```r
# at 1st check what is the directory we are working
getwd()

## [1] "C:/Users/SVMY/Documents/ICPC Project"
```

```r
# now we want to call data frame in our project
df=read.csv('icpc.csv')
head(df)

##   Year   Host World.Final.Ranking                                University
## 1 2020 Russia                   1           Nizhny Novgorod State University
## 2 2020 Russia                   2                  Seoul National University
## 3 2020 Russia                   3             St. Petersburg ITMO University
## 4 2020 Russia                   4 Moscow Institute of Physics and Technology
## 5 2020 Russia                   5                       University of Wroclaw
## 6 2020 Russia                   6                     University of Cambridge
##       Country Medal Percentage.Score num_of_top_100_last_20_year
## 1      Russia   Yes            80.00                           8
## 2 South Korea   Yes            73.33                          16
## 3      Russia   Yes            73.33                          21
## 4      Russia   Yes            73.33                          11
## 5      Poland   Yes            73.33                          12
## 6          UK   Yes            73.33                           5
```

```r
#How many years contained in this data set
Years=unique(df$Year)
Years

##  [1] 2020 2019 2018 2017 2016 2015 2014 2013 2012 2011
```

```r
#How many years contained in this data set
Years=unique(df$Year)
Years

##  [1] 2020 2019 2018 2017 2016 2015 2014 2013 2012 2011

# Host name for respected years
Host_names=c()
for(i in Years){
  Host_names=c(Host_names,unique(df[df$Year==i,'Host']))
}
Host_names

##  [1] "Russia"   "Portugal" "China"    "USA"      "Thailand" "Morocco"
##  [7] "Russia"   "Russia"   "Poland"   "USA"
```

```
for(i in 1:length(Years)){
  cat(Years[i],"=>",Host_names[i],"\n")
}
```

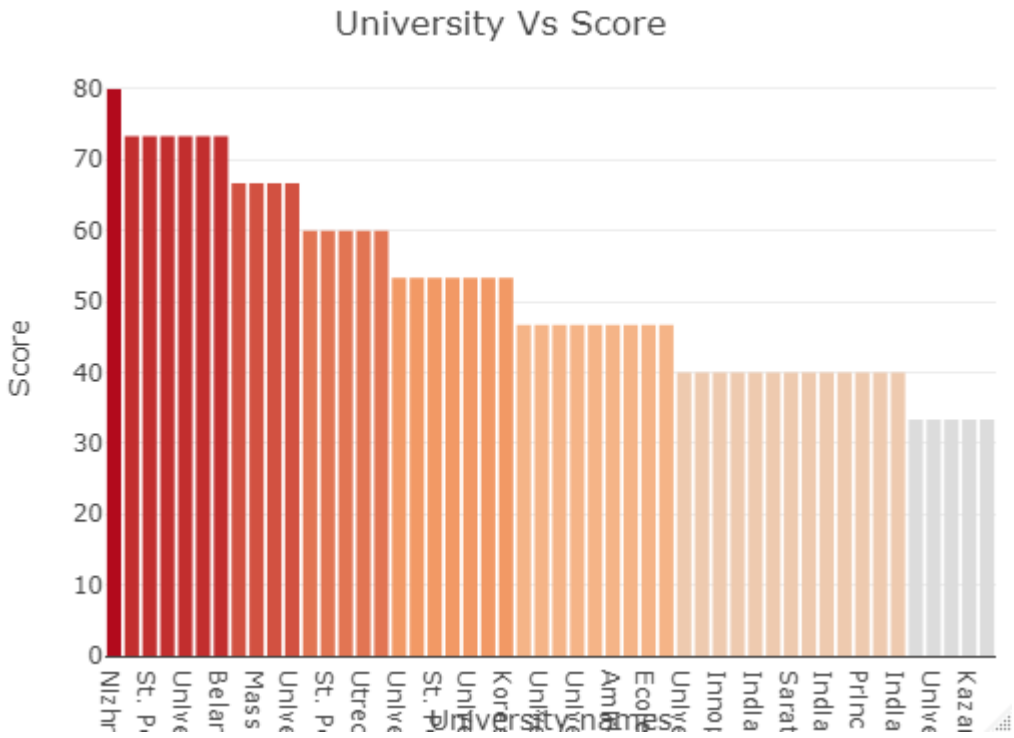
```
## 2020 => Russia
## 2019 => Portugal
## 2018 => China
## 2017 => USA
## 2016 => Thailand
## 2015 => Morocco
## 2014 => Russia
## 2013 => Russia
## 2012 => Poland
## 2011 => USA
```

```
#For 2020 and host country name is Russia
Ru=df[df$Year==2020,]
head(Ru)
```

```
##   Year   Host World.Final.Ranking                                University
## 1 2020 Russia                   1           Nizhny Novgorod State University
## 2 2020 Russia                   2                  Seoul National University
## 3 2020 Russia                   3              St. Petersburg ITMO University
## 4 2020 Russia                   4 Moscow Institute of Physics and Technology
## 5 2020 Russia                   5                       University of Wroclaw
## 6 2020 Russia                   6                     University of Cambridge
##       Country Medal Percentage.Score num_of_top_100_last_20_year
## 1      Russia   Yes            80.00                           8
## 2 South Korea   Yes            73.33                          16
## 3      Russia   Yes            73.33                          21
## 4      Russia   Yes            73.33                          11
## 5      Poland   Yes            73.33                          12
## 6          UK   Yes            73.33                           5
```

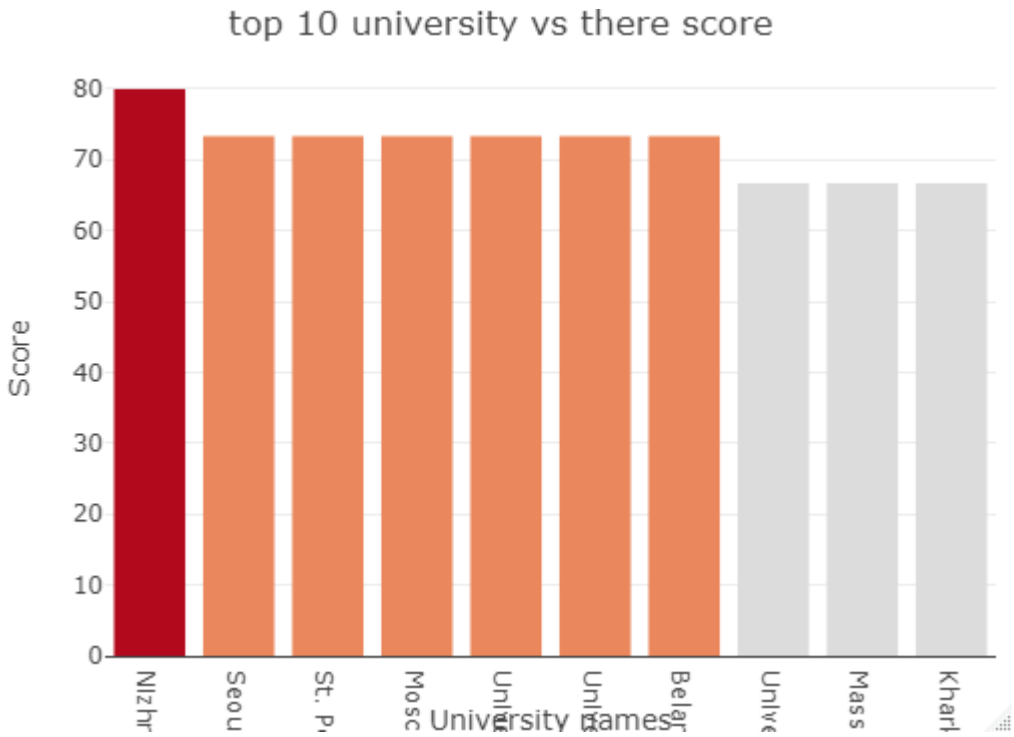
```
#bar diagram for university due to top rank
library(plotly)

un <- Ru$University
score <- Ru$Percentage.Score
data <- data.frame(un, score, stringsAsFactors = FALSE)
data$un <- factor(data$un, levels = unique(data$un)[order(data$score, decreasing = TRUE)])
fig=plot_ly(data, x = ~un, y = ~score, type = "bar",
        marker = list(color = ~score))
fig = fig %>% layout(title = "University Vs Score",
```

```
                      xaxis = list(title = "University names"),
                      yaxis = list(title = "Score"))
fig
```



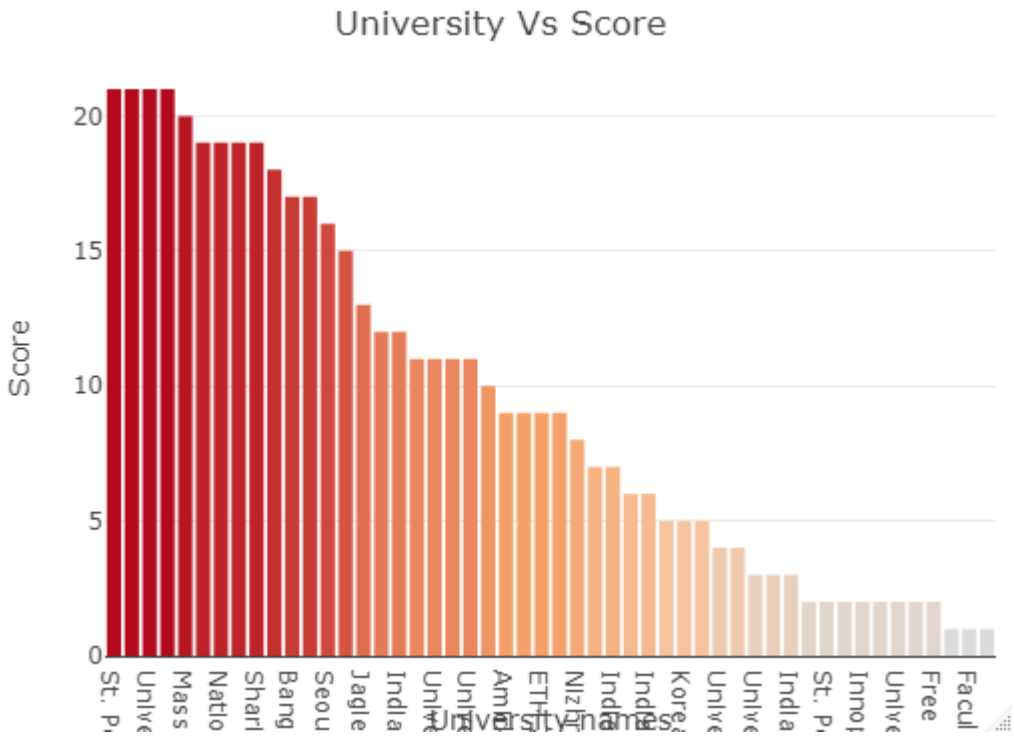
View

```
#Top 10 bar diagram
data10=head(data,10)
fig=plot_ly(data10,x=~un,y=~score,type = "bar",
            marker = list(color = ~score))
fig = fig %>% layout(title = "top 10 university vs there score",
                     xaxis = list(title = "University names"),
                     yaxis = list(title = "Score"))
fig
```



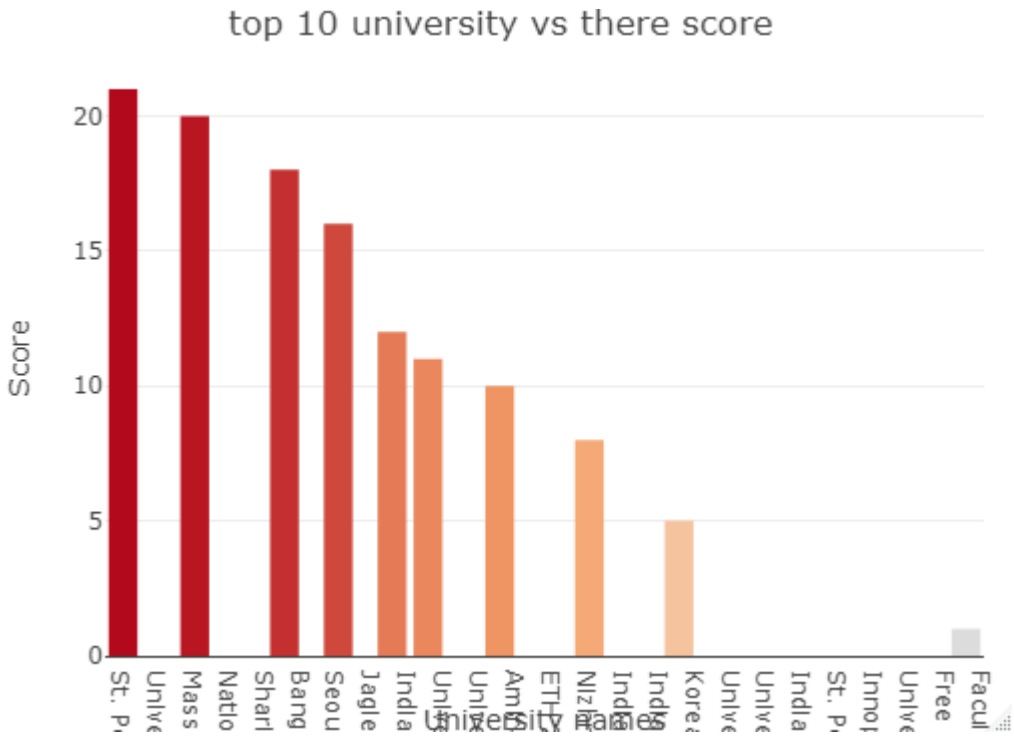
View

```
# bar diagram for top 100 university for 20 years
un <- Ru$University
score <- Ru$num_of_top_100_last_20_year
data100 <- data.frame(un, score, stringsAsFactors = FALSE)
data100$un <- factor(data100$un,
              levels = unique(data100$un)[order(data100$score, decreasing = TRUE)])
fig=plot_ly(data100, x = ~un, y = ~score, type = "bar",
            marker = list(color = ~score))
fig = fig %>% layout(title = "University Vs Score",
                     xaxis = list(title = "University names"),
                     yaxis = list(title = "Score"))
fig
```



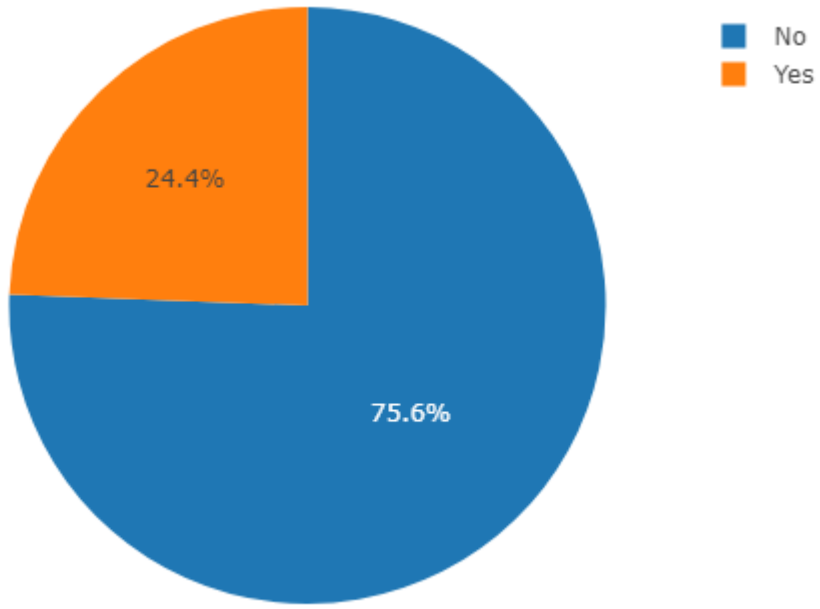
View

```
#Top 10 bar diagram for top 100 university for 20 years
data10010=head(data100,10)
fig=plot_ly(data10010,x=~un,y=~score,type = "bar",
            marker = list(color = ~score))
fig = fig %>% layout(title = "top 10 university vs there score",
                     xaxis = list(title = "University names"),
                     yaxis = list(title = "Score"))
fig
```



View

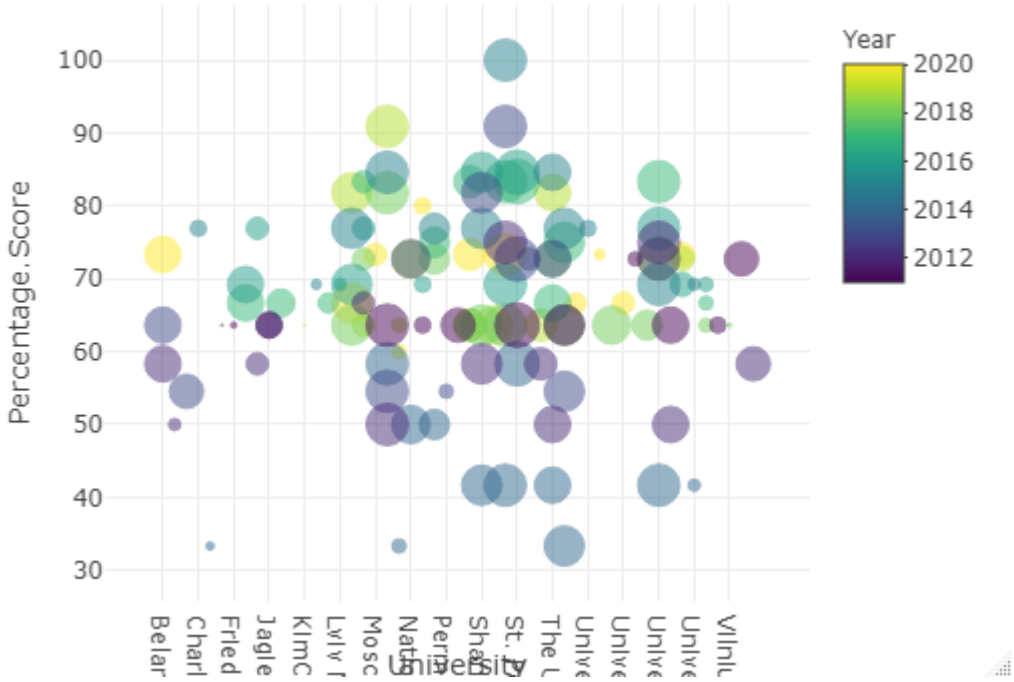
```
#Percent medal got university
fig <- plot_ly(df, labels = ~Medal, type = 'pie')
fig
```



View

```
#a bubble chart for that university which universities got the prize
Medal=df[df$Medal=='Yes',]
```

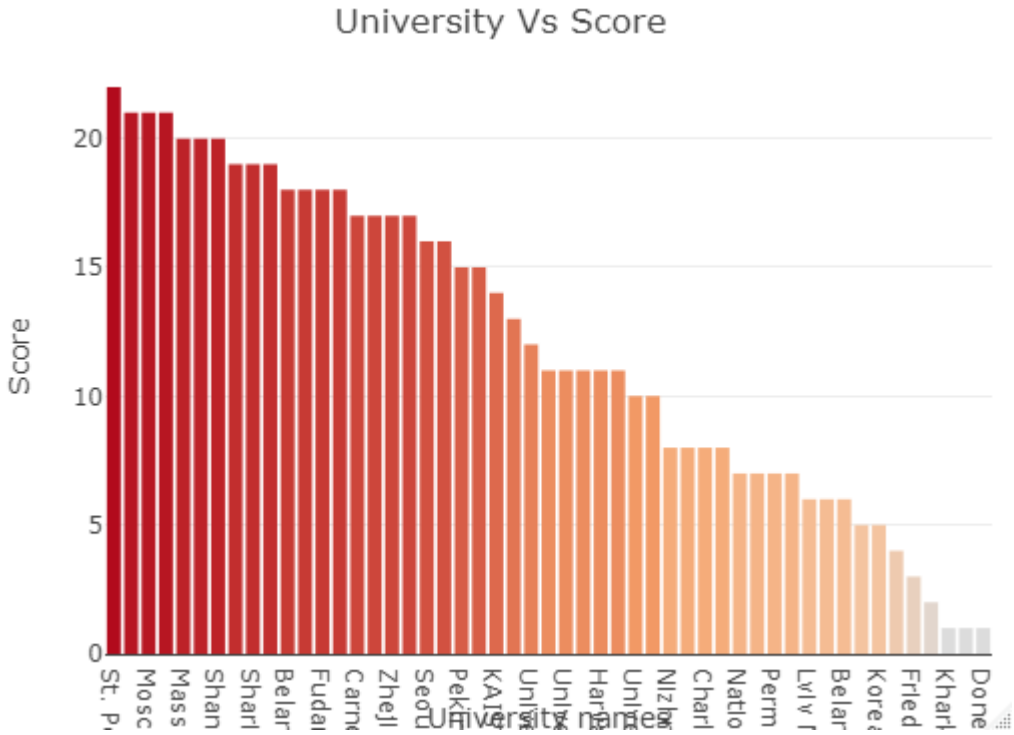
```
fig=plot_ly(Medal,x=~University,y=~Percentage.Score,
            type = "scatter",color = ~Year,
            mode="markers",
            marker = list(size = ~num_of_top_100_last_20_year, opacity = 0.5),
            hoverinfo="text",
            text=paste(Medal$University,"<br>",
                       Medal$Country,"<br>","score :",
                       Medal$Percentage.Score,"<br>",
                       "num of top 100 last 20 year",Medal$num_of_top_100_last_20_year))
fig
```



View

```
flt=Medal[,c('University',"num_of_top_100_last_20_year")]
flt=unique(flt)
```

```
# bar plot for sumvalue data frame
un <- flt$University
score <- flt$num_of_top_100_last_20_year
data <- data.frame(un, score, stringsAsFactors = FALSE)
data$un <- factor(data$un, levels = unique(data$un)[order(data$score, decreasing = TRUE)])
fig=plot_ly(data, x = ~un, y = ~score, type = "bar",
            marker = list(color = ~score))
fig = fig %>% layout(title = "University Vs Score",
                     xaxis = list(title = "University names"),
                     yaxis = list(title = "Score"))
fig
```



View

```
# We will able to predict the next champion and top most team
flt[flt$num_of_top_100_last_20_year==max(flt$num_of_top_100_last_20_year),"University"]

## [1] "St. Petersburg State University"
```

```
#another way
columns <- c("Year","Host","World.Final.Ranking","University","Country","Medal","Percentage.
myData = data.frame(matrix(nrow = 0, ncol = length(columns)))
colnames(myData) = columns
years_name=unique(df$Year)
for( i in 1:length(years_name)){
  y=head(df[df$Year==years_name[i],],5)
  myData=merge(y,myData,by=colnames(myData),all = TRUE)
}
head(myData)

##   Year   Host World.Final.Ranking                          University Country
## 1 2011    USA                   1                 Zhejiang University   China
## 2 2011    USA                   2 University of Michigan at Ann Arbor     USA
## 3 2011    USA                   3                 Tsinghua University   China
```

```
## 4 2011    USA                  4     St. Petersburg State University  Russia
## 5 2011    USA                  5    Nizhny Novgorod State University  Russia
## 6 2012 Poland                  1      St. Petersburg ITMO University  Russia
##   Medal Percentage.Score num_of_top_100_last_20_year
## 1   Yes            72.73                         17
## 2   Yes            72.73                          7
## 3   Yes            63.64                         20
## 4   Yes            63.64                         22
## 5   Yes            63.64                          8
## 6   Yes            75.00                         21
```

```
fig=plot_ly(myData,x=~World.Final.Ranking,y=~Percentage.Score,type = 'bar', frame = ~Year,
            hoverinfo="text",
            text=paste(myData$University,"<br>Rank:",myData$World.Final.Ranking,"<br>Country
fig <- fig %>%
  animation_button(
    x = 1, xanchor = "right", y = 0, yanchor = "bottom"
  )
fig <- fig %>%
  animation_slider(
    currentvalue = list(prefix = "YEAR ", font = list(color="red"))
  )
fig
```

View

```
#most frequency University in top 5
tt <- table(myData$University)
names(tt[tt==max(tt)])

## [1] "St. Petersburg ITMO University"  "St. Petersburg State University"
## [3] "University of Warsaw"

moscow=myData[myData$University=="St. Petersburg State University" ,]
moscow

##    Year     Host World.Final.Ranking                    University Country
## 4  2011      USA                   4 St. Petersburg State University  Russia
## 15 2013   Russia                   5 St. Petersburg State University  Russia
## 16 2014   Russia                   1 St. Petersburg State University  Russia
## 26 2016 Thailand                   1 St. Petersburg State University  Russia
## 34 2017      USA                   4 St. Petersburg State University  Russia
##    Medal Percentage.Score num_of_top_100_last_20_year
## 4    Yes            63.64                         22
## 15   Yes            72.73                         22
## 16   Yes            58.33                         22
## 26   Yes            84.62                         22
## 34   Yes            83.33                         22
```