รหัสวิชา 2-233-225
ชื่อวิชา วิทยาศาสตร์ข้อมูลเบื้องต้น
(Introduction to Data Science)
สัปดาห์ที่ 1

CS

รายละเอียดรายวิชา

ภาคเรียนที่ 2 ปีการศึกษา 2568

- ✓ จำนวนหน่วยกิต: 3(2-2-5)
- ✓ เวลาเรียน: วันพฤหัสบดี เวลา 08.00-12.00 น.
- ✓ อาจารย์ผู้สอน: ผศ.ดร.อรสา พัสดุ
- ✓ ห้องเรียน S901

CS

ข้อปฏิบัติในการเข้าชั้นเรียน

- ให้นักศึกษาทุกคนเช็คชื่อเข้าชั้นเรียน เวลา 08.00-08.15 น. หากเข้าชั้นเรียนเกิน 08.15 น. ถือว่า เข้าเรียนสาย
- การเข้าเรียน ต้องมาเรียนเกิน 80 %
- ประกาศเรื่องต่าง ๆ เกี่ยวกับวิชา ผ่าน google classroom ของรายวิชา*

• ช่องทางติดต่อกับอาจารย์ผู้สอน หากนักศึกษามีข้อสงสัยเพิ่มเติม

- ✓ สามารถติดต่ออาจารย์ผู้สอนผ่านทางอีเมล์ : orasa.p@mail.rmutk.ac.th
- ✓ เวลาที่สามารถติดต่อ วันศุกร์ เวลา 08.30-16.30 น.

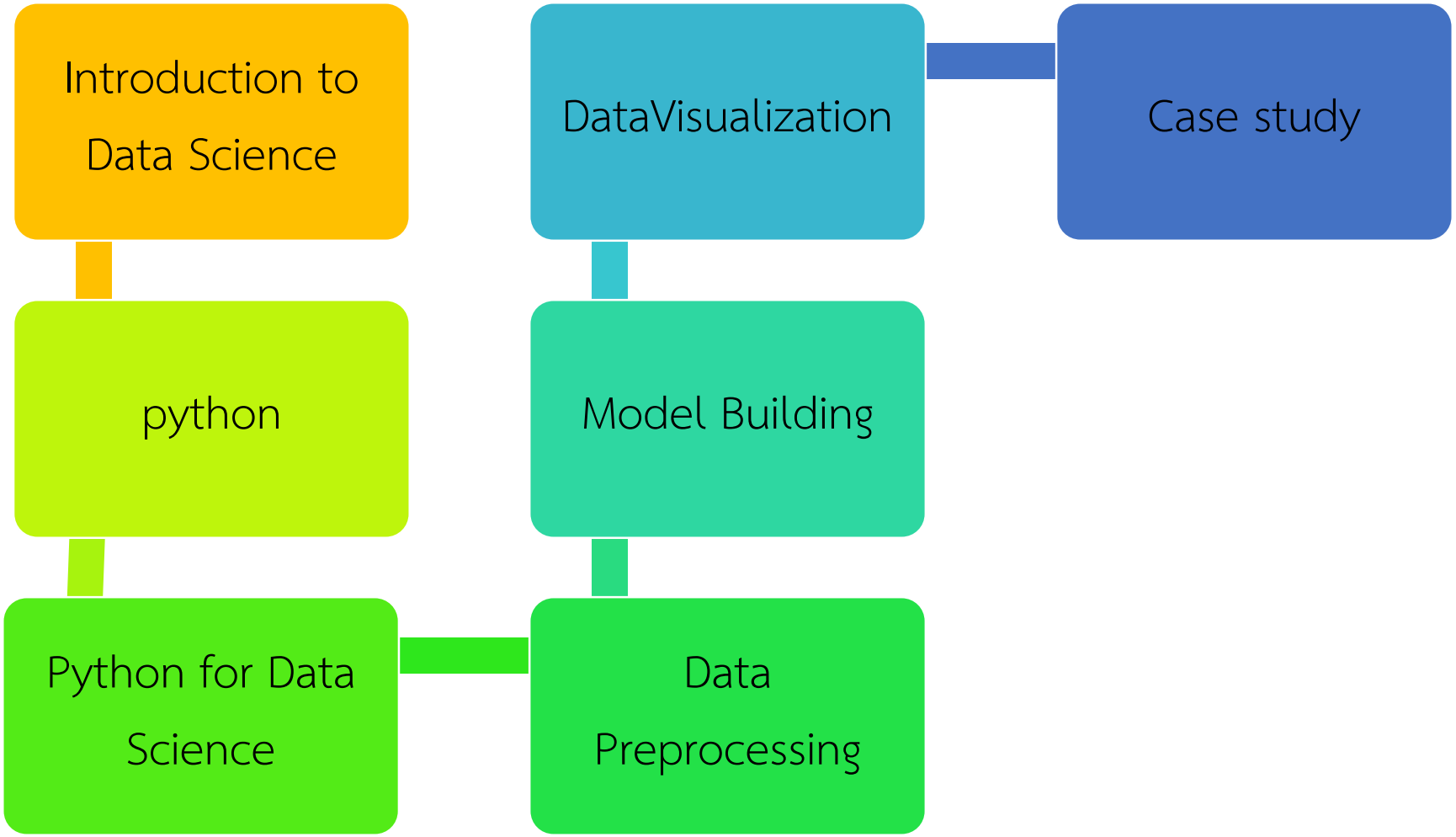
CS

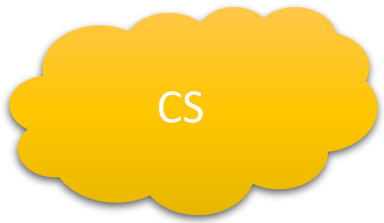
วัตถุประสงค์ของรายวิชา

- เพื่อให้นักศึกษามีความรู้ ความเข้าใจเกี่ยวกับพื้นฐาน องค์ประกอบ ประโยชน์และผลกระทบของการใช้เทคโนโลยีการทำวิทยาศาสตร์ข้อมูล
- เพื่อให้นักศึกษามีความรู้ในการคิดวิเคราะห์ปัญหาและสามารถแก้ปัญหาโดยใช้วิทยาศาสตร์ข้อมูล
- เพื่อให้นักศึกษามีความรู้ ความเข้าใจ เกี่ยวกับกรณีศึกษาด้านวิทยาศาสตร์ข้อมูล big data และกลยุทธ์ทางธุรกิจ แนวโน้มเทคโนโลยีการทำวิทยาศาสตร์ข้อมูลในอนาคต
- เพื่อให้นักศึกษามีความรู้จักกับเครื่องมือซอฟต์แวร์เพื่อการทำวิทยาศาสตร์ข้อมูลขั้นพื้นฐาน (Tutorials on data science tool)

CS

เนื้อหาของรายวิชา





หนังสือ

หนังสือ

Jake VanderPlas . Python Data Science Handbook . O'REILLY, 2017.

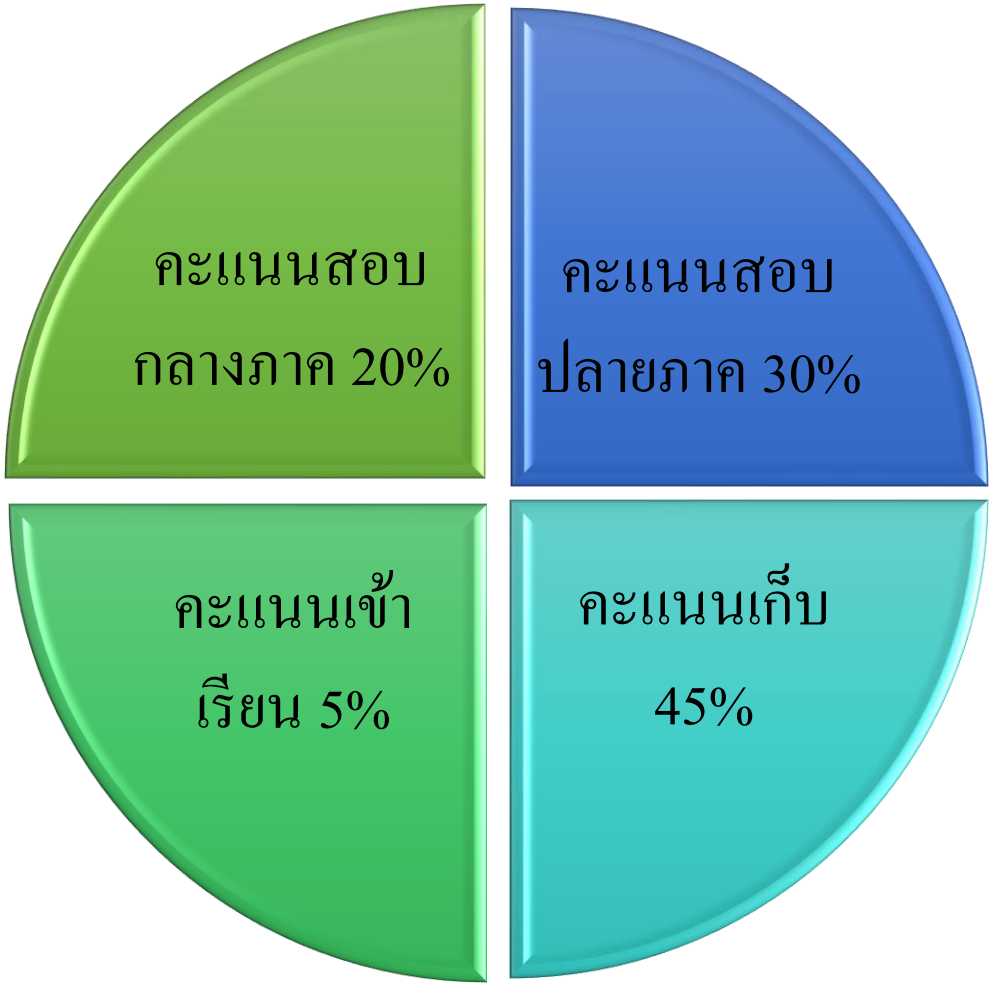
Jesús Rogel-Salazar. DATA SCIENCE AND ANALYTICS WITH PYTHON CRC Press Taylor & Francis Group, 2017.

อรพิน ประวัติบริสุทธิ์, Python สำหรับงาน Data Science Data Visualization และ Machine Learning, บริษัทโปรวิชั่น จำกัด, กรุงเทพ, 2564.

กอบเกียรติ สระอุบล, เรียนรู้ Data Science และ AI: Machine Learning ด้วย Python, หสม มีเดีย เนทเวิร์ค, พิมพ์ครั้งที่1, กรุงเทพ, 2563.

CS

การวัดผล


คะแนนสอบ
กลางภาค 20%
คะแนนสอบ
ปลายภาค 30%
คะแนนเข้า
เรียน 5%
คะแนนเก็บ
45%

CS

รายละเอียดคะแนน

รายละเอียด	คะแนน
สอบกลางภาค	20
สอบปลายภาค	30
คะแนนเก็บ -สอบย่อย -งานที่ได้รับมอบหมายเดี่ยว -งานกลุ่ม+รายงาน	 10 15 20
เข้าเรียน	5
รวม	100

CS

เกณฑ์ค่าระดับคะแนน

< 49 คะแนน (F)

65-69 คะแนน (C+)

50-54 คะแนน (D)

70-74 คะแนน (B)

55-59 คะแนน (D+)

75-79 คะแนน (B+)

60-64 คะแนน (C)

> 80 คะแนน (A)

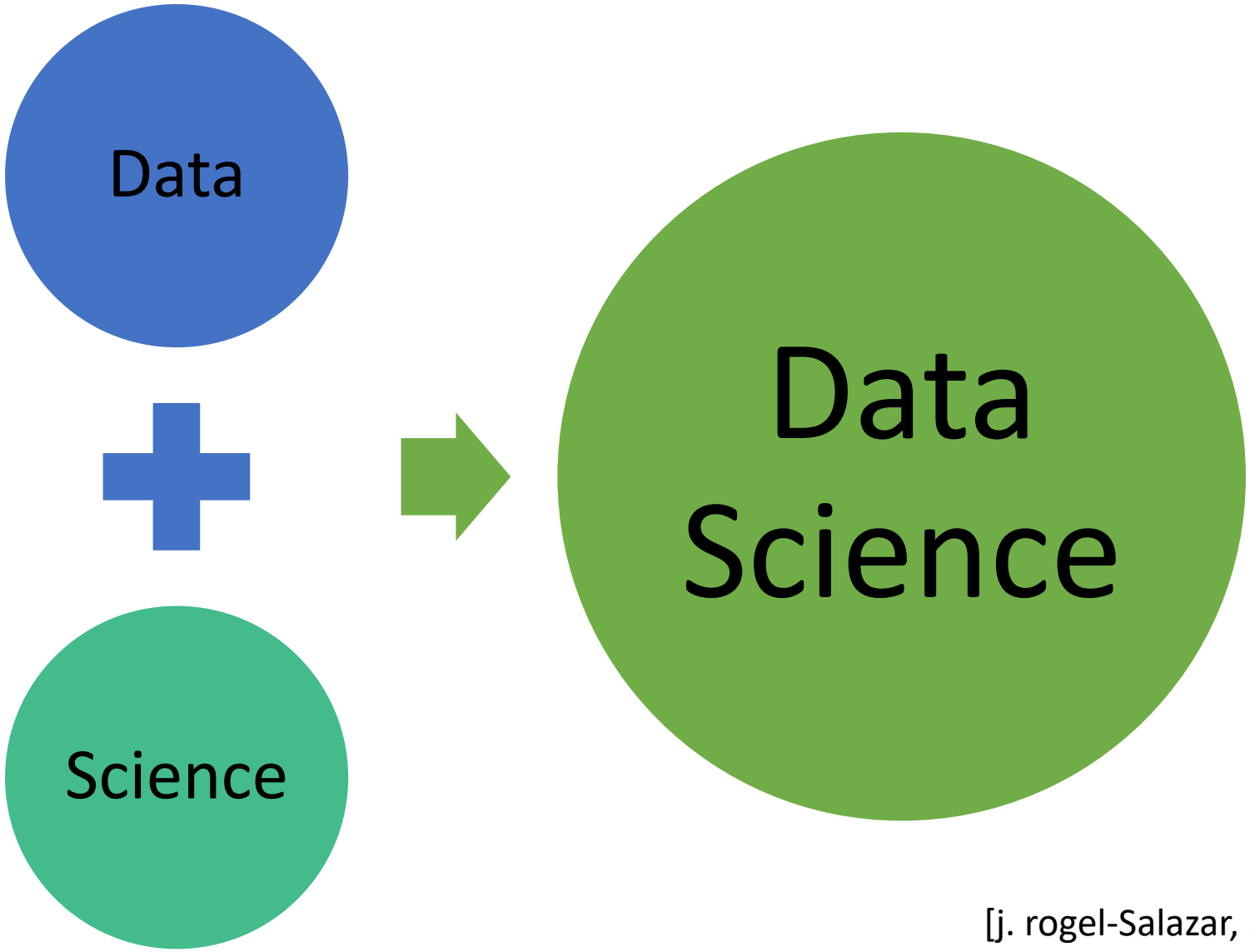
CS

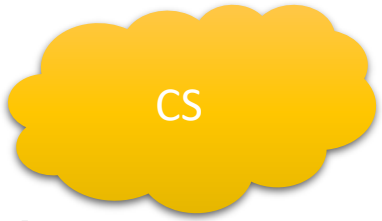
วัตถุประสงค์ของบทนี้

- เพื่อเข้าใจถึงภาพรวมของ Data Science
- เพื่อเข้าใจว่า Data Science คืออะไร
- เข้าใจถึงบทบาทของ Data scientist
- เข้าใจถึงการนำ Data science ไปประยุกต์ใช้ในการแก้ปัญหา
- เข้าใจถึงขั้นตอนในการทำ Data mining
- เข้าใจถึงเครื่องมือในการทำ Data science โดยใช้ Python และเครื่องมือต่าง ๆ

CS

ภาพรวมของ Data Science





[j. rogel-Salazar, 2017]

Data

LINE

Discord

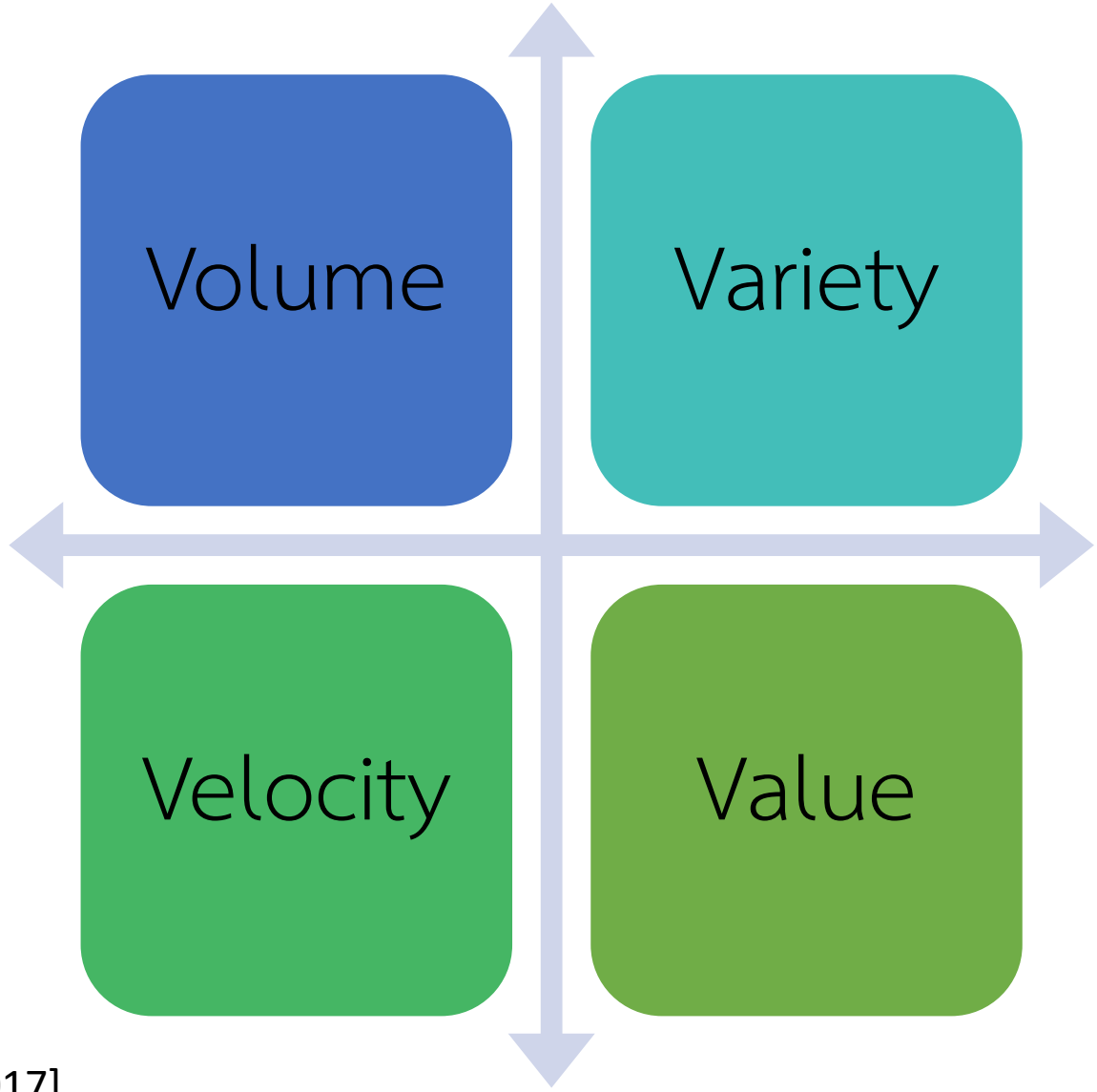
YouTube

twitter

Pantip
Learn, Share & Fun

CS

Big Data



CS

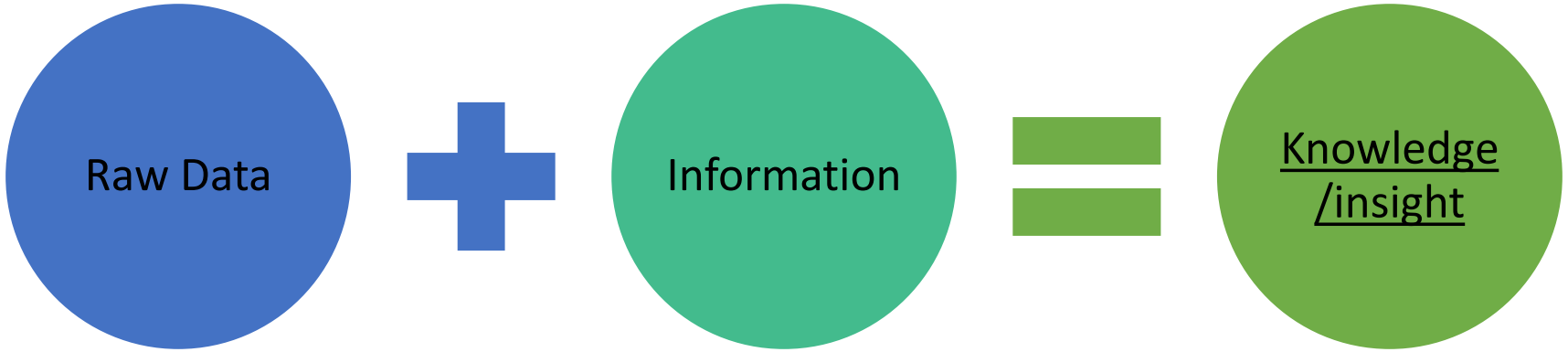
[j. rogel-Salazar, 2017]

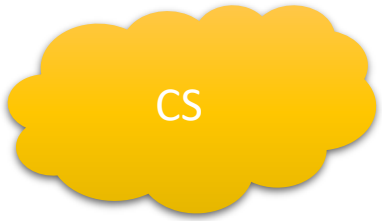
Data Science คืออะไร.....

Data Science คือการค้นพบสิ่งที่อาจไม่เคยรู้มาก่อนจากข้อมูลช่วยทำให้การบริหารจัดการและการตัดสินใจมีความแม่นยำมากขึ้น (กอบเกียรติ สระอุบล, 2563)

Data Science คือวิทยาศาสตร์ข้อมูลเป็นศาสตร์อย่างหนึ่งที่รวมเอาความรู้ด้านการเขียนโปรแกรม ด้านคณิตศาตร์ และสถิติมาประยุกต์ใช้เพื่อให้เกิดองค์ความรู้จากข้อมูล เพื่อให้ได้ข้อมูลที่มีมู,ค่าสามารถนำไปช่วยในการตัดสินใจวางแผนธุรกิจได้ (อรพิน ประวัติบริสุทธิ์, 2564)

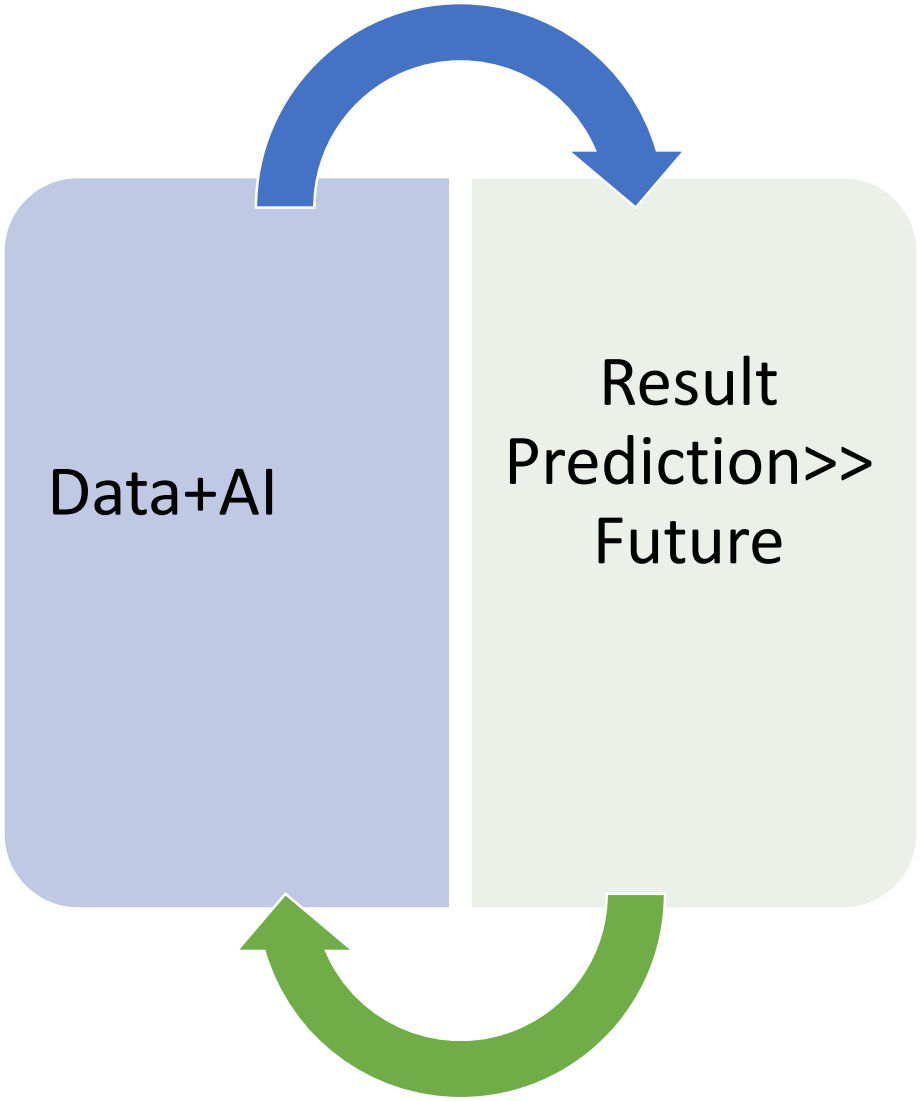
Data Science คืออะไร.....

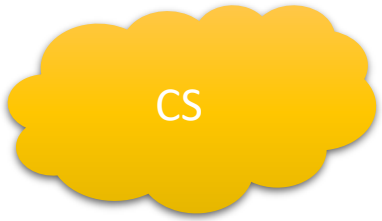




[j. rogel-Salazar, 2017]

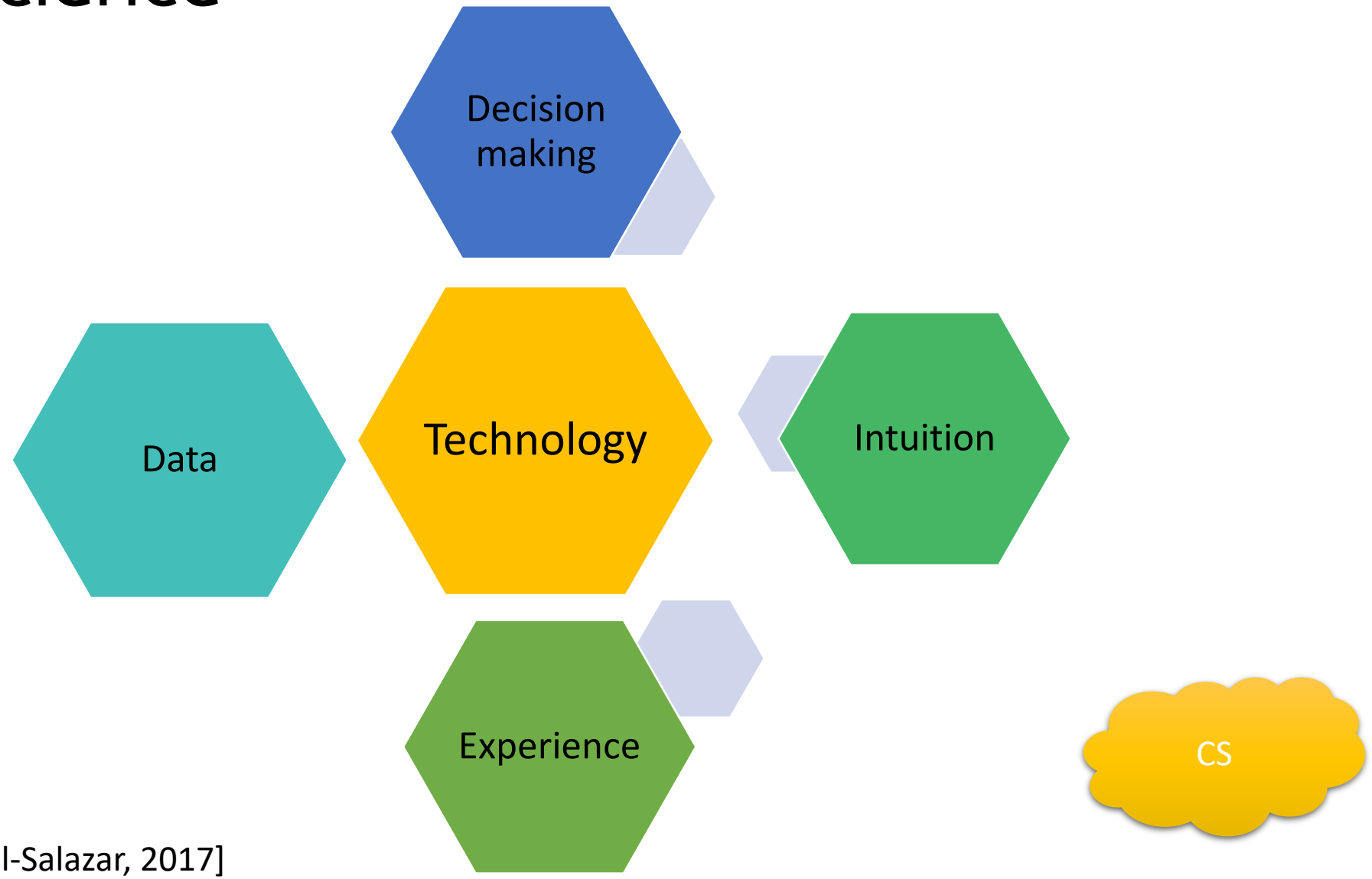
Data Science คืออะไร.....





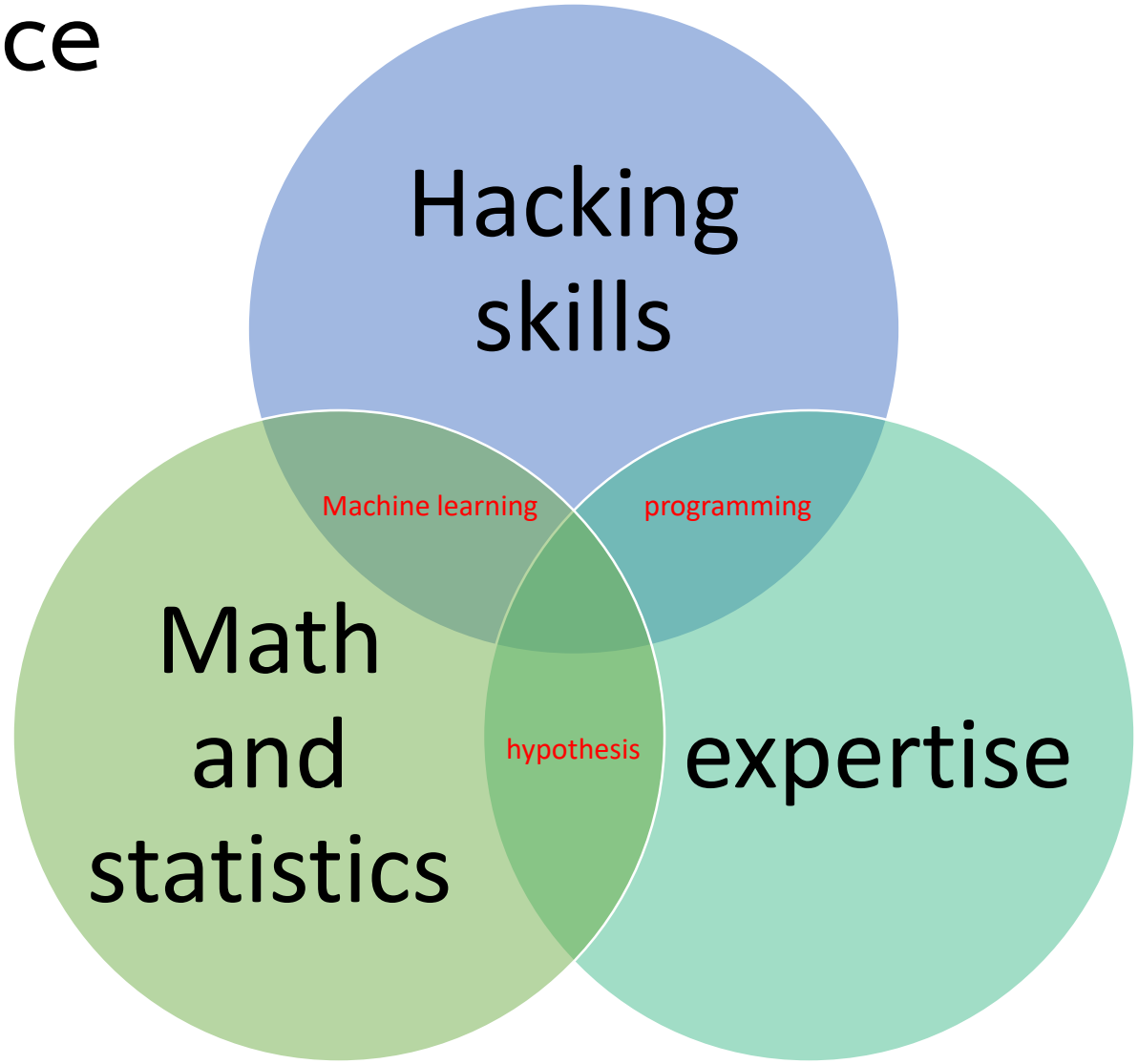
[j. rogel-Salazar, 2017]

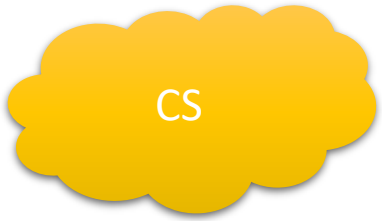
Data Science



[j. rogel-Salazar, 2017]

Data Science





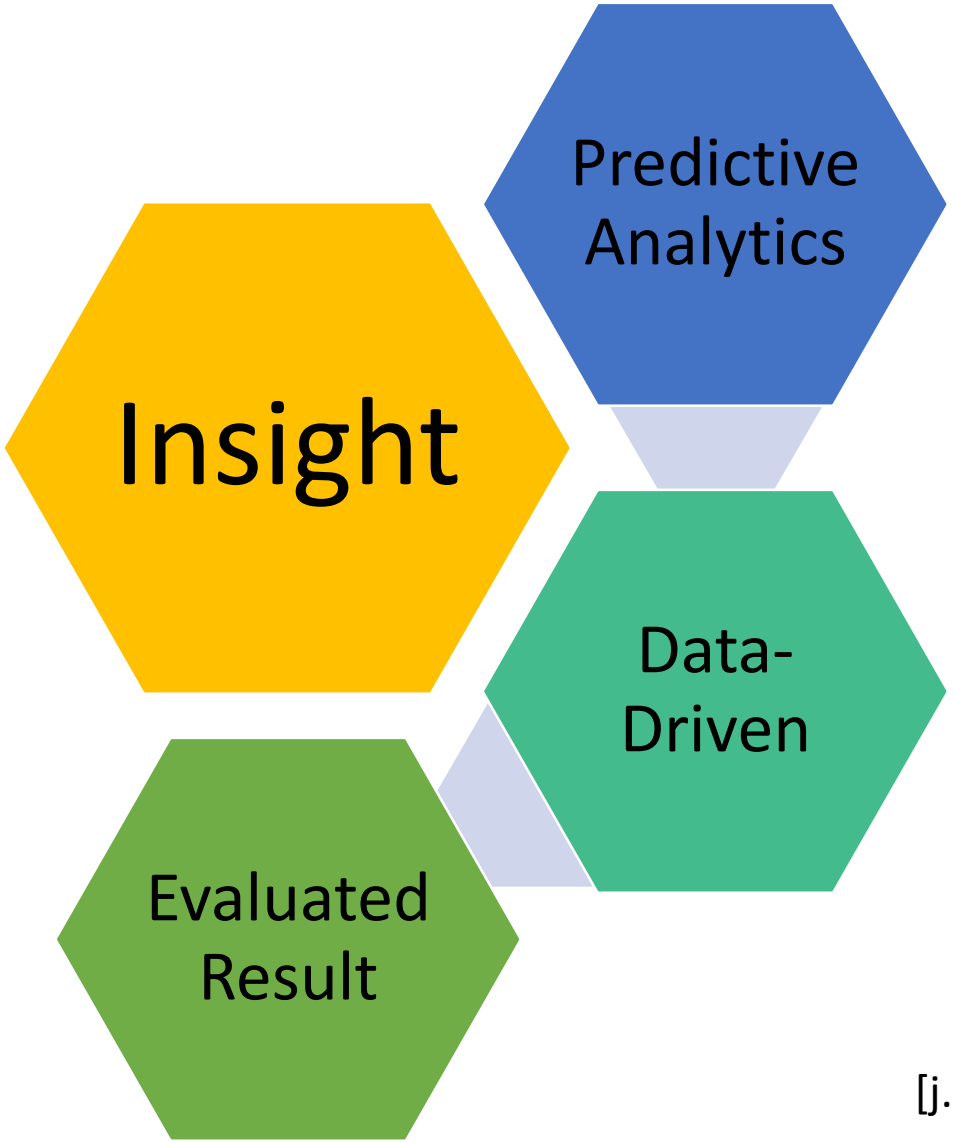
[j. rogel-Salazar, 2017]

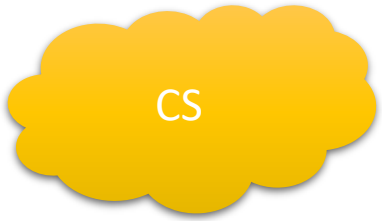
Data Visualization

สิทธินนท์ ดำรงศักดิ์และวศิน แสงเวช, 2566

CS

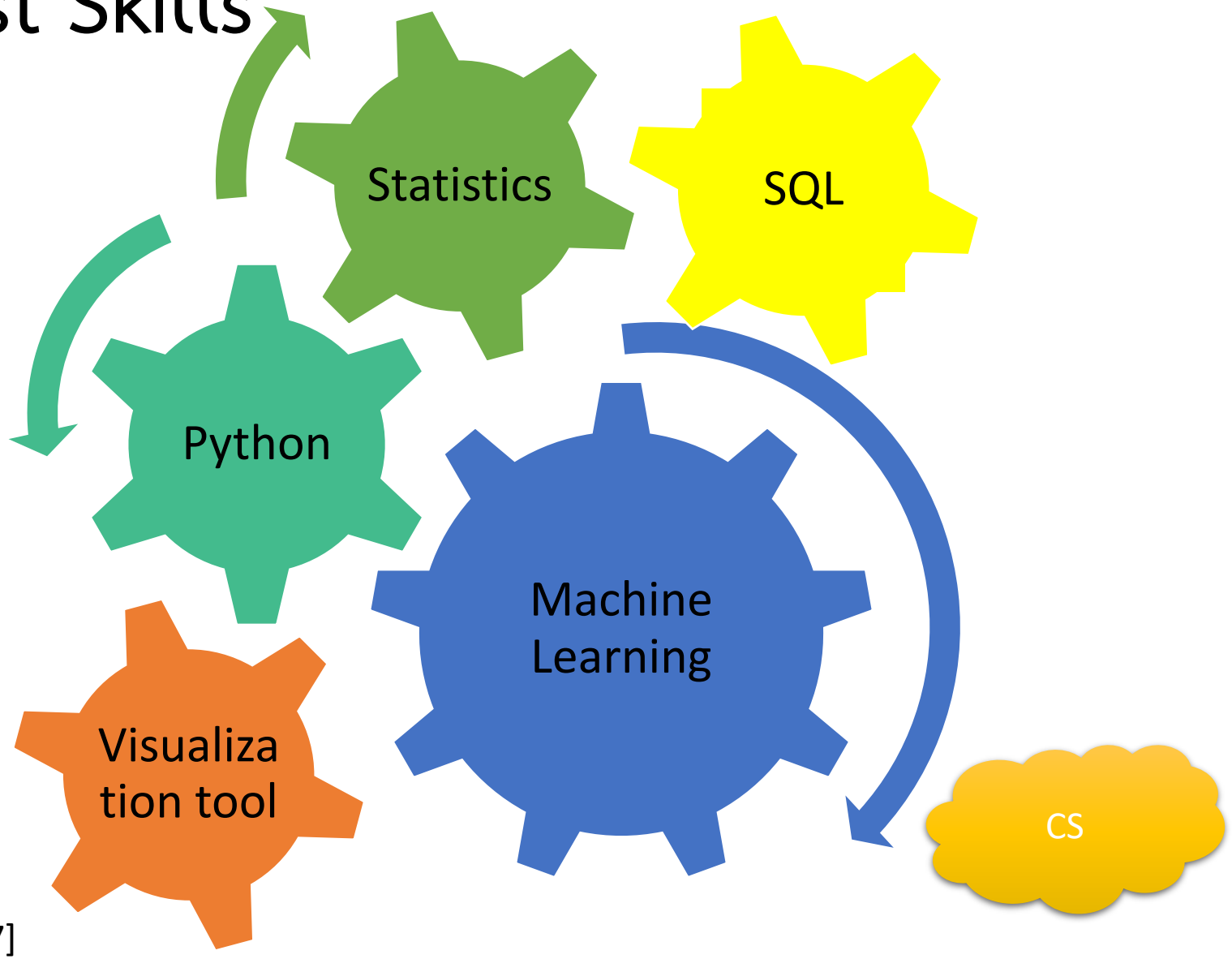
Who is Data Scientist?





[j. rogel-Salazar, 2017]

Data Scientist Skills



[j. rogel-Salazar, 2017]

Data Scientist Skills

ตั้งข้อสงสัย

Logic

เข้าใจธุรกิจ

CS

[j. rogel-Salazar, 2017]

Get requirement/เก็บรวบรวมข้อมูล

จัดการข้อมูล+Data Engineering

หน้าที่ของ Data Scientist

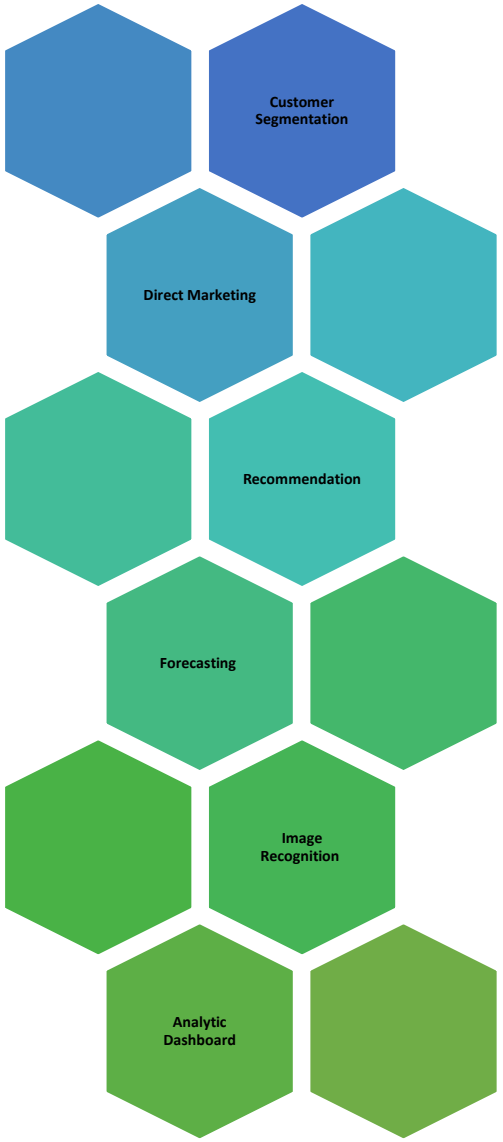
วิเคราะห์ข้อมูล (Predictive modeling)

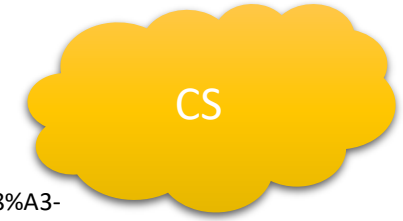
นำเสนอข้อมูล (Visualization)

CS

[j. rogel-Salazar, 2017]

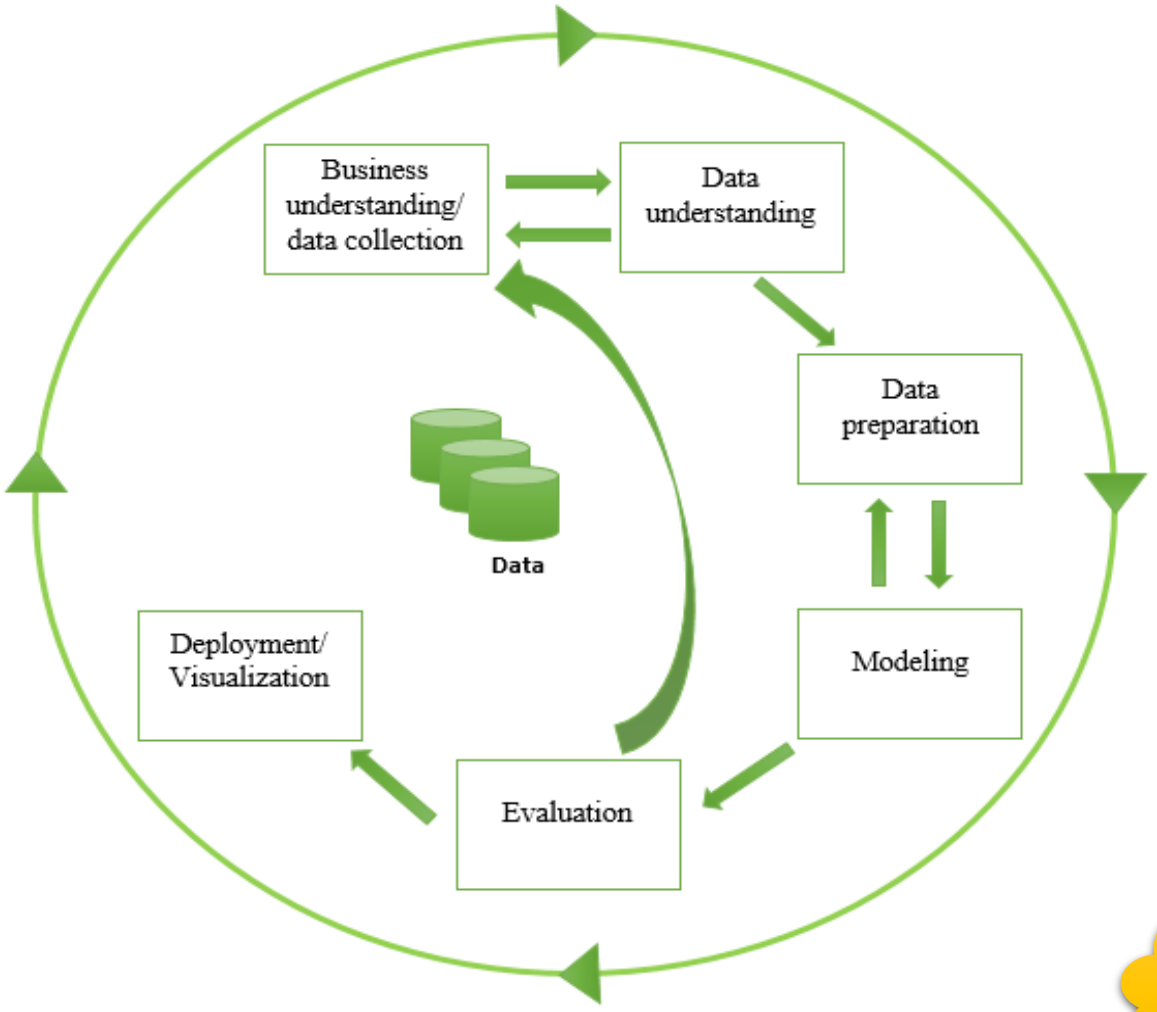
การนำ Data science ไปประยุกต์ใช้ในการแก้ปัญหา





https://medium.com/@thanyavuth/data-scientist-%E0%B8%97%E0%B8%B3%E0%B8%AD%E0%B8%B0%E0%B9%84%E0%B8%A3-%E0%B8%AD%E0%B8%A2%E0%B8%B2%E0%B8%81%E0%B9%80%E0%B8%9B%E0%B9%87%E0%B8%99%E0%B8%95%E0%B9%89%E0%B8%AD%E0%B8%87%E0%B8%97%E0%B8%B3%E0%B8%AD%E0%B8%A2%E0%B9%88%E0%B8%B2%E0%B8%87%E0%B9%84%E0%B8%A3-16f1e8028d72

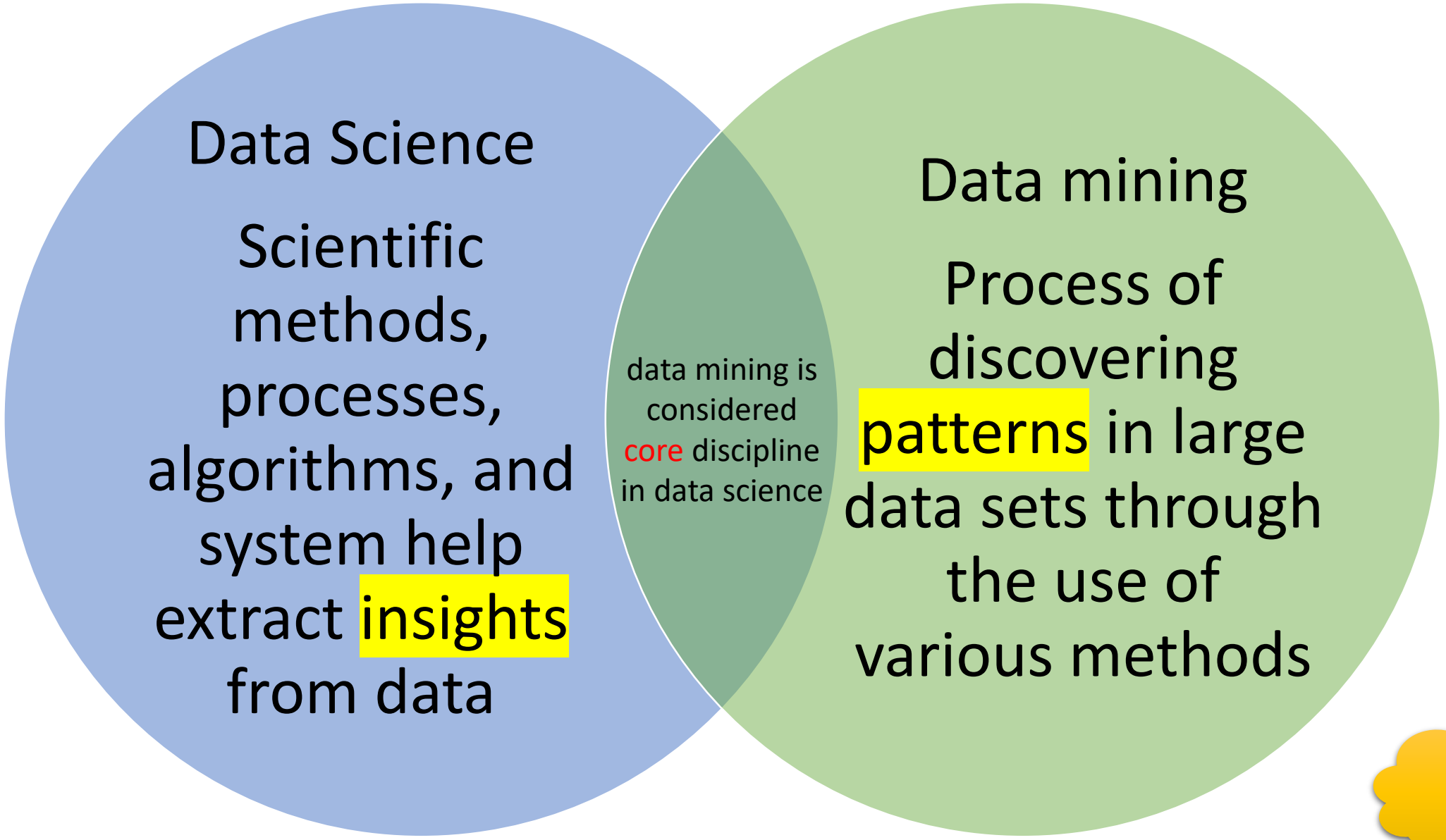
เหมืองข้อมูล



ขั้นตอนการทำเหมืองข้อมูล

ที่มา: Han, J. 2011

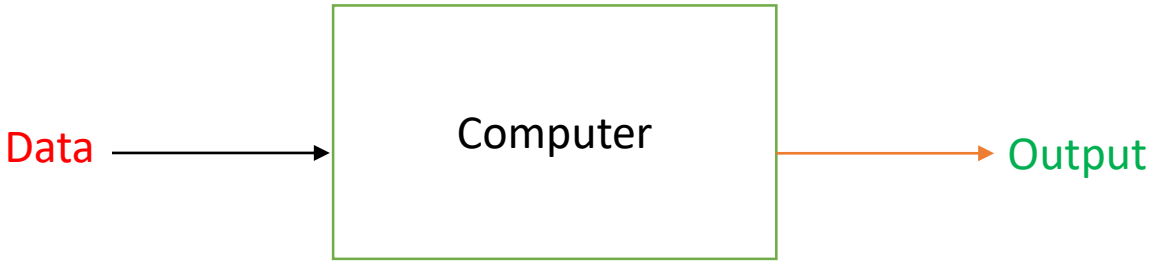
CS



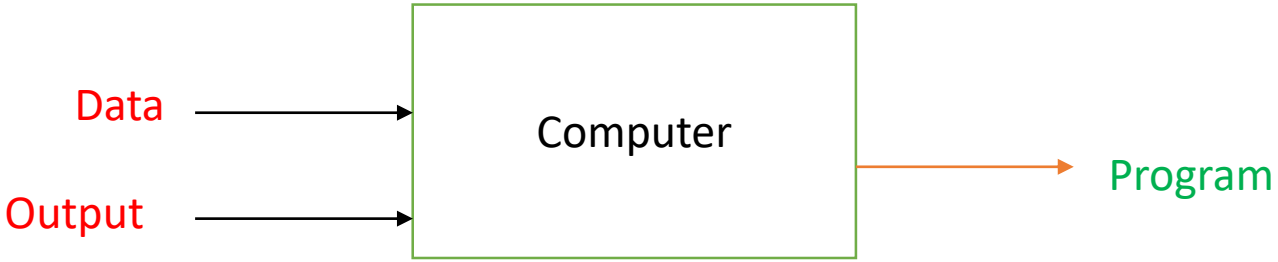
CS

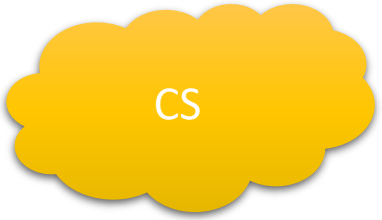
https://www.unite.ai/th/data-science-vs-data-mining-key-differences/

Traditional programming



Machine Learning

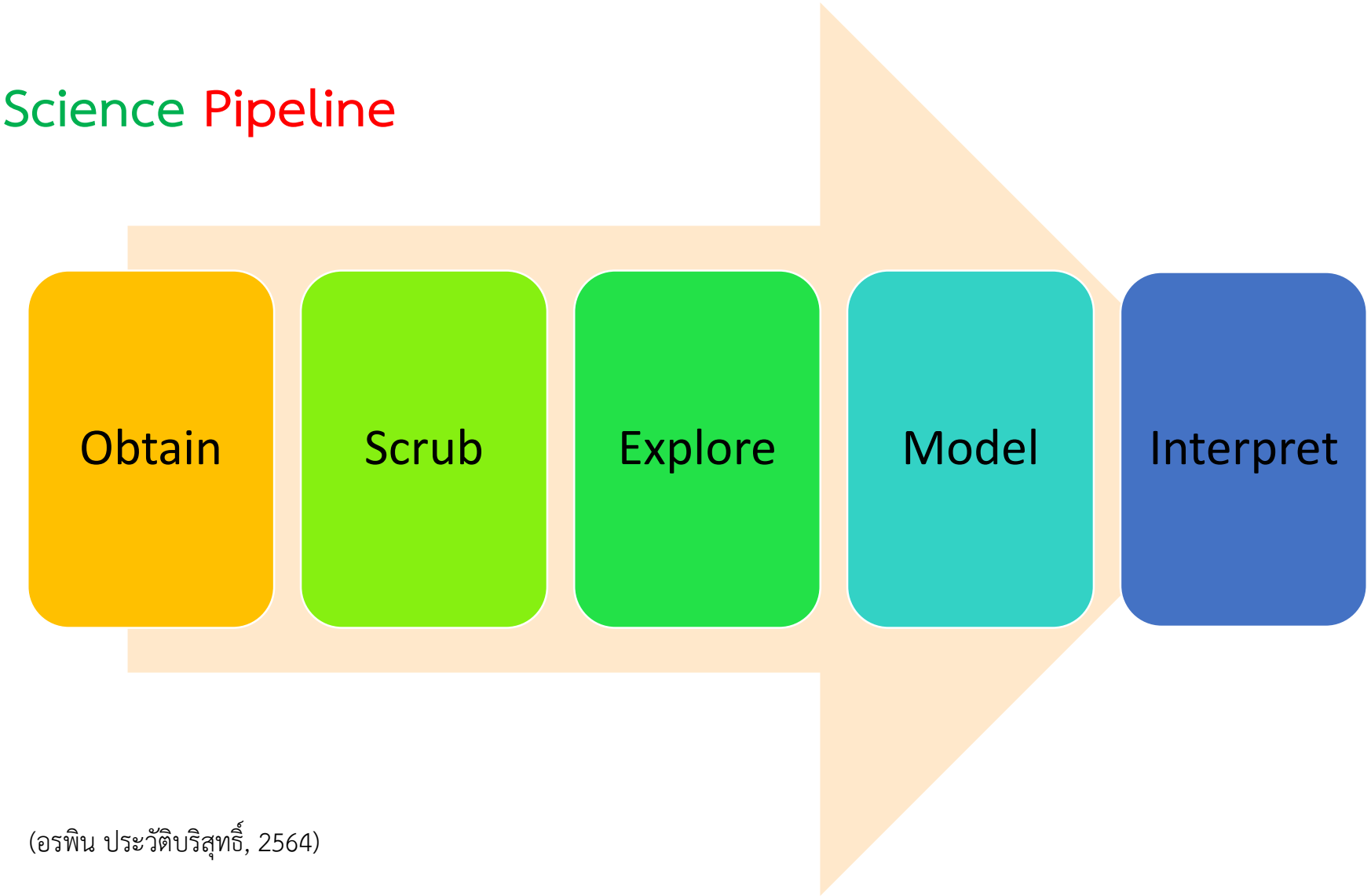




https://machinelearningmastery.com/basic-concepts-in-machine-learning/

ขั้นตอนการทำ Data Science

Data Science Pipeline



(อรพิน ประวัติบริสุทธิ์, 2564)

Data Science Tools

Programmers

python

R

sas

Non-Programmers

rapidminer

SQL

IBM SPSS

CS

การปรียบเทียบภาษา python และ R

python

- 1991
- Engineering
- Deep learning

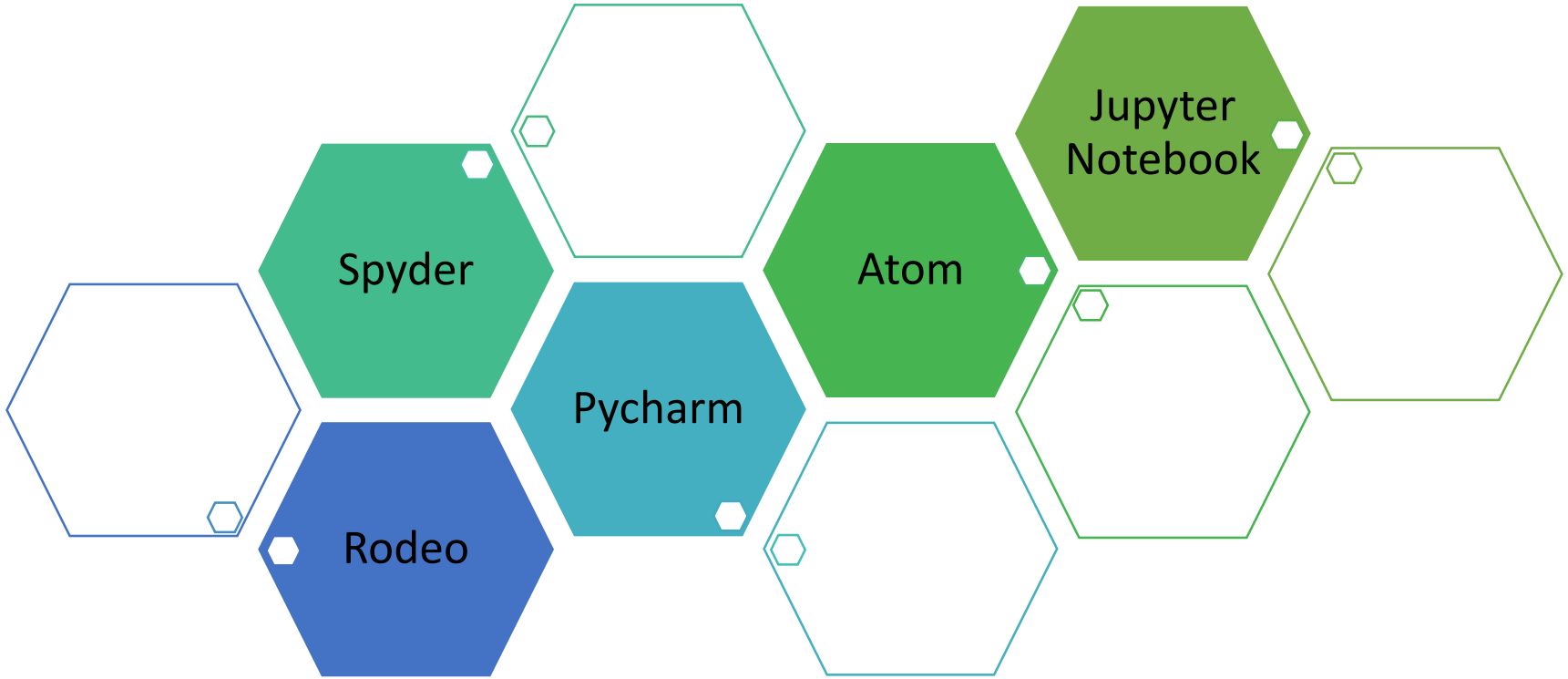
R

- 1995
- Statistics, research
- Data analysis
- Model building

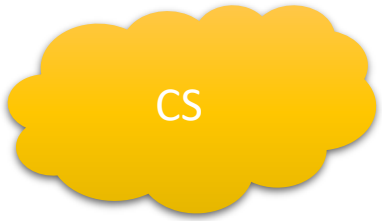
https://th.mort-sure.com/blog/difference-between-r-and-python/

CS

Python



https://www.datacamp.com/



Spyder VS Jupyter

CS

ความแตกต่างระหว่างภาษา python และภาษา Java

```python
print("Enter 'x' for exit.");
print("Enter marks obtained in 5 subjects: ");
mark1 = input();
if mark1 == 'x':
    exit();
else:
    mark1 = int(mark1);
    mark2 = int(input());
    mark3 = int(input());
    mark4 = int(input());
    mark5 = int(input());
    sum = mark1 + mark2 + mark3 + mark4 + mark5;
    average = sum/5;
    if(average>=91 and average<=100):
        print("Your Grade is A+");
    elif(average>=81 and average<=90):
        print("Your Grade is A");
    elif(average>=71 and average<=80):
        print("Your Grade is B+");
    elif(average>=61 and average<=70):
        print("Your Grade is B");
    elif(average>=51 and average<=60):
        print("Your Grade is C+");
    elif(average>=41 and average<=50):
        print("Your Grade is C");
    elif(average>=0 and average<=40):
        print("Your Grade is F");
    else:
        print("Strange Grade..!!");
```

https://codescracker.com/python/program/python-program-calculate-student-grade.htm

```java
public class CalculateStudentGrades {
   public static void main(String args[]) {
      Scanner sc = new Scanner(System.in);
      System.out.println("Enter average of your marks (less than 100)::");
      int average = sc.nextInt();
      char grade;

      if(average>=80){
         grade = 'A';
      }else if(average>=60 && average<80){
         grade = 'B';
      }
      else if(average>=40 && average<60){
         grade = 'C';
      }
      else {
         grade = 'D';
      }

      switch(grade) {
         case 'A' :
            System.out.println("Excellent!");
            break;
         case 'B' :
         case 'C' :
            System.out.println("Well done");
            break;
         case 'D' :
            System.out.println("You passed");
         case 'F' :
            System.out.println("Better try again");
            break;
         default :
         System.out.println("Invalid grade");
      }
      System.out.println("Your grade is " + grade);
   }
}
```

https://www.tutorialspoint.com/Java-program-to-calculate-student-grades

CS

Library สำหรับ Data Analysis

- NumPy สำหรับการจัดการข้อมูลในรูปแบบ array หลายมิติ
- SciPy ขยายมาจาก NumPy
- Matplotlib เป็นส่วนการแสดงผลในรูปแบบ visualization
- Jypyter เป็นเครื่องมือที่สร้างอยู่บน iPython แต่เป็น interative บน web browser ทำให้เราสามารถเขียน code, execute code และเอกสารต่าง ๆ อยู่ที่เดียวกัน

https://www.somkiat.cc/python-library-for-data-analysis/

CS

library for Python

- Pandas เป็น library สำหรับ Data Science เพื่อจัดการ data structure
- xarray เพื่อทำงานกับ array หลายมิติ
- scikit-learn เป็น library สำหรับ Machine Learning

https://docs.python.org/3/library/

CS

อ้างอิง

- http://www.zhanjunlang.com/resources/tutorial/Data%20Science%20from%20Scratch%20First%20Principles%20with%20Python.pdf
- Jesús Rogel-Salazar.DATA SCIENCE AND ANALYTICS WITH PYTHON. CRC Press Taylor & Francis Group, 2017.

CS