url= r"C:\Users\Rutuja Habib\Downloads\Iris.csv" df=pd.read_csv(url) #reading the file print(df)

_ →		Id	Sanall anathCm	SanalWidthCm	PetalLengthCm	DotalWidthCm	١
7,		Iu	Separteliguicii	Зератитиспсш	recarteingtiiciii	retaiwiuthtii	\
	0	1	5.1	3.5	1.4	0.2	
	1	2	4.9	3.0	1.4	0.2	
	2	3	4.7	3.2	1.3	0.2	
	3	4	4.6	3.1	1.5	0.2	
	4	5	5.0	3.6	1.4	0.2	
	145	146	6.7	3.0	5.2	2.3	
	146	147	6.3	2.5	5.0	1.9	
	147	148	6.5	3.0	5.2	2.0	
	148	149	6.2	3.4	5.4	2.3	
	149	150	5.9	3.0	5.1	1.8	

Species . Iris-setosa 0 Iris-setosa Iris-setosa 1 2 Iris-setosa Iris-setosa 3 4 145 Iris-virginica 146 Iris-virginica 147 Iris-virginica 148 Iris-virginica 149 Iris-virginica [150 rows x 6 columns]

df.describe()

_ →		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm
	count	150.000000	150.000000	150.000000	150.000000	150.000000
	mean	75.500000	5.843333	3.054000	3.758667	1.198667
	std	43.445368	0.828066	0.433594	1.764420	0.763161
	min	1.000000	4.300000	2.000000	1.000000	0.100000
	25%	38.250000	5.100000	2.800000	1.600000	0.300000
	50%	75.500000	5.800000	3.000000	4.350000	1.300000
	75%	112.750000	6.400000	3.300000	5.100000	1.800000
	max	150.000000	7.900000	4.400000	6.900000	2.500000

df.info()

</pre RangeIndex: 150 entries, 0 to 149 Data columns (total 6 columns):

Ducu	COTAMITS (COCAT	o coramis).	
#	Column	Non-Null Count	Dtype
0	Id	150 non-null	int64
1	SepalLengthCm	150 non-null	float64
2	SepalWidthCm	150 non-null	float64
3	PetalLengthCm	150 non-null	float64
4	PetalWidthCm	150 non-null	float64
5	Species	150 non-null	object
dtypes: float64(4),		int64(1), object	t(1)
	=	/B	

dtypes: float64(4), ir memory usage: 7.2+ KB

df.shape

→ (150, 6)

df.isnull()

Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species

0FalseFalseFalseFalseFalse1FalseFalseFalseFalseFalse

2FalseFalseFalseFalseFalse3FalseFalseFalseFalseFalse

4 False False False False False

147FalseFalseFalseFalseFalse148FalseFalseFalseFalseFalse

149 False False False False False

150 rows × 6 columns

df.isnull().sum()

Id 0
 SepalLengthCm 0
 SepalWidthCm 0
 PetalLengthCm 0
 PetalWidthCm 0
 Species 0

dtype: int64

Insert Null Values in the dataset

df.loc[5, 'petal.length']=np.nan
print(df)

Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm \ ₹ 5.1 1 3.5 1.4 0.2 1 2 4.9 3.0 1.4 0.2 2 3 4.7 3.2 1.3 0.2 3 4 4.6 3.1 1.5 0.2 4 5 5.0 3.6 1.4 0.2 6.7 5.2 145 146 3.0 2.3 146 147 6.3 2.5 5.0 1.9 147 148 3.0 5.2 2.0 6.5 148 149 6.2 3.4 5.4 2.3

3.0

5.1

1.8

Species petal.length 0 Iris-setosa Iris-setosa NaN 2 Iris-setosa NaN Iris-setosa 4 Iris-setosa NaN 145 Iris-virginica NaN Iris-virginica 146 NaN Iris-virginica NaN 147 148 Iris-virginica NaN 149 Iris-virginica NaN

5.9

[150 rows x 7 columns]

df.isnull().sum()

149 150

₹ Id 0 SepalLengthCm 0 SepalWidthCm 0 PetalLengthCm a PetalWidthCm 0 Species 0 petal.length 150 dtype: int64

df.loc[5, 'PetalLengthCm'] = np.nan
print(df)

Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm \
0 1 5.1 3.5 1.4 0.2

1 2 3	2	4.9	3.0	1.4	0.2
	3	4.7	3.2	1.3	0.2
	4	4.6	3.1	1.5	0.2
4	5	5.0	3.6	1.4	0.2
145	146	6.7	3.0	5.2	2.3
146	147	6.3	2.5	5.0	1.9
147	148	6.5	3.0	5.2	2.0
148	149	6.2	3.4	5.4	2.3
149	150	5.9	3.0	5.1	1.8

	Species	petal.length
0	Iris-setosa	NaN
1	Iris-setosa	NaN
2	Iris-setosa	NaN
3	Iris-setosa	NaN
4	Iris-setosa	NaN
145	Iris-virginica	NaN
146	Iris-virginica	NaN
147	Iris-virginica	NaN
148	Iris-virginica	NaN
149	Iris-virginica	NaN

[150 rows x 7 columns]

df.isnull().sum()

Id 0
SepalLengthCm 0
SepalWidthCm 0
PetalLengthCm 1
PetalWidthCm 0
Species 0
petal.length 150
dtype: int64

Fill NA Values with 0 or 1

df.fillna(1)

_		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	petal.length
	0	1	5.1	3.5	1.4	0.2	Iris-setosa	1.0
	1	2	4.9	3.0	1.4	0.2	Iris-setosa	1.0
	2	3	4.7	3.2	1.3	0.2	Iris-setosa	1.0
	3	4	4.6	3.1	1.5	0.2	Iris-setosa	1.0
	4	5	5.0	3.6	1.4	0.2	Iris-setosa	1.0
	145	146	6.7	3.0	5.2	2.3	Iris-virginica	1.0
	146	147	6.3	2.5	5.0	1.9	Iris-virginica	1.0
	147	148	6.5	3.0	5.2	2.0	Iris-virginica	1.0
	148	149	6.2	3.4	5.4	2.3	Iris-virginica	1.0
	149	150	5.9	3.0	5.1	1.8	Iris-virginica	1.0

150 rows × 7 columns

df.fillna(0)

~								
_		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	petal.length
	0	1	5.1	3.5	1.4	0.2	Iris-setosa	0.0
	1	2	4.9	3.0	1.4	0.2	Iris-setosa	0.0
	2	3	4.7	3.2	1.3	0.2	Iris-setosa	0.0
	3	4	4.6	3.1	1.5	0.2	Iris-setosa	0.0
	4	5	5.0	3.6	1.4	0.2	Iris-setosa	0.0
	145	146	6.7	3.0	5.2	2.3	Iris-virginica	0.0
	146	147	6.3	2.5	5.0	1.9	Iris-virginica	0.0
	147	148	6.5	3.0	5.2	2.0	Iris-virginica	0.0
	148	149	6.2	3.4	5.4	2.3	Iris-virginica	0.0
	149	150	5.9	3.0	5.1	1.8	Iris-virginica	0.0

150 rows × 7 columns

df.fillna(df.mean)

→ *		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	petal.length
	0	1	5.1	3.5	1.4	0.2	Iris-setosa	 d DataFrame.mean of Id Sepa
	1	2	4.9	3.0	1.4	0.2	Iris-setosa	 d DataFrame.mean of Id Sepa
	2	3	4.7	3.2	1.3	0.2	Iris-setosa	 bound method DataFrame.mean of Id Sepa
	3	4	4.6	3.1	1.5	0.2	Iris-setosa	 bound method DataFrame.mean of Id Sepa
	4	5	5.0	3.6	1.4	0.2	Iris-setosa	 bound method DataFrame.mean of Id Sepa
	145	146	6.7	3.0	5.2	2.3	Iris-virginica	 bound method DataFrame.mean of Id Sepa
	146	147	6.3	2.5	5.0	1.9	Iris-virginica	 d DataFrame.mean of Id Sepa
	147	148	6.5	3.0	5.2	2.0	Iris-virginica	 bound method DataFrame.mean of Id Sepa
	148	149	6.2	3.4	5.4	2.3	Iris-virginica	 d DataFrame.mean of Id Sepa
	149	150	5.9	3.0	5.1	1.8	Iris-virginica	 dound method DataFrame.mean of Id Sepa

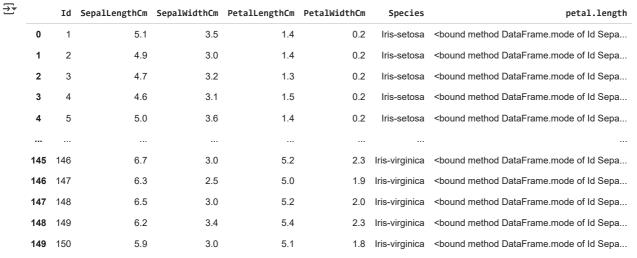
150 rows × 7 columns

df.fillna(df.median)

	•		,					
₹		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	Species	petal.length
	0	1	5.1	3.5	1.4	0.2	Iris-setosa	 dound method DataFrame.median of Id Se
	1	2	4.9	3.0	1.4	0.2	Iris-setosa	 dound method DataFrame.median of Id Se
	2	3	4.7	3.2	1.3	0.2	Iris-setosa	 dound method DataFrame.median of Id Se
	3	4	4.6	3.1	1.5	0.2	Iris-setosa	 dound method DataFrame.median of Id Se
	4	5	5.0	3.6	1.4	0.2	Iris-setosa	 dound method DataFrame.median of Id Se
	145	146	6.7	3.0	5.2	2.3	Iris-virginica	 dound method DataFrame.median of Id Se
	146	147	6.3	2.5	5.0	1.9	Iris-virginica	 dound method DataFrame.median of Id Se
	147	148	6.5	3.0	5.2	2.0	Iris-virginica	 dound method DataFrame.median of Id Se
	148	149	6.2	3.4	5.4	2.3	Iris-virginica	 dound method DataFrame.median of Id Se
	149	150	5.9	3.0	5.1	1.8	Iris-virginica	 dound method DataFrame.median of Id Se

150 rows × 7 columns

df.fillna(df.mode)



150 rows × 7 columns

Forward Fill

df.ffill

→ ▼	< hou	nd method NDFra	me ffill of	Id	SenalLength(m	SenalWidthCm	PetalLengthCm	PetalWidthCm	\
نک	0	1	5.1	3.5	1.4	0.2	r c carrengenem	i c caiwia ciiciii	`
	1	2	4.9	3.0	1.4	0.2			
	2	3	4.7	3.2	1.3	0.2			
	3	4	4.6	3.1	1.5	0.2			
	4	5	5.0	3.6	1.4	0.2			
			•••			•••			
	 145	146	6.7	3.0	5.2	2.3			
	146	147	6.3	2.5	5.0	1.9			
	147	148	6.5	3.0	5.2	2.0			
	148	149	6.2	3.4	5.4	2.3			
	149	150	5.9	3.0	5.1	1.8			
	149	150	3.9	5.0	3.1	1.0			
		Species	petal.length						
	0	Iris-setosa	. NaN						
	1	Iris-setosa	NaN						
	2	Iris-setosa	NaN						
	3	Iris-setosa	NaN						
	4	Iris-setosa	NaN						
	145	Iris-virginica	NaN						
	146	Iris-virginica							
	147	Iris-virginica							
	148	Iris-virginica							
	149	Iris-virginica							
		0							
	Γ150	rows x 7 colum	nsl>						
			- 3						

Backward Fill

149 Iris-virginica

df.bfill

_									
		nd method NDFra		Id			PetalLengthCm	PetalWidthCm	١
	0	1	5.1	3.5	1.4	0.2			
	1	2	4.9	3.0	1.4	0.2			
	2	3	4.7	3.2	1.3	0.2			
	3	4	4.6	3.1	1.5	0.2			
	4	5	5.0	3.6	1.4	0.2			
		• • •	• • •	• • •	• • •				
	145	146	6.7	3.0	5.2	2.3			
	146	147	6.3	2.5	5.0	1.9			
	147	148	6.5	3.0	5.2	2.0			
	148	149	6.2	3.4	5.4	2.3			
	149	150	5.9	3.0	5.1	1.8			
		Species	petal.length						
	0	Iris-setosa							
	1	Iris-setosa	NaN						
	2	Iris-setosa	NaN						
	3	Iris-setosa	NaN						
	4	Iris-setosa	NaN						
	145	Iris-virginica	NaN						
	146	Iris-virginica	NaN						
	147	Iris-virginica	NaN						
	148	Iris-virginica	NaN						

NaN

\

```
[150 rows x 7 columns]>
```

Drop na values

```
df.dropna()
```

Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm Species petal.length

df.loc[148, 'SepalLengthCm']=np.nan
print(df)

_							
Đ		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	,
	0	1	5.1	3.5	1.4	0.2	
	1	2	4.9	3.0	1.4	0.2	
	2	3	4.7	3.2	1.3	0.2	
	3	4	4.6	3.1	1.5	0.2	
	4	5	5.0	3.6	1.4	0.2	
	145	146	6.7	3.0	5.2	2.3	
	146	147	6.3	2.5	5.0	1.9	
	147	148	6.5	3.0	5.2	2.0	
	148	149	NaN	3.4	5.4	2.3	
	149	150	5.9	3.0	5.1	1.8	

Species petal.length Species po Iris-setosa Iris-setosa Iris-setosa Iris-setosa Iris-setosa 0 NaN 1 2 NaN 3 NaN 4 NaN 145 Iris-virginica NaN 146 Iris-virginica 147 Iris-virginica NaN 148 Iris-virginica NaN 149 Iris-virginica NaN

[150 rows x 7 columns]

df.loc[145, 'SepalLengthCm'] = np.nan
print(df)

∑ ₹		Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
_	0	1	5.1	3.5	1.4	0.2	
	1	2	4.9	3.0	1.4	0.2	
	2	3	4.7	3.2	1.3	0.2	
	3	4	4.6	3.1	1.5	0.2	
	4	5	5.0	3.6	1.4	0.2	
	145	146	NaN	3.0	5.2	2.3	
	146	147	6.3	2.5	5.0	1.9	
	147	148	6.5	3.0	5.2	2.0	
	148	149	NaN	3.4	5.4	2.3	
	149	150	5.9	3.0	5.1	1.8	

	Species	petal.length
0	Iris-setosa	NaN
1	Iris-setosa	NaN
2	Iris-setosa	NaN
3	Iris-setosa	NaN
4	Iris-setosa	NaN
145	Iris-virginica	NaN
146	Iris-virginica	NaN
147	Iris-virginica	NaN
148	Iris-virginica	NaN
149	Iris-virginica	NaN

[150 rows x 7 columns]

$\mathsf{df.ffill}$

⋺₹	<bou< th=""><th>nd method</th><th>NDFrame.ffill of</th><th>Id</th><th>SepalLengthCm</th><th>SepalWidthCm</th><th>PetalLengthCm</th><th>PetalWidthCm</th><th>\</th></bou<>	nd method	NDFrame.ffill of	Id	SepalLengthCm	SepalWidthCm	PetalLengthCm	PetalWidthCm	\
	0	1	5.1	3.5	1.4	0.2			
	1	2	4.9	3.0	1.4	0.2			
	2	3	4.7	3.2	1.3	0.2			
	3	4	4.6	3.1	1.5	0.2			
	4	5	5.0	3.6	1.4	0.2			
			• • •	• • •	• • •				
	145	146	NaN	3.0	5.2	2.3			
	146	147	6.3	2.5	5.0	1.9			
	147	148	6.5	3.0	5.2	2.0			

```
148 149
                        NaN
                                     3.4
                                                   5.4
                                                                 2.3
     149 150
                       5.9
                                     3.0
                                                   5.1
                                                                 1.8
                Species petal.length
     0
            Iris-setosa
            Iris-setosa
     1
     2
                                 NaN
            Iris-setosa
            Iris-setosa
                                 NaN
     3
     4
            Iris-setosa
                                 NaN
    145 Iris-virginica
                                 NaN
     146 Iris-virginica
                                 NaN
     147 Iris-virginica
                                 NaN
     148
        Iris-virginica
                                 NaN
     149 Iris-virginica
                                 NaN
     [150 rows x 7 columns]>
df.bfill
<bound method NDFrame.bfill of</pre>
                                      Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm \
     9
                       5.1
                                     3.5
           1
                                                   1.4
                                                                 0.2
                       4.9
                                     3.0
                                                   1.4
                                                                 0.2
     1
           2
     2
           3
                       4.7
                                     3.2
                                                   1.3
                                                                 0.2
     3
           4
                       4.6
                                     3.1
                                                   1.5
                                                                 0.2
     4
           5
                       5.0
                                     3.6
                                                   1.4
                                                                 0.2
     145 146
                       NaN
                                     3.0
                                                   5.2
                                                                 2.3
     146
         147
                       6.3
                                     2.5
                                                   5.0
                                                                 1.9
     147 148
                       6.5
                                     3.0
                                                   5.2
                                                                 2.0
     148 149
                       NaN
                                     3.4
                                                   5.4
                                                                 2.3
     149 150
                       5.9
                                     3.0
                                                   5.1
                                                                 1.8
               Species petal.length
    0
            Iris-setosa
                                 NaN
     1
            Iris-setosa
                                 NaN
     2
            Iris-setosa
                                 NaN
     3
            Iris-setosa
                                 NaN
           Iris-setosa
                                 NaN
     145 Iris-virginica
                                 NaN
     146 Iris-virginica
                                 NaN
     147 Iris-virginica
                                 NaN
     148 Iris-virginica
                                 NaN
    149 Iris-virginica
                                 NaN
     [150 rows x 7 columns]>
Mapping
mapping={'Iris-setosa':110,'Iris-virginica':111,'versicolor':112}
df['v_map']=df['Species'].map(mapping)
print(df)
∓
          Id SepalLengthCm SepalWidthCm PetalLengthCm PetalWidthCm \
                               3.5
                       5.1
                                                   1.4
                                                                 0.2
                       4.9
                                     3.0
                                                   1.4
           3
                       4.7
                                     3.2
                                                   1.3
                                                                 0.2
                        4.6
                                     3.1
                                                                 0.2
                                                   1.5
     4
           5
                       5.0
                                     3.6
                                                                 0.2
                                                   1.4
    145 146
                                     3.0
                                                   5.2
                                                                 2.3
                       NaN
     146
        147
                       6.3
                                     2.5
                                                   5.0
                                                                 1.9
     147
         148
                        6.5
                                     3.0
                                                   5.2
                                                                 2.0
     148
        149
                       NaN
                                     3.4
                                                   5.4
                                                                 2.3
     149 150
                                     3.0
                                                   5.1
```

Species petal.length v_map

Iris-setosa

NaN 110.0