Handling Missing Values



Dropna | Fillna | Handling Missing Values in Pandas

Drop Rows /columns temporarily or permanently.

```
- F 7
           Marks Grades
     Name
 Priyang
            98.0
                     NaN
  Aadhya
             NaN
                      AB
  Krisha
            99.0
                      AA
  Vedant
            87.0
                     NaN
  Parshv
            90.0
                      AC
  Mittal
             NaN
                      BA
 Archana
            82.0
                      BB
   NULL
             NaN
                     NaN
```

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(inplace=False)
print(sr)
```

| | Name | Marks | Grades |
|---|---------|-------|--------|
| 2 | Krisha | 99.0 | AA |
| 4 | Parshv | 90.0 | AC |
| 6 | Archana | 82.0 | ВВ |

Drop the Rows where atleast one value is missing.

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(axis=0,how="any")
print(sr)
```

| | Name | Marks | Grades |
|---|---------|-------|--------|
| 2 | Krisha | 99.0 | AA |
| 4 | Parshv | 90.0 | AC |
| 6 | Archana | 82.0 | BB |

Drop the Column where atleast one value is missing.

```
Name Marks Grades
Krisha 99.0 AA
Parshv 90.0 AC
Archana 82.0 BB
```

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(axis=1,how="any")
print(sr)
```

Drop Rows where All Values are missing.

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(axis=0,how="all")
print(sr)
```

| | Name | riai KS | braues |
|---|---------|---------|--------|
| 0 | Priyang | 98.0 | NaN |
| 1 | Aadhya | NaN | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | NaN |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | NaN | BA |
| 6 | Archana | 82.0 | BB |
| 7 | NULL | NaN | NaN |

Keep only the Rows with Atleast 2 Non NA values.

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(axis=0,thresh=2)
print(sr)
```

| | Name | marks | Grades |
|---|---------|-------|--------|
| 0 | Priyang | 98.0 | NaN |
| 1 | Aadhya | NaN | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | NaN |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | NaN | BA |
| 6 | Archana | 82.0 | BB |

Define in which Column to look for missing NA values

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.dropna(subset=["Marks","Name"])
print(sr)
```

| Grades | Marks | Name | |
|--------|-------|---------|---|
| NaN | 98.0 | Priyang | 0 |
| AA | 99.0 | Krisha | 2 |
| NaN | 87.0 | Vedant | 3 |
| AC | 90.0 | Parshv | 4 |
| BB | 82.0 | Archana | 6 |

Fill all NAN Values with 0.

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(0)
print(sr)
```

| | Name | Marks | Grades |
|---|---------|-------|--------|
| 0 | Priyang | 98.0 | Θ |
| 1 | Aadhya | 0.0 | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | 0 |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | 0.0 | BA |
| 6 | Archana | 82.0 | BB |
| 7 | NULL | 0.0 | 0 |

Fill all NAN Value elts in cols "Marks" and "Grades" with 97 and "AA" respectively.

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna({"Marks":97, "Grades":"AA"})
print(sr)
```

| | Name | Marks | Grades | |
|---|---------|-------|--------|--|
| 0 | Priyang | 98.0 | AA | |
| 1 | Aadhya | 97.0 | AB | |
| 2 | Krisha | 99.0 | AA | |
| 3 | Vedant | 87.0 | AA | |
| 4 | Parshv | 90.0 | AC | |
| 5 | Mittal | 97.0 | BA | |
| 6 | Archana | 82.0 | ВВ | |
| 7 | NULL | 97.0 | AA | |

Propagate non-null values forward and backward

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="ffill")
print(sr)
```

```
Marks Grades
          98.0
Priyang
                  NaN
 Aadhya
          98.0
                   AB
Krisha
          99.0
                   AA
          87.0
 Vedant
                   AA
          90.0
                   AC
Parshv
Mittal
          90.0
                   BA
Archana
          82.0
                   BB
  NULL
          82.0
                   BB
```

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="bfill")
print(sr)
```

| | Name | Marks | Grades |
|---|---------|-------|--------|
| 0 | Priyang | 98.0 | AB |
| 1 | Aadhya | 99.0 | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | AC |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | 82.0 | BA |
| 6 | Archana | 82.0 | BB |
| 7 | NULL | NaN | NaN |

Forward Filling

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="ffill",axis=0)
print(sr)
```

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="ffill",axis=1)
print(sr)
```

| | Name | Marks | Grades |
|---|---------|--------|--------|
| 0 | Priyang | 98.0 | NaN |
| 1 | Aadhya | 98.0 | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | AA |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | 90.0 | BA |
| 6 | Archana | 82.0 | ВВ |
| 7 | NULL | 82.0 | ВВ |
| | | | |
| | Name | Marks | Grades |
| 9 | Priyang | 98.0 | 98.0 |
| 1 | Aadhya | Aadhya | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | 87.0 |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | Mittal | BA |
| 5 | Archana | 82.0 | ВВ |
| 7 | NULL | NULL | NULL |

Backward Filling

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv
print(data)
sr=data.fillna(method="bfill",axis=0)
print(sr)
```

```
Priyang
         98.0
                  AB
 Aadhya
         99.0
                  AΒ
 Krisha
         99.0
                  AA
 Vedant
         87.0
                  АC
 Parshv
         90.0
                  АC
Mittal
         82.0
                  ΒA
         82.0
                  BB
Archana
  NULL
           NaN
                 NaN
```

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="bfill",axis=1)
print(sr)
```

| 0 | Priyang | 98.0 | NaN |
|---|---------|------|-----|
| 1 | Aadhya | AB | AB |
| 2 | Krisha | 99.0 | AA |
| 3 | Vedant | 87.0 | NaN |
| 4 | Parshv | 90.0 | AC |
| 5 | Mittal | BA | BA |
| 6 | Archana | 82.0 | BB |
| 7 | NULL | NaN | NaN |

Fill the first NaN elements

```
import pandas as pd
data=pd.read_csv("C:\\Users\\Dell\\Desktop\\minipro.csv")
print(data)
sr=data.fillna(method="ffill",limit=1)
print(sr)
         Priyang
                    98.0
                             NaN
          Aadhya
                    98.0
                              AB
          Krisha
                    99.0
                              AA
          Vedant
                    87.0
                              AA
          Parshv
                    90.0
                              AC
          Mittal
                    90.0
                              BΑ
                              BB
         Archana
                    82.0
           NULL
                              BB
                    82.0
```