```python
In [1]:   #Task:Mutiple Regression on data for India and Pakistan
          #Program By:Ayush Pandey
          #Email Id:1805290@kiit.ac.in
          #DATE:19-Oct-2021
          #Python Version:3.7
          #CAVEATS:None
          #LICENSE:None
```

```python
In [2]:   import pandas as pd
          import numpy as np
          import matplotlib.pyplot as plt
```

```python
In [3]:   #For reading the csv file
          df=pd.read_csv('data_2.csv')
```

```python
In [4]:   df.head()
```

Out[4]:

	Entity	Code	Year	Unsafe water source	Unsafe sanitation	No access to handwashing facility	Household air pollution from solid fuels	Non-exclusive breastfeeding	Discontinued breastfeeding	Child wasting	...	High systolic blood pressure	Smoking	Iron deficiency	Vitamin A deficiency
0	Afghanistan	AFG	1990	7554.049543	5887.747628	5412.314513	22388.49723	3221.138842	156.097553	22778.84925	...	28183.98335	6393.667372	726.431294	9344.131952
1	Afghanistan	AFG	1991	7359.676749	5732.770160	5287.891103	22128.75821	3150.559597	151.539851	22292.69111	...	28435.39751	6429.253320	739.245799	9330.182378
2	Afghanistan	AFG	1992	7650.437822	5954.804987	5506.657363	22873.76879	3331.349048	156.609194	23102.19794	...	29173.61120	6561.054957	873.485341	9769.844533
3	Afghanistan	AFG	1993	10270.731380	7986.736613	7104.620351	25599.75628	4477.006100	206.834451	27902.66996	...	30074.76091	6731.972560	1040.047422	11433.769490
4	Afghanistan	AFG	1994	11409.177110	8863.010065	8051.515953	28013.16720	5102.622054	233.930571	32929.00593	...	30809.49117	6889.328118	1101.764645	12936.955860

5 rows × 32 columns

```python
In [5]:   #Comparing the Death of India and Pakistan based on Unsafe water source,Unsafe sanitation,No access to handwashing facility
          df1=df.groupby(['Entity'],as_index=False).agg({'Unsafe water source':'sum','Unsafe sanitation':'sum','No access to handwashing facility':'sum'})
```

```python
In [6]:   #Storing all the data's where Entity is India and Pakistan
          India=df1[df1['Entity']=='India']
          Pakistan=df1[df1['Entity']=='Pakistan']
```

```python
In [7]:   x=['Unsafe water source','Unsafe sanitation','No access to handwashing facility']
          y=[India['Unsafe water source'].item(),India['Unsafe sanitation'].item(),India['No access to handwashing facility'].item()]
          y1=[Pakistan['Unsafe water source'].item(),Pakistan['Unsafe sanitation'].item(),Pakistan['No access to handwashing facility'].item()]
          X_axis = np.arange(len(x))
          plt.bar(X_axis - 0.2,y,0.4,label="India")
          plt.bar(X_axis + 0.2,y1,0.4,color="red",label="Palkistan")
          plt.xticks(rotation=90)
          plt.xlabel(['Unsafe water source','Unsafe sanitation','No access to handwashing facility'])
          plt.ylabel("No. of Death")
          plt.legend()
          plt.show()
```
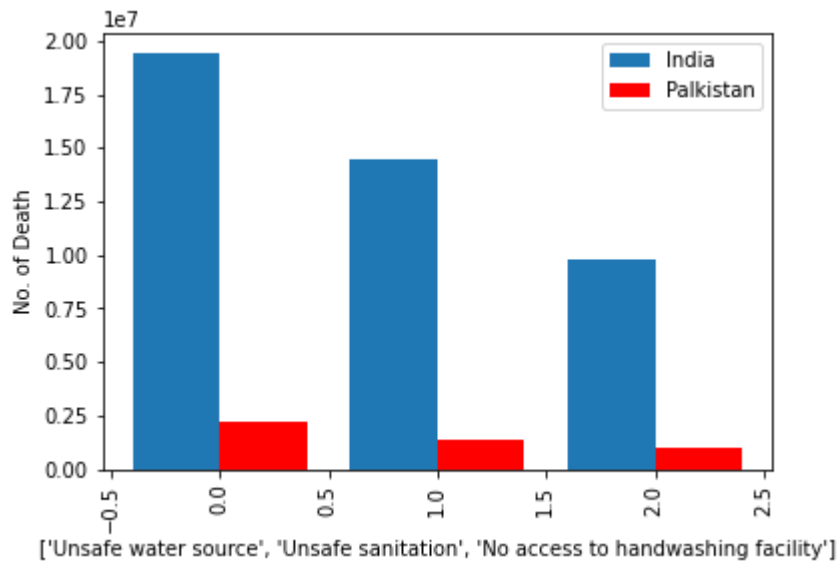


```python
In [8]:   India=df[df['Entity']=='India']
          Pakistan=df[df['Entity']=='Pakistan']
```

```python
In [9]:   from sklearn.linear_model import LinearRegression
```

```python
In [10]:  #For India Doing Mutiple regression
          #Independent
          x=India[['Unsafe water source','Unsafe sanitation']]
          #Depenent
          y=India['No access to handwashing facility']
          model=LinearRegression()
          model.fit(x,y)
          print(model.score(x,y)*100)
          print(model.coef_)

          #This shows that the Unsafe water souce is the main cause for no access to handwashing facility as the coefficient value is more than 50%.
```

```
99.87663882671968
[0.53941374 0.12691466]
```

```python
In [11]:  #For Pakistan Doing Mutiple regression
          #Independent
          x=Pakistan[['Unsafe water source','Unsafe sanitation']]
          #Depenent
          y=Pakistan['No access to handwashing facility']
          model=LinearRegression()
          model.fit(x,y)
          print(model.score(x,y)*100)
          print(model.coef_)

          #This shows that the Unsafe sanitation souce is the main cause for no access to handwashing facility as the coefficient value is around 40%.
```

```
99.83390966872696
[0.14601608 0.38339417]
```

```python
In [ ]:
```