Analysis of disgnostic referral completion

```
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

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

```
library(ggplot2)
```

Intro: I have implemented a method to analyze diagnostic referrrals. I output events to a file whenever an agent loses their diagnostic referral (by expiration or completion, deaths are excluded for simplicity).

This file records the timestep, length of referral at the time of referral loss, whether the referral expired, and whether the length from referral to diagnosis is less than or equal to 2 months/60 days.

In this document I will show the results from analyzing the original model parameters as well as a modified version of the model which increases the probability of diagnostic test completion in months 1 and 2 of the referral.

Loading in data:

```
bc_navigation_root <- '/project2/khanna7/bryanb/bc-navigation/dec9_navlength/bc-navigation/'
old_referral_data <- read.table(paste0(bc_navigation_root,"1a.dtdata"))
colnames(old_referral_data)<-c("time","diagnostic_referral_length","expired","within_2_months")

referral_data <- read.table(paste0(bc_navigation_root,"1.dtdata"))
colnames(referral_data)<-c("time","diagnostic_referral_length","expired","within_2_months")
```

Taking a peek at the data format:

```
old_referral_data[1:15,]
```

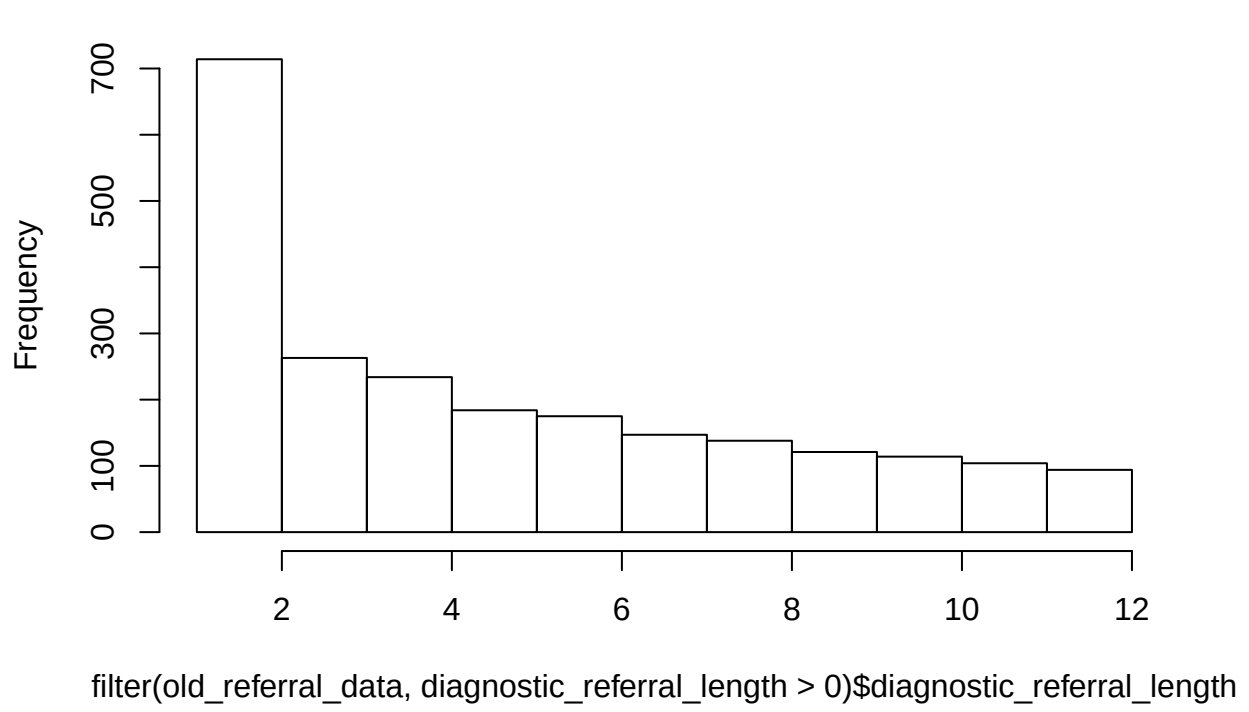
```
##    time diagnostic_referral_length expired within_2_months
## 1     1                          1       0               1
## 2     1                          1       0               1
## 3     1                          1       0               1
## 4     1                          1       0               1
## 5     1                          1       0               1
## 6     1                          1       0               1
## 7     1                          1       0               1
## 8     1                          1       0               1
## 9     2                          2       0               1
## 10    2                          2       0               1
## 11    2                          2       0               1
## 12    2                          2       0               1
## 13    2                          2       0               1
## 14    3                          3       0               0
## 15    3                          3       0               0
```

Note that we see 3570 and 3765 rows, corresponding to 3570/3765 agents losing their diagnostic referral through completion or expiration. The model was run for 360 months / 30 years.

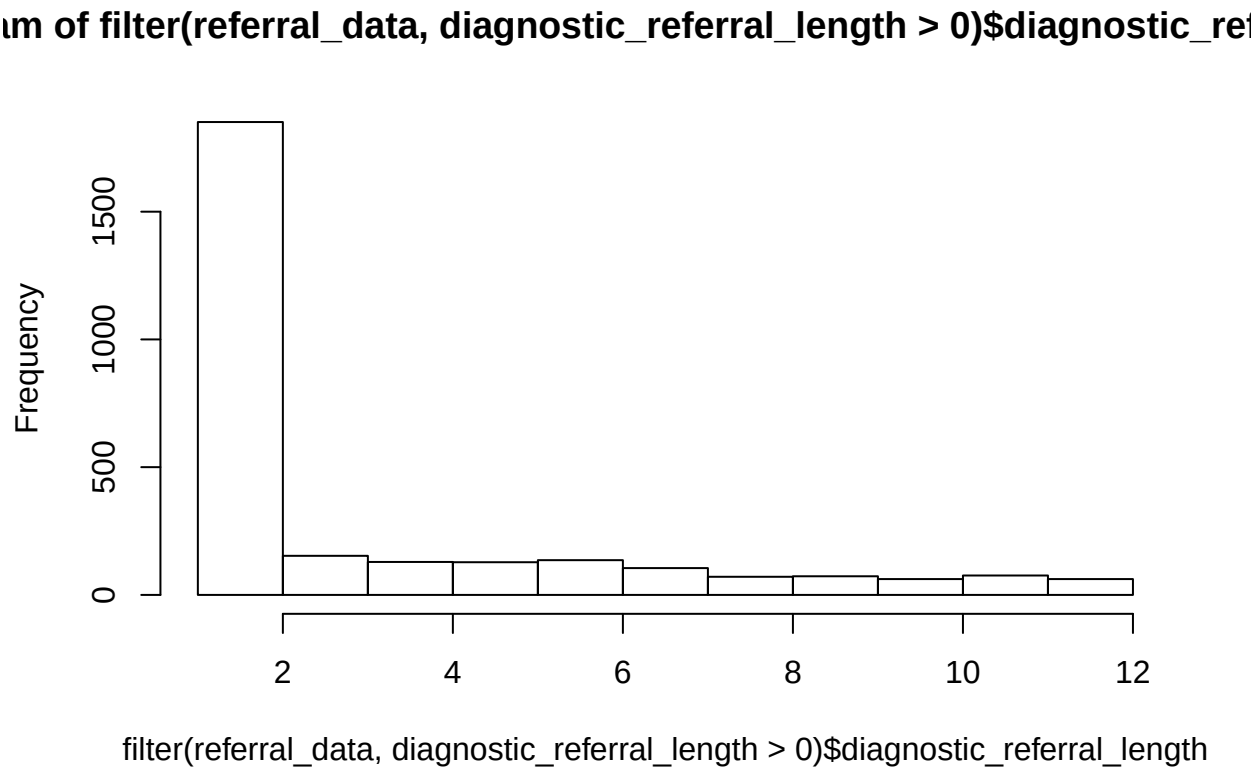
```
old_total <- 3570
new_total <- 3765
```

Plotting frequencies of length of referral (this is plotting the data with expirations filtered out):

```
hist(filter(old_referral_data, diagnostic_referral_length > 0)$diagnostic_referral_length)
```




```
hist(filter(referral_data, diagnostic_referral_length > 0)$diagnostic_referral_length)
```



These histograms show the agents who lost referrals due to completion. Dividing by the total number of events from earlier, we can figure out what portion of referrals got completed (ignore the 4, that's just the number of columns):

```
n_old_completed <- length(which(old_referral_data[2] > 0))
n_old_completed
```

```
## [1] 2288
```

```
old_completion_rate <- n_old_completed/old_total
old_completion_rate
```

```
## [1] 0.6408964
```

```
n_new_completed <- length(which(referral_data[2] > 0))
n_new_completed
```

```
## [1] 2846
```

```
new_completion_rate <- n_new_completed/new_total
new_completion_rate
```

```
## [1] 0.7559097
```

This shows that the original completion rate is 64% and the new is 75%.

We can compare this to other ratios, like the expiration rate (flagged with a length of -1):

```
n_old_expired <- length(which(old_referral_data[2] == -1))
n_old_expired
```

```
## [1] 1282
```

```
old_expiry_rate <- n_old_expired/old_total
old_expiry_rate
```

```
## [1] 0.3591036
```

```
n_new_expired <- length(which(referral_data[2] == -1))
n_new_expired
```

```
## [1] 919
```

```
new_expiry_rate <- n_new_expired/new_total
new_expiry_rate
```

```
## [1] 0.2440903
```

```
old_completion_rate + old_expiry_rate
```

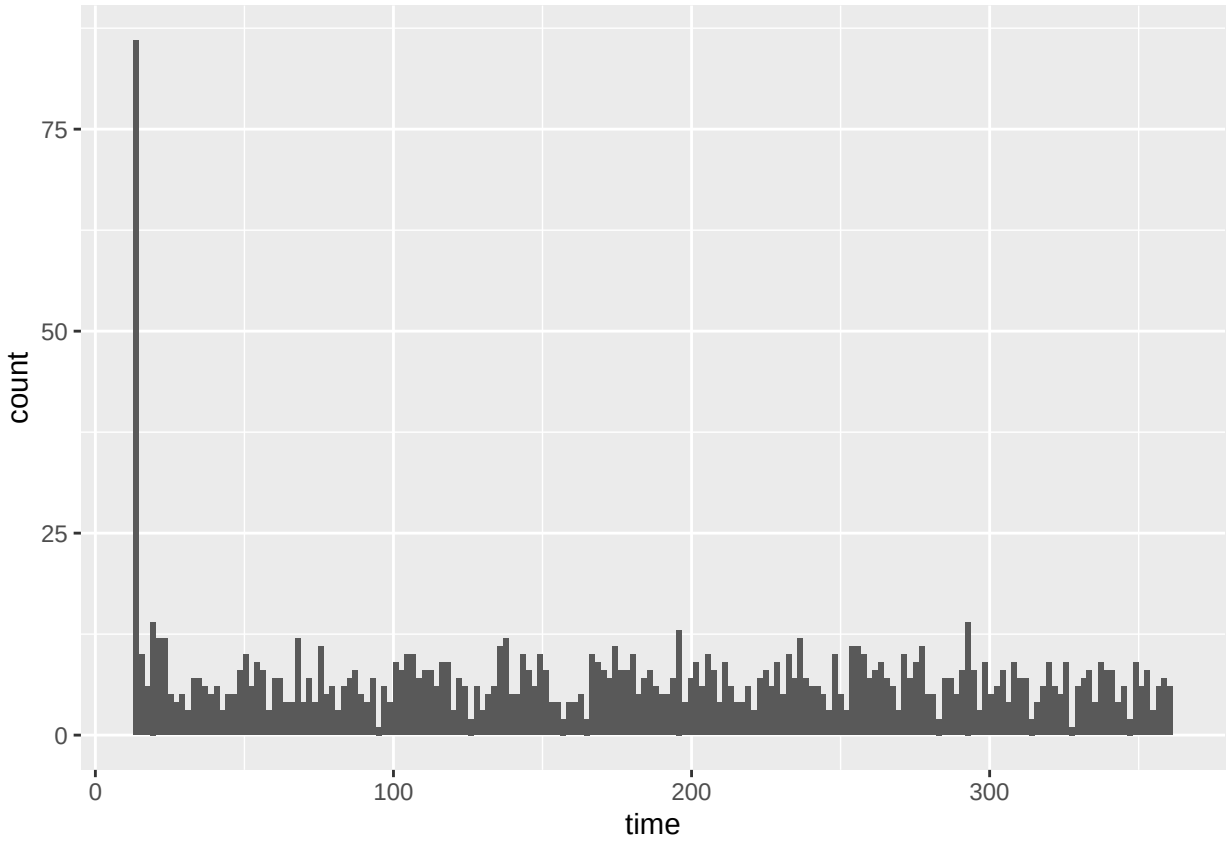
```
## [1] 1
```

```
new_completion_rate + new_expiry_rate
```

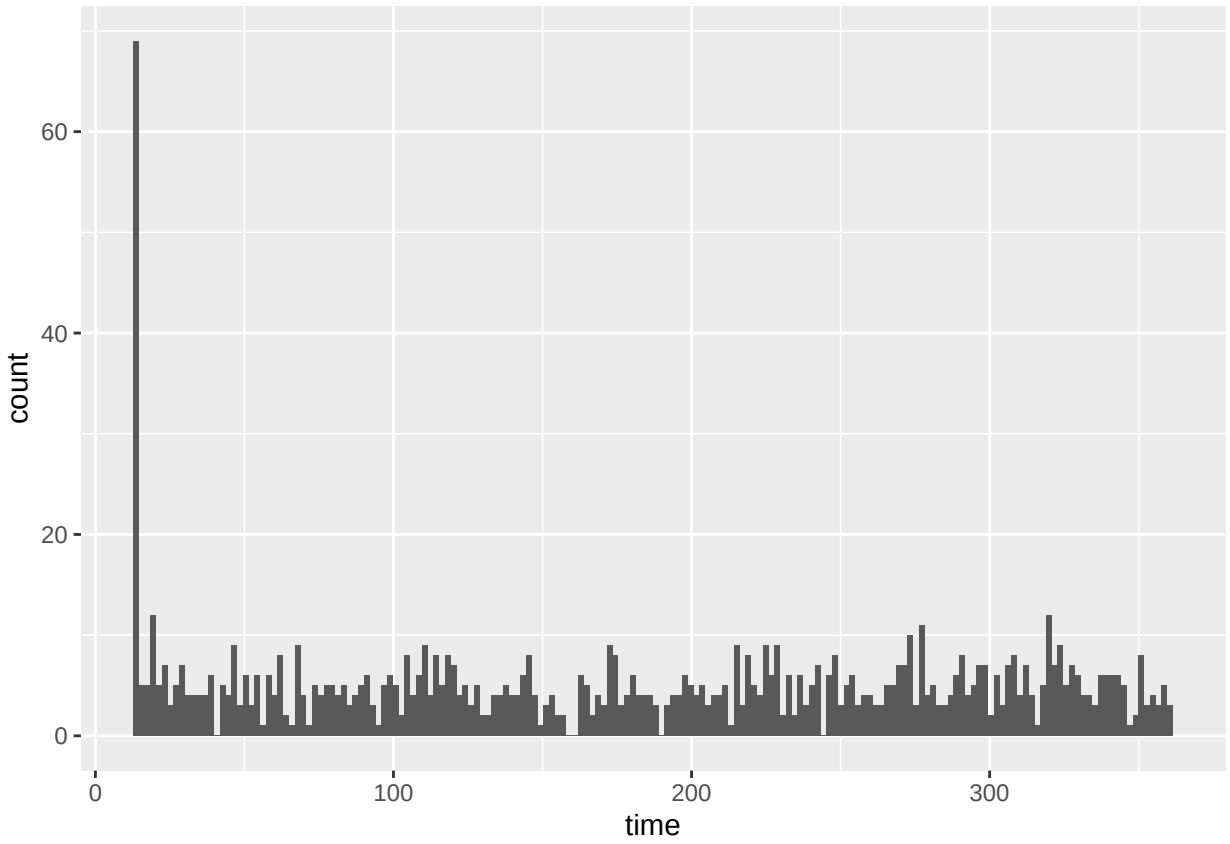
```
## [1] 1
```

Finally, look at expirations over time as a sanity check. Notice that months 1-12 contain no expirations(as expected):

```
ggplot(data=filter(old_referral_data, diagnostic_referral_length == -1))+
  geom_histogram(aes(x=time), bins = 180)
```




```
ggplot(data=filter(referral_data, diagnostic_referral_length == -1))+
  geom_histogram(aes(x=time),bins = 180)
```


60
40
20
0
count
0
100
200
300
time