

Business Analytics Session 5 Assignment

Problem Statement:

1. Consider only those participants who have all the data points

R Commands:

```
>library(readr)
```

```
>olympic_data <- read_csv("G:/Business Analytics_R_Acad glid/Assignments/Session 5  
Assignment1/olympic_data.csv")
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> library(readr)
> olympic_data <- read_csv("G:/Business Analytics_R_Acad glid/Assignments/session 5 Assignment1/olympic_data
.csv")
Parsed with column specification:
cols(
  Athlet = col_character(),
  Age = col_integer(),
  country = col_character(),
  Year = col_integer(),
  `close date of olympics` = col_character(),
  `sport name` = col_character(),
  gold = col_integer(),
  silver = col_integer(),
  Bronze = col_integer(),
  Total = col_integer()
)
> |
```

```
>View(olympic_data)
```

	Athlete	Age	Country	Year	Closing date of Olympics	Sport	Gold Medals	Silver Medals	Bronze Medals	Total Medals
1	Michael Phelps	23	United States	2008	08-24-08	Swimming	8	0	0	8
2	Michael Phelps	19	United States	2004	08-29-04	Swimming	6	0	2	8
3	Michael Phelps	27	United States	2012	08-12-2012	Swimming	4	2	0	6
4	Natalie Coughlin	25	United States	2008	08-24-08	Swimming	1	2	3	6
5	Aleksey Nemov	24	Russia	2000	10-01-2000	Gymnastics	2	1	3	6
6	Alicia Coutts	24	Australia	2012	08-12-2012	Swimming	1	3	1	5
7	Missy Franklin	17	United States	2012	08-12-2012	Swimming	4	0	1	5
8	Ryan Lochte	27	United States	2012	08-12-2012	Swimming	2	2	1	5
9	Allison Schmitt	22	United States	2012	08-12-2012	Swimming	3	1	1	5
10	Natalie Coughlin	21	United States	2004	08-29-04	Swimming	2	2	1	5
11	Ian Thorpe	17	Australia	2000	10-01-2000	Swimming	3	2	0	5
12	Dara Torres	33	United States	2000	10-01-2000	Swimming	2	0	3	5
13	Cindy Klassen	26	Canada	2006	02-26-06	Speed Skating	1	2	2	5
14	Nastia Liukin	18	United States	2008	08-24-08	Gymnastics	1	3	1	5
15	Marit Bjørgen	29	Norway	2010	02-28-10	Cross Country Skiing	3	1	1	5
16	Sun Yang	20	China	2012	08-12-2012	Swimming	2	1	1	4
17	Kirsty Coventry	24	Zimbabwe	2008	08-24-08	Swimming	1	3	0	4
18	Libby Lenton-Trickett	23	Australia	2008	08-24-08	Swimming	2	1	1	4

Showing 1 to 18 of 8,618 entries

```
>olympic_data<- olympic_data[complete.cases(olympic_data),]
```

```
>olympic_data
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> olympic_data<- olympic_data[complete.cases(olympic_data),]
> olympic_data
# A tibble: 8,613 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Natalie Coughlin 25 United States 2008 08-24-08 Swimming 1 2 3 6
5 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
6 Alicia Coutts 24 Australia 2012 08-12-2012 Swimming 1 3 1 5
7 Missy Franklin 17 United States 2012 08-12-2012 Swimming 4 0 1 5
8 Ryan Lochte 27 United States 2012 08-12-2012 Swimming 2 2 1 5
9 Allison Schmitt 22 United States 2012 08-12-2012 Swimming 3 1 1 5
10 Natalie Coughlin 21 United States 2004 08-29-04 Swimming 2 2 1 5
# ... with 8,603 more rows
>
```

```
>summary(olympic_data)
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> summary(olympic_data)
Athlete Length:8613 Min. :15.00 Country Length:8613 Min. :2000 Closing date of Olympics Length:8613 Length:8613 Min. :0.0000
Class :character 1st Qu.:23.00 Class :character 1st Qu.:2002 Class :character 1st Qu.:0.0000
Mode :character Median :26.00 Mode :character Median :2006 Mode :character Median :0.0000
Mean :26.41 Mean :2006 Mean :0.3649
3rd Qu.:29.00 3rd Qu.:2010 3rd Qu.:1.0000
Max. :61.00 Max. :2012 Max. :8.0000

Silver Medals Bronze Medals Total Medals
Min. :0.0000 Min. :0.0000 Min. :1.000
1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:1.000
Median :0.0000 Median :0.0000 Median :1.000
Mean :0.3629 Mean :0.3779 Mean :1.106
3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:1.000
Max. :3.0000 Max. :3.0000 Max. :8.000
>
```

2. Rank the participants in terms:

- Swimming
- Table Tennis
- Shooting
- Gymnastics
- Total Medal

R Commands:

```
>swimmingdata<- subset(olympic_data, Sport=="Swimming")
```

```
>swimmingdata
```

```
18:13 | (Top Level)
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> swimmingdata<- subset(olympic_data, Sport=="Swimming")
> swimmingdata
# A tibble: 487 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Natalie Coughlin 25 United States 2008 08-24-08 Swimming 1 2 3 6
5 Alicia Coutts 24 Australia 2012 08-12-2012 Swimming 1 3 1 5
6 Missy Franklin 17 United States 2012 08-12-2012 Swimming 4 0 1 5
7 Ryan Lochte 27 United States 2012 08-12-2012 Swimming 2 2 1 5
8 Allison Schmitt 22 United States 2012 08-12-2012 Swimming 3 1 1 5
9 Natalie Coughlin 21 United States 2004 08-29-04 Swimming 2 2 1 5
10 Ian Thorpe 17 Australia 2000 10-01-2000 Swimming 3 2 0 5
# ... with 477 more rows
>
```

```
>range(swimmingdata$`Total Medals`)
>rank_swimming<-swimmingdata[order(swimmingdata$`Total Medals`, decreasing = TRUE),]
>rank_swimming
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> range(swimmingdata$`Total Medals`)
[1] 1 8
> rank_swimming<-swimmingdata[order(swimmingdata$`Total Medals`, decreasing = TRUE),]
> rank_swimming
# A tibble: 487 x 10
  Athlete    Age Country Year `closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr>    <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Natalie Coughlin 25 United States 2008 08-24-08 Swimming 1 2 3 6
5 Alicia Coutts 24 Australia 2012 08-12-2012 Swimming 1 3 1 5
6 Missy Franklin 17 United States 2012 08-12-2012 Swimming 4 0 1 5
7 Ryan Lochte 27 United States 2012 08-12-2012 Swimming 2 2 1 5
8 Allison Schmitt 22 United States 2012 08-12-2012 Swimming 3 1 1 5
9 Natalie Coughlin 21 United States 2004 08-29-04 Swimming 2 2 1 5
10 Ian Thorpe 17 Australia 2000 10-01-2000 Swimming 3 2 0 5
# ... with 477 more rows
> |
```

```
>TTdata<- subset(olympic_data, Sport=="Table Tennis")
>TTdata
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> TTdata<- subset(olympic_data, Sport=="Table Tennis")
> TTdata
# A tibble: 67 x 10
  Athlete    Age Country Year `closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr>    <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Ding Ning 22 China 2012 08-12-2012 Table Tennis 1 1 0 2
2 Feng Tian wei 25 Singapore 2012 08-12-2012 Table Tennis 0 0 2 2
3 Li Xiaoxia 24 China 2012 08-12-2012 Table Tennis 2 0 0 2
4 Dmitriy Ovtcharov 23 Germany 2012 08-12-2012 Table Tennis 0 0 2 2
5 Wang Hao 28 China 2012 08-12-2012 Table Tennis 1 1 0 2
6 Zhang Jike 24 China 2012 08-12-2012 Table Tennis 2 0 0 2
7 Guo Yue 20 China 2008 08-24-08 Table Tennis 1 0 1 2
8 Ma Lin 28 China 2008 08-24-08 Table Tennis 2 0 0 2
9 Wang Hao 24 China 2008 08-24-08 Table Tennis 1 1 0 2
10 Wang Liqin 30 China 2008 08-24-08 Table Tennis 1 0 1 2
# ... with 57 more rows
> |
```

```
>range(TTdata$`Total Medals`)
>rank_TTdata<-TTdata[order(TTdata$`Total Medals`, decreasing = TRUE),]
>rank_TTdata
```

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> range(TTdata$`Total Medals`)
[1] 1 2
> rank_TTdata<-TTdata[order(TTdata$`Total Medals`, decreasing = TRUE),]
> rank_TTdata
# A tibble: 67 x 10
  Athlete    Age Country Year `closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr>    <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Ding Ning 22 China 2012 08-12-2012 Table Tennis 1 1 0 2
2 Feng Tian wei 25 Singapore 2012 08-12-2012 Table Tennis 0 0 2 2
3 Li Xiaoxia 24 China 2012 08-12-2012 Table Tennis 2 0 0 2
4 Dmitriy Ovtcharov 23 Germany 2012 08-12-2012 Table Tennis 0 0 2 2
5 Wang Hao 28 China 2012 08-12-2012 Table Tennis 1 1 0 2
6 Zhang Jike 24 China 2012 08-12-2012 Table Tennis 2 0 0 2
7 Guo Yue 20 China 2008 08-24-08 Table Tennis 1 0 1 2
8 Ma Lin 28 China 2008 08-24-08 Table Tennis 2 0 0 2
9 Wang Hao 24 China 2008 08-24-08 Table Tennis 1 1 0 2
10 Wang Liqin 30 China 2008 08-24-08 Table Tennis 1 0 1 2
# ... with 57 more rows
> |
```

```
>shootdata<- subset(olympic_data, Sport=="Shooting")
>shootdata
```

```

Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> shootdata<- subset(olympic_data, Sport=="Shooting")
> shootdata
# A tibble: 181 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int>
1 Niccolò Campriani 24 Italy 2012 08-12-2012 Shooting 1 1 0
2 Jin Jong-o 32 South Korea 2012 08-12-2012 Shooting 2 0 0
3 Olena Kostevych 27 Ukraine 2012 08-12-2012 Shooting 0 0 2
4 Jin Jong-o 28 South Korea 2008 08-24-08 Shooting 1 1 0
5 Katerina Kurková-Emmons 24 Czech Republic 2008 08-24-08 Shooting 1 1 0
6 Lyubov Galkina 31 Russia 2004 08-29-04 Shooting 1 1 0
7 Mariya Grozdeva 32 Bulgaria 2004 08-29-04 Shooting 1 0 1
8 Lee Bo-Na 23 South Korea 2004 08-29-04 Shooting 0 1 1
9 Mikhail Nestruev 35 Russia 2004 08-29-04 Shooting 1 1 0
10 Igor Basinsky 37 Belarus 2000 10-01-2000 Shooting 0 1 1
# ... with 171 more rows, and 1 more variables: `Total Medals` <int>
>

```

```

>range(shootdata$`Total Medals`)
>rank_shootdata<-shootdata[order(shootdata$`Total Medals`, decreasing = TRUE),]
>rank_shootdata

```

```

Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> range(shootdata$`Total Medals` )
[1] 1 2
> rank_shootdata<-shootdata[order(shootdata$`Total Medals`, decreasing = TRUE),]
> rank_shootdata
# A tibble: 181 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int>
1 Niccolò Campriani 24 Italy 2012 08-12-2012 Shooting 1 1 0
2 Jin Jong-o 32 South Korea 2012 08-12-2012 Shooting 2 0 0
3 Olena Kostevych 27 Ukraine 2012 08-12-2012 Shooting 0 0 2
4 Jin Jong-o 28 South Korea 2008 08-24-08 Shooting 1 1 0
5 Katerina Kurková-Emmons 24 Czech Republic 2008 08-24-08 Shooting 1 1 0
6 Lyubov Galkina 31 Russia 2004 08-29-04 Shooting 1 1 0
7 Mariya Grozdeva 32 Bulgaria 2004 08-29-04 Shooting 1 0 1
8 Lee Bo-Na 23 South Korea 2004 08-29-04 Shooting 0 1 1
9 Mikhail Nestruev 35 Russia 2004 08-29-04 Shooting 1 1 0
10 Igor Basinsky 37 Belarus 2000 10-01-2000 Shooting 0 1 1
# ... with 171 more rows, and 1 more variables: `Total Medals` <int>
>

```

```

> gymdata<- subset(olympic_data, Sport=="Gymnastics")
> gymdata

```

```

Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> gymdata<- subset(olympic_data, Sport=="Gymnastics")
> gymdata
# A tibble: 194 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
2 Nastia Liukin 18 United States 2008 08-24-08 Gymnastics 1 3 1 5
3 Aliya Mustafina 17 Russia 2012 08-12-2012 Gymnastics 1 1 2 4
4 Shawn Johnson 16 United States 2008 08-24-08 Gymnastics 1 3 0 4
5 Aly Raisman 18 United States 2012 08-12-2012 Gymnastics 2 0 1 3
6 Kohei Uchimura 23 Japan 2012 08-12-2012 Gymnastics 1 2 0 3
7 Zou Kai 24 China 2012 08-12-2012 Gymnastics 2 0 1 3
8 Cheng Fei 20 China 2008 08-24-08 Gymnastics 1 0 2 3
9 Yang Wei 28 China 2008 08-24-08 Gymnastics 2 1 0 3
10 Yang Yilin 15 China 2008 08-24-08 Gymnastics 1 0 2 3
# ... with 184 more rows
>

```

```

>range(gymdata$`Total Medals`)
>rank_gymdata<-gymdata[order(gymdata$`Total Medals`, decreasing = TRUE),]
>rank_gymdata

```

```

Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> range(gymdata$`Total Medals` )
[1] 1 6
> rank_gymdata<-gymdata[order(gymdata$`Total Medals`, decreasing = TRUE),]
> rank_gymdata
# A tibble: 194 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
2 Nastia Liukin 18 United States 2008 08-24-08 Gymnastics 1 3 1 5
3 Aliya Mustafina 17 Russia 2012 08-12-2012 Gymnastics 1 1 2 4
4 Shawn Johnson 16 United States 2008 08-24-08 Gymnastics 1 3 0 4
5 Aly Raisman 18 United States 2012 08-12-2012 Gymnastics 2 0 1 3
6 Kohei Uchimura 23 Japan 2012 08-12-2012 Gymnastics 1 2 0 3
7 Zou Kai 24 China 2012 08-12-2012 Gymnastics 2 0 1 3
8 Cheng Fei 20 China 2008 08-24-08 Gymnastics 1 0 2 3
9 Yang Wei 28 China 2008 08-24-08 Gymnastics 2 1 0 3
10 Yang Yilin 15 China 2008 08-24-08 Gymnastics 1 0 2 3
# ... with 184 more rows
>

```

```

>totaldata<- olympic_data

```


>totaldata

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> totaldata<- olympic_data
> totaldata
# A tibble: 8,613 x 10
  Athlete   Age Country Year `closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Natalie Coughlin 25 United States 2008 08-24-08 Swimming 1 2 3 6
5 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
6 Alicia Coutts 24 Australia 2012 08-12-2012 Swimming 1 3 1 5
7 Missy Franklin 17 United States 2012 08-12-2012 Swimming 4 0 1 5
8 Ryan Lochte 27 United States 2012 08-12-2012 Swimming 2 2 1 5
9 Allison Schmitt 22 United States 2012 08-12-2012 Swimming 3 1 1 5
10 Natalie Coughlin 21 United States 2004 08-29-04 Swimming 2 2 1 5
# ... with 8,603 more rows
> |
```

rank_totaldata<-totaldata[order(totaldata\$`Total Medals`, decreasing = TRUE),]

rank_totaldata

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> rank_totaldata<-totaldata[order(totaldata$`Total Medals`, decreasing = TRUE),]
> rank_totaldata
# A tibble: 8,613 x 10
  Athlete   Age Country Year `closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Natalie Coughlin 25 United States 2008 08-24-08 Swimming 1 2 3 6
5 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
6 Alicia Coutts 24 Australia 2012 08-12-2012 Swimming 1 3 1 5
7 Missy Franklin 17 United States 2012 08-12-2012 Swimming 4 0 1 5
8 Ryan Lochte 27 United States 2012 08-12-2012 Swimming 2 2 1 5
9 Allison Schmitt 22 United States 2012 08-12-2012 Swimming 3 1 1 5
10 Natalie Coughlin 21 United States 2004 08-29-04 Swimming 2 2 1 5
# ... with 8,603 more rows
> |
```

3. Rank the Categories in terms of Age.(Higher the Age,Higher the Rank)

R Commands:

>highage<-olympic_data[order(olympic_data\$Age, decreasing = TRUE),]

>highage

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> highage<-olympic_data[order(olympic_data$Age, decreasing = TRUE),]
> highage
# A tibble: 8,613 x 10
  Athlete   Age Country Year `close date of olympics` `Sport name` Gold Silver Bronze Total
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Ian Millar 61 Canada 2008 08-24-08 Equestrian 0 1 0 1
2 Mark Todd 56 New Zealand 2012 08-12-2012 Equestrian 0 0 1 1
3 Mac Cone 55 Canada 2008 08-24-08 Equestrian 0 1 0 1
4 Nick Skelton 54 Great Britain 2012 08-12-2012 Equestrian 1 0 0 1
5 Lesley Thompson-Willie 52 Canada 2012 08-12-2012 Rowing 0 1 0 1
6 Peter Charles 52 Great Britain 2012 08-12-2012 Equestrian 1 0 0 1
7 Peter Thomsen 51 Germany 2012 08-12-2012 Equestrian 1 0 0 1
8 Mary Thomson-King 51 Great Britain 2012 08-12-2012 Equestrian 0 1 0 1
9 Ilario Di Buò 51 Italy 2008 08-24-08 Archery 0 1 0 1
10 Andrew Nicholson 50 New Zealand 2012 08-12-2012 Equestrian 0 0 1 1
# ... with 8,603 more rows
> |
```

4. Identify Year wise top participants in terms of:

- Swimming
- Table Tennis
- Shooting
- Gymnastics
- Total Medal

>subset(olympic_data,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(olympic_data,!duplicated(Year))
# A tibble: 7 x 10
  Athlete Age Country Year `Close date of olympics` Sport name Gold silver Bronze Total
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
5 Cindy Klassen 26 Canada 2006 02-26-06 Speed Skating 1 2 2 5
6 Marit Bjørgen 29 Norway 2010 02-28-10 Cross Country Skiing 3 1 1 5
7 Ole Einar Bjørndalen 28 Norway 2002 02-24-02 Biathlon 4 0 0 4
```

>subset(swimmingdata,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(swimmingdata,!duplicated(Year))
# A tibble: 4 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Ian Thorpe 17 Australia 2000 10-01-2000 Swimming 3 2 0 5
```

>subset(TTdata,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(TTdata,!duplicated(Year))
# A tibble: 4 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Ding Ning 22 China 2012 08-12-2012 Table Tennis 1 1 0 2
2 Guo Yue 20 China 2008 08-24-08 Table Tennis 1 0 1 2
3 Zhang Yining 22 China 2004 08-29-04 Table Tennis 2 0 0 2
4 Kong Linghui 24 China 2000 10-01-2000 Table Tennis 1 1 0 2
```

>subset(shootdata,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(shootdata,!duplicated(Year))
# A tibble: 4 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Niccolò Campriani 24 Italy 2012 08-12-2012 Shooting 1 1 0 2
2 Jin Jong-o 28 South Korea 2008 08-24-08 Shooting 1 1 0 2
3 Lyubov Galkina 31 Russia 2004 08-29-04 Shooting 1 1 0 2
4 Igor Basinsky 37 Belarus 2000 10-01-2000 Shooting 0 1 1 2
```

>subset(gymdata,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(gymdata,!duplicated(Year))
# A tibble: 4 x 10
  Athlete Age Country Year `Closing date of Olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
2 Nastia Liukin 18 United States 2008 08-24-08 Gymnastics 1 3 1 5
3 Aliya Mustafina 17 Russia 2012 08-12-2012 Gymnastics 1 1 2 4
4 Marian Dragulescu 23 Romania 2004 08-29-04 Gymnastics 0 1 2 3
```

>subset(totaldata,!duplicated(Year))

```
Console G:/Business Analytics_R_Acad glid/Assignments/Session 5 Assignment1/
> subset(totaldata,!duplicated(Year))
# A tibble: 7 x 10
  Athlete Age Country Year `Closing date of olympics` Sport `Gold Medals` `Silver Medals` `Bronze Medals` `Total Medals`
  <chr> <int> <chr> <int> <chr> <chr> <int> <int> <int> <int>
1 Michael Phelps 23 United States 2008 08-24-08 Swimming 8 0 0 8
2 Michael Phelps 19 United States 2004 08-29-04 Swimming 6 0 2 8
3 Michael Phelps 27 United States 2012 08-12-2012 Swimming 4 2 0 6
4 Aleksey Nemov 24 Russia 2000 10-01-2000 Gymnastics 2 1 3 6
5 Cindy Klassen 26 Canada 2006 02-26-06 Speed Skating 1 2 2 5
6 Marit Bjørgen 29 Norway 2010 02-28-10 Cross Country Skiing 3 1 1 5
7 Ole Einar Bjørndalen 28 Norway 2002 02-24-02 Biathlon 4 0 0 4
```