ISYE 6501 HW 6

Question 9.1

I set my working directory up, imported the data, and scaled the data using the PCA function

> setwd("~/Desktop/ISYE 6501/ISYE HW 5/uscrime")

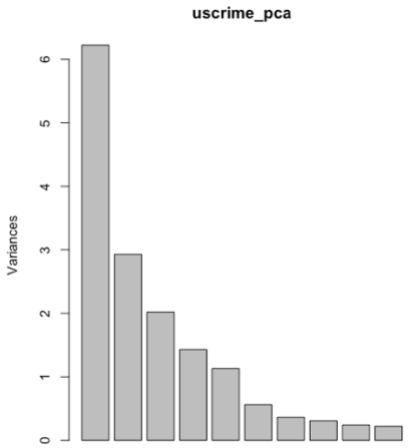
uscrime <- read.table("uscrime.txt", stringsAsFactors = FALSE, header = TRUE)

uscrime_pca <- prcomp(uscrime[,1:16], scale = TRUE)
summary(uscrime_pca)

```
Importance of components:
                          PC1    PC2    PC3     PC4     PC5     PC6     PC7     PC8     PC9    PC10
Standard deviation     2.4944 1.7111 1.4208 1.19585 1.06341 0.75087 0.60237 0.55503 0.49244 0.47036
Proportion of Variance 0.3889 0.1830 0.1262 0.08938 0.07068 0.03524 0.02268 0.01925 0.01516 0.01383
Cumulative Proportion  0.3889 0.5719 0.6981 0.78744 0.85812 0.89336 0.91603 0.93529 0.95044 0.96427
                          PC11    PC12    PC13    PC14    PC15    PC16
Standard deviation     0.43856 0.41777 0.29147 0.26063 0.21813 0.06584
Proportion of Variance 0.01202 0.01091 0.00531 0.00425 0.00297 0.00027
Cumulative Proportion  0.97629 0.98720 0.99251 0.99676 0.99973 1.00000
```

I used plot to visualize the variance, and then checked the attributes

plot(uscrime_pca)



attributes(uscrime_pca)

```
$names
[1] "sdev"     "rotation" "center"   "scale"    "x"

$class
[1] "prcomp"
```

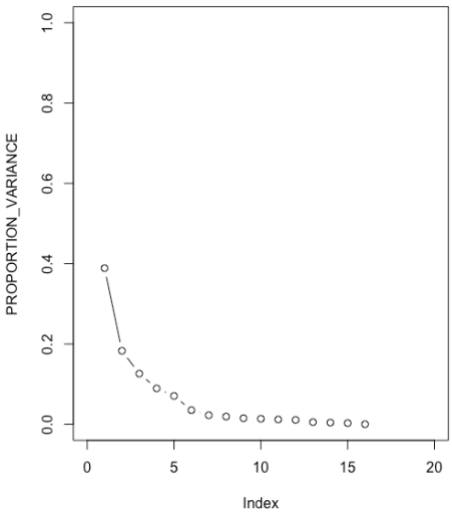
I used sdev squared to find the variance, and then variance to visualize the output

```
VARIANCE <- uscrime_pca$sdev^2
VARIANCE
```

```
[1] 6.222199347 2.928000127 2.018772756 1.430068772 1.130846060 0.563802258 0.362852350 0.308054904
[9] 0.242496934 0.221238989 0.192335693 0.174532062 0.084956874 0.067928691 0.047578814 0.004335367
```

I plotted the variance, the principal components increase with less variance, the first point contains 40% of the variance, the second point has 20%, and the third point is at 12%, so I used the first 5 principal components for the PCA

```
proportion_variance <- VARIANCE/sum(VARIANCE)
plot(proportion_variance, ylim = c(0,1), xlim = c(0,20), type = "b")
```




```
attributes(uscrime_pca$x)
```

```
$dim
[1] 47 16

$dimnames
$dimnames[[1]]
NULL

$dimnames[[2]]
 [1] "PC1"  "PC2"  "PC3"  "PC4"  "PC5"  "PC6"  "PC7"  "PC8"  "PC9"  "PC10" "PC11" "PC12" "PC13" "PC14"
[15] "PC15" "PC16"
```

```
uscrime_pca$x
```

```
           PC1         PC2         PC3         PC4         PC5         PC6          PC7         PC8
 [1,] -4.14470240 -1.28686914 -1.09049306 -0.48773442 -0.57542037  0.29463615  2.974612e-01  0.119074816
 [2,]  1.52730627  0.20521452  0.23170707 -1.83762098  0.85668057 -0.61048963 -3.941358e-01  0.134152780
 [3,] -4.27670702  0.18753174 -0.35183873  0.09333563 -0.56636069  0.61878339 -3.441300e-01  0.335525819
```

[4,] 4.41051751 -2.97003045 0.33605875 -1.71620948 0.52278262 0.83051828 -1.307531e-01 -1.552267788
[5,] 1.94524947 1.15708435 1.41675037 -0.66938364 -0.55204451 -0.45296947 -1.681356e-01 -0.527900611
[6,] 2.73005034 0.08721197 0.32384638 1.16416355 -1.85689752 -0.70846942 3.246225e-02 0.477269799
[7,] 0.26537172 -0.05239304 -0.89621940 -0.23991355 -1.26718631 -1.03315631 -4.925540e-01 0.986746107
[8,] 0.26688089 -1.81401931 0.73708812 -1.42350435 -0.99545113 -0.08283117 -1.234372e+00 0.534306380
[9,] -3.52611033 -0.98890238 1.33384083 -0.47795432 -0.33126232 0.19988352 4.447153e-01 0.022202434
[10,] 0.96551359 3.09518019 0.52185009 -0.41517056 0.54260694 0.44079287 2.428141e-01 -0.340734079
[11,] 2.96153230 -2.89080821 1.47876469 0.15520234 0.96059540 0.37326020 -2.018737e-01 0.630408651
[12,] 0.94602447 0.08446019 1.07195387 1.20068673 0.92003507 0.21201115 2.353033e-01 0.398577345
[13,] 0.26043407 1.25643288 1.68078912 1.78021214 0.98180047 0.97947184 8.090047e-02 0.851275561
[14,] 0.23108465 1.98746361 1.02226715 0.68710999 -0.30386997 0.50472546 -2.099029e-01 -0.006438083
[15,] -3.33331443 -0.12578706 -0.93044729 -0.48379933 0.01645702 0.05620962 -1.050479e+00 -0.325912128
[16,] -2.88473748 -2.27600287 -1.82383497 -0.73696119 -0.23473508 -0.17727555 -2.178105e-01 0.570103296
[17,] -0.48592796 1.74233405 -1.42750979 0.81307590 -0.88567809 -0.37142194 5.395064e-01 -0.774887404
[18,] -0.07103765 -0.46009987 -0.33294273 -1.36889911 -3.73225349 0.25094427 1.235043e+00 0.495830619
[19,] 2.81739799 -1.18119844 0.69347444 1.48959231 -1.47880403 -0.93498648 -1.706506e-01 -0.089496233
[20,] 2.45848881 -1.62427231 -2.86239148 -0.46089084 0.50866482 0.75180111 2.615344e-01 0.317950504
[21,] 1.02105442 0.38437075 0.15286579 1.40036602 1.81441094 0.41717718 -4.971461e-02 0.796759403
[22,] -5.73937141 -1.25186627 0.07823332 -0.15683524 0.27591768 0.35651472 1.410199e+00 0.494693355
[23,] 0.06356097 -0.72785569 0.70481886 0.37554206 0.17000567 0.26325515 -1.416816e+00 0.203459963
[24,] 1.32507927 2.01148375 -2.80991049 -1.02233073 0.76506993 0.05575549 1.634217e-02 -0.214465773
[25,] -0.01294690 2.79843146 1.82834917 0.85896360 -0.49770292 1.42396472 -1.004300e-01 0.470916970
[26,] 4.56399129 0.55396147 -0.33329063 -4.04945075 0.12548207 -0.18667918 3.252660e-01 0.969552345

[27,] 0.53124944 1.27986441 1.01778114 2.33569747 -0.49291192 -1.21470220 3.141233e-01 -0.018766361
[28,] 0.28530971 1.14717079 0.33670892 -1.21370884 1.07910916 -0.70538318 -1.224333e-05 -0.680727439
[29,] 3.69624495 -4.40998341 0.35534250 1.27750259 -0.98663819 0.86188986 2.853188e-01 -0.909778536
[30,] -4.11230623 -0.64275686 1.47577687 -0.37512612 -0.45458662 -0.31527359 5.005032e-01 -0.872101982
[31,] -1.00486560 1.98901943 -2.87730751 0.22203084 1.33040965 -0.34015232 4.535347e-02 -0.717398467
[32,] 1.61954661 -0.95139796 -0.77966922 1.25411832 0.29863400 1.34420119 6.118083e-01 0.090125829
[33,] -1.77222614 0.27214222 1.55283349 -0.35566161 -0.16350399 -0.32345127 -1.052560e+00 -0.088168703
[34,] 1.89390685 0.70190628 -0.78352414 0.25052294 -0.70471182 -0.52108435 -3.123382e-01 0.090447413
[35,] 1.47750484 -2.57195828 -2.12536958 2.19977459 0.95705504 0.55552630 4.980835e-01 -0.193153068
[36,] 1.27962355 -1.73187977 0.70977934 0.73132794 1.69141248 -2.47440272 3.326452e-01 0.437519756
[37,] -4.00977867 -0.35349518 2.55866310 -1.48592181 1.62341182 0.45031260 9.346830e-01 0.170848935
[38,] -0.97576772 2.97549025 -0.25409063 0.02940172 0.45091470 0.98894875 -8.083132e-01 -0.226638532
[39,] -3.29732035 -1.10709429 0.16982373 0.30100415 0.07641367 -0.36872388 -1.038009e+00 -0.115224037
[40,] -0.86916366 -1.32170810 0.87784509 -0.54004651 0.44501081 0.62369249 -6.789008e-01 -0.826701036
[41,] 1.39918932 1.89952034 1.85490580 0.50273261 0.93556475 -1.01214297 6.236917e-01 -0.343158369
[42,] -0.95131111 2.11645223 -1.69056479 0.86787970 -1.77794001 -0.28943187 5.677079e-02 -0.793447384
[43,] -2.36250762 -0.71784209 1.48906960 -0.70685936 0.02677089 -0.93792568 5.494348e-01 -0.603977006
[44,] 2.31016253 0.49356352 -0.60994632 -0.28550525 0.89085406 -0.53542714 5.661969e-01 -0.011991939
[45,] -3.10467207 -0.90264470 -3.51864372 1.21835963 0.93233896 -0.62869491 -4.352462e-01 0.342656181
[46,] 1.50169092 1.28646594 0.87908141 0.64407378 -1.35278364 0.77829116 -3.475787e-01 -0.038420066
[47,] 2.18080798 2.64810935 -1.39227454 -1.34318857 0.01233342 0.59250775 4.145552e-01 0.331350762

	PC9	PC10	PC11	PC12	PC13	PC14	PC15
[1,]	-0.21821990	-0.709569063	-0.132399755	-0.14269706	-0.076932292	0.212937655	-0.217802327

[2,] 0.10166884 -0.682786179 0.147625333 -0.15196379 -0.227810126 -0.166575495 -0.125053586
[3,] 0.05355273 -0.004354472 0.002212673 -0.48292356 0.026848333 -0.163255009 0.509291085
[4,] -0.55891531 -0.169806967 -0.030298654 0.06088706 -0.124985829 -0.012455061 -0.043235275
[5,] -0.11770492 -0.634146611 -0.196658830 -0.39142927 0.306818840 0.700118668 0.033255461
[6,] 0.75366786 0.290850396 -0.447372491 0.18986912 -0.305561895 0.220109370 -0.116773237
[7,] -0.83040855 0.584375720 -1.008103872 0.50033077 -0.451172646 -0.049030216 -0.017038315
[8,] -0.39879993 -0.103177447 -0.472799569 -0.33119599 0.422894218 -0.429638301 0.463419523
[9,] 0.11868696 -0.185075818 -0.048028679 0.14900156 -0.200879686 0.289160662 0.159682369
[10,] 0.03503179 0.222972368 -0.373346613 0.10210380 0.211052526 0.137044572 0.212201250
[11,] -0.15022518 0.072841138 0.561758941 -0.25008728 -0.802874148 0.008117164 0.159755927
[12,] -0.21074649 -0.027420900 0.183173150 0.28701391 -0.369947365 -0.070401578 0.037933881
[13,] -0.47772408 0.087985154 0.429848673 -0.43295255 0.451457242 0.046875186 -0.206857817
[14,] -0.38732044 -0.312094115 -0.051977911 0.12536055 -0.178671908 -0.470410415 -0.354098825
[15,] 0.10345470 0.459389331 1.118982712 0.60877411 0.298614942 -0.261911802 -0.228085247
[16,] 0.50232026 -0.424770822 -0.822379081 -0.08063531 -0.049489924 0.119417423 -0.113650830
[17,] -0.21531760 -0.753089917 0.439298746 -0.03646724 0.165523506 0.050351386 0.281909012
[18,] -0.34390436 0.577143134 0.912794933 0.04829574 -0.082070097 0.509798165 0.092334048
[19,] 0.52093457 -0.439247458 -0.430314790 0.41183068 0.728319345 -0.318258264 -0.291917189
[20,] 0.06392268 -0.524585734 0.073046031 0.72170824 -0.050023235 -0.149905815 -0.060086848
[21,] -0.24218969 0.294129719 0.085798648 -0.25344889 0.117749548 0.401143979 -0.028768923
[22,] -0.01784504 -0.333259221 -0.442440255 -1.21479396 0.077771234 -0.340462739 -0.197782600
[23,] 0.80199265 -0.803454855 0.276307645 -0.12284341 -0.480460922 0.283919998 -0.203457075
[24,] -0.13327210 -0.101227487 -0.621635340 -0.18993028 -0.018070132 0.214293230 0.220898220

[25,] -0.25577627 -0.270885960 0.318643543 -0.09307741 0.198735992 -0.175742606 -0.004080609
[26,] 0.43302914 0.292997556 0.635187521 -0.27982103 0.132317771 -0.183779283 0.087474287
[27,] 0.13624756 0.283148228 -0.225204930 -0.13121628 -0.213412382 -0.078529992 0.207282775
[28,] -0.01588145 -0.291066078 0.048314158 -0.47758690 -0.180326706 0.332421425 -0.464455244
[29,] 0.56475867 0.471991801 -0.065344123 -0.77615475 0.441046452 0.064474854 0.002467698
[30,] 0.14413883 0.360566207 0.224235352 0.04659698 -0.337399204 -0.329730988 0.028829502
[31,] 1.12804850 0.939177019 -0.016030646 -0.39771891 -0.236153412 -0.031618303 0.195207763
[32,] -0.34743734 -0.108112118 0.017147002 0.62907270 -0.301074748 -0.078055542 -0.022932474
[33,] -0.70344214 0.220245149 -0.022402478 0.10603258 0.166550154 0.288309274 0.122772833
[34,] -0.17540264 -0.160046878 -0.252803866 -0.41559504 0.062418098 -0.140295677 -0.235380665
[35,] -0.02457988 -0.012113711 -0.286796932 0.17997396 -0.006445855 -0.129989307 0.252838302
[36,] 0.44953651 -0.314565572 0.951797273 -0.01192037 0.261734541 -0.173995636 0.338774949
[37,] 0.96687648 -0.346732968 -0.451063425 1.32706398 0.366442480 0.239715711 0.086364850
[38,] 0.58605086 0.205741483 -0.114136414 -0.45759336 -0.214578830 -0.131124322 0.043025438
[39,] 0.16522413 0.044640528 0.001096932 0.20957016 0.056821866 -0.052780523 0.061112989
[40,] -0.79267098 0.691019977 -0.256919744 0.25876116 0.131006720 0.069013774 0.173889486
[41,] -0.65716305 -0.022958105 -0.392749685 0.25060306 -0.129629208 0.058476188 0.203381485
[42,] -0.28249589 -1.094446007 0.615845813 0.26545561 -0.228811680 -0.215972515 0.116818519
[43,] -0.26109081 1.031602906 0.008443255 -0.02750183 -0.178695910 -0.344946635 -0.426983874
[44,] -0.51050561 0.141883743 -0.257886098 -0.11325908 0.178426553 -0.015285426 -0.271753159
[45,] -0.61161919 0.591093230 0.522398756 0.17303398 0.382799111 0.479118210 -0.253484371
[46,] 1.25829843 0.636577361 -0.022406197 0.26475391 -0.024908782 0.160550544 -0.133328668
[47,] 0.05321667 0.328622316 -0.132456713 0.34671991 0.285037452 -0.371215988 -0.073914493

	PC16
[1,]	0.0247739196
[2,]	-0.0424578502
[3,]	0.0652545294
[4,]	0.0438745173
[5,]	-0.0642450800
[6,]	0.0043144032
[7,]	-0.0585628239
[8,]	-0.0054488182
[9,]	0.0402591721
[10,]	0.0195171983
[11,]	0.1425663109
[12,]	0.0303007475
[13,]	-0.0357506840
[14,]	-0.0001142815
[15,]	0.0314106616
[16,]	-0.0257220728
[17,]	0.0076631583
[18,]	0.0188316391
[19,]	0.0577127875
[20,]	-0.0106756140
[21,]	-0.0089449524
[22,]	-0.0217365775
[23,]	0.0361641203
[24,]	-0.0220322245
[25,]	-0.0559922453
[26,]	-0.1100084761
[27,]	0.0813100070
[28,]	-0.0386472199
[29,]	0.0132482183
[30,]	-0.0285851009
[31,]	0.0716593479
[32,]	-0.0799946618
[33,]	0.1202422644
[34,]	0.0725661033
[35,]	-0.1022407622
[36,]	-0.0528318002
[37,]	0.0675842618
[38,]	-0.0325903308
[39,]	-0.0976599959
[40,]	-0.0639360482
[41,]	-0.0821715728
[42,]	-0.0167652896
[43,]	-0.0566616118
[44,]	0.1419454312
[45,]	-0.0092155139

[46,] -0.1101431743
[47,] 0.1419359835

I selected the first 5 principal components for PCA below

```
principal_components <- uscrime_pca$x[,1:5]
summary(principal_components)
```

```
      PC1               PC2                PC3               PC4               PC5          
 Min.   :-5.7394   Min.   :-4.40998   Min.   :-3.5186   Min.   :-4.0495   Min.   :-3.73225  
 1st Qu.:-1.3885   1st Qu.:-1.14415   1st Qu.:-0.8399   1st Qu.:-0.6047   1st Qu.:-0.55920  
 Median : 0.2669   Median : 0.08446   Median : 0.3238   Median : 0.0294   Median : 0.07641  
 Mean   : 0.0000   Mean   : 0.00000   Mean   : 0.0000   Mean   : 0.0000   Mean   : 0.00000  
 3rd Qu.: 1.5734   3rd Qu.: 1.26815   3rd Qu.: 1.0200   3rd Qu.: 0.8360   3rd Qu.: 0.87377  
 Max.   : 4.5640   Max.   : 3.09518   Max.   : 2.5587   Max.   : 2.3357   Max.   : 1.81441  
```

```
uscrime_pca <- cbind(principal_components, uscrime[,16])
model1 <- lm(V6~., data = as.data.frame(uscrime_pca))
summary(model1)
```

```
Call:
lm(formula = V6 ~ ., data = as.data.frame(uscrime_pca))

Residuals:
     Min       1Q   Median       3Q      Max 
-305.496  -89.435    6.064   73.323  281.078 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  905.085     20.610  43.916  < 2e-16 ***
PC1           75.891      8.352   9.087 2.25e-11 ***
PC2          -92.650     12.175  -7.610 2.30e-09 ***
PC3           40.535     14.662   2.765   0.0085 ** 
PC4         -212.374     17.420 -12.191 3.22e-15 ***
PC5           51.545     19.590   2.631   0.0119 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 141.3 on 41 degrees of freedom
Multiple R-squared:  0.881,     Adjusted R-squared:  0.8665 
F-statistic: 60.74 on 5 and 41 DF,  p-value: < 2.2e-16
```

```
model1$coefficients
```

```
(Intercept)         PC1         PC2         PC3         PC4         PC5 
  905.08511    75.89073   -92.65004    40.53455  -212.37450    51.54475 
```

```
beta0 <- model1$coefficients[1]
betavector <- model1$coefficients[2:6]
alphavector <- uscrime_pca$rotation[,1:5]%*%betavector
og_alpha <- alphavector/sapply(uscrime[,1:16], sd)
og_beta0<- beta0-sum(alphavector*sapply(uscrime[,1:16], mean)/sapply (uscrime[,1:16], sd))
og_alpha
```

The original coefficients below

```
              [,1]
M       4.292046e+01
So      7.179246e+01
Ed      1.610427e+01
Po1     2.868444e+01
Po2     2.868298e+01
LF      1.235388e+03
M.F     2.865703e+01
Pop     7.690299e-01
NW      6.606744e+00
U1     -1.448960e+01
U2      3.138361e+01
Wealth  2.295557e-02
Ineq    6.972740e+00
Prob   -1.494492e+03
Time    2.217626e+00
Crime   4.261531e-01
```

```
estimates <- as.matrix(uscrime[,1:16]) %*% og_alpha + og_beta0
error_sumofsquares = sum((ESTIMATES-uscrime[,16])^2)
sstot = sum((us_crime[,16] - mean(uscrime[,16]))^2)
R_2 = 1 - error_sumofsquares/ssot
R_2
```

0.8810481

```
adjusted_R_2 = R_2 - (1-R_2)*5/(nrow(us_crime)-5-1)

adjusted_R_2
```

0.8665418

```
newcity <- data.frame (

  M= 14.0,

  So = 0,

  Ed = 10.0,

  Po1 = 12.0,

  Po2 = 15.5,

  LF = 0.640,

  M.F = 94.0,

  Pop = 150,
```

```
  NW = 1.1,
  U1 = 0.120,
  U2 = 3.6,
  Wealth = 3200,
  Ineq = 20.1,
  Prob = 0.04,
  Time = 39.0)
uscrime_pca <- prcomp(uscrime[,1:15], scale. = TRUE)
pred_newcity <- data.frame(predict(uscrime_pca, newcity))
pred_newcity_model <- predict(model1, pred_newcity)
pred_newcity_model
```

1457.421