

Assignment 2

ทดลองวิเคราะห์และนำเสนอข้อมูลจริงด้วย R Project

วิชา Probability and Statistics

ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ภาคการศึกษาที่ 2 ปีการศึกษา 2560

คิดเป็น 15% ของเกรดตัวอักษร

วัตถุประสงค์

1. เพื่อให้นักศึกษาประยุกต์ความรู้จากการเรียนมาวิเคราะห์ข้อมูลจริง
2. เพื่อให้นักศึกษาฝึกการนำเสนอผลการวิเคราะห์ให้เป็นรูปธรรมด้วยภาษา R

คำสั่ง

1. จับกลุ่มๆ ละ 2 - 3 คน
2. เสนอและจองชื่อข้อมูลที่จะศึกษา กลุ่มละ 1 ชุด จากลิงค์ข้างล่าง โดยกรอกรายละเอียดทาง Google Form ที่ <https://goo.gl/forms/3LY43nYVHBj3HRpw2> โดยแจ้งชื่อข้อมูล/Link และรหัส นศ ในกลุ่มตั้งแต่วันที่ 24.00 ของคืนวันที่ 3 มีนาคม และก่อนเวลา 24.00 ของคืนวันที่ 11 มีนาคม เสนอก่อนได้ก่อน และจะอนุมัติข้อมูลทางเว็บภาควิชาเรื่อยๆ ภายในวันที่ 17 มีนาคม ตามตัวอย่างเว็บไซต์ข้อมูลสาธารณะ ต่อไปนี้
 - a. <https://vincentarelbundock.github.io/Rdatasets/datasets.html>
 - b. <https://support.spatialkey.com/spatialkey-sample-csv-data/>
 - c. <http://www.stockhistoricaldata.com/historical-data-formats/csv-data-format>
 - d. <http://www.barchartmarketdata.com/sample-data-feeds>
 - e. [อื่นๆ](#)
3. ศึกษาการใช้งานภาษา R จากตัวอย่างจากลิงค์เหล่านี้
 - a. <http://sciso.sakaeo.buu.ac.th/scisobuusk/wp-content/uploads/2016/09/คู่มือการใช้งานโปรแกรมR-1.pdf>
 - b. <http://www.stou.ac.th/Offices/rdec/Lampang/main/pdf/km55/KM%20pg-r.pdf>
 - c. <https://www.scribd.com/doc/34869064/การใช้-โปรแกรมR>

- d. http://rtutorialseries.blogspot.com/2009/10/r-tutorial-series-introduction-to-r_11.html
 - e. <http://r-statistics.co/Linear-Regression.html>
 - f. อื่นๆ
4. หาค่าพื้นฐานทางสถิติของ คู่ตัวแปรต้นและตัวแปรตามจำนวน 1 คู่ ที่ นศ สนใจ เช่น ค่าเฉลี่ย ค่ามัธยฐาน ค่าฐานนิยม ค่าเบี่ยงเบนมาตรฐาน (Standard Deviation) และ อื่นๆ ในรูปตารางสรุปจากข้อมูลที่ได้รับอนุมัติ
 5. เลือกวาดกราฟ เช่น Histogram, Polygon, Cumulative Histogram, Ogive, Stem & Leave, Box Plot จากข้อ 4 ที่แสดงลักษณะเด่นของข้อมูลได้ชัดเจน แล้วบันทึกเป็นไฟล์ภาพหรือไฟล์แอนิเมชัน
 6. วาดกราฟเส้นตรง Linear Regression ซ้อนบน Scatter Plot ของชุดข้อมูลในข้อ 4 ด้วยคำสั่งใน R Project โดยตัวแปรต้นบนแกนนอน ตัวแปรตามบนแกนตั้ง แล้วบันทึกเป็นไฟล์ภาพหรือไฟล์แอนิเมชัน
 7. จัดทำ PowerPoint หรือ Keynote ความยาว 12-20 แผ่น ประกอบด้วย
 - a. หน้าแรก ประกอบด้วย ชื่อข้อมูล/Link ชื่อ/รหัสนักศึกษา ชื่อ/รหัสวิชา ภาค/ปี การศึกษา ภาควิชา คณะ และสถาบัน
 - b. สารบัญ
 - c. รายละเอียดคร่าวๆ ของข้อมูลว่าเป็นข้อมูลอะไร มีกี่คอลัมน์ เก็บในช่วงเวลาไหน จำนวนเท่าไร เป็นต้น
 - d. บอกชื่อตัวแปรต้น และ ตัวแปรตามที่น่าสนใจ พร้อมตารางและกราฟที่ได้จากข้อ 4-6
 - e. ผลการวิเคราะห์ความเชื่อมโยงระหว่างตัวแปรต้นและตัวแปรตาม que เลือกศึกษา
 8. ทำการ Upload ไฟล์ทาง Google Form นี้ ก่อนเวลา 01.00 ของคืนวันที่ 2 พฤษภาคม 2561 หากส่งช้าปรับคะแนน ชั่วโมงละ 10 คะแนน

การให้คะแนน รวม 100 คะแนน แบ่งเป็น

- | | |
|--|----------|
| ● ปริมาณและความซับซ้อนของข้อมูลที่เสนอเลือก | 20 คะแนน |
| ● ความสมบูรณ์ของการคำนวณและทฤษฎีที่ใช้ | 20 คะแนน |
| ● ความสวยงามและกราฟที่วาด | 20 คะแนน |
| ● ความเหมาะสมของคำอธิบายแต่ละสไลด์เชิงวิชาการ | 20 คะแนน |
| ● ความเหมาะสมของการนำเสนอ เช่น ขนาดตัวอักษร สีเส้น | 20 คะแนน |