

# 이력서

## 프로필



**임수인**

**SOOIN LIM**



**연락처**

**010-7158-1555**



**이메일**

**juniper0306@naver.com**



**GitHub**

**<https://github.com/tndls-dl>**

**Portfolio**

**<https://tndls-dl.github.io>**

## 한줄 소개

끊임없는 학습과 프로젝트 경험을 통해 데이터를 단순한 수치가 아닌 문제 해결의 무기로 전환하는 데이터 분석가입니다.

Python·SQL 기반 데이터 처리 역량과 꾸준한 기록 습관을 바탕으로, 복잡한 데이터 속에서 실행 가능한 인사이트를 도출하는 것을 목표로 하고 있습니다.

## 학위

**세종대학교 전자정보통신공학 전공**

**학사 (2020.03 ~ 2024.02)**

**수강 과목:** 고급프로그래밍입문-C, 프로그래밍, 데이터분석과도표작성 등

**학점:** 3.53/4.5

## 역량/스킬 [Skill]

---

### 프로그래밍

- Python
- SQL
- Git

### 데이터 분석

- Pandas/NumPy
- 통계 분석

### 머신러닝 & AI

- Scikit-learn, Classification, Regression
- API (FastAPI, OpenAI API)
- Agent Framework

### 시각화

- Matplotlib
- Seaborn

### 프로젝트 / 협업

- Slack
- Notion
- GitHub

## 어학 / 자격

---

### TOEIC Speaking

- 등급: Intermediate High (150점)
- 기관: YBM 어학시험
- 취득일: 2024.11

### TOEIC

- 점수: 885점
  - 기관: YBM 어학시험
  - 취득일: 2023.11
- 

### PCCE Lv.2

- 기관: 프로그래머스
- 취득일: 2025.10

### 컴퓨터 활용능력 2급

- 기관: 대한상공회의소
- 취득일: 2025.09

### SQLD (SQL 개발자)

- 기관: 한국데이터산업진흥원
- 취득일: 2025.09

### 6-시그마 (GB)

- 기관: 한국자격검정평가진흥원
- 취득일: 2025.04

### ADsP (데이터분석 준전문가)

- 기관: 한국데이터산업진흥원
- 취득일: 2025.03

### 자동차운전면허 2종 보통

- 기관: 경상남도경찰청
  - 취득일: 2021.08
-

## 교육이수사항

### AI 기반 데이터 분석가 양성과정

- 기간: 2025.06 ~ 2025.09
- 기관: 청년취업사관학교 SeSAC 서초캠퍼스
- 주요 학습: Python, SQL, Pandas/Numpy, EDA, 시각화(Matplotlib/Seaborn), 머신러닝 기초
- 프로젝트:
  - Ingrevia 챗봇 (화장품 전성분 기반 맞춤형 추천 에이전트 개발)
  - 영화 데이터 분석 (2005~2019, 흥행 요인·시각화·보고서 작성)

### 엑셀로 배우는 반도체 데이터분석

- 기간: 2024.10 ~ 2024.11
- 기관: 윈스펙
- 주요 학습: Cp/Cpk, 가설검정, 분산분석, 요인배치법, Excel 데이터 분석 실습

## 학습 기록

### GitHub TIL (Today I Learned)

- 꾸준한 학습 습관과 기술 성장 과정 정리
- 데이터 분석, Python, SQL, 머신러닝 등 일일 학습 내용 기록
- 링크: <https://github.com/tndls-dl/TIL>

## 프로젝트

### Ingrevia - 화장품 성분 기반 맞춤형 추천 챗봇

프로젝트 기간: 2025.09.15 - 2025.09.29

#### 프로젝트 배경 및 목표

- 리뷰 평점이 아닌 제품 전성분 분석 기반 맞춤형 스킨케어 제품 추천 시스템 개발
- INCI 표준명 기반 성분 통일 및 효능·유해성 분석을 통한 신뢰성 있는 추천 제공

#### 활용 데이터

- 화해 앱, 올리브영, 브랜드 공식 홈페이지에서 스킨케어 제품 약 120개 수집
- 한국화장품협회 성분 사전을 활용한 INCI 표준명 매핑
- EWG 등급 데이터 연동을 통한 유해성 점수 산출

#### 주요 수행 내용

- 전성분 데이터 수집 및 INCI 표준명 기반 성분 통일 작업
- 효능 성분·주의 성분 매핑 및 유해성 점수 분석 로직 구현
- LangChain 및 LangGraph 기반 대화형 추천 챗봇 설계
- LangSmith를 통한 대화 흐름 추적 및 노드별 검증
- Streamlit을 활용한 직관적인 웹 인터페이스 구현

#### 분석 및 구현 결과

- 피부 타입·고민·제품 카테고리 기반 개인화 추천 시스템 완성
- LLM을 활용한 성분 분석 및 추천 이유 자동 생성 기능 구현
- 후속 질문, 다중 고민 선택, 기초라인 추천 등 다양한 시나리오 지원
- 성분 배지 클릭 시 구글 검색 연동으로 정보 확장성 확보

#### 사용 기술 및 역량

- Python, LangChain, LangGraph, OpenAI GPT-4o
- Streamlit, Pandas, LangSmith
- 데이터 수집·전처리, API 연동, 대화형 AI 설계, 웹 애플리케이션 개발

# 한국 영화 데이터 분석 (2005-2019)

프로젝트 기간: 2025.08.18 - 2025.08.29

## 프로젝트 배경 및 목표

- 15년간 한국 영화 데이터(약 900편)를 분석해 흥행 요인을 규명하고 인사이트 도출

## 활용 데이터

- KOFIC API(관객·수익), 네이버 평점, 수동 입력 데이터
- 장르·감독·배우 메타데이터
- TMDB/IMDB API 시도 후 한국 영화 정보 부족으로 KOFIC API로 전환

## 주요 수행 내용

- 데이터 수집 실패 대응: 대체 데이터 확보 및 다중 소스 병합(KOFIC+네이버+수동입력)
- 결측치 처리 및 파생변수 생성(개봉월, 런타임 구간 등)
- 장르별 트렌드 분석, 평점-흥행 상관분석, 감독·배우 네트워크 분석 수행

## 분석 결과

- 장르별 흥행 공식 도출(사극=명절·대작, 스릴러=겨울, 가족영화=연말)
- 네티즌 평점과 흥행의 약한 상관관계 확인, 비평가 평점은 작품성 지표로 기능
- 감독·배우 협업 네트워크가 흥행에 미치는 브랜드 효과 검증
- 데이터 수집 및 정의의 중요성 확인

## 사용 기술 및 역량

- Python (Pandas, Matplotlib, Seaborn), API 연동 및 데이터 수집, EDA 및 데이터 전처리

# 기타 분석 실습 (Analysis Practice)

- 음악 데이터베이스 기반 고급 SQL 쿼리 분석 및 비즈니스 인사이트
- 채널별 마케팅 ROI 최적화 분석(LTV/CAC)
- FastAPI 기반 텔레그램 챗봇 개발: GPT 모델 연동 및 웹훅 시스템
- Titanic 데이터 분석: 고급 전처리 및 생존 예측 모델 비교
- 레드와인 품질 예측: 모델 규제 및 성능 최적화
- LLM Agent 구축: 도구 활용형 AI 학습 조교 시스템 개발