

Data Science Pathway Test

Part II: Data Analytics and Big Data Pathway Test

- This test/quiz can be taken ONLY ONCE.
- Correct answers will NOT be revealed after submission.

จงเลือกคำตอบที่ถูกต้องเพียง 1 ข้อ

1 กรณีศึกษา

องค์กรขนาดใหญ่แห่งหนึ่งทำธุรกิจเกี่ยวกับการให้บริการรถรับส่งโดยจองผ่านช่องทางออนไลน์บนมือถือ ซึ่งประสบความสำเร็จอย่างสูงจนมีผู้ใช้บริการเป็นจำนวนมาก และยังมีข้อมูลขนาดใหญ่ที่เข้ามาในระบบตลอดเวลา เช่น ชื่อผู้ใช้บริการ รายละเอียดการจ่ายเงิน หมายเลขทะเบียนรถ และสถานที่รับส่ง เป็นต้น ซึ่งในปัจจุบันข้อมูลขนาดใหญ่เหล่านี้จะถูกเก็บอยู่ในฐานข้อมูลขององค์กรโดยใช้ Super Computer จำนวน 1 เครื่องในการทำงาน นอกจากนั้นองค์กรยังมีช่องทางในการรับข้อมูลข่าวสารและคำติชมจากผู้ใช้บริการใน Social Media Page ที่มีผู้เข้ามาติชมเป็นจำนวนมากอีกด้วย ในส่วนของการให้บริการจะมีการส่งตำแหน่งของผู้ใช้บริการเข้ามาในระบบเป็นระยะ ๆ ตั้งแต่สถานที่เรียก ระหว่างทาง ตลอดจนถึงที่หมาย

ทางด้านประธานเจ้าหน้าที่บริหาร หรือ CEO (Chief Executive Officer) ขององค์กรได้เล็งเห็นถึงปัญหา 3 ข้อหลัก คือ

- ในเชิงการให้บริการ คนขับรถที่ลงทะเบียนไว้ในระบบ ไม่ยินดีที่จะให้บริการผู้ใช้งานในช่วงเวลาที่รถติดในบางสถานที่เนื่องจากระบบนั้นใช้การคำนวณค่าเดินทางแบบ Fixed pricing (เป็นการ Fix ราคา จากจุดรับไปที่จุดส่ง) ซึ่งทำให้รายได้ของคนขับและองค์กรน้อยลงในช่วงรถติด ทาง CEO จึงต้องการนำระบบ Dynamic Pricing (เป็นการคิดราคาตามสภาพรถติดและระยะทางที่เกิดขึ้นจริง) เข้ามาใช้แทนระบบปัจจุบัน
- ในข้อมูลคำติชมที่มีการส่งเข้ามาจากผู้ใช้บริการ มีข้อมูลเป็นจำนวนมากและมีความหลากหลายของเนื้อหา ทำให้ไม่สามารถรู้ได้ว่าหน่วยงานใดควรรับปัญหานี้ไปแก้ไข
- ทาง CEO ต้องการจะขยายธุรกิจไปทั่วประเทศ จึงจะต้องหาระบบที่สามารถรองรับข้อมูลขนาดใหญ่ได้มากขึ้น

จากกรณีศึกษาที่ให้ ให้ท่านตอบคำถามแต่ละข้อ

จากข้อมูลขององค์กรนี้ได้รับมา ข้อใดต่อไปนี้เป็น Geospatial Data?

- ☐ a. Car location ตำแหน่งรถยนต์
- ☒ b. Car registry ทะเบียนรถ
- ☐ c. Customer location ตำแหน่งลูกค้า
- ☐ d. Street Traffic การจราจร

2 Geospatial Data ขององค์กรนี้เกิดขึ้นจากสิ่งใด?

- ☒ a. จากการสื่อสารกันระหว่างอุปกรณ์กับมนุษย์
- ☐ b. จากการสื่อสารกันระหว่างอุปกรณ์ต่าง ๆ ผ่านอินเทอร์เน็ต
- ☐ c. จากการเชื่อมต่อระบบอินเทอร์เน็ตเข้ากับอุปกรณ์ต่าง ๆ เช่น Laptop
- ☐ d. จากการเชื่อมต่อกันระหว่างอุปกรณ์ 2 อุปกรณ์ที่สามารถรับ Wifi ได้

3 จากข้อมูลสถานที่รับและส่งของลูกค้า ข้อมูลดังกล่าวจัดอยู่ใน V ใดบ้าง?

1. Volume
2. Variety
3. Velocity
4. Veracity

จงเลือกข้อที่ถูกต้อง

- ☐ a. 1 และ 2
- ☐ b. 1 และ 3
- ☐ c. 1,2 และ 3
- ☒ d. 1,2,3 และ 4
-

4 ในปัจจุบัน องค์กรได้รับคำติชมจำนวนมากจากผู้ใช้งานผ่านช่องทาง Social Media Page, หากคุณเป็นนักวิทยาศาสตร์ข้อมูล คุณต้องแนะนำวิธีให้กับทีมงาน คุณจะแนะนำวิธีใดในการแยกหัวข้อเป็นคำติชมด้านบริการ คำติชมด้านราคา คำติชมด้าน Mobile App ฯลฯ

- ☐ a. Classification
- ☒ b. Clustering
- ☐ c. Image processing
- ☐ d. Spatial analytics
-

5 ทาง CEO ได้มองเห็นว่าในอีก 1 ปีข้างหน้า นั้น องค์กรจะมีข้อมูลที่มากขึ้นจากจำนวน 1 TB ไปเป็นจำนวน 100 TB และคาดว่า จะมีการประมวลผลเพิ่มขึ้นเป็นจำนวนมาก หากคุณจะต้องแนะนำวิธีที่เหมาะสมให้กับ CEO คุณจะแนะนำวิธีใด?

- ☐ a. ขยายระบบโดยเพิ่ม Super computer ที่มีประสิทธิภาพในการประมวลผลอย่างมาก
- ☒ b. ขยายระบบโดยนำเทคโนโลยี Map-reduce มาใช้กับ Server หลาย ๆ เครื่อง
- ☐ c. ขยายระบบโดยใช้ Virtual machine
- ☐ d. ขยายระบบโดยใช้ Relational database
-

6 หากคุณเป็นทีมงาน Data Science ขององค์กร คุณจะเลือกใช้ Data Analytics วิธีใดในการคำนวณขนาดความต้องการของผู้ใช้บริการร่วมกับปริมาณคนขับที่ต้องการให้บริการเพื่อหาราคาที่เหมาะสมในแต่ละสถานการณ์?

- ☐ a. ทำ Predictive analysis ด้วยข้อมูลสองอย่างคือ จุดหมายปลายทาง และ การจราจร
- ☐ b. ทำ Survey เพื่อวิเคราะห์ความต้องการของผู้ใช้บริการและคนขับ
- ☒ c. ทำ Data exploration
- ☐ d. ทำ Data visualization
-

7 ข้อใดไม่ใช่ “valuable insight” ที่ได้จาก Social media ขององค์กร

- ☒ a. จำนวน feedback แยกตามความยาวข้อร้องเรียน
- ☐ b. การร้องเรียนเรื่องราคาจะมาจากคนในเมืองใหญ่เป็นจำนวนมาก
- ☐ c. การร้องเรียนเรื่องราคาจะมาจากผู้หญิงเป็นส่วนใหญ่
- ☐ d. เมื่อราคาเพิ่ม จะมีปริมาณการร้องเรียนก็จะเพิ่มขึ้นตาม
-

The latest submission was made on 28-Feb-2020 @ 12:19:29



1.12.6.1 | [Privacy Policy](#) updated on Jun 30th, 2017

The development of myCourseVille is partially supported by [Learning Innovation Center, Chulalongkorn University](#)