

처음 만난 데이터, AI와 함께 다시 읽다

경기도일자리재단 데이터팀

배경

데이터 기반 사회문제 해결능력 강화와 새로운 기회창출

공공 데이터 한계 극복

시의성있는 지역경제 모니터링

중장기 정책 수립

단기 진단이 아닌 심층분석 및 미래 예측

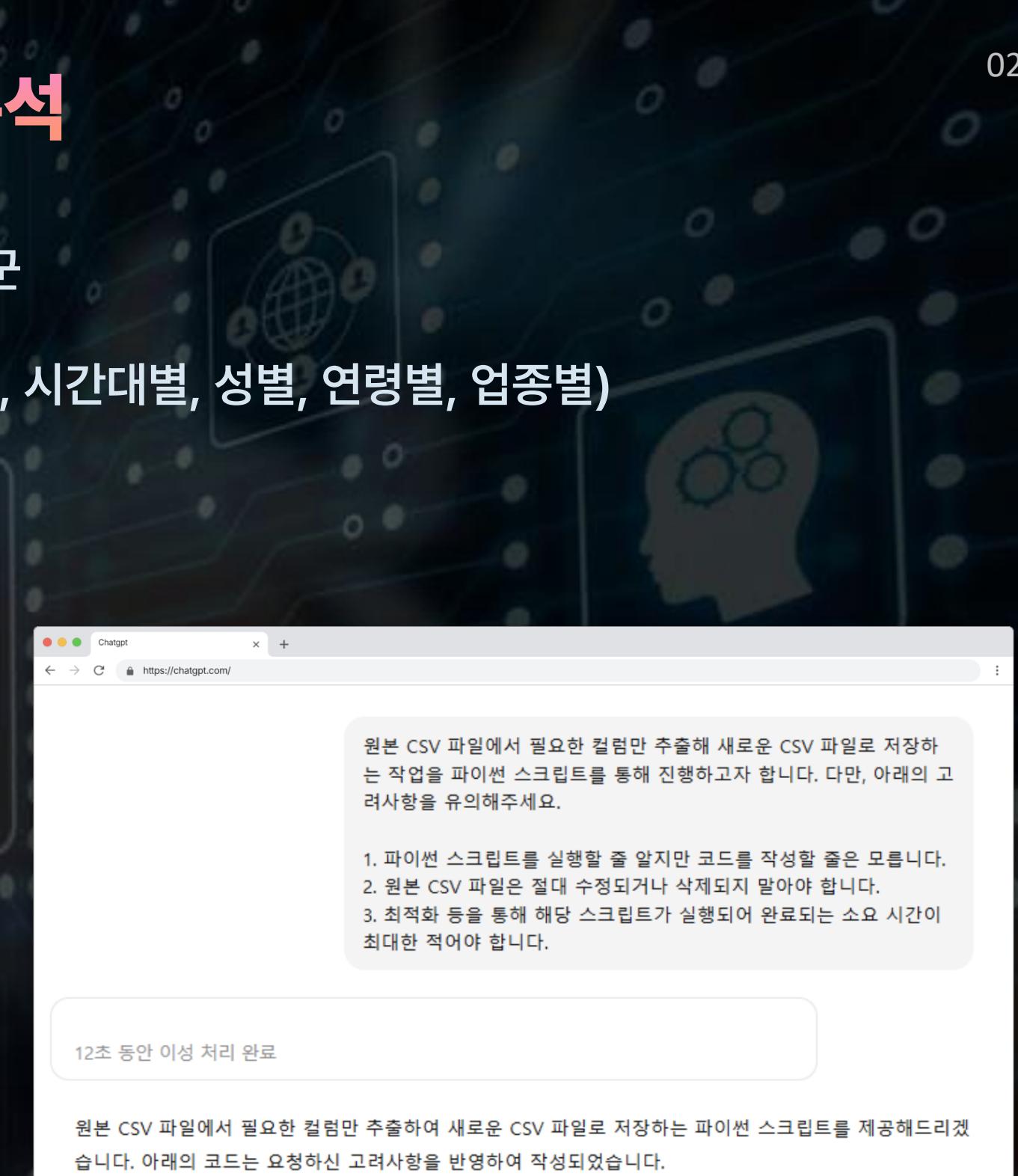
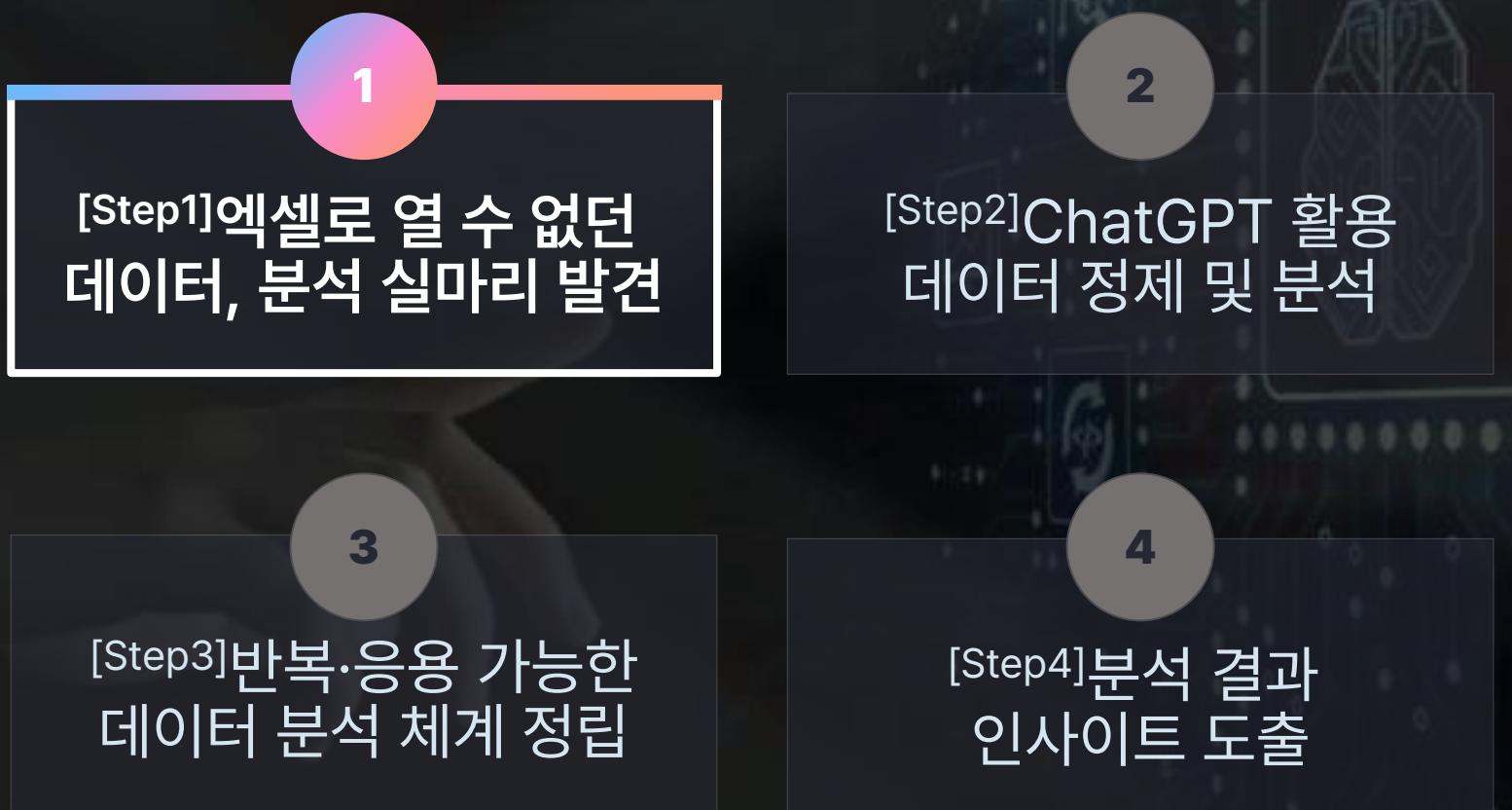
AI를 활용한 역량개발 능동적 대안

민간 카드사 데이터 자체 분석
(26억 Row, 28GB)

신속한 분석 요구사항 반영을 위한 자체 분석
애플리케이션 개발

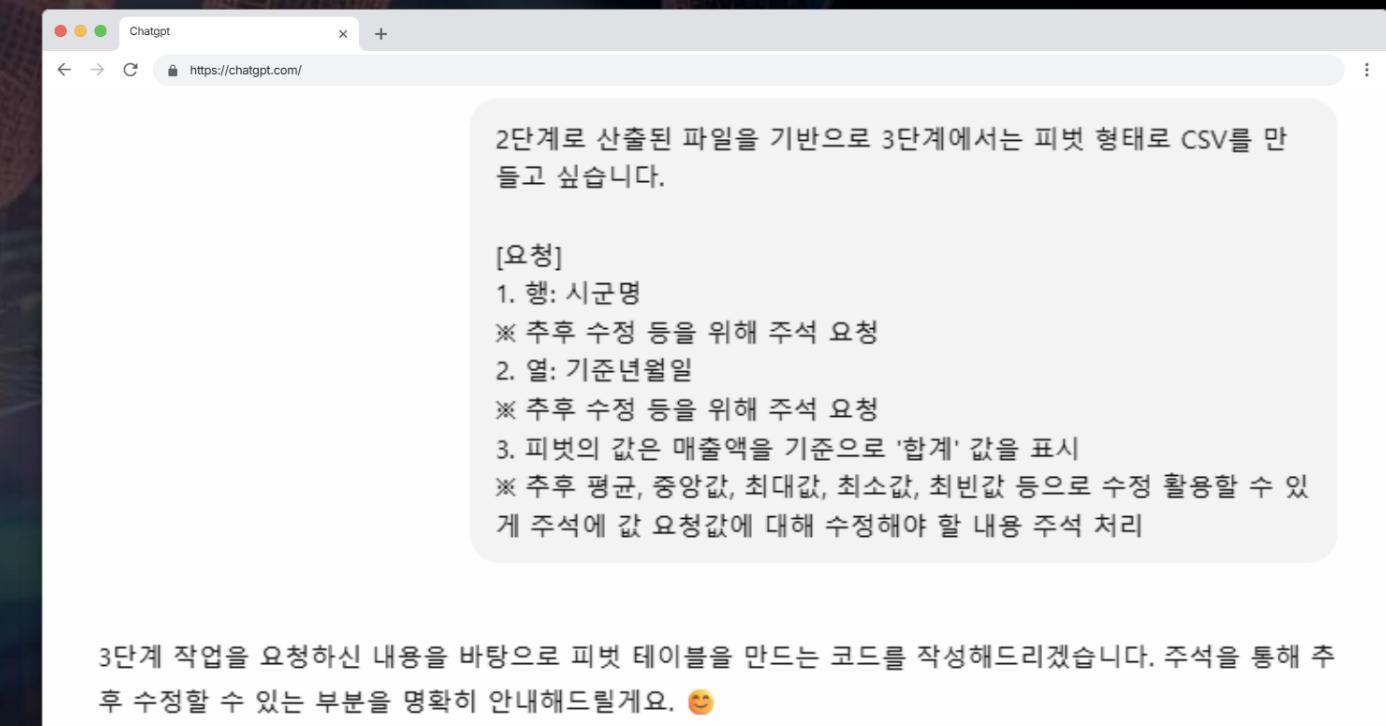
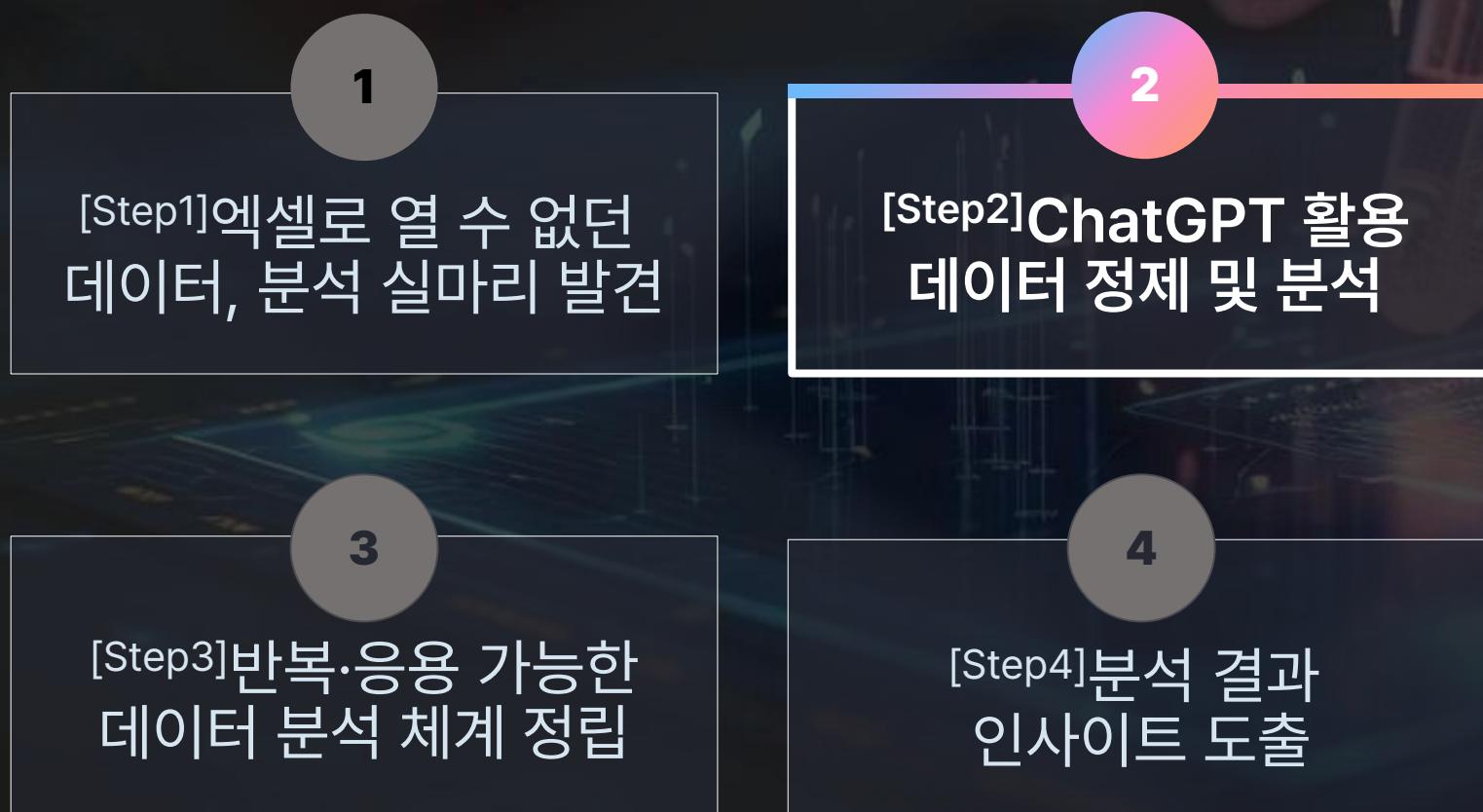
전략1 생성형 AI 기반 민간 카드 데이터 분석

- 분석범위: (시간) 25년 1월 기준, (공간) 경기도 및 31개 시·군
- 분석대상: 가맹점 수(신규, 휴업, 폐업), 매출액(일별, 요일별, 시간대별, 성별, 연령별, 업종별)
- 분석절차



전략1 생성형 AI 기반 민간 카드 데이터 분석

- 분석범위: (시간)25년 1월 기준, (공간) 경기도 및 31개 시·군
- 분석대상: 가맹점 수(신규, 휴업, 폐업), 매출액(일별, 요일별, 시간대별, 성별, 연령별, 업종별)
- 분석절차



전략1 생성형 AI 기반 민간 카드 데이터 분석

- 분석범위: (시간)25년 1월 기준, (공간) 경기도 및 31개 시·군
- 분석대상: 가맹점 수(신규, 휴업, 폐업), 매출액(일별, 요일별, 시간대별, 성별, 연령별, 업종별)
- 분석절차

1

[Step1] 엑셀로 열 수 없던
데이터, 분석 실마리 발견

2

[Step2] ChatGPT 활용
데이터 정제 및 분석

3

[Step3] 반복·응용 가능한
데이터 분석 체계 정립

4

[Step4] 분석 결과
인사이트 도출

```

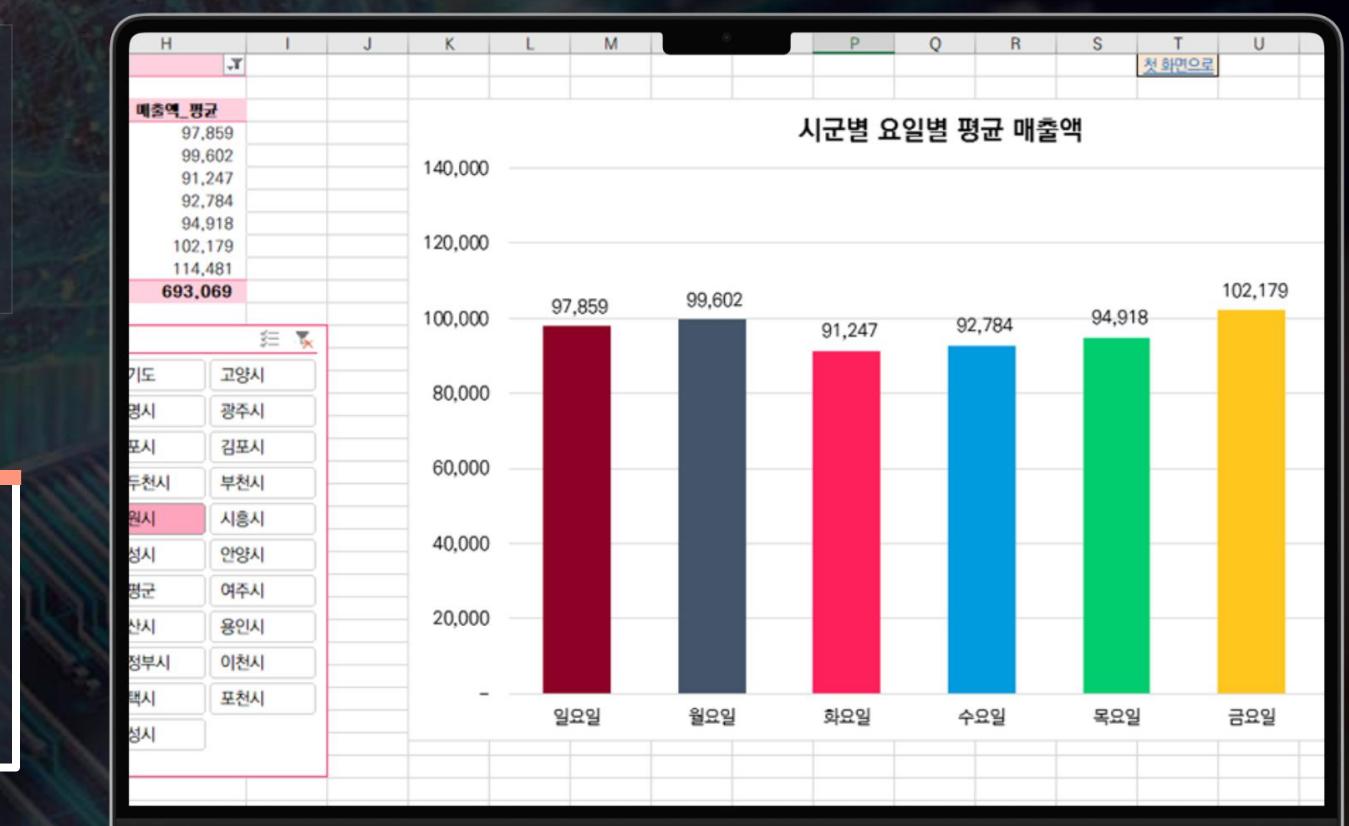
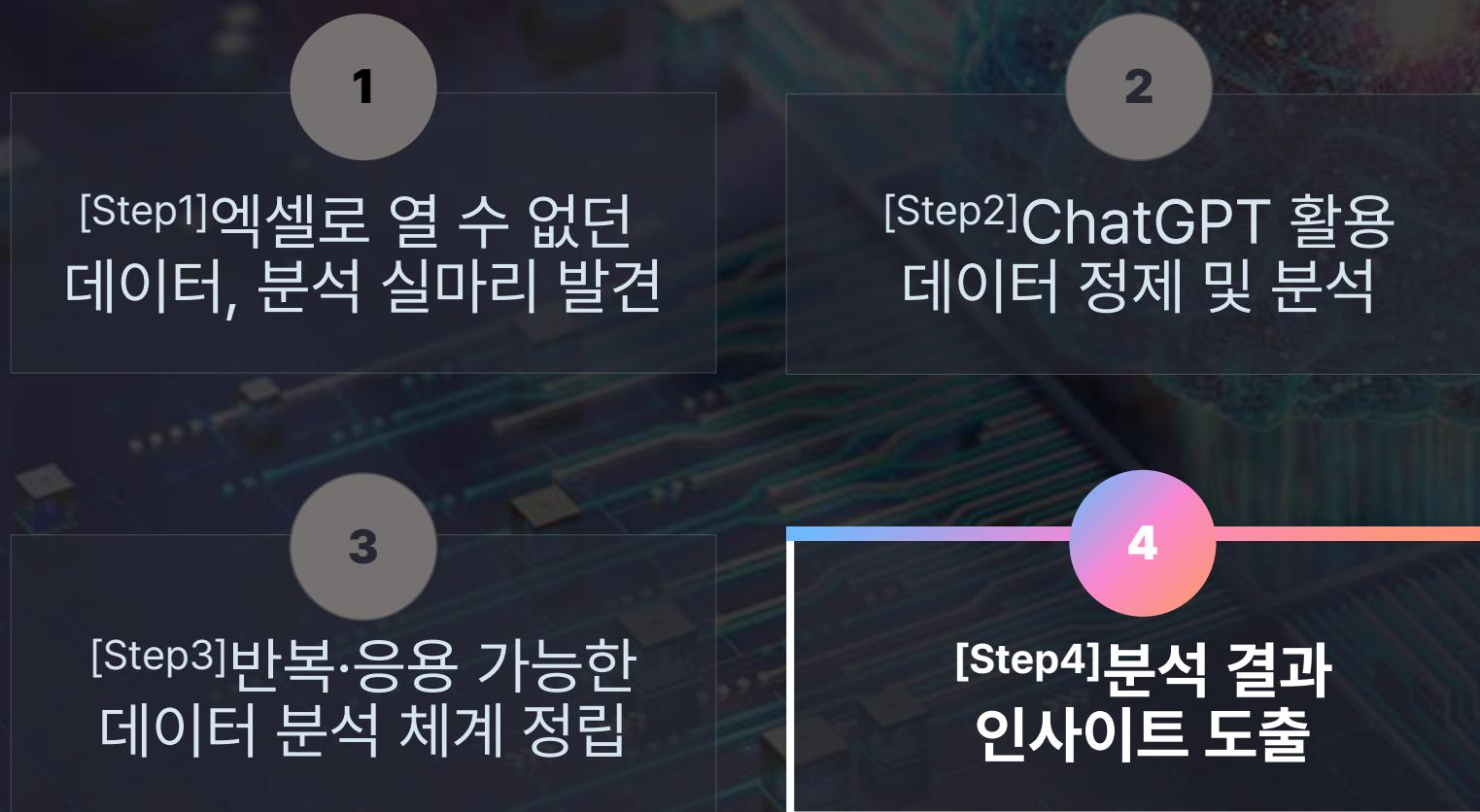
csv_diet&aggregator.py
1 print(f"▶ 경지원 인코딩: {encoding}")
2 print(f"▶ 경지원 구분자: '{delimiter}'")
3
4 # 자동 출력 파일명 생성 함수
5 def generate_output_path(input_filename, folder, prefix="최적화"):
6     base = os.path.splitext(input_filename)[0]
7     for i in range(1, 1000):
8         suffix = f"_0{i:02d}"
9         candidate = f"{folder}/{prefix}_{base}{suffix}.csv"
10        if not os.path.exists(candidate):
11            return candidate
12        raise Exception("▲ 저장 파일이 999개를 초과했습니다.")
13
14 # 헤더 명령(추출한 헤더가 없을 경우 반드시 추가 필요)
15 column_rename_map = {
16     "ta_ym": "기준년월",
17     "ta_ymd": "기준년월일",
18     "cty_rgn_no": "지역명",
19     "admi_cty_no": "행정동",
20 }
21
22 urls <-- csv_diet&aggregator
23
24 처리 결과 요약:
25 원본 행 (Row) : 115,837,143
26 필터링 행 (Row) : 115,837,143
27 원본 파일 크기 : 9.2 GB
28 결과 파일 크기 : 1.8 MB
29
30 다양한 통계 결과 파일:
31     data\card\데이터\카드4(3월없음)\집계결과_tbsh_gyeonggi_agrde_2405_002.csv
32
33 data > csv_diet&aggregator.py
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50

```

The screenshot shows the PyCharm IDE interface with a Python script named 'csv_diet&aggregator.py' open. The code is written in Python and performs several tasks: it prints file encoding and separator information, defines a function to generate output file paths, and contains a mapping for column renaming. Below the code editor, there's a summary of the data processing results, stating that the original file has 115,837,143 rows and the filtered result also has 115,837,143 rows, with both being 9.2 GB in size and the output being 1.8 MB. At the bottom, there are two tabs: 'urls' and 'csv_diet&aggregator'.

전략1 생성형 AI 기반 민간 카드 데이터 분석

- 분석범위: (시간) 25년 1월 기준, (공간) 경기도 및 31개 시·군
- 분석대상: 가맹점 수(신규, 휴업, 폐업), 매출액(일별, 요일별, 시간대별, 성별, 연령별, 업종별)
- 분석절차



전략2 맞춤형 데이터 분석 애플리케이션 개발

· 현황

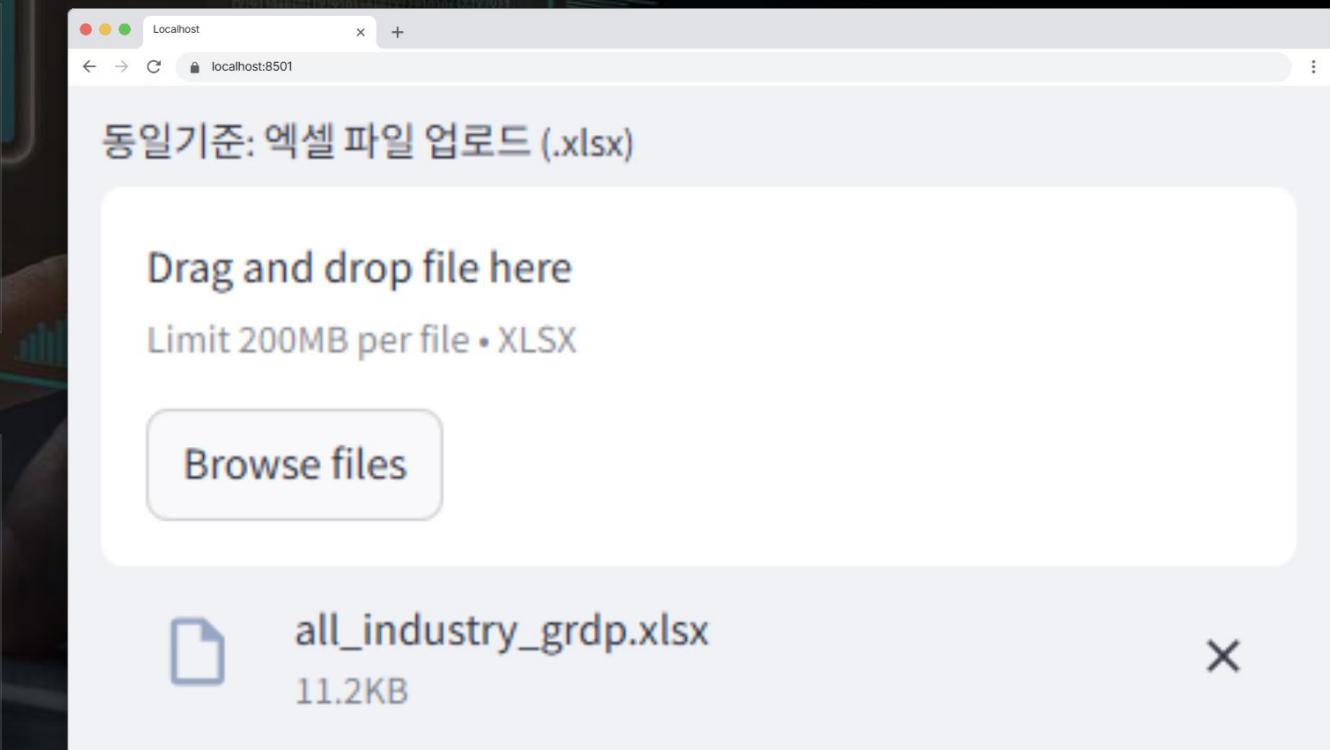
- 데이터 기반 의사결정시 다양한 관점의 데이터 분석 요구 증가
- 현존하는 분석 도구는 경기도 및 31개 시·군 맞춤형 일자리, 산업 분석 한계 존재· 주요기능

1
[Function1] 일자리 지표
엑셀 업로드
(경제활동인구, 고용률, GRDP)

3
[Function3] 일자리 지표의
시계열 추이와 성장률
그래프 생성

2
[Function2] 자동 세부항목 출력
및 선택 네비게이션 생성

4
[Function4] 미래예측 모델
(ARIMA, PROPHET) 적용
그래프 생성



전략2 맞춤형 데이터 분석 애플리케이션 개발

· 현황

- 데이터 기반 의사결정시 다양한 관점의 데이터 분석 요구 증가
- 현존하는 분석 도구는 경기도 및 31개 시·군 맞춤형 일자리, 산업 분석 한계 존재· 주요기능

1

[Function1] 일자리 지표
엑셀 업로드
(경제활동인구, 고용률, GRDP)

2

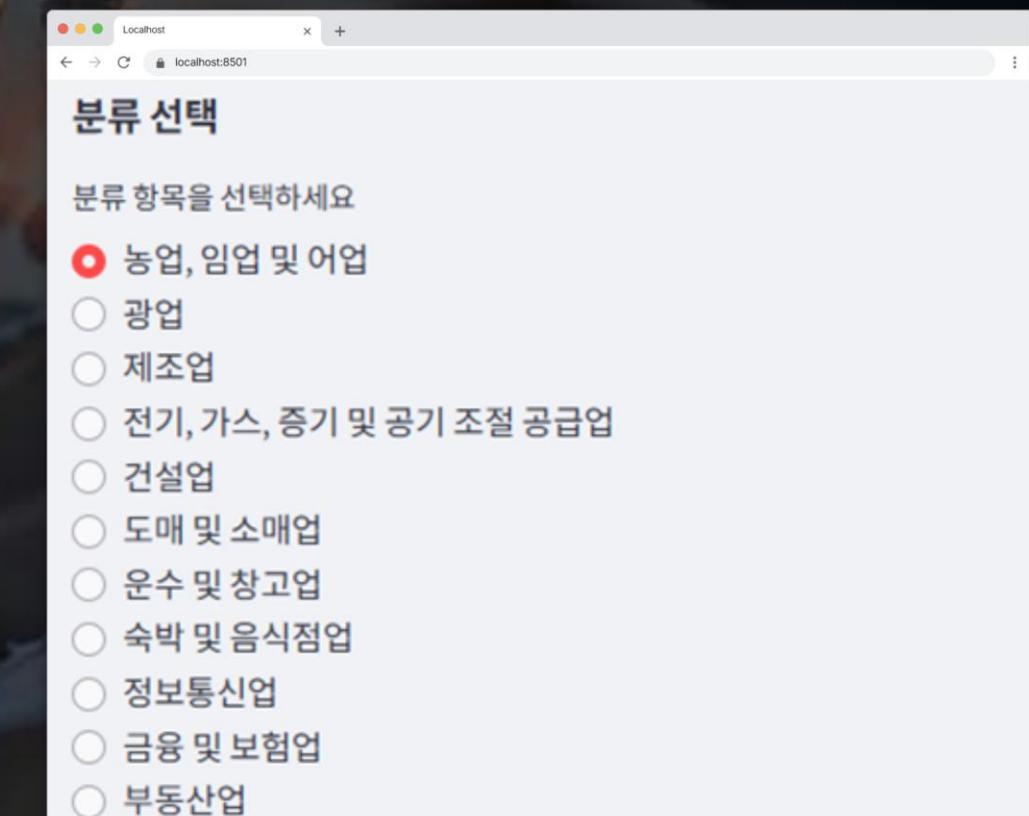
[Function2] 자동 세부항목 출력
및 선택 네비게이션 생성

3

[Function3] 일자리 지표의
시계열 추이와 성장률
그래프 생성

4

[Function4] 미래예측 모델
(ARIMA, PROPHET) 적용
그래프 생성



전략2 맞춤형 데이터 분석 애플리케이션 개발

· 현황

- 데이터 기반 의사결정시 다양한 관점의 데이터 분석 요구 증가
- 현존하는 분석 도구는 경기도 및 31개 시·군 맞춤형 일자리, 산업 분석 한계 존재· 주요기능

1

[Function1] 일자리 지표
엑셀 업로드
(경제활동인구, 고용률, GRDP)

2

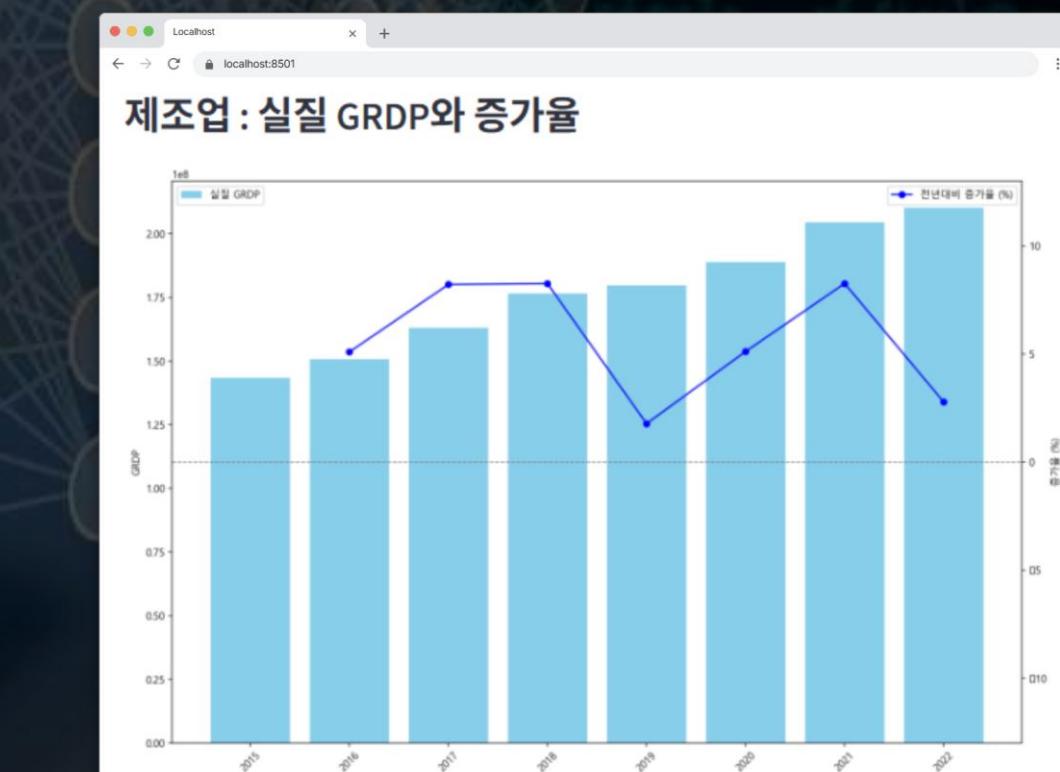
[Function2] 자동 세부항목 출력
및 선택 네비게이션 생성

3

[Function3] 일자리 지표의
시계열 추이와 성장률
그래프 생성

4

[Function4] 미래예측 모델
(ARIMA, PROPHET) 적용
그래프 생성



전략2 맞춤형 데이터 분석 애플리케이션 개발

· 현황

- 데이터 기반 의사결정시 다양한 관점의 데이터 분석 요구 증가
- 현존하는 분석 도구는 경기도 및 31개 시·군 맞춤형 일자리, 산업 분석 한계 존재· 주요기능

1

[Function1] 일자리 지표
엑셀 업로드
(경제활동인구, 고용률, GRDP)

2

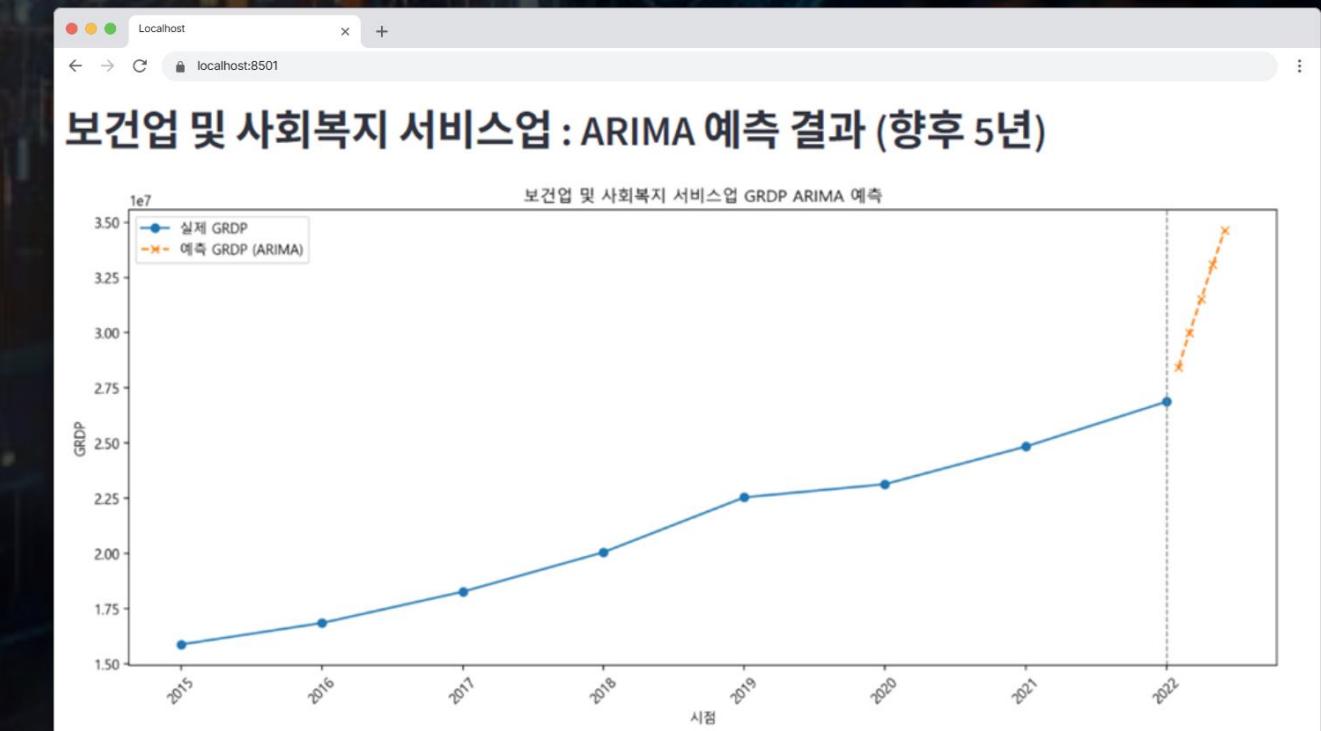
[Function2] 자동 세부항목 출력
및 선택 네비게이션 생성

3

[Function3] 일자리 지표의
시계열 추이와 성장률
그래프 생성

4

[Function4] 미래예측 모델
(ARIMA, PROPHET) 적용
그래프 생성



추진성과

정량적 성과

예산 2억 9천 4백만원 절감
행정 소요기간 4주 단축

정성적 성과

비전문가 주도의
데이터 분석 역량 내재화

민간 카드데이터 7종
32개 기술통계 분석 완료

반복 응용 가능한 템플릿 생성으로
지속 가능한 자산 전환

기대효과

적극성

외부에 의존하지 않고, 생성형 AI 활용하여
직접 분석 및 애플리케이션 개발 수행

창의성

Python, Orange3, Excel 등 다양한 분석도구와
OA프로그램을 혼합하여 최적화 분석 프로세스 완성

활용성

비전문가도 쉽게 Python 코드를 수정 및 확장,
누구나 분석 전데이터 전처리 및 집계 등이 가능한
실무 활용도 확보

확산성

별도 시스템이나 고가의 소프트웨어없이
데이터 분석 진행으로 지자체 및 공공기관에서
즉시 적용 및 벤치마킹 용이

THANK YOU

“AI가 인간을 대체하지는 않을 것이다.

그러나, AI를 활용하는 사람이 활용하지 않는 사람을 대체할 것이다.”

- 지니 로메티(前 IBM CEO) -