



มาตรฐานอาชีพและคุณวุฒิวิชาชีพ
Occupational Standard and Professional Qualifications

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล

จัดทำโดย สถาบันคุณวุฒิวิชาชีพ (องค์การมหาชน)

1. ชื่อมาตรฐานอาชีพ

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล

2. ประวัติการปรับปรุงมาตรฐาน

N/A

3. ทะเบียนอ้างอิง (Imprint)

N/A

4. ข้อมูลเบื้องต้น

มาตรฐานอาชีพและคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล มุ่งเน้นเพื่อส่งเสริมและสนับสนุนให้กลุ่มสาขาอาชีพจัดทำมาตรฐานอาชีพและคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล ให้สอดคล้องกับความต้องการของผู้ประกอบการ มีความเป็นสากลและเหมาะสมกับประเทศไทย เป็นที่ยอมรับทั้งภายในประเทศและระดับสากล โดยเฉพาะกลุ่มประเทศอาเซียน เพื่อรองรับประชาคมเศรษฐกิจอาเซียน เพื่อสร้างเครือข่ายการจัดทำ พัฒนา และเผยแพร่ มาตรฐานอาชีพและคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล และประชาสัมพันธ์ให้ระบบคุณวุฒิวิชาชีพและมาตรฐานอาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล เป็นที่รับรู้และยอมรับในทุกภาคส่วน มาตรฐานอาชีพและคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล ฉบับนี้จัดทำโดยการวิเคราะห์ข้อมูลเกี่ยวกับสาขาวิชาชีพวิทยาศาสตร์ข้อมูล ทั้งในประเทศและต่างประเทศ ร่วมกับวิธีเทคนิควิเคราะห์หน้าที่ (Functional Analysis) โดยจากการวิเคราะห์พบว่า มีหน้าที่หลัก (Key Function) 5 หน้าที่หลัก ซึ่งทุกหน้าที่งานมีความสำคัญต่อการปฏิบัติงานด้านวิทยาศาสตร์ข้อมูลทั้งสิ้น ทั้งนี้ในแต่ละหน้าที่งานยังประกอบไปด้วยหลากหลายอาชีพที่เกี่ยวข้อง อย่างไรก็ตามในมาตรฐานฉบับนี้ได้ดำเนินการจัดทำอาชีพที่มีความสำคัญเร่งด่วนเพื่อตอบสนองความต้องการของประเทศ โดยได้ดำเนินการจัดทำมาตรฐานอาชีพและคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล ทั้งสิ้น 5 อาชีพ 13 คุณวุฒิวิชาชีพตามรายละเอียดที่อยู่ในมาตรฐานอาชีพนี้

5. ประวัติการปรับปรุงมาตรฐานในแต่ละครั้ง

N/A

6. ครั้งที่

ครั้งที่ 1

7. คุณวุฒิวิชาชีพที่ครอบคลุม (Professional Qualifications included)

สาขาวิชาชีพอุตสาหกรรมดิจิทัล

สาขาวิทยาศาสตร์ข้อมูล

อาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3

8. คุณวุฒิวิชาชีพที่เกี่ยวข้อง (Related Professional Qualifications)

N/A

9. หน่วยสมรรถนะทั้งหมดในมาตรฐานอาชีพ (List of All Units of Competence within this Occupational Standards)

รหัสหน่วยสมรรถนะ	เนื้อหา
70204	สำรวจข้อมูล
70301	เลือกข้อมูลที่มีโครงสร้าง
70302	แปลงข้อมูลที่มีโครงสร้าง
70305	จัดรูปแบบข้อมูล

10. ระดับคุณวุฒิ

10.1 สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล อาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3

คุณลักษณะของผลการเรียนรู้ (Characteristics of Outcomes)

บุคคลที่มีคุณลักษณะของผลการเรียนรู้ที่พึงประสงค์ในอาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3 จะมีความรู้เกี่ยวกับ การวิเคราะห์และการประมวลผลแบบจำลอง (Analytics and Computational Modelling) การวิเคราะห์ความต้องการที่จำเป็นขององค์กร (Business Needs Analysis) การออกแบบ บริหารจัดการ และวิศวกรรมข้อมูล การบูรณาการระบบและข้อมูล การนำเสนอ รวมถึงการตระหนักถึงการศึกษาความปลอดภัย วิศวกรข้อมูลจะสนับสนุนการออกแบบ การพัฒนา และให้การบำรุงรักษาของเส้นทางกระแสข้อมูล และระบบการประมวลผลข้อมูลที่สนับสนุนการรวบรวมและการจัดเก็บข้อมูล ทั้งการประมวลผลแบบ Batch และ Real-time และทำการวิเคราะห์ข้อมูลข่าวสาร (Information) รวมถึงข้อมูลทั้งที่มีโครงสร้างที่ชัดเจน (Structured) ในลักษณะที่สามารถปรับขนาดได้ (Scalable) สามารถทำซ้ำได้ (Repeatable) และมีความปลอดภัย (Secure) โดยมีเป้าหมายหลักในการกำหนดแนวทางที่ดีที่สุดในการเก็บรวบรวมข้อมูล การประมวลผล รวมถึงการทำคลังข้อมูล และออกแบบแนวการเขียนภาษาโปรแกรมและทดสอบชุดข้อมูล ตลอดจนพัฒนาตามโครงสร้างพื้นฐานที่มี ซึ่งการรวบรวม การจัดการข้อมูล การวิเคราะห์และการนำเสนอข้อมูลขนาดใหญ่ขึ้นเพื่อการค้นหาความหมายเชิงลึก (Insights) จากข้อมูล (Information) ที่สามารถเข้าถึงจากแพลตฟอร์มที่หลากหลายได้ วิศวกรข้อมูลขอการทำงานเกี่ยวกับตัวเลขและการทำงานร่วมกับชุดข้อมูลขนาดใหญ่ มีความกระตือรือร้นในการทำความเข้าใจกระบวนการทางธุรกิจและการแก้ปัญหาที่เชื่อมโยงระหว่างฐานข้อมูลและสถาปัตยกรรมองค์กรสำหรับข้อมูลที่มีโครงสร้าง (Structured Data) มีเทคนิคในการประยุกต์หลักการในการแก้ปัญหาทางเทคนิคในงานควบคู่กับการใช้คู่มือ สามารถใช้สารสนเทศเพื่อควบคุมคุณภาพงานได้ โดยต้องมีความรับผิดชอบต่องานในหน้าที่ ให้การสนับสนุนผู้ร่วมงานตัดสินใจแก้ปัญหาในงานและรายงานผลการทำงานอย่างต่อเนื่อง และมีจริยธรรมในการประกอบอาชีพ

การเลื่อนระดับคุณวุฒิวิชาชีพ (Qualification Pathways)

- 1. คุณสมบัติของผู้ที่สามารถเข้ารับการประเมินคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล อาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3
 - ผู้ที่สำเร็จการศึกษาระดับประกาศนียบัตรวิชาชีพ (ปวช.) หรือ
 - ผู้ที่สำเร็จการศึกษาระดับมัธยมศึกษาตอนปลาย
- 2. ผู้ที่จะผ่านการประเมินและได้รับการรับรองคุณวุฒิวิชาชีพ สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล อาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3
 - ผ่านเกณฑ์การประเมินตามหน่วยสมรรถนะของอาชีพผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล ระดับ 3 จำนวน 4 หน่วย

หลักเกณฑ์การต่ออายุหนังสือรับรองมาตรฐานอาชีพ

N/A

กลุ่มบุคคลในอาชีพ (Target Group)

ผู้ทำงานในกลุ่มอุตสาหกรรมเทคโนโลยีสารสนเทศและการสื่อสารและดิจิทัลคอนเทนต์ ด้านวิทยาศาสตร์ข้อมูล เช่น นักวิเคราะห์ข้อมูล นักวิเคราะห์สถิติ นักการตลาด วิศวกรข้อมูล บุคคลที่สำเร็จการศึกษาในสาขาที่เกี่ยวข้อง หรือบุคคลที่สนใจในงานด้านวิทยาศาสตร์ข้อมูล

หน่วยสมรรถนะ (หน่วยสมรรถนะทั้งหมดของคุณวุฒิวิชาชีพนี้)

- 70204 สืบรวจข้อมูล
- 70301 เลือกข้อมูลที่มีโครงสร้าง
- 70302 แปลงข้อมูลที่มีโครงสร้าง
- 70305 จัดรูปแบบข้อมูล

ตารางแผนผังแสดงหน้าที่

1. ตารางแสดงหน้าที่ 1

ประกาศใช้ ณ 23/05/2566

ตาราง 1 : FUNCTIONAL MAP แสดง KEY PURPOSE , KEY ROLES , KEY FUNCTION

ความมุ่งหมายหลัก Key Purpose	บทบาทหลัก Key Roles		หน้าที่หลัก Key Function	
คำอธิบาย	รหัส	คำอธิบาย	รหัส	คำอธิบาย
พัฒนาศักยภาพของบุคลากรในสาขาอาชีพ ICT ให้สามารถแข่งขันและเป็นที่ยอมรับในระดับสากล	70	ปฏิบัติงานด้านวิทยาศาสตร์ข้อมูลได้ตามมาตรฐานอาชีพ	702	รวบรวมและศึกษาทำความเข้าใจข้อมูล
			703	จัดเตรียมข้อมูล

คำอธิบาย ตารางแผนผังแสดงหน้าที่เป็นแผนผังที่ใช้วิเคราะห์หน้าที่งานเพื่อให้ได้หน้าที่หลัก (Key Function)

2. ตารางแสดงหน้าที่ 1 (ต่อ)

ประกาศใช้ ณ 23/05/2566

ตาราง 2 : FUNCTIONAL MAP แสดง KEY FUNCTION , UNIT OF COMPETENCE , ELEMENT OF COMPETENCE

หน้าที่หลัก Key Function		หน่วยสมรรถนะ Unit of Competence		หน่วยสมรรถนะย่อย Element of Competence	
รหัส	คำอธิบาย	รหัส	คำอธิบาย	รหัส	คำอธิบาย
702	รวบรวมและศึกษาทำความเข้าใจข้อมูล	70204	สำรวจข้อมูล	70204.01	ตรวจสอบคำนิยามและความหมายของข้อมูลที่ได้รับ
				70204.02	ตรวจสอบคุณลักษณะข้อมูลที่ได้รับ
				70204.03	สรุปผลการสำรวจข้อมูล
				70204.04	สื่อสารในระหว่างการสำรวจข้อมูล
703	จัดเตรียมข้อมูล	70301	เลือกข้อมูลที่มีโครงสร้าง	70301.01	เข้าถึงข้อมูลที่มีของธุรกิจแบบมีโครงสร้าง (Structured Data)
				70301.02	เลือกข้อมูล (Select Data) ขององค์กรที่ต้องการใช้วิเคราะห์ที่มีโครงสร้าง
				70301.03	ปรับปรุงคุณภาพข้อมูล (Clean Data) ที่มีโครงสร้าง
		70302	แปลงข้อมูลที่มีโครงสร้าง	70302.01	ปรับปรุงคุณสมบัติข้อมูลเพื่อการวิเคราะห์แบบมีโครงสร้าง (Construct Data)
				70302.02	ควมรวมข้อมูล (Integrate Data) แบบมีโครงสร้าง
		70305	จัดรูปแบบข้อมูล	70305.01	ระบุข้อมูลและแบบจำลอง
				70305.02	จัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes)
				70305.03	จัดเรียงรายการข้อมูลใหม่ (Reordering Records)
				70305.04	เปลี่ยนค่าข้อมูล (Reformatted within-value)

คำอธิบาย

ตารางแผนผังแสดงหน้าที่ (ต่อ) เป็นแผนผังที่ใช้วิเคราะห์หน้าที่งานหลังจากได้หน้าที่หลัก (Key Function) เพื่อให้ได้ หน่วยสมรรถนะ (Unit of Competence) และหน่วยสมรรถนะย่อย (Element of Competence)

1. รหัสหน่วยสมรรถนะ

70204
2. ชื่อหน่วยสมรรถนะ

สำรวจข้อมูล
3. ทบทวนครั้งที่

1 / 2566
4. สร้างใหม่

☐

ปรับปรุง

☒

5. สำหรับชื่ออาชีพและรหัสอาชีพ (Occupational Classification)

นักวิเคราะห์ข้อมูล (Data Analyst) และผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล

6. คำอธิบายหน่วยสมรรถนะ (Description of Unit of Competency)

ผู้ผ่านสมรรถนะนี้จะมีความรู้เกี่ยวกับการสำรวจข้อมูล ทั้งความหมายและคุณลักษณะของข้อมูล โดยสามารถตรวจสอบความหมายของข้อมูลที่ได้รับโดยใช้ความรู้ทางธุรกิจ สามารถตรวจสอบคุณสมบัติพื้นฐานของข้อมูล และทำความเข้าใจเกี่ยวกับการกระจายของคุณลักษณะที่สำคัญของข้อมูล และเข้าใจความสัมพันธ์ข้อมูล

7. สำหรับระดับคุณวุฒิ

1	2	3	4	5	6	7	8
☐	☐	☒	☐	☐	☐	☐	☐

8. กลุ่มอาชีพ (Sector)

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล (Data Science)

9. ชื่ออาชีพและรหัสอาชีพอื่นที่หน่วยสมรรถนะนี้สามารถใช้ได้ (ถ้ามี)

N/A

10. ข้อกำหนดหรือกฎระเบียบที่เกี่ยวข้อง (Licensing or Regulation Related) (ถ้ามี)

N/A

11. สมรรถนะย่อยและเกณฑ์การปฏิบัติงาน (Elements and Performance Criteria)

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70204.01 ตรวจสอบคำนิยามและความหมายของข้อมูลที่ได้รับ	1. ระบุคำนิยามของข้อมูลได้ 2. ระบุความหมายของข้อมูลได้ 3. สรุปคำนิยามและความหมายของข้อมูลได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน
70204.02 ตรวจสอบคุณลักษณะข้อมูลที่ได้รับ	1. ระบุคุณลักษณะของข้อมูลที่ได้รับได้ 2. ระบุรายละเอียดคุณลักษณะของข้อมูลที่ได้รับได้ 3. สรุปคุณลักษณะของข้อมูลที่ได้รับได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน
70204.03 สรุปผลการสำรวจข้อมูล	1. ระบุรายละเอียดผลการสำรวจข้อมูลได้ 2. วิเคราะห์คุณลักษณะที่สำคัญของข้อมูลที่ได้รับได้ 3. ระบุรายละเอียดผลการวิเคราะห์รายละเอียดคุณลักษณะที่สำคัญของข้อมูลที่ได้รับได้ 4. สรุปผลการวิเคราะห์ละเอียดคุณลักษณะที่สำคัญของข้อมูลที่ได้รับได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน
70204.04 สื่อสารในระหว่างการสำรวจข้อมูล	1. รับและส่งสารในระหว่างการสำรวจข้อมูลได้ 2. สอบถามและตอบข้อมูลในระหว่างการสำรวจข้อมูลได้ 3. สามารถสื่อสารผ่านสื่อสารสนเทศช่องทางต่าง ๆ ได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน

12. ความรู้และทักษะก่อนหน้าที่จำเป็น (Pre-requisite Skill & Knowledge)

N/A

13. ทักษะและความรู้ที่ต้องการ (Required Skills and Knowledge)

(ก) ความต้องการด้านทักษะ

1. ทักษะในการสำรวจข้อมูล
2. ทักษะในการวิเคราะห์ข้อมูล และลักษณะของข้อมูล
3. ทักษะในการวิเคราะห์ทางสถิติทั่วไป และการวิเคราะห์เชิงบรรยาย (Statistical analysis and Descriptive analytics)
4. ทักษะในการวิเคราะห์เชิงปริมาณ (Quantitative Analytics)
5. ทักษะการใช้เครือข่ายทางสังคมและแหล่งข้อมูลแบบเปิด (Social network and open data)

(ข) ความต้องการด้านความรู้

1. ความรู้เกี่ยวกับภาพรวมธุรกิจ
2. ความรู้เกี่ยวกับการใช้ Business Analytics (BA) และ Business Intelligence (BI) ในการวิเคราะห์ข้อมูลโดยใช้เทคโนโลยีและการให้บริการที่เกี่ยวข้อง
3. ความรู้เกี่ยวกับการสำรวจข้อมูล
4. ความรู้เกี่ยวกับการวิเคราะห์ทางสถิติทั่วไป และการวิเคราะห์เชิงบรรยาย (Statistical analysis and Descriptive analytics)
5. ความรู้เกี่ยวกับการวิเคราะห์เชิงปริมาณ (Quantitative Analytics)
6. ความรู้เกี่ยวกับคุณลักษณะของข้อมูล
7. ความรู้เกี่ยวกับการใช้เครือข่ายทางสังคมและแหล่งข้อมูลแบบเปิด (Social network and open data)

14. หลักฐานที่ต้องการ (Evidence Guide)

(ก) หลักฐานการปฏิบัติงาน (Performance Evidence)

1. ใบรับรองการปฏิบัติงานจากสถานประกอบการ

(ข) หลักฐานความรู้ (Knowledge Evidence)

1. ใบรับรองการเข้ารับการฝึกอบรม
2. ใบประกาศนียบัตรวุฒิการศึกษา

(ค) คำแนะนำในการประเมิน

1. ผู้ประเมินตรวจประเมินเกี่ยวกับการสำรวจข้อมูล โดยพิจารณาจากร่องรอยหลักฐานที่เกี่ยวข้องทั้งหลักฐานการปฏิบัติงาน และหลักฐานความรู้

(ง) วิธีการประเมิน

1. พิจารณาตามหลักฐานการปฏิบัติงาน
2. พิจารณาตามหลักฐานความรู้

15. ขอบเขต (Range Statement)

(ก) คำแนะนำ

หน่วยสมรรถนะนี้เป็นการทดสอบ ประเมินการกำหนดวัตถุประสงค์ทางธุรกิจ โดยในการประเมินต้องคำนึงถึงข้อปฏิบัติดังต่อไปนี้

1. ผู้เข้ารับการประเมินสามารถแสดงความรู้ และความสามารถในการตรวจสอบค่านิยมและความหมายของข้อมูลที่ได้รับ การตรวจสอบคุณสมบัติพื้นฐานของข้อมูลที่ได้รับความสนใจเกี่ยวกับคุณลักษณะของข้อมูลที่ได้รับ

(ข) คำอธิบายรายละเอียด

การสำรวจข้อมูล หมายถึง การค้นและวิเคราะห์ข้อมูลจากศูนย์ข้อมูลหรือฐานข้อมูล

ซึ่งเป็นการตรวจสอบภาพรวมของข้อมูลที่มีและต้องทำความเข้าใจคุณลักษณะของข้อมูลให้ชัดเจน เพื่อให้ข้อมูลมีความถูกต้อง และมีคุณภาพ สำหรับการวิเคราะห์ข้อมูลเพื่อแก้ไข ปรับปรุง หรือพัฒนาธุรกิจหรืออุตสาหกรรมได้และมีประสิทธิภาพ

โดยมีประเภทของการสำรวจข้อมูล เช่น การสอบถามข้อมูลเชิงคุณลักษณะ (Attribute data query) หมายถึงการค้นข้อมูลโดยใช้ นิพจน์ทางตรรกศาสตร์ และการสอบถามข้อมูลเชิงพื้นที่ (Spatial data query) เป็นการสอบถามข้อมูลจากแผนที่ โดยการใช้คำสั่งสืบค้นข้อมูลจากระดับข้อมูล เป็นต้น

ขั้นตอนในการดำเนินการสำรวจข้อมูล 4 ขั้นตอนหลักๆ ดังนี้

- 1) ขั้นตอนการเตรียมงานและวางแผน (Planning) ซึ่งจะรวมเรื่องการสร้างแบบสอบถาม และการเลือกหน่วยตัวอย่างไว้ในขั้นตอนนี้
- 2) ขั้นตอนการเก็บรวบรวมข้อมูล (Collecting)
- 3) ขั้นตอนการประมวลผล (Processing & Analysis)
- 4) ขั้นตอนการนำเสนอผลการสำรวจ และจัดทำรายงาน (Presenting & Reporting)

ประเภทของข้อมูล (Data Type) ในการที่เราจะศึกษาเกี่ยวกับข้อมูล เราจำเป็นอย่างยิ่งที่จะต้องแยกประเภทของข้อมูลให้ได้ก่อน

เพื่อที่จะทำความเข้าใจธรรมชาติของข้อมูลประเภทนั้นๆ และเพื่อให้เราสามารถที่จะเลือกใช้เครื่องมือทางสถิติที่เหมาะสมในการวิเคราะห์หรืออธิบายข้อมูลนั้นๆ

เราสามารถแบ่งประเภทของข้อมูลได้ดังต่อไปนี้

1. ข้อมูลเชิงคุณภาพ (Qualitative or Attribute Data) เป็นข้อมูลที่ไม่สามารถวัดด้วยอุปกรณ์การวัดใดๆ ได้เลย แต่ได้มาจากการนับ การสังเกต เราจึงนิยามชนิดของข้อมูลที่ได้ว่าเป็นข้อมูลเชิงคุณภาพ แบ่งได้เป็นสองลักษณะย่อย ได้แก่แบ่งโดยใช้ลักษณะนาม (Nominal Scale) และการแบ่งโดยอันดับ (Ordinal Scale)
 2. ข้อมูลเชิงปริมาณ (Quantitative or Variable data) คือข้อมูลที่มีลักษณะเป็นตัวเลข สามารถแบ่งออกได้อีกเป็นสองชนิดย่อย ดังนี้
 - ข้อมูลแบบไม่ต่อเนื่อง (Discrete Data) ลักษณะที่สำคัญคือ เป็นตัวเลขที่ได้จากการนับ ค่าทศนิยมมีค่าได้จำกัด หรือไม่มีความหมาย
 - ข้อมูลแบบต่อเนื่อง (Continuous Data) ลักษณะที่สำคัญคือ ต้องใช้เครื่องมือวัดมา ค่าทศนิยมมีความหมายและไม่มีที่สิ้นสุด
 ขึ้นอยู่กับความละเอียดของเครื่องวัดว่าจะแยกแยะออกเป็นทศนิยมได้กี่ตำแหน่ง
- โดยในเทคนิคการวิเคราะห์และนำเสนอข้อมูล ประกอบไปด้วย
- 1) การวัดแนวโน้มเข้าสู่ศูนย์กลาง (Measure of location)
 - 2) การวัดการกระจายของข้อมูล (Measure of Dispersion)
 - 3) การนำเสนอข้อมูลด้วยแผนภูมิ (Graphical Data Presentation)
 - 4) การทดสอบความเป็นการกระจายแบบปกติ (Normality Test)

16. หน่วยสมรรถนะรวม (ถ้ามี)

N/A

17. อุตสาหกรรมร่วม/กลุ่มอาชีพร่วม (ถ้ามี)

N/A

18. รายละเอียดกระบวนการและวิธีการประเมิน (Assessment Description and Procedure)

18.1 เครื่องมือประเมินการตรวจสอบค่านิยมและความหมายของข้อมูลที่ได้รับตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการบริหารจัดการปฏิบัติงาน
2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.2 เครื่องมือประเมินการตรวจสอบคุณลักษณะข้อมูลที่ได้รับตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการบริหารจัดการปฏิบัติงาน
2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.3 เครื่องมือประเมินการสรุปผลการสำรวจข้อมูลตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการบริหารจัดการปฏิบัติงาน
2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.4 เครื่องมือประเมินการสื่อสารในระหว่างการสำรวจข้อมูลตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการบริหารจัดการปฏิบัติงาน
2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

1. รหัสหน่วยสมรรถนะ70301
2. ชื่อหน่วยสมรรถนะเลือกข้อมูลที่มีโครงสร้าง
3. ทบทวนครั้งที่1 / 2566
4. สร้างใหม่

☐

ปรับปรุง

☒

5. สำหรับชื่ออาชีพและรหัสอาชีพ (Occupational Classification)

ผู้ปฏิบัติงานด้านวิศวกรรมข้อมูล

6. คำอธิบายหน่วยสมรรถนะ (Description of Unit of Competency)

ผู้ที่ผ่านสมรรถนะนี้จะมีความรู้เกี่ยวกับการเลือกข้อมูลที่มีโครงสร้าง สามารถเลือกข้อมูลทั้งหมดหรือบางส่วนของข้อมูล (Select Data) จากข้อมูลที่มีโครงสร้างข้อมูลแบบฐานข้อมูลเชิงสัมพันธ์ (RDBMS) พร้อมทั้งเหตุผลและคำอธิบายประกอบการเลือกทั้งส่วนที่ใช้และไม่ใช้ข้อมูลที่ต้องการ รวมถึงการเตรียมข้อมูลเพื่อนำข้อมูลที่ไม่สมบูรณ์ออก ปรับคุณภาพของข้อมูล (Clean Data) ซึ่งข้อมูลและการดำเนินการมีความสัมพันธ์สอดคล้องกับเป้าหมายธุรกิจ

7. สำหรับระดับคุณวุฒิ

1	2	3	4	5	6	7	8
☐	☐	☒	☐	☐	☐	☐	☐

8. กลุ่มอาชีพ (Sector)

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล (Data Science)

9. ชื่ออาชีพและรหัสอาชีพอื่นที่หน่วยสมรรถนะนี้สามารถใช้ได้ (ถ้ามี)

N/A

10. ข้อกำหนดหรือกฎระเบียบที่เกี่ยวข้อง (Licensing or Regulation Related) (ถ้ามี)

N/A

11. สมรรถนะย่อยและเกณฑ์การปฏิบัติงาน (Elements and Performance Criteria)

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70301.01 เข้าถึงข้อมูลที่มีของธุรกิจแบบมีโครงสร้าง (Structured Data)	1. ระบุนิธการดำเนินการเพื่อเข้าถึงข้อมูลที่มีของธุรกิจได้ 2. ใช้เครื่องมือในการเข้าถึงข้อมูลที่มีของธุรกิจที่เป็นข้อมูลที่มีโครงสร้าง ที่รวมถึงข้อมูลแบบฐานข้อมูลเชิงสัมพันธ์ (RDBMS) ได้ 3. ใช้เครื่องมือในการเข้าถึงข้อมูลจัดเก็บในคลาวด์ (Cloud) ได้	การสาคิการปฏิบัติงาน ข้อสอบข้อเขียน
70301.02 เลือกข้อมูล (Select Data) ขององค์กรที่ต้องการใช้วิเคราะห์ที่มีโครงสร้าง	1. ระบุงานใจความสำคัญของข้อมูลในแต่ละส่วน ได้ 2. สามารถกำหนดเทคนิคเครื่องมือและข้อจำกัดสำหรับการเลือกใช้ส่วนของข้อมูลเพื่อการวิเคราะห์ได้ 3. สามารถกำหนดวิธีสุ่มตัวอย่างข้อมูลที่เป็นไปได้ ได้ 4. สามารถรวบรวม หรือนำออกข้อมูลตามความต้องการได้	การสาคิการปฏิบัติงาน ข้อสอบข้อเขียน

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70301.03 ปรับคุณภาพข้อมูล (Clean Data) ที่มีโครงสร้าง	1. ระบุคุณภาพข้อมูลที่ต้องการเพื่อการวิเคราะห์ได้ 2. ระบุความไม่สมบูรณ์ (Noise) ต่าง ๆ ของข้อมูลได้ 3. ใช้คำสั่ง หรือเครื่องมือเพื่อดำเนินการตรวจสอบความไม่สมบูรณ์ข้อมูลได้ 4. ใช้คำสั่ง หรือเครื่องมือเพื่อดำเนินการแก้ไข ปรับปรุงความไม่สมบูรณ์ข้อมูลได้ 5. ใช้คำสั่ง หรือเครื่องมือเพื่อปรับแต่งข้อมูลที่มีลักษณะเฉพาะให้ตรงกับ ความต้องการได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน

12. ความรู้และทักษะก่อนหน้าที่จำเป็น (Pre-requisite Skill & Knowledge)

N/A

13. ทักษะและความรู้ที่ต้องการ (Required Skills and Knowledge)

- (ก) ความต้องการด้านทักษะ
1. สามารถใช้วิธีการวิเคราะห์และวิธีทางสถิติเพื่อช่วยเตรียมข้อมูลและเลือกข้อมูลได้
 2. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผลข้อมูลขนาดใหญ่ได้
 3. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผล SQL หรือ NoSQL หรือที่เกี่ยวข้องได้
- (ข) ความต้องการด้านความรู้
1. เข้าใจธุรกิจและวิเคราะห์ข้อมูลให้สอดคล้องกับวัตถุประสงค์ธุรกิจ

14. หลักฐานที่ต้องการ (Evidence Guide)

- (ก) หลักฐานการปฏิบัติงาน (Performance Evidence)
1. ใบรับรองการปฏิบัติงานจากสถานประกอบการ
- (ข) หลักฐานความรู้ (Knowledge Evidence)
1. ใบรับรองการเข้ารับการฝึกอบรม
 2. ใบประกาศนียบัตรวุฒิการศึกษา
- (ค) คำแนะนำในการประเมิน
1. ผู้ประเมินตรวจสอบประเมินเกี่ยวกับการเลือกข้อมูลที่มีโครงสร้าง โดยพิจารณาจากร่องรอยหลักฐานที่เกี่ยวข้องทั้งหลักฐานการปฏิบัติงาน และหลักฐานความรู้
- (ง) วิธีการประเมิน
1. พิจารณาตามหลักฐานการปฏิบัติงาน
 2. พิจารณาตามหลักฐานความรู้

15. ขอบเขต (Range Statement)

- (ก) คำแนะนำ
- หน่วยสมรรถนะนี้เป็นการเลือกข้อมูล หรือส่วนของข้อมูลที่มีโครงสร้าง ที่รวมถึงฐานข้อมูลเชิงสัมพันธ์ที่จะนำมาใช้ในการวิเคราะห์ พร้อมทั้งเหตุผลและคำอธิบายประกอบทั้งส่วนที่ใช้และไม่ใช้ การนำออก (Export) รวมถึงการเตรียมข้อมูลเพื่อนำข้อมูลที่ไม่สมบูรณ์ออก ปรับคุณภาพของข้อมูล (Clean Data) โดยพิจารณาให้สอดคล้องสัมพันธ์กับเป้าหมายธุรกิจ คุณภาพของข้อมูล และข้อมูลด้านเทคนิค
- (ข) คำอธิบายรายละเอียด
1. มีความเข้าใจข้อมูลที่มีของธุรกิจ โดยการระบุชนิดข้อมูลตามโครงสร้างข้อมูลได้ เช่น
 - 1) ข้อมูลที่มีโครงสร้าง (Structured data)
 - 2) ข้อมูลที่ไม่มีโครงสร้าง (Unstructured data)ความแตกต่างระหว่าง Structured data และ Unstructured data มีดังนี้ Structured data จะหมายถึงข้อมูลที่มีการปรับแต่งเพื่อให้มีโครงสร้างชัดเจน มีความหมาย สามารถใช้งานได้ทันที เช่นข้อมูลที่จัดเก็บใน Database (ผ่านการ Normalization หรือ Meaning extraction แล้ว) ส่วน ซึ่งตรงกันข้ามกับ Unstructured data ข้อมูลที่มีโครงสร้าง (Structured) คือข้อมูลที่สามารถกำหนดเป็นรูปแบบของข้อมูล ที่ชัดเจนได้ เช่น Text, Numbers, money, percentages, Dates/Times, Booleans และ Email addresses เป็นต้น

2. การเข้าถึงข้อมูลที่มีของธุรกิจ เป็นการใช้อุปกรณ์ช่วยต่าง ๆ ในการเข้าถึงข้อมูลตามเทคโนโลยีที่เกี่ยวข้อง โดยการเข้าถึงอาจไม่ใช้อุปกรณ์ช่วยโดยเป็นการปฏิบัติตามขั้นตอนหรือคำสั่ง หรือใช้อุปกรณ์อื่นช่วย ซึ่งผู้ที่เข้าถึงจะต้องได้รับสิทธิ์การเข้าถึงข้อมูลนั้น ๆ หรือมอบหมายให้ผู้มีสิทธิ์เข้าถึงข้อมูลนั้นเป็นผู้ดำเนินการแทน เช่น

- 1) เข้าถึงฐานข้อมูลเชิงสัมพันธ์ (Relational Databases) หรือ SQL (Structured Query Language) เช่น MySQL, Oracle, IBM DB2, Sybase, MS SQL Server, Microsoft Azure, Mariadb และ PostgreSQL เป็นต้น
- 2) เข้าถึงข้อมูลที่จัดเก็บในคลาวด์ (Cloud) ทั้งขององค์กรเอง และใช้บริการภายนอกองค์กร
3. การปรับปรุงคุณภาพข้อมูล (Clean Data) สำหรับข้อมูลที่มีโครงสร้างแบบฐานข้อมูลเชิงสัมพันธ์ (RDBMS) โดยให้มาตรฐานคำสั่ง SQL หรือเครื่องมือช่วย (Tools) ที่เกี่ยวข้อง เช่น การลดจำนวนข้อมูลที่ผิดพลาด การแก้ไข เพิ่มเติมข้อมูล หรือเพิ่มจำนวนข้อมูลให้สมบูรณ์พร้อมใช้ เป็นต้น
4. คำสั่งและเครื่องมือพื้นฐานสำหรับการดำเนินการตามข้อ 2. และ 3. มีดังนี้

4.1 คำสั่งภาษา SQL (Structured Query Language)

ภาษา SQL ตามมาตรฐาน ANSI (American National Standards Institute) และเป็นที่ยอมรับของ ISO (International Organization for Standardization) ประกอบด้วย

1) ภาษานิยามข้อมูล (Data Definition Language: DDL) เป็นคำสั่งที่ใช้ในการสร้างฐานข้อมูล กำหนดโครงสร้างข้อมูลว่ามี Attribute ใด ชนิดของข้อมูล รวมทั้งการเปลี่ยนแปลงตาราง และการสร้างดัชนี เช่น CREATE, DROP และ ALTER

2) ภาษาจัดการข้อมูล (Data Manipulation Language: DML) เป็นคำสั่งที่ใช้ในการเรียกใช้ เพิ่ม ลบ และเปลี่ยนแปลงข้อมูลในตาราง เช่น SELECT, INSERT, UPDATE, DELETE โดยคำสั่ง SQL เป็นโครงสร้างภาษามาตรฐานเพื่อการจัดการระบบฐานข้อมูลเชิงสัมพันธ์

3) ภาษาควบคุมข้อมูล (Data Control Language: DCL) เป็นคำสั่งที่ใช้ในการกำหนดสิทธิการอนุญาต หรือ ยกเลิก การเข้าถึงฐานข้อมูล เพื่อป้องกันความปลอดภัยของฐานข้อมูล คำสั่ง: GRANT, REVOKE

ระบบจัดการฐานข้อมูล (DBMS) ที่สนับสนุนการใช้คำสั่ง SQL เช่น Oracle, DB2, MS-SQL, SAP Sybase ASE, Informix, MySQL และ MS-Access เป็นต้น

4) เครื่องมือช่วย (Tools) คือโปรแกรมที่ทำงานแทนชุดคำสั่ง SQL เพื่อช่วยให้การปฏิบัติงานง่ายและถูกต้องมากขึ้น โดยทั่วไประบบจัดการฐานข้อมูล (DBMS) จะมีระบบเครื่องมือช่วยของตนเอง แต่ก็มีโปรแกรมเครื่องมือที่สร้างขึ้นมาร่วมกับระบบจัดการฐานข้อมูล (DBMS) ต่าง ๆ เช่น SQLYog, MySQL Workbench, DBeaver, DataGrip, Navicat และ pgAdmin รวมถึงเครื่องมือสำหรับงานด้าน Data mining เช่น R, Python, Java, Weka, Rapidminer เป็นต้น

5) ข้อมูล Text File แบบมีโครงสร้างอย่างชัดเจน มีดังนี้

1) XML หรือ Extensible Markup Language คือภาษาที่ใช้ในการแสดงผลข้อมูล ซึ่งถูกออกแบบมาเพื่อเก็บข้อมูล ทั้งข้อมูลและโครงสร้างของข้อมูลนั้นๆ ไว้ด้วยกัน โดยภาษา XML มีโครงสร้างที่ประกอบด้วยแท็กเปิด และแท็กปิด เช่นเดียวกับภาษา HTML แต่ภาษา XML สามารถสร้างแท็กรวมทั้งกำหนดโครงสร้างของข้อมูลได้เองตามมาตรฐาน W3C (World Wide Web Consortium)

2) JSON ย่อมาจาก JavaScript Object Notation เป็น Standard format อย่างหนึ่งที่เป็น text และสามารถอ่านออกได้ด้วยตาเปล่า มาตรฐานของฟอร์แมต JSON ได้แก่ RFC 4627 มี Internet media type เป็น application/json และมีนามสกุลของไฟล์เป็น .json

3) Spreadsheet หรือแผ่นตารางทำการ เป็นลักษณะข้อมูลที่มีการจัดเรียงในลักษณะตารางสี่เหลี่ยม ที่ใช้ในการคำนวณเป็นหลักและสามารถใช้ในการเก็บข้อมูลได้

4) CSV (Comma-Separated Value) คือ Text File สำหรับเก็บข้อมูลแบบตาราง โดยใช้จุลภาค (,) แบ่งข้อมูลในแต่ละหลัก (Column) และใช้การเว้นบรรทัดแทนการแบ่งแถว (Row) ในกรณีที่ข้อมูลมีเครื่องหมายจุลภาค (,) อยู่ด้วย ให้ครอบข้อมูลด้วยเครื่องหมาย “ ”

5) ข้อมูลอื่น ๆ ที่มาจากโปรแกรมหรือระบบอื่น ๆ ที่มีลักษณะใกล้เคียงกับภาษา XML เช่น KML หรือ Keyhole Markup Language ซึ่งเป็น ภาษา XML notation ที่ใช้สำหรับ geographic annotation and visualization ที่แสดงผลแบบแผนที่ 2 มิติ และ 3 มิติ ซึ่ง KML พัฒนาโดย Google Earth

6) Cloud database เป็นฐานข้อมูลแบบใหม่ ที่ถูกปรับปรุงและสร้างขึ้นบนระบบ virtualized แบบเดียวกับ hybrid cloud, public cloud หรือ private cloud โดยสามารถขยายขนาดเพิ่มขึ้น (Scale) หรือ ปรับแต่ง resource ได้ตลอดเวลาตามความต้องการของระบบและใช้งาน การเข้าถึงทรัพยากรใน Cloud สามารถใช้เครื่องมือช่วยเช่น Cloud Control Panel เป็นต้น

16. หน่วยสมรรถนะรวม (ถ้ามี)

N/A

17. อุตสาหกรรมร่วม/กลุ่มอาชีพร่วม (ถ้ามี)

N/A

18. รายละเอียดกระบวนการและวิธีการประเมิน (Assessment Description and Procedure)

18.1 เครื่องมือประเมินการเข้าถึงข้อมูลแบบมีของธุรกิจตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.2 เครื่องมือประเมินการเลือกข้อมูล (Select Data) ขององค์กรที่ต้องการใช้วิเคราะห์ข้อมูลแบบมีโครงสร้างตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.3 เครื่องมือประเมินการปรับปรุงภาพข้อมูล (Clean Data) ที่มีโครงสร้างตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

1. รหัสหน่วยสมรรถนะ

70302
2. ชื่อหน่วยสมรรถนะ

แปลงข้อมูลที่มีโครงสร้าง
3. ทบทวนครั้งที่

1 / 2566
4. สร้างใหม่

☐

ปรับปรุง

☒

5. สำหรับชื่ออาชีพและรหัสอาชีพ (Occupational Classification)

วิศวกรรมข้อมูล (Data Engineer)

6. คำอธิบายหน่วยสมรรถนะ (Description of Unit of Competency)

ผู้ผ่านสมรรถนะนี้จะมีความรู้เกี่ยวกับการแปลงข้อมูลที่มีโครงสร้าง สามารถดำเนินการกับข้อมูลทั้งหมดหรือบางส่วนของข้อมูลที่ได้เลือกใช้ สำหรับข้อมูลที่มีโครงสร้าง รวมถึงข้อมูลแบบฐานข้อมูลเชิงสัมพันธ์ (RDBMS) ที่มีขององค์กร โดยดำเนินการข้อมูลได้แก่ การปรับปรุงรูปแบบให้เหมาะสม (Construct Data) และการรวบรวมข้อมูลเข้าด้วยกัน (Integrate Data) รวมทั้งข้อมูลด้านเทคนิคต่าง ๆ เช่น ข้อจำกัด รูปแบบข้อมูลที่ต้องการ ซึ่งมีความสัมพันธ์สอดคล้องกับเป้าหมายธุรกิจ

7. สำหรับระดับคุณวุฒิ

1	2	3	4	5	6	7	8
☐	☐	☒	☐	☐	☐	☐	☐

8. กลุ่มอาชีพ (Sector)

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล (Data Science)

9. ชื่ออาชีพและรหัสอาชีพอื่นที่หน่วยสมรรถนะนี้สามารถใช้ได้ (ถ้ามี)

N/A

10. ข้อกำหนดหรือกฎระเบียบที่เกี่ยวข้อง (Licensing or Regulation Related) (ถ้ามี)

N/A

11. สมรรถนะย่อยและเกณฑ์การปฏิบัติงาน (Elements and Performance Criteria)

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70302.01 ปรับปรุงคุณสมบัติข้อมูลเพื่อการวิเคราะห์แบบมีโครงสร้าง (Construct Data)	1. ระบุคุณลักษณะของข้อมูลที่เหมาะสมสำหรับการวิเคราะห์ได้ 2. ระบุคุณลักษณะของข้อมูลที่ต้องการเพิ่มเติมที่เหมาะสมสำหรับการวิเคราะห์ได้ 3. กำหนดเทคนิคที่เหมาะสมสำหรับการสร้างคุณลักษณะของข้อมูลที่ต้องการได้ 4. สามารถทำข้อมูลให้เป็นบรรทัดฐานคุณสมบัติข้อมูล (Attribute Normalized) ที่ต้องการได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน
70302.02 รวบรวมข้อมูล (Integrate Data) แบบมีโครงสร้าง	1. ระบุกลไกการทำงานของเครื่องมือสำหรับการรวบรวมข้อมูลได้ 2. ระบุเครื่องมือ หรือคำสั่งที่เหมาะสมสำหรับการรวบรวมข้อมูลได้ 3. เลือกใช้เครื่องมือ หรือคำสั่งสำหรับการรวบรวมข้อมูลได้ 4. สามารถรวบรวมข้อมูลที่ต้องการได้และถูกต้อง 5. สามารถใช้คำสั่ง SQL มาตรฐานปัจจุบันเพื่อการรวบรวมข้อมูลได้	การสังเกตการปฏิบัติงาน ข้อสอบข้อเขียน

12. ความรู้และทักษะก่อนหน้าที่จำเป็น (Pre-requisite Skill & Knowledge)

N/A

13. ทักษะและความรู้ที่ต้องการ (Required Skills and Knowledge)

(ก) ความต้องการด้านทักษะ

1. สามารถใช้วิธีการวิเคราะห์และวิธีทางสถิติเพื่อช่วยเตรียมข้อมูลและเลือกข้อมูลได้
2. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผลข้อมูลขนาดใหญ่ได้
3. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผล SQL หรือ NoSQL หรือที่เกี่ยวข้องได้

(ข) ความต้องการด้านความรู้

1. เข้าใจธุรกิจและวิเคราะห์ข้อมูลให้สอดคล้องกับวัตถุประสงค์ธุรกิจ

14. หลักฐานที่ต้องการ (Evidence Guide)

(ก) หลักฐานการปฏิบัติงาน (Performance Evidence)

1. ใบรับรองการปฏิบัติงานจากสถานประกอบการ

(ข) หลักฐานความรู้ (Knowledge Evidence)

1. ใบรับรองการเข้ารับการฝึกอบรม
2. ใบประกาศนียบัตรวุฒิการศึกษา

(ค) คำแนะนำในการประเมิน

1. ผู้ประเมินตรวจสอบประเมินเกี่ยวกับการแปลงข้อมูลที่มีโครงสร้าง โดยพิจารณาจากร่องรอยหลักฐานที่เกี่ยวข้องทั้งหลักฐานการปฏิบัติงาน และหลักฐานความรู้

(ง) วิธีการประเมิน

1. พิจารณาตามหลักฐานการปฏิบัติงาน
2. พิจารณาตามหลักฐานความรู้

15. ขอบเขต (Range Statement)

(ก) คำแนะนำ

หน่วยสมรรถนะนี้เป็นการเพิ่มเติมหรือปรับปรุงรูปแบบให้เหมาะสม (Construct Data) และการบูรณาการข้อมูลเข้าด้วยกัน (Integrate Data)

โดยพิจารณาให้สอดคล้องและสัมพันธ์กับเป้าหมายธุรกิจ คุณภาพของข้อมูลด้านเทคนิค

(ข) คำอธิบายรายละเอียด

1. คำสั่งและเครื่องมือพื้นฐานสำหรับการดำเนินการ Construct Data และ Integrate Data มีดังนี้

1.1 คำสั่งภาษา SQL (Structured Query Language)

ภาษา SQL ตามมาตรฐาน ANSI (American National Standards Institute) และเป็นที่ยอมรับของ ISO (International Organization for Standardization) ประกอบด้วย

1) ภาษานิยามข้อมูล (Data Definition Language: DDL) เป็นคำสั่งที่ใช้ในการสร้างฐานข้อมูล กำหนดโครงสร้างข้อมูลว่ามี Attribute ใด ชนิดของข้อมูล รวมทั้งการเปลี่ยนแปลงตาราง และการสร้างดัชนี เช่น CREATE, DROP และ ALTER

2) ภาษาจัดการข้อมูล (Data Manipulation Language: DML) เป็นคำสั่งที่ใช้ในการเรียกใช้ เพิ่ม ลบ และเปลี่ยนแปลงข้อมูลในตาราง เช่น SELECT, INSERT, UPDATE, DELETE โดยคำสั่ง SQL เป็นโครงสร้างภาษามาตรฐานเพื่อการจัดการบนระบบฐานข้อมูลเชิงสัมพันธ์

3) ภาษาควบคุมข้อมูล (Data Control Language: DCL) เป็นคำสั่งที่ใช้ในการกำหนดสิทธิการอนุญาต หรือ ยกเลิก การเข้าถึงฐานข้อมูล เพื่อป้องกันความปลอดภัยของฐานข้อมูล คำสั่ง: GRANT, REVOKE

ระบบจัดการฐานข้อมูล (DBMS) ที่สนับสนุนการใช้คำสั่ง SQL เช่น Oracle, DB2, MS-SQL, SAP Sybase ASE, Informix, MySQL และ MS-Access เป็นต้น

1. เครื่องมือช่วย (Tools) คือโปรแกรมที่ทำงานแทนชุดคำสั่ง SQL เพื่อช่วยให้การปฏิบัติงานง่ายและถูกต้องมากขึ้น โดยทั่วไประบบจัดการฐานข้อมูล (DBMS) จะมีระบบเครื่องมือช่วยของตนอยู่ แต่ก็มีโปรแกรมเครื่องมือที่สร้างขึ้นมาใช้ร่วมกับระบบจัดการฐานข้อมูล (DBMS) ต่าง ๆ เช่น SQLYog, MySQL Workbench, DBeaver, DataGrip, Navicat และ pgAdmin รวมถึงเครื่องมือสำหรับงานด้าน Data mining เช่น R, Python, Java, Weka, Rapidminer เป็นต้น

2. Cloud database เป็นฐานข้อมูลแบบใหม่ ที่ถูกปรับปรุงและสร้างขึ้นบนระบบ virtualized แบบเดียวกับ hybrid cloud, public cloud หรือ private cloud โดยสามารถขยายขนาดเพิ่มขึ้น (Scale) หรือ ปรับแต่ง resource ได้ตลอดเวลาตามความต้องการของระบบและผู้ใช้งาน การเข้าถึงทรัพยากรใน Cloud สามารถใช้เครื่องมือช่วยเช่น Cloud Control Panel เป็นต้น

16. หน่วยสมรรถนะร่วม (ถ้ามี)

N/A

17. อุตสาหกรรมร่วม/กลุ่มอาชีพร่วม (ถ้ามี)

N/A

18. รายละเอียดกระบวนการและวิธีการประเมิน (Assessment Description and Procedure)

18.1 เครื่องมือประเมินการปรับปรุงคุณสมบัติข้อมูลเพื่อการวิเคราะห์ (Construct Data) แบบมีโครงสร้างตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.2 เครื่องมือประเมินการดำเนินการรวบรวมข้อมูล (Integrate Data) แบบมีโครงสร้างตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

1. รหัสหน่วยสมรรถนะ

70305
2. ชื่อหน่วยสมรรถนะ

จัดรูปแบบข้อมูล
3. ทบทวนครั้งที่

1 / 2566
4. สร้างใหม่

☐

ปรับปรุง

☒

5. สำหรับชื่ออาชีพและรหัสอาชีพ (Occupational Classification)

วิศวกรรมข้อมูล (Data Engineer)

6. คำอธิบายหน่วยสมรรถนะ (Description of Unit of Competency)

สามารถดำเนินการกับข้อมูลต่าง ๆ ทั้งที่มีโครงสร้าง และไม่มีโครงสร้างที่ผ่านการเลือกและเปลี่ยนแปลงข้อมูลมาแล้ว โดยเป็นการปรับรูปแบบข้อมูล (Reformatted) ให้สอดคล้องกันกับข้อมูลที่เกี่ยวข้อง การจัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes) และการจัดเรียงลำดับรายการใหม่ (Reordering) เพื่อช่วยในการประมวลผล การเปลี่ยนค่าข้อมูล (Reformatted within-value) เพื่อให้ได้ข้อมูลที่สอดคล้องกันและดำเนินการต่อในขั้นต่อไปได้ ซึ่งมีความสัมพันธ์สอดคล้องกับเทคนิคของแบบจำลองและเครื่องมือที่ใช้

7. สำหรับระดับคุณวุฒิ

1	2	3	4	5	6	7	8
☐	☐	☒	☐	☐	☐	☐	☐

8. กลุ่มอาชีพ (Sector)

สาขาวิชาชีพอุตสาหกรรมดิจิทัล สาขาวิทยาศาสตร์ข้อมูล (Data Science)

9. ชื่ออาชีพและรหัสอาชีพอื่นที่หน่วยสมรรถนะนี้สามารถใช้ได้ (ถ้ามี)

N/A

10. ข้อกำหนดหรือกฎระเบียบที่เกี่ยวข้อง (Licensing or Regulation Related) (ถ้ามี)

N/A

11. สมรรถนะย่อยและเกณฑ์การปฏิบัติงาน (Elements and Performance Criteria)

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70305.01 ระบุข้อมูลและแบบจำลอง	1. ระบุโครงสร้างและความหมายข้อมูลได้ 2. ระบุข้อจำกัดของแบบจำลองที่ต้องใช้ ได้ 3. ระบุข้อจำกัดของเครื่องมือช่วยในการสร้างแบบจำลองและการวิเคราะห์ได้	ข้อสอบข้อเขียน
70305.02 จัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes)	1. ระบุหลักการจัดเรียงสดมภ์ข้อมูลใหม่ (Order) ได้ 2. ระบุข้อจำกัดของเครื่องมือช่วยในการสร้างแบบจำลองและการวิเคราะห์ที่มีผลต่ลำดับก่อนหลังของสดมภ์ข้อมูลได้ 3. สามารถใช้เครื่องมือจัดเรียงสดมภ์ข้อมูลใหม่ได้	การสาธิตการปฏิบัติงาน ข้อสอบข้อเขียน
70305.03 จัดเรียงรายการข้อมูลใหม่ (Reordering Records)	1. ระบุหลักการจัดเรียงรายการข้อมูลใหม่ (Order) ได้ 2. ระบุข้อจำกัดของเครื่องมือช่วยในการสร้างแบบจำลองและการวิเคราะห์ที่มีผลต่ลำดับก่อนหลังของรายการข้อมูลได้ 3. สามารถใช้เครื่องมือจัดเรียงรายการข้อมูลใหม่ได้	การสาธิตการปฏิบัติงาน ข้อสอบข้อเขียน

สมรรถนะย่อย (Element)	เกณฑ์ในการปฏิบัติงาน (Performance Criteria)	วิธีการประเมิน (Assessment)
70305.04 เปลี่ยนค่าข้อมูล (Reformatted within-value)	1. ระบุข้อกำหนดของเครื่องมือช่วยในการสร้างแบบจำลองและกา วิเคราะห์ที่มีผลต่อรูปแบบของข้อมูลได้ 2. สามารถปรับเปลี่ยนรูปแบบของข้อมูลได้ 3. สามารถใช้เครื่องมือช่วยปรับเปลี่ยนรูปแบบข้อมูลใหม่ได้ 4. สามารถแปลงข้อมูลให้อยู่ในช่วงเดียวกัน (Scale) ได้	การสาธิตการปฏิบัติงาน ข้อสอบข้อเขียน

12. ความรู้และทักษะก่อนหน้าที่จำเป็น (Pre-requisite Skill & Knowledge)

N/A

13. ทักษะและความรู้ที่ต้องการ (Required Skills and Knowledge)

(ก) ความต้องการด้านทักษะ

1. สามารถใช้วิธีการวิเคราะห์และวิธีทางสถิติเพื่อช่วยเตรียมข้อมูลและเลือกข้อมูลได้
2. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผลข้อมูลขนาดใหญ่ได้
3. สามารถใช้เทคนิคและเครื่องมือสำหรับประมวลผล SQL หรือ NoSQL หรือที่เกี่ยวข้องได้ เช่น SQL and NoSQL, Data Warehouse solutions, ETL (Extract, Transform, Load), OLTP และ OLAP เป็นต้น

(ข) ความต้องการด้านความรู้

1. เข้าใจธุรกิจและวิเคราะห์ข้อมูลให้สอดคล้องกับวัตถุประสงค์ธุรกิจ

14. หลักฐานที่ต้องการ (Evidence Guide)

(ก) หลักฐานการปฏิบัติงาน (Performance Evidence)

1. ใบรับรองการปฏิบัติงานจากสถานประกอบการ

(ข) หลักฐานความรู้ (Knowledge Evidence)

1. ใบรับรองการเข้ารับการฝึกอบรม
2. ใบประกาศนียบัตรวุฒิการศึกษา

(ค) คำแนะนำในการประเมิน

1. ผู้ประเมินตรวจสอบประเมินเกี่ยวกับการจัดรูปแบบข้อมูล โดยพิจารณาจากร่องรอยหลักฐานที่เกี่ยวข้องทั้งหลักฐานการปฏิบัติงาน และหลักฐานความรู้

(ง) วิธีการประเมิน

1. พิจารณาตามหลักฐานการปฏิบัติงาน
2. พิจารณาตามหลักฐานความรู้

15. ขอบเขต (Range Statement)

(ก) คำแนะนำ

1. หน่วยสมรรถนะนี้เป็นการปรับเปลี่ยนการจัดเรียงข้อมูล เพื่อให้เหมาะสมและสามารถดำเนินการได้กับเทคนิคแบบจำลองที่เลือกใช้ รวมทั้งเครื่องมือช่วย (Tools)

(ข) คำอธิบายรายละเอียด

1. มีความเข้าใจในโครงสร้างข้อมูลผ่านการแปลงข้อมูลแล้ว และสามารถดำเนินการปรับรูปแบบข้อมูล (Reformatted) จัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes) จัดเรียงลำดับรายการใหม่ (Reordering) เปลี่ยนค่าข้อมูล (Reformatted within-value) ได้อย่างเหมาะสม
2. การจัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes) บางเครื่องมือช่วยวิเคราะห์อาจต้องการการจัดเรียงสดมภ์ใหม่ หรือเพิ่มสดมภ์ที่เป็นคีย์หลักที่ไม่มีความซ้ำซ้อน เพื่อให้การประมวลผลของแบบจำลองมีความถูกต้อง
3. การจัดเรียงลำดับรายการใหม่ (Reordering) เพื่อให้การประมวลผลของแบบจำลองบางแบบจำลองมีความถูกต้อง จำเป็นต้องจัดเรียงรายการข้อมูลใหม่ โดยมีหลักการดังนี้
 - 1) Ascending คือ การเรียงข้อมูลจากน้อยไปหามาก เช่น 1 - 2 - 3 - หรือ A - Z เป็นต้น
 - 2) Descending คือ การเรียงข้อมูลจากมากไปหาน้อย เช่น 100 - 99 - 98 - หรือ Z - A เป็นต้น
4. การเปลี่ยนค่าข้อมูล (Reformatted within-value) เป็นการปรับเปลี่ยนรูปแบบ หรือค่าข้อมูลให้เหมาะสมกับแบบจำลอง และเครื่องมือที่ใช้

16. หน่วยสมรรถนะร่วม (ถ้ามี)

N/A

17. อุตสาหกรรมร่วม/กลุ่มอาชีพร่วม (ถ้ามี)

N/A

18. รายละเอียดกระบวนการและวิธีการประเมิน (Assessment Description and Procedure)

18.1 เครื่องมือประเมินการระบุข้อมูลและแบบจำลองตามข้อกำหนดมาตรฐาน

1. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.2 เครื่องมือประเมินการจัดเรียงสดมภ์ข้อมูลใหม่ (Rearranging Attributes) ตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.3 เครื่องมือประเมินการจัดเรียงรายการข้อมูลใหม่ (Reordering Records) ตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน

18.4 เครื่องมือประเมินการเปลี่ยนค่าข้อมูล (Reformatted within-value) ตามข้อกำหนดมาตรฐาน

1. แบบฟอร์มประเมินผลการการสาธิตการปฏิบัติงาน

2. ผลข้อสอบข้อเขียน

ดูรายละเอียดจากคู่มือประเมิน