CUBEAGENT · USE CASE BOOK

CubeAgent 활용 사례집

SQL 대신 문장으로.

쿼리를 작성하고 BI 화면을 개발해 답하던 분석 조직이
CubeAgent와 함께 일하는 방식을, **실제 조회 결과**로 보여드립니다.

카드 · 병원 · 통신 · 유통·판매

이 사례집의 흐름

한 문장 조회에서 스스로 쌓는 분석 지식까지, 6단계

앞 장의 다섯 가지 약속이 실제 업무에서 어떻게 작동하는지, 쉬운 단계부터 차례로 보여드립니다.



본 사례집의 모든 수치와 차트는 FORESOFT 데모 환경(카드·병원·통신·자전거 판매 샘플 데이터)에서 CubeAgent가 실제로 조회한 결과입니다. 금액 단위 $는 샘플 데이터의 원 단위(USD)를 그대로 표기했습니다.

FORESOFT

STEP 1 문장으로 추출·시각화 · 자유로운 질문

요청서와 SQL 대신, 한 문장이면 표와 차트가 나옵니다

"업종별 결제 금액 상위 10개" 같은 요청 하나에도 지금까지는 테이블을 찾고, 조인을 짜고, 결과를 검증하고, 화면을 만들어야 했습니다.

지금까지의 방식 BEFORE

1. **요건 정리** 현업 요청서 접수·해석
2. **테이블·코드 확인** 원천 테이블, 조인 키 파악
3. **SQL 작성·결과 검증** 지표 정의 확인, 수치 대조
4. **BI 화면 개발** Spotfire 시각화·배포
5. **수정 요청 반복** "업종 상세로도 볼 수 있을까요?"

작성해야 하는 쿼리 (예시)

```
SELECT m.업종,
       SUM(t.거래금액)  AS 결제금액,
       COUNT(*)         AS 결제건수
FROM   카드거래 t
JOIN   업종 m ON t.업종코드 = m.업종코드
WHERE  t.승인거절사유 IS NULL
GROUP  BY m.업종
ORDER  BY 결제금액 DESC
FETCH  FIRST 10 ROWS ONLY;
```

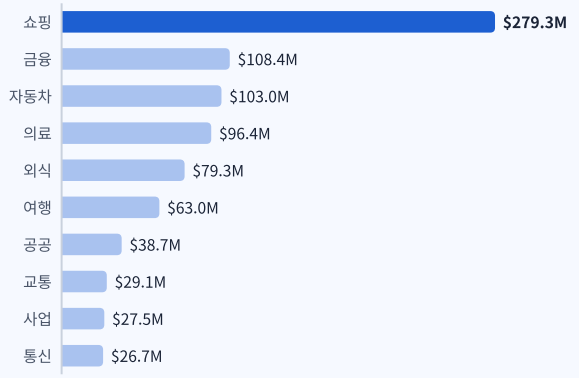
'결제 금액'에 거절 거래를 넣을지부터 담당자마다 다를 수 있습니다.

CubeAgent AFTER

나 **업종별 결제 금액 상위 10개 보여줘**

의도 RANKING · 측정값 .결제 금액 · 측정값 .결제 건수 · 차원 업종.업종

쇼핑이 $279.3M으로 1위이며, 2위 금융($108.4M)의 약 2.6배입니다. 상위 3개 업종(쇼핑·금융·자동차)이 상위 10개 합계의 58%를 차지합니다.



답변과 함께 남는 조회 정의 (실제 생성값)

```
{
  "intent": "RANKING",
  "measures": [".결제 금액", ".결제 건수"],
  "dimensions": ["업종.업종"],
  "intentParameters": { "rankCount": 10,
                        "rankDirection": "DESC" },
  "sort": [{ "field": ".결제 금액", "direction": "desc" }],
  "ccq": { "chartType": "bar", "xAxis": "업종.업종",
           "yAxis": ".결제 금액", "showLabels": true }
}
```

SQL이 아니라 '업무 언어로 된 조회 정의'

CubeAgent는 질문을 측정값·차원·의도로 구조화한 조회 정의를 만들고, 시맨틱 모델이 이를 실행합니다. 분석가는 SQL을 읽지 않고도 어떤 기준으로 답했는지 바로 검토할 수 있습니다.

차트 형태(ccq)까지 함께 정의되므로, 같은 질문은 언제 다시 물어도 같은 방식으로 조회되고 표현됩니다.

1 테이블 구조를 몰라도
질문은 업무 용어로. 시맨틱 모델이 측정값과 차원으로 번역합니다.

2 지표 정의는 모델이 보장
'결제 금액'의 계산식은 한 곳에서 정의되고, 누가 물어도 같습니다.

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

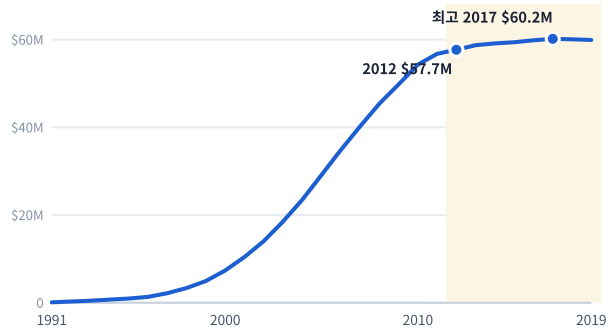
STEP 1 문장으로 추출·시각화 · 업종별 예시

데이터가 달라도 묻는 방식은 같습니다

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

카드 TREND

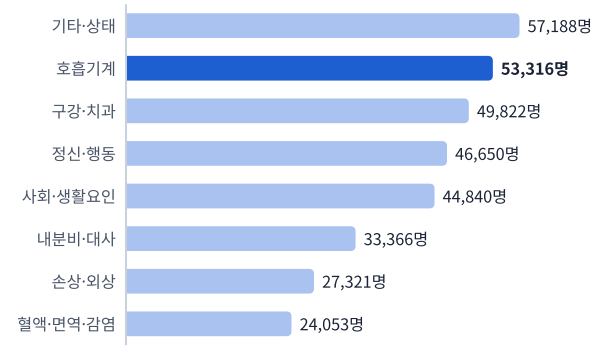
나 연도별 결제 금액 추이 보여줘



2012년을 기점으로 성장이 멈췄습니다. 전년 대비 증가율이 2011년 4.5%에서 2012년 1.6%로 꺾였고, 2014년 이후 연 ±1% 안에서 움직입니다.

병원 RANKING

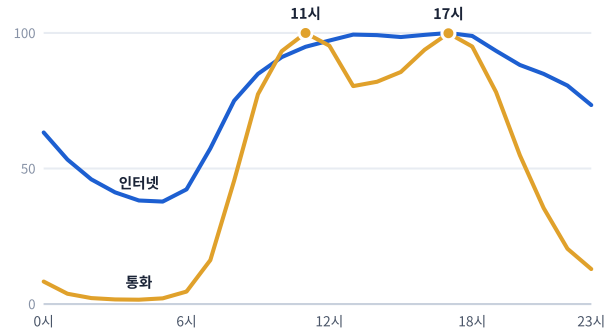
나 진단 환자가 많은 질환 대분류 순위는?



호흡기계가 실질 1위(53,316명)입니다. '기타·상태'를 빼면 호흡기·구강치과·정신행동 순입니다. 환자당 진단 건수는 사회·생활요인이 12.6건으로 가장 많습니다.

통신 DISTRIBUTION

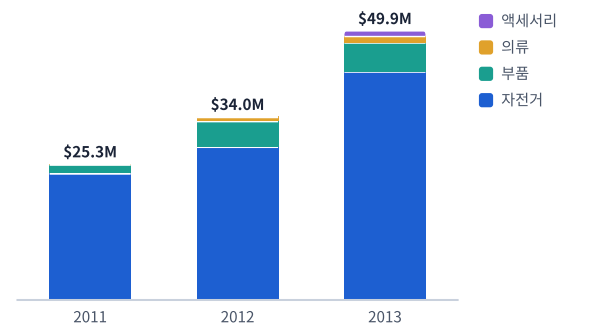
나 시간대별 인터넷·통화 트래픽 패턴 비교해줘



통화는 11시·17시 두 번의 피크, 인터넷은 13~18시 내내 고원을 유지합니다. 새벽 4시 통화는 피크의 1.6%까지 떨어지지만 인터넷은 38%가 남습니다.

유통·판매 COMPOSITION

나 연도별 제품 분류별 매출 구성 보여줘



매출의 80% 이상이 자전거입니다. 부품은 2012년 $4.7M으로 세 배 가까이 늘었고, 액세서리·의류는 금액은 작지만 2013년 주문 건수가 급증했습니다.

통신 차트는 두 측정값의 규모 차이가 커서 각 측정값의 최대 시간대를 100으로 환산해 비교했습니다. 유통·판매는 연중 일부만 포함된 2010·2014년을 제외했습니다.

STEP 2 대화로 좁혀 가기 · 자유로운 질문

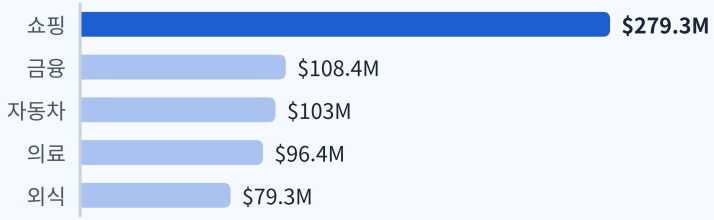
다음 질문이 떠오를 때마다 새 리포트를 요청하지 않아도 됩니다

분석은 한 번의 조회로 끝나지 않습니다. CubeAgent는 앞선 대화의 맥락(업종, 조건)을 이어받아 관점을 바꿔 가며 답합니다.

나 **업종별 결제 금액 순위 보여줘**

1

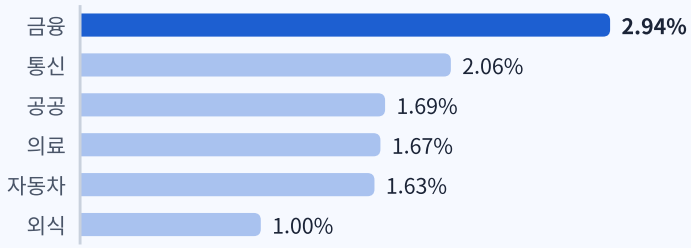
쇼핑이 1위 $279.3M.
금융 $108.4M, 자동차 $103.0M이 뒤를 잇습니다.



나 **그럼 승인 거절은 어느 업종이 많아?**

2

건수로는 쇼핑(130,414건), 비율로는 금융(2.94%)입니다.
거래량이 많은 업종이 건수 1위가 되기 쉬워 **.승인 거절 비율**로 비교했습니다. 금융은 외식(1.00%)의 약 2.9배입니다.



나 **금융은 왜 이렇게 거절이 많지?**

3

거절 10건 중 8건이 잔액 부족입니다.
금융 업종 거절 34,153건 중 잔액부족 27,322건(80.0%). 업종 상세로 보면 '금융서비스'가 33,495건으로 거의 전부이며, 거절 비율은 3.01%입니다.



POINT | 질문 3개로 **매출 → 리스크 → 원인**까지 관점이 바뀌었습니다. 기존 방식이었다면 리포트 3건을 요청하거나 쿼리 3개를 새로 짜야 했을 흐름입니다. 금액과 비율 중 어떤 지표로 비교할지도 에이전트가 판단하고 이유를 밝힙니다.

STEP 3 스스로 계획하는 복합 분석 · 복잡한 분석

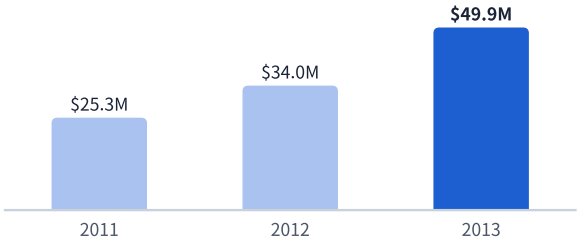
“왜?”라고 물으면 분석 단계를 스스로 설계합니다

나 **2013년 매출이 크게 늘었던데, 성장은 어디서 온 거야?**

분석 계획 — ① 전체 증가 규모 확인 → ② 채널별 기여 분해 → ③ 고객 수와 객단가로 분해 → ④ 제품별 변화 확인. 연중 일부만 있는 2010·2014년은 비교에서 제외합니다.

① 전체 규모

+46.7% $34.04M → $49.93M
이익률 9.82% → 12.57%



② 채널별 기여 — 증가분 $15.89M

66% 증가분의 3분의 2가 인터넷
(인터넷 +$10.51M, 대리점 +$5.38M)

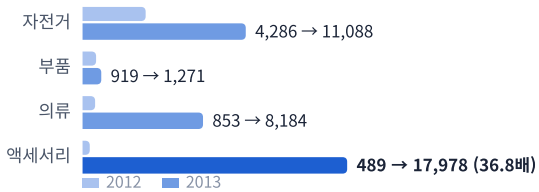


③ 인터넷 채널 분해 — 2012 → 2013

고객 수	인터넷 객단가	재구매율
5.4배	**−57%**	**15.3%**
3,255 → 17,429명	$1,787 → $768	0.43% → 15.29%

객단가는 반 이하로 줄었지만 고객이 다섯 배 넘게 늘어 인터넷 매출이 2.8배가 됐습니다.

④ 제품별 주문 건수 — 2012 → 2013



액세서리 주문이 36.8배로 뛰었습니다. 매출의 84.7%는 여전히 자전거입니다.

결론: 2013년 성장은 ‘고객 기반 확대형’입니다.

저가 액세서리·의류가 인터넷 신규 고객의 유입 창구가 됐고, 재구매율이 0.43%에서 15.29%로 올랐습니다. 매출의 중심은 여전히 자전거이므로, 다음으로는 **유입 고객이 자전거 구매로 전환되는 비율**을 확인해 볼 것을 제안합니다.

STEP 4 정의된 숫자, 검증되는 답 · 정확한 숫자

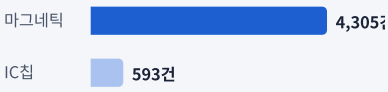
그럴듯한 숫자가 아니라 비교해도 되는 숫자로 답합니다

LLM이 원천 테이블을 직접 뒤지면 계산은 맞아도 해석이 틀리기 쉽습니다. CubeAgent는 시맨틱 모델의 지표 정의와 데이터 범위를 확인한 뒤 답합니다.

나 **결제 방식별로 부정 사용을 비교해줘. 마그네틱이 더 위험해?**

흔한 답 · 건수 합계

~~7.3배~~

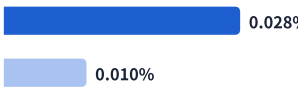


전체 기간 누적 건수

마그네틱 4,305건 vs IC칩 593건. 거래량이 다르니 건수 비교는 규모만 보여줍니다.

① 정의된 지표 · 전체 기간 비율

2.9배

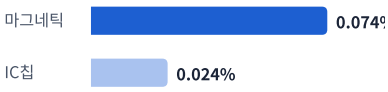


전체 기간 · 기간 불일치 포함

.부정 사용 비율(= 부정 사용 건수 ÷ 결제 건수)로 비교하면 0.028% vs 0.010%.

② 범위 점검 · 공통 기간 2015~2016

3.1배



2015~2016 · 두 방식 모두 측정

두 방식이 모두 측정된 기간만 보면 0.074% vs 0.024%. 이 기간 **건수는 오히려 IC칩이 많습니다**(580 vs 429건).

마그네틱의 부정 사용 비율이 IC칩의 약 3배입니다. 다만 전체 기간 누적 비교는 근거로 쓰기 어렵습니다. IC칩 거래는 2015년부터만 있고, 부정 사용 건수는 2010·2011·2017·2019년이 **0이 아니라 결측**이어서 두 방식의 분모 기간이 서로 다르기 때문입니다. 공통 기간 기준으로 답을 드립니다.

이 답을 가능하게 한 시맨틱 모델의 역할

A 지표 정의 단일화
'부정 사용 비율'의 분자·분모가 모델에 한 번만 정의됩니다. 사람마다, 화면마다 다른 계산식이 생기지 않습니다.

B 필터 전파 사전 검증
조회 전에 필터가 측정값의 원천 테이블까지 온전히 전달되는지 (FULL·PARTIAL) 확인해 조용히 틀린 숫자를 막습니다.

C 결측과 0의 구분
값이 없는 연도를 0으로 합산하지 않습니다. 데이터 범위의 한계를 답에 함께 밝힙니다.

구분	원천 테이블에 직접 SQL을 생성하는 방식	CubeAgent (시맨틱 모델 기반)
지표 계산식	질문할 때마다 LLM이 새로 작성	**모델에 정의된 측정값**을 그대로 사용
조인·필터	테이블 관계를 LLM이 추론	모델의 관계를 따르고, **필터 전파를 사전 검증**
결측 데이터	작성된 쿼리에 따라 0 또는 제외	결측을 구분하고 **답변에 한계로 명시**
결과 검토	생성된 SQL을 읽고 해석해야 함	측정값·차원 이름으로 된 **조회 정의**를 확인

STEP 5 스스로 찾고 검증하는 탐색 · 더 깊은 인사이트

차이를 발견하면 가설을 세우고 스스로 검증합니다

숙련된 분석가가 하던 일, 곧 “진짜 차이인가, 구성 차이인가?”를 따지는 과정을 에이전트가 단계별로 수행하고 기각된 가설까지 보고합니다.

나 **카드 브랜드별로 부정 사용 비율에 차이가 있어?**

① 관찰 발견

브랜드별 비율은 0.020~0.024%로 좁지만, **아멕스만 기준선의 0.86배**로 통계적으로 낮습니다(z = −2.51).

② 가설: 결제 방식 구성이 달라서? 기각

아멕스의 IC칩 비중은 26.6%로 다른 브랜드(27.1~28.9%)와 비슷합니다. 차이는 **마그네틱 결제에서만** 나타납니다(0.83배, z = −2.82). IC칩에서는 브랜드 간 차이가 없습니다.

③ 가설: 거래 시기 구성이 달라서? 기각

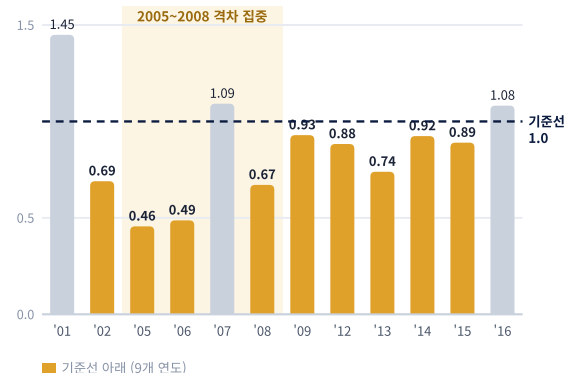
연도별 기준선으로 통제해도 0.83배(z = −3.03)로 유지되고, 비교 가능한 12개 연도 중 9개에서 기준선 아래입니다.

④ 추가 발견: 차이는 과거에 몰려 있다 핵심

격차는 2005~2008년(0.46~0.67배)에 집중되고, 2012~2016년에는 0.92배로 유의하지 않습니다.

아멕스 부정 사용 비율 ÷ 해당 연도 기준선

1.0 = 그해 전체 브랜드 평균. 4개 브랜드 모두 값이 있는 12개 연도만 표시



전체 기간 브랜드별 (기준선 대비)



디스커버는 연간 부정 건수가 1~27건으로 적어 연도별 배수의 변동을 표본 잡음으로 판단했습니다.

에이전트가 적용한 검증 방법

기대 건수 비교
기준선 비율로 기대 건수를 계산해 실제와 비교 (275건 vs 기대 319.9건)

z-점수
차이가 우연인지 판단하기 위해 표준화된 편차로 평가

다중 비교 보정
브랜드×결제방식 8개 셀을 동시에 보므로 본페로니 임계값(약 2.73) 적용

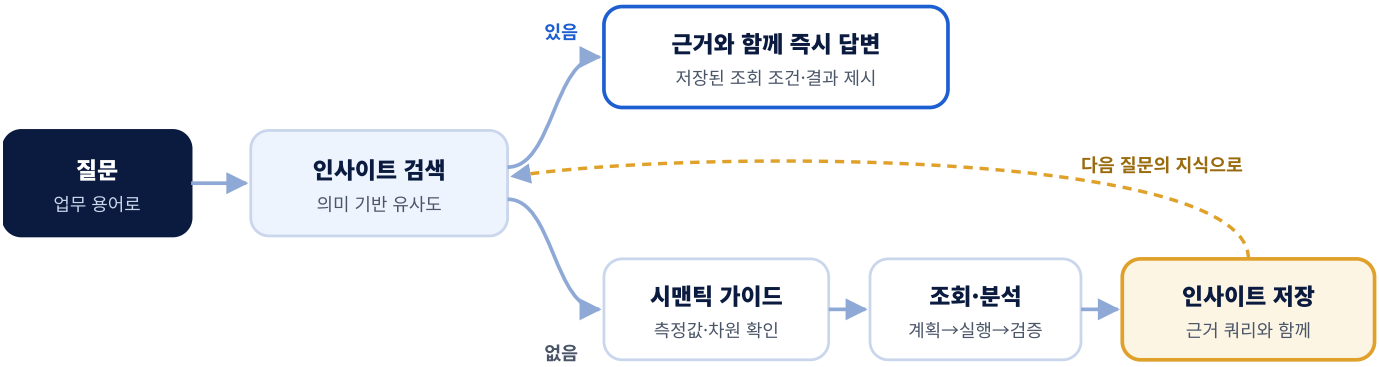
구성 효과 통제
결제 방식·연도별 기준선으로 표준화해 구성 차이를 제거

결론 “아멕스가 안전하다”는 결론은 **과거 마그네틱 시대의 현상**에 가깝습니다. 현재 리스크 정책의 근거로 삼기에는 약하다는 판단까지, 에이전트가 근거와 함께 제시했습니다.

STEP 6 쌓이는 분석 지식 · 더 깊은 인사이트

분석은 끝나도 지식은 남습니다

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



실제 사례 · 이 사례집을 만들며 던진 질문

INSIGHT 재사용

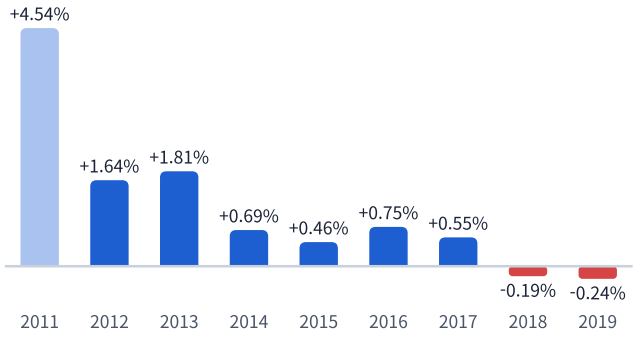
나 **월별 카드 이용금액 추이**

저장된 인사이트 #465 · 의미 유사도 0.77

연도별 결제 금액: 2012년 성장 둔화 전환, 2014년 이후 ±1.04% 안에서 유지, 최고 연도는 2017년

표현이 달라도('이용금액' ↔ '결제 금액') 의미로 찾아냅니다. 인사이트에는 당시의 조회 조건과 결과가 함께 저장돼 있어 바로 검증할 수 있습니다.

인사이트에 저장된 근거 · 전년 대비 결제 금액 증가율



1 같은 질문에 같은 답
한 번 검증한 결론을 재사용해 부서·담당자별로 답이 달라지지 않습니다.

2 인사이트가 인사이트를 부른다
앞 장의 결제 방식 분석은 이전 인사이트(#393, #394)를 근거로 인용했습니다.

3 사람이 바뀌어도 맥락은 남는다
분석가의 판단 과정이 조직의 자산으로 축적됩니다.

인사이트 하나에 저장되는 것

대표 질문과 제목
무엇을 물었고 무엇이 결론인지

요약과 해석
수치, 비교 기준, 해석상의 주의점

근거 조회 정의
측정값·차원·필터·의도 (다시 실행 가능)

결과 데이터
당시 조회된 행 그대로

차트 정의
차트 형태, 축, 레이블 방식

의미 벡터
표현이 달라도 찾아지는 의미 검색용

WITH BI 기존 BI와 함께 · 기존 BI 활용

Spotfire는 그대로, 개발 요청의 줄은 짧아집니다

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



지금: 모든 질문이 개발 요청으로

BEFORE

일회성 질문	요청	쿼리 작성	검증	화면 개발
후속 질문	재요청	쿼리 수정	화면 수정	
정기 보고	요청	쿼리 작성	화면 개발	운영

한 번 보고 끝날 질문도 쿼리 작성과 화면 개발을 거칩니다. 분석 인력의 시간이 반복 추출에 묶입니다.

CubeAgent 도입 후: 질문의 성격대로

AFTER

일회성 질문	문장으로 질문	답변		
후속 질문	이어서 질문	답변		
정기 보고	질문으로 검증	BI 화면 개발	운영	

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

분석 조직의 역할이 달라집니다

현업
요청서 대신 질문을 직접 던지고, 후속 질문까지 스스로 이어 갑니다.

분석가
추출 작업자에서 지표 설계자·검증자로. 시맨틱 모델과 인사이트의 품질을 관리합니다.

BI 개발
정말 반복해서 볼 화면에 집중합니다. 지표 정의는 같은 모델을 공유합니다.

도입 안내

우리 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
"업종별 결제 금액 상위 10개" → 표와 차트

STEP 2
대화로 좁혀 가기
매출 → 거절 비율 → 거절 사유, 질문 3개로

STEP 3
복합 분석
매출 +46.7%를 채널·고객·제품으로 분해

STEP 4
정확한 숫자
7.3배로 보이던 차이를 공통 기간 3.1배로 교정

STEP 5
스스로 탐색
가설 2개 기각, 과거 구간에 몰린 격차 발견

STEP 6
지식 축적
유사 질문에 저장된 인사이트로 즉시 답변

도입은 이렇게 진행됩니다

01 진단
데이터·질문 진단
현재 DW·Cube 구조와 부서별 대표 업무 질문을 함께 정리합니다.

02 모델링
시맨틱 모델 구성
기존 모델을 최대한 활용해 AI Ready 모델로 다듬고 핵심 지표를 정의합니다.

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

04 확산
운영·확대
업무별 스킬과 인사이트를 축적하고 부서와 데이터 영역을 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