

Part-B-Q3.knit

Part B Q3

Clear Memory & Set Data File Path

Put the path of the customer_churn file folder as argument for setwd(“customer_churn file folder path”)

```
rm(list = ls(all.names = TRUE)); gc();  
  
##          used (Mb) gc trigger (Mb) max used (Mb)  
## Ncells 511719 27.4    1133064 60.6    657980 35.2  
## Vcells 932566   7.2     8388608 64.0   1768021 13.5  
  
setwd("D:/uOttawa/First Semester/Fundamentals and Applied Data Science/Assignment 1/Assignment 1")
```

Load customer_churn.csv file Data in DataFrame

Load customer_churn.csv file Data in DataFrame.

```
# Specify the path to your CSV file  
csv_file_path <- "customer_churn.csv"  
# Load the data from the CSV file into a data frame  
customer_churn_data <- read.csv(csv_file_path)
```

Explore customer_churn data

```
head(customer_churn_data)  
  
##   Customer.ID Gender Age Married Number.of.Dependents      City Zip.Code  
## 1  0002-ORFBO Female  37     Yes                   0 Frazier Park   93225  
## 2  0003-MKNFE  Male  46     No                   0   Glendale   91206  
## 3  0004-TLHLJ  Male  50     No                   0   Costa Mesa   92627  
## 4  0011-IGKFF  Male  78     Yes                   0   Martinez   94553  
## 5  0013-EXCHZ Female  75     Yes                   0   Camarillo   93010  
## 6  0013-MHZWF Female  23     No                   3   Midpines   95345  
##   Latitude Longitude Number.of.Referrals Tenure.in.Months  Offer Phone.Service
```

## 1	34.82766	-118.9991	2	9	None	Yes
## 2	34.16251	-118.2039	0	9	None	Yes
## 3	33.64567	-117.9226	0	4	Offer E	Yes
## 4	38.01446	-122.1154	1	13	Offer D	Yes
## 5	34.22785	-119.0799	3	3	None	Yes
## 6	37.58150	-119.9728	0	9	Offer E	Yes
##	Avg.Monthly.Long.Distance.Charges		Multiple.Lines		Internet.Service	
## 1		42.39		No		Yes
## 2		10.69		Yes		Yes
## 3		33.65		No		Yes
## 4		27.82		No		Yes
## 5		7.38		No		Yes
## 6		16.77		No		Yes
##	Internet.Type		Avg.Monthly.GB.Download		Online.Security	
## 1		Cable		16	No	Yes
## 2		Cable		10	No	No
## 3		Fiber Optic		30	No	No
## 4		Fiber Optic		4	No	Yes
## 5		Fiber Optic		11	No	No
## 6		Cable		73	No	No
##	Device.Protection.Plan		Premium.Tech.Support		Streaming.TV	
## 1		No		Yes	Yes	No
## 2		No		No	No	Yes
## 3		Yes		No	No	No
## 4		Yes		No	Yes	Yes
## 5		No		Yes	Yes	No
## 6		No		Yes	Yes	Yes
##	Streaming.Music		Unlimited.Data		Contract	
## 1		No		Yes	One Year	Yes
## 2		Yes		No	Month-to-Month	No
## 3		No		Yes	Month-to-Month	Yes
## 4		No		Yes	Month-to-Month	Yes
## 5		No		Yes	Month-to-Month	Yes
## 6		Yes		Yes	Month-to-Month	Yes
##	Payment.Method		Monthly.Charge		Total.Charges	
## 1		Credit Card		65.6	593.30	0.00
## 2		Credit Card		-4.0	542.40	38.33
## 3		Bank Withdrawal		73.9	280.85	0.00
## 4		Bank Withdrawal		98.0	1237.85	0.00
## 5		Credit Card		83.9	267.40	0.00
## 6		Credit Card		69.4	571.45	0.00
##	Total.Extra.Data.Charges		Total.Long.Distance.Charges		Total.Revenue	
## 1		0			381.51	974.81
## 2		10			96.21	610.28
## 3		0			134.60	415.45
## 4		0			361.66	1599.51

```

## 5          0          22.14      289.54
## 6          0          150.93     722.38
## Customer.Status Churn.Category      Churn.Reason
## 1          Stayed
## 2          Stayed
## 3          Churned      Competitor Competitor had better devices
## 4          Churned Dissatisfaction      Product dissatisfaction
## 5          Churned Dissatisfaction      Network reliability
## 6          Stayed

str(customer_churn_data)

## 'data.frame':    5244 obs. of  38 variables:
## $ Customer.ID      : chr  "0002-ORFBO" "0003-MKNFE" "0004-TLHLJ" "0011-IGKFF" ...
## $ Gender            : chr  "Female" "Male" "Male" "Male" ...
## $ Age              : int   37 46 50 78 75 23 67 52 68 43 ...
## $ Married          : chr  "Yes" "No" "No" "Yes" ...
## $ Number.of.Dependents : int   0 0 0 0 3 0 0 0 1 ...
## $ City              : chr  "Frazier Park" "Glendale" "Costa Mesa" "Martinez" ...
## $ Zip.Code          : int  93225 91206 92627 94553 93010 95345 93437 94558 93063 95681 ...
## $ Latitude          : num   34.8 34.2 33.6 38 34.2 ...
## $ Longitude         : num  -119 -118 -118 -122 -119 ...
## $ Number.of.Referrals : int   2 0 0 1 3 0 1 8 0 3 ...
## $ Tenure.in.Months    : int   9 9 4 13 3 9 71 63 7 65 ...
## $ Offer             : chr  "None" "None" "Offer E" "Offer D" ...
## $ Phone.Service      : chr  "Yes" "Yes" "Yes" "Yes" ...
## $ Avg.Monthly.Long.Distance.Charges : num  42.39 10.69 33.65 27.82 7.38 ...
## $ Multiple.Lines     : chr  "No" "Yes" "No" "No" ...
## $ Internet.Service   : chr  "Yes" "Yes" "Yes" "Yes" ...
## $ Internet.Type      : chr  "Cable" "Cable" "Fiber Optic" "Fiber Optic" ...
## $ Avg.Monthly.GB.Download : int   16 10 30 4 11 73 14 7 21 14 ...
## $ Online.Security    : chr  "No" "No" "No" "No" ...
## $ Online.Backup      : chr  "Yes" "No" "No" "Yes" ...
## $ Device.Protection.Plan : chr  "No" "No" "Yes" "Yes" ...
## $ Premium.Tech.Support : chr  "Yes" "No" "No" "No" ...
## $ Streaming.TV       : chr  "Yes" "No" "No" "Yes" ...
## $ Streaming.Movies   : chr  "No" "Yes" "No" "Yes" ...
## $ Streaming.Music    : chr  "No" "Yes" "No" "No" ...
## $ Unlimited.Data     : chr  "Yes" "No" "Yes" "Yes" ...
## $ Contract           : chr  "One Year" "Month-to-Month" "Month-to-Month" "Month-to-Month" ...
## $ Paperless.Billing  : chr  "Yes" "No" "Yes" "Yes" ...
## $ Payment.Method     : chr  "Credit Card" "Credit Card" "Bank Withdrawal" "Bank Withdrawal" ...
## $ Monthly.Charge     : num   65.6 -4 73.9 98 83.9 ...
## $ Total.Charges      : num   593 542 281 1238 267 ...
## $ Total.Refunds      : num    0 38.3 0 0 0 ...
## $ Total.Extra.Data.Charges : int    0 10 0 0 0 0 0 20 0 0 ...
## $ Total.Long.Distance.Charges : num  381.5 96.2 134.6 361.7 22.1 ...

```

```
## $ Total.Revenue           : num  975 610 415 1600 290 ...
## $ Customer.Status         : chr   "Stayed" "Stayed" "Churned" "Churned" ...
## $ Churn.Category          : chr   "" "" "Competitor" "Dissatisfaction" ...
## $ Churn.Reason            : chr   "" "" "Competitor had better devices" "Product dissatisfaction" ...
```

Build Total Revenue Cube

that it's dimensions are • offer type (Offer) • internet service type (Internet.Type) • contract duration (Contract) • customer Status (Customer.Status)
so we can apply cube operations regarding those attributes perform adhoc quires.

```
Total.Revenue_cube <-
  tapply(customer_churn_data$Total.Revenue,
    customer_churn_data[,c("Offer", "Internet.Type", "Contract", "Customer.Status")],
    FUN=function(x){return(sum(x))})
```

a)

ROLLUP Total Revenue Cube & Collapse Customer Status Dimension

to aggregate revenue across different customer status in the cube

```
Summarized_Total.Revenue_cube <-
  apply(Total.Revenue_cube, c("Offer", "Internet.Type", "Contract"),
    FUN=function(x) {return(sum(x, na.rm=TRUE))})
```

Show The total revenue contribution

from a Two Year contract for each Offer by internet type by Cube Slice operation

```
Total.Rev.2yr_cont.per.off <- Summarized_Total.Revenue_cube[,,"Two Year"]
print(Total.Rev.2yr_cont.per.off)
```

```
##           Internet.Type
## Offer           Cable      DSL Fiber Optic
##   None      554678.24 1235674.43 2075682.41
##   Offer A  349750.71 720759.09 1474862.52
##   Offer B  164188.21 362671.78 473681.65
##   Offer C   10360.26 34148.92 80196.94
##   Offer D    5443.22 5592.51 22464.25
##   Offer E    1684.86 3019.54 6134.10
```

b)

Calculate Total Revenue From Offer B customers

```
Offer.B_Total.Revenue <- sum(Summarized_Total.Revenue_cube["Offer B",,])
```

Calculate Total Revenue From Offer B churned customers

that accepted Month-to-Month Contract for Cable internet service type

```
Offer.B_churn_mon.to.mon_cable_Total.Revenue <- sum(Total.Revenue_cube["Offer B","Cable","Month-to-Month","Churned"])
```

Calculate Total Revenue Precentage (for Offer B)

contributed by churned customers that accepted a Month-to-Month contract for
Cable service

```
total_rev_precent = Offer.B_churn_mon.to.mon_cable_Total.Revenue / Offer.B_Total.Revenue * 100  
print(paste(total_rev_precent,"%"))  
## [1] "0.451160910747686 %"
```