

「2026년 안양시 공공데이터 · AI 활용 대학생 경진대회」 제품 및 서비스 개발 사업계획서

✓ 평가자가 이해하기 쉽게 10페이지 이내 작성

참가자(팀)명	김형준
제품(서비스)명	안물 - 안양의 물가를 물어봐
서비스 주소	https://anyang2026.vercel.app
소스 저장소	https://github.com/HyeongJun02/AY-2026

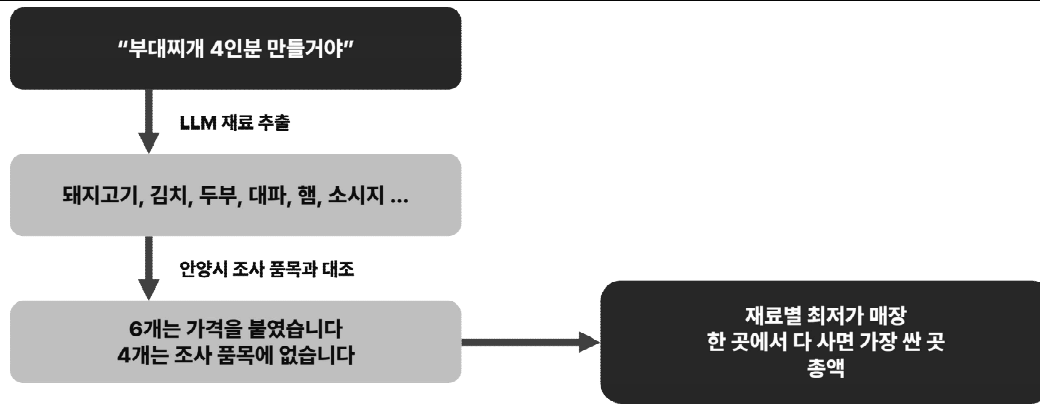
1. 제안 배경 및 출품작 소개(구체성, 서비스 경로 등)

○ 1-1. 안양시는 이미 답을 갖고 있지만, 시민은 그 답에 닿지 못한다.

- 안양시는 매달 생필품 44개 품목의 가격을 관내 22개 매장에서 조사해 공공데이터포털에 공개합니다. 2013년부터 이어진 조사입니다. 하지만, 시민이 실제로 사용하려면 44행 x 28열짜리 csv를 내려받아 직접 비교해야 합니다.
- 그 표 안에 있는 사실을 보면, 2026년 8월 기준 마른 멸치는 뉴코아아울렛에서 2,240원, 이마트(평촌점)에서 7,900원입니다. 규격이 같고, 둘 다 안양 안의 대형 유통이지만 약 3.5배의 가격 차이가 발생합니다.
- 이 사실은 공개되어 있지만, 아무도 모릅니다. 표를 열어 두 칸을 찾아 나란히 놓아야만 보이기 때문입니다.
- ‘공개’와 ‘활용 가능’은 다르다고 생각합니다. 이 격차를 메우는 것이 본 출품작의 목적입니다.

○ 1-2. 시민은 ‘품목’이 아니라 ‘요리’로 생각한다

- 기존 물가 서비스는 모두 ‘품목 조회형’입니다. 그러나 장을 보러 가는 사람의 머릿속에 먼저 떠오르는 것은 품목이 아니라 “오늘 저녁에 부대찌개를 해야지”입니다. 품목은 그 다음 순서에 따라 나옵니다.
- 안물은 그 순서를 그대로 따릅니다.



○ 1-3. 서비스 경로

- 하단 5개 탭으로 구성했습니다. 가운데 ‘장보기’가 메인 기능입니다.

경로	화면	하는 일
/	장보기	만들 요리를 입력 -> 재료 추출 -> 재료별 최저가 검색
/today	시세	예년과 다른 품목/오늘 가장 싼 매장/동네슈퍼 격차
/store, /search	찾기	매장별 가격지수 순위/품목별 가격 비교
/map	지도	공영주차장 107곳, 전통시장 9곳, 내 위치 기준
/cart	장바구니	체크리스트/남은 금액 계산
/store/[매장]	매장 상세	가격지수/위치 지도/파는 품목 42개/즐거찾기
/item/[품목]	품목 상세	13년 가격 추이/계절 판정/매장별 가격/장바구니 담기
/admin	(행정용)	이용 현황+물가 브리핑 대시보드/HTTP Basic 인증

○ 1-4. 시연 가능 여부

- 접수 마감 시점 배포 완료 상태이며, 별도 설치 없이 웹 브라우저로 즉시 시연 가능합니다. (모바일 웹앱 설계 및 데스크톱 브라우저 동작 O) - <https://anyang2026.vercel.app>

2. 학습 · 분석 도구 활용 및 AI 기술 적용 세부 내용

○ 2-1. 자연어 -> 재료 추출 (LLM, OpenAI gpt-4o-mini)

- 사용자가 자유롭게 쓴 문장에서 요리명 · 인분 · 재료 · 분량 · 조리법을 뽑습니다. Structured Outputs로 스키마를 강제해 파싱 실패를 차단했습니다.

- 겪은 문제와 해결한 방법

(가) 모델이 품목명 대신 규격을 반환하는 문제

=> 프롬프트에 카탈로그 구분

=> 서버 4단계 정규화 (완전일치, 괄호/하이픈 앞 어절, 규격 역인텍스, 부분일치)

(나) 한국인 기준이 아닌 결과

=> 프롬프트에 한국 가정식 기본값 명시

- => 분량, 불세기, 시간, 완료 신호를 반드시 포함하도록 강제
- => 요리를 분류 별로 나누어 받아 행정용 통계로 활용

○ 2-2. 계절성 학습과 이상 감지

- 배추는 원래 여름에 오르고, 상추는 겨울에 오릅니다. 정확한 판단을 위해서는 예년 같은 달과 전주어야 합니다.
- 안양시 생필품가격동향 29개 회차(2013-11 ~ 2026-08)를 수집했습니다. 이 중 2013-11 ~ 2015-11의 월별 연속 25회차에서 품목별 계절 지수를 학습했습니다.

$$S_i(m) = (\text{품목 } i \text{의 } m \text{월 평균가}) / (\text{품목 } i \text{의 전 구간 평균가})$$

$$n_i(m) = (S_i(m) / S_i(m-1) - 1) \times 100$$

검산: 배추 $S(7) = 1.1740, S(8) = 1.3818$

→ $n = +17.7\%$.

실제 +25.2% → 차이 +7.5%p → 정상범위

- 학습 결과는 농업적으로 타당합니다.

품목	1월	4월	7월	8월	11월	해석
배추	+0	-17	+17	+38	-22	여름 고랭지 고가, 11월 김장철 출하
호박	+29	+11	-24	-16	-5	겨울 시설재배, 고가
상추	+47	-32	-9	+0	+20	겨울 엽채류, 고가
쌀	+1	+0	-1	-1	-2	저장 작물, 계절성 없음

- 원본 ‘증감’ 열만 사용한다면 배추 +25.2%, 호박 +37.7%와 같은 무게로 경고되지만, 계절 보정을 하면 배추는 정상 범위가 되고, 오이가 예년보다 29%p 저렴하다는 정보가 드러납니다. 이는 기존 방식으로는 보이지 않던 신호입니다.

○ 2-3. 규격 불일치 자동 탐지

- 같은 품목이라도 매장마다 다른 규격으로 조사된 듯 보이는 셀이 있습니다. 이를 방지한다면 지수가 오염됩니다. 품목별 중앙값 대비 배율로 체크합니다.
 - (1) 1/2 ~ 2배 外 : 화면에 ‘규격 확인 필요’ 표시 (사용자에게 경고)
 - (2) 1/3 ~ 3배 外 : 지수 계산에서 제외 (통계 보호)
- 실제로, [간마늘 400g]의 22개 매장 값 중 19곳은 3,000~7,920원인데, 남부시장만 13,000원입니다. kg 당 32,500원으로 간마늘 소매가의 두세 배라, 같은 400g 기준으로 조사된 값으로 보기 어려웠습니다.

3. 출품작 핵심내용(공공데이터 활용 적정성, 기술성 등)

○ 공공데이터 활용 항목(필수)

※ 활용한 데이터의 명칭과 관련 URL 필수 기재

※ 안양시 공공데이터 활용(공공데이터포털)의 경우 하단 표 미기재 등 확인 불가일 경우 가중치 배점 부여 불가

제공 기관(플랫폼)명	데이터명	URL
공공데이터포털	경기도 안양시_생필품가격동향	https://www.data.go.kr/data/3045179/fileData.do
공공데이터포털	경기도 안양시_국민기초생활 수급자 동별 현황	https://www.data.go.kr/data/15114169/fileData.do
공공데이터포털	전국전통시장표준데이터	https://www.data.go.kr/data/15012894/standard.do
공공데이터포털	전국주차장정보표준데이터	https://www.data.go.kr/data/15012896/standard.do

○ 3-1. 데이터 파이프라인 - 29개 회차, 서식 3종 통합

- 13년치 csv의 정규화

문제	처리
서식 3종 혼재 (증감형 11열/기준일형 9열/2026년 28열)	열 위치를 세지 않고 “아는 매장 이름이 처음 나오는 자리부터가 매장 블록” 규칙으로 파싱
헤더 줄바꿈 (“증감 \n(%)“, 2013-12 ~ 2014-06)	줄 단위 분리를 버리고 따옴표 상태를 추적하는 전체 텍스트 파서로 교체
따옴표 안 쉼표 (고등어, “자반고등어(25cm), 1마리”)	44개 품목 중 10개의 매장 가격이 한 칸씩 밀림 -> 인용 인식 파서로 해결
2015-01 ~ 07 줄 끝 여분 쉼표	데이터기준일자가 매장 열로 오인됨 -> 꼬리 빈 열 제거 후 판정
파일명이 조사월이 아님	20231011은 공개일, 헤더는 2023-09-15 평균. 헤더 기준으로 파일명 교체
품목 개명	무 -> 무, 물오징어 -> 오징어, 커피믹스 -> 커피 통합. 단, 삼치 퇴출, 생밤 신규 같은 실제 조사 품목 교체는 통합 x

- 결과: 29개 회차, 최신 44개 품목, 33개 품목이 29회차 전 구간에 존재합니다. 5개 전통 시장(중앙, 남부, 박달, 관양, 호계)이 모든 회차에 있어 비교축이 됩니다.

○ 3-2. 가격지수 산출 - 단순 합계를 버린 이유

- 처음에는 44개 품목 가격을 단순 합산하여 ‘장바구니 총액’으로 매장을 줄 세웠습니다. 하지만 마늘처럼 규격 오기가 섞인 소수 품목(2,280원 ~ 13,000원)이 총액을 형클어 놓았습니다. 이에 절사 기하평균(Jevons 지수)으로 교체했습니다. 이는 소비자물가지수의 표준 방법론입니다.

$$\text{매장지수} = \exp\left(\frac{\sum \ln(\text{매장가}/\text{품목중앙값})}{n}\right) \times 100$$

로그를 취하므로 비율 편차가 대칭이 되고, 한 품목의 극단값이 전체를 끌고 가지 못합니다. 방법론이 결론을 바꾸는 사례였기에, 화면에도 산출 방식과 표본 수를 함께 표

기합니다.

(현재 매장 지수 범위: 87.6(뉴코아아울렛) ~ 120.4(롯데백화점 평촌점), 중앙값 100.)

○ 3-3. 안양시 데이터 2종 융합 - 동네슈퍼 x 기초생활수급자

- 2026-08 회차에는 동네슈퍼 11개 행정동 열이 있습니다. 이를 수급자 동별 현황과 결합했습니다.

발견 1) 같은 안양 안에서 동네슈퍼 물가가 17.6% 벌어집니다.

발견 2) 수급자 수와 가격지수 사이에 상관관계는 없습니다. ($r = -0.001$) 표본도 11개 동으로 상관을 논하기에 부족하기 때문에 본 서비스는 상관관계를 주장하지 않습니다. 대신 두 축을 나란히 놓아 우선순위만 제시합니다. (수급자 수와 가격지수가 중앙값 위인 동은 행정동이 먼저 살펴볼 후보이며 순서입니다.)

발견 3) 동네슈퍼 가격 조사는 11개 행정동에서만 이루어집니다. 나머지 17개 동네 사는 기초생활수급자 9,601명(2026년 기준 전체 16,948명의 57%)은 자기 동네 물가를 이 데이터로 확인할 방법이 없습니다. 수급자가 가장 많은 관양동(2,400명)이 그 중 하나입니다. 이는 서비스의 한계가 아닌, 안양시 조사 설계의 피드백이며, 관리 화면에 그대로 표기합니다.

○ 3-4. 기술 구성

구분	내용
프론트/백엔드	Next.js 16 (App Router, React 19, TypeScript, Tailwind CSS v4)
배포	Vercel
데이터베이스	Neon Postgres (이용 통계)
AI	OpenAI gpt-4o-mini : Structured Outputs
지도	Kakao Maps SDK (마커 클러스터링)
데이터 처리	Node.js 빌드 스크립트 -> 정적 JSON (build:prices, build:welfare, fetch:parking, fetch:markets)
접근 제어	Next.js proxy (구 middleware) - 관리 화면 HTTP Basic 인증

- 공공데이터는 빌드 시점에 JSON으로 굳힙니다.

런타임에 외부 API를 부르지 않으므로 API 장애/요청 한도/키 노출이 서비스 가용성에 영향을 주지 않습니다.

- 개인정보를 저장하지 않습니다.

이용 통계는 요리명/분류/재료명만 남기며 IP/식별자는 기록하지 않습니다. LLM 호출은 서버 라우트에서만 이루어져 API 키가 브라우저에 노출되지 않고 IP당 분 5회/일 40회로 제한합니다.

- 관리 화면은 잠급니다.

/admin과 기록 내보내기 API는 HTTP Basic 인증 뒤에 둡니다.

4. 기존 서비스와의 독창성(차별성)

○ vs 참가격 (한국소비자원)

	참가격(한국소비자원)	안물
조사 범위	전국 18개 권역 262개 점포 (대형 마트 119, 백화점 28, SSM 91, 전통시장 21, 편의점 3), 120개 품목 주 1회	안양시 22개 매장 (전통시장 5, 대형유통 6, 동네슈퍼 11개 동), 44개 품목
안양 내 커버리지	전통시장 21곳이 전국 기준이라 안양 개별 시장/동네슈퍼는 다루지 않음	안양 전통시장 5곳, 동네슈퍼 11개 행정동을 개별로
진입점	품목 검색	요리 입력
변동 해석	전월 대비 등락률	예년 같은 달과 비교한 계절 보정
결과 형태	가격 표	재료별 최저가 + 한 곳에서 다사는 최적 매장 + 장바구니
데이터 융합	가격 단일	가격 x 수급자 x 전통시장 x 주차장
행정 환류	x	조사 품목에 없는 재료를 집계해 시에 되돌림

- 참가격은 전국 서비스로서 훌륭하지만, 범위가 전국이라는 사실 자체가 지역 서비스의 빈자리를 만듭니다. 전통시장 21곳은 18개 권역에 나눠 담긴 수치라 안양시의 5개 전통시장과 11개 행정동 동네슈퍼는 여기에 없습니다.

- 핵심 차별점

(1) ‘올랐다’가 아닌, ‘이례적인가?’

안양시 13년 조사를 통해, 계절 요인과 실제 이상을 구분할 수 있도록 접근하여 비교합니다.

(2) 요리에서 출발하는 역방향 설계

품목명을 몰라도 쓸 수 있습니다. 재료 중 이미 집에 있는 것은 체크 해제할 수 있고, 조사 품목에 없는 재료는 따로 분리해서 보여줍니다.

(3) 시민의 질문이 행정 데이터로 환류

조사 품목을 어디로 넓혀야하는지에 대한 근거가 서비스 운영만으로 쌓입니다.

5. 출품작의 완성도

○ 5-1. 구현 범위

- 접수 마감 시점 기준 전 기능 동작하는 배포본이 존재합니다.
(<https://anyang2026.vercel.app>). 화면 9종, API 3종, 데이터 수집 스크립트 4종이 구현되어 있습니다.

○ 5-2. 검증 방법

- 정적 검증
모든 커밋 전 `tsc --noEmit`, `eslint`, `next build` 통과
- 실브라우저 검증
Chrome DevTools Protocol로 헤드리스 브라우저를 구동해 실제 키 입력/클릭/화면 전환을 재현하고, 스크린샷으로 확인합니다.
- 데이터 검증
파싱 결과의 계절 패턴이 농업적으로 타당한지 교차 확인합니다. (2-2) 색상 팔레트는 색각 이상 시뮬레이션을 포함한 5개의 검사를 통과시켰습니다.
- API 실검증
주차장 107곳, 전통시장 9곳을 실제 API 호출로 수신을 확인했습니다.
- 동작 실측
눈으로 보는 대신 값을 짭니다. 관리 화면 잠금, 장바구니 요약 높이, 즐겨찾기 왕복, 링크 무결성 등

○ 5-3. 사용자 경험 (UX)

- 모바일 우선
하단 5탭, 메인 기능은 가운데 원형 버튼.
- 화면 전환 애니메이션
React ViewTransition으로 내용만 교차시키고 헤더, 하단바는 고정합니다. 전환 중 클릭이 유실되지 않도록 오버레이의 포인터 이벤트를 해제했습니다.
- 접근성
색만으로 정보를 전달하지 않습니다. (경고는 아이콘+문구 병기, 막대는 방향+수치 병기). `prefers-reduced-motion`에서 모든 모션을 정지합니다. 모든 차트는 표로도 읽을 수 있습니다.

- 다크 모드

기기 설정을 따르되, 헤더에서 밝게/어둡게로 고정할 수 있습니다.

- 즐겨찾는 매장

‘가까운 시장 3곳’은 처음 가는 동네에서나 쓸모 있고, 평소 질문은 “늘 가는 그 시장은 얼마지?”입니다. 즐겨찾으면 거리와 무관하게 품목 상세에 함께 뜹니다.

- 정직한 표기. LLM 응답 대기 시 골격 화면(skeleton)으로 결과 모양을 미리 알리고, 가격은 “조사 기준 최저가이며 실제 판매가와 다를 수 있음”을 명시합니다.

6. 출품작의 발전 가능성

○ 6-1. 행정용 대시보드

- /admin 은 담당자가 바로 열어 쓰는 화면입니다.

기간(7일/30일/전체) 필터, 지표 6칸(검색/답기/매칭률/조사품목/조사매장/요리당 재료), 일자별 이용 추이, 많이 찾은 요리/분류/장바구니 순위, 그리고 전체 기록 CSV 내려받기(엑셀에서 한글이 깨지지 않도록 BOM 포함)를 담습니다.

- 설계에서 한 가지를 특별히 신경 썼습니다. 이용 기록이 0건이어도 화면이 비지 않습니다. 이벤트만 보여주는 대시보드는 도입 첫날 “아직 기록이 없습니다” 한 줄만 남아 쓸모를 증명하지 못합니다. 그래서 이용 기록과 무관하게 조사 데이터에서 바로 나오는 물가 브리핑(예년과 다른 품목 / 매장 지수 11곳 / 동네슈퍼 11개 동 x 수급자 / 조사 공백 17개 동)을 같은 화면에 함께 둡니다. 도입 즉시 쓸 수 있고, 이용이 쌓일수록 위쪽 절반이 채워집니다.

○ 6-2. 행정이 바로 쓸 수 있는 산출물

산출물	행정 활용
시민이 찾았으나 조사 품목에 없는 재료 순위	조사 품목 44개 확대 근거
조사되지 않는 행정동 17곳 (수급자 9,601명/전체의 57%)	조사 대상 동 확대 근거
동네슈퍼 동별 가격지수 (90.7~106.7)	물가 안정 시책 대상지 선정
계절 보정 이상 감지	물가 대책 발동 시점 판단 - 배추 여름 급등에 대응하는 낭비 줄이기
시민 요리 검색 분류 통계	지역 식생활 파악

○ 6-3. 데이터 제공 방식에 대한 제안

- 수집 과정에서 확인한 개선점을 안양시에 제안합니다.

(1) 서식 통일.

13년간 서식이 3번 바뀌어 시계열 결합에 별도 파서가 필요했습니다.

(2) 파일명에 조사월 표기.

현재 일부 파일명은 공개일이라 조사월과 최대 2개월 어긋납니다.

(3) 2016~2021년 회차 공개.

현재 공백 구간입니다. 이 구간이 열리면 계절 지수의 신뢰도가 크게 오릅니다.

(4) 품목 코드 부여. 개명(무→무우) 시 시계열이 끊깁니다.

○ 6-4. 확장 계획

- 단기: 공백 구간 회차를 확보해 계절 지수 재학습. 관측 수가 달마다 2~3개에서 8~10개로 늘면 예측 구간 추정이 가능해집니다.
- 중기: 알림 - 관심 품목이 예년 대비 이례적으로 쌀 때 통지.
- 장기: 동일 구조를 다른 기초지자체로 이식. 생필품 가격을 조사/공개하는 지자체는 안양시 외에도 있으며, 파이프라인이 서식 변화에 견디도록 설계되어 있어 데이터 소스 교체만으로 확장 가능합니다.

7. 기타 참고 자료

○ 7-1. 본 출품작이 밝히는 한계

(아래 내용은 모두 서비스 화면에도 표기되어 있습니다.)

1. 계절 지수의 표본이 얇다 (같은 달 관측 2~3개)
1. 2015~2022년 조사 공백
2. 계절 비교 가능 품목은 44개 중 38개
3. 수급자 수와 동네슈퍼 가격의 상관관계는 없다 ($r=-0.001$)

○ 7-2. 데이터 규모 요약

항목	수량
생필품가격동향 회차	29회 (2013-11 ~ 2026-08)
조사 품목 (최신 회차)	44개
전 구간 존재 품목	33개
조사 매장 (최신 회차)	22곳 (전통시장 5, 대형유통 6, 동네슈퍼 11)
기초생활수급자 데이터	2015~2026년, 31개 행정동, 16,948명 (2026)
동네슈퍼 조사 동 / 미조사 동	11곳 / 17곳 (수급자 9,601명)
공영주차장	107곳
전통시장	9곳

○ 7-3. 확인 방법

- 서비스: <https://anyang2026.vercel.app>
- 소스 코드: <https://github.com/HyeongJun02/AY-2026>
- 원본 CSV는 저장소 data/prices/ 에 조사될 파일명으로 정리되어 있어, 가공 전후를 대조하실 수 있습니다.

/admin과 기록 내보내기는 서비스가 모은 자료를 다루므로 잠가두었습니다. /admin 페이지 진입 시 아래 정보를 입력해주세요.

주소	https://anyang2026.vercel.app/admin
아이디	anyang
비밀번호	siCurCQwupqV

○ 7-4. 제품 스크린샷

