missing_data_R

```
dataset = read.csv(file='Data.csv')
dataset
```

```
##    Country Age Salary Purchased
## 1   France  44  72000        No
## 2    Spain  27  48000       Yes
## 3  Germany  30  54000        No
## 4    Spain  38  61000        No
## 5  Germany  40     NA       Yes
## 6   France  35  58000       Yes
## 7    Spain  NA  52000        No
## 8   France  48  79000       Yes
## 9  Germany  50  83000        No
## 10  France  37  67000       Yes
```

```
dataset$Age = ifelse(is.na(dataset$Age),
                     ave(dataset$Age, FUN = function(x) mean(x,na.rm = TRUE)),
                     dataset$Age)
dataset
```

```
##    Country      Age Salary Purchased
## 1   France 44.00000  72000        No
## 2    Spain 27.00000  48000       Yes
## 3  Germany 30.00000  54000        No
## 4    Spain 38.00000  61000        No
## 5  Germany 40.00000     NA       Yes
## 6   France 35.00000  58000       Yes
## 7    Spain 38.77778  52000        No
## 8   France 48.00000  79000       Yes
## 9  Germany 50.00000  83000        No
## 10  France 37.00000  67000       Yes
```

```
dataset$Salary = ifelse(is.na(dataset$Salary),
                     ave(dataset$Salary, FUN = function(x) mean(x,na.rm= TRUE)),
                     dataset$Salary)
dataset
```

```
##    Country      Age   Salary Purchased
## 1   France 44.00000 72000.00        No
## 2    Spain 27.00000 48000.00       Yes
## 3  Germany 30.00000 54000.00        No
## 4    Spain 38.00000 61000.00        No
## 5  Germany 40.00000 63777.78       Yes
## 6   France 35.00000 58000.00       Yes
## 7    Spain 38.77778 52000.00        No
## 8   France 48.00000 79000.00       Yes
## 9  Germany 50.00000 83000.00        No
## 10  France 37.00000 67000.00       Yes
```