

4. Implement a python program to demonstrate

1) Importing Datasets

2) Cleaning the Data

3) Data frame manipulation using Numpy

```
#lab4
```

```
import pandas as pd
```

```
import numpy as np
```

```
#import dataset
```

```
df = pd.read_csv("C:\\Users\\lenovo\\Desktop\\BL-Flickr-  
Images-Book.csv", encoding = "ISO-8859-1")
```

```
df
```

```
##Check the output and writedown
```

```
#cleaning data
```

```
to_drop = ['Edition Statement',  
           'Corporate Author',  
           'Corporate Contributors',  
           'Former owner',  
           'Engraver',  
           'Contributors',  
           'Issuance type',  
           'Shelfmarks']
```

```
#cleaning data
```

```
df.drop(to_drop, inplace=True, axis=1)
```

```
df.head()
```

```
#Check out the output
```

```
#2 loc operation
```

```
df.loc[1905:, 'Date of Publication'].head(10)
```

```
#check out the output
```

```
#cleaning remove null values
```

```
df = df[df['Date of Publication'].notnull()]
```

```
df.head()
```

```
#check out the output
```

```
#datarame manipulation
```

```
#get all publications from London
```

```
ndf = np.array(df)
```

```
ndf = ndf[ndf[:,1] == 'London']
```

```
ndf[0:10]
```

```
#convert np to df
```

```
idf = pd.DataFrame(ndf)
idf.head()
```

```
#lab5
```

```
#importing numpy
```

```
import numpy as np
```

```
#creating an array a
```

```
a = np.array( [[ 1, 2, 3, 4],
[ 5, 6, 7,8],
[9,10,11,12]] )
```

```
#printing array a
```

```
print ("Array is:\n",a)
```

```
#we can also print the other attributes like dimensions,shape and
size of an array
```

```
print ("Dimensions of a are:", a.ndim)
```

```
print ("Shape of a is", a.shape)
```

```
print ("Size of a is", a.size)
```

```
# array manipulation
```

```
#creating an array a zeros square array of dimensions 2X2
```

```
zeros_array = np.zeros([2,2], dtype = int)
```

```
print ("Array zeros is:", zeros_array)
```

```
ones_array = np.ones([2,2], dtype = int)
```

```
print ("Array ones is :", ones_array)
```

```
#copying content from ones_array to zeros
```

Prog5

#5 implement the python program demonstrate the following using num f
ile

#a)array manipulation,searching,sorting and splitting

```
import numpy as npy
print("Array Manipulation:")
print("accessing using index")
arr=npy.array(npy.arange(4,24))
for i in range(20):
    print(arr[i])
print("Array Slicing:")
print("All elements:",arr[0::])
print("Index from other end:",arr[::-1])
print("Alternate elemets starting from index 7:",arr[7::2])
print("Reshaping array:")
arr1=arr.reshape(4,5)
print(arr1)

print("Vertical stacking:")
a = npy.array([[12,4,5],[22,35,58],[45,33,21]])
b = npy.array([[1,30,55],[2,5,8],[5,3,1]])
print(npy.vstack((a,b)))

print("Searching in an array:")
x = npy.where(a == 35)
print("position of x=",x)

print("Sorting Array a:",npy.sort(a))

print("Splitting Array a:")
a1,a2,a3,a4=npy.split(arr,[4,10,16])
```

```
print(a1,a2,a3,a4)
```

```
print("Splitting along vertical axis into 3 parts:")  
a5=npy.vsplit(a,3)  
print(a5)
```

```
#b)broadcasting
```

```
import numpy as npy  
arr1=npy.array([2,3,4,6,5,9]).reshape(2,3)  
print("arr1\n",arr1)
```

```
arr2=npy.array([100,200,300]).reshape(1,3)  
print("arr2\n",arr2)
```

```
print("sum\n")  
print(arr1+arr2)
```

```
M = npy.ones((2,3))  
a = npy.arange(3)  
print("M\n",M)  
print("a\n",a)  
print("M+a\n",M+a)
```

```
a = npy.arange(3).reshape((3,1))  
b = npy.arange(3)  
print("a is:\n",a)  
print("b is:\n",b)  
print("a+b is:\n",a+b)
```

```
M = npy.ones((3, 2))  
a = npy.arange(3)
```

```
#plotting numpy arrays
```

```
import numpy as np
import matplotlib.pyplot as plt
```

```
#Computer x and y coordinates for points on sine and cosine curves
x=np.arange(0,3*np.pi, 0.1)
print("x=",x)
```

```
y_sine = np.sin(x)
y_cosine = np.cos(x)
```

```
#plot the points using matplotlib
plt.plot(x, y_sine)
plt.plot(x, y_cosine)
plt.xlabel('x values')
plt.ylabel('y values')
plt.title('plot of sine and cosine functions')
plt.legend(['sine', 'cosine'])
```

```
plt.show()
```

Prog6

#6. Implement a python program to demonstrate Data visualization
#with various Types of Graphs using matplotlib

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
cars_data = pd.read_csv('Toyota.csv', index_col=0, na_values=["??", "????"])

#removing missing values from the dataframe
cars_data.dropna(axis=0, inplace=True)
#scatter plot
plt.scatter(cars_data['Age'], cars_data['Price'], c='red')
plt.title('Scatter plot of Price vs Age of the cars')
plt.xlabel('Age(months)')
plt.ylabel('Price(Euros)')
plt.show()
#scatter plot
plt.scatter(cars_data['Age'], cars_data['Price'], c='red')
plt.title('Scatter plot of Price vs Age of the cars')
plt.xlabel('Age(months)')
plt.ylabel('Price(Euros)')
plt.show()
plt.hist(cars_data['KM'], color='green', edgecolor='white', bins=5)
plt.title('Histogram of Kilometer')
plt.xlabel('Kilometer')
plt.ylabel('Frequency')
plt.show()
#Bar Plot (for categorical variables)
import numpy as np
from matplotlib import pyplot as plt
counts=[979, 120, 12]
fuelType = ('Petrol', 'Diesel', 'CNG')
index = np.arange(len(fuelType))
```



```
plt.bar(index, counts,color=['red','blue', 'cyan'])  
plt.title('Bar plot of fuel types')  
plt.xlabel('Fuel Types')  
plt.ylabel('frequency')  
plt.xticks(index, fuelType, rotation = 45)  
plt.show()
```