Computation of summary statistics using R

The Olympic Games: *India first participated at the Olympic Games in 1900, with a lone athlete (Norman Pritchard) winning two medals in athletics. The nation first sent a team to the Summer Olympic Games in 1920 and has participated in every Summer Games since then. India has also competed at several Winter Olympic Games since 1964. Indian athletes have won a total of 26 medals, mostly in field hockey, but this is changing. For a period of time, India's men's field hockey team was dominant in Olympic competition, winning eleven medals in twelve Olympiads between 1928 and 1980, including six successive gold medals from 1928–1956. India gave its best performance in the 2012 London Olympics by winning 2 silver medals and 4 bronze medals.*

Source:**http://olympics.wikia.com/wiki/India**

The dataset: **120-years-of-olympic-history-athletes-and-results**.This is a historical dataset on the modern Olympic Games, including all the Games from Athens 1896 to Rio 2016 taken from **sports-reference** and was made available in the CSV format on **Kaggle**. Note that the Winter and Summer Games were held in the same year up until 1992. After that, they staggered them such that Winter Games occur on a four-year cycle starting with 1994, then Summer in 1996, then Winter in 1998, and so on.
Note: Ignore tuples with NA where necessary.

Problem Statement 1:

Who are the medal winners from India? (Print only their names as output.) Find the sum of all the medals won in different Sports starting from 1960. Which Sport (among Indian awardees) has won the highest number of medals?

Problem Statement 2:

Find the players who have participated in more than one edition of the Olympics. Find the players who have also won at multiple appearances. Find the player with the highest number of medals won in the same edition of the Olympics and display all relevant details.

Problem Statement 3:

For every country, for every edition of the Olympics, find the number of male and female participants from every country. Find the mean, median, mode, IQR, SD of the same separately for the two. Draw a box plot of the distribution and compare the two. Find the value of the 90th Percentile for each.

Problem Statement 4:

Plot the histogram of Medal winners over the years for India (total number of players who have won medals against Year). What is the nature of the graph? What inferences can we draw from the same?
The age of a winner of a gold medal in Olympics in Athletics has a strong correlation with the height of the athlete. Comment on the same.

Problem Statement 5:

What is the importance of sampling and why is it used? What are the differences between stratified sampling and cluster sampling? Explain with an example