Inferential Data Analysis

Ian Dinwoodie

2021-01-24

Loading the Data

Load the raw data and verify its dimensions and structure.

```
df <- readRDS('../data/tidy.Rds')
dim(df)
```

```
## [1] 1023   35
```

```
summary(df)
```

```
##  acq_12_wo_or_less    age_yrs        neutered       train_6mo_or_less
##  Mode :logical     Min.   : 1.000   Mode :logical   Mode :logical
##  FALSE:449         1st Qu.: 4.000   FALSE:132       FALSE:529
##  TRUE :557         Median : 7.000   TRUE :891       TRUE :494
##  NA's :17          Mean   : 7.131
##                    3rd Qu.:10.000
##                    Max.   :19.000
##
##  train_class_count train_technique aggression      fear_anxiety
##  1-3 : 49          punish: 54      Mode :logical   Mode :logical
##  4-6 :120          reward:440      FALSE:474       FALSE:310
##  7-9 : 72          NA's  :529      TRUE :549       TRUE :713
##  10+ :242
##  NA's:540
##
##
##   jumping         barking         coprophagia     compulsion
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:793       FALSE:806       FALSE:642       FALSE:769
##  TRUE :230       TRUE :217       TRUE :381       TRUE :254
##
##
##
##
##  rep_materials   hyperactive     destructive      escape
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:595       FALSE:907       FALSE:892       FALSE:793
##  TRUE :428       TRUE :116       TRUE :131       TRUE :230
##
##
##
##
##   mounting                                  owner_id   train_1_3_mo
```

```
##  Mode :logical   3ea182741999dd54cb902c478ba2704c:  8   FALSE:248
##  FALSE:833       1b9b35f5434de88ff7f3ff4b0e371d48:  7   TRUE :234
##  TRUE :190       796cf2f6f66cf06329ecc6067d7419f0:  6   NA's :541
##                  a5069b3d48cbac2d77080428c7d8d315:  6
##                  f9968086714b82f1c1c87019d1187507:  6
##                  0d29a6dde9e38788ba6a480bf902fb53:  4
##                  (Other)                        :986
##  train_4_mo  train_5_6_mo train_start_age    male         device_used
##  FALSE:267   FALSE:256    1-3 mo:234      Mode :logical   Mode :logical
##  TRUE :215   TRUE :226    4 mo  :130      FALSE:526       FALSE:62
##  NA's :541   NA's :541    5-6 mo:118      TRUE :497       TRUE :432
##                           NA's  :541                      NA's :529
##
##
##
##  buckle_collar   martingale      slip_collar     shock_collar
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:259       FALSE:404       FALSE:449       FALSE:485
##  TRUE :235       TRUE :90        TRUE :45        TRUE :9
##  NA's :529       NA's :529       NA's :529       NA's :529
##
##
##
##   harness        head_halter     choke_collar    prong_collar
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:345       FALSE:468       FALSE:467       FALSE:461
##  TRUE :149       TRUE :26        TRUE :27        TRUE :33
##  NA's :529       NA's :529       NA's :529       NA's :529
##
##
##
##  house_soiling   adj_train_technique punish_device
##  Mode :logical   punish:178          FALSE:316
##  FALSE:225       reward:316          TRUE :178
##  TRUE :798       NA's  :529          NA's :529
##
##
##
##
```

Control vs Experimental Group

Overview

The first question we wanted to answer was if training at a young age (i.e., puppy training) would have an impact on the likelihood of a dog having certain behavior problems. The behavior problems we are exploring are:

- Aggression
- Barking (excessively)
- Compulsion
- Coprophagia
- Destructive behavior
- Escaping/running away

- Fear/anxiety
- House soiling
- Hyperactivity/overactivity
- Mounting
- Problematic jumping
- Rolling in repulsive materials

We initialize a vector to hold these outcomes.

```
outcomes <- c(
  'aggression',
  'fear_anxiety',
  'jumping',
  'barking',
  'coprophagia',
  'compulsion',
  'house_soiling',
  'rep_materials',
  'hyperactive',
  'destructive',
  'escape',
  'mounting'
)
outcomes <- sort(outcomes)
```

Fischer Exact

We check for correlation between the predictor (attending puppy training) and each outcome (the presence of a specific behavior problem). Here we perform the Fisher Exact test with a Benjamini-Hochberg correction.

```
pred <- 'train_6mo_or_less'

idx <- NULL
p_values <- NULL
odds_ratios <- NULL
for (outcome in outcomes) {
  tbl <- table(df[, pred], df[, outcome], dnn=c(pred, outcome))
  fisher <- fisher.test(tbl)
  idx <- c(idx, outcome)
  p_values <- c(p_values, fisher$p.value)
  odds_ratios <- c(odds_ratios, fisher$estimate[[1]])
}

# Correct for the possibility of Type I errors.
p_values <- p.adjust(p_values, method='BH')

# Form a result data frame.
df_out <- data.frame(outcome=idx, p_value=p_values,
                     odds_ratio=odds_ratios)

add_sig_columns <- function(df) {
  df$level <- ''
  df[df$p_value <= .05, 'level'] <- '*'
  df[df$p_value <= .01, 'level'] <- '**'
  df[df$p_value <= .001, 'level'] <- '***'
```

```r
  df$dir <- ''
  df[df$odds_ratio < 1, 'dir'] <- '-'
  df[df$odds_ratio > 1, 'dir'] <- '+'

  return (df)
}

df_out <- add_sig_columns(df_out)
print(knitr::kable(df_out))
```

```
##
##
## |outcome       |   p_value| odds_ratio|level |dir |
## |:-------------|---------:|----------:|:-----|:---|
## |aggression    | 0.0000110|  0.5468702|***   |-   |
## |barking       | 0.0011152|  0.5811859|**    |-   |
## |compulsion    | 0.0019028|  0.6183370|**    |-   |
## |coprophagia   | 0.8685428|  1.0343473|      |+   |
## |destructive   | 0.0000230|  0.4074870|***   |-   |
## |escape        | 0.0007635|  0.5773186|***   |-   |
## |fear_anxiety  | 0.0000107|  0.5073318|***   |-   |
## |house_soiling | 0.0000932|  1.9058972|***   |+   |
## |hyperactive   | 0.5762927|  0.8549112|      |-   |
## |jumping       | 0.9402809|  1.0211929|      |+   |
## |mounting      | 0.8254266|  0.9313191|      |-   |
## |rep_materials | 0.3806390|  1.1616283|      |+   |
```

So we see that there appears to be an impact from the puppy training, but this fails to account for other factors that might be at play.

Binary Logistic Regression

Let's consider what factors may also come in to play:

- Age
- Sex
- Neuter status
- Acquisition of the dog at 12 w.o. or less

We'll perform logistic regression to determine the impact of these factors. To perform logistic regression we'll need to ensure out data subsets have enough responses (`n >= 10`) for each possible answer to be included in the model.

```r
apply_min_xtab <- function(df, outcome, cutoff=10)
{
  drops <- NULL
  for (col in names(df)) {
    if (col == outcome) next
    if (is.integer(df[,col])) next

    xtab <- table(df[,col], df[,outcome])
    if (min(xtab) < cutoff) {
      drops <- c(drops, col)
      break
    }
```

```
  }

  if (length(drops) > 0) {
    cat('\nDropped from model due to insufficient responses:\n')
    cat(drops)
    cat('\n')
  }

  return(df[, !(names(df) %in% drops)])
}
```

Now we perform the logistic regression.

```
pred <- 'train_6mo_or_less'
glm_attribs <- c(
  'age_yrs',
  'male',
  'neutered',
  'acq_12_wo_or_less'
)
for (outcome in outcomes) {
  cat(paste(replicate(80, '-'), collapse=''))
  cat(paste0('\n', outcome, '\n'))
  f <- as.formula(paste0(outcome, '~', '.'))

  df_tmp <- df[,c(outcome, pred, glm_attribs)]
  df_tmp <- apply_min_xtab(df_tmp, outcome)

  glm_fit <- glm(f, data=df_tmp, family='binomial')
  print(summary(glm_fit))
  print(exp(cbind(OR=coef(glm_fit), suppressMessages(confint(glm_fit)))))
  cat('\nVIF:\n')
  print(car::vif(glm_fit))
  cat('\n')
}
```

```
## --------------------------------------------------------------------------------
## aggression
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.5243  -1.1906   0.8893   1.0691   1.4894
##
## Coefficients:
##                         Estimate Std. Error z value Pr(>|z|)
## (Intercept)              0.28272    0.24369   1.160  0.24599
## train_6mo_or_lessTRUE -0.33812    0.15540  -2.176  0.02957 *
## age_yrs                  0.01044    0.01836   0.569  0.56945
## maleTRUE               -0.38260    0.12909  -2.964  0.00304 **
## neuteredTRUE             0.33652    0.20861   1.613  0.10671
## acq_12_wo_or_lessTRUE -0.28142    0.15785  -1.783  0.07461 .
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1389.7  on 1005  degrees of freedom
## Residual deviance: 1353.2  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1365.2
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)           1.3267295 0.8220011 2.1391212
## train_6mo_or_lessTRUE 0.7131094 0.5256385 0.9671128
## age_yrs               1.0104973 0.9747964 1.0475891
## maleTRUE              0.6820870 0.5292792 0.8780857
## neuteredTRUE          1.4000695 0.9317361 2.1134251
## acq_12_wo_or_lessTRUE 0.7547084 0.5537604 1.0286825
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.456646          1.132307          1.005151          1.181134
## acq_12_wo_or_less
##          1.478621
##
## --------------------------------------------------------------------------------
## barking
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.8319  -0.7457  -0.6305  -0.5255   2.0868
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.41966    0.32023  -4.433 9.28e-06 ***
## train_6mo_or_lessTRUE -0.38767    0.19117  -2.028   0.0426 *
## age_yrs                0.01086    0.02195   0.495   0.6207
## maleTRUE              -0.15796    0.15679  -1.007   0.3137
## neuteredTRUE           0.36269    0.28644   1.266   0.2054
## acq_12_wo_or_lessTRUE -0.11351    0.18923  -0.600   0.5486
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1033.4  on 1005  degrees of freedom
## Residual deviance: 1018.2  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1030.2
##
## Number of Fisher Scoring iterations: 4
```

```
##
##                             OR     2.5 %    97.5 %
## (Intercept)           0.2417966 0.1263175 0.4453051
## train_6mo_or_lessTRUE 0.6786391 0.4658726 0.9864198
## age_yrs               1.0109212 0.9682326 1.0553144
## maleTRUE              0.8538850 0.6272429 1.1604477
## neuteredTRUE          1.4371952 0.8361914 2.5849831
## acq_12_wo_or_lessTRUE 0.8926993 0.6153013 1.2928592
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.451750          1.106139          1.004433          1.147991
## acq_12_wo_or_less
##          1.470526
##
## ------------------------------------------------------------------------------
## compulsion
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9357  -0.7621  -0.6970  -0.5621   1.9612
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -0.987582   0.287977  -3.429 0.000605 ***
## train_6mo_or_lessTRUE -0.439158   0.180197  -2.437 0.014806 *
## age_yrs                0.007552   0.020877   0.362 0.717560
## maleTRUE              -0.413480   0.149038  -2.774 0.005532 **
## neuteredTRUE           0.199969   0.252728   0.791 0.428804
## acq_12_wo_or_lessTRUE  0.067512   0.179886   0.375 0.707435
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1125.9  on 1005  degrees of freedom
## Residual deviance: 1108.1  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1120.1
##
## Number of Fisher Scoring iterations: 4
##
##                             OR     2.5 %    97.5 %
## (Intercept)           0.3724764 0.2092693 0.6488176
## train_6mo_or_lessTRUE 0.6445788 0.4521277 0.9168961
## age_yrs               1.0075802 0.9670719 1.0496242
## maleTRUE              0.6613446 0.4929637 0.8845964
## neuteredTRUE          1.2213645 0.7529144 2.0343650
## acq_12_wo_or_lessTRUE 1.0698430 0.7518487 1.5228228
##
## VIF:
```

```
## train_6mo_or_less           age_yrs              male          neutered
##          1.463794          1.119520          1.004774          1.167963
## acq_12_wo_or_less
##          1.485732
##
## ------------------------------------------------------------------------------
## coprophagia
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.0736  -1.0043  -0.9518   1.3541   1.7361
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.193790   0.262279  -4.552 5.32e-06 ***
## train_6mo_or_lessTRUE  0.145809   0.159478   0.914  0.36057
## age_yrs               -0.003865   0.018579  -0.208  0.83520
## maleTRUE               0.103133   0.131219   0.786  0.43189
## neuteredTRUE           0.699608   0.229208   3.052  0.00227 **
## acq_12_wo_or_lessTRUE -0.043516   0.160791  -0.271  0.78667
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1331.8  on 1005  degrees of freedom
## Residual deviance: 1319.8  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1331.8
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.3030705 0.1795343 0.5028319
## train_6mo_or_lessTRUE 1.1569747 0.8468106 1.5831738
## age_yrs               0.9961424 0.9604391 1.0330605
## maleTRUE              1.1086388 0.8571880 1.4340226
## neuteredTRUE          2.0129633 1.2960477 3.1900610
## acq_12_wo_or_lessTRUE 0.9574173 0.6980785 1.3118437
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.481648          1.119302          1.003988          1.158537
## acq_12_wo_or_less
##          1.495042
##
## ------------------------------------------------------------------------------
## destructive
##
## Dropped from model due to insufficient responses:
## neutered
```

```
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.7319  -0.6191  -0.4187  -0.3747   2.3888
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.23782    0.25788  -4.800 1.59e-06 ***
## train_6mo_or_lessTRUE -0.51529    0.24468  -2.106  0.03520 *
## age_yrs               -0.02882    0.02650  -1.088  0.27675
## maleTRUE               0.08612    0.19282   0.447  0.65515
## acq_12_wo_or_lessTRUE -0.69455    0.23599  -2.943  0.00325 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 762.91  on 1005  degrees of freedom
## Residual deviance: 734.18  on 1001  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 744.18
##
## Number of Fisher Scoring iterations: 5
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.2900167 0.1736145 0.4776854
## train_6mo_or_lessTRUE 0.5973292 0.3677804 0.9611914
## age_yrs               0.9715941 0.9219466 1.0229972
## maleTRUE              1.0899347 0.7464088 1.5916759
## acq_12_wo_or_lessTRUE 0.4992980 0.3123305 0.7886628
##
## VIF:
## train_6mo_or_less           age_yrs              male acq_12_wo_or_less
##          1.427742          1.046708          1.001757          1.416973
##
## ---------------------------------------------------------------------------------
## escape
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.9128  -0.7640  -0.6618  -0.4341   2.2029
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.66065    0.33404  -4.971 6.65e-07 ***
## train_6mo_or_lessTRUE -0.37035    0.18698  -1.981   0.0476 *
## age_yrs                0.01900    0.02144   0.886   0.3755
## maleTRUE              -0.25087    0.15404  -1.629   0.1034
```

```
## neuteredTRUE           0.67752    0.30315   2.235   0.0254 *
## acq_12_wo_or_lessTRUE -0.08995    0.18484  -0.487   0.6265
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1069.4  on 1005  degrees of freedom
## Residual deviance: 1045.9  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1057.9
##
## Number of Fisher Scoring iterations: 4
##
##                              OR      2.5 %    97.5 %
## (Intercept)           0.1900159 0.09590775 0.3575299
## train_6mo_or_lessTRUE 0.6904927 0.47797735 0.9955305
## age_yrs               1.0191821 0.97715913 1.0629328
## maleTRUE              0.7781246 0.57456530 1.0515062
## neuteredTRUE          1.9689945 1.11583653 3.6896827
## acq_12_wo_or_lessTRUE 0.9139781 0.63560214 1.3126952
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.444170          1.090144          1.004798          1.120096
## acq_12_wo_or_less
##          1.458664
##
## ------------------------------------------------------------------------------
## fear_anxiety
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8427  -1.0271   0.6817   0.8731   1.4068
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)            0.370507   0.257060   1.441   0.1495
## train_6mo_or_lessTRUE -0.301046   0.171596  -1.754   0.0794 .
## age_yrs               -0.003274   0.020485  -0.160   0.8730
## maleTRUE              -0.138119   0.142894  -0.967   0.3338
## neuteredTRUE           1.131487   0.212636   5.321 1.03e-07 ***
## acq_12_wo_or_lessTRUE -0.426562   0.177709  -2.400   0.0164 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1234.4  on 1005  degrees of freedom
## Residual deviance: 1164.6  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
```

```
## AIC: 1176.6
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %   97.5 %
## (Intercept)           1.4484692 0.8748448 2.399263
## train_6mo_or_lessTRUE 0.7400436 0.5283411 1.035892
## age_yrs               0.9967312 0.9575231 1.037661
## maleTRUE              0.8709947 0.6579036 1.152352
## neuteredTRUE          3.1002623 2.0481048 4.719249
## acq_12_wo_or_lessTRUE 0.6527494 0.4602027 0.924253
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.438509          1.159379          1.006252          1.197763
## acq_12_wo_or_less
##          1.459568
##
## --------------------------------------------------------------------------------
## house_soiling
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.0094   0.5513   0.6130   0.7709   0.9528
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)            0.85347    0.29375   2.905  0.00367 **
## train_6mo_or_lessTRUE  0.44441    0.19021   2.336  0.01947 *
## age_yrs               -0.02300    0.02174  -1.058  0.29005
## maleTRUE               0.06972    0.15463   0.451  0.65205
## neuteredTRUE           0.23007    0.25843   0.890  0.37332
## acq_12_wo_or_lessTRUE  0.30159    0.18844   1.600  0.10950
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1056.8  on 1005  degrees of freedom
## Residual deviance: 1037.1  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1049.1
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %   97.5 %
## (Intercept)           2.3477817 1.3315667 4.223219
## train_6mo_or_lessTRUE 1.5595664 1.0751866 2.268068
## age_yrs               0.9772576 0.9364673 1.019879
## maleTRUE              1.0722128 0.7920439 1.452888
## neuteredTRUE          1.2586928 0.7484575 2.068435
```

```
## acq_12_wo_or_lessTRUE 1.3520086 0.9356246 1.959872
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.462884          1.133238          1.004504          1.206817
## acq_12_wo_or_less
##          1.493502
##
## ------------------------------------------------------------------------------
## hyperactive
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.6367  -0.5092  -0.4631  -0.4091   2.3717
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.436380   0.370745  -3.874 0.000107 ***
## train_6mo_or_lessTRUE -0.138624   0.247948  -0.559 0.576104
## age_yrs               -0.062339   0.029797  -2.092 0.036430 *
## maleTRUE              -0.297379   0.205315  -1.448 0.147504
## neuteredTRUE           0.005863   0.316359   0.019 0.985213
## acq_12_wo_or_lessTRUE -0.073545   0.250962  -0.293 0.769481
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 694.42  on 1005  degrees of freedom
## Residual deviance: 687.11  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 699.11
##
## Number of Fisher Scoring iterations: 5
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.2377871 0.1124937 0.4835239
## train_6mo_or_lessTRUE 0.8705556 0.5353117 1.4166413
## age_yrs               0.9395645 0.8855840 0.9954723
## maleTRUE              0.7427622 0.4944892 1.1079837
## neuteredTRUE          1.0058807 0.5517892 1.9198790
## acq_12_wo_or_lessTRUE 0.9290941 0.5666768 1.5176505
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.488189          1.130048          1.002832          1.198299
## acq_12_wo_or_less
##          1.522316
##
## ------------------------------------------------------------------------------
## jumping
```

```
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1539  -0.7625  -0.5875  -0.3695   2.2592
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -0.27125    0.28758  -0.943    0.346
## train_6mo_or_lessTRUE -0.09794    0.18977  -0.516    0.606
## age_yrs               -0.17924    0.02449  -7.320 2.47e-13 ***
## maleTRUE              -0.08511    0.15616  -0.545    0.586
## neuteredTRUE           0.39494    0.24336   1.623    0.105
## acq_12_wo_or_lessTRUE -0.09519    0.19292  -0.493    0.622
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1071.9  on 1005  degrees of freedom
## Residual deviance: 1011.3  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1023.3
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.7624269 0.4310765 1.3336231
## train_6mo_or_lessTRUE 0.9067054 0.6249065 1.3159186
## age_yrs               0.8359013 0.7960370 0.8763064
## maleTRUE              0.9184067 0.6757009 1.2469317
## neuteredTRUE          1.4842931 0.9293855 2.4178007
## acq_12_wo_or_lessTRUE 0.9091969 0.6222056 1.3265495
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.483850          1.127837          1.002282          1.189305
## acq_12_wo_or_less
##          1.515056
##
## --------------------------------------------------------------------------------
## mounting
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.8162  -0.7754  -0.5006  -0.4733   2.2252
##
## Coefficients:
##                         Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)           -1.267e+00  3.197e-01  -3.963 7.39e-05 ***
## train_6mo_or_lessTRUE  8.297e-02  2.022e-01   0.410    0.682
## age_yrs                8.095e-05  2.362e-02   0.003    0.997
## maleTRUE              -1.002e+00  1.756e-01  -5.706 1.16e-08 ***
## neuteredTRUE           2.548e-01  2.774e-01   0.919    0.358
## acq_12_wo_or_lessTRUE -1.189e-01  2.042e-01  -0.582    0.560
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 966.16  on 1005  degrees of freedom
## Residual deviance: 929.79  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 941.79
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.2816298 0.1480531 0.5202611
## train_6mo_or_lessTRUE 1.0865139 0.7313773 1.6172351
## age_yrs               1.0000810 0.9546437 1.0473487
## maleTRUE              0.3671550 0.2585949 0.5153114
## neuteredTRUE          1.2902306 0.7606592 2.2666129
## acq_12_wo_or_lessTRUE 0.8878848 0.5940726 1.3240177
##
## VIF:
## train_6mo_or_less           age_yrs              male          neutered
##          1.498365          1.117336          1.004566          1.165667
## acq_12_wo_or_less
##          1.525220
##
## --------------------------------------------------------------------------------
## rep_materials
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.3584  -1.0426  -0.8772   1.2531   1.8072
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.49472    0.25657  -5.826 5.69e-09 ***
## train_6mo_or_lessTRUE  0.18499    0.15772   1.173  0.24084
## age_yrs                0.03952    0.01847   2.140  0.03238 *
## maleTRUE               0.42282    0.12983   3.257  0.00113 **
## neuteredTRUE           0.54060    0.21939   2.464  0.01373 *
## acq_12_wo_or_lessTRUE  0.19634    0.15938   1.232  0.21800
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
```

```
##
##     Null deviance: 1367.8  on 1005  degrees of freedom
## Residual deviance: 1338.7  on 1000  degrees of freedom
##   (17 observations deleted due to missingness)
## AIC: 1350.7
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)           0.2243103 0.1345916 0.3684359
## train_6mo_or_lessTRUE 1.2032013 0.8834144 1.6402157
## age_yrs               1.0403115 1.0034078 1.0788137
## maleTRUE              1.5262538 1.1838531 1.9696989
## neuteredTRUE          1.7170341 1.1235367 2.6596628
## acq_12_wo_or_lessTRUE 1.2169354 0.8903594 1.6639116
##
## VIF:
## train_6mo_or_less           age_yrs              male           neutered
##          1.478271          1.130542          1.002118          1.169577
## acq_12_wo_or_less
##          1.491070
```

Investigating Training Specifics

At this point we have identified that puppy training has a significant impact on certain behavior problems. Now we want to investigate the impact of certain training factors.

Isolating the Experimental Data Set

First, we isolate the experimental group (those who attended puppy training) in to their own data set.

```
df_exp <- df %>%
  filter(train_6mo_or_less == TRUE)
summary(df_exp)
```

```
##  acq_12_wo_or_less    age_yrs        neutered       train_6mo_or_less
##  Mode :logical     Min.   : 1.000   Mode :logical   Mode:logical
##  FALSE:78          1st Qu.: 3.000   FALSE:103       TRUE:494
##  TRUE :410         Median : 6.000   TRUE :391
##  NA's :6           Mean   : 6.368
##                    3rd Qu.: 9.000
##                    Max.   :16.000
##
##  train_class_count train_technique aggression      fear_anxiety
##  1-3 : 49          punish: 54      Mode :logical   Mode :logical
##  4-6 :120          reward:440      FALSE:267       FALSE:186
##  7-9 : 72                          TRUE :227       TRUE :308
##  10+ :242
##  NA's: 11
##
##
##   jumping         barking        coprophagia     compulsion
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:382       FALSE:412       FALSE:308       FALSE:394
```

```
##  TRUE :112       TRUE :82        TRUE :186       TRUE :100
##
##
##
##
##  rep_materials   hyperactive     destructive      escape
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:278       FALSE:442       FALSE:455       FALSE:407
##  TRUE :216       TRUE :52        TRUE :39        TRUE :87
##
##
##
##
##   mounting                                   owner_id   train_1_3_mo
##  Mode :logical   3ea182741999dd54cb902c478ba2704c:  8   FALSE:248
##  FALSE:405       249e300dbb0ad0fe1be2fee5d1a3eadd:  4   TRUE :234
##  TRUE :89        30f24317ad30eb964fd7d4c0b9053a5f:  4   NA's : 12
##                  465f724d7f9d1903ffe9ef1230a2054b:  4
##                  7dd1f8eacb783aa0ec257424f46a3361:  4
##                  84155e784cae7d62097ef477c17422c9:  4
##                  (Other)                         :466
##  train_4_mo  train_5_6_mo train_start_age    male         device_used
##  FALSE:267   FALSE:256    1-3 mo:234      Mode :logical   Mode :logical
##  TRUE :215   TRUE :226    4 mo  :130      FALSE:248       FALSE:62
##  NA's : 12   NA's : 12    5-6 mo:118      TRUE :246       TRUE :432
##                           NA's  : 12
##
##
##
##  buckle_collar   martingale      slip_collar     shock_collar
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:259       FALSE:404       FALSE:449       FALSE:485
##  TRUE :235       TRUE :90        TRUE :45        TRUE :9
##
##
##
##
##   harness        head_halter     choke_collar    prong_collar
##  Mode :logical   Mode :logical   Mode :logical   Mode :logical
##  FALSE:345       FALSE:468       FALSE:467       FALSE:461
##  TRUE :149       TRUE :26        TRUE :27        TRUE :33
##
##
##
##
##  house_soiling   adj_train_technique punish_device
##  Mode :logical   punish:178          FALSE:316
##  FALSE:81        reward:316          TRUE :178
##  TRUE :413
##
##
##
##
```

Binary Logistic Regression

We want to answer the following questions about the training:

- Did training in the 1-3 month period produce a better outcome than the 4-6 month period?
- Did training technique (reward vs. punishment) have an impact on the outcome?
- Did the number of sessions have an impact on the outcome?
- Did the choice of restraining device have an impact on the outcome?

We will need to expand the independent variables used for the model to answer these questions.

```
df_exp <- df_exp %>%
  mutate(reward = ifelse(
    is.na(train_technique), NA, ifelse(
      train_technique == 'reward', TRUE, FALSE)))

devices <- c(
  'buckle_collar',
  'martingale',
  'slip_collar',
  'shock_collar',
  'harness',
  'head_halter',
  'choke_collar',
  'prong_collar'
)
training_params <- c(
  'train_1_3_mo',
  'train_4_mo',
  'train_5_6_mo',
  'reward',
  'train_class_count'
)
glm_attribs <- c(
  glm_attribs,
  devices,
  training_params
)

print(glm_attribs)
```

```
##  [1] "age_yrs"           "male"              "neutered"
##  [4] "acq_12_wo_or_less" "buckle_collar"     "martingale"
##  [7] "slip_collar"       "shock_collar"      "harness"
## [10] "head_halter"       "choke_collar"      "prong_collar"
## [13] "train_1_3_mo"      "train_4_mo"        "train_5_6_mo"
## [16] "reward"            "train_class_count"
```

Now we can build an evaluate our models for the various outcomes.

```
for (outcome in outcomes) {
  cat(paste(replicate(80, '-'), collapse=''))
  cat(paste0('\n', outcome, '\n'))
  f <- as.formula(paste0(outcome, '~', '.'))

  df_tmp <- df_exp[,c(outcome, glm_attribs)]
  df_tmp <- apply_min_xtab(df_tmp, outcome)
```

```r
  # If necessary, drop columns due to separability problems.
  if (outcome == 'destructive') {
    df_tmp <- subset(df_tmp, select=-c(shock_collar))
  }

  glm_fit <- glm(f, data=df_tmp, family='binomial')
  print(summary(glm_fit))
  print(exp(cbind(OR=coef(glm_fit), suppressMessages(confint(glm_fit)))))
  cat('\nVIF:\n')
  print(car::vif(glm_fit))
  cat('\n')
}
```

```
## --------------------------------------------------------------------------------
## aggression
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.8274  -1.0685  -0.8825   1.2283   1.5859
##
## Coefficients:
##                          Estimate Std. Error z value Pr(>|z|)
## (Intercept)              0.375434   0.544741   0.689   0.4907
## age_yrs                  0.003273   0.029978   0.109   0.9131
## maleTRUE                -0.376721   0.192757  -1.954   0.0507 .
## neuteredTRUE             0.230137   0.258571   0.890   0.3734
## acq_12_wo_or_lessTRUE -0.228781   0.292247  -0.783   0.4337
## buckle_collarTRUE        0.113402   0.203505   0.557   0.5774
## martingaleTRUE          -0.176716   0.263677  -0.670   0.5027
## slip_collarTRUE          0.111531   0.346274   0.322   0.7474
## harnessTRUE              0.150413   0.217818   0.691   0.4899
## head_halterTRUE         -0.437611   0.443730  -0.986   0.3240
## choke_collarTRUE         0.819801   0.460730   1.779   0.0752 .
## prong_collarTRUE         0.475241   0.437541   1.086   0.2774
## train_1_3_moTRUE        -0.039302   0.231290  -0.170   0.8651
## train_4_moTRUE           0.049367   0.200354   0.246   0.8054
## train_5_6_moTRUE        -0.105160   0.208459  -0.504   0.6139
## rewardTRUE              -0.455424   0.349988  -1.301   0.1932
## train_class_count.L     -0.190385   0.240111  -0.793   0.4278
## train_class_count.Q     -0.014559   0.241751  -0.060   0.9520
## train_class_count.C     -0.029090   0.223876  -0.130   0.8966
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 646.58  on 468  degrees of freedom
```

```
## Residual deviance: 628.11  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 666.11
##
## Number of Fisher Scoring iterations: 4
##
##                               OR     2.5 %    97.5 %
## (Intercept)           1.4556225 0.5003839 4.2554579
## age_yrs               1.0032785 0.9458725 1.0640534
## maleTRUE              0.6861075 0.4693509 0.9998718
## neuteredTRUE          1.2587719 0.7598596 2.0978293
## acq_12_wo_or_lessTRUE 0.7955031 0.4474612 1.4108040
## buckle_collarTRUE     1.1200821 0.7521611 1.6716435
## martingaleTRUE        0.8380180 0.4970849 1.4012737
## slip_collarTRUE       1.1179888 0.5638269 2.2067025
## harnessTRUE           1.1623142 0.7582417 1.7829650
## head_halterTRUE       0.6455767 0.2599245 1.5102692
## choke_collarTRUE      2.2700482 0.9386995 5.8168828
## prong_collarTRUE      1.6084013 0.6853527 3.8613057
## train_1_3_moTRUE      0.9614601 0.6107837 1.5142284
## train_4_moTRUE        1.0506060 0.7089613 1.5562562
## train_5_6_moTRUE      0.9001805 0.5974398 1.3540003
## rewardTRUE            0.6341790 0.3170655 1.2590260
## train_class_count.L   0.8266410 0.5155780 1.3267050
## train_class_count.Q   0.9855467 0.6129965 1.5845185
## train_class_count.C   0.9713291 0.6262632 1.5085679
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.282824  1        1.132618
## male              1.038389  1        1.019014
## neutered          1.267420  1        1.125798
## acq_12_wo_or_less 1.308716  1        1.143991
## buckle_collar     1.157141  1        1.075705
## martingale        1.123447  1        1.059928
## slip_collar       1.110889  1        1.053987
## harness           1.136737  1        1.066179
## head_halter       1.048534  1        1.023979
## choke_collar      1.148032  1        1.071462
## prong_collar      1.225087  1        1.106837
## train_1_3_mo      1.495116  1        1.222749
## train_4_mo        1.109771  1        1.053457
## train_5_6_mo      1.209798  1        1.099908
## reward            1.247220  1        1.116790
## train_class_count 1.245438  3        1.037259
##
## --------------------------------------------------------------------------------
## barking
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
```

```
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.0415  -0.6132  -0.5026  -0.3922   2.4956
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -2.93106    0.79223  -3.700 0.000216 ***
## age_yrs               0.01913    0.04147   0.461 0.644619
## maleTRUE             -0.29032    0.26714  -1.087 0.277135
## neuteredTRUE          0.48711    0.37606   1.295 0.195215
## acq_12_wo_or_lessTRUE  0.44385    0.43445   1.022 0.306949
## buckle_collarTRUE    -0.40922    0.28435  -1.439 0.150104
## martingaleTRUE        0.07953    0.35624   0.223 0.823331
## slip_collarTRUE       0.36985    0.43590   0.848 0.396183
## harnessTRUE          -0.01407    0.29887  -0.047 0.962456
## head_halterTRUE      -1.49449    1.04398  -1.432 0.152276
## choke_collarTRUE      0.52436    0.54847   0.956 0.339044
## prong_collarTRUE      0.57425    0.54676   1.050 0.293592
## train_1_3_moTRUE      0.05767    0.31319   0.184 0.853892
## train_4_moTRUE        0.31891    0.27498   1.160 0.246151
## train_5_6_moTRUE      0.06107    0.28387   0.215 0.829664
## rewardTRUE            0.43049    0.48396   0.890 0.373729
## train_class_count.L  -0.30321    0.33728  -0.899 0.368664
## train_class_count.Q  -0.27175    0.33719  -0.806 0.420284
## train_class_count.C   0.46197    0.31182   1.482 0.138467
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 405.58  on 468  degrees of freedom
## Residual deviance: 384.11  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 422.11
##
## Number of Fisher Scoring iterations: 5
##
##                               OR      2.5 %    97.5 %
## (Intercept)          0.05334027 0.01067494 0.2409907
## age_yrs              1.01931287 0.93928572 1.1055759
## maleTRUE             0.74802413 0.44057897 1.2596565
## neuteredTRUE         1.62760399 0.79801387 3.5216760
## acq_12_wo_or_lessTRUE 1.55869491 0.68639752 3.8289113
## buckle_collarTRUE    0.66416494 0.37734780 1.1549709
## martingaleTRUE       1.08278337 0.52329444 2.1324710
## slip_collarTRUE      1.44751186 0.58826231 3.3036279
## harnessTRUE          0.98603045 0.54148833 1.7552876
## head_halterTRUE      0.22436290 0.01225700 1.1382737
## choke_collarTRUE     1.68938552 0.53791869 4.7598612
## prong_collarTRUE     1.77580191 0.56692471 4.9816154
## train_1_3_moTRUE     1.05937048 0.57376187 1.9670901
## train_4_moTRUE       1.37562361 0.80198886 2.3648686
## train_5_6_moTRUE     1.06297235 0.60720670 1.8541205
```

```
## rewardTRUE            1.53800443 0.62242814 4.2227068
## train_class_count.L   0.73844387 0.39152413 1.4910625
## train_class_count.Q   0.76204492 0.38815722 1.4711930
## train_class_count.C   1.58719209 0.87819653 3.0095940
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.246158  1        1.116315
## male              1.037048  1        1.018356
## neutered          1.206674  1        1.098487
## acq_12_wo_or_less 1.251631  1        1.118763
## buckle_collar     1.143567  1        1.069377
## martingale        1.144600  1        1.069860
## slip_collar       1.114087  1        1.055503
## harness           1.130626  1        1.063309
## head_halter       1.017416  1        1.008670
## choke_collar      1.251491  1        1.118701
## prong_collar      1.262757  1        1.123724
## train_1_3_mo      1.436375  1        1.198488
## train_4_mo        1.110241  1        1.053680
## train_5_6_mo      1.174272  1        1.083638
## reward            1.330981  1        1.153682
## train_class_count 1.292615  3        1.043706
##
## ---------------------------------------------------------------------------------
## compulsion
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1333  -0.7084  -0.5681  -0.4278   2.2835
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.60503    0.69299  -2.316  0.02055 *
## age_yrs                0.00489    0.03748   0.130  0.89619
## maleTRUE              -0.74745    0.24809  -3.013  0.00259 **
## neuteredTRUE           0.13906    0.32447   0.429  0.66823
## acq_12_wo_or_lessTRUE  0.05868    0.35840   0.164  0.86994
## buckle_collarTRUE      0.10544    0.25423   0.415  0.67832
## martingaleTRUE        -0.24867    0.34422  -0.722  0.47003
## slip_collarTRUE        0.27552    0.41024   0.672  0.50183
## harnessTRUE            0.20434    0.26777   0.763  0.44540
## head_halterTRUE       -0.66063    0.64687  -1.021  0.30712
## choke_collarTRUE       0.25924    0.54697   0.474  0.63553
## prong_collarTRUE       1.08148    0.47279   2.287  0.02217 *
## train_1_3_moTRUE      -0.11022    0.28734  -0.384  0.70129
## train_4_moTRUE        -0.13426    0.25059  -0.536  0.59213
## train_5_6_moTRUE       0.29975    0.26348   1.138  0.25525
```

```
## rewardTRUE             0.26208    0.44089   0.594  0.55222
## train_class_count.L   -0.31755    0.29049  -1.093  0.27432
## train_class_count.Q   -0.20092    0.29650  -0.678  0.49800
## train_class_count.C    0.06612    0.27233   0.243  0.80817
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 469.93  on 468  degrees of freedom
## Residual deviance: 446.85  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 484.85
##
## Number of Fisher Scoring iterations: 4
##
##                               OR      2.5 %    97.5 %
## (Intercept)           0.2008839 0.04997078 0.7623563
## age_yrs               1.0049020 0.93330023 1.0813614
## maleTRUE              0.4735733 0.28861325 0.7652114
## neuteredTRUE          1.1491971 0.61529256 2.2071493
## acq_12_wo_or_lessTRUE 1.0604381 0.53158551 2.1794386
## buckle_collarTRUE     1.1112044 0.67468680 1.8323066
## martingaleTRUE        0.7798349 0.38522405 1.4971304
## slip_collarTRUE       1.3172185 0.56609669 2.8684510
## harnessTRUE           1.2267120 0.72063066 2.0646218
## head_halterTRUE       0.5165235 0.11694835 1.6082502
## choke_collarTRUE      1.2959440 0.41269382 3.6294496
## prong_collarTRUE      2.9490324 1.14715457 7.4354223
## train_1_3_moTRUE      0.8956371 0.50901974 1.5750632
## train_4_moTRUE        0.8743655 0.53255522 1.4256995
## train_5_6_moTRUE      1.3495260 0.80434771 2.2658798
## rewardTRUE            1.2996340 0.56568152 3.2233658
## train_class_count.L   0.7279284 0.41773756 1.3147182
## train_class_count.Q   0.8179775 0.45379540 1.4582678
## train_class_count.C   1.0683548 0.63091163 1.8432980
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.261390  1        1.123116
## male              1.048230  1        1.023831
## neutered          1.239231  1        1.113208
## acq_12_wo_or_less 1.312278  1        1.145547
## buckle_collar     1.151800  1        1.073219
## martingale        1.114260  1        1.055585
## slip_collar       1.104698  1        1.051046
## harness           1.133616  1        1.064714
## head_halter       1.036461  1        1.018067
## choke_collar      1.178939  1        1.085790
## prong_collar      1.329945  1        1.153233
## train_1_3_mo      1.458146  1        1.207537
## train_4_mo        1.104256  1        1.050836
## train_5_6_mo      1.236826  1        1.112127
## reward            1.357236  1        1.165005
```

```
## train_class_count 1.291849  3        1.043603
##
## --------------------------------------------------------------------------------
## coprophagia
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.4699  -0.9862  -0.7144   1.2138   2.0349
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -0.13542    0.57514  -0.235 0.813853
## age_yrs              -0.04851    0.03136  -1.547 0.121835
## maleTRUE             -0.03761    0.20120  -0.187 0.851705
## neuteredTRUE          1.08728    0.29787   3.650 0.000262 ***
## acq_12_wo_or_lessTRUE -0.59222    0.29718  -1.993 0.046285 *
## buckle_collarTRUE     0.30773    0.21552   1.428 0.153335
## martingaleTRUE       -0.46683    0.28820  -1.620 0.105275
## slip_collarTRUE       0.10948    0.36521   0.300 0.764358
## harnessTRUE          -0.33384    0.23196  -1.439 0.150092
## head_halterTRUE       0.02429    0.43763   0.055 0.955745
## choke_collarTRUE     -0.49160    0.49909  -0.985 0.324629
## prong_collarTRUE      0.12392    0.44722   0.277 0.781713
## train_1_3_moTRUE      0.21906    0.24906   0.880 0.379102
## train_4_moTRUE       -0.33485    0.21033  -1.592 0.111385
## train_5_6_moTRUE     -0.08970    0.22060  -0.407 0.684283
## rewardTRUE           -0.29551    0.36290  -0.814 0.415470
## train_class_count.L  -0.01087    0.25020  -0.043 0.965361
## train_class_count.Q  -0.05981    0.25215  -0.237 0.812495
## train_class_count.C  -0.33398    0.23223  -1.438 0.150393
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 620.67  on 468  degrees of freedom
## Residual deviance: 582.88  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 620.88
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %    97.5 %
## (Intercept)          0.8733472 0.2809540 2.6919439
## age_yrs              0.9526439 0.8953128 1.0126630
## maleTRUE             0.9630860 0.6485407 1.4283226
## neuteredTRUE         2.9661863 1.6783716 5.4149481
## acq_12_wo_or_lessTRUE 0.5530956 0.3077731 0.9892912
```

```
## buckle_collarTRUE     1.3603367 0.8917601 2.0781975
## martingaleTRUE        0.6269880 0.3514687 1.0918907
## slip_collarTRUE       1.1156946 0.5355843 2.2590879
## harnessTRUE           0.7161681 0.4520638 1.1240170
## head_halterTRUE       1.0245834 0.4256001 2.4072038
## choke_collarTRUE      0.6116447 0.2162624 1.5679029
## prong_collarTRUE      1.1319251 0.4624170 2.7057307
## train_1_3_moTRUE      1.2449081 0.7647103 2.0335675
## train_4_moTRUE        0.7154447 0.4726418 1.0791699
## train_5_6_moTRUE      0.9142048 0.5919098 1.4073414
## rewardTRUE            0.7441532 0.3653929 1.5264207
## train_class_count.L   0.9891932 0.6075210 1.6263567
## train_class_count.Q   0.9419406 0.5730655 1.5431430
## train_class_count.C   0.7160689 0.4535738 1.1291591
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.277412  1        1.130227
## male              1.026755  1        1.013289
## neutered          1.237394  1        1.112382
## acq_12_wo_or_less 1.318134  1        1.148100
## buckle_collar     1.178161  1        1.085431
## martingale        1.126870  1        1.061541
## slip_collar       1.127361  1        1.061773
## harness           1.138521  1        1.067015
## head_halter       1.050240  1        1.024812
## choke_collar      1.150090  1        1.072423
## prong_collar      1.236483  1        1.111973
## train_1_3_mo      1.572400  1        1.253954
## train_4_mo        1.092118  1        1.045044
## train_5_6_mo      1.229532  1        1.108842
## reward            1.296479  1        1.138630
## train_class_count 1.262836  3        1.039660
##
## ---------------------------------------------------------------------------------
## destructive
##
## Dropped from model due to insufficient responses:
## neutered
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -0.7767  -0.4079  -0.3251  -0.2651   2.6837
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.697174   0.991816  -1.711   0.0870 .
## age_yrs               -0.005883   0.054787  -0.107   0.9145
## maleTRUE              -0.001966   0.375730  -0.005   0.9958
## acq_12_wo_or_lessTRUE -0.825572   0.516301  -1.599   0.1098
## buckle_collarTRUE     -0.398816   0.399889  -0.997   0.3186
```

```
## martingaleTRUE        -0.024245   0.506273  -0.048   0.9618
## slip_collarTRUE       -0.590648   0.778975  -0.758   0.4483
## harnessTRUE           -0.161295   0.435538  -0.370   0.7111
## head_halterTRUE        0.881874   0.691411   1.275   0.2021
## choke_collarTRUE       1.437466   0.640824   2.243   0.0249 *
## prong_collarTRUE      -0.899521   1.110215  -0.810   0.4178
## train_1_3_moTRUE       0.221488   0.509597   0.435   0.6638
## train_4_moTRUE        -0.119693   0.416765  -0.287   0.7740
## train_5_6_moTRUE      -0.325218   0.431437  -0.754   0.4510
## rewardTRUE             0.112655   0.671229   0.168   0.8667
## train_class_count.L   -0.069224   0.483106  -0.143   0.8861
## train_class_count.Q   -0.426429   0.462634  -0.922   0.3567
## train_class_count.C   -0.425866   0.397785  -1.071   0.2844
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 238.79  on 468  degrees of freedom
## Residual deviance: 224.14  on 451  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 260.14
##
## Number of Fisher Scoring iterations: 6
##
##                              OR      2.5 %    97.5 %
## (Intercept)           0.1832006 0.02431287  1.225324
## age_yrs               0.9941338 0.89111444  1.105717
## maleTRUE              0.9980356 0.47528330  2.097369
## acq_12_wo_or_lessTRUE 0.4379842 0.15843030  1.225759
## buckle_collarTRUE     0.6711141 0.30026142  1.456856
## martingaleTRUE        0.9760462 0.33240236  2.490650
## slip_collarTRUE       0.5539684 0.08434547  2.087263
## harnessTRUE           0.8510409 0.34625472  1.942981
## head_halterTRUE       2.4154216 0.51289352  8.394027
## choke_collarTRUE      4.2100157 1.11673064 14.324153
## prong_collarTRUE      0.4067645 0.02072995  2.465045
## train_1_3_moTRUE      1.2479319 0.45281749  3.430504
## train_4_moTRUE        0.8871931 0.37658283  1.973527
## train_5_6_moTRUE      0.7223696 0.29718710  1.649918
## rewardTRUE            1.1192458 0.32688269  4.750168
## train_class_count.L   0.9331180 0.38691315  2.715203
## train_class_count.Q   0.6528361 0.24891065  1.574591
## train_class_count.C   0.6532040 0.29769839  1.443160
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.138061  1        1.066800
## male              1.041447  1        1.020513
## acq_12_wo_or_less 1.510577  1        1.229055
## buckle_collar     1.137011  1        1.066307
## martingale        1.125176  1        1.060743
## slip_collar       1.062154  1        1.030609
## harness           1.129719  1        1.062882
```

```
## head_halter       1.103055  1        1.050264
## choke_collar      1.411966  1        1.188262
## prong_collar      1.105616  1        1.051483
## train_1_3_mo      1.905362  1        1.380349
## train_4_mo        1.227481  1        1.107917
## train_5_6_mo      1.315932  1        1.147141
## reward            1.361419  1        1.166798
## train_class_count 1.357348  3        1.052241
##
## ----------------------------------------------------------------------------
## escape
##
## Dropped from model due to insufficient responses:
## slip_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1832  -0.6418  -0.5531  -0.4303   2.3929
##
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -1.87556    0.72520  -2.586   0.0097 **
## age_yrs               -0.03072    0.03930  -0.782   0.4344
## maleTRUE              -0.34760    0.25336  -1.372   0.1701
## neuteredTRUE           0.83352    0.38133   2.186   0.0288 *
## acq_12_wo_or_lessTRUE -0.11236    0.37806  -0.297   0.7663
## buckle_collarTRUE      0.29204    0.26234   1.113   0.2656
## martingaleTRUE        -0.03517    0.34881  -0.101   0.9197
## shock_collarTRUE      -0.06995    1.16145  -0.060   0.9520
## harnessTRUE           -0.17265    0.29031  -0.595   0.5520
## head_halterTRUE        0.80097    0.46980   1.705   0.0882 .
## choke_collarTRUE       0.54636    0.54468   1.003   0.3158
## prong_collarTRUE       0.08690    0.54953   0.158   0.8743
## train_1_3_moTRUE       0.37088    0.31019   1.196   0.2318
## train_4_moTRUE        -0.09863    0.26255  -0.376   0.7072
## train_5_6_moTRUE       0.02241    0.27579   0.081   0.9352
## rewardTRUE            -0.19802    0.44528  -0.445   0.6565
## train_class_count.L   -0.41035    0.29682  -1.382   0.1668
## train_class_count.Q    0.04525    0.31245   0.145   0.8848
## train_class_count.C    0.13604    0.30038   0.453   0.6506
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 434.74  on 468  degrees of freedom
## Residual deviance: 417.28  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 455.28
##
## Number of Fisher Scoring iterations: 4
```

```
##
##                            OR       2.5 %     97.5 %
## (Intercept)           0.1532690 0.03557839 0.6166672
## age_yrs               0.9697488 0.89710804 1.0469257
## maleTRUE              0.7063833 0.42766559 1.1577372
## neuteredTRUE          2.3014056 1.12368432 5.0671558
## acq_12_wo_or_lessTRUE 0.8937244 0.43120642 1.9154563
## buckle_collarTRUE     1.3391590 0.80181430 2.2488564
## martingaleTRUE        0.9654433 0.47322758 1.8727800
## shock_collarTRUE      0.9324427 0.04534613 6.7727200
## harnessTRUE           0.8414289 0.46946393 1.4711640
## head_halterTRUE       2.2277049 0.84638669 5.4572804
## choke_collarTRUE      1.7269505 0.55315228 4.8212415
## prong_collarTRUE      1.0907910 0.34378276 3.0458340
## train_1_3_moTRUE      1.4490150 0.79162981 2.6822779
## train_4_moTRUE        0.9060755 0.53882487 1.5126356
## train_5_6_moTRUE      1.0226671 0.59394224 1.7570205
## rewardTRUE            0.8203568 0.35163873 2.0415985
## train_class_count.L   0.6634156 0.37535116 1.2108178
## train_class_count.Q   1.0462898 0.56515809 1.9357047
## train_class_count.C   1.1457277 0.64444079 2.1102052
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.241771  1        1.114348
## male              1.033774  1        1.016747
## neutered          1.199772  1        1.095341
## acq_12_wo_or_less 1.308786  1        1.144022
## buckle_collar     1.116229  1        1.056517
## martingale        1.132317  1        1.064104
## shock_collar      1.118047  1        1.057378
## harness           1.116981  1        1.056873
## head_halter       1.065392  1        1.032178
## choke_collar      1.214980  1        1.102261
## prong_collar      1.251383  1        1.118652
## train_1_3_mo      1.560223  1        1.249089
## train_4_mo        1.086692  1        1.042445
## train_5_6_mo      1.230059  1        1.109080
## reward            1.407324  1        1.186307
## train_class_count 1.301495  3        1.044898
##
## --------------------------------------------------------------------------------
## fear_anxiety
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.1631  -1.1128   0.6847   0.9323   1.8367
##
```

```
## Coefficients:
##                       Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -0.295749   0.583622  -0.507    0.612
## age_yrs              -0.006624   0.032016  -0.207    0.836
## maleTRUE             -0.163034   0.204644  -0.797    0.426
## neuteredTRUE          1.181618   0.267541   4.417    1e-05 ***
## acq_12_wo_or_lessTRUE -0.408449   0.331267  -1.233    0.218
## buckle_collarTRUE    -0.247954   0.215781  -1.149    0.251
## martingaleTRUE       -0.199382   0.273688  -0.729    0.466
## slip_collarTRUE      -0.250162   0.367634  -0.680    0.496
## harnessTRUE           0.254063   0.233974   1.086    0.278
## head_halterTRUE      -0.232444   0.443044  -0.525    0.600
## choke_collarTRUE      0.720168   0.483117   1.491    0.136
## prong_collarTRUE      0.440187   0.478753   0.919    0.358
## train_1_3_moTRUE     -0.122317   0.242738  -0.504    0.614
## train_4_moTRUE        0.189073   0.213249   0.887    0.375
## train_5_6_moTRUE      0.295118   0.222542   1.326    0.185
## rewardTRUE            0.395585   0.367040   1.078    0.281
## train_class_count.L  -0.354792   0.258389  -1.373    0.170
## train_class_count.Q  -0.341077   0.261176  -1.306    0.192
## train_class_count.C  -0.084310   0.247394  -0.341    0.733
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 627.37  on 468  degrees of freedom
## Residual deviance: 573.63  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 611.63
##
## Number of Fisher Scoring iterations: 4
##
##                             OR     2.5 %   97.5 %
## (Intercept)           0.7439744 0.2362902 2.341720
## age_yrs               0.9933978 0.9328211 1.057835
## maleTRUE              0.8495621 0.5679431 1.267988
## neuteredTRUE          3.2596435 1.9393348 5.546194
## acq_12_wo_or_lessTRUE 0.6646801 0.3412753 1.257608
## buckle_collarTRUE     0.7803958 0.5104536 1.190830
## martingaleTRUE        0.8192366 0.4791400 1.404831
## slip_collarTRUE       0.7786750 0.3792602 1.613257
## harnessTRUE           1.2892527 0.8172293 2.048283
## head_halterTRUE       0.7925941 0.3355734 1.940253
## choke_collarTRUE      2.0547790 0.8150494 5.500366
## prong_collarTRUE      1.5529973 0.6223025 4.120343
## train_1_3_moTRUE      0.8848675 0.5496392 1.425579
## train_4_moTRUE        1.2081287 0.7960014 1.838449
## train_5_6_moTRUE      1.3432854 0.8690906 2.082023
## rewardTRUE            1.4852524 0.7210943 3.061204
## train_class_count.L   0.7013194 0.4177573 1.155661
## train_class_count.Q   0.7110040 0.4261019 1.189339
## train_class_count.C   0.9191459 0.5634255 1.489689
##
```

```
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.293048  1        1.137123
## male              1.038117  1        1.018880
## neutered          1.274560  1        1.128964
## acq_12_wo_or_less 1.246259  1        1.116360
## buckle_collar     1.154295  1        1.074381
## martingale        1.119378  1        1.058007
## slip_collar       1.120367  1        1.058474
## harness           1.117439  1        1.057090
## head_halter       1.053560  1        1.026431
## choke_collar      1.169008  1        1.081207
## prong_collar      1.254651  1        1.120112
## train_1_3_mo      1.458085  1        1.207512
## train_4_mo        1.111991  1        1.054510
## train_5_6_mo      1.215526  1        1.102509
## reward            1.296428  1        1.138608
## train_class_count 1.218547  3        1.033492
##
## -----------------------------------------------------------------------------
## house_soiling
##
## Dropped from model due to insufficient responses:
## slip_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -2.4160   0.4096   0.5105   0.6291   1.1955
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)           -0.023875   0.715962  -0.033   0.9734
## age_yrs                0.025041   0.041071   0.610   0.5421
## maleTRUE              -0.094439   0.260554  -0.362   0.7170
## neuteredTRUE           0.434842   0.337667   1.288   0.1978
## acq_12_wo_or_lessTRUE  0.274056   0.382225   0.717   0.4734
## buckle_collarTRUE      0.006564   0.270295   0.024   0.9806
## martingaleTRUE        -0.529588   0.330254  -1.604   0.1088
## shock_collarTRUE       0.380955   1.162205   0.328   0.7431
## harnessTRUE           -0.159271   0.292503  -0.545   0.5861
## head_halterTRUE        0.265267   0.655854   0.404   0.6859
## choke_collarTRUE       0.103688   0.570560   0.182   0.8558
## prong_collarTRUE      -0.047604   0.583453  -0.082   0.9350
## train_1_3_moTRUE       0.361359   0.331696   1.089   0.2760
## train_4_moTRUE         0.303461   0.286968   1.057   0.2903
## train_5_6_moTRUE       0.480018   0.296874   1.617   0.1059
## rewardTRUE             0.615075   0.442018   1.392   0.1641
## train_class_count.L    0.515414   0.311921   1.652   0.0985 .
## train_class_count.Q    0.178654   0.311530   0.573   0.5663
## train_class_count.C   -0.289380   0.290497  -0.996   0.3192
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 418.85  on 468  degrees of freedom
## Residual deviance: 397.89  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 435.89
##
## Number of Fisher Scoring iterations: 5
##
##                             OR     2.5 %    97.5 %
## (Intercept)           0.9764080 0.2400267  4.020001
## age_yrs               1.0253572 0.9465177  1.112335
## maleTRUE              0.9098833 0.5446687  1.517109
## neuteredTRUE          1.5447183 0.7898596  2.983300
## acq_12_wo_or_lessTRUE 1.3152886 0.6094332  2.751398
## buckle_collarTRUE     1.0065861 0.5920237  1.713677
## martingaleTRUE        0.5888477 0.3116602  1.144758
## shock_collarTRUE      1.4636813 0.2025072 30.149306
## harnessTRUE           0.8527651 0.4838399  1.529228
## head_halterTRUE       1.3037791 0.4101984  5.836549
## choke_collarTRUE      1.1092549 0.3860300  3.748191
## prong_collarTRUE      0.9535110 0.3238613  3.301521
## train_1_3_moTRUE      1.4352780 0.7533097  2.784729
## train_4_moTRUE        1.3545394 0.7785353  2.412116
## train_5_6_moTRUE      1.6161031 0.9122310  2.937948
## rewardTRUE            1.8497960 0.7578953  4.340206
## train_class_count.L   1.6743321 0.8849223  3.036709
## train_class_count.Q   1.1956068 0.6493297  2.217304
## train_class_count.C   0.7487273 0.4150701  1.305257
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.320303  1        1.149044
## male              1.040006  1        1.019807
## neutered          1.333954  1        1.154969
## acq_12_wo_or_less 1.324050  1        1.150674
## buckle_collar     1.112675  1        1.054834
## martingale        1.125101  1        1.060708
## shock_collar      1.119854  1        1.058231
## harness           1.137556  1        1.066563
## head_halter       1.051429  1        1.025392
## choke_collar      1.180608  1        1.086558
## prong_collar      1.275229  1        1.129260
## train_1_3_mo      1.672977  1        1.293436
## train_4_mo        1.220398  1        1.104716
## train_5_6_mo      1.302378  1        1.141218
## reward            1.438563  1        1.199401
## train_class_count 1.258133  3        1.039013
##
## ----------------------------------------------------------------------------
## hyperactive
##
```

```
## Dropped from model due to insufficient responses:
## acq_12_wo_or_less
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1216  -0.4856  -0.3791  -0.2822   2.5533
##
## Coefficients:
##                      Estimate Std. Error z value Pr(>|z|)
## (Intercept)         -1.382373   0.738587  -1.872   0.0613 .
## age_yrs             -0.061715   0.051823  -1.191   0.2337
## maleTRUE            -0.533756   0.328489  -1.625   0.1042
## neuteredTRUE        -0.142442   0.409681  -0.348   0.7281
## buckle_collarTRUE    0.284451   0.335263   0.848   0.3962
## martingaleTRUE       0.506873   0.399078   1.270   0.2040
## slip_collarTRUE      0.951895   0.465278   2.046   0.0408 *
## shock_collarTRUE     0.229088   1.183946   0.193   0.8466
## harnessTRUE          0.481076   0.343868   1.399   0.1618
## head_halterTRUE      0.620291   0.618141   1.003   0.3156
## choke_collarTRUE    -0.629790   0.863666  -0.729   0.4659
## prong_collarTRUE     0.905975   0.594848   1.523   0.1277
## train_1_3_moTRUE    -0.005089   0.361211  -0.014   0.9888
## train_4_moTRUE      -0.102434   0.332408  -0.308   0.7580
## train_5_6_moTRUE    -0.001012   0.347173  -0.003   0.9977
## rewardTRUE          -0.669746   0.506994  -1.321   0.1865
## train_class_count.L -0.574674   0.381133  -1.508   0.1316
## train_class_count.Q -0.324865   0.395249  -0.822   0.4111
## train_class_count.C  0.415263   0.366557   1.133   0.2573
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 315.37  on 474  degrees of freedom
## Residual deviance: 288.91  on 456  degrees of freedom
##   (19 observations deleted due to missingness)
## AIC: 326.91
##
## Number of Fisher Scoring iterations: 5
##
##                            OR      2.5 %   97.5 %
## (Intercept)         0.2509822 0.05626527 1.036010
## age_yrs             0.9401511 0.84782077 1.039514
## maleTRUE            0.5863984 0.30329499 1.107074
## neuteredTRUE        0.8672379 0.39386700 1.982158
## buckle_collarTRUE   1.3290325 0.68896489 2.582203
## martingaleTRUE      1.6600924 0.73772743 3.566901
## slip_collarTRUE     2.5906154 0.99289133 6.287462
## shock_collarTRUE    1.2574533 0.05975202 9.725324
## harnessTRUE         1.6178145 0.81630767 3.164988
## head_halterTRUE     1.8594685 0.48486981 5.770521
```

```
## choke_collarTRUE    0.5327038 0.07090973 2.392983
## prong_collarTRUE    2.4743443 0.72244221 7.656427
## train_1_3_moTRUE    0.9949238 0.48547210 2.016309
## train_4_moTRUE      0.9026377 0.46577639 1.726732
## train_5_6_moTRUE    0.9989890 0.50182412 1.972496
## rewardTRUE          0.5118385 0.19422587 1.443392
## train_class_count.L 0.5628885 0.27310966 1.240596
## train_class_count.Q 0.7226251 0.32897880 1.573810
## train_class_count.C 1.5147683 0.76150113 3.262333
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.317444  1        1.147800
## male              1.066816  1        1.032868
## neutered          1.316697  1        1.147474
## buckle_collar     1.152925  1        1.073744
## martingale        1.151159  1        1.072921
## slip_collar       1.116643  1        1.056713
## shock_collar      1.152666  1        1.073623
## harness           1.135152  1        1.065435
## head_halter       1.091692  1        1.044841
## choke_collar      1.125777  1        1.061026
## prong_collar      1.408341  1        1.186735
## train_1_3_mo      1.339131  1        1.157208
## train_4_mo        1.113497  1        1.055224
## train_5_6_mo      1.237475  1        1.112419
## reward            1.449073  1        1.203775
## train_class_count 1.371328  3        1.054039
##
## ---------------------------------------------------------------------
## jumping
##
## Dropped from model due to insufficient responses:
## slip_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.4996  -0.7297  -0.4981  -0.1727   2.5978
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)            0.199759   0.724868   0.276  0.78287
## age_yrs               -0.242080   0.043024  -5.627 1.84e-08 ***
## maleTRUE              -0.516313   0.248722  -2.076  0.03791 *
## neuteredTRUE           0.234723   0.310149   0.757  0.44917
## acq_12_wo_or_lessTRUE -0.319237   0.351404  -0.908  0.36363
## buckle_collarTRUE      0.173825   0.252027   0.690  0.49038
## martingaleTRUE         0.006702   0.330736   0.020  0.98383
## shock_collarTRUE      -0.517250   1.209982  -0.427  0.66903
## harnessTRUE            0.100596   0.264059   0.381  0.70323
## head_halterTRUE        0.888589   0.499913   1.777  0.07549 .
```

```
## choke_collarTRUE      -1.687648   1.061577  -1.590  0.11189
## prong_collarTRUE       0.738166   0.553195   1.334  0.18208
## train_1_3_moTRUE      -0.317923   0.293404  -1.084  0.27856
## train_4_moTRUE        -0.048166   0.251748  -0.191  0.84827
## train_5_6_moTRUE       0.083423   0.263761   0.316  0.75179
## rewardTRUE             0.157336   0.520978   0.302  0.76265
## train_class_count.L    0.598523   0.362072   1.653  0.09832 .
## train_class_count.Q   -0.966764   0.336529  -2.873  0.00407 **
## train_class_count.C   -0.148642   0.273071  -0.544  0.58621
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 498.81  on 468  degrees of freedom
## Residual deviance: 432.58  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 470.58
##
## Number of Fisher Scoring iterations: 6
##
##                              OR       2.5 %    97.5 %
## (Intercept)           1.2211086 0.286112525 4.9706239
## age_yrs               0.7849937 0.719531959 0.8520434
## maleTRUE              0.5967164 0.364079014 0.9673550
## neuteredTRUE          1.2645581 0.694207958 2.3491345
## acq_12_wo_or_lessTRUE 0.7267032 0.366161387 1.4589373
## buckle_collarTRUE     1.1898471 0.726491008 1.9553935
## martingaleTRUE        1.0067248 0.516998014 1.9008747
## shock_collarTRUE      0.5961579 0.027559333 5.0024194
## harnessTRUE           1.1058303 0.654964421 1.8479213
## head_halterTRUE       2.4316964 0.879531520 6.3800938
## choke_collarTRUE      0.1849540 0.009931756 0.9854140
## prong_collarTRUE      2.0920958 0.671883627 6.0319441
## train_1_3_moTRUE      0.7276585 0.408341157 1.2937916
## train_4_moTRUE        0.9529754 0.580048686 1.5595992
## train_5_6_moTRUE      1.0870015 0.647029760 1.8242246
## rewardTRUE            1.1703886 0.438596236 3.4461565
## train_class_count.L   1.8194300 0.938839478 3.9788669
## train_class_count.Q   0.3803119 0.189345389 0.7178206
## train_class_count.C   0.8618779 0.505440061 1.4797722
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.297995  1        1.139296
## male              1.072358  1        1.035547
## neutered          1.251266  1        1.118600
## acq_12_wo_or_less 1.347749  1        1.160926
## buckle_collar     1.114124  1        1.055521
## martingale        1.148263  1        1.071570
## shock_collar      1.120197  1        1.058394
## harness           1.134647  1        1.065198
## head_halter       1.068272  1        1.033572
## choke_collar      1.038209  1        1.018925
```

```
## prong_collar      1.279440  1        1.131124
## train_1_3_mo      1.499762  1        1.224648
## train_4_mo        1.097760  1        1.047740
## train_5_6_mo      1.218849  1        1.104015
## reward            1.364158  1        1.167972
## train_class_count 1.298939  3        1.044555
##
## -----------------------------------------------------------------------------
## mounting
##
## Dropped from model due to insufficient responses:
## slip_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.1867  -0.6938  -0.4786  -0.3039   2.4602
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
## (Intercept)          -1.4998406  0.7561272  -1.984   0.0473 *
## age_yrs               0.0376959  0.0396628   0.950   0.3419
## maleTRUE             -1.2483584  0.2763443  -4.517 6.26e-06 ***
## neuteredTRUE          0.3322148  0.3500520   0.949   0.3426
## acq_12_wo_or_lessTRUE -0.2126047  0.3711237  -0.573   0.5667
## buckle_collarTRUE     0.1978781  0.2620727   0.755   0.4502
## martingaleTRUE       -0.0049849  0.3464501  -0.014   0.9885
## shock_collarTRUE     -0.3129693  1.1981644  -0.261   0.7939
## harnessTRUE           0.5632694  0.2779762   2.026   0.0427 *
## head_halterTRUE      -1.7602473  1.0439690  -1.686   0.0918 .
## choke_collarTRUE     -0.6323080  0.6847490  -0.923   0.3558
## prong_collarTRUE     -0.0292057  0.5974019  -0.049   0.9610
## train_1_3_moTRUE     -0.1942244  0.3026190  -0.642   0.5210
## train_4_moTRUE        0.3871708  0.2647509   1.462   0.1436
## train_5_6_moTRUE      0.0422359  0.2708905   0.156   0.8761
## rewardTRUE           -0.1845292  0.5064406  -0.364   0.7156
## train_class_count.L  -0.0627794  0.3359793  -0.187   0.8518
## train_class_count.Q   0.0003757  0.3270400   0.001   0.9991
## train_class_count.C   0.1269765  0.2968234   0.428   0.6688
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 443.92  on 468  degrees of freedom
## Residual deviance: 403.36  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 441.36
##
## Number of Fisher Scoring iterations: 5
##
##                              OR       2.5 %    97.5 %
```

```
## (Intercept)          0.2231657 0.048516332 0.9515661
## age_yrs              1.0384154 0.960549433 1.1225743
## maleTRUE             0.2869755 0.163918931 0.4863359
## neuteredTRUE         1.3940523 0.713812111 2.8364679
## acq_12_wo_or_lessTRUE 0.8084757 0.393481787 1.6968564
## buckle_collarTRUE    1.2188139 0.729398422 2.0429723
## martingaleTRUE       0.9950275 0.492047927 1.9281125
## shock_collarTRUE     0.7312724 0.034060064 5.7858261
## harnessTRUE          1.7564056 1.014730249 3.0263376
## head_halterTRUE      0.1720023 0.009398555 0.8740413
## choke_collarTRUE     0.5313640 0.113632113 1.8039310
## prong_collarTRUE     0.9712167 0.273627369 2.9480148
## train_1_3_moTRUE     0.8234731 0.453952508 1.4920972
## train_4_moTRUE       1.4728081 0.876863877 2.4824643
## train_5_6_moTRUE     1.0431405 0.611648073 1.7736521
## rewardTRUE           0.8314957 0.319248847 2.3715838
## train_class_count.L  0.9391506 0.498993855 1.8903607
## train_class_count.Q  1.0003757 0.519536202 1.8889306
## train_class_count.C  1.1353903 0.639620436 2.0593044
##
## VIF:
##                      GVIF Df GVIF^(1/(2*Df))
## age_yrs           1.252838  1        1.119302
## male              1.030600  1        1.015185
## neutered          1.187056  1        1.089521
## acq_12_wo_or_less 1.322571  1        1.150031
## buckle_collar     1.092780  1        1.045361
## martingale        1.132735  1        1.064300
## shock_collar      1.125242  1        1.060774
## harness           1.148907  1        1.071871
## head_halter       1.016548  1        1.008240
## choke_collar      1.143288  1        1.069246
## prong_collar      1.281878  1        1.132201
## train_1_3_mo      1.437720  1        1.199050
## train_4_mo        1.114935  1        1.055905
## train_5_6_mo      1.166836  1        1.080202
## reward            1.441323  1        1.200551
## train_class_count 1.297137  3        1.044314
##
## ------------------------------------------------------------------------------
## rep_materials
##
## Dropped from model due to insufficient responses:
## shock_collar
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.7258  -1.0188  -0.7019   1.1504   1.9700
##
## Coefficients:
##                        Estimate Std. Error z value Pr(>|z|)
```

```
## (Intercept)           -1.092738   0.569546  -1.919 0.055033 .
## age_yrs                0.114907   0.031583   3.638 0.000274 ***
## maleTRUE               0.356415   0.198404   1.796 0.072429 .
## neuteredTRUE           0.493642   0.278023   1.776 0.075808 .
## acq_12_wo_or_lessTRUE -0.322648   0.298260  -1.082 0.279357
## buckle_collarTRUE      0.221394   0.209676   1.056 0.291020
## martingaleTRUE         0.085154   0.268791   0.317 0.751393
## slip_collarTRUE       -0.426186   0.370147  -1.151 0.249570
## harnessTRUE            0.346729   0.224765   1.543 0.122922
## head_halterTRUE       -0.123533   0.441467  -0.280 0.779612
## choke_collarTRUE      -0.851448   0.507506  -1.678 0.093403 .
## prong_collarTRUE      -0.721600   0.470098  -1.535 0.124784
## train_1_3_moTRUE       0.028952   0.238844   0.121 0.903518
## train_4_moTRUE        -0.026246   0.206152  -0.127 0.898691
## train_5_6_moTRUE       0.023108   0.213813   0.108 0.913937
## rewardTRUE            -0.297861   0.368359  -0.809 0.418737
## train_class_count.L   -0.198232   0.248252  -0.799 0.424574
## train_class_count.Q    0.009413   0.248755   0.038 0.969816
## train_class_count.C    0.004317   0.229917   0.019 0.985018
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 642.22  on 468  degrees of freedom
## Residual deviance: 598.32  on 450  degrees of freedom
##   (25 observations deleted due to missingness)
## AIC: 636.32
##
## Number of Fisher Scoring iterations: 4
##
##                              OR     2.5 %   97.5 %
## (Intercept)           0.3352972 0.1086536 1.018722
## age_yrs               1.1217696 1.0552447 1.194579
## maleTRUE              1.4282007 0.9689594 2.110645
## neuteredTRUE          1.6382728 0.9564892 2.853521
## acq_12_wo_or_lessTRUE 0.7242291 0.4021245 1.298225
## buckle_collarTRUE     1.2478146 0.8274158 1.884330
## martingaleTRUE        1.0888844 0.6405951 1.842386
## slip_collarTRUE       0.6529950 0.3086533 1.329030
## harnessTRUE           1.4144327 0.9111067 2.201915
## head_halterTRUE       0.8837924 0.3668592 2.105663
## choke_collarTRUE      0.4267963 0.1481667 1.107458
## prong_collarTRUE      0.4859742 0.1863547 1.194797
## train_1_3_moTRUE      1.0293756 0.6445094 1.646127
## train_4_moTRUE        0.9740951 0.6500055 1.459801
## train_5_6_moTRUE      1.0233766 0.6725900 1.556797
## rewardTRUE            0.7424048 0.3586248 1.530794
## train_class_count.L   0.8201797 0.5040861 1.339250
## train_class_count.Q   1.0094570 0.6193422 1.645617
## train_class_count.C   1.0043266 0.6401120 1.579025
##
## VIF:
##                       GVIF Df GVIF^(1/(2*Df))
```

```
## age_yrs           1.284161  1        1.133208
## male              1.032681  1        1.016209
## neutered          1.218125  1        1.103687
## acq_12_wo_or_less 1.301826  1        1.140976
## buckle_collar     1.152545  1        1.073567
## martingale        1.121574  1        1.059044
## slip_collar       1.109824  1        1.053482
## harness           1.145375  1        1.070222
## head_halter       1.047263  1        1.023359
## choke_collar      1.172813  1        1.082965
## prong_collar      1.232045  1        1.109975
## train_1_3_mo      1.495330  1        1.222837
## train_4_mo        1.101154  1        1.049359
## train_5_6_mo      1.194565  1        1.092962
## reward            1.308008  1        1.143682
## train_class_count 1.253524  3        1.038378
```

Trends in Dog Age

Impact of Age on Training Attendance

In our exploratory data analysis we saw a trend that seemed to indicate that younger dogs were more likely to have attended puppy training than older dogs. If true, this would seem to suggest that puppy training is becoming more popular over time. To start, we introduce the question to be answered:

- Were younger dogs more likely to attend more sessions?

Binomial Logistic Regression

To answer this, we build a regression model where the predictor is the dog's age and the outcome is whether or not they attended puppy training.

```
pred <- 'age_yrs'
outcome <- 'train_6mo_or_less'
df_tmp <- df[,c(pred, outcome)]
df_tmp <- apply_min_xtab(df_tmp, outcome)
f <- as.formula(paste0(outcome, '~', '.'))
glm_fit <- glm(f, data=df_tmp, family='binomial')
summary(glm_fit)
```

```
##
## Call:
## glm(formula = f, family = "binomial", data = df_tmp)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -1.4381  -1.1073  -0.8542   1.1560   1.6378
##
## Coefficients:
##             Estimate Std. Error z value Pr(>|z|)
## (Intercept)  0.70359    0.13922   5.054 4.33e-07 ***
## age_yrs     -0.10885    0.01756  -6.198 5.73e-10 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## (Dispersion parameter for binomial family taken to be 1)
##
##     Null deviance: 1417.0  on 1022  degrees of freedom
## Residual deviance: 1376.7  on 1021  degrees of freedom
## AIC: 1380.7
##
## Number of Fisher Scoring iterations: 4
```

```
print(exp(cbind(OR=coef(glm_fit), suppressMessages(confint(glm_fit)))))
```

```
##                    OR     2.5 %    97.5 %
## (Intercept) 2.0210008 1.5411546 2.6609099
## age_yrs     0.8968673 0.8662085 0.9279871
```

Linear Regression

Factors plot poorly, so let's extract the probability of attendance for each age range and fit it with a linear regression model.

```
p_attend_vec <- NULL
max_age <- max(df$age_yrs)
for (i in 1:max_age) {
  df_tmp <- df %>%
    filter(age_yrs == i) %>%
    select(train_6mo_or_less)
  p <- sum(df_tmp$train_6mo_or_less)/length(df_tmp$train_6mo_or_less)
  p_attend_vec <- c(p_attend_vec, p)
}

df_age <- data.frame(age=c(1:max_age), p_attend=p_attend_vec)
head(df_age)
```

```
##   age  p_attend
## 1   1 0.7142857
## 2   2 0.5925926
## 3   3 0.6901408
## 4   4 0.4507042
## 5   5 0.4505495
## 6   6 0.5465116
```

```
# Remove the outlier.
outlier <- df_age[15,]
df_age <- df_age[-c(15),]

lm_fit <- lm(p_attend~age, data=df_age)
summary(lm_fit)
```

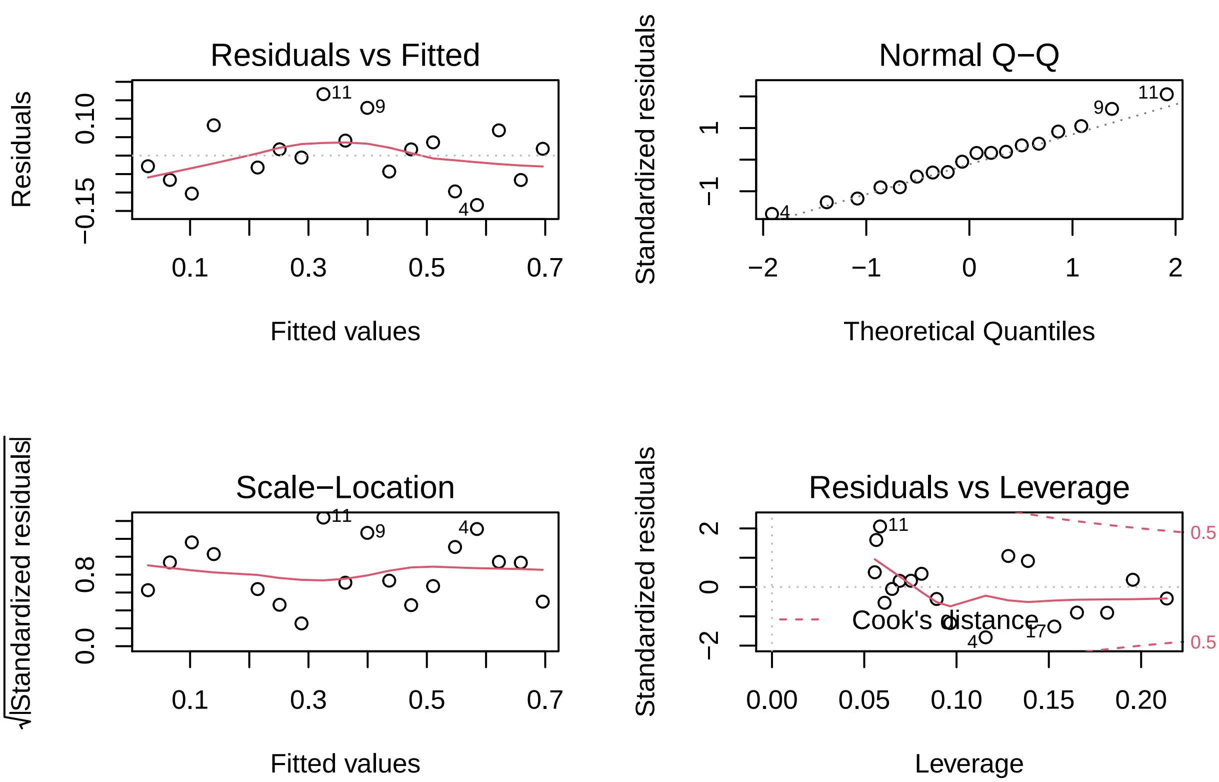
```
##
## Call:
## lm(formula = p_attend ~ age, data = df_age)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.133965 -0.060226  0.005833  0.039471  0.166517
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
```

```
## (Intercept)  0.73289    0.03976   18.43 3.36e-12 ***
## age         -0.03705    0.00356  -10.41 1.57e-08 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.08301 on 16 degrees of freedom
## Multiple R-squared:  0.8713, Adjusted R-squared:  0.8633
## F-statistic: 108.3 on 1 and 16 DF,  p-value: 1.571e-08
```

```r
print(confint(lm_fit))
```

```
##                  2.5 %      97.5 %
## (Intercept)  0.6485991  0.81717689
## age         -0.0446021 -0.02950727
```

```r
old_par <- par(mfrow=c(2,2))
plot(lm_fit)
```




```r
par(old_par)
```

TODO: observations

Polynomial Regression

```r
plm_fit <- lm(p_attend~poly(age, 3), data=df_age)
summary(plm_fit)
```

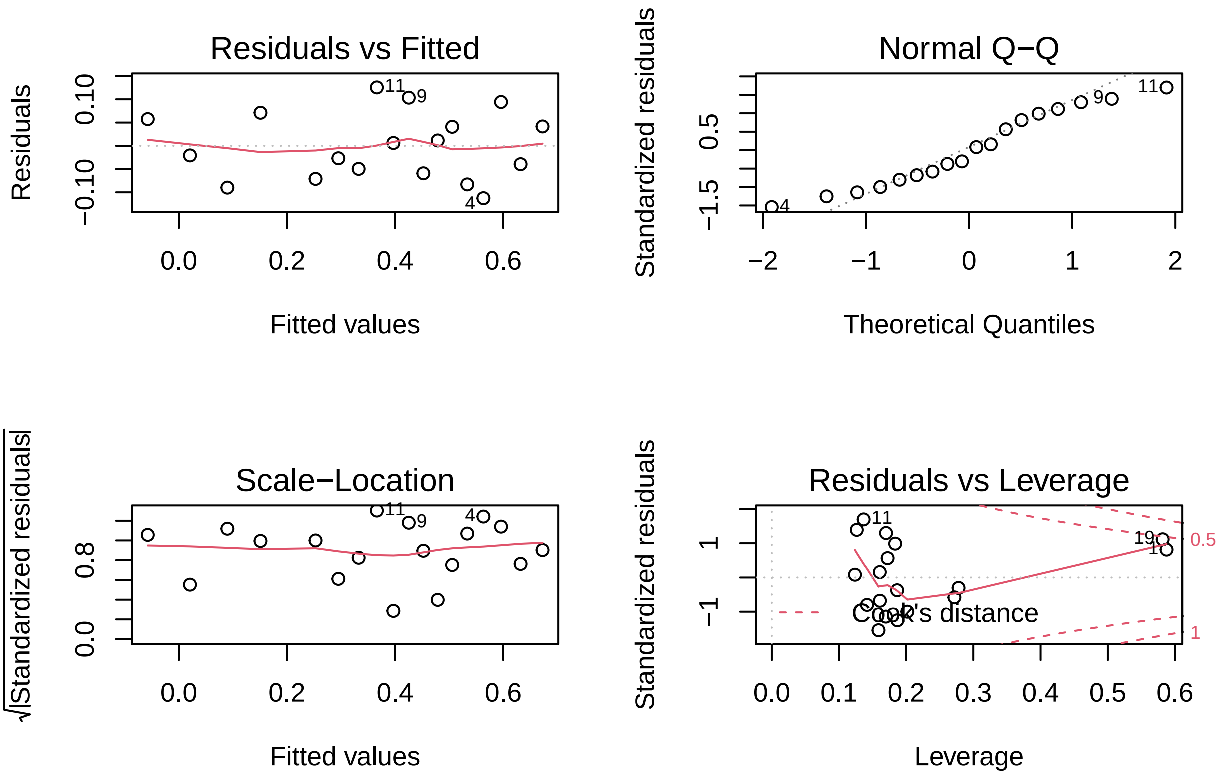
```
##
```

```
## Call:
## lm(formula = p_attend ~ poly(age, 3), data = df_age)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.11262 -0.05671 -0.00726  0.05349  0.12554
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)    0.37263    0.01875  19.869 1.18e-11 ***
## poly(age, 3)1 -0.86395    0.07957 -10.858 3.34e-08 ***
## poly(age, 3)2 -0.12730    0.07957  -1.600    0.132
## poly(age, 3)3 -0.07353    0.07957  -0.924    0.371
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07957 on 14 degrees of freedom
## Multiple R-squared:  0.8965, Adjusted R-squared:  0.8744
## F-statistic: 40.44 on 3 and 14 DF,  p-value: 3.804e-07
```

```r
print(confint(plm_fit))
```

```
##                    2.5 %      97.5 %
## (Intercept)    0.3324096  0.41285860
## poly(age, 3)1 -1.0346054 -0.69328933
## poly(age, 3)2 -0.2979551  0.04336095
## poly(age, 3)3 -0.2441917  0.09712441
```

```r
old_par <- par(mfrow=c(2,2))
plot(plm_fit)
```



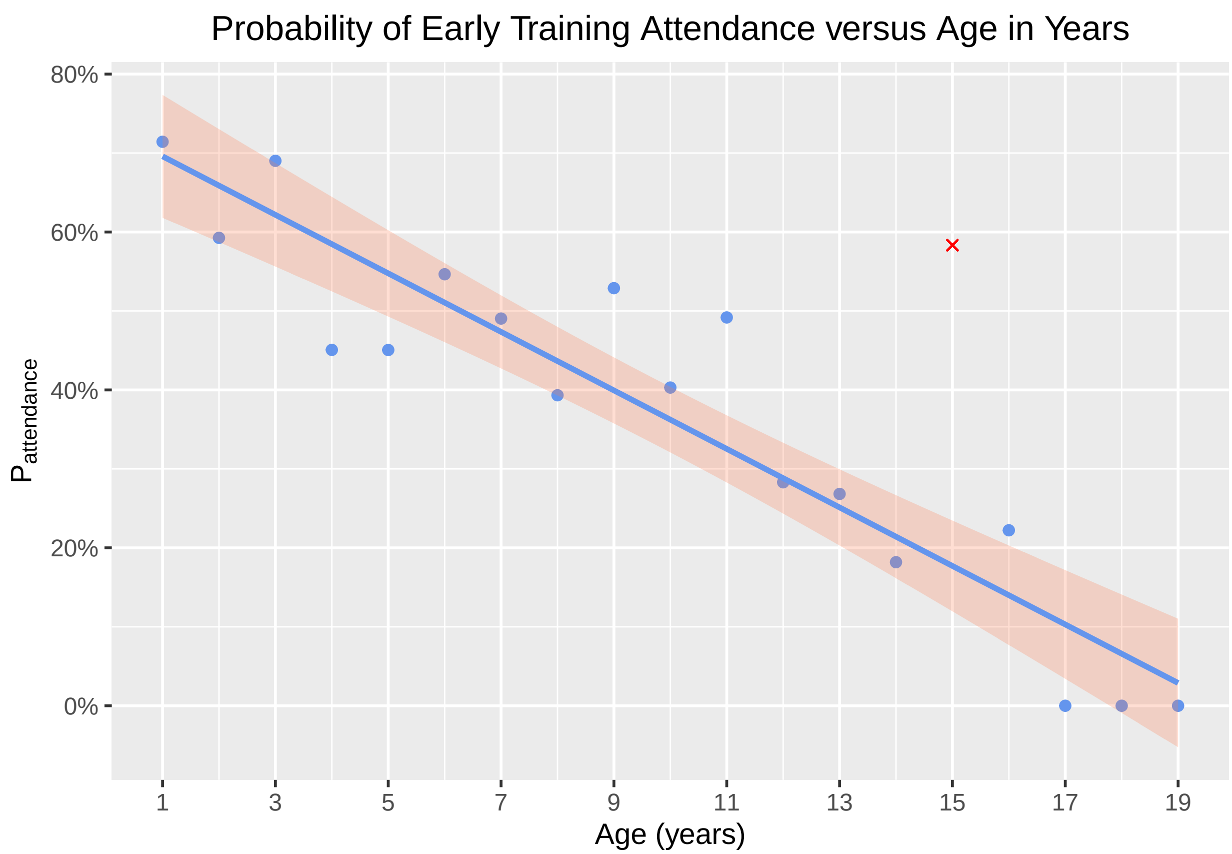

```
par(old_par)
```

The slight increase in explained variance from the polynomial model does not justify the significant increase in model complexity.

Visualizing the Trend

Now that we have the age and probability of attendance in a data frame and have verified a significant fit, let's use it to create out visual.

```
df_age %>%
  ggplot(aes(x=age, y=p_attend)) +
  geom_point(col='cornflowerblue') +
  stat_smooth(method='lm', col='cornflowerblue', fill='coral',
              formula=y~poly(x, 1), alpha=.25) +
  labs(x='Age (years)', y=expression('P'[attendance])) +
  scale_y_continuous(labels=scales::percent) +
  geom_point(x=outlier$age, y=outlier$p_attend, colour='red', shape=4) +
  scale_x_continuous(breaks=seq(1, 20, by=2)) +
  ggtitle('Probability of Early Training Attendance versus Age in Years') +
  theme(plot.title=element_text(hjust = 0.5))
```



Impact of Age on Problematic Jumping

We saw in our inferential analysis that increased age was correlated with a decreated probability for jumping up. We want to visualize this trend, so let's create a data frame with the probability of jumping up for each year of age.

```
p_jump_vec <- NULL
for (i in 1:max_age) {
  df_tmp <- df %>%
    filter(age_yrs == i) %>%
    select(jumping)
  p <- sum(df_tmp$jumping)/length(df_tmp$jumping)
  p_jump_vec <- c(p_jump_vec, p)
}

df_p_jump <- data.frame(age=c(1:max_age), p_jump=p_jump_vec)
head(df_p_jump)
```

```
##   age    p_jump
## 1   1 0.3928571
## 2   2 0.4197531
## 3   3 0.3380282
## 4   4 0.1971831
## 5   5 0.2857143
## 6   6 0.2325581
```

Now we create a visualization for this trend.

```r
df_p_jump %>%
  ggplot(aes(x=age, y=p_jump)) +
  geom_point(col='cornflowerblue') +
  stat_smooth(method='lm', col='cornflowerblue', fill='coral',
              formula=y~poly(x, 1), alpha=.25) +
  labs(x='Age (years)', y=expression('P'[jumping])) +
  scale_y_continuous(labels=scales::percent) +
  geom_point(x=outlier$age, y=outlier$p_attend, colour='red', shape=4) +
  scale_x_continuous(breaks=seq(1, 20, by=2)) +
  ggtitle('Probability of Problematic Jumping versus Age in Years') +
  theme(plot.title=element_text(hjust = 0.5))
```