13.1

1. Use the given link Data Set

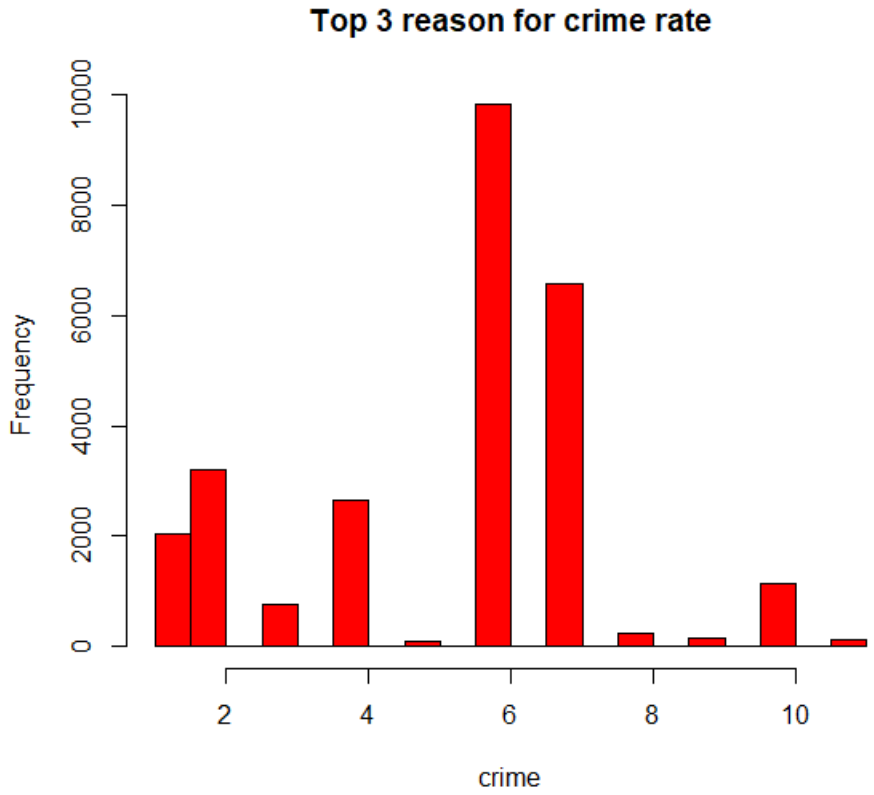
Answer the below questions:

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

```
b. yeast <- read.csv("C:/Users/satish/Desktop/yeast.txt", sep="")
c. >   View(yeast)
d. > View(yeast)
e. > cor1 <- cor.test(yeast$mcg, yeast$gvh, method = "pearson")
f. > cor1
g.
h.   Pearson's product-moment correlation
i.
j. data:  yeast$mcg and yeast$gvh
k. t = 27.526, df = 1482, p-value < 2.2e-16
l. alternative hypothesis: true correlation is not equal to 0
m. 95 percent confidence interval:
n.  0.5469332 0.6143348
o. sample estimates:
p.       cor
q. 0.5816314
r.
s. > cor2<- cor.test(yeast$gvh, yeast$alm, method = "pearson")
t. > cor2
u.
v.   Pearson's product-moment correlation
w.
x. data:  yeast$gvh and yeast$alm
y. t = -10.873, df = 1482, p-value < 2.2e-16
z. alternative hypothesis: true correlation is not equal to 0
aa.  95 percent confidence interval:
bb.   -0.3182836 -0.2240126
cc.  sample estimates:
dd.      cor
ee.  -0.2718
fff.
gg.  > cor3<- cor.test(yeast$alm, yeast$mit, method = "pearson")
hh.  > cor3
ii.
jj.         Pearson's product-moment correlation
kk.
ll.  data:  yeast$alm and yeast$mit
mm.  t = 2.3011, df = 1482, p-value = 0.02152
nn.  alternative hypothesis: true correlation is not equal to 0
oo.  95 percent confidence interval:
pp.   0.008809399 0.110219289
qq.  sample estimates:
rr.       cor
ss.  0.0596683
tt.
uu.  > cor4<- cor.test(yeast$mit, yeast$erl, method = "pearson")
vv.  > cor4
ww.
xx.         Pearson's product-moment correlation
yy.
zz.  data:  yeast$mit and yeast$erl
aaa. t = -0.22832, df = 1482, p-value = 0.8194
bbb. alternative hypothesis: true correlation is not equal to 0
ccc. 95 percent confidence interval:
ddd.  -0.05679921  0.04496851
eee. sample estimates:
fff.          cor
ggg. -0.005930705
hhh.
iii. > cor5<- cor.test(yeast$erl, yeast$mcg, method = "pearson")
jjj. > cor5
```

```
kkk.
lll.        Pearson's product-moment correlation
mmm.
nnn. data:  yeast$erl and yeast$mcg
ooo. t = 2.5046, df = 1482, p-value = 0.01237
ppp. alternative hypothesis: true correlation is not equal to 0
qqq. 95 percent confidence interval:
rrr.  0.01408241 0.11542587
sss. sample estimates:
ttt.        cor
uuu. 0.06492154
```

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



1. LARCENY-NON VEHICLE 2, LARCENY-VEHICLE 3. AUTO THEFT

c. Which all attributes have correlation with crime rate?

```
> cor11<- cor.test(COBRA.YTD2017$offense_id, COBRA.YTD2017$beat, method =
"pearson")
> cor11

        Pearson's product-moment correlation

data:  COBRA.YTD2017$offense_id and COBRA.YTD2017$beat
t = -3.6751, df = 26757, p-value = 0.0002383
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.03443370 -0.01048248
sample estimates:
        cor
-0.02246131
```

```
> cor12<- cor.test(COBRA.YTD2017$offense_id, COBRA.YTD2017$MinOfucr,
method = "pearson")
> cor12

        Pearson's product-moment correlation

data:  COBRA.YTD2017$offense_id and COBRA.YTD2017$MinOfucr
t = -38.827, df = 26757, p-value < 2.2e-16
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.2422611 -0.2195757
sample estimates:
       cor
-0.2309498

> cor13<- cor.test(COBRA.YTD2017$offense_id,
COBRA.YTD2017$MaxOfnum_victims, method = "pearson")
> cor13

        Pearson's product-moment correlation

data:  COBRA.YTD2017$offense_id and COBRA.YTD2017$MaxOfnum_victims
t = 2.5494, df = 26682, p-value = 0.0108
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 0.003607442 0.027598557
sample estimates:
       cor
0.01560525

> cor14<- cor.test(COBRA.YTD2017$offense_id, COBRA.YTD2017$loc_type,
method = "pearson")
> cor14

        Pearson's product-moment correlation

data:  COBRA.YTD2017$offense_id and COBRA.YTD2017$loc_type
t = 0.12292, df = 23413, p-value = 0.9022
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.01200549  0.01361189
sample estimates:
         cor
0.0008033295
```