import numpy as np

**Import the python libraries**

import pandas as pd

import matplotlib.pyplot as plt

**Read the excel file in to dataframe**

class AspProject:

Country = pd.read\_excel('IMVA.xls', sheet\_name='IMVA')

df\_Countries = Country[['Periods', "Americas", "Canada", "USA", "Other Markets In Americas", "Oceania", "Australia", "New Zealand", "Other Markets In Oceania", "Africa", "Egypt", "Mauritius", "South Africa (Rep Of)", "Other Markets In Africa"]]

**Print specific country**

**A list that will display the specific country**

print(df\_Countries.columns)

new = df\_Countries['Periods'].str.split(' ', n=1, expand=True)

df\_Countries = df\_Countries.assign(Year=new[0])

**This splits the periods to year and month from the excel**

print('->>>>><<<<<-Split Year->>>>><<<<<-')

df\_Countries['Year'] = pd.to\_numeric(df\_Countries['Year'])

cs1 = df\_Countries['Year'].dtypes

**Print data**

print("Year type is", cs1)

print("->>>>><<<<<-dropped periods->>>>><<<<<-")

print(df\_Countries.drop(['Periods'], axis=1))

df\_Countries1 = df\_Countries[(df\_Countries['Year'] >= 2006) & (df\_Countries['Year'] <= 2017)]

print('->>>>><<<<<-df\_Countries1->>>>><<<<<-')

**show the first three country**

**show the last three country**

print(df\_Countries1.head(3))

print(df\_Countries1.tail(3))

df\_Countries2 = df\_Countries1[["Americas", "Canada", "USA", "Other Markets In Americas", "Oceania", "Australia", "New Zealand", "Other Markets In Oceania", "Africa", "Egypt", "Mauritius", "South Africa (Rep Of)", "Other Markets In Africa"]]

**List to calculator data**

print('-------df\_Countries2-----')

**Print data**

print(df\_Countries2)

print('sorted')

**Create new dataframe to read specific columns**

df\_Countries3 = df\_Countries2.replace(',', '', regex=True)

df\_Countries4 = df\_Countries3.replace('na', '0', regex=True)

print('->>>>><<<<<-df\_Countries4->>>>><<<<<-')

print(df\_Countries4)

df\_Countries5 = df\_Countries4.astype(int)

**Sort the df\_Countries4**

print(df\_Countries5.dtypes)

psNotSorted=df\_Countries5.sum()

psSorted = df\_Countries5.sum().sort\_values(ascending=False)

print('->>>>><<<<<-sorting->>>>><<<<<-')

print(psSorted)

**Print the top 3 country**

csCountries = psSorted

print('->>>>><<<<<-Top 3 Countries->>>>><<<<<-')

top3countries = psSorted.head(3)

print(top3countries)

total=top3countries.values.sum()

mean=round(top3countries.values.mean(),2)

**Print the total and top 3 country value**

print("The total no. of visitors for the top 3 countries is ",total)

print("The mean value for the top 3 countries is ",mean)

indexAll = np.arange(len(csCountries.index))

plt.figure(figsize=(10, 10))

**Print the graph for the country**

plt.title('All other Countries from(Period:2006-2017)')

plt.xlabel('Countries', fontsize=8)

plt.ylabel('No. of Travellers (in thousands)', fontsize=8)

plt.xticks(indexAll, csCountries.index, fontsize=6, rotation=30)

plt.bar(csCountries.index, csCountries.values / 1000)

plt.show()

indexAll = np.arange(len(csCountries.index))

**Print the graph for the top 3 country**

plt.figure(figsize=(10, 10))

plt.xlabel('Countries', fontsize=8)

plt.ylabel('No. of Travellers (in thousands)', fontsize=8)

plt.xticks(indexAll, csCountries.index, fontsize=6, rotation=30)

plt.bar(top3countries.index, top3countries.values / 1000)

plt.show()