# 

#Create an empty DataFrame

# importing the pandas library

import pandas as pd

df = pd.DataFrame()

print (df)

# Create a DataFrame using List:

# importing the pandas library

import pandas as pd

# a list of strings

x = ['Python', 'Pandas']

# Calling DataFrame constructor on list

df = pd.DataFrame(x)

print(df)

#Create a DataFrame from Dict of ndarrays/ Lists

# importing the pandas library

import pandas as pd

info = {'ID' :[101, 102, 103],'Department':['B.Sc','B.Tech','M.Tech',]}

df = pd.DataFrame(info)

print (df)

# Create a DataFrame from Dict of Series:

# importing the pandas library

import pandas as pd

info = {'one' : pd.Series([1, 2, 3, 4, 5, 6], index=['a', 'b', 'c','d', 'e', 'f']),

'two' : pd.Series([1, 2, 3, 4, 5, 6, 7, 8], index=['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h'])}

d1 = pd.DataFrame(info)

print (d1)

# Column Selection

# importing the pandas library

import pandas as pd

info = {'one' : pd.Series([1, 2, 3, 4, 5, 6], index=['a', 'b', 'c', 'd', 'e', 'f']),

'two' : pd.Series([1, 2, 3, 4, 5, 6, 7, 8], index=['a', 'b', 'c', 'd', 'e', 'f', 'g', 'h'])}

d1 = pd.DataFrame(info)

print (d1 ['one'])

# Data Preprocessing

# Importing the libraries

import numpy as np

import matplotlib.pyplot as plt

import pandas as pd

# Load data from Local drive

from google.colab import files

uploaded = files.upload()

# Importing the dataset

dataset = pd.read\_csv("Data.csv")

print(dataset)

dataset

# Mounting Google Drive

from google.colab import drive

drive.mount('/content/drive')

# Importing the dataset

dataset = pd.read\_csv('/content/drive/My Drive/Data.csv')

data="https://raw.githubusercontent.com/deepakmoud/datasets-for-experiment/main/Data.csv"

# Importing the dataset

dataset = pd.read\_csv(data)

print(dataset)

dataset

#Print Total number of Rows & columns in dataset

print(dataset.shape)

#Print Information about data

dataset.info()

#Print Information Datatypes

datatypes = dataset.dtypes

print(datatypes)

#Count total number of classes in Data

class\_counts = dataset.groupby('Purchased').size()

print(class\_counts)

#Count total number of classes in Data

class\_counts = dataset.groupby('Country').size()

print(class\_counts)

dataset.isnull().sum()

dataset.notnull()

#1: Dropping rows with at least 1 null value.

# using dropna() function

df=dataset.dropna()

print(df)

# Dropping columns with at least 1 null value.

df= dataset.dropna(axis = 0, how ='any')

print(df)

# Dropping columns with at least 1 null value.

df= dataset.dropna(axis = 1, how ='any')

print(df)

# Filling null values with a single value

# filling missing value using fillna()

df=dataset.fillna(0)

print(df)

# Filling null values with the previous ones

# filling missing value using fillna()

print(dataset)

print("--------------after padding---------------------------")

df= dataset.fillna(method ='pad')

print(df)

# Code #3: Filling null value with the next ones

print(dataset)

print("----------------------after backward fill---------------------")

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

print(df)

df['Salary'].fillna(df['Salary'].median(), inplace=True)

print(df)

df['Age'].fillna(df['Age'].median(), inplace=True)

print(df)