

Online shoppers Dataset

Martha Irungu

9/11/2020

Ia) Specifying the question

The objective of this study is to support the brand's Sales and Marketing team of Kira Plastinina (a Russian brand) to understand their customer's behavior from the Ecommerce customer data that they have collected over the past year.

b)Defining the Metrics for success

To meet the objective of the study we will need to do the following:

- i) Perform clustering stating insights drawn from your analysis and visualizations.
- ii) Upon implementation, provide comparisons between i.e. K-Means clustering vs Hierarchical clustering highlighting the strengths and limitations of each approach in the context of your analysis.
- iii) Based on the analysis and clustering implemented, provide the brand's Sales and Marketing team of Kira Plastinina with the characteristics of customer groups of their products.

c) Understanding the context

Kira Plastinina is a Russian brand that is sold through a defunct chain of retail stores in Russia, Ukraine, Kazakhstan, Belarus, China, Philippines, and Armenia. This fashion brand owner was born in Moscow whose father invested in the business for her love of fashion. In 2007, the first Kira Plastinina store opened in Moscow, Plastinina introduced her first collection and became one of the youngest fashion designers in the world. Since then, the company has opened over 300 stores in Russia and Common Wealth of Independent States. In 2008, the Company made an unsuccessful attempt to enter the U.S. market. Throughout her career, Plastinina has presented her fashion collections during Rome, Milan, New York and Moscow fashion weeks. Her brand has been worn by many celebrities including Paris Hilton,[3] Lindsay Lohan,[4] Georgia May Jagger,[5] Karlie Kloss,[6] Rowan Blanchard, Lyndsy Fonseca, Victoria Justice, and many others.

This being a global teenage fashion designer, and having ecommerce options where customers can login, select and make a decision to make an order to purchase the fashion wear, it is critical for the entrepreneur to understand the characteristics of her customers. We will be investigating factors like region they come from, the browser, informational duration among other factors to establish how many clusters they belong to and what are the characteristics of specific clusters.

d) Recording the experimental design

The following steps were implemented

- 1.) Business Understanding.
- 2.) Reading the data.
- 3.) Perform Data Cleaning
- 4.) Perform Exploratory Data analysis (Univariate, Bivariate & Multivariate)
- 5.) Implementing the Solution
- 6.) Challenge the Solution
- 7.) Follow up Questions

e) Data Relevance

The data provided for this study was collected in the past year. It consists of variables that can define characteristics of customers that visit the Ecommerce website. This is relevant data as it relates to online shoppers who visit the site. The brand exists as per link below: The data can therefore be relied on to help us establish the characteristics of customers who visit the site and their intentions.

https://en.wikipedia.org/wiki/Kira_Plastinina

2) Previewing and reading the data

#Loading the data and previewing the head

```
library("data.table")
online_shoppers <- fread("/Users/marthairungu/desktop/online_shoppers_intention.csv")
head(online_shoppers)
```

##	Administrative	Administrative_Duration	Informational	Informational_Duration	
## 1:	0	0	0		
## 2:	0	0	0		
## 3:	0	-1	0		
## 4:	0	0	0		
## 5:	0	0	0		
## 6:	0	0	0		
##	ProductRelated	ProductRelated_Duration	BounceRates	ExitRates	PageValues
## 1:	1	0.000000	0.20000000	0.2000000	0

```
## 2:      2      64.000000  0.00000000 0.1000000 0
## 3:      1     -1.000000  0.20000000 0.2000000 0
## 4:      2      2.666667  0.05000000 0.1400000 0
## 5:     10     627.500000  0.02000000 0.0500000 0
## 6:     19    154.216667  0.01578947 0.0245614 0
##   SpecialDay Month OperatingSystems Browser Region TrafficType
## 1:      0   Feb           1           1      1          1
## 2:      0   Feb           2           2      1          2
## 3:      0   Feb           4           1      9          3
## 4:      0   Feb           3           2      2          4
## 5:      0   Feb           3           3      1          4
## 6:      0   Feb           2           2      1          3
##           VisitorType Weekend Revenue
## 1: Returning_Visitor   FALSE   FALSE
## 2: Returning_Visitor   FALSE   FALSE
## 3: Returning_Visitor   FALSE   FALSE
## 4: Returning_Visitor   FALSE   FALSE
## 5: Returning_Visitor    TRUE   FALSE
## 6: Returning_Visitor   FALSE   FALSE
```

#Checking the dimension of the dataset

```
dim(online_shoppers)
```

```
## [1] 12330    18
```

#The dataset has 12,330 observations and 18 variables

#Checking the structure of the dataset

```
str(online_shoppers)
```

```
## Classes 'data.table' and 'data.frame':  12330 obs. of  18 variables:
## $ Administrative      : int  0 0 0 0 0 0 0 1 0 0 ...
## $ Administrative_Duration: num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ Informational        : int  0 0 0 0 0 0 0 0 0 0 ...
## $ Informational_Duration : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ ProductRelated       : int  1 2 1 2 10 19 1 1 2 3 ...
## $ ProductRelated_Duration: num  0 64 -1 2.67 627.5 ...
## $ BounceRates           : num  0.2 0 0.2 0.05 0.02 ...
## $ ExitRates             : num  0.2 0.1 0.2 0.14 0.05 ...
## $ PageValues            : num  0 0 0 0 0 0 0 0 0 0 ...
## $ SpecialDay            : num  0 0 0 0 0 0 0.4 0 0.8 0.4 ...
## $ Month                 : chr  "Feb" "Feb" "Feb" "Feb" ...
## $ OperatingSystems      : int  1 2 4 3 3 2 2 1 2 2 ...
## $ Browser               : int  1 2 1 2 3 2 4 2 2 4 ...
## $ Region                : int  1 1 9 2 1 1 3 1 2 1 ...
## $ TrafficType           : int  1 2 3 4 4 3 3 5 3 2 ...
## $ VisitorType           : chr  "Returning_Visitor" "Returning_Visitor" "
Returning_Visitor" "Returning_Visitor" ...
## $ Weekend              : logi  FALSE FALSE FALSE FALSE TRUE FALSE ...
```

```
## $ Revenue          : logi FALSE FALSE FALSE FALSE FALSE ...
## - attr(*, ".internal.selfref")=<externalptr>
```

#The variables have datatypes in interger,number, character and logical datatypes. We will convert the varibales as appropriate as we analyse the data.

#Checking the summary of the dataset

```
summary(online_shoppers)
```

```
## Administrative      Administrative_Duration Informational
## Min.   : 0.000      Min.   : -1.00      Min.   : 0.000
## 1st Qu.: 0.000      1st Qu.:  0.00      1st Qu.: 0.000
## Median : 1.000      Median :  8.00      Median : 0.000
## Mean   : 2.318      Mean   : 80.91      Mean   : 0.504
## 3rd Qu.: 4.000      3rd Qu.: 93.50      3rd Qu.: 0.000
## Max.   :27.000      Max.   :3398.75      Max.   :24.000
## NA's   :14          NA's   :14          NA's   :14
## Informational_Duration ProductRelated      ProductRelated_Duration
## Min.   : -1.00      Min.   :  0.00      Min.   : -1.0
## 1st Qu.:  0.00      1st Qu.:  7.00      1st Qu.: 185.0
## Median :  0.00      Median : 18.00      Median :  599.8
## Mean   : 34.51      Mean   : 31.76      Mean   : 1196.0
## 3rd Qu.:  0.00      3rd Qu.: 38.00      3rd Qu.: 1466.5
## Max.   :2549.38      Max.   :705.00      Max.   :63973.5
## NA's   :14          NA's   :14          NA's   :14
## BounceRates          ExitRates          PageValues          SpecialDay
## Min.   :0.000000      Min.   :0.00000      Min.   : 0.000      Min.   :0.00000
## 1st Qu.:0.000000      1st Qu.:0.01429      1st Qu.: 0.000      1st Qu.:0.00000
## Median :0.003119      Median :0.02512      Median : 0.000      Median :0.00000
## Mean   :0.022152      Mean   :0.04300      Mean   : 5.889      Mean   :0.06143
## 3rd Qu.:0.016684      3rd Qu.:0.05000      3rd Qu.: 0.000      3rd Qu.:0.00000
## Max.   :0.200000      Max.   :0.20000      Max.   :361.764      Max.   :1.00000
## NA's   :14          NA's   :14
## Month                OperatingSystems      Browser                Region
## Length:12330          Min.   :1.000      Min.   : 1.000      Min.   :1.000
## Class :character      1st Qu.:2.000      1st Qu.: 2.000      1st Qu.:1.000
## Mode  :character      Median :2.000      Median : 2.000      Median :3.000
##                        Mean   :2.124      Mean   : 2.357      Mean   :3.147
##                        3rd Qu.:3.000      3rd Qu.: 2.000      3rd Qu.:4.000
##                        Max.   :8.000      Max.   :13.000      Max.   :9.000
##
## TrafficType          VisitorType          Weekend          Revenue
## Min.   : 1.00      Length:12330      Mode :logical      Mode :logical
## 1st Qu.: 2.00      Class :character  FALSE:9462          FALSE:10422
## Median : 2.00      Mode  :character  TRUE :2868           TRUE :1908
## Mean   : 4.07
## 3rd Qu.: 4.00
## Max.   :20.00
##
```

#The summary of the numerical variables is as tabulated

3)Data Cleaning

#Getting column names

```
colnames(online_shoppers)
```

```
## [1] "Administrative"      "Administrative_Duration"
## [3] "Informational"      "Informational_Duration"
## [5] "ProductRelated"    "ProductRelated_Duration"
## [7] "BounceRates"       "ExitRates"
## [9] "PageValues"        "SpecialDay"
## [11] "Month"             "OperatingSystems"
## [13] "Browser"           "Region"
## [15] "TrafficType"       "VisitorType"
## [17] "Weekend"           "Revenue"
```

#For ease of working with the data, we will change column names

```
names(online_shoppers)[1]<- 'admin'
names(online_shoppers)[2]<- 'admin_dur'
names(online_shoppers)[3]<- 'info'
names(online_shoppers)[4]<- 'info_dur'
names(online_shoppers)[5]<- 'prod'
names(online_shoppers)[6]<- 'prod_dur'
names(online_shoppers)[7]<- 'bouncerrates'
names(online_shoppers)[8]<- 'exitrates'
names(online_shoppers)[9]<- 'pagevalues'
names(online_shoppers)[10]<- 'specialday'
names(online_shoppers)[11]<- 'month'
names(online_shoppers)[12]<- 'ops_systems'
names(online_shoppers)[13]<- 'browser'
names(online_shoppers)[14]<- 'region'
names(online_shoppers)[15]<- 'traffic_type'
names(online_shoppers)[16]<- 'visitor_type'
names(online_shoppers)[17]<- 'weekend'
names(online_shoppers)[18]<- 'revenue'
```

#Confirming the variable names have been changed

```
colnames(online_shoppers)
```

```
## [1] "admin"      "admin_dur"  "info"      "info_dur"  "prod"
## [6] "prod_dur"   "bouncerrates" "exitrates" "pagevalues" "speciald
ay"
## [11] "month"      "ops_systems" "browser"   "region"    "traffic_
type"
## [16] "visitor_type" "weekend"    "revenue"
```

#Description of the variables

#Administrative","Administrative Duration","Informational","Informational Duration","Product Related" and "Product Related Duration" represents the number of different types of pages visited by the visitor in that session and total time spent in each of these page categories.

#The "Bounce Rate", "Exit Rate" and "Page Value" features represent the metrics measured by "Google Analytics" for each page in the e-commerce site.

#The value of the "Bounce Rate" feature for a web page refers to the percentage of visitors who enter the site from that page and then leave ("bounce") without triggering any other requests to the analytics server during that session.

#The value of the "Exit Rate" feature for a specific web page is calculated as for all pageviews to the page, the percentage that was the last in the session.

#The "Page Value" feature represents the average value for a web page that a user visited before completing an e-commerce transaction.

#The "Special Day" feature indicates the closeness of the site visiting time to a specific special day (e.g. Mother's Day, Valentine's Day) in which the sessions are more likely to be finalized with the transaction. The value of this attribute is determined by considering the dynamics of e-commerce such as the duration between the order date and delivery date. For example, for Valentine's day, this value takes a nonzero value between February 2 and February 12, zero before and after this date unless it is close to another special day, and its maximum value of 1 on February 8.

#Month-month visited #visitor type-Type of visitor;new, returning or other #Revenue-customer with revenue or without

#Checking for missing values

```
colSums(is.na(online_shoppers))
```

##	admin	admin_dur	info	info_dur	prod	prod_dur
##	14	14	14	14	14	14
##	bouncerrates	exitrates	pagevalues	specialday	month	ops_sys
##	14	14	0	0	0	0
##	browser	region	traffic_type	visitor_type	weekend	revenue
##	0	0	0	0	0	0

#We note that our dataset has missing values in the specific columns as per the summary. Most of the columns have 14 missing records.

```
head(online_shoppers)
```

```
##      admin admin_dur info info_dur prod      prod_dur bouncerrates exitrates
## 1:      0          0   0          0   1      0.000000  0.20000000 0.2000000
## 2:      0          0   0          0   2     64.000000  0.00000000 0.1000000
## 3:      0         -1   0         -1   1     -1.000000  0.20000000 0.2000000
## 4:      0          0   0          0   2      2.666667  0.05000000 0.1400000
## 5:      0          0   0          0  10    627.500000  0.02000000 0.0500000
## 6:      0          0   0          0  19   154.216667  0.01578947 0.0245614
##      pagevalues specialday month ops_systems browser region traffic_type
## 1:          0          0   Feb              1        1        1          1
## 2:          0          0   Feb              2        2        1          2
## 3:          0          0   Feb              4        1        9          3
## 4:          0          0   Feb              3        2        2          4
## 5:          0          0   Feb              3        3        1          4
## 6:          0          0   Feb              2        2        1          3
##      visitor_type weekend revenue
## 1: Returning_Visitor FALSE FALSE
## 2: Returning_Visitor FALSE FALSE
## 3: Returning_Visitor FALSE FALSE
## 4: Returning_Visitor FALSE FALSE
## 5: Returning_Visitor TRUE  FALSE
## 6: Returning_Visitor FALSE FALSE
```

#Dealing with missing values #Since the number of records are not too many most of the columns have 14 out of the 12,330 records, we will opt to omit them.

```
clean_online_shoppers <- na.omit(online_shoppers)
colSums(is.na(clean_online_shoppers))

##      admin      admin_dur      info      info_dur      prod      prod_
dur
##          0          0          0          0          0
0
## bouncerrates      exitrates      pagevalues      specialday      month ops_syst
ems
##          0          0          0          0          0
0
##      browser      region traffic_type visitor_type      weekend      reve
nue
##          0          0          0          0          0
0
```

#Changing the variables to the right datatypes

```
clean_online_shoppers$month<-as.factor(clean_online_shoppers$month)
clean_online_shoppers$ops_systems <-as.factor(clean_online_shoppers$ops_syste
ms)
clean_online_shoppers$browser<-as.factor(clean_online_shoppers$browser)
clean_online_shoppers$region<-as.factor(clean_online_shoppers$region)
clean_online_shoppers$traffic_type<-as.factor(clean_online_shoppers$traffic_t
ype)
clean_online_shoppers$visitor_type<-as.factor(clean_online_shoppers$visitor_t
```

```

type)
clean_online_shoppers$weekend<-as.factor(clean_online_shoppers$weekend)
clean_online_shoppers$revenue<-as.factor(clean_online_shoppers$revenue)
str(clean_online_shoppers)

## Classes 'data.table' and 'data.frame': 12316 obs. of 18 variables:
## $ admin      : int  0 0 0 0 0 0 0 1 0 0 ...
## $ admin_dur   : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ info        : int  0 0 0 0 0 0 0 0 0 0 ...
## $ info_dur    : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ prod        : int  1 2 1 2 10 19 1 1 2 3 ...
## $ prod_dur    : num  0 64 -1 2.67 627.5 ...
## $ bouncerrates : num  0.2 0 0.2 0.05 0.02 ...
## $ exitrates   : num  0.2 0.1 0.2 0.14 0.05 ...
## $ pagevalues  : num  0 0 0 0 0 0 0 0 0 0 ...
## $ specialday  : num  0 0 0 0 0 0 0.4 0 0.8 0.4 ...
## $ month       : Factor w/ 10 levels "Aug","Dec","Feb",...: 3 3 3 3 3 3 3 3 3 3
## $ ops_systems : Factor w/ 8 levels "1","2","3","4",...: 1 2 4 3 3 2 2 1 2 2
## $ browser     : Factor w/ 13 levels "1","2","3","4",...: 1 2 1 2 3 2 4 2 2 4
## $ region      : Factor w/ 9 levels "1","2","3","4",...: 1 1 9 2 1 1 3 1 2 1
## $ traffic_type: Factor w/ 20 levels "1","2","3","4",...: 1 2 3 4 4 3 3 5 3 2
## $ visitor_type: Factor w/ 3 levels "New_Visitor",...: 3 3 3 3 3 3 3 3 3 3
## $ weekend      : Factor w/ 2 levels "FALSE","TRUE": 1 1 1 1 2 1 1 2 1 1
## $ revenue     : Factor w/ 2 levels "FALSE","TRUE": 1 1 1 1 1 1 1 1 1 1
## - attr(*, ".internal.selfref")=<externalptr>

```

#We note that Month variable has 10 levels, Operating system 8 levels, Browser 13 levels, Region 9 levels, Traffic Type 20 levels, visitor type 3 levels, Weekend and revenue variables have 2 levels each.

#Checking for duplicates

```

duplicated_rows <- clean_online_shoppers[duplicated(clean_online_shoppers),]
duplicated_rows

##      admin admin_dur info info_dur prod prod_dur bouncerrates exitrates
## 1:      0          0    0          0    1          0          0.2        0.2
## 2:      0          0    0          0    1          0          0.2        0.2
## 3:      0          0    0          0    1          0          0.2        0.2
## 4:      0          0    0          0    1          0          0.2        0.2
## 5:      0          0    0          0    1          0          0.2        0.2
## ---
## 113:    0          0    0          0    1          0          0.2        0.2

```



```
## 114:      0      0      0      0      1      0      0.2      0.2
## 115:      0      0      0      0      1      0      0.2      0.2
## 116:      0      0      0      0      1      0      0.2      0.2
## 117:      0      0      0      0      1      0      0.2      0.2
##      pagevalues specialday month ops_systems browser region traffic_type
## 1:      0      0      Feb      1      1      1      3
## 2:      0      0      Feb      3      2      3      3
## 3:      0      0      Mar      1      1      1      1
## 4:      0      0      Mar      2      2      4      1
## 5:      0      0      Mar      3      2      3      1
## ---
## 113:      0      0      Dec      1      1      1      2
## 114:      0      0      Dec      1      1      4      1
## 115:      0      0      Dec      1      1      1      3
## 116:      0      0      Dec      1     13      9     20
## 117:      0      0      Dec      8     13      9     20
##      visitor_type weekend revenue
## 1: Returning_Visitor FALSE FALSE
## 2: Returning_Visitor FALSE FALSE
## 3: Returning_Visitor TRUE  FALSE
## 4: Returning_Visitor FALSE FALSE
## 5: Returning_Visitor FALSE FALSE
## ---
## 113:      New_Visitor FALSE FALSE
## 114: Returning_Visitor TRUE  FALSE
## 115: Returning_Visitor FALSE FALSE
## 116: Returning_Visitor FALSE FALSE
## 117:      Other    FALSE FALSE
```

#We note that our dataset has no duplicates

4)Univariate, Bivariate Analysis

#measures of central tendency and dispersion

```
library(psych)
data(clean_online_shoppers)

## Warning in data(clean_online_shoppers): data set 'clean_online_shoppers' not
## found

describe(clean_online_shoppers)

##      vars      n    mean      sd median trimmed      mad min      ma
## x
## admin      1 12316    2.32    3.32    1.00    1.63    1.48    0    27.0
## 0
## admin_dur  2 12316   80.91   176.86    8.00   42.19   11.86   -1  3398.7
## 5
## info      3 12316    0.50    1.27    0.00    0.18    0.00    0    24.0
```

```

0
## info_dur      4 12316   34.51  140.83   0.00    3.60   0.00  -1 2549.3
8
## prod          5 12316   31.76   44.49  18.00   22.78  19.27   0  705.0
0
## prod_dur      6 12316 1196.04 1914.37 599.77  821.41 743.05  -1 63973.5
2
## bouncerrates  7 12316    0.02    0.05   0.00    0.01   0.00   0   0.2
0
## exitrates     8 12316    0.04    0.05   0.03    0.03   0.02   0   0.2
0
## pagevalues    9 12316    5.90    18.58   0.00    1.30   0.00   0  361.7
6
## specialday   10 12316    0.06    0.20   0.00    0.00   0.00   0   1.0
0
## month*       11 12316    6.16    2.37   7.00    6.35   1.48   1  10.0
0
## ops_systems* 12 12316    2.12    0.91   2.00    2.06   0.00   1   8.0
0
## browser*     13 12316    2.36    1.72   2.00    2.00   0.00   1  13.0
0
## region*      14 12316    3.15    2.40   3.00    2.79   2.97   1   9.0
0
## traffic_type* 15 12316    4.07    4.02   2.00    3.22   1.48   1  20.0
0
## visitor_type* 16 12316    2.72    0.69   3.00    2.90   0.00   1   3.0
0
## weekend*      17 12316    1.23    0.42   1.00    1.17   0.00   1   2.0
0
## revenue*     18 12316    1.15    0.36   1.00    1.07   0.00   1   2.0
0
##
##      range  skew kurtosis    se
## admin      27.00  1.96    4.69  0.03
## admin_dur 3399.75  5.61   50.48  1.59
## info       24.00  4.03   26.89  0.01
## info_dur  2550.38  7.57   76.18  1.27
## prod       705.00  4.34   31.17  0.40
## prod_dur  63974.52  7.26  137.03 17.25
## bouncerrates  0.20  2.95    7.75  0.00
## exitrates   0.20  2.15    4.04  0.00
## pagevalues  361.76  6.38   65.53  0.17
## specialday   1.00  3.30    9.89  0.00
## month*      9.00 -0.83   -0.37  0.02
## ops_systems* 7.00  2.07   10.45  0.01
## browser*    12.00  3.24   12.73  0.02
## region*     8.00  0.98   -0.15  0.02
## traffic_type* 19.00  1.96    3.48  0.04
## visitor_type* 2.00 -2.06    2.28  0.01
## weekend*     1.00  1.27   -0.40  0.00
## revenue*    1.00  1.91    1.64  0.00

```

#The mean, standard deviation, median, min, max, range, skew, kurtosis of the numeric variables are as tabulated. We note that product duration and admin duration have very high standard deviation, meaning the datapoints vary greatly.

#plotting boxplots for all the numerical variables

```
par(mfrow=c(2,2))
boxplot((clean_online_shoppers$admin),horizontal=TRUE,col='light blue', main=
'boxplot of Administrative')
boxplot((clean_online_shoppers$admin_dur), horizontal=TRUE,col='light green',
main='boxplot of Administrative Duration')
boxplot((clean_online_shoppers$info),horizontal=TRUE,col='light blue', main='
boxplot of informational')
boxplot((clean_online_shoppers$info_dur), horizontal=TRUE,col='light green',
main='boxplot of Informational Duration')
```

boxplot of Administrative boxplot of Administrative Duration



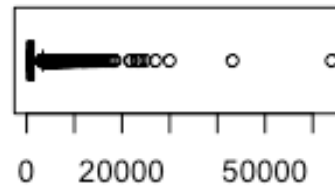
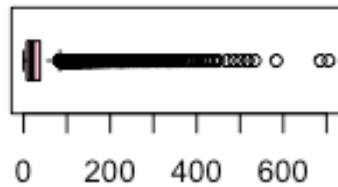
boxplot of informational boxplot of Informational Duration



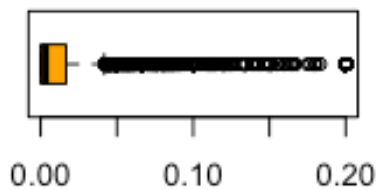
```
boxplot((clean_online_shoppers$prod),horizontal=TRUE, col='pink',main='boxplo
t of Product Related')
boxplot((clean_online_shoppers$prod_dur),horizontal=TRUE, col='black', main='
boxplot of Product Related Duration')
boxplot((clean_online_shoppers$bouncerates), horizontal=TRUE,col='orange', ma
in='boxplot of BounceRates')
```

```
boxplot((clean_online_shoppers$exitrates),horizontal=TRUE,col='light blue', main='boxplot of ExitRates')
```

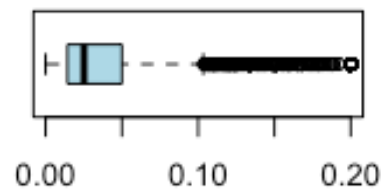
boxplot of Product Related



boxplot of BounceRates

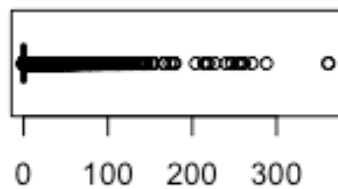


boxplot of ExitRates

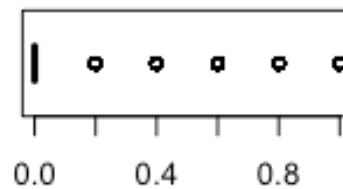


```
boxplot((clean_online_shoppers$pagevalues),horizontal=TRUE, col='pink',main='
boxplot of pagevalues')
boxplot((clean_online_shoppers$`specialday`), horizontal = TRUE, col = 'blue'
, main = "boxplot of SpecialDay")
#boxplot((clean_online_shoppers$`region`), horizontal = TRUE, col = 'purple',
main = "boxplot of region")
#boxplot((clean_online_shoppers$`browser`), horizontal = TRUE, col = 'red', m
ain = "boxplot of Browser")
#boxplot((clean_online_shoppers$`traffic_type`), horizontal = TRUE, col = 'ye
llow', main = "boxplot of TrafficType")
#boxplot((clean_online_shoppers$ops_systems), horizontal =TRUE,col= 'green',
main = 'boxplot of Operating System')
```

boxplot of pagevalues



boxplot of SpecialDay



#We observe that most of our data has outliers. We will opt not to remove them, since we are trying to understand the characteristics of customers who shop from Ecommerce, leaving them might help us to unearth the different patterns/clusters of the online shoppers.

#Checking the number of True and False values represented in revenue variable as this is our class label

```
revenue_table <- table(clean_online_shoppers$revenue)
revenue_table
```

```
##
## FALSE  TRUE
## 10408  1908
```

#We have 1,908 customers with revenue and 10,408 without revenue

#Finding correlation amongst the numeric variables #fetching all the numerical variables from the advertising dataset

```
admin<-clean_online_shoppers$admin
admin_dur<-clean_online_shoppers$admin_dur
info<-clean_online_shoppers$info
info_dur<-clean_online_shoppers$info_dur
prod<-clean_online_shoppers$prod
```

```

prod_dur<-clean_online_shoppers$prod_dur
bouncerrates<-clean_online_shoppers$bouncerrates
exitrates<-clean_online_shoppers$exitrates
pagevalues<-clean_online_shoppers$pagevalues
specialday<-clean_online_shoppers$specialday
#ops_systems<-clean_online_shoppers$ops_systems

```

#Creating a dataset with numeric variables

```

numeric_variables<- data.frame(admin, admin_dur, info, info_dur,prod,prod_dur
,bouncerrates,exitrates,pagevalues,specialday)

```

```

head(numeric_variables)    #previewing the dataframe

```

```

##  admin admin_dur info info_dur prod  prod_dur bouncerrates exitrates
## 1      0          0    0          0    1  0.000000  0.20000000 0.2000000
## 2      0          0    0          0    2 64.000000  0.00000000 0.1000000
## 3      0         -1    0         -1    1 -1.000000  0.20000000 0.2000000
## 4      0          0    0          0    2  2.666667  0.05000000 0.1400000
## 5      0          0    0          0   10 627.500000  0.02000000 0.0500000
## 6      0          0    0          0   19 154.216667  0.01578947 0.0245614
##  pagevalues specialday
## 1          0          0
## 2          0          0
## 3          0          0
## 4          0          0
## 5          0          0
## 6          0          0

```

#Previewing the correlation matrix

```

corr <- round(cor(numeric_variables), 1)
head(corr[, 1:10])    #previewing the matrix

```

```

##          admin admin_dur info info_dur prod prod_dur bouncerrates exitrate
s
## admin      1.0      0.6  0.4      0.3  0.4      0.4      -0.2      -0.
3
## admin_dur  0.6      1.0  0.3      0.2  0.3      0.4      -0.1      -0.
2
## info       0.4      0.3  1.0      0.6  0.4      0.4      -0.1      -0.
2
## info_dur   0.3      0.2  0.6      1.0  0.3      0.3      -0.1      -0.
1
## prod       0.4      0.3  0.4      0.3  1.0      0.9      -0.2      -0.
3
## prod_dur   0.4      0.4  0.4      0.3  0.9      1.0      -0.2      -0.
3
##          pagevalues specialday
## admin          0.1      -0.1
## admin_dur      0.1      -0.1
## info           0.0       0.0

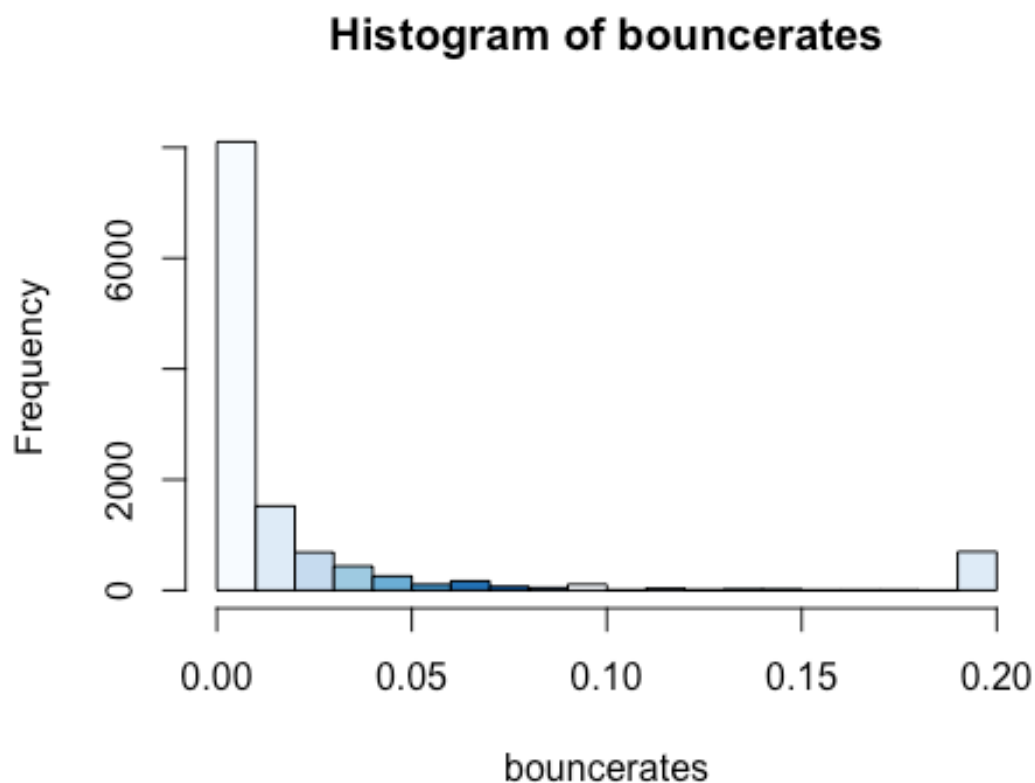
```

```
## info_dur      0.0      0.0
## prod         0.1      0.0
## prod_dur      0.1      0.0
```

#We observe that admin has a positive correlation of 0.6 with admin duration and 0.4 with Product duration. Info is positively correlated with info duration at 0.6, production and production duration have a positive correlation of 0.9

```
install.packages('ggplot2') library(ggplot2)
```

```
hist(clean_online_shoppers$bouncerrates, col=blues9,breaks=25,xlab="bouncerrates",main="Histogram of bouncerrates")
```

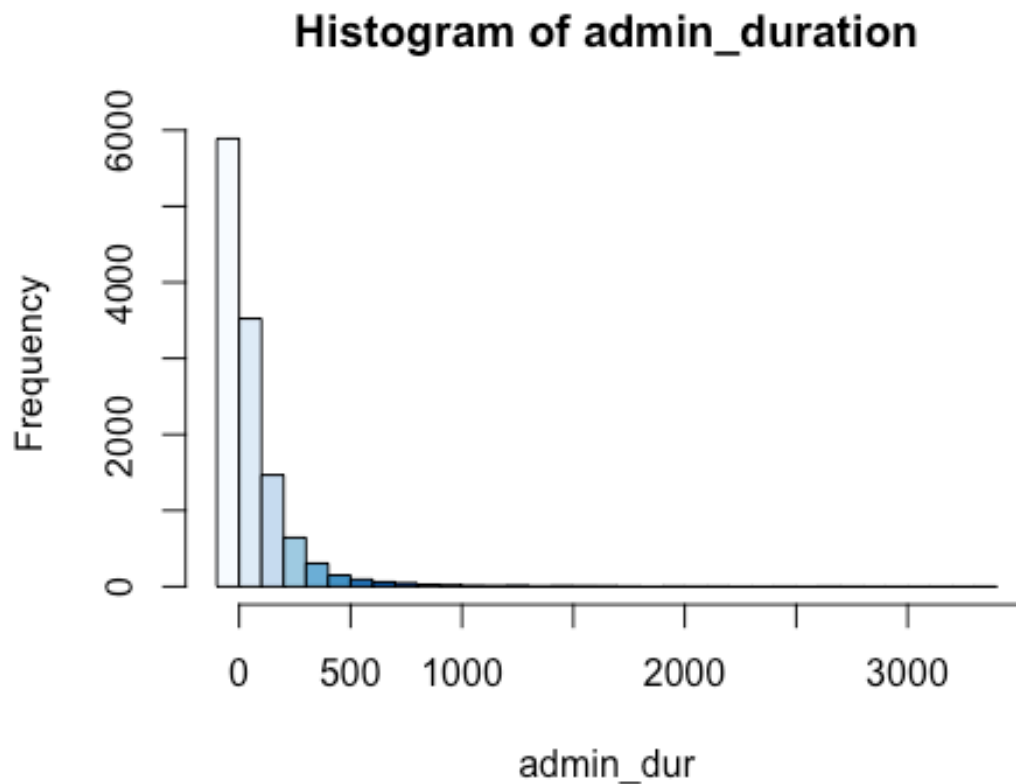


#We

observe that bounce rates is positively skewed to the right

#Histogram of admin duration

```
hist(clean_online_shoppers$admin_dur, col=blues9, breaks=25,xlab='admin_dur',main="Histogram of admin_duration")
```



#Admin

duration is positively skewed

```
install.packages('ggplot2') library(ggplot2)
```

```
install.packages('ggplot')
```

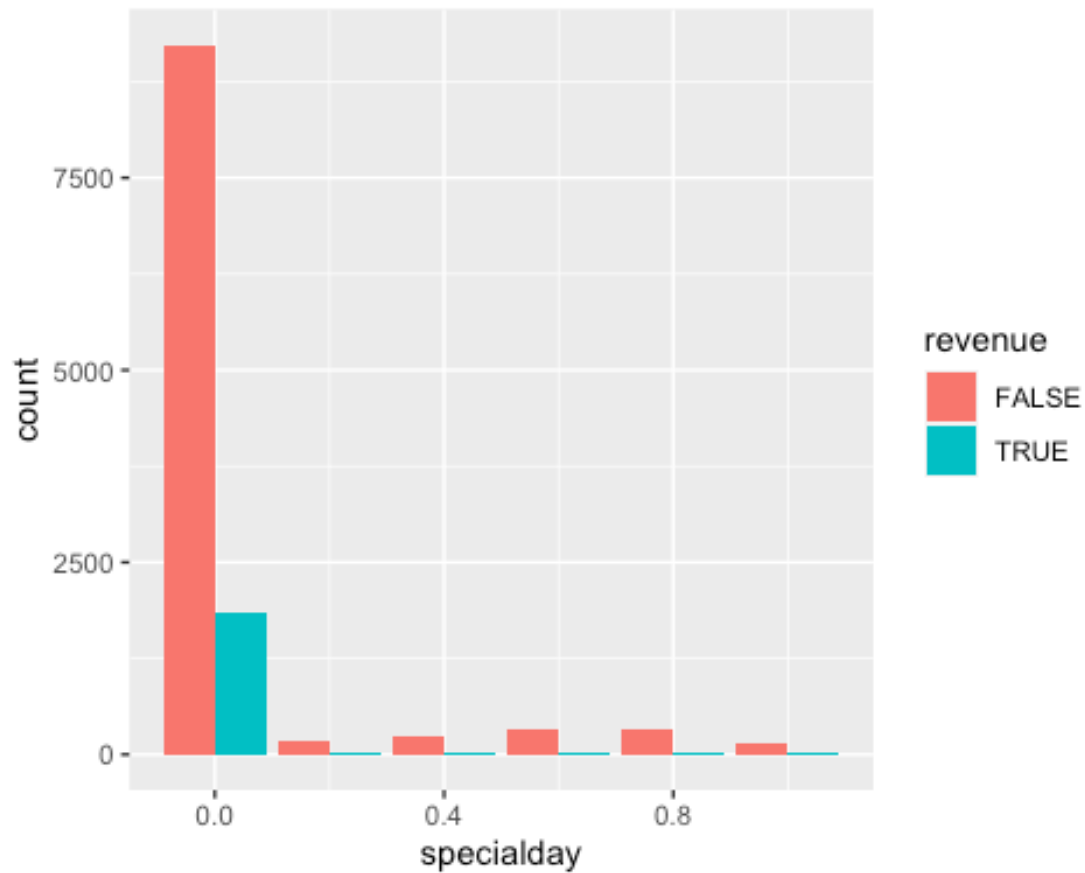
```
#Relationship between specialday and revenue
```

```
library(ggplot2)

##
## Attaching package: 'ggplot2'

## The following objects are masked from 'package:psych':
##
##   %+%, alpha

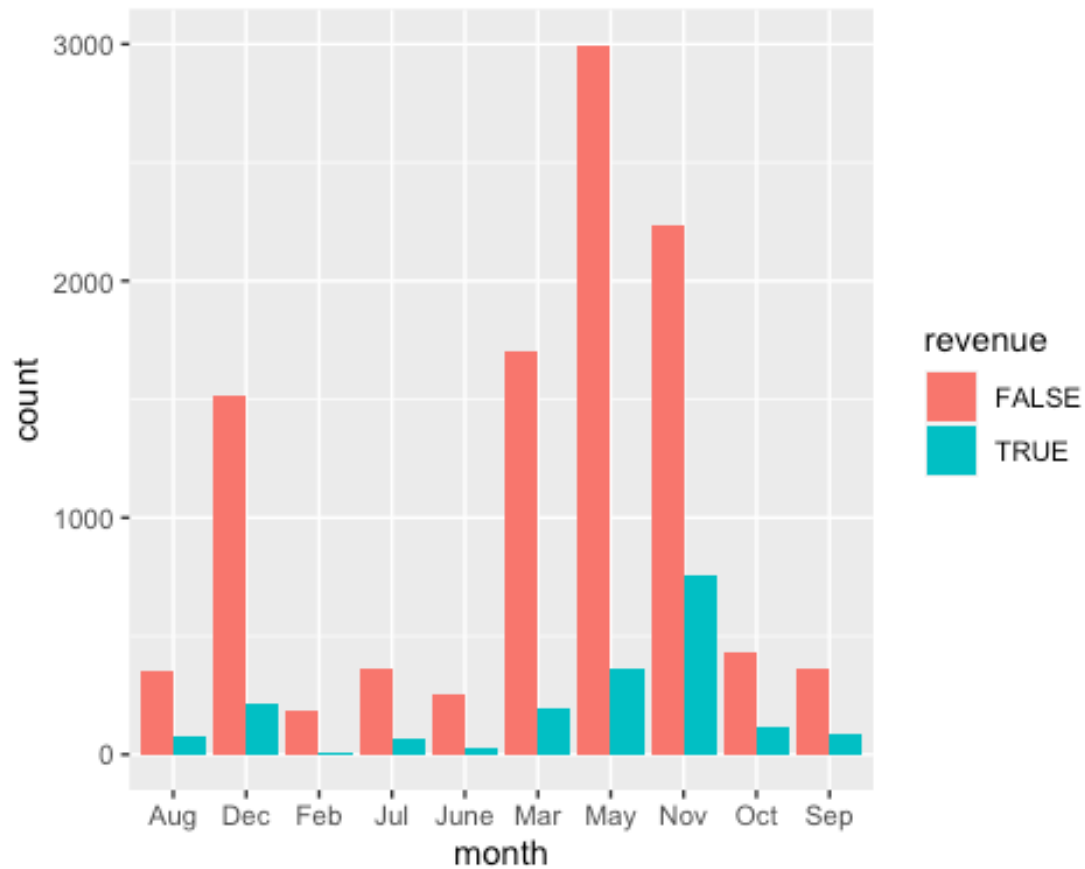
ggplot(data = clean_online_shoppers) +
  geom_bar(mapping = aes(x = specialday, fill = revenue), position = "dodge")
```

#We
note that specialday with zero values have higher revenue than those with non zero values.
#

#Checking the relationship between month and the class label revenue

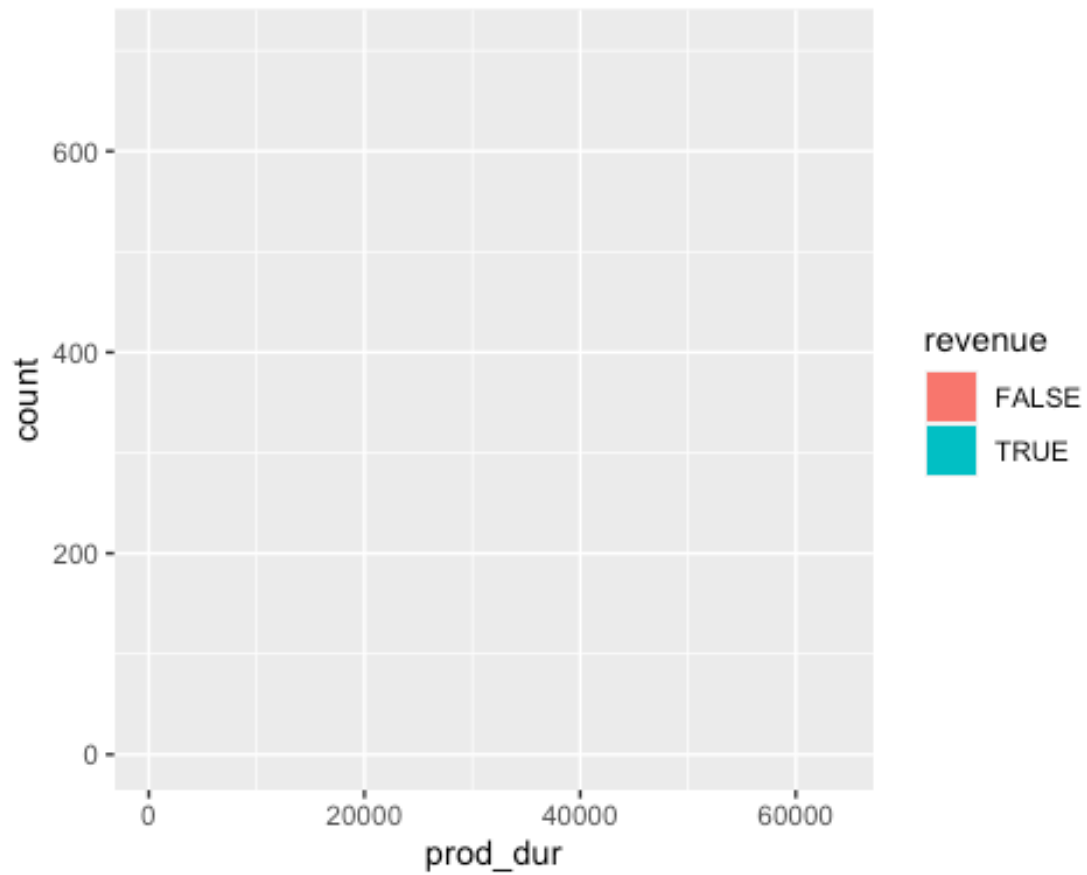
```
ggplot(data = clean_online_shoppers) +  
  geom_bar(mapping = aes(x = month, fill = revenue), position = "dodge")
```



#The month of November have the highest revenue, while the month of may registered highest non revenue.

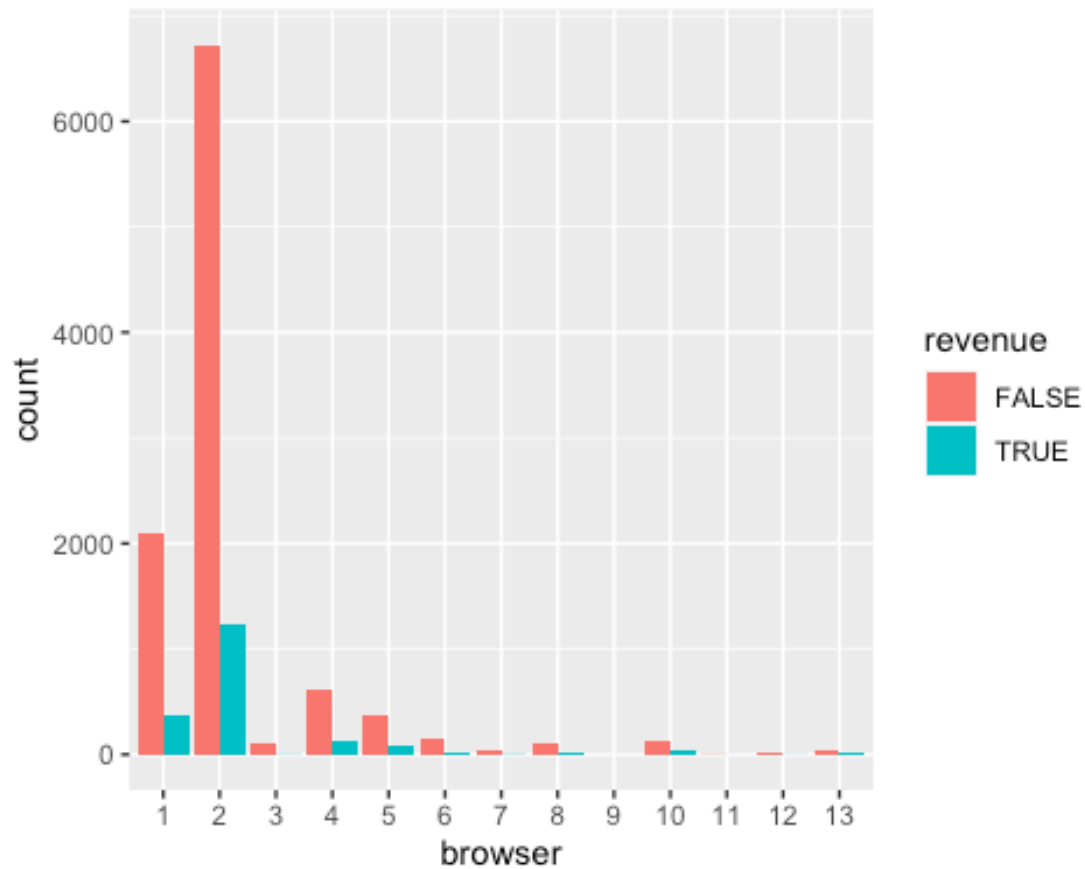
#Checking the relationship between product duration and the class label revenue

```
ggplot(data = clean_online_shoppers) +  
  geom_bar(mapping = aes(x = prod_dur, fill = revenue), position = "dodge")
```



#Checking the relationship between browser and the class label revenue

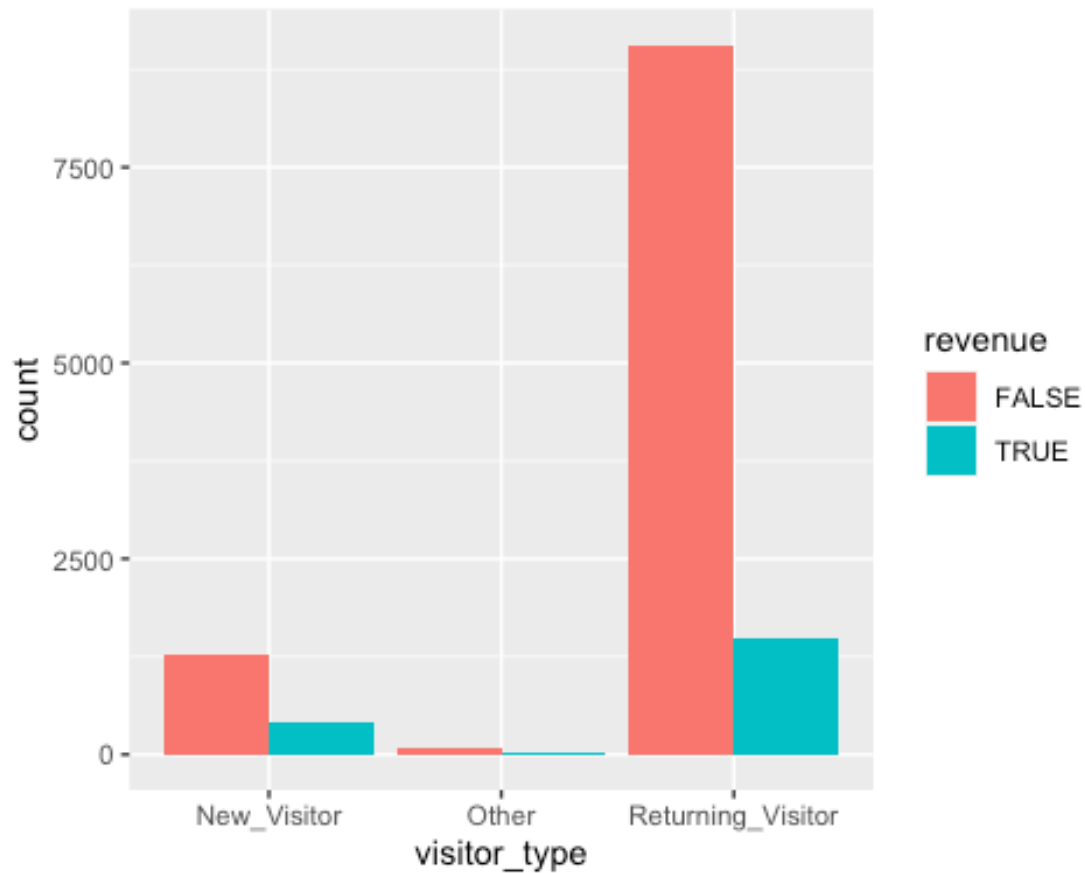
```
ggplot(data = clean_online_shoppers) +  
  geom_bar(mapping = aes(x = browser, fill = revenue), position = "dodge")
```



#We observe that browser values of 2 registered the highest non revenue and revenue status.

#Checking the relationship between visitor type and the class label revenue

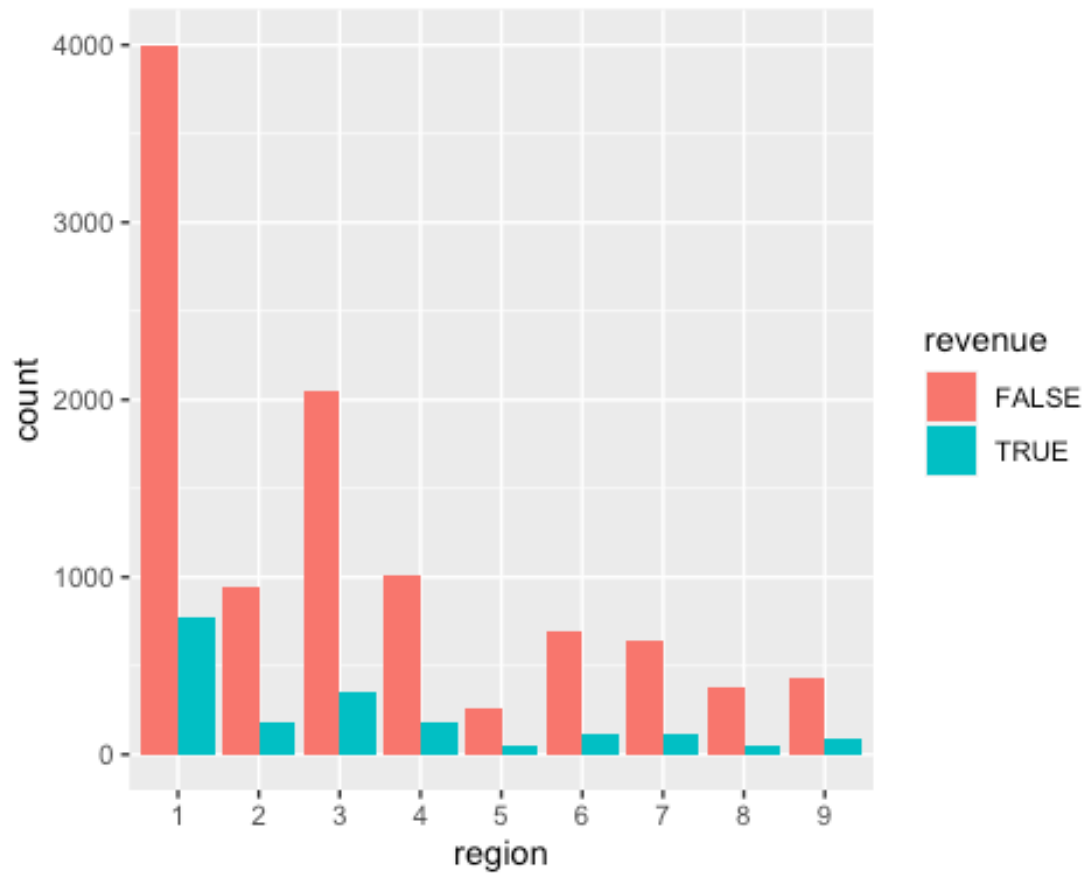
```
ggplot(data = clean_online_shoppers) +  
  geom_bar(mapping = aes(x = visitor_type, fill = revenue), position = "dodge"  
  )
```



#We observe that returning visitor type registered the highest non revenue and highest revenue status.

#Checking the relationship between region and the class label revenue

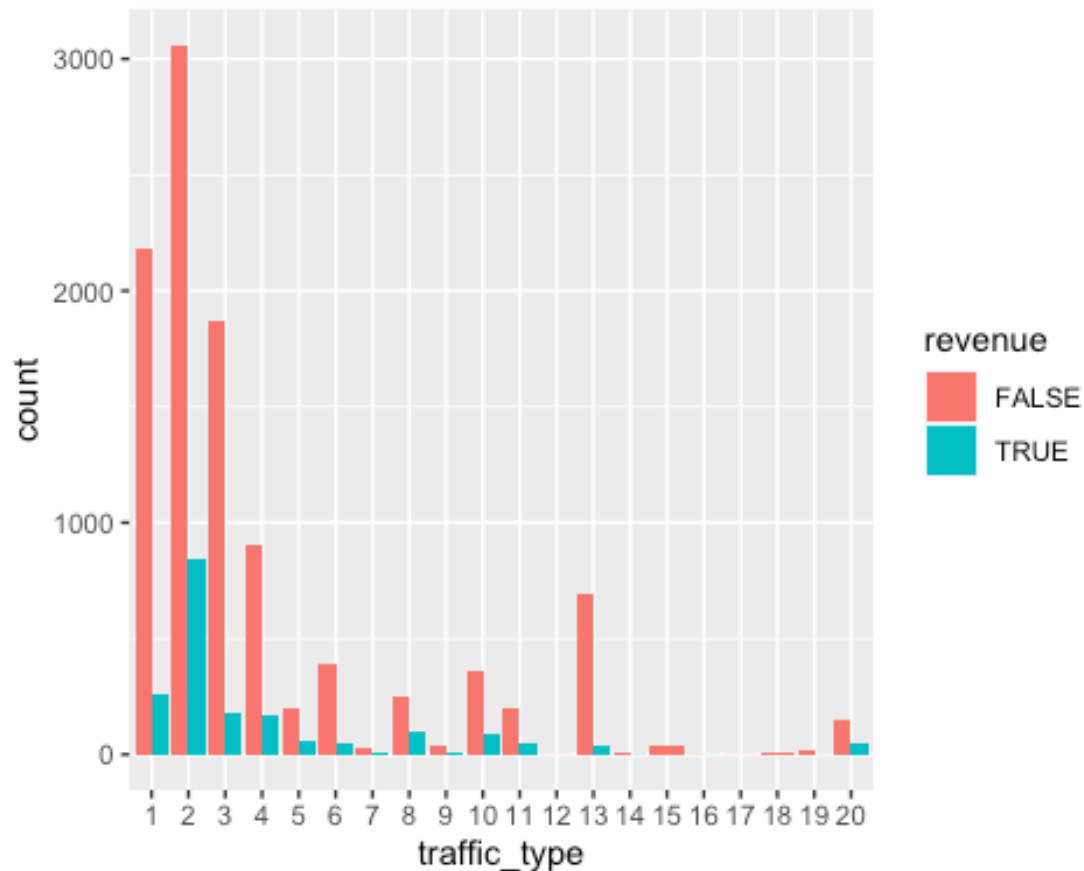
```
ggplot(data = clean_online_shoppers) +  
  geom_bar(mapping = aes(x = region, fill = revenue), position = "dodge")
```



#The region with value of 1 registered highest values of non revenue and revenue status.

#Checking the relationship between traffic type and the class label revenue

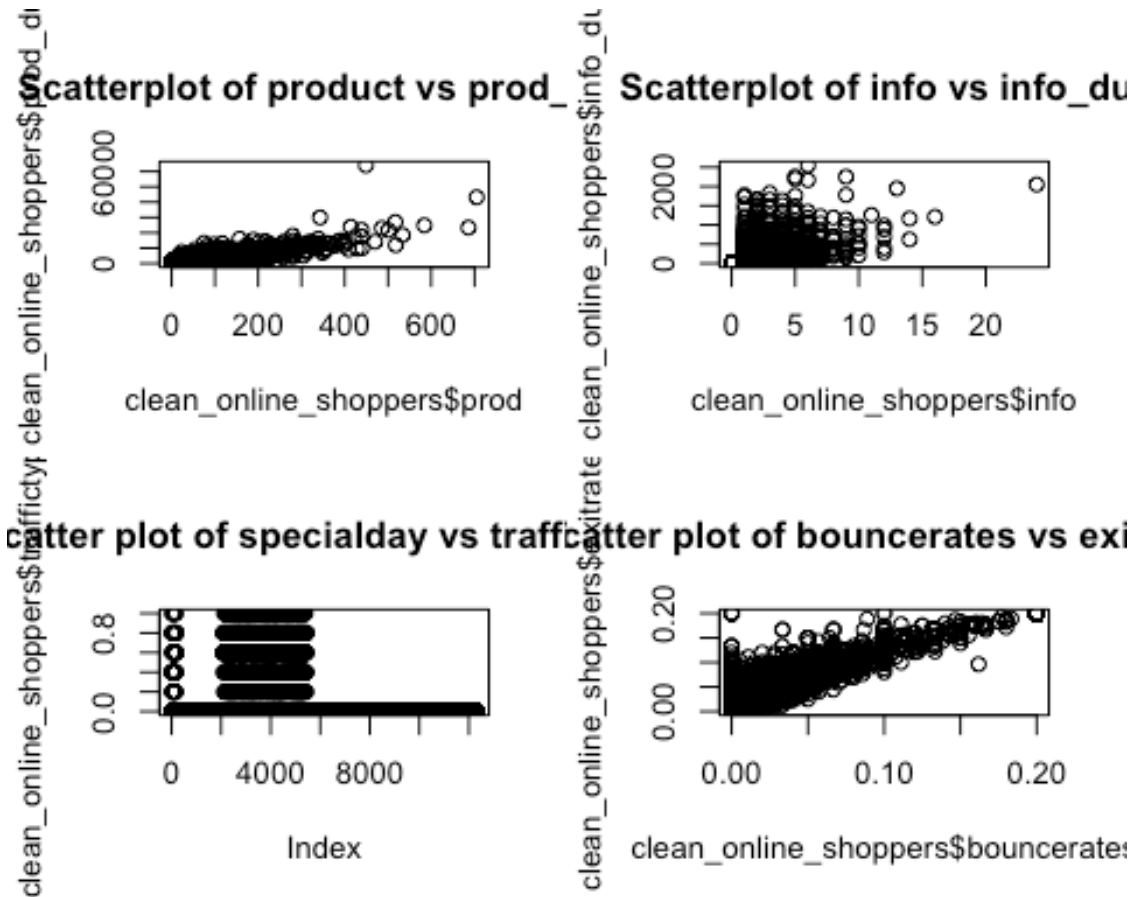
```
library(ggplot2)
ggplot(data = clean_online_shoppers) +
  geom_bar(mapping = aes(x = traffic_type, fill = revenue), position = "dodge")
```



#We observe the the traffic type with a value of 2 registered the highest non revenue status and highest revenue status.

#scatter plots to show the relationship of various variables

```
par(mfrow=c(2,2))
plot(clean_online_shoppers$prod,clean_online_shoppers$prod_dur, main="Scatter plot of product vs prod_dur")
plot(clean_online_shoppers$info,clean_online_shoppers$info_dur, main="Scatter plot of info vs info_dur")
plot(clean_online_shoppers$specialday, clean_online_shoppers$traffictype, main="Scatter plot of specialday vs traffictype")
plot(clean_online_shoppers$bouncerates, clean_online_shoppers$exitrates, main="Scatter plot of bouncerates vs exitrates")
```



#We observe a positive correlation between bounce rates and exit rates and product and product duration.

5) implementing unsupervised learning algorithm

#a) Unsupervised learning-with kmeans clustering K-means clustering is a clustering algorithm that is commonly used for partitioning a given data set into a set of k groups (i.e. k clusters), where k represents the number of groups pre-specified. The algorithm tries to find groups by minimizing the distance between the observations, called local optimal solutions. The distances are measured based on the coordinates of the observations

Advantages of kmeans

#Easy to implement #With a large number of variables, K-Means may be computationally faster than hierarchical clustering (if K is small). #K-Means may produce Higher clusters than hierarchical clustering #An instance can change cluster (move to another cluster) when the centroids are recomputed.

Disadvantages kmeans

#Difficult to predict the number of clusters (K-Value) #Initial seeds have a strong impact on the final results #Sensitive to scale: rescaling your datasets (normalization or standardization) will completely change results.

#checking stucture of the data

```
str(clean_online_shoppers)
```

```
## Classes 'data.table' and 'data.frame': 12316 obs. of 18 variables:
## $ admin      : int  0 0 0 0 0 0 0 1 0 0 ...
## $ admin_dur   : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ info        : int  0 0 0 0 0 0 0 0 0 0 ...
## $ info_dur    : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ prod        : int  1 2 1 2 10 19 1 1 2 3 ...
## $ prod_dur    : num  0 64 -1 2.67 627.5 ...
## $ bouncerrates : num  0.2 0 0.2 0.05 0.02 ...
## $ exitrates   : num  0.2 0.1 0.2 0.14 0.05 ...
## $ pagevalues  : num  0 0 0 0 0 0 0 0 0 0 ...
## $ specialday  : num  0 0 0 0 0 0 0.4 0 0.8 0.4 ...
## $ month       : Factor w/ 10 levels "Aug","Dec","Feb",...: 3 3 3 3 3 3 3 3 3 3
## $ ops_systems : Factor w/ 8 levels "1","2","3","4",...: 1 2 4 3 3 2 2 1 2 2
## $ browser     : Factor w/ 13 levels "1","2","3","4",...: 1 2 1 2 3 2 4 2 2 4
## $ region      : Factor w/ 9 levels "1","2","3","4",...: 1 1 9 2 1 1 3 1 2 1
## $ traffic_type: Factor w/ 20 levels "1","2","3","4",...: 1 2 3 4 4 3 3 5 3 2
## $ visitor_type: Factor w/ 3 levels "New_Visitor",...: 3 3 3 3 3 3 3 3 3 3
## $ weekend      : Factor w/ 2 levels "FALSE","TRUE": 1 1 1 1 2 1 1 2 1 1
## $ revenue     : Factor w/ 2 levels "FALSE","TRUE": 1 1 1 1 1 1 1 1 1 1
## - attr(*, ".internal.selfref")=<externalptr>
```

#changing the datatype of variables to numeric for ease of manipulation.

```
clean_online_shoppers$month<-as.numeric(clean_online_shoppers$month)
clean_online_shoppers$ops_systems <-as.numeric(clean_online_shoppers$ops_systems)
clean_online_shoppers$browser<-as.numeric(clean_online_shoppers$browser)
clean_online_shoppers$region<-as.numeric(clean_online_shoppers$region)
clean_online_shoppers$traffic_type<-as.numeric(clean_online_shoppers$traffic_type)
clean_online_shoppers$visitor_type<-as.numeric(clean_online_shoppers$visitor_type)
clean_online_shoppers$weekend<-as.numeric(clean_online_shoppers$weekend)
str(clean_online_shoppers)

## Classes 'data.table' and 'data.frame': 12316 obs. of 18 variables:
## $ admin      : int  0 0 0 0 0 0 0 1 0 0 ...
## $ admin_dur   : num  0 0 -1 0 0 0 -1 -1 0 0 ...
## $ info        : int  0 0 0 0 0 0 0 0 0 0 ...
## $ info_dur    : num  0 0 -1 0 0 0 -1 -1 0 0 ...
```

```
## $ prod      : int  1 2 1 2 10 19 1 1 2 3 ...
## $ prod_dur   : num  0 64 -1 2.67 627.5 ...
## $ bouncerrates : num  0.2 0 0.2 0.05 0.02 ...
## $ exitrates   : num  0.2 0.1 0.2 0.14 0.05 ...
## $ pagevalues  : num  0 0 0 0 0 0 0 0 0 0 ...
## $ specialday  : num  0 0 0 0 0 0 0.4 0 0.8 0.4 ...
## $ month       : num  3 3 3 3 3 3 3 3 3 3 ...
## $ ops_systems : num  1 2 4 3 3 2 2 1 2 2 ...
## $ browser     : num  1 2 1 2 3 2 4 2 2 4 ...
## $ region      : num  1 1 9 2 1 1 3 1 2 1 ...
## $ traffic_type: num  1 2 3 4 4 3 3 5 3 2 ...
## $ visitor_type: num  3 3 3 3 3 3 3 3 3 3 ...
## $ weekend      : num  1 1 1 1 2 1 1 2 1 1 ...
## $ revenue     : Factor w/ 2 levels "FALSE","TRUE": 1 1 1 1 1 1 1 1 1 1 ..
.
## - attr(*, ".internal.selfref")=<externalptr>
```

#We need to remove the class label from the dataset

```
new_data<-clean_online_shoppers[, -18]
new.class<-clean_online_shoppers[, "revenue"]
head(new_data)

##      admin admin_dur info info_dur prod  prod_dur bouncerrates exitrates
## 1:      0          0   0          0   1    0.000000  0.20000000 0.2000000
## 2:      0          0   0          0   2    64.000000  0.00000000 0.1000000
## 3:      0         -1   0         -1   1    -1.000000  0.20000000 0.2000000
## 4:      0          0   0          0   2     2.666667  0.05000000 0.1400000
## 5:      0          0   0          0  10   627.500000  0.02000000 0.0500000
## 6:      0          0   0          0  19  154.216667  0.01578947 0.0245614
##      pagevalues specialday month ops_systems browser region traffic_type
## 1:           0           0     3             1       1       1           1
## 2:           0           0     3             2       2       1           2
## 3:           0           0     3             4       1       9           3
## 4:           0           0     3             3       2       2           4
## 5:           0           0     3             3       3       1           4
## 6:           0           0     3             2       2       1           3
##      visitor_type weekend
## 1:              3       1
## 2:              3       1
## 3:              3       1
## 4:              3       1
## 5:              3       2
## 6:              3       1
```

#normalize the dataset so that all the variables are on the same scale

```
normalize<- function(x) {
  return((x-min(x)) / (max(x)-min(x)))
}
```

```
#normalizing specific variables and printing normalized data
```

```
new_data$admni <-normalize(new_data$admin)
new_data$admni_dur <-normalize(new_data$admin_dur)
new_data$info <-normalize(new_data$info)
new_data$info_dur <-normalize(new_data$info_dur)
new_data$prod <-normalize(new_data$prod)
new_data$prod_dur <-normalize(new_data$prod_dur)
new_data$bouncerrates <-normalize(new_data$bouncerrates)
new_data$exitrates <-normalize(new_data$exitrates)
new_data$pagevalues<-normalize(new_data$pagevalues)
new_data$specialday<-normalize(new_data$specialday)
new_data$month<-normalize(new_data$month)
new_data$ops_systems<-normalize(new_data$ops_systems)
new_data$browser<-normalize(new_data$browser)
new_data$region<-normalize(new_data$region)
new_data$traffic_type<-normalize(new_data$traffic_type)
new_data$visitor_type<-normalize(new_data$visitor_type)
new_data$weekend<-normalize(new_data$weekend)
head(new_data)
```

```
##      admin admin_dur info      info_dur      prod      prod_dur bouncerrates
## 1:      0          0    0 0.0003920992 0.001418440 1.563122e-05 1.000000000
## 2:      0          0    0 0.0003920992 0.002836879 1.016029e-03 0.000000000
## 3:      0         -1    0 0.0000000000 0.001418440 0.000000e+00 1.000000000
## 4:      0          0    0 0.0003920992 0.002836879 5.731448e-05 0.250000000
## 5:      0          0    0 0.0003920992 0.014184397 9.824223e-03 0.100000000
## 6:      0          0    0 0.0003920992 0.026950355 2.426226e-03 0.07894737
##      exitrates pagevalues specialday      month ops_systems      browser region
## 1: 1.000000          0          0 0.2222222 0.0000000 0.00000000 0.000
## 2: 0.500000          0          0 0.2222222 0.1428571 0.08333333 0.000
## 3: 1.000000          0          0 0.2222222 0.4285714 0.00000000 1.000
## 4: 0.700000          0          0 0.2222222 0.2857143 0.08333333 0.125
## 5: 0.250000          0          0 0.2222222 0.2857143 0.16666667 0.000
## 6: 0.122807          0          0 0.2222222 0.1428571 0.08333333 0.000
##      traffic_type visitor_type weekend admni      admni_dur
## 1: 0.00000000          1          0    0 0.0002941393
## 2: 0.05263158          1          0    0 0.0002941393
## 3: 0.10526316          1          0    0 0.0000000000
## 4: 0.15789474          1          0    0 0.0002941393
## 5: 0.15789474          1          1    0 0.0002941393
## 6: 0.10526316          1          0    0 0.0002941393
```

```
#Applying the K-means clustering algorithm with no. of centroids(k)=7
```

```
result<- kmeans(new_data,7, nstart=50)
result
```

```
## K-means clustering with 7 clusters of sizes 12, 7943, 63, 2532, 167, 434,
1165
##
```

```

## Cluster means:
##      admin  admin_dur      info  info_dur      prod  prod_dur
## 1 10.333333 2338.31577 0.215277778 0.285345020 0.27907801 0.24103369
## 2  0.5521843   6.70507 0.009888161 0.006381637 0.03190659 0.01302984
## 3  8.7619048 1371.87706 0.089947090 0.056206336 0.10136215 0.05585223
## 4  4.2231438   97.04505 0.030098078 0.019050993 0.05735886 0.02281341
## 5  8.4491018  770.77139 0.059131737 0.039417553 0.09971546 0.04708384
## 6  7.6866359  432.00364 0.063844086 0.045435631 0.08744975 0.03984282
## 7  6.9047210  228.98505 0.049821173 0.033708397 0.07887742 0.03229279
##  bouncerates  exitrates  pagevalues  specialday      month  ops_systems
## 1  0.05342082 0.12450145 0.009554398 0.00000000 0.4907407   0.1666667
## 2  0.15332465 0.27548063 0.012236106 0.07934030 0.5663687   0.1623532
## 3  0.02952891 0.11163893 0.029945841 0.01587302 0.6102293   0.1496599
## 4  0.03257716 0.10705784 0.023219504 0.02922591 0.5807881   0.1560032
## 5  0.03479814 0.09755568 0.025859348 0.03113772 0.6180971   0.1608212
## 6  0.03864945 0.10508635 0.023027308 0.04055300 0.5888377   0.1593153
## 7  0.03323520 0.10167956 0.024400774 0.02523605 0.5960897   0.1595340
##   browser      region  traffic_type  visitor_type  weekend      admni
## 1 0.08333333 0.1041667   0.1578947   0.9166667 0.2500000 0.38271605
## 2 0.11637081 0.2633766   0.1664557   0.8862520 0.2206975 0.02045127
## 3 0.13227513 0.2023810   0.1478697   0.8650794 0.2063492 0.32451499
## 4 0.10907056 0.2821880   0.1489149   0.7782385 0.2547393 0.15641273
## 5 0.09780439 0.2881737   0.1525370   0.9011976 0.2215569 0.31292970
## 6 0.11443932 0.2724654   0.1574097   0.8732719 0.2465438 0.28469022
## 7 0.10085837 0.2746781   0.1597470   0.8364807 0.2635193 0.25573041
##   admni_dur
## 1 0.688084644
## 2 0.002266364
## 3 0.403817063
## 4 0.028838901
## 5 0.227008276
## 6 0.127363376
## 7 0.067647635
##
## Clustering vector:
##      [1] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2
##      [37] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 7 2 2 4 2 2 2
##      [73] 2 2 2 2 5 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 4
##     [109] 2 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 6 2 2 2 2 2 4 2 2 2 4 2
##     [145] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2
##     [181] 2 2 2 2 2 2 2 2 7 4 2 2 3 4 2 2 2 4 4 2 4 2 2 7 2 4 7 2 2 4 2 2 2 2
##     [217] 4 2 2 2 2 2 2 4 5 2 4 2 2 2 2 5 2 4 4 2 4 4 2 2 4 2 2 2 2 2 2 2 4
##     [253] 2 2 2 4 2 3 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 3 2 6

```

2 2 2
[289] 7 2 2 2 2 2 2 2 2 7 2 2 7 2 2 2 2 2 2 2 4 2 2 2 2 7 4 2 6 4 2 2
2 2 4
[325] 5 4 4 2 4 4 2 2 2 4 2 2 2 2 2 2 2 2 2 2 7 2 7 4 2 4 2 4 4 4 2 2
2 2 2
[361] 2 2 2 4 2 5 2 2 2 2 7 2 2 2 2 2 4 4 2 4 2 2 2 2 7 2 2 2 2 2 4 2 2
7 2 4
[397] 7 2 2 7 2 2 7 4 7 2 5 2 7 4 2 7 7 7 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2
2 4 2
[433] 2 4 2 2 2 2 2 2 2 2 2 4 2 7 7 4 4 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2
4 7 4
[469] 2 2 2 2 2 2 2 4 4 5 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 6 2 4 3 2
2 7 4
[505] 2 4 2 2 2 2 6 6 2 6 2 2 2 4 2 4 2 4 2 2 4 4 2 2 2 2 2 2 2 2 2 2 2
2 2 2
[541] 2 2 2 2 4 4 2 4 4 2 2 2 4 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2
2 4 7
[577] 2 2 2 4 2 2 2 2 4 2 4 2 2 2 2 2 2 7 2 4 2 2 2 2 4 4 4 4 2 2 4 4 2
2 7 7
[613] 6 4 4 2 7 2 7 6 4 4 2 2 2 4 2 2 2 2 7 2 7 2 2 7 2 2 2 4 4 5 2 2 2
2 2 2
[649] 2 2 2 2 2 2 2 4 2 2 2 2 2 4 4 2 2 4 2 4 4 2 2 2 2 2 6 2 2 2 2 2 2
4 2 2
[685] 2 2 2 2 2 2 2 2 4 2 4 4 4 2 6 2 2 7 4 2 6 2 7 4 4 2 2 2 2 2 2 2 7
4 2 7
[721] 2 2 6 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 4 2 7 2 2 2 2 2 4 2 2 2 2 4 7
4 2 7
[757] 2 2 2 2 7 2 2 2 2 2 2 2 2 4 5 2 2 6 4 2 2 2 7 2 2 2 2 4 2 2 2 2 4
2 2 2
[793] 4 7 4 2 4 2 2 7 2 2 4 2 2 2 2 2 4 4 7 2 2 2 2 4 6 4 2 2 2 2 2 2 2
2 7 2
[829] 7 2 4 2 2 2 4 7 4 4 4 2 4 2 4 2 4 4 4 2 7 4 2 7 2 6 4 4 2 2 2 2 2
2 2 6
[865] 2 2 4 2 2 7 2 7 2 7 4 2 2 7 2 2 2 2 7 4 5 4 2 4 2 2 2 4 2 2 7 2 2
2 2 2
[901] 2 2 2 2 2 4 5 2 2 4 6 2 4 2 2 2 7 2 7 7 2 4 2 4 7 4 2 2 2 2 2 3 4
2 4 2
[937] 2 2 2 7 4 6 2 2 7 2 4 2 3 2 2 2 2 2 2 4 2 2 2 2 2 2 4 2 2 4 7 2 2
2 7 4
[973] 7 2 2 4 2 2 2 7 2 4 2 2 2 2 7 2 2 2 2 4 2 2 2 2 7 2 2 2 2 6 4 7 2
2 2 2
[1009] 2 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 4 2 2 2 2 2 2 2 2 2 2
4 4 7
[1045] 2 2 2 2 2 4 2 7 2 2 7 2 2 2 7 2 2 2 2 4 2 2 2 4 2 2 2 7 2 2 2 2
2 2 2
[1081] 2 2 2 4 2 2 2 2 4 2 2 2 2 2 2 2 7 2 2 2 2 2 7 2 2 5 4 2 4 2 2 7 2 6
2 2 2
[1117] 4 2 2 2 2 4 2 2 4 2 7 2 2 2 7 4 2 7 2 6 2 2 4 4 7 4 2 2 2 2 7 2 2
2 2 4
[1153] 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 7 4 4 4 2 2 6 2 4

2 2 2
[1189] 2 4 2 2 2 2 2 2 7 2 2 2 2 5 4 2 4 4 2 2 2 2 2 2 5 2 2 4 2 2 2 2 2
2 2 2
[1225] 2 7 2 4 2 2 2 2 2 2 2 2 2 2 2 4 2 4 2 2 2 4 4 2 2 2 2 2 4 2 2 2 2
2 4 2
[1261] 2 6 2 2 2 2 6 5 2 2 5 3 2 2 4 7 2 4 4 2 2 2 2 2 2 2 2 2 2 2 2 2 2
2 2 7
[1297] 3 5 2 2 4 2 2 2 2 4 4 2 2 2 2 2 4 2 7 7 4 4 7 2 2 2 4 7 4 2 2 2 2
2 5 2
[1333] 2 2 2 2 2 4 2 2 3 2 2 2 4 3 2 7 4 2 2 7 2 7 2 4 5 2 2 2 2 2 4 4 2
2 4 2
[1369] 2 2 2 4 2 7 2 2 2 2 4 2 2 4 4 2 2 2 2 4 2 2 2 4 2 2 2 2 2 2 2 4 4
4 2 2
[1405] 4 2 2 2 4 2 4 7 4 2 4 2 2 4 2 7 2 4 2 2 2 4 2 4 2 4 2 2 2 4 7 2 4
2 2 2
[1441] 2 2 7 4 2 2 2 2 2 2 2 2 2 2 2 4 2 3 7 2 2 2 4 2 4 4 2 2 2 4 2 2 4
2 7 4
[1477] 2 4 2 4 4 7 6 2 2 6 4 7 2 2 4 2 2 2 2 2 2 2 2 2 5 4 4 6 2 2 2 2 4 2
6 2 2
[1513] 2 2 6 2 7 2 2 6 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 7 4 2 4 2 2 2 4 2 4
2 7 2
[1549] 2 2 2 2 2 2 2 2 2 2 2 2 2 4 4 1 2 2 2 2 2 2 6 2 2 7 2 7 2 4 2 2 4 2
4 2 2
[1585] 2 2 2 2 2 2 2 2 6 2 2 3 2 2 2 2 2 4 2 2 2 2 2 4 2 2 4 2 2 7 2 2 4
2 7 7
[1621] 2 2 2 2 2 7 2 2 2 4 2 2 2 7 5 2 7 4 2 2 4 2 2 2 2 2 4 2 2 2 2 4 2
2 7 2
[1657] 7 2 2 7 2 2 2 2 2 4 2 7 3 2 2 6 2 2 4 2 2 2 2 2 4 2 2 2 4 2 2 2 2
4 2 2
[1693] 2 4 4 4 2 6 2 4 4 2 2 2 2 2 2 2 2 2 2 2 2 4 2 4 2 4 2 2 7 4 7 2
2 2 2
[1729] 2 7 2 2 2 2 2 2 2 7 2 2 2 4 2 4 2 4 2 2 2 2 2 6 4 2 2 2 2 2 2 2
2 2 4
[1765] 2 2 4 4 2 2 4 2 6 4 2 2 2 4 2 2 2 2 2 4 2 2 7 2 4 2 2 4 2 2 2 4 2
2 2 6
[1801] 4 2 4 2 2 2 4 2 2 2 2 2 3 4 2 2 7 4 2 2 2 2 2 4 2 2 2 5 2 2 2 2 2
2 2 7
[1837] 7 2 2 6 2 6 4 2 4 6 4 2 2 2 4 4 2 2 2 4 2 2 2 2 4 2 2 2 2 7 2 4
2 4 2
[1873] 2 2 2 4 2 2 2 2 2 2 2 2 2 7 2 2 2 2 7 6 2 4 4 4 2 2 2 2 2 4 2 7 6
2 2 2
[1909] 2 5 2 7 2 7 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 4 4 2 4 2 7 2 4 2 2 2 2
2 2 2
[1945] 2 2 2 7 6 2 4 2 2 4 2 2 2 2 2 2 4 2 4 4 2 6 2 2 2 7 2 2 4 2 4 2 2
2 5 7
[1981] 4 2 2 2 2 2 5 2 2 7 6 4 2 2 4 2 4 2 7 2 4 4 7 2 2 4 2 4 2 2 2 2 2
2 2 2
[2017] 2 4 2 2 7 7 4 4 2 2 7 7 6 2 2 4 5 2 2 2 2 2 4 2 2 2 2 2 2 4 2 2 2
2 4 7
[2053] 7 2 2 2 4 7 4 2 4 2 2 2 2 2 4 7 2 2 2 2 4 4 2 2 2 2 2 2 7 2 2 2 2

4 4 2
[2089] 2 2 6 2 2 2 2 2 7 4 6 6 2 2 4 2 2 2 2 2 2 2 2 4 4 2 4 2 2 2 7 4
2 2 2
[2125] 2 2 2 7 2 2 2 4 2 2 2 2 2 5 2 2 4 2 2 2 2 2 2 4 2 2 2 4 4 2 4 2 7
2 2 2
[2161] 7 7 2 2 2 6 2 2 4 2 2 4 4 2 2 5 2 2 2 2 4 2 2 2 2 2 4 7 4 6 2 2 2
2 2 6
[2197] 6 2 2 2 7 4 2 4 4 2 4 2 2 2 2 2 2 2 6 4 2 4 7 2 2 4 2 2 2 4 2 2 2
2 2 4
[2233] 2 7 7 2 2 2 2 2 7 4 7 6 4 2 2 2 2 4 2 2 2 2 2 2 2 4 2 2 2 4 4 2 7
7 2 2
[2269] 2 2 2 7 2 2 2 4 2 2 2 4 2 2 2 2 2 2 2 4 2 2 2 2 4 2 2 2 4 2 6 7 4
2 4 2
[2305] 2 2 4 2 2 2 2 2 2 2 2 7 2 2 2 2 2 2 4 4 4 4 2 2 2 4 2 2 2 4 4 4 4
2 2 2
[2341] 2 2 2 2 4 2 4 2 2 2 2 2 2 7 4 2 4 7 6 2 4 2 2 4 2 2 4 2 2 2 2 2 2 2
2 2 2
[2377] 7 4 6 2 2 2 4 2 7 2 6 6 2 2 2 2 2 4 2 2 2 2 4 2 2 7 6 2 2 2 4 2 4
2 2 2
[2413] 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2 4 3 2 2 2 2 4 2 2 6 2 2 7
2 4 2
[2449] 2 2 2 2 4 2 2 4 4 2 2 6 5 2 2 5 2 4 2 2 4 2 2 4 2 4 2 2 7 4 7 2 2
2 2 2
[2485] 6 2 2 2 4 2 4 2 4 2 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2 4 4 2 2 2 2
4 2 2
[2521] 2 2 4 2 4 5 4 2 2 7 2 2 2 2 2 2 2 2 2 2 4 2 4 2 7 4 7 6 7 2 2 7 4
6 2 4
[2557] 2 2 2 2 2 2 2 2 2 4 7 2 2 2 2 4 2 2 4 2 4 5 2 2 2 2 2 2 2 2 7 2 2
4 2 4
[2593] 2 6 2 4 2 2 2 2 2 4 2 6 6 2 2 2 2 2 7 2 7 4 2 2 2 2 7 4 2 4 2 7 7
2 4 2
[2629] 2 2 2 2 4 2 2 4 2 7 2 4 2 2 7 2 2 2 4 4 2 2 6 4 2 4 2 4 2 2 4 2 4
2 4 2
[2665] 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 6 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2 4 2
7 4 2
[2701] 2 2 2 2 4 2 2 2 2 2 2 2 7 6 2 4 2 2 4 2 4 2 4 2 2 7 2 2 4 2 2 2 4
4 2 2
[2737] 7 2 2 4 2 4 2 4 2 7 2 2 2 2 2 2 4 7 2 2 2 2 7 2 2 2 4 6 2 2 4 2 2 7
7 2 2
[2773] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2 2 2 2 2 2 2 2 2 2 2 4 2
2 2 7
[2809] 4 2 5 2 2 4 2 2 2 2 2 2 4 2 2 2 2 7 2 2 4 2 6 2 2 2 2 2 4 2 2 2 2
4 3 2
[2845] 2 7 4 2 4 4 2 4 4 2 2 2 2 2 2 2 2 2 2 7 2 2 7 2 4 4 2 2 4 2 2 4
4 2 2
[2881] 7 2 2 2 2 2 2 4 4 4 2 6 2 4 2 2 2 2 7 2 2 2 2 2 2 2 2 2 2 2 2 4 2
2 2 4
[2917] 4 2 2 7 2 2 2 2 7 6 2 6 4 2 7 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2
2 4 2
[2953] 2 7 7 7 2 2 4 2 7 2 2 2 2 2 2 2 2 4 2 2 2 2 4 2 2 2 2 7 4 2 7 7 7 2

2 2 4
[2989] 2 2 2 2 4 2 2 2 2 4 2 2 2 4 4 2 2 4 2 4 2 4 4 2 2 2 2 6 2 4 4 7 2
2 4 2
[3025] 4 7 2 2 2 2 2 4 2 2 2 7 2 2 2 2 4 2 2 2 2 4 2 2 2 2 2 2 4 2 2 7
2 2 4
[3061] 4 2 6 2 7 4 2 2 2 4 2 2 2 7 2 4 2 4 4 2 5 2 2 2 2 2 2 2 2 2 4 3 2
2 2 4
[3097] 2 2 2 2 2 2 2 2 4 4 2 4 6 2 2 4 2 2 2 7 2 2 2 2 2 2 2 2 2 4 4 2 2
5 2 2
[3133] 4 2 6 7 2 2 2 7 4 7 6 2 6 2 4 2 4 2 7 2 2 2 4 7 2 4 7 2 2 2 2 2 2
4 2 2
[3169] 2 2 4 3 4 2 2 2 2 6 2 2 2 2 4 4 2 2 4 2 7 2 7 4 4 4 2 2 2 2 2 7 2
7 2 4
[3205] 2 6 4 2 2 2 4 2 2 3 5 4 2 2 2 2 4 4 4 7 2 2 6 2 2 2 4 2 7 2 7 2 2
2 2 2
[3241] 2 2 2 2 2 2 6 2 2 2 2 2 4 2 5 2 2 4 2 2 2 2 2 4 2 2 7 2 2 2 2 2 4
2 2 4
[3277] 2 6 2 7 2 2 2 2 2 2 2 2 2 2 2 2 2 6 4 4 2 2 2 2 7 2 2 4 2 2 4 2
2 4 2
[3313] 2 2 4 7 2 2 2 2 4 2 7 2 5 2 2 7 4 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2
6 4 2
[3349] 2 2 4 6 2 2 2 4 7 2 7 2 4 7 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 7 2 2
2 2 4
[3385] 6 4 2 2 2 2 2 2 2 7 2 2 4 4 2 2 2 2 2 2 2 4 2 6 2 2 2 2 2 2 7 4
2 2 2
[3421] 2 2 2 2 2 4 2 4 2 4 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2
2 2 2
[3457] 2 4 2 2 4 2 2 2 4 2 2 2 4 2 2 2 4 2 2 4 2 7 2 7 2 2 2 2 2 4 7 2 2
4 2 6
[3493] 2 2 2 5 4 2 7 2 2 4 2 4 2 2 2 7 2 2 7 2 4 2 2 2 4 2 2 2 4 2 4 4 2
2 2 7
[3529] 7 2 4 2 2 2 7 2 2 6 2 2 2 2 2 2 2 2 2 4 4 2 4 2 2 2 2 2 2 6 2 2 2
2 2 2
[3565] 6 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 5 2 2 4 2 2 6 4 7 2 2 2 2 2 2
4 2 4
[3601] 2 4 2 2 4 7 2 2 2 2 2 2 2 2 7 2 6 2 2 4 2 2 2 2 4 2 2 2 7 2 4 2 4
2 2 2
[3637] 2 2 4 2 3 4 2 2 2 4 2 2 2 2 2 7 2 7 2 2 2 2 2 2 2 4 2 2 2 4 2 7 4 4
2 4 2
[3673] 4 2 2 7 2 4 4 2 2 4 2 4 2 2 2 2 2 7 2 2 4 4 4 2 2 2 3 7 2 4 4 2 4
4 2 2
[3709] 2 2 2 2 2 2 2 6 2 2 7 2 2 2 7 2 7 2 2 2 2 2 2 4 2 2 4 2 4 2 4 2 2
2 2 4
[3745] 2 5 2 2 2 2 7 4 2 2 2 2 2 2 4 2 2 2 2 2 2 4 4 6 7 2 2 2 2 2 2 2 4
2 2 2
[3781] 2 4 2 4 2 2 7 4 4 4 2 2 6 4 4 2 2 5 6 2 2 2 2 2 2 2 2 2 4 2 2 4 2
2 4 2
[3817] 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 4 2 2 7 4 2 2 2 4 2 2 2 2
7 2 4
[3853] 4 2 2 2 2 4 4 2 2 2 7 2 6 2 4 2 4 2 5 2 2 2 4 7 2 2 2 2 2 7 6 2 7

2 2 2
[3889] 2 2 2 7 2 2 2 2 2 2 6 2 4 2 4 2 2 2 4 2 4 2 2 2 2 2 2 2 2 2 4 2 2 2
4 7 4
[3925] 6 4 2 2 2 2 4 2 2 2 2 2 2 2 4 2 7 2 2 2 2 2 2 2 2 2 2 2 4 4 2 2 2
2 2 2
[3961] 2 2 7 2 2 2 2 4 2 2 2 2 2 2 6 2 4 4 2 2 2 4 2 4 4 4 2 2 2 2 4 2 2
2 2 4
[3997] 2 2 4 4 2 2 2 2 4 2 2 2 2 4 2 4 4 6 2 2 7 2 2 2 7 4 2 4 2 2 2 2 4
2 2 2
[4033] 2 2 2 6 2 2 4 2 2 2 4 2 2 5 4 2 7 2 2 7 2 4 4 2 4 2 7 2 2 4 4 2 2
2 2 4
[4069] 2 2 7 2 2 2 2 2 2 2 2 4 2 2 2 2 7 2 2 2 7 2 2 5 2 2 5 4 2 7 3 2 6
7 2 2
[4105] 4 2 5 2 2 2 2 2 2 2 2 2 4 7 4 7 6 2 2 4 2 2 7 2 2 2 2 2 7 2 2 2 4
4 2 4
[4141] 4 2 2 7 2 2 2 2 4 2 2 4 2 2 6 4 2 2 7 5 2 2 2 4 2 2 4 7 2 2 7 4 2
2 2 4
[4177] 2 4 2 3 2 4 2 2 2 6 2 2 2 2 2 2 6 4 2 2 2 2 2 7 2 2 2 2 2 2 4 2 4
2 2 2
[4213] 2 4 2 2 7 2 2 2 6 2 2 2 2 2 2 2 2 2 2 2 4 4 7 2 2 4 4 7 2 2 2 4 4
4 7 2
[4249] 4 5 2 2 4 2 4 2 2 2 7 7 2 5 2 2 6 4 2 2 4 2 2 5 2 2 2 2 2 2 2 2 2
7 2 2
[4285] 2 6 2 7 2 2 2 2 2 2 2 2 2 2 7 2 2 2 6 2 2 2 7 7 2 7 2 2 5 2 2 7 2
4 4 4
[4321] 2 2 4 2 2 2 2 7 2 2 2 2 2 4 7 2 2 2 7 6 2 2 2 4 4 4 2 2 2 2 2 2 2
2 4 2
[4357] 4 4 5 2 2 2 2 2 3 4 4 2 2 4 4 2 2 2 7 4 2 2 2 4 2 2 2 2 5 2 4 7 2
2 4 2
[4393] 4 2 7 2 2 2 7 2 2 2 2 6 2 4 2 4 2 7 4 2 2 2 4 2 4 4 6 4 4 2 2 4 4
4 2 4
[4429] 4 7 4 2 2 2 2 4 2 7 2 4 2 2 4 2 2 2 2 6 2 2 2 2 4 2 7 2 2 6 2 2
2 2 2
[4465] 2 4 2 2 2 2 7 7 2 2 7 2 7 2 2 5 2 2 5 2 2 2 4 4 4 2 2 7 7 2 2 1 2
4 2 4
[4501] 2 7 4 2 7 2 2 6 2 2 2 2 2 2 2 2 2 4 4 4 2 2 2 4 2 2 2 4 2 6 2 4 2 4
2 2 4
[4537] 2 2 2 2 2 2 5 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 4 5 2 7 2 7 2 4 2 2
2 2 7
[4573] 2 4 2 2 2 6 2 7 2 2 2 2 7 2 2 5 2 7 2 2 2 2 4 4 2 4 2 6 2 2 4 2 2
2 2 2
[4609] 4 4 2 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 4 5 4 2 3 2 2 2 7 2 2 2 2 2
7 2 2
[4645] 4 2 2 2 2 7 7 2 2 2 4 2 4 4 2 2 4 2 2 2 7 7 6 7 2 4 2 2 2 2 2 2 2
2 2 2
[4681] 2 2 7 7 2 2 7 2 4 2 2 2 2 4 4 2 4 2 2 2 2 2 4 2 2 2 2 2 2 2 5 2
2 4 2
[4717] 2 2 7 2 4 2 2 2 2 2 4 2 4 2 2 2 2 2 2 4 4 4 4 2 2 2 2 2 4 2 4 2 2
2 2 2
[4753] 2 2 2 4 4 2 2 2 4 2 2 2 2 2 2 2 4 2 2 4 2 4 2 4 2 2 4 2 7 2 4 2 4 4

2 2 2
[4789] 6 2 2 4 2 4 4 4 2 4 2 2 2 2 6 2 2 2 4 4 2 2 2 2 2 2 5 2 2 6 4 4 2
2 6 2
[4825] 2 2 2 4 2 2 4 7 4 2 4 2 7 2 2 2 7 2 2 4 2 2 2 2 2 2 2 2 2 2 2 5
5 4 2
[4861] 7 2 2 2 4 7 7 2 2 2 1 2 2 2 4 6 2 2 4 2 2 2 2 6 2 4 7 2 4 4 6 2 4
2 2 2
[4897] 2 2 2 2 2 2 2 4 4 2 4 2 4 2 2 2 2 4 7 2 2 2 4 4 2 2 6 7 2 2 4 2 2
2 2 5
[4933] 2 2 2 4 4 2 7 4 2 2 2 2 2 2 2 2 2 4 2 2 4 7 4 2 2 2 2 2 2 2 2 2
2 2 4
[4969] 2 2 2 2 4 6 2 2 4 2 4 4 6 2 2 2 7 2 2 2 2 2 2 3 7 7 2 7 5 2 2 2 2 2
2 2 2
[5005] 2 7 2 2 2 2 2 2 2 6 2 2 6 4 2 2 2 2 2 2 2 7 2 2 4 2 4 7 2 4 2 2 2
7 2 2
[5041] 2 4 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2 4 2 2 4 2 2 2 2 2 2 2 2 2 4 2
2 4 4
[5077] 7 5 2 2 4 2 2 2 2 7 4 2 6 2 2 4 2 2 2 2 2 7 2 2 7 2 4 2 2 2 2 2 2
7 2 2
[5113] 2 4 2 2 4 2 7 2 2 4 2 2 2 4 4 2 4 2 2 7 6 2 2 7 2 4 1 4 2 2 2 2 4
4 2 4
[5149] 2 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2 7 2 2 2 4 4 2 4 4 4 7 2 2 2 2 2 6
4 2 7
[5185] 2 2 4 2 2 5 2 2 2 7 7 2 2 2 2 2 4 4 4 7 4 4 2 2 7 2 2 7 4 2 2 2 2
2 2 2
[5221] 2 4 2 7 2 7 2 7 2 2 2 2 2 2 2 4 2 7 2 2 2 2 2 2 2 2 2 2 2 7 2 2 2
4 2 2
[5257] 2 2 4 2 2 7 2 4 2 2 4 7 2 2 2 2 2 2 2 7 2 2 2 2 2 4 2 4 2 7 2 2 2
2 2 2
[5293] 2 2 2 2 2 2 4 2 2 2 2 2 4 4 2 4 4 2 2 2 2 2 2 2 2 2 2 6 2 4 2 2 2
4 2 2
[5329] 2 4 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 7 2 2 2 4 4 2 4 4 4 2 7 2 2 7
2 2 2
[5365] 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 4 2 2 7 4 2 4 2 5 2 2 4 2 2 2 7 2 2
2 2 7
[5401] 2 2 2 4 2 4 2 2 2 2 2 7 2 2 2 2 2 2 2 7 4 2 2 2 4 6 2 2 4 4 2 7
7 2 4
[5437] 2 4 2 4 2 2 2 4 5 2 4 4 2 2 4 4 2 7 7 2 2 2 2 4 2 4 4 7 2 2 2 2 7
2 2 2
[5473] 4 2 3 2 2 2 4 4 2 4 2 5 2 4 7 2 2 2 2 2 4 4 2 6 4 4 2 2 2 4 2 6 2
5 4 2
[5509] 2 4 2 4 7 4 2 2 7 4 7 4 4 7 7 7 2 4 2 2 2 4 2 4 5 2 2 2 6 2 2 2 6
6 4 2
[5545] 2 2 2 5 2 2 2 4 2 2 2 6 2 5 4 2 2 7 2 4 2 2 2 2 2 4 2 2 2 2 2 2 2
2 7 2
[5581] 2 2 4 6 2 4 2 2 4 2 4 2 2 4 7 2 7 2 2 2 4 2 4 6 2 2 2 2 2 2 2 2 2
4 2 4
[5617] 4 4 6 4 4 4 2 5 2 2 4 2 4 4 4 4 2 2 2 2 6 2 7 2 2 4 7 2 2 4 2 2 2
2 2 4
[5653] 2 2 4 2 2 2 2 4 4 2 2 2 2 2 2 2 4 2 5 4 2 2 2 4 6 2 2 4 2 2 2 7 2

2 4 2
[5689] 4 4 2 2 2 7 4 7 4 7 2 4 2 4 7 2 4 4 2 2 4 4 2 2 2 4 7 2 7 2 4 2 2
7 2 2
[5725] 2 2 2 2 2 2 4 6 4 2 7 2 7 2 2 2 3 7 4 6 7 4 2 2 4 2 2 2 4 4 2 7 2
2 4 4
[5761] 2 7 1 4 4 4 2 4 4 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 2 4 7
2 2 7
[5797] 2 2 7 7 2 6 2 6 4 2 2 2 2 4 4 2 4 5 4 7 2 7 4 7 2 2 7 2 4 2 2 2 2
2 4 5
[5833] 2 2 2 2 4 2 2 4 4 4 2 4 2 2 2 6 4 2 4 2 2 2 2 2 2 2 2 2 4 2 2 6 2
4 2 2
[5869] 2 6 2 7 2 2 2 2 4 4 2 4 2 7 2 2 7 4 7 2 5 7 2 2 2 2 7 7 2 2 7 4 7
2 7 7
[5905] 7 2 2 4 4 2 2 2 7 2 2 4 2 2 2 2 4 4 2 6 4 5 4 2 3 2 2 2 4 2 7 2 2
7 2 4
[5941] 4 2 2 4 2 2 2 4 2 4 2 7 4 2 2 7 7 6 7 2 2 2 4 2 7 2 2 4 2 7 2 2 2
2 2 4
[5977] 6 2 2 6 4 2 2 2 6 2 2 2 4 4 4 5 4 6 2 7 6 4 4 2 2 2 4 2 4 2 4 2 7
2 2 4
[6013] 2 2 4 7 7 2 2 2 2 4 2 2 4 2 2 4 2 2 2 2 4 2 2 2 4 7 2 5 2 2 2 2 4
4 4 4
[6049] 2 4 2 4 3 5 2 2 2 4 5 2 4 2 4 2 4 2 4 2 4 4 2 4 2 2 2 4 4 2 7 2 4
2 4 2
[6085] 2 4 2 2 6 7 4 4 2 2 2 2 2 4 7 4 2 2 7 2 2 6 2 2 2 2 2 6 4 4 2 7 2
2 2 4
[6121] 2 2 7 2 2 7 2 2 2 4 4 2 7 2 2 4 2 7 4 2 2 2 7 2 2 4 4 6 7 2 2 1 4
4 2 2
[6157] 4 2 4 4 2 4 4 2 3 4 2 4 4 2 4 2 7 7 2 2 7 7 2 2 7 2 2 2 2 2 2 2 2
2 6 4
[6193] 7 4 4 4 2 2 4 5 5 2 2 4 4 2 6 4 4 4 2 7 2 2 7 7 2 2 2 2 4 4 2 2 7
4 4 7
[6229] 7 4 2 2 2 7 7 2 4 2 4 7 2 2 2 2 7 2 2 7 2 2 2 2 4 2 2 2 6 6 2 7 2
2 2 4
[6265] 2 2 2 4 2 2 2 2 4 2 4 2 2 2 2 6 2 4 7 7 7 7 4 4 2 4 7 4 4 2 2 4 4
4 2 4
[6301] 2 2 4 2 4 4 4 2 7 2 6 4 4 2 7 2 2 2 4 2 2 2 4 4 2 2 2 4 2 2 2 2 2
6 7 2
[6337] 4 2 2 4 4 4 2 2 7 2 4 2 4 2 2 2 2 6 2 2 7 7 2 2 2 2 2 2 2 4 4 4 2
2 4 2
[6373] 2 4 2 4 2 2 2 4 4 2 2 2 2 2 2 4 2 7 2 4 2 2 4 2 2 6 4 7 2 2 2 2 4
2 2 4
[6409] 4 7 2 6 4 2 4 4 2 2 2 2 2 4 2 2 4 2 2 2 2 4 2 4 4 4 4 2 5 2 7 4 6
4 4 4
[6445] 2 4 2 2 5 2 4 2 2 7 2 2 2 4 2 4 6 2 4 7 4 4 2 2 4 2 7 2 7 6 6 4 2
4 2 4
[6481] 4 7 2 2 2 4 7 2 7 2 4 2 4 2 2 2 4 4 4 5 2 2 2 4 4 2 2 4 4 2 2 2 2
2 4 2
[6517] 4 7 2 2 4 4 4 2 2 2 2 2 7 7 2 6 2 4 4 6 4 2 2 2 4 7 2 7 4 4 2 7 4
2 2 4
[6553] 4 4 7 2 2 4 4 4 2 2 2 2 2 4 2 2 7 7 4 2 2 2 2 2 4 7 7 7 2 2 7 2 7

2 2 5
[6589] 6 6 4 2 2 7 2 2 2 2 2 7 2 2 2 2 4 6 2 2 2 2 2 7 2 7 2 2 2 2 2 7 2
4 2 2
[6625] 4 7 7 4 2 2 7 3 2 5 2 7 2 2 2 2 4 2 4 2 4 2 2 7 2 4 2 4 7 2 2 2 4
7 4 2
[6661] 6 4 2 2 7 2 2 2 4 2 7 5 2 6 4 7 2 7 2 4 2 2 2 2 2 4 2 2 2 4 2 2 4
6 4 2
[6697] 7 7 2 7 4 4 2 2 2 6 2 2 4 2 4 4 4 6 2 4 2 2 2 2 2 2 2 7 2 4 4 4 4
2 4 2
[6733] 2 2 7 4 2 4 4 4 2 2 2 6 2 4 4 2 7 4 7 4 4 2 2 4 4 6 4 4 2 4 4 4 4
4 7 6
[6769] 4 2 2 2 4 4 7 7 2 7 4 7 2 2 2 4 4 7 4 2 6 4 4 4 2 2 2 7 2 6 2 2 2
4 4 7
[6805] 2 2 7 2 4 2 7 2 2 2 2 2 2 5 2 4 4 2 2 2 4 2 2 4 7 2 7 4 4 2 4 7 4
2 3 2
[6841] 4 4 2 2 2 2 4 4 4 4 2 2 2 2 7 3 7 7 2 2 2 2 7 7 2 2 2 2 7 7 7 4 4
2 4 2
[6877] 2 4 7 6 4 2 2 4 3 2 7 2 4 2 4 4 7 4 2 4 2 7 2 4 7 4 2 2 7 2 2 6 2
2 4 4
[6913] 2 4 2 2 4 2 4 4 4 2 4 4 4 6 2 6 7 2 2 2 2 2 7 2 4 4 2 2 2 2 5 4 6
6 2 4
[6949] 2 4 2 2 2 4 2 7 2 7 4 2 2 2 2 2 2 7 2 6 4 4 4 2 2 4 2 4 2 4 7 2 2
4 4 2
[6985] 6 4 2 2 2 4 2 2 4 6 2 4 2 6 2 2 2 4 2 2 4 2 4 2 7 2 2 2 4 2 2 2 4
2 2 2
[7021] 2 2 6 4 7 6 7 7 2 2 2 7 2 2 4 2 2 2 4 7 2 4 4 6 4 4 2 4 2 7 2 4 2
4 2 4
[7057] 2 2 7 7 4 2 2 2 2 6 7 4 7 2 2 2 4 6 4 4 4 7 4 2 6 2 6 2 2 2 2 4 2
7 2 2
[7093] 4 2 2 2 4 2 7 4 4 4 2 2 6 7 2 6 4 2 2 2 2 7 7 6 4 4 2 4 6 7 6 2 2
4 4 2
[7129] 7 7 7 7 7 7 4 4 2 2 4 2 4 2 4 2 2 2 5 2 4 2 7 2 2 6 7 7 2 2 2 4 6
4 7 2
[7165] 4 4 2 7 4 4 2 2 6 2 2 4 2 2 7 2 6 4 4 7 2 2 4 2 2 4 2 2 4 7 2 2 2
4 2 2
[7201] 2 2 2 2 2 7 4 7 2 2 7 7 7 7 2 2 2 2 2 7 2 7 4 2 2 2 6 2 2 2 2 7 4
2 4 4
[7237] 2 4 4 2 4 2 2 2 7 2 4 4 2 6 4 2 2 7 6 2 4 2 6 2 2 2 2 2 2 2 4 7 4
4 6 2
[7273] 2 2 2 4 7 6 4 6 4 4 2 2 2 4 2 7 2 7 2 4 2 7 2 7 4 2 2 2 6 6 2 2 2
5 6 7
[7309] 2 4 2 4 4 4 2 4 4 7 2 2 4 4 2 2 2 2 5 4 2 2 2 4 2 2 2 4 5 4 2 5 2
2 2 4
[7345] 2 2 2 4 5 6 2 2 4 2 2 2 4 6 4 4 4 4 2 5 2 2 2 4 4 7 7 2 2 6 4 4 4
4 4 4
[7381] 4 4 2 7 6 4 7 2 3 2 2 2 2 2 7 2 7 7 4 7 6 2 4 7 2 4 2 6 4 4 7 4 2
2 4 2
[7417] 2 4 4 2 2 2 2 7 2 2 2 7 2 5 2 4 4 2 4 7 2 4 7 2 2 2 2 7 7 2 2 2 2
2 2 2
[7453] 2 4 4 4 2 7 4 4 2 2 2 2 7 7 4 2 2 2 7 4 7 2 2 2 2 2 7 4 7 2 2 2 2

4 4 2
[7489] 7 7 6 7 4 2 6 7 2 7 2 2 2 7 2 2 4 2 7 2 4 6 2 2 7 4 4 2 2 4 4 2 2
4 2 4
[7525] 7 7 2 4 7 4 2 4 4 4 2 4 6 4 4 6 2 2 7 4 4 4 2 2 2 6 2 2 2 2 2 2 2
4 7 7
[7561] 4 4 7 7 2 7 2 7 2 4 4 2 2 2 4 7 4 2 2 2 2 2 4 4 4 7 2 2 2 2 2 2 4
7 2 2
[7597] 4 7 7 2 4 6 4 4 2 2 2 2 4 4 2 4 2 4 4 4 6 4 4 2 2 2 2 4 2 2 4 2 2
4 2 2
[7633] 7 5 2 7 4 4 2 2 4 4 2 2 7 2 7 7 3 4 2 2 4 2 4 6 2 4 4 4 2 2 7 2 2
4 4 2
[7669] 2 2 2 2 4 2 2 6 4 2 7 2 2 4 2 4 7 4 2 4 2 7 2 2 4 2 5 4 2 2 2 4 4
2 2 2
[7705] 4 4 4 2 2 2 4 4 2 4 4 4 6 4 7 2 2 7 3 2 4 6 2 2 2 4 2 2 6 7 4 4 7
4 2 2
[7741] 2 2 2 6 7 2 2 2 7 7 2 4 4 2 7 4 2 2 2 7 4 4 4 2 4 4 7 2 4 2 4 7 4
4 4 6
[7777] 2 7 4 4 2 2 4 2 4 2 2 4 5 7 4 2 2 7 2 4 2 4 4 2 4 4 2 2 4 7 4 4 2
2 7 2
[7813] 2 2 7 7 7 2 7 4 2 4 2 2 4 2 5 4 4 7 2 2 6 2 4 4 4 2 4 2 4 4 6 6 4
6 7 4
[7849] 4 2 7 2 2 2 2 4 7 4 4 4 7 2 4 6 4 2 2 2 2 4 4 2 2 2 4 4 4 2 7 7 4
2 2 2
[7885] 4 4 6 2 2 4 4 4 2 2 2 2 4 7 2 2 4 4 2 2 4 2 4 2 2 2 4 6 2 7 2 2 4
2 4 4
[7921] 2 4 2 4 4 2 4 4 6 2 6 2 7 2 4 2 7 2 2 7 2 2 4 7 2 4 4 4 2 2 2 4 2
4 4 2
[7957] 7 5 2 4 6 4 5 4 7 7 4 2 2 2 2 2 2 2 2 4 2 4 2 2 7 2 2 2 4 4 2 2 2
4 4 2
[7993] 6 2 2 4 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 6 2 4 2 2 2 2 4 2 2 6 2 2
2 7 7
[8029] 2 2 4 4 2 2 2 2 2 2 2 7 4 2 4 2 2 2 2 2 2 2 2 2 6 2 2 2 2 1 2 4 2
2 2 2
[8065] 4 2 2 2 2 2 2 2 2 4 4 2 2 4 4 5 5 3 2 4 2 2 2 2 7 2 2 2 7 2 2 2 4
4 2 2
[8101] 2 2 4 6 2 7 6 2 7 2 2 7 2 2 2 4 2 2 2 2 2 2 4 2 2 4 2 2 2 2 2 4 4
2 4 2
[8137] 2 4 2 2 2 4 2 6 2 2 2 4 2 2 2 2 2 2 4 2 2 2 4 6 2 4 4 2 2 2 2 2 4
2 2 6
[8173] 2 2 4 2 7 2 2 4 6 2 2 2 2 7 2 2 2 7 4 2 6 2 2 2 2 2 4 2 2 2 4 4 2
2 2 2
[8209] 2 2 2 6 7 2 4 7 4 2 4 4 7 4 2 2 7 2 7 2 2 2 4 7 2 2 4 2 2 2 4 4 7
2 6 2
[8245] 2 2 4 2 2 2 2 2 4 2 2 5 4 2 2 2 7 2 2 4 4 2 2 6 4 2 4 2 2 5 2 2 4
7 2 2
[8281] 2 2 2 4 2 2 2 5 4 2 2 2 4 2 5 4 2 2 4 7 2 2 2 2 2 2 4 4 2 2 2 2 2
7 4 2
[8317] 4 2 4 2 2 2 2 2 2 4 2 2 2 4 6 4 6 2 2 6 2 4 2 2 7 2 4 2 2 2 2 4 2
2 2 2
[8353] 4 2 3 2 2 6 4 2 2 7 2 4 2 4 2 4 6 5 2 4 2 2 2 6 4 2 2 2 2 2 2 4 4

2 2 2
[8389] 2 2 2 2 7 2 2 2 2 2 7 2 2 2 7 4 7 2 2 4 2 4 2 2 2 2 2 2 2 2 2 4 2
7 4 2
[8425] 4 2 2 2 2 2 2 2 7 2 2 4 2 4 4 4 7 4 2 2 4 2 4 2 2 2 2 2 2 4 2 2 2
2 2 2
[8461] 2 2 2 2 4 4 2 4 7 2 2 4 2 6 4 2 4 2 2 2 2 2 7 6 4 2 2 2 4 2 2 7 7
4 2 2
[8497] 4 4 2 2 4 7 2 4 2 7 2 7 2 4 2 2 2 4 2 4 2 2 4 4 2 7 2 2 2 2 2 2 4
4 2 2
[8533] 2 4 3 2 4 2 2 2 2 2 2 7 2 2 4 2 2 2 2 2 2 2 2 2 5 2 2 2 6 2 2 2 2
4 2 2
[8569] 4 2 6 7 7 4 6 2 2 2 2 5 7 2 7 2 2 2 4 2 7 2 2 2 2 2 7 7 2 2 4 2 2
2 2 2
[8605] 7 4 2 2 2 7 2 2 2 4 2 5 4 2 6 6 2 2 2 4 2 2 2 2 2 2 2 2 4 2 2 2 2 2
2 4 2
[8641] 2 6 2 2 4 4 6 2 4 2 7 4 4 2 2 2 4 2 7 6 2 2 2 2 5 2 2 4 2 7 2 4 2
7 4 2
[8677] 2 4 2 2 5 2 7 2 2 4 2 4 2 2 2 2 7 3 2 4 2 2 2 2 7 2 2 2 2 2 2 4 2
2 2 2
[8713] 2 2 2 2 2 2 2 4 4 2 2 4 6 4 7 4 2 4 2 2 2 4 2 4 4 2 2 2 2 2 4 2 2
2 2 7
[8749] 2 2 2 4 2 2 4 4 7 2 2 2 2 2 2 7 7 4 2 2 2 2 3 4 2 4 4 4 2 4 4 2 2
2 2 2
[8785] 6 4 2 2 2 6 2 2 2 2 4 2 2 2 2 2 4 2 7 2 2 2 6 2 2 2 4 4 7 2 2 2 3
2 2 6
[8821] 4 2 7 2 2 4 4 2 6 4 2 2 2 2 2 2 7 2 2 2 2 7 2 2 4 2 2 4 2 2 2 2 7
2 2 4
[8857] 7 2 4 2 7 2 2 7 4 4 2 2 2 2 2 7 2 2 2 2 2 2 4 4 4 7 2 2 2 4 4 4 2
2 2 2
[8893] 7 4 7 2 2 2 2 4 2 2 4 2 2 4 2 4 2 2 2 7 2 2 4 4 2 2 6 2 2 2 2 4 2
2 2 2
[8929] 6 2 2 2 2 2 2 2 2 4 4 2 2 2 2 5 2 2 2 2 2 2 2 6 2 2 2 7 2 2 5 2 2
2 7 4
[8965] 2 2 4 2 2 2 2 4 2 2 2 4 2 2 2 7 2 4 6 2 2 7 2 2 2 2 2 7 2 2 2 2 2
7 2 2
[9001] 4 4 4 6 7 7 2 2 3 2 2 2 2 4 2 2 2 7 4 2 2 4 2 2 2 2 4 4 7 5 2 2 2
2 2 2
[9037] 7 2 4 2 7 2 2 4 5 4 2 2 2 2 2 2 4 2 5 2 2 4 4 2 4 2 2 7 2 2 2 4
2 2 2
[9073] 2 2 2 4 2 2 2 2 2 4 4 2 2 7 2 4 6 2 7 2 6 2 4 2 4 2 2 5 7 6 2 2 2
2 2 2
[9109] 6 2 7 2 2 2 4 2 4 2 2 2 2 4 2 2 2 2 4 2 4 2 2 2 7 7 2 2 4 2 2 7 6
2 7 2
[9145] 2 2 4 2 2 7 4 4 2 2 6 2 2 2 2 2 2 2 2 7 4 2 2 2 2 2 2 2 2 2 2 2 2
2 4 2
[9181] 2 7 4 7 2 4 2 2 7 2 6 2 4 2 4 4 6 4 2 7 7 4 6 2 6 2 4 2 2 7 7 2 2
2 2 4
[9217] 2 2 2 2 2 4 4 4 1 2 2 2 7 2 2 2 2 2 7 6 7 4 4 4 7 2 2 4 4 2 4 2 4
7 2 2
[9253] 4 2 2 2 2 7 2 4 2 2 2 4 2 7 4 2 2 7 2 2 2 2 2 2 2 2 7 4 4 4 2 2 4 7

4 2 4
[9289] 2 2 2 2 2 7 2 2 2 6 2 2 7 2 2 2 4 2 2 2 7 4 7 2 2 2 2 2 2 2 7 2 2
4 7 2
[9325] 2 2 4 2 2 2 7 2 4 2 4 2 4 4 2 3 2 4 2 2 2 2 2 2 2 2 2 2 4 4 7 2 2
4 4 2
[9361] 4 2 2 2 4 2 2 2 4 4 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2 5 7 2 4 2 2 2 2
2 2 2
[9397] 2 2 4 2 2 4 2 2 4 4 2 7 2 2 2 2 2 2 2 2 2 4 4 4 2 2 2 2 2 2 2 2
2 2 2
[9433] 2 2 5 2 2 2 4 2 2 2 2 7 4 2 7 2 4 2 2 4 5 7 2 2 2 2 2 2 2 4 4 2 2
2 2 4
[9469] 2 5 4 7 4 4 4 2 4 7 2 2 2 2 2 2 2 6 2 2 2 7 2 2 2 7 2 7 2 4 2 2 2
2 2 2
[9505] 2 2 4 2 2 2 2 4 4 2 2 2 4 4 2 2 2 2 2 4 4 2 2 2 6 2 7 2 2 4 4 2
2 2 2
[9541] 2 2 2 2 7 4 2 2 3 6 7 2 2 2 2 2 4 7 4 4 2 4 5 2 7 2 2 2 7 2 4 7 4
2 2 4
[9577] 2 2 4 5 6 2 2 7 4 2 2 6 2 4 2 2 2 2 4 2 4 4 2 2 4 2 6 2 2 4 2 2 2
2 2 2
[9613] 2 2 2 4 2 2 2 3 2 6 2 2 2 6 2 2 4 2 4 2 4 2 2 7 2 7 4 7 7 2 7 2 2
2 4 2
[9649] 1 2 4 2 2 2 2 4 4 2 2 2 7 2 2 2 2 2 2 2 4 7 2 2 2 2 7 4 6 2 4 2 2
5 2 4
[9685] 4 2 2 2 2 2 7 5 2 7 4 7 2 2 4 6 2 2 7 2 2 7 7 4 4 4 2 7 2 7 2 2 4
3 7 6
[9721] 2 2 1 7 2 7 4 2 4 2 4 2 6 2 4 2 2 4 2 2 2 7 4 2 2 4 2 4 4 2 7 2 6
2 4 2
[9757] 2 7 2 2 4 2 2 4 2 2 7 2 2 2 7 4 2 2 4 2 7 2 2 4 2 2 7 2 2 2 2 2 3
2 4 2
[9793] 7 2 7 2 4 2 2 2 4 2 2 2 7 2 4 2 2 2 2 2 2 2 4 2 4 2 4 2 2 2 2 6
2 2 2
[9829] 2 2 2 6 2 4 7 7 2 7 2 2 2 2 2 2 4 2 4 2 2 2 2 2 2 2 4 6 2 2 4 5 2
2 2 2
[9865] 2 2 4 2 2 4 2 4 4 7 2 2 2 2 6 2 4 2 2 2 2 2 2 2 2 2 4 4 2 2 2 2
4 2 7
[9901] 6 6 2 2 2 6 7 2 4 4 2 2 7 2 2 2 6 2 2 7 7 4 2 2 2 6 4 2 2 2 2 2
2 2 2
[9937] 2 4 4 2 2 2 4 2 7 6 2 2 7 4 4 2 2 2 4 6 2 3 2 2 2 5 7 4 7 2 2 7 6
2 2 4
[9973] 4 2 2 2 2 2 4 2 2 2 2 7 4 5 2 2 7 4 4 2 2 2 2 4 2 7 4 7 2 2 2 2
2 2 4
[10009] 2 2 4 2 2 2 2 2 2 7 2 7 7 7 2 2 2 2 4 2 4 2 2 2 4 2 2 2 4 4 2 4 2
7 2 4
[10045] 2 4 2 2 2 2 2 2 7 2 2 2 4 2 2 4 2 4 4 7 2 2 2 4 2 2 4 2 2 7 2 4 4
2 2 2
[10081] 2 2 2 2 2 4 2 6 4 2 2 7 4 2 2 2 4 2 2 4 4 2 2 7 7 4 4 2 4 4 2 2 2
2 2 2
[10117] 2 4 2 7 3 2 4 6 2 4 2 2 2 2 4 4 2 2 2 2 2 2 2 2 4 4 2 2 4 2 7 2 2
5 4 2
[10153] 2 2 4 4 2 4 2 4 4 2 7 2 2 4 5 2 2 4 4 2 7 2 2 4 2 2 2 2 7 4 2 5 7

2 2 2
[10189] 7 7 2 2 2 2 5 2 2 2 6 2 2 4 2 2 2 4 7 6 2 2 2 4 4 2 7 2 2 2 2 2 2
7 2 2
[10225] 2 2 2 4 2 2 6 7 2 4 2 2 2 7 2 2 4 4 4 2 4 4 2 2 2 2 2 2 7 2 2 2 2
2 2 2
[10261] 6 2 2 2 2 4 7 4 2 2 2 2 2 2 6 2 2 2 2 2 7 2 4 2 5 2 2 2 6 2 2 2 2
4 2 2
[10297] 4 2 7 3 2 2 2 2 6 7 2 4 2 7 4 2 2 5 2 7 2 2 2 6 6 2 2 2 2 2 2 2
2 2 2
[10333] 2 2 4 2 5 4 4 2 4 2 2 7 6 2 6 4 6 4 4 6 2 2 2 2 2 2 5 2 2 2 2 4 4
2 2 4
[10369] 4 4 4 4 4 7 2 2 4 2 7 2 2 2 2 2 2 4 4 4 2 2 4 7 4 2 2 4 2 2 4 2 4
2 2 2
[10405] 2 3 4 4 2 2 2 2 2 2 2 2 7 4 4 2 4 7 2 2 7 2 7 2 4 2 2 2 7 2 2 2 2
4 2 2
[10441] 7 4 4 2 2 6 7 2 2 2 2 4 4 2 2 7 2 2 6 2 2 2 2 2 2 2 2 4 2 2 2 7 2
2 2 2
[10477] 2 4 7 2 2 4 2 2 2 6 2 2 4 2 7 2 2 7 2 6 7 2 2 2 4 6 2 7 2 2 2 4 6
2 2 4
[10513] 2 2 2 2 7 2 2 2 2 4 4 2 7 4 2 2 4 2 2 4 2 4 2 2 4 2 4 7 2 4 4 2 4
2 2 7
[10549] 2 2 2 2 2 7 2 2 2 7 2 2 2 2 7 2 2 2 2 7 2 2 2 2 4 2 2 2 2 4 4 2
2 2 2
[10585] 2 2 2 4 4 2 2 2 7 2 2 2 2 2 6 2 2 2 2 2 2 7 7 2 7 2 2 7 2 6 2 2
2 4 2
[10621] 2 2 2 4 4 2 3 6 4 2 7 2 2 2 4 2 2 2 2 2 2 2 2 6 4 2 2 7 2 6 6 5 4
2 7 2
[10657] 2 2 4 4 7 2 2 2 4 7 2 2 6 7 2 2 2 2 2 7 2 2 7 2 2 2 4 2 2 2 2 2 2
2 2 4
[10693] 2 2 4 6 2 2 4 2 4 3 2 7 2 2 4 2 7 2 2 2 2 2 2 2 2 7 2 2 4 2 2 2 2
2 2 2
[10729] 2 2 2 2 2 2 4 7 2 2 7 2 4 2 7 2 4 2 6 2 4 2 4 2 7 2 2 2 2 2 4 2 4
7 4 4
[10765] 4 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 7 2 6 2 2 2 7 2 2 2
4 2 2
[10801] 2 2 4 2 2 2 2 2 2 2 2 2 2 7 2 2 2 2 2 6 2 4 2 2 4 4 2 2 2 2 2 2
2 2 2
[10837] 2 2 4 2 2 2 2 2 4 2 2 2 4 7 2 2 4 2 2 2 2 2 2 4 2 7 2 7 4 2 2 4 7
4 4 4
[10873] 2 6 4 6 2 2 4 2 4 2 2 2 2 2 2 2 4 2 7 2 2 2 6 2 2 2 7 4 2 2 5 2 2
2 2 2
[10909] 2 2 2 2 2 6 2 2 7 4 2 2 7 4 2 5 2 2 2 6 2 2 2 2 2 4 7 2 2 2 2 2 4
6 7 2
[10945] 2 2 2 2 2 2 2 2 2 2 2 2 4 2 4 2 2 2 2 5 7 2 2 2 7 2 4 6 2 7 2 2 2
7 2 4
[10981] 2 6 2 2 2 4 2 4 2 4 4 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 7 7 7
2 2 2
[11017] 2 4 2 7 7 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 2 2 2 7 2 4 7 2 4 2 2
2 2 2
[11053] 2 2 7 7 2 6 2 2 2 2 2 4 2 2 2 2 2 2 2 7 2 4 2 7 4 2 2 7 2 2 2 2 2

6 2 2
[11089] 7 6 2 4 2 2 7 2 2 2 2 2 2 2 7 2 5 7 4 2 7 2 7 2 2 4 4 2 2 2 2 2 4
2 2 7
[11125] 2 2 2 2 2 2 2 2 2 2 7 2 2 4 4 4 2 2 2 4 2 2 2 7 2 2 7 2 2 2 3 4 2
7 2 2
[11161] 2 2 2 2 2 2 2 2 4 2 2 2 7 2 2 4 2 2 7 4 4 2 7 2 4 4 2 4 2 2 7 2 2
7 2 2
[11197] 2 2 2 2 2 2 2 2 7 7 2 2 2 2 2 2 2 2 7 4 2 2 3 2 2 2 2 2 4 2 2 2 2
2 4 4
[11233] 2 4 4 2 5 2 6 4 6 6 6 7 6 2 2 2 4 2 6 2 4 2 2 7 7 2 2 2 7 2 4 4 4
2 6 2
[11269] 2 2 4 2 2 2 2 5 4 2 2 7 2 4 2 2 2 4 2 7 4 2 2 2 2 2 5 2 2 7 2 6 7
4 7 4
[11305] 2 4 2 4 2 2 2 2 2 2 2 2 2 2 2 2 4 7 2 2 2 2 2 2 4 7 6 2 7 6 2 4 2
6 2 2
[11341] 2 2 7 2 4 2 2 2 2 2 7 2 2 7 2 2 2 7 2 7 2 4 2 2 2 2 2 2 2 4 2 4 2
2 2 2
[11377] 2 4 2 7 4 2 2 2 2 4 2 2 2 2 2 6 7 4 7 2 2 4 2 2 2 2 2 3 2 2 2 2 4
2 4 7
[11413] 2 2 4 2 4 2 2 2 2 4 4 2 2 7 2 2 7 2 2 2 7 4 4 2 7 4 2 2 2 2 4 2 2
2 4 2
[11449] 2 2 2 2 2 2 7 4 4 4 4 2 2 2 3 6 4 7 2 4 2 4 7 2 4 2 2 2 2 2 4 2 4
7 2 4
[11485] 2 7 2 2 2 2 7 2 2 4 2 4 2 7 4 2 4 2 2 4 2 4 2 2 2 2 2 2 7 2 2 2 4
4 4 4
[11521] 2 2 2 5 2 2 2 7 7 2 2 4 2 2 4 2 2 7 2 2 4 7 2 7 2 2 2 7 2 2 2 2 2
2 2 2
[11557] 2 4 2 4 5 4 2 2 2 6 2 2 4 4 4 7 2 2 2 2 2 2 2 7 7 2 4 4 7 2 7 6 2
7 2 2
[11593] 2 2 2 7 2 6 2 4 2 2 4 2 2 7 2 2 2 2 2 4 2 2 4 4 2 2 4 4 2 2 2 2 4
4 7 2
[11629] 4 6 2 2 2 2 4 2 4 2 2 4 2 4 2 2 2 2 2 2 2 2 2 2 2 4 2 2 2 2 2 2 7
2 4 2
[11665] 2 2 2 2 7 2 4 2 2 2 2 2 2 4 2 7 4 2 4 2 2 4 2 4 4 2 2 2 2 2 2 2 2
2 1 4
[11701] 2 2 4 2 2 2 2 2 7 2 2 2 2 2 2 7 2 2 2 2 2 5 2 2 4 4 7 4 2 2 2 2 2
2 4 2
[11737] 2 2 2 7 4 2 4 2 2 2 2 2 2 2 2 7 2 4 2 2 7 2 2 4 2 4 2 2 2 7 7 2 2 4
4 2 2
[11773] 2 7 2 2 2 2 2 2 7 6 2 2 2 4 2 2 2 2 2 2 2 4 2 2 2 2 2 2 2 2 4 2 2 2
2 2 2
[11809] 7 4 2 4 2 2 2 7 4 2 2 4 4 2 2 4 2 5 4 2 2 2 2 7 2 4 4 2 4 7 6 4 6
5 4 6
[11845] 2 2 2 7 4 2 2 2 6 2 7 7 4 6 2 2 2 2 4 5 2 4 2 2 2 2 2 2 2 4 4 6 2 4
4 2 4
[11881] 2 2 7 4 7 2 2 2 2 7 2 2 4 7 2 2 2 2 2 2 4 2 4 2 2 2 2 2 2 2 2 2 4
4 2 4
[11917] 2 4 2 2 2 2 7 4 2 2 2 2 2 2 2 2 4 2 4 2 2 2 2 2 2 2 2 2 2 2 2 7 2
2 5 6
[11953] 2 2 2 2 2 4 2 2 2 4 2 4 4 2 2 2 4 6 2 2 6 4 7 2 2 2 4 2 7 2 2 2 2

```

4 2 4
## [11989] 4 2 2 4 4 2 2 4 2 2 2 6 2 2 2 1 2 2 2 2 2 7 2 2 2 5 2 2 2 2 4 6 2
2 6 2
## [12025] 4 2 7 2 6 2 2 7 2 2 7 2 4 2 7 4 2 2 2 2 4 2 2 2 7 2 2 7 2 7 2 4 4
2 2 2
## [12061] 4 4 2 4 5 2 2 2 2 4 4 2 2 2 2 4 2 4 2 7 6 5 7 2 6 2 4 2 4 7 2 2 4
2 2 4
## [12097] 2 4 2 4 4 4 2 2 2 2 4 2 6 2 2 2 2 2 4 4 7 7 2 4 2 2 2 2 2 7 2 2 2
2 2 4
## [12133] 2 4 7 4 2 4 2 2 2 2 2 2 7 2 2 2 2 2 2 4 2 4 4 5 4 2 2 2 2 2 3 2 3
7 2 2
## [12169] 4 2 3 2 2 2 6 4 4 4 6 2 5 4 2 2 4 4 2 2 2 2 6 2 2 4 2 7 2 7 2 2 2
2 2 2
## [12205] 2 2 2 7 2 2 2 2 2 4 2 2 6 2 2 6 2 4 7 2 2 2 2 2 7 2 5 2 2 7 2 2 2
2 2 2
## [12241] 2 4 2 2 2 4 7 4 2 2 2 6 7 2 4 4 4 4 4 4 2 2 4 2 7 2 4 4 2 2 2 6 2
7 4 2
## [12277] 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 2 7 2 4 2 2 4 2 2 2 2 4 2 2 2 2 4
4 2 4
## [12313] 2 2 4 2
##
## Within cluster sum of squares by cluster:
## [1] 2160830 1461535 2402503 2513972 2692651 2459740 2520197
## (between_SS / total_SS = 95.8 %)
##
## Available components:
##
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withi
nss"
## [6] "betweenss"    "size"         "iter"         "ifault"

```

#Previewing the no. of records in each cluster

```

result$size
## [1] 12 7943 63 2532 167 434 1165

```

#Installing packages

```

library(fpc)
library(dbSCAN)

##
## Attaching package: 'dbSCAN'

## The following object is masked from 'package:fpc':
##
## dbSCAN

if(!require(devtools)) install.packages("devtools")

## Loading required package: devtools

```

```
## Loading required package: usethis

devtools::install_github("kassambara/factoextra")

## Skipping install of 'factoextra' from a github remote, the SHA1 (1689fc74)
has not changed since last install.
## Use `force = TRUE` to force installation

library(factoextra)

## Welcome! Want to learn more? See two factoextra-related books at https://g
oo.gl/ve3WBa
```

#Installing the package

```
#install.packages("factoextra")
```

#visualizing the data with k=7

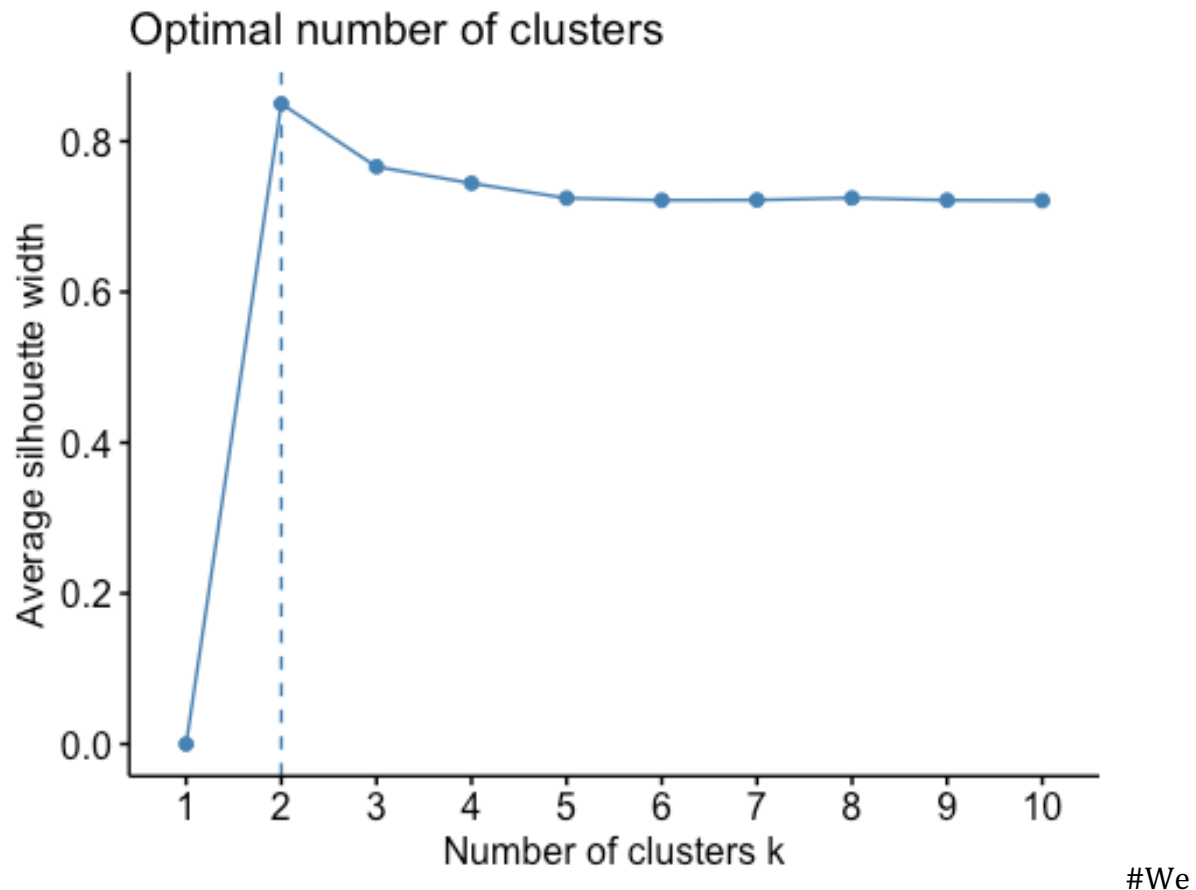
```
library(factoextra)
fviz_cluster(result, data=new_data,ggtheme =theme_bw())
```



#Determining the optimal number of clusters #Using silhouette method
install.packages("cluster")

#Visualizing the optimal number of k

```
library(cluster)
library(NbClust)
fviz_nbclust(x=new_data, FUNcluster=kmeans, method ="silhouette")
```



observe that 2 clusters are the optimal clusters

#kmeans with optimal clusters(2)

```
new_datak2<-kmeans(new_data, 2, nstart=50)
new_datak2

## K-means clustering with 2 clusters of sizes 11789, 527
##
## Cluster means:
##      admin admin_dur      info  info_dur      prod  prod_dur bouncerat
## 1 2.059547   52.4777 0.01892301 0.01233418 0.04272152 0.01738792  0.11400
## 2 8.094877  716.8517 0.06744149 0.04944220 0.09725867 0.04831182  0.03811
##      exitrates pagevalues specialday      month ops_systems  browser      regio
## 1 0.2198421 0.01591171 0.06270252 0.5725017  0.1605128 0.1131351 0.268587
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
## [7057] 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1
1 1 1
## [7093] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1
1 1 1
## [7129] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7165] 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7201] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1
1 1 1
## [7237] 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7273] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1
2 2 1
## [7309] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 1 1 2 1
1 1 1
## [7345] 1 1 1 1 2 2 1 1 1 1 1 1 1 2 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1
1 1 1
## [7381] 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7417] 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7453] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7489] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7525] 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 2 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1
1 1 1
## [7561] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7597] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7633] 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7669] 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1
1 1 1
## [7705] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1
1 1 1
## [7741] 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7777] 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7813] 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 2 1 1 1 1 1 1 1 1 1 1 1 2 2 1
1 1 1
## [7849] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
## [7885] 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 2 1 1 1 1 1 1 1
1 1 1
## [7921] 1 1 1 1 1 1 1 1 2 1 2 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1
1 1 1
```

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

```
## [1] "cluster"      "centers"      "totss"        "withinss"     "tot.withi
nss"
## [6] "betweenss"    "size"         "iter"         "ifault"
```

#Visualize with 2 clusters

```
fviz_cluster(new_datak2, data=new_data, ggtheme=theme_bw())
```



#Observations on kmeans: we observe that it is very easy to implement. When we applied 7 clusters it was difficult to distinguish the clusters, but when we calculated the optimal value of k, it's very clear the datapoints of each cluster. We conclude that based on this the online customers have been clustered into 2 groups/clusters.

#b) Implementing unsupervised learning using Hierarchical clustering Hierarchical clustering builds a hierarchy of clusters i.e. tree-type structure based on hierarchy. There are two types: Agglomerative and Divisive.

Advantages

#We do not need to specify the number of clusters required for the algorithm.

#Hierarchical clustering outputs a hierarchy, i.e. a structure that is more informative than the unstructured set of flat clusters returned by k-means. #It is also easy to implement.

Below are the limitations of the hierarchical clustering technique;

#There is no mathematical objective for Hierarchical clustering. #High space and time complexity for Hierarchical clustering. Hence this clustering algorithm cannot be used when we have huge data

#Installing foreign install.packages("foreign")

#loading the library

```
library(foreign)
```

#Before hierarchical clustering, we will compute some descriptive statistics

```
desc_stats <- data.frame(  
  Min = apply(new_data, 2, min),      # minimum  
  Med = apply(new_data, 2, median),   # median  
  Mean = apply(new_data, 2, mean),    # mean  
  SD = apply(new_data, 2, sd),        # Standard deviation  
  Max = apply(new_data, 2, max)       # Maximum  
)  
desc_stats <- round(desc_stats, 1)  
head(desc_stats)
```

	Min	Med	Mean	SD	Max
admin	0	1	2.3	3.3	27.0
admin_dur	-1	8	80.9	176.9	3398.8
info	0	0	0.0	0.1	1.0
info_dur	0	0	0.0	0.1	1.0
prod	0	0	0.0	0.1	1.0
prod_dur	0	0	0.0	0.0	1.0

#We note that admin duration has high mean and maximum value compared to other variables. #we therefore need to standardize the variables(i.e., scaled) to make them comparable. this is # transforming the variables such that they have mean zero and standard deviation one.

#Scaling the data

```
new_data <- scale(new_data)  
head(new_data)
```

	admin	admin_dur	info	info_dur	prod	prod_dur
[1,]	-0.6975533	-0.4574578	-0.3966145	-0.2450294	-0.6914734	-0.6247671
[2,]	-0.6975533	-0.4574578	-0.3966145	-0.2450294	-0.6689966	-0.5913358
[3,]	-0.6975533	-0.4631119	-0.3966145	-0.2521304	-0.6914734	-0.6252895
[4,]	-0.6975533	-0.4574578	-0.3966145	-0.2450294	-0.6689966	-0.6233742
[5,]	-0.6975533	-0.4574578	-0.3966145	-0.2450294	-0.4891823	-0.2969835
[6,]	-0.6975533	-0.4574578	-0.3966145	-0.2450294	-0.2868911	-0.5442099

	bouncerrates	exitrates	pagevalues	specialday	month	ops_systems
[1,]	3.67247746	3.2352400	-0.3173633	-0.309001	-1.334201	-1.2332048
[2,]	-0.45743910	1.1745443	-0.3173633	-0.309001	-1.334201	-0.1361914
[3,]	3.67247746	3.2352400	-0.3173633	-0.309001	-1.334201	2.0578354

```
## [4,]  0.57504004  1.9988226 -0.3173633 -0.309001 -1.334201  0.9608220
## [5,] -0.04444744  0.1441964 -0.3173633 -0.309001 -1.334201  0.9608220
## [6,] -0.13139305 -0.3800157 -0.3173633 -0.309001 -1.334201 -0.1361914
##      browser      region traffic_type visitor_type  weekend      admni
## [1,] -0.7901988 -0.8941841 -0.76292777  0.4080401 -0.5505615 -0.6975533
## [2,] -0.2081361 -0.8941841 -0.51445574  0.4080401 -0.5505615 -0.6975533
## [3,] -0.7901988  2.4360812 -0.26598370  0.4080401 -0.5505615 -0.6975533
## [4,] -0.2081361 -0.4779009 -0.01751167  0.4080401 -0.5505615 -0.6975533
## [5,]  0.3739266 -0.8941841 -0.01751167  0.4080401  1.8161802 -0.6975533
## [6,] -0.2081361 -0.8941841 -0.26598370  0.4080401 -0.5505615 -0.6975533
##      admni_dur
## [1,] -0.4574578
## [2,] -0.4574578
## [3,] -0.4631119
## [4,] -0.4574578
## [5,] -0.4574578
## [6,] -0.4574578
```

#We will use the R function `hclust()` for hierarchical clustering #First we use the `dist()` function to compute the Euclidean distance between observations, #d will be the first argument in the `hclust()` function dissimilarity matrix

```
d <- dist(new_data, method = "euclidean")
```

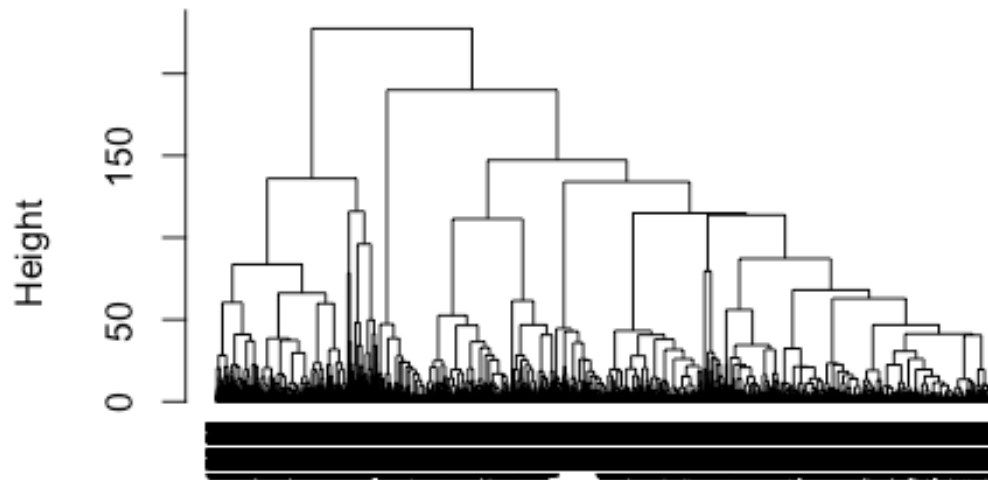
#We then hierarchical clustering using the Ward's method

```
res.hc <- hclust(d, method = "ward.D2" )
```

#plot the obtained dendrogram

```
plot(res.hc, cex = 0.6, hang = -1)
```

Cluster Dendrogram



d
hclust (*, "ward.D2")

#Observations of hierarchical clustering: This algorithm has limitations with huge data. As it is very difficult to identify which variables are in what cluster and how variables are clustered. Based on the dendrogram, we should have at least 10 clusters yet when we used kmeans, we established the optimal numbers of clusters are 2. You get very different results with each approach. Though both are very easy to implement.

#c)Implementing using DBSCAN

DBSCAN (Density-Based Spatial Clustering of Applications with Noise) is a type of clustering algorithm that focuses more on the proximity and density of observations to form clusters. This algorithm is commonly used to identify clusters of any shape in a data set containing noise and outliers.

The algorithm is not sensitive to outliers/noise, its is therefore the best for our dataset as it had many outliers.It is also applied in customer segmentation problem and this study is one of those.

Limitations #It does not work well when dealing with clusters of varying densities. #It also does not work well with high dimensionality data.

Importing the required package

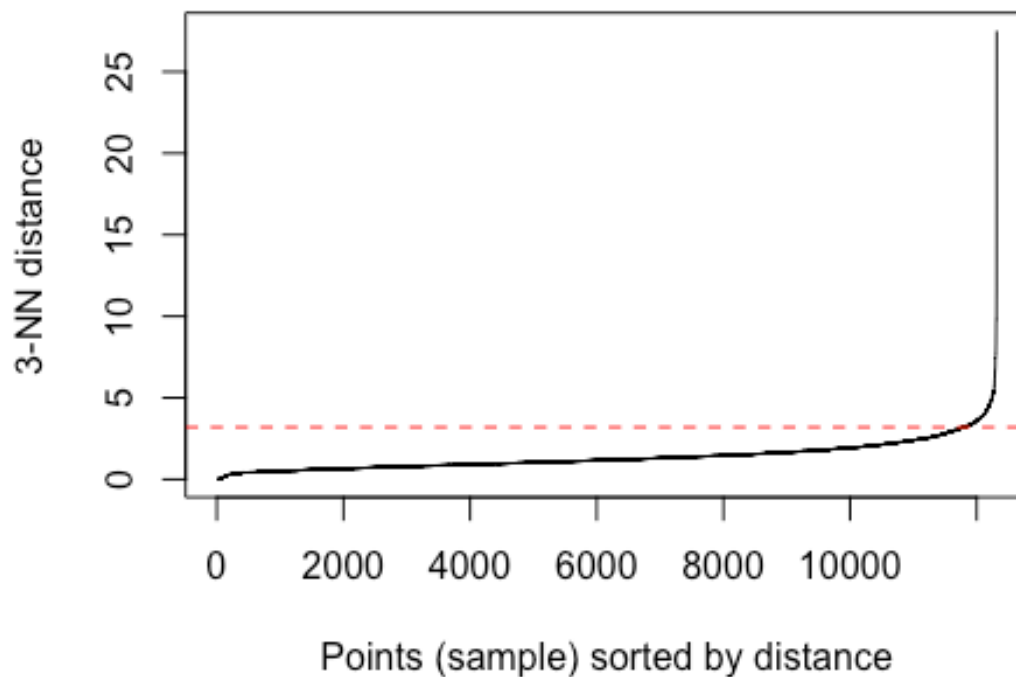
```
install.packages("dbscan")
```

Loading the required library

```
library(dbscan)
```

#Determining the optimal value of eps

```
kNNdistplot(new_data, k = 3)  
abline(h=3.2, col = "red", lty=2)
```



#We

note that the optimal value of eps is 3.2

#Applying our DBSCAN algorithm #We want minimum 4 points with in a distance of eps(0.4)

```
db <- dbscan(new_data, eps = 0.4, minPts = 4)  
print(db)
```

```
## DBSCAN clustering for 12316 objects.
```

```
## Parameters: eps = 0.4, minPts = 4
```

```
## The clustering contains 74 cluster(s) and 11753 noise points.
```

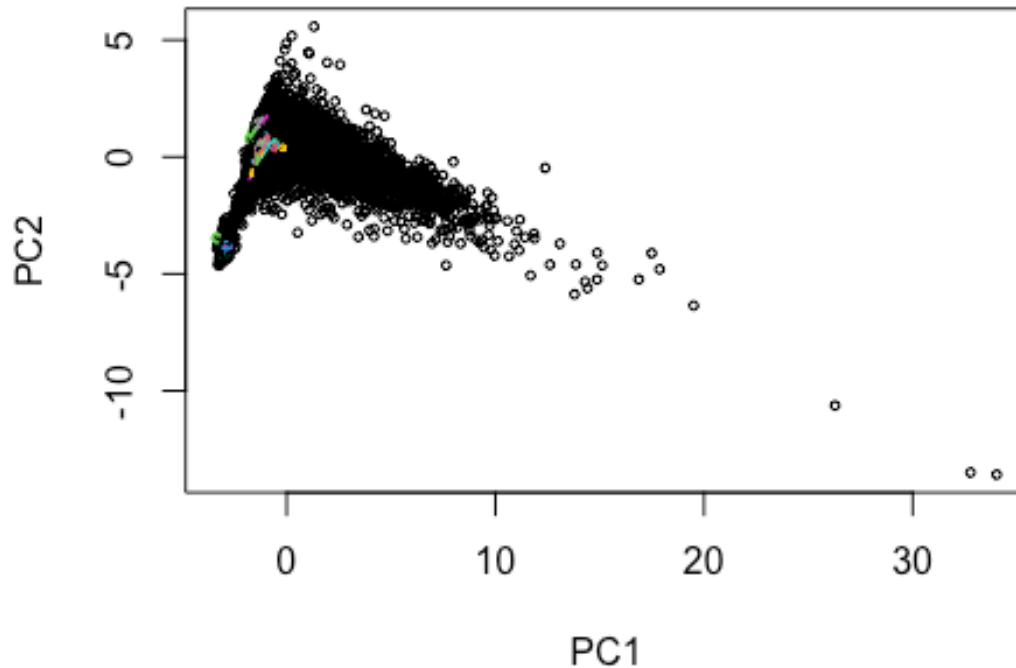
```
##
##      0      1      2      3      4      5      6      7      8      9     10     11
12
## 11753     12     52      4      8     20     11      4      8      4      4      4
10
##      13     14     15     16     17     18     19     20     21     22     23     24
25
##      13     12      6      6      4      6      4      4      4      4      4      3
7
##      26     27     28     29     30     31     32     33     34     35     36     37
38
##      4      4      5      5      5      5     24      4      4     12      5     12
7
##      39     40     41     42     43     44     45     46     47     48     49     50
51
##      4      5      4      4      4     35      4     16      9      5     20     10
9
##      52     53     54     55     56     57     58     59     60     61     62     63
64
##      6      7      7      7      4      7      5      4      4      5      4     10
4
##      65     66     67     68     69     70     71     72     73     74
##      5      5      5      5      6      4      4      4      4      4
##
## Available fields: cluster, eps, minPts
```

#We also plot our clusters as shown #The dataset and cluster method of dbscan is used to plot the clusters.

```
hullplot(new_data,db$cluster)
```

```
## Warning in hullplot(new_data, db$cluster): Not enough colors. Some colors
will
## be reused.
```


Convex Cluster Hulls



#From this we observe that the clusters have been 2 components leaving out the noisy data.

6)Challenge the solution Having applied the 3 algorithmic approaches, this gave us a feel of how the results of each approach turned out. We believe DBSCAN and kmeans with optimal value of k did a good job in clustering the customer data into 2 clusters. We were able to achieve the grouping of data points into distinct non-overlapping subgroups.

7)Follow up questions 1. Did we have the right data? Yes, we had the right data. 2. what could improve? The hierarchical clustering was not able to bring out the insights the variables in each cluster.