# 문제는 통계야
# - 빅데이터 시대의 데이터 문해력

강의계획서, 2026 봄학기

Course Home Page: eTL

## 강의목표

21대 국회의원선거에서는 정말 선거부정이 있었던 것일까? 여론조사 결과는 왜 조사기관별로 차이가 많이 나는가? 코로나-19 백신은 정말로 위험한가? 연령대별로 접종여부에 관한 기준은 어떻게 제시되는가? 이 수업을 통해서 이 모든 것에 대한 대답을 찾을 수 있다. 이 과목을 통해서 수강생들은 데이터 리터러시에 위한 기본지식을 함양하고 데이터를 통한 의사결정능력과 데이터를 전달하는 미디어를 올바르게 이해하는 능력을 배양하게 될 것이다.

- 정보의 홍수시대에 올바르고 필요한 정보를 취사선택하는 방법을 대해서 배운다.
- 데이터 기반 주장의 타당성 여부를 검토할 수 있고 데이터를 통한 의사 결정을 하는 법을 배운다.
- 여론조사를 비롯한 각종 미디어의 데이터 관련기사의 이해와 문제점에 대해서 알아본다.
- 상관관계와 인과추론의 차이점에 대해서 알아본다
- 데이터를 활용하여 본인의 주장을 올바르게 전달할 수 있는 능력을 배양한다

## 연락처/강의 시간

**담당교수:**

장원철 25동 323호 wcjang@snu.ac.kr

**조교:**

TBA TBA TBA

**강의** 월수 14:00-15:15 28동 103호

**면담시간**

| | | |
|---|---|---|
| 장원철 | TBA | 25동 323호 |
| TBA | TBA | TBA |

## 추천도서

- 닉 폴슨, 제임스 스콧 지음, 노태복 번역 (2020). 수학의 쓸모: 불확실한 미래에서 보통 사람들도 답을 얻는 방법. 더퀘스트. ISBN 9791165210991
- 데이비드 스피겔헬터 지음, 양병찬 번역 (2025). 불확실성에 맞서는 기술. 생각의 힘. ISBN 9791193166994
- 레일라 슈넵스, 코랄리 콜메즈 지음, 김일선 번역 (2020) 법정에 선 수학 - 수학이 판결을 뒤바꾼 세기의 재판 10. 아날로그. ISBN 9791187147589
- 칼 벅스트롬, 제빈 웨스트 지음, 박선령 번역 (2021). 똑똑하게 생존하기: 거짓과 기만 속에서 살아가는 현대인을 위한 헛소리 까발리기의 기술. 안드로메디안. ISBN 9791135497902
- 톰 치버스 지음, 홍한결 번역 (2025) 모든 것은 예측 가능하다: 진단검사에서 뇌의 작동원리까지, 세상을 설명하는 베이즈 정리의 놀라운 힘. 김영사. ISBN 979117332014903410
- 김용대 (2021). 데이터 과학자의 사고법: 더 나은 선택을 위한 통계학적 통찰의 힘. 김영사. ISBN 9788934990475
- 정성규 (2022). 수학보다 데이터 문해력. EBS Books. ISBN 9788954799539

## 평가방법

수업평가는 수업참여 및 출석 (30%)와 기말고사 (30%) 그리고 4번의 과제 (40%)를 통해 이루어진다. **총 4회의 과제는 다음과 같은 여론조사 분석, 인과추론 사례, 데이터 오남용 탐색, 공공데이터 해석 등으로 구성되며 주제당 2-3페이지 분량의 보고서를 제출하여야 한다.**

1. **여론 조사의 함정: 실제 언론 보도 중 여론조사 기사를 하나 선택하여, 표본 설계 · 질문 문항 · 보도 문체를 비판적으로 분석하라**

2. **인과추론과 데이터 편향: "흡연과 폐암", "부모의 학력과 자녀 성적" 등 사례 중 하나를 선택해 심슨의 역설 · 콜라이더 편향 개념을 적용하여 설명하라.**

3. **데이터 오남용 사례: SNS나 뉴스 기사에서 발견한 통계적 오류(잘못된 그래프, 상관/인과 혼동 등)를 수집 · 분석하라.**

4. **데이터 기반 의사결정과 전달 : 공공데이터(Open Data) 중 관심 주제를 선택하여, 데이터 기반 주장과 그 한계를 설명하라.**

과제는 마감일 자정까지 eTL에 통해서 제출해야한다. 늦게 제출한 숙제는 어떤사유 (천재지변포함)라도 받지않는다. 단 집안사정/건강상의 이유에 한하여 미제출한 숙제의 점수를 본인의 나머지 숙제점수의 평균으로 대체 가능하다. 학기말 정당한 사유없이 학점을 올려달라고 부탁할 경우 수업참여도 점수를 0점으로 처리하여 불이익을 받을 수 있다. 마찬가지로 재수강을 이유로 학점변경을 요청할 경우 청탁금지법 위반으로 허용하지 않는다.

## 숙제 협력에 관한 규칙

수업을 듣는 동료 학생들과 같이 숙제를 하는 것은 장려한다. 하지만 답안지는 반드시 본인 스스로 작성해야 되며 도움을 받을시에는 도움을 준 학생의 이름을 숙제제출시 명시해야한다. 그렇지 않을 경우 같은 답안을 제출한 숙제는 모두 0점 처리한다.

## 생성형 AI 도구 활용 방침

본 강의에서는 생성형 AI를 학습 보조 도구로 활용할 수 있습니다. 과제 수행에 필요한 아이디어 및 주제 탐색, 자료분석 등의 목적으로 생성형 AI를 사용할 수 있습니다. 다만, 생성형 AI를 과제 수행에 사용한 경우, 사용도구명, 사용 과정 및 생성 내용을 과제에 명시해야 합니다.

## 장애학생에 대한 지원서비스 안내

<table>
<tr><td rowspan="2">장애유형</td><td colspan="2">지원서비스</td></tr>
<tr><td>강의수강관련</td><td>과제 및 평가관련</td></tr>
<tr><td>시각장애</td><td>교재재작 (디지털교재, 점자교재, 확대교재등)<br>대필도우미 허용</td><td rowspan="5">과제제출기한 연장<br>과제제출방식 및 응답방식의 조정<br>평가시간 연장<br>평가 문항제시 및 응답방식의 조정<br>별도 고사실 제공</td></tr>
<tr><td>지체장애</td><td>교재 제작 (디지털교재)<br>대필도우미 및 수업보조도우미 허용</td></tr>
<tr><td>청각장애</td><td>대필및 문자통역 도우미 활동허용<br>강의 녹취허용</td></tr>
<tr><td>건강장애</td><td>질병등으로 인한 결석에 대한 출석 인정<br>대필도우미 허용</td></tr>
<tr><td>학습장애</td><td>대필도우미 허용</td></tr>
<tr><td>지적장애<br>자폐성장애</td><td>대필도우미 및 수업멘토 허용</td><td>개별화 과제 제출및 대체평가 실시</td></tr>
</table>

본 강의를 수강하는 장애학생들에게는 이상의 지원 서비스 이외에도 장애학생 개개인의 특성과 요구에 따라, 지도교수 및 장애학생지원센터와의 상담을 통하여 적절한 수준의 지원 서비스를 제공합니다. 장애학생에 대한 지원서비스와 관련하여 문의사항이 있는 학생들은 담당교수 혹은 장애학생지원센터(02-880-8787)로 문의바랍니다.

# 강의계획

| Date | Topic | Reading | 비고 |
|---|---|---|---|
| 3/2 | 대체공휴일 | | |
| 3/4 | 강의소개 | 데이터 문해력이란? | |
| 3/9 | 여론조사 | 들쑥날쑥 여론조사 정말 믿어야 하나? | |
| 3/11 | 선거부정논란 | 김어준의 K값과 21대 국회의원 부정선건 논란 | |
| 3/16 | 통계조작논란 | 통계청 가계동향조사와 KBS 세대인식 집중조사 사례 연구 | |
| 3/18 | 인과그래프로 알아보는 통계편향 | Collider 편향과 심슨의 역설 | |
| 3/23 | 다양한 통계편향 | 그 집 자제분은 서울대 어떻게 들어갔어요? | |
| 3/25 | 회귀분석과 예측 | 부모의 키로 자녀의 키를 예측할 수 있을까 | 과제 1 |
| 3/30 | 추천시스템 | 오늘 밤에는 무슨 영화를 볼까? | |
| 4/1 | 랜덤이란? | 카사노바의 프랑스 로또는 공정했을까? | **실험 1** |
| 4/6 | 조건부확률 | 몬티홀 문제와 죄수의 딜레마 | |
| 4/8 | 인공지능과 베이즈 법칙 | 베이즈법칙은 인공지능을 어떻게 바꾸었는가? | |
| 4/13 | 백신과 임상시험 | 임상시험은 어떻게 진행되어야 하나? | |
| 4/15 | 관측연구와 인과관계 | 흡연이 폐암을 유발한다는 것을 어떻게 밝혀냈을까? | |
| 4/17 | 연구윤리 | 멘델의 데이터 조작과 브라이언 완싱크의 p값 해킹 | 과제 2 |
| 4/22 | 개인정보보호와 재현가능연구 | 제시카 알바는 왜 택시 팁을 주지 않았을까? | |
| 4/24 | 취합검사와 다중비교 | 코로나-19에 실제로 감염된 사람은 몇명일까? | |
| 4/29 | 통계와 법 | 헤롤드 시프먼의 연쇄살인은 막을 수 있었을까? | **실험 2** |
| 5/4 | 디지털 인문학 | 양식측정학과 비틀즈의 in My Life | |
| 5/6 | 네트워크 | 예송논쟁과 공동법안발의 | |
| 5/11 | 머니볼 | 메이저리그 구단은 어떻게 데이터 혁명을 이끌어 냈을까? | |
| 5/13 | 데이터 시각화 | 뉴욕타임스와 파이낸셜 타임스의 시각화 사례연구 | 과제 3 |
| 5/18 | 빅데이터와 통계 | Google Flu Trend로 알아보는 빅데이터의 흥망성쇠 | **실험 3** |
| 5/20 | 데이터 저널리즘 | 데이터 저널리즘은 단독(아무도 관심없음)과 속보(확인하지 않음)의 경쟁을 막을 수 있을까? | |
| 5/25 | 누가누가 잘하나? | 바둑기사의 랭킹은 어떻게 정해지는 걸까? | |
| 5/27 | 대체공휴일 | | |
| 6/1 | 웹서치 | 구글은 어떻게 웹페이지 순위를 정할까? | |
| 6/3 | 튜링과 에니그마 | 튜링은 어떻게 독일군의 암호를 해독한 것일까? | **실험 4** |
| 6/8 | 지방선거 | | |
| 6/10 | 정규분포와 포아송분포 | 통계학에서 많이 사용되는 분포들의 역사와 배경에 대해 알아보자 | 과제 4 |
| 6/15 | 기말고사 | | |

## 수업시간에 준수해야 할 사항

- 강의실 첫 6번째줄까지 착선한 사람들에 한해서 전자기기 사용이 허용된다. 6번째 줄 이후에 착석할 경우 수업시간에 어떤 전자기기 사용도 허용되지 않는다.
- 수업시간에 다른 사람과 대화 등으로 수업을 방해하지 않는다.
- 수업에 늦을 경우 다른 학생들에게 방해되지 않게 조용하게 들어와 앉는다.
- 전자기기(휴대폰, 아이패드, 컴퓨터 등)는 강의노트를 보는 용도 이외 수업시간 사용을 금한다.
- 특별한 이유없이 결석을 잦을 경우 수업참여 점수가 0점처리된다.