



กิจกรรมประจำชุดวิชา
99202 การวิเคราะห์ข้อมูล
ภาคการศึกษาที่ 1 ปีการศึกษา 2566

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยสุโขทัยธรรมมาธิราช

คำนำ

เนื่องด้วยมหาวิทยาลัยสุโขทัยธรรมมาธิราช มุ่งให้ผู้เรียนได้มีส่วนร่วมในกระบวนการศึกษาเล่าเรียนครบวงจร ตั้งแต่ก่อนเรียน ระหว่างเรียน และหลังจากเรียนเสร็จสิ้นไปแล้ว โดยจัดระบบการประเมินครบทั้ง 3 ส่วน ทั้งการประเมินก่อนเรียน ระหว่างเรียน และประเมินผลสุดท้าย

การประเมินกิจกรรมเป็นส่วนหนึ่งของการวัดผลสัมฤทธิ์ทางการเรียนสุดท้าย จึงกำหนดให้นักศึกษาทำกิจกรรมภาคปฏิบัติตามที่กำหนดให้ โดยมีวัตถุประสงค์เพื่อให้นักศึกษามีความสามารถ ดังนี้

1. สรุปหรือประมวลเนื้อหาสาระของเอกสารการสอนทั้งชุดวิชาหรือกลุ่มเนื้อหา กลุ่มใดกลุ่มหนึ่ง
2. ประยุกต์ความรู้จากเอกสารการสอนเพื่อจัดทำโครงการพัฒนางานอย่างใดอย่างหนึ่งที่นักศึกษาทำ
3. พัฒนาระบบ โครงการ ชิ้นงาน ฯลฯ ตามกระบวนการหรือขั้นตอนที่แสดงไว้ในหน่วยใดหน่วยหนึ่งของเอกสารการสอน
4. คิด วิเคราะห์ นำเสนอข้อมูลและความคิดในเชิงสร้างสรรค์

นอกจากนี้การทำกิจกรรมประจำชุดวิชายังทำให้นักศึกษาได้ศึกษาเอกสารการสอนตั้งแต่ต้นภาคการศึกษา และจากการวิจัยพบว่านักศึกษาที่ทำกิจกรรมจะมีโอกาสสอบผ่านในปลายภาคมากกว่านักศึกษาที่ไม่ทำกิจกรรม

คณะกรรมการบริหารชุดวิชาวิเคราะห์ข้อมูล ขอให้นักศึกษาทุกท่านประสบความสำเร็จในการศึกษาชุดวิชานี้ และสามารถนำความรู้ไปเป็นประโยชน์ต่อการดำเนินชีวิตและการประกอบอาชีพสืบไป

คณะกรรมการบริหารชุดวิชา

99202 การวิเคราะห์ข้อมูล

1. การประเมินผล

เกณฑ์การให้คะแนนกิจกรรมพิจารณาจากการตอบที่ตรงประเด็นคำถาม การครอบคลุมความถูกต้องของคำตอบ ความชัดเจนของการนำเสนอ และความละเอียดประณีตของชิ้นงาน

มหาวิทยาลัยไม่บังคับให้นักศึกษาทุกคนต้องทำกิจกรรม นักศึกษาอาจเลือกทำหรือไม่ทำก็ได้ โดยการประเมินปลายภาคสำหรับชุดวิชานี้ แบ่งออกเป็น 2 กรณี

กรณีที่ 1 นักศึกษาทำกิจกรรม ในกรณีนี้มหาวิทยาลัยแบ่งคะแนนออกเป็น 2 ส่วน ส่วนแรกจากคะแนนสอบปลายภาคคิดร้อยละ 80 และส่วนที่สองจากคะแนนกิจกรรมคิดร้อยละ 20 โดยคะแนนกิจกรรมจะนำไปใช้ในการประเมินทั้งการสอบไล่และสอบซ่อม นักศึกษาที่ไม่ได้ส่งกิจกรรมในการสอบไล่ จะส่งกิจกรรมเพื่อเป็นคะแนนในการสอบซ่อมไม่ได้

กรณีที่ 2 นักศึกษาไม่ทำกิจกรรม ในกรณีนี้มหาวิทยาลัยประเมินผลจากการสอบปลายภาคเพียงอย่างเดียว

ในการประเมินผลปลายภาค นักศึกษากลุ่มที่ทำกิจกรรมและไม่ทำกิจกรรม จะได้รับการประเมินผลโดยใช้ข้อสอบฉบับเดียวกัน นักศึกษากลุ่มที่ทำกิจกรรมมีคะแนนเต็ม 80 คะแนน ส่วนนักศึกษากลุ่มที่ไม่ทำกิจกรรมจะมีคะแนนเต็ม 100 คะแนน

สำหรับนักศึกษาที่ทำกิจกรรม มหาวิทยาลัยจะพิจารณาให้นักศึกษาได้ประโยชน์สูงสุด โดยการนำคะแนนสอบปลายภาคของนักศึกษาเพียงอย่างเดียว มาเปรียบเทียบกับการคิดคะแนนสอบปลายภาครวมกับคะแนนกิจกรรม แล้วนำคะแนนส่วนที่มากกว่าไปใช้ในการตัดสินผลการสอบให้แก่นักศึกษา ดังตัวอย่างต่อไปนี้

ตัวอย่างที่ 1 นักศึกษาได้คะแนนกิจกรรม 18 คะแนน และทำข้อสอบได้ 70 ข้อ (คิดเป็น $\frac{70}{120} \times 80$ เท่ากับ 46.67 คะแนน) นักศึกษาจะได้คะแนนกิจกรรมรวมกับคะแนนสอบปลายภาค $18 + 46.67$ เท่ากับ 64.67 คะแนน นั่นคือ “สอบผ่าน” กรณีที่นักศึกษาไม่ได้ทำกิจกรรมประจำชุดวิชา จะคิดคะแนนจากการสอบปลายภาคเพียงอย่างเดียว นักศึกษาทำข้อสอบได้ 70 ข้อ จะได้ $70 \times \frac{100}{120}$ เท่ากับ 58.33 คะแนน นั่นคือ “สอบไม่ผ่าน”

ดังนั้นมหาวิทยาลัยจะเลือกให้นักศึกษาได้คะแนน 64.67 คะแนน นั่นคือ “สอบผ่าน”

ตัวอย่างที่ 2 นักศึกษาได้คะแนนกิจกรรม 13 คะแนน และทำข้อสอบได้ 92 ข้อ (คิดเป็น $\frac{92}{120} \times 80$ เท่ากับ 61.33 คะแนน) นักศึกษาจะได้คะแนนกิจกรรมรวมกับคะแนนสอบปลายภาค $13 + 61.33$ เท่ากับ 74.33 คะแนน กรณีคิดคะแนนจากการสอบปลายภาคเพียงอย่างเดียว นักศึกษาจะได้ $92 \times \frac{100}{120}$ เท่ากับ 76.67 คะแนน

ดังนั้นมหาวิทยาลัยจะเลือกให้นักศึกษาได้คะแนนมากที่สุด ได้แก่ 76.67 คะแนน นั่นคือ “สอบผ่าน”

2. การส่งกิจกรรมประจำชุดวิชา

ให้นักศึกษาดำเนินการดังนี้

1. ให้นักศึกษาส่งกิจกรรมประจำชุดวิชาฉบับจริงไปยังมหาวิทยาลัยและสำเนากิจกรรมที่ทำเสร็จแล้ว เก็บไว้เป็นหลักฐาน 1 ชุด
2. การส่งกิจกรรมประจำชุดวิชาภายใน วันที่ 30 ตุลาคม 2566
3. ให้จัดทำหน้าปกรายงานให้มีข้อความตามตัวอย่างที่แนบมา
4. ส่งกิจกรรมที่ทำเสร็จเรียบร้อยแล้วด้วยตนเอง ณ สำนักบริการการศึกษา มหาวิทยาลัยสุโขทัยธรรมาธิราช หรือส่งทางไปรษณีย์ลงทะเบียน โดยเจ้าหน้าที่ของดังนี้

ศูนย์บริการการสอนทางไปรษณีย์
สำนักบริการการศึกษา
มหาวิทยาลัยสุโขทัยธรรมาธิราช
ตำบลบางพูด อำเภอปากเกร็ด
จังหวัดนนทบุรี 11120

(กิจกรรมประจำชุดวิชา 99202 การวิเคราะห์ข้อมูล
สาขาวิชาวิทยาศาสตร์และเทคโนโลยี)

หรือส่งผ่านช่องทางออนไลน์ของเว็บไซต์ที่มหาวิทยาลัยกำหนดไว้

5. นักศึกษาสามารถตรวจสอบว่าสำนักบริการการศึกษาได้รับกิจกรรมที่นักศึกษาส่งไปแล้วหรือยัง โดยโทรศัพท์สอบถามที่หมายเลข 0-2982-9633 หรือโทรศัพท์ติดต่อสำนักบริการการศึกษา หมายเลข 0-2-504-7621 หรือโทรศัพท์ติดต่อศูนย์สารสนเทศ หมายเลข 0-2504-7788 หรือ e-mail: ic.proffice@stou.ac.th หรือตรวจสอบผ่านระบบสารสนเทศได้ที่ <https://regis.stou.ac.th/STOU/login.jsp>

3. เนื้อหากิจกรรม

คำชี้แจง ให้นักศึกษาตอบคำถามใน 15 กิจกรรมต่อไปนี้ในกระดาษ A4 โดยให้นักศึกษาทำกิจกรรมด้วยตนเอง และเขียนด้วยลายมือตนเองเท่านั้น ถ้าตรวจสอบได้ว่าการลอกกันหรือไม่ได้ใช้ความรู้ของตนเองในการตอบคำถาม จะไม่ตรวจให้คะแนนไม่ว่ากรณีใดๆ กิจกรรมประจำชุดวิชานี้ มีจำนวน 15 ข้อ คะแนนเต็ม 60 คะแนน

กิจกรรมที่ 1 หน่วยที่ 1 ข้อมูลและการวิเคราะห์ข้อมูล

ข้อที่ 1.1 การแบ่งข้อมูลตามมาตราการวัด (measurement scales) สามารถแบ่งได้กี่ระดับ อะไรบ้าง อธิบาย พร้อมยกตัวอย่างมาพอสังเขป (2 คะแนน)

ข้อที่ 1.2 จงอธิบายเปรียบเทียบความแตกต่างระหว่างการวิเคราะห์แบบพหุการณ และการวิเคราะห์แบบให้คำแนะนำมาพอสังเขป (2 คะแนน)

กิจกรรมที่ 2 หน่วยที่ 2 วิทยาศาสตร์ข้อมูลและนักวิทยาศาสตร์ข้อมูล

ข้อที่ 2.1 จงอธิบายศาสตร์ที่เกี่ยวข้องกับวิทยาศาสตร์ข้อมูลมาพอสังเขป (2 คะแนน)

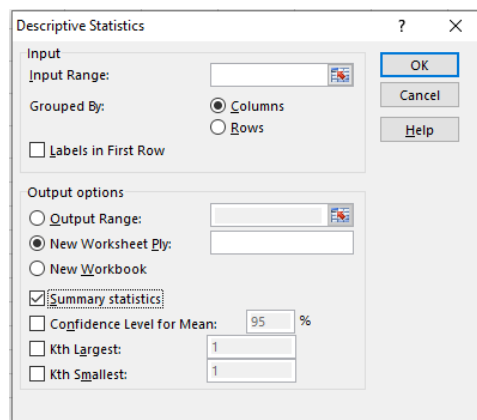
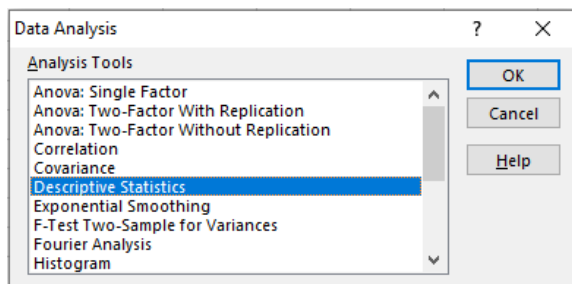
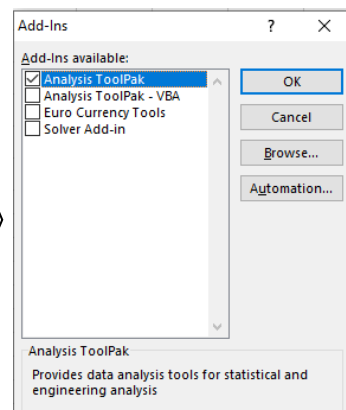
ข้อที่ 2.2 กระบวนการวิทยาศาสตร์ข้อมูลคืออะไร มีกี่ขั้นตอน อะไรบ้าง พร้อมทั้งอธิบายหน้าที่การทำงานของแต่ละขั้นตอนมาพอเข้าใจ (2 คะแนน)

กิจกรรมที่ 3 หน่วยที่ 3 เครื่องมือในการวิเคราะห์ข้อมูล

ข้อที่ 3.1 จงยกตัวอย่างความสามารถด้านการวิเคราะห์ข้อมูลของโปรแกรมภาษาอาร์ (R) และโปรแกรมภาษาไพธอน (Python) มาโปรแกรมละ 3 ข้อ (2 คะแนน)

ข้อที่ 3.2 หากต้องการวิเคราะห์ข้อมูลด้านล่าง ด้วยโปรแกรม MS Excel ในการหาค่าสถิติพื้นฐาน โดยการเลือก Analysis Toolpak และ Descriptive Statistics เมื่อคลิกให้แสดงข้อมูล Summary statistics จะแสดงข้อมูลทางสถิติใดบ้าง (2 คะแนน)

	A	B	C	D
1	ตารางแสดงยอดขายสินค้าของร้านปาริชาติ			
2	วันที่	สินค้า	จำนวน	ยอดขาย
3	2/7/2564	คุกกี้	24	840.00
4	3/7/2564	คุกกี้	35	1,225.00
5	9/7/2564	บราวนี่	46	1,840.00
6	11/7/2564	ครัวซองต์	28	1,260.00
7	19/7/2564	ครัวซองต์	59	2,655.00



กิจกรรมที่ 4 หน่วยที่ 4 การเก็บรวบรวมและการเตรียมข้อมูลก่อนการประมวลผล

ข้อที่ 4.1 การสุ่มตัวอย่างโดยไม่ใช้ความน่าจะเป็น มีลักษณะเป็นอย่างไร แบ่งออกเป็นกี่ประเภท อะไรบ้าง อธิบายแต่ละประเภทมาพอสังเขป (2 คะแนน)

ข้อที่ 4.2 อธิบายเทคนิคการแปลงข้อมูลหรือการปรับเปลี่ยนรูปแบบข้อมูล (data transformation) พร้อมยกตัวอย่างประกอบ (2 คะแนน)

กิจกรรมที่ 5 หน่วยที่ 5 จินตทัศน์ข้อมูล

ข้อที่ 5.1 ระบบจินตทัศน์ข้อมูล (data visualization) ที่มีความน่าสนใจ ควรมีคุณสมบัติอะไรบ้าง บอกมาอย่างน้อย 3 ประการ (2 คะแนน)

ข้อที่ 5.2 แผนภาพวิเคราะห์ข้อความเหมาะกับการนำมาใช้งานในลักษณะใด อธิบายพร้อมยกตัวอย่างประกอบ (2 คะแนน)

กิจกรรมที่ 6 หน่วยที่ 6 สถิติพื้นฐาน

ข้อที่ 6. ผลการสอบวิชาการวิเคราะห์ข้อมูลของนักศึกษาวิชาเอกเทคโนโลยีสารสนเทศและการสื่อสาร กลุ่มที่ 2 จำนวน 10 คน ได้คะแนนสอบดังต่อไปนี้ 51, 75, 42, 53, 76, 69, 36, 75, 49 และ 80 จงหาค่าต่อไปนี้ พร้อมแสดงวิธีการคำนวณโดยใช้สูตร (4 คะแนน)

6.1 จงหาค่าเปอร์เซ็นต์ไทล์ที่ 75

6.2 จงหาค่าเฉลี่ยเลขคณิตของคะแนนสอบ

6.3 จงหาค่ามัธยฐาน

6.4 จงหาค่าฐานนิยม

กิจกรรมที่ 7 หน่วยที่ 7 การวิเคราะห์ข้อมูลทางสถิติด้วยการทดสอบค่าทีและการวิเคราะห์ความแปรปรวน

ข้อที่ 7. ในการสอบวิชาฐานข้อมูลเบื้องต้น อาจารย์กำหนดเกณฑ์มาตรฐานคือ 60 คะแนน พบว่า คะแนนสอบเฉลี่ยของนักศึกษา 25 คน คือ 64 คะแนน ส่วนเบี่ยงเบนมาตรฐานคือ 2 จงหาค่า t (4 คะแนน)

กิจกรรมที่ 8 หน่วยที่ 8 การวิเคราะห์ค่าสหสัมพันธ์และการวิเคราะห์การถดถอย

ข้อที่ 8. จากตารางการวิเคราะห์ความแปรปรวนของการวิเคราะห์การถดถอยเชิงเส้นด้านล่าง จงคำนวณหาค่าต่อไปนี้ (4 คะแนน)

แหล่งความแปรปรวน (Sources)	องศาอิสระ (df)	Sum Square (SS)	Mean Square (MS)	F
การถดถอย (Regression)	1	SSR = 105	MSR = ?	MSR/MSE = ?
ความคลาดเคลื่อน (Error)	10	SSE = 15	MSE = ?	
ผลรวม (Total)	11	SST = ?		

8.1 SST _____

8.2 MSR _____

8.3 MSE _____

8.4 $F = MSR/MSE$ _____

กิจกรรมที่ 9 หน่วยที่ 9 การวิเคราะห์ข้อมูลเชิงคุณภาพ

ข้อที่ 9. เทคนิคการเก็บรวบรวมข้อมูลเชิงคุณภาพ มีกี่วิธี อะไรบ้าง อธิบายลักษณะของข้อมูลที่ได้จากการเก็บข้อมูลแต่ละวิธีมาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 10 หน่วยที่ 10 การเรียนรู้ของเครื่องแบบมีผู้สอน

ข้อที่ 10. บอกอัลกอริทึมการจำแนกข้อมูลแบบมีผู้สอนที่นิยมใช้มาอย่างน้อย 2 อัลกอริทึม อธิบายวิธีการทำงาน พร้อมทั้งระบุข้อดีและข้อจำกัดของแต่ละอัลกอริทึมมาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 11 หน่วยที่ 11 การเรียนรู้ของเครื่องแบบไม่มีผู้สอน

ข้อที่ 11. อธิบายลักษณะการทำงานของอัลกอริทึมการจัดกลุ่มข้อมูลต่อไปนี้มาพอเข้าใจ (4 คะแนน)

- เคมีนส์ (K-means)

- ไฮราคิเคิลคลัสเตอร์ริง (hierarchical clustering)

กิจกรรมที่ 12 หน่วยที่ 12 ข้อมูลขนาดใหญ่

ข้อที่ 12. คุณลักษณะของข้อมูลขนาดใหญ่ ด้วยหลักการ 5V มีอะไรบ้าง จงอธิบายมาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 13 หน่วยที่ 13 เทคโนโลยีและเครื่องมือในการวิเคราะห์ข้อมูลขนาดใหญ่

ข้อที่ 13. จงอธิบายความหมายและคุณสมบัติของฐานข้อมูลโนเอสคิวแอล (NoSQL) มาพอเข้าใจ (4 คะแนน)

กิจกรรมที่ 14 หน่วยที่ 14 ตัวอย่างการวิเคราะห์ข้อมูลทางธุรกิจ

ข้อที่ 14. วิธีการวิเคราะห์ข้อมูลทางธุรกิจ ตามกระบวนการขั้นตอนของ CRISP-DM มีกี่ขั้นตอน อะไรบ้าง อธิบายมาพอสังเขป (4 คะแนน)

กิจกรรมที่ 15 หน่วยที่ 15 การประยุกต์ใช้เครื่องมือในการวิเคราะห์ข้อมูลทางธุรกิจ

ข้อที่ 15. อธิบายขั้นตอนการสร้างโมเดลในกรณีศึกษาการประยุกต์ใช้เครื่องมือในการวิเคราะห์ข้อมูลสำหรับการ
หากฎความสัมพันธ์ (Association Rule) ด้วยโปรแกรม RapidMiner มาพอสังเขป (4 คะแนน)

This image shows a blank sheet of white paper with horizontal ruling lines. The lines are evenly spaced and extend across the width of the page. There are no margins, text, or other markings on the paper.

ปกรายงาน

กิจกรรมประจำชุดวิชา
99202 การวิเคราะห์ข้อมูล
ภาคการศึกษาที่ 1 ปีการศึกษา 2566

สาขาวิชาวิทยาศาสตร์และเทคโนโลยี

ชื่อนักศึกษา.....

รหัสประจำตัวนักศึกษา

--	--	--	--	--	--	--	--	--	--

ที่อยู่

.....

โทรศัพท์ (ถ้ามี)

ข้าพเจ้าขอยอมรับการตัดสินผลคะแนนภาคปฏิบัติจากผู้ประเมินเป็นที่สุด

ลงชื่อ.....

(.....)