Synopsis

Problem Statement

The tennis players that are widely regarded as being the greatest of all time are Rafeal Nadal , Roger Federer and Novak Djocovik .The fact that these players have won 60 Grand slams in total among themselves in the last 15 years speaks volumes about their complete domination of the Tennis World in the current era. Never before in modern history, any sport has so completely been overshadowed by just 2-3 players as has been done by these three in tennis.

Though there is no doubt about these three being the greatest, however, the debate of the greatest amongst these three is still not settled with the jury being still out on this. What makes it even more interesting is the fact that each of them has won 20 grand slams each, thus picking one out of them is even more difficult. Though statistics do not always give the complete picture when comparing the performance of players, that is the only objective way of identifying the greatest when it comes to tennis players.

Hence the problem statement is “To figure out the Greatest Tennis Player of All Time by means of statistical analysis and data management (through graphs and charts.)”

Objective and Scope

The objective or aim of this project is to highlight the contribution and domination of three of the greatest sportsmen of this era in the World of tennis and to try and figure out the better of these three by comparing their performance in the Grand slams by means of statistical analysis.

We, by means of this project, have tried to use the features of the python pyplot to bring out the individual and comparable performance of these players by means of bar graphs and pie charts in different grand slams along with the performance of the rest of the tennis players.

Hardware and software requirement

Hardware Required:

* A Personal Computer
* Processor : intel i3 or above
* RAM : 2 GB or above
* Internet Connectivity
* Printer

Software Required:

* Operating System : Windows 10
* Coding Language  : Python
* Python Pandas Library
* Python Matplotlib Library
* Python Pyplot module

Pyplot functions used

The statistical analysis of the records set by the tennis players has been done using bar graphs , line charts , histograms etc. These graphs have been created by using the various functions that the pyplot module of the python matplotlib library has provided us . The pyplot functions used in the project are as follows:

* plt.plot()
* plt.bar()
* plt.xlabel()
* plt.ylabel()
* plt.title()
* plt.xlim()
* plt.ylim()
* plt.barh()

**Contribution in Real World**

This project can have many applications in the real world, some of them are given below:

1. As this project is an analysis of the statistics of tennis players it will be of help to the writers of sports magazines who can use it to predict which player will win/lose a certain match in certain conditions using the record of the previous matches played by that player in those conditions.
2. This project also provides tennis fans with the necessary statistical information about their favourite tennis players in the most visually appealing way, helping them debate among themselves regarding who is the best tennis player.

**CODES**

import pandas as pd

import matplotlib.pyplot as plt

import numpy as np

def menu():

print("-----------------------------------------Tennis Statistical Analysis----------------------------------------------")

print('\n')

print("Choose the player you want to see the analysis of:")

print("1.Roger Federer")

print("2.Rafael Nadal")

print("3.Novak Djokovic")

print("4.Combined Analysis of all three")

ch=int(input("Enter your choice[1-4]="))

if ch==1:

while True:

print("----------------------------------------------Roger Federer Tennis Statistical Analysis--------------------------------------------------")

print("1.Hard Courts")

print("2.Clay Courts")

print("3.Grass Courts")

print("4.Return to main menu")

ch1 = int(input("Enter your choice[1-4]="))

print("\n")

if ch1 ==1:

print("a.Display Bar chart")

print("b.Display Line Chart")

print("c.Return to menu")

ch2 = input("Enter your choice[a-c]=")

if ch2 == "a":

n1 = ["R16","R32","QF","SF","F","W"]

n2 = [8,4,3,11,3,11]

plt.bar(n1,n2,color = "red" , width = 0.5)

plt.xlabel("Won/Round Evicted")

plt.ylabel("No of times")

plt.title("Result on Hard Courts")

plt.show()

if ch2 == "b":

n1 = [2007, 2008 ,2009, 2010,2011 ,2012 ,2013, 2014 ,2015, 2016 ,2017, 2018 ,2019 ,2020, 2021]

n2 = [88.0,77.3, 78.3,87.0,86.8,85.4,71.8,88.7,86.7,80.0,90.9,81.8,81.6,83.3 ,50.0]

plt.plot(n1,n2,"r+" , linestyle = "solid" , markeredgecolor = "b",markersize = 5)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Hard Courts year-wise")

plt.show()

if ch2 == "c":

menu()

if ch1 == 2:

print("a.Display Horizontal Bar chart")

print("b.Display Line Chart")

print("c.Return to menu")

ch2 = input("Enter your choice[a-c]=")

if ch2 == "a":

n1 = ["R128","R16","R32","QF","SF","F","W"]

n2 = [3,3,1,4,3,4,1]

plt.barh(n1,n2,color = "blue" , height = 0.5)

plt.xlabel("Won/Round Evicted")

plt.ylabel("No of times")

plt.title("Result on Clay Courts")

plt.show()

if ch2 == "b":

n1 = [2007,2008,2009, 2010,2011,2012 ,2013,2014 ,2015, 2016,2019,2021]

n2 = [84.2 ,84.0,90.0,71.4,75.0 ,83.3 ,70.6 ,66.7 ,76.5 ,60.0,81.8,75.0]

plt.plot(n1,n2,"r+" , linestyle = "solid" , markeredgecolor = "y",markersize = 5)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Clay Courts year-wise")

plt.show()

if ch2 == "c":

menu()

if ch1 ==3:

print("a.Display Bar chart")

print("b.Display Line Chart")

print("c.Return to menu")

ch2 = input("Enter your choice[a-c]=")

if ch2 == "a":

n1 = ["R128","R64","QF","SF","F","W"]

n2 = [3,1,5,1,4,8]

plt.bar(n1,n2,color = "yellow" , width = 0.5)

plt.xlabel("Won/Round Evicted")

plt.ylabel("No of times")

plt.title("Result on Grass Courts")

plt.show()

if ch2 == "b":

n1 = [2007, 2008 ,2009, 2010,2011 ,2012 ,2013, 2014 ,2015, 2016 ,2017, 2018 ,2019 , 2021]

n2 = [100.0,91.7,100.0,80.0,85.7,88.2,83.3,90.0,91.7,76.9,92.3,85.7,91.7,71.4]

plt.plot(n1,n2,"b+" , linestyle = "solid" , markeredgecolor = "g",markersize = 5)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Grass Courts year-wise")

plt.show()

if ch2 == "c":

menu()

if ch1==4:

menu()

if ch==2:

while True:

print("-------------------------------------------------------------------Rafael Nadal Tennis Statistical Analysis--------------------------------------------------------------------------------")

print("1.Hard Courts")

print("2.Clay Courts")

print("3.Grass Courts")

print("4.Return to main menu")

ch3 = int(input("Enter your choice[1-4]="))

print("\n")

if ch3 == 1:

print("a.Display Horizontal Bar Chart")

print("b.Display Line Chart")

print("c.Return to menu")

ch4 = input("Enter your choice[a-c]=")

if ch4 =="a":

n3 = ["R128","R64","R32","R16","QF","SF","F","W"]

n4 = [1,2,3,3,8,4,5,5]

plt.barh(n3,n4, color = "green" , height = 0.5)

plt.xlabel("Won/Round Evicted")

plt.ylabel("No of times")

plt.title("Result on Hard Courts")

plt.show()

if ch4 =="b":

n3 =[2007, 2008 ,2009, 2010,2011 ,2012 ,2013, 2014 ,2015, 2016 ,2017, 2018 ,2019,2020 ,2021]

n4 =[72.1,82.1,77.8,81.6,75.0,85.0,90.0,76.9,71.4 ,64.3,81.6,87.5,91.2,75.0,71.4]

plt.plot(n3,n4,"r+" , linestyle = "solid" , markeredgecolor = "g",markersize = 5)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Hard Courts year-wise")

plt.show()

if ch4 =="c":

menu()

if ch3 ==2:

print("a.Display Pie Chart")

print("b.Display Line Chart")

print("c.Return to menu")

ch4 = input("Enter your choice[a-c]=")

if ch4 =="a":

n3 = [2007, 2008,2009,2010,2011,2012 ,2013,2014 ,2015,2016 ,2017,2018 ,2019,2020 ,2021]

n4 =[96.9,96.0,92.3,100.0,93.3,95.8,95.1,89.3,81.3,84.0,96.0,96.3,87.5,90.0,86.4]

plt.pie(n4,labels = n3 ,autopct = "%5.2f%%")

plt.title("Result on C Courts")

plt.show()

if ch4 == "b":

n3 = ["R32","R16","QF","SF","W"]

n4 = [1,1,1,1,13]

plt.plot(n3,n4 , "r+" , linestyle = "solid" , markeredgecolor = "black" ,markersize = 5)

plt.ylabel("years")

plt.xlabel("Round reached")

plt.title("Performance on Clay Courts year-wise")

plt.show()

if ch4 =="c":

menu()

if ch3 ==3:

print("a.Display Horizontal Bar Chart")

print("b.Display Pie Chart")

print("c.Return to menu")

ch4 = input("Enter your choice[a-c]=")

if ch4 =="a":

n3 = ["R128","R64","R32","R16","QF","SF","F","W"]

n4 = [1,3,1,2,0,2,3,2]

plt.barh(n3,n4,color ="magenta" , height = 0.7)

plt.xlabel("Round reached")

plt.ylabel("No of times reached")

plt.title("Result on Grass Courts")

plt.show()

if ch4 =="b":

n3 = [2007,2008,2010,2011,2012,2014,2015,2017,2018,2019]

n4 = [80.0,100.0,90.0,80.0,50.0,60.0,71.4,75.0,83.3,83.3]

plt.pie(n4 , labels = n3 , autopct = "%6.2f")

plt.title("Performance on Grass Courts year-wise")

plt.show()

if ch4 =="c":

menu()

if ch3 ==4:

menu()

if ch==3:

while True:

print("---------------------------------------------------------------------Novak Djokovic Tennis Statistical Analysis----------------------------------------------------------------")

print("1.Hard Courts")

print("2.Clay Courts")

print("3.Grass Courts")

print("4.Return to main menu")

ch3 = int(input("Enter your choice[1-4]="))

print("\n")

if ch3 ==1:

print("a.Display bar chart")

print("b.Display line chart")

print("c.Return to menu")

chh = input("Enter your choice:")

if chh == "a":

n5 = ["R128","R64","R32","R16","QF","SF","F","W"]

n6 = [2,1,2,4,3,3,6,12]

plt.bar(n5,n6,color = "cyan" , width = 0.8)

plt.xlabel("Won/Round Evicted")

plt.ylabel("No of times")

plt.title("Result on Hard Courts")

plt.show()

if chh == "b":

n5 = [2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021]

n6 = [78.2,78.2,82.8,78.2,90.2,90.9,91.4,87.0,92.2,88.7,80.0,83.8,81.4,88.2,88.2]

plt.plot(n5,n6,"b.", linestyle = "dashed" , markerfacecolor = "k" , markersize = 4 , linewidth = 2)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Hard Courts year-wise")

plt.show()

if chh == "c":

menu()

if ch3 ==2:

print("a.Display bar chart")

print("b.Display pie chart")

print("c.Return to menu")

chh = input("Enter your choice:")

if chh == "a":

n5 =["R64","R32","QF","SF","F","W"]

n6 =[1,1,4,5,4,2]

plt.bar(n5,n6 , color = "yellow" , width = 0.8)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Clay Courts year-wise")

plt.show()

if chh == "b":

n5 =[2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2020,2021]

n6 =[79.2,84.2,73.9,75.0,94.4,80.0,83.3,87.5,94.1,88.9,75.0,68.8,83.3,91.7,85.7]

plt.pie(n6 , labels = n5 , autopct = "%6.2f")

plt.title("Result on Clay Courts")

plt.show()

if chh == "c":

menu()

if ch3 ==3:

print("a.Display horizontal bar chart")

print("b.Display line chart")

print("c.Return to menu")

chh = input("Enter your choice:")

if chh == "a":

n5 =["R64","R32","R16","QF","SF","F","W"]

n6 =[1,2,1,2,3,1,6]

plt.barh(n5,n6,color = "green" , height = 0.6)

plt.ylabel("Won/Round Evicted")

plt.xlabel("No of times reached")

plt.title("Result on Grass Courts")

plt.show()

if chh == "b":

n5 =[2007,2008,2009,2010,2011,2012,2013,2014,2015,2016,2017,2018,2019,2021]

n6 =[75.0,71.4,80.0,75.0,100.0,75.0,85.7,100.0,100.0,66.7,88.9,91.7,100.0,100.0]

plt.plot(n5,n6, "g\*" , linestyle = "dashed" , markeredgecolor = "red" , markersize = 3 , linewidth = 3)

plt.xlabel("years")

plt.ylabel("Performance calculated in Percentage")

plt.title("Performance on Grass Courts year-wise")

plt.show()

if chh == "c":

menu()

if ch3 == 4:

menu()

if ch==4:

print

print("\n")

print("1.Display Multiple Bar Chart")

print("2.Return to Menu")

ch1 = int(input("Enter your Choice(1-2):"))

if ch1==1:

n1 =["Roger Federer" ,"Rafael Nadal" , "Novak Djokovic"]

n2 =[8,2,6]

n3 =[1,13,2]

n4 =[5,4,3]

n5 =[6,2,9]

X = np.arange(len(n1))

plt.bar(n1,n2,label = "Wimbledon" , color = "green",width = 0.15)

plt.bar(X+0.15,n3,label = "French Open" ,color = "red" ,width = 0.15)

plt.bar(X+0.30,n4,label = "US Open" , color = "blue", width = 0.15)

plt.bar(X+0.45,n5,label = "Australia Open", color = "yellow" , width = 0.15)

plt.legend(loc = "upper left")

plt.xlabel("Names of Players")

plt.ylabel("No of Wins")

plt.title("The Big 3 by Grand Slam Wins")

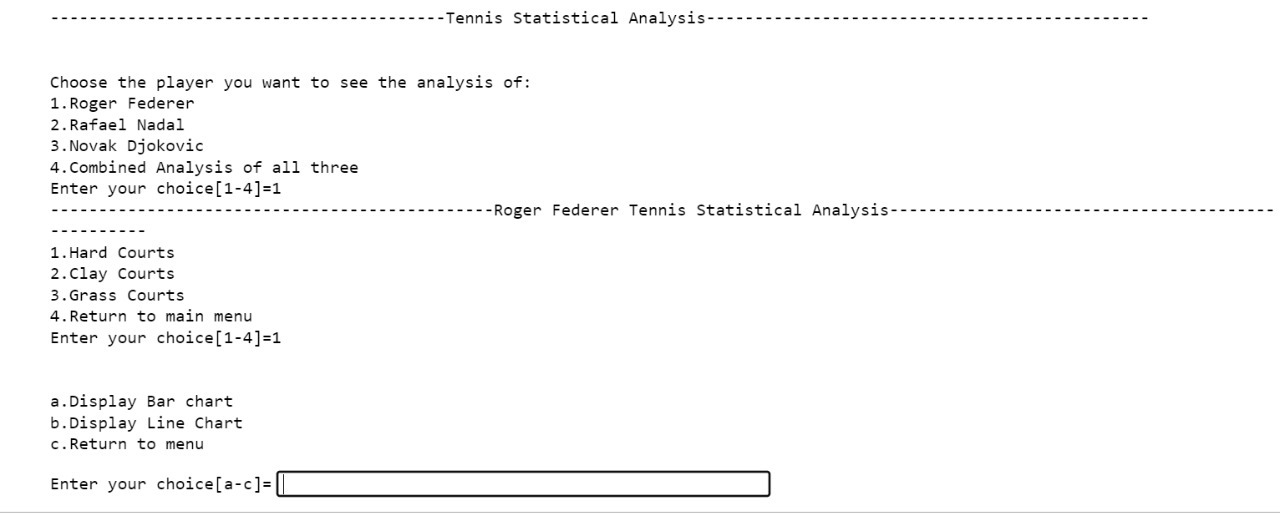
plt.show()

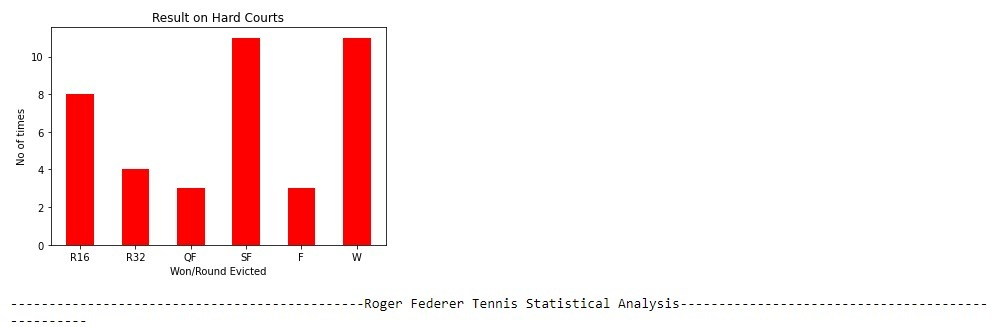
if ch==2:

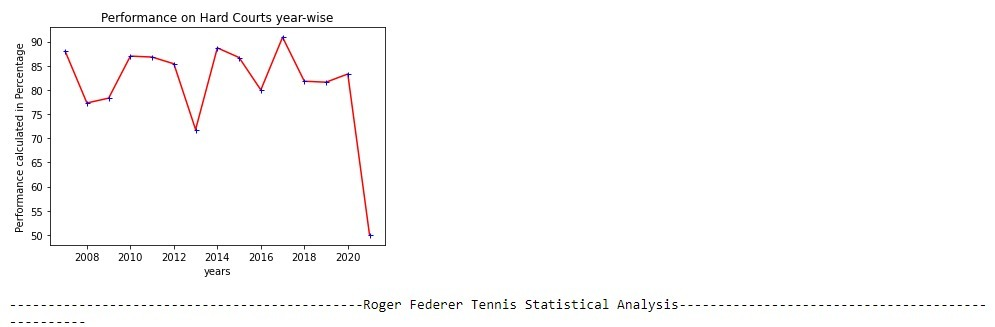
menu()

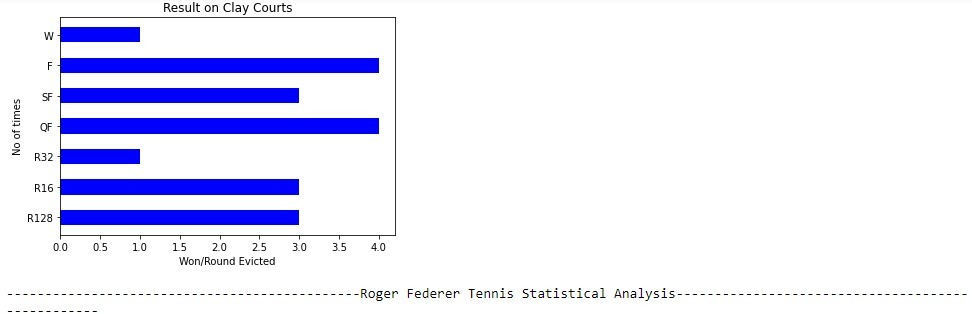
menu()

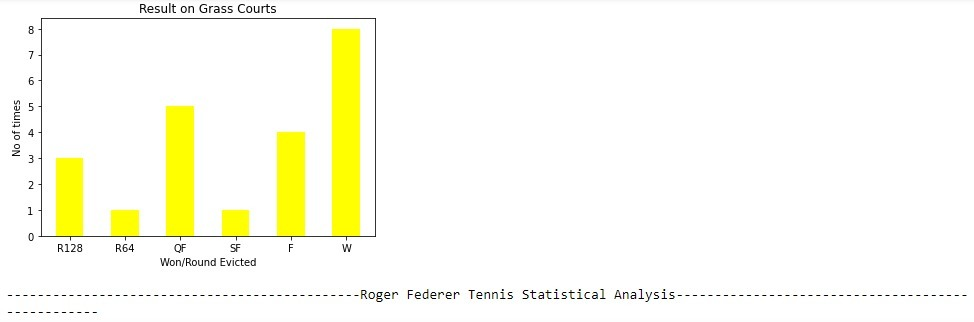
**SCREENSHOTS**

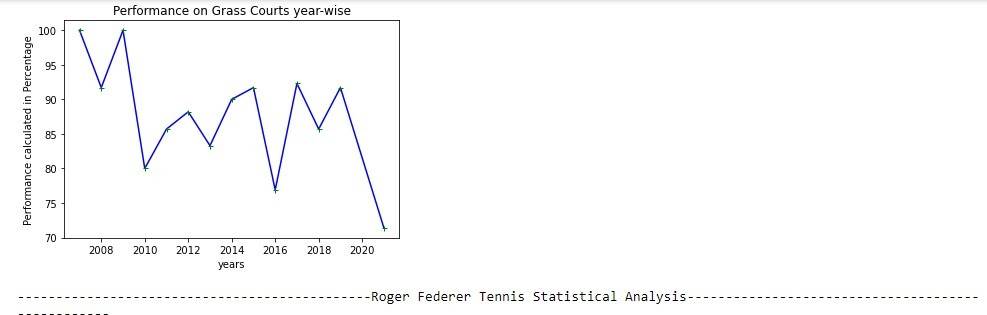


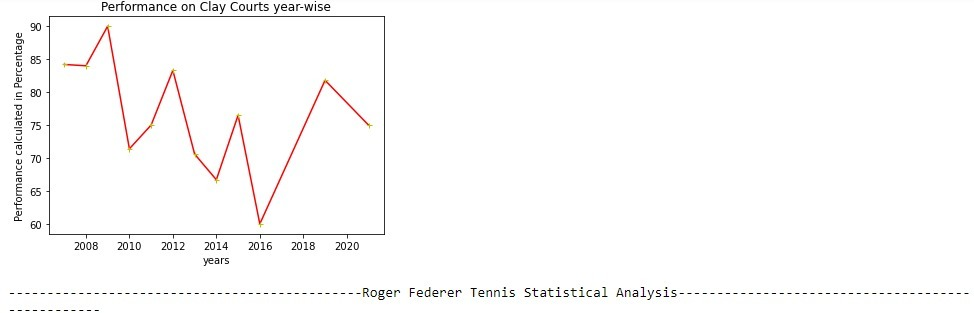


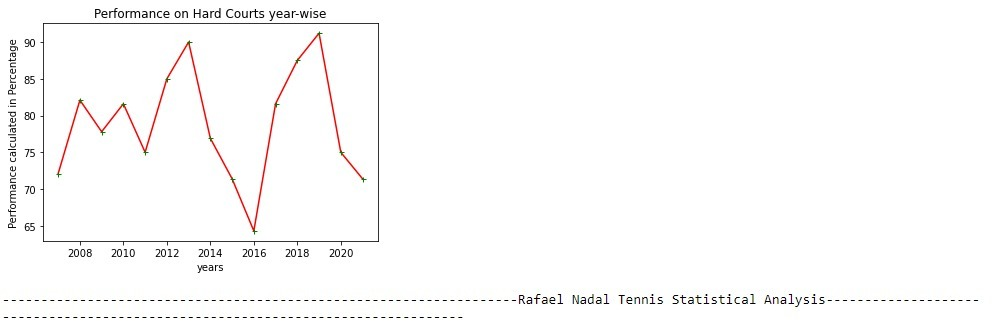


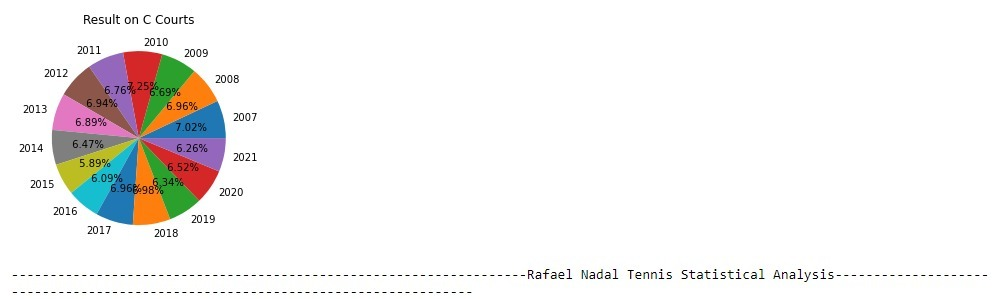


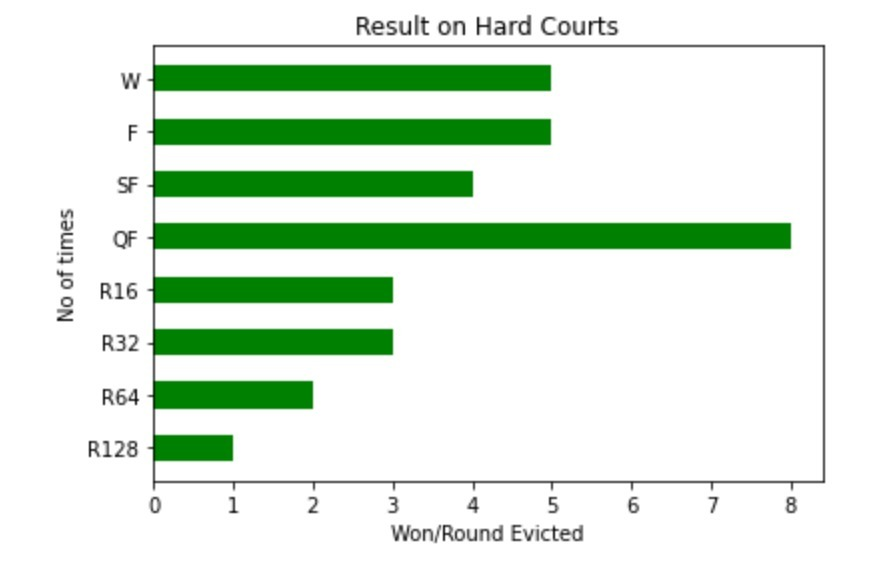


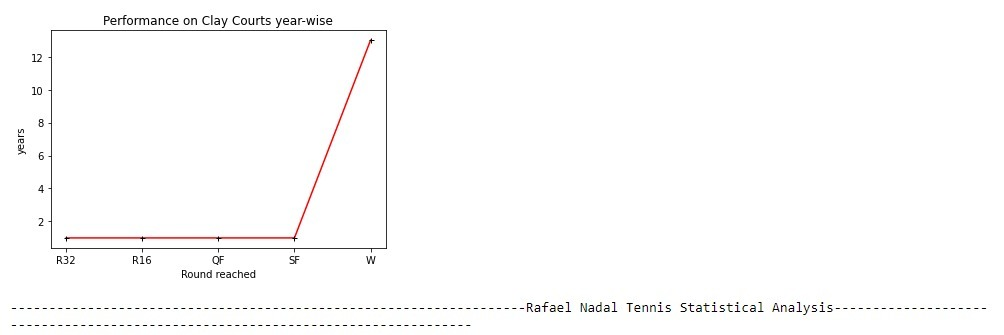


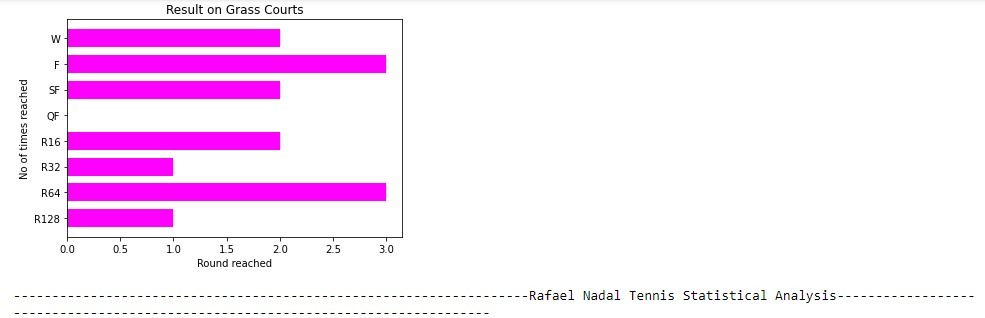


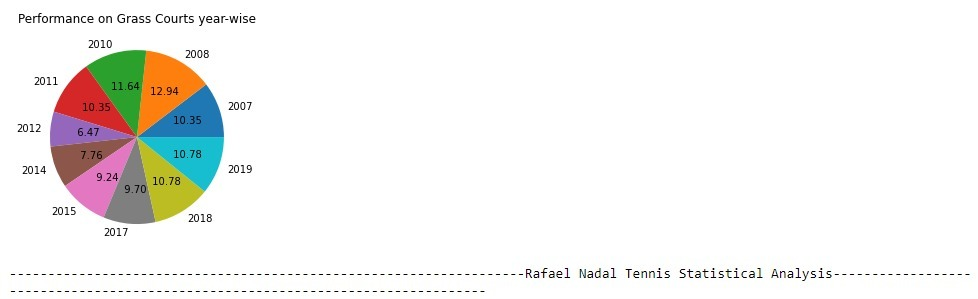


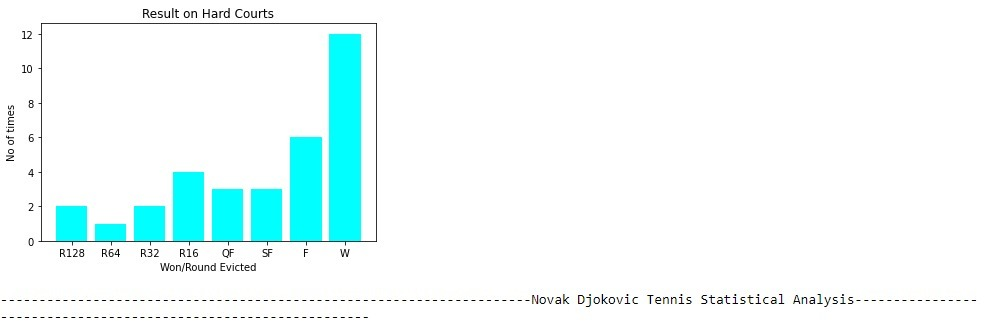


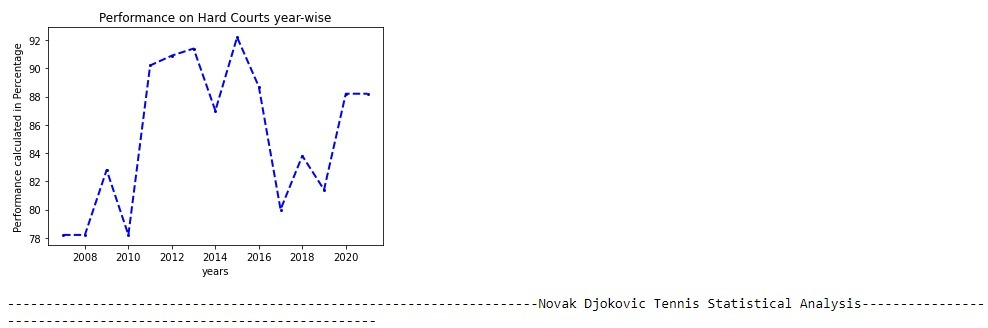


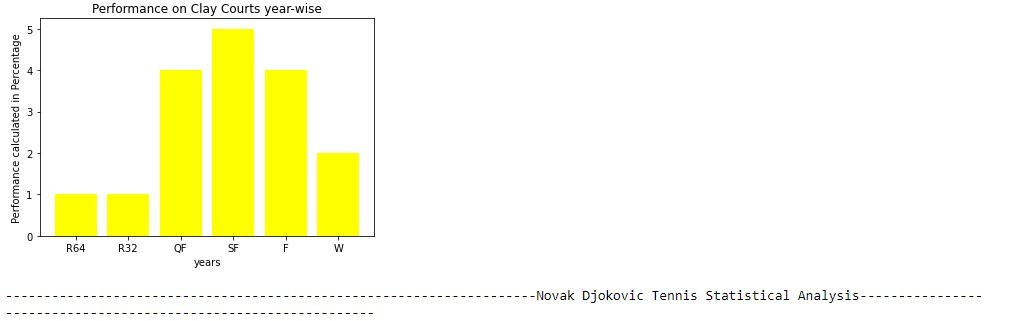


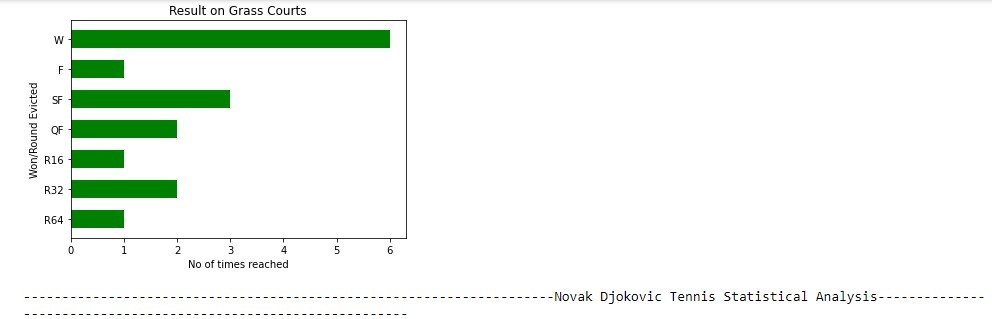


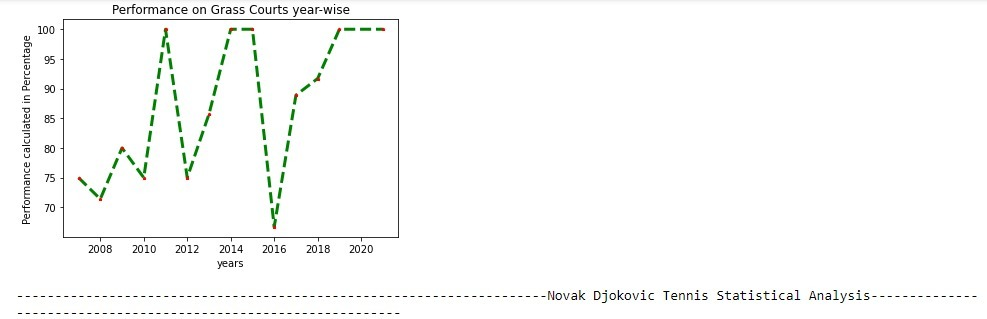


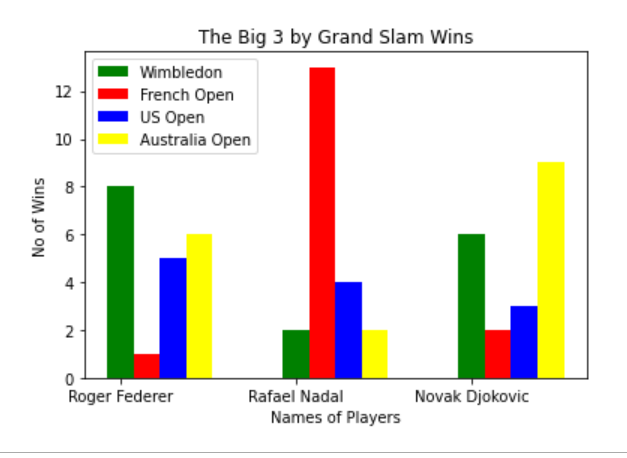












**FUTURE ENHANCEMENTS**

This project can be enhanced further in the future in the following ways:

* 1. This project could include the performance of the players in other tournaments as well like the Olympics, the Davis Cup, the Indian Wells Masters, the Miami Masters etc.
  2. This project could include other graphs and charts which depict the performance of the players using other criteria like their first serve won percentage, second serve won percentage, break points ratio etc
  3. We have created this project using the performance of the players from 2007 to 2021 so this project can always be enhanced and updated further in the future as new tournaments will keep happening and new records will keep getting created over the years.