

Homework 2

- ใช้ Column :

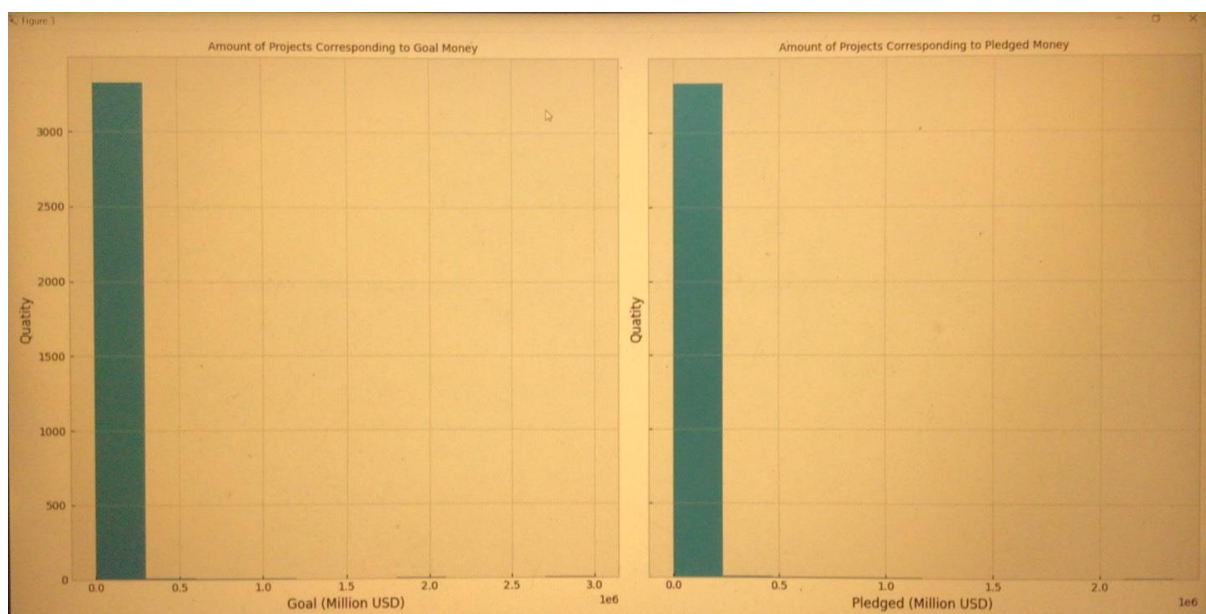
- Goal : จำนวนที่ต้องผู้สร้างโปรเจคต้องการ , หน่วย Million USD
- Pledged : จำนวนเงินที่ได้รับบริจาค , หน่วย Million USD

- ค่าทางสถิติต่างๆ

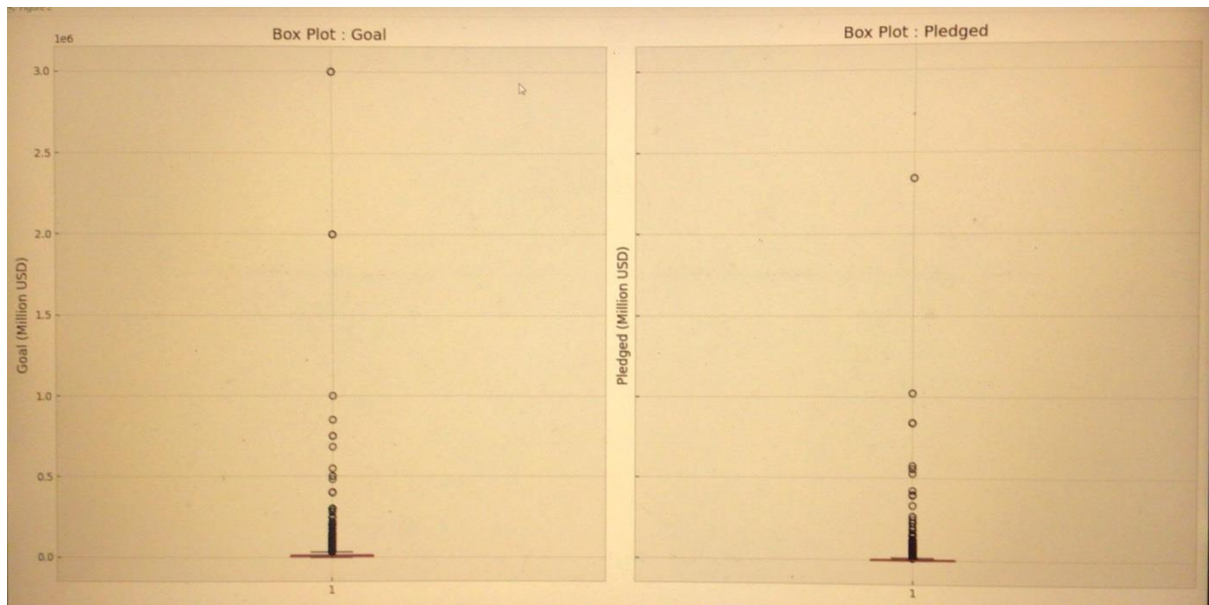
	Goal	Pledged
Mean	21699.36 USD	9125.37 USD
Median	5000.0 USD	787.0 USD
Mode	5000.0 USD	0 USD
SSD	100961.71	56502.18
Sample Variance	10193267039.86	3192495868.74

- กราฟต่างๆ

Histogram



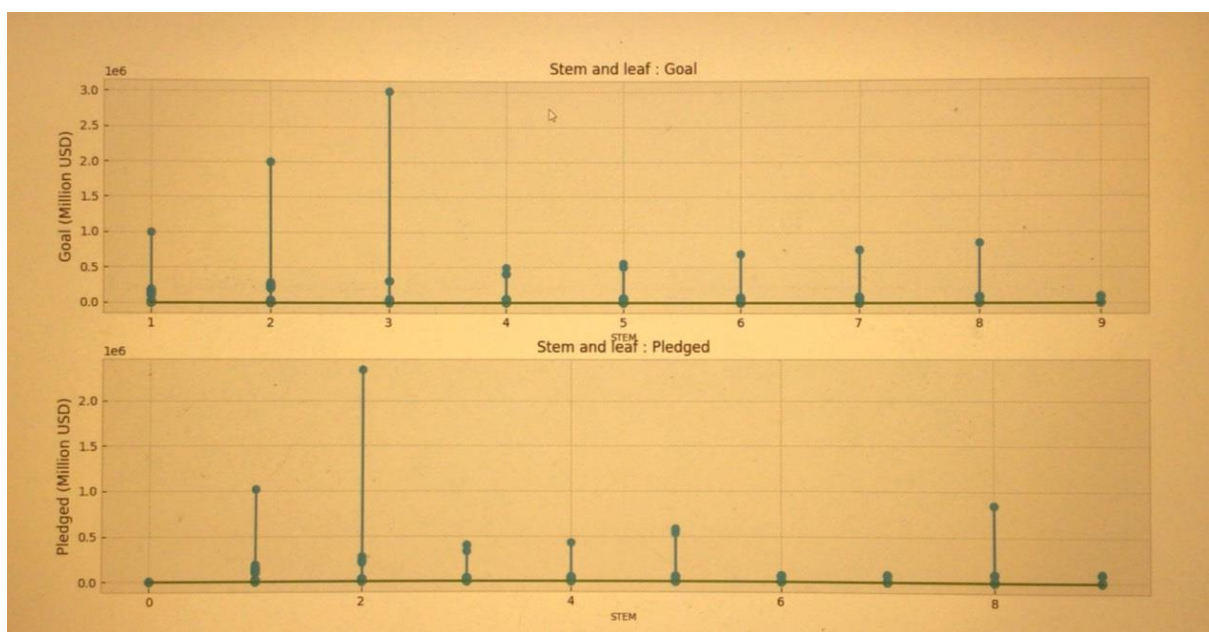
Box Plot



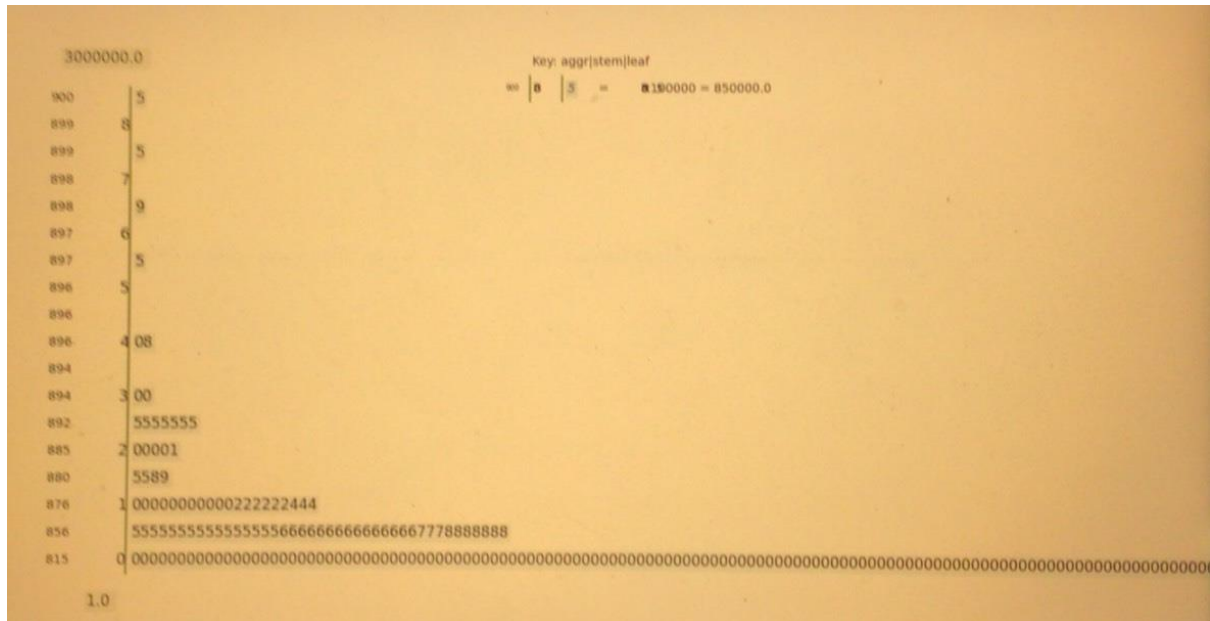
Outlier

- Goal : Outlier ของ Goal จะอยู่ในช่วง $\sim 0.1 - 3.0$ Million USD
- Pledged : Outlier ของ Pledged จะอยู่ในช่วง $\sim 0.05 - 2.3$ Million USD

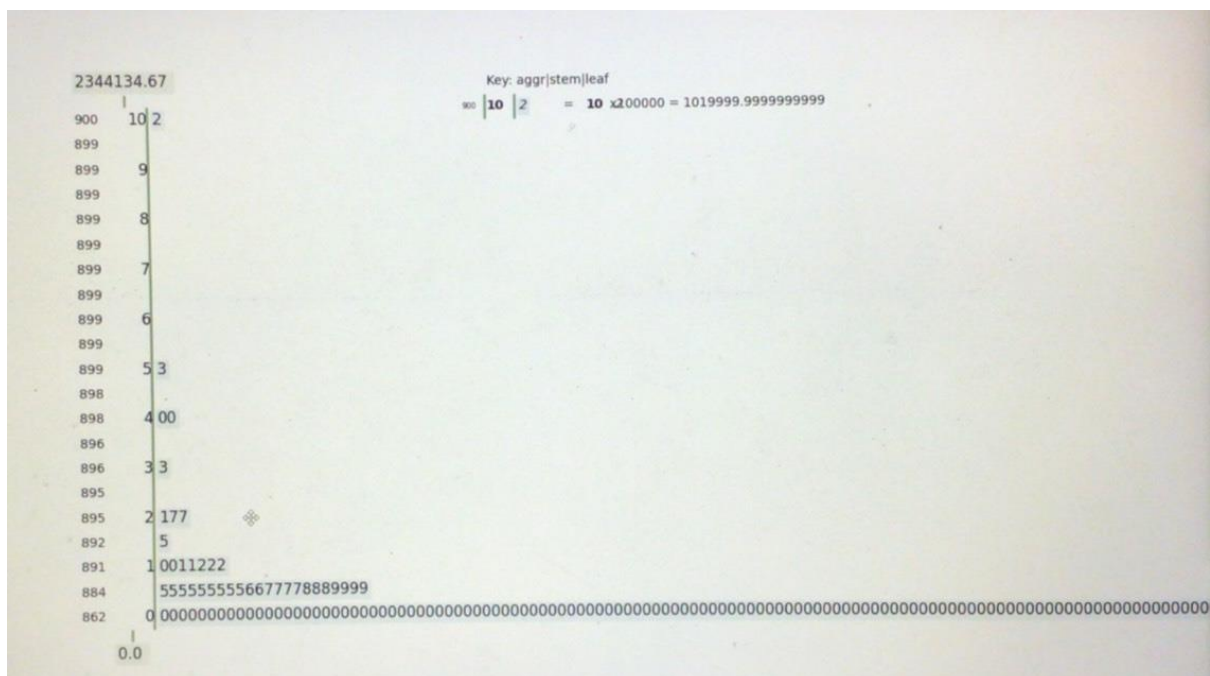
Stem & Leaves



Goal : Stem & Leaves



Pledged : Stem & Leaves



Scatter



- ตัวแปรต้น : Goal (จำนวนเงินที่ผู้สร้างโปรเจกต้องการ)
- ตัวแปรตาม : Pledged (จำนวนเงินที่ผู้สร้างโปรเจกได้รับการบริจาค)
- เหตุผล

เหตุผลที่ใช้ Goal เป็นตัวแปรต้นเพราะต้องการศึกษาว่า ปริมาณเงินที่ผู้สร้างโปรเจกขอรับบริจาคจะมีผลต่อเงินที่ได้รับบริจาคหรือไม่ เช่น หากตั้งจำนวนเงินที่ต้องการสูง จะส่งผลให้ผู้บริจาคสมทบทุนน้อยลงหรือไม่

บทวิเคราะห์ข้อมูลจากกราฟ

จากข้อมูลที่น่ามาจะเห็นได้ว่า หากจำนวนเงินที่ผู้สร้างโปรเจกขอมีปริมาณอยู่ในช่วง 0-30000 USD โปรเจกนั้นจะมีโอกาสระดมทุนสำเร็จสูง และจำนวนเงินที่ได้รับบริจาคจะเกินจำนวนเงินที่ขอไปในระดับหนึ่งหรือในบางกรณีจะเกินไปเป็นจำนวนมาก แต่เมื่อในจำนวนเงินที่ขออยู่ในช่วง 30000 USD ขึ้นไป โอกาสที่โปรเจกนั้นจะระดมทุนสำเร็จนั้นน้อยมากหรือแทบจะไม่มีเลย เพราะจำนวนเงินที่ได้รับบริจาคกลับเป็นไปในทางตรงกันข้าม กล่าวอีกอย่างคือ ยิ่งจำนวนเงินเป้าหมายมาก จำนวนเงินที่ได้รับจะน้อย

เราจึงสามารถวิเคราะห์ได้ว่า ปริมาณเงินที่ผู้สร้างโปรเจกขอรับบริจาคนั้นส่งผลต่อการระดมทุนอย่างเห็นได้ชัด โดยจะมีความสัมพันธ์แบบแปรผกผัน คือ ยิ่งจำนวนเงินเป้าหมายมาก จำนวนเงินที่ได้รับกลับมีน้อย

Source Code : Python

```
import matplotlib.pyplot as plt
import pandas as pd
import statistics as stc
import stemgraphic as stg

plt.style.use('bmh')
df = pd.read_csv('Data.csv')

x = df['goal']
y = df['pledged']

goal = x.to_list()
pledged = y.to_list()

#Goal Statistic Values
print("Goal Statistics Values")
#Goal Mean
goal_mean = stc.mean(goal)
print("Goal Mean : " + str(goal_mean) + ' USD')

#Goal Median
goal_median = stc.median(goal)
print("Goal Median : " + str(goal_median) + ' USD')

#Goal Mode
goal_mode = stc.mode(goal)
print("Goal Mode : " + str(goal_mode) + ' USD')

#Goal Sample Standard Deviation
goal_ssd = stc.stdev(goal)
print("Goal SSD : " + str(goal_ssd) + ' USD')

#Goal Sample Variance of Data
goal_variance = stc.variance(goal)
print("Goal Sample Variance : " + str(goal_variance) + ' USD')

print()
print()

#Pledged Statistic Values
print("Pledged Statistics Values")
#Pledged Mean
pledged_mean = stc.mean(pledged)
print('Pledged Mean : ' + str(pledged_mean) + ' USD')

#Pledged Median
```



```

pledged_median = stc.median(pledged)
print('Pledged Median : ' + str(pledged_median) + ' USD')

#Pledged Mode
pledged_mode = stc.mode(pledged)
print('Pledged Mode : ' + str(pledged_mode) + ' USD')

#Pledged Sample Standard Deviation
pledged_ssd = stc.stdev(pledged)
print('Pledged SSD : ' + str(pledged_ssd) + ' USD')

#Pledged Sample Variance of Data
pledged_variance = stc.variance(pledged)
print('Pledged Sample Variance : ' + str(pledged_variance) + ' USD')

#Graph Plotting

#Scatter
plt.title("Kickstarter Project : Amount of Goal & Pledge")
plt.xlabel('Goal (Million USD)',fontsize = 15)
plt.ylabel('Pledged (Million USD)',fontsize = 15)
plt.scatter(x,y)

#Subplot
fig, box = plt.subplots(1, 2, sharey=True, tight_layout=True)
fig, hist = plt.subplots(1, 2, sharey=True, tight_layout=True)
fig, stem = plt.subplots(2)

#BoxPlot
#Goal
box[0].set_title("Box Plot : Goal")
box[0].set_ylabel('Goal (Million USD)')
box[0].boxplot(x)
#Pledged
box[1].set_title("Box Plot : Pledged")
box[1].set_ylabel('Pledged (Million USD)')
box[1].boxplot(y)

#Histogram
#Goal
hist[0].set_title('Amount of Projects Corresponding to Goal Money',fontsize=10)
hist[0].set_ylabel("Quatity")
hist[0].set_xlabel("Goal (Million USD)")
hist[0]
hist[0].hist(x)
#Pledged

```

```
hist[1].set_title('Amount of Projects Corresponding to Pledged Money',fontsize
=10)
hist[1].set_ylabel("Quatity")
hist[1].set_xlabel("Pledged (Million USD)")
hist[1].hist(y)

#Stem

#set X axis
#Goal
# goal_axis = []
# for ele in x :
#     goal_axis.append(int(str(ele)[:1]))

# stem[0].set_title('Stem and leaf : Goal',fontsize=12)
# stem[0].set_xlabel('STEM',fontsize=8)
# stem[0].set_ylabel('Goal (Million USD)',fontsize=12)
# stem[0].stem(goal_axis,x)
stem[0] = stg.stem_graphic(x)

# #Pledged
# pledged_axis =[]
# for ele in y:
#     pledged_axis.append(int(str(ele)[:1]))
# stem[1].set_title('Stem and leaf : Pledged',fontsize=12)
# stem[1].set_xlabel('STEM',fontsize=8)
# stem[1].set_ylabel('Pledged (Million USD)',fontsize=12)
# stem[1].stem(pledged_axis,y)
stem[1] = stg.stem_graphic(y)

plt.show()
```