# >>> Working with excel files / same is for csv files

Excel experiment

import pandas as pd

import numpy as np

def main\_function():

# Loading excel data

userdata = pd.read\_excel("profile.xlsx")

# Shape of the data file . it is rows x column

print(userdata.shape)

# Display column names

print(userdata.columns)

# Display length of column names

print(len(userdata.columns))

# Display nth row of the dataframe

nthrow = userdata.iloc[0]

print(nthrow)

userdata = userdata.sort\_values('Year')

# get list of values from specific column

year = userdata["Year"].values

print(year)

citations = userdata["Cites"].values

# >>> Ploting charts

import matplotlib.pyplot as plt

import numpy as np

def addlabels(x, y):

for i in range(len(x)):

plt.text(i, y[i], y[i], ha='center')

xpoints = np.array(year)

ypoints = np.array(citations)

plt.grid(color = 'green', linestyle = '--', linewidth = 0.5)

plt.bar(xpoints, ypoints)

addlabels(xpoints, ypoints)

plt.title("Citations per year")

# giving X and Y labels

plt.xlabel("Year")

plt.ylabel("Citations")

plt.show()

papers=[str('['+str(x)+']') for x in range(len(citations))]

xpoints = np.array(papers)

ypoints = np.array(citations)

plt.grid(color = 'green', linestyle = '--', linewidth = 0.5)

plt.bar(xpoints, ypoints)

addlabels(xpoints, ypoints)

plt.title("Citations per publication")

# giving X and Y labels

plt.xlabel("Year")

plt.ylabel("Citations")

plt.show()

if \_\_name\_\_ == '\_\_main\_\_':

main\_function()