import pandas as pd

Import python libararies

country = pd.read\_excel('IMVA.xls', sheet\_name = "IMVA")  
print(country)

Read 'IMVA.xls' into dataframe

print(country.columns)

Print out header columns

asia5 = country[['Periods','Brunei Darussalam','Indonesia','Malaysia','Myanmar','Philippines','Thailand','Vietnam','China','Hong Kong SAR','Taiwan','Japan','South Korea','Bangladesh','India','Pakistan','Sri Lanka','Iran','Israel','Kuwait','Saudi Arabia','United Arab Emirates']]  
print(asia5.columns)  
print(asia5)

Print selected header in new dataframe and print out columns of the selected country

date = asia5['Periods'].str.split(' ', n = 1, expand = True)  
print(date)

Split periods column and print the new splited column

asia5 = asia5.assign(year=date[0])

Assign new column

set1 = asia5[(asia5['year']) >= str(2011)]  
print(set1)  
  
print(set1.head(3))  
print(set1.tail(3))  
  
set2 = set1[['Brunei Darussalam','Indonesia','Malaysia','Myanmar','Philippines','Thailand','Vietnam','China','Hong Kong SAR','Taiwan','Japan','South Korea','Bangladesh','India','Pakistan','Sri Lanka','Iran','Israel','Kuwait','Saudi Arabia','United Arab Emirates']]

Show the selected data year and print the first 3 and last 3

To make a new dataframe without periods columns

asia = set2.replace(',','', regex=True)  
asia111 = asia.replace('na','0', regex=True)  
print(asia)

Replace na to 0 and remove ‘,’

asia1 = asia111.astype(int)  
print(asia1.dtypes)  
psNotSorted=asia1.sum()  
print(psNotSorted)  
psSorted = psNotSorted.sort\_values(ascending=False)  
print(psSorted)  
print(psSorted.head(3))  
print('The total no. of visitors for the top 3 countries is',sum(psSorted.head(3)))  
print('The mean value for the top 3 countries is',round(sum(psSorted.head(3))/len(psSorted.head(3)),2))

Sort the country from descending order

Show the sum and mean for the top 3 country

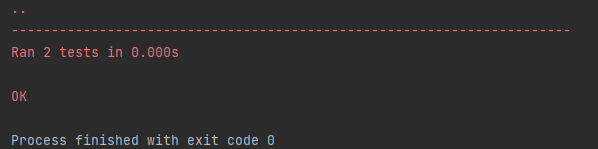
import numpy as np  
import matplotlib.pyplot as plt

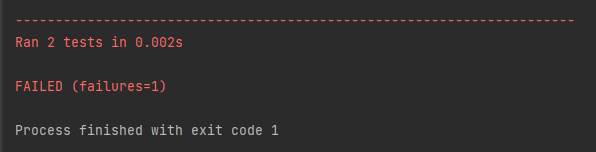
Import libraries

ps = psNotSorted.sort\_values(ascending=False)  
index = np.arange(len(ps.index))  
plt.bar(ps.index, ps.values/2500)  
plt.xlabel("Countries", fontsize=10)  
plt.ylabel("No. of Travellers (In thousands)", fontsize=10)  
plt.xticks(index, ps.index, fontsize=10, rotation=80)  
plt.title("All other countries from (Period:2011 - 2020) ", size=12)  
plt.show();  
  
top3 = psSorted.head(3)  
  
ps = top3.sort\_values(ascending=False)  
index = np.arange(len(ps.index))  
plt.xlabel('Countries', fontsize=10)  
plt.ylabel('No. of Travellers(in thousands)', fontsize=10)  
plt.xticks(index, ps.index, fontsize=10, rotation=45)  
plt.title('Top 3 Asia Countries from (Period:2011 - 2020)')  
plt.bar(ps.index, ps.values/1000)  
plt.show();

Show top 3 countries bar graph

Show all countries bar graph





import unittest  
import group5\_ASP\_Project as prog  
  
class MyTestCase(unittest.TestCase):  
 def test\_total(self):  
 data = prog.top3  
 result = prog.Test.sum(data)  
 self.assertEqual(result, 60923003)  
  
 def test\_mean(self):  
 data = prog.top3  
 result = round(prog.Test.mean(data), 2)  
 self.assertEqual(result, 20307667.6)  
  
  
if \_\_name\_\_ == '\_\_main\_\_':  
 unittest.main()  
  
'Brunei Darussalam','Indonesia','Malaysia','Myanmar','Philippines','Thailand','Vietnam','China','Hong Kong SAR','Taiwan','Japan','South Korea','Bangladesh','India','Pakistan','Sri Lanka','Iran','Israel','Kuwait','Saudi Arabia','United Arab Emirates'