**High value customers identification for an E-Commerce company.**

Q1) Use the clustering methodology to segment customers into groups:  
 Use the following clustering algorithms:

K means

Hierarchical

Ans) The below mention is the code in R which gives an overview of the K-Means

print("High Value Customer Identification")

ecom\_data<-read.csv("C:/Users/shiva/Desktop/Shivani/SimpliLearn/Data Science with R/Data Science with R Projects/Ecommerce.csv")

print(ecom\_data)

str(ecom\_data)

print("K-Means Algorithm")

set.seed(110)

# Data Cleaning

del\_vars<-names(ecom\_data)%in%c("InvoiceNo","StockCode","Description","InvoiceDate","Country")

cluster\_up<-kmeans(ecom\_data,3,iter.max=10)

ecom\_data\_num<-ecom\_data[!del\_vars]

ecom\_data\_num<-na.omit(ecom\_data\_num)

View(ecom\_data\_num)

cluster\_up<-kmeans(ecom\_data\_num,3,iter.max=10)

str(cluster\_up)

ecom\_data\_num<-cbind(ecom\_data\_num,clusternum=cluster\_up$cluster)

View(ecom\_data\_num)

The output of the K-Means is mention below:

str(ecom\_data)

'data.frame': 541909 obs. of 8 variables:

$ InvoiceNo : chr "536365" "536365" "536365" "536365" ...

$ StockCode : chr "85123A" "71053" "84406B" "84029G" ...

$ Description: chr "WHITE HANGING HEART T-LIGHT HOLDER" "WHITE METAL LANTERN" "CREAM CUPID HEARTS COAT HANGER" "KNITTED UNION FLAG HOT WATER BOTTLE" ...

$ Quantity : int 6 6 8 6 6 2 6 6 6 32 ...

$ InvoiceDate: chr "29-Nov-16" "29-Nov-16" "29-Nov-16" "29-Nov-16" ...

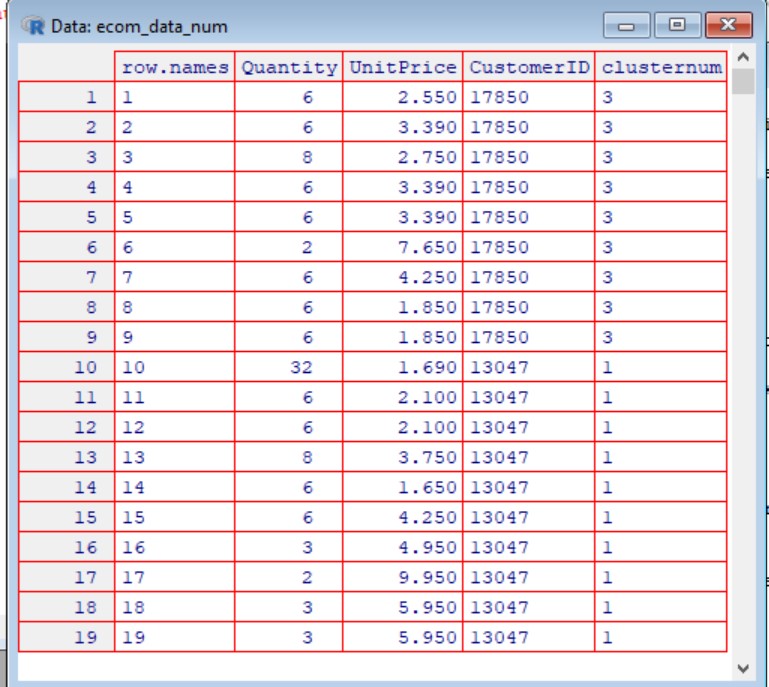
$ UnitPrice : num 2.55 3.39 2.75 3.39 3.39 7.65 4.25 1.85 1.85 1.69 ...

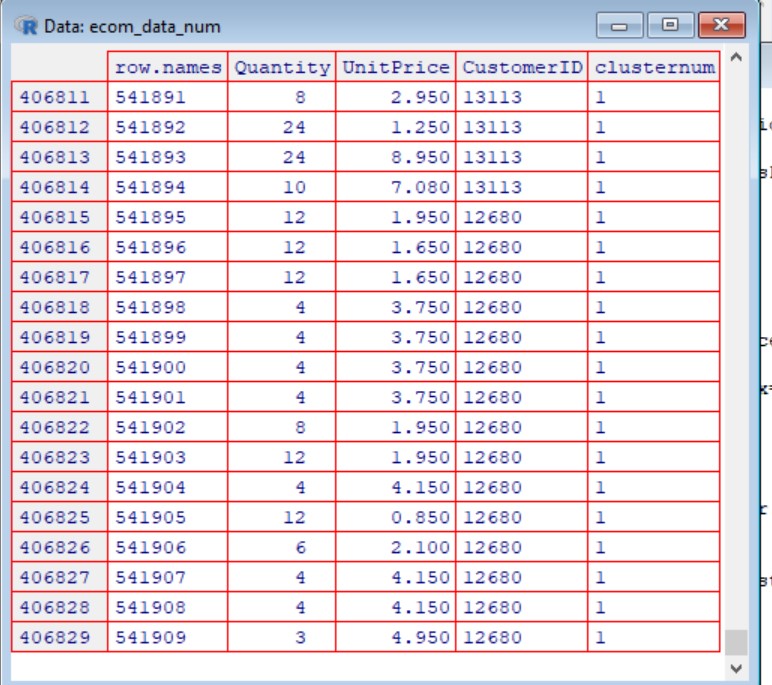
$ CustomerID : int 17850 17850 17850 17850 17850 17850 17850 17850 17850 13047 ...

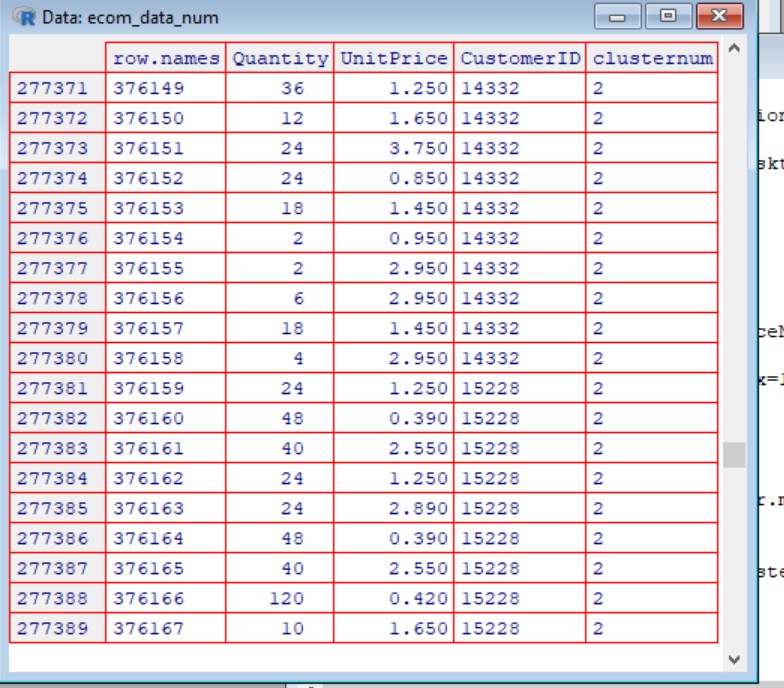
$ Country : chr "United Kingdom" "United Kingdom" "United Kingdom" "United Kingdom" ...

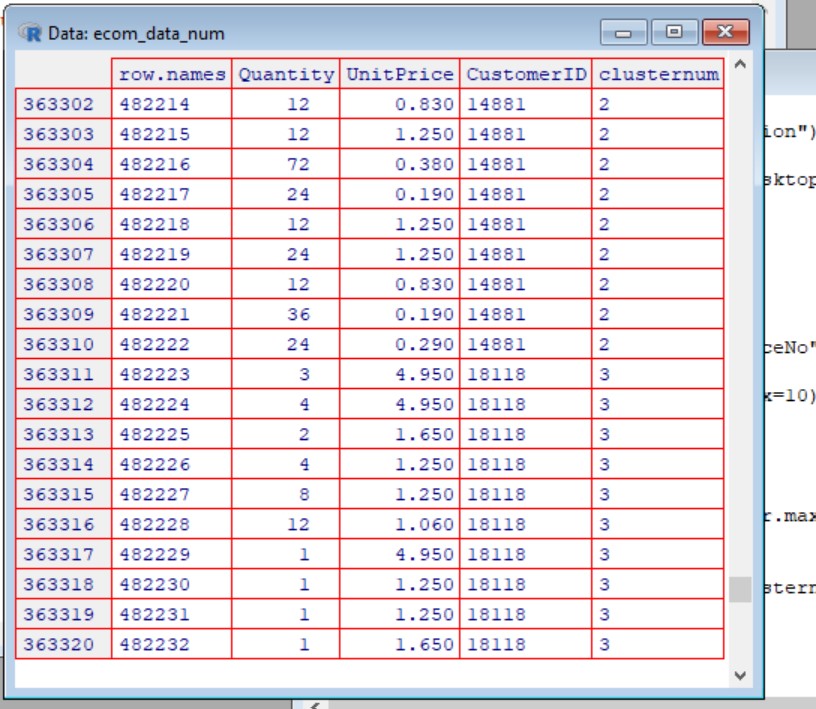
>

The above mention output represent the structure of the ecommerce data where the datatypes of various columns are listed accordingly.









The above mention screenshots represents the data of the ecommerce company where the data is segregated via K-Means Algorithm into three different clusters. Here, the data is divided into 3 clusters and maximum number of iterations is upto 10

The below mention is the code in R which gives an overview of the Hierarchical Algorithm

print("High Value Customer Identification")

ecom\_data<-read.csv("C:/Users/shiva/Desktop/Shivani/SimpliLearn/Data Science with R/Data Science with R Projects/Ecommerce.csv")

print(ecom\_data)

ecom\_data<-na.omit(ecom\_data)

head(ecom\_data)

ecom\_data<-ecom\_data[,-1]

ecom\_data<-ecom\_data[,-2]

ecom\_data<-ecom\_data[,-3]

ecom\_data<-ecom\_data[,-5]

ecom\_data<-ecom\_data[,-8]

str(ecom\_data)

ecom\_data<- as.data.frame(scale(ecom\_data))

print(ecom\_data)

ecom\_data<-head(ecom\_data)

summary(ecom\_data)

hclust\_avg <- hclust(dist\_mat, method = 'average')

plot(hclust\_avg)

The below mention are the output of the Hierarchical Algorithm

ecom\_data<-na.omit(ecom\_data)

> head(ecom\_data)

InvoiceNo StockCode Description Quantity InvoiceDate

1 536365 85123A WHITE HANGING HEART T-LIGHT HOLDER 6 29-Nov-16

2 536365 71053 WHITE METAL LANTERN 6 29-Nov-16

3 536365 84406B CREAM CUPID HEARTS COAT HANGER 8 29-Nov-16

4 536365 84029G KNITTED UNION FLAG HOT WATER BOTTLE 6 29-Nov-16

5 536365 84029E RED WOOLLY HOTTIE WHITE HEART. 6 29-Nov-16

6 536365 22752 SET 7 BABUSHKA NESTING BOXES 2 29-Nov-16

UnitPrice CustomerID Country

1 2.55 17850 United Kingdom

2 3.39 17850 United Kingdom

3 2.75 17850 United Kingdom

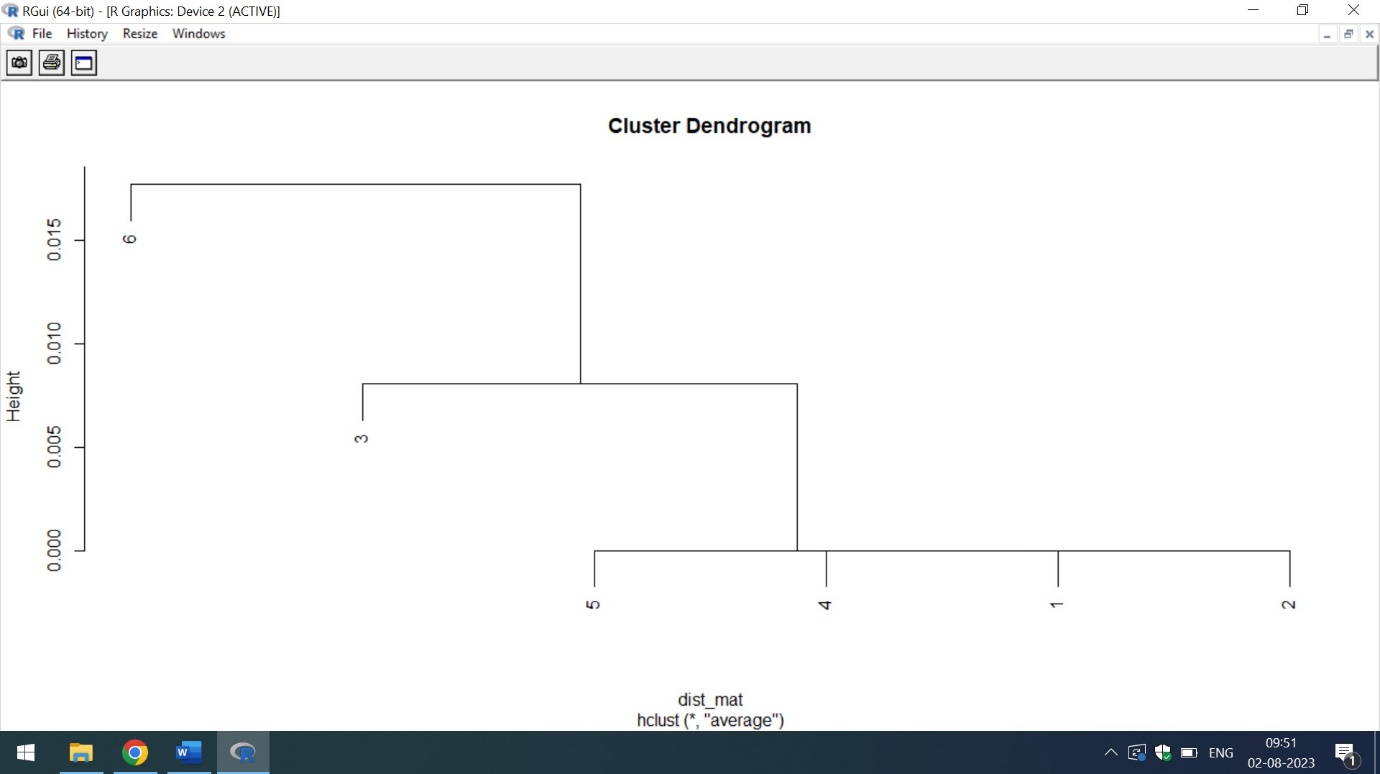
4 3.39 17850 United Kingdom

5 3.39 17850 United Kingdom

6 7.65 17850 United Kingdom

>

The above mention screenshot represents the top 6 rows of ecommerce data



From the above mention cluster Dendogram,it is clear that highest value of quantities of each product per transaction is 6. It simply means that Customers of various countries have purchased maximum of 6 products

Q2) Identify the right number of customer segments.

Ans) The right number of customer segments can be identified by K-Means Clustering Method.Below mentioned is the code in R which gives the list of customerids identified by K-Means

print("High Value Customer Identification")

ecom\_data<-read.csv("C:/Users/shiva/Desktop/Shivani/SimpliLearn/Data Science with R/Data Science with R Projects/Ecommerce.csv")

print(ecom\_data)

str(ecom\_data)

# From K-Means identify the number of valued customers

print("K-Means Algorithm")

set.seed(110)

# Data Cleaning

del\_vars<-names(ecom\_data)%in%c("InvoiceNo","StockCode","Description","InvoiceDate","Country")

cluster\_up<-kmeans(ecom\_data,3,iter.max=10)

ecom\_data\_num<-ecom\_data[!del\_vars]

ecom\_data\_num<-na.omit(ecom\_data\_num)

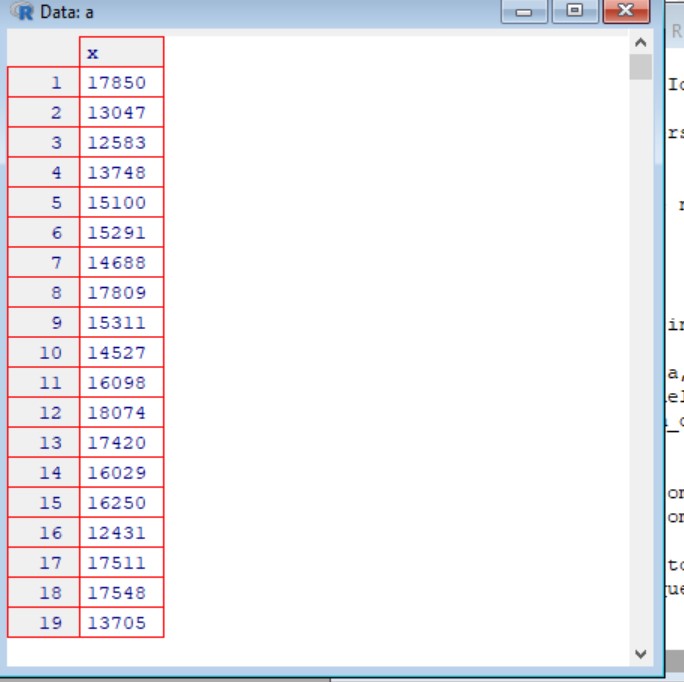
View(ecom\_data\_num)

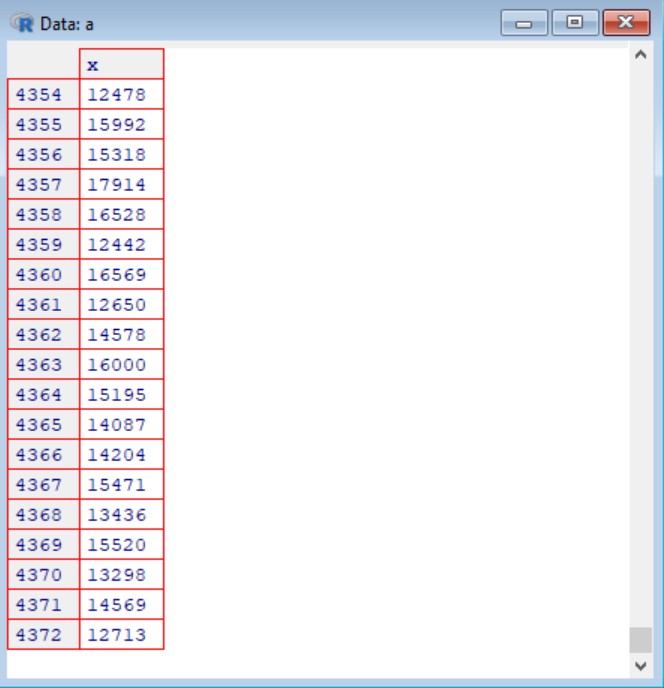
a<-print(ecom\_data\_num$CustomerID)

View(a) # List of the CustomerID's with duplicate values

a<-unique(ecom\_data\_num$CustomerID)

View(a) # List of the unique CustomerID's





The above mention screenshot represent the list of some of the customerID’s which are identified via K-Means Algorithm

Q3) Provide the number of customers who are highly valued.

Ans) By using K-Means Algorithm the number of highly valued customers can be identified.

The below mention is the code in R

print("High Value Customer Identification")

ecom\_data<-read.csv("C:/Users/shiva/Desktop/Shivani/SimpliLearn/Data Science with R/Data Science with R Projects/Ecommerce.csv")

print(ecom\_data)

str(ecom\_data)

# From K-Means identify the number of valued customers

print("K-Means Algorithm")

set.seed(110)

# Data Cleaning

del\_vars<-names(ecom\_data)%in%c("InvoiceNo","StockCode","Description","InvoiceDate","Country")

cluster\_up<-kmeans(ecom\_data,3,iter.max=10)

ecom\_data\_num<-ecom\_data[!del\_vars]

ecom\_data\_num<-na.omit(ecom\_data\_num)

View(ecom\_data\_num)

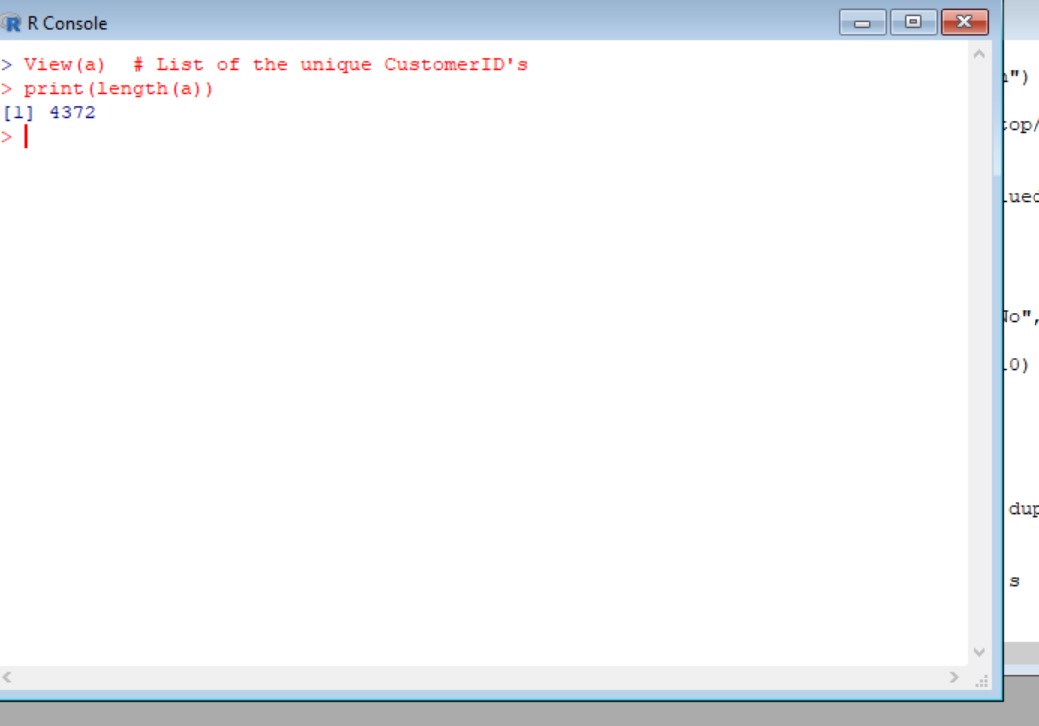
a<-print(ecom\_data\_num$CustomerID)

View(a) # List of the CustomerID's with duplicate values

a<-unique(ecom\_data\_num$CustomerID)

View(a) # List of the unique CustomerID's

print(length(a))



The above mention screenshot highlights that the total count of customers who are highly valued are 4372. Hence, the total count is 4372. This count is identified via K-Means

Q5)  Identify the clustering algorithm that gives maximum accuracy and explains robust clusters.

Ans) Naïve Bayes is one of the clustering Algorithm which gives the maximum accuracy.

Below attached is the code in R. The analysis is done on the columns of **InvoiceNo,Quantity,UnitPrice and CustomerID of UK customers only.**

print("High Value Customers")

ecom\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Ecommerce/main/MyEcommerce.csv")

print(ecom\_data)

View(ecom\_data)

ecom\_data<-na.omit(ecom\_data)

any(is.na(ecom\_data))

# Convert admit to factor

ecom\_data$CustomerID<-sapply(ecom\_data$CustomerID,factor)

# Build the model

tree\_model<-naiveBayes(CustomerID~.,data=ecom\_data)

print(tree\_model)

summary(tree\_model)

The below mention is the output of NaiveBayes

print(tree\_model)

Naive Bayes Classifier for Discrete Predictors

Call:

naiveBayes.default(x = X, y = Y, laplace = laplace)

A-priori probabilities:

Y

17850 13047 13748 15100 15291

8.621690e-04 5.416190e-04 7.737414e-05 1.658017e-05 3.012065e-04

14688 17809 15311 14527 16098

9.920470e-04 1.768552e-04 6.883535e-03 2.793759e-03 1.851453e-04

18074 17420 16029 16250 17511

3.592371e-05 8.290087e-05 7.571613e-04 6.632069e-05 2.973378e-03

17548 13705 13747 13408 13767

4.697716e-05 7.737414e-05 2.763362e-06 1.384444e-03 1.102582e-03

17924 13448 15862 15513 16218

1.050078e-04 5.499091e-04 4.062142e-04 8.676957e-04 2.487026e-04

14045 14307 17908 17920 12838

1.658017e-05 4.974052e-04 1.602750e-04 1.923300e-03 3.398936e-04

13255 16583 18085 13758 13694

3.868707e-05 3.868707e-05 8.013750e-05 3.233134e-04 1.616567e-03

15983 14849 17968 16210 17897

3.122599e-04 1.083238e-03 2.348858e-04 3.398936e-04 3.177867e-04

17377 16552 17181 17951 14729

1.157849e-03 4.697716e-05 6.355733e-05 5.803061e-05 1.961987e-04

12748 15012 12868 17572 14078

1.282753e-02 3.537104e-04 2.956798e-04 3.316035e-05 1.796185e-04

14001 15525 14237 17905 15485

1.105345e-04 4.974052e-04 2.487026e-05 9.395432e-05 2.376492e-04

16955 15350 15605 18144 15922

7.184742e-05 1.381681e-05 1.851453e-04 1.188246e-04 3.316035e-05

14594 15165 16456 17841 17346

1.464582e-04 7.461078e-05 3.067332e-04 2.205992e-02 1.389971e-03

17643 17873 13093 12921 13468

2.210690e-05 1.271147e-04 4.697716e-04 2.047651e-03 8.455888e-04

17760 16928 16048 16274 14496

6.908406e-05 5.885962e-04 2.487026e-05 1.851453e-04 5.250388e-05

14696 16539 17025 13777 17690

3.675272e-04 1.243513e-04 2.487026e-05 6.051763e-04 7.129475e-04

12947 17460 18229 14142 17069

2.846263e-04 1.022444e-04 4.531914e-04 6.079397e-05 4.255578e-04

13065 14606 16835 15235 13576

3.868707e-05 7.687674e-03 2.763362e-05 3.951608e-04 3.592371e-04

18011 13090 15694 14741 13715

7.737414e-05 4.449013e-04 2.183056e-04 1.630384e-04 2.984431e-04

14092 17732 12855 15752 17855

5.996496e-04 4.974052e-05 8.290087e-06 1.138505e-03 4.697716e-05

14047 17925 13941 17017 14135

9.395432e-05 2.763362e-06 1.298780e-04 7.405811e-04 3.702905e-04

13108 15601 13418 14766 15658

2.763362e-05 1.144032e-03 8.676957e-04 3.785806e-04 1.547483e-04

14388 14901 18041 15955 15070

5.278022e-04 2.708095e-04 1.309834e-03 5.112220e-04 2.763362e-06

16244 15111 14390 16546 15260

2.514660e-04 2.321224e-04 8.041384e-04 8.566423e-05 2.155423e-04

13305 14491 14060 15923 16752

2.238323e-04 4.697716e-05 5.609625e-04 5.803061e-05 2.487026e-05

17287 15363 12915 15544 15738

3.426569e-04 4.697716e-05 6.079397e-05 7.212375e-04 4.863518e-04

16042 17381 15827 14180 13117

2.072522e-04 3.039698e-04 3.702905e-04 6.853138e-04 2.100155e-04

16916 17964 14466 17235 16510

3.951608e-04 2.873897e-04 1.989621e-04 8.013750e-05 3.592371e-05

17802 15107 17976 14449 15838

4.062142e-04 5.526724e-05 1.796185e-04 7.129475e-04 4.670082e-04

16781 17547 13491 16186 17685

5.250388e-05 5.526724e-06 2.376492e-04 6.051763e-04 3.592371e-04

17581 15732 13138 15823 17567

1.249040e-03 3.868707e-05 1.740918e-04 4.697716e-05 4.891151e-04

15061 16203 15640 15574 16770

1.132979e-03 9.948104e-05 1.367864e-03 4.642449e-04 4.034509e-04

17838 17228 14829 17412 14031

3.785806e-04 5.747793e-04 1.243513e-04 1.823819e-04 7.129475e-04

14775 12971 15834 17659 15299

1.658017e-04 8.511156e-04 7.626880e-04 4.449013e-04 3.316035e-05

15646 13958 14443 16995 13402

9.948104e-05 1.381681e-05 2.708095e-04 2.763362e-06 1.077711e-04

18168 17757 14625 13011 13798

3.868707e-04 2.050415e-03 1.575116e-04 8.290087e-06 1.213116e-03

15384 14264 13295 16754 16634

1.188246e-04 5.250388e-05 3.039698e-05 5.526724e-06 2.487026e-05

18239 14576 13145 14395 14865

2.431759e-04 2.763362e-06 1.934354e-05 9.754669e-04 1.105345e-05

15093 16150 17552 14236 17961

3.094966e-04 3.758173e-04 1.409315e-04 4.145043e-05 6.880772e-04

14573 17135 17396 14213 12967

6.659703e-04 1.436948e-04 7.461078e-05 1.381681e-05 9.119095e-05

14679 15240 18225 13370 16883

2.763362e-06 2.293591e-04 7.903216e-04 1.215879e-04 1.050078e-04

12841 16905 17967 16891 14589

1.213116e-03 5.775427e-04 1.326414e-04 5.609625e-04 8.290087e-06

14680 17884 14083 13013 14210

8.842759e-04 3.233134e-04 5.001686e-04 6.134664e-04 2.846263e-04

16477 16013 17949 16926 17787

4.034509e-04 4.172677e-04 2.183056e-04 3.868707e-05 3.702905e-04

14723 17954 17819 15373 16140

7.682147e-04 1.351284e-03 1.354047e-04 3.785806e-04 1.022444e-04

17198 17238 15769 14396 14898

2.652828e-04 1.564063e-03 4.062142e-04 2.017254e-04 3.288401e-04

16725 16455 13081 15545 17243

4.200311e-04 2.735729e-04 2.931927e-03 2.238323e-04 1.124688e-03

15465 15708 13089 16033 13838

2.873897e-04 8.207186e-04 5.131564e-03 3.183393e-03 4.808250e-04

15351 18055 15038 18109 13069

6.438634e-04 1.003100e-03 3.675272e-04 1.254566e-03 1.298780e-03

16241 14800 16839 16168 16931

1.829346e-03 1.083238e-03 8.234819e-04 2.735729e-04 2.481499e-03

16814 13269 14813 17062 14810

8.980927e-04 8.842759e-04 8.290087e-05 1.961987e-04 7.046574e-04

16353 18118 13831 13506 17796

2.597561e-04 3.548157e-03 2.044888e-04 8.013750e-05 2.265957e-04

15716 18156 15221 16983 17059

3.260767e-04 3.868707e-04 3.039698e-05 4.117410e-04 1.425895e-03

16327 17211 17026 15570 15880

8.815126e-04 5.885962e-04 1.105345e-05 9.119095e-04 2.929164e-04

16718 15808 17858 16393 14161

1.243513e-04 5.803061e-04 5.499091e-04 9.892837e-04 2.984431e-04

17863 16411 16402 15023 17402

4.089776e-04 1.243513e-04 1.658017e-04 1.555773e-03 1.630384e-04

15426 15894 15867 14506 15555

8.290087e-04 1.132979e-04 1.279437e-03 8.732225e-04 2.556110e-03

13034 15889 15953 16143 16422

1.022444e-04 2.569927e-04 2.846263e-04 3.260767e-04 1.102582e-03

12747 15502 14081 17965 14404

2.846263e-04 6.963673e-04 4.310845e-04 9.091462e-04 2.873897e-04

13579 13174 15987 17677 13652

6.355733e-05 8.676957e-04 3.868707e-04 8.870393e-04 1.740918e-04

17428 16161 16858 18219 14293

9.478332e-04 1.387208e-03 3.592371e-05 3.177867e-04 1.685651e-04

14748 16638 17674 13094 15899

2.542293e-04 2.625194e-04 7.461078e-05 8.290087e-05 1.105345e-05

15898 17691 17227 15039 15919

2.404125e-04 5.250388e-05 1.298780e-04 4.167150e-03 3.951608e-04

13497 15860 14189 15304 14867

5.526724e-05 3.923974e-04 7.654513e-04 7.184742e-05 1.022444e-04

14344 16719 15048 15301 14708

1.961987e-04 5.609625e-04 2.708095e-04 6.825505e-04 8.013750e-05

16898 14825 18113 17596 15078

4.504280e-04 3.343668e-04 2.763362e-06 7.461078e-04 1.246276e-03

14085 17860 13155 15028 16919

1.246276e-03 1.602750e-04 2.348858e-04 1.326414e-04 9.036195e-04

17259 17019 14487 16722 15882

3.426569e-04 6.466268e-04 4.006875e-04 6.300466e-04 6.632069e-05

14667 16710 16550 15984 17682

1.591697e-03 1.110872e-03 7.212375e-04 7.599246e-04 3.288401e-04

16658 17068 15356 17223 16817

2.210690e-05 8.207186e-04 1.434185e-03 1.354047e-04 2.625194e-04

13030 17191 15194 14409 16654

2.735729e-04 6.549169e-04 9.008561e-04 1.519849e-04 8.566423e-05

16861 16138 13481 13495 17519

2.210690e-05 2.763362e-06 3.094966e-04 1.436948e-04 2.293591e-04

17218 17812 14215 15602 17320

7.571613e-04 3.896341e-04 3.012065e-04 1.492216e-04 2.017254e-04

16782 12913 17869 13564 14850

5.443824e-04 1.381681e-04 1.105345e-04 2.348858e-04 6.079397e-05

15021 13963 17450 17091 17470

1.323651e-03 9.671768e-05 9.699401e-04 5.471457e-04 2.155423e-04

17616 15649 14907 17735 16059

8.290087e-06 5.526724e-06 4.642449e-04 1.906720e-03 4.974052e-04

17720 15018 14733 13756 14176

1.381681e-04 8.842759e-05 5.968862e-04 3.951608e-04 8.842759e-05

13319 15353 18119 17491 14282

1.381681e-03 1.740918e-04 8.290087e-05 3.039698e-04 4.310845e-04

13569 15498 14673 13488 16221

2.348858e-04 1.445238e-03 2.044888e-04 1.179956e-03 2.680461e-04

15002 15965 15024 17675 13769

3.177867e-04 3.371302e-04 4.283211e-04 1.992384e-03 3.951608e-04

16904 16086 13786 13880 13173

1.961987e-03 1.105345e-04 1.934354e-05 4.559548e-04 1.188246e-04

15858 15180 14739 16293 17419

5.278022e-04 1.658017e-05 1.685651e-04 3.316035e-04 5.747793e-04

16775 16306 16950 17591 12839

4.366112e-04 1.215879e-04 1.160612e-04 4.614815e-04 8.732225e-04

13807 16125 15081 18102 16365

1.934354e-05 3.592371e-05 9.671768e-05 1.196536e-03 1.409315e-04

17870 14821 16863 13078 13267

7.737414e-04 5.526724e-06 7.737414e-05 1.050078e-03 6.189931e-04

13050 15628 14796 16779 13599

1.166139e-03 4.338479e-04 3.219317e-03 9.505966e-04 7.295276e-04

13329 18077 14744 17406 15750

5.250388e-05 6.272832e-04 7.737414e-05 3.205500e-04 8.151919e-04

13983 13842 13199 14032 17706

3.094966e-04 5.692526e-04 3.316035e-04 4.200311e-04 1.091528e-03

16081 16525 14952 17975 16122

6.272832e-04 5.526724e-04 4.117410e-04 8.207186e-04 1.050078e-04

13126 15973 16558 16713 16893

3.923974e-04 1.658017e-05 1.309834e-03 1.746445e-03 1.243513e-04

15529 18043 16016 17722 17324

1.862506e-03 3.343668e-04 6.549169e-04 3.316035e-04 1.381681e-04

17341 14051 15916 16233 16385

5.056953e-04 5.941229e-04 4.421380e-04 6.908406e-05 1.630384e-04

15032 17213 18061 13113 16553

1.519849e-04 4.531914e-04 1.271147e-04 7.737414e-04 2.376492e-04

13455 14217 17262 14299 18259

1.464582e-04 2.956798e-04 3.316035e-05 4.476647e-04 1.160612e-04

17676 16679 18269 15298 18193

2.183056e-04 8.290087e-06 2.210690e-05 5.360923e-04 4.974052e-05

13523 12870 15005 15581 15680

4.891151e-04 5.526724e-06 3.205500e-03 4.172677e-04 2.155423e-04

13649 17894 13270 14618 14866

6.355733e-05 5.333289e-04 2.763362e-06 4.697716e-05 2.846263e-04

13077 13115 13531 13848 13140

3.094966e-04 3.647638e-04 6.079397e-05 1.381681e-05 1.271147e-04

15713 17999 15347 17340 13984

5.526724e-05 2.569927e-04 1.381681e-05 1.171666e-03 4.697716e-05

15996 14030 16252 16596 16700

3.620005e-04 6.714970e-04 5.803061e-05 3.316035e-05 8.621690e-04

16011 17655 16717 14243 12963

5.858328e-04 2.569927e-04 6.521535e-04 4.089776e-04 2.652828e-04

15079 14437 12942 15660 18092

1.326414e-04 1.658017e-05 2.072522e-04 2.597561e-04 2.487026e-04

15358 18071 13854 13369 17857

1.550246e-03 6.908406e-05 3.150233e-04 5.250388e-05 1.547483e-04

17411 14987 14355 17526 14896

1.188246e-04 1.105345e-05 3.039698e-05 6.355733e-05 1.188246e-04

14440 13240 14479 16065 17430

1.519849e-04 1.851453e-04 4.974052e-05 2.044888e-04 3.868707e-05

16520 15945 15181 15856 13102

4.338479e-04 5.526724e-06 4.974052e-05 1.804476e-03 8.566423e-04

14205 17702 15271 13982 15288

1.077711e-04 2.210690e-04 7.599246e-04 3.177867e-04 2.431759e-04

17978 16795 16519 17816 13021

3.316035e-05 1.823819e-04 3.592371e-05 4.697716e-05 3.702905e-04

16255 17671 17646 16350 12872

7.571613e-04 1.685651e-04 1.492216e-04 1.658017e-04 1.823819e-04

17339 18230 16499 15192 14441

2.487026e-05 2.459392e-04 6.079397e-05 1.409315e-04 1.630384e-04

15279 16579 15380 17442 17950

5.803061e-05 2.763362e-06 1.022444e-04 3.868707e-04 1.989621e-04

12826 17456 17307 15811 13327

2.597561e-04 4.697716e-05 2.763362e-06 3.841074e-04 3.012065e-04

17700 13136 15535 15329 15211

4.034509e-04 3.841074e-04 2.155423e-04 1.713285e-04 1.436948e-04

15596 16163 16907 17576 14713

2.321224e-04 1.934354e-05 4.172677e-04 1.152322e-03 9.423065e-04

14702 13230 13969 17530 15805

9.616501e-04 1.691178e-03 1.749208e-03 1.088765e-03 8.483522e-04

17696 17076 17827 16265 17880

3.481836e-04 6.908406e-05 6.079397e-04 7.184742e-04 7.184742e-05

13000 17790 14419 14082 17912

8.566423e-05 5.250388e-04 5.692526e-04 8.013750e-05 1.105345e-04

15723 16656 14037 14256 14543

1.132979e-04 2.210690e-04 7.737414e-05 1.492216e-04 2.956798e-04

16062 18116 16019 18178 18065

1.768552e-04 1.044551e-03 4.559548e-04 2.735729e-04 3.813440e-04

13148 12989 14107 15881 16727

9.588867e-04 1.879086e-04 3.426569e-04 8.842759e-05 1.271147e-04

15514 13787 17865 17220 16565

6.024130e-04 8.566423e-05 1.074948e-03 1.304307e-03 8.290087e-06

16873 14062 15179 18016 14448

7.986117e-04 3.067332e-04 2.625194e-04 3.398936e-04 5.526724e-05

15503 17251 16923 14159 17126

1.934354e-04 1.077711e-04 1.923300e-03 3.349195e-03 7.240009e-04

15224 14577 17236 17372 16367

4.421380e-05 3.398936e-04 2.901530e-04 6.991306e-04 3.592371e-04

12875 17917 16010 17085 13668

5.526724e-06 2.597561e-04 2.376492e-04 5.222755e-04 1.384444e-03

14684 17188 15547 15998 14413

1.163376e-03 2.459392e-04 1.939880e-03 1.406551e-03 6.355733e-05

18004 13959 16665 15454 15044

1.105345e-04 2.542293e-04 2.487026e-05 6.908406e-05 1.273910e-03

17937 14608 17618 14505 18037

1.188246e-04 5.526724e-05 2.155423e-04 2.218980e-03 5.250388e-05

13198 13225 15361 15689 14290

6.328100e-04 9.395432e-05 1.934354e-05 2.487026e-05 2.846263e-04

13209 13999 14498 15854 16701

9.091462e-04 5.858328e-04 5.333289e-04 3.868707e-04 4.255578e-04

16927 17990 14544 17338 17231

1.658017e-05 8.013750e-05 3.150233e-04 2.091865e-03 1.508796e-03

17799 14565 16726 17303 14639

1.132979e-03 3.564737e-04 5.112220e-04 1.464582e-04 4.835884e-04

17611 17179 17524 17293 17890

1.608277e-03 2.210690e-04 1.658017e-05 3.730539e-04 2.597561e-04

17375 17371 17969 15159 13211

1.961987e-04 2.404125e-04 3.868707e-05 1.989621e-03 1.547483e-04

15777 14525 13324 13802 13948

6.079397e-05 8.234819e-04 2.127789e-04 1.823819e-04 8.290087e-06

13008 14206 16202 16442 18075

4.145043e-05 5.803061e-05 4.974052e-05 4.808250e-04 4.227944e-04

15468 16570 14524 18141 16810

8.013750e-05 3.675272e-04 5.609625e-04 2.763362e-06 4.421380e-05

14961 17041 16458 16083 16262

8.759858e-04 1.961987e-04 5.581992e-04 8.566423e-05 8.290087e-06

17194 14057 14871 13178 18044

8.013750e-05 8.041384e-04 2.708095e-04 7.322910e-04 3.537104e-04

16812 16191 14298 17158 14415

4.145043e-05 4.172677e-04 4.531914e-03 1.851453e-04 1.691178e-03

16899 16531 16556 16805 16503

3.454203e-04 2.127789e-04 5.720160e-04 2.708095e-04 2.376492e-04

17739 13238 12829 13097 12867

1.326414e-04 8.842759e-05 3.316035e-05 1.207589e-03 1.522613e-03

15392 13846 14401 16145 17629

2.348858e-04 1.602750e-04 5.056953e-04 6.549169e-04 3.426569e-04

14150 13496 13611 16954 17454

7.184742e-05 9.948104e-05 2.100155e-04 1.492216e-04 1.685651e-04

17449 17230 15950 15512 17525

7.709781e-04 2.017254e-04 2.846263e-04 1.934354e-05 2.487026e-05

13458 14426 16003 13092 13426

1.088765e-03 7.295276e-04 2.846263e-04 7.184742e-05 4.393746e-04

15101 17033 13221 14201 17692

2.487026e-05 7.737414e-05 1.823819e-04 1.077711e-04 1.298780e-04

15822 16134 17551 18050 14913

9.119095e-05 1.050078e-04 1.188246e-04 1.381681e-05 2.404125e-04

15641 13124 17214 15456 15062

3.288401e-04 6.493901e-04 2.873897e-04 1.658017e-04 1.575116e-04

12928 15727 17980 15290 17752

2.321224e-04 8.345354e-04 8.290087e-05 5.858328e-04 2.763362e-06

15615 17044 13829 17633 13266

1.185482e-03 2.763362e-05 2.763362e-06 1.989621e-04 5.581992e-04

17820 15664 17189 17496 16607

1.381681e-05 5.250388e-05 1.685651e-04 2.763362e-05 4.808250e-04

13317 13890 17001 15246 16996

3.730539e-04 3.150233e-04 4.670082e-04 9.119095e-05 1.132979e-04

15482 16037 14841 16670 16985

3.509470e-04 1.077711e-04 3.813440e-04 7.240009e-04 3.343668e-04

16705 16470 14414 15563 12980

7.847949e-04 4.725349e-04 2.708095e-04 3.233134e-04 6.632069e-05

16746 14878 16743 17146 12877

1.249040e-03 3.426569e-04 5.333289e-04 9.671768e-05 4.255578e-04

15394 14621 15379 14309 17288

2.348858e-04 2.652828e-04 5.416190e-04 3.398936e-04 3.923974e-04

14709 15157 17744 16316 15620

5.443824e-04 3.675272e-04 3.785806e-04 4.310845e-04 2.873897e-04

15789 16560 13534 17580 14777

6.908406e-05 1.602750e-04 9.450699e-04 1.243513e-04 8.290087e-06

15759 14373 18176 15366 14735

1.934354e-05 1.105345e-05 9.119095e-05 1.215879e-04 5.664893e-04

16271 18223 15532 16816 16684

3.702905e-04 8.262453e-04 7.267643e-04 7.737414e-05 7.765048e-04

14334 16945 14643 15197 14220

4.504280e-04 6.880772e-04 9.395432e-05 9.948104e-05 5.803061e-05

13874 18260 17848 17392 14985

4.974052e-05 3.868707e-04 1.409315e-04 1.713285e-04 2.404125e-04

13004 14560 12931 13627 14769

1.130215e-03 2.100155e-04 2.818629e-04 2.597561e-04 3.023118e-03

15780 17315 16711 15204 17888

8.013750e-05 1.345757e-03 6.079397e-05 2.487026e-05 2.238323e-04

17007 15160 13033 13869 14146

1.464582e-04 1.105345e-05 3.039698e-05 8.483522e-04 6.770237e-04

16885 17913 14472 17358 16686

2.790996e-04 1.630384e-04 7.765048e-04 1.630384e-04 1.624857e-03

18245 17343 18256 16153 12944

4.891151e-04 4.697716e-05 1.105345e-05 2.459392e-04 8.290087e-05

15312 14099 14113 13104 17204

1.326414e-04 6.963673e-04 1.105345e-04 5.526724e-05 1.851453e-04

13027 17621 17090 14119 15527

7.184742e-05 7.461078e-05 1.768552e-04 2.763362e-06 4.587181e-04

18097 15460 15034 15208 17306

3.012065e-04 1.602750e-04 1.105345e-03 1.188246e-04 1.906720e-04

14737 13029 16270 13487 17400

4.697716e-05 2.763362e-05 1.492216e-04 9.395432e-05 5.885962e-04

17593 13953 14533 14944 17239

5.941229e-04 5.526724e-05 1.796185e-04 4.752983e-04 4.421380e-05

14532 15562 15587 13922 13165

1.768552e-04 2.763362e-06 2.487026e-05 1.934354e-05 1.160612e-04

13304 13313 16875 16282 12957

5.803061e-05 2.155423e-04 3.398936e-04 3.039698e-05 6.742604e-04

14462 13397 13680 15164 17457

6.797871e-04 1.077711e-04 5.968862e-04 2.846263e-04 5.250388e-05

14732 14221 14553 13253 17389

4.117410e-04 5.360923e-04 3.122599e-04 9.948104e-05 6.189931e-04

15841 16112 13629 16600 15661

1.381681e-04 4.421380e-05 1.630384e-04 1.713285e-04 1.381681e-05

16532 15266 14653 16771 17585

1.105345e-05 1.658017e-05 6.549169e-04 7.046574e-04 5.913595e-04

15046 13082 15464 18226 18062

5.720160e-04 2.072522e-04 4.338479e-04 6.770237e-04 7.461078e-05

17368 14258 15719 13368 17634

3.537104e-04 3.094966e-04 2.592034e-03 8.013750e-05 8.207186e-04

16676 16424 18283 13764 13656

2.542293e-04 1.658017e-05 2.089102e-03 5.747793e-04 1.354047e-04

14312 13280 15845 16094 13187

2.404125e-04 2.348858e-04 3.868707e-04 1.851453e-04 1.022444e-04

16567 14828 16551 16889 16473

9.671768e-05 4.697716e-05 8.842759e-05 3.730539e-04 5.250388e-05

13750 15607 15139 17707 18198

6.632069e-05 3.592371e-05 3.592371e-05 1.381681e-05 4.587181e-04

18179 17284 13233 14889 14672

2.487026e-04 8.124285e-04 1.188246e-04 2.763362e-05 1.685651e-04

13593 13162 13183 17095 12939

3.564737e-04 2.265957e-04 2.597561e-04 2.127789e-04 1.298780e-04

15611 17114 15059 16326 17128

1.243513e-04 3.730539e-04 6.604436e-04 8.842759e-04 3.868707e-05

14761 17555 16735 16833 17800

3.039698e-05 2.929164e-04 3.537104e-04 1.298780e-04 2.873897e-04

16748 14755 16549 13565 15799

1.906720e-04 6.549169e-04 2.710858e-03 1.050078e-04 2.708095e-04

18032 13451 14794 18171 16279

1.381681e-05 1.226933e-03 2.265957e-04 1.547483e-04 3.012065e-04

17811 16850 14321 15326 17849

2.409652e-03 5.250388e-05 3.398936e-04 8.013750e-05 7.184742e-05

13672 14229 13107 16497 16303

4.421380e-05 1.188246e-04 1.658017e-04 3.758173e-04 4.614815e-04

15281 13473 14320 17742 16351

4.145043e-04 6.079397e-05 1.160612e-04 4.145043e-05 2.210690e-05

13144 15346 15370 14295 16034

8.290087e-06 4.697716e-05 2.625194e-04 1.934354e-05 6.079397e-05

17365 15189 14669 18095 15518

1.293254e-03 6.411000e-04 4.697716e-04 5.803061e-05 6.079397e-04

15749 14209 15569 17569 13784

4.145043e-05 2.431759e-04 8.566423e-05 8.013750e-05 8.566423e-05

13862 16791 16235 14270 16655

3.923974e-04 8.898026e-04 1.326414e-04 8.013750e-05 7.212375e-04

16359 16554 15939 14587 14715

1.934354e-04 2.044888e-04 2.127789e-04 1.061131e-03 1.961987e-04

14185 14040 14514 15171 14167

8.290087e-06 6.659703e-04 2.017254e-04 1.934354e-05 1.436948e-04

[ reached getOption("max.print") -- omitted 2950 entries ]

Conditional probabilities:

If,I look at the above mentione output,the Apriori Probabilites of different customer IDs are listed accordingly

Q5) If the number of observations is loaded in one of the clusters, break down that cluster further using the clustering algorithm. [ hint: Here loaded means if any cluster has more number of data points as compared to other clusters then split that clusters by increasing the number of clusters and observe, compare the results with previous results.]

Ans) The below mention is the clustering Algorithm which gives an overview of various Clusters. The code below is mentioned in R and the analysis is done on on the columns of **InvoiceNo,Quantity,UnitPrice and CustomerID of UK customers only.**

ecom\_data<-read.csv("https://raw.githubusercontent.com/shivanipriya89/Ecommerce/main/MyEcommerce.csv")

print(ecom\_data)

View(ecom\_data)

str(ecom\_data)

ecom\_data$InvoiceNo<-as.integer(ecom\_data$InvoiceNo)

ecom\_data<-na.omit(ecom\_data)

str(ecom\_data)

ecom\_data<- scale(ecom\_data)

k2<-kmeans(ecom\_data, centers = 2, nstart = 25)

str(k2)

fviz\_cluster(k2, data =ecom\_data)

k3<-kmeans(ecom\_data, centers = 3, nstart = 25)

fviz\_cluster(k3, data =ecom\_data)

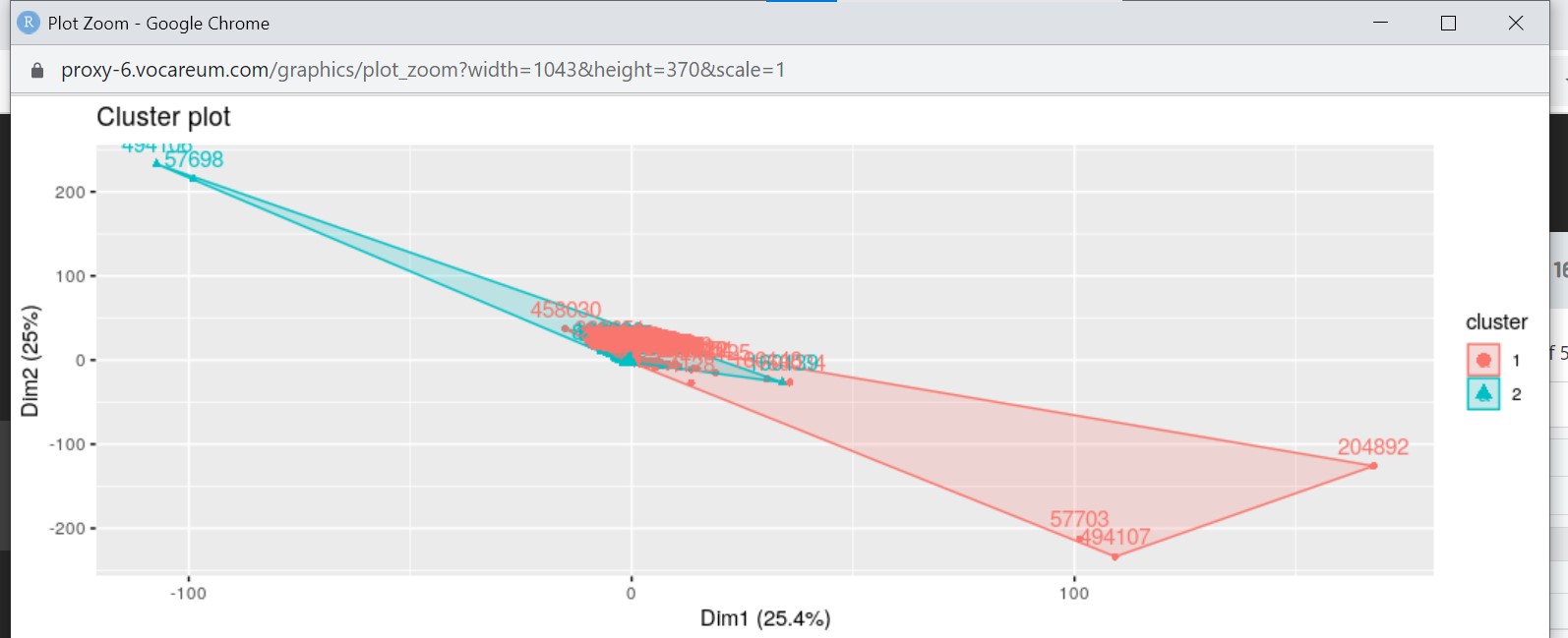
str(k3)

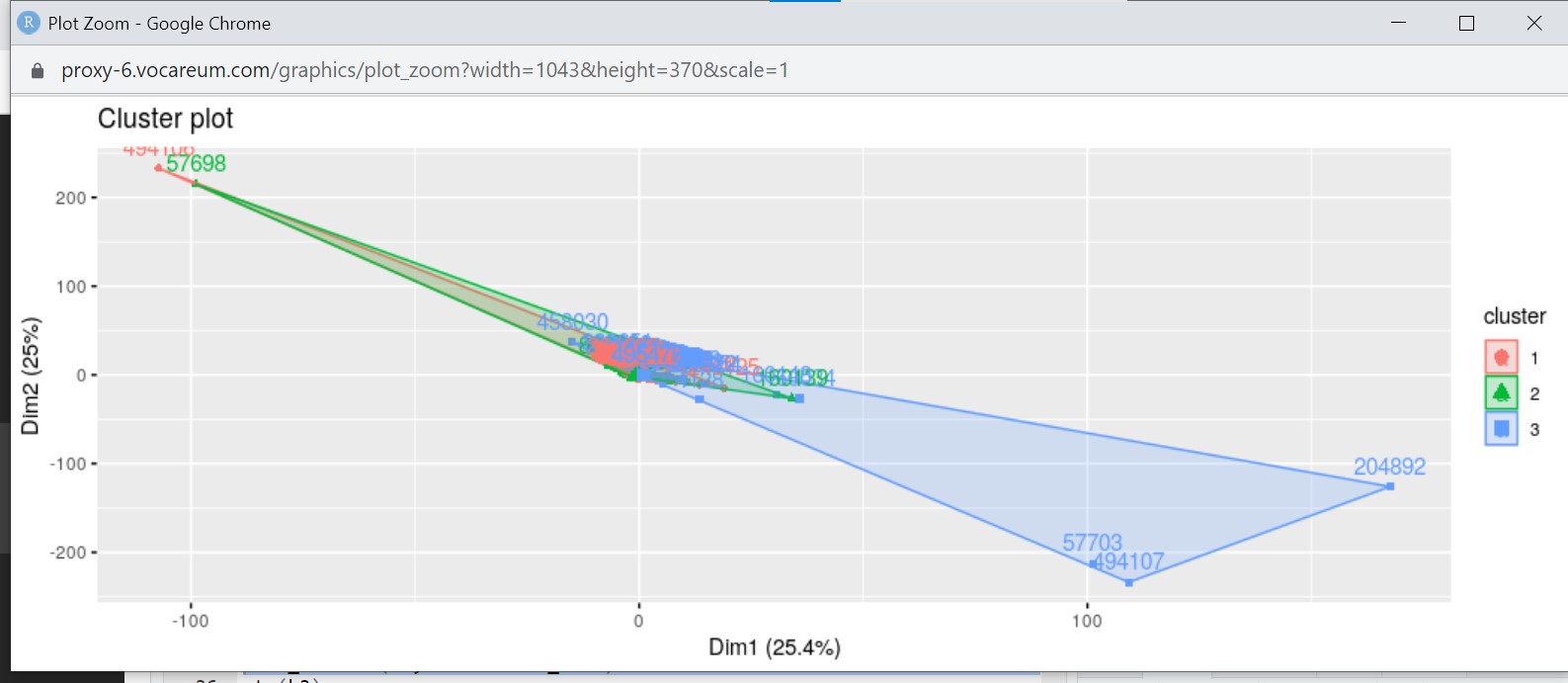
k4<-kmeans(ecom\_data, centers = 4, nstart = 25)

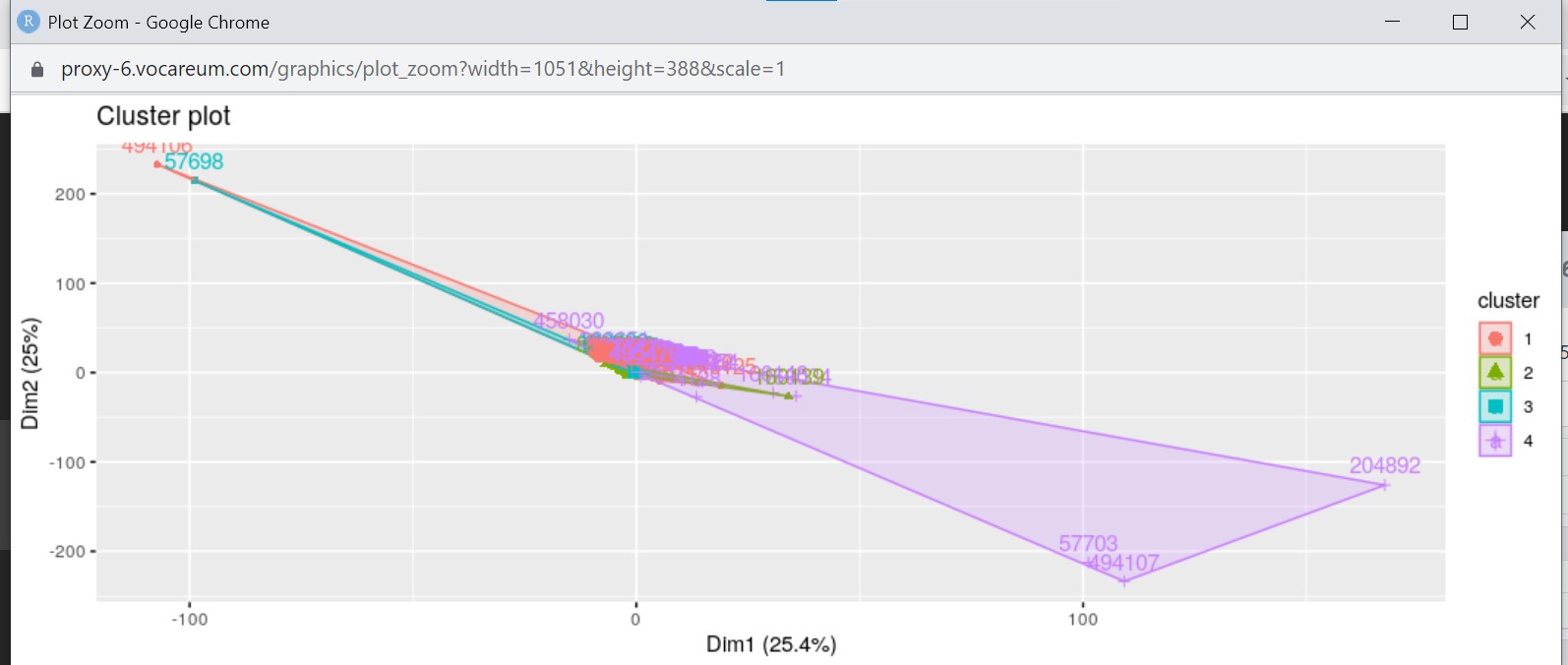
fviz\_cluster(k4, data =ecom\_data)

str(k4)

Below attached are the screenshots of various clusters







The above mention screenshot represents the datapoints of various Clusters