Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

Aim: Case Study: Analysis of the Shark Tank US dataset

IDE: R Studio

Dataset:

Download the dataset of Shark Tank US from kaggle. Link of dataset:
https://www.kaggle.com/datasets/thirumani/shark-tank-us-dataset/download?datasetVersionNumber=17

Program:

Write R script code and perform the following analysis:

1. Which season is having the overall highest deal in terms of the amount?
2. Enlist episodes for each season having the highest deal in terms of the amount?
3. Which are the top-10 deals in the shark tank?
4. Top-3 Industries with highest deals in the shark tank?
5. Which are the top-5 cities with the maximum number of entrepreneurs?
6. Which are the top-3 states that has got maximum number of deals?
7. Industry wise count the total number of startups who pitched in shark tank
8. Count the number of pitchers who are male, female and belongs to mixed team
9. Find the maximum amount requested by a pitcher in each industrial segment
10. Find the maximum equity received by a shark in each industrial segment
11. Find the total amount invested by each shark throughout shark tank
12. Name the startups with Top-15 investments
13. Find number of deals having [1,2,3,4,5] sharks included in the deal
14. Which are the top-3 industries where "Kevin O Leary" is more interested to invest into?
15. Which are the 3 least favoured industries by the sharks?

Required Functions:-

```
# Black Box

Solution = function(vector,num){
  freq_table = table(vector)
  ranked_vect = names(sort(freq_table,decreasing = TRUE))
  ranked_vect = ranked_vect[nzchar(ranked_vect)]
  print(ranked_vect[1:num])
}
```

Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
# Vector Sum

Vector_Sum = function(vector){
  Sum = 0

  for(i in 1:length(vector)){
    Sum = Sum + vector[i]
  }

  return (Sum)
}
```

```
A$Total.Deal.Amount[is.na(A$Total.Deal.Amount)] = 0
Season = A$Season.Number[which(A$Total.Deal.Amount == max(A$Total.Deal.Amount))]
```

```
season_unique = unique(A$Season.Number)
price_Season = c()

for(sn in season_unique){
  price = c()
  i = 1
  while(i<NROW(A)){
    if(A$Season.Number[i] == sn){
      price = append(price,A$Total.Deal.Amount[i])
    }
    i = i + 1
  }
  price[is.na(price)] = 0
  price_Season = append(price_Season,max(price))
}

Episodes = c()

Final_Episode = c()
for(i in 1:10){
  Episodes = append(Episodes,A$Episode.Number[which(price_Season[i] == A$Total.Deal.Amount)][[1]])
}
```

```
Name = c()
Highest_deal_Season = sort(A$Total.Deal.Amount,decreasing = TRUE)
Highest_deal_Season = unique(Highest_deal_Season)
Highest_deal_Season = Highest_deal_Season[1:10]

for( i in 1:10){
  Name = append(Name,A$Startup.Name[which(A$Total.Deal.Amount == Highest_deal_Season[i])][[1]])
}
```

Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
Industry = c()

for( i in 1:3){
  Industry = append(Industry,A$Industry[which(A$Total.Deal.Amount == Highest_deal_Season[i])][[1]])
}
```

```
Solution(A$Pitchers.City,5)
```

```
States = c()

for(i in 1:nrow(A)){
  if(A$Got.Deal[i] == 1){
    States = append(States,A$Pitchers.State[i])
  }
}

Solution(States,3)
```

```
Industry_Table = table(A$Industry)
Data_Frame = as.data.frame(Industry_Table)
Data_Frame <- Data_Frame[order(Data_Frame$Freq),]
print(Data_Frame)
```

```
Team_Table = table(A$Pitchers.Gender)
Team_Data = data.frame(Team_Table)
```

Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
Industry = unique(A$Industry)

Max_Amount = c()
A$Original.Ask.Amount[is.na(A$Original.Ask.Amount)] = 0

for(i in 1:length(Industry)){
  Industry_Amount = c()
  for(j in 1:nrow(A)){
    if((A$Industry[j] == Industry[i])){
      Industry_Amount = append(Industry_Amount,A$Original.Ask.Amount[j])
    }
  }
  Max_Amount = append(Max_Amount,max(Industry_Amount))
}

Amount_Industry = data.frame(Industry,Max_Amount)
```

```
Max_Equity = c()
A$Total.Deal.Equity[is.na(A$Total.Deal.Equity)] = 0

for(i in 1:length(Industry)){  #pick one industry
  Industry_Eqity = c()
  for(j in 1:nrow(A)){
    if((A$Industry[j] == Industry[i])){
      Industry_Eqity = append(Industry_Eqity,A$Total.Deal.Equity[j])
    }
  }
  Max_Equity = append(Max_Equity,max(Industry_Eqity))
}

Equity_Industry = data.frame(Industry,Max_Equity)
```

Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
# Barbara_Corcora
Barbara_Corcora_Amount = A$Barbara.Corcoran.Investment.Amount[!is.na(A$Barbara.Corcoran.Investment.Amount)]
Barbara_Corcoran = Vector_Sum(Barbara_Corcora_Amount)

# Mark_Cuban
Mark_Cuban_Amount = A$Mark.Cuban.Investment.Amount[!is.na(A$Mark.Cuban.Investment.Amount)]
Mark_Cuban = Vector_Sum(Mark_Cuban_Amount)

# Lori Greine

Lori_Greine_Amount = A$Lori.Greiner.Investment.Amount[!is.na(A$Lori.Greiner.Investment.Amount)]
Lori_Greine = Vector_Sum(Lori_Greine_Amount)

# Robert Herjavec

Robert_Herjavec_Amount = A$Robert.Herjavec.Investment.Amount[!is.na(A$Robert.Herjavec.Investment.Amount)]
Robert_Herjavec = Vector_Sum(Robert_Herjavec_Amount)

# Daymond John

Daymond_John_Amount = A$Daymond.John.Investment.Amount[!is.na(A$Daymond.John.Investment.Amount)]
Daymond_John = Vector_Sum(Daymond_John_Amount)

# Kevin O Leary

Kevin_O_Leary_Amount = A$Kevin.O.Leary.Investment.Amount[!is.na(A$Kevin.O.Leary.Investment.Amount)]
Kevin_O_Leary = Vector_Sum(Kevin_O_Leary_Amount)
```

```
Data = data.frame(Shark_Name = c("Barbara Corcora","Mark Cuban","Lori Greine","Robert Herjavec","Daymond John","Kevin O Leary"),
                  c(Barbara_Corcoran,Mark_Cuban,Lori_Greine,Robert_Herjavec,Daymond_John,Kevin_O_Leary))
```

```
Name_1 = c()
Highest_deal_Season_1 = sort(A$Total.Deal.Amount,decreasing = TRUE)
Highest_deal_Season_1 = unique(Highest_deal_Season_1)
Highest_deal_Season_1 = Highest_deal_Season_1[1:15]

for( i in 1:15){
  Name_1 = append(Name_1,A$Startup.Name[which(A$Total.Deal.Amount == Highest_deal_Season_1[i])][[1]])
}
```

Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
Sharks_No = c(0,0,0,0,0)

for(i in 1:length(Sharks)){
  if(Sharks[i] == 1){
    Sharks_No[1] = Sharks_No[1]  + 1
  } else if(Sharks[i] == 2){
    Sharks_No[2] = Sharks_No[2]  + 1
  } else if(Sharks[i] == 3){
    Sharks_No[3] = Sharks_No[3]  + 1
  } else if(Sharks[i] == 4){
    Sharks_No[4] = Sharks_No[4]  + 1
  } else if(Sharks[i] == 5){
    Sharks_No[5] = Sharks_No[5]  + 1
  }
}

Frame = data.frame(No_Of_Sharks = 1:5,No_Of_Start_Up = Sharks_No)
```

```
A$Kevin.O.Leary.Investment.Amount[is.na(A$Kevin.O.Leary.Investment.Amount)] = 0
Industry_Kevin = c()

for(i in 1:length(A$Kevin.O.Leary.Investment.Amount)){
  if(A$Kevin.O.Leary.Investment.Amount[i] > 0){
    Industry_Kevin = append(Industry_Kevin,A$Industry[i])
  }
}

Solution(Industry_Kevin,3)
```

```
Industry = A$Industry[A$Total.Deal.Amount != 0]

Least_Industry = table(Industry)
Ranked_Least_Industry = names(sort(Least_Industry,decreasing = FALSE))
Ranked_Least_Industry = Ranked_Least_Industry[-1]
```

Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

Output:

```
> A$Total.Deal.Amount[is.na(A$Total.Deal.Amount)] = 0
> Season = A$Season.Number[which(A$Total.Deal.Amount == max(A$Total.Deal.Amount))]
> Season
[1] 6
>
```

```
> Episodes
 [1] 10  8  2 16  6 27  6  6  8 28
>
```

```
> Name
 [1] "AirCar"                      "SynDaverLabs"                "Zipz"
 [4] "TenThirtyOneHauntedHayrides" "RuggedManiac"                "XCraft"
 [7] "Trunkster"                   "HyConn"                      "TheRedDressBoutique"
[10] "HDYRSushiBars"
```

```
> Industry
[1] "Travel"            "Health/Wellness"   "Food and Beverage"
```

```
> Solution(A$Pitchers.City,5)
[1] "Los Angeles"   "New York"      "San Francisco" "Chicago"       "Austin"
```

```
> Solution(States,3)
[1] "CA" "TX" "FL"
```

Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
                     Var1 Freq
1              Automotive   12
2       Business Services   17
3      Children/Education   93
4          Fashion/Beauty  167
5  Fitness/Sports/Outdoors   93
6       Food and Beverage  182
7         Green/CleanTech   11
8         Health/Wellness   36
9          Lifestyle/Home  140
10    Media/Entertainment   23
11           Pet Products   41
12          Software/Tech   60
13                 Travel   11
14        Uncertain/Other    9
```

```
> Team_Data
        Var1 Freq
1     Female  221
2       Male  535
3 Mixed Team  139
>
```

```
> Amount_Industry
                   Industry Max_Amount
1          Health/Wellness    3000000
2        Food and Beverage    2500000
3        Business Services    1200000
4           Lifestyle/Home    2000000
5            Software/Tech    2000000
6       Children/Education    1500000
7                Automotive     500000
8          Fashion/Beauty    2500000
9      Media/Entertainment    5000000
10 Fitness/Sports/Outdoors    3000000
11             Pet Products     750000
12        Green/CleanTech    2000000
13                 Travel    5000000
14        Uncertain/Other     640000
```

Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
> Equity_Industry
                  Industry Max_Equity
1          Health/Wellness         55
2        Food and Beverage        100
3        Business Services         50
4           Lifestyle/Home        100
5            Software/Tech         50
6       Children/Education        100
7               Automotive        100
8           Fashion/Beauty        100
9      Media/Entertainment        100
10 Fitness/Sports/Outdoors         70
11            Pet Products         60
12         Green/CleanTech         35
13                  Travel         50
14         Uncertain/Other        100
>
```

```
> Data
       Shark_Name
1 Barbara Corcora
2      Mark Cuban
3     Lori Greine
4 Robert Herjavec
5    Daymond John
6   Kevin O Leary
  c.Barbara_Corcoran..Mark_Cuban..Lori_Greine..Robert_Herjavec..
1                                                      11740000
2                                                      37686667
3                                                      24525000
4                                                      25378166
5                                                      15809000
6                                                      17265833
>
```

Marwadi University	Marwadi University Faculty of Technology Department of Information and Communication Technology	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

```
> Name_1
 [1] "AirCar"                    "SynDaverLabs"
 [3] "Zipz"                      "TenThirtyOneHauntedHayrides"
 [5] "RuggedManiac"              "XCraft"
 [7] "Trunkster"                 "HyConn"
 [9] "TheRedDressBoutique"       "HDYRSushiBars"
[11] "sunscreeenr"               "FirstDefenseNasalScreen"
[13] "ZinePak"                   "Fixed"
[15] "EmazingLights"
>
```

```
> Frame
  No_Of_Sharks No_Of_Start_Up
1            1            360
2            2            111
3            3             13
4            4              1
5            5              5
>
```

```
> Solution(Industry_Kevin,3)
[1] "Lifestyle/Home"     "Food and Beverage"  "Children/Education"
>
```

```
> Ranked_Least_Industry
 [1] "Travel"                 "Business Services"
 [3] "Green/CleanTech"        "Automotive"
 [5] "Media/Entertainment"    "Health/Wellness"
 [7] "Pet Products"           "Software/Tech"
 [9] "Children/Education"     "Fitness/Sports/Outdoors"
[11] "Lifestyle/Home"         "Fashion/Beauty"
[13] "Food and Beverage"
>
```

Marwadi University	**Marwadi University** **Faculty of Technology** **Department of Information and Communication Technology**	
Subject: Introduction to R and R Studio (01CT0106)	**Aim: Case Study: Analysis of the Shark Tank US dataset**	
Experiment: 15	**Date: 04/05/2023**	**Enrollment No: 92200133030**

Observation and Learnings:

1. Explain the dataset

2. Write your answer and your inference for each of the questions mentioned in “Program” section