

Use the given link below:

<https://archive.ics.uci.edu/ml/machine-learning-databases/communities>

Note: As the dataset provided at the above link was incomplete i.e. headers were missing, So I Have used Crimes data of Atlanta till 2017

```
library(readr)
library(data.table)

COBRA.YTD2017 <- read_csv("G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA
YTD2017.csv")

View(COBRA.YTD2017)
str(COBRA.YTD2017)
summary(COBRA.YTD2017)
sum(is.na(COBRA.YTD2017))

# finding the missing values
library(VIM)
missingvalue_plot <- aggr(COBRA.YTD2017, col=c('navyblue','red'), numbers=TRUE, sortVars=TRUE,
labels=names(COBRA.YTD2017), cex.axis=.7, gap=3, ylab=c("Histogram of missing data","Pattern"))

COBRA_YTD<-COBRA.YTD2017[complete.cases(COBRA.YTD2017), ]

missingvalue_plot <- aggr(COBRA_YTD, col=c('green','yellow'), numbers=TRUE, sortVars=TRUE,
labels=names(COBRA_YTD), cex.axis=.7, gap=3, ylab=c("Histogram of missing data","Pattern"))
```

```

> library(readr)
> library(data.table)
data.table 1.11.8 Latest news: r-databatable.com
> COBRA.YTD2017 <- read_csv("G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv")
Parsed with column specification:
cols(
  .default = col_character(),
  MI_PRINX = col_double(),
  offense_id = col_double(),
  occur_time = col_time(format = ""),
  poss_time = col_time(format = ""),
  beat = col_double(),
  dispo_code = col_double(),
  Maxofnum_victims = col_double(),
  loc_type = col_double(),
  x = col_double(),
  y = col_double()
)

```

See spec(...) for full column specifications.

Warning: 9 parsing failures.

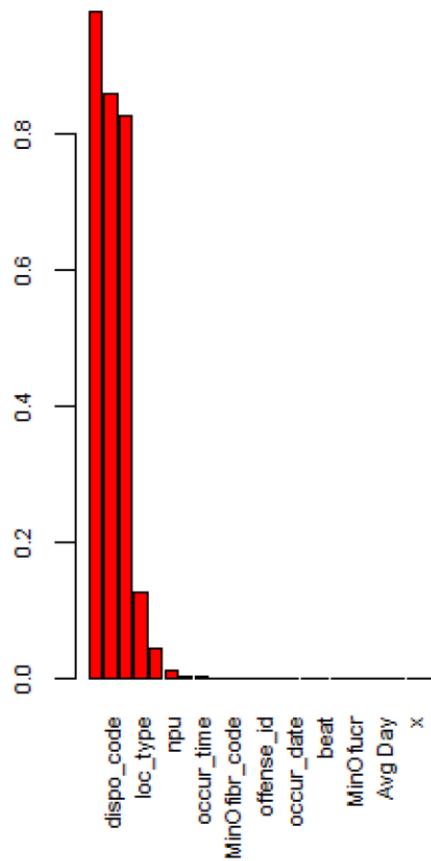
row	col	expected	actual	file
3239	dispo_code	a double	COS	'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
7945	dispo_code	a double	ADM	'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
8527	dispo_code	a double	ADM	'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
10145	dispo_code	a double	ADM	'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
11912	dispo_code	a double	ADM	'G:/DATA ANALYTICS/DATA/crime-in-atlanta-2017/COBRA-YTD2017.csv'
.....				

See problems(...) for more details.

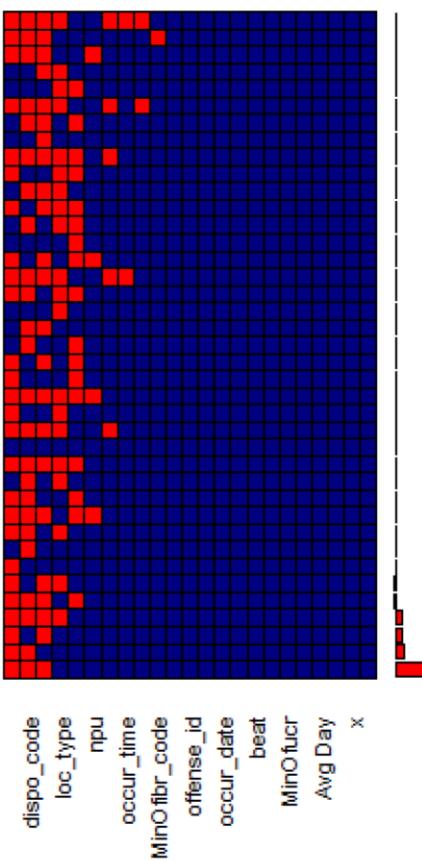
	MI_PRINX	offense_id	rpt_date	occur_date	occur_time	poss_date	poss_time	beat	apt_office_prefix	apt_office_num	location	MinOffcr	MinOffb
1	8924155	173650072	12/31/2017	12/30/2017	23:15:00	12/31/2017	00:30:00	510	NA	NA	43 JESSE HILL JR DR NE	0640	2305
2	8924156	173650102	12/31/2017	12/18/2017	13:00:00	12/30/2017	22:00:00	501	NA	NA	1169 ATLANTIC DR NW	0640	2305
3	8924157	173650144	12/31/2017	12/30/2017	22:01:00	12/31/2017	01:00:00	303	NA	NA	633 PRYOR ST SW	0640	2305
4	8924158	173650149	12/31/2017	12/30/2017	20:00:00	12/31/2017	01:06:00	507	NA	NA	333 NELSON ST SW	0640	2305
5	8924159	173650159	12/31/2017	12/31/2017	00:41:00	12/31/2017	00:48:00	409	NA	NA	2348 CASCADE RD SW	0640	2305
6	8924160	173650180	12/31/2017	12/30/2017	23:00:00	12/31/2017	01:26:00	612	NA	NA	1245 GLENWOOD AVE SE	0650	2304
7	8924161	173650230	12/31/2017	12/31/2017	01:55:00	12/31/2017	01:59:00	605	NA	13	351 CHEROKEE AVE SE	0311	1212
8	8924162	173650241	12/31/2017	12/31/2017	00:00:00	12/31/2017	02:00:00	603	NA	NA	461 PONCE DE LEON AVE NE	0640	2305
9	8924163	173650295	12/31/2017	12/30/2017	00:00:00	12/31/2017	03:02:00	605	NA	NA	437 MEMORIAL DR SE	0640	2305
10	8924164	173650389	12/31/2017	12/31/2017	00:00:00	12/31/2017	03:34:00	304	NA	8	1053 LINAM ST SE	0531	2202A
11	8924165	173650449	12/31/2017	12/31/2017	00:40:00	12/31/2017	04:10:00	303	NA	NA	683 PRYOR ST SW	0710	2404
12	8924166	173650562	12/31/2017	12/31/2017	00:00:00	12/31/2017	05:53:00	104	NA	NA	192 CHICAMAUGA AVE SW	0640	2305


```
> sum(is.na(COBRA.YTD2017))
[1] 76194
> |
```

Histogram of missing data



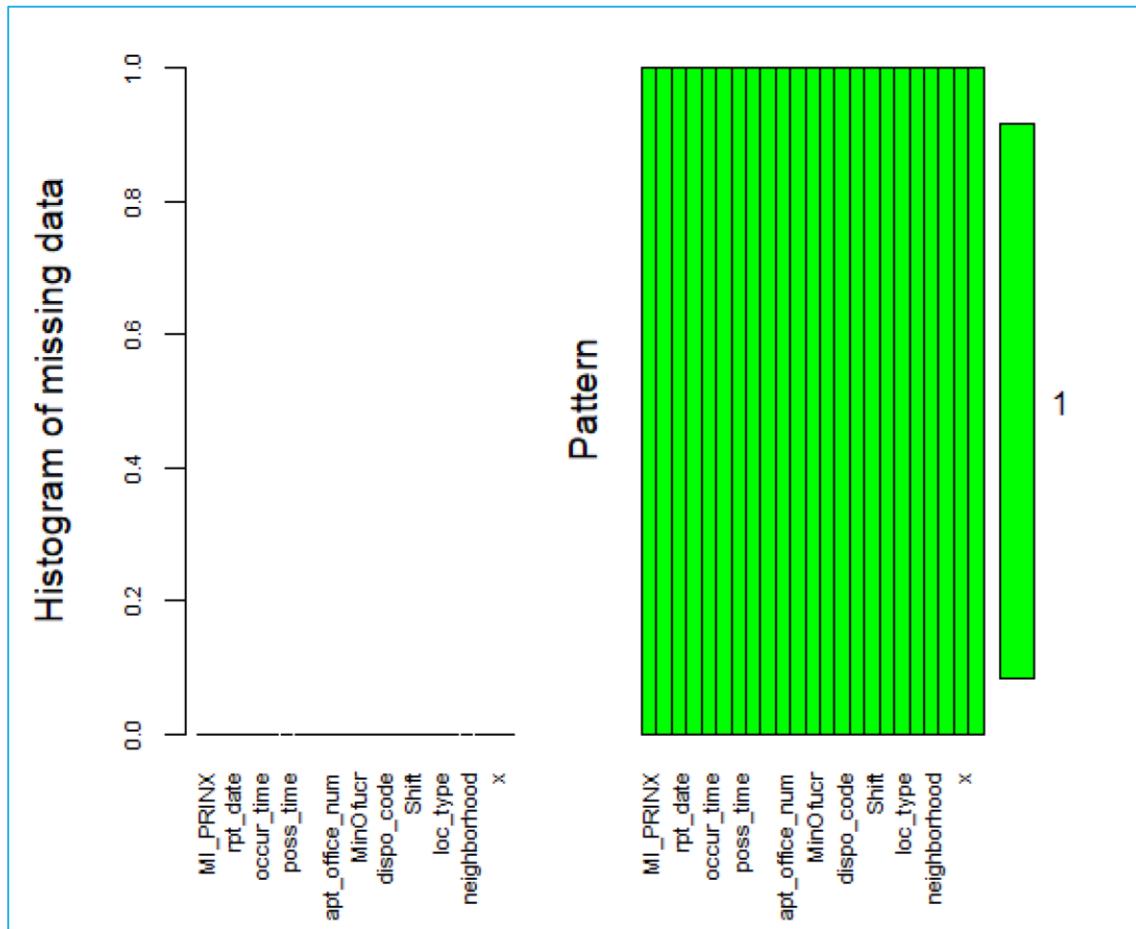
Pattern



```
> missingvalue_plot <- aggr(COBRA_YTD, col=c('black','yellow'), numbers=TRUE, sortVars=TRUE  
sing data","Pattern"))
```

Variables sorted by number of missings:

variable	Count
MI_PRINX	0
offense_id	0
rpt_date	0
occur_date	0
occur_time	0
poss_date	0
poss_time	0
beat	0
apt_office_prefix	0
apt_office_num	0
location	0
MinOfucr	0
MinOfibr_code	0
dispo_code	0
Maxofnum_victims	0
shift	0
Avg Day	0
loc_type	0
uc2_literal	0
neighborhood	0
npu	0
x	0
y	0



Perform the below operations:

- Find out top 5 attributes having highest correlation (select only Numeric features)

#a. Find out top 5 attributes having highest correlation (select only Numeric features).

```
fit<-lm(beat~MinOfucr+MaxOfnum_victims+loc_type+neighborhood+x+y,data =COBRA.YTD2017,
na.action = na.omit)
fit
summary(fit)
fit1<-lm(formula=MinOfucr~beat+MaxOfnum_victims+loc_type+neighborhood+x+y,data
=COBRA.YTD2017)
fit1
summary(fit1)

vif(fit)
vif(fit1)
vif(fit)>5
vif(fit1)>5
```

```
> fit<-lm(beat~MinOfucr+MaxOfnum_victims+loc_type+neighborhood+x+y,data =COBRA.YTD2017, na.action = na.omit)
> fit

call:
lm(formula = beat ~ MinOfucr + MaxOfnum_victims + loc_type +
neighborhood + x + y, data = COBRA.YTD2017, na.action = na.omit)

Coefficients:
              (Intercept)          MinOfucr0220          MinOfucr0311
                2.995e+02           -1.754e+01           4.436e+03
          MinOfucr0312           4.435e+03           4.436e+03
          MinOfucr0315           4.438e+03           4.435e+03
          MinOfucr0321           4.436e+03           4.437e+03
          MinOfucr0324           4.436e+03           4.436e+03
          MinOfucr0331           4.449e+03           4.435e+03
          MinOfucr0334           4.435e+03           4.435e+03
          MinOfucr0337           4.436e+03           4.432e+03
          MinOfucr0343           4.440e+03           4.435e+03
          MinOfucr0347           4.435e+03           4.434e+03
MinOfucr0220           -1.754e+01           4.436e+03
MinOfucr0313            4.436e+03           4.437e+03
MinOfucr0316            4.435e+03           4.438e+03
MinOfucr0322            4.437e+03           4.436e+03
MinOfucr0325            4.436e+03           4.434e+03
MinOfucr0332            4.435e+03           4.435e+03
MinOfucr0335            4.435e+03           4.438e+03
MinOfucr0341            4.432e+03           4.439e+03
MinOfucr0344            4.435e+03           4.429e+03
MinOfucr0410            4.434e+03           4.438e+03
```

```

> summary(fit)

Call:
lm(formula = beat ~ Minofucr + Maxofnum_victims + loc_type +
neighborhood + x + y, data = COBRA.YTD2017, na.action = na.omit)

Residuals:
    Min      1Q  Median      3Q     Max 
-502.97 -0.79   0.05   1.11  400.21 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.995e+02 3.177e+00 94.280 < 2e-16 ***
Minofucr0220 -1.754e+01 9.950e+00 -1.770 0.07667 .  
Minofucr0311 4.436e+03 2.674e+03 1.659 0.09715 .  
Minofucr0312 4.435e+03 2.674e+03 1.659 0.09717 .  
Minofucr0313 4.436e+03 2.674e+03 1.659 0.09714 .  
Minofucr0314 4.437e+03 2.674e+03 1.660 0.09702 .  
Minofucr0315 4.438e+03 2.674e+03 1.660 0.09698 .  
Minofucr0316 4.435e+03 2.674e+03 1.659 0.09720 .  
Minofucr0317 4.438e+03 2.674e+03 1.660 0.09700 .  
Minofucr0321 4.436e+03 2.674e+03 1.659 0.09709 .  
Minofucr0322 4.437e+03 2.674e+03 1.659 0.09707 .  
Minofucr0323 4.436e+03 2.674e+03 1.659 0.09717 .  
Minofucr0324 4.436e+03 2.674e+03 1.659 0.09712 .  
Minofucr0325 4.436e+03 2.674e+03 1.659 0.09715 .  
Minofucr0327 4.434e+03 2.674e+03 1.658 0.09730 .  
Minofucr0331 4.449e+03 2.674e+03 1.664 0.09615 .  
Minofucr0332 4.435e+03 2.674e+03 1.659 0.09715 .  
Minofucr0333 4.435e+03 2.674e+03 1.659 0.09719 .  
Minofucr0334 4.435e+03 2.674e+03 1.659 0.09722 .  
Minofucr0335 4.435e+03 2.674e+03 1.658 0.09724 .  
Minofucr0336 4.438e+03 2.674e+03 1.660 0.09693 .  
Minofucr0337 4.436e+03 2.674e+03 1.659 0.09714 .  
Minofucr0341 4.432e+03 2.674e+03 1.657 0.09744 .  
Minofucr0342 4.439e+03 2.674e+03 1.660 0.09688 .  
Minofucr0343 4.440e+03 2.674e+03 1.661 0.09679 .  
Minofucr0344 4.435e+03 2.674e+03 1.659 0.09722 .  
Minofucr0345 4.429e+03 2.674e+03 1.656 0.09768 .  
Minofucr0347 4.435e+03 2.674e+03 1.659 0.09718 .  
Minofucr0410 4.434e+03 2.674e+03 1.658 0.09731 .  
Minofucr0420 4.438e+03 2.674e+03 1.660 0.09699 .  
Minofucr0430 4.436e+03 2.674e+03 1.659 0.09711 .  
Minofucr0440 4.440e+03 2.674e+03 1.660 0.09683 .  
Minofucr0511 4.435e+03 2.674e+03 1.659 0.09720 .  
Minofucr0512 4.435e+03 2.674e+03 1.659 0.09717 .  
Minofucr0521 4.435e+03 2.674e+03 1.659 0.09717 .  
Minofucr0522 4.438e+03 2.674e+03 1.660 0.09700 .  
Minofucr0531 4.434e+03 2.674e+03 1.658 0.09727 .  
Minofucr0532 4.440e+03 2.674e+03 1.660 0.09684 .  
Minofucr0610 4.436e+03 2.674e+03 1.659 0.09713 .  
Minofucr0620 4.428e+03 2.674e+03 1.656 0.09775 .  
Minofucr0630 4.436e+03 2.674e+03 1.659 0.09711 .  
Minofucr0640 4.436e+03 2.674e+03 1.659 0.09714 .  
Minofucr0650 4.436e+03 2.674e+03 1.659 0.09712 .  
Minofucr0660 4.436e+03 2.674e+03 1.659 0.09709 .  
Minofucr0670 4.435e+03 2.674e+03 1.659 0.09721 .  
Minofucr0680 4.443e+03 2.674e+03 1.662 0.09659 .

```

Console

```

neighborhoodGrant Park           3.034e+02  2.792e+00 108.654 < 2e-16 ***
neighborhoodGreen Acres Valley   1.034e+02  1.183e+01  8.744 < 2e-16 ***
neighborhoodGreen Forest Acres   1.038e+02  9.532e+00 10.891 < 2e-16 ***
neighborhoodGreenbriar          1.017e+02  3.669e+00 27.723 < 2e-16 ***
neighborhoodGreenbriar Village   1.054e+02  9.953e+00 10.589 < 2e-16 ***
neighborhoodGrove Park          -1.867e+02 3.108e+00 -60.074 < 2e-16 ***
neighborhoodHammond Park        1.245e+00  3.379e+00  0.368 0.71261
neighborhoodHanover West        -8.289e+01 1.192e+01 -6.956 3.59e-12 ***
neighborhoodHarland Terrace     6.973e+01  3.438e+00 20.282 < 2e-16 ***
neighborhoodHarris Chiles       -1.898e+02 3.977e+00 -47.723 < 2e-16 ***
neighborhoodHarvel Homes Community -1.957e+02 1.832e+01 -10.681 < 2e-16 ***
neighborhoodHeritage Valley     1.056e+02  7.222e+00 14.622 < 2e-16 ***
neighborhoodHigh Point          1.494e+00  6.541e+00  0.228 0.81933
neighborhoodHills Park          -8.706e+01 4.349e+00 -20.020 < 2e-16 ***
neighborhoodHome Park           2.077e+02  2.970e+00 69.910 < 2e-16 ***
neighborhoodHorseshoe Community 1.050e+02  1.826e+01  5.750 9.07e-09 ***
neighborhoodHunter Hills         -1.900e+02 3.301e+00 -57.576 < 2e-16 ***
neighborhoodHuntington          9.959e+01  1.531e+01  6.503 8.02e-11 ***
neighborhoodInman Park          3.063e+02  3.092e+00 99.063 < 2e-16 ***
neighborhoodIvan Hill           1.064e+02  8.624e+00 12.336 < 2e-16 ***
neighborhoodJoyland             2.876e+00  5.018e+00  0.573 0.56647
neighborhoodJust Us              -1.945e+02 2.549e+01 -7.628 2.48e-14 ***
neighborhoodKings Forest         1.038e+02  5.123e+00 20.258 < 2e-16 ***
neighborhoodKingswood            -7.802e+01 1.846e+01 -4.227 2.38e-05 ***
neighborhoodKirkwood             3.136e+02  3.882e+00 80.783 < 2e-16 ***
neighborhoodKnight Park/Howell station -1.871e+02 5.181e+00 -36.112 < 2e-16 ***
neighborhoodLake Claire          3.132e+02  4.936e+00 63.462 < 2e-16 ***
neighborhoodLake Estates         1.017e+02  2.577e+01  3.945 7.99e-05 ***
neighborhoodLakewood             3.189e+00  4.945e+00  0.645 0.51903
neighborhoodLakewood Heights     1.256e+00  2.813e+00  0.447 0.65524
neighborhoodLaurens Valley       1.028e+02  1.819e+01  5.651 1.62e-08 ***
neighborhoodLeila Valley          9.150e-01  4.786e+00  0.191 0.84838
neighborhoodLenox                -7.186e+01 4.580e+00 -15.690 < 2e-16 ***
neighborhoodLincoln Homes        -1.777e+02 6.240e+00 -28.485 < 2e-16 ***
neighborhoodLindbergh/Morosgo    -7.476e+01 3.924e+00 -19.052 < 2e-16 ***
neighborhoodLindridge/Martin Manor -7.303e+01 4.154e+00 -17.579 < 2e-16 ***
neighborhoodLoring Heights       -8.428e+01 3.594e+00 -23.452 < 2e-16 ***
neighborhoodMagnum Manor         1.026e+02  9.038e+00 11.351 < 2e-16 ***
neighborhoodMargaret Mitchell    -8.366e+01 1.108e+01 -7.549 4.57e-14 ***
neighborhoodMarietta Street Artery 2.108e+02  3.381e+00 62.332 < 2e-16 ***
neighborhoodMays                  1.069e+02  4.777e+00 22.377 < 2e-16 ***
neighborhoodMeadowbrook Forest   1.037e+02  8.465e+00 12.245 < 2e-16 ***
neighborhoodMechanicsville      3.271e+00  2.492e+00  1.312 0.18938
neighborhoodMellwood             -5.193e+01 1.834e+01 -2.832 0.00463 **
neighborhoodMemorial Park        -8.163e+01 1.830e+01 -4.461 8.20e-06 ***
neighborhoodMidtown              2.107e+02  2.809e+00 75.008 < 2e-16 ***
neighborhoodMidwest Cascade      1.085e+02  5.999e+00 18.085 < 2e-16 ***
neighborhoodMonroe Heights        -1.788e+02 5.409e+00 -33.066 < 2e-16 ***
neighborhoodMorningside/Lenox Park -7.618e+01 3.660e+00 -20.816 < 2e-16 ***
[ reached getoption("max.print") -- omitted 90 rows ]
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 25.39 on 22091 degrees of freedom
(4378 observations deleted due to missingness)
Multiple R-squared:  0.9774,    Adjusted R-squared:  0.9771
F-statistic: 3306 on 289 and 22091 DF,  p-value: < 2.2e-16

```

```

> fit1<-lm(formula=Minofucr~beat+MaxOfnum_victims+loc_type+neighborhood+x+y,data =COBRA.YTD2017)
> fit1

Call:
lm(formula = Minofucr ~ beat + MaxOfnum_victims + loc_type +
    neighborhood + x + y, data = COBRA.YTD2017)

Coefficients:
(Intercept)                                beat
2.218e+02                                    7.126e-03
loc_type                                     -4.508e-02
neighborhoodAlmond Park                     neighborhoodAdams Park
-1.903e+01                                    -9.749e+00
neighborhoodArden/Habersham                  neighborhoodAmar Heights
-3.377e+00                                    -1.164e-01
neighborhoodArlington Estates                neighborhoodArdmore
-3.531e+01                                    2.239e+01
neighborhoodAtkins Park                      neighborhoodAshley Courts
3.540e+00                                     -8.837e+00
neighborhoodAtlantic Station                 neighborhoodAtlanta Industrial Park
1.603e+01                                     4.923e+01
neighborhoodBaker Hills                      neighborhoodAudobon Forest
-2.243e+01                                    8.961e+00
neighborhoodBankhead/Bolton                  neighborhoodBakers Ferry
3.446e+01                                     -7.371e+00
neighborhoodBen Hill Acres                   neighborhoodBeecher Hills
-1.669e+01                                    4.918e+00
neighborhoodBen Hill Forest                 neighborhoodBen Hill Forest
                                         1.236e+02
                                         MaxOfnum_victims
                                         -1.229e+01
neighborhoodAdamsville                      neighborhoodAdams Park
                                         -9.269e+00
neighborhoodAnsley Park                     neighborhoodAmar Heights
                                         -1.197e+00
neighborhoodArgonne Forest                  neighborhoodArdmore
                                         6.209e+01
neighborhoodAshview Heights                 neighborhoodAshley Courts
                                         -3.203e+01
neighborhoodAtlanta University Center        neighborhoodAtlanta Industrial Park
                                         6.809e-01
neighborhoodAudobon Forest West             neighborhoodAudobon Forest
                                         5.861e+01
neighborhoodBankhead                         neighborhoodBakers Ferry
                                         -7.140e+01
neighborhoodBen Hill                         neighborhoodBeecher Hills
                                         -3.544e+00
neighborhoodBen Hill Pines                  neighborhoodBen Hill Forest
                                         -4.936e+00

```

```

> summary(fit1)

Call:
lm(formula = Minofucr ~ beat + MaxOfnum_victims + loc_type +
neighborhood + x + y, data = COBRA.YTD2017)

Residuals:
    Min      1Q  Median      3Q     Max 
-333.92   -35.53   20.95   61.95  429.07 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2.218e+02  1.461e+01 15.184 < 2e-16 ***
beat         7.126e-03  2.641e-02  0.270  0.787314    
MaxOfnum_victims -1.229e+01  1.003e+00 -12.255 < 2e-16 ***
loc_type      -4.508e-02  4.229e-02 -1.066  0.286378    
neighborhoodAdams Park -9.749e+00  1.633e+01 -0.597  0.550649    
neighborhoodAdamsville -9.269e+00  1.117e+01 -0.830  0.406568    
neighborhoodAlmond Park -1.903e+01  1.808e+01 -1.052  0.292635    
neighborhoodAmal Heights -1.164e-01  2.291e+01 -0.005  0.995947    
neighborhoodAnsley Park -1.197e+00  1.866e+01 -0.064  0.948858    
neighborhoodArden/Habersham -3.377e+00  4.704e+01 -0.072  0.942767    
neighborhoodArdmore 2.239e+01  2.280e+01  0.982  0.326053    
neighborhoodArgonne Forest 6.209e+01  3.833e+01  1.620  0.105302    
neighborhoodArlington Estates -3.531e+01  2.595e+01 -1.361  0.173611    
neighborhoodAshley courts -8.837e+00  1.854e+01 -0.477  0.633634    
neighborhoodAshview Heights -3.203e+01  1.351e+01 -2.371  0.017734 *  
neighborhoodAtkins Park 3.540e+00  5.926e+01  0.060  0.952370    
neighborhoodAtlanta Industrial Park 4.923e+01  2.278e+01  2.161  0.030671 *  
neighborhoodAtlanta University Center 6.809e-01  1.396e+01  0.049  0.961095    
neighborhoodAtlantic Station 1.603e+01  1.448e+01  1.107  0.268284    
neighborhoodAudobon Forest 8.961e+00  2.658e+01  0.337  0.735972    
neighborhoodAudobon Forest west 5.861e+01  3.888e+01  1.507  0.131752    
neighborhoodBaker Hills -2.243e+01  2.084e+01 -1.076  0.281915    
neighborhoodBakers Ferry -7.371e+00  4.557e+01 -0.162  0.871524    
neighborhoodBankhead -7.140e+01  1.342e+01 -5.322  1.04e-07 ***
neighborhoodBankhead/Bolton 3.446e+01  2.491e+01  1.383  0.166544    
neighborhoodBeecher Hills 4.918e+00  2.909e+01  0.169  0.865739    
neighborhoodBen Hill -3.544e+00  2.230e+01 -0.159  0.873721    
neighborhoodBen Hill Acres -1.669e+01  2.347e+01 -0.711  0.476929    
neighborhoodBen Hill Forest 1.236e+02  7.181e+01  1.722  0.085147 . 
neighborhoodBen Hill Pines -4.936e+00  4.275e+01 -0.115  0.908080    
neighborhoodBen Hill Terrace 1.217e+00  2.194e+01  0.055  0.955768    
neighborhoodBenteen Park -2.507e+01  1.990e+01 -1.260  0.207583    
neighborhoodBerkeley Park 2.098e+01  1.298e+01  1.616  0.106162    
neighborhoodBetmar Lavilla 2.739e+01  1.687e+01  1.623  0.104539    
neighborhoodBlair Villa/Poole creek 2.762e+01  1.720e+01  1.606  0.108317    
neighborhoodBlandtown 3.626e+01  1.273e+01  2.848  0.004406 **  
neighborhoodBolton 1.308e+01  1.527e+01  0.857  0.391730    
neighborhoodBolton Hills 2.209e+00  4.210e+01  0.052  0.958159    
neighborhoodBoulder Park -2.241e+01  4.180e+01 -0.536  0.591895    
neighborhoodBoulevard Heights 4.861e+00  1.947e+01  0.250  0.802878    
neighborhoodBrandon 5.670e+01  3.161e+01  1.794  0.072852 . 
neighborhoodBrentwood -7.443e+01  4.012e+01 -1.855  0.063549 . 
neighborhoodBriar Glen 4.873e+01  4.238e+01  1.150  0.250262    
neighborhoodBrookhaven 4.735e+00  3.674e+01  0.129  0.897448    
neighborhoodBrookview Heights 9.546e+00  2.162e+01  0.441  0.658876    
neighborhoodBrookwood -8.286e+00  2.136e+01 -0.388  0.698026

```

neighborhoodMorningside/Lenox Park	1.639e+01	1.445e+01	1.135	0.256477
neighborhoodMozley Park	-2.716e+01	1.429e+01	-1.900	0.057434 .
neighborhoodMt. Gilead Woods	-4.631e+01	3.944e+01	-1.174	0.240313
neighborhoodMt. Paran Parkway	2.164e+01	1.013e+02	0.213	0.830942
neighborhoodMt. Paran/Northside	-1.300e+01	2.964e+01	-0.439	0.660913
neighborhoodNiskey Cove	3.107e+01	7.154e+01	0.434	0.664004
neighborhoodNiskey Lake	6.724e+01	5.117e+01	1.314	0.188860
neighborhoodNorth Buckhead	1.867e+01	1.819e+01	1.026	0.304853
neighborhoodNorwood Manor	-1.205e+01	1.809e+01	-0.666	0.505368
neighborhoodOak Cliff	3.736e+01	3.895e+01	0.959	0.337453
neighborhoodOakland	4.533e+00	2.489e+01	0.182	0.855482
neighborhoodOakland City	-2.498e+01	1.108e+01	-2.255	0.024165 *
neighborhoodOld Fairburn Village	4.873e+01	1.004e+02	0.485	0.627339
neighborhoodOld Fourth Ward	-2.104e+00	1.343e+01	-0.157	0.875535
neighborhoodOld Gordon	-1.341e+01	2.382e+01	-0.563	0.573585
neighborhoodOrchard Knob	2.727e+00	1.970e+01	0.138	0.889909
neighborhoodOrmewood Park	-2.503e+01	1.470e+01	-1.702	0.088719 .
neighborhoodPaces	3.021e+01	2.446e+01	1.235	0.216919
neighborhoodPeachtree Battle Alliance	5.466e+01	3.078e+01	1.776	0.075811 .
neighborhoodPeachtree Heights East	2.207e+01	3.130e+01	0.705	0.480826
neighborhoodPeachtree Heights West	1.580e+01	1.856e+01	0.852	0.394469
neighborhoodPeachtree Hills	2.217e+01	1.988e+01	1.115	0.264844
neighborhoodPeachtree Park	-1.739e+00	2.034e+01	-0.085	0.931870
neighborhoodPenelope Neighbors	-9.972e+01	3.041e+01	-3.279	0.001043 **
neighborhoodPeoplestown	-2.815e+01	1.172e+01	-2.403	0.016289 *
neighborhoodPerkerson	1.361e+01	1.191e+01	1.143	0.253221
neighborhoodPeyton Forest	2.500e+01	3.278e+01	0.763	0.445761
neighborhoodPiedmont Heights	3.562e-01	1.482e+01	0.024	0.980827
neighborhoodPine Hills	-2.234e+00	1.860e+01	-0.120	0.904364
neighborhoodPittsburgh	-2.804e+00	1.025e+01	-0.274	0.784408
neighborhoodPleasant Hill	3.624e+01	4.720e+01	0.768	0.442645
neighborhoodPolar Rock	4.473e+00	2.027e+01	0.221	0.825378
neighborhoodPomona Park	1.397e+01	7.127e+01	0.196	0.844614
neighborhoodPoncey-Highland	2.103e+01	1.548e+01	1.358	0.174405
neighborhoodPrinceton Lakes	5.334e+01	1.607e+01	3.320	0.000902 ***
neighborhoodRandall Mill	2.865e+01	2.346e+01	1.221	0.221968
neighborhoodRebel Valley Forest	-7.690e+00	1.941e+01	-0.396	0.692056
neighborhoodReynoldstown	-1.769e+01	1.593e+01	-1.111	0.266663
neighborhoodRidgecrest Forest	-1.680e+01	3.149e+01	-0.533	0.593810
neighborhoodRidgedale Park	-5.404e+00	2.641e+01	-0.205	0.837852
neighborhoodRidgewood Heights	1.853e+01	3.132e+01	0.591	0.554223
neighborhoodRiverside	1.021e+01	1.471e+01	0.694	0.487558
neighborhoodRockdale	2.647e+01	1.817e+01	1.457	0.145122
neighborhoodRosedale Heights	-2.206e+01	1.861e+01	-1.185	0.235855
neighborhoodRue Royal	-1.624e+02	7.170e+01	-2.265	0.023504 *
neighborhoodSandlewood Estates	1.123e+01	3.141e+01	0.357	0.720790
neighborhoodScotts Crossing	-2.423e+01	1.787e+01	-1.356	0.175150
neighborhoodSherwood Forest	-3.203e+01	5.168e+01	-0.620	0.535337
neighborhoodSouth Atlanta	-2.876e+01	1.279e+01	-2.247	0.024619 *
[reached getoption("max.print") -- omitted 42 rows]				

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 99.8 on 22139 degrees of freedom				
(4378 observations deleted due to missingness)				
Multiple R-squared: 0.147, Adjusted R-squared: 0.1377				
F-statistic: 15.83 on 241 and 22139 DF, p-value: < 2.2e-16				

```

> vif(fit)
          GVIF Df GVIF^(1/(2*Df))
Minofucr      NaN 49      NaN
Maxofnum_victims  NaN  1      NaN
Loc_type       NaN  1      NaN
neighborhood    NaN 236      NaN
x              NaN  1      NaN
y              NaN  1      NaN
> vif(fit1)
          GVIF Df GVIF^(1/(2*Df))
beat        4.411556e+01  1     6.641955
Maxofnum_victims 1.017245e+00  1     1.008586
Loc_type      1.025699e+00  1     1.012768
neighborhood   3.705006e+03 236     1.017562
x            1.883397e+05  1    433.981262
y            1.884420e+05  1    434.099027
> vif(fit)>5
          GVIF Df GVIF^(1/(2*Df))
Minofucr      NA  TRUE      NA
Maxofnum_victims  NA FALSE      NA
Loc_type       NA FALSE      NA
neighborhood    NA  TRUE      NA
x              NA FALSE      NA
y              NA FALSE      NA
> vif(fit1)>5
          GVIF Df GVIF^(1/(2*Df))
beat        TRUE FALSE      TRUE
Maxofnum_victims FALSE FALSE      FALSE
Loc_type      FALSE FALSE      FALSE
neighborhood   TRUE  TRUE      FALSE
x            TRUE FALSE      TRUE
y            TRUE FALSE      TRUE
>

```

- a) P-values are very important because, we can consider linear model to be statistically significant only.
- b) When both these p-values are less than the pre-determined statistical determined level, which is ideally 0.05.
- c) This is visually interpreted by the significance stars at the end of the row.
- d) The more stars beside the variable's p-value, the more significant the variable.
- e) When there is a p-value, there is a null and alternative hypothesis associated with it.
- f) Null and Alternate Hypothesis
- g) In Linear regression, the Null Hypothesis is that the coefficients associated with the variables is equal to zero.
- h) The alternate hypothesis is that the coefficients are not equal to zero.
- i) There exists a relationship between the independent variable in question and the dependent variable.

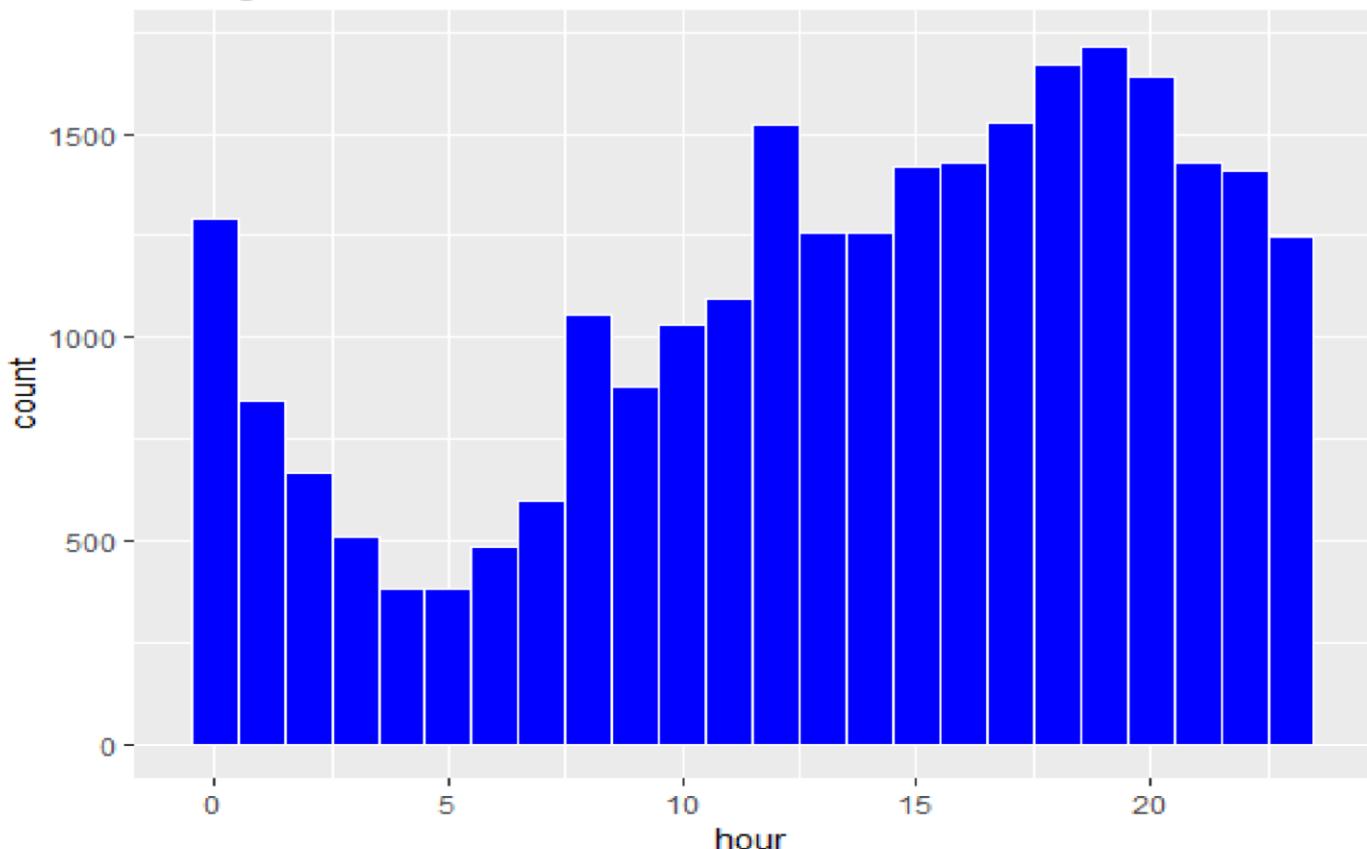
b) Find out top 3 reasons for having more crime in a city.

#b. Find out top 3 reasons for having more crime in a city.

```
library(ggplot2)
COBRA.YTD2017$hour <- sub(":.*", "", COBRA.YTD2017$occur_time)
COBRA.YTD2017$hour <- as.numeric(COBRA.YTD2017$hour)
ggplot(aes(x = hour), data = COBRA.YTD2017) + geom_histogram(bins = 24, color='white', fill='blue')
+
  ggttitle('Histogram of Crime Time')
```

```
UC2<-table(COBRA.YTD2017$`UC2 Literal`)
hist(UC2)
```

Histogram of Crime Time

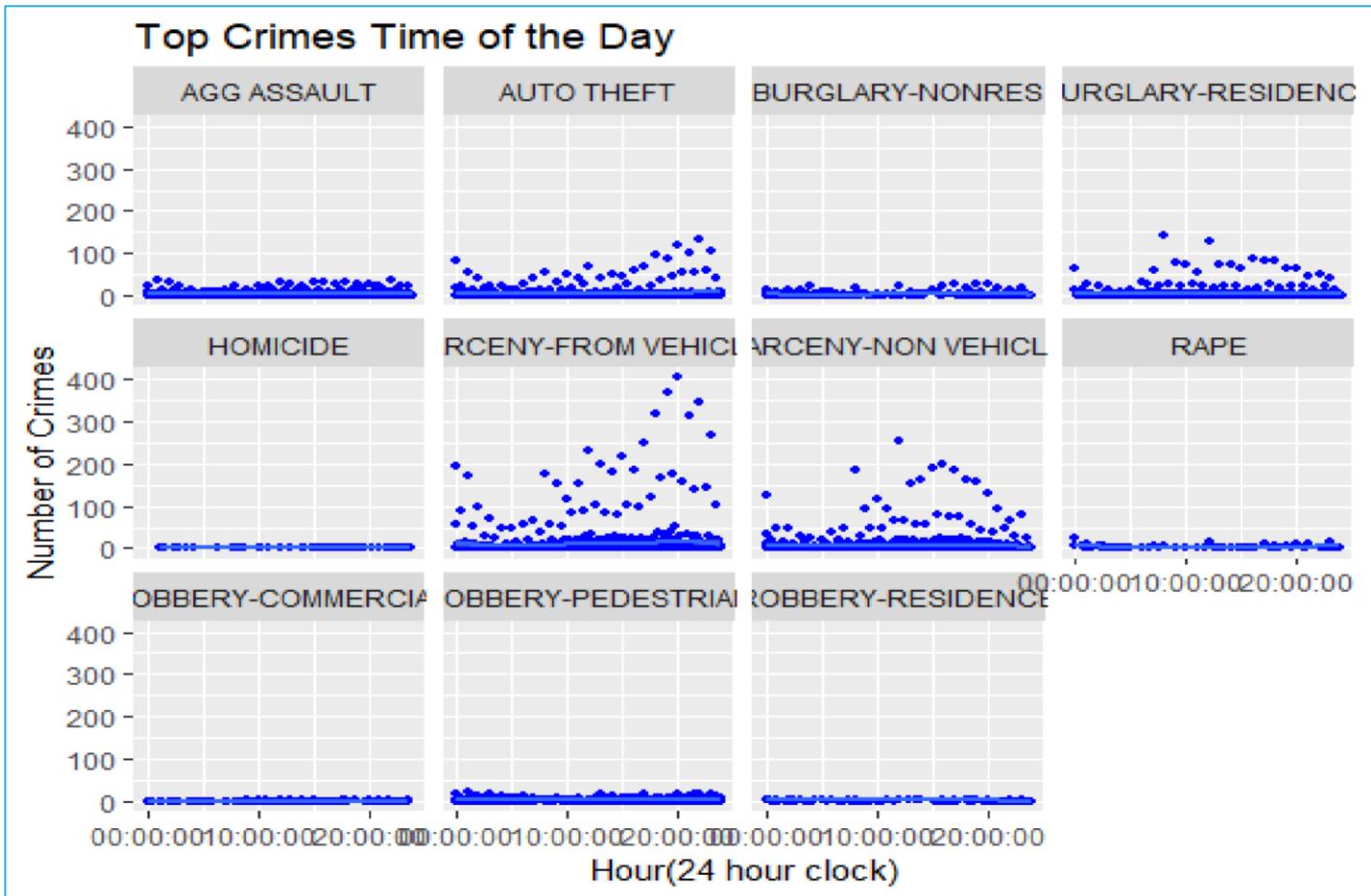


```

library(dplyr)
COBRA <- COBRA.YTD2017 %>% group_by(`UC2 Literal`, occur_time) %>%
  summarise(total = n())

ggplot(aes(x = occur_time, y = total), data = COBRA) + geom_point(colour="blue", size=1) +
  geom_smooth(method="loess") + xlab('Hour(24 hour clock)') +
  ylab('Number of Crimes') + ggtitle('Top Crimes Time of the Day') + facet_wrap(~`UC2 Literal`)

```



#Downtown and midtown are the most common locations where crimes take place, followed by Old Fourth Ward and West End. larceny theft are the top crimes in Atlanta followed by aggravated assault

```

library(knitr)
library(kableExtra)

kable(count(COBRA.YTD2017, COBRA.YTD2017$`UC2 Literal`, sort=TRUE), "html",
col.names=c("Crime Type", "Frequency")) %>%
  kable_styling(bootstrap_options="striped", full_width=FALSE)

```

Crime Type	Frequency
LARCENY-FROM VEHICLE	9840
LARCENY-NON VEHICLE	6589
AUTO THEFT	3197
BURGLARY-RESIDENCE	2635
AGG ASSAULT	2024
ROBBERY-PEDESTRIAN	1126
BURGLARY-NONRES	758
RAPE	226
ROBBERY-COMMERCIAL	157
ROBBERY-RESIDENCE	132
HOMICIDE	75

c) Which all attributes have high correlation with crime rate?

#c. Which all attributes have correlation with crime rate?

```
library(ggplot2)
```

```
library(corrplot)
```

pairs(COBRA.YTD2017)

```
rank1<-sample(COBRA.YTD2017[1:100,22:23], 20, replace=T)
```

```
rank2<-sample(COBRA.YTD2017[1:100,22:23], 20, replace=T)
```

`cbind(rank1,rank2)`

plot(rank1, rank2)

```
cor(rank1,rank2, method="spearman")
```

```
cor(rank1,rank2, method="pearson")
```

```

> #c. which all attributes have correlation with crime rate?
> library(ggplot2)
> library(corrplot)
> rank1<-sample(COBRA.YTD2017[1:100,22:23], 20, replace=T)
> rank2<-sample(COBRA.YTD2017[1:100,22:23], 20, replace=T)
> cbind(rank1,rank2)

      x     y     y     y     y     y     x     x     x     x     x     x     x     x     x     x     x     x     x     x     x     x     x     x     y
1 -84.38013 33.75582 33.75582 33.75582 33.75582 -84.38013 33.75582 -84.38013 -84.38013 33.75582 -84.38013 -84.38013 33.75582
2 -84.39745 33.78674 33.78674 33.78674 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674
3 -84.39486 33.73760 33.73760 33.73760 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760
4 -84.39887 33.75156 33.75156 33.75156 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156
5 -84.46522 33.72146 33.72146 33.72146 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146
6 -84.34660 33.74006 33.74006 33.74006 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006
7 -84.37373 33.74505 33.74505 33.74505 33.74505 -84.37373 33.74505 -84.37373 33.74505 -84.37373 33.74505 -84.37373 33.74505
8 -84.37190 33.77303 33.77303 33.77303 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303
9 -84.37285 33.74639 33.74639 33.74639 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639
10 -84.38625 33.72579 33.72579 33.72579 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579
11 -84.39495 33.73616 33.73616 33.73616 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616
12 -84.43192 33.74974 33.74974 33.74974 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974
13 -84.40766 33.68243 33.68243 33.68243 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243
14 -84.39276 33.71691 33.71691 33.71691 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691
15 -84.38742 33.73048 33.73048 33.73048 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048
16 -84.46574 33.70087 33.70087 33.70087 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087
17 -84.43107 33.75831 33.75831 33.75831 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831
18 -84.47527 33.79671 33.79671 33.79671 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671
19 -84.38031 33.70678 33.70678 33.70678 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678
20 -84.36983 33.84746 33.84746 33.84746 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746
21 -84.33868 33.75757 33.75757 33.75757 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757
22 -84.38023 33.77067 33.77067 33.77067 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067
23 -84.40795 33.70369 33.70369 33.70369 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369
24 -84.43089 33.75437 33.75437 33.75437 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437
25 -84.42057 33.74647 33.74647 33.74647 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647

      y     y     x     x     y     x     y     x     y     y     x     y     x     y     x     x     x     x     x     x     x     x     x     x     x     x     y
1  33.75582 33.75582 -84.38013 33.75582 -84.38013 33.75582 -84.38013 33.75582 -84.38013 33.75582 -84.38013 33.75582 -84.38013 33.75582
2  33.78674 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674 -84.39745 33.78674
3  33.73760 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760 -84.39486 33.73760
4  33.75156 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156 -84.39887 33.75156
5  33.72146 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146 -84.46522 33.72146
6  33.74006 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006 -84.34660 33.74006
7  33.75405 33.75405 -84.37373 33.75405 -84.37373 33.75405 -84.37373 33.75405 -84.37373 33.75405 -84.37373 33.75405 -84.37373 33.75405
8  33.77303 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303 -84.37190 33.77303
9  33.74639 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639 -84.37285 33.74639
10 33.72579 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579 -84.38625 33.72579
11 33.73616 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616 -84.39495 33.73616
12 33.74974 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974 -84.43192 33.74974
13 33.68243 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243 -84.40766 33.68243
14 33.71691 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691 -84.39276 33.71691
15 33.73048 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048 -84.38742 33.73048
16 33.70087 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087 -84.46574 33.70087
17 33.75831 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831 -84.43107 33.75831
18 33.79671 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671 -84.47527 33.79671
19 33.70678 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678 -84.38031 33.70678
20 33.84746 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746 -84.36983 33.84746
21 33.75757 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757 -84.33868 33.75757
22 33.77067 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067 -84.38023 33.77067
23 33.70369 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369 -84.40795 33.70369
24 33.75437 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437 -84.43089 33.75437
25 33.74647 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647 -84.42057 33.74647
```

	y	y	x	x	y	x	y	x	x
1	33.75582	33.75582	-84.38013	33.75582	-84.38013	33.75582	-84.38013	-84.38013	-84.38013
2	33.78674	33.78674	-84.39745	33.78674	-84.39745	33.78674	-84.39745	-84.39745	-84.39745
3	33.73760	33.73760	-84.39486	33.73760	-84.39486	33.73760	-84.39486	-84.39486	-84.39486
4	33.75156	33.75156	-84.39887	33.75156	-84.39887	33.75156	-84.39887	-84.39887	-84.39887
5	33.72146	33.72146	-84.46522	33.72146	-84.46522	33.72146	-84.46522	-84.46522	-84.46522
6	33.74006	33.74006	-84.34660	33.74006	-84.34660	33.74006	-84.34660	-84.34660	-84.34660
7	33.74505	33.74505	-84.37373	33.74505	-84.37373	33.74505	-84.37373	-84.37373	-84.37373
8	33.77303	33.77303	-84.37190	33.77303	-84.37190	33.77303	-84.37190	-84.37190	-84.37190
9	33.74639	33.74639	-84.37285	33.74639	-84.37285	33.74639	-84.37285	-84.37285	-84.37285
10	33.72579	33.72579	-84.38625	33.72579	-84.38625	33.72579	-84.38625	-84.38625	-84.38625
11	33.73616	33.73616	-84.39495	33.73616	-84.39495	33.73616	-84.39495	-84.39495	-84.39495
12	33.74974	33.74974	-84.43192	33.74974	-84.43192	33.74974	-84.43192	-84.43192	-84.43192
13	33.68243	33.68243	-84.40766	33.68243	-84.40766	33.68243	-84.40766	-84.40766	-84.40766
14	33.71691	33.71691	-84.39276	33.71691	-84.39276	33.71691	-84.39276	-84.39276	-84.39276
15	33.73048	33.73048	-84.38742	33.73048	-84.38742	33.73048	-84.38742	-84.38742	-84.38742
16	33.70087	33.70087	-84.46574	33.70087	-84.46574	33.70087	-84.46574	-84.46574	-84.46574
17	33.75831	33.75831	-84.43107	33.75831	-84.43107	33.75831	-84.43107	-84.43107	-84.43107
18	33.79671	33.79671	-84.47527	33.79671	-84.47527	33.79671	-84.47527	-84.47527	-84.47527
19	33.70678	33.70678	-84.38031	33.70678	-84.38031	33.70678	-84.38031	-84.38031	-84.38031
20	33.84746	33.84746	-84.36983	33.84746	-84.36983	33.84746	-84.36983	-84.36983	-84.36983
21	33.75757	33.75757	-84.33868	33.75757	-84.33868	33.75757	-84.33868	-84.33868	-84.33868
22	33.77067	33.77067	-84.38023	33.77067	-84.38023	33.77067	-84.38023	-84.38023	-84.38023
23	33.70369	33.70369	-84.40795	33.70369	-84.40795	33.70369	-84.40795	-84.40795	-84.40795
24	33.75437	33.75437	-84.43089	33.75437	-84.43089	33.75437	-84.43089	-84.43089	-84.43089
25	33.74647	33.74647	-84.42057	33.74647	-84.42057	33.74647	-84.42057	-84.42057	-84.42057
	[reached 'max' / getopt("max.print") -- omitted 75 rows]								

> |

