

```

import pandas as pd
df = pd.read_csv('Titanic-Dataset.csv')
df.head()

df.describe()

df.isnull().sum()

len(df)

missing_percentage = df.isnull().mean() * 100
columns_with_missing_30 = missing_percentage[missing_percentage > 30]
print("Columns with more than 30% missing values:")
print(columns_with_missing_30)

missing_percentage = df.isnull().mean() * 100
threshold = 30

columns_to_drop = missing_percentage[missing_percentage > threshold].index

df_cleaned = df.drop(columns=columns_to_drop)

print(f"Columns dropped: {list(columns_to_drop)}")
print(f"Remaining columns: {df_cleaned.columns}")

data = pd.DataFrame({
    'A': [1, 2, None, 4],
    'B': [5, None, 7, 8],
    'C': ['cat', None, 'dog', 'cat']
})

data['A'].fillna(data['A'].mean(), inplace=True)

data['B'].fillna(data['B'].median(), inplace=True)

data['C'].fillna(data['C'].mode()[0], inplace=True)
print(data)

data = pd.DataFrame({
    'A': [1, None, 3, None, 5],

```

```
'B': [None, 6, None, 8, None]
})
data_forward_fill = data.fillna(method='ffill')
data_backward_fill = data.fillna(method='bfill')
print(data_forward_fill)
print(data_backward_fill)


from sklearn.impute import KNNImputer

data = pd.DataFrame({
    'A': [1, 2, None, 4],
    'B': [5, None, 7, 8],
    'C': [9, 10, 11, None]
})

imputer = KNNImputer(n_neighbors=2)
data_imputed_knn = pd.DataFrame(imputer.fit_transform(data),
                                columns=data.columns)
print(data_imputed_knn)
```