2247153

MCA171 – Python Programming

Practice Lab: Numpy / Pandas &amp; Matplotlib

It is mandatory to Use IDLE to complete this exercise. Upload .py file and output file.

1. Import the given “basketball.csv” file as dataframe using pandas and perform the

following:

a) Find the number of values present in the dataset

b) Display all the columns in the dataset

c) How many different schools have participated?

d) What is the average winning percentage?

e) Which school in which year has maximum winning percentage?

f) Use appropriate graph for showing the year\_wise opponent\_points of a school.

g) Use histogram for presenting total number of participation of every school in

games, its corresponding wins and losses.

h) Are there any null values present in the dataset? If yes, how will you deal with

it? Explain with appropriate code.

import pandas as pd #importing pandas as pd

import numpy as np

db=pd.read\_csv(r'C:\Users\nandi\Downloads\basketball.csv') #importing the csv file

print("Total values in the dataframe: \n",db.count()) #count() to find the total of values

print("Columns in the dataset: \n",db.columns)

print("Schools that participated are:\n ",db.school.unique()) #it prints the names of all unique values whereas nunique prints the total count

dm=db["win\_percentage"].mean()#it will give the average

print("Mean of win percentage is: \n",dm)

dmax=db["school"].max()#max() function will give the name of the school with maximum value

print("Maximum is: \n",dmax)

import matplotlib.pyplot as plt

gr=db.plot.bar[x="year”,y=”opponent\_points”]



