QUESTION_1:

In [112]:
```
myName = "Ayush Dadhich"
myEmail = "adadhich@hawk.iit.edu"
myName
myEmail
```

'Ayush Dadhich'

'adadhich@hawk.iit.edu'

QUESTION_2:

In [113]:
```
myVector = sample(99:10000)
myVector
```

```
5013  7890  4547  5222  6635  7888  9850  2959  335  6105  1541  5996  4976
3667  3496  2771  1731  7277  6659  3963  431  1561  1242  5563  9913  1614
9903  5667  8652  1671  3742  7317  2637  4554  3995  7850  7858  5302  294
7194  1450  5919  7316  7567  9161  6431  5006  2924  5042  7979  2316  9142
6204  9806  6403  2833  6208  1868  838  4667  9388  6812  7376  991  5240
6570  8260  3933  7469  4101  7694  6475  7013  1966  8673  6059  9448  8667
2774  4906  5583  5756  1289  9632  1124  6644  7265  9164  5716  8345  3927
4155  1013  9970  2216  2318  4354  2829  6355  3531  655  4643  6955  2881
2710  2228  516  7319  2475  5920  8507  4606  3133  2466  2265  4551  3514
1866  8751  5437  2063  8071  4507  7968  3805  8533  1925  2440  3467  5380
4567  2585  7805  8003  7931  3324  4583  6415  8993  5044  2492  355  1914
8741  7632  6963  5292  237  4112  9554  1854  3734  951  9818  325  7363
5728  400  3953  8010  2037  6521  4088  571  8563  9908  7250  6950  6117
318  4361  5698  5736  2117  4680  2055  9300  2620  577  7661  9192  6088
8965  2054  526  3969  5252  5251  3251  5084  2204  7928  1582  9426  5972
8680  3961  640  6262  7985  7212  5183  8541  8871  7349  1969  2083  5026
3336  1143  9715  8923  1792  7306  9707  4750  6883  1604  380  9347  7784
```

In [114]:
```
new.mySumFunc = function() {
    print(sum(myVector))
}
```

In [115]:
```
new.mySumFunc()
```

```
[1] 50000149
```

In [116]:
```
myVect = sort(myVector)
```

In [117]:
```
new.myMinFunc = function() {
    print(myVect[1])
}
```

In [118]:
```
new.myMinFunc()
```

```
[1] 99
```

In [119]:
```
new.myMaxFunc = function() {
     print(myVect[9902])
}
```

In [120]:
```
new.myMaxFunc()
```

```
[1] 10000
```

In [121]:
```
new.myMedianFunc = function() {
     print(median(myVect))
}
```

In [122]:
```
new.myMedianFunc()
```

```
[1] 5049.5
```

QUESTION_3:

In [123]:
```
new.divis = function(num){
    if(num %% 127 == 0){
        return(TRUE)
    }
    else{
        return(FALSE)
    }
}
```

In [124]:
```
new.divis(80)
```

FALSE

In [125]:
```
new.divis(127*5)
```

TRUE

QUESTION_4:

In [126]:
```
countDivis = 0
for (i in myVector){
    if(i %% 127 == 0){
        countDivis = countDivis + 1
    }
}

countDivis
```

78

QUESTION_5:

In [127]:
```
names = c("Kermit Chacko",
          "Eleonore Chien",
          "Genny Layne",
          "Willene Chausse",
          "Taylor Lyttle",
          "Tillie Vowell",
          "Carlyn Tisdale",
          "Antione Roddy",
          "Zula Lapp",
          "Delphia Strandberg",
          "Barry Brake",
          "Warren Hitchings",
          "Krista Alto",
          "Stephani Kempf",
          "Sebastian Esper",
          "Mariela Hibner",
          "Torrie Kyler")
```

In [128]:
```
names[length(names)-8]
```

'Zula Lapp'

QUESTION_6:

In [129]:
```
countLastNameStartsWithL = 0
for (i in names){
    if(grepl(" L", i)){
        countLastNameStartsWithL = countLastNameStartsWithL + 1
    }
}
countLastNameStartsWithL
```

3

QUESTION_7:

In [130]:
```
new.nameMap = function(first_name){
    for(i in names){
        if(grepl(first_name, i)){
            print(gsub(" ", "", sub("^[^  ]*", "", i), fixed = TRUE))
        }
    }
}
```

In [131]:
```
new.nameMap("Krista")
```

```
[1] "Alto"
```

QUESTION_8:

In [132]:
```
df = read.csv(file="C:/Users/Admin/Desktop/adult.csv", header=TRUE, sep=",")
```

In [133]:
```
df
```

Age	Workclass	Fnlwgt	Education	Education_num	Martial_status	Occupation	Relationship	
39	State-gov	77516	Bachelors	13	Never-married	Adm-clerical	Not-in-family	\
50	Self-emp-not-inc	83311	Bachelors	13	Married-civ-spouse	Exec-managerial	Husband	\
38	Private	215646	HS-grad	9	Divorced	Handlers-cleaners	Not-in-family	\
53	Private	234721	11th	7	Married-civ-spouse	Handlers-cleaners	Husband	E
28	Private	338409	Bachelors	13	Married-civ-spouse	Prof-specialty	Wife	E
37	Private	284582	Masters	14	Married-civ-spouse	Exec-managerial	Wife	\
49	Private	160187	9th	5	Married-spouse-absent	Other-service	Not-in-family	E

QUESTION_9:

In [134]:
```
df[df$Workclass == " Private", "workSector"] = "Private"
df[df$Workclass == " State-gov", "workSector"] = "Government"
df[df$Workclass == " Federal-gov", "workSector"] = "Government"
df[df$Workclass == " Local-gov", "workSector"] = "Government"
df[df$Workclass == " Self-emp-not-inc", "workSector"] = "selfEmployed"
df[df$Workclass == " Self-emp-inc", "workSector"] = "selfEmployed"
df[c("workSector")][is.na(df[c("workSector")])] <- "Other"
```

In [135]:
```
df[, c("Workclass", "workSector")]
```

Workclass	workSector
State-gov	Government
Self-emp-not-inc	selfEmployed
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Self-emp-not-inc	selfEmployed
Private	Private
Private	Private
Private	Private
State-gov	Government
Private	Private
Private	Private
Private	Private
Private	Private
Self-emp-not-inc	selfEmployed
Private	Private
Private	Private
Self-emp-not-inc	selfEmployed
Private	Private
Private	Private
Federal-gov	Government
Private	Private
Private	Private
Local-gov	Government
Private	Private
?	Other
Private	Private
Private	Private
...	...
?	Other
Private	Private
Private	Private
Private	Private

Workclass	workSector
Private	Private
Private	Private
Private	Private
Private	Private
?	Other
State-gov	Government
?	Other
?	Other
Local-gov	Government
Private	Private
Local-gov	Government
Private	Private
Private	Private
Self-emp-not-inc	selfEmployed
State-gov	Government
Self-emp-not-inc	selfEmployed
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Private	Private
Self-emp-inc	selfEmployed

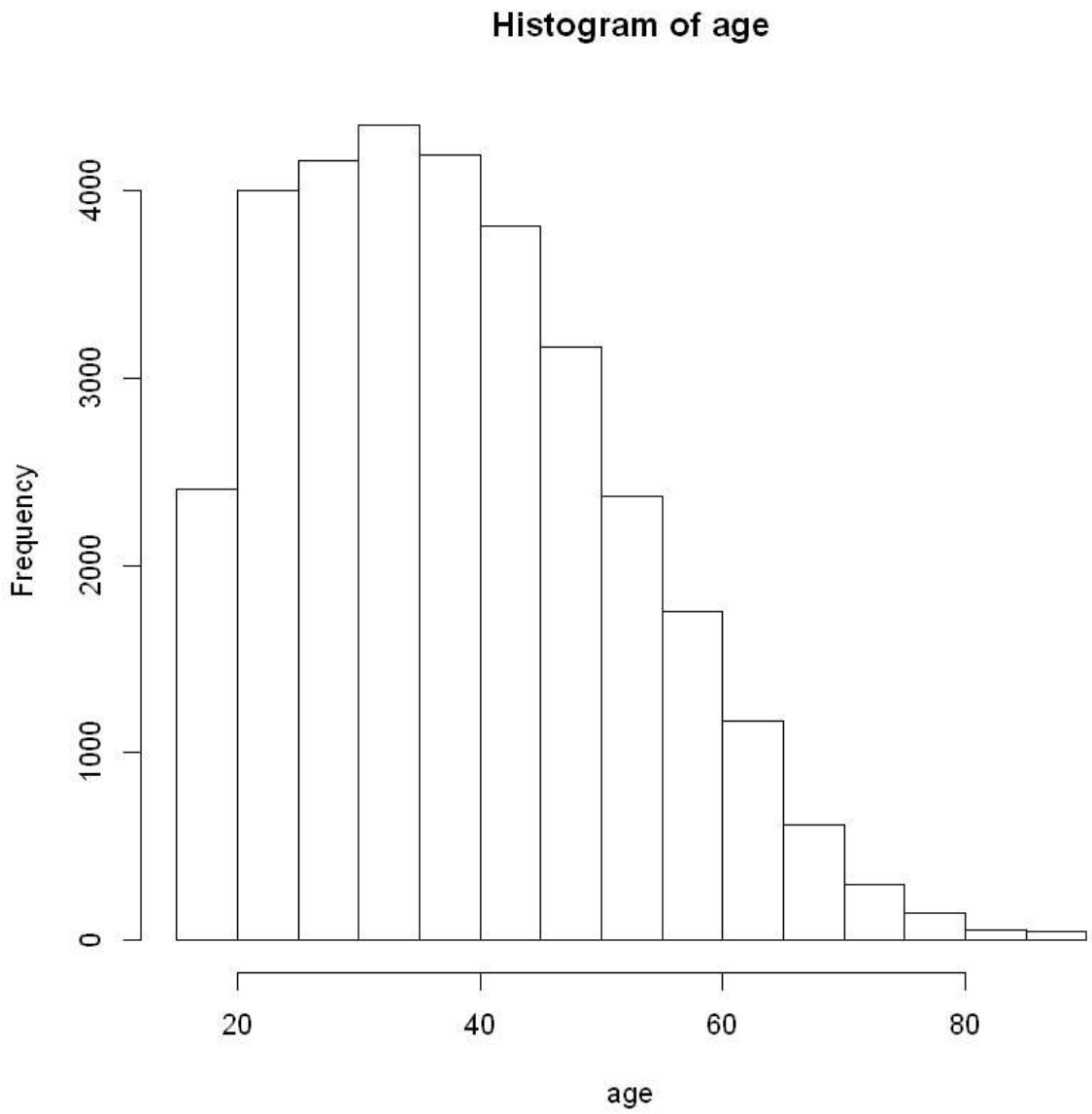
In [136]:
```
sort(table(df["workSector"]), decreasing = TRUE)
```

```
  Private Government selfEmployed     Other 
    22696      4351      3657      1857 
```

QUESTION_10:

In [137]:
```
age = as.numeric(unlist(df["Age"]))
```

In [138]:
```
hist(age)
```



QUESTION_11:

In [139]:
```
head(sort(tapply(df$Hours_per_week, df$Occupation, mean), decreasing = TRUE), 3)
```

Farming-fishing	46.989939637827
Exec-managerial	44.9877029021151
Transport-moving	44.6562304320601

QUESTION_12:

In [140]:
```
sort(tapply(df$Hours_per_week, df$workSector, mean), decreasing = TRUE)
```

selfEmployed	45.7634673229423
Government	40.4881636405424
Private	40.2670955234403
Other	31.9122240172321

In [141]:
```
a = tapply(df$Capital_gain, df$workSector, mean)
a
```

Government	816.587680992875
Other	603.611739364566
Private	889.217791681354
selfEmployed	2798.40224227509

In [142]:
```
b = tapply(df$Capital_loss, df$workSector, mean)
b
```

Government	102.452309813836
Other	60.0732364028002
Private	80.0087240042298
selfEmployed	128.382827454197

In [143]:
```
sort(a - b)
```

Other	543.538502961766
Government	714.135371179039
Private	809.209067677124
selfEmployed	2670.01941482089

CONCLUSION: After analysing the given data, I conclude that the statement of my friend was wrong which is "to make more money, you have to work longer hours".

QUESTION_13:

In [144]:
```
new.charCombos = function(myTestString, z){
    vector = c()
    for(i in 1:nchar(myTestString)){
        if(i+(z-1) <= nchar(myTestString)){
            vector = append(vector, substr(myTestString, i, i+(z-1)))
        }
    }
    table(vector)
}
```

In [145]:
```
new.charCombos("abcbcb", 2)
```

```
vector
ab bc cb
 1  2  2
```

In [146]:
```
new.charCombos("abcbcb", 3)
```

```
vector
abc bcb cbc
  1   2   1
```

QUESTION_14:

In [147]:
```
bigListOfWords <- readLines('https://raw.githubusercontent.com/dwyl/english-words
```

In [152]:
```
allQwords = NULL
allQuwords = NULL

for (i in 1:length(bigListOfWords)){
    if (grepl("q",bigListOfWords[i]) || grepl("Q",bigListOfWords[i])){
        if(!(grepl("q ",bigListOfWords[i]) || grepl("q[^[:alnum:]]",
                                                    bigListOfWords[i]))){
            allQwords<-append(allQwords,new.charCombos(bigListOfWords[i],
                                                    nchar(bigListOfWords[i])))
        }
    }
    if (grepl("qu",bigListOfWords[i]) || grepl("QU",bigListOfWords[i])
        || grepl("Qu",bigListOfWords[i])){
        allQuwords<-append(allQuwords,new.charCombos(bigListOfWords[i],
                                                    nchar(bigListOfWords[i])))
    }
}
pctQU = (length(allQuwords)/length(allQwords))*100
print(pctQU)
```

```
[1] 96.79212
```

QUESTION_15:

In [151]:
```
all_letters = NULL
for (i in 1:length(bigListOfWords)){
    let = unlist(strsplit(bigListOfWords[i], ""))
    for (k in 1:length(let)){
        if (grepl("q", let[k],ignore.case = TRUE)){
            if(!(grepl("u", let[k+1],ignore.case = TRUE)) && k != length(let)
               && !(grepl("[^[:alpha:]]",let[k+1],ignore.case = TRUE)))
                all_letters<-append(allletters,let[k+1])
            }
        }
    }
lowercase = unlist(lapply(all_letters, tolower))
top5 = head(sort(table(lowercase),decreasing = TRUE),5)
print(top5)
```

```
lowercase
 a  i  s  r  e
40 23 17 14 11
```

In [ ]: