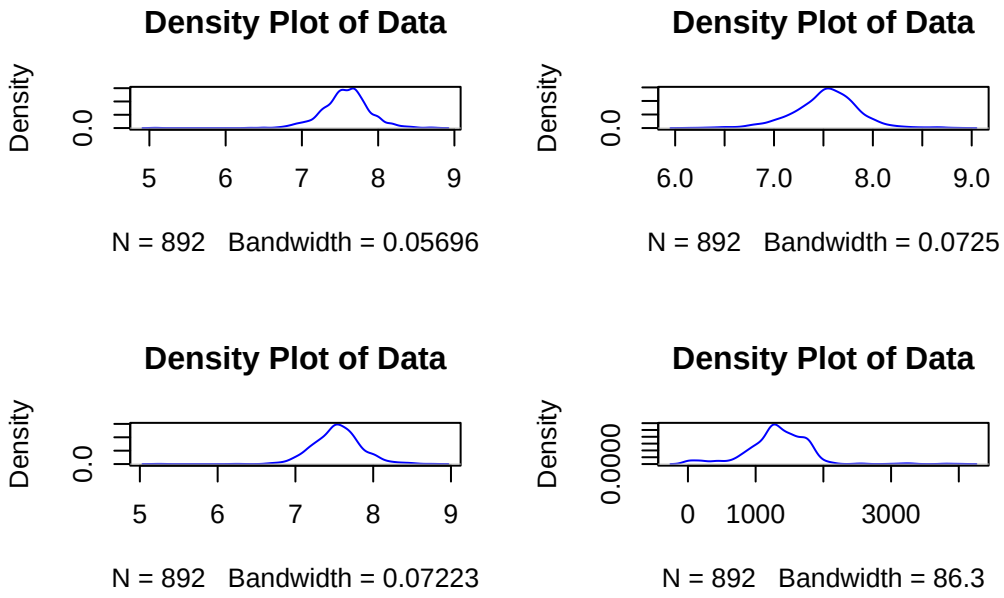
2 Data preparing & Cleaning

```
#data cleaning
data<-read.csv("dataset11.csv")
data<-na.omit(data)
data$Qualityclass_dummy<-ifelse(data$Qualityclass=="Good",1,0) #for "Good"=1
 ↪  "Poor"=0
data$Qualityclass <- as.factor(data$Qualityclass)
data$harvested <- as.factor(data$harvested)
data$category_two_defects<-as.factor(data$category_two_defects)
#Outlier in altitude_mean_meters
sum(data$altitude_mean_meters>8848)+sum(data$altitude_mean_meters<0)
```

```
[1] 2
```

```r
data=data%>%
  filter(data$altitude_mean_meters<8848 & data$altitude_mean_meters>0)

#standarized
par(mfrow=c(2,2))
plot(density(data$aroma), col = "blue", main = "Density Plot of Data")
plot(density(data$flavor), col = "blue", main = "Density Plot of Data")
plot(density(data$acidity), col = "blue", main = "Density Plot of Data")
plot(density(data$altitude_mean_meters), col = "blue", main = "Density Plot
 ↪  of Data")
```

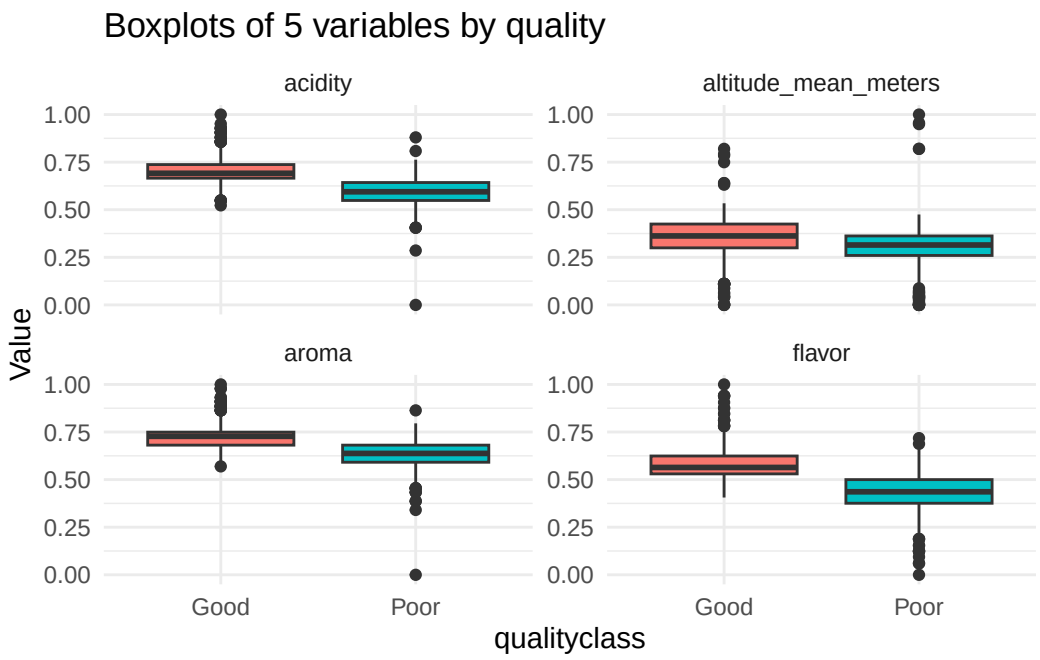



```r
#Min-Max standarized
min_max_norm=function(x){
  return((x-min(x))/(max(x)-min(x)))
}
data[,2:4]=lapply(data[,2:4],min_max_norm)
data$altitude_mean_meters=min_max_norm(data$altitude_mean_meters)
```

in this dataset aroma,flavor,acidity and altitude_mean_meters are continuous variables, category_two_defects, Qualityclass and harvested are categorical variables. #Exploratory Data Analysis # Illuminating visualizations of the data

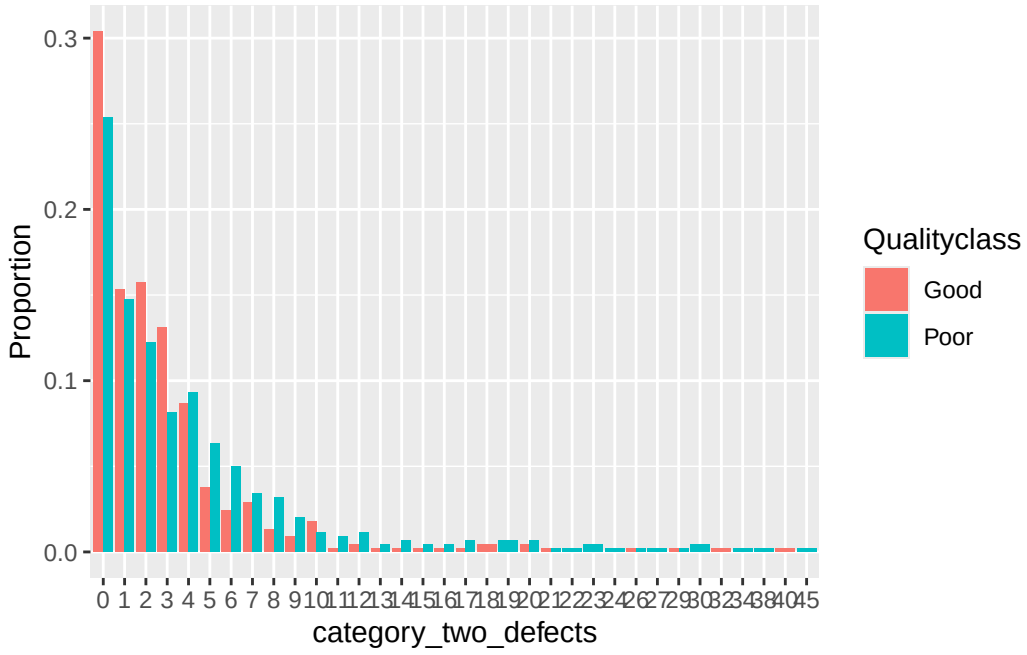
```r
library(tidyr)
#change formula
data_long <- data %>%
  pivot_longer(cols = c(aroma,flavor,acidity,altitude_mean_meters),
               names_to = "Variable",
               values_to = "Value")
library(ggplot2)

#boxplot
#continuous
ggplot(data = data_long, aes(x = Qualityclass, y = Value, fill =
 ↪  Qualityclass)) +
  geom_boxplot() +
  facet_wrap(~Variable, scales = "free_y") +
  theme_minimal() +
  labs(title = "Boxplots of 5 variables by quality",
       x = "qualityclass",
       y = "Value") +
  theme(legend.position = "none")
```

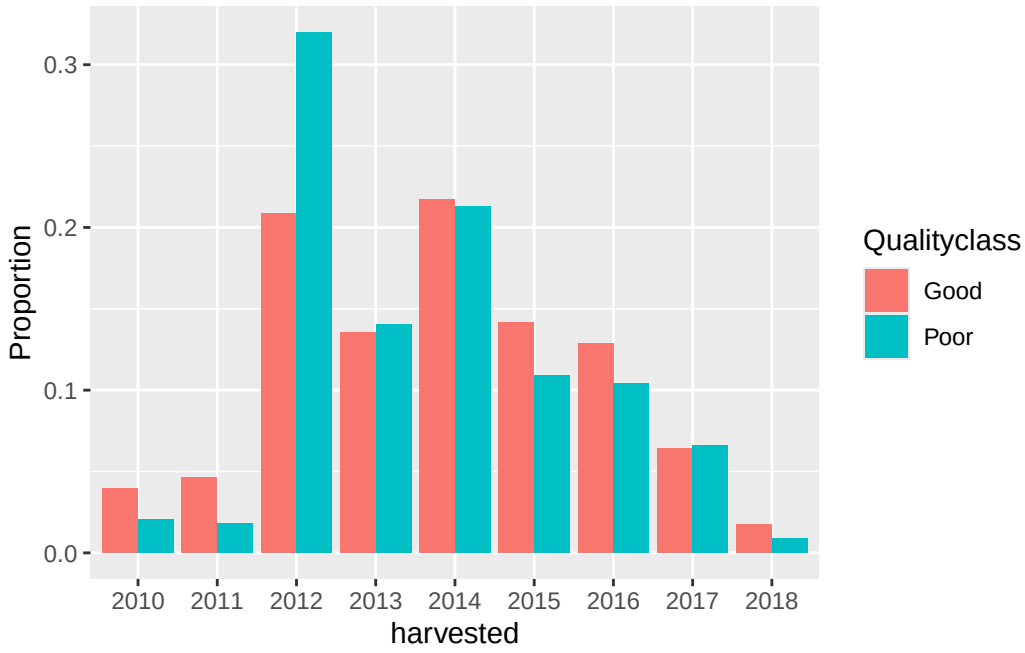



```r
#categorical
#category_two_defects
```

```
ggplot(data, aes(x=category_two_defects ,  y = ..prop.., group=Qualityclass,
 ↪  fill=Qualityclass)) +
    geom_bar(position="dodge", stat="count") +
    labs(y = "Proportion")
```




```
#harvested
ggplot(data, aes(x=harvested ,  y = ..prop.., group=Qualityclass,
 ↪  fill=Qualityclass)) +
    geom_bar(position="dodge", stat="count") +
    labs(y = "Proportion")
```



#Numerical Summaries

```
#summary
summary(data)
```

```
 country_of_origin      aroma            flavor          acidity
 Length:892         Min.   :0.0000   Min.   :0.0000   Min.   :0.0000
 Class :character   1st Qu.:0.6376   1st Qu.:0.4361   1st Qu.:0.5943
 Mode  :character   Median :0.6812   Median :0.5301   Median :0.6429
                    Mean   :0.6797   Mean   :0.5098   Mean   :0.6535
                    3rd Qu.:0.7275   3rd Qu.:0.5940   3rd Qu.:0.7143
                    Max.   :1.0000   Max.   :1.0000   Max.   :1.0000

 category_two_defects altitude_mean_meters   harvested   Qualityclass
 0      :249          Min.   :0.0000       2012   :235   Good:451
 1      :134          1st Qu.:0.2747       2014   :192   Poor:441
 2      :125          Median :0.3274       2013   :123
 3      : 95          Mean   :0.3302       2015   :112
 4      : 80          3rd Qu.:0.3997       2016   :104
 5      : 45          Max.   :1.0000       2017   : 58
 (Other):164                               (Other): 68
 Qualityclass_dummy
 Min.   :0.0000
```

```
 1st Qu.:0.0000
 Median :1.0000
 Mean   :0.5056
 3rd Qu.:1.0000
 Max.   :1.0000
```

```
#####need to add a table for summary use gt()
```

#Formal Data Analysis #Model fitted original model

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 \cdot \text{aroma} + \beta_2 \cdot \text{flavor} + \beta_3 \cdot \text{acidity} + \beta_4 \cdot \text{defects} + \beta_5 \cdot \text{meters} + \beta_6 \cdot \text{harvested}$$

(each variable need to be explain)

```
#model fitted for original
model <- glm(Qualityclass_dummy ~
 ↪  aroma+flavor+acidity+category_two_defects+altitude_mean_meters+harvested,
 ↪  data = data, family = binomial(link = "logit"))
summary(model)
```

```
Call:
glm(formula = Qualityclass_dummy ~ aroma + flavor + acidity +
    category_two_defects + altitude_mean_meters + harvested,
    family = binomial(link = "logit"), data = data)

Coefficients:
                          Estimate Std. Error z value Pr(>|z|)
(Intercept)             -3.247e+01  2.778e+00 -11.691  < 2e-16 ***
aroma                    1.810e+01  2.735e+00   6.618 3.64e-11 ***
flavor                   2.025e+01  2.494e+00   8.120 4.65e-16 ***
acidity                  1.376e+01  2.533e+00   5.432 5.56e-08 ***
category_two_defects1   -1.527e-01  3.649e-01  -0.418  0.67572
category_two_defects2   -8.175e-02  3.825e-01  -0.214  0.83075
category_two_defects3    4.523e-01  4.053e-01   1.116  0.26444
category_two_defects4    4.698e-01  4.452e-01   1.055  0.29131
category_two_defects5    6.914e-02  5.687e-01   0.122  0.90323
category_two_defects6   -5.740e-01  6.216e-01  -0.924  0.35574
category_two_defects7   -1.043e+00  8.714e-01  -1.197  0.23123
category_two_defects8    2.646e-03  8.002e-01   0.003  0.99736
```

```
category_two_defects9  -7.248e-01  8.452e-01  -0.858  0.39117
category_two_defects10  1.499e+00  9.563e-01   1.568  0.11694
category_two_defects11  5.490e-01  1.324e+00   0.415  0.67838
category_two_defects12  3.154e+00  1.915e+00   1.647  0.09964 .
category_two_defects13 -8.069e-01  2.143e+00  -0.377  0.70652
category_two_defects14  6.154e-01  1.684e+00   0.365  0.71484
category_two_defects15 -1.016e+00  1.489e+00  -0.682  0.49504
category_two_defects16  2.320e+00  6.870e+00   0.338  0.73556
category_two_defects17  1.791e+00  1.357e+00   1.320  0.18675
category_two_defects18  1.306e+01  1.536e+03   0.009  0.99321
category_two_defects19 -1.128e+01  1.284e+03  -0.009  0.99299
category_two_defects20  3.653e+00  3.341e+00   1.093  0.27422
category_two_defects21  2.356e+00  6.836e+00   0.345  0.73034
category_two_defects22 -1.048e+01  2.400e+03  -0.004  0.99652
category_two_defects23 -1.407e+01  1.312e+03  -0.011  0.99144
category_two_defects24 -5.145e+00  2.400e+03  -0.002  0.99829
category_two_defects26 -3.092e-01  3.468e+00  -0.089  0.92895
category_two_defects27 -3.219e+00  2.400e+03  -0.001  0.99893
category_two_defects29  1.180e+00  1.288e+01   0.092  0.92701
category_two_defects30 -7.186e+00  1.686e+03  -0.004  0.99660
category_two_defects32  1.495e+01  2.400e+03   0.006  0.99503
category_two_defects34 -7.445e+00  2.400e+03  -0.003  0.99752
category_two_defects38 -6.458e+00  2.400e+03  -0.003  0.99785
category_two_defects40  1.388e+01  2.400e+03   0.006  0.99538
category_two_defects45 -9.804e+00  2.400e+03  -0.004  0.99674
altitude_mean_meters    2.696e+00  1.033e+00   2.611  0.00902 **
harvested2011          -1.790e-01  1.121e+00  -0.160  0.87312
harvested2012          -8.377e-01  9.387e-01  -0.892  0.37220
harvested2013          -2.441e-01  9.480e-01  -0.258  0.79678
harvested2014           8.853e-02  9.450e-01   0.094  0.92535
harvested2015          -4.161e-01  9.534e-01  -0.436  0.66252
harvested2016           3.715e-01  9.871e-01   0.376  0.70668
harvested2017           1.888e-01  9.839e-01   0.192  0.84779
harvested2018           1.378e+00  1.284e+00   1.073  0.28331
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1236.46  on 891  degrees of freedom
Residual deviance:  495.99  on 846  degrees of freedom
AIC: 587.99
```

```
Number of Fisher Scoring iterations: 15
```

```
summ(model)
```

```
MODEL INFO:
Observations: 892
Dependent Variable: Qualityclass_dummy
Type: Generalized linear model
  Family: binomial
  Link function: logit

MODEL FIT:
²(45) = 740.47, p = 0.00
Pseudo-R² (Cragg-Uhler) = 0.75
Pseudo-R² (McFadden) = 0.60
AIC = 587.99, BIC = 808.49

Standard errors:MLE
------------------------------------------------------------------
                               Est.      S.E.   z val.      p
---------------------------- -------- --------- -------- ------
(Intercept)                  -32.47      2.78   -11.69   0.00
aroma                         18.10      2.73     6.62   0.00
flavor                        20.25      2.49     8.12   0.00
acidity                       13.76      2.53     5.43   0.00
category_two_defects1         -0.15      0.36    -0.42   0.68
category_two_defects2         -0.08      0.38    -0.21   0.83
category_two_defects3          0.45      0.41     1.12   0.26
category_two_defects4          0.47      0.45     1.06   0.29
category_two_defects5          0.07      0.57     0.12   0.90
category_two_defects6         -0.57      0.62    -0.92   0.36
category_two_defects7         -1.04      0.87    -1.20   0.23
category_two_defects8          0.00      0.80     0.00   1.00
category_two_defects9         -0.72      0.85    -0.86   0.39
category_two_defects10         1.50      0.96     1.57   0.12
category_two_defects11         0.55      1.32     0.41   0.68
category_two_defects12         3.15      1.92     1.65   0.10
category_two_defects13        -0.81      2.14    -0.38   0.71
category_two_defects14         0.62      1.68     0.37   0.71
category_two_defects15        -1.02      1.49    -0.68   0.50
category_two_defects16         2.32      6.87     0.34   0.74
category_two_defects17         1.79      1.36     1.32   0.19
```

```
category_two_defects18          13.06   1535.80     0.01   0.99
category_two_defects19         -11.28   1284.07    -0.01   0.99
category_two_defects20           3.65      3.34     1.09   0.27
category_two_defects21           2.36      6.84     0.34   0.73
category_two_defects22         -10.48   2399.54    -0.00   1.00
category_two_defects23         -14.07   1312.39    -0.01   0.99
category_two_defects24          -5.14   2399.54    -0.00   1.00
category_two_defects26          -0.31      3.47    -0.09   0.93
category_two_defects27          -3.22   2399.55    -0.00   1.00
category_two_defects29           1.18     12.88     0.09   0.93
category_two_defects30          -7.19   1686.23    -0.00   1.00
category_two_defects32          14.95   2399.54     0.01   1.00
category_two_defects34          -7.44   2399.54    -0.00   1.00
category_two_defects38          -6.46   2399.54    -0.00   1.00
category_two_defects40          13.88   2399.54     0.01   1.00
category_two_defects45          -9.80   2399.54    -0.00   1.00
altitude_mean_meters             2.70      1.03     2.61   0.01
harvested2011                   -0.18      1.12    -0.16   0.87
harvested2012                   -0.84      0.94    -0.89   0.37
harvested2013                   -0.24      0.95    -0.26   0.80
harvested2014                    0.09      0.94     0.09   0.93
harvested2015                   -0.42      0.95    -0.44   0.66
harvested2016                    0.37      0.99     0.38   0.71
harvested2017                    0.19      0.98     0.19   0.85
harvested2018                    1.38      1.28     1.07   0.28
-----------------------------------------------------------------
```

find p-value of category_two_defects and altitude_mean_meters and harvested are higher than 0.05

```
#del category_two_defects variale
model_ctd <- glm(Qualityclass_dummy ~
 ↪  aroma+flavor+acidity+altitude_mean_meters+harvested, data = data, family
 ↪  = binomial(link = "logit"))
summary(model_ctd)
```

```
Call:
glm(formula = Qualityclass_dummy ~ aroma + flavor + acidity +
    altitude_mean_meters + harvested, family = binomial(link = "logit"),
    data = data)
```

```
Coefficients:
                      Estimate Std. Error z value Pr(>|z|)
(Intercept)          -31.90576    2.70450 -11.797  < 2e-16 ***
aroma                 18.43367    2.69614   6.837 8.08e-12 ***
flavor                19.27574    2.37645   8.111 5.01e-16 ***
acidity               13.40118    2.46435   5.438 5.39e-08 ***
altitude_mean_meters   2.43429    0.97737   2.491   0.0128 *
harvested2011         -0.05273    1.08890  -0.048   0.9614
harvested2012         -0.58227    0.90349  -0.644   0.5193
harvested2013         -0.19932    0.92158  -0.216   0.8288
harvested2014          0.10122    0.91885   0.110   0.9123
harvested2015         -0.40804    0.92349  -0.442   0.6586
harvested2016          0.44626    0.95346   0.468   0.6398
harvested2017          0.22480    0.95680   0.235   0.8142
harvested2018          1.66547    1.25028   1.332   0.1828
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1236.46  on 891  degrees of freedom
Residual deviance:  513.89  on 879  degrees of freedom
AIC: 539.89

Number of Fisher Scoring iterations: 7
```

```
summ(model_ctd)
```

```
MODEL INFO:
Observations: 892
Dependent Variable: Qualityclass_dummy
Type: Generalized linear model
  Family: binomial
  Link function: logit

MODEL FIT:
 ²(12) = 722.58, p = 0.00
Pseudo-R² (Cragg-Uhler) = 0.74
Pseudo-R² (McFadden) = 0.58
AIC = 539.89, BIC = 602.20

Standard errors:MLE
```

	Est.	S.E.	z val.	p
(Intercept)	-31.91	2.70	-11.80	0.00
aroma	18.43	2.70	6.84	0.00
flavor	19.28	2.38	8.11	0.00
acidity	13.40	2.46	5.44	0.00
altitude_mean_meters	2.43	0.98	2.49	0.01
harvested2011	-0.05	1.09	-0.05	0.96
harvested2012	-0.58	0.90	-0.64	0.52
harvested2013	-0.20	0.92	-0.22	0.83
harvested2014	0.10	0.92	0.11	0.91
harvested2015	-0.41	0.92	-0.44	0.66
harvested2016	0.45	0.95	0.47	0.64
harvested2017	0.22	0.96	0.23	0.81
harvested2018	1.67	1.25	1.33	0.18

AIC decreased

```
#del altitude_mean_meters variale
model_0 <- glm(Qualityclass_dummy ~
 ↪  aroma+flavor+acidity+altitude_mean_meters, data = data, family =
 ↪  binomial(link = "logit"))
summary(model_0)
```

```
Call:
glm(formula = Qualityclass_dummy ~ aroma + flavor + acidity +
    altitude_mean_meters, family = binomial(link = "logit"),
    data = data)

Coefficients:
                    Estimate Std. Error z value Pr(>|z|)
(Intercept)          -31.0176     2.4204 -12.815  < 2e-16 ***
aroma                 17.0856     2.5335   6.744 1.54e-11 ***
flavor                18.7270     2.2919   8.171 3.06e-16 ***
acidity               14.0064     2.4087   5.815 6.06e-09 ***
altitude_mean_meters   1.8539     0.9269   2.000   0.0455 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1236.46  on 891  degrees of freedom
Residual deviance:  528.38  on 887  degrees of freedom
AIC: 538.38

Number of Fisher Scoring iterations: 7
```

```r
summ(model_0)
```

```
MODEL INFO:
Observations: 892
Dependent Variable: Qualityclass_dummy
Type: Generalized linear model
  Family: binomial
  Link function: logit

MODEL FIT:
²(4) = 708.08, p = 0.00
Pseudo-R² (Cragg-Uhler) = 0.73
Pseudo-R² (McFadden) = 0.57
AIC = 538.38, BIC = 562.35

Standard errors:MLE
------------------------------------------------------------
                             Est.   S.E.   z val.      p
------------------------- -------- ------ -------- ------
(Intercept)                -31.02   2.42   -12.82   0.00
aroma                       17.09   2.53     6.74   0.00
flavor                      18.73   2.29     8.17   0.00
acidity                     14.01   2.41     5.81   0.00
altitude_mean_meters         1.85   0.93     2.00   0.05
------------------------------------------------------------
```

AIC decreased

```r
#del harvested variable
model1 <- glm(Qualityclass_dummy ~ aroma+flavor+acidity, data = data, family
 ↪  = binomial(link = "logit"))
summary(model1)
```

```
Call:
glm(formula = Qualityclass_dummy ~ aroma + flavor + acidity,
    family = binomial(link = "logit"), data = data)

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept)  -30.711      2.398 -12.806  < 2e-16 ***
aroma         17.647      2.531   6.972 3.14e-12 ***
flavor        18.328      2.256   8.123 4.55e-16 ***
acidity       14.202      2.388   5.948 2.72e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for binomial family taken to be 1)

    Null deviance: 1236.46  on 891  degrees of freedom
Residual deviance:  532.32  on 888  degrees of freedom
AIC: 540.32

Number of Fisher Scoring iterations: 7
```

```
summ(model1)
```

```
MODEL INFO:
Observations: 892
Dependent Variable: Qualityclass_dummy
Type: Generalized linear model
  Family: binomial
  Link function: logit

MODEL FIT:
²(3) = 704.14, p = 0.00
Pseudo-R² (Cragg-Uhler) = 0.73
Pseudo-R² (McFadden) = 0.57
AIC = 540.32, BIC = 559.50

Standard errors:MLE
------------------------------------------------
                      Est.   S.E.   z val.      p
----------------- -------- ------ -------- ------
(Intercept)         -30.71   2.40   -12.81   0.00
```

```
aroma                17.65   2.53     6.97   0.00
flavor               18.33   2.26     8.12   0.00
acidity              14.20   2.39     5.95   0.00
-------------------------------------------------
```

```
####### need a table for summ
```

aroma flavor acidity significant this three varibales will be saved and AIC decreased to min. optimization model (final model)

$$\ln\left(\frac{p}{1-p}\right) = \alpha + \beta_1 \cdot \text{aroma} + \beta_2 \cdot \text{flavor} + \beta_3 \cdot \text{acidity} + \beta_4 \cdot \text{harvested}$$

########model1

```
levels(data$Qualityclass) #base on "good"
```

```
[1] "Good" "Poor"
```

```
#for original model
mod1coefs <- round(coef(model), 2)
library(knitr)
confint(model) %>%
  kable()
```

	2.5 %	97.5 %
(Intercept)	-38.1854338	-27.2711993
aroma	12.8982348	23.6359497
flavor	15.5380009	25.3372050
acidity	8.8898953	18.8349218
category_two_defects1	-0.8694362	0.5639448
category_two_defects2	-0.8312538	0.6710736
category_two_defects3	-0.3364056	1.2559917
category_two_defects4	-0.3960145	1.3537865
category_two_defects5	-1.0634771	1.1720588
category_two_defects6	-1.8281503	0.6249114
category_two_defects7	-2.6861160	0.7122507
category_two_defects8	-1.5780035	1.5897154
category_two_defects9	-2.5195019	0.8658394

	2.5 %	97.5 %
category_two_defects10	-0.3228813	3.4391795
category_two_defects11	-2.6381797	3.0713435
category_two_defects12	-0.5229555	6.4090166
category_two_defects13	-4.8848764	2.6926501
category_two_defects14	-2.9873634	3.7271785
category_two_defects15	-4.3434076	2.3157523
category_two_defects16	-4.0840962	8.8195512
category_two_defects17	-1.4210880	4.4165181
category_two_defects18	-227.4871279	NA
category_two_defects19	NA	154.8072880
category_two_defects20	-0.8322924	8.4077228
category_two_defects21	-4.0348970	8.8394569
category_two_defects22	NA	477.9443385
category_two_defects23	NA	191.2042648
category_two_defects24	NA	482.2000562
category_two_defects26	-5.3327893	4.8223827
category_two_defects27	NA	482.8918280
category_two_defects29	-7.1169297	9.5218187
category_two_defects30	NA	254.7657979
category_two_defects32	-473.8564449	NA
category_two_defects34	NA	480.4074862
category_two_defects38	NA	480.9510968
category_two_defects40	-474.8441203	NA
category_two_defects45	NA	478.4965206
altitude_mean_meters	0.6578071	4.7241363
harvested2011	-2.4669518	1.9719426
harvested2012	-2.8197380	0.9105193
harvested2013	-2.2390861	1.5293781
harvested2014	-1.9001820	1.8585887
harvested2015	-2.4223610	1.3662790
harvested2016	-1.6916394	2.2290361
harvested2017	-1.8647301	2.0434207
harvested2018	-1.1584341	3.9710165

```
#for optimization model
mod1coefs1 <- round(coef(model1), 2)
library(knitr)
confint(model1) %>%
  kable()
```

	2.5 %	97.5 %
(Intercept)	-35.658890	-26.24300
aroma	12.821469	22.75872
flavor	14.056373	22.91486
acidity	9.609975	18.98674

log-odds

```
mod.coef.logodds<-model1 %>%
                  summary() %>%
                  coef()
```

```
plot_model(model1, show.values = TRUE, transform = NULL,
           title = "Log-Odds (quality-good)", show.p = FALSE)
```

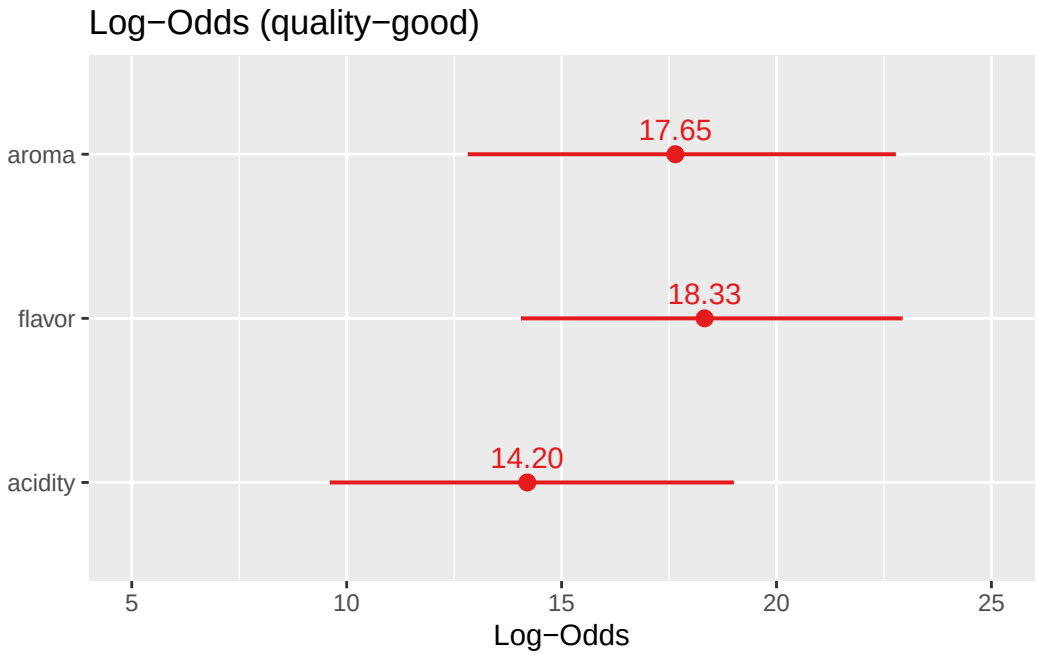


Figure 1: the log-odds of explanatory variables for quality good

```
data<- data%>%
      mutate(logodds.good = predict(model1))
```

odds

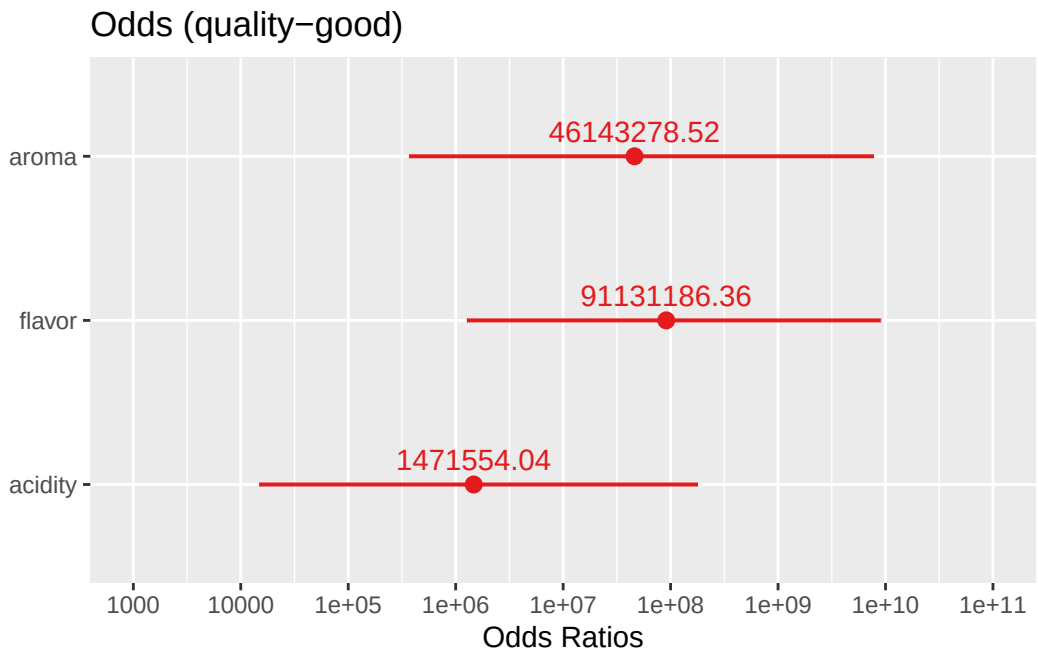
```
model1 %>%
 coef() %>%
  exp()
```

```
 (Intercept)        aroma       flavor      acidity
4.593956e-14 4.614328e+07 9.113119e+07 1.471554e+06
```

```
#check value
exp(coef(model1))
```

```
 (Intercept)        aroma       flavor      acidity
4.593956e-14 4.614328e+07 9.113119e+07 1.471554e+06
```

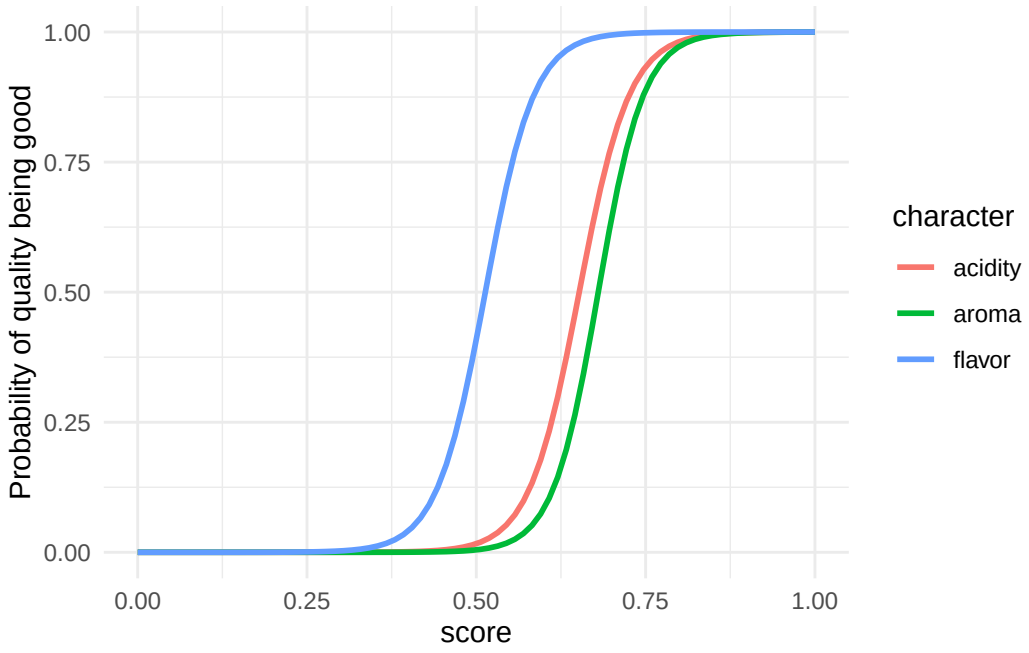
```
#odd ratrio for quality good
plot_model(model1, show.values = TRUE,
           title = "Odds (quality-good)", show.p = FALSE)
```



```r
data<- data%>%
      mutate(odds.good = exp(logodds.good))
data<- data%>%
      mutate(prob.good = fitted(model1))
```

probability continuous

```r
#aroma/acidity/flavor prob
library(ggplot2)
library(tidyr)
library(dplyr)
data_long1 <- data %>%
 pivot_longer(cols = c(aroma, flavor,acidity), names_to = "Type", values_to =
  ↪  "Value")
# plot
ggplot(data = data_long1, aes(x =Value, y =prob.good, color = Type)) +
 geom_smooth(method = "glm", method.args = list(family = "binomial"), se =
  ↪  FALSE) +
 labs(x = "score", y = "Probability of quality being good", color =
  ↪  "character") +
 theme_minimal()
```



```
library(grid)
library(gridExtra)
#plot

p1=plot_model(model1, type = "pred",terms = "aroma" ,title = "Aroma",
              axis.title = c("aroma", "Prob. of quality being good"))
p2=plot_model(model1, type = "pred", terms="flavor",title = "Flavor",
              axis.title = c("flavor", "Prob. of quality being good"))
p3=plot_model(model1, type = "pred",terms = "acidity", title = "Acidity",
              axis.title = c("acidity", "Prob. of quality being good"))
#merge
grid.arrange(p1,p2,p3,nrow=1)
```