# Assignment (13.2) 20- Jan 2018

Sol:-

temp <- aggregate(crime$crime, by=list(crime$crime, crime$time.tag), FUN=length)

names(temp) <- c("crime", "time.tag", "count")

library(plyr)

temp <- ddply(crime, .(crime, day), summarise, count = length(date))

library(doBy)

temp <- summaryBy(Case ~ crime + month, data = crime, FUN= length)

                    names(temp)[3] <- 'count'

crime.agg <- ddply(crime, .(crime, Arrest, Beat, date, `X Coordinate`, `Y Coordinate`, time.tag, day, month),

                   summarise, count=length(date), .progress='text')

beats <- sort(unique(crime.agg$Beat))

dates <- sort(as.character(unique(crime.agg$date)))

temp <- expand.grid(beats, dates)

names(temp) <- c("Beat", "date")

model.data <- aggregate(crime.agg[, c('count', 'Arrest')], by=

                        list(crime.agg$Beat, as.character(crime.agg$date)), FUN=sum)

names(model.data) <- c("Beat", "date", "count", "Arrest")

model.data <- merge(temp, model.data, by= c('Beat', 'date'), all.x= TRUE)

View(model.data)

model.data$count[[is.na](http://is.na/" \t "_blank)(model.data$count)] <- 0

model.data$Arrest[[is.na](http://is.na/" \t "_blank)(model.data$Arrest)] <- 0

model.data$day <- weekdays(as.Date(model.data$date), abbreviate= TRUE)

model.data$month <- months(as.Date(model.data$date), abbreviate= TRUE)

pastDays <- function(x) {c(0, rep(1, x))}

model.data$past.crime.1 <- ave(model.data$count, model.data$Beat,

                               FUN=function(x) filter(x, pastDays(1), sides= 1))

model.data$past.crime.7 <- ave(model.data$count, model.data$Beat,

                               FUN=function(x) filter(x, pastDays(7), sides= 1))

model.data$past.crime.30 <- ave(model.data$count, model.data$Beat,

                               FUN=function(x) filter(x, pastDays(30), sides= 1))

meanNA <- function(x){mean(x, na.rm= TRUE)}

model.data$past.crime.1 <- ifelse([is.na](http://is.na/" \t "_blank)(model.data$past.crime.1),

                                    meanNA(model.data$past.crime.1), model.data$past.crime.1)

model.data$past.crime.7 <- ifelse([is.na](http://is.na/" \t "_blank)(model.data$past.crime.7),

                                    meanNA(model.data$past.crime.7), model.data$past.crime.7)

model.data$past.crime.30 <- ifelse([is.na](http://is.na/" \t "_blank)(model.data$past.crime.30),

                                     meanNA(model.data$past.crime.30), model.data$past.crime.30)

model.data$past.arrest.30 <- ave(model.data$Arrest, model.data$Beat,

                                 FUN= function(x) filter(x, pastDays(30), sides= 1))

model.data$past.arrest.30 <- ifelse([is.na](http://is.na/" \t "_blank)(model.data$past.arrest.30),

                                    meanNA(model.data$past.arrest.30), model.data$past.arrest.30)

model.data$policing <- ifelse(model.data$past.crime.30 == 0, 0,

                              model.data$past.arrest.30/model.data$past.crime.30)

model.data$crime.trend <- ifelse(model.data$past.crime.30 == 0, 0,

                                 model.data$past.crime.7/model.data$past.crime.30)

model.data$season <- as.factor(ifelse(model.data$month %in% c("Mar", "Apr", "May"), "spring",

                                     ifelse(model.data$month %in% c("Jun", "Jul", "Aug"), "summer",

                                            ifelse(model.data$month %in% c("Sep", "Oct","Nov"), "fall", "winter"))))

model.cor <- cor(model.data[, c("count", "past.crime.1", "past.crime.7",

                                 "past.crime.30","policing", "crime.trend")])

model.cor

psych::cor.plot(model.cor)

mean(model.data$count)

var(model.data$count)

A measure used to indicate the extent to which two random variables change in tandem is known as covariance.

#A measure used to represent how strongly two random variables are related #known as correlation

#Covariance is nothing but a measure of correlation. On the contrary,

#correlation refers to the scaled form of covariance

#The value of correlation takes place between -1 and +1.

#Conversely, the value of covariance lies between -∞ and +∞

#Covariance is affected by the change in scale, i.e. if all the value of one variable is multiplied

#by a constant and all the value of another variable are multiplied, by a similar or different constant, then the covariance is changed.

#As against this, correlation is not influenced by the change in scale

#Correlation is dimensionless, i.e. it is a unit-free measure of the relationship between variables. Unlike covariance,

#where the value is obtained by the product of the units of the two variables