

PROGRAM:

```
#import necessary libraries
import pandas as pd
import numpy as np
#importing the dataset
dataset = pd.read_csv('/empdata.csv')
#converting the dataset into dataframe
df = pd.DataFrame(dataset)
df
#remove the rows with null values
new_df = df.dropna()
new_df
#calculate mean value of attribute
x = df['Age '].mean()
#replace mean value in place of null values
df["Age "].fillna(x, inplace = True)
df
df["salary"].fillna(25000, inplace = True)
df
#remove duplicates
df.drop_duplicates(inplace = True)
df
```

OUTPUT:

```
#  
#
```



	Emp_id	Name	Age	salary	city
0	101	Poovarasana	19	300000.0	erode
1	102	praveen	19	NaN	chennai
2	103	prakash	20	50000.0	coimbatore
3	104	prasad	19	40000.0	chennai
4	105	jp	22	NaN	coimbatore



```
df.drop_duplicates(inplace = True)  
df
```



	Emp_id	Name	Age	salary	city
0	101	Poovarasana	19	300000.0	erode
1	102	praveen	19	25000.0	chennai
2	103	prakash	20	50000.0	coimbatore
3	104	prasad	19	40000.0	chennai
4	105	jp	22	25000.0	coimbatore

