|  |
| --- |
| Title: “IST687 – Final Project | 2016 Olympics in Rio de Janeiro” |
| Name: Sathish Kumar Rajendiran |
| Group Number: 3 |
| Members: Deanna Zarrillo, Jeff Thomson, Sathish Kumar Rajendiran |
| Date: 05/12/2020 |

Exercise: 2016 Olympics, Rio de Janeiro, Brazil

Countries: Number of Players: Male: Female: Medal Category: Gold, Silver, Bronze Games: Venues Played:

￼ 1. Problem formulation 2. Data cleanup 3. Pre-processing 4. Train-test Split 5. Model Building 6. Accuracy & Validation 7. Prediction 8. Confusion Matrix

Descriptive Analysis Predictive Analysis Prescrptive Analysis

# Install necessary packages

# import libraries  
  
install.packages(pkgs=c("ggplot2","reshape2","ggeasy","dplyr","data.table","RSQLite","sqldf","reshape"),repos = "http://cran.us.r-project.org")

##   
## The downloaded binary packages are in  
## /var/folders/\_z/ltmjkt4156b37rsk7cgvj7l80000gn/T//RtmpHoS05M/downloaded\_packages

library(ggplot2)  
library(ggeasy)  
library(reshape2)  
library(data.table)

##   
## Attaching package: 'data.table'

## The following objects are masked from 'package:reshape2':  
##   
## dcast, melt

library(stats)  
library(dplyr)

##   
## Attaching package: 'dplyr'

## The following objects are masked from 'package:data.table':  
##   
## between, first, last

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

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

library(RSQLite)  
library(sqldf)

## Loading required package: gsubfn

## Loading required package: proto

library(reshape)

##   
## Attaching package: 'reshape'

## The following object is masked from 'package:dplyr':  
##   
## rename

## The following object is masked from 'package:data.table':  
##   
## melt

## The following objects are masked from 'package:reshape2':  
##   
## colsplit, melt, recast

library(tidyverse)

## ── Attaching packages ─────────────────────────────────────── tidyverse 1.3.0 ──

## ✓ tibble 3.0.1 ✓ purrr 0.3.4  
## ✓ tidyr 1.0.2 ✓ stringr 1.4.0  
## ✓ readr 1.3.1 ✓ forcats 0.5.0

## ── Conflicts ────────────────────────────────────────── tidyverse\_conflicts() ──  
## x dplyr::between() masks data.table::between()  
## x tidyr::expand() masks reshape::expand()  
## x dplyr::filter() masks stats::filter()  
## x dplyr::first() masks data.table::first()  
## x dplyr::lag() masks stats::lag()  
## x dplyr::last() masks data.table::last()  
## x reshape::rename() masks dplyr::rename()  
## x purrr::transpose() masks data.table::transpose()

library(rvest)

## Loading required package: xml2

##   
## Attaching package: 'rvest'

## The following object is masked from 'package:purrr':  
##   
## pluck

## The following object is masked from 'package:readr':  
##   
## guess\_encoding

library(magrittr)

##   
## Attaching package: 'magrittr'

## The following object is masked from 'package:purrr':  
##   
## set\_names

## The following object is masked from 'package:tidyr':  
##   
## extract

library(ggmap)

## Google's Terms of Service: https://cloud.google.com/maps-platform/terms/.

## Please cite ggmap if you use it! See citation("ggmap") for details.

##   
## Attaching package: 'ggmap'

## The following object is masked from 'package:magrittr':  
##   
## inset

library(stringr)  
library(RSQLite)  
library(viridis)

## Loading required package: viridisLite

library(hrbrthemes)

## NOTE: Either Arial Narrow or Roboto Condensed fonts are required to use these themes.

## Please use hrbrthemes::import\_roboto\_condensed() to install Roboto Condensed and

## if Arial Narrow is not on your system, please see https://bit.ly/arialnarrow

#1. Import raw data into dataframes 1. Download files 2. Create data frames 3. Review meta data

#import csv files into data frames  
  
 athletes\_fpath <- "/Users/sathishrajendiran/Documents/olympic-games/athletes.csv"  
 countries\_fpath <- "/Users/sathishrajendiran/Documents/olympic-games/countries.csv"  
 events\_fpath <- "/Users/sathishrajendiran/Documents/olympic-games/events.csv"  
  
 # function readFiles  
 readFiles <- function(fpath)  
 {  
 dftemp <- data.frame(read.csv(fpath),stringsAsFactors=FALSE)  
 return(dftemp)  
 }  
   
 athletes <- readFiles(athletes\_fpath)   
 countries <- readFiles(countries\_fpath)   
 events <- readFiles(events\_fpath)   
  
   
# function printDataInfo  
 printDataInfo <- function(myData)  
 {  
 strinfo <- str(myData)  
 cat("str:",strinfo,"\n")  
   
 colnamesinfo <- colnames(myData)  
 cat("colnames:",colnamesinfo,"\n")  
   
 diminfo <- dim(myData)  
 cat("dim:",diminfo,"\n")  
   
 nrowinfo <- nrow(myData)  
 cat("nrow:",nrowinfo,"\n")  
   
 nrowsinfo <- myData[1:3,]  
 return(nrowsinfo)  
 }  
   
# printDataInfo  
 printDataInfo(athletes) #dim: 11538 11

## 'data.frame': 11538 obs. of 11 variables:  
## $ id : int 736041664 532037425 435962603 521041435 33922579 173071782 266237702 382571888 87689776 997877719 ...  
## $ name : Factor w/ 11517 levels "A Jesus Garcia",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ nationality: Factor w/ 207 levels "AFG","ALB","ALG",..: 60 103 34 120 142 11 199 11 60 62 ...  
## $ sex : Factor w/ 2 levels "female","male": 2 1 2 2 2 2 2 2 1 1 ...  
## $ dob : Factor w/ 5596 levels "","01/01/1969",..: 4503 4151 2325 32 5087 424 2450 4186 5340 3194 ...  
## $ height : num 1.72 1.68 1.98 1.83 1.81 1.8 2.05 1.93 1.8 1.65 ...  
## $ weight : int 64 56 79 80 71 67 98 100 62 54 ...  
## $ sport : Factor w/ 28 levels "aquatics","archery",..: 3 10 3 23 8 25 26 1 3 3 ...  
## $ gold : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ silver : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ bronze : int 0 0 1 0 0 0 1 0 0 0 ...  
## str:   
## colnames: id name nationality sex dob height weight sport gold silver bronze   
## dim: 11538 11   
## nrow: 11538

## id name nationality sex dob height weight  
## 1 736041664 A Jesus Garcia ESP male 10/17/1969 1.72 64  
## 2 532037425 A Lam Shin KOR female 09/23/1986 1.68 56  
## 3 435962603 Aaron Brown CAN male 05/27/1992 1.98 79  
## sport gold silver bronze  
## 1 athletics 0 0 0  
## 2 fencing 0 0 0  
## 3 athletics 0 0 1

printDataInfo(countries) #dim: 201 4

## 'data.frame': 201 obs. of 4 variables:  
## $ country : Factor w/ 201 levels "Afghanistan",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ code : Factor w/ 201 levels "AFG","AHO","ALB",..: 1 3 4 11 5 6 7 8 9 10 ...  
## $ population : int 32526562 2889167 39666519 55538 70473 25021974 91818 43416755 3017712 103889 ...  
## $ gdp\_per\_capita: num 594 3945 4206 NA NA ...  
## str:   
## colnames: country code population gdp\_per\_capita   
## dim: 201 4   
## nrow: 201

## country code population gdp\_per\_capita  
## 1 Afghanistan AFG 32526562 594.3231  
## 2 Albania ALB 2889167 3945.2176  
## 3 Algeria ALG 39666519 4206.0312

printDataInfo(events) #dim: 306 6

## 'data.frame': 306 obs. of 6 variables:  
## $ id : int 701492 305278 708010 729643 567019 607924 519818 778941 795006 93436 ...  
## $ sport : Factor w/ 28 levels "aquatics","archery",..: 1 1 1 1 1 1 1 1 1 1 ...  
## $ discipline: Factor w/ 50 levels "archery","artistic gymnastics",..: 3 3 3 3 8 8 8 8 9 9 ...  
## $ name : Factor w/ 273 levels "10m Air Pistol Men",..: 187 196 52 63 188 197 53 64 189 198 ...  
## $ sex : Factor w/ 3 levels "female","male",..: 1 1 2 2 1 1 2 2 1 1 ...  
## $ venues : Factor w/ 33 levels "Beach Volleyball Arena",..: 17 17 17 17 17 17 17 17 17 17 ...  
## str:   
## colnames: id sport discipline name sex venues   
## dim: 306 6   
## nrow: 306

## id sport discipline name sex  
## 1 701492 aquatics backstroke Women's 100m Backstroke female  
## 2 305278 aquatics backstroke Women's 200m Backstroke female  
## 3 708010 aquatics backstroke Men's 100m Backstroke male  
## venues  
## 1 Olympic Aquatics Stadium  
## 2 Olympic Aquatics Stadium  
## 3 Olympic Aquatics Stadium

#2. Data Cleaning Process

1. Remove NA Characters  
2. Change Data type   
3. Remove unwanted columns  
4. Add columns (Copy + calculated)  
5. Rename Columns  
6. Re-arrange Columns  
7. ?

2.1 Athletes Dataframe

# 2. Data Cleaning Process | 2.1 Athletes Dataframe  
  
# Preview the data  
   
 printDataInfo(athletes)

## 'data.frame': 11538 obs. of 11 variables:  
## $ id : int 736041664 532037425 435962603 521041435 33922579 173071782 266237702 382571888 87689776 997877719 ...  
## $ name : Factor w/ 11517 levels "A Jesus Garcia",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ nationality: Factor w/ 207 levels "AFG","ALB","ALG",..: 60 103 34 120 142 11 199 11 60 62 ...  
## $ sex : Factor w/ 2 levels "female","male": 2 1 2 2 2 2 2 2 1 1 ...  
## $ dob : Factor w/ 5596 levels "","01/01/1969",..: 4503 4151 2325 32 5087 424 2450 4186 5340 3194 ...  
## $ height : num 1.72 1.68 1.98 1.83 1.81 1.8 2.05 1.93 1.8 1.65 ...  
## $ weight : int 64 56 79 80 71 67 98 100 62 54 ...  
## $ sport : Factor w/ 28 levels "aquatics","archery",..: 3 10 3 23 8 25 26 1 3 3 ...  
## $ gold : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ silver : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ bronze : int 0 0 1 0 0 0 1 0 0 0 ...  
## str:   
## colnames: id name nationality sex dob height weight sport gold silver bronze   
## dim: 11538 11   
## nrow: 11538

## id name nationality sex dob height weight  
## 1 736041664 A Jesus Garcia ESP male 10/17/1969 1.72 64  
## 2 532037425 A Lam Shin KOR female 09/23/1986 1.68 56  
## 3 435962603 Aaron Brown CAN male 05/27/1992 1.98 79  
## sport gold silver bronze  
## 1 athletics 0 0 0  
## 2 fencing 0 0 0  
## 3 athletics 0 0 1

# 2.1.1 [ No need now]  
 # Data format of dob column is as factor  
 # Lets change it to yyyy-mm-dd format  
   
 # athletes$dob = as.Date(athletes$dob,"%m/%d/%y")  
  
# 2.1.2  
 # Check if the NAs in athletes and replace them   
  
 clnames <- colnames(athletes)[colSums(is.na(athletes)) > 0]  
 clnames

## [1] "height" "weight"

# create subset of dataframe rows having NAs  
 na\_athletes <- athletes[rowSums(is.na(athletes)) > 0,]  
 na\_athletes # 680- rows

## id name nationality sex  
## 13 258556239 Abbas Qali IOA male  
## 29 349871091 Abdelhafid Benchabla ALG male  
## 31 23564778 Abdelkader Chadi ALG male  
## 38 934545704 Abdelrahman Salah Orabi Abdelgawwad EGY male  
## 48 469953606 Abdoullah Bamoussa ITA male  
## 51 325809293 Abdul Omar GHA male  
## 53 262868423 Abdulaziz Alshatti IOA male  
## 54 101781750 Abdulkadir Abdullayev AZE male  
## 57 897549624 Abdullah Hel Baki BAN male  
## 58 153457 Abdullahi Shehu NGR male  
## 60 937153550 Abdulrahman Al Faihan IOA male  
## 63 884912875 Abeku Gyekye Jackson GHA male  
## 67 157637514 Ablaikhan Zhussupov KAZ male  
## 73 140663745 Achraf Kharroubi MAR male  
## 95 222063859 Adama Jammeh GAM male  
## 109 75289002 Adilbek Niyazymbetov KAZ male  
## 112 456381770 Adlan Abdurashidov RUS male  
## 130 593328448 Adriana Araujo BRA female  
## 140 577601340 Afa Ismail MDV female  
## 166 443237540 Ahmad Alafasi IOA male  
## 175 9418987 Ahmed Attellesey LBA male  
## 181 64299896 Ahmed Gebrel PLE male  
## 247 959689266 Akossiwa Claire Ayivon TOG female  
## 248 152032519 Akua Obeng-Akrofi GHA female  
## 269 184326688 Albert Ramon Ramirez VEN male  
## 272 670559048 Albert Selimov AZE male  
## 278 331799449 Alberto Ezequiel Melian ARG male  
## 280 566773050 Alberto Ignacio Palmetta ARG male  
## 323 646633611 Aleksei Kurbatov RUS male  
## 351 527236860 Alex Amankwah GHA male  
## 366 107240268 Alex Rose SAM male  
## 389 37893138 Alexander Lerionka Sampao KEN male  
## 445 448880966 Alexi Pappas GRE female  
## 459 461686848 Ali Elghrari LBA male  
## 460 668278260 Ali Eren Demirezen TUR male  
## 553 791528062 Alonzo Russell BAH male  
## 555 601227072 Alphonce Felix Simbu TAN male  
## 578 398682051 Amalie Iuel NOR female  
## 587 557518345 Amanda Kurtovic NOR female  
## 605 193520300 Ambroise Uwiragiye RWA male  
## 608 39468987 Amel Bouderra FRA female  
## 625 509681869 Aminath Shajan MDV female  
## 629 389896499 Aminu Umar NGR male  
## 637 657073926 Amnat Ruenroeng THA male  
## 682 711141982 Anas Beshr EGY male  
## 692 971263231 Anastasia Rodionova AUS female  
## 696 901391372 Anastasiia Beliakova RUS female  
## 728 289946138 Anders Pedersen NOR male  
## 801 153735693 Andrei Zamkovoi RUS male  
## 802 957911430 Andreia Bandeira BRA female  
## 831 149520623 Andrew Fisher BRN male  
## 853 295536761 Andrey Zeits KAZ male  
## 856 730965977 Andrii Khripta UKR male  
## 860 506743542 Andrique Allisop SEY male  
## 909 54089758 Anilda Thomas IND female  
## 952 550413879 Anna Laurell Nash SWE female  
## 982 8185560 Anne Cairns SAM female  
## 1017 547860838 Anthony Barbar LIB male  
## 1067 824387423 Antonio Vargas USA male  
## 1069 565509368 Antony Fowler GBR male  
## 1076 479087750 Anvar Yunusov TJK male  
## 1086 706967900 Arajik Marutjan GER male  
## 1088 481716571 Aram Avagyan ARM male  
## 1093 756028808 Arashi Morisaka JPN male  
## 1096 554149372 Are Hansen NOR male  
## 1097 269857952 Are Strandli NOR male  
## 1106 978867868 Ariane Fortin CAN female  
## 1114 520698565 Arina Rodionova AUS female  
## 1120 720999336 Arlen Lopez CUB male  
## 1133 725424114 Aron Kifle ERI male  
## 1138 361362188 Arslanbek Achilov TKM male  
## 1140 552816016 Artem Chebotarev RUS male  
## 1142 172069479 Artem Harutyunyan GER male  
## 1150 216948866 Arthur Biyarslanov CAN male  
## 1160 311001055 Artur Hovhannisyan ARM male  
## 1192 667381951 Ashleigh Simon RSA female  
## 1227 915895832 Atheyna Bylon PAN female  
## 1249 571398392 Augusto Soares TLS male  
## 1270 950658256 Aya Takeuchi JPN female  
## 1285 296010311 Ayouba Traore MLI male  
## 1299 316544316 Baboloki Thebe BOT male  
## 1306 417882039 Bakhodir Jalolov UZB male  
## 1358 409808283 Batuhan Gozgec TUR male  
## 1372 687126240 Beatrice Gyaman GHA female  
## 1373 33337656 Beatrice Kamuchanga Alice COD female  
## 1393 252644554 Bektemir Melikuziev UZB male  
## 1430 678983494 Benjamin Kiplagat UGA male  
## 1441 381218659 Benny Muziyo Muziyo ZAM male  
## 1443 544056432 Benson Gicharu Njangiru KEN male  
## 1445 574621304 Berik Abdrakhmanov KAZ male  
## 1477 579106443 Bianca Stuart BAH female  
## 1485 543026073 Bilel Mhamdi TUN male  
## 1494 877888215 Bin Lv CHN male  
## 1502 131536621 Birhanu Balew BRN male  
## 1504 112486635 Birzhan Zhakypov KAZ male  
## 1535 813532025 Boe Warawara VAN male  
## 1541 309321530 Boitumelo Masilo BOT male  
## 1583 327920403 Bralon Taplin GRN male  
## 1606 262561474 Brendan Emmet Irvine IRL male  
## 1654 290572119 Brooke Stratton AUS female  
## 1681 326008552 Bunturabie Jalloh SLE female  
## 1702 100675394 Cale Simmons USA male  
## 1725 384002345 Camilla Herrem NOR female  
## 1737 495694072 Cancan Ren CHN female  
## 1762 639279380 Carlos Andres Mina ECU male  
## 1769 955217464 Carlos Eduardo Quipo Pilataxi ECU male  
## 1780 318016892 Carlos Zenon Balderas Jr. USA male  
## 1790 692509182 Carmiesha Cox BAH female  
## 1791 386843805 Carmine Tommasone ITA male  
## 1816 944589858 Carvin Nkanata KEN male  
## 1847 71279109 Caylee Watson ISV female  
## 1857 82592097 Cedric Dubler AUS male  
## 1859 437255555 Ceiber David Avila COL male  
## 1894 523414424 Changgeun Lee KOR male  
## 1916 792508512 Charles Albert Shone Conwell USA male  
## 1946 534915497 Charly Coronel Suarez PHI male  
## 1948 172605040 Chatchai Butdee THA male  
## 1952 284105289 Cheick Sallah Junior Cisse CIV male  
## 1992 885083820 Chinzorig Baatarsukh MGL male  
## 1993 775422570 Chirine Njeim LIB female  
## 2012 82546289 Chouaib Bouloudinats ALG male  
## 2024 110660329 Chris Langridge GBR male  
## 2047 461859304 Christian Mbilli FRA male  
## 2048 530803618 Christian Nassif Djidagui CAF male  
## 2056 511915390 Christian Zimmermann PLE male  
## 2065 524363491 Christine Botlogetswe BOT female  
## 2085 804679484 Christopher George TTO male  
## 2098 141489782 Chu-En Lai TPE male  
## 2128 175558091 Claressa Maria Shields USA female  
## 2155 212851704 Clemente Russo ITA male  
## 2190 362762819 Corey Ollivierre GRN male  
## 2230 492543946 Cy Thompson ISV male  
## 2257 429516855 Daisuke Narimatsu JPN male  
## 2303 119961114 Daniah Hagul LBA female  
## 2305 161929163 Daniel Akpeyi NGR male  
## 2308 2865880 Daniel Asenov BUL male  
## 2329 738563409 Daniel Jason Lewis AUS male  
## 2349 815557568 Daniel Teklehaimanot ERI male  
## 2384 76405574 Daniyar Yeleussinov KAZ male  
## 2411 967665587 Daria Ustinova RUS female  
## 2415 977866181 Dariga Shakimova KAZ female  
## 2462 231429156 David Graf GER male  
## 2483 845122813 David Oliver Joyce IRL male  
## 2498 661068918 David van der Colff BOT male  
## 2502 382116248 Davilson dos Santos Morais CPV male  
## 2513 996704082 Deajah Stevens USA female  
## 2554 598284598 Denis Petrashov KGZ male  
## 2565 556851081 Denys Solonenko UKR male  
## 2573 368262898 Derek Sua SAM male  
## 2623 316835810 Dieudonne Wilfred Seyi Ntsengue CMR male  
## 2658 158299289 Dival Forele Malonga Dzalamou CGO male  
## 2676 744288493 Dmytro Mytrofanov UKR male  
## 2679 608020695 Doaa Elghobashy EGY female  
## 2703 531161609 Donghyen Shin KOR male  
## 2720 690232757 Dorcas Gyimah GHA female  
## 2761 578322269 Dwight Lewis VEN male  
## 2764 348929342 Dylan Carter TTO male  
## 2771 733134503 Dzmitry Asanau BLR male  
## 2790 684979118 Edgar Ramon Munoz VEN male  
## 2810 706338154 Edvald Boasson Hagen NOR male  
## 2829 293548179 Efe Ajagba NGR male  
## 2835 184575330 Ehsan Rouzbahani IRI male  
## 2836 181030927 Ei Ei Thet MYA female  
## 2843 974051662 Eimantas Stanionis LTU male  
## 2846 381623095 Ejowvokoghene Oduduru NGR male  
## 2864 271247492 Ekaterina Tunguskova UZB female  
## 2878 969535461 Eldred Henry IVB male  
## 2880 205100723 Eleanor Bezzina MLT female  
## 2887 222839663 Elena Aniushina RUS female  
## 2892 137876590 Elena Galiabovitch AUS female  
## 2910 15776102 Elias Eliseo Emigdio Abarca MEX male  
## 2912 592765424 Elie Konki FRA male  
## 2978 434777707 Elmo Jankari FIN male  
## 2984 786667098 Eloi Imaniraguha RWA male  
## 2990 31920396 Elshod Rasulov UZB male  
## 2993 821149858 Elvin Mamishzada AZE male  
## 3009 468726382 Emerric Kpegba TOG male  
## 3026 542195691 Emilie Hegh Arntzen NOR female  
## 3030 838354626 Emily Gielnik AUS female  
## 3058 102727571 Emmanuel Daniel NGR male  
## 3059 545212161 Emmanuel Dasor GHA male  
## 3062 364575203 Emmanuel Matadi LBR male  
## 3068 450857033 Endry Jose Saavedra VEN male  
## 3072 957160839 Enkh-Amar Kharkhuu MGL male  
## 3075 3678441 Enrico Lacruz NED male  
## 3108 184812086 Erik Pfeifer GER male  
## 3110 476898873 Erik Tysse NOR male  
## 3131 738743977 Erislandy Savon CUB male  
## 3132 626115327 Erkin Adylbek Uulu KGZ male  
## 3147 11916654 Espen Kofstad NOR male  
## 3160 317848923 Estelle Mossely FRA female  
## 3178 910783696 Eun Ju Lee KOR female  
## 3198 509893044 Evaldas Petrauskas LTU male  
## 3213 94836813 Evelyn Cipriano CUB female  
## 3230 93776802 Evgeny Tishchenko RUS male  
## 3233 664096617 Evita Leter SUR female  
## 3243 340239700 Ezinne Okparaebo NOR female  
## 3259 391451002 Fabiano Joseph TAN male  
## 3271 683014073 Fabrice Dabla TOG male  
## 3282 203217392 Fahem Hammachi ALG male  
## 3283 442721 Faicel Jaballah TUN male  
## 3309 457192080 Fatma El Sharnouby EGY female  
## 3312 930422859 Faye Husain IOA female  
## 3313 471409647 Faye Njie GAM male  
## 3314 454377713 Fazliddin Gaibnazarov UZB male  
## 3335 426841600 Felipe AMARAL BRA male  
## 3342 619195032 Felipe Carmo BRA male  
## 3366 481744691 Fengkai Yu CHN male  
## 3386 175585720 Fernando Daniel Martinez ARG male  
## 3393 201491381 Fernando Prass BRA male  
## 3404 155085294 Filip Hrgovic CRO male  
## 3405 797627457 Filip Ingebrigtsen NOR male  
## 3433 166844486 Flings Owusu-Agyapong GHA female  
## 3462 861278504 Francelina Cabral TLS female  
## 3514 464256317 Freddy Figueroa ECU male  
## 3541 590308057 Gabriel Maestre VEN male  
## 3557 697532444 Gabriella Doueihy LIB female  
## 3561 463517978 Gabrielle Adcock GBR female  
## 3577 866485713 Galal Yafai GBR male  
## 3579 711261984 Galia Dvorak ESP female  
## 3583 781481729 Gan-Erdene Gankhuyag MGL male  
## 3584 897822303 Ganapathi Krishnan IND male  
## 3590 538703926 Gaone Leaname Maotoanong BOT male  
## 3595 249661772 Gary Russell USA male  
## 3690 827807718 Ghirmay Ghebreslassie ERI male  
## 3712 784102100 Gina Bass GAM female  
## 3760 56023329 Goitom Kifle ERI male  
## 3778 702595268 Grace Claxton PUR female  
## 3829 905148205 Guido Vianello ITA male  
## 3843 504342645 Gunn-Rita Dahle Flesjaa NOR female  
## 3856 833430291 Gustavo Vernes BRA male  
## 3860 796100413 Gwanghyeok Lee KOR male  
## 3880 177699282 Hafsatu Kamara SLE female  
## 3895 9093283 Hakan Eresker QAT male  
## 3901 779373345 Halimah Nakaayi UGA female  
## 3919 672961939 Hamza Touba GER male  
## 3960 326027910 Hansol Kim KOR male  
## 3991 542430668 Hasanboy Dusmatov UZB male  
## 3992 741535794 Hasnaa Lachgar MAR female  
## 3994 101286340 Hassan Amzile FRA male  
## 3998 97563452 Hassan Mohamed Mahmoud EGY male  
## 3999 173152080 Hassan Ndam Njikam CMR male  
## 4000 147777095 Hassan Saada MAR male  
## 4001 644229404 Hassan Saaid MDV male  
## 4005 796467757 Hassen Chaktami TUN male  
## 4006 101394753 Havard Haukenes NOR male  
## 4022 136391039 Heather Olver GBR female  
## 4026 677773211 Heba Allejji SYR female  
## 4028 335558674 Hector Luis Garcia Mora DOM male  
## 4030 869886406 Hedda Hynne NOR female  
## 4044 528568211 Heidi Loke NOR female  
## 4047 42902988 Hela Ayari TUN female  
## 4079 136465545 Henrik Christiansen NOR male  
## 4080 706752580 Henrik Ingebrigtsen NOR male  
## 4129 83035378 Hiskel Tewelde ERI male  
## 4149 816766723 Hosam Hussein Bakr Abdin EGY male  
## 4156 61502 Hovhannes Bachkov ARM male  
## 4165 170401459 Hrvoje Sep CRO male  
## 4192 66237669 Hurshid Tojibaev UZB male  
## 4196 2983955 Hussein Iashaish JOR male  
## 4228 909245021 Iakiv Khammo UKR male  
## 4236 372106126 Iaroslava Iakushina RUS female  
## 4241 564969731 Ibrahim Nishwan MDV male  
## 4245 59350988 Ida Alstad NOR female  
## 4260 877477413 Ignacio Perrin ARG male  
## 4272 295973567 Igor Pawel Jakubowski POL male  
## 4307 695779236 Ilyas Abbadi ALG male  
## 4309 587168078 Iman Essa Jasim BRN female  
## 4312 712856100 Imoh Ezekiel NGR male  
## 4313 229646007 Imre Balazs Bacskai HUN male  
## 4328 771770373 Ingeborg Lovnes NOR female  
## 4333 464022421 Ingrit Lorena Valencia Victoria COL female  
## 4335 22127464 Inika McPherson USA female  
## 4337 846310345 Inkululeko Suntele LES male  
## 4340 927315526 Inna Kashyna UKR female  
## 4381 587220722 Irma Testa ITA female  
## 4413 846218313 Isabelle Pedersen NOR female  
## 4432 938396110 Ismail Kamara SLE male  
## 4452 666382432 Ivan Dychko KAZ male  
## 4492 860885364 Izzy Joachim VIN female  
## 4510 637078881 Jackson Kiprop UGA male  
## 4514 144248521 Jacob Araptany UGA male  
## 4522 194290880 Jacob Kiplimo UGA male  
## 4526 807399236 Jacob Rozani RSA male  
## 4542 519260746 Jahvid Best LCA male  
## 4559 654667926 Jakson Vicent Monasterio VEN male  
## 4568 118360942 Jamal Wilson BAH male  
## 4599 531741106 Jamila Lunkuse UGA female  
## 4631 71008427 Janet Amponsah GHA female  
## 4651 565862528 Jared Jarvis ANT male  
## 4657 797374623 Jarlinson Pantano Gomez COL male  
## 4680 42049937 Jason Eric Whateley AUS male  
## 4694 532207877 Javid Chalabiyev AZE male  
## 4711 25728763 Jaysuma Saidy Ndure NOR male  
## 4716 344886355 Jean Paulo Fernandes Filho BRA male  
## 4736 246975227 Jeffery Gibson BAH male  
## 4758 188078227 Jennifer Chieng FSM female  
## 4786 974287219 Jereem Richards TTO male  
## 4792 357577386 Jeremy Dodson SAM male  
## 4835 803585591 Jessica Morrison AUS female  
## 4859 427460731 Jeyvier Jesus Cintron PUR male  
## 4881 715962035 Jianguan Hu CHN male  
## 4890 24639304 Jiawei Zhang CHN male  
## 4928 350253517 Jinson Johnson IND male  
## 4953 818759516 Joahnys Argilagos CUB male  
## 4999 921707760 Joe Joyce GBR male  
## 5006 751008649 Joedison Teixeira BRA male  
## 5023 25172268 Johanna Umurungi RWA female  
## 5040 299026310 John Obi Mikel NGR male  
## 5063 197342499 Jonas Junias Jonas NAM male  
## 5125 728621076 Jorge Luis Vivas COL male  
## 5179 384555082 Joselito Velazquez MEX male  
## 5183 872493079 Joseph Cordina GBR male  
## 5188 610541912 Joseph Ward IRL male  
## 5194 213543405 Josh Kelly GBR male  
## 5199 267338735 Joshua Buatsi GBR male  
## 5201 27887256 Joshua E Tibatemwa UGA male  
## 5204 894182503 Joshua Kiprui Cheptegei UGA male  
## 5215 910659584 Josue Deprez HAI male  
## 5227 17067669 Juan Carlos Carrillo COL male  
## 5240 517352231 Juan Nogueira BRA male  
## 5243 585918319 Juan Pablo Romero MEX male  
## 5258 264395591 Judith Mbougnade CAF female  
## 5299 338982011 Juliao Neto BRA male  
## 5311 950043572 Juliet Chekwel UGA female  
## 5320 333152040 Julio Cesar Castillo ECU male  
## 5325 912637672 Julio Cesar la Cruz CUB male  
## 5333 837661821 Jun Shan CHN male  
## 5342 184749002 Junhua Yin CHN female  
## 5354 957010520 Jurgen Themen SUR male  
## 5394 319193694 Kairat Yeraliyev KAZ male  
## 5402 743593230 Kamil Kuczynski POL male  
## 5407 790596898 Kamongwa Salukombo Makorobondo COD male  
## 5408 372258149 Kamran Shakhsuvarly AZE male  
## 5419 553259590 Kanika Beckles GRN female  
## 5427 717634951 Kara Chad CAN female  
## 5429 788211955 Karabo Sibanda BOT male  
## 5441 339808080 Kari Aalvik Grimsbo NOR female  
## 5476 130728600 Karoline Bjerkeli Grovdal NOR female  
## 5482 687827213 Karsten Warholm NOR male  
## 5502 952109931 Katarzyna Niewiadoma POL female  
## 5507 750498579 Kate Foo kune MRI female  
## 5553 810072016 Katie Taylor IRL female  
## 5562 152567636 Katrine Lunde NOR female  
## 5577 55365531 Kaya Adwoa Forson GHA female  
## 5585 279377870 Kaylin Swart RSA female  
## 5629 576431902 Kellys Arias COL female  
## 5659 347348758 Kennedy Katende UGA male  
## 5660 392095745 Kennedy St Pierre MRI male  
## 5708 828792898 Khadija Mardi MAR female  
## 5728 953232269 Khushbir Kaur IND female  
## 5765 240927532 Kingsley Madu NGR male  
## 5767 656286014 Kira Stepanova RUS female  
## 5771 106579503 Kirani James GRN male  
## 5786 134565248 Kirsty Gilmour GBR female  
## 5794 187677131 Kjetil Borch NOR male  
## 5797 184949058 Klaudia Jans-Ignacik POL female  
## 5803 189267095 Kodo Nakano PHI male  
## 5838 461399217 Krishan Vikas IND male  
## 5844 7457125 Kristian Blummenfelt NOR male  
## 5849 978961185 Kristian Ruth NOR male  
## 5868 957396077 Kristoffer Brun NOR male  
## 5918 141935257 Kyle Greaux TTO male  
## 5952 598281407 Lalit Mathur IND male  
## 5956 486290067 Lanece Clarke BAH female  
## 5979 857905937 Lars Petter Nordhaug NOR male  
## 6001 996823500 Laura Giombini ITA female  
## 6046 272927492 Lauren Smith GBR female  
## 6055 821849198 Laurent Jr. Clayton ISV male  
## 6067 523134751 Lawrence Okolie GBR male  
## 6069 205316686 Lazaro Jorge Alvarez CUB male  
## 6083 495484356 Leandro Blanc ARG male  
## 6088 327424646 Lebenya Nkoka LES male  
## 6107 308992513 Leinier Eunice Pero CUB male  
## 6130 888575799 Leonel de los Santos Nunez DOM male  
## 6156 237908464 Lexi Weeks USA female  
## 6227 644836357 Lindolfo Delgado MEX male  
## 6228 934402077 Lindon Victor GRN male  
## 6242 485795887 Linn-Kristin Riegelhuth Koren NOR female  
## 6309 554615022 Lorenzo Sotomayor Collazo AZE male  
## 6409 138487719 Luis Cabrera VEN male  
## 6419 920379695 Luis Lopez ESA male  
## 6420 851144245 Luis Martin Arcon VEN male  
## 6470 966057256 Luke Bezzina MLT male  
## 6483 905348917 Luvo Manyonga RSA male  
## 6486 409442789 Lydia Jele BOT female  
## 6560 514748025 Mahaman Smaila CMR male  
## 6561 219888752 Mahammadrasul Majidov AZE male  
## 6571 782024861 Mahmoud Abdelaal EGY male  
## 6580 284015405 Maia Agerup NOR female  
## 6626 258262021 Malin Westerheim NOR female  
## 6630 759848453 Malwina Kopron POL female  
## 6641 119339549 Mamina Kone CIV female  
## 6651 878393556 Mandy Bujold CAN female  
## 6663 761454606 Manoj Kumar IND male  
## 6677 621320289 Manuel Cappai ITA male  
## 6711 627766210 Marcelo Alberto Acosta Jimenez ESA male  
## 6752 22207553 Marcus Duncan TTO male  
## 6771 780545996 Margaret Nyairera Wambui KEN female  
## 6786 297444162 Mari Molid NOR female  
## 6824 82454552 Maria Mollestad NOR female  
## 6854 406892958 Mariam Kromah LBR female  
## 6867 267541302 Mariana Ramalho BRA female  
## 6873 321680601 Marianne Skarpnord NOR female  
## 6948 780954637 Marit Malm Frafjord NOR female  
## 7005 786699109 Marlo Javier Delgado ECU male  
## 7083 300217673 Martyna Mikolajczak POL female  
## 7097 104377411 Maryan Muse SOM female  
## 7116 566847633 Masbah Ahmmed BAN male  
## 7129 454434360 Matelita Buadromo FIJ female  
## 7150 152701098 Mathias Tulyoongeleni Hamunyela NAM male  
## 7151 791421238 Mathieu Albert Daniel Bauderlique FRA male  
## 7271 919598411 Mayada Sayyad PLE female  
## 7281 128985176 Mazoon Al-Alawi OMA female  
## 7286 647839865 Meaghan Volker AUS female  
## 7299 463812760 Mehboob Ali PAK male  
## 7303 510649353 Mehmet Nadir Unal TUR male  
## 7321 559761961 Meli Malani FIJ male  
## 7356 726189235 Mercy Cherono KEN female  
## 7364 447162714 Merven Clair MRI male  
## 7392 360677958 Michael John Conlan IRL male  
## 7400 876833914 Michael O,Reilly IRL male  
## 7434 297451205 Michel Borges BRA male  
## 7438 760953520 Michele Scartezzini ITA male  
## 7493 81613069 Mihai Nistor ROU male  
## 7507 2710032 Mikaela Joslin Mayer USA female  
## 7516 109189780 Mikhail Dauhaliavets BLR male  
## 7517 217263774 Mikhail Dovgalyuk RUS male  
## 7536 825689633 Milagro Mena CRC female  
## 7556 806789845 Milly Clark AUS female  
## 7578 44722504 Minggang Zhao CHN male  
## 7586 215136501 Minsoo Park KOR male  
## 7593 793097599 Mira Potkonen FIN female  
## 7616 401258599 Misael Uziel Rodriguez MEX male  
## 7622 355342299 Misha Aloian RUS male  
## 7629 852265607 Mitchell Iles-Crevatin AUS male  
## 7654 219235142 Mohamed Daud Mohamed SOM male  
## 7655 310689410 Mohamed Elhadi Elkawisah LBA male  
## 7658 464902728 Mohamed Flissi ALG male  
## 7659 662693282 Mohamed Hamout MAR male  
## 7662 173845600 Mohamed Hrezi LBA male  
## 7681 855226703 Mohammad Mahfizur Rahman BAN male  
## 7686 416345678 Mohammed Abukhousa PLE male  
## 7691 207078605 Mohammed Arjaoui MAR male  
## 7696 200493972 Mohammed Rabii MAR male  
## 7704 421422970 Molly Goodman AUS female  
## 7745 511776561 Moroke Jeremia Mokhotho LES male  
## 7749 757235037 Moses Martin Kurong UGA male  
## 7757 61885653 Mpi Anauel Ngamissengue CGO male  
## 7760 111349032 Muenfuh Sincere NGR male  
## 7763 545267756 Muhammad Ali GBR male  
## 7772 501615561 Mulern Jean HAI female  
## 7779 703714683 Munyo Solomon Mutai UGA male  
## 7784 411476785 Murodjon Akhmadaliev UZB male  
## 7798 462238804 Mykola Butsenko UKR male  
## 7840 397860366 Najima Parveen PAK female  
## 7842 414759552 Namakoe Nkhasi LES male  
## 7856 834851644 Naomi Flood AUS female  
## 7858 268877038 Naomi Ruele BOT female  
## 7862 230104226 Naomy Grand Pierre HAI female  
## 7869 188253954 Narek Abgaryan ARM male  
## 7885 131987947 Natalia Lovtcova RUS female  
## 7940 440261722 Nathan Byukusenge RWA male  
## 7946 920855247 Nathon Allen JAM male  
## 7961 328956324 Ndifreke Udo NGR male  
## 7962 467504641 Nebiat Habtemariam ERI female  
## 7971 251197905 Nelia Martins TLS female  
## 8029 134647499 Nico Miguel Hernandez USA male  
## 8033 410663669 Nicola Adams GBR female  
## 8078 768436561 Nien-Chin Chen TPE female  
## 8080 970010589 Nigel Paul TTO male  
## 8081 481034512 Nigina Sharipova UZB female  
## 8091 805742930 Nikita Lobintsev RUS male  
## 8118 80208991 Nikolas Sylvester VIN male  
## 8128 677508603 Nils Jakob Hoff NOR male  
## 8149 95604146 Nirmla IND female  
## 8155 399376832 Njisane Phillip TTO male  
## 8179 483970756 Nora Mork NOR female  
## 8188 742257212 Nouchka Fontijn NED female  
## 8206 395961774 Nyakisi Adero UGA female  
## 8211 703472773 Obada Alkasbeh JOR male  
## 8212 381088762 Odbayar Ganbaatar MGL male  
## 8213 2156402 Odd Arne Brekne NOR male  
## 8218 598707040 Oghenekaro Etebo NGR male  
## 8222 800670857 Okechukwu Azubuike NGR male  
## 8228 602321842 Olaf Tufte NOR male  
## 8233 76211755 Ole Kristian Bryhn NOR male  
## 8277 341366636 Olga Zabelinskaya RUS female  
## 8295 929295902 Olivia Ekpone NGR female  
## 8310 299832039 Oluwafemi Ajayi NGR male  
## 8315 168982138 Olzhas Sattibayev KAZ male  
## 8322 671572172 Omar McLeod JAM male  
## 8329 980936428 Onder Sipal TUR male  
## 8334 47810803 Onkabetse Nkobolo BOT male  
## 8338 338111921 Onur Sipal TUR male  
## 8341 10035386 Oreoluwa Cherebin GRN female  
## 8363 324446347 Osman Kamara SLE male  
## 8371 349670408 Otgondalai Dorjnyambuu MGL male  
## 8441 443908875 Pap D. Jonga GAM male  
## 8447 318649476 Parvenn Rana IND male  
## 8448 851811760 Parviz Baghirov AZE male  
## 8454 52489905 Pat McCormack GBR male  
## 8470 703130921 Patrick Barnes IRL male  
## 8476 319404194 Patrick Lourenco BRA male  
## 8504 674380639 Paul Omba Biongolo FRA male  
## 8540 286853003 Pavel Kastramin BLR male  
## 8547 663676373 Pavel Sozykin RUS male  
## 8558 196109685 Pawel Juraszek POL male  
## 8567 34305365 Peamwilai Laopeam THA female  
## 8583 202596543 Pedrya Seymour BAH female  
## 8602 923556859 Peruth Chemutai UGA female  
## 8622 788615502 Peter Mullenberg NED male  
## 8623 730243242 Peter Mungai Warui KEN male  
## 8636 211340599 Petr Khamukov RUS male  
## 8653 729088638 Phetetso Monese LES male  
## 8674 330465212 Phillip Kipyeko UGA male  
## 8747 327752493 Popoola Saliu NGR male  
## 8759 534519239 Prenam Pesse TOG female  
## 8768 513769796 Qais Ashfaq GBR male  
## 8773 497917027 Qian Li CHN female  
## 8781 212028981 Qianxun Hu CHN male  
## 8797 910962257 R. Mohan Kumar IND male  
## 8830 825109201 Rafael Andrade BRA male  
## 8853 913614600 Ragna Agerup NOR female  
## 8858 296185766 Raijieli Daveua FIJ female  
## 8862 818043080 Rajeev Ram USA male  
## 8863 465151965 Rajiv Ouseph GBR male  
## 8900 595425086 Raul Curiel MEX male  
## 8914 372384369 Rayton Nduku Okwiri KEN male  
## 8916 35093364 Rebeca Quinteros Ortiz ESA female  
## 8934 874150353 Reda Benbaziz ALG male  
## 8988 248931578 Rexford Tullius ISV male  
## 9029 582543706 Richardson Hitchins HAI male  
## 9070 58169707 Robeilys Peinado VEN female  
## 9071 921399403 Robeisy Ramirez CUB male  
## 9073 371852297 Robenilson de Jesus BRA male  
## 9124 938744484 Robson Conceicao BRA male  
## 9138 480321541 Rodrick Kuku COD male  
## 9148 992517041 Rogen Ladon PHI male  
## 9186 106685753 Ronald Forbes CAY male  
## 9188 244164442 Ronald Musagala UGA male  
## 9190 115983009 Ronald Serugo UGA male  
## 9196 151465997 Roniel Iglesias CUB male  
## 9209 686173584 Rose Chelimo BRN female  
## 9215 287037985 Rosefelo Siosi SOL male  
## 9255 334846524 Rufat Huseynov AZE male  
## 9286 822684944 Rustam Tulaganov UZB male  
## 9310 788805861 Ryan Held USA male  
## 9346 966020626 Sadiq Umar NGR male  
## 9352 436856323 Safa Saidani TUN female  
## 9358 627433948 Saidi Juma Makula TAN male  
## 9379 11169898 Sally Yee FIJ female  
## 9382 646360332 Salome Nyirarukundo RWA female  
## 9395 826848760 Sam Groth AUS male  
## 9406 144477923 Samantha Birch AUS female  
## 9426 992495912 Samuel Carmona Heredia ESP male  
## 9446 740295602 Sandeep Kumar IND male  
## 9473 15622309 Sangmyeong Ham KOR male  
## 9479 385623198 Sanna Solberg NOR female  
## 9500 451333617 Sapana Sapana IND female  
## 9519 764316889 Sara Ramadhani TAN female  
## 9526 717386326 Sarah Banting AUS female  
## 9542 950602968 Sarah Ourahmoune FRA female  
## 9573 458458780 Saturday Erimuya NGR male  
## 9574 742022352 Saud Alzaabi UAE male  
## 9581 649115184 Savannah Marshall GBR female  
## 9587 436141773 Saylom Ardee THA male  
## 9645 23409142 Selcuk Eker TUR male  
## 9657 743927516 Senbere Teferi ETH female  
## 9670 270678554 Serge Michel GER male  
## 9730 294616056 Shakhobidin Zoirov UZB male  
## 9731 295291162 Shakhram Giyasov UZB male  
## 9734 682806875 Shakur Stevenson USA male  
## 9737 551301673 Shamoli Ray BAN female  
## 9755 353324453 Shaquania Dorsett BAH female  
## 9763 509787602 Sharon Firisua SOL female  
## 9768 891682431 Shavez Hart BAH male  
## 9782 425110382 Shelley Marie Watts AUS female  
## 9810 229032173 Shirin Akter BAN female  
## 9812 652257369 Shiva Thapa IND male  
## 9852 219470376 Signe Marie Fidge Store NOR female  
## 9869 578898700 Simeon Chamov BUL male  
## 9873 910071832 Simon Clarke AUS male  
## 9908 203750302 Simplice Fotsala CMR male  
## 9923 100364594 Sisila Seavula FIJ female  
## 9953 496243437 Sofiane Guitone FRA male  
## 9954 147169766 Sofiane Oumiha FRA male  
## 9968 956809483 Sondre Nordstad Moen NOR male  
## 9974 388561246 Sonia Aktar BAN female  
## 10000 510174906 Soren Opti SUR male  
## 10011 306264862 Soukphaxay Sithisane LAO male  
## 10013 298080547 Souleymane Diop Cissokho FRA male  
## 10032 670857265 Stan Okoye NGR male  
## 10033 217337729 Stanimira Petrova BUL female  
## 10035 574714431 Stanley Amuzie NGR male  
## 10079 305158735 Stella Chesang UGA female  
## 10084 891602037 Stephan de Freitas Barcha BRA male  
## 10106 82771185 Stephen Kiprotich UGA male  
## 10111 769819070 Stephen Newbold BAH male  
## 10119 590851106 Steven Gardiner BAH male  
## 10120 917340321 Steven Gerard Donnelly IRL male  
## 10127 581997901 Stig-Andre Berge NOR male  
## 10131 945304583 Stine Bredal Oftedal NOR female  
## 10140 54125373 Su Oh AUS female  
## 10161 67222900 Sunayna Wahi SUR female  
## 10185 819679922 Suzann Pettersen NOR female  
## 10186 756706296 Suzanne Hearn AUS female  
## 10189 167125216 Sven Erik Bystrom NOR male  
## 10192 599845580 Sven Martin Skagestad NOR male  
## 10205 237017632 Svitlana Stanko-Klymenko UKR female  
## 10250 808567905 Tales Cerdeira BRA male  
## 10271 845644172 Tameka Williams SKN female  
## 10338 849242377 Taylor Ellis-Watson USA female  
## 10358 675969120 Teofimo Andres Lopez Rivera HON male  
## 10360 654886163 Teray Smith BAH male  
## 10383 649148509 Tetyana Kob UKR female  
## 10384 159812118 Tewelde Estifanos ERI male  
## 10385 807447097 Teymur Mammadov AZE male  
## 10388 210040104 Thadius Katua PNG male  
## 10462 962042754 Thulasi Tharumalingam QAT male  
## 10484 939411758 Tigest Getent BRN female  
## 10485 608439897 Tigist Gashaw BRN female  
## 10486 775831149 Tigist Tufa ETH female  
## 10516 114387888 Timothy Toroitich UGA male  
## 10527 390366123 Tina Skaar NOR female  
## 10533 975655821 Ting Ying Huang TPE female  
## 10540 727998449 Tiril Bue NOR female  
## 10581 324400159 Tomasz Jablonski POL male  
## 10600 724384082 Tonje Angelsen NOR female  
## 10609 34099719 Tony Victor James Yoka FRA male  
## 10611 764236472 Tore Navrestad NOR male  
## 10626 244396184 Trevor Barry BAH male  
## 10635 49752494 Tsendbaatar Erdenebat MGL male  
## 10636 646516993 Tsepang Sello LES female  
## 10637 132965685 Tsepo Mathibelle LES male  
## 10657 300389824 Tuvshinbat Byamba MGL male  
## 10665 77544133 Tynia Gaither BAH female  
## 10675 617399912 Uijo Hwang KOR male  
## 10698 39407530 Usman Muhammed NGR male  
## 10713 655768386 Vaidas Kariniauskas LTU male  
## 10716 839835613 Valdivia BRA male  
## 10733 125173784 Valentino Manfredonia ITA male  
## 10764 937539913 Vasilii Egorov RUS male  
## 10771 235882258 Vassiliy Levit KAZ male  
## 10775 655301496 Vegard Stake Laengen NOR male  
## 10791 386589042 Veronica Kristiansen NOR female  
## 10815 70687418 Victor Rodriguez VEN male  
## 10876 983925141 Vincenzo Mangiacapre ITA male  
## 10882 989120045 Violah Cheptoo Lagat KEN female  
## 10889 930876031 Vita Heine NOR female  
## 10894 806574554 Vitaly Dunaytsev RUS male  
## 10917 229367719 Vladimir Margaryan ARM male  
## 10919 108548232 Vladimir Morozov RUS male  
## 10920 610704045 Vladimir Nikitin RUS male  
## 10937 204414092 Volodymyr Matviichuk UKR male  
## 10943 539224149 Waheed Abdulridha Waheed Karaawi IRQ male  
## 10947 422425972 Walid Mohamed EGY male  
## 10964 450932930 Wei Liu CHN male  
## 11015 213332524 William Ekong NGR male  
## 11029 630860281 Winnie Nanyondo UGA female  
## 11032 404382792 Winston Hill FIJ male  
## 11047 640901977 Wuileixis De Jesus Rivas Espinoza VEN male  
## 11049 332172731 Wuttichai Masuk THA male  
## 11122 380260850 Yamil Alberto Peralta ARG male  
## 11127 76793158 Yana Alekseevna AZE female  
## 11180 893193167 Yasnier Toledo CUB male  
## 11209 550672127 Yemane Haileselassie ERI male  
## 11221 5964702 Yesenia Miranda ESA female  
## 11239 390260480 Yigal Kopinsky SUR male  
## 11261 540772965 Yodgoroy Mirzaeva UZB female  
## 11262 607223536 Yoel Segundo Finol VEN male  
## 11292 693270892 Yosbany Veitia CUB male  
## 11303 942438156 Youba Sissokho Ndiaye ESP male  
## 11371 86099624 Yulia Efimova RUS female  
## 11399 57329234 Yurberjen Herney Martinez COL male  
## 11466 980866226 Zhaina Shekerbekova KAZ female  
## 11469 405483135 Zhanibek Alimkhanuly KAZ male  
## 11514 352983859 Zohir Kedache ALG male  
## 11515 424586991 Zohra Ez Zahraoui MAR female  
## 11518 495365659 Zoltan Adam Harcsa HUN male  
## dob height weight sport gold silver bronze  
## 13 10/11/1992 NA NA aquatics 0 0 0  
## 29 09/26/1986 1.86 NA boxing 0 0 0  
## 31 12/12/1986 1.78 NA boxing 0 0 0  
## 38 10/09/1987 1.85 NA boxing 0 0 0  
## 48 06/08/1986 NA NA athletics 0 0 0  
## 51 10/03/1993 NA NA boxing 0 0 0  
## 53 10/30/1990 NA NA fencing 0 0 0  
## 54 07/17/1988 1.88 NA boxing 0 0 0  
## 57 08/01/1989 NA NA shooting 0 0 0  
## 58 03/12/1993 1.70 NA football 0 0 1  
## 60 06/24/1986 NA NA shooting 0 0 0  
## 63 04/12/2000 NA NA aquatics 0 0 0  
## 67 01/10/1997 1.75 NA boxing 0 0 0  
## 73 09/25/1990 1.60 NA boxing 0 0 0  
## 95 06/10/1993 NA NA athletics 0 0 0  
## 109 05/19/1989 1.80 NA boxing 0 1 0  
## 112 07/31/1990 1.72 NA boxing 0 0 0  
## 130 11/04/1981 1.67 NA boxing 0 0 0  
## 140 11/01/1993 NA NA athletics 0 0 0  
## 166 01/10/1983 NA NA shooting 0 0 0  
## 175 07/30/1995 NA NA aquatics 0 0 0  
## 181 01/22/1991 NA NA aquatics 0 0 0  
## 247 08/11/1996 NA NA rowing 0 0 0  
## 248 07/10/1996 NA NA athletics 0 0 0  
## 269 05/07/1992 1.82 NA boxing 0 0 0  
## 272 04/05/1986 1.71 NA boxing 0 0 0  
## 278 01/02/1990 1.70 NA boxing 0 0 0  
## 280 04/05/1990 1.75 NA boxing 0 0 0  
## 323 05/09/1994 NA NA cycling 0 0 0  
## 351 03/02/1992 NA NA athletics 0 0 0  
## 366 11/17/1991 NA NA athletics 0 0 0  
## 389 12/31/1996 NA 67 athletics 0 0 0  
## 445 03/28/1990 NA NA athletics 0 0 0  
## 459 01/31/1997 NA NA archery 0 0 0  
## 460 04/02/1990 1.89 NA boxing 0 0 0  
## 553 02/08/1992 NA NA athletics 0 0 1  
## 555 02/14/1992 NA 55 athletics 0 0 0  
## 578 04/17/1994 NA NA athletics 0 0 0  
## 587 07/25/1991 1.75 NA handball 0 0 1  
## 605 12/31/1980 NA NA athletics 0 0 0  
## 608 03/26/1989 1.63 NA basketball 0 0 0  
## 625 10/29/1993 NA NA aquatics 0 0 0  
## 629 03/06/1995 1.72 NA football 0 0 1  
## 637 12/18/1979 1.64 NA boxing 0 0 0  
## 682 07/19/1993 NA NA athletics 0 0 0  
## 692 05/12/1982 NA NA tennis 0 0 0  
## 696 05/01/1993 1.73 NA boxing 0 0 1  
## 728 05/04/1992 NA NA sailing 0 0 0  
## 801 07/04/1987 1.82 NA boxing 0 0 0  
## 802 05/03/1987 1.69 NA boxing 0 0 0  
## 831 12/15/1991 NA NA athletics 0 0 0  
## 853 12/14/1986 NA NA cycling 0 0 0  
## 856 11/29/1986 NA NA cycling 0 0 0  
## 860 06/02/1993 1.75 NA boxing 0 0 0  
## 909 05/06/1993 NA NA athletics 0 0 0  
## 952 02/12/1980 1.84 NA boxing 0 0 0  
## 982 01/11/1981 NA NA canoe 0 0 0  
## 1017 11/18/1992 NA NA aquatics 0 0 0  
## 1067 08/15/1996 1.68 NA boxing 0 0 0  
## 1069 03/10/1991 1.79 NA boxing 0 0 0  
## 1076 02/01/1987 1.69 NA boxing 0 0 0  
## 1086 08/15/1992 1.80 NA boxing 0 0 0  
## 1088 01/18/1991 1.73 NA boxing 0 0 0  
## 1093 07/02/1996 1.76 NA boxing 0 0 0  
## 1096 01/16/1982 NA NA shooting 0 0 0  
## 1097 08/18/1988 NA NA rowing 0 0 1  
## 1106 11/20/1984 1.75 NA boxing 0 0 0  
## 1114 12/15/1989 NA NA tennis 0 0 0  
## 1120 02/21/1993 1.78 NA boxing 1 0 0  
## 1133 02/20/1998 NA NA athletics 0 0 0  
## 1138 07/01/1993 1.77 NA boxing 0 0 0  
## 1140 10/26/1988 1.84 NA boxing 0 0 0  
## 1142 08/13/1990 1.72 NA boxing 0 0 1  
## 1150 04/22/1995 1.74 NA boxing 0 0 0  
## 1160 03/23/1996 1.60 NA boxing 0 0 0  
## 1192 05/11/1989 1.58 NA golf 0 0 0  
## 1227 04/06/1989 1.75 NA boxing 0 0 0  
## 1249 08/22/1986 1.65 NA athletics 0 0 0  
## 1270 08/05/1986 NA NA rugby sevens 0 0 0  
## 1285 08/09/1982 1.80 NA judo 0 0 0  
## 1299 03/18/1997 NA NA athletics 0 0 0  
## 1306 07/08/1994 1.98 NA boxing 0 0 0  
## 1358 12/20/1990 1.72 NA boxing 0 0 0  
## 1372 02/17/1987 NA 53 athletics 0 0 0  
## 1373 11/20/1997 NA NA athletics 0 0 0  
## 1393 04/08/1996 1.77 NA boxing 0 1 0  
## 1430 03/04/1989 NA NA athletics 0 0 0  
## 1441 11/08/1992 NA NA boxing 0 0 0  
## 1443 05/03/1985 1.68 NA boxing 0 0 0  
## 1445 06/20/1986 1.65 NA boxing 0 0 0  
## 1477 05/17/1988 NA NA athletics 0 0 0  
## 1485 01/27/1992 1.76 NA boxing 0 0 0  
## 1494 10/18/1994 1.65 NA boxing 0 0 0  
## 1502 02/27/1996 NA NA athletics 0 0 0  
## 1504 07/07/1984 1.63 NA boxing 0 0 0  
## 1535 01/26/1995 NA NA boxing 0 0 0  
## 1541 08/05/1995 NA NA athletics 0 0 0  
## 1583 05/08/1992 NA NA athletics 0 0 0  
## 1606 05/17/1996 1.67 NA boxing 0 0 0  
## 1654 07/12/1993 NA NA athletics 0 0 0  
## 1681 05/10/1998 NA 79 aquatics 0 0 0  
## 1702 02/05/1991 NA NA athletics 0 0 0  
## 1725 10/08/1986 1.67 NA handball 0 0 1  
## 1737 01/26/1988 1.68 NA boxing 0 0 1  
## 1762 04/30/1992 1.87 NA boxing 0 0 0  
## 1769 05/17/1990 1.56 NA boxing 0 0 0  
## 1780 08/24/1996 1.76 NA boxing 0 0 0  
## 1790 05/16/1995 NA NA athletics 0 0 0  
## 1791 03/30/1984 1.67 NA boxing 0 0 0  
## 1816 05/06/1991 NA NA athletics 0 0 0  
## 1847 10/10/1994 NA NA aquatics 0 0 0  
## 1857 01/13/1995 NA NA athletics 0 0 0  
## 1859 05/26/1989 1.63 NA boxing 0 0 0  
## 1894 08/30/1993 NA NA football 0 0 0  
## 1916 11/02/1997 1.76 NA boxing 0 0 0  
## 1946 08/14/1988 1.68 NA boxing 0 0 0  
## 1948 03/26/1985 1.66 NA boxing 0 0 0  
## 1952 09/19/1993 1.90 NA taekwondo 1 0 0  
## 1992 09/21/1991 1.74 NA boxing 0 0 0  
## 1993 10/04/1984 NA NA athletics 0 0 0  
## 2012 01/08/1987 1.82 NA boxing 0 0 0  
## 2024 05/02/1985 1.80 NA badminton 0 0 1  
## 2047 04/26/1995 1.75 NA boxing 0 0 0  
## 2048 01/01/1994 NA NA aquatics 0 0 0  
## 2056 12/12/1961 NA NA equestrian 0 0 0  
## 2065 10/01/1995 NA NA athletics 0 0 0  
## 2085 12/25/1983 NA NA judo 0 0 0  
## 2098 07/23/1996 1.68 NA boxing 0 0 0  
## 2128 03/17/1995 1.76 NA boxing 1 0 0  
## 2155 07/27/1982 1.81 NA boxing 0 0 0  
## 2190 03/16/1997 NA NA aquatics 0 0 0  
## 2230 07/01/1988 NA NA sailing 0 0 0  
## 2257 12/14/1989 1.72 NA boxing 0 0 0  
## 2303 02/07/1999 NA NA aquatics 0 0 0  
## 2305 03/08/1986 1.87 NA football 0 0 1  
## 2308 05/17/1997 1.63 NA boxing 0 0 0  
## 2329 12/18/1993 1.78 NA boxing 0 0 0  
## 2349 11/10/1988 NA NA cycling 0 0 0  
## 2384 03/13/1991 1.78 NA boxing 1 0 0  
## 2411 08/29/1998 NA NA aquatics 0 0 0  
## 2415 11/20/1988 1.72 NA boxing 0 0 1  
## 2462 01/12/1989 1.91 NA boxing 0 0 0  
## 2483 02/12/1987 1.71 NA boxing 0 0 0  
## 2498 04/29/1997 NA NA aquatics 0 0 0  
## 2502 02/03/1989 1.88 NA boxing 0 0 0  
## 2513 05/19/1995 NA NA athletics 0 0 0  
## 2554 02/01/2000 NA NA aquatics 0 0 0  
## 2565 10/25/1992 1.79 NA boxing 0 0 0  
## 2573 12/29/1987 NA NA judo 0 0 0  
## 2623 01/23/1998 1.83 NA boxing 0 0 0  
## 2658 04/18/1995 NA NA boxing 0 0 0  
## 2676 11/08/1989 1.73 NA boxing 0 0 0  
## 2679 11/08/1996 NA NA volleyball 0 0 0  
## 2703 09/23/1989 1.66 NA gymnastics 0 0 0  
## 2720 02/02/1992 NA NA athletics 0 0 0  
## 2761 10/07/1987 1.98 NA basketball 0 0 0  
## 2764 01/30/1996 NA NA aquatics 0 0 0  
## 2771 05/18/1996 1.76 NA boxing 0 0 0  
## 2790 12/22/1983 1.84 NA boxing 0 0 0  
## 2810 05/17/1987 NA NA cycling 0 0 0  
## 2829 04/22/1994 1.90 NA boxing 0 0 0  
## 2835 06/23/1988 1.84 NA boxing 0 0 0  
## 2836 12/13/1992 NA NA aquatics 0 0 0  
## 2843 08/17/1994 1.73 NA boxing 0 0 0  
## 2846 10/07/1996 NA 70 athletics 0 0 0  
## 2864 05/10/1988 NA NA athletics 0 0 0  
## 2878 09/18/1994 NA 95 athletics 0 0 0  
## 2880 03/25/1977 1.54 NA shooting 0 0 0  
## 2887 12/08/1993 NA NA canoe 0 0 0  
## 2892 11/13/1989 1.77 NA shooting 0 0 0  
## 2910 06/13/1991 1.65 NA boxing 0 0 0  
## 2912 04/06/1992 1.70 NA boxing 0 0 0  
## 2978 10/13/1992 NA NA equestrian 0 0 0  
## 2984 01/01/1995 NA NA aquatics 0 0 0  
## 2990 03/07/1986 1.84 NA boxing 0 0 0  
## 2993 12/17/1991 1.65 NA boxing 0 0 0  
## 3009 05/29/1999 NA NA aquatics 0 0 0  
## 3026 01/01/1994 1.83 NA handball 0 0 1  
## 3030 05/13/1992 NA NA football 0 0 0  
## 3058 12/17/1993 1.74 NA football 0 0 1  
## 3059 09/14/1995 NA NA athletics 0 0 0  
## 3062 04/15/1991 NA NA athletics 0 0 0  
## 3068 05/14/1991 1.80 NA boxing 0 0 0  
## 3072 08/05/1992 1.70 NA boxing 0 0 0  
## 3075 08/31/1993 1.81 NA boxing 0 0 0  
## 3108 01/22/1987 1.91 NA boxing 0 0 0  
## 3110 12/04/1980 NA NA athletics 0 0 0  
## 3131 07/21/1990 1.92 NA boxing 0 0 1  
## 3132 02/14/1991 1.90 NA boxing 0 0 0  
## 3147 08/11/1987 NA NA golf 0 0 0  
## 3160 08/19/1992 1.68 NA boxing 1 0 0  
## 3178 03/05/1999 1.48 NA gymnastics 0 0 0  
## 3198 03/19/1992 1.64 NA boxing 0 0 0  
## 3213 12/24/1995 NA NA athletics 0 0 0  
## 3230 07/15/1991 1.96 NA boxing 1 0 0  
## 3233 07/05/1995 NA NA aquatics 0 0 0  
## 3243 03/03/1988 NA NA athletics 0 0 0  
## 3259 12/24/1985 NA 50 athletics 0 0 0  
## 3271 11/20/1992 NA NA athletics 0 0 0  
## 3282 03/07/1992 1.78 NA boxing 0 0 0  
## 3283 05/01/1988 1.96 NA judo 0 0 0  
## 3309 11/18/1997 NA NA athletics 0 0 0  
## 3312 10/20/1994 NA NA aquatics 0 0 0  
## 3313 11/23/1993 NA NA judo 0 0 0  
## 3314 06/16/1991 1.69 NA boxing 1 0 0  
## 3335 10/09/1990 NA NA equestrian 0 0 0  
## 3342 03/12/1997 NA NA football 0 0 0  
## 3366 03/13/1995 1.93 NA boxing 0 0 0  
## 3386 07/18/1991 1.65 NA boxing 0 0 0  
## 3393 07/09/1978 NA NA football 0 0 0  
## 3404 06/04/1992 1.97 NA boxing 0 0 1  
## 3405 04/20/1993 NA NA athletics 0 0 0  
## 3433 10/16/1988 NA NA athletics 0 0 0  
## 3462 03/23/1985 1.65 NA cycling 0 0 0  
## 3514 11/26/1994 1.92 NA judo 0 0 0  
## 3541 09/22/1986 1.79 NA boxing 0 0 0  
## 3557 04/30/1999 NA NA aquatics 0 0 0  
## 3561 09/30/1990 1.67 NA badminton 0 0 0  
## 3577 12/11/1992 1.58 NA boxing 0 0 0  
## 3579 04/01/1988 NA NA table tennis 0 0 0  
## 3583 03/29/1993 1.62 NA boxing 0 0 0  
## 3584 06/29/1989 NA NA athletics 0 0 0  
## 3590 05/07/1991 NA NA athletics 0 0 0  
## 3595 06/14/1996 1.73 NA boxing 0 0 0  
## 3690 11/14/1995 NA NA athletics 0 0 0  
## 3712 05/03/1995 NA NA athletics 0 0 0  
## 3760 12/03/1993 NA NA athletics 0 0 0  
## 3778 08/19/1993 NA NA athletics 0 0 0  
## 3829 05/09/1994 1.98 NA boxing 0 0 0  
## 3843 02/10/1973 NA NA cycling 0 0 0  
## 3856 03/24/1993 NA NA football 0 0 0  
## 3860 09/11/1995 NA NA football 0 0 0  
## 3880 12/07/1991 NA NA athletics 0 0 0  
## 3895 07/05/1994 1.71 NA boxing 0 0 0  
## 3901 10/16/1994 NA NA athletics 0 0 0  
## 3919 11/06/1991 1.70 NA boxing 0 0 0  
## 3960 12/29/1995 1.65 NA gymnastics 0 0 0  
## 3991 06/24/1993 1.56 NA boxing 1 0 0  
## 3992 09/27/1989 1.64 NA boxing 0 0 0  
## 3994 06/05/1988 1.83 NA boxing 0 0 0  
## 3998 02/10/1984 NA NA athletics 0 0 0  
## 3999 02/18/1984 1.82 NA boxing 0 0 0  
## 4000 01/02/1994 1.83 NA boxing 0 0 0  
## 4001 03/04/1992 NA NA athletics 0 0 0  
## 4005 12/14/1988 1.80 NA boxing 0 0 0  
## 4006 04/22/1990 NA NA athletics 0 0 0  
## 4022 03/15/1986 1.70 NA badminton 0 0 0  
## 4026 01/20/1997 NA NA table tennis 0 0 0  
## 4028 11/01/1991 1.73 NA boxing 0 0 0  
## 4030 03/13/1990 NA NA athletics 0 0 0  
## 4044 12/12/1982 1.73 NA handball 0 0 1  
## 4047 08/26/1994 NA 52 judo 0 0 0  
## 4079 10/09/1996 NA NA aquatics 0 0 0  
## 4080 02/24/1991 NA NA athletics 0 0 0  
## 4129 09/15/1986 NA NA athletics 0 0 0  
## 4149 10/26/1985 1.80 NA boxing 0 0 0  
## 4156 12/02/1992 1.73 NA boxing 0 0 0  
## 4165 02/26/1986 1.75 NA boxing 0 0 0  
## 4192 11/13/1989 1.72 NA boxing 0 0 0  
## 4196 08/06/1995 1.80 NA boxing 0 0 0  
## 4228 06/11/1994 1.88 NA judo 0 0 0  
## 4236 06/24/1993 1.70 NA boxing 0 0 0  
## 4241 06/12/1997 NA NA aquatics 0 0 0  
## 4245 06/13/1985 1.72 NA handball 0 0 1  
## 4260 01/20/1985 1.72 NA boxing 0 0 0  
## 4272 08/06/1992 1.84 NA boxing 0 0 0  
## 4307 10/21/1992 1.85 NA boxing 0 0 0  
## 4309 07/09/1997 NA NA athletics 0 0 0  
## 4312 10/24/1993 1.71 NA football 0 0 1  
## 4313 01/29/1988 1.77 NA boxing 0 0 0  
## 4328 09/05/1992 NA NA athletics 0 0 0  
## 4333 09/03/1988 1.60 NA boxing 0 0 1  
## 4335 09/29/1986 NA NA athletics 0 0 0  
## 4337 04/30/1994 NA NA boxing 0 0 0  
## 4340 09/27/1991 NA NA athletics 0 0 0  
## 4381 12/28/1997 1.74 NA boxing 0 0 0  
## 4413 01/27/1992 NA NA athletics 0 0 0  
## 4432 02/14/1997 NA NA athletics 0 0 0  
## 4452 08/11/1990 2.06 NA boxing 0 0 1  
## 4492 05/11/2000 NA NA aquatics 0 0 0  
## 4510 10/20/1986 NA NA athletics 0 0 0  
## 4514 02/11/1994 NA NA athletics 0 0 0  
## 4522 11/14/2000 NA NA athletics 0 0 0  
## 4526 01/24/1988 NA NA athletics 0 0 0  
## 4542 01/30/1989 NA NA athletics 0 0 0  
## 4559 12/31/1991 NA NA rowing 0 0 0  
## 4568 09/01/1988 NA NA athletics 0 0 0  
## 4599 06/01/1997 NA NA aquatics 0 0 0  
## 4631 04/12/1993 1.71 NA athletics 0 0 0  
## 4651 08/29/1994 NA NA athletics 0 0 0  
## 4657 11/19/1988 NA NA cycling 0 0 0  
## 4680 11/18/1990 1.96 NA boxing 0 0 0  
## 4694 05/17/1992 1.70 NA boxing 0 0 0  
## 4711 01/01/1984 NA NA athletics 0 0 0  
## 4716 10/26/1995 NA NA football 0 0 0  
## 4736 08/15/1990 NA NA athletics 0 0 0  
## 4758 04/29/1986 1.61 NA boxing 0 0 0  
## 4786 01/13/1994 1.89 NA athletics 0 0 0  
## 4792 08/30/1987 NA NA athletics 0 0 0  
## 4835 05/18/1992 NA NA rowing 0 0 0  
## 4859 02/08/1995 1.71 NA boxing 0 0 0  
## 4881 05/11/1993 1.68 NA boxing 0 0 1  
## 4890 01/08/1989 1.72 NA boxing 0 0 0  
## 4928 03/15/1991 NA NA athletics 0 0 0  
## 4953 01/11/1997 1.52 NA boxing 0 0 1  
## 4999 09/19/1985 1.97 NA boxing 0 1 0  
## 5006 01/28/1994 1.72 NA boxing 0 0 0  
## 5023 04/07/1996 NA NA aquatics 0 0 0  
## 5040 04/22/1987 1.73 NA football 0 0 1  
## 5063 11/24/1993 1.81 NA boxing 0 0 0  
## 5125 01/22/1988 1.72 NA boxing 0 0 0  
## 5179 09/30/1993 1.64 NA boxing 0 0 0  
## 5183 12/01/1991 1.75 NA boxing 0 0 0  
## 5188 10/30/1993 1.86 NA boxing 0 0 0  
## 5194 03/07/1994 1.78 NA boxing 0 0 0  
## 5199 03/14/1993 1.85 NA boxing 0 0 1  
## 5201 09/10/1996 NA NA aquatics 0 0 0  
## 5204 09/12/1996 NA NA athletics 0 0 0  
## 5215 10/30/1981 NA NA judo 0 0 0  
## 5227 10/10/1992 1.84 NA boxing 0 0 0  
## 5240 05/01/1988 1.86 NA boxing 0 0 0  
## 5243 01/30/1990 1.80 NA boxing 0 0 0  
## 5258 07/11/1998 1.50 NA boxing 0 0 0  
## 5299 08/16/1981 1.65 NA boxing 0 0 0  
## 5311 03/25/1990 NA NA athletics 0 0 0  
## 5320 05/10/1988 1.83 NA boxing 0 0 0  
## 5325 08/11/1989 1.82 NA boxing 1 0 0  
## 5333 08/07/1994 1.72 NA boxing 0 0 0  
## 5342 08/27/1990 1.67 NA boxing 0 1 0  
## 5354 10/26/1985 NA NA athletics 0 0 0  
## 5394 11/08/1990 1.65 NA boxing 0 0 0  
## 5402 03/23/1985 NA NA cycling 0 0 0  
## 5407 08/13/1988 NA NA athletics 0 0 0  
## 5408 12/06/1992 1.80 NA boxing 0 0 1  
## 5419 10/03/1991 NA NA athletics 0 0 0  
## 5427 01/09/1996 NA NA equestrian 0 0 0  
## 5429 07/02/1998 NA NA athletics 0 0 0  
## 5441 01/04/1985 1.80 NA handball 0 0 1  
## 5476 06/14/1990 NA NA athletics 0 0 0  
## 5482 02/28/1996 NA NA athletics 0 0 0  
## 5502 09/29/1994 NA NA cycling 0 0 0  
## 5507 03/29/1993 NA NA badminton 0 0 0  
## 5553 07/01/1986 1.62 NA boxing 0 0 0  
## 5562 03/30/1980 1.81 NA handball 0 0 1  
## 5577 03/19/2002 NA NA aquatics 0 0 0  
## 5585 09/30/1994 NA NA football 0 0 0  
## 5629 07/03/1989 NA 150 athletics 0 0 0  
## 5659 03/15/1985 1.76 NA boxing 0 0 0  
## 5660 10/23/1992 1.80 NA boxing 0 0 0  
## 5708 02/01/1991 1.83 NA boxing 0 0 0  
## 5728 07/09/1993 NA NA athletics 0 0 0  
## 5765 12/12/1995 1.71 NA football 0 0 1  
## 5767 11/12/1993 NA NA canoe 0 0 0  
## 5771 09/01/1992 NA 66 athletics 0 1 0  
## 5786 09/21/1993 1.68 NA badminton 0 0 0  
## 5794 02/14/1990 NA NA rowing 0 0 1  
## 5797 09/24/1984 NA NA tennis 0 0 0  
## 5803 03/08/1993 NA 81 judo 0 0 0  
## 5838 02/10/1992 1.77 NA boxing 0 0 0  
## 5844 02/14/1994 NA NA triathlon 0 0 0  
## 5849 07/23/1985 NA NA sailing 0 0 0  
## 5868 04/07/1988 NA NA rowing 0 0 1  
## 5918 04/26/1988 NA NA athletics 0 0 0  
## 5952 12/18/1994 NA NA athletics 0 0 0  
## 5956 11/04/1987 NA NA athletics 0 0 0  
## 5979 05/14/1984 NA NA cycling 0 0 0  
## 6001 01/04/1989 1.83 NA volleyball 0 0 0  
## 6046 09/26/1991 1.70 NA badminton 0 0 0  
## 6055 07/18/1990 NA NA boxing 0 0 0  
## 6067 12/16/1992 1.95 NA boxing 0 0 0  
## 6069 01/28/1991 1.73 NA boxing 0 0 1  
## 6083 05/02/1993 1.65 NA boxing 0 0 0  
## 6088 10/19/1982 NA NA athletics 0 0 0  
## 6107 11/24/1992 1.92 NA boxing 0 0 0  
## 6130 12/14/1994 1.70 NA boxing 0 0 0  
## 6156 11/20/1996 NA NA athletics 0 0 0  
## 6227 12/31/1994 1.74 NA boxing 0 0 0  
## 6228 02/28/1993 NA NA athletics 0 0 0  
## 6242 08/01/1984 1.75 NA handball 0 0 1  
## 6309 02/16/1985 1.85 NA boxing 0 1 0  
## 6409 05/20/1995 1.76 NA boxing 0 0 0  
## 6419 01/18/1994 NA NA athletics 0 0 0  
## 6420 06/01/1992 1.79 NA boxing 0 0 0  
## 6470 06/07/1995 NA NA athletics 0 0 0  
## 6483 01/08/1991 NA NA athletics 0 1 0  
## 6486 06/22/1990 NA NA athletics 0 0 0  
## 6560 02/28/1986 1.78 NA boxing 0 0 0  
## 6561 09/27/1986 1.90 NA boxing 0 0 0  
## 6571 01/01/1992 1.76 NA boxing 0 0 0  
## 6580 06/22/1995 NA NA sailing 0 0 0  
## 6626 11/10/1993 NA NA shooting 0 0 0  
## 6630 11/16/1994 NA NA athletics 0 0 0  
## 6641 12/27/1988 1.73 NA taekwondo 0 0 0  
## 6651 07/25/1987 1.60 NA boxing 0 0 0  
## 6663 12/10/1986 1.75 NA boxing 0 0 0  
## 6677 10/09/1992 1.68 NA boxing 0 0 0  
## 6711 07/11/1996 NA NA aquatics 0 0 0  
## 6752 12/04/1986 1.77 NA athletics 0 0 0  
## 6771 09/15/1995 NA 66 athletics 0 0 1  
## 6786 08/08/1990 1.78 NA handball 0 0 1  
## 6824 07/23/1992 NA NA sailing 0 0 0  
## 6854 01/01/1994 NA NA athletics 0 0 0  
## 6867 08/17/1987 NA NA rugby sevens 0 0 0  
## 6873 02/11/1986 NA NA golf 0 0 0  
## 6948 11/25/1985 1.82 NA handball 0 0 1  
## 7005 05/06/1993 1.84 NA boxing 0 0 0  
## 7083 05/12/1991 NA NA rowing 0 0 0  
## 7097 01/01/1997 NA NA athletics 0 0 0  
## 7116 03/11/1995 NA NA athletics 0 0 0  
## 7129 01/15/1996 NA NA aquatics 0 0 0  
## 7150 10/15/1992 1.60 NA boxing 0 0 0  
## 7151 07/03/1989 1.85 NA boxing 0 0 1  
## 7271 10/26/1992 NA NA athletics 0 0 0  
## 7281 11/14/1997 NA NA athletics 0 0 0  
## 7286 10/20/1990 NA NA rowing 0 0 0  
## 7299 04/10/1990 NA NA athletics 0 0 0  
## 7303 01/13/1993 1.81 NA boxing 0 0 0  
## 7321 11/17/1996 NA 94 aquatics 0 0 0  
## 7356 05/07/1991 NA 51 athletics 0 0 0  
## 7364 07/02/1993 1.78 NA boxing 0 0 0  
## 7392 11/19/1991 1.70 NA boxing 0 0 0  
## 7400 04/30/1993 1.80 NA boxing 0 0 0  
## 7434 06/16/1991 1.82 NA boxing 0 0 0  
## 7438 01/10/1992 NA NA cycling 0 0 0  
## 7493 11/05/1990 1.80 NA boxing 0 0 0  
## 7507 07/04/1990 1.76 NA boxing 0 0 0  
## 7516 05/18/1990 1.77 NA boxing 0 0 0  
## 7517 06/03/1995 NA NA aquatics 0 0 0  
## 7536 04/30/1993 NA NA cycling 0 0 0  
## 7556 03/01/1989 NA NA athletics 0 0 0  
## 7578 05/30/1988 1.85 NA boxing 0 0 0  
## 7586 11/21/1994 1.63 NA gymnastics 0 0 0  
## 7593 11/17/1980 1.65 NA boxing 0 0 1  
## 7616 04/07/1994 1.78 NA boxing 0 0 1  
## 7622 08/23/1988 1.65 NA boxing 0 1 0  
## 7629 03/25/1999 1.78 NA shooting 0 0 0  
## 7654 03/01/1996 NA NA athletics 0 0 0  
## 7655 03/08/1987 NA NA judo 0 0 0  
## 7658 02/13/1990 1.74 NA boxing 0 0 0  
## 7659 12/11/1993 1.67 NA boxing 0 0 0  
## 7662 10/28/1991 NA NA athletics 0 0 0  
## 7681 05/15/1993 NA NA aquatics 0 0 0  
## 7686 12/30/1992 NA NA athletics 0 0 0  
## 7691 06/06/1987 1.86 NA boxing 0 0 0  
## 7696 07/13/1993 1.80 NA boxing 0 0 1  
## 7704 02/19/1993 NA NA rowing 0 0 0  
## 7745 11/12/1990 NA NA boxing 0 0 0  
## 7749 07/07/1994 NA NA athletics 0 0 0  
## 7757 02/03/1996 1.70 NA boxing 0 0 0  
## 7760 04/28/1998 1.70 NA football 0 0 1  
## 7763 06/20/1996 1.73 NA boxing 0 0 0  
## 7772 09/25/1992 NA NA athletics 0 0 0  
## 7779 10/22/1992 NA NA athletics 0 0 0  
## 7784 11/02/1994 1.65 NA boxing 0 0 1  
## 7798 06/25/1991 1.70 NA boxing 0 0 0  
## 7840 12/20/1990 NA NA athletics 0 0 0  
## 7842 01/10/1993 NA NA athletics 0 0 0  
## 7856 04/17/1986 1.75 NA canoe 0 0 0  
## 7858 01/13/1997 NA NA aquatics 0 0 0  
## 7862 04/16/1997 NA NA aquatics 0 0 0  
## 7869 01/06/1992 1.66 NA boxing 0 0 0  
## 7885 04/14/1988 NA NA aquatics 0 0 0  
## 7940 08/08/1980 NA NA cycling 0 0 0  
## 7946 10/28/1995 NA NA athletics 0 1 0  
## 7961 08/15/1998 1.73 NA football 0 0 1  
## 7962 12/29/1978 NA NA athletics 0 0 0  
## 7971 07/09/1998 1.50 NA athletics 0 0 0  
## 8029 01/04/1996 1.66 NA boxing 0 0 1  
## 8033 10/26/1982 1.64 NA boxing 1 0 0  
## 8078 05/10/1997 1.69 NA boxing 0 0 0  
## 8080 06/27/1989 1.94 NA boxing 0 0 0  
## 8081 08/10/1995 NA NA athletics 0 0 0  
## 8091 11/21/1988 NA NA aquatics 0 0 0  
## 8118 01/20/2000 NA NA aquatics 0 0 0  
## 8128 02/05/1985 NA NA rowing 0 0 0  
## 8149 07/15/1995 NA NA athletics 0 0 0  
## 8155 05/29/1991 NA NA cycling 0 0 0  
## 8179 04/05/1991 1.69 NA handball 0 0 1  
## 8188 11/09/1987 1.80 NA boxing 0 1 0  
## 8206 07/02/1986 NA NA athletics 0 0 0  
## 8211 07/30/1994 1.66 NA boxing 0 0 0  
## 8212 08/20/1989 1.60 NA judo 0 0 0  
## 8213 09/01/1984 NA NA shooting 0 0 0  
## 8218 11/09/1995 1.72 NA football 0 0 1  
## 8222 04/19/1997 1.70 NA football 0 0 1  
## 8228 04/27/1976 NA NA rowing 0 0 1  
## 8233 05/01/1989 NA NA shooting 0 0 0  
## 8277 05/10/1980 NA NA cycling 0 1 0  
## 8295 01/05/1993 NA 59 athletics 0 0 0  
## 8310 01/29/1996 1.72 NA football 0 0 1  
## 8315 05/02/1988 1.60 NA boxing 0 0 0  
## 8322 04/25/1994 NA NA athletics 1 0 0  
## 8329 05/01/1987 1.75 NA boxing 0 0 0  
## 8334 07/22/1993 NA NA athletics 0 0 0  
## 8338 03/17/1989 1.72 NA boxing 0 0 0  
## 8341 12/24/1997 NA NA aquatics 0 0 0  
## 8363 12/31/1987 NA 78 aquatics 0 0 0  
## 8371 01/28/1988 1.70 NA boxing 0 0 1  
## 8441 07/01/1997 NA NA aquatics 0 0 0  
## 8447 10/12/1992 NA NA wrestling 0 0 0  
## 8448 02/10/1994 1.77 NA boxing 0 0 0  
## 8454 06/08/1995 1.82 NA boxing 0 0 0  
## 8470 04/09/1987 1.63 NA boxing 0 0 0  
## 8476 07/02/1993 1.64 NA boxing 0 0 0  
## 8504 12/28/1995 1.89 NA boxing 0 0 0  
## 8540 07/12/1991 1.76 NA boxing 0 0 0  
## 8547 NA NA sailing 0 0 0  
## 8558 10/08/1994 NA NA aquatics 0 0 0  
## 8567 10/20/1983 1.63 NA boxing 0 0 0  
## 8583 05/29/1995 NA NA athletics 0 0 0  
## 8602 07/10/1999 NA NA athletics 0 0 0  
## 8622 12/30/1987 1.83 NA boxing 0 0 0  
## 8623 04/22/1981 NA NA boxing 0 0 0  
## 8636 07/15/1991 1.72 NA boxing 0 0 0  
## 8653 09/22/1984 NA NA cycling 0 0 0  
## 8674 01/01/1995 NA NA athletics 0 0 0  
## 8747 08/07/1994 1.70 NA football 0 0 1  
## 8759 12/31/1997 NA NA athletics 0 0 0  
## 8768 03/10/1993 1.70 NA boxing 0 0 0  
## 8773 06/06/1990 1.80 NA boxing 0 0 1  
## 8781 09/18/1987 1.78 NA boxing 0 0 0  
## 8797 12/14/1996 NA NA athletics 0 0 0  
## 8830 05/07/1986 1.68 NA gymnastics 0 0 0  
## 8853 06/22/1995 NA NA sailing 0 0 0  
## 8858 05/30/1992 NA 69 rugby sevens 0 0 0  
## 8862 03/18/1984 NA NA tennis 0 1 0  
## 8863 08/30/1986 1.91 NA badminton 0 0 0  
## 8900 12/06/1995 1.77 NA boxing 0 0 0  
## 8914 03/26/1986 1.80 NA boxing 0 0 0  
## 8916 08/28/1997 NA NA aquatics 0 0 0  
## 8934 09/05/1993 1.85 NA boxing 0 0 0  
## 8988 03/10/1987 NA NA aquatics 0 0 0  
## 9029 09/26/1997 NA NA boxing 0 0 0  
## 9070 11/26/1997 NA NA athletics 0 0 0  
## 9071 12/20/1993 1.62 NA boxing 1 0 0  
## 9073 09/24/1987 1.66 NA boxing 0 0 0  
## 9124 10/25/1988 1.75 NA boxing 1 0 0  
## 9138 04/06/1986 NA 66 judo 0 0 0  
## 9148 11/10/1993 1.65 NA boxing 0 0 0  
## 9186 04/05/1985 NA NA athletics 0 0 0  
## 9188 12/16/1992 NA NA athletics 0 0 0  
## 9190 09/05/1984 NA NA boxing 0 0 0  
## 9196 08/14/1988 1.82 NA boxing 0 0 0  
## 9209 07/12/1989 NA NA athletics 0 0 0  
## 9215 08/23/1996 NA NA athletics 0 0 0  
## 9255 04/25/1997 1.70 NA boxing 0 0 0  
## 9286 10/08/1991 1.86 NA boxing 0 0 1  
## 9310 06/27/1995 NA NA aquatics 1 0 0  
## 9346 02/02/1997 1.72 NA football 0 0 1  
## 9352 05/26/1990 NA NA table tennis 0 0 0  
## 9358 08/01/1994 NA 54 athletics 0 0 0  
## 9379 04/10/2001 NA 84 table tennis 0 0 0  
## 9382 12/20/1997 NA NA athletics 0 0 0  
## 9395 10/19/1987 NA NA tennis 0 0 0  
## 9406 06/06/1981 1.61 NA equestrian 0 0 0  
## 9426 05/28/1996 1.62 NA boxing 0 0 0  
## 9446 05/01/1986 NA NA athletics 0 0 0  
## 9473 11/10/1995 1.68 NA boxing 0 0 0  
## 9479 06/16/1990 1.78 NA handball 0 0 1  
## 9500 01/02/1988 NA NA athletics 0 0 0  
## 9519 12/30/1987 NA 45 athletics 0 0 0  
## 9526 11/09/1993 NA NA rowing 0 0 0  
## 9542 01/21/1982 1.57 NA boxing 0 1 0  
## 9573 01/10/1998 1.71 NA football 0 0 1  
## 9574 08/07/1988 NA NA athletics 0 0 0  
## 9581 05/19/1991 1.81 NA boxing 0 0 0  
## 9587 07/07/1986 1.75 NA boxing 0 0 0  
## 9645 06/12/1991 1.69 NA boxing 0 0 0  
## 9657 05/03/1995 NA NA athletics 0 0 0  
## 9670 09/10/1988 1.81 NA boxing 0 0 0  
## 9730 03/03/1993 1.69 NA boxing 1 0 0  
## 9731 07/07/1993 1.74 NA boxing 0 1 0  
## 9734 06/28/1997 1.73 NA boxing 0 1 0  
## 9737 04/05/1994 NA NA archery 0 0 0  
## 9755 09/16/1997 NA NA athletics 0 0 0  
## 9763 12/15/1993 NA NA athletics 0 0 0  
## 9768 09/06/1992 NA NA athletics 0 0 0  
## 9782 08/10/1987 1.64 NA boxing 0 0 0  
## 9810 10/12/1994 NA NA athletics 0 0 0  
## 9812 12/08/1993 1.69 NA boxing 0 0 0  
## 9852 08/23/1995 NA NA wrestling 0 0 0  
## 9869 12/24/1990 1.80 NA boxing 0 0 0  
## 9873 07/18/1986 NA NA cycling 0 0 0  
## 9908 05/09/1989 1.54 NA boxing 0 0 0  
## 9923 11/15/1995 NA NA athletics 0 0 0  
## 9953 03/27/1989 1.86 NA rugby sevens 0 0 0  
## 9954 12/23/1994 1.74 NA boxing 0 1 0  
## 9968 01/12/1991 NA NA athletics 0 0 0  
## 9974 07/15/1997 NA NA aquatics 0 0 0  
## 10000 05/16/1997 NA NA badminton 0 0 0  
## 10011 05/01/1996 NA NA judo 0 0 0  
## 10013 07/04/1991 1.79 NA boxing 0 0 1  
## 10032 04/10/1991 1.98 NA basketball 0 0 0  
## 10033 12/16/1990 1.70 NA boxing 0 0 0  
## 10035 02/28/1996 1.71 NA football 0 0 1  
## 10079 12/01/1996 NA NA athletics 0 0 0  
## 10084 10/27/1989 NA NA equestrian 0 0 0  
## 10106 02/27/1989 NA NA athletics 0 0 0  
## 10111 08/05/1994 NA NA athletics 0 0 1  
## 10119 09/12/1995 1.88 NA athletics 0 0 1  
## 10120 09/07/1988 1.80 NA boxing 0 0 0  
## 10127 07/20/1983 NA NA wrestling 0 0 1  
## 10131 09/25/1991 1.68 NA handball 0 0 1  
## 10140 05/23/1996 1.68 NA golf 0 0 0  
## 10161 08/14/1990 NA NA athletics 0 0 0  
## 10185 04/07/1981 NA NA golf 0 0 0  
## 10186 02/13/1956 NA NA equestrian 0 0 0  
## 10189 01/21/1992 NA NA cycling 0 0 0  
## 10192 01/13/1995 NA NA athletics 0 0 0  
## 10205 05/13/1976 NA NA athletics 0 0 0  
## 10250 01/21/1987 NA NA aquatics 0 0 0  
## 10271 08/31/1989 NA NA athletics 0 0 0  
## 10338 05/06/1993 NA NA athletics 1 0 0  
## 10358 07/30/1997 1.76 NA boxing 0 0 0  
## 10360 09/28/1994 1.88 NA athletics 0 0 0  
## 10383 10/25/1987 1.61 NA boxing 0 0 0  
## 10384 11/02/1981 NA NA athletics 0 0 0  
## 10385 01/11/1993 1.96 NA boxing 0 0 0  
## 10388 11/04/1997 1.67 NA boxing 0 0 0  
## 10462 10/24/1992 1.78 NA boxing 0 0 0  
## 10484 07/07/1997 NA NA athletics 0 0 0  
## 10485 12/25/1996 NA NA athletics 0 0 0  
## 10486 01/26/1987 NA NA athletics 0 0 0  
## 10516 10/10/1991 NA NA athletics 0 0 0  
## 10527 08/31/1993 NA NA taekwondo 0 0 0  
## 10533 05/29/1990 1.60 NA cycling 0 0 0  
## 10540 04/26/1993 NA NA sailing 0 0 0  
## 10581 12/29/1988 1.78 NA boxing 0 0 0  
## 10600 01/17/1990 NA NA athletics 0 0 0  
## 10609 04/27/1992 2.00 NA boxing 1 0 0  
## 10611 02/19/1996 NA NA cycling 0 0 0  
## 10626 06/14/1983 NA NA athletics 0 0 0  
## 10635 10/16/1996 1.63 NA boxing 0 0 0  
## 10636 02/23/1997 NA NA athletics 0 0 0  
## 10637 06/30/1991 NA NA athletics 0 0 0  
## 10657 03/27/1987 1.74 NA boxing 0 0 0  
## 10665 03/16/1993 1.58 NA athletics 0 0 0  
## 10675 08/28/1992 NA NA football 0 0 0  
## 10698 03/02/1994 1.69 NA football 0 0 1  
## 10713 11/16/1993 1.97 NA basketball 0 0 0  
## 10716 10/04/1994 NA NA football 0 0 0  
## 10733 09/29/1989 1.80 NA boxing 0 0 0  
## 10764 09/16/1993 1.60 NA boxing 0 0 0  
## 10771 02/24/1988 1.85 NA boxing 0 1 0  
## 10775 02/07/1989 NA NA cycling 0 0 0  
## 10791 07/10/1990 1.75 NA handball 0 0 1  
## 10815 03/27/1995 1.67 NA boxing 0 0 0  
## 10876 01/17/1989 1.70 NA boxing 0 0 0  
## 10882 03/01/1989 1.65 NA athletics 0 0 0  
## 10889 11/21/1984 NA NA cycling 0 0 0  
## 10894 04/12/1992 1.74 NA boxing 0 0 1  
## 10917 03/08/1991 1.72 NA boxing 0 0 0  
## 10919 06/16/1992 NA NA aquatics 0 0 0  
## 10920 03/25/1990 1.65 NA boxing 0 0 1  
## 10937 12/29/1982 1.71 NA boxing 0 0 0  
## 10943 05/22/1983 1.78 NA boxing 0 0 0  
## 10947 08/22/1993 1.67 NA boxing 0 0 0  
## 10964 11/27/1987 1.75 NA boxing 0 0 0  
## 11015 09/01/1993 1.75 NA football 0 0 1  
## 11029 08/23/1993 NA NA athletics 0 0 0  
## 11032 09/17/1993 NA NA boxing 0 0 0  
## 11047 08/27/1990 1.74 NA wrestling 0 0 0  
## 11049 03/16/1990 1.75 NA boxing 0 0 0  
## 11122 07/16/1991 1.92 NA boxing 0 0 0  
## 11127 10/30/1987 1.69 NA boxing 0 0 0  
## 11180 09/15/1989 1.75 NA boxing 0 0 0  
## 11209 02/21/1998 NA NA athletics 0 0 0  
## 11221 03/26/1994 NA NA athletics 0 0 0  
## 11239 10/16/1985 NA NA judo 0 0 0  
## 11261 04/22/1996 1.60 NA boxing 0 0 0  
## 11262 09/21/1996 1.68 NA boxing 0 0 1  
## 11292 03/12/1992 1.60 NA boxing 0 0 0  
## 11303 11/07/1991 1.82 NA boxing 0 0 0  
## 11371 04/03/1992 NA NA aquatics 0 2 0  
## 11399 11/01/1991 1.65 NA boxing 0 1 0  
## 11466 12/17/1989 1.54 NA boxing 0 0 0  
## 11469 04/01/1993 1.82 NA boxing 0 0 0  
## 11514 03/02/1986 1.78 NA boxing 0 0 0  
## 11515 11/18/1983 1.65 NA boxing 0 0 0  
## 11518 11/20/1992 1.84 NA boxing 0 0 0

# nasql\_athletes <- sqldf("select name,nationality,dob,height,weight from na\_athletes order by name")  
 # nasql\_athletes <- sqldf("select name,nationality,dob,height,weight from na\_athletes where dob is null order by name")  
 # nasql\_athletes  
   
   
# 2.1.3   
 # Replace dob, height & weight column's NA values  
   
 # athletes[is.na(athletes$dob),]  
 # athletes[which(athletes$dob==""),]  
 athletes[athletes$nationality=="RUS" & athletes$name=="Pavel Sozykin",]$dob <-"12/25/1987" # Update DOB of "Pavel Sozykin" born 25 December 1987  
   
 #[ No need for now]  
 # athletes[athletes$nationality=="HUN" & athletes$name=="Adam Decker",]$dob <-"1984-02-29" # Update DOB of "Adam DECKER" born 29 Feb 1984  
 # athletes[athletes$nationality=="UKR" & athletes$name=="Artem Morozov",]$dob <-"1980-02-29" # Update DOB of "Artem Morozov" born 29 Feb 1980  
 # athletes[athletes$nationality=="GBR" & athletes$name=="Hannah Mills",]$dob <-"1988-02-29" # Update DOB of "Hannah Mills" born 29 February 1988  
 # athletes[athletes$nationality=="SRB" & athletes$name=="Jovana Crnogorac",]$dob <-"1992-02-29" # Update DOB of "Jovana Crnogorac" born 29 February 1992  
 # athletes[athletes$nationality=="IND" & athletes$name=="Prakash Nanjappa",]$dob <-"1976-02-29" # Update DOB of "Prakash Nanjappa" born 29 February 1976  
 # athletes[athletes$nationality=="IND" & athletes$name=="Prakash Nanjappa",]  
 # athletes[athletes$nationality=="RUS" & athletes$name=="Pavel Sozykin" ,] # validate other RUS athletes  
   
# 2.1.4  
   
 # Check if the NAs in column "weight" and replace them by the mean value of this column  
 athletes$weight[is.na(athletes$weight)] <- mean(athletes$weight, na.rm=TRUE)  
  
# 2.1.5  
   
 # Check if the NAs in column "height" and replace them by the mean value of this column  
 athletes$height[is.na(athletes$height)] <- mean(athletes$height, na.rm=TRUE)   
   
 # Check if the NAs in athletes and replace them | Verify if NAs has been replaced   
 # na\_clnames <- colnames(athletes)[colSums(is.na(athletes)) > 0]  
 # na\_clnames  
 #   
   
# 2.1.5   
   
 #id Column in athletes doesnt mean much. lets remove it  
 athletes <- athletes[,!grepl("id",colnames(athletes))]  
 # athletes[1:3,]  
   
# 2.1.6   
   
 #Rename name,nationality columns   
 colnames(athletes) <- c("fullname","country\_code","sex","dob","height","weight","sport","gold","silver","bronze")  
 # athletes[1:3,]  
  
# 2.1.7   
   
 # Add columns age,tot\_medals,weight\_lbs,bmi\_metric  
 # lets add Age as a calculated column to hold age of athletes.   
 # This is by simply performing a data difference between athletes' dob and Jan 1st of 2016.  
 # athletes$age <- 0  
   
 athletes$age <- round(as.numeric(as.Date(as.character("01/01/2016"), format="%m/%d/%Y") - as.Date(as.character(athletes$dob), format="%m/%d/%Y"))/365,0)  
   
 # lets calculate total number of medals by summing all medal categories (gold,silver and bronze)  
 athletes$tot\_medals = athletes$gold+athletes$silver+athletes$bronze  
  
 # convert players wieght into lbs  
 athletes$weight\_lbs = 2.2 \* athletes$weight  
   
 # calculate BMI using metric formula  
 height\_m2 <- ifelse(athletes$height >0, athletes$height \* athletes$height,0)  
 athletes$bmi\_metric = round(athletes$weight/height\_m2,1)  
   
 # athletes[1:3,]  
  
# 2.1.8   
   
 # re-order the columns  
  
 ath\_col\_order <- c("fullname","sex","country\_code","sport","dob","age","height","weight","weight\_lbs","bmi\_metric","gold","silver","bronze","tot\_medals")  
 athletes <- athletes[, ath\_col\_order]  
 athletes[1:3,]

## fullname sex country\_code sport dob age height weight  
## 1 A Jesus Garcia male ESP athletics 10/17/1969 46 1.72 64  
## 2 A Lam Shin female KOR fencing 09/23/1986 29 1.68 56  
## 3 Aaron Brown male CAN athletics 05/27/1992 24 1.98 79  
## weight\_lbs bmi\_metric gold silver bronze tot\_medals  
## 1 140.8 21.6 0 0 0 0  
## 2 123.2 19.8 0 0 0 0  
## 3 173.8 20.2 0 0 1 1

# athletes[which(athletes$age<0),]

#Age validation  
# athletes[which(athletes$age>0),]  
# athletes[which(athletes$age<0),]  
# # athletes$age  
# athletes[which(athletes$dob==""),]  
# athletes[is.na(athletes$age),]  
#   
# athletes[athletes$country\_code=="RUS" & athletes$fullname=="Pavel Sozykin",]  
# athletes$age <- round(as.numeric(as.Date(as.character("01/01/2016"), format="%m/%d/%Y") - as.Date(as.character(athletes$dob), format="%m/%d/%Y"))/365,0)  
# athletes  
  
 summary(athletes$age)

## Min. 1st Qu. Median Mean 3rd Qu. Max.   
## 13.00 22.00 26.00 26.24 29.00 62.00

head(athletes[order(athletes$age),])

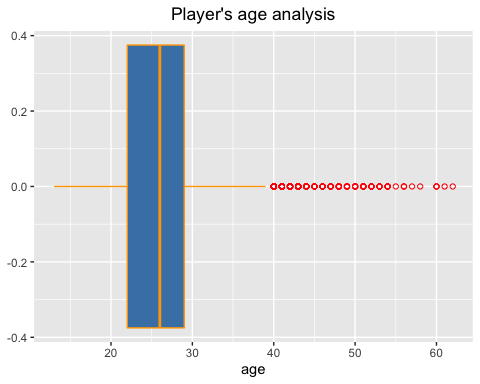
## fullname sex country\_code sport dob age height  
## 655 Ana Iulia Dascal female ROU aquatics 09/12/2002 13 1.83  
## 3599 Gaurika Singh female NEP aquatics 11/26/2002 13 1.55  
## 232 Ajna Kesely female HUN aquatics 09/10/2001 14 1.65  
## 716 Anastasiya Tyurina female TJK aquatics 09/27/2001 14 1.65  
## 2432 Darya Semyonova female TKM aquatics 05/28/2002 14 1.70  
## 3306 Fatima Alkaramova female AZE aquatics 06/26/2002 14 1.75  
## weight weight\_lbs bmi\_metric gold silver bronze tot\_medals  
## 655 60 132.0 17.9 0 0 0 0  
## 3599 45 99.0 18.7 0 0 0 0  
## 232 54 118.8 19.8 0 0 0 0  
## 716 50 110.0 18.4 0 0 0 0  
## 2432 50 110.0 17.3 0 0 0 0  
## 3306 60 132.0 19.6 0 0 0 0

head(athletes[order(-athletes$age),])

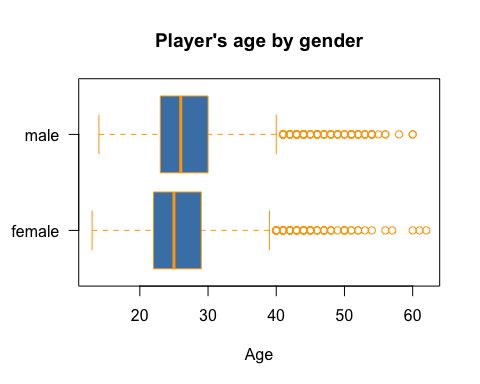
## fullname sex country\_code sport dob age height  
## 5300 Julie Brougham female NZL equestrian 05/20/1954 62 1.570000  
## 7093 Mary Hanna female AUS equestrian 12/01/1954 61 1.730000  
## 5046 John Whitaker male GBR equestrian 08/05/1955 60 1.610000  
## 6983 Mark Todd male NZL equestrian 03/01/1956 60 1.900000  
## 10186 Suzanne Hearn female AUS equestrian 02/13/1956 60 1.766282  
## 8019 Nick Skelton male GBR equestrian 12/30/1957 58 1.750000  
## weight weight\_lbs bmi\_metric gold silver bronze tot\_medals  
## 5300 48.0000 105.6000 19.5 0 0 0 0  
## 7093 63.0000 138.6000 21.0 0 0 0 0  
## 5046 70.0000 154.0000 27.0 0 0 0 0  
## 6983 78.0000 171.6000 21.6 0 0 0 0  
## 10186 72.0682 158.5501 23.1 0 0 0 0  
## 8019 76.0000 167.2000 24.8 1 0 0 1

#Charts definition

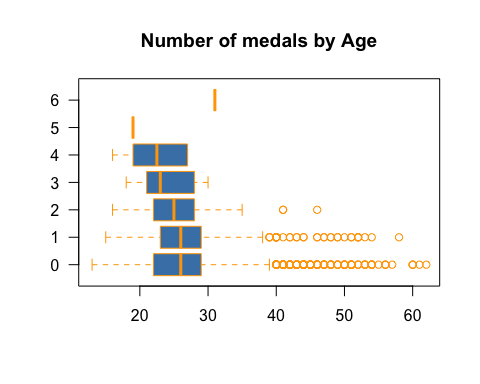
# barplot(athletes$dob)  
 hcolor <- c("orange")  
 hfill <- c("steelblue")  
 htitle <- c("values distribution")  
 theme <-theme(plot.title = element\_text(hjust = 0.5),axis.title = element\_text())  
   
 #valdiation of player's age across   
 ggAgeboxplot <- ggplot(athletes,aes(age)) +geom\_boxplot(fill = "steelblue", colour = "orange",outlier.colour = "red", outlier.shape = 1,na.rm=TRUE)+ggtitle("Player's age analysis")  
   
 ggAgeboxplot+theme



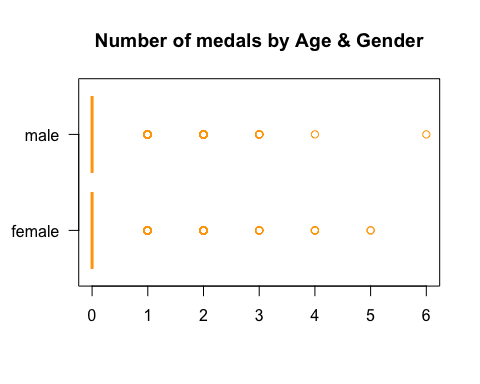
#Age by gender  
boxplot(age~sex,  
 data = athletes, main="Player's age by gender",   
 ylab="",  
 xlab = "Age",  
 border="orange",   
 col="steelblue",  
 freq=FALSE,  
 las=1,   
 breaks=5,  
 horizontal = TRUE,na.rm=TRUE)



boxplot(age~tot\_medals,  
 data = athletes, main="Number of medals by Age",   
 ylab="",  
 xlab = "",  
 border="orange",   
 col="steelblue",  
 freq=FALSE,  
 las=1,   
 breaks=5,  
 horizontal = TRUE,na.rm=TRUE)



boxplot(tot\_medals~sex,  
 data = athletes, main="Number of medals by Age & Gender",   
 ylab="",  
 xlab = "",  
 border="orange",   
 col="steelblue",  
 freq=FALSE,  
 las=1,   
 breaks=5,  
 horizontal = TRUE,na.rm=TRUE)



#2. Data Cleaning Process | 2.2 Countries Dataframe

1. Remove NA Characters  
2. Change Data type   
3. Remove unwanted columns  
4. Add columns (Copy + calculated)  
5. Rename Columns  
6. Re-arrange Columns  
7. ?

# Preview the data  
   
   
 printDataInfo(countries)

## 'data.frame': 201 obs. of 4 variables:  
## $ country : Factor w/ 201 levels "Afghanistan",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ code : Factor w/ 201 levels "AFG","AHO","ALB",..: 1 3 4 11 5 6 7 8 9 10 ...  
## $ population : int 32526562 2889167 39666519 55538 70473 25021974 91818 43416755 3017712 103889 ...  
## $ gdp\_per\_capita: num 594 3945 4206 NA NA ...  
## str:   
## colnames: country code population gdp\_per\_capita   
## dim: 201 4   
## nrow: 201

## country code population gdp\_per\_capita  
## 1 Afghanistan AFG 32526562 594.3231  
## 2 Albania ALB 2889167 3945.2176  
## 3 Algeria ALG 39666519 4206.0312

# 2.2.1   
   
 #Rename code to country\_code   
 colnames(countries) <- c("country","country\_code","population","gdp\_per\_capita")  
 # countries[1:3,]  
   
   
# 2.2.2  
   
 # Check if the NAs in athletes and replace them   
 nacountry\_clmns <- colnames(countries)[colSums(is.na(countries)) > 0]  
 nacountry\_clmns

## [1] "population" "gdp\_per\_capita"

# create subset of dataframe rows having NAs  
 na\_countries <- countries[rowSums(is.na(countries)) > 0,]  
 na\_countries # 25- rows

## country country\_code population gdp\_per\_capita  
## 4 American Samoa\* ASA 55538 NA  
## 5 Andorra AND 70473 NA  
## 10 Aruba\* ARU 103889 NA  
## 21 Bermuda\* BER 65235 NA  
## 28 British Virgin Islands IVB 30117 NA  
## 37 Cayman Islands\* CAY 59967 NA  
## 46 Cook Islands COK NA NA  
## 50 Cuba CUB 11389562 NA  
## 62 Eritrea ERI NA NA  
## 75 Guam GUM 169885 NA  
## 87 Iran IRI 79109272 NA  
## 97 Korea, North PRK 25155317 NA  
## 106 Libya LBA 6278438 NA  
## 107 Liechtenstein LIE 37531 NA  
## 117 Mauritania MTN 4067564 NA  
## 122 Monaco MON 37731 NA  
## 131 Netherlands Antilles\* AHO NA NA  
## 140 Palestine, Occupied Territories PLE NA NA  
## 142 Papua New Guinea PNG 7619321 NA  
## 148 Puerto Rico\* PUR 3474182 NA  
## 157 San Marino SMR 31781 NA  
## 177 Syria SYR 18502413 NA  
## 178 Taiwan TPE NA NA  
## 196 Venezuela VEN 31108083 NA  
## 198 Virgin Islands\* ISV 103574 NA

# nasql\_countries <- sqldf("select country,country\_code,population,gdp\_per\_capita from na\_countries order by country")  
 # nasql\_countries #25 rows  
  
# 2.2.3  
   
 # find out if any country\_code from athletes dataframe is missing in countries dataframe  
 countries[1:3,]

## country country\_code population gdp\_per\_capita  
## 1 Afghanistan AFG 32526562 594.3231  
## 2 Albania ALB 2889167 3945.2176  
## 3 Algeria ALG 39666519 4206.0312

athletes[1:3,]

## fullname sex country\_code sport dob age height weight  
## 1 A Jesus Garcia male ESP athletics 10/17/1969 46 1.72 64  
## 2 A Lam Shin female KOR fencing 09/23/1986 29 1.68 56  
## 3 Aaron Brown male CAN athletics 05/27/1992 24 1.98 79  
## weight\_lbs bmi\_metric gold silver bronze tot\_medals  
## 1 140.8 21.6 0 0 0 0  
## 2 123.2 19.8 0 0 0 0  
## 3 173.8 20.2 0 0 1 1

x <- data.frame(countries[,1:2],stringsAsFactors=FALSE)  
 y <- data.frame(athletes$country\_code,stringsAsFactors=FALSE)  
 colnames(y) <- c("country\_code")  
 z <- data.frame(unique(anti\_join(y,x, by="country\_code",),stringsAsFactors=FALSE))  
 z

## country\_code  
## 1 IOA  
## 5 ROU  
## 10 MNE  
## 11 SRB  
## 41 TTO  
## 45 ROT  
## 66 MHL  
## 84 KIR  
## 105 TUV  
## 123 SSD  
## 164 KOS

country\_code <- c("IOA","KIR","KOS","MHL","MNE","ROT","ROU","SRB","SSD","TTO","TUV")  
 country <- c("Individual Olympic athletes","Kiribati","Kosovo","Marshall Islands","Montenegro","Refugee Olympic Team","Romania","Serbia","South Sudan","Trinidad and Tobago","Tuvalu")  
 population <- rep(0, 11)  
 gdp\_per\_capita <- rep(0, 11)  
  
 missing\_countries <- data.frame(country\_code,country,population,gdp\_per\_capita,stringsAsFactors=FALSE)  
 missing\_countries

## country\_code country population gdp\_per\_capita  
## 1 IOA Individual Olympic athletes 0 0  
## 2 KIR Kiribati 0 0  
## 3 KOS Kosovo 0 0  
## 4 MHL Marshall Islands 0 0  
## 5 MNE Montenegro 0 0  
## 6 ROT Refugee Olympic Team 0 0  
## 7 ROU Romania 0 0  
## 8 SRB Serbia 0 0  
## 9 SSD South Sudan 0 0  
## 10 TTO Trinidad and Tobago 0 0  
## 11 TUV Tuvalu 0 0

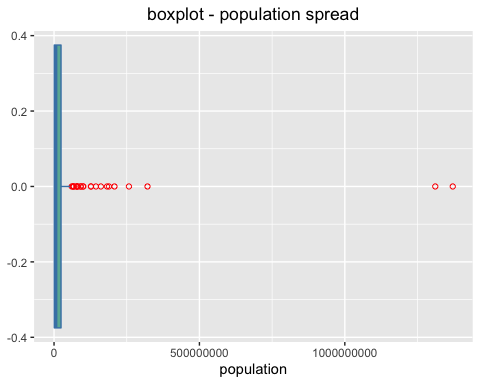
countries <- rbind(countries, missing\_countries)  
  
 countries[1:3,]

## country country\_code population gdp\_per\_capita  
## 1 Afghanistan AFG 32526562 594.3231  
## 2 Albania ALB 2889167 3945.2176  
## 3 Algeria ALG 39666519 4206.0312

clnames3 <- colnames(countries)[colSums(is.na(countries)) > 0]  
 clnames3

## [1] "population" "gdp\_per\_capita"

options(scipen=999)   
 #valdiation of player's age across   
  
 ggPopluationboxplot <- ggplot(countries,aes(population)) +geom\_boxplot(fill = "#69b3a2", colour = "steelblue",outlier.colour = "red", outlier.shape = 1,na.rm=TRUE,orientation=NA)+ggtitle("boxplot - population spread")  
   
 ggPopluationboxplot+theme



#3. Join Athletes and Countries data frame

Join Athletes and Countries data frame  
Re-arrange the columns in Athletes data frame

#3.1  
 #jojn Athletes and Countries dataframe by nationality  
   
 # printDataInfo(athletes)  
 # printDataInfo(countries)  
  
 athletes <- left\_join(athletes,countries, by="country\_code")  
  
#3.2  
 # re-order the columns  
 # colnames(athletes)  
   
 col\_order <- c("fullname","sex","country","country\_code","sport","dob","age","height","weight","weight\_lbs","bmi\_metric","gold","silver","bronze","tot\_medals","population","gdp\_per\_capita")  
 athletes <- athletes[, col\_order]  
 # athletes[1:3,]  
   
#3.3  
 # review for NAs  
 na\_athletes <- colnames(athletes)[colSums(is.na(athletes)) > 0]  
 # na\_athletes  
   
 #create subset of dataframe rows having NAs  
 # na\_athletes <- athletes[rowSums(is.na(athletes)) > 0,]  
 # na\_athletes #509 rows  
   
 # nasql\_athletes <- sqldf("select fullname,country,dob,height,weight,weight\_lbs,bmi\_metric,population,gdp\_per\_capita from na\_athletes order by fullname")  
 # nasql\_athletes  
   
 printDataInfo(athletes)

## 'data.frame': 11538 obs. of 17 variables:  
## $ fullname : Factor w/ 11517 levels "A Jesus Garcia",..: 1 2 3 4 5 6 7 8 9 10 ...  
## $ sex : Factor w/ 2 levels "female","male": 2 1 2 2 2 2 2 2 1 1 ...  
## $ country : Factor w/ 209 levels "Afghanistan",..: 170 99 36 122 132 11 192 11 170 65 ...  
## $ country\_code : Factor w/ 212 levels "AFG","ALB","ALG",..: 60 103 34 120 142 11 199 11 60 62 ...  
## $ sport : Factor w/ 28 levels "aquatics","archery",..: 3 10 3 23 8 25 26 1 3 3 ...  
## $ dob : Factor w/ 5596 levels "","01/01/1969",..: 4503 4151 2325 32 5087 424 2450 4186 5340 3194 ...  
## $ age : num 46 29 24 25 25 26 23 24 27 24 ...  
## $ height : num 1.72 1.68 1.98 1.83 1.81 1.8 2.05 1.93 1.8 1.65 ...  
## $ weight : num 64 56 79 80 71 67 98 100 62 54 ...  
## $ weight\_lbs : num 141 123 174 176 156 ...  
## $ bmi\_metric : num 21.6 19.8 20.2 23.9 21.7 20.7 23.3 26.8 19.1 19.8 ...  
## $ gold : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ silver : int 0 0 0 0 0 0 0 0 0 0 ...  
## $ bronze : int 0 0 1 0 0 0 1 0 0 0 ...  
## $ tot\_medals : int 0 0 1 0 0 0 1 0 0 0 ...  
## $ population : num 46418269 50617045 35851774 3554150 4595700 ...  
## $ gdp\_per\_capita: num 25832 27222 43249 1848 37808 ...  
## str:   
## colnames: fullname sex country country\_code sport dob age height weight weight\_lbs bmi\_metric gold silver bronze tot\_medals population gdp\_per\_capita   
## dim: 11538 17   
## nrow: 11538

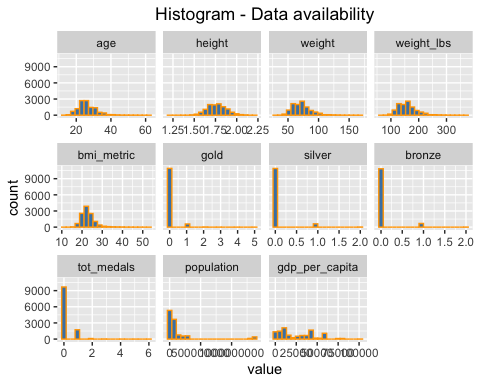
## fullname sex country country\_code sport dob age  
## 1 A Jesus Garcia male Spain ESP athletics 10/17/1969 46  
## 2 A Lam Shin female Korea, South KOR fencing 09/23/1986 29  
## 3 Aaron Brown male Canada CAN athletics 05/27/1992 24  
## height weight weight\_lbs bmi\_metric gold silver bronze tot\_medals population  
## 1 1.72 64 140.8 21.6 0 0 0 0 46418269  
## 2 1.68 56 123.2 19.8 0 0 0 0 50617045  
## 3 1.98 79 173.8 20.2 0 0 1 1 35851774  
## gdp\_per\_capita  
## 1 25831.58  
## 2 27221.52  
## 3 43248.53

# sort(unique(events$sport))  
  
 # Values in Events and Athlets are at different levels; So, join doesnt really work between these two dataframes.   
 # events[1:3,]  
   
 #jojn Athletes and events dataframe by sport & id  
  
 # athletes <- left\_join(athletes,events)  
 # athletes[1:3,]

# Plot histogram on countries data to analyze the data spread for missing data elements  
  
 hcolor <- c("orange")  
 hfill <- c("steelblue")  
 htitle <- c("Histogram - Data availability")  
 theme <-theme(plot.title = element\_text(hjust = 0.5),axis.title = element\_text())  
   
 gghist <- ggplot(data=melt(athletes),mapping = aes(x= value))

## Using fullname, sex, country, country\_code, sport, dob as id variables

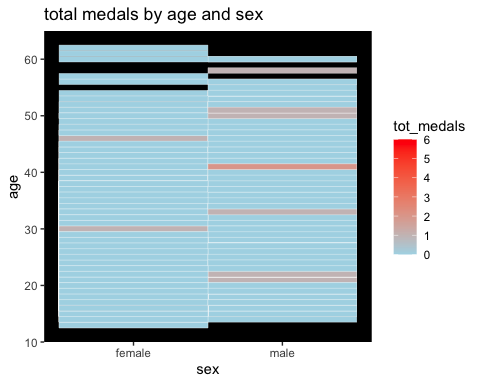
gghist+geom\_histogram(bins = 20,color=hcolor,fill=hfill,na.rm = TRUE)+facet\_wrap(~variable,scales = "free\_x")+ ggtitle(htitle) + theme



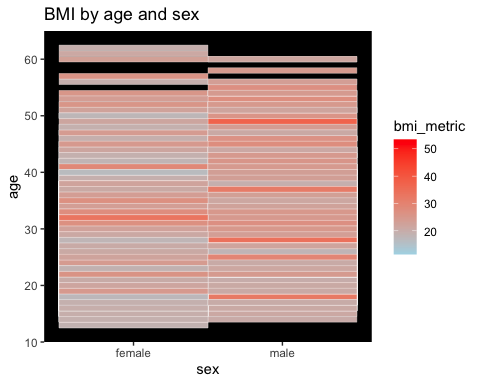
athletes[1:3,]

## fullname sex country country\_code sport dob age  
## 1 A Jesus Garcia male Spain ESP athletics 10/17/1969 46  
## 2 A Lam Shin female Korea, South KOR fencing 09/23/1986 29  
## 3 Aaron Brown male Canada CAN athletics 05/27/1992 24  
## height weight weight\_lbs bmi\_metric gold silver bronze tot\_medals population  
## 1 1.72 64 140.8 21.6 0 0 0 0 46418269  
## 2 1.68 56 123.2 19.8 0 0 0 0 50617045  
## 3 1.98 79 173.8 20.2 0 0 1 1 35851774  
## gdp\_per\_capita  
## 1 25831.58  
## 2 27221.52  
## 3 43248.53

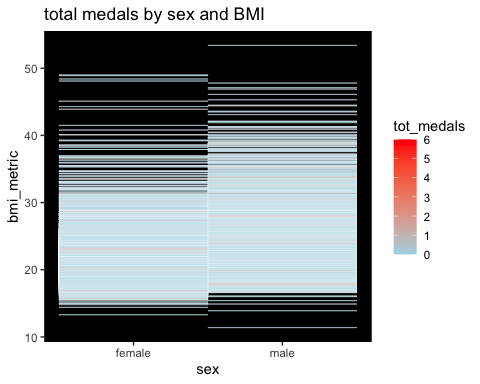
hmap <- ggplot(athletes, aes(x=sex, y=age))  
hmap <- hmap + geom\_tile(aes(fill=tot\_medals), color="white",na.rm=TRUE)   
hmap <- hmap + scale\_fill\_gradient(low= "lightblue", high="red") + ggtitle("total medals by age and sex")  
hmap <- hmap + theme(  
 panel.border = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank()  
)  
hmap <- hmap + theme(  
 panel.background = element\_rect(fill = "black")  
)   
hmap



h <- ggplot(athletes, aes(x=sex, y=age))  
h <- h + geom\_tile(aes(fill=bmi\_metric), color="white", na.rm=TRUE)  
h <- h + scale\_fill\_gradient(low="lightblue", high= "red")  
h <- h + theme(  
 panel.border = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank()  
)  
h <- h + theme(  
 panel.background = element\_rect(fill = "black")  
) + ggtitle("BMI by age and sex")  
h



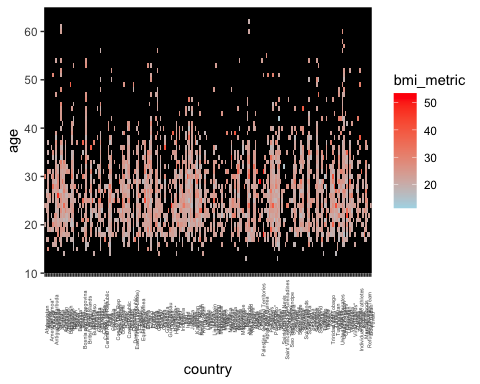
p2 <- ggplot(athletes, aes(x=sex, y=bmi\_metric))  
p2 <- p2 + geom\_tile(aes(fill=tot\_medals), color="white",na.rm=TRUE)   
p2 <- p2 + scale\_fill\_gradient(low= "lightblue", high="red") + ggtitle("total medals by sex and BMI")  
p2 <- p2 + theme(  
 panel.border = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank()  
)  
p2 <- p2 + theme(  
 panel.background = element\_rect(fill = "black")  
)   
p2



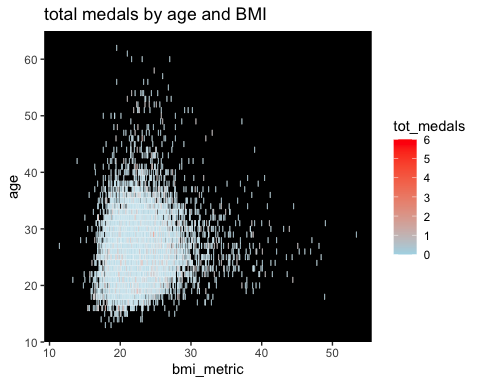
athletes[1:3,]

## fullname sex country country\_code sport dob age  
## 1 A Jesus Garcia male Spain ESP athletics 10/17/1969 46  
## 2 A Lam Shin female Korea, South KOR fencing 09/23/1986 29  
## 3 Aaron Brown male Canada CAN athletics 05/27/1992 24  
## height weight weight\_lbs bmi\_metric gold silver bronze tot\_medals population  
## 1 1.72 64 140.8 21.6 0 0 0 0 46418269  
## 2 1.68 56 123.2 19.8 0 0 0 0 50617045  
## 3 1.98 79 173.8 20.2 0 0 1 1 35851774  
## gdp\_per\_capita  
## 1 25831.58  
## 2 27221.52  
## 3 43248.53

m <- ggplot(athletes, aes(x=country, y=age))  
m <- m + geom\_tile(aes(fill=bmi\_metric), na.rm=TRUE)  
m <- m + scale\_fill\_gradient(low="lightblue", high= "red")   
m <- m + theme(axis.text.x = element\_text(size=4, angle = 90))  
m <- m+ theme(  
 panel.border = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank()  
)  
m <- m + theme(  
 panel.background = element\_rect(fill = "black")  
)  
m

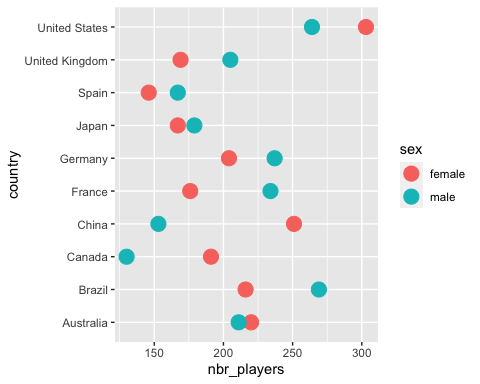


p <- ggplot(athletes, aes(x=bmi\_metric, y=age))  
p <- p + geom\_tile(aes(fill=tot\_medals), color="white",na.rm=TRUE)   
p <- p + scale\_fill\_gradient(low= "lightblue", high="red") + ggtitle("total medals by age and BMI")  
p <- p + theme(  
 panel.border = element\_blank(),  
 panel.grid.major = element\_blank(),  
 panel.grid.minor = element\_blank()  
)  
p <- p + theme(  
 panel.background = element\_rect(fill = "black")  
)   
p



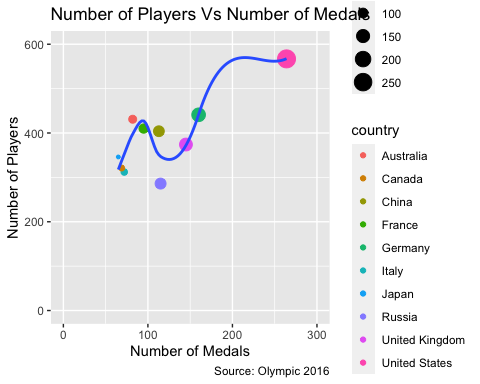
athletes <- na.omit(athletes) # Remove NAs from the data  
   
 sqlResults <- sqldf("select \* from athletes limit 5")  
 # sqlResults  
   
 nbrofplayers <- sqldf("select country, sex, count(fullname) nbr\_players from athletes group by country,sex order by nbr\_players desc " )  
 # nbrofplayers  
   
 topcountries <- sqldf("select country, count(fullname) nbr\_players from athletes group by country order by nbr\_players desc limit 10" )  
 # topcountries  
   
   
 toplevelcountries <- sqldf("select country, sum(gold) as nbr\_gold, sum(silver) as nbr\_silver,sum(bronze) as nbr\_bronze,sum(tot\_medals) as nbr\_medals from athletes group by country order by nbr\_medals desc limit 10" )  
 # toplevelcountries  
   
 topplayers <- sqldf("select fullname,sex, sum(gold) as nbr\_gold, sum(silver) as nbr\_silver,sum(bronze) as nbr\_bronze,sum(tot\_medals) as nbr\_medals from athletes group by fullname,sex order by nbr\_medals desc" )  
 # topplayers  
   
 topsports <- sqldf("select sport,count(fullname) nbr\_players,sum(tot\_medals) as nbr\_medals from athletes group by sport order by nbr\_medals desc,nbr\_players desc limit 10" )  
 # topsports  
   
 topmedals <- sqldf("select country, count(fullname) nbr\_players ,sum(tot\_medals) as tot\_medals from athletes group by country order by tot\_medals desc,nbr\_players desc limit 10" )  
 # topmedals  
   
 topgender <- sqldf("select sport,sex, sum(tot\_medals) as nbr\_medals from athletes group by sport,sex order by nbr\_medals desc" )  
 # topgender

# topcountries  
  
Countries <- topcountries$country [order (topcountries$nbr\_players, decreasing = TRUE)]  
# Countries  
  
x <- subset(nbrofplayers,nbrofplayers$country %in% Countries)  
   
ggplot (data = x,  
 aes (nbr\_players,country)) +  
 geom\_point (aes (color = sex), size = 5)



options(scipen=999) # turn-off scientific notation like 1e+48  
# theme\_set(theme\_bw()) # pre-set the bw theme.  
  
  
# Scatterplot  
gg <- ggplot(topmedals, aes(x=tot\_medals, y=nbr\_players)) +   
 geom\_point(aes(col=country, size=tot\_medals)) +   
 geom\_smooth(method="loess", se=F) +   
 xlim(c(0, 300)) +   
 ylim(c(0, 600)) +   
 labs( y="Number of Players",   
 x="Number of Medals",   
 title="Number of Players Vs Number of Medals",   
 caption = "Source: Olympic 2016")  
  
plot(gg)

## `geom\_smooth()` using formula 'y ~ x'



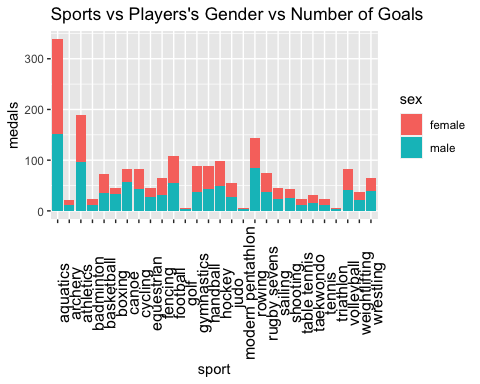
topgender <- sqldf("select sport,sex, sum(tot\_medals) as medals from athletes group by sport,sex order by medals desc" )  
topgender

## sport sex medals  
## 1 aquatics female 187  
## 2 aquatics male 151  
## 3 athletics male 97  
## 4 athletics female 93  
## 5 rowing male 84  
## 6 rowing female 60  
## 7 canoe male 56  
## 8 football female 54  
## 9 football male 54  
## 10 gymnastics female 51  
## 11 hockey male 50  
## 12 hockey female 49  
## 13 cycling male 44  
## 14 handball female 44  
## 15 handball male 44  
## 16 volleyball male 42  
## 17 volleyball female 41  
## 18 wrestling male 40  
## 19 cycling female 39  
## 20 gymnastics male 38  
## 21 rugby sevens male 38  
## 22 basketball female 36  
## 23 basketball male 36  
## 24 rugby sevens female 36  
## 25 boxing male 33  
## 26 fencing female 33  
## 27 fencing male 32  
## 28 judo male 28  
## 29 equestrian male 27  
## 30 judo female 27  
## 31 canoe female 26  
## 32 shooting male 26  
## 33 sailing male 24  
## 34 wrestling female 24  
## 35 sailing female 21  
## 36 weightlifting male 21  
## 37 equestrian female 18  
## 38 shooting female 18  
## 39 taekwondo male 16  
## 40 weightlifting female 16  
## 41 taekwondo female 15  
## 42 archery male 12  
## 43 badminton female 12  
## 44 badminton male 12  
## 45 boxing female 12  
## 46 table tennis male 12  
## 47 tennis male 12  
## 48 table tennis female 11  
## 49 tennis female 11  
## 50 archery female 9  
## 51 golf male 4  
## 52 modern pentathlon female 3  
## 53 modern pentathlon male 3  
## 54 triathlon female 3  
## 55 triathlon male 3  
## 56 golf female 2

str(topgender)

## 'data.frame': 56 obs. of 3 variables:  
## $ sport : Factor w/ 28 levels "aquatics","archery",..: 1 1 3 3 18 18 7 11 11 13 ...  
## $ sex : Factor w/ 2 levels "female","male": 1 2 2 1 2 1 2 1 2 1 ...  
## $ medals: int 187 151 97 93 84 60 56 54 54 51 ...

ggplot(topgender, aes(fill=sex, y=medals, x=sport)) + geom\_bar(position="stack", stat="identity")+  
 scale\_color\_brewer(palette = "Paired")+  
 ggtitle("Sports vs Players's Gender vs Number of Goals") +  
 theme(axis.text.x = element\_text(face="plain", color="black", size=12, angle=90))

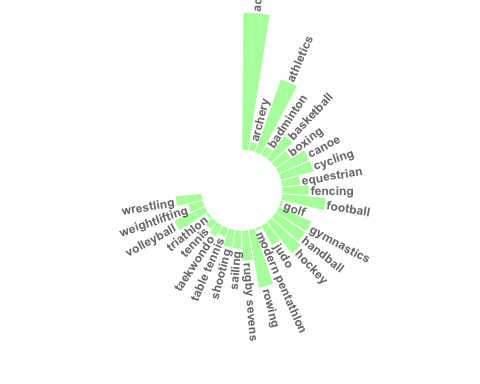


# Create dataset  
data <- data.frame(  
 individual=paste( "Mister ", seq(1,60), sep=""),  
 value=sample( seq(10,100), 60, replace=T)  
)  
 # data  
 # athletes  
data <- sqldf("select sport as individual, sum(tot\_medals) as value from athletes group by sport order by sport" )  
  
# data  
   
# Set a number of 'empty bar'  
empty\_bar <- 10  
   
# Add lines to the initial dataset  
to\_add <- matrix(NA, empty\_bar, ncol(data))  
colnames(to\_add) <- colnames(data)  
data <- rbind(data, to\_add)  
data

## individual value  
## 1 aquatics 338  
## 2 archery 21  
## 3 athletics 190  
## 4 badminton 24  
## 5 basketball 72  
## 6 boxing 45  
## 7 canoe 82  
## 8 cycling 83  
## 9 equestrian 45  
## 10 fencing 65  
## 11 football 108  
## 12 golf 6  
## 13 gymnastics 89  
## 14 handball 88  
## 15 hockey 99  
## 16 judo 55  
## 17 modern pentathlon 6  
## 18 rowing 144  
## 19 rugby sevens 74  
## 20 sailing 45  
## 21 shooting 44  
## 22 table tennis 23  
## 23 taekwondo 31  
## 24 tennis 23  
## 25 triathlon 6  
## 26 volleyball 83  
## 27 weightlifting 37  
## 28 wrestling 64  
## 29 <NA> NA  
## 30 <NA> NA  
## 31 <NA> NA  
## 32 <NA> NA  
## 33 <NA> NA  
## 34 <NA> NA  
## 35 <NA> NA  
## 36 <NA> NA  
## 37 <NA> NA  
## 38 <NA> NA

data$id <- seq(1, nrow(data))  
   
# Get the name and the y position of each label  
label\_data <- data  
number\_of\_bar <- nrow(label\_data)  
angle <- 90 - 360 \* (label\_data$id-0.5) /number\_of\_bar # I substract 0.5 because the letter must have the angle of the center of the bars. Not extreme right(1) or extreme left (0)  
label\_data$hjust <- ifelse( angle < -90, 1, 0)  
label\_data$angle <- ifelse(angle < -90, angle+180, angle)  
   
# Make the plot  
p <- ggplot(data, aes(x=as.factor(id), y=value)) + # Note that id is a factor. If x is numeric, there is some space between the first bar  
 geom\_bar(stat="identity", fill=alpha("green", 0.4)) +  
 ylim(-100,350) +  
 theme\_minimal() +  
 theme(  
 axis.text = element\_blank(),  
 axis.title = element\_blank(),  
 panel.grid = element\_blank(),  
 plot.margin = unit(rep(-1,4), "cm")   
 ) +  
 coord\_polar(start = 0) +   
 geom\_text(data=label\_data, aes(x=id, y=value+5, label=individual, hjust=hjust), color="black", fontface="bold",alpha=0.6, size=3.2, angle= label\_data$angle, inherit.aes = FALSE,na.rm = TRUE )   
   
p

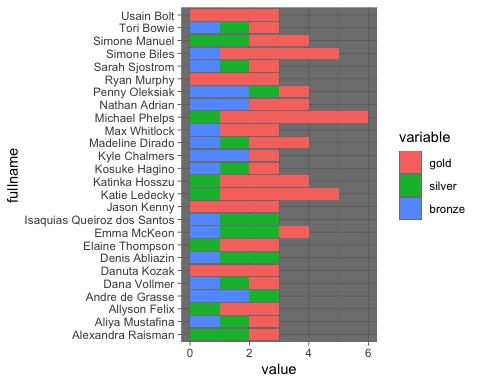
## Warning: Removed 10 rows containing missing values (position\_stack).



df1 <- sqldf("select fullname, gold, silver,bronze from athletes") #--where fullname like 'Michael Phelps  
# df1  
  
  
mdata <- melt(df1, id=c("fullname"))  
mdata <- mdata[order(-mdata$value,mdata$fullname),]  
mdata <- mdata[which(mdata$value>=1),]  
# mdata  
# mdata[mdata$fullname =="Michael Phelps",]  
  
  
# topcountries  
  
Countries <- topcountries$country [order (topcountries$nbr\_players, decreasing = TRUE)]  
# Countries  
  
x <- subset(nbrofplayers,nbrofplayers$country %in% Countries)  
   
p1 <- sqldf("select fullname,sum(tot\_medals) as medals from athletes group by fullname order by medals desc")  
p1 <- p1[which(p1$medals>=3),]  
  
  
players <- p1$fullname[order(p1$medals, decreasing = TRUE)]   
# players  
  
x <- subset(mdata,mdata$fullname %in% players)  
x

## fullname variable value  
## 7098 Michael Phelps gold 5  
## 5318 Katie Ledecky gold 4  
## 9475 Simone Biles gold 4  
## 2294 Danuta Kozak gold 3  
## 4489 Jason Kenny gold 3  
## 5322 Katinka Hosszu gold 3  
## 8911 Ryan Murphy gold 3  
## 10247 Usain Bolt gold 3  
## 11426 Alexandra Raisman silver 2  
## 515 Allyson Felix gold 2  
## 22773 Andre de Grasse bronze 2  
## 13467 Denis Abliazin silver 2  
## 2756 Elaine Thompson gold 2  
## 13952 Emma McKeon silver 2  
## 15263 Isaquias Queiroz dos Santos silver 2  
## 27724 Kyle Chalmers bronze 2  
## 6246 Madeline Dirado gold 2  
## 6953 Max Whitlock gold 2  
## 7598 Nathan Adrian gold 2  
## 29656 Nathan Adrian bronze 2  
## 30289 Penny Oleksiak bronze 2  
## 9482 Simone Manuel gold 2  
## 20511 Simone Manuel silver 2  
## 397 Alexandra Raisman gold 1  
## 495 Aliya Mustafina gold 1  
## 11524 Aliya Mustafina silver 1  
## 22553 Aliya Mustafina bronze 1  
## 11544 Allyson Felix silver 1  
## 11744 Andre de Grasse silver 1  
## 2198 Dana Vollmer gold 1  
## 13227 Dana Vollmer silver 1  
## 24256 Dana Vollmer bronze 1  
## 24496 Denis Abliazin bronze 1  
## 13785 Elaine Thompson silver 1  
## 2923 Emma McKeon gold 1  
## 24981 Emma McKeon bronze 1  
## 26292 Isaquias Queiroz dos Santos bronze 1  
## 16347 Katie Ledecky silver 1  
## 16351 Katinka Hosszu silver 1  
## 5590 Kosuke Hagino gold 1  
## 16619 Kosuke Hagino silver 1  
## 27648 Kosuke Hagino bronze 1  
## 5666 Kyle Chalmers gold 1  
## 17275 Madeline Dirado silver 1  
## 28304 Madeline Dirado bronze 1  
## 29011 Max Whitlock bronze 1  
## 18127 Michael Phelps silver 1  
## 8231 Penny Oleksiak gold 1  
## 19260 Penny Oleksiak silver 1  
## 9137 Sarah Sjostrom gold 1  
## 20166 Sarah Sjostrom silver 1  
## 31195 Sarah Sjostrom bronze 1  
## 31533 Simone Biles bronze 1  
## 10171 Tori Bowie gold 1  
## 21200 Tori Bowie silver 1  
## 32229 Tori Bowie bronze 1

v <- ggplot(x)+aes(x=fullname,y = value,fill =variable)+geom\_col()+ coord\_flip() + scale\_colour\_brewer(palette = "Set1")+ theme\_dark()  
v

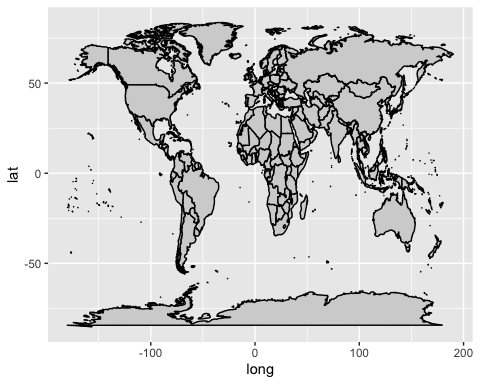


# options(scipen=999)   
# #valdiation of player's age across   
#   
#   
# ggPopluationboxplot <- ggplot(countries,aes(population)) +geom\_boxplot(fill = "#69b3a2", colour = "steelblue",outlier.colour = "red", outlier.shape = 1,na.rm=TRUE,orientation=NA,xmin = min(x), lower = quantile(x, 0.25), middle = median(x), upper = quantile(x, 0.75), xmax = max(x),stat = "identity")+ggtitle("boxplot - population spread")  
#   
# ggPopluationboxplot+theme

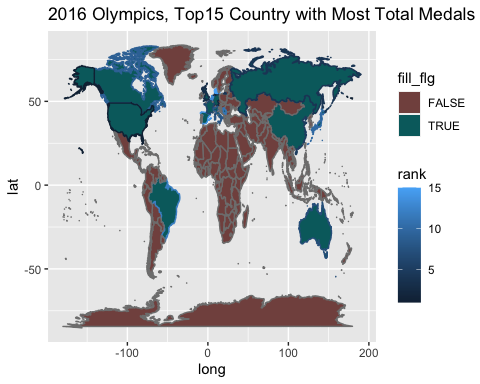
# Create new table that sums the total medals by country  
# Note: Some code used from this website: https://www.r-bloggers.com/how-to-make-a-global-map-in-r-step-by-step/  
  
Top15<- athletes %>%  
 group\_by(country) %>%  
 summarise(tot\_medals = sum(tot\_medals))

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

#Order from highest to lowest  
Top15<- Top15[order(-Top15$tot\_medals),]  
  
#Only take the top 15  
Top15<- Top15[1:15,]  
  
#create map showing total medals per country  
# Updates country in Top15 to Character  
Top15<- Top15 %>% mutate\_if(is.factor, as.character)  
  
#Update Country Names for United States to USA and United Kingdom to UK  
Top15$country <- recode(Top15$country,'United States' = 'USA','United Kingdom' = 'UK')  
  
Top15$rank<- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15) #add ranking order to use for colouring map by most medals  
  
#get world map  
world\_map<-map\_data("world")  
ggplot(world\_map, aes(x=long, y = lat, group=group))+  
 geom\_polygon(fill="lightgray", colour="black")



world.map.medals<-left\_join(world\_map,Top15,by=c('region'='country'))  
  
world.map.medals<-world.map.medals %>% mutate(fill\_flg = ifelse(is.na(rank),F,T))  
  
Top15map<- ggplot() +  
 geom\_polygon(data = world.map.medals, aes(x = long, y = lat, group = group, fill = fill\_flg,colour=rank)) +  
 scale\_fill\_hue(l=40,c=35) +  
 labs(title = '2016 Olympics, Top15 Country with Most Total Medals')  
Top15map



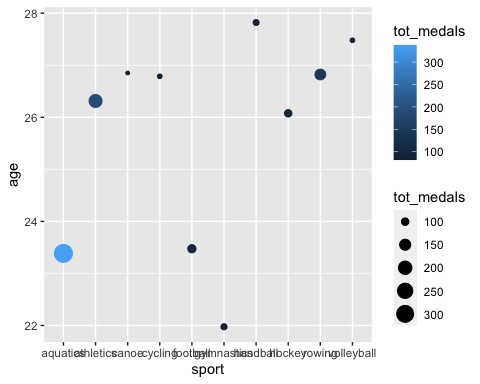
#Scatterplot of Sport, Age and Total Medals ranked by Top 10 sports with most total medals  
dfAgeSportMedals<- athletes %>%  
 group\_by(sport) %>%  
 summarise(age = mean(age),tot\_medals = sum(tot\_medals))

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

dfAgeSportMedals<- dfAgeSportMedals[order(-dfAgeSportMedals$tot\_medals),]  
dfAgeSportMedals<- dfAgeSportMedals[1:10,]  
dfAgeSportMedals

## # A tibble: 10 x 3  
## sport age tot\_medals  
## <fct> <dbl> <int>  
## 1 aquatics 23.4 338  
## 2 athletics 26.3 190  
## 3 rowing 26.8 144  
## 4 football 23.5 108  
## 5 hockey 26.1 99  
## 6 gymnastics 22.0 89  
## 7 handball 27.8 88  
## 8 cycling 26.8 83  
## 9 volleyball 27.5 83  
## 10 canoe 26.9 82

ggplot(dfAgeSportMedals,aes(x=sport,y=age))+geom\_point(aes(size=tot\_medals,color=tot\_medals))



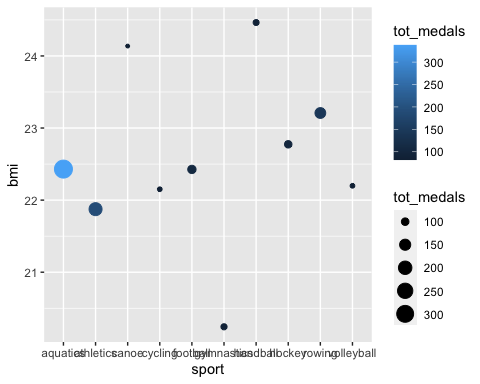
#Scatterplot of Sport, BMI and Total Medals ranked by Top 10 sports with most total medals  
dfbmiSportMedals<- athletes %>%  
 group\_by(sport) %>%  
 summarise(bmi = mean(bmi\_metric),tot\_medals = sum(tot\_medals))

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

dfbmiSportMedals<- dfbmiSportMedals[order(-dfbmiSportMedals$tot\_medals),]  
dfbmiSportMedals<- dfbmiSportMedals[1:10,]  
dfbmiSportMedals

## # A tibble: 10 x 3  
## sport bmi tot\_medals  
## <fct> <dbl> <int>  
## 1 aquatics 22.4 338  
## 2 athletics 21.9 190  
## 3 rowing 23.2 144  
## 4 football 22.4 108  
## 5 hockey 22.8 99  
## 6 gymnastics 20.2 89  
## 7 handball 24.5 88  
## 8 cycling 22.2 83  
## 9 volleyball 22.2 83  
## 10 canoe 24.1 82

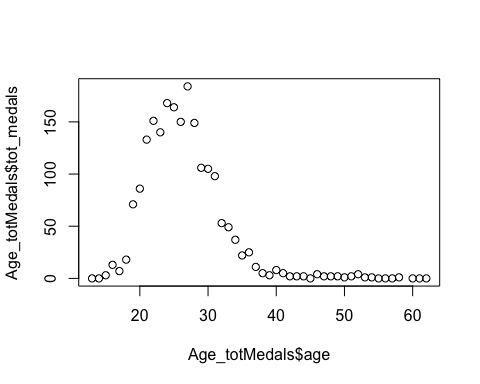
#Scatterplot of Sport, Age and Total Medals  
ggplot(dfbmiSportMedals,aes(x=sport,y=bmi))+geom\_point(aes(size=tot\_medals,color=tot\_medals))



#1.Regression Analysis- Age versus Total Medals  
  
Age\_totMedals<-athletes %>%  
 group\_by(age) %>%  
 summarise(tot\_medals = sum(tot\_medals))

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

Age\_totMedals<-Age\_totMedals[order(-Age\_totMedals$age),]  
  
# #Only take positive  
# Age\_totMedals<- Age\_totMedals[1:35,] # CAN DELETE ONCE INTEGRATE SATHISH's CODE THAT FIXES NEGATIVE AGES  
  
plot(Age\_totMedals$age,Age\_totMedals$tot\_medals)



Reg\_Age<-lm(formula = tot\_medals~age,data=Age\_totMedals)  
summary(Reg\_Age)

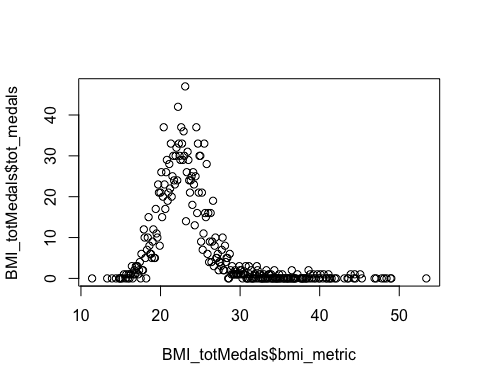
##   
## Call:  
## lm(formula = tot\_medals ~ age, data = Age\_totMedals)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -92.61 -25.78 -8.49 11.12 121.64   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 120.7043 20.0690 6.014 0.000000256 \*\*\*  
## age -2.1611 0.5055 -4.275 0.000092602 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 50.4 on 47 degrees of freedom  
## Multiple R-squared: 0.28, Adjusted R-squared: 0.2647   
## F-statistic: 18.28 on 1 and 47 DF, p-value: 0.0000926

# Assuming an alpha of 0.05, the p-value of 0.0371 indicates that age does have a correlation to total medals won

#2.Regression Analysis-BMI versus Total Medals  
  
#sathish ----  
BMI\_totMedals<-athletes %>%  
 group\_by(bmi\_metric) %>%  
 summarise(tot\_medals = sum(tot\_medals))

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

BMI\_totMedals<-BMI\_totMedals[order(-BMI\_totMedals$bmi\_metric),]  
  
plot(BMI\_totMedals$bmi\_metric,BMI\_totMedals$tot\_medals)



#sathish ----  
  
Reg\_BMI<-lm(formula = tot\_medals~bmi\_metric,data=BMI\_totMedals)  
summary(Reg\_BMI)

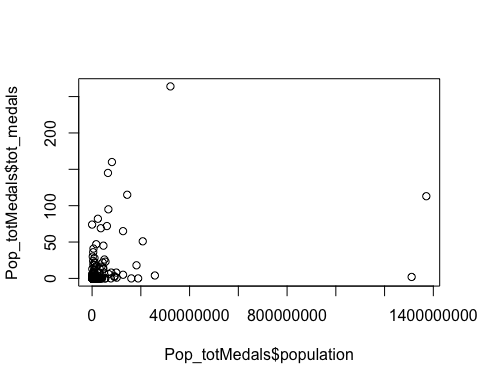
##   
## Call:  
## lm(formula = tot\_medals ~ bmi\_metric, data = BMI\_totMedals)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -16.911 -5.341 -2.860 3.025 36.527   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 23.18449 2.06210 11.243 < 0.0000000000000002 \*\*\*  
## bmi\_metric -0.55029 0.06903 -7.972 0.0000000000000469 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 9.776 on 265 degrees of freedom  
## Multiple R-squared: 0.1934, Adjusted R-squared: 0.1904   
## F-statistic: 63.55 on 1 and 265 DF, p-value: 0.00000000000004687

# Assuming an alpha of 0.05, the p-value of 0^-14 indicates that BMI does have a correlation to total medals won.

#3 Run regression for Population and Total Medals  
  
#sathish ----  
Pop\_totMedals<-athletes %>%  
 group\_by(population) %>%  
 summarise(tot\_medals = sum(tot\_medals))

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

Pop\_totMedals<-Pop\_totMedals[order(-Pop\_totMedals$population),]  
  
plot(Pop\_totMedals$population,Pop\_totMedals$tot\_medals)

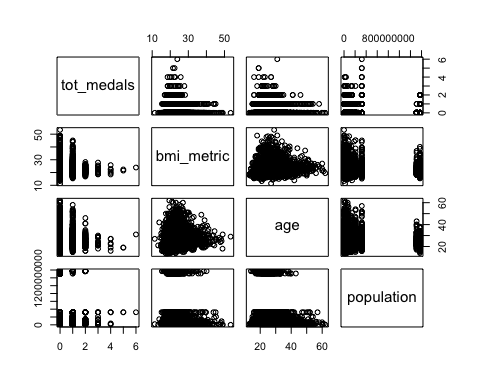


#sathish ----  
  
Reg\_Pop<-lm(formula = tot\_medals~population,data=Pop\_totMedals)  
summary(Reg\_Pop)

##   
## Call:  
## lm(formula = tot\_medals ~ population, data = Pop\_totMedals)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -94.410 -9.420 -8.797 -3.992 233.745   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 8.76825676781 2.38299570067 3.680 0.000313 \*\*\*  
## population 0.00000006685 0.00000001551 4.309 0.0000276 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 30.21 on 171 degrees of freedom  
## Multiple R-squared: 0.09797, Adjusted R-squared: 0.09269   
## F-statistic: 18.57 on 1 and 171 DF, p-value: 0.00002755

# Assuming an alpha of 0.05, the p-value of 0^-6 indicates that population does have a correlation to total medals won.Of note is that the p-value for BMI & population is smaller than age, so we will run the regression for age, BMI & population as out independent variables and analyze the output.

#multilinear Regression to predict toal medals from BMI, Age and Population  
  
#sathish -----  
pairs(~tot\_medals+bmi\_metric+age+population, data = athletes) # scotter plot



#sathish -----  
  
Reg\_Combined<-lm(formula = tot\_medals~bmi\_metric+age+population,data=athletes)  
summary(Reg\_Combined)

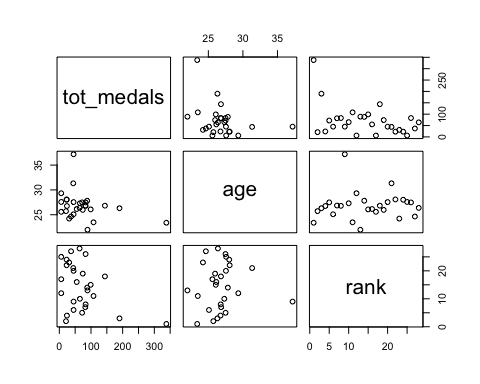
##   
## Call:  
## lm(formula = tot\_medals ~ bmi\_metric + age + population, data = athletes)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -0.3315 -0.1789 -0.1700 -0.1620 5.7979   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 0.09989395914488 0.03373030409318 2.962 0.00307 \*\*   
## bmi\_metric 0.00283950642852 0.00124074691677 2.289 0.02212 \*   
## age 0.00014515002964 0.00077227772509 0.188 0.85092   
## population 0.00000000009273 0.00000000001443 6.425 0.000000000138 \*\*\*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 0.4336 on 11025 degrees of freedom  
## Multiple R-squared: 0.00412, Adjusted R-squared: 0.003849   
## F-statistic: 15.21 on 3 and 11025 DF, p-value: 0.0000000007163

# From the regression output, population and BMI have a p-value below our alpha of 0.05 indicating they have a stronger correlation to total medals won than age.

#Multi-linear regression to predict total medals from the average age and sport  
  
  
  
All\_Sports<- athletes %>%  
 group\_by(sport) %>%  
 summarise(age = mean(age), tot\_medals = sum(tot\_medals))

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

All\_Sports$rank<- c(1,2,3,4,5,6,7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28) #add ranking order  
  
#sathish -----  
pairs(~tot\_medals+age+rank, data = All\_Sports) # scotter plot



#sathish -----  
  
  
agesport\_Reg<-lm(formula = tot\_medals~age+rank,data=All\_Sports)  
summary(agesport\_Reg)

##   
## Call:  
## lm(formula = tot\_medals ~ age + rank, data = All\_Sports)  
##   
## Residuals:  
## Min 1Q Median 3Q Max   
## -97.814 -25.201 -9.455 24.463 199.587   
##   
## Coefficients:  
## Estimate Std. Error t value Pr(>|t|)   
## (Intercept) 305.593 114.516 2.669 0.0132 \*  
## age -7.012 4.219 -1.662 0.1090   
## rank -3.196 1.417 -2.255 0.0332 \*  
## ---  
## Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1  
##   
## Residual standard error: 60.49 on 25 degrees of freedom  
## Multiple R-squared: 0.249, Adjusted R-squared: 0.1889   
## F-statistic: 4.144 on 2 and 25 DF, p-value: 0.0279

#

countrieswithmedals <- athletes %>% filter(tot\_medals>0) %>% group\_by(country) %>% summarize(medals=sum(tot\_medals,na.rm = TRUE), gold = sum(gold),silver = sum(silver),bronze = sum(bronze))

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

# countrieswithmedals  
  
countriesbyRank <- countrieswithmedals[order(-countrieswithmedals$medals),]  
countriesbyRank <- countriesbyRank %>% mutate(rank = 1:n())  
col\_order <- c("rank","country","medals","gold","silver","bronze")  
# countriesbyRank  
countriesbyRank <- countriesbyRank[, col\_order]  
  
Top5 <- data.frame(countriesbyRank[1:5,],stringsAsFactors = FALSE )  
Top5$country <- recode(Top5$country  
 ,'United States' = 'USA'  
 ,'United Kingdom' = 'UK'  
 )  
# Top5  
  
map.world <- map\_data("world")  
  
# glimpse(countriesbyRank)  
# as.factor(Top5$country) %>% levels()  
  
  
df.country\_points <- as.data.frame(Top5$country,stringsAsFactors = F)  
colnames(df.country\_points) <- c("country")  
df.country\_points$country <- as.character(df.country\_points$country)  
  
# ?register\_google  
  
register\_google("AIzaSyDWp5xj2USZ5eOYnL7V5sJMx0w76d1NCDY")  
   
df.country\_points <- mutate\_geocode(df.country\_points, country)

## Source : https://maps.googleapis.com/maps/api/geocode/json?address=USA&key=xxx

## Source : https://maps.googleapis.com/maps/api/geocode/json?address=Germany&key=xxx

## Source : https://maps.googleapis.com/maps/api/geocode/json?address=UK&key=xxx

## Source : https://maps.googleapis.com/maps/api/geocode/json?address=Russia&key=xxx

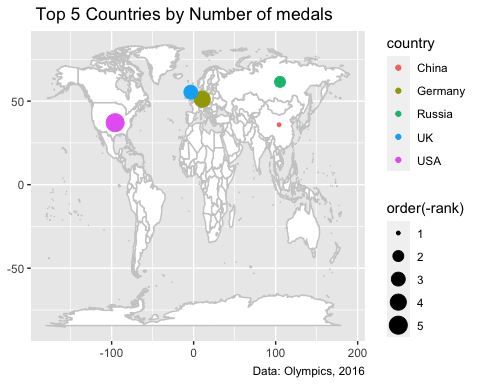
## Source : https://maps.googleapis.com/maps/api/geocode/json?address=China&key=xxx

# glimpse(df.country\_points)  
  
df.country\_points <- left\_join(df.country\_points,Top5)

## Joining, by = "country"

# df.country\_points  
  
  
map.world\_olympics <- left\_join(map.world, Top5, by = c('region' = 'country')) %>% mutate(fill\_flg = ifelse(is.na(rank),F,T))  
map.world\_olympics <- rename(map.world\_olympics, c("country"="region"))  
  
map.world\_olympics <- map.world\_olympics %>% mutate(fill\_flg = ifelse(fill\_flg=="TRUE", country, fill\_flg))  
  
map.world\_olympics$fill\_flg <- recode(map.world\_olympics$fill\_flg,'FALSE' = 'Others')  
  
#rename column name  
names(map.world\_olympics)[names(map.world\_olympics)=="fill\_flg"] <- "Countries"  
  
# map.world\_olympics  
# unique(map.world\_olympics[c("Countries")])  
# df.country\_points

#map with dotted plot  
top5.map <- ggplot() + geom\_polygon(data = map.world\_olympics, aes(x = long, y = lat, group = group),show.legend = FALSE,fill="white",color="#CCCCCC")   
top5.map <- top5.map + geom\_point(data = df.country\_points, aes(x = lon, y = lat,color=country,size=order(-rank)) )  
   
top5.map + labs(x = NULL,   
 y = NULL,   
 title = " Top 5 Countries by Number of medals",   
 caption = "Data: Olympics, 2016") + theme ()



#Colour Palette  
  
pal <- c(  
 "USA" = "red",  
 "Germany" = "orange",   
 "UK" = "yellow",   
 "Russia" = "forestgreen" ,  
 "China" ="blue",  
 "Others" ="#CCCCCC"  
)  
  
mapG <- ggplot() + geom\_polygon(data = map.world\_olympics, aes(x = long, y = lat, group = group,fill=Countries),show.legend = TRUE)+ scale\_fill\_manual(  
 values = pal,  
 limits = names(pal)  
 )  
mapG <- mapG +   
 labs(x = NULL,   
 y = NULL,   
 title = " Top 5 Countries by Number of medals",   
 caption = "Data: Olympics, 2016") +theme   
  
mapG

