



กรมควบคุมโรค
Department of Disease Control



DATA MANAGEMENT AND ANALYSIS COURSE: DATA SCIENCE AND DATA ENGINEER

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล
สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล



หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล
ได้ผ่านการตรวจประเมินและรับรองมาตรฐานผลิตภัณฑ์ เพื่อการเฝ้าระวัง ป้องกัน ควบคุมโรค
และภัยสุขภาพ กรมควบคุมโรค
ณ วันที่ 31 กรกฎาคม 2567

DCC 67008



หลักสูตรการจัดการและการวิเคราะห์ข้อมูล

ระดับพื้นฐาน (Basic) ระดับกลาง (Intermediate) และระดับสูง (Advance)

คณะที่ปรึกษา :

นายแพทย์ศุภณัฐ วังศานุพัทธ์
นายเอกชัย วิวรรณศิริรักษ์

นายสิงหนาท เรืองโอสถ
นายวริทธิ์ เต็นพิพัฒน์
นายศุภวิชญ์ สมสกุล

กองระบาดวิทยา
ผู้อำนวยการศูนย์นวัตกรรมซอฟต์แวร์และการประมวลผล
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี
มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี

บรรณาธิการ :

นายแพทย์พงษ์ธรชาติพิทักษ์
นางสาวรัฐลักษณ์ พิทักษ์จักรพิภพ

ผู้อำนวยการกองดิจิทัลเพื่อการควบคุมโรค
รองผู้อำนวยการกองดิจิทัลเพื่อการควบคุมโรค

คณะดำเนินการ :

นางสาวคุณกัญญ์ศศิ พิมพขันธ์
นางสาวรัฐชญา ไหลลั่น
นางสาวกรรฎทัย กิตติคุณ
นายธนดน พัฒนะโกไคย
นายอนุชิต สายศร
นายอนันทวีป ใจเมา
นางสาวธนาภรณ์ แสงบรรเจิดกุล
นางสาวนัทธ์ชนภัทร พรหมณะ
นายศรายุทธ ฤาข่าย

กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค
กองดิจิทัลเพื่อการควบคุมโรค

Citation :

Pimpakhan K., Saisorn A., Chaimao A., Lailon R., Kittikhun K., Promana N., et. al. Data management and analysis course: Data science and data engineer. Nonthaburi: Division of Digital Disease Control; 2024.

คำนำ

ปัจจุบันข้อมูลมีความสำคัญต่อการกำหนดทิศทางนโยบายสุขภาพ การพัฒนาระบบบริการสุขภาพ และการจัดการกับปัญหาสุขภาพทั้งในระดับชาติและระดับพื้นที่ เรียกว่าเทคโนโลยีดิจิทัลมีส่วนสำคัญในการเก็บสะสมมวลของข้อมูลได้อย่างไร้ขีดจำกัด ทำให้ข้อมูลมีขนาดใหญ่ (Scale) มีกระแสข้อมูลที่ส่งตลอดเวลา (Data Flow) และมีรูปแบบที่หลากหลาย (Forms) เมื่อเกิดข้อมูลปริมาณ การจะนำข้อมูลระดับ Big Data มาใช้ได้นั้น จำเป็นต้องใช้ทักษะและความรู้ในการบริหารจัดการและวิเคราะห์ข้อมูล กองดิจิทัลเพื่อการควบคุมโรค เล็งเห็นถึงความสำคัญของการพัฒนาศักยภาพการทำงานของบุคลากรสาธารณสุข ในการบริหารจัดการข้อมูลด้านสาธารณสุขเพื่อให้เกิดประสิทธิภาพอย่างสูงสุด โดยสามารถนำไปช่วยในการวางแผน สนับสนุนการตัดสินใจ ทำนาย คาดการณ์ สิ่งที่จะเกิดขึ้นในอนาคตได้อย่างถูกต้อง รวดเร็วทันต่อสถานการณ์

กองดิจิทัลเพื่อการควบคุมโรค จึงได้พัฒนาหลักสูตรการจัดการและการวิเคราะห์ข้อมูล (Data Management and Analysis) ฉบับนี้ขึ้น ซึ่งหลักสูตรประกอบด้วย 2 สาขา ได้แก่ สาขาวิทยาศาสตร์ข้อมูล (Data Science) และสาขาวิศวกรรมข้อมูล (Data Engineer) โดยแบ่งตามระยะการเรียนรู้เป็น 3 ระยะ ได้แก่ ระยะที่ 1 หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูลระดับปรับพื้นฐาน ระยะที่ 2 หลักสูตรการจัดการ และการวิเคราะห์ข้อมูลสาขาวิทยาศาสตร์ข้อมูลและสาขาวิศวกรรมข้อมูล และระยะที่ 3 หลักสูตรเตรียมความพร้อม สำหรับการนำเสนอผลงาน โดยมีวัตถุประสงค์เพื่อพัฒนาศักยภาพบุคลากรสาธารณสุขให้มีความรู้ ความเข้าใจ เกี่ยวกับงานวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล ทั้งในด้านการวิเคราะห์และประมวลผลข้อมูล สามารถใช้ความรู้ ทักษะ เทคโนโลยี และเครื่องมือในการวิเคราะห์ ออกแบบ สร้าง และประเมินผลงาน ในการแก้ปัญหาขององค์กร หรือนำมาใช้ประกอบการตัดสินใจและการแก้ปัญหาได้

กองดิจิทัลเพื่อการควบคุมโรค ขอขอบพระคุณผู้ที่มีส่วนเกี่ยวข้องทุกภาคส่วน ที่ร่วมแรง ร่วมใจพัฒนา หลักสูตรฉบับนี้จนสำเร็จลุล่วงเป็นอย่างดี และหวังเป็นอย่างยิ่งว่าหลักสูตรการจัดการและการวิเคราะห์ข้อมูล (Data Management and Analysis) ฉบับนี้ จะเป็นประโยชน์ต่อการนำไปปฏิบัติงานด้านสาธารณสุข ให้มีประสิทธิภาพต่อไป

กองดิจิทัลเพื่อการควบคุมโรค

มีนาคม พ.ศ. 2567

กิตติกรรมประกาศ

การจัดทำหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล สำเร็จลุล่วงเป็นอย่างดี เนื่องจากได้รับความกรุณาจากผู้บริหาร อาจารย์ที่ปรึกษา และคณะทำงาน ขอขอบคุณ ผู้ที่มีส่วนร่วมทุกท่านที่สละเวลามาร่วมปรึกษา ให้ข้อมูล เสนอความคิดเห็นที่มีประโยชน์ในการจัดทำหลักสูตรนี้ ให้มีความสมบูรณ์ชัดเจนถูกต้อง และเป็นประโยชน์ต่อการพัฒนาศักยภาพบุคลากร ให้มีคุณภาพสำหรับการดำเนินงาน ด้านการป้องกันและควบคุมโรคต่อไป

กองดิจิทัลเพื่อการควบคุมโรค

มีนาคม พ.ศ. 2567

สารบัญ

| | |
|---|----|
| คำนำ | ก |
| กิตติกรรมประกาศ | ข |
| สารบัญ | ค |
| หมวดที่ 1 ข้อมูลทั่วไป | 1 |
| 1. รหัสและชื่อหลักสูตร | 1 |
| 2. หน่วยงานที่รับผิดชอบ | 1 |
| 3. ที่มาและความสำคัญของหลักสูตร | 1 |
| 4. วัตถุประสงค์ | 2 |
| 5. ตัวชี้วัด | 2 |
| 6. กลุ่มเป้าหมาย | 3 |
| 7. คุณสมบัติวิทยากร | 3 |
| 8. สถานที่จัดการอบรม | 4 |
| หมวดที่ 2 ระบบการจัดการอบรม การดำเนินการ และโครงสร้างของหลักสูตร | 5 |
| 1. รูปแบบของหลักสูตร | 5 |
| 2. การดำเนินการหลักสูตร | 5 |
| 3. โครงสร้างรายวิชาในหลักสูตร | 7 |
| 4. คำอธิบายรายวิชา | 10 |
| 1) หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล | 10 |
| 2) หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูล | 13 |
| 3) หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิศวกรรมข้อมูล | 24 |
| 4) หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน | 36 |
| หมวดที่ 3 การประกันคุณภาพหลักสูตร | 38 |
| 1. การปรับปรุงหลักสูตร | 38 |

| | |
|---|----|
| 2. แผนพัฒนาปรับปรุงหลักสูตร..... | 38 |
| 3. การประเมินผู้เข้ารับการอบรม..... | 39 |
| 4. แผนการประเมินความพึงพอใจ | 42 |
| ภาคผนวก ก..... | 43 |
| คำสั่งแต่งตั้งคณะทำงานพัฒนาและจัดทำหลักสูตรการจัดการและการวิเคราะห์ข้อมูล | 43 |
| ภาคผนวก ข..... | 46 |
| แบบประเมินความพึงพอใจ..... | 46 |
| แบบประเมินการนำเสนอผลงาน..... | 49 |
| ตัวอย่างใบเกียรติบัตร | 50 |

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล

หมวดที่ 1 ข้อมูลทั่วไป

1. รหัสและชื่อหลักสูตร

1.1 ระบุรหัส : DMA

1.2 ชื่อหลักสูตร : (ภาษาไทย) หลักสูตรการจัดการและการวิเคราะห์ข้อมูล

สาขาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล

(ภาษาอังกฤษ) Data Management and Analysis Course :

Data Science and Data Engineer

2. หน่วยงานที่รับผิดชอบ

กรมควบคุมโรค กระทรวงสาธารณสุข

3. ที่มาและความสำคัญของหลักสูตร

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สำหรับบุคลากรสาธารณสุข เกิดขึ้นจากการเปลี่ยนแปลงของสภาพแวดล้อมทางสาธารณสุขในปัจจุบัน ที่ส่งผลให้บุคลากรสาธารณสุขต้องเผชิญกับปัญหาและสถานการณ์ใหม่ๆ ที่มีความซับซ้อนมากขึ้น ประกอบกับเทคโนโลยีสารสนเทศและการสื่อสารที่พัฒนาอย่างรวดเร็ว ทำให้มีข้อมูลขนาดใหญ่ (Big Data) เกิดขึ้นมากมาย ซึ่งข้อมูลเหล่านี้สามารถนำมาใช้ในการวิเคราะห์เพื่อแก้ปัญหาและสถานการณ์ต่างๆ ได้อย่างมีประสิทธิภาพ

ดังนั้น บุคลากรสาธารณสุขจึงจำเป็นต้องมีความรู้และทักษะในการจัดการและการวิเคราะห์ข้อมูล เพื่อนำข้อมูลมาใช้ประโยชน์ในการดำเนินงานด้านสาธารณสุขได้อย่างมีประสิทธิภาพ โดยหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สำหรับบุคลากรสาธารณสุข จะช่วยพัฒนาความรู้และทักษะที่จำเป็น ดังนี้

- ความรู้พื้นฐานเกี่ยวกับการจัดการและการวิเคราะห์ข้อมูล
- เทคนิคและวิธีการในการจัดการและการวิเคราะห์ข้อมูล
- การนำข้อมูลมาประยุกต์ใช้ในการทำงานด้านสาธารณสุข

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สำหรับบุคลากรสาธารณสุข มีความสำคัญต่อบุคลากรสาธารณสุข ดังนี้

- ช่วยเพิ่มขีดความสามารถในการจัดการและการวิเคราะห์ข้อมูล บุคลากรสาธารณสุขที่มีทักษะในการจัดการและการวิเคราะห์ข้อมูล จะสามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ และสามารถวิเคราะห์ข้อมูลเพื่อค้นหาความสัมพันธ์และแนวโน้มต่างๆ ที่ซ่อนอยู่ในข้อมูล ซึ่งสามารถนำไปใช้ในการแก้ปัญหาและสถานการณ์ต่างๆ ได้อย่างมีประสิทธิภาพ
- ช่วยยกระดับคุณภาพการให้บริการสาธารณสุข บุคลากรสาธารณสุขที่มีทักษะในการจัดการและการวิเคราะห์ข้อมูล ให้สามารถใช้ประโยชน์จากข้อมูลเพื่อพัฒนาคุณภาพการให้บริการสาธารณสุข

เช่น การวิเคราะห์ข้อมูลเพื่อวางแผนการรักษาพยาบาล การวิเคราะห์ข้อมูลเพื่อพัฒนาระบบ
เฝ้าระวังโรค เป็นต้น

- ช่วยส่งเสริมความก้าวหน้าทางวิชาการด้านสาธารณสุข บุคลากรสาธารณสุขที่มีทักษะ
ในการจัดการและการวิเคราะห์ข้อมูล จะสามารถมีส่วนร่วมในการวิจัยและพัฒนาด้านสาธารณสุข
ได้อย่างมีประสิทธิภาพ ซึ่งจะช่วยส่งเสริมความก้าวหน้าทางวิชาการด้านสาธารณสุข

การจัดการและการวิเคราะห์ข้อมูล หมายถึง กระบวนการในการเก็บรวบรวม จัดเก็บ การควบคุม
การเข้าถึงข้อมูล และดูแลรักษาข้อมูลอย่างมีประสิทธิภาพ เพื่อให้สามารถนำมาใช้ประโยชน์ได้อย่าง
เต็มที่ และนำข้อมูลที่ได้ออกมาค้นหาข้อมูลเชิงลึก แนวโน้ม และรูปแบบต่าง ๆ ที่ช่วยให้สามารถตัดสินใจ
และวางแผนได้ดียิ่งขึ้น การใช้เทคนิคต่าง ๆ เพื่อการคาดการณ์และการแนะนำแนวทางในการแก้ไข
ปัญหาหรือปรับปรุงกระบวนการ

4. วัตถุประสงค์

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล มีวัตถุประสงค์ ดังนี้

- 1) เพื่อให้บุคลากรสาธารณสุขมีความรู้ ความเข้าใจเกี่ยวกับงานวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล
ทั้งในด้านการจัดการและการวิเคราะห์ข้อมูลจนนำไปสู่การประมวลผลข้อมูลได้ถูกต้อง
- 2) เพื่อให้บุคลากรสาธารณสุขสามารถใช้ความรู้ ทักษะ เทคโนโลยี และเครื่องมือทางวิทยาศาสตร์ข้อมูล
และวิศวกรรมข้อมูลในการวิเคราะห์ ออกแบบ สร้างและประเมินผลงาน เพื่อแก้ปัญหาขององค์กรได้
- 3) เพื่อให้บุคลากรสาธารณสุขสามารถตัดสินใจและแก้ปัญหา โดยใช้ทักษะทางวิทยาศาสตร์ข้อมูลและ
วิศวกรรมข้อมูลด้วยตนเองได้
- 4) เพื่อให้บุคลากรสาธารณสุขสามารถสื่อสารผลงาน องค์กรความรู้และความเข้าใจเชิงทั่วไปและเชิงเทคนิคได้
- 5) เพื่อให้บุคลากรสาธารณสุขสามารถประยุกต์ใช้ความรู้เรื่องธรรมาภิบาลข้อมูลกับงานวิทยาศาสตร์ข้อมูล
และวิศวกรรมข้อมูล

5. ตัวชี้วัด

ตัวชี้วัดของหลักสูตรการจัดการและการวิเคราะห์ข้อมูล แบ่งออกเป็น 3 ระดับ ดังนี้

ระดับความรู้

- บุคลากรสาธารณสุขที่เข้าอบรม สามารถตอบคำถามเกี่ยวกับแนวคิดและหลักการพื้นฐาน
ของการจัดการและการวิเคราะห์ข้อมูล ได้ถูกต้องไม่น้อยกว่าร้อยละ 80
- บุคลากรสาธารณสุขที่เข้าอบรม สามารถตอบคำถามเกี่ยวกับเทคนิคและวิธีการในการจัดการ
และการวิเคราะห์ข้อมูล เช่น สถิติ กระแสงาน (Workflow) และการเรียนรู้ของเครื่อง
(Machine Learning) ได้ถูกต้องไม่น้อยกว่าร้อยละ 80

ระดับทักษะ

- บุคลากรสาธารณสุขที่เข้าอบรม สามารถประมวลผลข้อมูลได้อย่างมีประสิทธิภาพ
- บุคลากรสาธารณสุขที่เข้าอบรม สามารถวิเคราะห์ข้อมูลเพื่อค้นหาความสัมพันธ์และแนวโน้มต่างๆ ที่ซ่อนอยู่ในข้อมูลได้ถูกต้อง
- บุคลากรสาธารณสุขที่เข้าอบรม สามารถสร้างกระแสนงานเพื่อการประมวลผลข้อมูลได้
- บุคลากรสาธารณสุขที่เข้าอบรม สามารถนำเสนอข้อมูลในรูปแบบต่างๆ ได้เข้าใจง่าย

ระดับเจตคติ

- บุคลากรสาธารณสุขที่เข้าอบรม ตระหนักถึงความสำคัญของการจัดการและการวิเคราะห์ข้อมูล
- บุคลากรสาธารณสุขที่เข้าอบรม มีทัศนคติที่ดีต่อการใช้ข้อมูลในการตัดสินใจ

6. กลุ่มเป้าหมาย

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล ออกแบบมาเพื่อบุคลากรด้านสาธารณสุขทุกระดับ ที่ต้องการพัฒนาความรู้และทักษะในการจัดการและการวิเคราะห์ข้อมูล เพื่อนำไปใช้ประโยชน์ในการทำงานด้านสาธารณสุข ซึ่งสามารถรองรับผู้เข้าร่วมอบรม ในหลักสูตรได้ไม่เกิน 50 คน โดยคุณสมบัติของผู้เข้าร่วมอบรมในหลักสูตรควรมี ดังนี้

- 1) เป็นผู้ที่ทำงานด้านสาธารณสุข หรือสารสนเทศสาธารณสุข เช่น แพทย์ พยาบาล นักวิเคราะห์นโยบายและแผน นักวิชาการสาธารณสุข นักวิชาการคอมพิวเตอร์ หรือผู้ปฏิบัติงานทางด้านข้อมูล เป็นต้น
- 2) มีพื้นฐานความรู้ด้านการประมวลผลข้อมูลเบื้องต้น หรือมีประสบการณ์การทำงานด้านการวิเคราะห์ข้อมูล
- 3) มีความสนใจและต้องการพัฒนาความรู้และทักษะในการจัดการและการวิเคราะห์ข้อมูล

บุคลากรสาธารณสุขที่สมัครเข้าอบรมหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชา วิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล จะต้องผ่านการสอบหลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล ตามเกณฑ์คะแนนการคัดเลือกไม่น้อยกว่าร้อยละ 80 จึงจะได้เข้าอบรมอยู่ในหลักสูตรการจัดการและการวิเคราะห์ ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูลต่อไป

7. คุณสมบัติวิทยากร

- 1) จบการศึกษาระดับปริญญาตรีหรือสูงกว่า ในสาขาวิชาวิทยาการคอมพิวเตอร์ วิศวกรรมคอมพิวเตอร์ วิทยาการข้อมูล เทคโนโลยีสารสนเทศ หรือสาขาอื่นๆ ที่เกี่ยวข้อง
- 2) มีประสบการณ์การทำงานด้านวิทยาศาสตร์ข้อมูลหรือวิศวกรรมข้อมูลดังนี้
 - (1) จบการศึกษาระดับปริญญาตรี มีประสบการณ์ทำงานไม่น้อยกว่า 5 ปี
 - (2) จบการศึกษาระดับปริญญาโท มีประสบการณ์ทำงานไม่น้อยกว่า 3 ปี
 - (3) จบการศึกษาระดับปริญญาเอก มีประสบการณ์ทำงานไม่น้อยกว่า 2 ปี
- 3) มีประสบการณ์การสอนหรือจัดอบรมด้านวิทยาศาสตร์ข้อมูลหรือวิศวกรรมข้อมูลให้กับหน่วยงาน ราชการหรือเอกชน

8. สถานที่จัดการอบรม

การจัดการอบรมจะเป็นการจัดบรรยายในห้องอบรม ซึ่งห้องอบรมควรมีคุณสมบัติดังต่อไปนี้

- 1) มีขนาดเพียงพอสำหรับรองรับผู้เข้าอบรม โดยให้มีระยะห่างระหว่างโต๊ะและเก้าอี้เพียงพอให้ผู้เข้าอบรมสามารถมองเห็นวิทยากรและกระดานได้อย่างสะดวก
- 2) มีสภาพแวดล้อมที่เอื้อต่อการเรียนรู้ สะอาด สว่าง โปร่งสบาย และมีการระบายอากาศที่ดี
- 3) มีอุปกรณ์อำนวยความสะดวกในการเรียนการสอน ได้แก่
 - (1) โปรเจคเตอร์และฉาก หรือจอดิจิทัลพร้อมอุปกรณ์เชื่อมต่อ
 - (2) เครื่องขยายเสียง
 - (3) ปลั๊กไฟ (ให้เพียงพอกับผู้เข้าอบรม)
 - (4) อินเทอร์เน็ตไร้สาย

หมวดที่ 2 ระบบการจัดการอบรม การดำเนินการ และโครงสร้างของหลักสูตร

1. รูปแบบของหลักสูตร

หลักสูตรแกนกลาง (The Core Curriculum)

2. การดำเนินการหลักสูตร

- การจัดการอบรม

การจัดการอบรมเป็นรูปแบบหลักสูตรรายวิชา แบ่งออกเป็น 3 ระยะ ดังนี้

ระยะที่ 1 หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล 24 ชั่วโมง

(Pre-basic course on Data Management and Analysis)

ระยะที่ 2 หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล (Data Management and Analysis – Data Science and Data Engineer)

แบ่งเป็น หลักสูตรรายวิชาขั้นพื้นฐาน 36 ชั่วโมง

(Basic course)

หลักสูตรรายวิชาชั้นกลาง 42 ชั่วโมง

(Intermediate course)

หลักสูตรรายวิชาชั้นสูง 42 ชั่วโมง

(Advance course)

รวม 120 ชั่วโมง

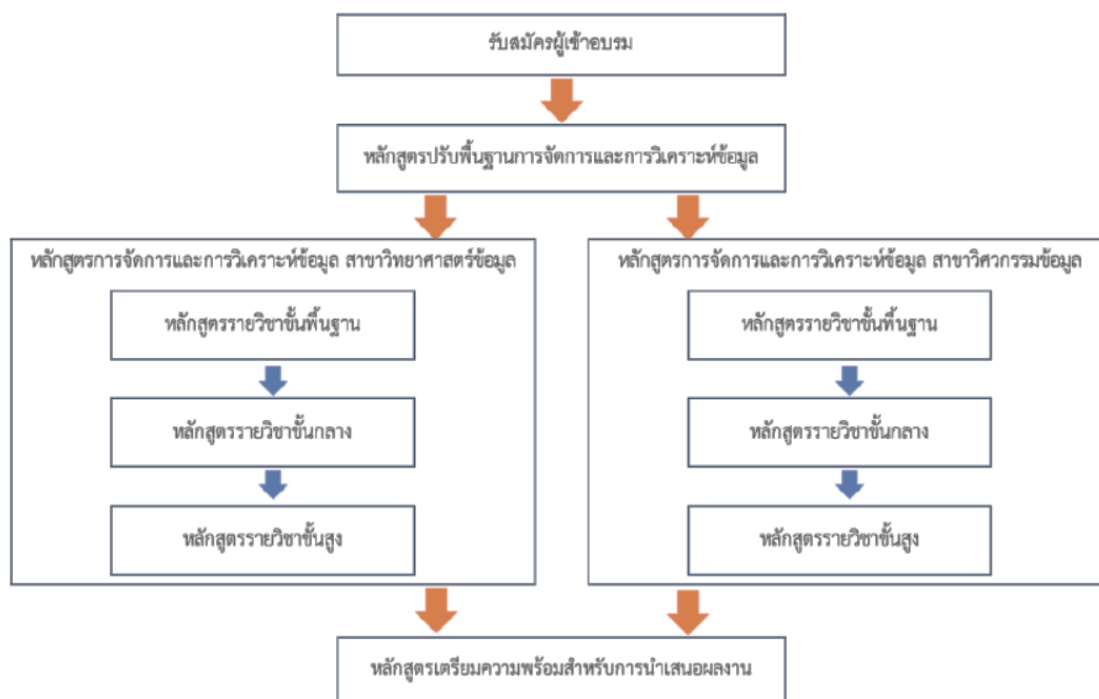
ระยะที่ 3 หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน 12 ชั่วโมง

(Preparation for Project Presentation)

รวม 3 ระยะ 156 ชั่วโมง

- โครงสร้างการจัดการอบรม

การจัดการอบรมเริ่มจากการรับสมัครผู้เข้าอบรม เพื่อเข้ารับการอบรมในระยะเวลาที่ 1 หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล หลังจากการอบรมในระยะเวลาที่ 1 แล้ว ผู้เข้าอบรมจะถูกแบ่งเข้ารับการอบรมในระยะเวลาที่ 2 ตามสาขาวิชาที่ได้เลือกไว้ หลังจากนั้นจะกลับมารับการอบรมรวมกันอีกครั้ง ในระยะเวลาที่ 3 หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน โดยจะมีการทดสอบหลังเรียน เพื่อผ่านเข้าสู่การอบรมรอบถัดไป



หลังจากรับการอบรมในแต่ละระยะ ผู้เข้ารับการอบรมจะต้องผ่านการทดสอบหลังเรียน เพื่อประเมินความรู้และความพร้อมสำหรับการเรียนในระยะเวลาถัดไป โดยการจัดการทดสอบหลังเรียน จะมีด้วยกัน 4 ครั้ง ดังนี้

1. การทดสอบหลังเรียน หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล
2. การทดสอบหลังเรียน หลักสูตรรายวิชาขั้นพื้นฐาน
3. การทดสอบหลังเรียน หลักสูตรรายวิชาขั้นกลาง
4. การทดสอบหลังเรียน หลักสูตรรายวิชาขั้นสูง

โดยผู้เข้ารับการอบรมจะต้องผ่านการประเมินตามเกณฑ์ที่ได้กำหนดไว้ในหมวดที่ 3 เกี่ยวกับการประเมินผู้เข้าอบรม จึงจะสามารถเข้ารับการอบรมในหลักสูตรถัดไปได้

3. โครงสร้างรายวิชาในหลักสูตร

ตาราง 2.1 หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล จำนวน 3 รายวิชา

| รายวิชา | รหัสวิชา | ชื่อวิชา | จำนวน ชั่วโมง |
|---------------------------------------|----------|---|---------------|
| 1. รายวิชาปรับพื้นฐาน (Pre-sessional) | | | 24 |
| 1 | PS100 | ความรู้เบื้องต้นเกี่ยวกับการใช้ข้อมูลประกอบการตัดสินใจ Introduction to Data-Driven Decision Making | 6 |
| 2 | PS101 | การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น Basic Python Programming for Data Analysis | 12 |
| 3 | PS102 | ฐานข้อมูลและระบบจัดการฐานข้อมูล Database and Relational Database Management System | 6 |

ตาราง 2.2 หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูล จำนวน 14 รายวิชา

| รายวิชา | รหัสวิชา | ชื่อวิชา | จำนวน ชั่วโมง |
|-------------------------------|----------|---|---------------|
| 1. รายวิชาขั้นพื้นฐาน (Basic) | | | 36 |
| 1 | DS100 | ความรู้เบื้องต้นเกี่ยวกับวิทยาศาสตร์ข้อมูล Introduction to Data Science | 6 |
| 2 | DS101 | การโปรแกรมภาษาไพธอนสำหรับวิทยาศาสตร์ข้อมูล Python Programming for Data Science | 12 |
| 3 | DS111 | การสำรวจข้อมูล Data Exploration | 6 |
| 4 | DS112 | การเตรียมข้อมูล Data Preparation | 6 |
| 5 | DS113 | การนำเสนอข้อมูล Data Visualization | 6 |

| | | | |
|--|-------|--|-----------|
| 2. รายวิชาชั้นกลาง (Intermediate) | | | 42 |
| 6 | DS201 | สถิติสำหรับวิทยาศาสตร์ข้อมูล Statistics for Data Science | 12 |
| 7 | DS211 | การนำเข้าข้อมูล Data Ingestion | 6 |
| 8 | DS212 | การเรียนรู้ของเครื่อง Machine Learning | 24 |
| 3. รายวิชาชั้นสูง (Advance) | | | 42 |
| 9 | DS301 | ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ Introduction to Big Data | 6 |
| 10 | DS311 | การบูรณาการกับระบบโครงสร้างพื้นฐาน Integration with Infrastructure | 6 |
| 11 | DS411 | การประมวลผลข้อมูลสารสนเทศภูมิศาสตร์ Geographic Information Processing | 6 |
| 12 | DS412 | การวิเคราะห์ข้อมูลอนุกรมเวลา Time Series Analysis | 12 |
| 13 | DS413 | การประมวลผลข้อความ Text Processing | 6 |
| 14 | DS414 | การเรียนรู้เชิงลึก Deep Learning | 6 |

ตาราง 2.3 หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูล จำนวน 16 รายวิชา

| รายวิชา | รหัสวิชา | ชื่อวิชา | จำนวน ชั่วโมง |
|--------------------------------------|----------|---|------------------|
| 1. รายวิชาขั้นพื้นฐาน (Basic) | | | 36 |
| 1 | DE100 | ความรู้เบื้องต้นเกี่ยวกับวิศวกรรมข้อมูล Introduction to Data Engineer | 6 |
| 2 | DE101 | การเขียนโปรแกรมภาษาไพธอนเบื้องต้นสำหรับวิศวกรข้อมูล Python Programming for Data Engineer | 12 |
| 3 | DE102 | ระบบปฏิบัติการ Linux เบื้องต้น Basic Linux Operating System | 6 |

| | | | |
|--|-------|--|-----------|
| 4 | DE103 | การสำรวจข้อมูล Data Exploration | 6 |
| 5 | DE104 | การเตรียมข้อมูล Data Preparation | 6 |
| 2. รายวิชาชั้นกลาง (Intermediate) | | | 42 |
| 6 | DE201 | ระบบการจัดการฐานข้อมูลเชิงสัมพันธ์ขั้นสูง Advance Relational Database Management System | 12 |
| 7 | DE202 | กระบวนการ ETL ETL Process | 18 |
| 8 | DE203 | การจัดการกระแสนงาน Workflow Management | 6 |
| 9 | DE204 | การจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล Managing development environment for data-related applications | 6 |
| 3. รายวิชาชั้นสูง (Advance) | | | 42 |
| 10 | DE301 | ไมโครเซอร์วิส Microservice | 6 |
| 11 | DE302 | การแลกเปลี่ยนข้อมูลและ API Gateway Data Exchange and API Gateway | 6 |
| 12 | DE303 | ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ Introduction to Big Data | 6 |
| 13 | DE304 | การนำเข้าข้อมูลในระบบข้อมูลขนาดใหญ่ Data Ingestion in Big Data | 6 |
| 14 | DE305 | การเตรียมข้อมูลในระบบข้อมูลขนาดใหญ่ Data Preparation in Big Data | 6 |
| 15 | DE306 | การให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่ Data Service in Big Data | 6 |
| 16 | DE307 | การออกแบบระบบข้อมูลขนาดใหญ่ Big Data Architectural Design | 6 |

ตาราง 2.4 หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน จำนวน 2 รายวิชา

| รายวิชา | รหัสวิชา | ชื่อวิชา | จำนวน ชั่วโมง |
|---------|----------|---|---------------|
| 1 | PR100 | การนำเสนอผลงานอย่างมีประสิทธิภาพ Pitching and Effective Presentation | 6 |
| 2 | PR101 | ธรรมาภิบาลข้อมูล Data Governance | 6 |

4. คำอธิบายรายวิชา

1. หลักสูตรปรับพื้นฐานการจัดการและการวิเคราะห์ข้อมูล

PS100 ความรู้เบื้องต้นเกี่ยวกับการใช้ข้อมูลประกอบการตัดสินใจ 6 ชั่วโมง

(Introduction to Data-Driven Decision Making)

ความรู้เบื้องต้นเกี่ยวกับการวิเคราะห์ข้อมูล รวมถึงการอภิปรายเทคนิคทั่วไปที่สำคัญสำหรับการวิเคราะห์และแปลงข้อมูลให้ได้มาซึ่งสารสนเทศที่มีความหมายจากชุดข้อมูลต่างๆ อธิบายหลักการที่เกี่ยวข้องกับการดำเนินการเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล การตัดสินใจจากข้อมูล โดยเน้นการคิดเชิงวิเคราะห์ บทบาทและขั้นตอนการทำงานของบุคลากรด้านข้อมูลในองค์กร ทั้งนักวิทยาศาสตร์ข้อมูล วิศวกรข้อมูล และนักวิเคราะห์ข้อมูล รวมถึงแนะนำตัวอย่างการใช้งานข้อมูลประกอบการตัดสินใจที่เกี่ยวข้องกับงานด้านสาธารณสุข

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการที่เกี่ยวข้องกับการเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล และการตัดสินใจจากข้อมูล
2. เพื่อให้ผู้เรียนสามารถคิดเชิงวิเคราะห์และตัดสินใจจากข้อมูล
3. เพื่อให้ผู้เรียนเข้าใจบทบาทและขั้นตอนการทำงานของบุคลากรด้านข้อมูลในองค์กร

ตาราง 2.5 เนื้อหาราย 3 ชั่วโมง ของรายวิชา PS100 ความรู้เบื้องต้นเกี่ยวกับการใช้ข้อมูลประกอบการตัดสินใจ

(Introduction to Data-Driven Decision Making)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|-------------|--|--------|
| 1-3 | <p>ความรู้เบื้องต้นเกี่ยวกับการวิเคราะห์ข้อมูล</p> <ul style="list-style-type: none"> - ความสำคัญของการนำข้อมูลไปใช้ประกอบการตัดสินใจ - ตัวอย่างการใช้งานข้อมูลประกอบการตัดสินใจที่เกี่ยวข้องกับงานด้านสาธารณสุข - แนะนำการนำข้อมูลไปใช้ประกอบการตัดสินใจ โดย Guest Speaker จากกระทรวงสาธารณสุข | บรรยาย |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|--------|
| 4-6 | หลักการที่เกี่ยวข้องกับการวิเคราะห์ข้อมูล <ul style="list-style-type: none"> - บทบาทและขั้นตอนการทำงานของบุคลากรด้านข้อมูล - ภาพรวมขั้นตอนการทำงานกับข้อมูล | บรรยาย |

PS101 การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น

12 ชั่วโมง

(Basic Python Programming for Data Analysis)

แนะนำให้ผู้รู้จักกับภาษาไพธอน หลักการเบื้องต้นของการเขียนโปรแกรม คำสั่งการทำงานพื้นฐาน เช่น การแสดงผล การรับข้อมูล การสร้างตัวแปร แนะนำให้ผู้รู้จักและทดลองใช้งานไลบรารีสำหรับการวิเคราะห์ข้อมูล โดยเรียนรู้จากกระบวนการทำงานวิเคราะห์ข้อมูลทั่วไป

วัตถุประสงค์

1. เพื่อให้ผู้เรียนรู้จักกับภาษาไพธอนและหลักการเบื้องต้นของการเขียนโปรแกรม
2. เพื่อให้ผู้เรียนรู้จักและทดลองใช้งานไลบรารีสำหรับการวิเคราะห์ข้อมูล
3. เพื่อให้ผู้เรียนเข้าใจกระบวนการทำงานวิเคราะห์ข้อมูลทั่วไป

ตาราง 2.6 เนื้อหาราย 3 ชั่วโมง ของรายวิชา PS101 การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น (Basic Python Programming for Data Analysis)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | แนะนำให้ผู้รู้จักกับภาษาไพธอน <ul style="list-style-type: none"> - ติดตั้งโปรแกรมสำหรับการเขียนภาษาไพธอน - หลักการเบื้องต้นของการเขียนโปรแกรม - คำสั่งในการทำงานพื้นฐาน | บรรยาย + ปฏิบัติ |
| 4-6 | การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น 1 <ul style="list-style-type: none"> - คำสั่งในการทำงานพื้นฐาน (ต่อ) - การใช้งานภาษาไพธอนในการวิเคราะห์ข้อมูล - แนะนำไลบรารีสำหรับการวิเคราะห์ข้อมูล - ขั้นตอนการใช้งานไลบรารีสำหรับการวิเคราะห์ข้อมูล | บรรยาย + ปฏิบัติ |
| 7-9 | การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น 2 <ul style="list-style-type: none"> - การใช้งาน Pandas สำหรับการทำงานกับข้อมูลตาราง <ul style="list-style-type: none"> ○ การอ่านไฟล์ข้อมูล ○ การสำรวจข้อมูล | บรรยาย + ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 10-12 | การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ข้อมูลเบื้องต้น 3 <ul style="list-style-type: none"> - การใช้งาน Pandas สำหรับการทำงานกับข้อมูลตาราง (ต่อ) <ul style="list-style-type: none"> ○ การแปลงข้อมูล ○ การบันทึกข้อมูล | บรรยาย + ปฏิบัติ |

PS102 ฐานข้อมูลและระบบจัดการฐานข้อมูล

6 ชั่วโมง

(Database and Relational Database Management System)

แนะนำหลักการของฐานข้อมูล รูปแบบของ Structured Query Language เบื้องต้น โดยจะเน้นไปที่การเลือกข้อมูลจากฐานข้อมูลที่มีอยู่แล้ว ลงมือปฏิบัติและเขียนโปรแกรมเพื่อใช้งานและจัดการฐานข้อมูลผ่านโปรแกรมจัดการฐานข้อมูล และผ่านไลบรารีบนภาษาไพธอน

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการพื้นฐานของฐานข้อมูล
2. เพื่อให้ผู้เรียนสามารถเขียนโปรแกรม SQL เพื่อเลือกข้อมูลจากฐานข้อมูลที่มีอยู่แล้ว
3. เพื่อให้ผู้เรียนสามารถใช้งานโปรแกรมจัดการฐานข้อมูล และไลบรารีบนภาษาไพธอนเพื่อจัดการฐานข้อมูล

ตาราง 2.7 เนื้อหาราย 3 ชั่วโมง ของรายวิชา PS102 ฐานข้อมูลและระบบจัดการฐานข้อมูล (Database and Relational Database Management System)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | ฐานข้อมูลคืออะไร <ul style="list-style-type: none"> - หลักการของฐานข้อมูล - รูปแบบของฐานข้อมูล - ติดตั้งโปรแกรมที่ใช้จัดการฐานข้อมูล - Structured Query Language เบื้องต้น การใช้งานฐานข้อมูล <ul style="list-style-type: none"> - การใช้งานฐานข้อมูลผ่านทางโปรแกรมจัดการฐานข้อมูล | บรรยาย + ปฏิบัติ |
| 4-6 | การใช้งานฐานข้อมูล <ul style="list-style-type: none"> - การใช้งานฐานข้อมูลผ่านการเขียนโปรแกรมภาษาไพธอน Workshop การดึงข้อมูลจากฐานข้อมูล | บรรยาย + ปฏิบัติ |

2. หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูล

DS100 ความรู้เบื้องต้นเกี่ยวกับวิทยาศาสตร์ข้อมูล

6 ชั่วโมง

(Introduction to Data Science)

ความรู้เบื้องต้นเกี่ยวกับการวิเคราะห์ข้อมูล รวมถึงการอภิปรายเทคนิคทั่วไปที่สำคัญสำหรับการวิเคราะห์และแปลงข้อมูลให้ได้มาซึ่งสารสนเทศที่มีความหมายจากชุดข้อมูลต่างๆ อธิบายหลักการที่เกี่ยวข้องกับการดำเนินการเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล การตัดสินใจจากข้อมูล โดยเน้นการคิดเชิงวิเคราะห์ บทบาทของนักวิทยาศาสตร์ข้อมูลในองค์กร ขั้นตอนการทำงานของนักวิทยาศาสตร์ข้อมูล ความเกี่ยวข้องกับบุคลากรข้อมูลอื่นๆ เครื่องมือของนักวิทยาศาสตร์ข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการที่เกี่ยวข้องกับการเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล และการตัดสินใจจากข้อมูล
2. เพื่อให้ผู้เรียนเข้าใจบทบาทและขั้นตอนการทำงานของนักวิทยาศาสตร์ข้อมูล

ตาราง 2.8 รายละเอียดเนื้อหาของรายวิชา DS100 ความรู้เบื้องต้นเกี่ยวกับวิทยาศาสตร์ข้อมูล (Introduction to Data Science)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|--------|
| 1-3 | ความรู้เบื้องต้นเกี่ยวกับวิทยาศาสตร์ข้อมูล - ภาพรวมของการทำงานและบทบาทของนักวิทยาศาสตร์ข้อมูล - ความเกี่ยวข้องกับบุคลากรด้านข้อมูลอื่นๆ | บรรยาย |
| 4-6 | งานที่นักวิทยาศาสตร์ข้อมูลรับผิดชอบ - เครื่องมือของนักวิทยาศาสตร์ข้อมูล - กระบวนการทำงานของนักวิทยาศาสตร์ข้อมูล | บรรยาย |

DS101 การโปรแกรมภาษาไพธอนสำหรับวิทยาศาสตร์ข้อมูล

12 ชั่วโมง

(Python Programming for Data Science)

การเขียนโปรแกรมภาษาไพธอนเพื่องานด้านวิทยาศาสตร์ข้อมูล แนะนำและทดลองใช้งานไลบรารีที่เกี่ยวข้องกับงานด้านวิทยาศาสตร์ข้อมูล เช่น Pandas, Numpy, SQL Alchemy และการเขียนโปรแกรมเป็นโมดูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนสามารถเขียนโปรแกรมภาษาไพธอนเพื่องานด้านวิทยาศาสตร์ข้อมูล
2. เพื่อให้ผู้เรียนรู้จักและทดลองใช้งานไลบรารีที่เกี่ยวข้องกับงานด้านวิทยาศาสตร์ข้อมูล เช่น Pandas, Numpy, SQL Alchemy

ตาราง 2.9 รายละเอียดเนื้อหาของรายวิชา DS101 การโปรแกรมภาษาไพธอนสำหรับวิทยาศาสตร์ข้อมูล (Python Programming for Data Science)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|---|
| 1-3 | <p>แนะนำไลบรารีสำหรับงานด้านวิทยาศาสตร์ข้อมูล</p> <ul style="list-style-type: none"> - คุณสมบัติของไลบรารีต่างๆ - ตัวอย่างการใช้งานในสถานการณ์จริง - การติดตั้งไลบรารี <p>Numpy</p> <ul style="list-style-type: none"> - หลักการทำงาน - การทำงานกับ Numpy Array และ Matrix - ฟังก์ชันอื่นๆ ที่มีประโยชน์ (Random Number Generator, Statistics, Linear Algebra) | <p>บรรยาย</p> <p style="text-align: center;">+</p> <p>ปฏิบัติ</p> |
| 4-6 | <p>Pandas</p> <ul style="list-style-type: none"> - ชนิดของข้อมูลใน Pandas (Series และ DataFrame) - การทำงานกับ Pandas Series - การทำงานกับ Pandas DataFrame <ul style="list-style-type: none"> ○ การอ่านข้อมูลเป็น Pandas DataFrame ○ การสำรวจ DataFrame เบื้องต้น | <p>บรรยาย</p> <p style="text-align: center;">+</p> <p>ปฏิบัติ</p> |
| 7-9 | <p>SQLAlchemy</p> <ul style="list-style-type: none"> - การเชื่อมต่อกับฐานข้อมูล - การดึงข้อมูลจากฐานข้อมูล - การเขียนข้อมูลลงฐานข้อมูล | <p>บรรยาย</p> <p style="text-align: center;">+</p> <p>ปฏิบัติ</p> |
| 10-12 | <p>Workshop การสร้างกระแสด้านประมวลผลข้อมูลอย่างง่าย</p> | <p>บรรยาย</p> <p style="text-align: center;">+</p> <p>ปฏิบัติ</p> |

DS111 การสำรวจข้อมูล

6 ชั่วโมง

(Data Exploration)

หลักการของการสำรวจข้อมูล ทำความเข้าใจถึงโครงสร้างของข้อมูล โดยใช้วิธีการสำรวจแบบต่างๆ ที่เหมาะสมกับชนิดของข้อมูล เช่น ตารางแจกแจงความถี่ แผนภูมิแจกแจงความถี่ แผนภูมิก่อร่าง แผนภูมิแท่ง แผนภูมิจุด และค่าสถิติ เช่น ค่าเฉลี่ยเลขคณิต มัชยฐาน ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน เพื่อใช้ในการ

ตรวจสอบข้อผิดพลาด การตรวจสอบสมมติฐาน การคัดเลือกตัวแปร และการวิเคราะห์ความสัมพันธ์ โดยเน้นไปที่การเขียนโปรแกรมเพื่อทำการสำรวจข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการสำรวจข้อมูล
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการสำรวจข้อมูล
3. เพื่อให้ผู้เรียนสามารถเขียนโปรแกรมเพื่อทำการสำรวจข้อมูล

ตาราง 2.10 รายละเอียดเนื้อหาของรายวิชา DS111 การสำรวจข้อมูล (Data Exploration)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|---------------------------------------|
| 1-3 | <p>แนะนำการสำรวจข้อมูล</p> <ul style="list-style-type: none"> - ขั้นตอนทั่วไปของการสำรวจข้อมูล - โลบรารีที่ใช้ในการสำรวจข้อมูล (Pandas, Matplotlib, Seaborn) <p>การสำรวจข้อมูลตัวแปรเดียว</p> <ul style="list-style-type: none"> - แบบ Non-graphical (Frequency table, Range, Mean, Trimmed mean, Median, Variance, Standard Deviation, Interquartile Range) - แบบ Graphical (Countplot, Histogram) | <p>บรรยาย</p> <p>+</p> <p>ปฏิบัติ</p> |
| 4-6 | <p>การสำรวจข้อมูลหลายตัวแปร</p> <ul style="list-style-type: none"> - แบบ Non-graphical (Frequency table, Correlation coefficient) - แบบ Graphical (Bar plot, Box plot, Line plot, Scatter plot) | <p>บรรยาย</p> <p>+</p> <p>ปฏิบัติ</p> |

DS112 การเตรียมข้อมูล

6 ชั่วโมง

(Data Preparation)

การเตรียมข้อมูลเพื่อให้พร้อมสำหรับการวิเคราะห์ ตารางหน่วยการวิเคราะห์ การทำความสะอาดข้อมูล การจัดการข้อมูลค่าว่าง การจัดการข้อมูลกลุ่มย่อย การแปลงข้อมูลในแนวกว้าง-ยาว การสกัดคุณลักษณะ การสรุปข้อมูล การแปลงชนิดของข้อมูล การเชื่อมตาราง โดยเน้นไปที่การเขียนโปรแกรม เพื่อทำการเตรียมข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการเตรียมข้อมูลสำหรับการวิเคราะห์
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการเตรียมข้อมูล

ตาราง 2.11 รายละเอียดเนื้อหาของรายวิชา DS112 การเตรียมข้อมูล (Data Preparation)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | ขั้นตอนทั่วไปของการเตรียมข้อมูล การทำความสะอาดข้อมูล <ul style="list-style-type: none"> - สาเหตุของความไม่สะอาดของข้อมูล - การทำความสะอาดข้อมูลแบบต่างๆ (ข้อมูลค่าว่าง ข้อมูลผิดพลาด ข้อมูลกลุ่มย่อย และข้อมูล Outlier) | บรรยาย + ปฏิบัติ |
| 4-6 | การเตรียมข้อมูล <ul style="list-style-type: none"> - การเตรียมข้อมูลแบบต่างๆ (การสกัดคุณลักษณะ การสรุปข้อมูล การแปลงชนิดของข้อมูล การเชื่อมตาราง การสร้างข้อมูลโปรไฟล์) | บรรยาย + ปฏิบัติ |

DS113 การนำเสนอข้อมูล

6 ชั่วโมง

(Data Visualization)

การนำเสนอข้อมูลที่วิเคราะห์แล้วในรูปแบบต่างๆ การเลือกใช้กราฟที่เหมาะสมกับข้อมูล หลักการออกแบบการแสดงผล การใช้งานโปรแกรมสำหรับสร้างการแสดงผลเบื้องต้น

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและเทคนิคในการนำเสนอข้อมูล
2. เพื่อให้ผู้เรียนสามารถเลือกใช้กราฟที่เหมาะสมกับข้อมูล

ตาราง 2.12 รายละเอียดเนื้อหาของรายวิชา DS113 การนำเสนอข้อมูล (Data Visualization)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | หลักการออกแบบการแสดงผล การเล่าเรื่องโดยใช้ข้อมูล ตัวอย่างการนำเสนอข้อมูล | บรรยาย |
| 4-6 | การใช้งานโปรแกรม Tableau เบื้องต้น <ul style="list-style-type: none"> - แนะนำโปรแกรมสำหรับสร้างการแสดงผล - ติดตั้งโปรแกรม Tableau - การใช้งาน Tableau เบื้องต้น <ul style="list-style-type: none"> ○ การเชื่อมต่อกับแหล่งข้อมูล ○ การสร้าง Visualization และ Dashboard | บรรยาย + ปฏิบัติ |

(Statistics for Data Science)

ความรู้เบื้องต้นเกี่ยวกับสถิติที่จำเป็นสำหรับนักวิทยาศาสตร์ข้อมูล ทฤษฎีบทความน่าจะเป็น ทฤษฎีบทของเบย์ สถิติเชิงพรรณนา (ค่าเฉลี่ย มัธยฐาน ฐานนิยม ความแปรปรวน และค่าเบี่ยงเบนมาตรฐาน) สถิติเชิงอนุมาน (การประเมินค่าพารามิเตอร์ในประชากร การทดสอบสมมติฐาน ความคลาดเคลื่อนในการทดสอบสมมติฐาน การประมาณแบบช่วง การประเมินคุณสมบัติและการควบคุมความผิดพลาดของการทดสอบสมมติฐาน) โดยทำการเขียนโปรแกรมควบคู่ไปกับการฟังบรรยาย

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจพื้นฐานความรู้ทางสถิติที่จำเป็นสำหรับนักวิทยาศาสตร์ข้อมูล
2. เพื่อให้ผู้เรียนสามารถตีความผลลัพธ์ทางสถิติได้อย่างถูกต้อง

ตาราง 2.13 รายละเอียดเนื้อหาของรายวิชา DS201 สถิติสำหรับวิทยาศาสตร์ข้อมูล (Statistics for Data Science)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | สถิติเบื้องต้น ชนิดของข้อมูล รูปแบบของตัวแปร การสุ่มตัวอย่าง | บรรยาย + ปฏิบัติ |
| 4-6 | ทฤษฎีความน่าจะเป็น ตัวแปรสุ่ม สถิติเชิงพรรณนา | บรรยาย + ปฏิบัติ |
| 7-9 | สถิติเชิงอนุมาน การทดสอบสมมติฐาน <ul style="list-style-type: none"> - z-test - t-test (one sample, two sample, paired) - chi-squared test - Analysis of Variance | บรรยาย + ปฏิบัติ |
| 10-12 | การวิเคราะห์ถดถอยเชิงเส้น <ul style="list-style-type: none"> - Single Linear Regression - Confidence Interval - Evaluation of Regression Model - Multiple Regression | บรรยาย + ปฏิบัติ |

DS211 การนำเข้าข้อมูล

6 ชั่วโมง

(Data Ingestion)

แหล่งข้อมูลและการนำข้อมูลเข้ามาประมวลผลจากแหล่งต่างๆ เช่น ไฟล์ ฐานข้อมูล API หรือหน้าเว็บ การจัดการกับข้อมูลรูปแบบต่างๆ การเลือกใช้วิธีที่เหมาะสมกับแหล่งข้อมูล โดยเน้นไปที่การเขียนโปรแกรมเพื่อนำข้อมูลเข้ามาประมวลผล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับแหล่งข้อมูลประเภทต่างๆ
2. เพื่อให้ผู้เรียนสามารถดึงข้อมูลจากแหล่งต่างๆ โดยเลือกใช้วิธีที่เหมาะสมกับแหล่งข้อมูล

ตาราง 2.14 รายละเอียดเนื้อหาของรายวิชา DS211 การนำเข้าข้อมูล (Data Ingestion)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | การนำเข้าข้อมูลจากไฟล์ การนำเข้าข้อมูลจากฐานข้อมูล | บรรยาย + ปฏิบัติ |
| 4-6 | การนำเข้าข้อมูลจาก API การนำเข้าข้อมูลจากหน้าเว็บ | บรรยาย + ปฏิบัติ |

DS212 การเรียนรู้ของเครื่อง

24 ชั่วโมง

(Machine Learning)

พื้นฐานทางทฤษฎีของการเรียนรู้ของเครื่องด้วยตัวอย่างและแบบฝึกหัดการลงมือเขียนโปรแกรม ขั้นตอนที่สำคัญสำหรับการสร้างแบบจำลองการเรียนรู้ของเครื่อง การเรียนรู้ของเครื่องในรูปแบบต่างๆ ทั้งการเรียนรู้แบบมีผู้สอนและไม่มีผู้สอน ต้นไม้ตัดสินใจ การวิเคราะห์ถดถอยเชิงเส้น ซัพพอร์ตเวกเตอร์แมชชีน การเรียนรู้แบบกลุ่ม โคจรข่ายประสาทเทียม การแบ่งกลุ่ม K-means การประเมินผลแบบจำลอง และการปรับแต่งพารามิเตอร์ การคัดเลือกตัวแปร การลดมิติของข้อมูล โดยเรียนรู้จากการเขียนโปรแกรม

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจพื้นฐานทางทฤษฎีของการเรียนรู้ของเครื่อง
2. เพื่อให้ผู้เรียนสามารถสร้างแบบจำลองการเรียนรู้ของเครื่องได้
3. เพื่อให้ผู้เรียนสามารถประเมินผลแบบจำลองและปรับแต่งพารามิเตอร์ได้

ตาราง 2.15 รายละเอียดเนื้อหาของรายวิชา DS212 การเรียนรู้ของเครื่อง (Machine Learning)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | พื้นฐานทางทฤษฎีของ Machine Learning ขั้นตอนการสร้างแบบจำลอง แนะนำไลบรารีสำหรับงาน Machine Learning | บรรยาย + ปฏิบัติ |
| 4-6 | แบบจำลองจำแนกกลุ่ม (Classification Model) - Decision Tree - Linear Discriminant Analysis - Support Vector Machine - Artificial Neural Network | บรรยาย + ปฏิบัติ |
| 7-9 | แบบจำลองประมาณค่า (Regression Model) - Linear Regression - Poisson Regression - Decision Tree Regression | บรรยาย + ปฏิบัติ |
| 10-12 | การเรียนรู้แบบกลุ่ม (Ensemble Model) - Bootstrap Aggregation - Boosting - Stacking | บรรยาย + ปฏิบัติ |
| 13-15 | แบบจำลองการแบ่งกลุ่มและความเหมือน (Clustering Model and Similarity) - Similarity measurement - K-Means Clustering - Variation of K-Means Clustering - Hierarchical Clustering | บรรยาย + ปฏิบัติ |
| 16-18 | Survival Analysis | บรรยาย + ปฏิบัติ |
| 19-21 | การตรวจสอบคุณภาพของแบบจำลอง - Model evaluation - Parameter Tuning | บรรยาย + ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 22-24 | การคัดเลือกตัวแปรและการลดมิติของข้อมูล <ul style="list-style-type: none"> - Forward selection - Backward selection - Regularization - Principal Component Analysis | บรรยาย + ปฏิบัติ |

DS301 ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่

6 ชั่วโมง

(Introduction to Big Data)

พื้นฐานของการวิเคราะห์ข้อมูลขนาดใหญ่ เอกลักษณ์ของข้อมูลขนาดใหญ่ ระบบจัดการข้อมูลขนาดใหญ่ Hadoop Ecosystem หลักการของการประมวลผลแบบกระจาย เครื่องมือต่างๆ ในระบบจัดการข้อมูลขนาดใหญ่ โดยเน้นให้เห็นถึงภาพรวมของการจัดการและประมวลผลข้อมูลขนาดใหญ่

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนได้รับประสบการณ์การใช้งานเครื่องมือต่างๆ ในระบบจัดการข้อมูลขนาดใหญ่

ตาราง 2.16 รายละเอียดเนื้อหาของรายวิชา DS301 ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ (Introduction to Big Data)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | แนะนำระบบข้อมูลขนาดใหญ่ เครื่องมือต่างๆ ในระบบข้อมูลขนาดใหญ่ Hadoop Distributed File System | บรรยาย |
| 4-6 | Apache Hive <ul style="list-style-type: none"> - การใช้งาน Apache Hive เบื้องต้น Apache Spark <ul style="list-style-type: none"> - Spark RDD - Spark Data Frame - Spark Machine Learning | บรรยาย + ปฏิบัติ |

DS311 การบูรณาการกับระบบโครงสร้างพื้นฐาน

6 ชั่วโมง

(Integration with Infrastructure)

การควบคุมเวอร์ชันของซอร์สโค้ด การควบคุมเวอร์ชันของไลบรารี การออกแบบกระแสนงานของแบบจำลอง การบันทึกแบบจำลองและการอ่านแบบจำลองมาใช้งาน

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับการควบคุมเวอร์ชันของซอร์สโค้ดและไลบรารี
2. เพื่อให้ผู้เรียนสามารถบันทึกและออกแบบจำลองมาใช้งานได้

ตาราง 2.17 รายละเอียดเนื้อหาของรายวิชา DS311 การบูรณาการกับระบบโครงสร้างพื้นฐาน (Integration with Infrastructure)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|--|------------------------|
| 1-3 | การควบคุมเวอร์ชันของซอร์สโค้ด การควบคุมเวอร์ชันของไลบรารี | บรรยาย + ปฏิบัติ |
| 4-6 | การออกแบบกระแสนงานของแบบจำลอง การบันทึกแบบจำลองและการอ่านแบบจำลองมาใช้งาน | บรรยาย + ปฏิบัติ |

DS411 การประมวลผลข้อมูลสารสนเทศภูมิศาสตร์

6 ชั่วโมง

(Geographic Information Processing)

คุณสมบัติของข้อมูลสารสนเทศภูมิศาสตร์ การทำงานกับข้อมูลสารสนเทศภูมิศาสตร์ เช่น Shapefile GeoJSON GeoDF การสร้างการแสดงผล การวิเคราะห์เชิงพื้นที่ การวิเคราะห์โครงข่ายและเส้นทาง การดึงข้อมูลเชิงพื้นที่ จาก Google Earth Engine เน้นไปที่ การเขียนโปรแกรมโดยใช้ไลบรารีสำหรับการวิเคราะห์ข้อมูลสารสนเทศภูมิศาสตร์

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจและทำงานกับข้อมูลสารสนเทศภูมิศาสตร์
2. เพื่อให้ผู้เรียนสามารถดึงข้อมูลเชิงพื้นที่และวิเคราะห์ข้อมูลสารสนเทศภูมิศาสตร์ได้

ตาราง 2.18 รายละเอียดเนื้อหาของรายวิชา DS411 การประมวลผลข้อมูลสารสนเทศภูมิศาสตร์ (Geographic Information Processing)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | คุณสมบัติของข้อมูลสารสนเทศภูมิศาสตร์ แนะนำไลบรารีสำหรับทำงานกับข้อมูลสารสนเทศภูมิศาสตร์ การอ่านและแสดงผลไฟล์ข้อมูลสารสนเทศภูมิศาสตร์ | บรรยาย + ปฏิบัติ |
| 4-6 | การดึงข้อมูลเชิงพื้นที่จาก Google Earth Engine การวิเคราะห์ข้อมูลสารสนเทศภูมิศาสตร์ <ul style="list-style-type: none"> - Geometric Operations - Spatial Analysis - Routing and Network Analysis | บรรยาย + ปฏิบัติ |

DS412 การวิเคราะห์ข้อมูลอนุกรมเวลา

12 ชั่วโมง

(Time Series Analysis)

คุณสมบัติของข้อมูลอนุกรมเวลา การปรับหน่วยเวลา Autocorrelation ขั้นตอนวิธีสำหรับการทำนายค่าในอนาคต เช่น ARMA, ARIMA, ARIMAX, SARIMA และ SARIMAX เน้นไปที่การเขียนโปรแกรมโดยใช้ไลบรารีสำหรับการวิเคราะห์ข้อมูลอนุกรมเวลา

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจคุณสมบัติของข้อมูลอนุกรมเวลา
2. เพื่อให้ผู้เรียนสามารถวิเคราะห์ข้อมูลอนุกรมเวลาโดยใช้แบบจำลองที่เหมาะสม

ตาราง 2.19 รายละเอียดเนื้อหาของรายวิชา DS412 การวิเคราะห์ข้อมูลอนุกรมเวลา (Time Series Analysis)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | คุณสมบัติของข้อมูลอนุกรมเวลา การแสดงผลข้อมูลอนุกรมเวลา การปรับหน่วยเวลา | บรรยาย + ปฏิบัติ |
| 4-6 | Time Series Models and Forecasting <ul style="list-style-type: none"> - Autocorrelation - Fourier Transform - ARMA - ARIMA - Exponential Smoothing | บรรยาย + ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| | - Holt-Winters | |
| 7-9 | Advance Time Series Model - Seasonal effects - Exogenous variables - Interrupted Time Series | บรรยาย + ปฏิบัติ |
| 10-12 | Model Evaluation and Accuracy Metrics - Forecast evaluation measures (MAE, RMSE) - In-sample vs. out-of-sample validation | บรรยาย + ปฏิบัติ |

DS413 การประมวลผลข้อความ

6 ชั่วโมง

(Text Processing)

คุณสมบัติของข้อมูลข้อความ การจัดการข้อมูลข้อความ การเปรียบเทียบข้อความ Regular Expression และ Fuzzy Matching การแทนที่ข้อความด้วยตัวเลข ระยะห่างโคไซน์ การจัดกลุ่มข้อความ เน้นไปที่การเขียนโปรแกรมโดยใช้ไลบรารีสำหรับการวิเคราะห์ข้อมูลข้อความ

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจคุณสมบัติของข้อมูลข้อความ
2. เพื่อให้ผู้เรียนมีทักษะการจัดการข้อมูลข้อความโดยเทคนิคที่เหมาะสม

ตาราง 2.20 รายละเอียดเนื้อหาของรายวิชา DS413 การประมวลผลข้อความ (Text Processing)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | คุณสมบัติของข้อมูลข้อความ Regular Expression Text Preprocessing | บรรยาย + ปฏิบัติ |
| 4-6 | Text Similarity - Levenshtein distance - Fuzzy string matching - Document-term matrix - Cosine similarity | บรรยาย + ปฏิบัติ |

(Deep Learning)

Gradient Descent โครงข่ายประสาทเทียม Backpropagation การออกแบบโครงสร้างของโครงข่ายประสาทเทียม โครงสร้างของโครงข่ายประสาทเทียมแบบต่างๆ Convolutional Neural Network และ Recurrent Neural Network

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการทำงานของโครงข่ายประสาทเทียม
2. เพื่อให้ผู้เรียนมีความรู้เกี่ยวกับโครงสร้างของโครงข่ายประสาทเทียมแบบต่างๆ

ตาราง 2.21 รายละเอียดเนื้อหาของรายวิชา DS414 การเรียนรู้เชิงลึก (Deep Learning)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|---|------------------------|
| 1-3 | พื้นฐานของการเรียนรู้เชิงลึก <ul style="list-style-type: none"> - Gradient Descent - Feed Forward Network - Backpropagation - Regularization | บรรยาย + ปฏิบัติ |
| 4-6 | แนะนำโครงสร้างของ Deep Learning แบบต่างๆ <ul style="list-style-type: none"> - Convolutional Neural Network - Recurrent Neural Network - โครงสร้างขั้นสูง (Encoder-Decoder, Generative Adversarial Network, Transformer, Attention) | บรรยาย + ปฏิบัติ |

3. หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิศวกรรมข้อมูล

DE100 ความรู้เบื้องต้นเกี่ยวกับวิศวกรรมข้อมูล

6 ชั่วโมง

(Introduction to Data Engineer)

ภาพรวมของการทำงานของบุคลากรด้านข้อมูลในองค์กร บทบาทของวิศวกรข้อมูลในองค์กร ขั้นตอนการทำงานของวิศวกรข้อมูล ความเกี่ยวข้องกับบุคลากรด้านข้อมูลอื่นๆ เครื่องมือของวิศวกรข้อมูล งานที่วิศวกรข้อมูลรับผิดชอบ เช่น การเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการทำงานที่เกี่ยวข้องกับการเก็บรวบรวมข้อมูล การแปลงข้อมูล การทำความสะอาดข้อมูล และการตัดสินใจจากข้อมูล
2. เพื่อให้ผู้เรียนเข้าใจบทบาทและขั้นตอนการทำงานของวิศวกรข้อมูล

ตาราง 2.22 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE100 ความรู้เบื้องต้นเกี่ยวกับวิศวกรรมข้อมูล (Introduction to Data Engineer)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|--------|
| 1-3 | ความรู้เบื้องต้นเกี่ยวกับวิศวกรรมข้อมูล <ul style="list-style-type: none"> - ภาพรวมของการทำงานและบทบาทของวิศวกรข้อมูล - ความเกี่ยวข้องกับบุคลากรด้านข้อมูลอื่นๆ | บรรยาย |
| 4-6 | งานที่วิศวกรข้อมูลรับผิดชอบ <ul style="list-style-type: none"> - เครื่องมือของวิศวกรข้อมูล - ภาพรวมกระบวนการ ETL เบื้องต้น - ความสำคัญของวิศวกรรมข้อมูลในงานสาธารณสุข | บรรยาย |

หมายเหตุ : หัวข้อ “ความสำคัญของวิศวกรรมข้อมูลในงานสาธารณสุข” เป็นหัวข้อบรรยายพิเศษ โดยวิทยากรจากกระทรวงสาธารณสุข เป็นระยะเวลา 30 นาที

DE101 การโปรแกรมภาษาไพธอนสำหรับวิศวกรรมข้อมูล
(Python Programming for Data Engineer)

12 ชั่วโมง

การเขียนโปรแกรมภาษาไพธอนเพื่องานด้านวิศวกรรมข้อมูล แนะนำและทดลองใช้งานไลบรารีที่เกี่ยวข้องกับงานด้านวิศวกรรมข้อมูล เช่น Pandas, SQL Alchemy

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการการเขียนโปรแกรมภาษาไพธอนเพื่อจัดประมวลผลข้อมูล
2. เพื่อให้ผู้เรียนมีความรู้ในการใช้ไลบรารีที่เกี่ยวข้องกับการประมวลผลข้อมูล

ตาราง 2.23 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE101 การโปรแกรมภาษาไพธอนสำหรับวิศวกรรมข้อมูล (Python Programming for Data Engineer)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | การใช้งาน Pandas <ul style="list-style-type: none"> - Pandas Data Structure - Creating Data Frame | บรรยาย + ปฏิบัติ |
| 4-6 | การใช้งาน Pandas (ต่อ) <ul style="list-style-type: none"> - การเรียกดูคุณสมบัติของ Data Frame - การจัดการ Data Frame | บรรยาย + ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 7-9 | การใช้งาน SQL Alchemy <ul style="list-style-type: none"> - การเชื่อมต่อและจัดการ Database - SQL Expressions - Executing Expression | บรรยาย + ปฏิบัติ |
| 10-12 | Workshop: การเขียนโปรแกรม Python เพื่อแก้โจทย์ปัญหา โดยใช้ Pandas และ SQL Alchemy | บรรยาย + ปฏิบัติ |

DE102 ระบบปฏิบัติการ Linux เบื้องต้น

6 ชั่วโมง

(Basic Linux Operating System)

ทำความเข้าใจกับระบบปฏิบัติการ Linux โครงสร้างของระบบไฟล์ การใช้งาน Command Line Tools เช่น การดูรายการไฟล์ การสร้างไฟล์ การแก้ไขไฟล์ การลบไฟล์ การตรวจสอบสถานะของเครื่อง การติดตั้งโปรแกรม การจัดการผู้ใช้งาน การจัดการสิทธิ์

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการของระบบปฏิบัติการ Linux
2. เพื่อให้ผู้เรียนมีความเข้าใจในการใช้คำสั่งของระบบปฏิบัติการ Linux

ตาราง 2.24 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE102 ระบบปฏิบัติการ Linux เบื้องต้น (Basic Linux Operating System)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | แนะนำระบบปฏิบัติการ Linux <ul style="list-style-type: none"> - โครงสร้างของระบบปฏิบัติการ Linux - Basic Commands | บรรยาย + ปฏิบัติ |
| 4-6 | Linux File Management <ul style="list-style-type: none"> - คำสั่งเกี่ยวกับการจัดการ File และ Directory Executing Linux Commands <ul style="list-style-type: none"> - คำสั่งตรวจสอบ Resource - คำสั่งเกี่ยวกับ Network - คำสั่งเกี่ยวกับ User Permission | บรรยาย + ปฏิบัติ |

หมายเหตุ : การบรรยายและปฏิบัติในวิชานี้ให้ใช้ระบบปฏิบัติการ Ubuntu เป็นกรณีศึกษา แต่สามารถปรับเปลี่ยนเป็นระบบปฏิบัติการ Linux อื่นๆ ได้ตามที่วิทยากร และ/หรือคณะผู้จัดการอบรม เห็นสมควร เพื่อให้เหมาะสมกับความต้องการของกระทรวงสาธารณสุข

DE103 การสำรวจข้อมูล

6 ชั่วโมง

(Data Exploration)

หลักการของการสำรวจข้อมูล ทำความเข้าใจถึงโครงสร้างของข้อมูล โดยใช้วิธีการสำรวจแบบต่างๆ ที่เหมาะสมกับชนิดของข้อมูล เช่น ตารางแจกแจงความถี่ แผนภูมิแจกแจงความถี่ แผนภูมิก่อร่าง แผนภูมิแท่ง แผนภูมิจุด และค่าสถิติ เช่น ค่าเฉลี่ยเลขคณิต มัชยฐาน ความแปรปรวน ส่วนเบี่ยงเบนมาตรฐาน เพื่อใช้ในการตรวจสอบข้อผิดพลาด การตรวจสอบสมมติฐาน การคัดเลือกตัวแปร และการวิเคราะห์ความสัมพันธ์ โดยเน้นไปที่การเขียนโปรแกรมเพื่อทำการสำรวจข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการสำรวจข้อมูล
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการสำรวจข้อมูล
3. เพื่อให้ผู้เรียนสามารถเขียนโปรแกรมเพื่อทำการสำรวจข้อมูล

ตาราง 2.25 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE103 การสำรวจข้อมูล (Data Exploration)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|---------------------------------------|
| 1-3 | <p>แนะนำการสำรวจข้อมูล</p> <ul style="list-style-type: none"> - ขั้นตอนทั่วไปของการสำรวจข้อมูล - โลจิกที่ใช้ในการสำรวจข้อมูล (Pandas, Matplotlib, Seaborn) <p>การสำรวจข้อมูลตัวแปรเดียว</p> <ul style="list-style-type: none"> - แบบ Non-graphical (Frequency table, Range, Mean, Trimmed mean, Median, Variance, Standard Deviation, Interquartile Range) - แบบ Graphical (Countplot, Histogram) | <p>บรรยาย</p> <p>+</p> <p>ปฏิบัติ</p> |
| 4-6 | <p>การสำรวจข้อมูลหลายตัวแปร</p> <ul style="list-style-type: none"> - แบบ Non-graphical (Frequency table, Correlation coefficient) - แบบ Graphical (Bar plot, Box plot, Line plot, Scatter plot) | <p>บรรยาย</p> <p>+</p> <p>ปฏิบัติ</p> |

DE104 การเตรียมข้อมูล

6 ชั่วโมง

(Data Preparation)

การเตรียมข้อมูลเพื่อให้พร้อมสำหรับการวิเคราะห์ ตารางหน่วยการวิเคราะห์ การทำความสะอาดข้อมูล การจัดการข้อมูลค่าว่าง การจัดการข้อมูลกลุ่มย่อย การแปลงข้อมูลในแนวกว้าง-ยาว การสกัด

คุณลักษณะ การสรุปข้อมูล การแปลงชนิดของข้อมูล การเชื่อมตาราง โดยเน้นไปที่การเขียนโปรแกรม เพื่อทำการเตรียมข้อมูล

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการเตรียมข้อมูลสำหรับการวิเคราะห์
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการเตรียมข้อมูล

ตาราง 2.26 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE104 การเตรียมข้อมูล (Data Preparation)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|---|------------------------|
| 1-3 | ขั้นตอนทั่วไปของการเตรียมข้อมูล การทำความสะอาดข้อมูล - สาเหตุของความไม่สะอาดของข้อมูล - การทำความสะอาดข้อมูลแบบต่างๆ (ข้อมูลค่าว่าง ข้อมูลผิดพลาด ข้อมูลกลุ่มย่อย และข้อมูล Outlier) | บรรยาย + ปฏิบัติ |
| 4-6 | การเตรียมข้อมูล - การเตรียมข้อมูลแบบต่างๆ (การสกัดคุณลักษณะ การสรุปข้อมูล การแปลงชนิดของข้อมูล การเชื่อมตาราง การสร้างข้อมูลโปรไฟล์) | บรรยาย + ปฏิบัติ |

DE201 ระบบการจัดการฐานข้อมูลเชิงสัมพันธ์ขั้นสูง (Advance Relational Database Management System) **12 ชั่วโมง**

การจัดการฐานข้อมูลเชิงสัมพันธ์ การติดตั้งฐานข้อมูล การสร้างฐานข้อมูล การสร้างตาราง เรียกดู แก้ไข และลบข้อมูล การออกแบบฐานข้อมูลเชิงสัมพันธ์

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการระบบจัดการฐานข้อมูล
2. เพื่อให้ผู้เรียนมีความรู้ในการใช้คำสั่งจัดการฐานข้อมูลที่มีความซับซ้อน

ตาราง 2.27 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE201 ระบบการจัดการฐานข้อมูลเชิงสัมพันธ์ขั้นสูง (Advance Relational Database Management System)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|---|------------------------|
| 1-3 | การติดตั้ง Relational Database Management System การจัดการสิทธิใน Database | บรรยาย + ปฏิบัติ |
| 4-6 | Advance SQL Statement - JOIN และ UNION - Aggregate Functions | บรรยาย + ปฏิบัติ |

| | | |
|-------|---|------------------------|
| 7-9 | Advance SQL Statement (ต่อ) - SUBQUERY - Writing Efficient Queries | บรรยาย + ปฏิบัติ |
| 10-12 | Database Design Process - Database Structure - Relationship between Entities - Database Normalization - Multidimensional Data | บรรยาย + ปฏิบัติ |

DE202 กระบวนการ ETL

18 ชั่วโมง

(ETL Process)

ทำความเข้าใจกับกระบวนการ ETL (Extract Transform Load) ทุกขั้นตอน การวิเคราะห์ความต้องการด้านข้อมูล การเก็บรวบรวมข้อมูลจากแหล่งต่างๆ การจัดการข้อมูลต่างรูปแบบ การทำความสะอาดข้อมูล การแปลงข้อมูล การจัดเก็บข้อมูล โดยเน้นไปที่ การเขียนโปรแกรม เพื่อสร้างกระบวนการ ETL ด้วยตนเอง

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการทำกระบวนการ ETL
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการทำกระบวนการ ETL
3. เพื่อให้ผู้เรียนสามารถเขียนโปรแกรมเพื่อทำกระบวนการ ETL

ตาราง 2.28 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE202 กระบวนการ ETL

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | ETL Process & Tools - What is ETL - How ETL Is Being Used - ETL Tools | บรรยาย + ปฏิบัติ |
| 4-6 | Workshop: ETL pipeline 1 (Basic) ETL: Extract - Extract ข้อมูลจาก Flat files - Extract ข้อมูลจาก Relational Database - Extract ข้อมูลจาก NoSQL - Extract จาก API | บรรยาย + ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 7-9 | ETL: Extract (ต่อ) <ul style="list-style-type: none"> - Extract ข้อมูลจาก JSON - Extract ข้อมูลจาก HTML (เน้น) - Extract ข้อมูลที่ TEXT - Extract ข้อมูลจาก MS Access | บรรยาย + ปฏิบัติ |
| 10-12 | ETL: Transform <ul style="list-style-type: none"> - Cleaning - Deduplication - Format revision - Key restructuring. | บรรยาย + ปฏิบัติ |
| 13-15 | ETL: Load <ul style="list-style-type: none"> - Load ข้อมูลเข้าเป็น Relational Database - Load ข้อมูลเข้าเป็น NoSQL - Load ข้อมูลเข้าเป็น Flat Files | บรรยาย + ปฏิบัติ |
| 16-18 | Workshop: ETL pipeline 2 (Advance) | บรรยาย + ปฏิบัติ |

DE203 การจัดการกระแสนงาน

6 ชั่วโมง

(Workflow Management)

ความสำคัญของกระแสนงาน กระแสนงานต่างๆ ในระบบข้อมูล การสร้างและควบคุมกระแสนงาน การตรวจสอบสถานะ การแก้ไขข้อผิดพลาด เน้นไปที่การเขียนโปรแกรมและการใช้เครื่องมือเพื่อสร้างและควบคุมการทำงานของกระแสนงาน โดยเรียนรู้จากตัวอย่างจริง

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการจัดการกระแสนงาน
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการจัดการกระแสนงาน

ตาราง 2.29 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE203 การจัดการกระแสนงาน (Workflow Management)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | Apache Airflow - Apache Airflow Concept - Airflow UI | บรรยาย + ปฏิบัติ |
| 4-6 | DAGs - File structure / Composing DAG Failure and Troubleshooting Workshop: Composing workflow | บรรยาย + ปฏิบัติ |

DE204 การจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล

6 ชั่วโมง

(Managing development environment for data-related applications)

ขั้นตอนของการพัฒนาโปรแกรม สภาพแวดล้อมในการพัฒนาโปรแกรม ชนิดของการพัฒนาโปรแกรม แบบต่างๆ การพัฒนาแบบ Local, Shared Remote และ Shared Virtualized การจัดการ Library ในภาษาไพธอน Virtual Environment การติดตั้งโปรแกรมสำหรับสร้าง Shared Environment โดยเน้นไปที่การทดลอง ทำด้วยตนเอง

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล

ตาราง 2.30 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE204 การจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล (Managing development environment for data-related applications)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | Development Environment Python environment and packages | บรรยาย + ปฏิบัติ |
| 4-6 | Workshop: Setting up JupyterHub | บรรยาย + ปฏิบัติ |

DE301 ไมโครเซอร์วิส

6 ชั่วโมง

(Microservices)

หลักการของไมโครเซอร์วิส สถาปัตยกรรมไมโครเซอร์วิส การออกแบบไมโครเซอร์วิสแบบต่างๆ การใช้งาน Docker สำหรับการสร้างไมโครเซอร์วิส การเขียน Dockerfile

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการสร้างไมโครเซอร์วิส
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการสร้างไมโครเซอร์วิส

ตาราง 2.31 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE301 ไมโครเซอร์วิส (Microservices)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|---|------------------------|
| 1-3 | แนวคิด Microservice และ API เทคโนโลยีและโครงสร้าง Microservice และ API | บรรยาย + ปฏิบัติ |
| 4-6 | ออกแบบ สร้าง และตรวจสอบ Microservice และ API | บรรยาย + ปฏิบัติ |

DE302 การแลกเปลี่ยนข้อมูลและ API Gateway

6 ชั่วโมง

(Data Exchange and API Gateway)

ทำความรู้จักกับ API แบบต่างๆ การออกแบบ API การสร้างและทดสอบ API โดยการเขียนโปรแกรม ภาษาไพธอน การใช้งานโปรแกรมสำหรับทดสอบ API การแลกเปลี่ยนข้อมูลแบบต่างๆ ทำความรู้จักกับ API Gateway รวมถึงการติดตั้งและใช้งานจริง

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการแลกเปลี่ยนข้อมูลและ API Gateway
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการแลกเปลี่ยนข้อมูลและ API Gateway

ตาราง 2.32 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE302 การแลกเปลี่ยนข้อมูลและ API Gateway (Data Exchange and API Gateway)

| ชั่วโมงที่ | เนื้อหา | รูปแบบ |
|------------|---|------------------------|
| 1-3 | แนวคิด Data Exchange Service แนวคิด API Gateway | บรรยาย + ปฏิบัติ |
| 4-6 | ติดตั้งและจัดการ Data Exchange Server และ API Gateway | บรรยาย + ปฏิบัติ |

หมายเหตุ : การบรรยายและปฏิบัติในวิชานี้ให้ใช้ซอฟต์แวร์ชื่อ Docker และ KONG เป็นกรณีศึกษา แต่สามารถปรับเปลี่ยนได้ตามที่วิทยาการ และ/หรือคณะผู้จัดการอบรม เห็นสมควรเพื่อให้เหมาะสมกับความต้องการของกระทรวงสาธารณสุข

DE303 ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่

6 ชั่วโมง

(Introduction to Big Data)

พื้นฐานของการวิเคราะห์ข้อมูลขนาดใหญ่ เอกลักษณะของข้อมูลขนาดใหญ่ ระบบจัดการข้อมูลขนาดใหญ่ Hadoop Ecosystem หลักการของการประมวลผลแบบกระจาย เครื่องมือต่างๆ ในระบบจัดการข้อมูลขนาดใหญ่ โดยเน้นให้เห็นถึงภาพรวมของการจัดการและประมวลผลข้อมูลขนาดใหญ่

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนได้รับประสบการณ์การใช้งานเครื่องมือต่างๆ ในระบบจัดการข้อมูลขนาดใหญ่

ตาราง 2.33 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE303 ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ (Introduction to Big Data)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | Big Data Concept Introduction to Hadoop | บรรยาย + ปฏิบัติ |
| 4-6 | Workshop: Hadoop Core (HDFS, MapReduce) - การใช้งานคำสั่ง HDFS เบื้องต้น - การเขียนโปรแกรม MapReduce เบื้องต้น | บรรยาย + ปฏิบัติ |

DE304 การนำเข้าข้อมูลในระบบข้อมูลขนาดใหญ่

6 ชั่วโมง

(Data Ingestion in Big Data)

พื้นฐานของการนำเข้าข้อมูล หลักการในการนำเข้าข้อมูลแบบต่างๆ การใช้งานเครื่องมือในระบบข้อมูลขนาดใหญ่เพื่อนำเข้าข้อมูล เช่น Sqoop และ Flume การออกแบบการนำเข้าข้อมูล โดยเน้นไปที่การทดลองทำการนำเข้าข้อมูลจริง

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจเกี่ยวกับการนำเข้าข้อมูลในระบบข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนสามารถดึงข้อมูลจากแหล่งต่างๆ โดยเลือกใช้วิธีที่เหมาะสมกับแหล่งข้อมูล

ตาราง 2.34 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE304 การนำเข้าข้อมูลในระบบข้อมูลขนาดใหญ่ (Data Ingestion in Big Data)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | การใช้งาน Flume <ul style="list-style-type: none"> - Architecture - การใช้งาน Flume การใช้งาน Sqoop <ul style="list-style-type: none"> - Import - Export | บรรยาย + ปฏิบัติ |
| 4-6 | การใช้งาน Web HDFS <ul style="list-style-type: none"> - File and Directory Operations | บรรยาย + ปฏิบัติ |

DE305 การเตรียมข้อมูลในระบบข้อมูลขนาดใหญ่

6 ชั่วโมง

(Data Preparation in Big Data)

การเตรียมข้อมูลเพื่อให้พร้อมสำหรับการวิเคราะห์ หลักการของการประมวลผลข้อมูลบนระบบข้อมูลขนาดใหญ่ โดยจะเน้นไปที่การใช้งานเครื่องมือเฉพาะของระบบข้อมูลขนาดใหญ่เพื่อการเตรียมข้อมูล ได้แก่ Apache Spark รวมถึงหลักการและการตั้งค่าต่างๆ

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการเตรียมข้อมูลในระบบข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการเตรียมข้อมูล

ตาราง 2.35 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE305 การเตรียมข้อมูลในระบบข้อมูลขนาดใหญ่ (Data Preparation in Big Data)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | Apache Spark <ul style="list-style-type: none"> - RDD - Transformations - Actions | บรรยาย + ปฏิบัติ |
| 4-6 | Apache Spark (ต่อ) <ul style="list-style-type: none"> - Programming with RDD - Deployment | บรรยาย + ปฏิบัติ |

DE306 การให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่

6 ชั่วโมง

(Data Service in Big Data)

เครื่องมือสำหรับให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่ หลักการของฐานข้อมูลบนระบบข้อมูลขนาดใหญ่ Schema-on-Read รูปแบบไฟล์ข้อมูลบนระบบข้อมูลขนาดใหญ่ ความแตกต่างกับระบบฐานข้อมูลธรรมดา โดยเรียนรู้ผ่านเครื่องมือ Hive, Beeline และ Hue

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการและวิธีการให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนสามารถใช้เครื่องมือและเทคนิคต่างๆ ในการให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่

ตาราง 2.36 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE306 การให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่ (Data Service in Big Data)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|------------------------|
| 1-3 | Hive - การจัดการฐานข้อมูล - การจัดการตารางข้อมูล | บรรยาย + ปฏิบัติ |
| 4-6 | Hive (ต่อ) - Built-In Operators - Built-In Functions - Views And Indexes - HiveQL | บรรยาย + ปฏิบัติ |

DE307 การออกแบบระบบข้อมูลขนาดใหญ่

6 ชั่วโมง

(Big Data Architectural Design)

ระบบ Hadoop Ecosystem ส่วนประกอบต่างๆ ของระบบ แนะนำโปรแกรมต่างๆ ในระบบ หลักการในการออกแบบ Lambda Architecture, Kappa Architecture กรณีศึกษาจากสถานการณ์จริง ทำการออกแบบระบบข้อมูลขนาดใหญ่จากสถานการณ์ที่กำหนดให้

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการการออกแบบระบบข้อมูลขนาดใหญ่
2. เพื่อให้ผู้เรียนสามารถออกแบบระบบฐานข้อมูลขนาดใหญ่

ตาราง 2.37 เนื้อหาราย 3 ชั่วโมง ของรายวิชา DE307 การออกแบบระบบข้อมูลขนาดใหญ่ (Big Data Architectural Design)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|------------------------|
| 1-3 | Big Data Components - ส่วนประกอบของ Big Data Architecture | บรรยาย + ปฏิบัติ |
| 4-6 | Big Data Architecture - รูปแบบ architecture - Example scenario - Real-world use case - Workshop – Big Data Architecture Design | บรรยาย + ปฏิบัติ |

4. หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน

PR100 การนำเสนอผลงานอย่างมีประสิทธิภาพ

6 ชั่วโมง

(Pitching and Effective Presentation)

การนำเสนอในรูปแบบ Pitching เป็นกระบวนการนำเสนอขนาดใหญ่ ที่มีแนวทางการนำเสนอที่แตกต่างจากการนำเสนอทั่วไป โดยเน้นให้ผู้ฟังเกิดความสนใจและมีส่วนร่วมกับเรื่องที่น่าสนใจ โดยเฉพาะอย่างยิ่งเป็นการเล่าเรื่อง (Storytelling) ได้อย่างครบถ้วนและกระชับในระยะเวลาสั้นๆ การนำเสนอในรูปแบบ Pitching มักใช้ในการนำเสนอโครงการหรือธุรกิจต่อกลุ่มเป้าหมายต่างๆ เช่น นักลงทุน ลูกค้า หรือพันธมิตรทางธุรกิจ เพื่อโน้มน้าวใจให้กลุ่มเป้าหมายเหล่านั้นสนใจและตัดสินใจสนับสนุนโครงการหรือธุรกิจนั้นๆ

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจโมเดลการนำเสนออย่างได้ผลและสามารถนำไปประยุกต์ใช้ได้จริง
2. เพื่อส่งเสริมและพัฒนาผู้เรียนให้เกิดแรงบันดาลใจต่อการนำเสนออย่างสร้างสรรค์
3. เพื่อให้ผู้เรียนสามารถใช้เทคนิคการนำเสนอเป็นเครื่องมือในการสื่อสารด้วยวัตถุประสงค์อื่นได้

ตาราง 2.38 รายละเอียดเนื้อหาของรายวิชา PR100 การนำเสนอผลงานอย่างมีประสิทธิภาพ (Pitching and Effective Presentation)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|--------------------|
| 1-3 | การนำเสนอแบบ Pitching - ทำความรู้จักกับการนำเสนอแบบ Pitching และแนวคิดพื้นฐานการ Pitching | บรรยาย/ ปฏิบัติ |

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|---|--------------------|
| | <ul style="list-style-type: none"> - เทคนิคการ Pitching ด้วย Pitch Deck, ส่วนประกอบและรายละเอียดของการนำเสนอ - กิจกรรมให้ผู้เข้าอบรมทดลองสร้าง Pitch Deck | |
| 4-6 | <p>การเพิ่มประสิทธิภาพการนำเสนอ</p> <ul style="list-style-type: none"> - รูปแบบ (Style) การนำเสนอรูปแบบต่างๆ - หลักคิดในการออกแบบการนำเสนอ/Slide นำเสนอ - เทคนิคและแนวทางในการเพิ่มประสิทธิภาพการนำเสนอ - กิจกรรมให้ผู้เข้าอบรมทดลองนำเสนอผลงานที่ได้เตรียมมา | บรรยาย/ ปฏิบัติ |

PR101 ธรรมาภิบาลข้อมูล

6 ชั่วโมง

(Data Governance)

หลักการของธรรมาภิบาลข้อมูล ปัญหาและข้อจำกัดของการใช้ข้อมูล เมตาดาต้าและการจัดการเมตาดาต้า การวิเคราะห์ผลกระทบและการไหลของข้อมูล ความมั่นคงปลอดภัยของข้อมูล บุคลากรข้อมูลที่เกี่ยวข้อง มาตรฐานที่เกี่ยวข้องกับการออกแบบ สถาปัตยกรรม และการบริหารจัดการคุณภาพข้อมูล เกณฑ์การประเมินคุณภาพของข้อมูล แนวทางการนำธรรมาภิบาลข้อมูลไปใช้ในองค์กร

วัตถุประสงค์

1. เพื่อให้ผู้เรียนเข้าใจหลักการสำคัญของธรรมาภิบาลข้อมูล
2. เพื่อให้ผู้เรียนสามารถนำมาตรฐานที่เกี่ยวข้องมาใช้กับงานจัดการข้อมูลในองค์กร
3. เพื่อให้ผู้เรียนมีแนวทางการนำธรรมาภิบาลข้อมูลไปใช้ในองค์กร

ตาราง 2.39 รายละเอียดเนื้อหาของรายวิชา PR101 ธรรมาภิบาลข้อมูล (Data Governance)

| ชั่วโมง ที่ | เนื้อหา | รูปแบบ |
|----------------|--|---------------------|
| 1-3 | <p>หลักการของธรรมาภิบาลข้อมูล</p> <ul style="list-style-type: none"> - การจัดการเมตาดาต้า - ความมั่นคงปลอดภัยของข้อมูล - ประเด็นด้านธรรมาภิบาลข้อมูลต่างๆ <p>กฎหมายและมาตรฐานที่เกี่ยวข้องกับธรรมาภิบาลข้อมูล</p> <ul style="list-style-type: none"> - พระราชบัญญัติคุ้มครองข้อมูลส่วนบุคคล พ.ศ. 2562 - มาตรฐาน ISO | บรรยาย |
| 4-6 | <p>แนวทางการนำธรรมาภิบาลข้อมูลไปใช้ในองค์กร</p> <p>Workshop ประเมินระดับของธรรมาภิบาลข้อมูลในองค์กร</p> | บรรยาย + ปฏิบัติ |

หมวดที่ 3 การประกันคุณภาพหลักสูตร

1. การปรับปรุงหลักสูตร

หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูล มุ่งเน้นให้ผู้เข้ารับการอบรมได้รับความรู้และทักษะที่จำเป็นในการพัฒนาตนเองและปฏิบัติงานได้อย่างมีประสิทธิภาพ เพื่อให้หลักสูตรมีคุณภาพ ควรมีการจัดการประชุมปรับปรุงหลักสูตรทุก ๆ 2 ปี เพื่อให้หลักสูตรมีเนื้อหาที่ทันสมัย สอดคล้องกับความต้องการขององค์กร บุคลากร และตลาดแรงงานที่มีการเปลี่ยนแปลง โดยตัวอย่างข้อบ่งชี้ที่ควรปรับปรุงจัดการปรับปรุงหลักสูตร เช่น

- วิธีการหรือเครื่องมือใหม่ๆ ในการวิเคราะห์ข้อมูล
- แนวโน้มของเทคโนโลยีที่เปลี่ยนไป
- ข้อกำหนดหรือข้อจำกัดที่ส่งผลกระทบต่องานการวิเคราะห์ข้อมูล เช่น กฎหมาย นโยบาย หรือความร่วมมือต่างๆ

โดยมีแนวทางในการปรับปรุง คือ

1. ปรับปรุงเนื้อหาหลักสูตรให้ทันสมัยและสอดคล้องกับสถานการณ์ปัจจุบัน หลักสูตรควรครอบคลุมเนื้อหาเกี่ยวกับการจัดการและการวิเคราะห์ข้อมูลในรูปแบบใหม่ๆ เช่น ปัญญาประดิษฐ์ (AI) และการเรียนรู้ของเครื่อง (Machine Learning) ซึ่งสามารถนำไปประยุกต์ใช้ในการทำงานด้านสาธารณสุขได้อย่างมีประสิทธิภาพ
2. ปรับปรุงรูปแบบการสอนให้น่าสนใจและมีส่วนร่วมของผู้เข้าอบรม หลักสูตรควรใช้รูปแบบการสอนที่หลากหลาย เช่น การสอนแบบบรรยาย การสอนแบบสาธิต การสอนแบบฝึกปฏิบัติ และการสอนแบบกลุ่ม เพื่อให้ผู้เข้าอบรมมีส่วนร่วมในการเรียนรู้และเข้าใจเนื้อหาได้อย่างลึกซึ้ง
3. ปรับปรุงการประเมินผลสัมฤทธิ์ของหลักสูตรให้ครอบคลุมและมีประสิทธิภาพ หลักสูตรควรมีการประเมินผลสัมฤทธิ์ที่ครอบคลุมทั้งความรู้ ทักษะ และทัศนคติของผู้เข้าอบรม เพื่อวัดผลลัพธ์การเรียนรู้อย่างแท้จริง

2. แผนพัฒนาปรับปรุงหลักสูตร

| แผนการพัฒนา/เปลี่ยนแปลง | กลยุทธ์ | หลักฐาน/ตัวบ่งชี้ |
|--|--|--|
| ปรับปรุงหลักสูตรให้สอดคล้องกับความต้องการของธุรกิจ และการเปลี่ยนแปลงของเทคโนโลยี | ปรับปรุงย่อยตามข้อเสนอแนะในรายงานผลการดำเนินการของหลักสูตร และปรับปรุงหลักสูตรทุก ๆ 2 ปี | - รายงานผลการประเมินการนำความรู้ไปใช้งานในการทำงาน - แผนการดำเนินงานหรือนโยบายขององค์กร |

| | | |
|---|--|--|
| ปรับปรุงหลักสูตรให้สอดคล้องกับ ผู้เข้าอบรม | ปรับปรุงย่อยตามข้อเสนอแนะใน รายงานผลการดำเนินการของ หลักสูตร | - รายงานผลความพึงพอใจของ ผู้เข้ารับการอบรม - รายงานผลประเมิน ความสามารถของผู้เข้ารับการ อบรม |
|---|--|--|

3. การประเมินผู้เข้ารับการอบรม

การประเมินผู้เข้ารับการอบรมเป็นกระบวนการสำคัญในการวัดผลสัมฤทธิ์ของหลักสูตรการอบรม เพื่อตอบวัตถุประสงค์ที่ตั้งไว้ โดยในการประเมินผู้เข้ารับการอบรม จะใช้วิธีการทดสอบก่อนและหลังเรียน ในแต่ละระยะของการอบรม ได้แก่ หลักสูตรปรับพื้นฐาน หลักสูตรขั้นพื้นฐาน หลักสูตรชั้นกลาง และหลักสูตรขั้นสูง รวมทั้งหมด 8 ครั้ง โดยแบบทดสอบในแต่ละหลักสูตรจะเป็นการทดสอบความรู้ที่ได้เรียนมาในหลักสูตรนั้นๆ โดยผู้เข้ารับการอบรมจะต้องได้รับผลการประเมินไม่น้อยกว่าร้อยละ 80 ถึงจะผ่านเกณฑ์และสามารถเข้ารับการอบรม ในหลักสูตรถัดไปได้ นอกจากนี้ หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน จะประเมินผู้เข้าอบรม ด้วยการทำชิ้นงาน (Project) แบบทีม (2-3 คน) หรือแบบเดี่ยว และนำเสนอผลงาน โดยผ่านการประเมิน จากผู้วิพากษ์ พร้อมจัดทำรายงานฉบับสมบูรณ์ จึงได้รับใบประกาศนียบัตร ผ่านการอบรมหลักสูตรการจัดการ และการวิเคราะห์ข้อมูล สาขาวิชาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล ทั้งนี้ ผู้อบรมที่ผ่านตามเกณฑ์ที่ หลักสูตรกำหนดจะได้รับใบประกาศนียบัตรเทียบสาขาวิทยาศาสตร์ข้อมูล (Data Science) ระดับ 5 และสาขา วิศวกรรมข้อมูล (Data Engineer) ระดับ 5 จากสถาบันคุณวุฒิวิชาชีพเช่นเดียวกัน ซึ่งการทดสอบจะมีรูปแบบ และเนื้อหาตาม

ตาราง 3.1 ดังนี้

ตาราง 3.1 รูปแบบและเนื้อหาของการทดสอบ

| ที่ | การทดสอบ | รูปแบบ | เนื้อหา | หมายเหตุ |
|-----|---|--|--|----------------------|
| 1 | ทดสอบก่อนและ หลังเรียน หลักสูตร ปรับพื้นฐานการ จัดการและการ วิเคราะห์ข้อมูล | ปรนัย 20 ข้อ การเขียน โปรแกรม 4 ข้อ | - ความรู้เบื้องต้นเกี่ยวกับการใช้ข้อมูลประกอบการ ตัดสินใจ 5 ข้อ - การเขียนโปรแกรมภาษาไพธอนสำหรับการวิเคราะห์ ข้อมูลเบื้องต้น 10 ข้อ - ฐานข้อมูลและระบบจัดการฐานข้อมูล 5 ข้อ - ข้อสอบเขียนโปรแกรม 4 ข้อ แบ่งเป็นความสามารถ ด้านวิทยาศาสตร์ข้อมูล และวิศวกรรมข้อมูล อย่างละ 2 ข้อ | สอบทุกคน |
| 2 | ทดสอบก่อนและ หลังเรียน หลักสูตร | ปรนัย 30 ข้อ | - ความรู้เบื้องต้นเกี่ยวกับวิทยาศาสตร์ข้อมูล 5 ข้อ - การโปรแกรมภาษาไพธอนสำหรับวิทยาศาสตร์ข้อมูล 10 ข้อ | สอบเฉพาะ สาขาวิชา |

| ที่ | การทดสอบ | รูปแบบ | เนื้อหา | หมายเหตุ |
|-----|---|---------------------------------------|---|-------------------------------------|
| | การจัดการและการวิเคราะห์ข้อมูล สาขาวิชา วิทยาศาสตร์ข้อมูล ชั้นพื้นฐาน | การเขียนโปรแกรม 1 ข้อ | <ul style="list-style-type: none"> - การสำรวจข้อมูล 5 ข้อ - การเตรียมข้อมูล 5 ข้อ - การนำเสนอข้อมูล 5 ข้อ - การเขียนโปรแกรมเพื่อสร้างกระบวนการประมวลผลข้อมูลเบื้องต้น 1 ข้อ | วิทยาศาสตร์ข้อมูล |
| 3 | ทดสอบก่อนและหลังเรียน หลักสูตร การจัดการและการวิเคราะห์ข้อมูล สาขาวิชา วิทยาศาสตร์ข้อมูล ชั้นกลาง | ปรนัย 35 ข้อ | <ul style="list-style-type: none"> - สถิติสำหรับวิทยาศาสตร์ข้อมูล 15 ข้อ - การนำเข้าข้อมูล 5 ข้อ - การเรียนรู้ของเครื่อง 15 ข้อ | สอบเฉพาะ สาขาวิชา วิทยาศาสตร์ข้อมูล |
| 4 | ทดสอบก่อนและหลังเรียน หลักสูตร การจัดการและการวิเคราะห์ข้อมูล สาขาวิชา วิทยาศาสตร์ข้อมูล ชั้นสูง | ปรนัย 30 ข้อ | <ul style="list-style-type: none"> - ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ 3 ข้อ - การบูรณาการกับระบบโครงสร้างพื้นฐาน 2 ข้อ - การประมวลผลข้อมูลสารสนเทศภูมิศาสตร์ 5 ข้อ - การวิเคราะห์ข้อมูลอนุกรมเวลา 10 ข้อ - การประมวลผลข้อความ 5 ข้อ - การเรียนรู้เชิงลึก 5 ข้อ | สอบเฉพาะ สาขาวิชา วิทยาศาสตร์ข้อมูล |
| 5 | ทดสอบก่อนและหลังเรียน หลักสูตร การจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิศวกรรมข้อมูลชั้นพื้นฐาน | ปรนัย 30 ข้อ การเขียนโปรแกรม 1 ข้อ | <ul style="list-style-type: none"> - ความรู้เบื้องต้นเกี่ยวกับวิศวกรรมข้อมูล 5 ข้อ - การโปรแกรมภาษาไพธอนสำหรับวิศวกรรมข้อมูล 10 ข้อ - ระบบปฏิบัติการ Linux เบื้องต้น 5 คะแนน - การสำรวจข้อมูล 5 ข้อ - การเตรียมข้อมูล 5 ข้อ - การเขียนโปรแกรมสร้างกระบวนการประมวลผลข้อมูล 1 ข้อ | สอบเฉพาะ สาขาวิชา วิศวกรรมข้อมูล |
| 6 | ทดสอบก่อนและหลังเรียน หลักสูตร การจัดการและการวิเคราะห์ข้อมูล สาขาวิชาวิศวกรรมข้อมูลชั้นกลาง | ปรนัย 35 ข้อ | <ul style="list-style-type: none"> - ระบบการจัดการฐานข้อมูลเชิงสัมพันธ์ขั้นสูง 10 ข้อ - กระบวนการ ETL 15 ข้อ - การจัดการกระแสนาน 5 ข้อ - การจัดการระบบสำหรับการพัฒนาโปรแกรมด้านข้อมูล 5 ข้อ | สอบเฉพาะ สาขาวิชา วิศวกรรมข้อมูล |

| ที่ | การทดสอบ | รูปแบบ | เนื้อหา | หมายเหตุ |
|-----|---|--|--|---|
| 7 | ทดสอบก่อนและหลังเรียน หลักสูตรการจัดการและการวิเคราะห์ข้อมูลสาขาวิชาวิศวกรรมข้อมูลขั้นสูง | ปรนัย 35 ข้อ | <ul style="list-style-type: none"> - ไมโครเซอร์วิส 5 ข้อ - การแลกเปลี่ยนข้อมูลและ API Gateway 5 ข้อ - ความรู้เบื้องต้นเกี่ยวกับข้อมูลขนาดใหญ่ 5 ข้อ - การนำเข้าข้อมูลในระบบข้อมูลขนาดใหญ่ 5 ข้อ - การเตรียมข้อมูลในระบบข้อมูลขนาดใหญ่ 5 ข้อ - การให้บริการข้อมูลในระบบข้อมูลขนาดใหญ่ 5 ข้อ - การออกแบบระบบข้อมูลขนาดใหญ่ 5 ข้อ | สอบเฉพาะสาขาวิชาวิศวกรรมข้อมูล |
| 8 | การนำเสนอผลงานผ่านผู้วิพากษ์ หลักสูตรเตรียมความพร้อมสำหรับการนำเสนอผลงาน | การนำเสนอผลงาน 1 ชิ้น | <ul style="list-style-type: none"> - ความเหมาะสมของการจัดเรียงเนื้อหานำเสนอ - ความชัดเจนของวัตถุประสงค์ - ความน่าเชื่อถือและความถูกต้องของผลงาน การเลือกใช้วิธีการหรือโมเดลที่เหมาะสม - ความน่าสนใจ/ความคิดสร้างสรรค์ของผลงาน - การถ่ายทอดเนื้อหา - การนำผลงานไปใช้ประโยชน์ - เวลาที่ใช้ในการนำเสนอ - ความสนใจผลงานของผู้วิพากษ์ - ความมีวินัยและตรงต่อเวลา | นำเสนอทั้งสาขาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล |
| 9 | ประเมินจบหลักสูตรการจัดการและการวิเคราะห์ข้อมูลสาขาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล | 1. การนำเสนอผลงาน 1 ชิ้น 2. รายงานฉบับสมบูรณ์ผลงาน (Project) รูปแบบ E-Book 1 เล่ม | <ul style="list-style-type: none"> - การมีวินัยตรงต่อเวลาเข้าอบรมทุกหลักสูตร - นำเสนอผลงานกับผู้วิพากษ์ - คะแนนสอบหลักการอบรมผ่านตามเกณฑ์ที่กำหนดไม่น้อยกว่าร้อยละ 80 อย่างน้อย 2 หลักสูตรขึ้นไป - รายงานผลงาน (Project) รูปแบบ E-Book | ใบประกาศนียบัตรผ่าน หรือเข้าร่วม อบรมหลักสูตรการจัดการและการวิเคราะห์ข้อมูลสาขาวิทยาศาสตร์ข้อมูลและวิศวกรรมข้อมูล |

4. แผนการประเมินความพึงพอใจ

1) วัตถุประสงค์

- (1) เพื่อประเมินระดับความพึงพอใจของผู้เข้าอบรมที่มีต่อหลักสูตรการจัดการและการวิเคราะห์ข้อมูล
- (2) เพื่อนำข้อมูลผลประเมินไปปรับปรุงหลักสูตรให้มีประสิทธิภาพยิ่งขึ้น

2) กลุ่มเป้าหมาย

ผู้เข้าอบรมหลักสูตรการจัดการและการวิเคราะห์ข้อมูล

3) วิธีการ

ประเมินผลหลังการอบรมโดยใช้แบบสอบถามความพึงพอใจ

4) เครื่องมือ

แบบสอบถามความพึงพอใจ

5) เนื้อหาของแบบสอบถาม

- (1) ข้อมูลส่วนบุคคลของผู้ตอบแบบสอบถาม เช่น เพศ อายุ ตำแหน่งงาน
- (2) ระดับความพึงพอใจในด้านต่างๆ ดังนี้
 - ด้านเนื้อหา
 - ด้านกิจกรรม
 - ด้านสื่อการเรียนรู้
 - ด้านวิทยากร
 - ด้านตารางเวลาของการอบรม
 - ด้านสภาพแวดล้อมของการอบรม
- (3) ข้อคิดเห็นและเสนอแนะเพิ่มเติม

6) การนำผลไปใช้

ผลของการประเมินจะถูกนำไปใช้วิเคราะห์หาจุดแข็ง และข้อปรับปรุงแก้ไขของหลักสูตร และนำไปใช้ในการปรับปรุงหลักสูตรให้มีประสิทธิภาพยิ่งขึ้น รวมถึงนำไปประกอบในการพัฒนาสื่อการสอน กิจกรรมการอบรม และสถานที่อบรมในอนาคต

ภาคผนวก ก

คำสั่งแต่งตั้งคณะทำงานพัฒนาและจัดทำหลักสูตรการจัดการและการวิเคราะห์ข้อมูล

(GEEKS : Data Engineer & Data Science)



คำสั่งกรมควบคุมโรค
ที่ ๒๗๗ / ๒๕๖๗

เรื่อง แต่งตั้งคณะทำงานพัฒนาและจัดทำหลักสูตรการจัดการและการวิเคราะห์ข้อมูล
(GEEKS : Data Engineer & Data Science)

ตามที่กรมควบคุมโรค โดยกองดิจิทัลเพื่อการควบคุมโรค ได้พัฒนาหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูลและสาขาวิทยาศาสตร์ข้อมูล (GEEKS : Data Engineer & Data Science) และดำเนินการจัดอบรมหลักสูตรการจัดการและการวิเคราะห์ข้อมูลระดับพื้นฐาน ระดับกลาง และระดับสูง โดยมีจุดมุ่งหมายเพื่อพัฒนาศักยภาพและเสริมสร้างทักษะให้กับบุคลากรทางการแพทย์และสาธารณสุข สำหรับปฏิบัติงานทางด้านการจัดการและการวิเคราะห์ข้อมูล นั้น

เพื่อพัฒนาหลักสูตรให้ทันสมัย เป็นปัจจุบัน และมีความเหมาะสมสำหรับนำไปจัดอบรมให้แก่บุคลากรที่ทำงานด้านสาธารณสุข หรือสารสนเทศสาธารณสุข เช่น แพทย์ พยาบาล นักวิเคราะห์นโยบายและแผน นักวิชาการสาธารณสุข นักระบาดวิทยา นักสถิติ นักวิชาการคอมพิวเตอร์ หรือนักเทคโนโลยีสารสนเทศ เป็นไปด้วยความเรียบร้อยและมีประสิทธิภาพ อาศัยอำนาจตามความในมาตรา ๓๒ แห่งพระราชบัญญัติระเบียบบริหารราชการแผ่นดิน พ.ศ. ๒๕๓๔ แก้ไขเพิ่มเติมโดยพระราชบัญญัติระเบียบบริหารราชการแผ่นดิน (ฉบับที่ ๕) พ.ศ. ๒๕๔๕ กรมควบคุมโรค จึงแต่งตั้งคณะทำงานพัฒนาและจัดทำหลักสูตรการจัดการและการวิเคราะห์ข้อมูล (GEEKS : Data Engineer & Data Science) โดยมีองค์ประกอบ หน้าที่และอำนาจ ดังนี้

๑. องค์ประกอบ

- | | |
|--|-------------------|
| ๑.๑ นายพงศ์ธร ชาติพิทักษ์ | ประธานคณะทำงาน |
| นายแพทย์เชี่ยวชาญ ผู้อำนวยการกองดิจิทัลเพื่อการควบคุมโรค | |
| ๑.๒ นางสาวรัฐลักษณ์ พิทักษ์จักรพิภพ | รองประธานคณะทำงาน |
| นักวิเคราะห์นโยบายและแผนชำนาญการ รองผู้อำนวยการกองดิจิทัลเพื่อการควบคุมโรค | |
| ๑.๓ นายเอกชัย วิวรรณศิริรักษ์ | คณะทำงาน |
| ผู้อำนวยการศูนย์นวัตกรรมซอฟต์แวร์และการประมวลผล มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี | |
| ๑.๔ นายสิงหนาท เรืองโอสถ | คณะทำงาน |
| นักคอมพิวเตอร์ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี | |
| ๑.๕ นายวริทธิ์ เต็นพิพัฒน์ | คณะทำงาน |
| นักคอมพิวเตอร์ภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี | |
| ๑.๖ นายศุภวิชญ์ สมสกุล | คณะทำงาน |
| นักวิจัยภาควิชาวิศวกรรมคอมพิวเตอร์ มหาวิทยาลัยเทคโนโลยีพระจอมเกล้าธนบุรี | |

๑.๗ นาย...

- | | |
|--|---------------------------------|
| ๑.๗ นายศุภณัฐ วงศานุพัทธ์ นายแพทย์ชำนาญการ กองระบาดวิทยา | คณะทำงาน |
| ๑.๘ นางสาวคุณกัญญ์ศศิ พิมพพันธ์ นักวิชาการสาธารณสุขชำนาญการ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน |
| ๑.๙ นางสาวรัญชนา ไหล่ลั่น นักวิชาการสาธารณสุขปฏิบัติการ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน |
| ๑.๑๐ นางสาวกนกอร งามนิก นักวิชาการสาธารณสุขปฏิบัติการ สำนักงานคณะกรรมการควบคุมเครื่องดื่มแอลกอฮอล์ | คณะทำงาน |
| ๑.๑๑ นายอนันทวีป ใจเมา นักวิชาการคอมพิวเตอร์ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน |
| ๑.๑๒ นายธนดน พัฒนะโกไคย นักวิชาการคอมพิวเตอร์ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน |
| ๑.๑๓ นายอนุชิต สายสร นักวิชาการคอมพิวเตอร์ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน |
| ๑.๑๔ นางสาวสิริสุดา สุษะลา นักเทคโนโลยีสารสนเทศ กองโรคไม่ติดต่อ | คณะทำงาน |
| ๑.๑๕ นางสาวภรณ์ทิพย์ จันทะนาเขต นักวิชาการคอมพิวเตอร์ กองโรคจากการประกอบอาชีพและสิ่งแวดล้อม | คณะทำงาน |
| ๑.๑๖ นายธีรยุทธ ชูพุทธพงศ์ นักวิชาการคอมพิวเตอร์ กองโรคจากการประกอบอาชีพและสิ่งแวดล้อม | คณะทำงาน |
| ๑.๑๗ นางสาวกรรฤทัย กิตติคุณ นักวิชาการคอมพิวเตอร์ปฏิบัติการ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน และเลขานุการ |
| ๑.๑๘ นางสาวมยุรี ต้นตยาพงศ์ ผู้จัดการโครงการ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน และผู้ช่วยเลขานุการ |
| ๑.๑๙ นางสาวธนาภรณ์ แสงบรรเจิดกุล เจ้าหน้าที่ประสานงานโครงการ กองดิจิทัลเพื่อการควบคุมโรค | คณะทำงาน และผู้ช่วยเลขานุการ |

๒. หน้าที...

๒. หน้าที่และอำนาจ

๒.๑ วางแผนและจัดทำกรอบเนื้อหาหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูลและสาขาวิทยาศาสตร์ข้อมูล (GEEKS : Data Engineer & Data Science)

๒.๒ พิจารณาให้ข้อเสนอแนะและปรับปรุงรายละเอียดเนื้อหาหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูลและสาขาวิทยาศาสตร์ข้อมูล (GEEKS : Data Engineer & Data Science) ให้สอดคล้องกับมาตรฐานหลักสูตรของกรมควบคุมโรค


๒.๓ วางแผนการประเมินผลความพึงพอใจของหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูลและสาขาวิทยาศาสตร์ข้อมูล (GEEKS : Data Engineer & Data Science)

๒.๔ ให้การสนับสนุนและส่งเสริมการพัฒนาหลักสูตรและการจัดอบรมหลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิศวกรรมข้อมูลและสาขาวิทยาศาสตร์ข้อมูล (GEEKS : Data Engineer & Data Science) ให้เป็นไปโดยเรียบร้อยและมีประสิทธิภาพ

๒.๕ ปฏิบัติหน้าที่อื่นๆ ตามที่ได้รับมอบหมาย

ทั้งนี้ ตั้งแต่บัดนี้เป็นต้นไป

สั่ง ณ วันที่ ๑๒ กุมภาพันธ์ พ.ศ. ๒๕๖๗



(นายนิติ เทศานุกรักษ์)
รองอธิบดี ปฏิบัติราชการแทน
อธิบดีกรมควบคุมโรค

ภาคผนวก ข

แบบประเมินความพึงพอใจการอบรมเชิงปฏิบัติการ

หลักสูตร การจัดการและการวิเคราะห์ข้อมูลระดับ Basic Intermediate Advance

คำชี้แจง แบบประเมินนี้มีวัตถุประสงค์เพื่อสอบถามความพึงพอใจของผู้เรียนในการอบรมเชิงปฏิบัติการ แบ่งเป็น 6 ตอน ดังนี้

ส่วนที่ 1 ข้อมูลทั่วไป

ส่วนที่ 2 ด้านวิทยากร

ส่วนที่ 3 ความเข้าใจในเนื้อหา

ส่วนที่ 4 ความเหมาะสมของเนื้อหาและเวลา

ส่วนที่ 5 สื่อ อุปกรณ์การสอนและสถานที่จัดอบรม

ส่วนที่ 6 เจตคติต่อหลักสูตรฯ

ส่วนที่ 1 ข้อมูลทั่วไปผู้เข้าอบรม

1. ประเภทหน่วยงาน

- ส่วนกลาง
- ส่วนภูมิภาค

2. เพศ

- ชาย
- หญิง

3. ระดับการศึกษา

- ต่ำกว่าปริญญาตรี
- ปริญญาตรี
- ปริญญาโท
- ปริญญาเอก

4. ตำแหน่ง

- แพทย์
- นักวิชาการคอมพิวเตอร์

- นักวิชาการสาธารณสุข
- อื่น

5. ท่านทำงานในตำแหน่งปัจจุบัน เป็นระยะเวลา ปี

- น้อยกว่า 1 ปี
- 1-3 ปี
- 3-6 ปี
- มากกว่า 6ปี

| ลำดับ | หัวข้อ | ก่อนการอบรม | | | | | หลังการอบรม | | | | |
|--|---|-------------|---|---|---|---|-------------|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 1 | 2 | 3 | 4 | 5 |
| ส่วนที่ 2 ด้านวิทยากร | | | | | | | | | | | |
| 1. | การถ่ายทอดความรู้ของวิทยากร | | | | | | | | | | |
| 2. | ความรู้ ความสามารถ และประสบการณ์ของวิทยากร | | | | | | | | | | |
| 3. | การอธิบายเนื้อหาได้ชัดเจนและตรงประเด็น | | | | | | | | | | |
| 4. | การจัดลำดับเนื้อหาวิชาการเชื่อมโยงกันเหมาะสม | | | | | | | | | | |
| 5. | การใช้ภาษาที่เหมาะสมและเข้าใจง่าย | | | | | | | | | | |
| ส่วนที่ 3 ความเข้าใจในเนื้อหาการสอน | | | | | | | | | | | |
| 1. | หลักการพื้นฐานของการประมวลผลของ Apache Spark | | | | | | | | | | |
| 2. | Apache Spark workshop (ชุดข้อมูล 43 files) | | | | | | | | | | |
| 3. | การสร้างโมเดลในการวิเคราะห์ข้อมูลด้วย Apache Spark | | | | | | | | | | |
| 4. | Model deployment workshop | | | | | | | | | | |
| 5. | Introduction to Business intelligence and BI Tools | | | | | | | | | | |
| 6. | Introduction to Tableau and Data Preparation | | | | | | | | | | |
| 7. | Creating basic charts and graphs | | | | | | | | | | |
| 8. | Table Calculation | | | | | | | | | | |
| 9. | Advanced Chart forms and tools | | | | | | | | | | |
| 10. | Creating and publishing Dashboard | | | | | | | | | | |
| 11. | Dashboard Design Project and Presentation (ชุดข้อมูล 43 files) | | | | | | | | | | |

| หัวข้อ | | ระดับความพึงพอใจ | | | | |
|---|--|------------------|---|---|---|---|
| | | 5 | 4 | 3 | 2 | 1 |
| ส่วนที่ 4 ความเหมาะสมของเนื้อหาและเวลา | | | | | | |
| 1. | ระยะเวลาที่สอนมีความสอดคล้องกับปริมาณเนื้อหา | | | | | |
| 2. | ระยะเวลาในการฝึกปฏิบัติมีความเหมาะสมทำให้มีความเข้าใจมากขึ้น | | | | | |
| ส่วนที่ 5 สื่อ อุปกรณ์การสอนและสถานที่จัดอบรม | | | | | | |
| 1. | จอและสไลด์มีความคมชัด ง่ายต่อการมองเห็น | | | | | |
| 2. | เอกสารประกอบการเรียนมีความเหมาะสม | | | | | |
| 3. | เสียงมีคุณภาพดี ชัดเจน | | | | | |
| 4. | มีแสงสว่างเหมาะสม | | | | | |
| 5. | สัญญาณไวไฟ | | | | | |

ส่วนที่ 6 เจตคติต่อการจัดการและการวิเคราะห์ข้อมูล หลังเรียนหลักสูตรผ่านทั้งการบรรยาย และการปฏิบัติ

| ข้อความถาม | ระดับเจตคติ | | | | |
|---|-------------|---|---|---|---|
| | 1 | 2 | 3 | 4 | 5 |
| 1. หลังจากได้รับการอบรม ท่านมีความรู้เพิ่มมากขึ้นเพียงใด | | | | | |
| 2. ท่านสามารถนำความรู้ที่ได้รับไปประยุกต์ใช้ในการปฏิบัติงานได้ | | | | | |
| 3. ท่านสามารถนำความรู้ไปเผยแพร่/ถ่ายทอดได้ | | | | | |
| 4. ความพึงพอใจภาพรวมของการอบรม | | | | | |
| 5. หลักสูตรนี้ มีเนื้อหาที่เหมาะสมกับการปฏิบัติงานของท่าน | | | | | |
| 6. ท่านมีความสนใจศึกษาต่อ หากมีการจัดการอบรมในปีหน้าต่อไป | | | | | |
| 7. ท่านสามารถนำเทคนิคและวิธีการในการจัดการและการวิเคราะห์ข้อมูลไปใช้ได้ | | | | | |

เนื้อหาที่ต้องการให้มีเพิ่มเติมในหลักสูตร และข้อเสนอแนะเพิ่มเติม

.....

.....

.....

.....

ขอบคุณค่ะ

แบบประเมินการนำเสนอผลงาน หลักสูตรการจัดการและการวิเคราะห์ข้อมูล สาขาวิทยาศาสตร์ข้อมูล
และวิศวกรรมข้อมูล

ทีมที่.....

คำชี้แจง

1. แบบประเมิน ทั้งหมด 10 ข้อ โปรดให้ข้อมูลครบทุกข้อ เพื่อความสมบูรณ์ของการประเมินผลงาน
2. โปรดใส่คะแนนในช่อง คะแนนที่ได้ ในแต่ละหัวข้อประเมิน

| หัวข้อประเมิน | คะแนนเต็ม (คะแนน) | คะแนนที่ได้ (คะแนน) |
|--|----------------------|------------------------|
| 1. ความเหมาะสมของการจัดเรียงเนื้อหาเสนอ | 10 | |
| 2. ความชัดเจนของวัตถุประสงค์ | 10 | |
| 3. ความน่าเชื่อถือและความถูกต้องของผลงาน การเลือกใช้ วิธีการหรือโมเดลที่เหมาะสม | 10 | |
| 4. ความน่าสนใจ/ความคิดสร้างสรรค์ของผลงาน | 10 | |
| 5. การถ่ายทอดเนื้อหา | 10 | |
| 6. การนำผลงานไปใช้ประโยชน์ | 10 | |
| 7. เวลาที่ใช้ในการนำเสนอ | 10 | |
| 8. ความสนใจผลงานของผู้วิพากษ์ (เลือก 0, 20) | 20 | |
| 9. ความมีวินัยและตรงต่อเวลา (ตลอดหลักสูตรไม่เข้าสาย) | 10 | 10 |
| รวม | 100 | |

ข้อคิดเห็นเพิ่มเติม

.....

.....

.....

.....

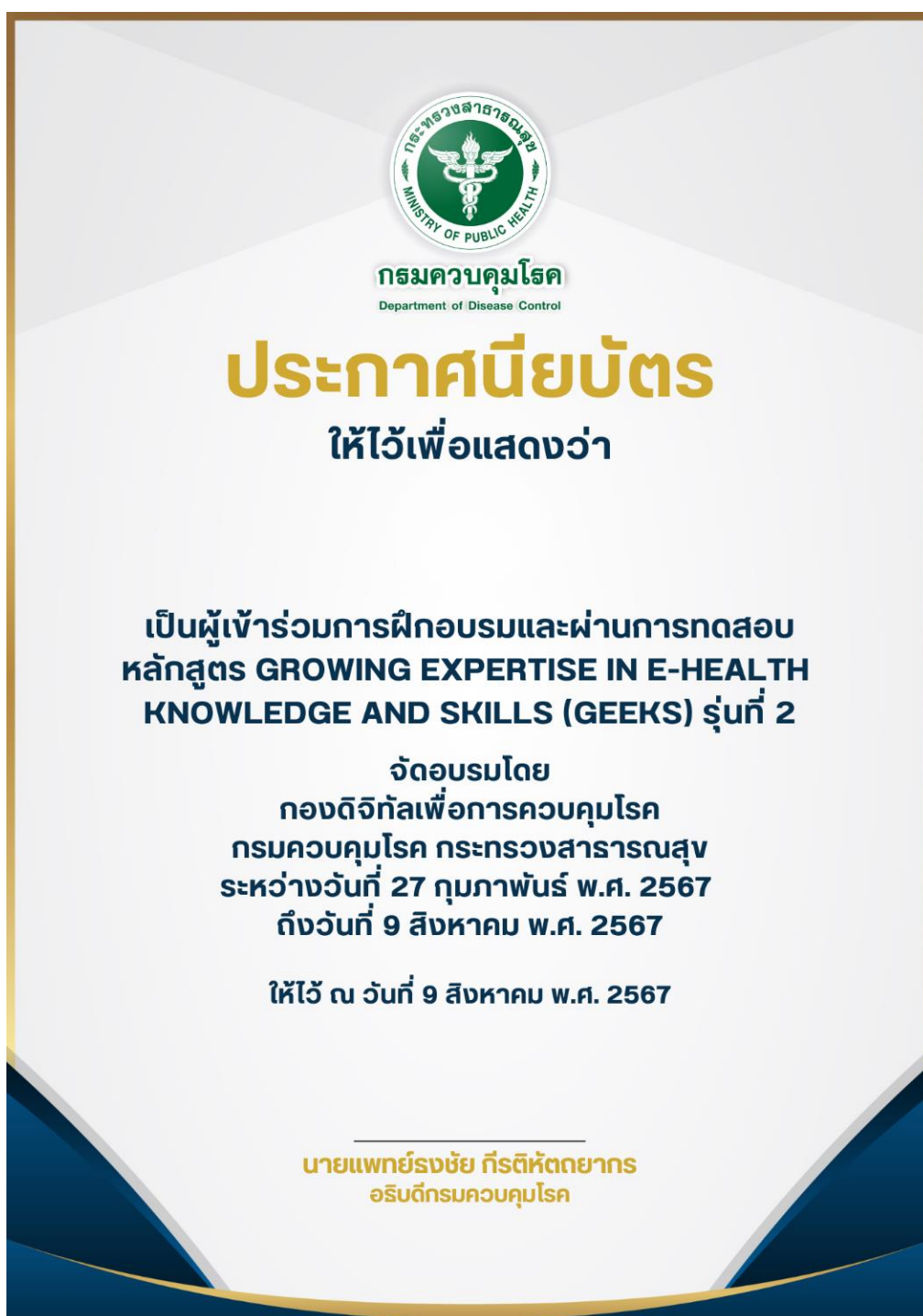
ลงชื่อ.....(ผู้ประเมิน)

(.....)

วันที่...../...../.....

ตัวอย่างใบเกียรติบัตร

เมื่อเรียนจบหลักสูตรการจัดการและการวิเคราะห์ข้อมูล ระดับ Advance ทั้งสาขาวิทยาศาสตร์ข้อมูล และวิศวกรรมข้อมูล ผู้เข้าอบรมจะได้รับเกียรติบัตร ซึ่งมอบในวันสุดท้ายของการอบรม ดังตัวอย่างนี้





กรมควบคุมโรค
Department of Disease Control

ประกาศนียบัตร

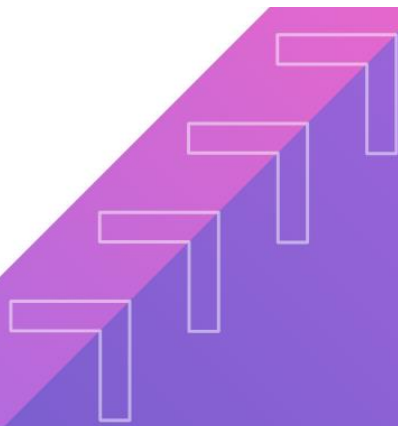
ให้ไว้เพื่อแสดงว่า

เป็นผู้เข้าร่วมการฝึกอบรมหลักสูตร
GROWING EXPERTISE IN E-HEALTH
KNOWLEDGE AND SKILLS (GEEKS) รุ่นที่ 2

จัดอบรมโดย
กองดิจิทัลเพื่อการควบคุมโรค
กรมควบคุมโรค กระทรวงสาธารณสุข
ระหว่างวันที่ 27 กุมภาพันธ์ พ.ศ. 2567
ถึงวันที่ 9 สิงหาคม พ.ศ. 2567

ให้ไว้ ณ วันที่ 9 สิงหาคม พ.ศ. 2567

นายแพทย์รังชัย กิรติหัตถยากร
อธิบดีกรมควบคุมโรค



กรมควบคุมโรค
Department of Disease Control

DATA MANAGEMENT AND ANALYSIS COURSE: