PreCheck Donkey data - Judas efficiency

Carlo Pacioni

Load libraries

```
# Load libraries
options(java.parameters = "-Xmx8024m")
library(XLConnect, quietly = T)
```

```
## XLConnect 0.2-15 by Mirai Solutions GmbH [aut],
##   Martin Studer [cre],
##   The Apache Software Foundation [ctb, cph] (Apache POI),
##   Graph Builder [ctb, cph] (Curvesapi Java library)
```

```
## http://www.mirai-solutions.com
## https://github.com/miraisolutions/xlconnect
```

```
library(ggplot2, quietly = T)
library(data.table, quietly = T)
```

Read data

```
data.path <- "../Data/"
judas.master <- data.table(readWorksheetFromFile(
  file = file.path(data.path, "Tracking_History_Judas_with habitat_MZ.xlsx"),
                       sheet="Judas_Tracking_History"))
names(judas.master)
```

```
##  [1] "Tracking_ID"  "REGION"       "SHIRE"        "AREA"
##  [5] "JUDAS_ID"     "EVENT_DATE"   "LAT"          "LONG"
##  [9] "SEARCHED"     "EVENT_ID"     "EVENT"        "N_FERALS"
## [13] "N_JUDAS"      "ACTION"       "Habitat.Type" "LAT_ORG"
## [17] "LONG_ORG"
```

Check formatting

```
sapply(judas.master, class)
```

```
## $Tracking_ID
## [1] "numeric"
##
## $REGION
## [1] "character"
##
## $SHIRE
## [1] "character"
##
## $AREA
## [1] "character"
##
## $JUDAS_ID
## [1] "character"
##
## $EVENT_DATE
## [1] "POSIXct" "POSIXt"
##
```

```
## $LAT
## [1] "numeric"
##
## $LONG
## [1] "character"
##
## $SEARCHED
## [1] "character"
##
## $EVENT_ID
## [1] "numeric"
##
## $EVENT
## [1] "character"
##
## $N_FERALS
## [1] "numeric"
##
## $N_JUDAS
## [1] "numeric"
##
## $ACTION
## [1] "character"
##
## $Habitat.Type
## [1] "character"
##
## $LAT_ORG
## [1] "numeric"
##
## $LONG_ORG
## [1] "character"
```

LONG is character while it should be numeric. Checked what values are not numeric.

```
judas.master[is.na(sapply(judas.master[, LONG], as.numeric)), ]
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion

## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
##    Tracking_ID
## 1:       20149
## 2:       24575
##
## 1: KIMBERLEY
## 2: KIMBERLEY
##    SHIRE        AREA JUDAS_ID EVENT_DATE       LAT LONG SEARCHED EVENT_ID
## 1:    HC Mabel Downs     MD03 2003-08-15 -17.20017 NULL      YES    16788
## 2:    HC   Nicholson    NIC06 2003-11-08 -17.01800 NULL      YES    20025
##    EVENT N_FERALS N_JUDAS ACTION Habitat.Type   LAT_ORG LONG_ORG
## 1: FERAL        8       1  FREED         NULL -17.20017     NULL
## 2: ALONE        0       1  FREED         NULL -17.01800     NULL
```

```
# Number dead
judas.master[, sum(ACTION == "DEAD")]
```

```
## [1] 7
```

```
# Number collared
judas.master[, sum(ACTION == "COLLARED")]
```

```
## [1] 1721
```

```
# Number of judas in DB
judas.master[, length(unique(JUDAS_ID))]
```

```
## [1] 1721
```

```
# Rm useless spaces in REGION
judas.master[, REGION := sub(pattern = " +", replacement = "", x = REGION)]
# Number by region/shire
judas.master[, .(njudas=length(unique(JUDAS_ID)),
                 start.date=min(EVENT_DATE), end.date=max(EVENT_DATE)), by=REGION]
```

```
##       REGION njudas start.date   end.date
## 1: KIMBERLEY   1385 1994-07-12 2017-11-14
## 2:   PILBARA    336 1998-05-20 2017-11-03
```

```
judas.master[, .(njudas=length(unique(JUDAS_ID)),
                 start.date=min(EVENT_DATE), end.date=max(EVENT_DATE)),
                 by=c("REGION", "SHIRE")]
```

```
##       REGION SHIRE njudas start.date   end.date
## 1: KIMBERLEY    EK    513 1997-10-21 2017-11-14
## 2: KIMBERLEY    WK    451 1994-07-12 2017-06-20
## 3:   PILBARA    PB    336 1998-05-20 2017-11-03
## 4: KIMBERLEY    HC    421 1995-10-26 2017-11-14
```

Check whether there are animals that moved between regions, shires or property during the study (it seems none, is this correct?)

```
# build a function where IDbyLoc is a data.table with one col reporting locations
   # (e.g. shire), and the second, named 'IDs', is the judas' ids within each location
check.migration <- function(IDbyLoc, location) {
  areas <- unique(IDbyLoc[[location]])
  nareas <- length(areas)
  # Unique IDs
  un.IDs <- IDbyLoc[, unique(IDs)]
  names(un.IDs) <- IDbyLoc[, unique(IDs)]
  setkeyv(IDbyLoc, location)
  l2 <- vector(mode = "list", length = nareas)
  # For each area checks what ID are present
  for (a in areas) {
    l2[[a]] <- vector(length = length(un.IDs))
    names(l2[[a]]) <- un.IDs
    for (u in un.IDs) {
      l2[[a]][u] <- sum(IDbyLoc[a, IDs] %in% un.IDs[u]) > 0
    }
  }
  # Combined in a data.frame where locations are columns
  d <- do.call(cbind, args = l2)
  # sum row-wise, if the same ID is present in more than one location the sum>1
  keep.row <- apply(d, MARGIN = 1, sum)
  d <- cbind(un.IDs, data.table(d))
```

```
  keep.row <- keep.row > 1
  return(d[keep.row,])
}

# Create IDbyLoc
IDbyShire <- judas.master[, .(IDs=unique(JUDAS_ID)), by=SHIRE]
check.migration(IDbyShire, location = "SHIRE")
```

```
## Empty data.table (0 rows) of 5 cols: un.IDs,EK,WK,PB,HC
```

```
IDbyRegion <- judas.master[, .(IDs=unique(JUDAS_ID)), by=REGION]
check.migration(IDbyRegion, location = "REGION")
```

```
## Empty data.table (0 rows) of 3 cols: un.IDs,KIMBERLEY,PILBARA
```

```
IDbyArea <- judas.master[, .(IDs=unique(JUDAS_ID)), by=AREA]
check.migration(IDbyArea, location = "AREA")
```

```
## Empty data.table (0 rows) of 109 cols: un.IDs,Doongan,Mitchell Plateau,Prince Regent,Drysdale Nat Pa
```

Clean up and generating variables

```
# Rm found dead because they are out of the program
judas.cleaned <- judas.master[ACTION != "DEAD", ]

# Rm not found because they do not contribute
judas.cleaned <- judas.cleaned[ACTION != "NONE", ]

# Cross check totals are still the same
judas.cleaned[, .(njudas=length(unique(JUDAS_ID)),
                  start.date=min(EVENT_DATE), end.date=max(EVENT_DATE)),
              by=c("REGION", "SHIRE")]
```

```
##       REGION SHIRE njudas start.date   end.date
## 1: KIMBERLEY    EK    513 1997-10-21 2017-11-14
## 2: KIMBERLEY    WK    451 1994-07-12 2017-06-20
## 3:   PILBARA    PB    336 1998-05-20 2017-11-03
## 4: KIMBERLEY    HC    421 1995-10-26 2017-11-14
```

```
# Rm judas with no long
keep <- !is.na(sapply(judas.cleaned[, LONG], as.numeric))
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
## Warning in lapply(X = X, FUN = FUN, ...): NAs introduced by coercion
```

```
judas.cleaned <- judas.cleaned[keep, ]
judas.cleaned[is.na(sapply(judas.cleaned[, LONG], as.numeric)), ]
```

```
## Empty data.table (0 rows) of 17 cols: Tracking_ID,REGION,SHIRE,AREA,JUDAS_ID,EVENT_DATE...
```

```
judas.cleaned[, LONG := as.numeric(LONG)]

# Set start and end date
judas.cleaned[, start.date := min(EVENT_DATE), by=JUDAS_ID]
judas.cleaned[, end.date := max(EVENT_DATE), by=JUDAS_ID]

# Cross check start.date matches collared
```

```
judas.cleaned[ACTION == "COLLARED", date.coll := EVENT_DATE, by=JUDAS_ID]
judas.cleaned[ACTION == "COLLARED", sum(start.date != date.coll, na.rm = T)]
```

```
## [1] 3
```

Some animals have been collared after their start date. Is this a collar replacement?

```
judas.cleaned[start.date != date.coll, ]
```

```
##    Tracking_ID    REGION SHIRE        AREA JUDAS_ID EVENT_DATE       LAT
## 1:       18898 KIMBERLEY    WK     Leopold     LP11 1996-07-16 -17.58333
## 2:       19091 KIMBERLEY    WK     Leopold     LP22 2006-05-12 -17.81228
## 3:       19862 KIMBERLEY    WK Mt Anderson     MA10 1996-10-22 -17.99817
##        LONG SEARCHED EVENT_ID EVENT N_FERALS N_JUDAS   ACTION Habitat.Type
## 1: 125.2833      YES    15886 ALONE        0       1 COLLARED       HLLEWS
## 2: 125.4543      YES    16024 ALONE        0       1 COLLARED       UPLEWG
## 3: 123.8610      YES    16550 ALONE        0       1 COLLARED       SADPWG
##      LAT_ORG    LONG_ORG start.date   end.date  date.coll
## 1: -17.58333 125.2833333 1996-03-18 2000-12-12 1996-07-16
## 2: -17.81228   125.45425 1996-08-30 2010-09-16 2006-05-12
## 3: -17.99817     123.861 1996-07-17 1997-11-24 1996-10-22
```

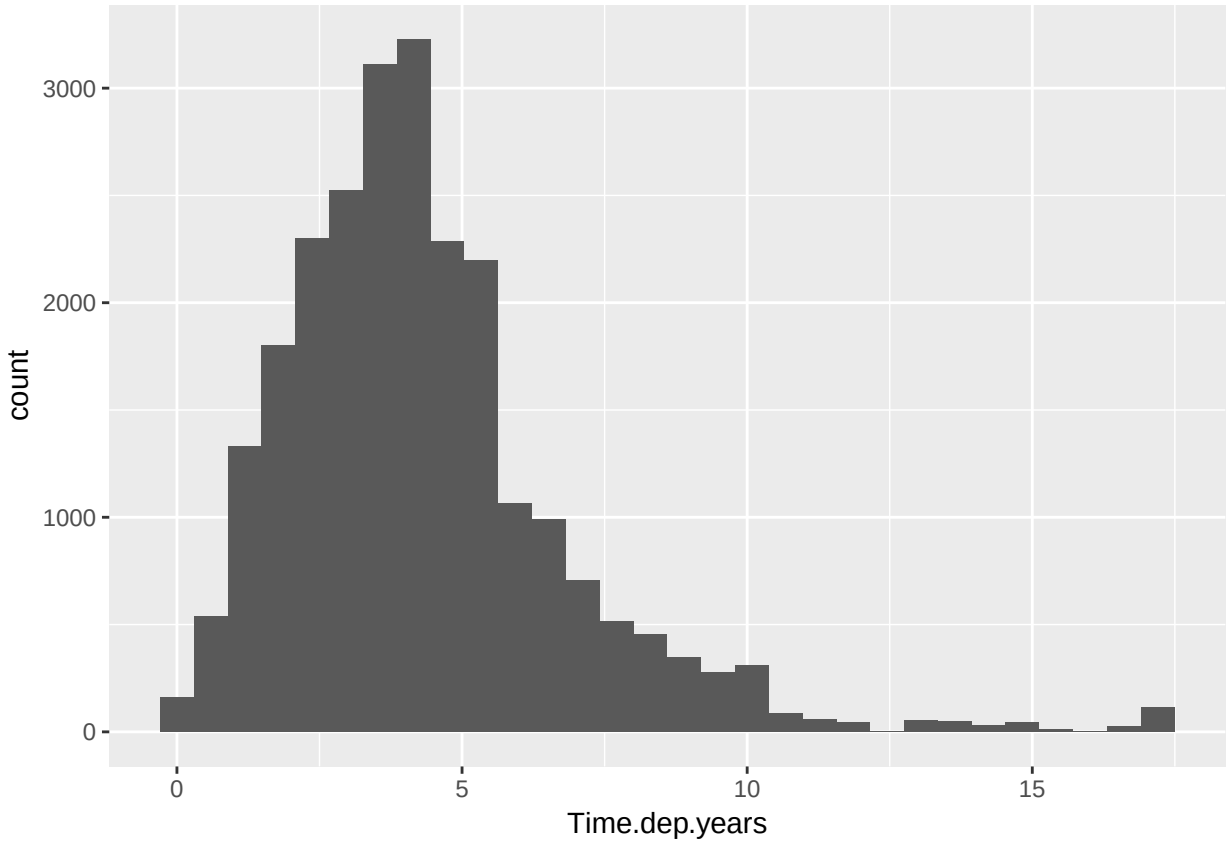
```
# Rm date.coll
judas.cleaned[, date.coll := NULL]
```

Let's have a look at the length in the program for each donkey"

```
# Length in the program
judas.cleaned[, Time.deployment := difftime(end.date, start.date, units="weeks")]
judas.cleaned[, Time.dep.years := round(as.numeric(Time.deployment) / 52, 2)]

ggplot(judas.cleaned, aes(Time.dep.years)) + geom_histogram()
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
summary(judas.cleaned[, Time.dep.years])
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   2.680   3.940   4.313   5.270  17.200
```

Some animals have end.date set on the day they have been collared (listed below). I'm assumed that these were animals that were collared on start.date and searched and not found, is this correct? Some other aniamls were tracked for several years, it is correct?

```
# List of animals that have been deployed for 0
judas.cleaned[Time.deployment == 0, .SD,
              .SDcols=c("SHIRE", "JUDAS_ID", "EVENT_DATE", "EVENT_ID", "EVENT",
                        "ACTION")]
```

```
##     SHIRE JUDAS_ID EVENT_DATE EVENT_ID       EVENT   ACTION
##  1:    WK    ANA06 2011-08-26      133 FERAL/JUDAS COLLARED
##  2:    PB    COR17 2017-08-24     4707       ALONE COLLARED
##  3:    EK     CR25 2004-08-04     5007       ALONE COLLARED
##  4:    WK    DEF25 2017-06-20     6289       ALONE COLLARED
##  5:    EK    DNP75 2013-10-07     7664       FERAL COLLARED
##  6:    EK    ELL17 2001-12-06     9048       ALONE COLLARED
##  7:    WK    ELZ14 2005-07-12     9863       FERAL COLLARED
##  8:    WK    ELZ19 2007-09-07     9923       FERAL COLLARED
##  9:    WK    GIB25 2005-07-12    11154       FERAL COLLARED
## 10:    WK    GIB36 2014-06-23    11267       FERAL COLLARED
## 11:    WK    GIB38 2017-06-20    11272       FERAL COLLARED
## 12:    PB  H-ETH01 2017-05-26    12256       ALONE COLLARED
## 13:    PB    HIL33 2005-05-10    12674       FERAL COLLARED
```

```
## 14:    WK     KM18 1996-09-24    13807       ALONE COLLARED
## 15:    WK     LA03 1996-10-30    14522       FERAL COLLARED
## 16:    PB    MEE04 2005-10-28    17165       ALONE COLLARED
## 17:    HC    ORD03 1998-11-12    20396       FERAL COLLARED
## 18:    HC   ORD06C 1998-11-14    20405       ALONE COLLARED
## 19:    HC    ORD18 1999-03-17    20575       FERAL COLLARED
## 20:    HC    ORD29 1999-10-09    20705       FERAL COLLARED
## 21:    HC   ORD39C 2005-05-18    20829       ALONE COLLARED
## 22:    HC    ORD46 2006-07-31    20878       FERAL COLLARED
## 23:    WK    PNT19 2017-06-20    21191       ALONE COLLARED
## 24:    WK    PNT20 2017-06-20    21192       FERAL COLLARED
## 25:    PB    ROY02 2016-11-04    21383       ALONE COLLARED
## 26:    HC     SO10 1997-05-08    22299       FERAL COLLARED
## 27:    HC     SO25 2003-05-08    22500       ALONE COLLARED
## 28:    HC    SPC13 2012-06-04    22677       FERAL COLLARED
## 29:    HC     TA17 2005-05-18    23058       ALONE COLLARED
## 30:    HC     TA21 2006-07-31    23080       FERAL COLLARED
## 31:    PB    VCL30 2017-08-24    23653       ALONE COLLARED
## 32:    PB    WAN07 2008-11-28    23890       FERAL COLLARED
## 33:    HC     WR07 2000-05-24    24109       FERAL COLLARED
## 34:    HC     WR28 2005-05-18    24366       ALONE COLLARED
## 35:    HC     WR32 2005-05-18    24395       ALONE COLLARED
## 36:    WK    XCK01 2003-08-06    24776       FERAL COLLARED
##     SHIRE JUDAS_ID EVENT_DATE EVENT_ID       EVENT   ACTION
```

```
judas.cleaned <- judas.cleaned[Time.deployment>0,]

# List of animals that have been deployed for > 10 yrs
judas.cleaned[Time.dep.years > 10,  unique(JUDAS_ID)]
```

```
##  [1] "BEV11" "HAR17" "HAR22" "BEV18" "COR12" "GIB17" "HAR23" "MHO02"
##  [9] "WAN03" "MX12"  "THE01" "GIB11" "MHO26" "LP22"  "MA01"  "MX03"
## [17] "MX10"  "NIC06" "NIC13" "ORD11" "DNP01" "DNP10" "MX04"  "NIC12"
## [25] "ORD37" "SPC06"
```

Check whether there are judas with < 5 data points and rm

```
locs <- judas.cleaned[, .N, by=JUDAS_ID]
locs[, sum(N<6)]
```

```
## [1] 295
```

```
IDs.rm <- locs[N<6, JUDAS_ID]
judas.cleaned <- judas.cleaned[!JUDAS_ID %in% IDs.rm, ]
```

Calculates the home range centre as arithmetic mean of coordinate and compute deviation from the centre on the X and Y

```
# Home Range centres
judas.cleaned[, ':='(HRlat=mean(LAT), HRlong=mean(LONG)), by=JUDAS_ID]

calc.latlong.dist<- function(xy1,xy2)
{
  # uses spherical law of cosines to calculate distance between two lat/long
  # coordinates in decimal degrees
  R <- 6371 # Earths radius
  xy1 <- (pi * xy1)/180 # radians
```

```
  xy2 <- (pi * xy2)/180
  D <- acos(sin(xy1[,1])*sin(xy2[,1]) + cos(xy1[,1])*cos(xy2[,1])*cos(xy2[,2]-xy1[,2]))
  return(R*D)
}

# Calculate deviations from HRcentres
judas.cleaned[, xdev:=calc.latlong.dist(judas.cleaned[, .(LAT, HRlong)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, ydev:=calc.latlong.dist(judas.cleaned[, .(HRlat, LONG)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, summary(xdev)]
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
##    0.000    1.349    3.058    7.304    5.987 1929.792
```

```
judas.cleaned[, summary(ydev)]
```

```
##      Min.   1st Qu.    Median      Mean   3rd Qu.      Max.
##     0.000     1.383     3.138    23.322     6.214 12695.024
```

Something is clearly wrong as some animals have moved ~ 2,000 km north or south and >12,000 km east or west. Let's check what happened:

```
judas.cleaned[ydev>12000,]
```

```
##     Tracking_ID    REGION SHIRE          AREA JUDAS_ID EVENT_DATE LAT
##  1:        1413 KIMBERLEY    HC Bedford Downs     BE25 2005-05-18   0
##  2:       22304 KIMBERLEY    WK    Mornington    MOR06 2005-07-27   0
##  3:       24661 KIMBERLEY    HC     Nicholson    NIC09 2005-07-27   0
##  4:       24663 KIMBERLEY    HC     Nicholson    NIC10 2005-05-18   0
##  5:       25506 KIMBERLEY    HC   Ord Reserve    ORD30 2005-05-18   0
##  6:       25766 KIMBERLEY    HC Osmond Valley     OV02 2005-05-18   0
##  7:       28327 KIMBERLEY    HC     Tableland     TA04 2006-04-14   0
##  8:       28378 KIMBERLEY    HC     Tableland     TA10 2005-05-18   0
##  9:       28422 KIMBERLEY    HC     Tableland     TA18 2005-05-18   0
## 10:       29258 KIMBERLEY    HC Violet Valley     VV09 2012-06-04   0
##      LONG SEARCHED EVENT_ID EVENT N_FERALS N_JUDAS   ACTION Habitat.Type
##  1: 0e+00      YES     2535 ALONE        0       1 COLLARED         NULL
##  2: 0e+00      YES    18381 FERAL        1       1     SHOT         NULL
##  3: 0e+00      YES    20094 FERAL        7       1    FREED         NULL
##  4: 0e+00      YES    20096 ALONE        0       1 COLLARED         NULL
##  5: 0e+00      YES    20719 ALONE        0       1    FREED         NULL
##  6: 0e+00      YES    20921 ALONE        0       1 COLLARED         NULL
##  7: 3e-05      YES    22972 FERAL        8       1    FREED         NULL
##  8: 0e+00      YES    23018 ALONE        0       1    FREED         NULL
##  9: 0e+00      YES    23059 ALONE        0       1 COLLARED         NULL
## 10: 0e+00      YES    23754 ALONE        0       1    FREED         NULL
##     LAT_ORG LONG_ORG start.date   end.date Time.deployment Time.dep.years
##  1:       0        0 2005-05-18 2007-08-14  116.8571 weeks           2.25
##  2:       0        0 2000-01-20 2005-07-27  287.8631 weeks           5.54
##  3:       0        0 2002-11-11 2006-02-13  170.0000 weeks           3.27
##  4:       0        0 2005-05-18 2009-09-21  226.7143 weeks           4.36
##  5:       0        0 2000-06-14 2009-09-22  483.8571 weeks           9.30
##  6:       0        0 2005-05-18 2008-12-02  184.8512 weeks           3.55
##  7:       0 3.00E-05 2001-04-17 2007-08-14  330.0000 weeks           6.35
##  8:       0        0 2003-08-17 2008-10-10  268.7083 weeks           5.17
```

```
##  9:       0        0 2005-05-18 2009-09-14  225.7143 weeks           4.34
## 10:       0        0 2009-06-22 2015-06-15  312.0000 weeks           6.00
##         HRlat   HRlong     xdev     ydev
##  1: -15.98715 116.5202 1777.690 12195.98
##  2: -16.08704 119.8453 1788.797 12509.31
##  3: -15.85523 114.5482 1763.021 12014.34
##  4: -16.44344 119.0143 1828.427 12395.04
##  5: -17.35504 123.1675 1929.792 12695.02
##  6: -15.78906 117.6339 1755.663 12322.70
##  7: -16.19924 118.3331 1801.273 12352.62
##  8: -15.61380 115.5262 1736.175 12131.67
##  9: -15.90962 116.3858 1769.069 12189.84
## 10: -15.60577 116.4137 1735.283 12219.69
```

Okay some animals have coordinates 0 (somewhere in central Africa!). I guess that these were meant to be NA, removed for now and check again.

```
judas.cleaned <- judas.cleaned[LAT < 0,]
judas.cleaned[, ':='(HRlat=mean(LAT), HRlong=mean(LONG)), by=JUDAS_ID]
judas.cleaned[, xdev:=calc.latlong.dist(judas.cleaned[, .(LAT, HRlong)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, ydev:=calc.latlong.dist(judas.cleaned[, .(HRlat, LONG)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, summary(xdev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.336   3.025   4.710   5.890 536.515
```

```
judas.cleaned[, summary(ydev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.370   3.106   4.948   6.102 303.858
```

This is better, but still the distibution is very odd... 75% moved less than 6 km but max value is 300 or 500 km?

```
judas.cleaned[xdev > 200,]
```

```
##    Tracking_ID    REGION SHIRE        AREA JUDAS_ID EVENT_DATE       LAT
## 1:         182 KIMBERLEY    HC Alice Downs     AL08 2001-08-27 -13.02167
## 2:         205 KIMBERLEY    HC Alice Downs     AL10 2001-10-05 -12.86583
## 3:        1906 KIMBERLEY    WK     Bohemia     BO01 1998-10-20 -16.92050
## 4:       12573 KIMBERLEY    WK  Gibb River    GIB01 2001-07-25 -19.50611
## 5:       22313 KIMBERLEY    WK  Mornington    MOR07 2003-08-13 -12.02733
## 6:       27873 KIMBERLEY    HC   Spring Ck    SPC03 2002-09-30 -18.83867
## 7:       29873 KIMBERLEY    HC  Wood River     WR08 2003-02-23 -18.55600
##        LONG SEARCHED EVENT_ID EVENT N_FERALS N_JUDAS ACTION Habitat.Type
## 1: 127.7625      YES     1554 ALONE        0       1  FREED         NULL
## 2: 128.1767      YES     1576 ALONE        0       1  FREED         NULL
## 3: 127.9307      YES     2892 ALONE        0       1   SHOT       HRPEWS
## 4: 126.3203      YES    10662 FERAL       15       1  FREED        LRBNV
## 5: 126.2355      YES    18389 FERAL        6       1  FREED         NULL
## 6: 128.8273      YES    22580 FERAL        3       1  FREED       SADASS
## 7: 127.1332      YES    24123 FERAL        8       1  FREED       HLLEWS
##      LAT_ORG    LONG_ORG start.date   end.date Time.deployment
## 1: -13.02167    127.7625 1997-11-17 2001-08-27  197.0060 weeks
## 2: -12.86583 128.1766667 2001-05-30 2003-05-08  101.1429 weeks
```

```
## 3: -16.92050 127.9306667 1996-11-01 1998-10-20  102.5774 weeks
## 4: -19.50611 126.3202778 2000-07-20 2005-06-10  255.1429 weeks
## 5: -12.02733    126.2355 2000-12-21 2007-08-14  346.7202 weeks
## 6: -18.83867 128.8273333 2000-05-20 2003-05-09  154.8571 weeks
## 7: -18.55600 127.1331667 2000-05-25 2003-11-09  180.4226 weeks
##    Time.dep.years     HRlat   HRlong     xdev        ydev
## 1:           3.79 -17.80583 127.7533 531.9751   0.9717219
## 2:           1.95 -17.44817 128.1956 509.5322   2.0095998
## 3:           1.97 -18.73713 126.3842 202.0003 162.8414946
## 4:           4.91 -16.63870 126.2741 318.8421   4.9223408
## 5:           6.67 -16.85233 126.2954 536.5150   6.3746490
## 6:           2.98 -17.03041 128.8208 201.0689   0.6937959
## 7:           3.47 -16.74767 127.1376 201.0775   0.4761914
```

```
judas.cleaned[, summary(LAT)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##  -22.62  -18.26  -17.07  -17.47  -15.98  -12.03
```

Some donkey have LAT<-14 and are in the middle of the Timor sea! Other have done a sudden move of >200 km... is that possible?

For now I keep only entries where xdev and ydev is less than ~20 km

```
judas.cleaned <- judas.cleaned[xdev < 20,]
judas.cleaned[, ':='(HRlat=mean(LAT), HRlong=mean(LONG)), by=JUDAS_ID]
judas.cleaned[, xdev:=calc.latlong.dist(judas.cleaned[, .(LAT, HRlong)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, ydev:=calc.latlong.dist(judas.cleaned[, .(HRlat, LONG)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, summary(xdev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.262   2.834   3.901   5.455  26.140
```

```
judas.cleaned[, summary(ydev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.315   3.012   4.577   5.867 303.858
```

```
judas.cleaned <- judas.cleaned[ydev < 20,]
judas.cleaned[, ':='(HRlat=mean(LAT), HRlong=mean(LONG)), by=JUDAS_ID]
judas.cleaned[, xdev:=calc.latlong.dist(judas.cleaned[, .(LAT, HRlong)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, ydev:=calc.latlong.dist(judas.cleaned[, .(HRlat, LONG)],
                                        judas.cleaned[, .(HRlat, HRlong)])]
judas.cleaned[, summary(xdev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.232   2.790   3.844   5.392  26.685
```

```
judas.cleaned[, summary(ydev)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##   0.000   1.218   2.777   3.852   5.327  23.810
```

Save final dataset and csv with sample size

```
dir.create(file.path(data.path, "Analysis"))
```

```
## Warning in dir.create(file.path(data.path, "Analysis")): '..\Data\
## \Analysis' already exists
```

```
save(judas.cleaned, file = file.path(data.path, "judas.cleaned.rda"))
descr.fin <- judas.cleaned[, .(njudas=length(unique(JUDAS_ID)),
                 start.date=min(EVENT_DATE), end.date=max(EVENT_DATE)),
                 by=c("REGION", "SHIRE")]
descr.fin
```

```
##       REGION SHIRE njudas start.date   end.date
## 1: KIMBERLEY    EK    433 1997-10-21 2017-11-14
## 2: KIMBERLEY    WK    376 1994-07-12 2017-06-20
## 3:   PILBARA    PB    253 1998-05-20 2017-11-01
## 4: KIMBERLEY    HC    323 1995-10-26 2017-06-20
```

```
descr.fin[, sum(njudas)]
```

```
## [1] 1385
```

```
write.csv(descr.fin, file = file.path(data.path, "Analysis", "descr.fin.csv"),
          row.names = F)

ntrack.events <- judas.cleaned[, .N, by=JUDAS_ID]
ntrack.events[, summary(N)]
```

```
##    Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##    1.00   10.00   15.00   16.45   21.00   60.00
```