# Maxwell Final

## Robert Maxwell

## 12/9/2020

```r
library(tidyverse)
```

```
## -- Attaching packages -------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
```

```
## -- Conflicts ----------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(lubridate) #to fix my dates
```

```
##
## Attaching package: 'lubridate'
```

```
## The following objects are masked from 'package:base':
##
##     date, intersect, setdiff, union
```

```r
library(readxl) #to read XLS files
library(biogeo) #to convert from degrees to decimal lat/long
```

## Loading Dataset and cleaning dates

I hand coded my data from the 1967 report, codenamed "Red Baron" that covered every air to air encounter over North Vietnam. I coded this in Excel, which allows for an ease of legibility and use-friendlyness. However, because the majority of work on this project will be done in QGIS, which plays much better with .CSV Files, I will read the file from the .xls, and use lubridate to fix the dates, and output a CSV file.

```r
nv_dataset <- read_xlsx("maxwell_red_baron.xlsx")
#now that I have my dataframe, I will use "lubridate" to fix the dates
nv_dataset$Date <- as_date(nv_dataset$Date)
```

## Making my lat/long workable

My chosen GIS program QGIS only accepts Lat/long in decimal form. My source is only in degree form. . . ..I will now mutate two new columns to make workable decimal coordinates. I will then select out the original imputs, leaving a cleaner decimal lat and long.

```r
nv_dataset<- nv_dataset %>%
  mutate(lat_dec = dms2dd(Lat, 'Lat Min', 0, 'Lat Card')) %>%
  mutate(long_dec = dms2dd(Long, 'Long Min', 0, 'Lat Card'))
nv_dataset<- nv_dataset %>%
  select(-Lat, -'Lat Min', -'Lat Card', -Long, -'Long Min', -'Long Card')
structure(nv_dataset)
```

```
## # A tibble: 40 x 15
##    Date       'Route Pack' 'US AC Number' 'US AC Type' 'NVAP AC Number'
##    <date>     <chr>               <dbl> <chr>        <chr>
##  1 1965-04-03 4                       4 F-8          3
##  2 1965-04-09 Yankee                  4 F-4          4
##  3 1965-06-04 5                       2 F-4          4
##  4 1965-06-17 5                       2 F-4          4
##  5 1965-07-10 6A                      4 F-4          2
##  6 1965-10-06 6B                      2 F-4          3
##  7 1965-12-23 1                       2 F-4          2
##  8 1966-04-23 6A                      4 F-4          4
##  9 1966-04-25 5                       2 F-4          2
## 10 1966-04-30 5                       2 F-4          4
## # ... with 30 more rows, and 10 more variables: 'NVAP Airfraft Type' <chr>, 'US
## #   Pilot Branch' <chr>, 'US AC Damaged' <dbl>, 'US AC Destroyed' <dbl>, 'NVAF
## #   AC Damaged' <dbl>, 'NVAF AC Destroyed' <dbl>, 'US Weapon Used' <chr>, 'NVAF
## #   Weapon Used' <chr>, lat_dec <dbl>, long_dec <dbl>
```

## Now Making my CSV File

```r
write_csv(nv_dataset, "nv_dataset.csv")
```