

전주시 사회적기업 빅데이터 분석



2025학년도 2학기 SW 캡스톤디자인 경진대회

오픈소스 URL : <https://github.com/streamlit/streamlit>, <https://github.com/pandas-dev/pandas>, <https://github.com/numpy/numpy>, <https://github.com/plotly/plotly.py>

팀 명 페이로드
지도교수 박영진

팀 원 양수빈(문헌정보학과, 4), 박세정(컴퓨터공학부, 4), 박희진(IT지능정보공학과, 4), 전서연(IT지능정보공학과, 4)
산업체 선도소프트

개발 동기 및 목적

- 본 프로젝트는 전주시 사회적 경제 기업의 매출, 고용 구조, 지역별 분포 등 핵심 지표를 체계적으로 파악할 수 있는 의사결정 지원 도구의 필요성에서 출발하였다. 기존에는 기업 유형별 특징, 인증 현황, 취약계층 고용률 등의 정보가 여러 자료에 분산되어 있어 정책 활용이 어렵고 분석의 일관성이 부족했다.
- 이에 본 시스템은 전주시 사회경제기업 데이터를 통합하여 현황 분석 → 유형별 비교 → 연도별 추세 분석 → 리스크 탐지 → 우선지원 기업 추천 → AI 기반 정책 분석까지 하나의 플랫폼에서 제공하는 것을 목표로 하였다.
- 이를 통해 담당자가 지역 사회경제 생태계를 직관적으로 이해하고, 지원 필요성이 높은 기업을 효율적으로 식별하며, 정책 수립 과정에서 활용 가능한 정량적 근거를 제공하는 것을 핵심 개발 목적으로 설정하였다.

주요 기술

1. 데이터 처리 및 분석 기술
 - Python: 분석 로직 구현, 데이터 처리, API 연동, UI 렌더링 등 전반적 시스템 구성에 활용.
 - Pandas / NumPy: 결측치 처리, 데이터 타입 변환, 표준스키마 적용, 통계값 계산 등 고품질 전처리를 위해 활용.
 - GeoJSON 기반 공간 데이터 처리: 전주시 행정동 경계를 추출하여 지도 시각화용 레이어 생성.
2. 시각화 및 인터페이스 기술
 - Streamlit: 대시보드 UI 구성, 사용자 입력 처리, 페이지간 이동 등 웹 기반 인터랙션을 구현. Tabs 및 Columns 구조를 활용하여 반응형 레이아웃을 구성하고 기능 확장성을 확보함.
 - Plotly: 지도(Choropleth), 산점도, 라인 차트, 막대 그래프 등 다양한 시각화 모듈을 사용하여 분석 결과를 직관적으로 표현함.
3. 외부 API 및 AI 기술
 - Kakao 주소 API: 기업 주소 → 위도·경도 변환 및 행정동 추출을 통해 정확한 공간 분석 기반을 확보.
 - OpenAI API: 기업 데이터 기반 인사이드 도출, 정책 대안 생성, 시나리오 시뮬레이션 등 고도화된 정책 분석 자동화를 구현.

개발 내용

1. 데이터 통합 및 표준화
 - 사회적기업, 마을기업, 자활기업, 협동조합 등 여러 유형의 데이터를 하나의 표준 스키마로 통합하였다.
 - 결측치, 중복, 형식 오류를 정제하여 분석 가능한 형태로 정리하였다.
2. 지도 기반 공간 데이터 구축
 - Kakao Address API로 기업 주소를 위·경도, 행정동 정보로 변환하였다.
 - GeoJSON을 이용해 전주시 행정동 경계를 적용하고, 기업 분포·매출·취약계층 고용을 지도에 시각화하였다.
3. 분석 로직 및 자동 분류 기능 개발
 - 인증/예비기업 자동 분류, 예비기업 유효기간 계산 및 임박 기업 표시 기능을 구현하였다.
 - 연도별·누적 인증 추세를 계산하고, 매출·고용률·SVI를 결합한 우선지원 대상 추천 알고리즘을 설계하였다.
4. 연도별 추세·리스크 탐지 기능
 - 샘플데이터로 매출·고용·취약계층 고용률의 연도별 변화율을 분석하였다.
 - 주요 지표의 급감 기업을 자동 탐지하는 리스크 분석 기능을 구현하였다.
5. AI 기반 정책 분석 기능
 - OpenAI API를 활용해 인사이드 도출, 정책 대안 생성, 시나리오 시뮬레이션을 자동화하였다.
 - 실시간 타이핑 방식으로 분석 결과를 출력하여 과정과 결과를 직관적으로 확인할 수 있도록 하였다.

결과 및 분석

- 본 프로젝트를 통해 구축된 플랫폼은 전주시 사회경제 기업의 구조를 정량적·시각적으로 이해할 수 있도록 지원하며, 다음과 같은 성과를 도출하였다.
1. 현황 파악의 효율성 향상
 - 지도 기반 분포, 업종 구조, 매출·고용 지표를 통합 제공하여 지역별·유형별 현황을 즉시 파악할 수 있게 되었다.
 2. 정책 의사결정의 근거 제시
 - 기업 유형 간 매출·고용·SVI 분석을 통해 정책 우선순위 설정 및 자원 배분 판단에 활용 가능한 근거를 제공한다.
 3. 리스크 관리 및 선제적 대응 가능
 - 연도별 변화율 기반 리스크 탐지 기능을 통해 위험 기업을 조기 발견할 수 있어 행정적 모니터링 효율성이 크게 증가하였다.
 4. 우선지원 및 인증 관리 자동화
 - 예비기업 유효기간 자동 계산, 임박 기업 하이라이팅, 우선지원 대상 추천 기능을 통해 담당자의 업무 부담을 실질적으로 감소시켰다.
 5. AI 기반 정책지원 고도화
 - AI 분석 모듈을 통해 정책 인사이드·대안·시뮬레이션을 자동 생성함으로써 정책 보고서 작성 효율성을 크게 향상시켰다.
- 종합적으로, 본 프로젝트는 사회경제기업 데이터를 활용한 실무형 정책지원 플랫폼으로서 높은 활용 가능성을 보여주며, 향후 행정 데이터 연동 및 분석 모델 고도화를 통해 더욱 발전된 정책지원 시스템으로 확장될 수 있는 기반을 마련하였다.