

의약품 처방 데이터 분석 프로젝트

1. 서론

1.1 분석 배경

의약품 처방 데이터는 의료 시장과 환자의 건강 트렌드를 분석하는 데 중요한 역할을 한다. 본 프로젝트는 국민건강보험공단에서 제공한 처방 데이터를 활용하여 주요 처방 패턴을 분석하고, 향후 처방 추세를 예측하는 데 목적을 둔다. 또한, 의료 데이터 분석 경험을 쌓고, 데이터 시각화 및 활용 방안을 모색하는 과정으로서 진행되었다. 이를 통해 의료 데이터를 다루는 역량을 향상시키고, 향후 데이터 기반 의사결정에 대한 이해도를 높이는 것을 목표로 한다.

1.2 프로젝트 목적

본 프로젝트의 목표는 다음과 같다:

- 성별, 연령대, 지역별 처방 패턴 분석
- 특정 약물군 (예: 항히스타민제)별 월별 처방 추이 분석
- 머신러닝 기반 예측 모델을 활용한 향후 처방 추이 예측
- 의료 데이터 분석 및 Power BI 를 활용한 데이터 시각화 경험 습득
- 의료 데이터 활용 방안을 모색하여 데이터 기반 의사결정 능력 향상

2. 데이터 개요

2.1 데이터 수집 과정

HIRA(건강보험심사평가원)의 보건의료빅데이터개방시스템에서 유의미한 처방 데이터를 확보하려 했으나, 활용 가능한 처방 데이터가 충분하지 않아 공공데이터포털을 통해 국민건강보험공단의 의약품 처방정보 데이터를 활용하였다. 해당 데이터는 매년 100만 명의 표본 데이터를 기반으로 하며, 연간 약 3,000만 건의 처방 기록을 포함하고 있다.

처음에는 2019년부터 2023년까지의 5년 치 데이터를 활용하려 했으나, 2022년 이후 데이터는 ATC 코드로, 2021년 이전 데이터는 주성분 코드로 표기되어 있어 데이터 일관성이 맞지 않았다.

이에 따라 보다 정확한 처방 분석을 위해 ATC 코드가 제공되는 2022년 이후 데이터를 사용하려 했으나, 존재하지 않는 ATC 코드(예: 1018A, 2229A 등)가 포함된 것을 발견하였다. 이는 주성분 코드가 ATC 코드로 변환되는 과정에서 임의의 코드로 변환되면서 생긴 오류였고, 데이터 신뢰성 및 일치성 문제로 인해 2022년 이후 데이터 활용이 어렵다고 판단하였다. 따라서 2019년부터 2021년까지의 3년 치 데이터를 분석 대상으로 선정하였다.

2.2 데이터 전처리 과정

데이터 전처리는 2019년, 2020년, 2021년 데이터를 통합하고, 분석을 위한 일관된 포맷으로 정리하는 과정이 포함되었다.

1. 파일 로드 및 칼럼 정리

- 각 연도의 처방 데이터는 세 개의 CSV 파일로 나누어져 있었으며, 이를 하나로 병합하였다.
- 2019~2020년과 2021년 데이터의 칼럼명이 상이하여 통일하는 과정을 거쳤다.

2. 데이터 필터링 및 변환

- 요양개시일자(처방 날짜) 컬럼을 변환하여 처방월(YYYY-MM) 형태로 생성하였다.
- 불필요한 칼럼(기준년도, 가입자 일련번호 등) 제거
- 주성분코드가 한 자리만 포함된 이상치를 제거하였다.
- 1회투약량, 1일투약량, 총투여일수 값이 0 이하인 데이터를 제거하였다.
- 성별코드, 연령대코드, 시도코드 등의 범주형 데이터는 정수형으로 변환하였다.

3. 데이터 병합 및 저장

- 전처리된 2019~2021년 데이터를 하나의 데이터셋으로 병합하였다.
- 병합된 데이터를 processed_prescription_2019_2021.csv 파일로 저장하였다.

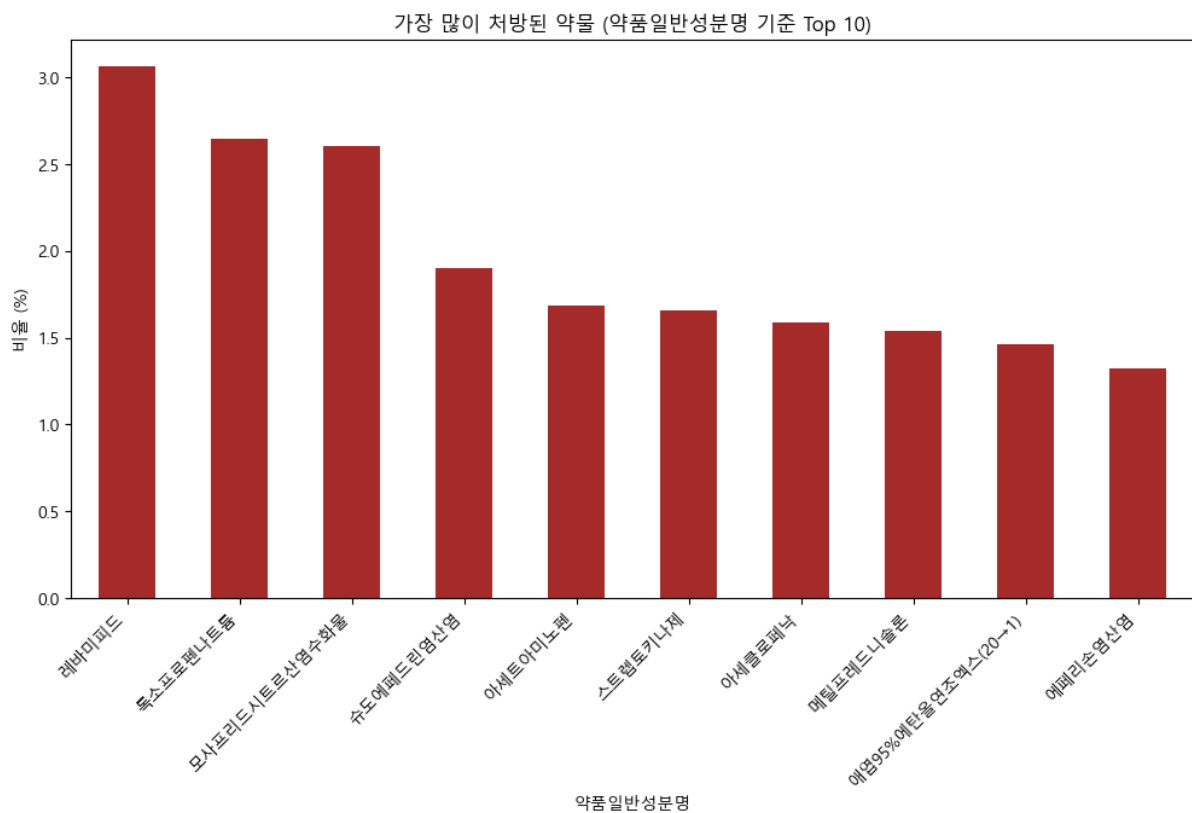
3. 데이터 분석 결과

본 장에서는 2019년부터 2021년까지의 의약품 처방 데이터를 분석한 결과를 정리한다. 분석은 가장 많이 처방된 주요 성분을 도출하는 것부터 시작하여, 성별, 연령대, 지역별 처방 패턴을 중심으로 진행되었다. 또한 연도별 및 월별 처방 트렌드를 확인하여 특정 약물의 계절성과 환자군별 특성을 파악하고, 향후 처방 경향을 예측할 수 있는 기반을 마련하였다.

3.1 가장 많이 처방된 약물

분석 결과

- 전체 처방 데이터에서 가장 많이 처방된 주성분 Top 10을 선정하여 비교 분석하였다.
- 성별, 연령대, 지역별로 가장 많이 처방된 약물 순위를 도출하고, 각 환자군에서의 차이를 분석하였다.
- 처방 비율을 백분율로 계산하여 특정 성분의 처방 비중이 얼마나 높은지 확인하였다.



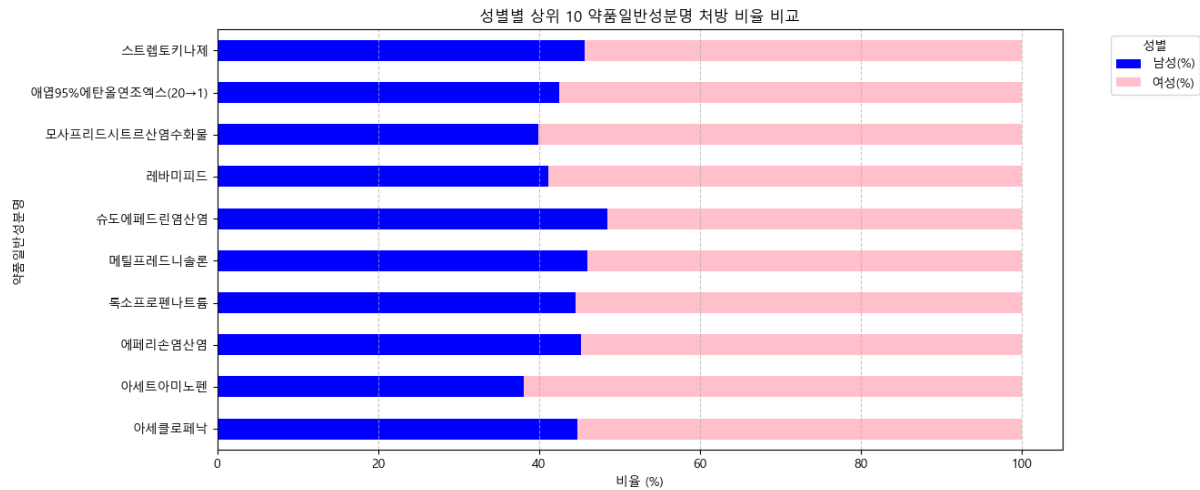
분석 결과를 보면, **소화기 관련 약물**(레바미피드, 모사프리드)과 **진통/소염제**(록소프로펜, 아세클로페낙, 아세트아미노펜, 메틸프레드니솔론)가 주요하게 처방되고 있음을 알 수 있다. 이는 만성 소화기 질환이나 근골격계 질환이 흔한 질환이라는 점을 시사한다.

또한, **혈전용해제**(스트렙토키나제)와 **비충혈 완화제**(슈도에페드린염산염)가 포함되어 있어, 혈액 순환 개선 및 감기 관련 증상 완화를 위한 처방도 주요한 비중을 차지하고 있음을 확인할 수 있다.

3.2 성별별 처방 약물 비교

분석 결과

- 성별에 따라 처방되는 약물의 차이를 분석하였다.
- 전체 처방량에서 남성과 여성 각각의 비율을 계산하여 비교하였다.
- 진통제 및 소화기 관련 약물이 성별에 따라 처방 비율에 차이가 있는지 검토하였다.



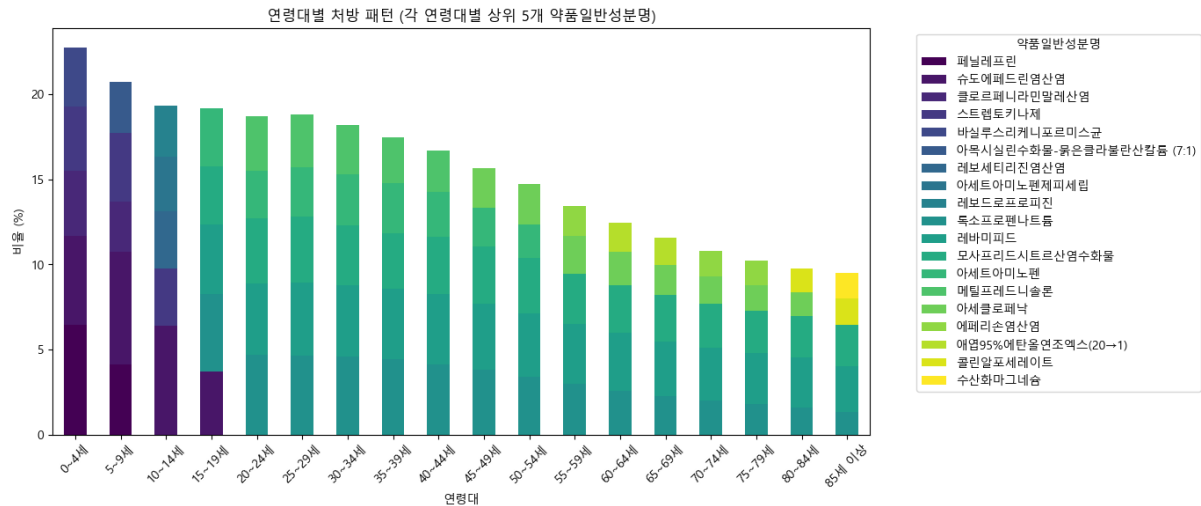
분석 결과, 대부분의 주요 약물에서 여성이 더 높은 처방 비율을 보였다. 특히 아세트아미노펜 (61.96%)의 경우 여성의 처방 비율이 매우 높았으며, 이는 생리통 완화 목적으로 널리 사용되기 때문으로 추정된다. 또한 모사프리드(60.11%), 레바미피드(58.87%), 에엠펜95%에탄올연조엑스 (57.54%) 등도 여성의 처방 비율이 높은 것으로 나타났다.

반면, 남성이 상대적으로 높은 처방 비율을 보인 약물은 슈도에페드린염산염(48.50%)과 메틸프레드니솔론(45.97%) 등이 있었다. 이는 감기 및 비출혈 완화 목적의 처방이 남성에게 조금 더 빈번하게 이루어질 가능성을 시사한다.

3.3 연령대별 처방 약물 비교

분석 결과

- 연령대별로 주요하게 처방되는 상위 5개 약물의 분포를 분석하였다.
- 특정 연령대에서 선호되는 약물과 연령 증가에 따른 처방 패턴 변화를 확인하였다.
- 처방 약물의 다양성이 연령대에 따라 어떻게 변하는지 분석하였다.



분석 결과, 연령대별 처방 약물에는 뚜렷한 차이가 존재하였다. 0~9세 영유아 연령대에서는 페닐레프린, 슈도에페드린, 클로르페니라민과 같은 감기약 및 비충혈 완화제가 주요하게 처방되었다. 이는 해당 연령대에서 감기나 호흡기 질환이 흔하게 발생하는 것과 관련이 있을 것으로 보인다.

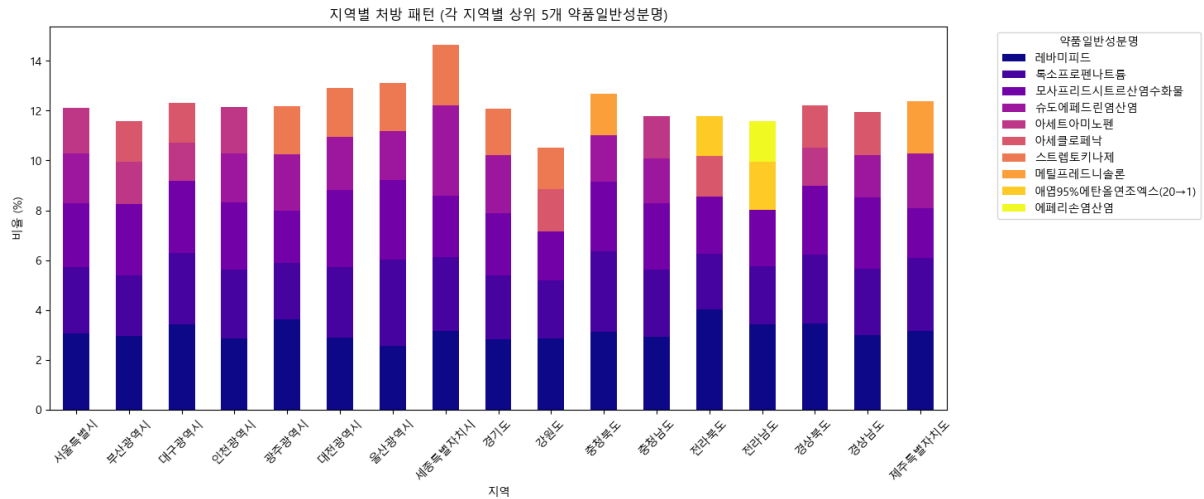
반면, 50대 이후 연령대에서는 레바미피드, 모사프리드, 아세트아미노펜, 메틸프레드니솔론 등의 소화기 관련 약물과 진통제가 더 높은 처방 비율을 보였다. 특히 85세 이상 연령대에서는 수산화마그네슘과 같은 변비 치료제가 주요하게 처방되며, 이는 노년층에서 소화기계 기능 저하와 관련이 있는 것으로 해석할 수 있다.

또한, 연령대가 어릴수록 상위 5개 약물의 비율이 전체 처방에서 차지하는 비중이 높았다. 즉, 특정 약물이 해당 연령대에서 집중적으로 처방된다는 것을 의미한다. 하지만 연령이 증가할수록 처방되는 주요 약물의 비중이 줄어들며, 이는 연령이 높아질수록 처방 약물의 종류가 더 다양해진다는 것을 보여준다. 이는 노년층에서 만성 질환이나 복합적인 건강 문제로 인해 다양한 약물이 처방되는 경향과 연관이 있을 것으로 해석된다.

3.4 지역별 처방 약물 비교

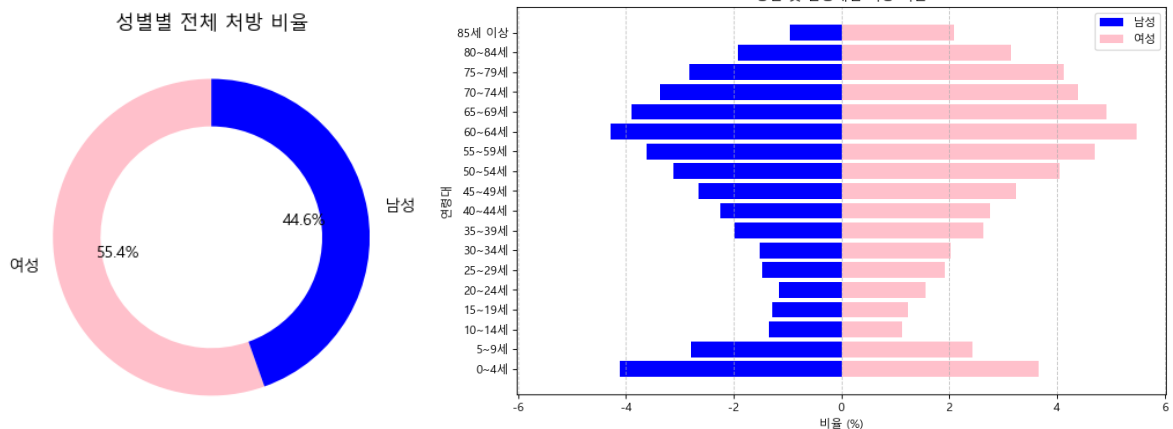
분석 결과

- 지역별로 주요하게 처방되는 상위 5개 약물을 비교하였다.
- 지역 간 주요 처방 약물의 차이가 있는지 분석하였다.
- 처방 약물의 다양성과 특정 지역에서의 처방 패턴을 확인하였다.



분석 결과, 지역별 처방 약물의 패턴에는 큰 차이가 없었다. 대부분의 지역에서 레바미피드, 록소프로펜나트륨, 모사프리드시트르산염수화물, 아세트아미노펜 등의 약물이 공통적으로 많이 처방되었다. 이는 전국적으로 주요 처방 약물이 크게 차이가 나지 않으며, 표준화된 의료 가이드라인에 따라 처방이 이루어지고 있음을 시사한다. 전반적으로 지역별 처방 패턴은 대체로 유사하며, 이는 한국의 보건의료 시스템이 전국적으로 균일하게 운영되고 있다는 점을 보여준다.

3.5 성별/연령대/지역별 처방 비율 분석



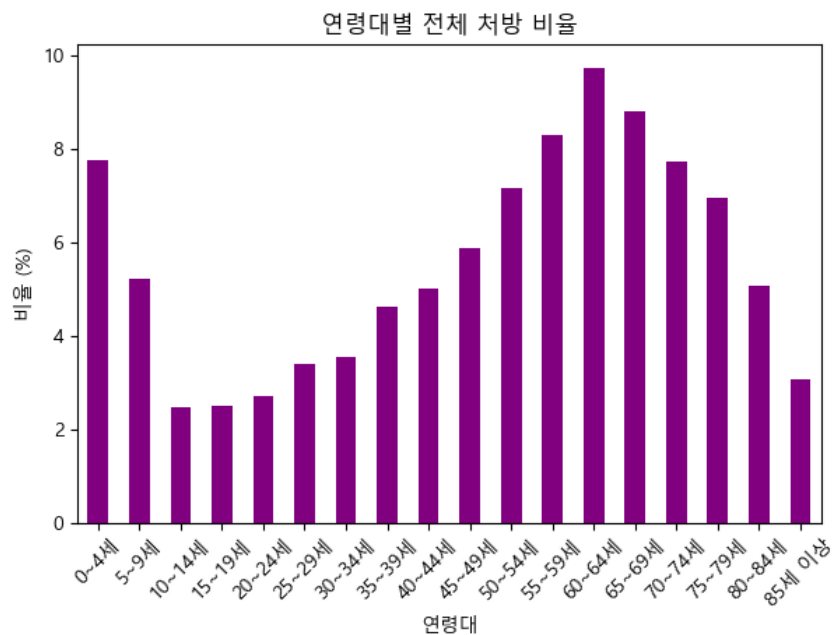
3.5.1 성별 처방 비율 분석

분석 결과, 여성이 전체 처방량의 55.4%를 차지하며, 남성(44.6%)보다 더 높은 처방 비율을 보였다.

이는 여성의 생리통, 골다공증, 철분 결핍 등의 관리와 관련된 의료 이용 빈도가 높은 것과 관련이 있을 가능성이 있다.

성별 및 연령대별 처방 패턴을 나타내는 피라미드 그래프를 보면, 영유아기와 아동기(0~14세)를 제외한 대부분의 연령대에서 여성의 처방 비율이 남성보다 높게 나타났다.

이는 여성의 건강 관리 및 질환 예방 목적의 의료 이용 빈도가 전반적으로 높을 수 있음을 시사한다.



3.5.2 연령대별 처방 비율 분석

연령대별 전체 처방 비율을 살펴보면, 영유아 및 아동기 때 처방 비율이 높게 나타났으며, 청소년기부터 처방 비율이 감소하는 경향을 보였다.

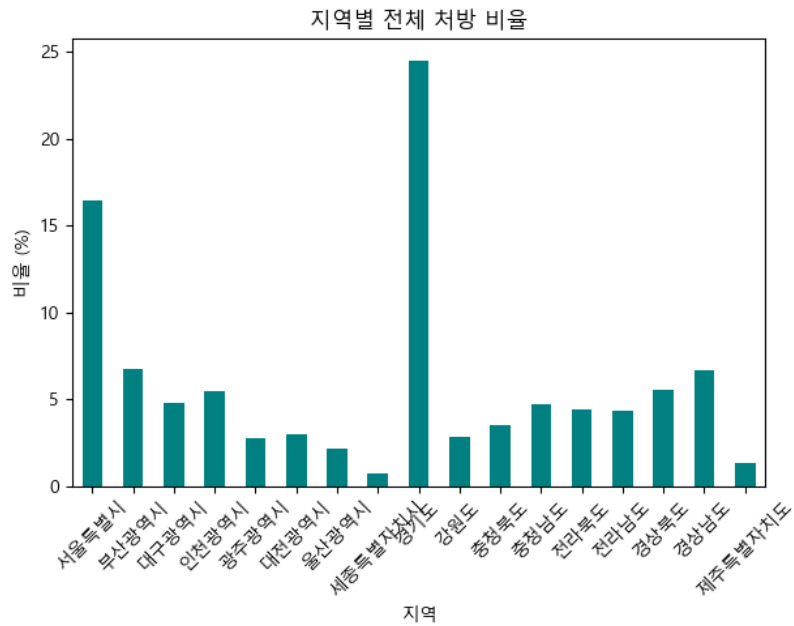
이는 영유아 및 아동기에 감염성 질환(감기, 폐렴 등)과 면역력 관련 질환으로 인해 의료기관을 자주 방문하는 경향이 반영된 결과로 해석할 수 있다.

이후 청소년기(15~19세)와 20대 초반(20~24세)에는 처방 비율이 가장 낮아지는 경향을 보였으며, 이후 연령이 증가함에 따라 처방 비율도 점진적으로 증가하는 패턴을 나타냈다.

특히, 50~60대 연령층에서 처방 비율이 가장 높았으며, 이는 고혈압, 당뇨병, 관절염 등의 만성 질환 관리와 관련이 있을 가능성이 높다.

70대 이후부터는 처방 비율이 다소 감소하는 경향을 보이지만, 여전히 높은 수준을 유지하고 있다.

연령이 증가할수록 질병의 관리와 치료를 위한 처방 빈도가 증가한다는 점을 확인할 수 있다.



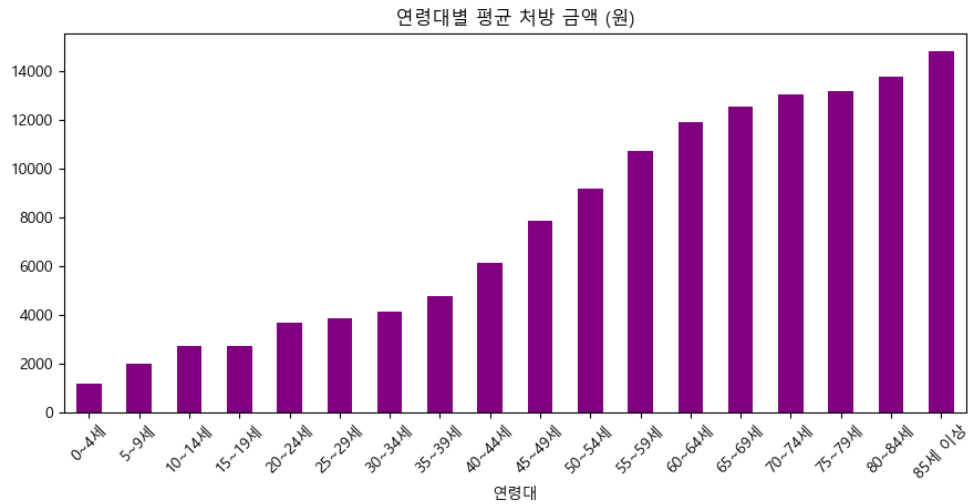
3.5.3 지역별 처방 비율 분석

지역별 전체 처방 비율을 비교한 결과, 경기도에서 가장 높은 처방 비율을 보였으며, 다음으로 서울 특별시와 인구 규모가 큰 광역시(부산, 대구, 인천 등)가 상대적으로 높은 처방 비율을 나타냈다.

전반적으로 **지역별 처방량은 인구 규모와 비례**하는 경향을 보였으며, 이는 지역별 의료 이용 빈도나 건강보험 청구 데이터의 규모가 인구 크기에 따라 달라지는 것을 반영한 것으로 보인다.

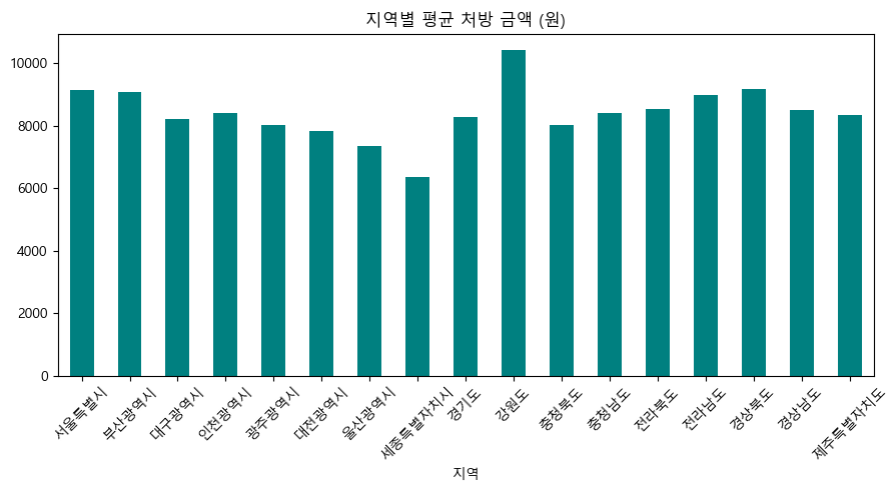
3.6 평균 처방 금액 비교

처방 데이터의 평균 처방 금액을 분석한 결과, 전반적인 **평균 처방 금액은 약 8,400원** 정도로 나타났다.



3.6.1 연령대별 평균 처방 금액

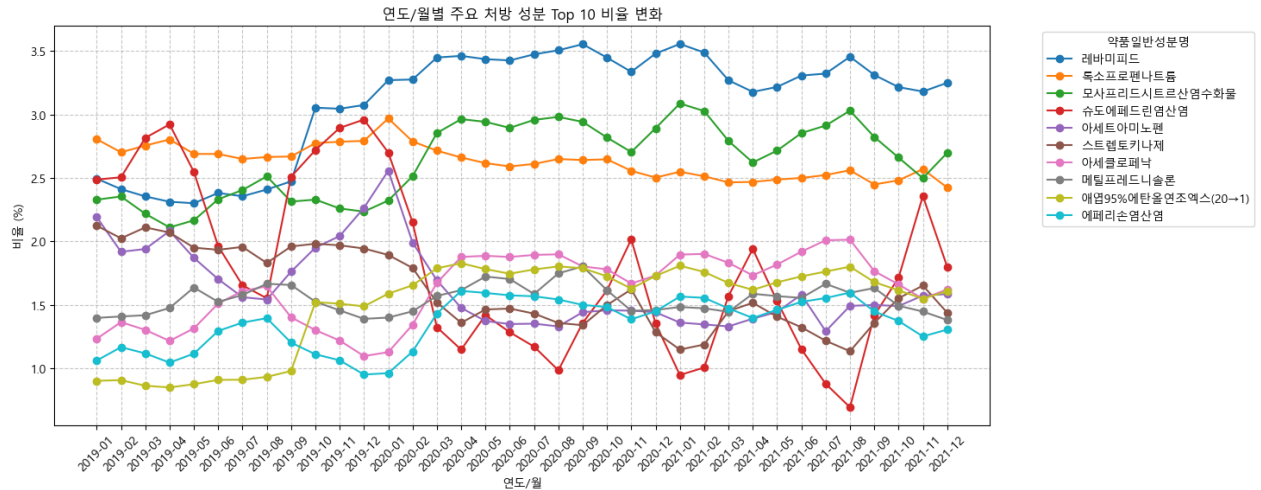
연령대별 평균 처방 금액을 살펴보면, **연령이 증가할수록 처방 금액도 증가**하는 추세를 보였다. 이는 고령층에서 만성질환 관리 및 복합 처방이 많아지는 경향과 관련이 있는 것으로 해석된다.



3.6.2 지역별 평균 처방 금액

지역별로는 대부분 유사한 처방 금액을 기록했으나, 세종특별자치시에서 평균 처방 금액이 유독 낮은 특징을 보였다. 이는 세종시의 평균 연령이 낮다는 점이 영향을 미친 것으로 보인다.

3.7 연도/월별 주요 처방 성분 Top 10 비율 변화



분석 결과

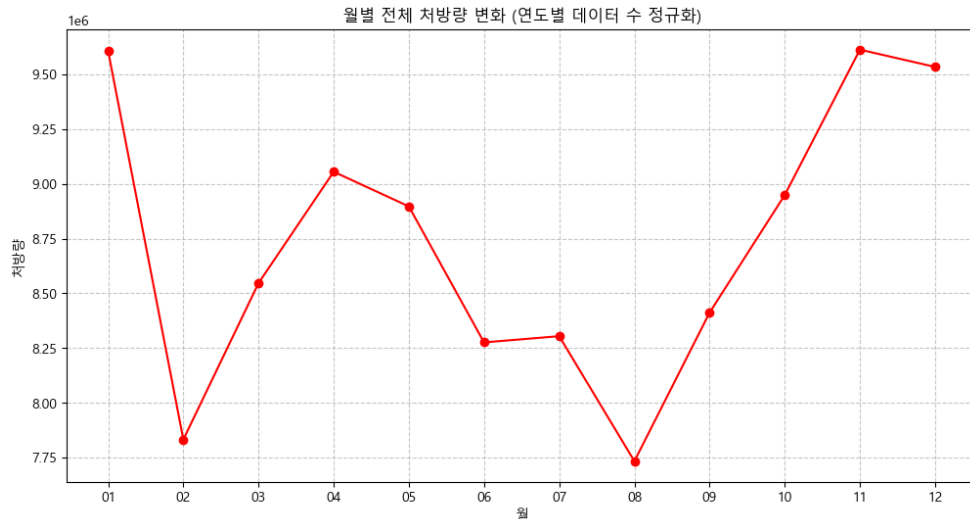
- 2019년부터 2021년까지 각 월별 주요 처방 성분 Top 10의 비율 변화를 분석하였다.
- 대부분의 주요 처방 성분(레바미피드, 독소프로펜나트륨, 모사프리드 등)은 월별 처방 비율이 비교적 일정하게 유지되었다.
- 특정 약물은 계절성 영향을 크게 받아 처방 비율이 주기적으로 변하는 경향을 보였다.

처방 비율이 주기적으로 변하는 경향을 보이는 주요 처방 성분에 슈도에페드린염산염이 있다. 슈도에페드린염산염은 감기약 및 비염 완화제로 사용되며, 계절성 알레르기와의 관련이 깊다. 처방 비율이 봄철과 가을철에 증가하는 경향을 보였으며, 이는 계절성 알레르기 및 감기 유행과 관련이 있는 것으로 해석할 수 있다.

또한, 2019년 말 코로나19 발생 이후 마스크 착용이 일반화되면서, 감기약의 처방 비율이 이전보다 낮아지는 경향을 보였다. 이는 마스크 착용과 사회적 거리두기로 인해 감염성 질환(감기, 독감 등)의 발생이 감소했기 때문으로 분석된다.

반면, 위장 보호제 및 소화기 관련 약물(레바미피드, 모사프리드)의 처방 비율은 계절성과 큰 관계 없이 일정한 수준을 유지하는 모습을 보였다. 이는 만성적인 위장 질환 관리와 관련된 처방이 지속적으로 이루어지기 때문으로 보인다.

3.8 월별 전체 처방량 변화 분석



분석 결과

- 월별 전체 처방량 변화를 분석하기 위해 연도별 데이터를 정규화한 후, 동일한 월의 처방량을 합산하여 추이를 살펴보았다.
- 특정 연도별 처방량 차이를 비교하는 것은 데이터 신뢰성 문제로 인해 배제하고, 월별 패턴을 중심으로 분석을 진행하였다.

전반적으로 계절성 요인에 따른 처방량 변화가 일정 부분 영향을 미치는 것으로 판단된다. 특히 **환절기인 2월 말~4월 초와 8월 말~10월 초 구간에서 처방량이 증가하는** 모습을 보인다.

3.9 약물군별 처방 비율 변화 분석

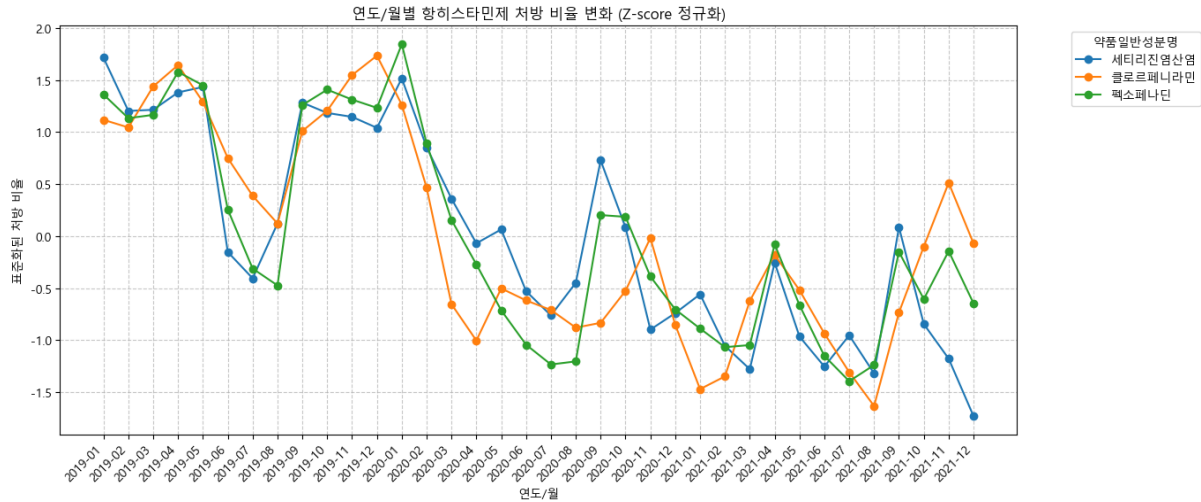
분석 목적

동일한 약물군에 속하는 성분들이 처방 비율 변화에서 유사한 추세를 보이는지 비교하고, 이를 통해 특정 시기의 처방 경향을 분석하는 것이 목표다. 비교를 위해 계절성 질환 치료에 사용되는 **항히스타민제**, 여름철 주로 발생하는 **항진균제**, 그리고 동일한 계열에 속하는 **세팔로스포린계 항생제**의 처방 비율 변화를 비교하였다.

분석 방법

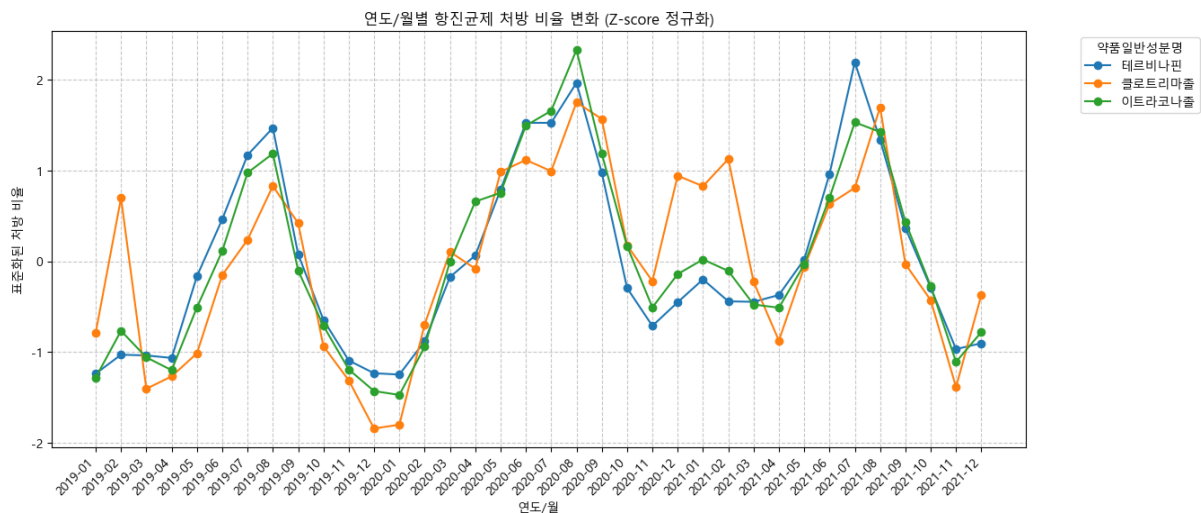
- 약물별 처방량을 연도/월별로 집계하여 **Z-score 정규화**를 적용하여 스케일 차이를 조정하였다.
- 이를 통해 절대 처방량이 다르더라도 각 성분이 시간에 따라 변화하는 패턴을 비교할 수 있도록 하였다.

3.9.1 항히스타민제 처방 비율 변화



- 분석 대상: 세티리진염산염, 클로르페니라민, 펙소페나딘
- 분석 결과:
 - 봄(4월)과 가을(10~11월)에 처방률이 증가하는 패턴을 보인다.
 - 2020년 4월에는 처방률 증가가 크게 나타나지 않음 → 이는 코로나19 이후 미세 입자 차단 마스크 착용 증가로 인해 계절성 알레르기 발병률이 줄어든 영향으로 추정됨.
 - 이후 전체적인 처방률이 감소하는 경향을 보이며, 이는 코로나 이후 사회적 거리두기 및 마스크 착용 습관의 지속적인 영향을 반영하는 것으로 보인다.

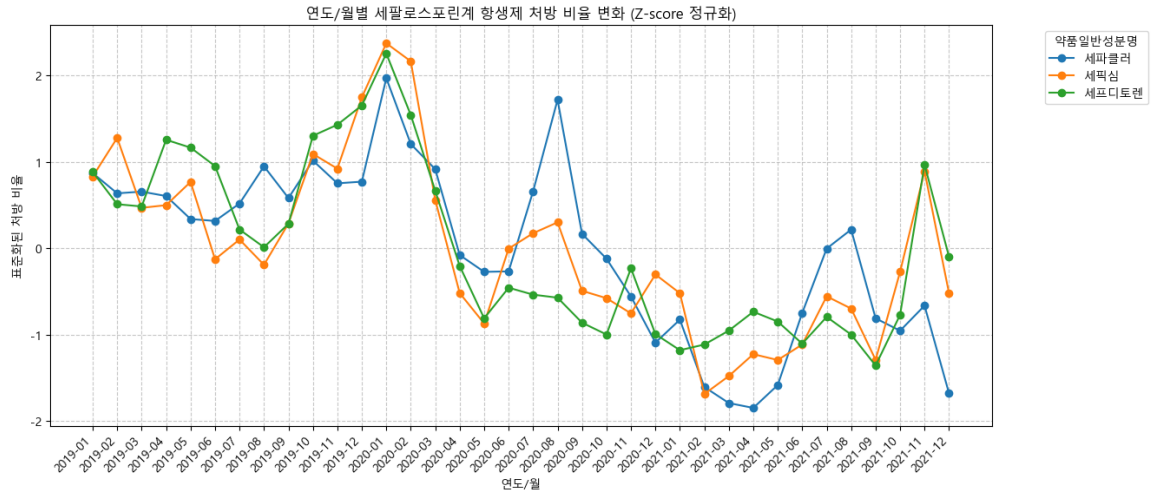
3.9.2 항진균제 처방 비율 변화



- 분석 대상: 테르비나핀, 클로트리마졸, 이트라코나졸

- 분석 결과:
 - 뚜렷한 계절성을 보이며, 매년 8월(여름철)에 처방률이 증가하는 경향이 관찰됨.
 - 여름철 높은 온도와 습도로 인해 피부 곰팡이 감염증이 증가하면서 해당 약물들의 처방이 급증하는 것으로 해석할 수 있다.

3.9.3 세팔로스포린계 항생제 처방 비율 변화



- 분석 대상: 세파클러, 세픽심, 세프디토렌
- 분석 결과:
 - 처방률이 크게 요동치는 항히스타민제나 항진균제와 달리, 세팔로스포린계 항생제는 비교적 완만한 변화 패턴을 보인다.
 - 하지만, 비슷한 시기에 처방 비율이 동반 증가하거나 감소하는 경향이 관찰됨 → 동일한 계열의 항생제들이 대체적으로 같은 시기에 사용되는 경향이 있음을 시사.

4. 데이터예측 모델 성능 분석

본 장에서는 제공된 데이터를 바탕으로 미래의 처방률을 예측하는 다양한 모델을 사용하고, 실제 데이터와 비교하여 각 모델의 성능을 평가하는 과정을 다룬다. 처방률 예측의 대상은 **항진균제의 일종인 이트라코나졸(itraconazole)**이며, 2019년부터 2021년까지의 처방 데이터를 바탕으로 2022년의 1년간 처방률을 예측하는 실험을 수행하였다.

4.1 평가 데이터 준비

예측 모델을 학습하기 위해 2019년부터 2021년까지의 이트라코나졸 처방 데이터를 학습 데이터로 활용하였다. 그러나, 2022년의 처방 데이터는 주성분 코드가 아닌 ATC 코드로 표기되어 있어 기존 데이터와의 일치성이 맞지 않았다. 이에 따라, **주성분 코드와 ATC 코드 간의 매핑을 수행하여 2022년의 실제 처방 데이터를 정리하는 과정**이 필요했다. 이를 통해 예측 모델의 결과를 검증할 수 있도록 평가 데이터를 구축하였다.

4.2 예측 모델 선정

시계열 예측을 위해 대표적인 세 가지 모델을 선정하였다.

1. ARIMA(AutoRegressive Integrated Moving Average)

- 과거 데이터의 자기상관성을 활용하여 미래 값을 예측하는 통계적 모델
- 단기 및 장기 추세를 반영하는 데 효과적

2. Prophet(Facebook Prophet)

- 계절성 및 추세를 자동으로 학습하는 시계열 모델
- 주기적인 패턴을 분석하는 데 강점

3. LSTM(Long Short-Term Memory Neural Network)

- 딥러닝 기반 순환 신경망(RNN) 모델로, 장기 의존성을 고려한 학습 가능
- 비선형적인 패턴까지 포착할 수 있으나 데이터량이 많아야 효과적

4.3 성능 평가 기준

모델 성능을 평가하기 위해 다음과 같은 지표를 사용하였다.

- MAE(Mean Absolute Error): 예측값과 실제값의 절대 오차 평균
- RMSE(Root Mean Squared Error): 오차 제곱의 평균을 루트 씌운 값으로, 오차가 클수록 패널티를 부여하는 지표

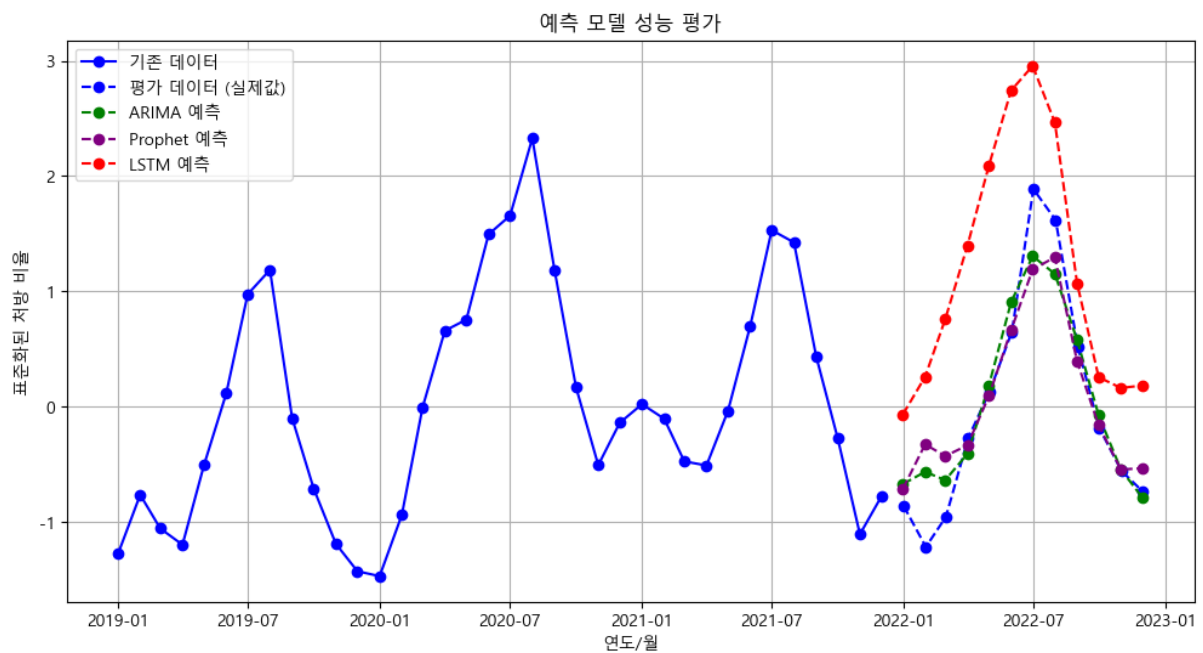
4.4 모델 성능 비교 결과

모델별 예측 성능은 다음과 같다.

모델명	MAE	RMSE	순위
ARIMA	0.2409	0.3196	1
Prophet	0.2561	0.3816	2
LSTM	1.1883	1.3069	3

- ARIMA 모델이 가장 낮은 MAE(0.2409)와 RMSE(0.3196)를 기록하며 가장 정확한 예측 성능을 보였다.
- Prophet 모델은 ARIMA보다 다소 높은 오차를 보였지만, 계절성과 추세를 반영하는 데 유용하였다.
- LSTM 모델은 상대적으로 높은 오차를 기록하였으며, 이는 학습 데이터가 충분하지 않았기 때문으로 판단된다.

4.5 예측 결과 시각화



다음 그래프는 각 모델이 예측한 2022년 이트라코나졸 처방률과 실제 데이터를 비교한 것이다.

1. 실제 처방 데이터(파란색)와 모델 예측값 비교
2. ARIMA 모델(초록색), Prophet 모델(보라색), LSTM 모델(빨간색)
3. ARIMA가 실제 데이터와 가장 유사한 패턴을 보이며, Prophet은 다소 높은 편차를, LSTM은 큰 오차를 나타냄

5. 결론 및 향후 활용 방안

5.1 프로젝트 결론

본 프로젝트는 의료 데이터 분석을 통한 처방 패턴 탐색과 처방 예측 모델 성능 분석을 목표로 진행되었다. 특히, 가장 많이 처방된 약물의 트렌드를 파악하고, 특정 약물군(예: 항히스타민제)의 계절적 처방 변화를 분석하였으며, 머신러닝 기반 예측 모델을 적용하여 향후 처방률을 예측하는 과정까지 수행하였다.

프로젝트를 통해 다양한 데이터 분석 기법을 익히고, 의료 데이터의 구조와 활용 가능성에 대한 깊은 이해를 쌓을 수 있었다. 특히, 의료 데이터 처리 과정과 시각화를 경험할 수 있는 기회였다. 또한, 추가적으로 데이터 시각화 툴 사용 능력을 기르기 위해 Power BI를 활용하여 데이터 시각화를 진행하며, 복잡한 데이터를 직관적으로 해석할 수 있도록 구성하는 과정도 수행하였다.

시계열 예측 실험에서는 ARIMA 모델이 가장 우수한 성능을 보였으며, Prophet 모델은 일정한 패턴을 따르는 데이터에서 활용 가능성이 있음을 확인하였다. 반면, LSTM 모델은 데이터 양이 충분하지 않거나 시계열 특성을 명확히 학습하지 못하는 경우 성능이 저하될 수 있음을 확인할 수 있었다.

본 프로젝트를 통해 의료업계의 처방 현황을 데이터 기반으로 분석할 수 있는 가능성을 탐색하였으며, 다양한 분석 방법을 적용하여 실무에서 활용 가능한 데이터 분석 및 의사결정 지원 시스템 구축의 필요성을 인식하게 되었다.

5.2 향후 프로젝트 활용 방안

본 프로젝트를 바탕으로 향후 의료 데이터 분석 및 자동화 모델 개발을 위한 여러 방향을 고려할 수 있다.

1. 대화형 AI 기반 데이터 분석 시스템 개발

- 인공지능 API와 결합하여 사용자가 원하는 데이터를 대화형으로 요청하면 자동으로 데이터 분석과 보고서 작성을 수행하는 모델을 구축하고자 한다.
- 이를 통해 사용자는 별도의 코드 작성 없이 자연어 질의를 통해 데이터 분석 결과를 확인할 수 있으며, 신속한 의료 데이터 활용이 가능할 것으로 기대된다.

2. 추가적인 머신러닝 기법 적용 및 성능 개선

- 이번 연구에서는 시계열 예측 모델(ARIMA, Prophet, LSTM)을 중심으로 분석하였으나, 향후 XGBoost, Transformer 기반 모델 등 다양한 기법을 적용하여 성능

을 개선할 수 있다.

3. 실시간 데이터 분석 및 모니터링 시스템 구축

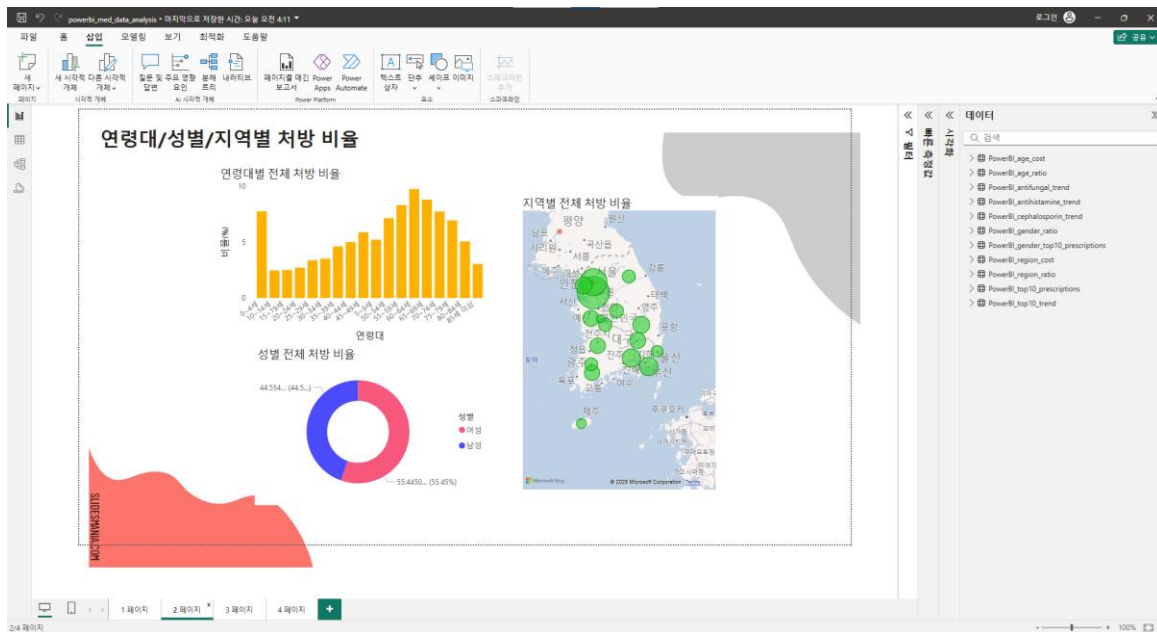
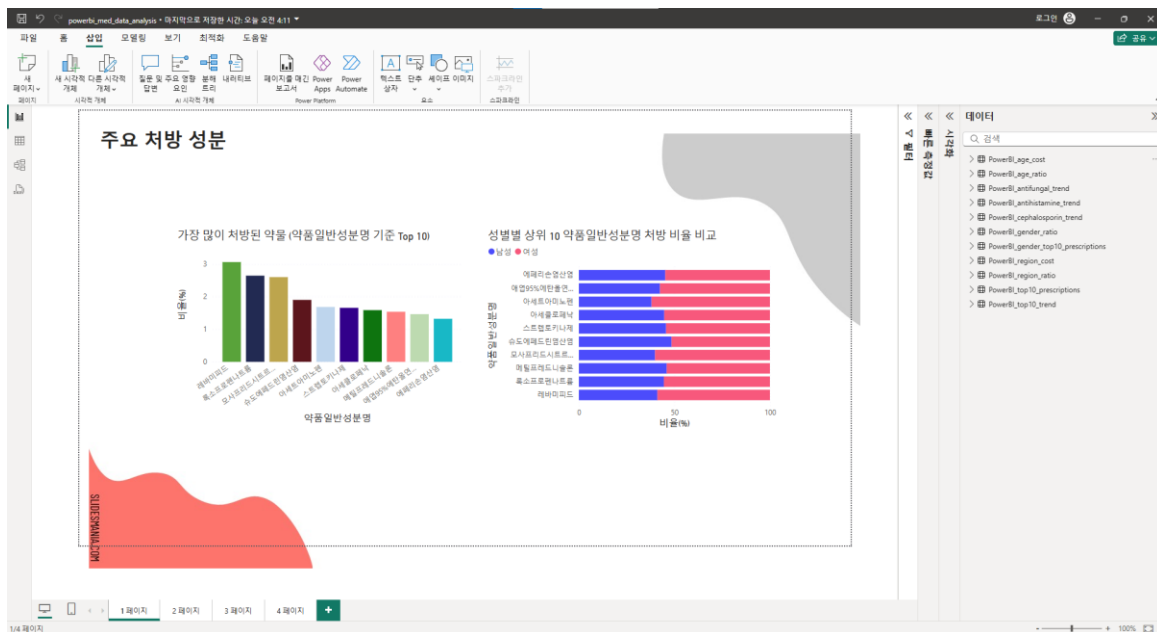
- 현재 프로젝트에서는 2019~2021년 데이터를 활용하였지만, 실시간 데이터 스트리밍을 적용하여 지속적인 처방 패턴 모니터링이 가능하도록 발전시킬 수도 있다.

<부록>

Power BI를 활용한 데이터 시각화

본 프로젝트에서는 Python의 matplotlib.pyplot (plt)을 이용하여 데이터 분석과 시각화를 수행하였으며, 이를 통해 예측 모델 성능 분석, 처방 패턴 분석, 성별 및 연령대별 처방 비율 비교, 지역별 처방 특성 탐색 등의 다양한 분석을 진행하였다.

그러나, 데이터 시각화 및 보고서 작성 과정에서 보다 직관적인 인터페이스를 제공하고 현업에서 주로 사용되는 데이터 분석 툴인 Power BI를 추가적으로 활용하여 데이터 시각화를 보완하였다.



데이터 전처리 코드 링크 : https://chungjuwon.github.io/prescription-data-analysis/data_preprocessing.html

데이터 분석 코드 링크 : https://chungjuwon.github.io/prescription-data-analysis/EDA_Prescription_Analysis.html

데이터 예측 모델 성능 분석 코드 링크 : https://chungjuwon.github.io/prescription-data-analysis/timeseries_forecasting_models.html