

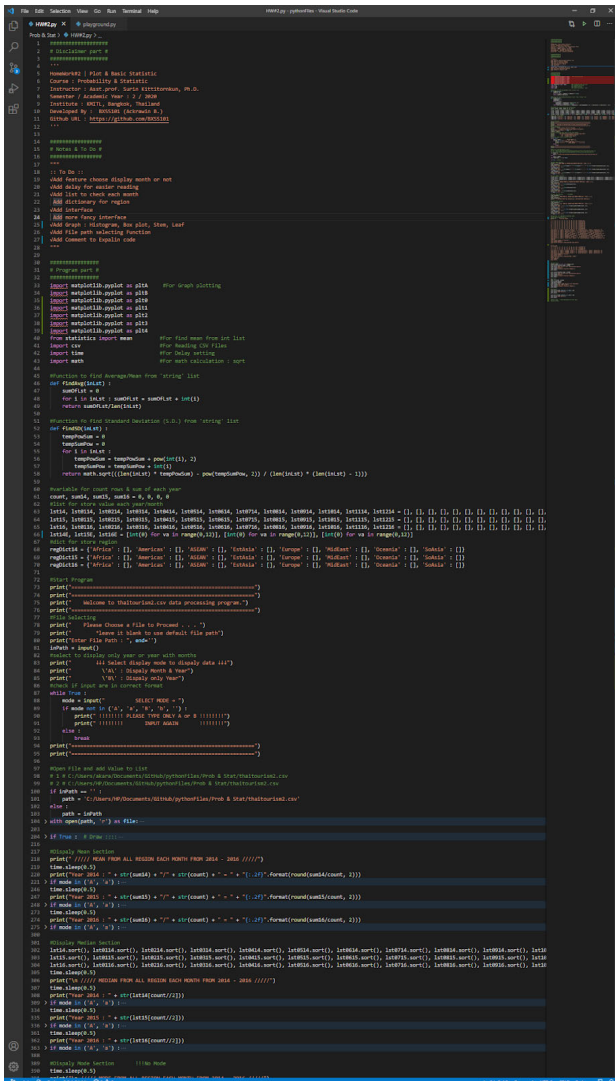
จากงาน HW#1 ได้ข้อมูลจำนวนนักท่องเที่ยวจากแต่ละภูมิภาคเป็น 5 คอลัมน์ ได้แก่ Region, Month, Yr2014, Yr2015 และ YR2016 โดยมีข้อมูลที่เป็นตัวเลข 3 คอลัมน์ซึ่งสามารถนำมาคำนวณค่าต่างๆตามที่ได้รับมอบหมายจากงานHW#2ได้ และสามารถนำมาแสดงผลเป็นกราฟแบบต่างๆได้ ซึ่งต้องคำนวณค่าออกมาดังนี้ : Mean, และแสดงผลกราฟดังนี้ : XY Scatter, Hi

	A	B	C	D	E
1	Region	Month	Yr2014	Yr2015	Yr2016
22	Americas	9	66664	98434	83152
23	Americas	10	90863	98434	106492
24	Americas	11	105350	123942	144750
25	Americas	12	131182	144527	168917
26	ASEAN	1	416820	586095	646833
27	ASEAN	2	437269	563927	672561
28	ASEAN	3	505648	639947	717634
29	ASEAN	4	539183	651714	712668
30	ASEAN	5	527748	695168	741764
31	ASEAN	6	482304	699993	722209

ได้ ซึ่งต้องคำนวณค่าออกมาดังนี้ : Mean, Median, Mode, S.D.  
และแสดงผลกราฟดังนี้ : XY Scatter, Histogram, Box Plot, Stem & Leaf

จากข้อมูลในไฟล์csvผู้จัดทำได้เลือก  
ข้อมูลที่จะนำมาเปรียบเทียบ 2 คอลัมน์ (  
ตามโจทย์) ได้แก่คอลัมน์ Yr2015 และ  
Yr2016 โดยเลือกใช้ภาษา python ใน  
การพัฒนาโปรแกรมเพื่อประมวลผลข้อมูลและ  
แสดงผลกราฟรูปแบบต่างๆ ส่วนของการ  
ประมวลผลข้อมูลมีการใช้ library ต่างๆ  
เพื่อช่วยการทำงานดังนี้ library  
statistic และ library math เพื่อ  
ช่วยการคำนวณค่า library csv เพื่อใช้  
เปิดอ่านไฟล์csv และ library  
matplotlib เพื่อแสดงกราฟรูปแบบต่างๆ

จากข้อมูล มีค่าบางอย่างที่ไม่เหมาะกับการ  
หาค่าหรือไม่สามารถหาค่าได้ เนื่องจาก  
ข้อมูลมีการกระจายตัวสูง และข้อมูลไม่มีการ  
ซ้ำกัน เช่น ไม่สามารถหาModeได้ ไม่มา  
ารถplot stem & leafได้



## ส่วนค่าสถิติพื้นฐาน

```

C:\WINDOWS\system32\cmd.exe - python "C:\Users\HP\Documents\GitHub\pythonFiles\Prob & Stat\HW#2.py"
C:\Users\HP>python "C:\Users\HP\Documents\GitHub\pythonFiles\Prob & Stat\HW#2.py"

=====
Welcome to thaitourism2.csv data processing program.
=====
Please Choose a File to Proceed . . .
*leave it blank to use default file path
Enter File Path :
Enter Display Mode :
Select display mode to display data
'A' : Display Month & Year
'B' : Display only Year
SELECT MODE -> B

=====

Year 2014 : 24809683/97 = 255769.93
Year 2015 : 30137563/97 = 310696.53
Year 2016 : 32588303/97 = 335961.89

Year 2014 : 97737
Year 2015 : 113419
Year 2016 : 121219

Year 2014 : 267973.94
Year 2015 : 348212.80
Year 2016 : 380694.29

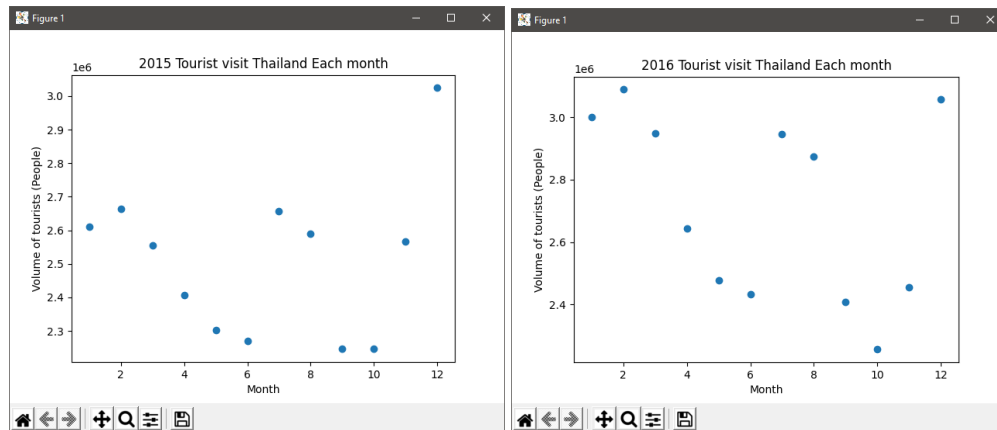
. . . Plotting Graphs . . .
  
```

ภาพหน้าจอของโปรแกรม

จากภาพของโปรแกรมจะแสดงค่าต่างๆของแต่ละคอลัมน์ โดยมีทั้งหมด 4 ค่าด้วยกันได้แก่ Mean, Median, Mode และ S.D. ของทั้งปี 2014 - 2016 โดยค่าทั้งหมดคำนวณมาจากจำนวนนักท่องเที่ยวในแต่ละเดือนของปีนั้นๆ

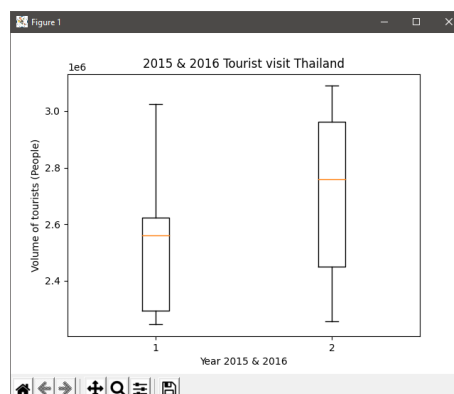
- **Mean** แสดงค่าเฉลี่ยของจำนวนนักท่องเที่ยว ซึ่งในปี 2016 มีนักท่องเที่ยวเฉลี่ยต่อเดือนมากที่สุดที่ 335,961.89 คน และยังมีนักท่องเที่ยวรวมทั้งปีมากที่สุดที่ 32,588,303 คน รองลงมาคือปี 2015 และ 2014 ตามลำดับ
- **Median** แสดงค่ามัธยฐานหรือค่ากลางของข้อมูลเมื่อมีการเรียงลำดับน้อยไปมากแล้ว ในปี 2016 มีค่าอยู่ที่ 121,279 คน ปี 2015 มีค่า 113,419 คน
- **Mode** แสดงค่าฐานนิยมหรือค่าที่ซ้ำกันมากที่สุด ซึ่งในข้อมูลชุดนี้ค่าแต่ละช่องไม่มีซ้ำกันจึงไม่มีฐานนิยม
- **S.D.** แสดงค่าเบี่ยงเบนมาตรฐานหมายถึงข้อมูลมีการกระจายตัวมากน้อยเพียงใด

## ส่วนกราฟ



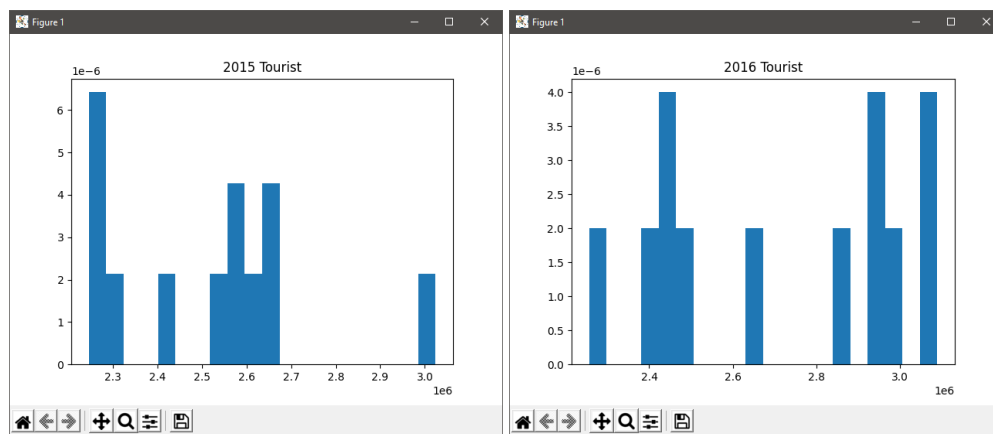
รูปแสดงผลลัพธ์ของโปรแกรม : XY (Scatter Plot)

จากกราฟจะเห็นได้ชัดว่าข้อมูลมีการกระจายตัวมาก แต่จะมีทิศทางข้อมูลที่คล้ายๆกัน เช่น เดือน10มีค่าต่ำทั้งสองกราฟ เดือน12 มีค่าพุ่งขึ้นสูง เป็นต้น



รูปแสดงผลลัพธ์ของโปรแกรม : Box Plot

จากกราฟจะเห็นว่าข้อมูลปี2015มีการกระจุกตัวที่ต่ำกว่าค่าของปี2016 แสดงให้เห็นว่านักท่องเที่ยวมีการเข้ามาเที่ยวในประเทศไทยปี2016มากกว่าในหลายช่วง



รูปแสดงผลลัพธ์ของโปรแกรม : Histogram

## ส่วน Coding

เนื่องจากโค้ดทั้งหมดของโปรแกรมมีความยาวมาก เพราะผู้จัดทำใส่ฟังก์ชันพิเศษเพิ่มเติมเข้าไปคือสามารถดูค่าต่างๆรายเดือนได้ด้วยนอกจากดูรายปี ดังนั้นจึง Capture ส่วนที่ใช้คำนวณที่สำคัญกับโจทย์มาเพื่อพิจารณา หาดต้องการดูเพิ่มเติมสามารถใช้ GitHub เพื่ออ่านโค้ดเพิ่มเติม : [View in GitHub](#)  
 ดูคลิปการทำงานของโปรแกรมเพิ่มเติม : [View in Drive\(KMITL account\)](#)

```
#Display Mean Section
print(" ///// MEAN FROM ALL REGION EACH MONTH FROM 2014 - 2016 /////")
time.sleep(0.5)
print("Year 2014 : " + str(sum14) + "/" + str(count) + " = " + "{:.2f}".format(round(sum14/count, 2)))
if mode in ('A', 'a') : ...
time.sleep(0.5)
print("Year 2015 : " + str(sum15) + "/" + str(count) + " = " + "{:.2f}".format(round(sum15/count, 2)))
if mode in ('A', 'a') : ...
time.sleep(0.5)
print("Year 2016 : " + str(sum16) + "/" + str(count) + " = " + "{:.2f}".format(round(sum16/count, 2)))
if mode in ('A', 'a') : ...
```

ส่วนนี้ใช้สำหรับหาค่าเฉลี่ย

```

#Display Median Section
lst14.sort(), lst0114.sort(), lst0214.sort(), lst0314.sort(), lst0414.sort()
lst15.sort(), lst0115.sort(), lst0215.sort(), lst0315.sort(), lst0415.sort()
lst16.sort(), lst0116.sort(), lst0216.sort(), lst0316.sort(), lst0416.sort()
time.sleep(0.5)
print("\n ///// MEDIAN FROM ALL REGION EACH MONTH FROM 2014 - 2016 /////")
time.sleep(0.5)
print("Year 2014 : " + str(lst14[count//2]))
if mode in ('A', 'a') : ...
time.sleep(0.5)
print("Year 2015 : " + str(lst15[count//2]))
if mode in ('A', 'a') : ...
time.sleep(0.5)
print("Year 2016 : " + str(lst16[count//2]))
if mode in ('A', 'a') : ...

```

ส่วนนี้ใช้สำหรับค่ามัธยฐาน

```

#Function fo find Standard Deviation (S.D.) from 'string' list
def findSD(inLst) :
    tempPowSum = 0
    tempSumPow = 0
    for i in inLst :
        tempPowSum = tempPowSum + pow(int(i), 2)
        tempSumPow = tempSumPow + int(i)
    return math.sqrt(((len(inLst) * tempPowSum) - pow(tempSumPow, 2)) / (len(inLst) * (len(inLst) - 1)))

```

ส่วนนี้ใช้สำหรับค่าส่วนเบี่ยงเบนมาตรฐาน

```
#Scatter plot
months_range = [ i for i in range(1,13)]
plt0.scatter(months_range, lst15E)
plt0.title('2015 Tourist visit Thailand Each month')
plt0.xlabel('Month')
plt0.ylabel('Volume of tourists (People)')
plt0.show()

plt1.scatter(months_range, lst16E)
plt1.title('2016 Tourist visit Thailand Each month')
plt1.xlabel('Month')
plt1.ylabel('Volume of tourists (People)')
plt1.show()

#Box plot
data = [lst15E, lst16E]
fig = pltB.figure()
pltB.boxplot(data)
pltB.title('2015 & 2016 Tourist visit Thailand')
plt1.xlabel('Year 2015 & 2016')
plt1.ylabel('Volume of tourists (People)')
pltB.show()

#Histogram
plt2.hist(lst15E, density = 1, bins = 20)
plt2.title('2015 Tourist')
plt2.show()

plt3.hist(lst16E, density = 1, bins = 20)
plt3.title('2016 Tourist')
plt3.show()
```

ส่วนนี้สำหรับการplotกราฟ