**How to identify perfectly correlated variables?**

# Create model predicting the classification column

model\_praise <- glm(classification ~ ., data=df\_praise, family = "binomial")

model\_summary\_praise <- summary(model\_praise)

# In the output see if there are aliased variables

# drop these as they may have perfect multicollinearity

aliased\_vars\_praise <- names(which(model\_summary\_praise$aliased))

df\_praise\_unaliased <- df\_praise[, !(names(df\_praise) %in% aliased\_vars\_praise)]

# Re-run the regression

model\_praise <- glm(classification ~ ., data=df\_praise\_unaliased, family = "binomial")

model\_summary\_praise <- summary(model\_praise)

summary(model\_praise)