In [ ]:

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

In [ ]:

```
df=pd.DataFrame(data={
    'feature_1':[np.nan,3,6,9,12,15,np.nan],
    'feature_2':[100,np.nan,200,300,np.nan,np.nan,600],
    'feature_3':[1000,500,2000,3000,4000,6000,8000,],
})
df
```

In [ ]:

```
df.isnull()
```

In [ ]:

```
df.isnull().sum()
```

In [ ]:

```
df.fillna(method='pad',limit=1)
```

In [ ]:

```
df.fillna(method='pad',limit=2)
```

In [ ]:

```
df.fillna(method='bfill')
```

In [ ]:

```
df.fillna(method='ffill')
```

In [ ]:

```
df.dropna(axis=0)  # drops all records with any missing value
```

In [ ]:

```
df.dropna(axis=1)  # drops all features with missing values and keep only features whic
h are non-missing
```

In [ ]:

```
df.dropna(thresh=int(df.shape[0] * .9), axis=1) # give threshold that >10% missing data
be dropped
```

In [ ]:

```
df.shape[0] * .9
```

In [ ]:

```
df['feature_1'].fillna(df['feature_1'].mean())
```

In [ ]:

```
df['feature_2'].interpolate()
```

In [ ]:

```

```