9.5/10

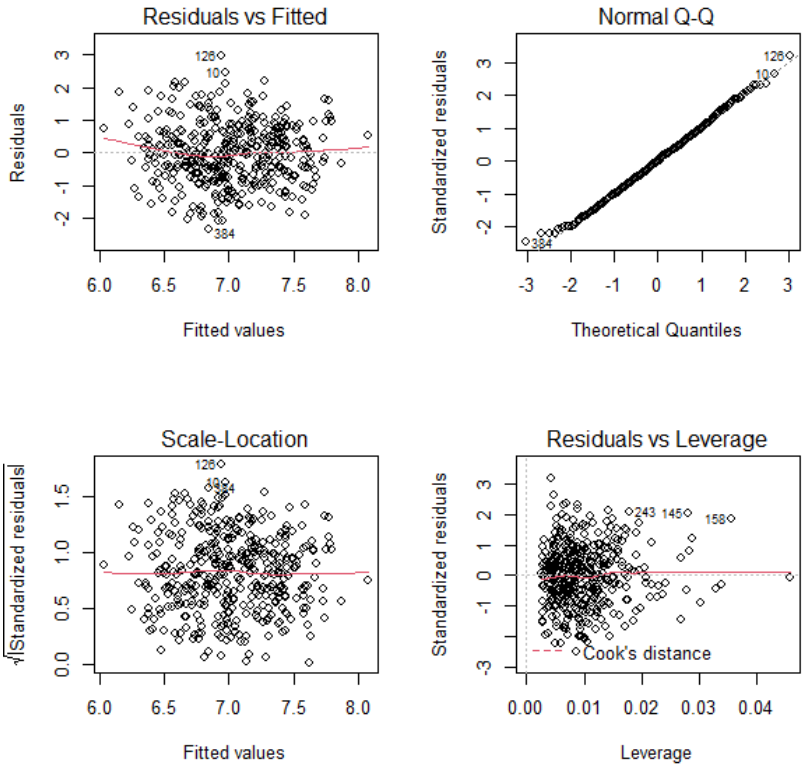
a) The coefficients estimated through lm command are:

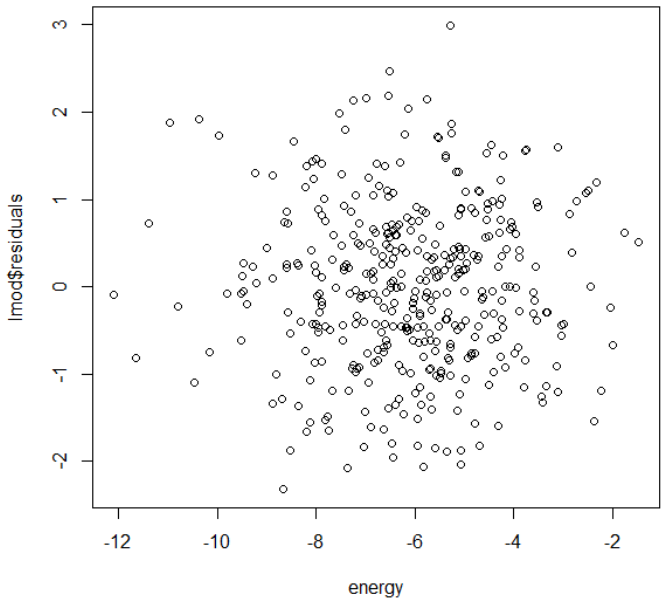
```
(Intercept)    loudness      energy       tempo
 9.18676786  0.09930218  0.07258300 -0.00888967
```

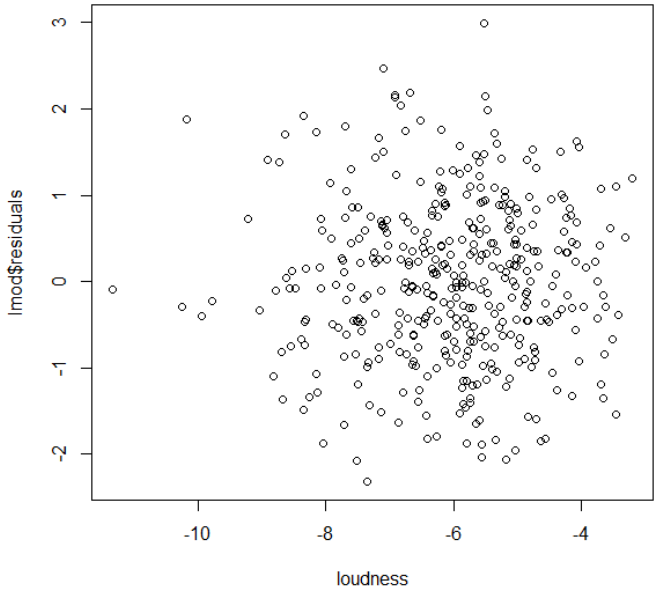
Sigma is : 0.935.

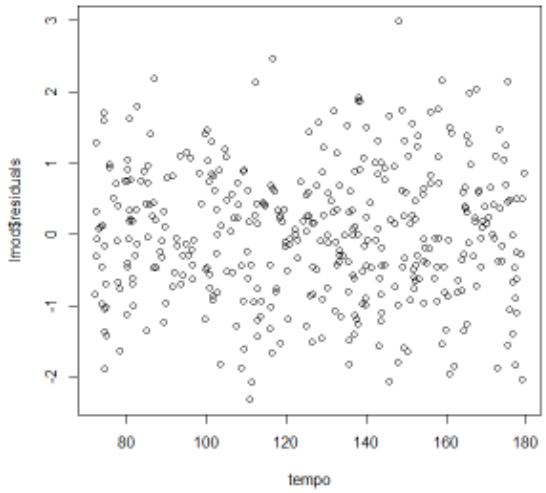
b) The model assumptions are normality, independence from the regressors, and heteroskedasticity of the error terms (eps) -0.5

We check these by plotting the residuals against the fitted values and against the regressors and by using Shapiro.test function. They seem all ok.




lmod$residuals
3
2
1
0
-1
-2
-12
-10
-8
-6
-4
-2
energy


lmod$residuals
3
2
1
0
-1
-2
-10
-8
-6
-4
loudness



c) By using the linearHypothesis command we cannot both discard energy and loudness since the pvalue is under 5%.

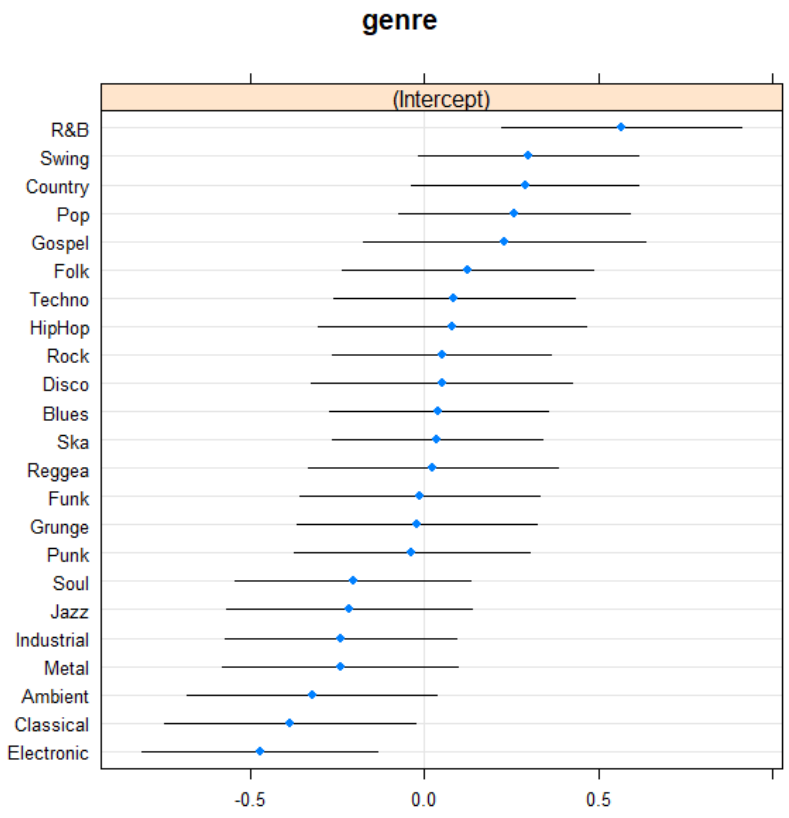
d) At level 5% we can say that energy is not statistically relevant and has the highest pvalue between the four coefficients.
Therefore we discard it and through the anova command we can confirm that the reduced model does not loose effectiveness at level 5%
Also loudness becomes more significantly different from 0.

```
(Intercept)     loudness        tempo
 9.182772292  0.170266165 -0.008962479
```

Sigma : 0.937

e) The PVRE is 0.1034443

f) R&B is the genre with highest danceability.


genre
(Intercept)
R&B
Swing
Country
Pop
Gospel
Folk
Techno
HipHop
Rock
Disco
Blues
Ska
Reggea
Funk
Grunge
Punk
Soul
Jazz
Industrial
Metal
Ambient
Classical
Electronic
-0.5
0.0
0.5