

แบบฝึกหัดบทที่ 1

1. จงอธิบายข้อมูลประเภทต่างๆ ต่อไปนี้ พร้อมยกตัวอย่างข้อมูลในแต่ละประเภท

- ข้อมูลที่มีโครงสร้าง (structured data)

ตอบ ข้อมูลที่มีโครงสร้างคือข้อมูลที่จัดเก็บในรูปแบบที่เป็นระเบียบและสามารถจัดการได้ง่ายในฐานข้อมูลที่เป็นตารางและมักมีรูปแบบที่แน่นอน เช่น ชื่อคอลัมน์ ประเภทของข้อมูล

- ข้อมูลที่ไม่มีโครงสร้าง (unstructured data)

ตอบ ข้อมูลที่ไม่มีโครงสร้างคือข้อมูลที่ไม่ได้จัดเก็บในรูปแบบที่เป็นระเบียบ ไม่สามารถจัดเก็บในตารางหรือตารางข้อมูลได้ง่าย ข้อมูลเหล่านี้มักจะเป็นข้อความ รูปภาพ วิดีโอ หรือไฟล์เสียง

- ข้อมูลกึ่งโครงสร้าง (semi-structured data)

ตอบ ข้อมูลกึ่งโครงสร้างคือข้อมูลที่ไม่ได้จัดเก็บในรูปแบบที่เป็นระเบียบอย่างเต็มที่ แต่มีองค์ประกอบที่สามารถจัดเก็บในรูปแบบโครงสร้างได้ ข้อมูลเหล่านี้มักอยู่ในรูปแบบ XML, JSON หรือเอกสาร HTML ที่มีแท็กเพื่อกำหนดโครงสร้างบางส่วน

2. จงอธิบายถึงความแตกต่างของ Database และ Data Warehouse

ตอบ **Database** เน้นการจัดเก็บและจัดการข้อมูล เหมาะสำหรับการทำงานที่ต้องการการเพิ่ม ลบ และแก้ไขข้อมูลอย่างรวดเร็ว มักใช้กับระบบที่ต้องการความเร็วสูงและการเข้าถึงข้อมูลที่แม่นยำในทันที ในขณะที่ **Data Warehouse** ออกแบบมาเพื่อการวิเคราะห์และรายงานข้อมูลระยะยาว มุ่งเน้นการวิเคราะห์เชิงลึกและการตัดสินใจทางธุรกิจบนพื้นฐานของข้อมูลย้อนหลังและข้อมูลที่รวบรวมจากแหล่งต่าง ๆ

3. จงอธิบายถึงความสัมพันธ์ระหว่าง Big Data และ Data Lake

ตอบ Data Lake ถูกออกแบบมาเพื่อเก็บข้อมูลขนาดใหญ่ที่ไม่สามารถจัดการได้โดยระบบฐานข้อมูลแบบดั้งเดิม สามารถเก็บข้อมูลทุกประเภทได้ ไม่ว่าจะเป็นข้อมูลแบบมีโครงสร้าง, กึ่งโครงสร้าง, และไม่มีโครงสร้าง และยังสามารถประมวลผลข้อมูลขนาดใหญ่ได้อย่างมีประสิทธิภาพ Data Lake จึงเป็นเครื่องมือที่สำคัญในการจัดเก็บและจัดการ Big Data เพื่อให้สามารถนำข้อมูลเหล่านี้ไปประมวลผลและวิเคราะห์ได้อย่างมีประสิทธิภาพ