

ProbStat_HW2

ชื่อข้อมูล : อุบัติเหตุบนโครงข่ายถนนของกระทรวงคมนาคม 2562

คอลัมน์ที่เลือกใช้วิเคราะห์ในครั้งนี้ : จำนวนผู้บาดเจ็บในอุบัติเหตุ ,จำนวนผู้เสียชีวิตในอุบัติเหตุ

ตัวอย่าง ข้อมูลที่ใช้ในการวิเคราะห์

ปีที่เกิดเหตุ	วันที่เกิดเหตุ	เวลา	รถคันที่ 1	จำนวนผู้เสียชีวิต	รวมจำนวนผู้บาดเจ็บ
2019	01/01/2019	00:00	รถจักรยานยนต์	0	2
2019	01/01/2019	00:03	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	2
2019	01/01/2019	00:05	รถจักรยานยนต์	1	0
2019	01/01/2019	00:20	รถจักรยานยนต์	0	1
2019	01/01/2019	00:25	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	0
2019	01/01/2019	00:30	รถปิคอัพบรรทุก 4 ล้อ	0	2
2019	01/01/2019	00:30	รถจักรยานยนต์	0	2
2019	01/01/2019	00:35	รถจักรยานยนต์	1	0
2019	01/01/2019	00:40	รถจักรยานยนต์	3	0
2019	01/01/2019	00:45	รถจักรยานยนต์	0	1
2019	01/01/2019	00:45	รถจักรยานยนต์	1	0
2019	01/01/2019	00:45	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	0	1
2019	01/01/2019	01:00	รถจักรยานยนต์	1	0
2019	01/01/2019	01:04	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:15	รถจักรยานยนต์	0	2
2019	01/01/2019	01:20	รถจักรยานยนต์	1	1
2019	01/01/2019	01:20	รถจักรยานยนต์	0	1
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	5
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	0	0
2019	01/01/2019	01:30	รถจักรยานยนต์	1	1
2019	01/01/2019	01:30	รถปิคอัพบรรทุก 4 ล้อ	0	1
2019	01/01/2019	01:30	รถยนต์นั่งส่วนบุคคล/รถยนต์นั่งสาธารณะ	1	0

รูปที่ 1 แสดงตัวอย่างข้อมูลที่ใช้ในการวิเคราะห์

การวิเคราะห์ค่าสถิติพื้นฐาน

```
=====
จำนวนอุบัติเหตุทั้งหมด : 19966 ครั้ง
จำนวนผู้บาดเจ็บทั้งหมด : 18191 คน
จำนวนผู้เสียชีวิตทั้งหมด : 3258 คน

-----

(median) มัธยฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : 1 คน
(mode) ฐานนิยม จำนวนผู้บาดเจ็บในอุบัติเหตุ : 0 คน
(mean) ค่าเฉลี่ย จำนวนผู้บาดเจ็บในอุบัติเหตุ : 0.91 คน
(SD) ส่วน เบี่ยงเบนมาตรฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : 1.70

(median) มัธยฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0 คน
(mode) ฐานนิยม จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0 คน
(mean) ค่าเฉลี่ย จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0.16 คน
(SD) ส่วน เบี่ยงเบนมาตรฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : 0.49
=====
```

รูปที่ 2 แสดงผลการคำนวณค่าสถิติพื้นฐานจากโปรแกรม

จากการคำนวณค่าสถิติพื้นฐาน โดยการเขียนโปรแกรมด้วยภาษา Python ได้ผลลัพธ์ดังนี้

สถิติพื้นฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ

ค่ามัธยฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 1 คน

ค่าฐานนิยมของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 0 คน

ค่าเฉลี่ยของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 0.91 คน

ส่วนเบี่ยงเบนมาตรฐานของจำนวนผู้บาดเจ็บในอุบัติเหตุ มีค่าเท่ากับ 1.70 คน

สถิติพื้นฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ

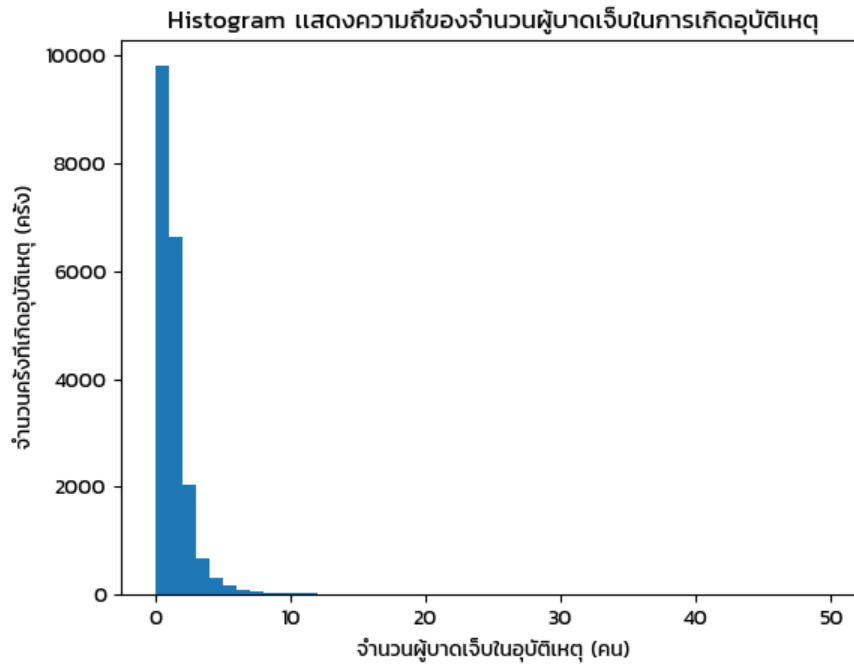
ค่ามัธยฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0 คน

ค่าฐานนิยมของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0 คน

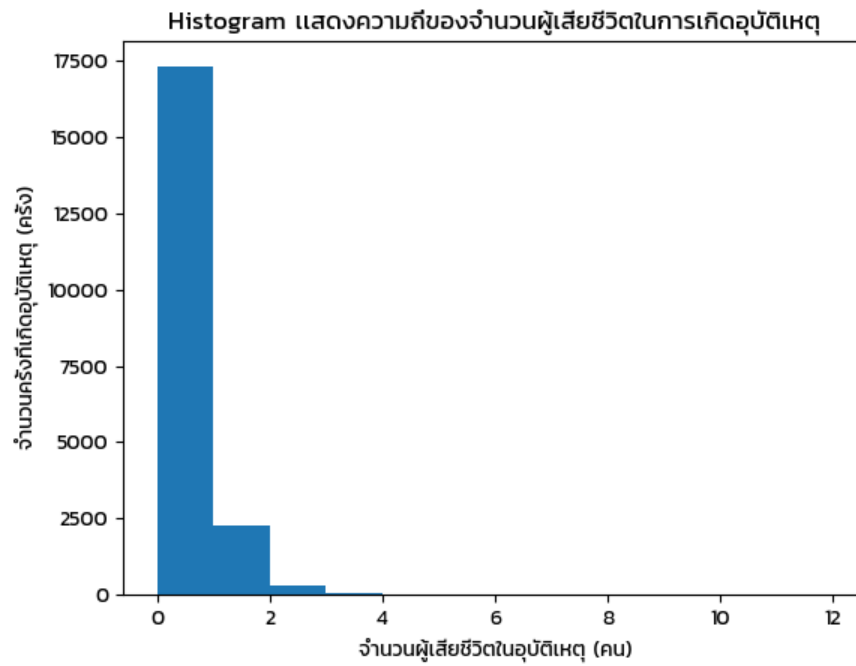
ค่าเฉลี่ยของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0.16 คน

ส่วนเบี่ยงเบนมาตรฐานของจำนวนผู้เสียชีวิตในอุบัติเหตุ มีค่าเท่ากับ 0.49 คน

การวิเคราะห์ Histogram



จาก Histogram แสดงความถี่ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ พบว่า โดยส่วนใหญ่แล้ว การเกิดอุบัติเหตุบนท้องถนนในแต่ละครั้ง จะมีจำนวนผู้บาดเจ็บ 0-1 คน



จาก Histogram แสดงความถี่ของจำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุ พบว่า โดยส่วนใหญ่แล้ว การเกิดอุบัติเหตุบนท้องถนนในแต่ละครั้ง จะมีจำนวนผู้เสียชีวิต 0-1 คน

การวิเคราะห์ XY Scatter Plot

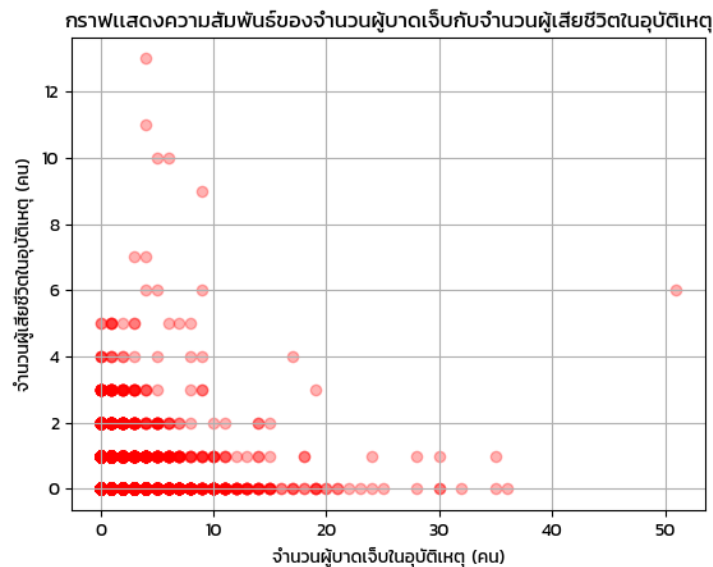
ตัวแปรที่เลือกในการทำ XY Scatter Plot

ตัวแปรต้น : จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)

ตัวแปรตาม : จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)

เหตุผลในการเลือกตัวแปร

เนื่องจาก XY Scatter Plot เป็นกราฟสำหรับแสดงแนวโน้มและความสัมพันธ์ของข้อมูลชนิด Numeric ต่อข้อมูลชนิด Numeric ซึ่งจำนวนผู้บาดเจ็บในอุบัติเหตุ และจำนวนผู้เสียชีวิตในอุบัติเหตุ นั้นเป็นข้อมูลชนิด Numeric ประกอบกับเพื่อที่จะสังเกตแนวโน้มและความสัมพันธ์ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ และจำนวนผู้เสียชีวิตในอุบัติเหตุ แต่ละครั้ง



Source Code ของโปรแกรม

```

1 import csv
2 import statistics
3 import pandas as pd
4 import matplotlib.pyplot as plt
5
6 plt.rcParams['font.family'] = 'kanit'
7
8 csv_path = "../datasets/ubatiehtuthaangthnn.csv"
9
10 with open(csv_path, encoding='utf-8') as csv_file:
11     csv_reader = csv.reader(csv_file)
12
13     dead_num_list = []
14     patient_num_list = []
15
16     accident_num = 0
17     for row in csv_reader:
18
19         if accident_num != 0:
20             vehicle, dead_num, patient_num = row[10], int(row[15]), int(row[16])
21
22             dead_num_list.append(dead_num)
23             patient_num_list.append(patient_num)
24
25             accident_num += 1
26
27     ## Result
28     print("=====\n")
29
30     print("จำนวนอุบัติเหตุทั้งหมด : {} ครั้ง".format(accident_num))
31     print("จำนวนผู้บาดเจ็บทั้งหมด : {} คน".format(sum(patient_num_list)))
32     print("จำนวนผู้เสียชีวิตทั้งหมด : {} คน".format(sum(dead_num_list)))
33
34     print("\n-----\n")
35
36     print("(median) มัธยฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : {} คน".format(statistics.median(patient_num_list)))
37     print("(mode) ฐานนิยม จำนวนผู้บาดเจ็บในอุบัติเหตุ : {} คน".format(statistics.mode(patient_num_list)))
38     print("(mean) ค่าเฉลี่ย จำนวนผู้บาดเจ็บในอุบัติเหตุ : {:.2f} คน".format(statistics.mean(patient_num_list)))
39     print("(SD) ส่วนเบี่ยงเบนมาตรฐาน จำนวนผู้บาดเจ็บในอุบัติเหตุ : {:.2f}".format(statistics.stdev(patient_num_list)))
40
41     print()
42
43     print("(median) มัธยฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : {} คน".format(statistics.median(dead_num_list)))
44     print("(mode) ฐานนิยม จำนวนผู้เสียชีวิตในอุบัติเหตุ : {} คน".format(statistics.mode(dead_num_list)))
45     print("(mean) ค่าเฉลี่ย จำนวนผู้เสียชีวิตในอุบัติเหตุ : {:.2f} คน".format(statistics.mean(dead_num_list)))
46     print("(SD) ส่วนเบี่ยงเบนมาตรฐาน จำนวนผู้เสียชีวิตในอุบัติเหตุ : {:.2f}".format(statistics.stdev(dead_num_list)))
47
48     print("\n=====")
49
50     ## Histogram
51     plt.hist(patient_num_list, bins=range(max(patient_num_list)))
52     plt.title("Histogram แสดงความถี่ของจำนวนผู้บาดเจ็บในการเกิดอุบัติเหตุ")
53     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
54     plt.ylabel("จำนวนครั้งที่เกิดอุบัติเหตุ (ครั้ง)")
55     plt.show()
56     plt.hist(dead_num_list, bins=range(max(dead_num_list)))
57     plt.title("Histogram แสดงความถี่ของจำนวนผู้เสียชีวิตในการเกิดอุบัติเหตุ")
58     plt.xlabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
59     plt.ylabel("จำนวนครั้งที่เกิดอุบัติเหตุ (ครั้ง)")
60     plt.show()
61
62     ## XY Scatter Plot
63     plt.scatter(patient_num_list, dead_num_list, color='red', alpha=0.3)
64     plt.title("กราฟแสดงความสัมพันธ์ของจำนวนผู้บาดเจ็บกับจำนวนผู้เสียชีวิตในอุบัติเหตุ")
65     plt.xlabel("จำนวนผู้บาดเจ็บในอุบัติเหตุ (คน)")
66     plt.ylabel("จำนวนผู้เสียชีวิตในอุบัติเหตุ (คน)")
67     plt.grid()
68     plt.show()

```