In [2]:
```
# Import packages
# read the data

import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
```

In [ ]:
```
file_path="C:\\Users\\kurre\\OneDrive\\Documents\\Naresh IT\\datafiles\\Vis
visa_df=pd.read_csv(file_path)
visa_df
```

In [3]:
```
dict1={'Name':['Ram','sita','Laxman',np.nan],'Age':[30,31,np.nan,33],'City'
pd.DataFrame(dict1)
```

Out[3]:

	Name	Age	City
0	Ram	30.0	NaN
1	sita	31.0	hyd
2	Laxman	NaN	pune
3	NaN	33.0	chenni

In [4]:
```
dict1={'Name':['Ram','sita','Laxman',None],'Age':[30,31,None,33],'City':[No
pd.DataFrame(dict1)
```

Out[4]:

	Name	Age	City
0	Ram	30.0	None
1	sita	31.0	hyd
2	Laxman	NaN	pune
3	None	33.0	chenni

In [6]:
```
dict1={'Name':['Ram','sita','Laxman','Null'],'Age':[30,31,'NUll',33],'City'
pd.DataFrame(dict1)
```

Out[6]:

	Name	Age	City
0	Ram	30	Null
1	sita	31	hyd
2	Laxman	NUll	pune
3	Null	33	chenni

In [7]:
```
dict1={'Name':['Ram','sita','Laxman',np.nan],'Age':[30,31,np.nan,33],'City'
df=pd.DataFrame(dict1)
df
```

Out[7]:

	Name	Age	City
0	Ram	30.0	NaN
1	sita	31.0	hyd
2	Laxman	NaN	pune
3	NaN	33.0	chenni

$Method - 1$

Fill with a random numbers to all Null values

In [9]:
```
df.isnull()
```

Out[9]:

	Name	Age	City
0	False	False	True
1	False	False	False
2	False	True	False
3	True	False	False

In [10]:
```
df.isnull().sum()
```

Out[10]:
```
Name    1
Age     1
City    1
dtype: int64
```

In [14]:
```
df.isnull().sum().count()
```

Out[14]:
```
3
```

In [11]:
```
df.isnull().count()
```

Out[11]:
```
Name    4
Age     4
City    4
dtype: int64
```

In [12]:
```
df.isna()
```

Out[12]:

	Name	Age	City
0	False	False	True
1	False	False	False
2	False	True	False
3	True	False	False

$fillna$

In [17]:
```
df.fillna(40,inplace=True)
df
```

Out[17]:

	Name	Age	City
0	Ram	30.0	40
1	sita	31.0	hyd
2	Laxman	40.0	pune
3	40	33.0	chenni

In [18]:
```
df.dtypes
```

Out[18]:
```
Name      object
Age      float64
City      object
dtype: object
```

$method - 2$

fill the random value by using column wise

In [3]:
```
###################### Read the data again ##########################
dict1={'Name':['Ram','Sita','Laxman',np.nan],
       'Age':[30,31,np.nan,33],
       'City':[np.nan,'Hyd','Pune','Chennai']}

df=pd.DataFrame(dict1)
df
```

Out[3]:

	Name	Age	City
0	Ram	30.0	NaN
1	Sita	31.0	Hyd
2	Laxman	NaN	Pune
3	NaN	33.0	Chennai

In [4]:
```
df['Name'].fillna("Raheem",inplace=True)
df['Age'].fillna(32,inplace=True) # based on data type
df['City'].fillna("Blr",inplace=True)
df
```

Out[4]:

	Name	Age	City
0	Ram	30.0	Blr
1	Sita	31.0	Hyd
2	Laxman	32.0	Pune
3	Raheem	33.0	Chennai

$method - 3$

- pad
- bfill
- backfill
- ffill

In [6]:
```
################## Read the data again #################
dict1={'Name':['Ram','Sita','Laxman',np.nan],
       'Age':[30,31,np.nan,33],
       'City':[np.nan,'Hyd','Pune','Chennai']}

df=pd.DataFrame(dict1)
df
```

Out[6]:

	Name	Age	City
0	Ram	30.0	NaN
1	Sita	31.0	Hyd
2	Laxman	NaN	Pune
3	NaN	33.0	Chennai

In [28]:
```
dict1={'Name':['Ram','sita','Laxman',np.nan],'Age':[30,31,np.nan,33],'City'
df=pd.DataFrame(dict1)
df
```

Out[28]:

	Name	Age	City
0	Ram	30.0	NaN
1	sita	31.0	hyd
2	Laxman	NaN	pune
3	NaN	33.0	chenni

In [7]:
```
print("========original===========")
print(df)

print("======pad===========")
print(df.fillna(method='pad'))

print("======ffill===========")
print(df.fillna(method='ffill'))

print("======bfill===========")
print(df.fillna(method='bfill'))

print("======backfill===========")
print(df.fillna(method='backfill'))
```

```
========original===========
     Name   Age     City
0     Ram  30.0      NaN
1    Sita  31.0      Hyd
2  Laxman   NaN     Pune
3     NaN  33.0  Chennai
======pad===========
     Name   Age     City
0     Ram  30.0      NaN
1    Sita  31.0      Hyd
2  Laxman  31.0     Pune
3  Laxman  33.0  Chennai
======ffill===========
     Name   Age     City
0     Ram  30.0      NaN
1    Sita  31.0      Hyd
2  Laxman  31.0     Pune
3  Laxman  33.0  Chennai
======bfill===========
     Name   Age     City
0     Ram  30.0      Hyd
1    Sita  31.0      Hyd
2  Laxman  33.0     Pune
3     NaN  33.0  Chennai
======backfill===========
     Name   Age     City
0     Ram  30.0      Hyd
1    Sita  31.0      Hyd
2  Laxman  33.0     Pune
3     NaN  33.0  Chennai
```

In [8]:
```
print("========original==========")
print(df)

print("======pad==========")
print(df.fillna(method='pad'))

print("======pad==========")
print(df.fillna(method='pad',axis=1))
```

```
========original==========
     Name   Age     City
0     Ram  30.0      NaN
1    Sita  31.0      Hyd
2  Laxman   NaN     Pune
3     NaN  33.0  Chennai
======pad==========
     Name   Age     City
0     Ram  30.0      NaN
1    Sita  31.0      Hyd
2  Laxman  31.0     Pune
3  Laxman  33.0  Chennai
======pad==========
     Name     Age     City
0     Ram    30.0     30.0
1    Sita    31.0      Hyd
2  Laxman  Laxman     Pune
3     NaN    33.0  Chennai
```

In [ ]:
```
```

$method - 4$

- mean
- median
- mode

- mode is used for categorical data
- mean and median is used for numerical data

In [39]:
```
dict1={'Name':['Ram','sita','Laxman',np.nan],'Age':[30,31,np.nan,33],'City'
df=pd.DataFrame(dict1)
df
```

Out[39]:

	Name	Age	City
0	Ram	30.0	NaN
1	sita	31.0	hyd
2	Laxman	NaN	pune
3	NaN	33.0	chenni

In [40]:
```
mean_age=df['Age'].mean()
df['Age'].fillna(mean_age)
```

Out[40]:
```
0    30.000000
1    31.000000
2    31.333333
3    33.000000
Name: Age, dtype: float64
```

In [41]:
```
median_age=df['Age'].median()
df['Age'].fillna(median_age)
```

Out[41]:
```
0    30.0
1    31.0
2    31.0
3    33.0
Name: Age, dtype: float64
```

In [42]:
```
mode_name=df['Name'].mode()
df['Name'].fillna(mode_name)
```

Out[42]:
```
0       Ram
1      sita
2    Laxman
3       NaN
Name: Name, dtype: object
```

- Mode is used for categorical data
- Mean and median is used for numerical data

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: