

## 1. 상관관계 분석

- 주제 : 서울과 부산의 전기차 보급률 증가와 미세먼지(PM2.5) 농도 변화 간의 연도별 관계 분석
- 가설 : 서울과 부산 모두에서 전기차 보급률이 높아질수록 초미세먼지(PM2.5) 농도가 낮아질 것이다. 특히, 전기차 보급 정책이 더 적극적으로 시행된 서울에서 감소 폭이 더 클 것이다.
- 사용할 데이터
  - 산업통상자원부 : 전기차 등록 현황 (연도별·시도별) (2018~2024년, 시도별 전기차 등록 대수, 충전소 수 포함)  
<https://www.data.go.kr/data/15083033/fileData.do>
  - 환경공단 AirKorea : 대기환경 연평균 통계 (PM2.5) (서울·부산의 연도별 초미세먼지 농도)  
[https://www.airkorea.or.kr/web/last\\_amb\\_hour\\_data?pMENU\\_NO=123](https://www.airkorea.or.kr/web/last_amb_hour_data?pMENU_NO=123)
  - 서울시 전기차 보급 정책 자료 (보조금, 충전 인프라 확충)  
<https://data.seoul.go.kr/dataList/OA-15500/S/1/datasetView.do>
  - 부산광역시 친환경차 보급 현황 및 지원정책 자료  
<https://www.busan.go.kr/ghgreduction/evcar01>

## 2. 확률분포 분석

- 주제 : 항공편 지연시간의 분포 분석
- 가설 : 항공편 출발 지연시간은 정규분포가 아닌 오른쪽 꼬리가 긴(Log-normal) 형태의 비대칭 분포를 따를 것이다. 또한 짧은 지연이 다수를 차지하고, 장시간 지연은 드물게 발생할 것이다.
- 사용할 데이터
  - 국토교통부 항공정보포털시스템 (항공편 운항정보 통계) (국내 주요 공항별 항공편 지연·취소 현황, 출발 지연시간 등) <https://www.airportal.go.kr/knowledge/statsnew/airline.jsp>
  - DAICON 항공 지연 예측 AI 경진대회 데이터셋 (Open Dataset) (항공편별 출발 예정/실제 시각, 지연시간, 기상 변수 포함) <https://dacon.io/competitions/open/235539/data>

### 3. 예측 모형 모델링 (회귀분석 / 의사결정나무)

- 주제 : GTX 창릉역 개통 시 고양 창릉 신도시 LH 분양 아파트의 민간 매매가격 예측
- 가설 : 전용면적, 층수, 역세권 거리 등 물리적 요인은 아파트의 민간 매매가격에 유의미한 영향을 미친다.
- 사용할 데이터
  - LH 고양창릉 분청약 보도자료 (A4, S5, S6 블록 분양가) (2025년 1월 기준 LH 공식 보도자료 및 분양공고문)  
<https://www.lh.or.kr/lhinfo/contents.do?menuId=1905120401>
  - 국토교통부 실거래가 공개시스템 (인근 덕양구 실거래 내역)  
<https://rt.molit.go.kr/>
  - 지도 기반 GTX 창릉역 거리 데이터 (공공지도 API 활용) (역세권 거리 0.5~2.0km 가정 데이터 산출용)  
<https://map.kakao.com/>