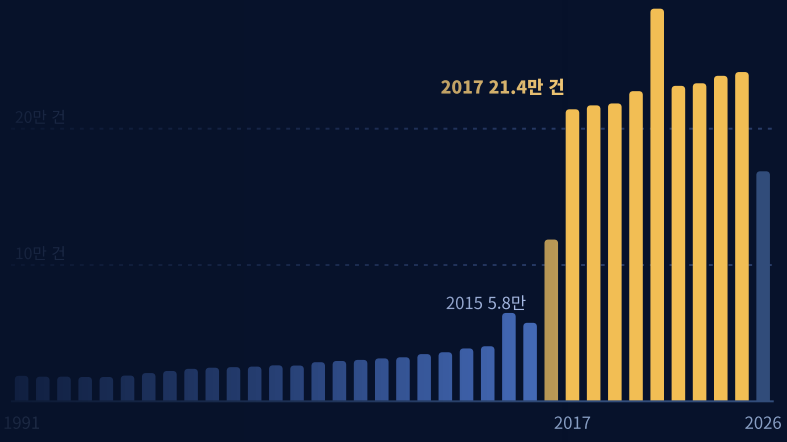
병원 데모 데이터 · 환자 57,188명 · 진료 340만 건 · 36년

CUBEAGENT · USE CASE BOOK · HEALTHCARE

CubeAgent 활용 사례집 — 병원 데이터 편

진료 기록에 문장으로 묻다.



진료·진단·처방·보험 청구가 얽힌 병원 데이터.
원무·심사·QI 담당자가 SQL 없이 질문만으로
분석한 과정을 **실제 조회 결과**로 보여드립니다.

57,188명	340만 건	198만 건	264만 건	$92.4억
환자	진료	진단	처방	총 청구금액

연도별 진료 건수 1991~2026 (2026년은 9월까지)

이 사례집의 흐름

한 문장 조회에서 스스로 쌓는 분석 지식까지, 6단계

STEP 1 문장으로 추출·시각화
한 문장 → 표와 차트
· 자유로운 질문 p.2

STEP 2 대화로 좁혀 가기
맥락을 잇는 후속 질문
· 자유로운 질문 p.4

STEP 3 스스로 계획하는 복합 분석
"왜?"를 단계로 분해
· 복잡한 분석 p.5

STEP 4 정의된 숫자, 검증되는 답
이상치·범위 점검
· 정확한 숫자 p.6

STEP 5 스스로 찾고 검증하는 탐색
가설 수립과 기각
· 더 깊은 인사이트 p.7

STEP 6 쌓이는 분석 지식
인사이트 저장·재사용
· 더 깊은 인사이트 p.8

본 사례집의 모든 수치와 차트는 FORESOFT 데모 환경의 병원 샘플 데이터(실제 환자가 아닌 데모용 진료 기록)에서 CubeAgent가 실제로 조회한 결과입니다. 금액 단위 $는 원천 데이터의 USD를 그대로 표기했습니다. 의료기관 45,454곳, 의료진 45,786명, 보험자 10곳의 기록이 포함돼 있습니다.

FORESOFT

STEP 1 문장으로 추출·시각화 · 자유로운 질문

원무 리포트 요청 대신, 한 문장이면 표와 차트가 나옵니다

"진료유형별 청구금액" 하나에도 진료·보험 청구 테이블을 조인하고, 유형 코드를 해석하고, 건수와 금액을 따로 집계해 화면을 만들어야 했습니다.

지금까지의 방식 BEFORE

1. **요건 정리** 원무·심사팀 요청서 접수
2. **테이블·코드 확인** 진료·청구 테이블, 유형 코드 파악
3. **SQL 작성·결과 검증** 청구금액 정의 확인, 수치 대조
4. **BI 화면 개발** Spotfire 시각화·배포
5. **수정 요청 반복** "입원만 진료사유별로 볼 수 있을까요?"

작성해야 하는 쿼리 (예시)

```
SELECT e.encounter_class,
       COUNT(*)              AS 진료건수,
       SUM(e.total_claim)    AS 청구금액,
       AVG(e.total_claim)    AS 건당금액
FROM   encounters e
LEFT   JOIN payers p
       ON e.payer_id = p.id
GROUP  BY e.encounter_class
ORDER  BY 청구금액 DESC;
```

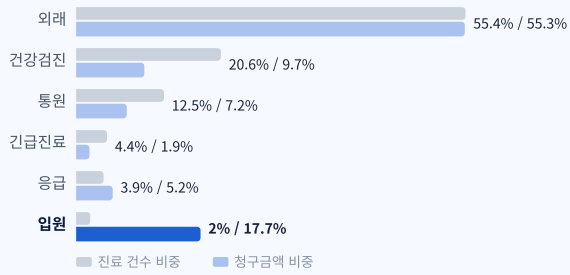
'청구금액'이 기본 진료비인지 총 청구액인지부터 담당자마다 다를 수 있습니다.

CubeAgent AFTER

나 **진료유형별 진료 건수와 청구금액 보여줘**

의도 RANKING · 측정값 .진료 건수 · 측정값 .총 청구금액 · 차원 진료.진료유형

입원은 진료 건수의 2.0%지만 청구금액의 17.7%입니다. 건당 $24,216으로 외래($2,706)의 약 9배입니다. 금액 1위는 외래로 $51.0억(55.3%)입니다.



답변과 함께 남는 조회 정의 (실제 생성값)

```
{
  "intent": "RANKING",
  "measures": [".진료 건수", ".총 청구금액",
               ".평균 진료비", ".보험 보장률"],
  "dimensions": ["진료.진료유형"],
  "sort": [{ "field": ".총 청구금액",
             "direction": "desc" }],
  "ccq": { "chartType": "bar", "xAxis": "진료.진료유형",
           "yAxis": ".총 청구금액", "showLabels": true }
}
```

SQL이 아니라 '업무 언어로 된 조회 정의'

CubeAgent는 질문을 측정값·차원·의도로 구조화하고, 시맨틱 모델이 이를 실행합니다. 원무 담당자도 SQL을 읽지 않고 어떤 기준으로 답했는지 바로 검토할 수 있습니다.

'.평균 진료비', '.보험 보장률' 같은 파생 지표도 모델에 한 번 정의되어 있어, 누가 물어도 같은 계산식으로 답합니다.

1 테이블 구조를 몰라도
진료·진단·처방·보험자 테이블의 관계는 시맨틱 모델이 알고 있습니다.

2 지표 정의는 모델이 보장
보장률·환자 부담 추정액의 계산식은 한 곳에서 정의되고, 누가 물어도 같습니다.

3 근거까지 투명하게
어떤 측정값·차원·필터로 조회했는지 답과 함께 남습니다.

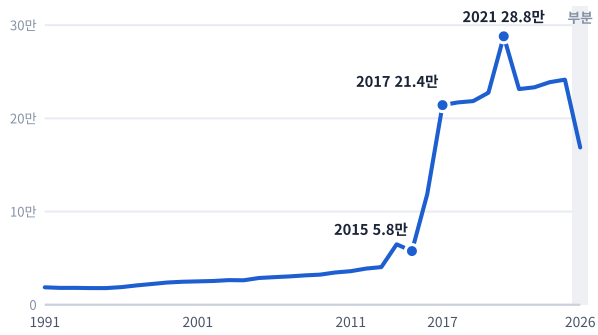
STEP 1 문장으로 추출·시각화 · 부서별 예시

원무·임상·보험, 묻는 방식은 같습니다

추이·순위·분포·비교. 분석 의도에 맞는 차트 형태도 CubeAgent가 고릅니다.

원무 TREND

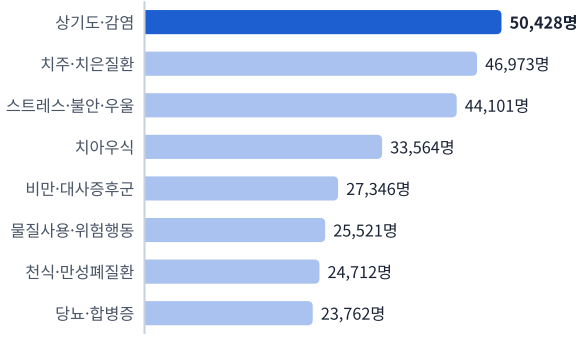
나 연도별 진료 건수 추이 보여줘



2015년 5.8만 건에서 2017년 21.4만 건으로 3.7배 뛴 뒤, 2025년까지 연 21~24만 건에서 안정됐습니다. 2021년은 28.8만 건으로 튀었습니다.

임상 RANKING

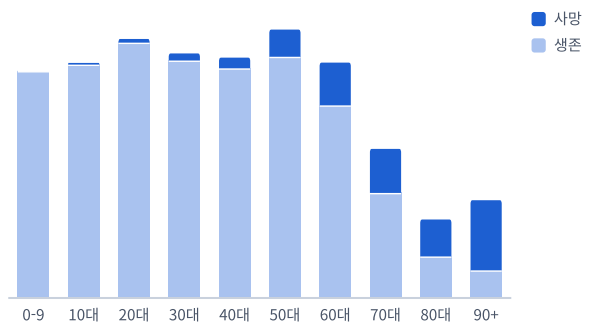
나 진단 환자가 많은 질환 순위는?



상기도 감염이 50,428명으로 1위입니다. 검진·고용 같은 비임상 코드는 '질환분석대상여부' 플래그로 자동 제외하고 임상 질환만 순위를 매겼습니다.

환자 구성 DISTRIBUTION

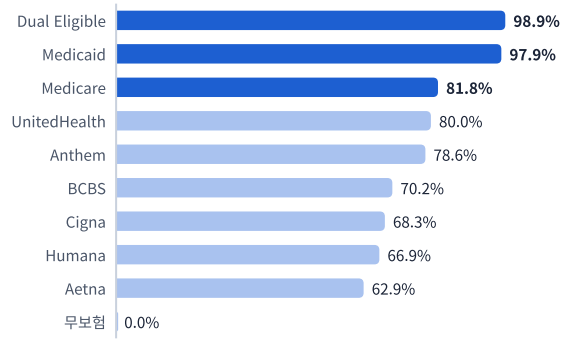
나 연령대별 환자 수와 사망 환자 수



60대부터 사망 환자 비중이 가파르게 오릅니다(18.7% → 70대 30.2% → 80대 48.0%). 90세 이상이 80대보다 많은 역전은 STEP 4에서 따로 점검합니다.

보험 심사 COMPARISON

나 보험자별 보장률 비교해줘



Medicaid·Dual Eligible은 98% 안팎, 민간은 63~80%입니다. Medicare는 81.8%입니다. 무보험(NO_INSURANCE) 진료 환자는 21,497명으로 보험자 중 가장 많습니다.

연도별 추이는 달력에 매핑된 1991~2026년만 표시했습니다(1991년 이전 진료 275,801건 제외, STEP 4 참조). 2026년은 9월까지의 부분 연도입니다.

STEP 2 대화로 좁혀 가기 · 자유로운 질문

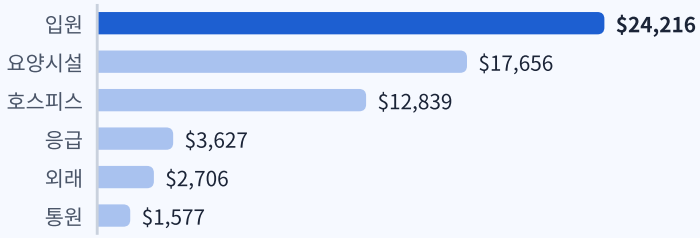
다음 질문이 떠오를 때마다 새 리포트를 요청하지 않아도 됩니다

CubeAgent는 앞선 대화의 조건(입원, 진료사유)을 이어받아 필터를 쌓아 가며 답합니다. 질문 세 개로 청구금액의 쏠림을 환자 단위까지 좁혀 봅니다.

나 **진료유형별로 건당 진료비가 얼마야?**

1

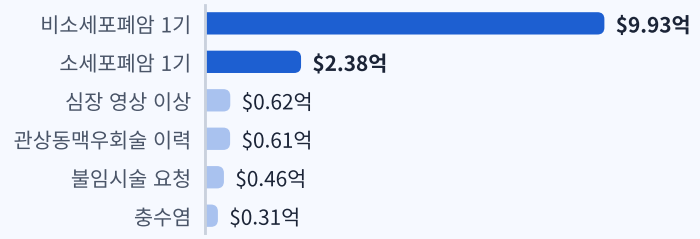
입원이 건당 $24,216으로 가장 비쌉니다.
요양시설 $17,656, 호스피스 $12,839가 뒤를 잇습니다. 입원 전체 청구는 $16.4억입니다.



나 **입원 청구금액은 어떤 진료사유에서 나와?**

2

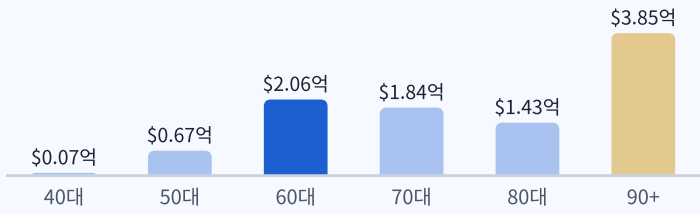
60.6%가 비소세포폐암 1기입니다.
$9.93억, 입원 13,831건. 소세포폐암 1기($2.38억)까지 합치면 **폐암 두 종류가 입원 청구의 75.1%**를 차지합니다.



나 **그 폐암 입원 환자는 몇 명이고, 어떤 연령대야?**

3

407명이 1인당 평균 34회 입원했습니다.
60대 85명($2.06억)이 가장 많고, 보장률은 50대 86.7%에서 80대 78.1%로 떨어집니다. **'90세 이상' 159명은 연령 산정에 문제가 있어 STEP 4에서 따로 확인합니다.**



POINT 질문 3개로 **진료유형 → 진료사유 → 환자**까지 좁혀, 입원 청구의 60%가 환자 407명에게 몰려 있다는 사실에 도달했습니다. 기존 방식이라면 리포트 3건을 요청해야 했을 흐름이며, 이상해 보이는 연령 구간은 에이전트가 먼저 짚었습니다.

STEP 3 스스로 계획하는 복합 분석 · 복잡한 분석

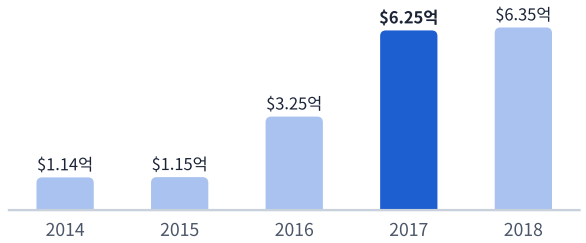
“왜?”라고 물으면 분석 단계를 스스로 설계합니다

나 **2017년에 청구금액이 갑자기 늘었던데, 어디서 온 거야?**

분석 계획 — ① 증가 규모와 시점 확인 → ② 진료유형별 기여 분해 → ③ 최대 기여 유형을 환자 수·1인당 건수·건당 금액으로 분해 → ④ 증가 이후 수준이 유지되는지 확인.

① 증가 규모 — 2015 → 2017

5.4배 $1.15억 → $6.25억
진료 57,592 → 214,239건



② 유형별 기여 — 증가분 $5.10억

76.5% 증가분의 4분의 3이 외래
(외래 +$3.91억, 그 외 +$1.20억)

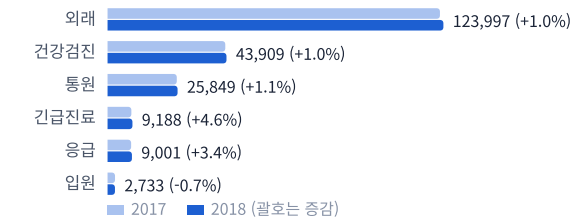


③ 외래 청구 11.9배 분해 — 2015 → 2017

진료 환자 수	1인당 진료	건당 청구액
4.15배	**1.12배**	**2.57배**
7,496 → 31,099명	3.52 → 3.95건	$1,353 → $3,473

4.15 × 1.12 × 2.57 = 11.9배. 환자 유입과 건당 금액 상승이 동시에 일어났고, 1인당 방문 횟수는 거의 그대로입니다.

④ 이후 수준 — 2017 vs 2018 진료 건수



모든 유형의 차이가 5% 이내입니다. 2017년에 새로운 수준으로 올라선 뒤 유지됐습니다.

결론: 2017년 증가는 ‘외래 환자 유입 + 건당 금액 상승’이 겹친 계단형 변화입니다.

환자 수와 건당 금액이 한 해에 동시에 바뀐 것은 진료 행태보다 **수가 체계나 데이터 수집 범위가 바뀌었을 가능성**을 시사합니다. 연도 비교 리포트는 2017년 이후 구간으로 나눠 볼 것을 제안하며, 원천 시스템 변경 이력을 확인해 볼 것을 권합니다.

POINT 질문 하나에 에이전트가 **4단계 분석 계획**을 세우고, 증가분을 유형별 기여와 **환자 수 × 1인당 건수 × 건당 금액**으로 분해해 곱이 맞는지 검산까지 했습니다. 숫자의 원인이 진료 행태인지 데이터 변화인지 구분해 볼 다음 질문도 함께 제안합니다.

STEP 4 정의된 숫자, 검증되는 답 · 정확한 숫자

합계가 이상하면 보고하기 전에 먼저 멈춥니다

병원 데이터에는 관리 코드, 범위 밖 날짜, 입력 오류가 섞여 있습니다. 합계를 그대로 옮기면 틀린 결론이 보고서에 실립니다. CubeAgent는 숫자의 크기와 범위를 먼저 점검합니다.

나 **약제비가 가장 많이 나간 약품 순위 보여줘**

그대로 옮긴 답

~~70.6~~ %

"알테플라제 주사가 전체 약제비의 70.6% ($121.6억)를 차지합니다."



전체 약제비 $172.3억 · 그 외 품목 합계 $50.7억

1위 품목의 금액은 데이터 이상치로 판단해 분리했습니다.

처방 1건당 금액
$2,312 만
526건 · 환자 526명 · 1회 주사

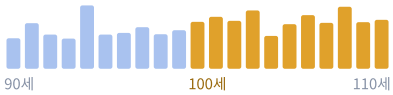
전체 약제비 vs 총 청구금액
$172 억 > $92억
약제비가 전체 청구보다 큼

이 품목을 빼면 1위는 **클로피도그렐 $7.1억(7,438명)**입니다. 원천 시스템의 단가·조제횟수 산식 확인을 권합니다.

이 사례집을 만들며 에이전트가 함께 찾아낸 점검 항목

범위 밖 날짜 8.1%
275,801 건
진료일이 달력(1991~) 밖이라 연도가 비어 있습니다. 연도별 합계에서 조용히 빠지지 않도록 따로 표기했습니다.

연령 산정 오류 58.1%
100~110 세
'90세 이상' 2,769명 중 1,610명이 100세 이상이고, 생존 110세도 34명입니다. 사망 후에도 기준일 연령이 계속 늘어난 것으로 보입니다.



통화 기호 불일치 ₩ → $
₩518,029,583
Medicare '.환자 부담 추정액'이 ₩로 표시되지만, 값은 청구액 $28.4억 × (1 − 보장률 81.8%)와 일치합니다. 실제 단위는 USD입니다.

이 답을 가능하게 한 시맨틱 모델의 역할

A 지표 정의 단일화
'.보험 보장률', '.환자 부담 추정액'의 산식이 모델에 있어, 표시 단위가 틀려도 값을 역산해 검증할 수 있습니다.

B 필터 사전 점검
조회 전에 필터 품질을 확인합니다. 예: '진료사유'는 41%가 비어 있어 (HIGH_NULL) 사유별 합계가 전체와 다릅니다.

C 크기·범위 상식 점검
부분이 전체보다 크거나, 1건 금액이 비현실적이거나, 구간이 역전되면 보고 전에 멈추고 밝힙니다.

구분	원천 테이블에 직접 SQL을 생성하는 방식	CubeAgent (시맨틱 모델 기반)
이상치	SUM 결과를 그대로 순위로 제시	건당 금액·전체 대비 비중으로 **상식 점검 후 분리**
날짜 범위	달력 조인에서 빠진 행이 조용히 누락	미매핑 건수를 **별도 행으로 드러냄**
관리 코드	검진·고용 코드까지 질환 순위에 포함	'질환분석대상여부' 규칙으로 **임상 질환만** 집계
단위·서식	표시된 기호를 그대로 신뢰	측정값 산식으로 **역산해 단위 검증**

STEP 5 스스로 찾고 검증하는 탐색 · 더 깊은 인사이트

통념을 물으면 가설로 바꿔 스스로 검증합니다

"무보험 환자는 응급실에 몰린다"는 흔한 가정입니다. CubeAgent는 이를 비교 가능한 지표로 바꾸고, 기각된 가설과 새로 드러난 차이를 함께 보고합니다.

나 **무보험 환자는 응급실을 더 많이 이용해?**

① 관찰: 무보험 환자의 규모 발견

무보험 진료 환자는 **21,497명으로 보험자 중 가장 많고**, 청구 $9.54억 전액이 본인 부담입니다. 보험자 10곳의 환자 부담 추정액 중 36.9%입니다.

② 가설: 응급실 이용 비중이 높다? 기각

건수가 아니라 **전체 진료 중 응급 비중**으로 비교하면 무보험 3.61%, 민간 3.62%로 같습니다. 오히려 공공보험이 4.19%로 가장 높습니다.

③ 대안 확인: 이용 구성이 다른가? 차이 작음

외래 54.3% vs 52.0%, 건강검진 25.3% vs 24.9%로 민간보험과 구성이 거의 같습니다. 입원만 2.55% vs 1.57%로 조금 높습니다.

④ 새로 드러난 차이: 응급 1건의 금액 핵심

응급 1건당 청구액은 **무보험 $4,885로 민간 $3,839의 1.27배**입니다. 이용 빈도가 아니라 한 번의 비용이 다릅니다.

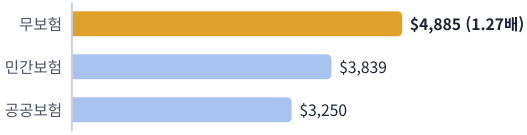
전체 진료 중 응급 진료 비중

보험 소유형태별 · 진료 건수 기준



응급 진료 1건당 청구액

보험 소유형태별 · 총 청구금액 ÷ 진료 건수



공공보험은 Medicare·Medicaid·Dual Eligible, 민간은 Humana 등 6곳입니다. 한 환자가 여러 보험자에 걸칠 수 있어 환자 수는 보험자 간 합산하지 않았습니다.

에이전트가 적용한 검증 방법

비율로 비교
집단 규모가 달라 건수 대신 전체 진료 중 비중으로 비교

대안 가설 점검
응급 외 진료유형 구성까지 확인해 다른 설명을 배제

중복 집계 주의
보험자별 환자 수 합이 전체 환자를 넘는 점을 확인하고 합산 금지

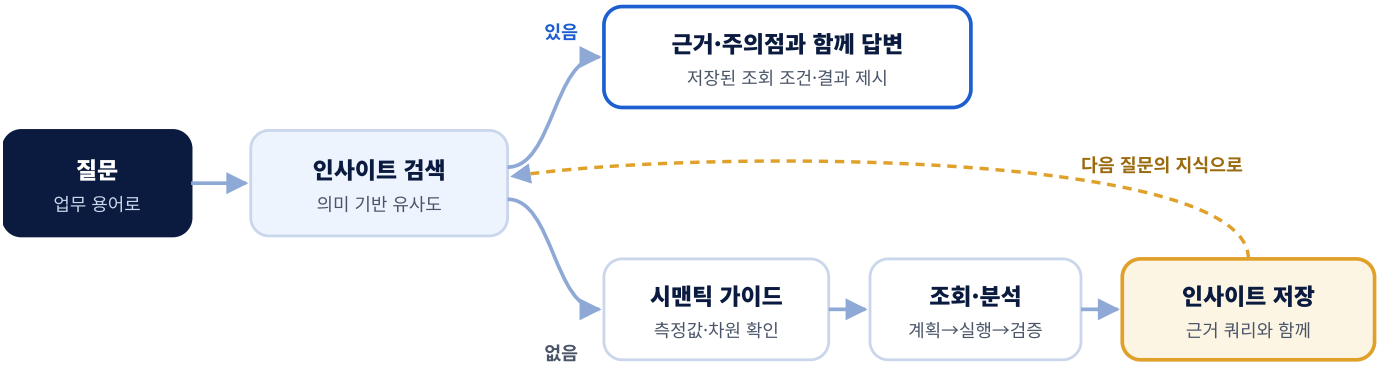
한계 명시
건당 금액 차이의 원인(중증도 구성)은 추가 분석 대상으로 분리

결론 | 무보험 환자의 부담은 **응급실을 더 자주 찾아서가 아니라, 같은 이용에 대한 비용을 전액 본인이 지기 때문**입니다. 다음 단계로 무보험 응급 진료의 진료사유 구성을 확인해 건당 금액 차이의 원인을 볼 것을 제안했습니다.

STEP 6 쌓이는 분석 지식 · 더 깊은 인사이트

앞사람이 빠진 함정을 다음 사람은 피해 갑니다

검증을 마친 분석 결과는 근거 쿼리와 해석상의 주의점까지 인사이트로 저장됩니다. 다음 질문은 이 지식에서 먼저 답을 찾고, 없을 때만 새로 분석합니다.



실제 사례 · 이 사례집의 첫 질문

나 **질환대분류별 진단 환자 수**

이전 분석에서 저장된 인사이트 5건 발견 · 최고 유사도 0.86

#324 “기타·상태 57,188명은 전체 환자 수와 같은 일괄 관리코드이므로 질환 부담으로 해석할 수 없음”

‘사회·생활요인’은 취업·교육 같은 비임상 코드이고, 대분류 합계 430,378은 환자 중복 집계라는 주의점도 함께 돌아왔습니다. 새 분석가는 같은 함정을 다시 밟지 않고 **임상 질환만 거른 순위(#327)**로 바로 시작합니다.

INSIGHT 재사용

인사이트에 저장된 근거 · 질환대분류별 진단 환자 수



1 주의점까지 재사용
숫자뿐 아니라 “이렇게 해석하면 안 된다”는 판단이 함께 저장됩니다.

2 이번 작업도 지식으로
이 사례집을 만들며 저장한 인사이트 10건(#760~#769)이 다음 질문의 출발점이 됩니다.

3 사람이 바뀌어도 맥락은 남는다
원무·심사 담당자가 바뀌어도 데이터의 함정은 조직에 남습니다.

인사이트 하나에 저장되는 것

대표 질문과 제목	요약·주의점	근거 조회 정의	결과 데이터	차트 정의	의미 벡터
무엇을 물었고 무엇이 결론인지	수치, 관리코드·중복 집계 경고	다시 실행 가능한 조회 조건	당시 조회된 행 그대로	차트 형태, 축, 레이블 방식	표현이 달라도 찾는 의미 검색용

WITH BI 기존 BI와 함께 · 기존 BI 활용

경영 대시보드는 그대로, 원무·심사 질문은 바로 답합니다

CubeAgent는 BI를 대체하지 않습니다. 이미 구축한 의료 DW와 시맨틱 모델 위에서 정형 대시보드와 비정형 질문의 역할을 나눕니다.



지금: 모든 질문이 전산 요청으로

BEFORE

심사 대응	요청 → 전산팀 추출 → 검증 → 엑셀 가공
후속 질문	재요청 → 쿼리 수정 → 재추출
정기 보고	요청 → 쿼리 작성 → 화면 개발 → 운영

심사 대응, 병상 회의, 적정성 평가 자료처럼 한 번 보고 끝날 질문도 전산팀 추출과 화면 개발을 거칩니다.

CubeAgent 도입 후: 질문의 성격대로

AFTER

심사 대응	문장으로 질문 → 답변
후속 질문	이어서 질문 → 답변
정기 보고	질문으로 검증 → BI 화면 개발 → 운영

일회성·탐색 질문은 문장으로 바로 답하고, 매일 볼 지표만 대시보드로 개발합니다.

병원 조직의 역할이 달라집니다

원무·심사·QI 현업
요청서 대신 질문을 직접 던지고, 청구·재원·질환 관점을 바꿔 가며 후속 질문을 이어 갑니다.

의료정보·분석팀
추출 작업자에서 지표 설계자·검증자로. 관리코드 제외 규칙과 지표 정의를 모델에 담습니다.

BI 개발
경영·병상 현황처럼 매일 볼 화면에 집중합니다. 지표 정의는 같은 모델을 공유합니다.

도입 안내

우리 병원 데이터로, 우리 질문으로 확인해 보십시오

사례집의 6단계는 모두 같은 원리 위에 있습니다. 잘 정의된 시맨틱 모델, 그 위에서 계획하고 검증하는 에이전트, 그리고 쌓이는 지식입니다.

STEP 1
문장으로 추출·시각화
"진료유형별 청구금액" → 입원 2.0%가 청구 17.7%

STEP 2
대화로 좁혀 가기
진료유형 → 진료사유 → 환자, 폐암 407명에 도달

STEP 3
복합 분석
2017년 청구 5.4배를 유형·환자·단가로 분해

STEP 4
정확한 숫자
약제비 70.6% 이상치 등 4가지 함정을 먼저 발견

STEP 5
스스로 탐색
"무보험은 응급실" 가설 기각, 건당 비용 차이 발견

STEP 6
지식 축적
이전 분석의 관리코드 경고를 재사용해 바로 시작

도입은 이렇게 진행됩니다

01 진단
데이터·질문 진단
EMR·원무 DW 구조와 원무·심사·QI 부서의 대표 질문을 함께 정리합니다.

02 모델링
시맨틱 모델 구성
진료·진단·처방·청구 지표를 정의하고, 관리코드 제외 같은 규칙을 모델에 담습니다.

03 PoC
실제 질문으로 검증
대표 질문에 대한 답을 기존 원무 리포트 수치와 대조해 정확도를 확인합니다.

04 확산
운영·확대
부서별 스킬과 인사이트를 축적하고 임상·경영 영역으로 넓힙니다.

POC · DEMO 문의

쿼리 한 줄 없이, 다음 회의의 질문에 답해 보세요.

Tel 02-2088-1501, 010-9733-2366 | Fax 02-6918-0898
E-mail sspark@foresoft.co.kr | Web www.foresoft.co.kr
08501 서울시 금천구 가산디지털2로 184, 벽산디지털밸리2차 1109호

FORESOFT
데이터, 그리고 더 나은 내일