

LOGO

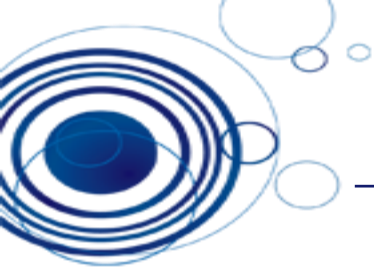
# BigData Engineering

1주차: Introduction

강의 : 신경섭

11101001110000111110101110010101010011001010011010111101001110000111110101110010101010011001010011010111101001110  
11100110000011011101001011101011111010101010010101111100110000011011101001011101011111010101010010101111100110000  
110100111001101010101001011011111010100110101001010111010011100110101010100101101111010100110101001010111010011100



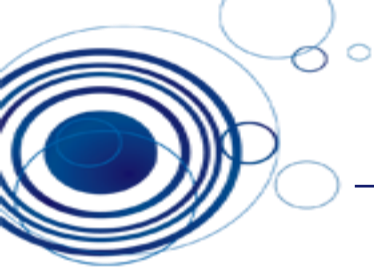


# Introduction

---

- **Objectives**
  - Introduction to data science
  - Python for data analysis
- 이론 2시간, 실습 1시간



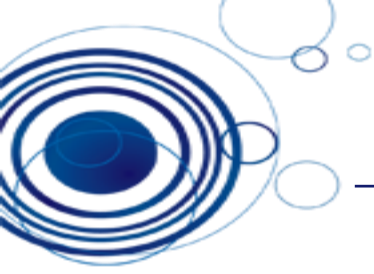


# Evaluation Strategy

---

- Midterm: 30%
- Final: 40%
- Project/homework: 20%
- Attendance: 10%



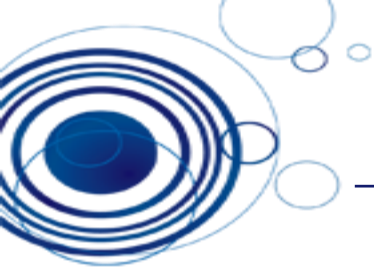


# Textbook

---

- Basically “handout”
- Refer “Python for Data Analysis”



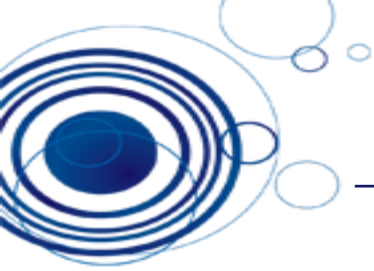


# Why Python?

---

- Easy to learn
- Full featured
  - Data acquisition, cleaning, databases, computing, and more
- Strong Data Science Libraries
  - SciPy Ecosystem

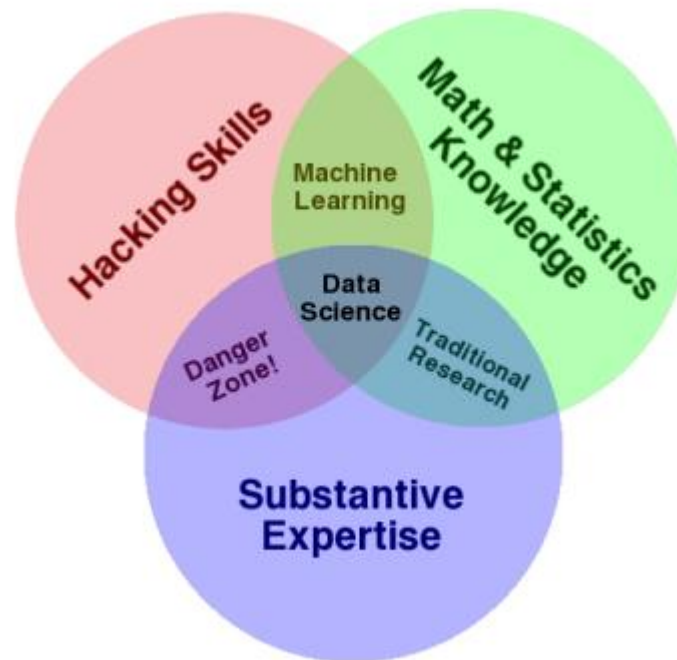


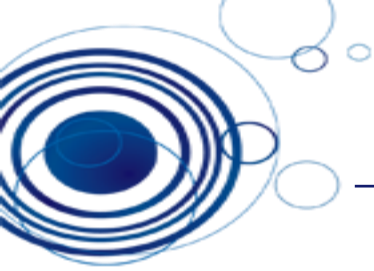


# Data Science

---

- Drew Conway's perspective





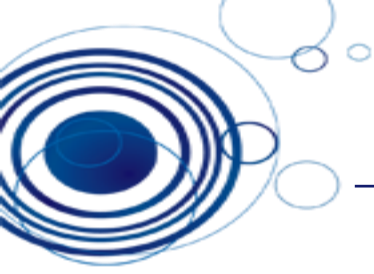
# Data Science

---

- David Donoho's perspective
  1. Data Exploration and Preparation
  2. Data Representation and Transformation
  3. Computing with Data
  4. Data Modeling
  5. Data Visualization and Presentation
  6. Science about Data Science

[ref] David Donoho, "50 years of data science"





# Python Functions

---

- Basic python functions
- Refer video

