**Map Reduce**

**Olympic Medals & Athletes Data Set [2000-2012]**

**Data Set Description**

Column 1: Athlete Name: Name of Athlete

Column 2: Age: Age of respective Athlete

Column 3: Country: Country of Athlete

Column 4: Year: Year of Olympic game

Column 4: Closing Ceremony Date: Closing Ceremony Date of Game

Column 5: Sport: Sports Athletes Participated (Swimming, High-Jump)

Column 6: Gold Medals: Gold Medals won by athlete

Column 7: Silver Medals: Silver Medals won by athlete

Column 8: Bronze Medals: Bronze Medals won by athlete

Column 9: Total Medals: Total Medals won by athlete

**Sample Dataset**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Athlete | Age | Country | Year | Closing Ceremony Date | Sport | Gold Medals | Silver Medals | Bronze Medals | Total Medals |
| Michael Phelps | 23 | United States | 2008 | 24-08-2008 | Swimming | 8 | 0 | 0 | 8 |
| Michael Phelps | 19 | United States | 2004 | 29-08-2004 | Swimming | 6 | 0 | 2 | 8 |
| Michael Phelps | 27 | United States | 2012 | 12-08-2012 | Swimming | 4 | 2 | 0 | 6 |
| Natalie Coughli | 25 | United States | 2008 | 24-08-2008 | Swimming | 1 | 2 | 3 | 6 |

1) List the players participated in each year

2) Count the unique players participated in each year

3) Youngest and Oldest athlete who participated in each year

4) Total medals won by each country for each year

5) Athlete won the maximum number of medals in each year

6) Count the number of players participated on basis of Sport (Swimming, Cycling, etc..)

7) List the athletes participated consecutively more than 2 times

8) Count the total number of Gold, Silver, Bronze medals won by each country for each year?

9) List the maximum number of medals won by Country on each year?

10) List the country that doesn’t own any medals on each year?

11) Name the oldest and youngest player who won medal for each year?

12) List the sport who contributed Maximum no of medals to country for each year

**PIG**

**Movie Information and ratings from users**

**Data Set Description**

u.data: This is a tab separated list of user id | item id | rating | timestamp.

u.item: This is a tab separated list of movie id | movie title | release date | video release date IMDb URL | unknown | Action | Adventure | Animation |Children's | Comedy | Crime | Documentary | Drama | Fantasy | Film-Noir | Horror | Musical | Mystery | Romance | Sci-Fi | Thriller | War | Western |

u.genre: A list of the genres.

u.user: This is a tab separated list of user id | age | gender | occupation | zip code

u.occupation: A list of the occupations

1) Movies released in the year between 1990 to 2000?

2) Top 10 Movies have highest overall rating

3) List movies which are of genre war

4) List movies which have more than one genre

5) List movies which are of genre Sci-fi and released in the year of 2000

6) List the single word named movie which are of genre Romance

7) List movies which their rating are more than 5

8) List the movies which have rating from user educators?

9) Count the number of users on bases of occupation?

10) Count the no of M/F users?

11) Count the movies rated by female user who are of Student?

12) Count the movies rated?

**Hive**

1) List the Movies which have title of Number?

2) List the Movies which are released in the day of 1st Jan?

3) List the Movies rated by the user 210?

4) Movies which are rated on the same timestamp?

5) Is the user of Scientist who likes Romance Movies?

6) Is the user of above age 40 who rated Animation Movies?

7) Is the user of below age 15 who rated Horror Movies?

8) Name the documentary movie which received high rating?

10) Name the Romance Movie which received low rating?

11) Movies which are released before 1950's which are of genre Film-Noir?