Homework 1 Bonus

Hamed

2/9/2020

#### (a). Read the data into a data frame in R

pdcalls=read.csv("jerseycitypdservicecalls2017csv.csv",header = T)  
attach(pdcalls)  
head(pdcalls,6)

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000003 W 01/01/2017 00:05:17 1 01/01/2017 00:07:00  
## 4 17-000004 E 01/01/2017 00:05:52 1 01/01/2017 00:08:00  
## 5 17-000005 S 01/01/2017 00:06:55 1 01/01/2017 00:08:00  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/0001 00:00:00 B21.00  
## 5 01/01/0001 00:00:00 B23.20  
## 6 01/01/2017 00:23:00 C14.20  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 MERCHANT/CUSTOMER DISPUTE; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 OTHER INTER/CONFLICT; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address city  
## 1 PH 2 W601 1 40 JONES ST; JERSEY CITY  
## 2 PH 2 W101 0 40 JONES ST; JERSEY CITY  
## 3 PH 5 W401 1 20 TONNELE AVE; 3B JERSEY CITY  
## 4 PH 5 E501 1 2 FAIRMOUNT AVE; JERSEY CITY  
## 5 PH 2 S201 1 105 WILKINSON AVE; JERSEY CITY  
## 6 911 4 W201 1 2 WILMOT AVE; 13 JERSEY CITY  
## latitude longitude geo.count geo.error  
## 1 40.73022 -74.06236 1 S5HPNTSCZA  
## 2 40.73022 -74.06236 1 S5HPNTSCZA  
## 3 40.73093 -74.06704 1 S5HPNTSCZA  
## 4 40.71735 -74.06103 1 S5HPNTSCZA  
## 5 40.70592 -74.08002 1 S5HPNTSCZA  
## 6 0.00000 0.00000 0 Not Found

#### (b). Find the number of variables and number of rows of data, The first number is for rows and the second for variables

dim(pdcalls)

## [1] 107335 18

#### (c). Inspect the data by printing out the first six rows.

head(pdcalls,6)

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000003 W 01/01/2017 00:05:17 1 01/01/2017 00:07:00  
## 4 17-000004 E 01/01/2017 00:05:52 1 01/01/2017 00:08:00  
## 5 17-000005 S 01/01/2017 00:06:55 1 01/01/2017 00:08:00  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/0001 00:00:00 B21.00  
## 5 01/01/0001 00:00:00 B23.20  
## 6 01/01/2017 00:23:00 C14.20  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 MERCHANT/CUSTOMER DISPUTE; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 OTHER INTER/CONFLICT; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address city  
## 1 PH 2 W601 1 40 JONES ST; JERSEY CITY  
## 2 PH 2 W101 0 40 JONES ST; JERSEY CITY  
## 3 PH 5 W401 1 20 TONNELE AVE; 3B JERSEY CITY  
## 4 PH 5 E501 1 2 FAIRMOUNT AVE; JERSEY CITY  
## 5 PH 2 S201 1 105 WILKINSON AVE; JERSEY CITY  
## 6 911 4 W201 1 2 WILMOT AVE; 13 JERSEY CITY  
## latitude longitude geo.count geo.error  
## 1 40.73022 -74.06236 1 S5HPNTSCZA  
## 2 40.73022 -74.06236 1 S5HPNTSCZA  
## 3 40.73093 -74.06704 1 S5HPNTSCZA  
## 4 40.71735 -74.06103 1 S5HPNTSCZA  
## 5 40.70592 -74.08002 1 S5HPNTSCZA  
## 6 0.00000 0.00000 0 Not Found

#### (d). How many rows are there with missing values?

sum(!complete.cases(pdcalls))

## [1] 55

#### (e). Remove the rows with missing data

dim(pdcalls) #Before removing the missing data

## [1] 107335 18

pdcalls <- pdcalls[complete.cases(pdcalls), ]  
  
dim(pdcalls) #After removing the missing data

## [1] 107280 18

#### (f). Check to see if there are any duplicate rows. If there are duplicate rows then remove them.

which(duplicated(pdcalls)) #check for duplicates

## integer(0)

dim(pdcalls) #dimension of the dataframe before removing duplicates

## [1] 107280 18

pdcalls=pdcalls[!duplicated(pdcalls), ] #remove duplicates  
dim(pdcalls)

## [1] 107280 18

#### (g). Sort the data by call type

head(pdcalls) #Before the sorting

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000003 W 01/01/2017 00:05:17 1 01/01/2017 00:07:00  
## 4 17-000004 E 01/01/2017 00:05:52 1 01/01/2017 00:08:00  
## 5 17-000005 S 01/01/2017 00:06:55 1 01/01/2017 00:08:00  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/0001 00:00:00 B21.00  
## 5 01/01/0001 00:00:00 B23.20  
## 6 01/01/2017 00:23:00 C14.20  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 MERCHANT/CUSTOMER DISPUTE; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 OTHER INTER/CONFLICT; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address city  
## 1 PH 2 W601 1 40 JONES ST; JERSEY CITY  
## 2 PH 2 W101 0 40 JONES ST; JERSEY CITY  
## 3 PH 5 W401 1 20 TONNELE AVE; 3B JERSEY CITY  
## 4 PH 5 E501 1 2 FAIRMOUNT AVE; JERSEY CITY  
## 5 PH 2 S201 1 105 WILKINSON AVE; JERSEY CITY  
## 6 911 4 W201 1 2 WILMOT AVE; 13 JERSEY CITY  
## latitude longitude geo.count geo.error  
## 1 40.73022 -74.06236 1 S5HPNTSCZA  
## 2 40.73022 -74.06236 1 S5HPNTSCZA  
## 3 40.73093 -74.06704 1 S5HPNTSCZA  
## 4 40.71735 -74.06103 1 S5HPNTSCZA  
## 5 40.70592 -74.08002 1 S5HPNTSCZA  
## 6 0.00000 0.00000 0 Not Found

head(pdcalls[order(call.type),]) #After sorting

## event.number district time.received shift time.dispatched  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## 13 17-000018 W 01/01/2017 00:20:40 1 01/01/2017 00:23:00  
## 15 17-000020 S 01/01/2017 00:23:52 1 01/01/2017 00:25:00  
## 16 17-000021 W 01/01/2017 00:25:23 1 01/01/2017 00:28:00  
## 21 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## 22 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## time.arrived callcode  
## 6 01/01/2017 00:23:00 C14.20  
## 13 01/01/2017 00:28:00 H31.00  
## 15 01/01/0001 00:00:00 F16.00  
## 16 01/01/2017 00:34:00 A12.00  
## 21 01/01/2017 00:38:00 F18.00  
## 22 01/01/2017 00:38:00 F18.00  
## call.code.description  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 13 PERSON DOWN/CAUSE UNKNOWN; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 15 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 16 ASSAULT NO WEAPON; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 21 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 22 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address  
## 6 911 4 W201 1 2 WILMOT AVE; 13   
## 13 911 2 W101 1 EMORY ST & MONTICELLO AVE;   
## 15 911 3 S501 1 12 WEGMAN CT;   
## 16 911 5 W201 1 205 MONTICELLO (MONTICELLO AVE);   
## 21 911 2 S601 0 213 FULTON AVE;   
## 22 911 2 S201 1 213 FULTON AVE;   
## city latitude longitude geo.count geo.error  
## 6 JERSEY CITY 0.00000 0.00000 0 Not Found  
## 13 JERSEY CITY 40.71907 -74.07142 1 SX  
## 15 JERSEY CITY 40.70072 -74.07793 1 S5HPNTSCZA  
## 16 JERSEY CITY 40.72182 -74.06989 1 S5HPNTSCZA  
## 21 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA  
## 22 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA

#### (h). Create a new data frame called calls911 by filtering the original dataset for the 911 calls. Print out the first six rows and check if the filtering worked. How many 911 calls were there?

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

calls911=filter(pdcalls,call.type=='911') #create the dataframe  
head(calls911)

## event.number district time.received shift time.dispatched  
## 1 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## 2 17-000018 W 01/01/2017 00:20:40 1 01/01/2017 00:23:00  
## 3 17-000020 S 01/01/2017 00:23:52 1 01/01/2017 00:25:00  
## 4 17-000021 W 01/01/2017 00:25:23 1 01/01/2017 00:28:00  
## 5 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## 6 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## time.arrived callcode  
## 1 01/01/2017 00:23:00 C14.20  
## 2 01/01/2017 00:28:00 H31.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/2017 00:34:00 A12.00  
## 5 01/01/2017 00:38:00 F18.00  
## 6 01/01/2017 00:38:00 F18.00  
## call.code.description  
## 1 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 2 PERSON DOWN/CAUSE UNKNOWN; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 ASSAULT NO WEAPON; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 6 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address  
## 1 911 4 W201 1 2 WILMOT AVE; 13   
## 2 911 2 W101 1 EMORY ST & MONTICELLO AVE;   
## 3 911 3 S501 1 12 WEGMAN CT;   
## 4 911 5 W201 1 205 MONTICELLO (MONTICELLO AVE);   
## 5 911 2 S601 0 213 FULTON AVE;   
## 6 911 2 S201 1 213 FULTON AVE;   
## city latitude longitude geo.count geo.error  
## 1 JERSEY CITY 0.00000 0.00000 0 Not Found  
## 2 JERSEY CITY 40.71907 -74.07142 1 SX  
## 3 JERSEY CITY 40.70072 -74.07793 1 S5HPNTSCZA  
## 4 JERSEY CITY 40.72182 -74.06989 1 S5HPNTSCZA  
## 5 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA  
## 6 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA

# Number of 911 calls  
nrow(calls911)

## [1] 23098

#### (i). Create a new variable (column) called dispatchduration by subtracting time.received from time.dispatched.

library(dplyr)  
pdcalls$time\_R<-as.POSIXct(strptime(pdcalls$time.received,  
 format = "%m/%d/%Y %H:%M:%OS",tz = "EST"))  
  
pdcalls$time\_D<-as.POSIXct(strptime(pdcalls$time.dispatched,  
 format = "%m/%d/%Y %H:%M:%OS", tz = "EST"))  
  
pdcalls$dispatchduration<-difftime(pdcalls$time\_D, pdcalls$time\_R,  
 units = "secs")  
  
  
head(pdcalls)

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000003 W 01/01/2017 00:05:17 1 01/01/2017 00:07:00  
## 4 17-000004 E 01/01/2017 00:05:52 1 01/01/2017 00:08:00  
## 5 17-000005 S 01/01/2017 00:06:55 1 01/01/2017 00:08:00  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/0001 00:00:00 B21.00  
## 5 01/01/0001 00:00:00 B23.20  
## 6 01/01/2017 00:23:00 C14.20  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 MERCHANT/CUSTOMER DISPUTE; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 OTHER INTER/CONFLICT; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address city  
## 1 PH 2 W601 1 40 JONES ST; JERSEY CITY  
## 2 PH 2 W101 0 40 JONES ST; JERSEY CITY  
## 3 PH 5 W401 1 20 TONNELE AVE; 3B JERSEY CITY  
## 4 PH 5 E501 1 2 FAIRMOUNT AVE; JERSEY CITY  
## 5 PH 2 S201 1 105 WILKINSON AVE; JERSEY CITY  
## 6 911 4 W201 1 2 WILMOT AVE; 13 JERSEY CITY  
## latitude longitude geo.count geo.error time\_R  
## 1 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 2 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 3 40.73093 -74.06704 1 S5HPNTSCZA 2017-01-01 00:05:17  
## 4 40.71735 -74.06103 1 S5HPNTSCZA 2017-01-01 00:05:52  
## 5 40.70592 -74.08002 1 S5HPNTSCZA 2017-01-01 00:06:55  
## 6 0.00000 0.00000 0 Not Found 2017-01-01 00:09:15  
## time\_D dispatchduration  
## 1 2017-01-01 00:06:00 47 secs  
## 2 2017-01-01 00:06:00 47 secs  
## 3 2017-01-01 00:07:00 103 secs  
## 4 2017-01-01 00:08:00 128 secs  
## 5 2017-01-01 00:08:00 65 secs  
## 6 2017-01-01 00:13:00 225 secs

#### (j). Now check if there are missing values in the newly created column and also check for dispatch durations that are negative or zero. This is garbage data so remove these rows.

#Missing values  
sum(!complete.cases(pdcalls$dispatchduration))

## [1] 0

#Negative values  
nrow(pdcalls[pdcalls$dispatchduration<0,])

## [1] 3185

# Zero values  
nrow(pdcalls[pdcalls$dispatchduration==0,])

## [1] 384

# Remove negative and zero values  
pdcalls\_clean=pdcalls[pdcalls$dispatchduration>0,]  
dim(pdcalls\_clean)

## [1] 103711 21

head(pdcalls\_clean)

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000003 W 01/01/2017 00:05:17 1 01/01/2017 00:07:00  
## 4 17-000004 E 01/01/2017 00:05:52 1 01/01/2017 00:08:00  
## 5 17-000005 S 01/01/2017 00:06:55 1 01/01/2017 00:08:00  
## 6 17-000006 W 01/01/2017 00:09:15 1 01/01/2017 00:13:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/0001 00:00:00 F16.00  
## 4 01/01/0001 00:00:00 B21.00  
## 5 01/01/0001 00:00:00 B23.20  
## 6 01/01/2017 00:23:00 C14.20  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 SUSPICIOUS PERSON/CONDITION; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 MERCHANT/CUSTOMER DISPUTE; TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 OTHER INTER/CONFLICT; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## 6 DAMAGE MOTOR VEHICLE; IN PROGRESS AND/OR ACTOR ON SCENE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address city  
## 1 PH 2 W601 1 40 JONES ST; JERSEY CITY  
## 2 PH 2 W101 0 40 JONES ST; JERSEY CITY  
## 3 PH 5 W401 1 20 TONNELE AVE; 3B JERSEY CITY  
## 4 PH 5 E501 1 2 FAIRMOUNT AVE; JERSEY CITY  
## 5 PH 2 S201 1 105 WILKINSON AVE; JERSEY CITY  
## 6 911 4 W201 1 2 WILMOT AVE; 13 JERSEY CITY  
## latitude longitude geo.count geo.error time\_R  
## 1 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 2 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 3 40.73093 -74.06704 1 S5HPNTSCZA 2017-01-01 00:05:17  
## 4 40.71735 -74.06103 1 S5HPNTSCZA 2017-01-01 00:05:52  
## 5 40.70592 -74.08002 1 S5HPNTSCZA 2017-01-01 00:06:55  
## 6 0.00000 0.00000 0 Not Found 2017-01-01 00:09:15  
## time\_D dispatchduration  
## 1 2017-01-01 00:06:00 47 secs  
## 2 2017-01-01 00:06:00 47 secs  
## 3 2017-01-01 00:07:00 103 secs  
## 4 2017-01-01 00:08:00 128 secs  
## 5 2017-01-01 00:08:00 65 secs  
## 6 2017-01-01 00:13:00 225 secs

#### (k). Find the average (mean) dispatch duration using the new variable (column) you created above.

mean(pdcalls\_clean$dispatchduration)

## Time difference of 95.01306 secs

#### (l). Find the average (mean) dispatch duration by call type

library(doBy)  
summaryBy(dispatchduration~call.type, data=pdcalls\_clean, FUN=mean)

## call.type dispatchduration.mean  
## 1 911 131.31167  
## 2 IC 265.50000  
## 3 MVS 55.15983  
## 4 PH 113.00391  
## 5 SI 45.95293  
## 6 STA 119.40000

#### (m). How many rows contain the word GUNSHOTS in the call code description column.

length(grep("GUNSHOTS", pdcalls$call.code.description))

## [1] 1348

#### (n). Now create a data frame called gunshotdata that has just the rows that contain the word GUNSHOTS in the call code description

library(dplyr)  
gunshotdata=dplyr::filter(pdcalls, grepl('GUNSHOTS', call.code.description))  
head(gunshotdata)

## event.number district time.received shift time.dispatched  
## 1 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 2 17-000002 W 01/01/2017 00:05:13 1 01/01/2017 00:06:00  
## 3 17-000011 S 01/01/2017 00:15:55 1 01/01/2017 00:17:00  
## 4 17-000011 S 01/01/2017 00:15:55 1 01/01/2017 00:17:00  
## 5 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## 6 17-000027 S 01/01/2017 00:33:17 1 01/01/2017 00:34:00  
## time.arrived callcode  
## 1 01/01/2017 00:20:00 F18.00  
## 2 01/01/2017 00:20:00 F18.00  
## 3 01/01/2017 00:21:00 F18.00  
## 4 01/01/2017 00:21:00 F18.00  
## 5 01/01/2017 00:38:00 F18.00  
## 6 01/01/2017 00:38:00 F18.00  
## call.code.description  
## 1 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 2 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 3 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 4 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 5 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## 6 GUNSHOTS FIRED HEARD ( NO WITNESS / VICTIM ); TIME ELAPSED NOT APPLICABLE; INJURY NOT APPLICABLE  
## call.type priority unit.id is.primary address  
## 1 PH 2 W601 1 40 JONES ST;   
## 2 PH 2 W101 0 40 JONES ST;   
## 3 PH 2 S601 1 SEAVIEW AVE & KENNEDY BLVD;   
## 4 PH 2 S101 0 SEAVIEW AVE & KENNEDY BLVD;   
## 5 911 2 S601 0 213 FULTON AVE;   
## 6 911 2 S201 1 213 FULTON AVE;   
## city latitude longitude geo.count geo.error time\_R  
## 1 JERSEY CITY 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 2 JERSEY CITY 40.73022 -74.06236 1 S5HPNTSCZA 2017-01-01 00:05:13  
## 3 JERSEY CITY 40.69556 -74.09901 2 SX 2017-01-01 00:15:55  
## 4 JERSEY CITY 40.69556 -74.09901 2 SX 2017-01-01 00:15:55  
## 5 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA 2017-01-01 00:33:17  
## 6 JERSEY CITY 40.70488 -74.08734 1 S5HPNTSCZA 2017-01-01 00:33:17  
## time\_D dispatchduration  
## 1 2017-01-01 00:06:00 47 secs  
## 2 2017-01-01 00:06:00 47 secs  
## 3 2017-01-01 00:17:00 65 secs  
## 4 2017-01-01 00:17:00 65 secs  
## 5 2017-01-01 00:34:00 43 secs  
## 6 2017-01-01 00:34:00 43 secs