

แบบฝึกหัด Pandas: Time Series

โดย อ.ดร.เสกฐวิทย์ เกิดผล

1. เปิดไฟล์ `italy_earthquakes_from_2016-08-24_to_2016-11-30.csv`
2. ปรับคอลัมน์ Time ให้เป็น Timestamp และตั้งให้เป็น index
3. เวลาที่เกิดแผ่นดินไหวแรงที่สุดคือวันเวลาไหน
4. ในวันที่เกิดแผ่นดินไหวแรงที่สุด มีแผ่นดินไหวกี่ครั้ง
5. แสดงข้อมูลตั้งแต่ 7 ถึง 9 โมงเช้าในวันที่เกิดแผ่นดินไหวแรงที่สุด
6. แผ่นดินไหวที่แรงที่สุดกับแรงเป็นที่ 2 ห่างกันเท่าไร ตอบเป็นวัน-ชั่วโมง-นาที
7. คำนวณความแรงของแผ่นดินไหวที่เปลี่ยนไปในแต่ละครั้ง ของข้อมูลในข้อ 5.
หมายเหตุ: ใช้คำสั่ง `shift()` ช่วยเลื่อนแถวของคอลัมน์ Magnitude เพื่อคำนวณความแตกต่าง
8. มีแผ่นดินไหวเกิดขึ้นกี่ครั้ง ในแต่ละวัน
9. ถ้าแยกเวลาเป็นช่วง ช่วงละหนึ่งชั่วโมง มีกี่ช่วงที่ไม่มีแผ่นดินไหวเกิดขึ้นเลยในชั่วโมงนั้น
10. ถ้าแยกเวลาเป็นช่วง ช่วงละ 5 นาที ช่วงเวลาไหนมีค่าเฉลี่ยความรุนแรงของแผ่นดินไหวสูงที่สุด และค่าเฉลี่ยนั้นมีค่าเท่าไร
11. ระยะเวลาที่ไม่มีแผ่นดินไหวเลยนานที่สุดคือเท่าไร
12. ระยะเวลาที่ไม่มีแผ่นดินไหวเลยนานที่สุดคือช่วงไหน ระบุเป็นเวลาเริ่มต้นและเวลาจบ
13. สมมติว่าเรานิยามแผ่นดินไหวที่รุนแรงนั้นต้องมี Magnitude มากกว่าการไหวครั้งก่อนและครั้งหลังอย่างน้อย 2.0 จงคัดเลือกเหตุการณ์แผ่นดินไหวที่รุนแรงออกมา
14. จากตารางข้อ 13. ให้เพิ่มคอลัมน์ใหม่ชื่อ `hour_count` แสดงจำนวนแผ่นดินไหวในชั่วโมงที่มีเหตุการณ์แผ่นดินไหวรุนแรงด้วย
15. แสดงจำนวนแผ่นดินไหวแยกตาม 24 ชั่วโมง (เอาจำนวนแผ่นดินไหวของแต่ละชั่วโมง 0-23 ของทุกวันรวมกัน แสดงผลตามชั่วโมง)
16. คำนวณค่าเฉลี่ยของความรุนแรงแผ่นดินไหวที่มากที่สุดในช่วง 3 วัน โดยหาความแรงของแผ่นดินไหวที่มากที่สุดในแต่ละวันก่อน แล้วใช้ `sliding window` หาค่าเฉลี่ยของทุก ๆ 3 วัน