

แบบฝึกหัดทบทวนก่อนสอบปลายภาค ประจำภาคเรียนที่ 1 ปีการศึกษา 2560
วิชา ภาษาการเขียนโปรแกรม รหัส 02-212-213 วท.บ. วิทยาการคอมพิวเตอร์

(1) จงเขียนอธิบายถึงข้อมูล (data)

- 1.1 ข้อมูล (data) คือ
- 1.2 ข้อมูล (data) ต่างจากสารสนเทศ (information) อย่างไร
- 1.3 คำกล่าว “ข้อมูลจะไม่มี ความหมายถ้าหากไม่นำมาทำการเชื่อมโยง” หมายความว่าอย่างไร

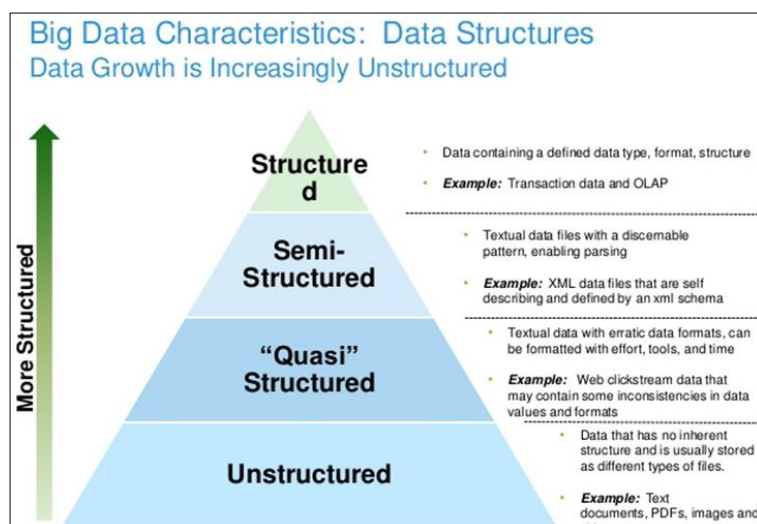
ปัจจุบันประเทศไทยก้าวเข้าสู่ยุคของ big data ที่มีองค์ประกอบของ 3 Vs ดังแสดงใน ภาพ1 ข้างล่างนี้ ได้แก่ ภาวะที่องค์กรมีข้อมูลมหาศาลและซับซ้อน ทั้งปริมาณ (volume) ความหลากหลายในประเภทและแหล่งนำเข้าข้อมูล (variety) และมีการปรับปรุงเปลี่ยนแปลงอย่างรวดเร็ว (velocity) ให้ตอบคำถามข้อ (2)



ภาพ1: 3 Vs (จากอินเทอร์เน็ต 2017)

(2) ให้สน.เขียนอธิบายย่อย่อ 2.1 - 2.3 ยกตัวอย่างองค์กร/บริษัท จากกรณีศึกษาที่ได้เรียนมา ทราบหรือสังเกตเห็นในชีวิตประจำวัน

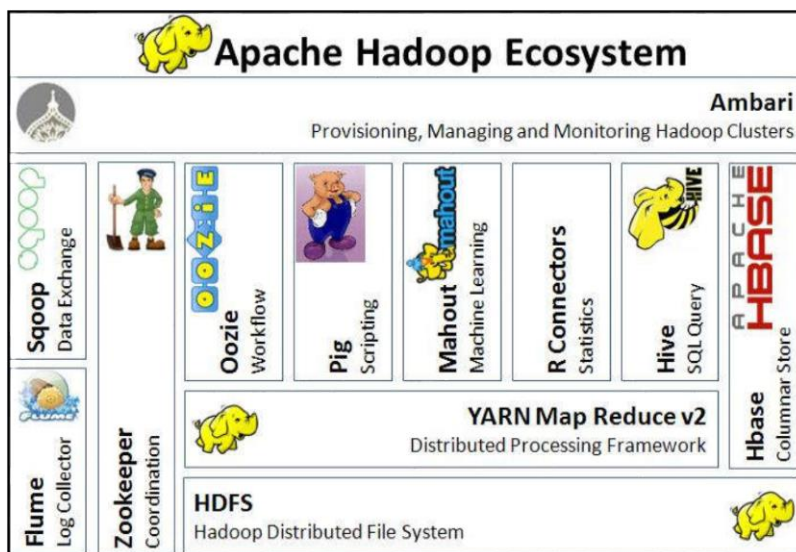
- 2.1 ตัวอย่างได้แก่องค์กร/บริษัท อะไร
 - 2.2 ข้อมูลที่จัดเก็บในองค์กร/บริษัทดังกล่าว ถือเป็น big data เพราะอะไร
 - 2.3 องค์กร/บริษัทดังกล่าว สามารถนำข้อมูล big data นั้นมาใช้งาน หรือช่วยแก้ปัญหาในองค์กร/บริษัท ได้อย่างไร
- (3) ภาพ2 ข้างล่างนี้แสดงถึงประเภทของข้อมูล (ตามโครงสร้าง) ในยุค big data ให้เขียนอธิบายในย่อย่อ 3.1 - 3.3 ว่าหมายถึงอะไร พร้อมยกตัวอย่างประกอบมาพอเข้าใจ



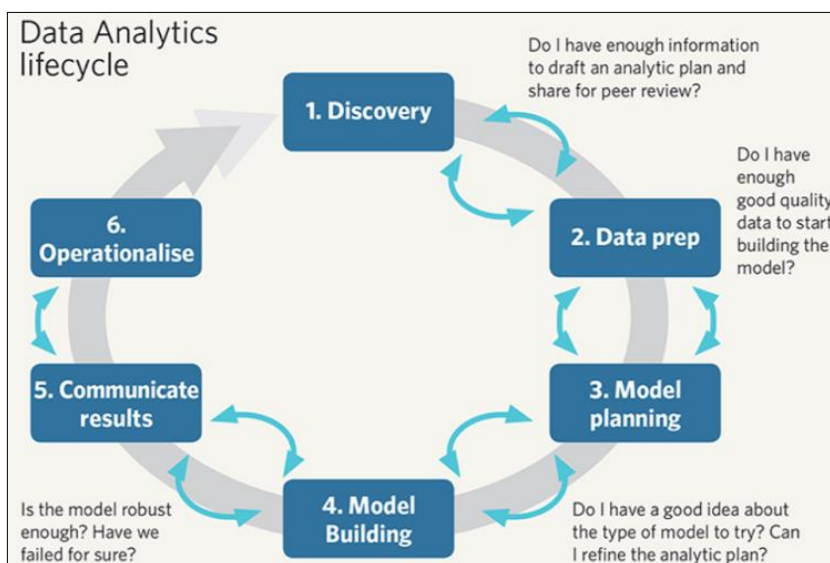
ภาพ2: big data structure (โดยEMC/Dell)

(4) จงบอกเทคโนโลยี เครื่องมือ หรือซอฟต์แวร์ ที่นำมาช่วยในการวิเคราะห์ข้อมูลขนาดมหาศาล (big data) ที่สน. ทราบมา สัก 2 ตัวอย่าง พร้อมทั้งอธิบายประกอบมาพอเข้าใจ

(5) Hadoop เป็นเฟรมเวิร์กที่นำมาใช้กับ big data ประกอบไปด้วย 2 ส่วนองค์ประกอบสำคัญได้แก่ ส่วนสำรองข้อมูล (data storage) สำหรับ big data และ MapReduce ในส่วนวิเคราะห์ข้อมูล (Analytics) ให้อธิบาย 2 ส่วนดังกล่าว โดยใช้ ภาพ 3 ประกอบ



ภาพ3: hadoop for big data (โดยEMC/Dell)



ภาพ4: Data analytics lifecycle (โดยEMC/Dell)

ให้แนศ. ใช้ ภาพ4 ข้างบนนี้ตอบคำถามในกระบวนการวิเคราะห์ข้อมูลในข้อ (6) 6.1 – 6.6

(6) ใช้วงจรการวิเคราะห์ข้อมูลจาก ภาพ4 ข้างบนมาเขียนอธิบายแต่ละกระบวนการของวงจรวิเคราะห์ข้อมูล

- 6.1 Data Discovery ได้แก่กระบวนการ / คำสั่ง R ในกระบวนการนี้ได้แก่
- 6.2 Data preparation ได้แก่กระบวนการ / คำสั่ง R ในกระบวนการนี้ได้แก่
- 6.3 Model planning ได้แก่กระบวนการ / คำสั่ง R ในกระบวนการนี้ได้แก่
- 6.4 Model building ได้แก่กระบวนการ / คำสั่ง R ในกระบวนการนี้ได้แก่
- 6.5 Communication results ได้แก่กระบวนการ / ผลลัพธ์ที่วิเคราะห์ข้อมูลได้แก่
- 6.6 Operationalize ได้แก่กระบวนการ