Breeds impact on adoption

Matthew Wilcox, Randy Otoo, Anthony Barrios, Arianne PozziBarba

Libraries

```
library(tidyverse)
```

```
## Warning: package 'ggplot2' was built under R version 4.3.2

## Warning: package 'dplyr' was built under R version 4.3.2

## Warning: package 'stringr' was built under R version 4.3.2

## Warning: package 'lubridate' was built under R version 4.3.2

## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## v purrr     1.0.2
## -- Conflicts ------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

```
library(knitr)
```

```
## Warning: package 'knitr' was built under R version 4.3.2
```

```
library(kableExtra)
```

```
## Warning: package 'kableExtra' was built under R version 4.3.2

##
## Attaching package: 'kableExtra'
##
## The following object is masked from 'package:dplyr':
##
##     group_rows
```

```
library(Rgraphviz)
```

```
## Loading required package: graph
## Loading required package: BiocGenerics
##
## Attaching package: 'BiocGenerics'
##
## The following objects are masked from 'package:lubridate':
##
##     intersect, setdiff, union
```

```
##
## The following objects are masked from 'package:dplyr':
##
##     combine, intersect, setdiff, union
##
## The following objects are masked from 'package:stats':
##
##     IQR, mad, sd, var, xtabs
##
## The following objects are masked from 'package:base':
##
##     anyDuplicated, aperm, append, as.data.frame, basename, cbind,
##     colnames, dirname, do.call, duplicated, eval, evalq, Filter, Find,
##     get, grep, grepl, intersect, is.unsorted, lapply, Map, mapply,
##     match, mget, order, paste, pmax, pmax.int, pmin, pmin.int,
##     Position, rank, rbind, Reduce, rownames, sapply, setdiff, sort,
##     table, tapply, union, unique, unsplit, which.max, which.min
##
##
## Attaching package: 'graph'
##
## The following object is masked from 'package:stringr':
##
##     boundary
##
## Loading required package: grid
```

Import and Clean Data

```r
animal_intake.df <- read_csv("Louisville_Metro_ky_-_Animal_Service_Intake_and_Outcome.csv")
```

```
## Rows: 54494 Columns: 18
## -- Column specification ----------------------------------------------------
## Delimiter: ","
## chr (17): kennel, animalid, jurisdiction, intype, insubtype, indate, surreas...
## dbl  (1): ObjectId
##
## i Use `spec()` to retrieve the full column specification for this data.
## i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```r
dog_intake.df <- animal_intake.df %>% filter(animaltype == "DOG") %>%
  mutate(indate = str_remove(indate, " .*"),
         outdate = str_remove(outdate, " .*"),
         indate = as.Date(indate, format = "%Y/%m/%d"),
         outdate = as.Date(outdate, format = "%Y/%m/%d"),
         timeHeld = outdate - indate
         )%>%
  filter(intype == "STRAY",
         outtype != "DISPOSAL") %>%
  select(outtype, sex, bites, petsize, color, breed, timeHeld)
```

Table 1: Frequency Table of Pitbull and Adoption

	Is Pitbull	Not Pitbull
ADOPTION	1455	3147
EUTH	336	334

Probability

```r
pitbull_adopt_euth <- dog_intake.df %>%
  select(outtype, breed) %>%
  filter(
    outtype == "ADOPTION" | outtype == "EUTH"
  ) %>%
  mutate(breed = case_when(
    grepl("PIT BULL", breed) ~ "Is Pitbull",
    TRUE ~ "Not Pitbull"
  ))%>% group_by(outtype, breed) %>%
  summarise(n = n()) %>%
  pivot_wider(names_from = breed, values_from = n) %>%
  column_to_rownames(., var = "outtype")
```

```
## `summarise()` has grouped output by 'outtype'. You can override using the
## `.groups` argument.
```

```r
pitbull_adopt_euth %>% kbl(caption = "Frequency Table of Pitbull and Adoption") %>%  kable_classic_2(ful
```

```r
# if (!require("BiocManager", quietly = TRUE))
#     install.packages("BiocManager")
#
# BiocManager::install("Rgraphviz")
library("Rgraphviz")
```

```
## [1] "P"                    "Is Pitbull"
## [3] "Not Pitbull"          "IS Pitbull & Adopted"
## [5] "Is Pitbull & Euthanized'" "Not Pitbull & Adopted"
## [7] "Not Pitbull & Euthanized"

## $P
## [1] "Is Pitbull"  "Not Pitbull"
##
## $`Is Pitbull`
## [1] "IS Pitbull & Adopted"    "Is Pitbull & Euthanized'"
##
## $`Not Pitbull`
## [1] "Not Pitbull & Adopted"    "Not Pitbull & Euthanized"
##
## $`IS Pitbull & Adopted`
## character(0)
##
## $`Is Pitbull & Euthanized'`
## character(0)
##
## $`Not Pitbull & Adopted`
## character(0)
```

```
## 
## $`Not Pitbull & Euthanized`
## character(0)
```
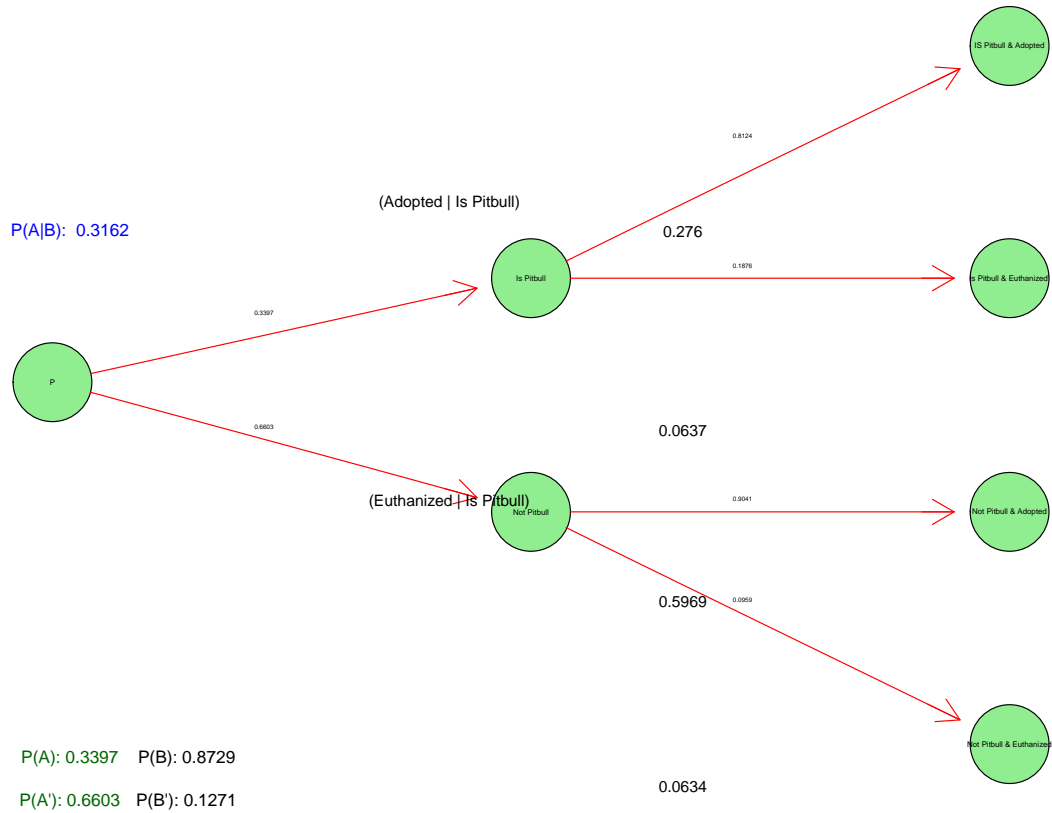
P(A|B): 0.3162

(Adopted | Is Pitbull)

0.276

0.3397

0.0637

0.6603

(Euthanized | Is Pitbull)

Is Pitbull & Adopted

Is Pitbull & Euthanized

Not Pitbull & Adopted

0.5969

Not Pitbull & Euthanized

P(A): 0.3397   P(B): 0.8729

P(A'): 0.6603   P(B'): 0.1271

0.0634

Figure 1: Your Graph Caption

https://www.harrysurden.com/wordpress/archives/292 refrenced to help build the tree diagram.