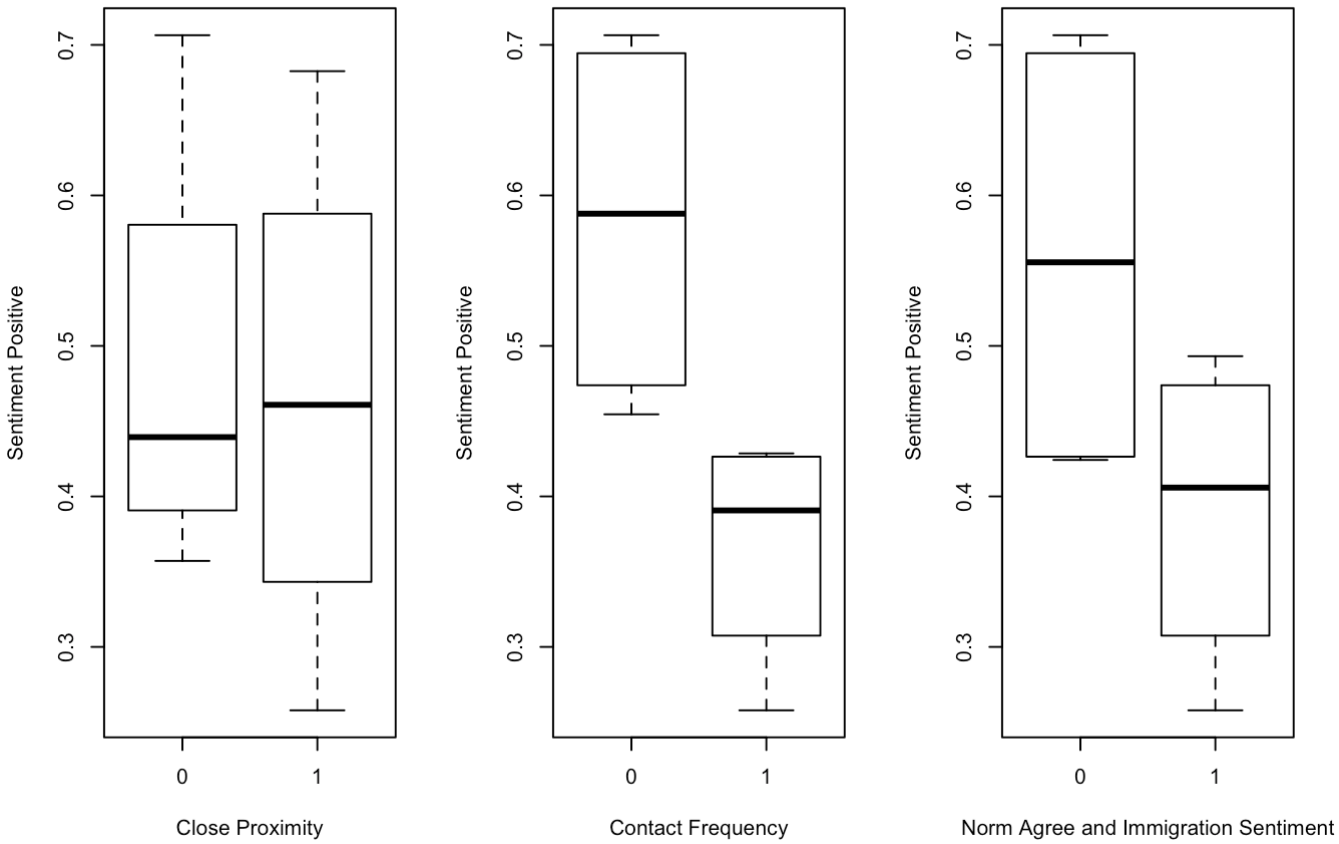
Housing Dataset Doubts

On comparing sentiment percentage with respect to different variable is see strange results boxplots.
Question : Is it okay to compare sentiment percentage with respect to differnt variables?.

```
## 'data.frame':    8 obs. of  8 variables:
##  $ ProximityClose       : int  0 0 0 0 1 1 1 1
##  $ ContactFrequent      : int  0 0 1 1 0 0 1 1
##  $ NormaFavorable       : int  0 1 0 1 0 1 0 1
##  $ Sentiment.Favorable  : int  77 30 14 15 43 36 27 41
##  $ Sentiment.Unfavorable: int  32 36 19 27 20 37 36 118
##  $ total                : int  109 66 33 42 63 73 63 159
##  $ pos_sent             : num  0.706 0.455 0.424 0.357 0.683 ...
##  $ neg_sent             : num  0.294 0.545 0.576 0.643 0.317 ...
```



Interpretation for boxplot 1

Living close to immigrants generally does not have a huge effect on sentiment. Question : Here personally I would assume that if I lived close the immigrants, I would have high of low sentiment.

Interpretation for boxplot 2

By being in contact with immigrant one tends to have a lower positive sentiment when compared to not being in contact.

Interpretation for boxplot 3

People who do not agree to norms generally tend to have a higher positive sentiment to immigration. Question : I would assume the opposite. I am not sure if I am interpretting the data correctly.

Interpretting model

```
m = cbind(Sentiment.Favorable,Sentiment.Unfavorable)
model = glm(m ~ ProximityClose+ContactFrequent+NormaFavorable, family = "binomial")
summary(model)
```

```
##
## Call:
## glm(formula = m ~ ProximityClose + ContactFrequent + NormaFavorable,
##     family = "binomial")
##
## Deviance Residuals:
##       1        2        3        4        5        6        7        8
##  0.3183  -0.8017  -0.4872   1.0446   0.1935   0.2362  -0.2167  -0.1578
##
## Coefficients:
##                  Estimate Std. Error z value Pr(>|z|)
## (Intercept)       0.81143    0.16106   5.038 4.70e-07 ***
## ProximityClose1  -0.09816    0.18441  -0.532    0.595
## ContactFrequent1 -0.94585    0.18113  -5.222 1.77e-07 ***
## NormaFavorable1  -0.79599    0.17665  -4.506 6.61e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 69.8269  on 7  degrees of freedom
## Residual deviance:  2.2378  on 4  degrees of freedom
## AIC: 46.948
##
## Number of Fisher Scoring iterations: 3
```

```
pchisq(2.2378, 8 - 3)
```

```
## [1] 0.1846424
```

Question : How to use deviance and chisqure to decide if the model is good / Interpret the goodnes s of fit.