

# read\_large.R

*jiayuan*

*Wed Dec 2 10:03:43 2015*

```
library(ff)
```

```
## Loading required package: bit
## Attaching package bit
## package:bit (c) 2008-2012 Jens Oehlschlaegel (GPL-2)
## creators: bit bitwhich
## coercion: as.logical as.integer as.bit as.bitwhich which
## operator: ! & | xor != ==
## querying: print length any all min max range sum summary
## bit access: length<- [ [<- [[ [<-
## for more help type ?bit
##
## Attaching package: 'bit'
##
## The following object is masked from 'package:base':
##
##     xor
##
## Attaching package ff
## -getOption("fftempdir")=="/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T//RtmpgMDtam"
##
## -getOption("ffextension")== "ff"
##
## -getOption("ffdrop")==TRUE
##
## -getOption("fffinonexit")==TRUE
##
## -getOption("ffpagesize")==65536
##
## -getOption("ffcaching")== "mmnoflush" -- consider "ffeachflush" if your system stalls on large writes
##
## -getOption("ffbatchbytes")==16777216 -- consider a different value for tuning your system
##
## -getOption("ffmaxbytes")==536870912 -- consider a different value for tuning your system
##
## Attaching package: 'ff'
##
## The following objects are masked from 'package:bit':
##
##     clone, clone.default, clone.list
##
## The following objects are masked from 'package:utils':
##
##     write.csv, write.csv2
##
## The following objects are masked from 'package:base':
```

```
##
##      is.factor, is.ordered
```

```
library(ffbase)
```

```
##
## Attaching package: 'ffbase'
##
## The following objects are masked from 'package:ff':
##
##      [.ff, [<- .ff
##
## The following objects are masked from 'package:base':
##
##      %in%, table
```

```
library(tabplot)
```

```
setwd("/Users/jiayuan/Documents/MA881/hw12")
#https://data.cityofboston.gov/City-Services/311-Service-Requests/awu8-dc52
requests <- read.csv.ffdf(file="311__Service_Requests-2.csv", header=TRUE, VERBOSE=TRUE,
                          first.rows=10000, next.rows=50000, colClasses=NA)
```

```
## read.table.ffdf 1..10000 (10000)  csv-read=2.631sec ffdwrite=0.315sec
## read.table.ffdf 10001..60000 (50000)  csv-read=14.695sec ffdwrite=0.82sec
## read.table.ffdf 60001..110000 (50000)  csv-read=12.474sec ffdwrite=1.068sec
## read.table.ffdf 110001..160000 (50000)  csv-read=12.586sec ffdwrite=1.195sec
## read.table.ffdf 160001..210000 (50000)  csv-read=12.368sec ffdwrite=1.084sec
## read.table.ffdf 210001..260000 (50000)  csv-read=12.534sec ffdwrite=1.426sec
## read.table.ffdf 260001..310000 (50000)  csv-read=13.732sec ffdwrite=1.938sec
## read.table.ffdf 310001..360000 (50000)  csv-read=12.989sec ffdwrite=2.025sec
## read.table.ffdf 360001..410000 (50000)  csv-read=12.425sec ffdwrite=2.009sec
## read.table.ffdf 410001..460000 (50000)  csv-read=13.105sec ffdwrite=1.975sec
## read.table.ffdf 460001..510000 (50000)  csv-read=14.798sec ffdwrite=2.426sec
## read.table.ffdf 510001..560000 (50000)  csv-read=13.81sec ffdwrite=1.833sec
## read.table.ffdf 560001..610000 (50000)  csv-read=16.369sec ffdwrite=2.428sec
## read.table.ffdf 610001..660000 (50000)  csv-read=18.675sec ffdwrite=5.333sec
## read.table.ffdf 660001..698057 (38057)  csv-read=13.394sec ffdwrite=2.85sec
## csv-read=196.585sec ffdwrite=28.725sec TOTAL=225.31sec
```

```
requests <- read.csv.ffdf(file="https://dl.dropboxusercontent.com/u/382061/311__Service_Requests.csv",
                          first.rows=10000, next.rows=50000, colClasses=NA)
```

```
## read.table.ffdf 1..10000 (10000)  csv-read=3.685sec ffdwrite=0.645sec
## read.table.ffdf 10001..60000 (50000)  csv-read=9.688sec ffdwrite=1.591sec
## read.table.ffdf 60001..110000 (50000)  csv-read=10.28sec ffdwrite=3.079sec
## read.table.ffdf 110001..160000 (50000)  csv-read=15.961sec ffdwrite=1.482sec
## read.table.ffdf 160001..210000 (50000)  csv-read=9.551sec ffdwrite=0.962sec
## read.table.ffdf 210001..260000 (50000)  csv-read=11.085sec ffdwrite=4.175sec
## read.table.ffdf 260001..310000 (50000)  csv-read=10.605sec ffdwrite=4.267sec
## read.table.ffdf 310001..360000 (50000)  csv-read=16.815sec ffdwrite=7.78sec
## read.table.ffdf 360001..410000 (50000)  csv-read=12.66sec ffdwrite=3.364sec
```

```
## read.table.ffdf 410001..460000 (50000) csv-read=12.158sec ffd-f-write=3.493sec
## read.table.ffdf 460001..510000 (50000) csv-read=12.643sec ffd-f-write=2.288sec
## read.table.ffdf 510001..560000 (50000) csv-read=11.917sec ffd-f-write=2.648sec
## read.table.ffdf 560001..610000 (50000) csv-read=15.624sec ffd-f-write=4.086sec
## read.table.ffdf 610001..660000 (50000) csv-read=12.991sec ffd-f-write=3.804sec
## read.table.ffdf 660001..710000 (50000) csv-read=14.69sec ffd-f-write=5.147sec
## read.table.ffdf 710001..722778 (12778) csv-read=2.61sec ffd-f-write=4.691sec
## csv-read=182.963sec ffd-f-write=53.502sec TOTAL=236.465sec
```

```
dim(requests)
```

```
## [1] 722778      33
```

```
names(requests)
```

```
## [1] "CASE_ENQUIRY_ID"      "OPEN_DT"
## [3] "TARGET_DT"           "CLOSED_DT"
## [5] "OnTime_Status"       "CASE_STATUS"
## [7] "CLOSURE_REASON"      "CASE_TITLE"
## [9] "SUBJECT"              "REASON"
## [11] "TYPE"                 "QUEUE"
## [13] "Department"          "SubmittedPhoto"
## [15] "ClosedPhoto"         "Location"
## [17] "fire_district"       "pwd_district"
## [19] "city_council_district" "police_district"
## [21] "neighborhood"        "neighborhood_services_district"
## [23] "ward"                 "precinct"
## [25] "land_usage"           "LOCATION_STREET_NAME"
## [27] "LOCATION_ZIPCODE"      "Property_Type"
## [29] "Property_ID"         "LATITUDE"
## [31] "LONGITUDE"           "Source"
## [33] "Geocoded_Location"
```

```
head(requests)
```

```
## ffd (all open) dim=c(722778,6), dimorder=c(1,2) row.names=NULL
## ffd virtual mapping
##           PhysicalName VirtualVmode PhysicalVmode  AsIs
## CASE_ENQUIRY_ID CASE_ENQUIRY_ID      double      double FALSE
## OPEN_DT          OPEN_DT             integer      integer FALSE
## TARGET_DT        TARGET_DT            integer      integer FALSE
## CLOSED_DT        CLOSED_DT            integer      integer FALSE
## OnTime_Status    OnTime_Status         integer      integer FALSE
## CASE_STATUS      CASE_STATUS           integer      integer FALSE
##           VirtualIsMatrix PhysicalIsMatrix PhysicalElementNo
## CASE_ENQUIRY_ID      FALSE              FALSE              1
## OPEN_DT              FALSE              FALSE              2
## TARGET_DT            FALSE              FALSE              3
## CLOSED_DT            FALSE              FALSE              4
## OnTime_Status        FALSE              FALSE              5
## CASE_STATUS          FALSE              FALSE              6
##           PhysicalFirstCol PhysicalLastCol PhysicalIsOpen
```

## CASE_ENQUIRY_ID	1	1	TRUE
## OPEN_DT	1	1	TRUE
## TARGET_DT	1	1	TRUE
## CLOSED_DT	1	1	TRUE
## OnTime_Status	1	1	TRUE
## CASE_STATUS	1	1	TRUE

## fdfd data

##	CASE_ENQUIRY_ID	OPEN_DT
## 1	101000295613	07/01/2011 01:32:33 AM
## 2	101000295614	07/01/2011 02:18:31 AM
## 3	101000295615	07/01/2011 02:28:04 AM
## 4	101000295616	07/01/2011 03:03:48 AM
## 5	101000295617	07/01/2011 03:12:31 AM
## 6	101000295618	07/01/2011 03:43:07 AM
## 7	101000295619	07/01/2011 05:35:41 AM
## 8	101000295620	07/01/2011 06:24:38 AM
## :	:	:
## 722771	101001640653	11/22/2015 09:23:00 PM
## 722772	101001640654	11/22/2015 09:28:00 PM
## 722773	101001640655	11/22/2015 09:59:00 PM
## 722774	101001640656	11/22/2015 10:00:16 PM
## 722775	101001640658	11/22/2015 10:04:40 PM
## 722776	101001640659	11/22/2015 10:30:08 PM
## 722777	101001640661	11/22/2015 11:03:00 PM
## 722778	101001640662	11/22/2015 11:57:24 PM
##	TARGET_DT	CLOSED_DT
## 1	07/15/2011 01:32:32 AM	07/01/2011 05:06:58 AM
## 2	07/05/2011 02:18:31 AM	07/01/2011 06:12:43 AM
## 3	08/15/2011 02:28:04 AM	08/01/2011 03:21:46 PM
## 4	07/15/2011 03:03:48 AM	07/22/2011 04:13:45 PM
## 5		07/01/2011 06:12:38 AM
## 6		07/05/2011 09:16:10 AM
## 7	08/15/2011 05:35:41 AM	10/17/2011 12:33:28 PM
## 8	07/15/2011 06:24:38 AM	08/23/2011 10:33:03 AM
## :	:	:
## 722771	12/01/2015 08:30:00 AM	
## 722772	11/25/2015 08:30:00 AM	
## 722773		
## 722774	11/24/2015 08:30:00 AM	11/22/2015 10:50:07 PM
## 722775	11/24/2015 08:30:00 AM	
## 722776	11/23/2015 10:30:06 PM	
## 722777	11/27/2015 08:30:00 AM	
## 722778	11/24/2015 08:30:00 AM	
##	OnTime_Status	CASE_STATUS
## 1	ONTIME	Closed
## 2	ONTIME	Closed
## 3	ONTIME	Closed
## 4	OVERDUE	Closed
## 5	ONTIME	Closed
## 6	ONTIME	Closed
## 7	OVERDUE	Closed
## 8	OVERDUE	Closed
## :	:	:
## 722771	ONTIME	Open

```
## 722772 ONTIME      Open
## 722773 ONTIME      Open
## 722774 ONTIME      Closed
## 722775 ONTIME      Open
## 722776 ONTIME      Open
## 722777 ONTIME      Open
## 722778 ONTIME      Open
```

```
str(requests[1:10,])
```

```
## 'data.frame':  10 obs. of  33 variables:
## $ CASE_ENQUIRY_ID      : num  1.01e+11 1.01e+11 1.01e+11 1.01e+11 1.01e+11 ...
## $ OPEN_DT              : Factor w/ 701229 levels "07/01/2011 01:01:28 PM",...: 15 35 38 51 5
## $ TARGET_DT            : Factor w/ 394017 levels "", "01/01/2013 01:13:08 PM",...: 2189 477 5
## $ CLOSED_DT            : Factor w/ 636836 levels "", "01/02/2013 03:18:07 PM",...: 503 507 7
## $ OnTime_Status        : Factor w/ 3 levels "", "ONTIME", "OVERDUE": 2 2 2 3 2 2 3 3 2 3
## $ CASE_STATUS          : Factor w/ 2 levels "Closed", "Open": 1 1 1 1 1 1 1 1 1 1
## $ CLOSURE_REASON       : Factor w/ 234270 levels "", "Case Closed",...: 1958 2295 1771 1398 6
## $ CASE_TITLE           : Factor w/ 6825 levels ".", "Abandoned Bicycle",...: 357 316 222 357
## $ SUBJECT              : Factor w/ 19 levels "Inspectional Services",...: 5 5 6 5 5 2 6 5 5
## $ REASON               : Factor w/ 63 levels "Abandoned Bicycle",...: 22 19 20 22 12 14 20 1
## $ TYPE                 : Factor w/ 220 levels "Abandoned Bicycle",...: 102 91 51 102 31 58 5
## $ QUEUE                : Factor w/ 167 levels "BTDT_Abandoned Bicycle",...: 81 76 5 84 46 2
## $ Department           : Factor w/ 18 levels "BTDT", "BWSC",...: 9 9 1 9 9 4 1 9 9 5
## $ SubmittedPhoto       : logi  NA NA NA NA NA NA NA ...
## $ ClosedPhoto          : logi  NA NA NA NA NA NA NA ...
## $ Location             : Factor w/ 111677 levels "", "0 Blake St Mattapan MA 02126",...: 6
## $ fire_district        : int   8 12 NA 12 7 9 4 9 7 1
## $ pwd_district         : Factor w/ 14 levels "", "02", "03", "04",...: 7 6 6 8 3 10 10 11 7 9
## $ city_council_district : int   3 5 6 5 3 6 8 7 3 1
## $ police_district      : Factor w/ 24 levels "", "A1", "A15",...: 6 13 13 12 7 11 10 11 7 4
## $ neighborhood        : Factor w/ 24 levels "", "Allston", "Allston / Brighton",...: 13 18 2
## $ neighborhood_services_district: int   7 10 12 10 7 11 14 13 7 1
## $ ward                : Factor w/ 54 levels "", "01", "02", "03",...: 37 41 41 38 33 30 42 31
## $ precinct            : int  1714 2008 2017 1816 1309 1006 2102 1103 1602 107
## $ land_usage           : Factor w/ 17 levels "", "A", "C", "CC",...: 1 11 1 12 1 12 1 1 3 13
## $ LOCATION_STREET_NAME : Factor w/ 109756 levels "", "0 Blake St",...: 6749 3653 6557 2380 6
## $ LOCATION_ZIPCODE     : int   NA 2131 NA 2136 NA 2130 NA NA 2122 2128
## $ Property_Type        : Factor w/ 3 levels "", "Address", "Intersection": 3 2 3 2 3 2 3 3 2
## $ Property_ID          : int   5512 87461 3522 55708 4182 51077 617 1853 60876 274500
## $ LATITUDE             : num   42.3 42.3 42.3 42.2 42.3 ...
## $ LONGITUDE            : num  -71.1 -71.1 -71.2 -71.1 -71.1 ...
## $ Source               : Factor w/ 7 levels "Citizens Connect App",...: 3 5 5 5 2 2 5 2 3 5
## $ Geocoded_Location    : Factor w/ 96601 levels "(0, 0)", "(42.2333, -71.1323)",...: 553 129
```

```
ontime <- requests$OnTime_Status
length(ontime)
```

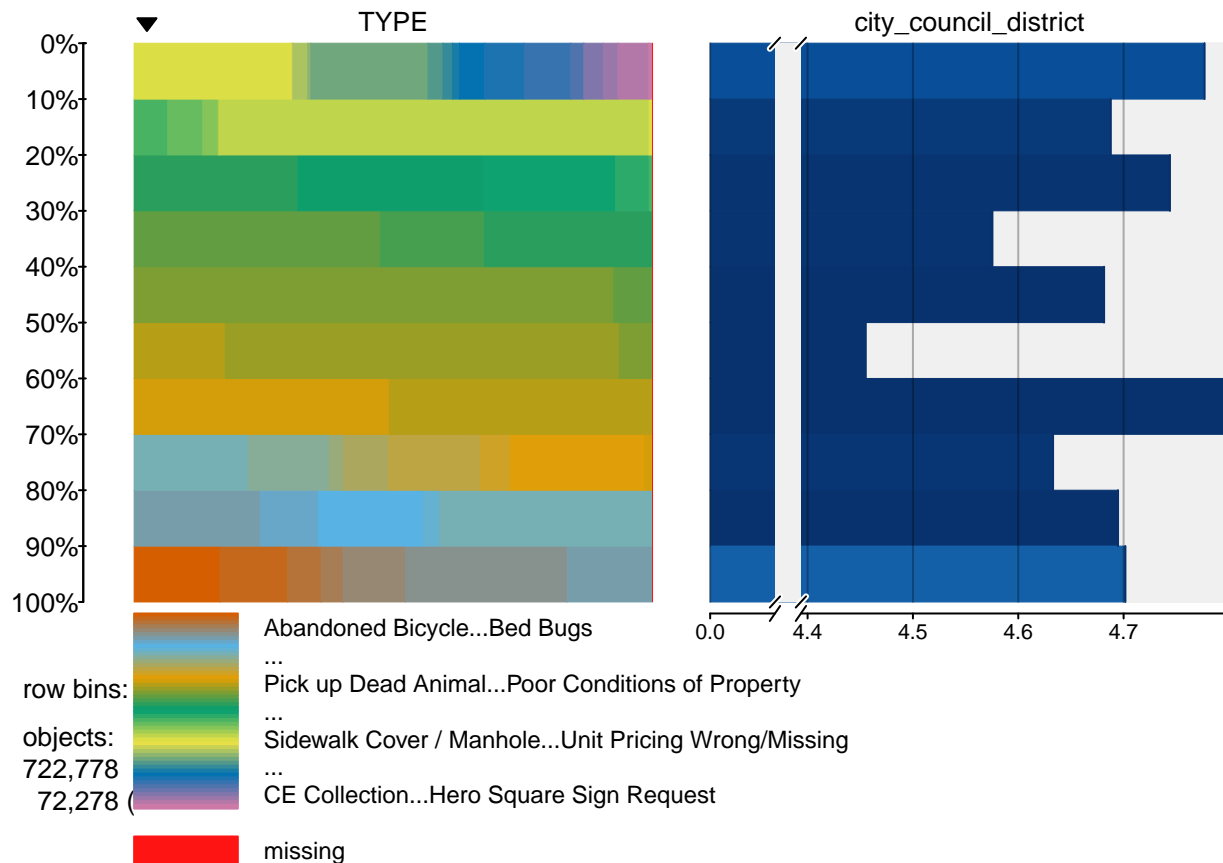
```
## [1] 722778
```

```
colnames(requests)
```

```
## [1] "CASE_ENQUIRY_ID"      "OPEN_DT"
## [3] "TARGET_DT"           "CLOSED_DT"
## [5] "OnTime_Status"       "CASE_STATUS"
## [7] "CLOSURE_REASON"      "CASE_TITLE"
## [9] "SUBJECT"             "REASON"
## [11] "TYPE"                "QUEUE"
## [13] "Department"          "SubmittedPhoto"
## [15] "ClosedPhoto"         "Location"
## [17] "fire_district"       "pwd_district"
## [19] "city_council_district" "police_district"
## [21] "neighborhood"        "neighborhood_services_district"
## [23] "ward"                "precinct"
## [25] "land_usage"          "LOCATION_STREET_NAME"
## [27] "LOCATION_ZIPCODE"      "Property_Type"
## [29] "Property_ID"         "LATITUDE"
## [31] "LONGITUDE"           "Source"
## [33] "Geocoded_Location"
```

```
tableplot(requests,select = c(11,19),nBins=10)
```

```
## Warning in grid.Call(L_convert, x, as.integer(whatfrom),
## as.integer(whatto), : font width unknown for character 0x9
```



```
p <- requests$precinct
str(p)
```

```
## list()
## - attr(*, "physical")=Class 'ff_pointer' <externalptr>
## .. attr(*, "vmode")= chr "integer"
## .. attr(*, "maxlength")= int 722778
## .. attr(*, "pattern")= chr "ffdf"
## .. attr(*, "filename")= chr "/private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/"
## .. attr(*, "pagesize")= int 65536
## .. attr(*, "finalizer")= chr "close"
## .. attr(*, "finonexit")= logi TRUE
## .. attr(*, "readonly")= logi FALSE
## .. attr(*, "caching")= chr "mmnoflush"
## - attr(*, "virtual")= list()
## .. attr(*, "Length")= int 722778
## .. attr(*, "Symmetric")= logi FALSE
## .. attr(*, "class")= chr "virtual"
## - attr(*, "class") = chr [1:2] "ff_vector" "ff"
```

```
length(p)
```

```
## [1] 722778
```

```
head(p)
```

```
## [1] 1714 2008 2017 1816 1309 1006
```

```
levels.ff(ontime)
```

```
## [1] "" "ONTIME" "OVERDUE"
```

```
sum(ontime[1:722778]=="ONTIME")
```

```
## [1] 568475
```

```
ffsave(requests, file="req")
```

```
## [1] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513dd83ae0.
## [2] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513f0167d8.
## [3] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517548b638.
## [4] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685178b352af.
## [5] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685144f99d11.
## [6] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515fb6ea77.
## [7] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685161db5bb8.
## [8] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851862bf39.f
## [9] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851af83b7c.f
## [10] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68512f094f84.
## [11] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851c437b3c.f
```

```
## [12] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851224fa66e.
## [13] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511b37915b.
## [14] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d27fb52.
## [15] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68516be4f445.
## [16] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851461155a.f
## [17] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511c0c9f5.f
## [18] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851180af4b9.
## [19] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517748aa03.
## [20] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514391e823.
## [21] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851201f747d.
## [22] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517118cb04.
## [23] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685112b8ad9e.
## [24] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511a8e69ac.
## [25] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517bba8d2.
## [26] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d48b274.
## [27] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514fbc0d84.
## [28] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851431b7e01.
## [29] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685141ed9612.
## [30] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515517998e.
## [31] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851626547.ff
## [32] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513be71883.
## [33] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685140fa5f2e.
```

```
reqs <- fflload("req")
```

```
## Warning in FUN(X[[i]], ...): did not overwrite object 'requests'
```

```
# Start the clock!
ptm <- proc.time()
sum(ontime[1:722778]=="ONTIME")
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685144f99d11.ff
```

```
## [1] 568475
```

```
# Stop the clock
proc.time() - ptm
```

```
## user system elapsed
## 0.109 0.012 0.202
```

```
system.time(sum(ontime[1:722778]=="ONTIME"))
```

```
## user system elapsed
## 0.129 0.003 0.131
```

```
length(requests$precinct)
```

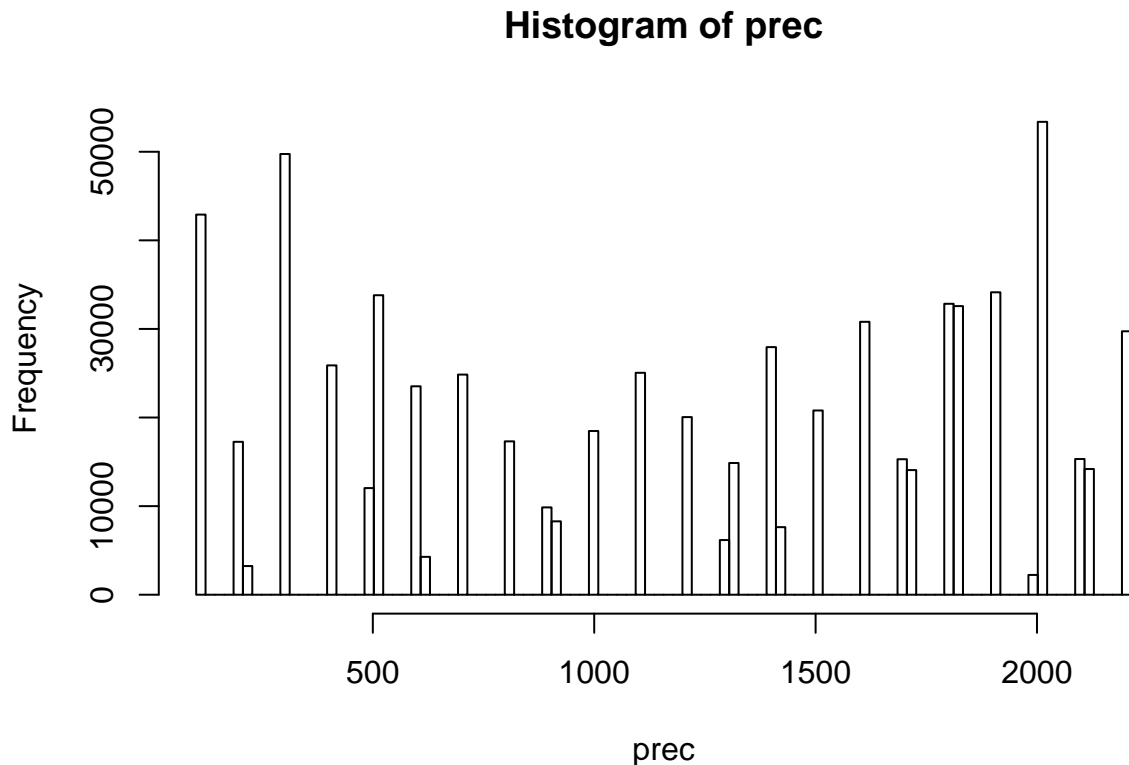
```
## [1] 722778
```



```
any(requests$precinct,NA)
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511a8e69ac.ff  
## [1] TRUE
```

```
prec <- requests$precinct[!is.na(requests$precinct)]  
hist(prec, plot = TRUE)
```



```
#####  
str(requests[1:10,])
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513dd83ae0.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513f0167d8.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517548b638.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685178b352af.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515fb6ea77.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685161db5bb8.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851862bf39.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851af83b7c.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68512f094f84.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851c437b3c.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851224fa66e.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511b37915b.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d27fb52.ff  
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68516be4f445.ff
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf6851461155a.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68511c0c9f5.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf6851180af4b9.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68517748aa03.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68514391e823.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf6851201f747d.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68517118cb04.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf685112b8ad9e.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68517bbba8d2.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68515d48b274.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68514fbc0d84.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf6851431b7e01.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf685141ed9612.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68515517998e.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf6851626547.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68513be71883.ff
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf685140fa5f2e.ff
```

```
## 'data.frame': 10 obs. of 33 variables:
```

```
## $ CASE_ENQUIRY_ID : num 1.01e+11 1.01e+11 1.01e+11 1.01e+11 1.01e+11 ...
## $ OPEN_DT : Factor w/ 701229 levels "07/01/2011 01:01:28 PM",...: 15 35 38 51 5
## $ TARGET_DT : Factor w/ 394017 levels "", "01/01/2013 01:13:08 PM",...: 2189 477 5
## $ CLOSED_DT : Factor w/ 636836 levels "", "01/02/2013 03:18:07 PM",...: 503 507 7
## $ OnTime_Status : Factor w/ 3 levels "", "ONTIME", "OVERDUE": 2 2 2 3 2 2 3 3 2 3
## $ CASE_STATUS : Factor w/ 2 levels "Closed", "Open": 1 1 1 1 1 1 1 1 1 1
## $ CLOSURE_REASON : Factor w/ 234270 levels "", "Case Closed",...: 1958 2295 1771 1398 6
## $ CASE_TITLE : Factor w/ 6825 levels ".", "Abandoned Bicycle",...: 357 316 222 357
## $ SUBJECT : Factor w/ 19 levels "Inspectional Services",...: 5 5 6 5 5 2 6 5 5
## $ REASON : Factor w/ 63 levels "Abandoned Bicycle",...: 22 19 20 22 12 14 20 1
## $ TYPE : Factor w/ 220 levels "Abandoned Bicycle",...: 102 91 51 102 31 58 5
## $ QUEUE : Factor w/ 167 levels "BTDT_Abandoned Bicycle",...: 81 76 5 84 46 2
## $ Department : Factor w/ 18 levels "BTDT", "BWSC",...: 9 9 1 9 9 4 1 9 9 5
## $ SubmittedPhoto : logi NA NA NA NA NA NA ...
## $ ClosedPhoto : logi NA NA NA NA NA NA ...
## $ Location : Factor w/ 111677 levels "", "0 Blake St Mattapan MA 02126",...: 6
## $ fire_district : int 8 12 NA 12 7 9 4 9 7 1
## $ pwd_district : Factor w/ 14 levels "", "02", "03", "04",...: 7 6 6 8 3 10 10 11 7 9
## $ city_council_district : int 3 5 6 5 3 6 8 7 3 1
## $ police_district : Factor w/ 24 levels "", "A1", "A15",...: 6 13 13 12 7 11 10 11 7 4
## $ neighborhood : Factor w/ 24 levels "", "Allston", "Allston / Brighton",...: 13 18 2
## $ neighborhood_services_district: int 7 10 12 10 7 11 14 13 7 1
## $ ward : Factor w/ 54 levels "", "01", "02", "03",...: 37 41 41 38 33 30 42 31
## $ precinct : int 1714 2008 2017 1816 1309 1006 2102 1103 1602 107
## $ land_usage : Factor w/ 17 levels "", "A", "C", "CC",...: 1 11 1 12 1 12 1 1 3 13
## $ LOCATION_STREET_NAME : Factor w/ 109756 levels "", "0 Blake St",...: 6749 3653 6557 2380 6
## $ LOCATION_ZIPCODE : int NA 2131 NA 2136 NA 2130 NA NA 2122 2128
## $ Property_Type : Factor w/ 3 levels "", "Address", "Intersection": 3 2 3 2 3 2 3 3 2
## $ Property_ID : int 5512 87461 3522 55708 4182 51077 617 1853 60876 274500
## $ LATITUDE : num 42.3 42.3 42.3 42.2 42.3 ...
## $ LONGITUDE : num -71.1 -71.1 -71.2 -71.1 -71.1 ...
## $ Source : Factor w/ 7 levels "Citizens Connect App",...: 3 5 5 5 2 2 5 2 3 5
## $ Geocoded_Location : Factor w/ 96601 levels "(0, 0)", "(42.2333, -71.1323)",...: 553 129
```

```
# $ OnTime_Status: Factor w/ 3 levels "", "ONTIME", "OVERDUE": 2 2 2 3 2 2 3 3 2 3
```

```
ontime <- requests$OnTime_Status  
length(ontime)
```

```
## [1] 722778
```

```
str(ontime)
```

```
## list()  
## - attr(*, "physical")=Class 'ff_pointer' <externalptr>  
## .. attr(*, "vmode")= chr "integer"  
## .. attr(*, "maxlength")= int 722778  
## .. attr(*, "pattern")= chr "ffdf"  
## .. attr(*, "filename")= chr "/private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/  
## .. attr(*, "pagesize")= int 65536  
## .. attr(*, "finalizer")= chr "close"  
## .. attr(*, "finonexit")= logi TRUE  
## .. attr(*, "readonly")= logi FALSE  
## .. attr(*, "caching")= chr "mmnoflush"  
## - attr(*, "virtual")= list()  
## .. attr(*, "Length")= int 722778  
## .. attr(*, "Symmetric")= logi FALSE  
## .. attr(*, "Levels")= chr [1:3] "" "ONTIME" "OVERDUE"  
## .. attr(*, "ramclass")= chr "factor"  
## .. attr(*, "class")= chr "virtual"  
## - attr(*, "class")= chr [1:2] "ff_vector" "ff"
```

```
length(ontime) #722778
```

```
## [1] 722778
```

```
head(ontime)
```

```
## [1] ONTIME ONTIME ONTIME OVERDUE ONTIME ONTIME  
## Levels: ONTIME OVERDUE
```

```
levels.ff(ontime)
```

```
## [1] "" "ONTIME" "OVERDUE"
```

```
sum(ontime[1:722778]=="ONTIME") #568475
```

```
## [1] 568475
```

```
sum(ontime[1:722778]=="OVERDUE") #153713
```

```
## [1] 153713
```

```
sum(ontime[1:722778]=="") #590
```

```
## [1] 590
```

```
568475+153713+590 #722778
```

```
## [1] 722778
```

```
ffsave(requests, file="req")
```

```
## [1] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513dd83ae0.
## [2] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513f0167d8.
## [3] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517548b638.
## [4] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685178b352af.
## [5] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685144f99d11.
## [6] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515fb6ea77.
## [7] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685161db5bb8.
## [8] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851862bf39.f
## [9] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851af83b7c.f
## [10] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68512f094f84.
## [11] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851c437b3c.f
## [12] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851224fa66e.
## [13] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511b37915b.
## [14] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d27fb52.
## [15] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68516be4f445.f
## [16] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851461155a.f
## [17] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511c0c9f5.f
## [18] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851180af4b9.
## [19] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517748aa03.
## [20] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514391e823.
## [21] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851201f747d.
## [22] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517118cb04.
## [23] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685112b8ad9e.
## [24] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511a8e69ac.
## [25] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517bbba8d2.
## [26] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d48b274.
## [27] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514fbc0d84.
## [28] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851431b7e01.
## [29] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685141ed9612.
## [30] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515517998e.
## [31] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851626547.ff
## [32] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513be71883.
## [33] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685140fa5f2e.
```

```
reqs <- fffload("req")
```

```
## Warning in FUN(X[[i]], ...): did not overwrite object 'requests'
```

```
# Start the clock!
ptm <- proc.time()
sum(ontime[1:722778]=="ONTIME")
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf685144f99d11.ff
```

```
## [1] 568475
```

```
# Stop the clock  
proc.time() - ptm
```

```
##      user  system elapsed  
##    0.093   0.024   0.543
```

```
system.time(sum(ontime[1:722778]=="ONTIME"))
```

```
##      user  system elapsed  
##    0.085   0.002   0.087
```

```
# The variable city_council_district is not a factor.  
length(requests$city_council_district)
```

```
## [1] 722778
```

```
str(requests$city_council_district)
```

```
## list()  
## - attr(*, "physical")=Class 'ff_pointer' <externalptr>  
## .. attr(*, "vmode")= chr "integer"  
## .. attr(*, "maxlength")= int 722778  
## .. attr(*, "pattern")= chr "ffdf"  
## .. attr(*, "filename")= chr "/private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/  
## .. attr(*, "pagesize")= int 65536  
## .. attr(*, "finalizer")= chr "close"  
## .. attr(*, "finonexit")= logi TRUE  
## .. attr(*, "readonly")= logi FALSE  
## .. attr(*, "caching")= chr "mmnoflush"  
## - attr(*, "virtual")= list()  
## .. attr(*, "Length")= int 722778  
## .. attr(*, "Symmetric")= logi FALSE  
## .. attr(*, "class")= chr "virtual"  
## - attr(*, "class") = chr [1:2] "ff_vector" "ff"
```

```
head(requests$city_council_district)
```

```
## opening ff /private/var/folders/9m/5v4_k6dx1mx_szfqkpgmh6gh0000gn/T/RtmpgMDtam/ffdf68517748aa03.ff
```

```
## [1] 3 5 6 5 3 6
```

```
any(requests$city_council_district,NA)
```

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

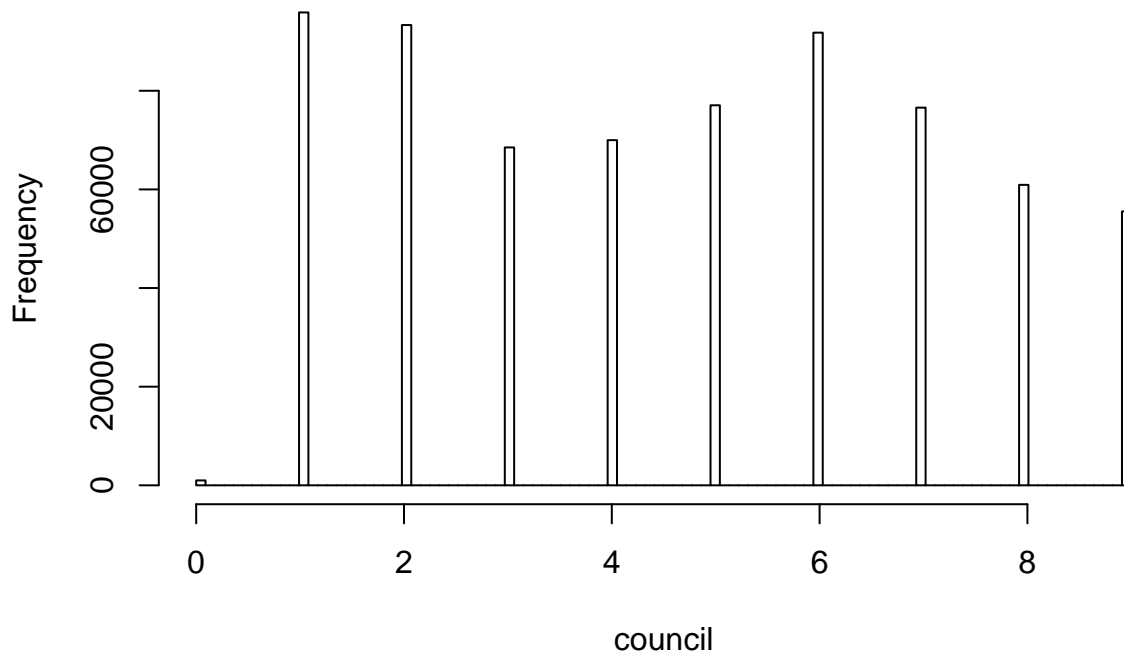
```

council <- requests$city_council_district[!is.na(requests$city_council_district)]
hist(council, plot = TRUE)

# Start the clock!
ptm <- proc.time()
council <- requests$city_council_district[!is.na(requests$city_council_district)]
hist(council, plot = TRUE)

```

**Histogram of council**



```

# Stop the clock
proc.time() - ptm

```

```

##      user  system elapsed
## 0.349   0.052   0.481

```

```
ffsave(requests, file="req")
```

```

## [1] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513dd83ae0.
## [2] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513f0167d8.
## [3] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517548b638.
## [4] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685178b352af.
## [5] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685144f99d11.
## [6] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515fb6ea77.
## [7] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685161db5bb8.
## [8] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851862bf39.f
## [9] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851af83b7c.f
## [10] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68512f094f84.
## [11] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851c437b3c.f

```

```
## [12] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851224fa66e.
## [13] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511b37915b.
## [14] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d27fb52.
## [15] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68516be4f445.
## [16] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851461155a.f
## [17] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511c0c9f5.f
## [18] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851180af4b9.
## [19] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517748aa03.
## [20] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514391e823.
## [21] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851201f747d.
## [22] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517118cb04.
## [23] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685112b8ad9e.
## [24] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68511a8e69ac.
## [25] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68517bba8d2.
## [26] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515d48b274.
## [27] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68514fbc0d84.
## [28] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851431b7e01.
## [29] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685141ed9612.
## [30] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68515517998e.
## [31] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf6851626547.ff
## [32] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf68513be71883.
## [33] " adding: private/var/folders/9m/5v4_k6dx1mx_szfqpkmh6gh0000gn/T/RtmpgMDtam/ffdf685140fa5f2e.
```

```
reqs <- fflload("req")
```

```
## Warning in FUN(X[[i]], ...): did not overwrite object 'requests'
```