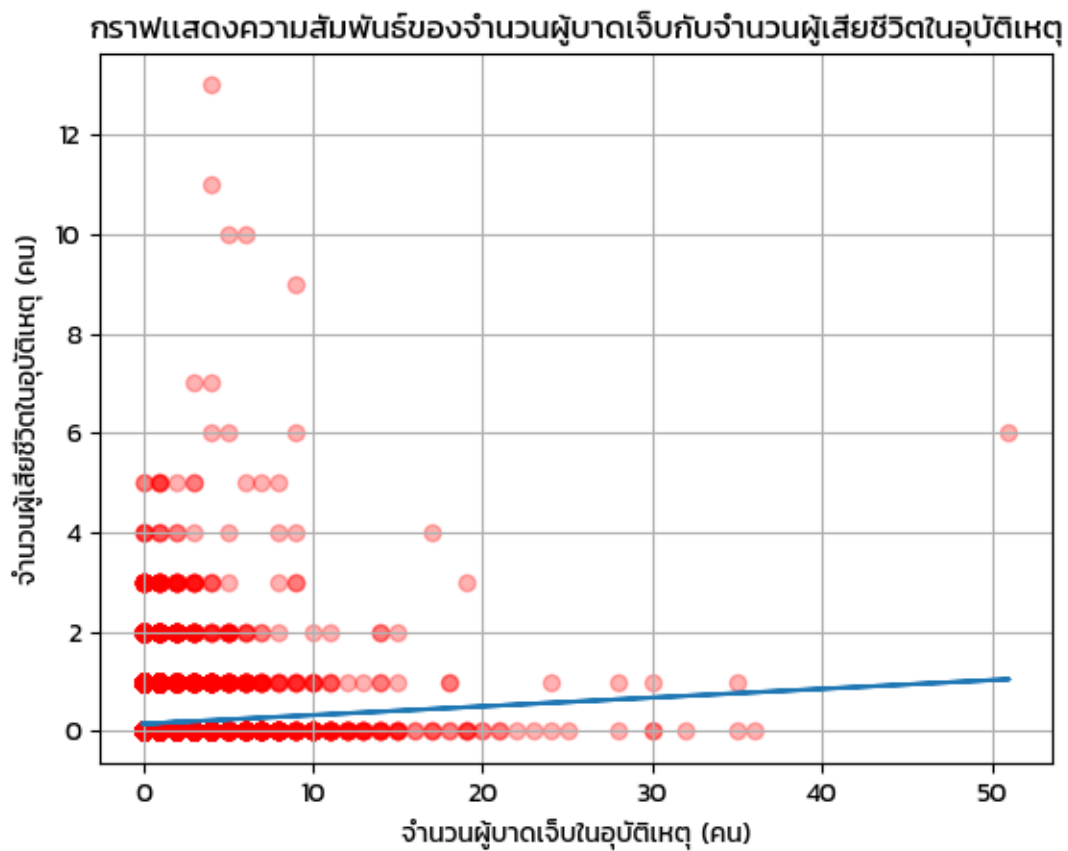


ProbStat_HW5

ชื่อข้อมูล : อุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562

คอลัมน์ที่เลือกใช้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ , จำนวนผู้เสียชีวิตในอุบัติเหตุ

การวิเคราะห์ความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ โดย Linear Regression



ผลการวิเคราะห์

```
=====
สมการ Linear Regression : Y = (0.017600744672421075*X) + 0.14714875300095107
ค่าสัมประสิทธิ์สหสัมพันธ์ (r) : 0.061045859658720183
ค่าคลาดเคลื่อนมาตรฐาน : 0.0020368131663988463
=====
```

รูปที่ 1 แสดงผลการวิเคราะห์ความสัมพันธ์ระหว่างจำนวนผู้บาดเจ็บ และจำนวนผู้เสียชีวิตในอุบัติเหตุ

จากผลการวิเคราะห์ความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ ได้ผลลัพธ์ ดังนี้ สมการ Linear Regression คือ $Y = (0.02 * X) + 0.15$ และค่าสัมประสิทธิ์สหสัมพันธ์ (r) คือ 0.06 โดยมีค่าคลาดเคลื่อนมาตรฐาน คือ 0.002 ซึ่งกล่าวได้ว่า ความสัมพันธ์ระหว่างจำนวนผู้บาดเจ็บและจำนวนผู้เสียชีวิตในอุบัติเหตุ นั้น มีความสัมพันธ์ในทิศทางเดียวกัน หรือกล่าวได้ว่าการเกิดอุบัติเหตุทางท้องถนนในแต่ละครั้งนั้น “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมากด้วยเช่นกัน” ทั้งนี้เนื่องจากค่าสัมประสิทธิ์สหสัมพันธ์ (r) มีค่าเพียง 0.06 ซึ่งกล่าวได้ว่าความสัมพันธ์ที่กล่าวมาข้างต้นนั้นมีความสัมพันธ์กันค่อนข้างต่ำ ซึ่งก็คือในการเกิดอุบัติเหตุแต่ละครั้งนั้น “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมาก” อาจไม่สามารถกล่าวได้อย่างมั่นใจมากนัก ทั้งนี้ค่าความคลาดเคลื่อนในการวิเคราะห์นี้อยู่ที่ 0.002 ซึ่งก็คือการวิเคราะห์นี้ค่อนข้างแม่นยำ

สรุปได้ว่าความสัมพันธ์ของจำนวนผู้บาดเจ็บ กับจำนวนผู้เสียชีวิตในอุบัติเหตุ มีความสัมพันธ์กันค่อนข้างต่ำ โดยลักษณะความสัมพันธ์ คือ “หากจำนวนผู้บาดเจ็บมีจำนวนมาก แล้วจะมีจำนวนผู้เสียชีวิตมาก”

Source Code ของโปรแกรม

```
1 import csv
2 import statistics
3 import pandas as pd
4 import matplotlib.pyplot as plt
5 import numpy as np
6 import scipy.stats
7
8 plt.rcParams['font.family'] = 'kanit'
9
10 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
11
12 with open(csv_path, encoding='utf-8') as csv_file:
13     csv_reader = csv.reader(csv_file)
14
15     dead_num_list = []
16     patient_num_list = []
17
18     accident_num = 0
19     for row in csv_reader:
20
21         if accident_num != 0:
22             vehicle, dead_num, patient_num = row[10], int(row[15]), int(row[16])
23
24             dead_num_list.append(dead_num)
25             patient_num_list.append(patient_num)
26
27         accident_num += 1
28
29     dead_num_list = np.array(dead_num_list)
30     patient_num_list = np.array(patient_num_list)
31
32     ## Linear Regression Solve
33     (slope, intercept, r_value, p_value, std_err) = scipy.stats.linregress(patient_num_list, dead_num_list)
34
35     ## XY Scatter Plot
36     plt.scatter(patient_num_list, dead_num_list, color='red', alpha=0.3)
37     plt.plot(patient_num_list, slope*patient_num_list + intercept)
38     plt.title("กราฟแสดงความสัมพันธ์ของจำนวนผู้บาดเจ็บกับจำนวนผู้เสียชีวิตในอุบัติเหตุ")
39     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
40     plt.ylabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
41     plt.grid()
42     plt.show()
43
44     print("=====\n")
45     print("สมการ Linear Regression : Y = ({}*X) + {}".format(slope, intercept))
46     print("ค่าสัมประสิทธิ์สหสัมพันธ์ (r) : {}".format(r_value))
47     print("ค่าคลาดเคลื่อนมาตรฐาน : {}".format(std_err))
48     print("\n=====")
```