Firstly preprocess the data

Step1:-go to kaggle and find that data set that have null values

Step2: df=pd.read\_csv("/kaggle/input/data-science-job-salaries/ds\_salaries.csv")

Step3:df

Step4:df.info()

import numpy as np

import pandas as pd

df=pd.read\_csv('D:\\Study\\Du\_Mca\\3rd\_semester\\Data\_Science\_with\_python\\New folder\\ds\_salaries.csv')

print(df.head())

df.info()

df.head(5)

df.tail(5)

df.describe()

df.shape

df.dtypes

df.index

df.count

df["salary"].unique()

df.duplicated