# Summer Olympic Medals

ESSOUNAINI SOUFIANE

##Synopsis:

Using this dataset downloaded from Kaggle.com, we are trying to get a clearer idea of the evolution of different countries at the olympics especially our country Morocco throughout the years and also to see which are the countries who have dominated the olympics as far as the total number of medals harvested.

##1.Processing Data

First, we import packages which are going to help us in our analysis:

library(dplyr)

## Warning: package 'dplyr' was built under R version 3.6.3

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':  
##   
## filter, lag

## The following objects are masked from 'package:base':  
##   
## intersect, setdiff, setequal, union

library(ggplot2)

## Warning: package 'ggplot2' was built under R version 3.6.3

Then, we are going to read file in a table format:

data <-read.csv("Summer-Olympic-medals-1976-to-2008.csv")  
head(data)

## City Year Sport Discipline Event Athlete  
## 1 Montreal 1976 Aquatics Diving 3m springboard KÖHLER, Christa  
## 2 Montreal 1976 Aquatics Diving 3m springboard KOSENKOV, Aleksandr  
## 3 Montreal 1976 Aquatics Diving 3m springboard BOGGS, Philip George  
## 4 Montreal 1976 Aquatics Diving 3m springboard CAGNOTTO, Giorgio Franco  
## 5 Montreal 1976 Aquatics Diving 10m platform WILSON, Deborah Keplar  
## 6 Montreal 1976 Aquatics Diving 10m platform LOUGANIS, Gregory  
## Gender Country\_Code Country Event\_gender Medal  
## 1 Women GDR East Germany W Silver  
## 2 Men URS Soviet Union M Bronze  
## 3 Men USA United States M Gold  
## 4 Men ITA Italy M Silver  
## 5 Women USA United States W Bronze  
## 6 Men USA United States M Silver

##2.Top 20 gold medal winning countries:

data\_ordered\_by <- subset(data,data$Medal=="Gold")  
data\_ordered\_by$Medal="1"  
head(data\_ordered\_by)

## City Year Sport Discipline Event  
## 3 Montreal 1976 Aquatics Diving 3m springboard  
## 7 Montreal 1976 Aquatics Diving 10m platform  
## 9 Montreal 1976 Aquatics Diving 10m platform  
## 12 Montreal 1976 Aquatics Diving 3m springboard  
## 13 Montreal 1976 Aquatics Swimming 4x100m freestyle relay  
## 18 Montreal 1976 Aquatics Swimming 4x100m medley relay  
## Athlete Gender Country\_Code Country Event\_gender  
## 3 BOGGS, Philip George Men USA United States M  
## 7 VAYTSEKHOVSKAYA, Elena Women URS Soviet Union W  
## 9 DIBIASI, Klaus Men ITA Italy M  
## 12 CHANDLER, Jennifer Kay Women USA United States W  
## 13 BABASHOFF, Shirley Farber Women USA United States W  
## 18 ANKE, Hannelore Women GDR East Germany W  
## Medal  
## 3 1  
## 7 1  
## 9 1  
## 12 1  
## 13 1  
## 18 1

data\_ordered\_by$Medal<-as.numeric(data\_ordered\_by$Medal)

data3<-data\_ordered\_by %>%  
 group\_by(Country) %>%  
 summarise(somme\_Gold=sum(Medal),na.rm=TRUE) %>%  
 arrange(desc(somme\_Gold))%>%  
 select(Country,somme\_Gold)

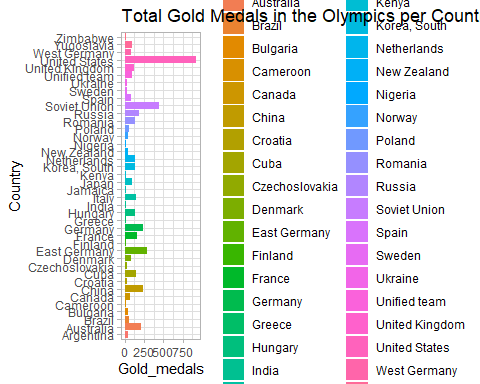
## `summarise()` ungrouping output (override with `.groups` argument)

x<-seq(1:87)  
  
data31<-cbind(x,data3)  
colnames(data31)<-c("Rank","Country","Gold\_medals")  
data31[1:20,]

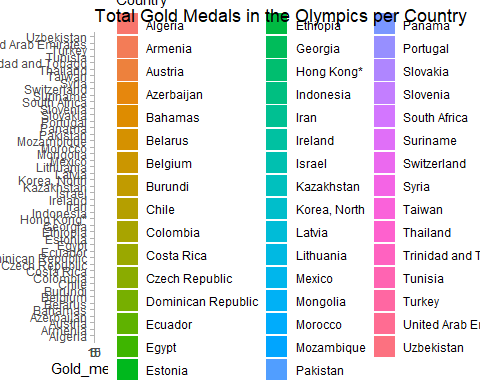
## Rank Country Gold\_medals  
## 1 1 United States 928  
## 2 2 Soviet Union 439  
## 3 3 East Germany 286  
## 4 4 Germany 237  
## 5 5 China 234  
## 6 6 Australia 216  
## 7 7 Russia 192  
## 8 8 France 154  
## 9 9 Cuba 152  
## 10 10 Italy 145  
## 11 11 Korea, South 140  
## 12 12 Netherlands 137  
## 13 13 Romania 135  
## 14 14 Hungary 129  
## 15 15 United Kingdom 122  
## 16 16 Japan 94  
## 17 17 Unified team 92  
## 18 18 Yugoslavia 90  
## 19 19 Spain 87  
## 20 20 West Germany 84

##3.Histograms showing each country’s total number of gold medals won from 1976 until 2008:

data4<-data31[1:40,]  
data5<-data31[41:87,]  
g<-ggplot(data4,aes(y=Country,x=Gold\_medals))  
g+geom\_bar(stat="identity",aes(fill=Country)) +theme\_light()+labs(title = "Total Gold Medals in the Olympics per Country",xlab="Sum of Gold Medals",ylab="Country")



g<-ggplot(data5,aes(y=Country,x=Gold\_medals))  
g+geom\_bar(stat="identity",aes(fill=Country)) +theme\_light()+labs(title = "Total Gold Medals in the Olympics per Country",xlab="Sum of Gold Medals",ylab="Country")



##4.Top silver medal winning countries:

data\_Silver <- subset(data,data$Medal=="Silver")  
data\_Silver$Medal="1"  
head(data\_Silver)

## City Year Sport Discipline Event  
## 1 Montreal 1976 Aquatics Diving 3m springboard  
## 4 Montreal 1976 Aquatics Diving 3m springboard  
## 6 Montreal 1976 Aquatics Diving 10m platform  
## 11 Montreal 1976 Aquatics Diving 10m platform  
## 14 Montreal 1976 Aquatics Swimming 400m freestyle  
## 15 Montreal 1976 Aquatics Swimming 4x100m freestyle relay  
## Athlete Gender Country\_Code Country Event\_gender  
## 1 KÖHLER, Christa Women GDR East Germany W  
## 4 CAGNOTTO, Giorgio Franco Men ITA Italy M  
## 6 LOUGANIS, Gregory Men USA United States M  
## 11 KNAPE-LINDBERGH, Ulrika Women SWE Sweden W  
## 14 SHAW, Timothy Andrew Men USA United States M  
## 15 PRIEMER, Petra Women GDR East Germany W  
## Medal  
## 1 1  
## 4 1  
## 6 1  
## 11 1  
## 14 1  
## 15 1

data\_Silver$Medal<-as.numeric(data\_Silver$Medal)  
class(data\_Silver$Medal)

## [1] "numeric"

data\_Sil<-data\_Silver %>%  
 group\_by(Country) %>%  
 summarise(somme\_Silver=sum(Medal),na.rm=TRUE) %>%  
 arrange(desc(somme\_Silver))%>%  
 select(Country,somme\_Silver)

## `summarise()` ungrouping output (override with `.groups` argument)

x<-seq(1:106)  
  
data\_Sil1<-cbind(x,data\_Sil)  
colnames(data\_Sil1)<-c("Rank","Country","Silver\_medals")  
data\_Sil1

## Rank Country Silver\_medals  
## 1 1 United States 583  
## 2 2 Soviet Union 285  
## 3 3 Australia 270  
## 4 4 China 252  
## 5 5 Russia 206  
## 6 6 East Germany 190  
## 7 7 Korea, South 186  
## 8 8 Germany 176  
## 9 9 Spain 165  
## 10 10 Italy 163  
## 11 11 Romania 157  
## 12 12 United Kingdom 157  
## 13 13 Netherlands 140  
## 14 14 Brazil 137  
## 15 15 West Germany 135  
## 16 16 Poland 113  
## 17 17 Japan 112  
## 18 18 Canada 111  
## 19 19 France 110  
## 20 20 Sweden 110  
## 21 21 Cuba 109  
## 22 22 Hungary 104  
## 23 23 Bulgaria 100  
## 24 24 Yugoslavia 86  
## 25 25 Unified team 65  
## 26 26 Norway 58  
## 27 27 Nigeria 38  
## 28 28 Ukraine 38  
## 29 29 Argentina 37  
## 30 30 Switzerland 37  
## 31 31 Jamaica 34  
## 32 32 Greece 32  
## 33 33 Croatia 30  
## 34 34 Czechoslovakia 29  
## 35 35 Taiwan 26  
## 36 36 Belarus 25  
## 37 37 Kenya 21  
## 38 38 New Zealand 21  
## 39 39 Denmark 20  
## 40 40 Czech Republic 18  
## 41 41 Austria 17  
## 42 42 Finland 17  
## 43 43 Paraguay 17  
## 44 44 Kazakhstan 16  
## 45 45 Belgium 15  
## 46 46 Serbia 15  
## 47 47 Iceland 14  
## 48 48 Indonesia 14  
## 49 49 Peru 14  
## 50 50 Korea, North 11  
## 51 51 Slovakia 11  
## 52 52 Turkey 11  
## 53 53 Mexico 10  
## 54 54 South Africa 10  
## 55 55 Bahamas 9  
## 56 56 Latvia 9  
## 57 57 Iran 6  
## 58 58 Ireland 6  
## 59 59 Slovenia 6  
## 60 60 Trinidad and Tobago 6  
## 61 61 Ethiopia 5  
## 62 62 Mongolia 5  
## 63 63 Portugal 5  
## 64 64 Uzbekistan 5  
## 65 65 Lithuania 4  
## 66 66 Morocco 4  
## 67 67 Namibia 4  
## 68 68 Thailand 4  
## 69 69 Zimbabwe 4  
## 70 70 Azerbaijan 3  
## 71 71 Estonia 3  
## 72 72 Malaysia 3  
## 73 73 Moldova 3  
## 74 74 Singapore 3  
## 75 75 Algeria 2  
## 76 76 Chile 2  
## 77 77 Colombia 2  
## 78 78 Egypt 2  
## 79 79 Georgia 2  
## 80 80 Hong Kong\* 2  
## 81 81 Tanzania 2  
## 82 82 Venezuela 2  
## 83 83 Vietnam 2  
## 84 84 Armenia 1  
## 85 85 Costa Rica 1  
## 86 86 Cote d'Ivoire 1  
## 87 87 Dominican Republic 1  
## 88 88 Ecuador 1  
## 89 89 Independent Olympic Participants (1992) 1  
## 90 90 India 1  
## 91 91 Israel 1  
## 92 92 Kyrgyzstan 1  
## 93 93 Netherlands Antilles\* 1  
## 94 94 Philippines 1  
## 95 95 Puerto Rico\* 1  
## 96 96 Saudi Arabia 1  
## 97 97 Senegal 1  
## 98 98 Sri Lanka 1  
## 99 99 Sudan 1  
## 100 100 Syria 1  
## 101 101 Tajikistan 1  
## 102 102 Tonga 1  
## 103 103 Uganda 1  
## 104 104 Uruguay 1  
## 105 105 Virgin Islands\* 1  
## 106 106 Zambia 1

##5.Top bronze medal winning countries:

data\_Bronze <- subset(data,data$Medal=="Bronze")  
data\_Bronze$Medal="1"  
head(data\_Bronze)

## City Year Sport Discipline Event  
## 2 Montreal 1976 Aquatics Diving 3m springboard  
## 5 Montreal 1976 Aquatics Diving 10m platform  
## 8 Montreal 1976 Aquatics Diving 3m springboard  
## 10 Montreal 1976 Aquatics Diving 10m platform  
## 16 Montreal 1976 Aquatics Swimming 1500m freestyle  
## 17 Montreal 1976 Aquatics Swimming 400m individual medley  
## Athlete Gender Country\_Code Country Event\_gender  
## 2 KOSENKOV, Aleksandr Men URS Soviet Union M  
## 5 WILSON, Deborah Keplar Women USA United States W  
## 8 POTTER-MCINGVALE, Cynthia Women USA United States W  
## 10 ALEINIK, Vladimir Men URS Soviet Union M  
## 16 HOLLAND, Stephen Men AUS Australia M  
## 17 SMIRNOV, Andrei Men URS Soviet Union M  
## Medal  
## 2 1  
## 5 1  
## 8 1  
## 10 1  
## 16 1  
## 17 1

data\_Bronze$Medal<-as.numeric(data\_Bronze$Medal)  
class(data\_Silver$Bronze)

## [1] "NULL"

data\_Brn<-data\_Bronze %>%  
 group\_by(Country) %>%  
 summarise(somme\_Bronze=sum(Medal),na.rm=TRUE) %>%  
 arrange(desc(somme\_Bronze))%>%  
 select(Country,somme\_Bronze)

## `summarise()` ungrouping output (override with `.groups` argument)

x<-seq(1:108)  
  
data\_Brn1<-cbind(x,data\_Brn)  
colnames(data\_Brn1)<-c("Rank","Country","Bronze\_medals")  
data\_Brn1

## Rank Country Bronze\_medals  
## 1 1 United States 481  
## 2 2 Australia 312  
## 3 3 Soviet Union 297  
## 4 4 Germany 278  
## 5 5 Russia 240  
## 6 6 China 193  
## 7 7 Romania 190  
## 8 8 United Kingdom 188  
## 9 9 France 185  
## 10 10 Japan 182  
## 11 11 Italy 178  
## 12 12 Netherlands 151  
## 13 13 East Germany 150  
## 14 14 Hungary 135  
## 15 15 Korea, South 128  
## 16 16 West Germany 126  
## 17 17 Brazil 125  
## 18 18 Canada 117  
## 19 19 Bulgaria 107  
## 20 20 Yugoslavia 102  
## 21 21 Poland 98  
## 22 22 Cuba 88  
## 23 23 Ukraine 78  
## 24 24 Spain 76  
## 25 25 Argentina 70  
## 26 26 Unified team 66  
## 27 27 Sweden 55  
## 28 28 Belarus 53  
## 29 29 New Zealand 51  
## 30 30 Denmark 50  
## 31 31 Norway 46  
## 32 32 Lithuania 42  
## 33 33 Jamaica 38  
## 34 34 Pakistan 33  
## 35 35 Czechoslovakia 29  
## 36 36 Switzerland 28  
## 37 37 Nigeria 25  
## 38 38 Greece 23  
## 39 39 Mexico 22  
## 40 40 Belgium 20  
## 41 41 Chile 19  
## 42 42 Finland 19  
## 43 43 Croatia 18  
## 44 44 Kenya 17  
## 45 45 Korea, North 17  
## 46 46 Turkey 15  
## 47 47 Kazakhstan 14  
## 48 48 Serbia 14  
## 49 49 Czech Republic 13  
## 50 50 Ghana 13  
## 51 51 Ethiopia 12  
## 52 52 Indonesia 12  
## 53 53 Taiwan 12  
## 54 54 Georgia 11  
## 55 55 Slovenia 11  
## 56 56 Morocco 10  
## 57 57 Thailand 10  
## 58 58 Azerbaijan 9  
## 59 59 Algeria 8  
## 60 60 Austria 8  
## 61 61 Iran 8  
## 62 62 Slovakia 8  
## 63 63 Uzbekistan 8  
## 64 64 Armenia 7  
## 65 65 Mongolia 7  
## 66 66 Portugal 7  
## 67 67 South Africa 7  
## 68 68 Estonia 6  
## 69 69 Venezuela 6  
## 70 70 Colombia 5  
## 71 71 Israel 5  
## 72 72 Egypt 4  
## 73 73 India 4  
## 74 74 Ireland 4  
## 75 75 Puerto Rico\* 4  
## 76 76 Trinidad and Tobago 4  
## 77 77 Bahamas 3  
## 78 78 Latvia 3  
## 79 79 Malaysia 3  
## 80 80 Moldova 3  
## 81 81 Costa Rica 2  
## 82 82 Iceland 2  
## 83 83 Independent Olympic Participants (1992) 2  
## 84 84 Kyrgyzstan 2  
## 85 85 Philippines 2  
## 86 86 Qatar 2  
## 87 87 Afghanistan 1  
## 88 88 Barbados 1  
## 89 89 Bermuda\* 1  
## 90 90 Cameroon 1  
## 91 91 Djibouti 1  
## 92 92 Dominican Republic 1  
## 93 93 Eritrea 1  
## 94 94 Guyana 1  
## 95 95 Kuwait 1  
## 96 96 Lebanon 1  
## 97 97 Macedonia 1  
## 98 98 Mauritius 1  
## 99 99 Mozambique 1  
## 100 100 Saudi Arabia 1  
## 101 101 Suriname 1  
## 102 102 Syria 1  
## 103 103 Tajikistan 1  
## 104 104 Togo 1  
## 105 105 Tunisia 1  
## 106 106 Uganda 1  
## 107 107 Zambia 1  
## 108 108 Zimbabwe 1

##6.Number of medals won by Morocco in the Olympics from 1976 until 2008:

data\_Morocco<-subset(data,Country=="Morocco")  
data\_Morocco

## City Year Sport Discipline Event  
## 3072 Los Angeles 1984 Athletics Athletics 400m hurdles  
## 3140 Los Angeles 1984 Athletics Athletics 5000m  
## 4549 Seoul 1988 Athletics Athletics 10000m  
## 4566 Seoul 1988 Athletics Athletics 800m  
## 4777 Seoul 1988 Boxing Boxing 54 - 57kg (featherweight)  
## 6101 Barcelona 1992 Athletics Athletics 1500m  
## 6106 Barcelona 1992 Athletics Athletics 10000m  
## 6429 Barcelona 1992 Boxing Boxing 51 - 54kg (bantamweight)  
## 7816 Atlanta 1996 Athletics Athletics 5000m  
## 7964 Atlanta 1996 Athletics Athletics 10000m  
## 9768 Sydney 2000 Athletics Athletics 3000m steeplechase  
## 9833 Sydney 2000 Athletics Athletics 400m hurdles  
## 9849 Sydney 2000 Athletics Athletics 1500m  
## 9907 Sydney 2000 Athletics Athletics 5000m  
## 10088 Sydney 2000 Boxing Boxing 54 - 57kg (featherweight)  
## 11800 Athens 2004 Athletics Athletics 800m  
## 11835 Athens 2004 Athletics Athletics 1500m  
## 11838 Athens 2004 Athletics Athletics 5000m  
## 13872 Beijing 2008 Athletics Athletics 800m  
## 13930 Beijing 2008 Athletics Athletics marathon  
## Athlete Gender Country\_Code Country Event\_gender Medal  
## 3072 EL MOUTAWAKEL, Nawal Women MAR Morocco W Gold  
## 3140 AOUITA, Said Men MAR Morocco M Gold  
## 4549 BOUTAYEB, Moulay Brahim Men MAR Morocco M Gold  
## 4566 AOUITA, Said Men MAR Morocco M Bronze  
## 4777 ACHIK, Abdelhak Men MAR Morocco M Bronze  
## 6101 EL BASIR, Rachid Men MAR Morocco M Silver  
## 6106 SKAH, Khalid Men MAR Morocco M Gold  
## 6429 ACHIK, Mohamed Men MAR Morocco M Bronze  
## 7816 BOULAMI, Khalid Men MAR Morocco M Bronze  
## 7964 HISSOU, Salah Men MAR Morocco M Bronze  
## 9768 EZZINE, Ali Men MAR Morocco M Bronze  
## 9833 BIDOUANE, Nouzha Women MAR Morocco W Bronze  
## 9849 EL GUERROUJ, Hicham Men MAR Morocco M Silver  
## 9907 LAHLAFI, Brahim Men MAR Morocco M Bronze  
## 10088 TAMSAMANI, Tahar Men MAR Morocco M Bronze  
## 11800 BENHASSI, Hasna Women MAR Morocco W Silver  
## 11835 EL GUERROUJ, Hicham Men MAR Morocco M Gold  
## 11838 EL GUERROUJ, Hicham Men MAR Morocco M Gold  
## 13872 BENHASSI, Hasna Women MAR Morocco W Bronze  
## 13930 GHARIB, Jaouad Men MAR Morocco M Silver

data\_Morocco$Medal=="1"

## [1] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE  
## [13] FALSE FALSE FALSE FALSE FALSE FALSE FALSE FALSE

data\_Morocco$Medal<-as.numeric(data\_Morocco$Medal)  
data\_Moroccoperyear<-data\_Morocco %>%  
 group\_by(Year)%>%  
 summarise(Somme=sum(Medal),na.rm=TRUE) %>%  
 arrange(Somme) %>%  
 select(Year,Somme)

## `summarise()` ungrouping output (override with `.groups` argument)

data\_Moroccoperyear

## # A tibble: 7 x 2  
## Year Somme  
## <int> <dbl>  
## 1 1996 4  
## 2 1984 6  
## 3 2008 6  
## 4 1988 7  
## 5 1992 9  
## 6 2004 10  
## 7 2000 12

##7.Plot shows medals won by Morocco in the Olympics from 1976 until 2008:

g<-ggplot(data\_Moroccoperyear,aes(Year,Somme))  
g+geom\_point()+geom\_line(colour="green",alpha=1)+labs(title="Medals won by Morocco from 1976 until 2008 ")

