

# SimpleImputer Cheatsheet

## 1. Library Import Karna

python

 Copy code

```
from sklearn.impute import SimpleImputer
```

## 2. Imputer Create Karna

- **Missing Values (NaNs) Ke Liye:**

python

 Copy code

```
imputer = SimpleImputer(missing_values=np.nan, strategy='mean')
```

- **missing\_values=np.nan:** Kya replace karna hai (default hai NaN).
- **strategy='mean':** Kaise replace karna hai (options: 'mean', 'median', 'most\_frequent', 'constant').

## 3. Imputer Ko Fit Karna

python

 Copy code

```
imputer.fit(data)
```

- **data:** Aapka input data jahan NaNs hain (numpy array ya pandas DataFrame).

## 4. Data Ko Transform Karna

python

 Copy code

```
data_transformed = imputer.transform(data)
```

- **data\_transformed:** NaNs ko replace kiya gaya hai.

## 5. Fit Aur Transform Ko Ek Saath Karna

python

 Copy code

```
data_transformed = imputer.fit_transform(data)
```

- **fit\_transform:** Ek hi step mein fit aur transform dono karta hai.

## 6. Strategies Samjhaana

- **Mean:** NaNs ko column ka mean value se replace karta hai.

python

 Copy code

```
imputer = SimpleImputer(strategy='mean')
```

- **Median:** NaNs ko column ka median value se replace karta hai.

python

 Copy code

```
imputer = SimpleImputer(strategy='median')
```

- **Most Frequent:** NaNs ko column ke sabse frequent value se replace karta hai.

python

 Copy code

```
imputer = SimpleImputer(strategy='most_frequent')
```

- **Constant:** NaNs ko ek constant value se replace karta hai.

python

 Copy code

```
imputer = SimpleImputer(strategy='constant', fill_value=0)
```

## 7. ColumnTransformer Ke Saath Use Karna

python

 Copy code

```
from sklearn.compose import ColumnTransformer trf = ColumnTransformer([ ('imputer1',
SimpleImputer(strategy='mean'), ['Age']), ('imputer2',
SimpleImputer(strategy='median'), ['Fare']) ], remainder='passthrough')
```

- **ColumnTransformer:** Different columns pe alag-alag imputation strategy apply kar sakte hain.

## 8. DataFrames Handle Karna

- **Pandas DataFrame Example:**

python

 Copy code

```
import pandas as pd data = pd.DataFrame({ 'Age': [25, 30, np.nan, 35], 'Fare':
[100, np.nan, 50, 70] }) imputer = SimpleImputer(strategy='mean') data[['Age',
'Fare']] = imputer.fit_transform(data[['Age', 'Fare']])
```

## Notes

- **Yaad Rakhein:** Hamesha imputer ko training data pe fit karein aur transform training aur testing data pe dono pe karein.
- **Scikit-learn Documentation:** Aur details ke liye, official documentation yahan check karein.

Yeh cheatsheet SimpleImputer ko use karne ke basics cover karta hai. Happy coding!