Final project

Kerinna Good and Josh Wagner

Introduction and Data

```
library(tidyverse)
library(tidymodels)
library(survival)

acquisitions <- read_csv("data/acquisitions.csv")
all_cetaceans <- read_csv("data/allCetaceanData.csv")
survival_rates <- read_csv("data/survivalRates.csv")
```

```
all_cetaceans2 <- all_cetaceans %>%
  rename(transferdate = transferDate,
         transfertype = transfer,
         currentlocation = currently) %>%
  mutate(statusDate2 = ifelse(status == "Alive", "2017-05-07",
                              all_cetaceans$statusDate)) %>%
  transform(statusyear = as.numeric(substr(statusDate2, 1, 4))) %>%
  mutate(age = as.numeric(statusyear - birthYear),
         originyear = as.numeric(substr(originDate, 1, 4)),
         captivity = (statusyear - originyear) / age,
         status2 = case_when(status %in% c("Miscarriage", "Escaped (Found Dead)",
                                           "Died", "Stillbirth") ~ 1,
                             status %in% c("Escaped", "Released", "Released?",
                                           "Unknown", "Alive") ~ 0),
         acquisition2 = case_when(acquisition %in% c("Born", "Miscarriage",
                                                     "Stillbirth") ~ "Born",
                                  acquisition %in% c("Capture", "Rescue") ~
                                    "Wild",
                                  acquisition %in% c("Unknown") ~ "Unknown"),
         species2 = ifelse(species == "Bottlenose", "Bottlenose", "Other"),
```

```
         error1 = ((originyear - birthYear) < 0),
         error2 = ((statusyear - birthYear) < 0),
         transfer = ifelse(originLocation != currentlocation, 1, 0)) %>%
  filter(error1 == FALSE, error2 == FALSE) %>%
  subset(select = -c(id, name, mother, father, region, COD, ...1, entryDate,
                     notes, statusDate, transfers, error1, error2, status,
                     acquisition, species))

all_cetaceans2 <- all_cetaceans2[, c(16, 1, 2, 15, 3, 4, 12, 5, 6, 17, 8, 14, 9,
                                     10, 11, 13)]
```

Data dictionary for `cetacean_deaths` data set

Variable	Class	Description
species	character	Species of animal (whale or dolphin)
sex	character	Sex of animal
accuracy	character	Accuracy of an animals' birthdate
acquisition	character	Method through which an animal was brought into captivity
birthYear	integer	Year of birth
origindate	integer	Date that an animal was born into captivity, rescued, or captured
originyear	integer	Year that an animal was born into captivity, rescued, or captured
originlocation	character	Location that an animal originated from
transfers	character	List of facilities an animal was transferred between
transferdate	integer	Date of transfer into the US; NA for animals born in the US
transfertype	character	Whether an animal was transferred within the US or outside the US
currentlocation	character	Location of animal at date of recorded status
status	character	Whether an animal was dead or alive at date of recorded status
status2	binary	Binary indicator for status (dead or alive)
statusDate2	integer	Date of recorded status
statusyear	integer	Year of recorded status
age	integer	Age of animal at date of recorded status
captivity	integer	Percentage of animal's life lived in captivity

Our data comes from and article in The Pudding titled "Free Willy and Flipper by the Numbers.

Methodology

```
coxph(formula = Surv(birthYear, statusyear, status2) ~ species2 + sex + accuracy +
        acquisition2 + captivity + transfertype, data = all_cetaceans2)
```

```
Warning in Surv(birthYear, statusyear, status2): Stop time must be > start time,
NA created
```

```
Call:
coxph(formula = Surv(birthYear, statusyear, status2) ~ species2 +
    sex + accuracy + acquisition2 + captivity + transfertype,
    data = all_cetaceans2)

                          coef  exp(coef)   se(coef)       z      p
species2Other       -3.207e-01  7.256e-01  1.294e-01  -2.479 0.0132
sexM                 1.636e-01  1.178e+00  9.177e-02   1.783 0.0746
sexU                -2.035e-01  8.158e-01  2.817e-01  -0.723 0.4700
accuracye           -9.956e-01  3.695e-01  1.006e+00  -0.989 0.3225
acquisition2Unknown  1.969e+00  7.167e+00  1.068e+00   1.844 0.0652
acquisition2Wild     1.325e+00  3.761e+00  1.005e+00   1.319 0.1873
captivity           -8.786e+01  6.992e-39  7.732e+00 -11.363 <2e-16
transfertypeUS       2.712e-01  1.312e+00  7.197e-01   0.377 0.7063

Likelihood ratio test=177  on 8 df, p=< 2.2e-16
n= 1003, number of events= 519
   (1 observation deleted due to missingness)
```

```
#coxph(formula = Surv(birthYear, statusyear, status2) ~ species, data = all_cetaceans2)
```

Results

Discussion