# Biased QC

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | ***Neutral*** | | ***Biased*** | |
|  | ***Characteristic*** | ***Frequency*** | ***Characteristic*** | ***Frequency*** |
| ***Gender*** | 0 | 1490 | 0 | 1512 |
| 1 | 1510 | 1 | 1488 |
| ***Ethnicity*** | 0 | 1501 | 0 | 661 |
| 1 | 1499 | 1 | 2339\*\* |
| \*\*Considering that Scotland has a very high white population percentage, it is not strange to see that the division is different for biased compared to that of neutral. | | | | |

# Modelling

Overall warning: predictions from a rank-deficient fit may be misleading

## OPTION 1

|  |  |
| --- | --- |
| ***Neutral*** | ***Biased*** |
|  |  |

## OPTION 2

|  |  |
| --- | --- |
| ***Neutral*** | ***Biased*** |
|  |  |

## OPTION 3

|  |  |  |
| --- | --- | --- |
|  | ***Neutral*** | ***Biased*** |
| ***Train*** |  |  |
| ***Test*** |  |  |

Error in model.frame.default(Terms, newdata, na.action = na.action, xlev = object$xlevels) :

factor ETHNIC\_GROUP has new levels 07

## OPTION 4

|  |  |  |
| --- | --- | --- |
|  | ***Neutral*** | ***Biased*** |
| ***Train*** |  |  |
| ***Test*** |  |  |
|  |  |  |

SOURCE: <https://towardsdatascience.com/modelling-binary-logistic-regression-using-tidymodels-library-in-r-part-1-c1bdce0ac055>