Summary Statistics:

Generate summary statistics of the dataset to identify missing values. By using methods like .describe(), .info(), or .isnull().sum() in pandas, you can quickly see which columns have missing values and how many are missing.

Example:

python

Copy code

import pandas as pd

df = pd.read\_csv('data.csv')

print(df.info())

print(df.isnull().sum())

Visual Inspection:

Use visualizations to identify missing values.

Visualization libraries like matplotlib or seaborn can create heatmaps to visualize missing values.

Example:

python

Copy code

import seaborn as sns

import matplotlib.pyplot as plt

sns.heatmap(df.isnull(), cbar=False)

plt.show()

Boolean Indexing:

Use boolean indexing to filter and identify rows with missing values.

By creating a boolean mask, you can filter out rows or columns with missing values.

Example:

python

Copy code

missing\_data = df[df.isnull().any(axis=1)]

print(missing\_data)

Missing Value Indicators:

Description: Identify missing values using placeholders or specific indicators.

How it Works: Some datasets use placeholders like 'NaN', 'None', or specific strings to indicate missing values. These can be identified and counted.

Example:

python

Copy code

placeholder\_count = df[df == 'None'].count()

print(placeholder\_count)

Data Profiling Tools:

Description: Use data profiling tools and libraries that automatically detect missing values.

How it Works: Tools like pandas-profiling or missingno provide comprehensive reports and visualizations for missing values in the dataset.

Example:

python

Copy code

import pandas\_profiling

profile = pandas\_profiling.ProfileReport(df)

profile.to\_file("output.html")