

**PART B ->MAP-  
REDUCE**

# DATASET INFORMATION

**Dataset Name : Titanic Data Set**

**Description :** Dataset consists of the below mentioned information

Column Name	Column Description
PassengerId	Ticket Id of passengers
Survived	0 if true 1 if died
PClass	Passenger ticket class
Name	Name of the pax
Sex	Gender
Age	Age
Sibsp	Ship specific info
Parch	Ship specific
Ticket	Ticket info
Fare	Ticket fare
Cabin	Alloted cabin
Embarked	Boarded or not

## Dataset Source And Analysis Information

Source : <https://drive.google.com/file/d/0ByJLBTmJojjzNmV0dk1EMmwwQ1U/view>

Technology used : Hadoop framework

What will be the result of our analysis? we will find the average age of males and females who died in the Titanic tragedy and number of people died or survived in each class with their genders and ages.

**THANK YOU !**