

서울대학교 데이터사이언스대학원

2024년 여름방학 부트캠프 오리엔테이션

- 2024년 7월 18일 목요일
- 43-2동 B102호



SEOUL NATIONAL UNIVERSITY
GRADUATE SCHOOL OF DATASCIENCE

01

데이터사이언스대학원 소개



데이터사이언스란?

- 대용량 데이터의 명확하고 깊은 이해를 통해 통찰력(insight)을 갖추고 복잡한 문제를 해결하기 위한 학문 분야

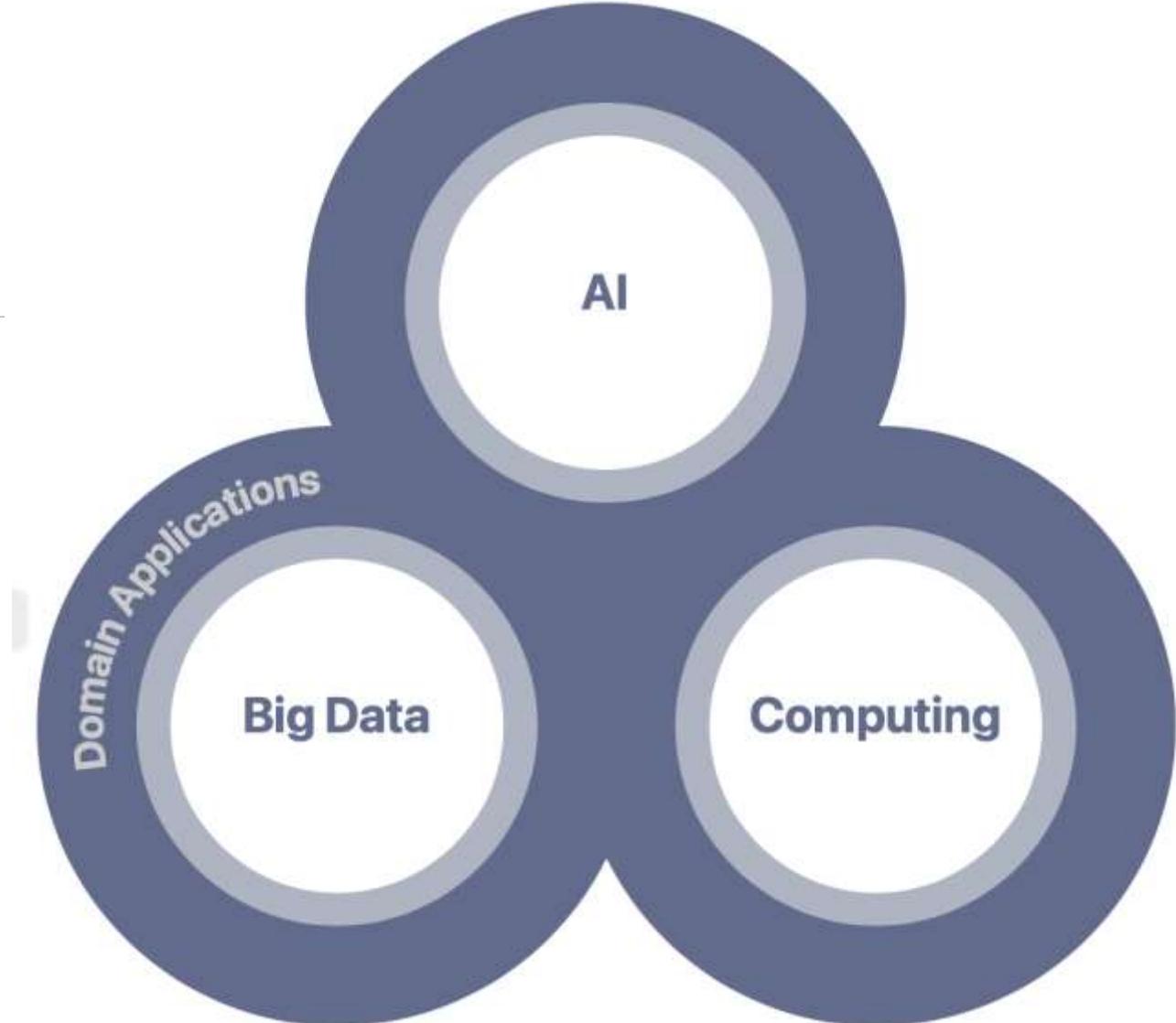
데이터사이언스란?

- 다학제적인 훈련과 여러가지 도구를 필요로 하여 기존의 여러 세부 학문 분야(ABC)가 융합되어 새롭게 탄생한 분야
 - **AI models & algorithms** : 대용량 데이터의 기각화, 분석, 모델링, 추론
 - **Big data technology** : 대용량 데이터의 획득, 저장, 정제, 관리, 보호
 - **Computing technology** : 자동화 및 효율적이고 빠른 처리
- ABC를 통한 핵심원리의 배양은, 문제 영역에 대한 지식과 핵심원리를 적용하여 주어진 문제를 해결하고 (Domain application) 가치를 창출하는 기반

데이터사이언스의 ABCD

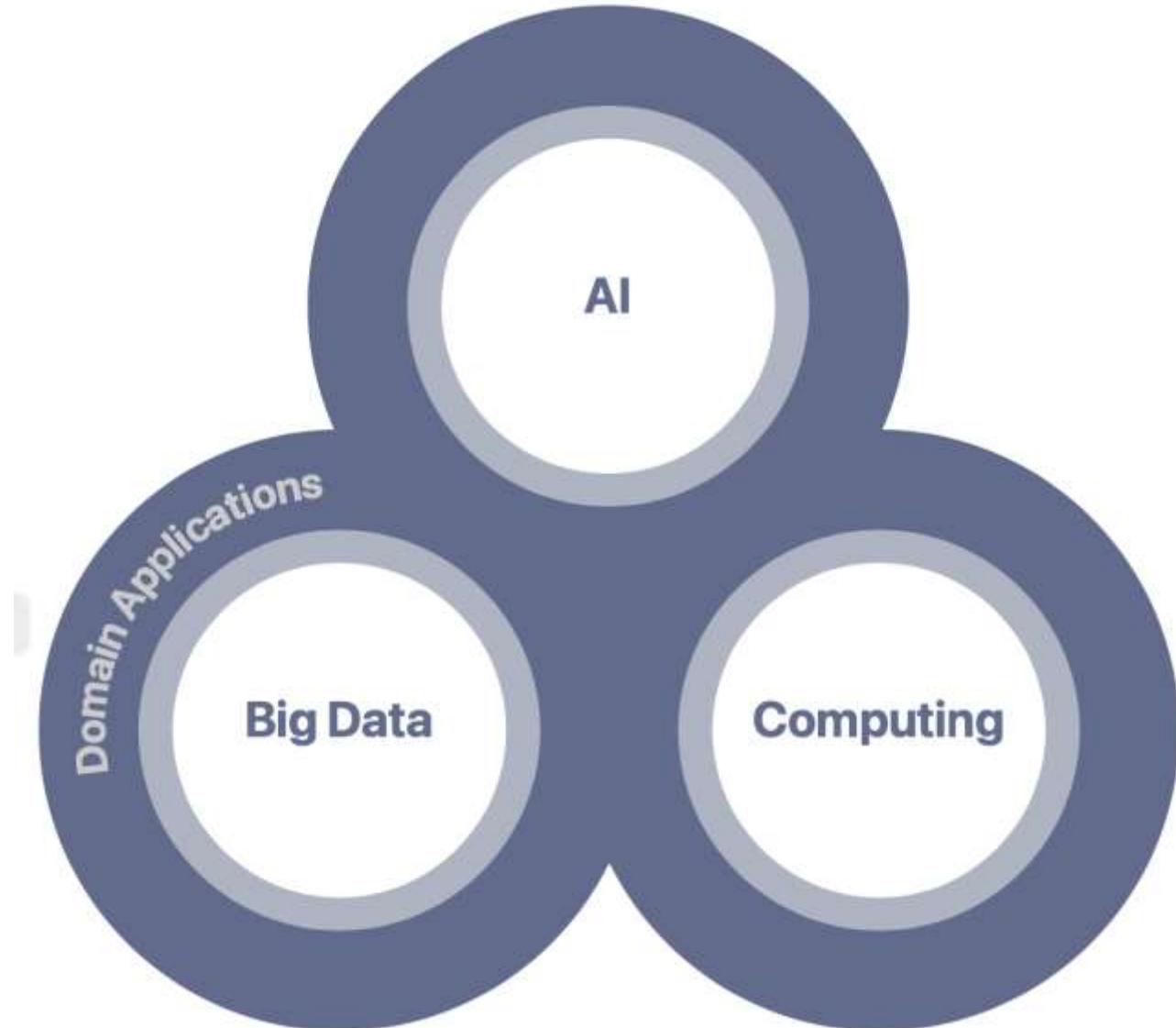
데이터사이언스

- 대용량 데이터의 명확하고 깊은 이해를 통해 통찰력(insight)을 갖고 복잡한 문제를 해결하기 위한 학문 분야
- AI, Big Data, Computing이 핵심이 되는 세 축
- 핵심원리를 적용 분야에 응용(Domain application)



AI 기술 발전을 견인하는 요소

- AI 모델 학습/추론 기술 (A)
- 대규모 학습데이터 (B)
- 컴퓨터시스템의 성능 향상 (C)
- AI 기술 적용 분야 응용 (D)



초거대 AI (Large Language Models)

- 끝에서 끝까지의 전체 기술 스택 필요
 - 학습 데이터 수집
 - 학습 데이터 처리
 - AI 응용
 - AI 모델 및 알고리즘
 - SW 및 HW 플랫폼
- Microsoft와 OpenAI의 전략적 협업에 의해 탄생
- 데이터사이언스대학원은 전체 기술 스택을 완비



초거대 AI 모델 및 플랫폼 최적화 센터

- 과학기술정보통신부 선도연구센터(ERC)
 - 2023년 6월 1일
- Center for optimizing Hyperscale AI Models and Platforms
- 데이터사이언스대학원의 교수진과 학생들이 주축



Seoul National University
Data Science Day

초거대 AI 모델 및 플랫폼 최적화 센터* 개소식
· 과학기술정보통신부 선도연구센터(ERC) | 센터장: 이재진 교수

“차세대 인공지능을 향한 대도약”

2023. 7. 20(목) 9:30 AM-11:50 AM
서울대학교 문화관 대강당

초청 연사
Prof. Andrew Ng
앤드류 응 교수

Founder, DeepLearning.AI
Founder & CEO, Landing AI
General Partner, AI Fund
Chairman & Co-Founder, Coursera
Adjunct Professor, Stanford University

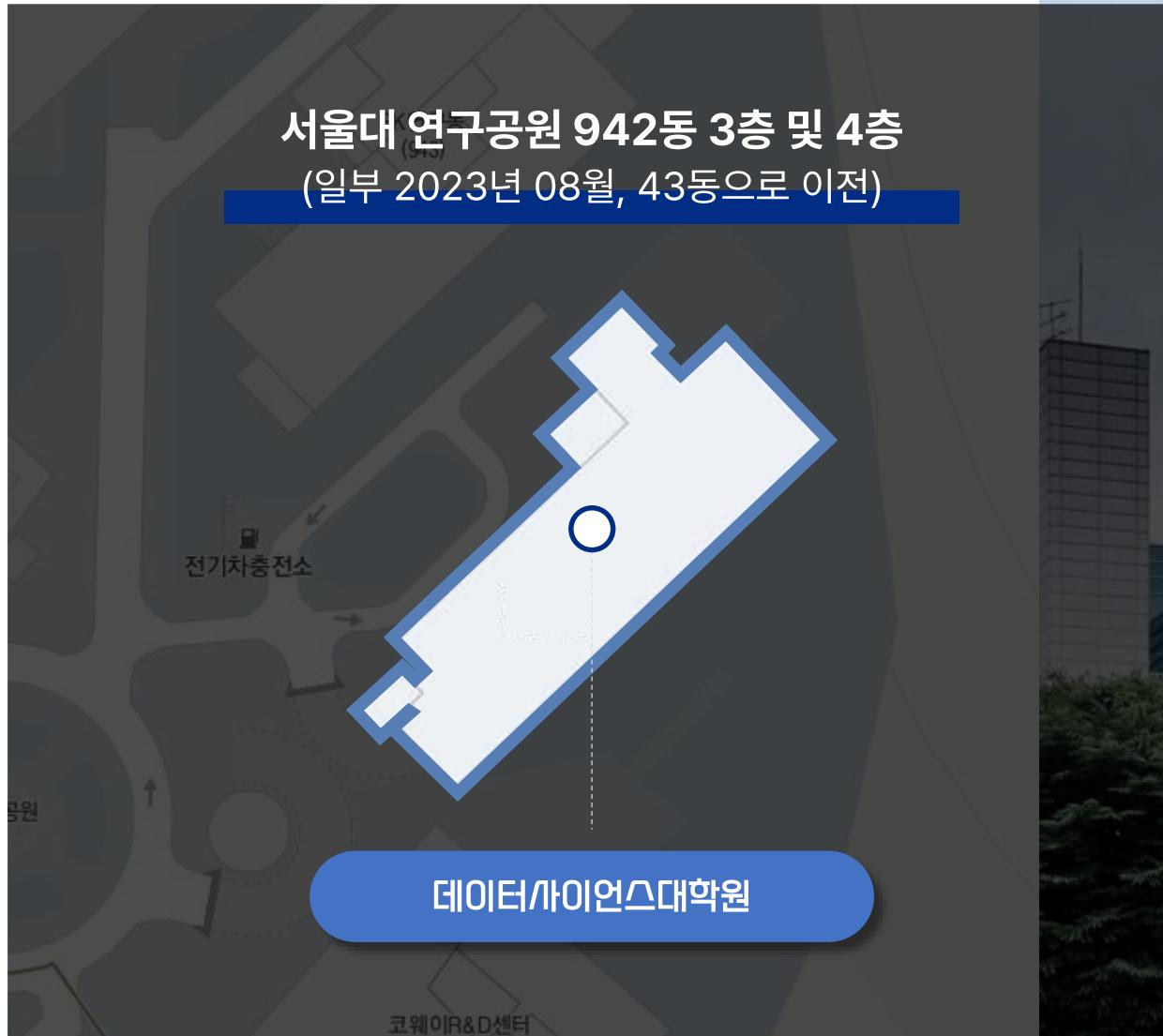
 Seoul National University
Graduate School of Data Science

 NRF 한국연구재단

 과학기술정보통신부
Ministry of Science and ICT

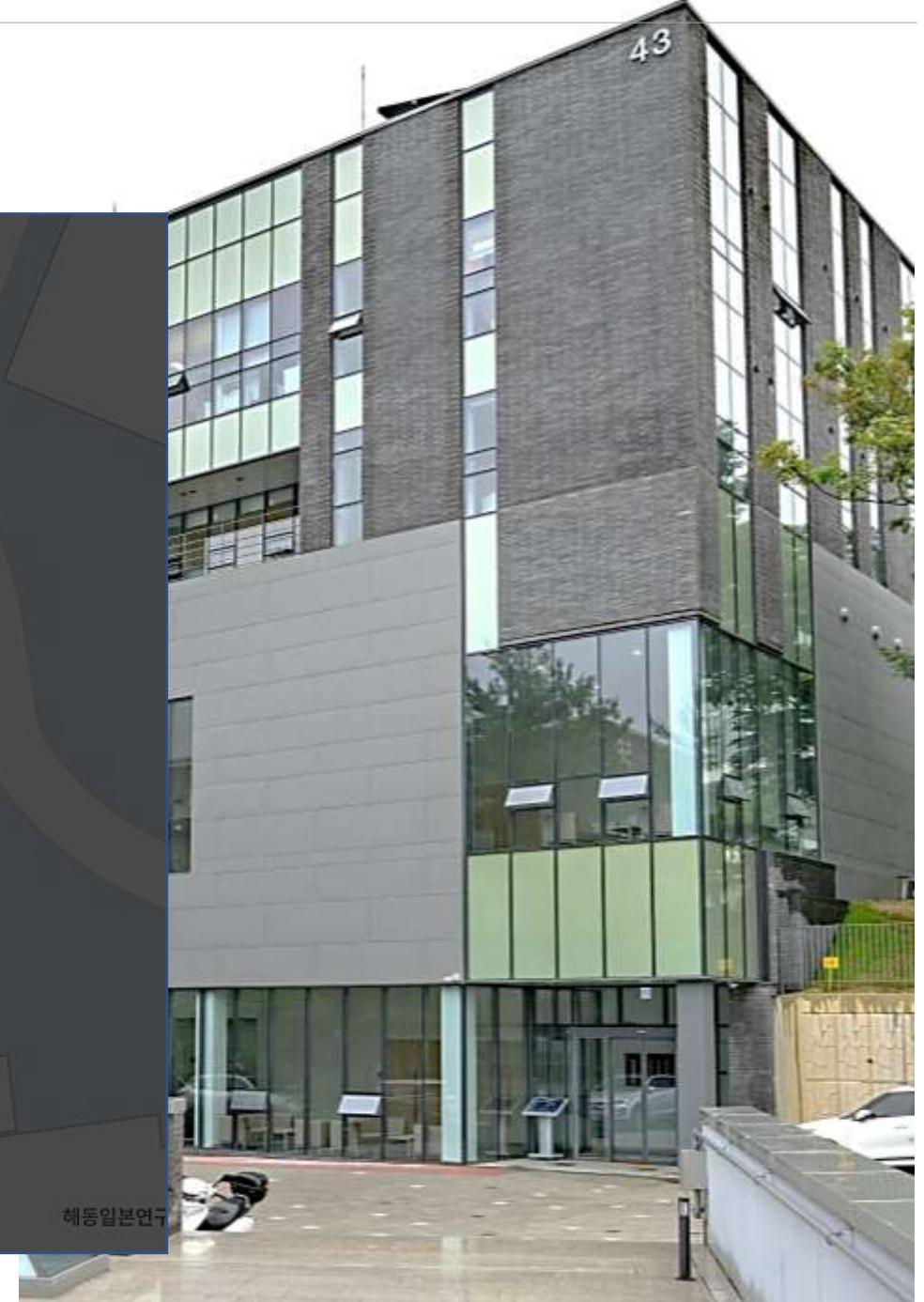


서울대 데이터사이언스 대학원 위치



데이터사이언스 대학원 위치 ②

서울대 연구공원 43동 4층,5층 / 43-2동 일부



인재상

역사·박사 과정 운영

세계를 선도하는
데이터사이언스
석사·박사 과정 운영

01

교육·연구 과정 개발

데이터사이언스
혁신인재 양성을 위한
교육·연구 과정 개발

02

교육·연구의 허브 역할

데이터사이언스
교육·연구의 허브 역할로
교육·연구 성과의 확산

03



주어진 어떤 분야의 문제도
데이터사이언스의 핵심 원리를
접목·활용하여 해결하고
데이터 기반 혁신을 이끌 수 있는

창의인재 양성





SEOUL NATIONAL UNIVERSITY
GRADUATE SCHOOL OF DATASCIENCE

02

부트캠프 소개



데이터사이언스 부트캠프

- ▶ 데이터사이언스 전공에 필요한 프로그래밍, 컴퓨팅, 수학 및 확률/통계에 대한 균형 잡힌 기초 강의 제공
- ▶ 강좌 목록

1 데이터사이언스를 위한 **Python** 프로그래밍과 데이터 구조

2 데이터사이언스를 위한 **시스템과 C 프로그래밍**

3 데이터사이언스를 위한 **수학**

4 데이터사이언스를 위한 **확률과 통계**

- ▶ Bootcamp 과목들은 4개의 강좌로 세분화되어 필요한 수업을 선택해 수강 가능
- ▶ 각 강좌당 3주 이내(주5일)로 오프라인 및 온라인 하이브리드 수업 진행

2024 여름방학 부트캠프 스케줄

			7/17 (수)	7/18 (목)	7/19 (금) 43-2동 B102호
10:00 ~ 12:00			공휴일	오리엔테이션	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 1회차
14:00 ~ 16:00					
	7/22 (월) 43-2동 B102호	7/23 (화) 43-2동 B102호	7/24 (수) 43-2동 B102호	7/25 (목) 43-2동 B102호	7/26 (금) 43-2동 B102호
10:00 ~ 12:00	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 2회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 3회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 4회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 5회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 6회차
14:00 ~ 16:00					
	7/29 (월) 43-2동 B102호	7/30 (화) 43-2동 B102호	7/31 (수) 43-2동 B102호	8/1 (목) 43-2동 B102호	8/2 (금) - 18동 101호
10:00 ~ 12:00	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 7회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 8회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 9회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) 10회차	데이터사이언스를 위한 Python 프로그래밍과 데이터 구조 (조요한) Exam
14:00 ~ 16:00	데이터사이언스를 위한 수학 (정형수) 1회차	데이터사이언스를 위한 수학 (정형수) 2회차	데이터사이언스를 위한 수학 (정형수) 3회차	데이터사이언스를 위한 수학 (정형수) 4회차	데이터사이언스를 위한 수학 (정형수) 5회차
	8/5 (월) - 18동 101호	8/6 (화) - 18동 101호	8/7 (수) - 18동 101호	8/8 (목) - 18동 101호	8/9 (금)
10:00 ~ 12:00	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 1회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 2회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 3회차	데이터사이언스를 위한 수학 (정형수) 9회차	
14:00 ~ 16:00	데이터사이언스를 위한 수학 (정형수) 6회차	데이터사이언스를 위한 수학 (정형수) 7회차	데이터사이언스를 위한 수학 (정형수) 8회차	데이터사이언스를 위한 수학 (정형수) 10회차	데이터사이언스를 위한 수학 (정형수) Exam
	8/12 (월) 43-2동 B102호	8/13 (화) 43-2동 B102호	8/14 (수) 43-2동 B102호	8/15 (목)	8/16 (금) 43-2동 B102호
10:00 ~ 12:00	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 4회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 5회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 6회차	공휴일	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 7회차
14:00 ~ 16:00					
	8/19 (월) 43-2동 B102호	8/20 (화) 43-2동 B102호	8/21 (수) 43-2동 B102호	8/22 (목) 43-2동 B102호	8/23 (금) 43-2동 B102호
10:00 ~ 12:00			데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 8회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 9회차	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) 10회차
14:00 ~ 16:00	데이터사이언스를 위한 확률과 통계 (오민환) 1회차	데이터사이언스를 위한 확률과 통계 (오민환) 2회차	데이터사이언스를 위한 확률과 통계 (오민환) 3회차	데이터사이언스를 위한 확률과 통계 (오민환) 4회차	데이터사이언스를 위한 확률과 통계 (오민환) 5회차
	8/26 (월) 43-2동 B102호	8/27 (화) 43-2동 B102호	8/28 (수) 43-2동 B102호	8/29 (목)	8/30 (금)
10:00 ~ 12:00	데이터사이언스를 위한 시스템과 C 프로그래밍 (김형신) Exam	데이터사이언스를 위한 확률과 통계 (오민환) 7회차	데이터사이언스를 위한 확률과 통계 (오민환) 9회차	데이터사이언스를 위한 확률과 통계 (오민환) Exam	수료식
14:00 ~ 16:00	데이터사이언스를 위한 확률과 통계 (오민환) 6회차	데이터사이언스를 위한 확률과 통계 (오민환) 8회차	데이터사이언스를 위한 확률과 통계 (오민환) 10회차		

교수진



정형수 교수

Data Lakehouse Systems
for Data Science



김형신 교수

Edge/Mobile AI,
Distributed Machine Learning,
Internet of Things



오민환 교수

Reinforcement learning,
Statistical machine learning



조요한 교수

Natural Language Processing,
Conversational AI

데이터사이언스 부트캠프 교과과정

1 데이터사이언스를 위한 Python 프로그래밍과 데이터 구조



- Abstraction, Hello Python, Memory Models, and Functions
Function Memory Model, Function Design, and Strings
- Control Structures, Modules, and Classes, Lists and Loops
- Sets, Tuples, Dictionaries, and Mutability, File I/O
- Object-Oriented Programming
- Computational Complexity, Searching, and Sorting
- MergeSort, Algorithm Design, Testing, and Debugging
- Data Structures: Arrays, Linked Lists, Stacks, and Queues
- Data Structures: Trees
- Data Structures: Graphs
- Data Structures: Hash Tables

데이터사이언스 부트캠프 교과과정

2 데이터사이언스를 위한 시스템과 C 프로그래밍



강의 by
김형신 교수

- Bits, Data Types, Operations, Semi-conductor & Logic Gates
- Von Neumann Model and Machine codes, Great Ideas in Computer Architecture
- Hello C, Variables and Operators
- Control structures
- Functions
- Pointers
- Arrays
- I/O
- Structures
- Linked lists

데이터사이언스 부트캠프 교과과정

3

데이터사이언스를 위한 수학



- Linear Algebra - Matrices, Basic operations
- Linear Algebra - Basis, Rank, Linear Mappings
- Linear Algebra - Norms, Inner Products & Orthogonality
- Linear Algebra - Projections & Gram-Schmidt Orthogonalization
- Linear Algebra - Eigenvalues, Eigenvectors, Eigenspaces, Diagonalization
- Linear Algebra - Singular Value Decomposition & Matrix Approximation
- Vector Calculus - Differentiation, Partial Differentiation & Gradients
- Vector Calculus - Gradients of Vector-Valued Functions
- Vector Calculus - Backpropagation & Automatic Differentiation

데이터사이언스 부트캠프 교과과정

4

데이터사이언스를 위한 확률과 통계



- Probability
- Random Variable
- Expectation & Variance
- Convergence
- Statistical Inference
- Probability Density Function & Cumulative Distribution Function
- Bootstrap
- Parametric Inference
- Hypothesis Testing and p-value
- Bayesian Inference

데이터사이언스 부트캠프 이수 기준

- **일반 수료**: 평가 항목(과제, 퀴즈) 합산점수로 결정
 - **우등 수료**: 평가 항목(과제, 퀴즈, **시험**) 합산점수로 결정
-
- ✓ “데이터사이언스를 위한 **Python 프로그래밍과 데이터 구조**” 및 “데이터사이언스를 위한 **시스템과 C 프로그래밍**” 강좌 모두 **우등 수료**인 경우: “컴퓨팅의 기초” 통과로 간주
- 상응과목: 데이터사이언스를 위한 컴퓨팅의 기초
- ✓ “데이터사이언스를 위한 **수학**” 및 “데이터사이언스를 위한 **확률과 통계**” 강좌 모두 **우등 수료**인 경우, 서울대학교 데이터사이언스대학원 입학 시 “수학과 통계의 기초” 통과로 간주
- 상응과목: 데이터사이언스를 위한 수학과 통계의 기초

데이터사이언스 부트캠프 FAQ

Q: 타 대학 또는 단과대학(학과) 학생인데 참여 가능한가요?

A: 가능합니다.

Q: 매 수업 대면 출석을 해야 하는가요?

A: 출석 확인을 하지 않으며, 수료요건에 없습니다.

Q: 하이브리드 수업이면 모든 회차의 수업에 온라인으로 참여 가능한가요?

A: 가능합니다.

Q: 시험은 대면인가요? 온라인인가요?

A: 온라인으로 진행 예정입니다.

Q: 부트캠프 수료가 데이터사이언스대학원 입시에 영향을 주는가요? 부트캠프 수료를 못하면 입시에 불이익이 있나요?

A: 아닙니다. 데이터사이언스대학원 입시는 부트캠프 수료 여부 상관없이 지원 가능합니다.



Seoul National University
Graduate School of Data Science

질의응답

