숫자로 읽는 기업

성장생과 수익성 분석부터 예측 모델링까지

강의 전체 목표 및 로드맵

강의 목표

- ☑ 데이터 기반 투자 사고방식 훈련
- 감(感) 이 아닌, 숫자와 데이터로 기업을 분석
- 재무제표 기반의 기업 경쟁력 평가
- 기업의 성장성과 수익성을 데이터로 해석하는 눈 키우기
- ☑ 데이터 분석부터 대시보드까지 완성
- 실제 5,000개 기업 비식별 재무 데이터셋 활용
- 데이터 수집 → 저장 → 분석 → 시각화 → 대시보드 완성
- Colab + DuckDB + Streamlit 으로 현업형 분석 파이프라 인 구축
- ☑ 예측 모델링까지 확장
- 과거 데이터를 기반으로 미래 매출 및 수익성 예측
- 분류/회귀 모델을 활용하여 투자 의사결정 보조

학습 기대 효과

- ☞ 기업 재무 분석 역량 확보
- 매출 증가율, ROE, 부채비율 등 주요 지표의 계산과 해석 능력
- ☞ 데이터 기반 투자 판단
- 재무 지표로 기업의 현재 위치와 성장 가능성 진단
- ◎ 실무형 분석 환경 익히기
- Colab 과 DuckDB 를 활용하여 대용량 데이터 처리
- Streamlit 으로 실시간 대시보드 구현 및 공유
- ◎ 예측 모델을 통한 투자 전략 보완
- 머신러닝 모델로 성장성 예측 및 리스크 관리

☑ 투자 성공률을 높이는 재무 분석

- 시장에는 수많은 정보가 존재합니다.
- 뉴스, 테마, 전문가 의견에 휘둘리지 않고 객관적인 숫자로 기업을 분석하는 능력이 투자 성공의 핵심.
- 재무제표는 기업이 투자자에게 보내는 공식적인 신호입니다.
- 숫자는 거짓말하지 않습니다. 기업의 매출, 이익, 자산, 부채 등 모든 것이 기업의 "현재" 를 설명합니다.

☑ 숫자는 기업의 이야기다

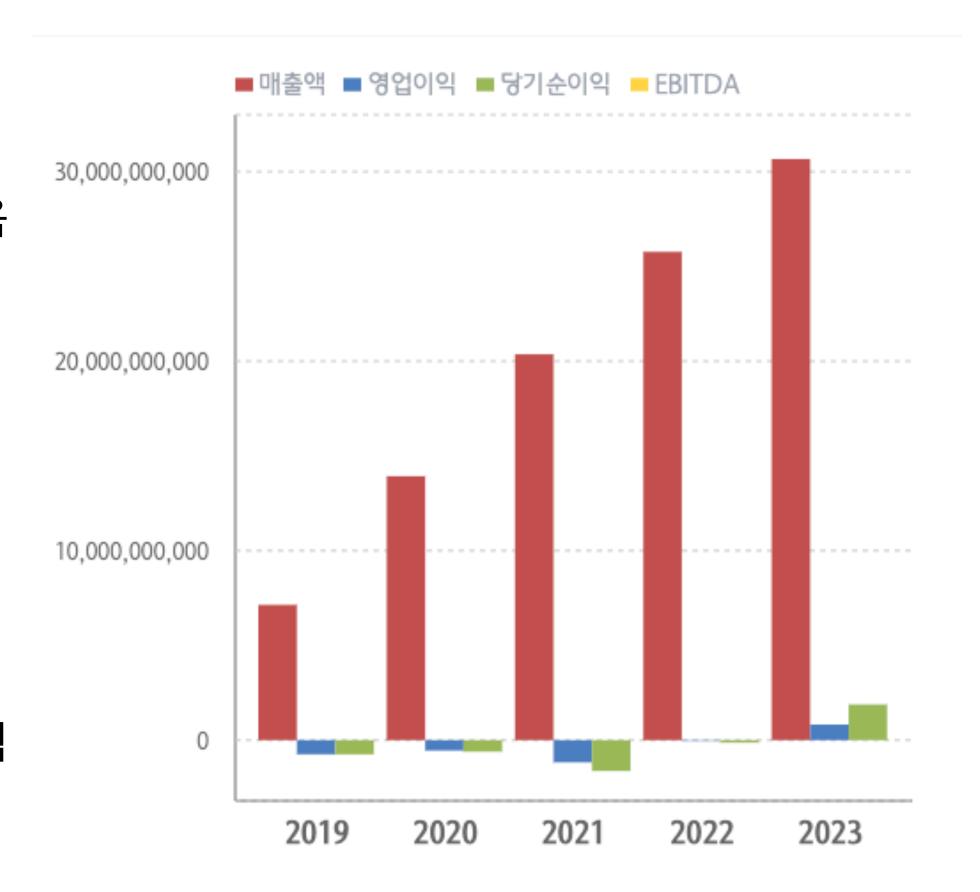
- 매출이 증가하는 기업
 - → 시장 점유율 확장, 성장 가능성
- 영업이익률이 상승하는 기업
 - → 경쟁력 강화, 효율적인 경영
- 부채가 과도하게 증가하는 기업
 - → 재무 위험 신호
- 이익잉여금이 꾸준히 증가하는 기업
 - → 장기적인 안정성

☑ 시장 심리 vs 기업 내재가치

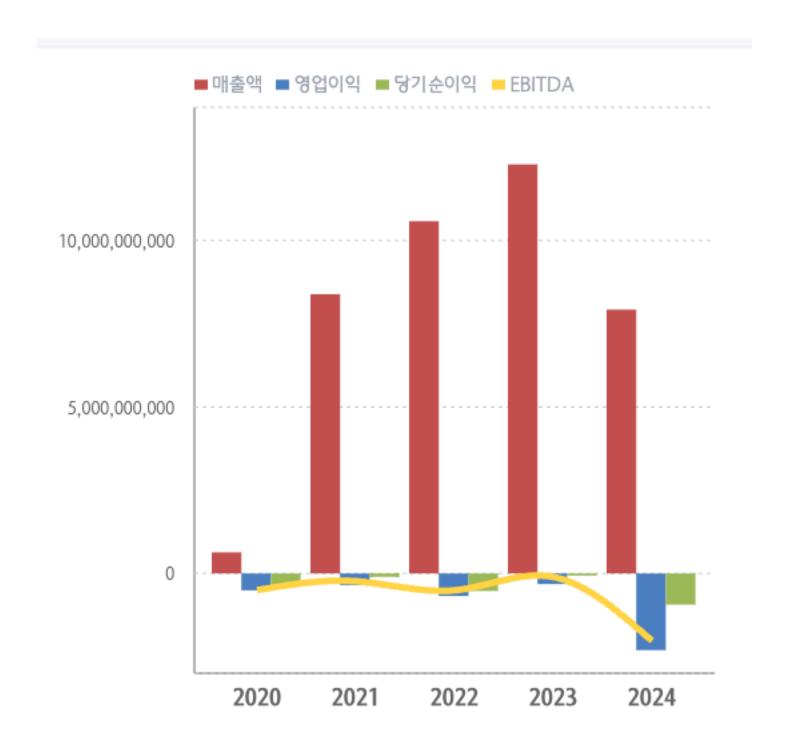
구분	시장 심리	기업 내재가치
정의	투자자들의 기대감, 공포, 탐욕 등 감정적 요소	기업이 실제로 창출하는 가치 (매출, 이익 등)
특징	단기 변동성에 큰 영향을 줌	장기적인 기업 가치를 결정
지표	공포탐욕지수, VIX, 투자자 심리지수 등	매출액, 영업이익, ROE, 부채비율 등
투자 전략	변동성 활용 (단기 트레이딩)	펀더멘털 분석 기반 (장기 투자)

☑ 실제 사례: "쿠팡, 적자 기업의 성공적 상장"

- 쿠팡은 상장 당시까지 지속적인 영업손실 이 있었음에도 불구하고, 미국 뉴욕 증시에서 100조 원 이상의 기업가치 를 평가받았습니다.
- 투자자들이 쿠팡에 주목한 이유는 무엇일까요?
 - 폭발적인 매출 성장
 - 높은 시장 점유율
 - 압도적인 물류 인프라 투자
- 재무제표상 이익은 적자였지만, 성장성 지표 가 투자자의 신뢰를 얻은 대표적 사례입니다.



- ☑ 실제 사례: "LG에너지솔루션, 배터리 시장의 수익성과 성장성"
- 전기차 배터리 시장의 세계적 강자인 LG에너지솔루션
- 매출 급증 (2020 → 2023 매출 약 19배 이상 증가)
- 영업이익률 개선 → 원가 절감, 규모의 경제 효과
- 전기차 시장 성장과 함께 기업도 동반 성장하는 이상적인 케이스



- ☑ 실제 사례: "WeWork, 매출 성장 속의 경고 신호"
- WeWork 도 한때 유니콘으로 주목
- 폭발적인 매출 성장에도 불구하고, 과도한 부채와 비용 구조 문제
- 투자자들은 부채비율 과 수익성 악화를 간과했고, 결국 기업가치 붕괴

손익계산서 (Income Statement)

- 일정 기간 동안 얼마나 벌었고, 얼마나 썼는지 보여주는 표
- 수익과 비용을 통해 경영 성과 를 나타냄

구성

- 매출액
- 매출원가
- **매출총이익** = 매출액 매출원가
- 판관비 (판매비와 관리비)
- 영업이익 = 매출총이익 판관비
- 영업외수익 / 영업외비용
- 법인세 차감 전 순이익
- 당기순이익 = 최종 이익

손익계산서					*단위 천
					✔ 전체 열
항목	2019	2020	2021	2022	2023
매출액 ▼	114,702,975	153,774,566	199,073,166	215,665,838	215,401,057
매출원가 ▼	7,044,436	4,318,772	1,566,217	1,337,103	4,304,143
매출총이익	107,658,539	149,455,794	197,506,949	214,328,735	211,096,914
판관비 ▼	113,307,424	145,036,760	213,018,633	248,617,019	238,037,158
기본주당순손익	-	-12,143	-54,178	-	-19,951
EBITDA	-4,459,191	5,854,443	-13,639,009	-31,729,179	-23,881,096
영업이익	-5,648,885	4,419,034	-15,511,685	-34,288,283	-26,940,244
영업외수익 ▼	1,116,431	1,081,813	3,926,324	83,791,263	7,941,627
영업외비용 ▼	42,497,856	20,604,357	54,737,372	10,017,731	9,180,512
법인세차감전순이익	-47,030,310	-15,103,509	-66,322,732	39,485,248	-28,179,129
계속사업손익법인세비용	-	-285,822	-	-1,023,859	722,347
계속사업이익(손실)	-47,030,310	-14,817,687	-66,322,732	40,509,107	-28,901,476
당기순이익 ▼	-47,030,310	-14,817,687	-66,322,732	40,509,107	-28,901,476
기본주당계속사업손익	-	-12,143	-54,178	-	-19,951

[&]quot;손익계산서는 기업이 얼마를 벌었고, 비용을 어떻게 관리하는지 보여줍니다. 가장 직관적인 '기업의 성적표' 입니다."

기업의 '성적표'

재무상태표 (Balance Sheet)

- 특정 시점에서 **무엇을 가지고 있고** (자산), 얼마나 빚이 있으며 (부채), 남는 **자본이 얼마**인지 보여주는 표
- **자산** = **부채** + **자본** 의 등식

구성

- 자산
 - 유동자산 (현금, 재고, 매출채권 등)
 - 비유동자산 (토지, 건물, 장비 등)
- 부채
 - 유동부채 (1년 이내 상환)
 - 비유동부채 (장기 부채)
- 자본
 - 자본금
 - 자본잉여금
 - 이익잉여금

재무상태표					*단위 천 5
					✔ 전체 열기
항목	2019	2020	2021	2022	2023
유동자산 ▼	57,833,946	66,892,569	53,899,023	193,367,349	167,738,086
비유동자산 ▼	4,710,041	24,629,788	32,471,338	68,983,574	70,858,694
자산총계	62,543,987	91,522,356	86,370,361	262,350,923	238,596,780
유동부채 ▼	97,339,606	138,625,843	198,285,491	171,686,785	178,391,055
비유동부채 ▼	1,890,838	1,471,699	1,361,339	2,878,062	2,662,417
부채총계	99,230,444	140,097,542	199,646,830	174,564,847	181,053,472
자본금 ▼	609,437	610,562	614,710	721,531	726,842
자본잉여금 ▼	69,322,349	73,113,396	71,813,416	223,183,085	226,634,329
자본조정 ▼	1,368,471	-38,721	4,579,465	6,794,948	4,699,501
기타포괄손익누계액 ▼	2,460	-491	14,611	5,197,586	1,247,782
이익잉여금 ▼	-107,989,173	-122,259,931	-190,298,672	-148,111,073	-175,765,146
당기순이익	-47,030,310	-14,817,687	-66,322,732	40,509,107	-28,901,476
자본총계	-36,686,457	-48,575,186	-113,276,469	87,786,076	57,543,308
부채와자본총계	62,543,987	91,522,356	86,370,361	262,350,923	238,596,780
Net Debt	-2,939,416	-33,432,110	-14,516,727	-1,749,472	-10,676,607

"재무상태표를 보면, 이 기업이 부채로 얼마나 의존하는지, 현금은 충분한지, 기업 체력이 보입니다."

기업의 '건강검진표'

자본변동표 (Statement of Changes in Equity)

- 일정 기간 동안 기업 자본이 어떻게 변화했는지 보여주는 표
- 기업이 얼마나 이익을 쌓았고, 배당했으며, 추가 자본을 유치 했는지 파악 가능

구성

- 기초 자본
- 당기순이익 (손익계산서에서 연결)
- 배당금 지급
- 기타 포괄손익
- 기말 자본

자본변동표						*단위 천 원
구분	자본금	자본잉여금	자본조정	기타포괄손익누계액	이익잉여금(결손금)	합계
당기순이익(손실)	-	-	-	-	40,509,107	40,509,107
기타	-	-	-	-	1,678,492	1,678,492
2022 당기말	721,531	223,144,362	6,833,670	5,197,586	-148,111,073	87,786,076
기초자본금	721,531	-	-	-	-	721,531
주식매수선택권	5,311	3,451,245	-1,954,915	-	-	1,501,641
기초자본잉여금	-	223,144,362	-	-	-	223,144,362
기타	-	-140,532	-	-	1,247,403	1,106,871
기초자본조정	-	-	6,833,670	-	-	6,833,670
기초기타포괄손익누계액	-	-	-	5,197,586	-	5,197,586
유가증권평가손실(-)	-	-	-	-3,949,804	-	-3,949,804
처분후이익잉여금(결손금)	-	-	-	-	-148,111,073	-148,111,073
당기순이익(손실)	-	-	-	-	-28,901,476	-28,901,476
2023 당기말	726,842	226,455,075	4,878,755	1,247,782	-175,765,146	57,543,308

"자본변동표를 보면, 이익이 어떻게 쌓였고, 배당이나 자본 확충이 어떻게 이루어졌는지 알 수 있습니다."

자본의 '흐름도'

◈ 재무제표 간 연결고리

손익계산서

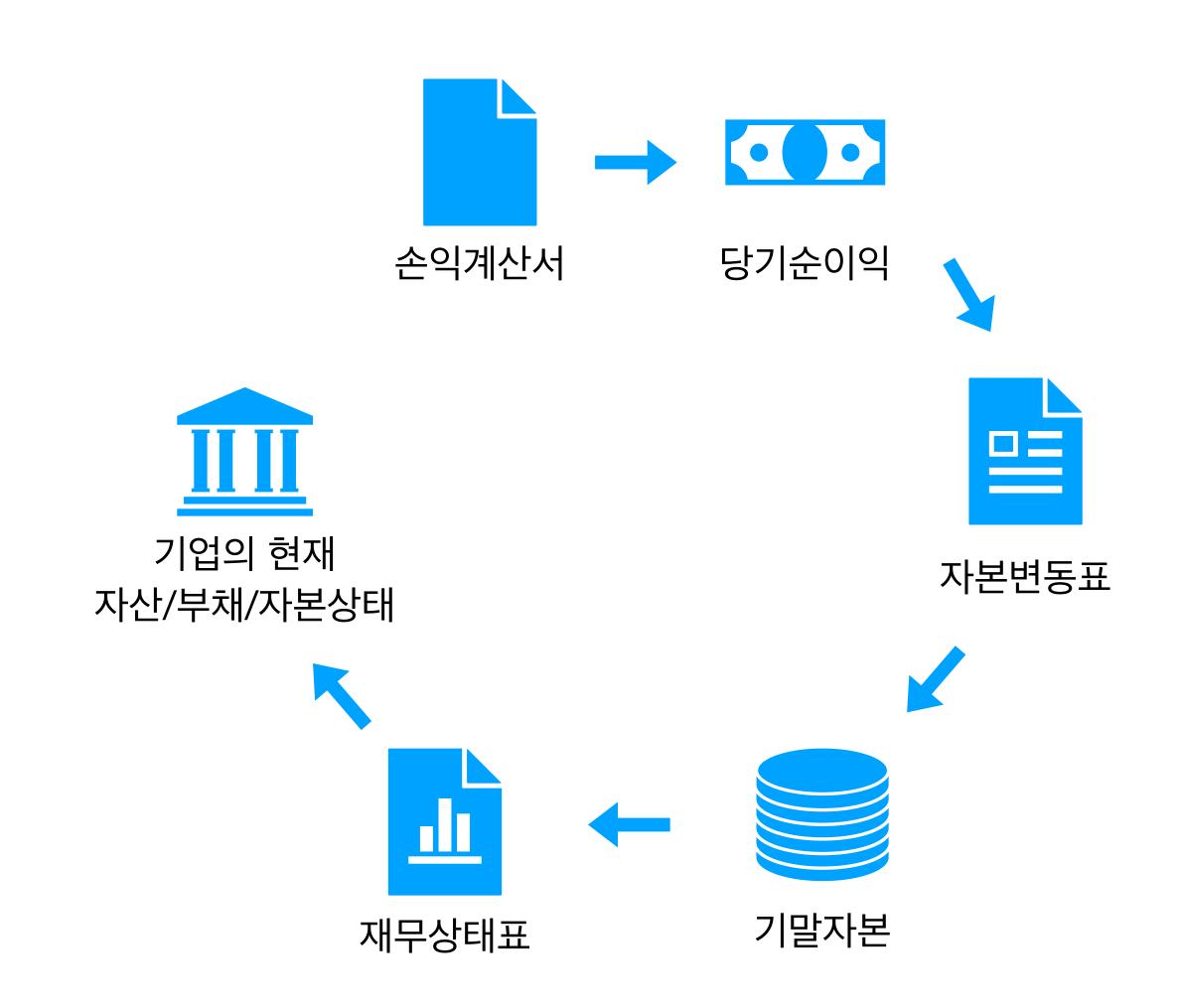
• 기간 내 수익성과 비용을 통해 당기순이익 계산

자본변동표

- 손익계산서에서 넘어온 당기순이익이 반영
- 배당금 등 자본 변동 요소 포함

재무상태표

• 자본변동표에서 최종 자본이 재무상태표의 자본으로 연결



설망성 지표 (Growth Indicators)

"기업이 얼마나 빠르게 성장하고 있는가?"

지표	계산식	의미
매출 증가율	(당기 매출 - 전기 매출) ÷ 전기 매출 × 100	기업의 시장 확장 속도
영업이익 증가율	(당기 영업이익 - 전기 영업이익) ÷ 전기 영업이익 × 100	수익 기반의 성장 정도
총자산 증가율	(당기 자산 - 전기 자산) ÷ 전기 자산 × 100	사업 확장 속도

수익성 지표 (Profitability Indicators)

"기업이 얼마나 효율적으로 이익을 내고 있는가?"

지표	계산식	의미
영업이익률	영업이익 ÷ 매출액 × 100	영업활동 수익성
순이익률	당기순이익 ÷ 매출액 × 100	최종 수익성
ROE (자기자본이익률)	당기순이익 ÷ 자본총계 × 100	투자 대비 수익성
ROA (총자산이익률)	당기순이익 ÷ 총자산 × 100	자산 효율성

"수익성이 뛰어난 기업은 위기 시기에도 이익을 유지합니다. 성장하는 기업이라면 수익성까지 확보되는지 반드시 확인해야 합니다."

③ 안정성 지표 (Stability Indicators)

"기업이 재무적으로 얼마나 튼튼한가?"

지표	계산식	의미
부채비율	부채총계 ÷ 자본총계 × 100	재무 구조 안정성
유동비율	유동자산 ÷ 유동부채 × 100	단기 지급 능력
이자보상배율	영업이익 ÷ 이자비용	이자 상환 능력 (높을수록 안정)

4 활동성 지표 (Efficiency / Activity Indicators)

"기업이 자산을 얼마나 효율적으로 운용하는가?"

지표	계산식	의미
자산회전율	매출액 ÷ 자산총계	자산 활용 효율성
재고자산회전율	매출원가 ÷ 평균재고자산	재고 관리 효율성
매출채권회전율	매출액 ÷ 매출채권	외상 매출 관리 효율성

★ 네 가지 축의 통합 평가

왜 4대 축이 중요한가?

기업은 단순히 한두 가지 지표로 평가할 수 없습니다.

단순 매출 성장만 봐서는 안 되고, 높은 이익률이 있어도 부채가 지나치게 많다면 위험할 수 있습니다. "균형 잡힌 시각" 을 가지기 위해 4대 축을 기준으로 분석합니다.

축	높은 경우	낮은 경우
성장성	시장 리더, 미래 기대	시장 정체, 성장성 부족
수익성	견고한 수익 창출	외형 성장에도 수익 악화
안정성	불황에도 생존 가능	부채 과다, 재무 위험
활동성	자산 효율적 운용	자금 흐름 문제 가능성



"기업이 시장에서 얼마나 빠르게 성장하고 있는가?"

정의

기업의 외형적 확장성 과 시장 점유율 증가를 보여주는 지표

주요 지표

- 매출 증가율 (당기 매출 - 전기 매출) ÷ 전기 매출 × 100
- 영업이익 증가율 (당기 영업이익 - 전기 영업이익) ÷ 전기 영업이익 × 100

해석포인트

- 빠른 매출 성장은 시장 내 점유율 확대를 의미
- 영업이익이 함께 증가하면, 외형 성장 + 수익성 개선 동시 진행

주의 사항

- 무리한 외형 성장 → 비용 증가 → 수익성 악화 가능성
- '허수' 성장인지, 실질적인 성장인지 분별 필요

✓ 수익성 (Profitability)

"성장하는 기업이 얼마나 효율적으로 돈을 벌고 있는가?"

정의

기업이 자원을 얼마나 효율적으로 활용해 수익을 창출하는지 보여주는 지표

주요 지표

• 영업이익률

영업이익 ÷ 매출액 × 100

• 순이익률

당기순이익 ÷ 매출액 × 100

• ROE (자기자본이익률)

당기순이익 ÷ 자본총계 × 100

해석 포인트

- 영업이익률은 본업 이 얼마나 효율적인지
- 순이익률은 모든 비용 및 세금 포함 최종 수익성
- ROE 는 투자자 관점 에서 기업의 수익 창출력

주의 사항

- 수익성이 높아도 성장성이 낮으면 한계 시장일 가능성
- 일회성 이익인지 지속 가능한 수익인지 구분

☑ 성장성 + 수익성 = 기업의 기본 경쟁력

분류	특징	전략적 시사점
고성장 + 고수익	이상적인 기업, 시장 리더 가능성	공격적인 투자 매력
고성장 + 저수익	성장 초기 기업, 시장 확장 중	향후 수익성 개선 필요
저성장 + 고수익	안정적 수익 기반	배당주 투자 등 안정적 수익 추구
저성장 + 저수익	구조적 문제 기업 가능성	투자 유의 필요

기업의 체력과 효율성을 읽다: 안정성 & 활동성

위기에도 버틸 수 있는, 자산을 잘 활용하는 기업 찾기



"위기 상황에서도 기업이 얼마나 버틸 수 있는가?"

정의

기업이 외부 충격 (금리 인상, 경기 침체 등) 에서 재무적으로 얼마나 견고한지 평가하는 지표

주요 지표

• 부채비율 (Debt Ratio)

부채총계 ÷ 자본총계 × 100

- → 기업이 자본보다 얼마나 많은 부채를 사용하고 있는지
- 유동비율 (Current Ratio)

유동자산 ÷ 유동부채 × 100

- → 단기 채무 상환 능력
- 이자보상배율 (Interest Coverage Ratio)

영업이익 ÷ 이자비용

→ 영업이익으로 이자비용을 얼마나 충당할 수 있는지

해석 포인트

- 부채비율이 높으면 위험. 일반적으로 100% 이하가 바람직
- 유동비율 100% 이상이면 단기 지급능력 양호
- 이자보상배율은 3배 이상이면 양호

주의 사항

- 낮은 부채가 무조건 좋은 것은 아님 (저성장 가능성)
- 업종마다 적정 부채비율 다름 (ex. 금융업은 부채비율 높음)

기업의 체력과 효율성을 읽다: 안정성 & 활동성

위기에도 버틸 수 있는, 자산을 잘 활용하는 기업 찾기



☑ 활동성 (Efficiency / Activity)

"기업이 자산을 얼마나 효율적으로 사용하고 있는가?"

정의

기업이 보유 자산을 얼마나 효과적으로 운용하여 매출로 연결시키는지 평 가하는 지표

주요 지표

- 자산회전율 (Total Asset Turnover)
- 매출액 ÷ 자산총계
- → 보유 자산 대비 매출 생성 효율성
- 재고자산회전율 (Inventory Turnover)
- 매출원가 ÷ 평균 재고자산
- → 재고가 얼마나 빠르게 판매되고 있는지
- 매출채권회전율 (Receivables Turnover)
- 매출액 ÷ 매출채권
- → 외상 매출이 얼마나 잘 회수되고 있는지

해석 포인트

- 높은 자산회전율: 자산 효율적 활용
- 재고자산회전율 높으면 재고 과잉 없음
- 매출채권회전율 높으면 외상 매출 관리 잘됨

주의 사항

- 너무 높은 회전율은 재고 부족 우려
- 업종 특성 고려 (유통업은 회전율 높고, 제조업은 낮음)

☑ 안정성 + 활동성 = 기업의 내재 체력

분류	특징	전략적 시사점
고안정성 + 고활동성	위기에도 버티며, 자산 효율성도 높음	장기 투자 매력
고안정성 + 저활동성	재무는 건전하지만 성장 모멘텀 약함	수익성 개선 필요
저안정성 + 고활동성	공격적 경영, 성장 추구	위험 감수형 기업
저안정성 + 저활동성	재무구조 취약 + 비효율적 자산 운용	투자 유의 필요

숫자로 읽는 기업

"실제 기업 10,000개 데이터로 배우는 재무 분석 실습"

☑ 데이터셋 개요

- 총 10,000개 이상 기업
- 다양한 업종
 - 응용 소프트웨어 개발 및 공급업
 - 전자상거래 소매업
 - 시스템 소프트웨어 개발 및 공급업 등
- 분석 기간: 최근 5년 (2019 ~ 2023)
- 기업별 데이터
 - 기업명, 사업자번호, 법인번호
 - 기준연도, 기업 상태 (정상/휴업/폐업 등)

☑ 포함된 주요 재무 항목

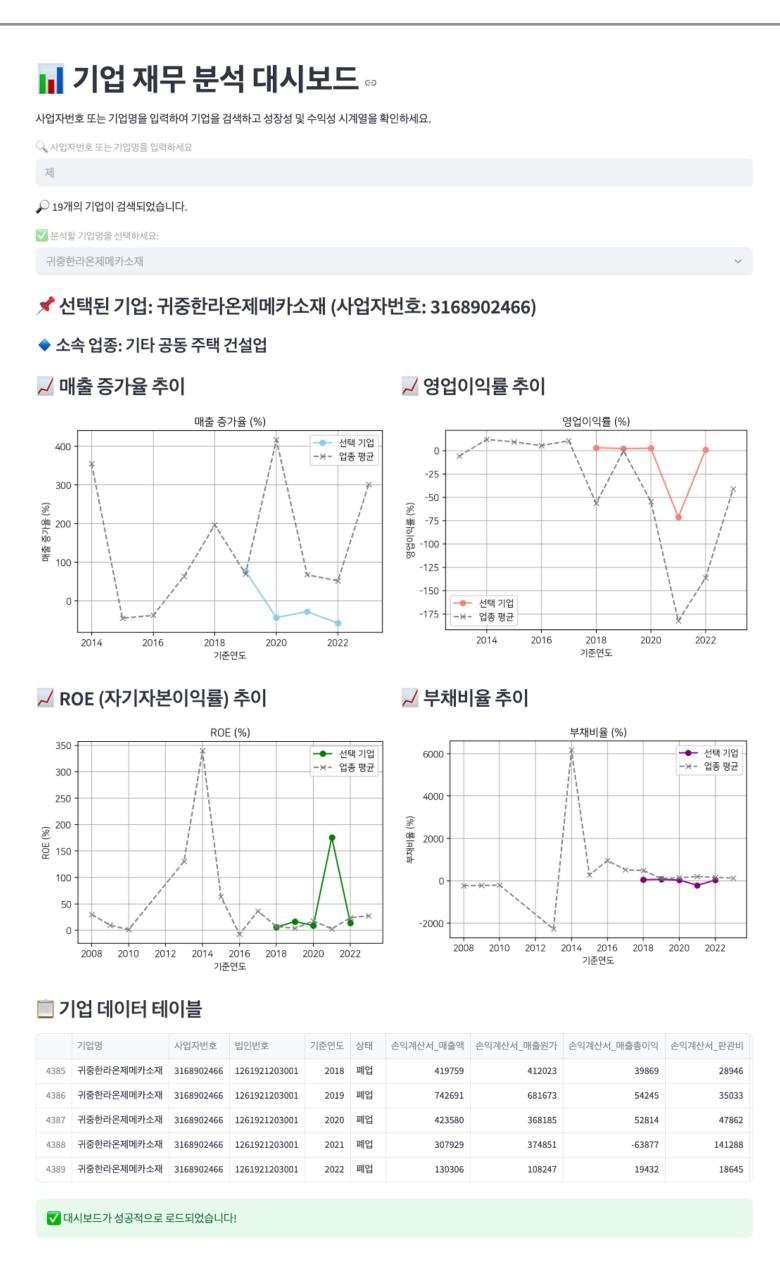
- ▲ 손익계산서 (Income Statement)
 - 매출액
 - 매출원가
 - 매출총이익
 - 판관비
 - 영업이익
 - 영업외수익/비용
 - 당기순이익
- **11** 기타
 - 근로자수 / 입사자수 / 퇴사자수
 - 업종명

- 제무상태표 (Balance Sheet)
 - 유동자산 / 비유동자산 / 총자산
 - 유동부채 / 비유동부채 / 총부채
 - 자본금 / 자본잉여금 / 이익잉여금 / 총자본
- 자본변동표 (Statement of Changes in Equity)
 - 자본조정
 - 기타포괄손익누계액

☑ 대시보드 목표

- 사업자번호, 기업명 기반 검색
- 4가지 분석 차트 제공
 - 매출 증가율, 영업 이일률, 자기자본이익률, 부채비율
 - 개별 기업 수치, 산업 평균 수치
- 기업 재무데이터 제공
 - 기업명, 사업자번호, 법인번호
 - 기준연도, 기업 상태 (정상/휴업/폐업 등)

업종 · 기업 · 연도별 재무 지표를 활용한 데이터 기반 분석



☑ 라이브러리 설치

☑ 구글 드라이브 연동

☑ 구글 드라이브 연동

```
# ▼ 한글 폰트 설치 및 설정 (Colab 재시작 시마다 실행 필요)
import matplotlib.pyplot as plt
import matplotlib.font_manager as fm
import os
# 폰트 설치
!sudo apt-get install -y fonts-nanum
# 설치된 Nanum 폰트 경로 확인
font_dirs = ['/usr/share/fonts/truetype/nanum/']
font_files = fm.findSystemFonts(fontpaths=font_dirs)
# 폰트 매니저에 폰트 추가
for font_file in font_files:
   fm.fontManager.addfont(font_file)
# 폰트 이름 확인
nanum_font = fm.FontProperties(fname=font_files[0]).get_name()
# 폰트 설정
plt.rcParams['font.family'] = nanum_font
plt.rcParams['axes.unicode_minus'] = False
print(f"♥️ 한글 폰트 설치 및 설정 완료! 적용 폰트: {nanum_font}")
```

☑ 데이터 DB에 저장.

```
# 2. 데이터 로드 및 DuckDB 저장
print("\n∰ 기업 재무 데이터 로드 및 DuckDB 저장 시작!")
import pandas as pd
import matplotlib.pyplot as plt
import duckdb
# ▼ CSV 파일 경로 (Google Drive)
file_path = '/content/drive/MyDrive/company_sample_10000.csv'
# 🚺 데이터 로드
print("🃤 CSV 파일에서 데이터 로딩 중...")
company_df = pd.read_csv(file_path, encoding='utf-8-sig')
print(f"♥️데이터 로드 완료! 총 {len(company_df)} 행")
# V DuckDB 연결
con = duckdb.connect(database='company_data.duckdb', read_only=False)
# 🚺 DuckDB 저장
print("槽 DuckDB 에 데이터 저장 중...")
con.execute("CREATE OR REPLACE TABLE company_data AS SELECT * FROM company_df")
print("♥ DuckDB 저장 완료!")
con.close()
```

☑ 데이터 로딩

```
# ✔ 데이터 로드 및 DuckDB 연결
import pandas as pd
import duckdb

print("\n  DuckDB 연결 및 데이터 로딩 중...")

con = duckdb.connect(database='company_data.duckdb', read_only=False)

company_df = con.execute("SELECT * FROM company_data").df()

con.close()

print("  데이터 로딩 및 DuckDB 연결 종료!")
```

☑ 성장성,수익성 계산

```
# 🗸 성장성 계산: 매출 증가율 & 영업이익 증가율
print("\n# 성장성 계산 중...")
company_df = company_df.sort_values(by=['기업명', '기준연도'])
company_df['손익계산서_매출액'] = pd.to_numeric(company_df['손익계산서_매출액'], errors='coerce')
company_df['손익계산서_영업이익'] = pd.to_numeric(company_df['손익계산서_영업이익'], errors='coerce')
company_df['매출증가율'] = company_df.groupby('기업명')['손익계산서_매출액'].pct_change() * 100
company_df['영업이익증가율'] = company_df.groupby('기업명')['손익계산서_영업이익'].pct_change() * 100
print("√ 성장성 계산 완료!")
# 🗸 수익성 계산: 영업이익률
print("\n# 수익성 계산 중...")
company_df['영업이익률'] = (company_df['손익계산서_영업이익'] / company_df['손익계산서_매출액']) * 100
print("√ 수익성 계산 완료!")
# 🗸 계산된 데이터 DuckDB 에 저장
print("\n DuckDB 에 갱신된 데이터 저장 중...")
con = duckdb.connect(database='company_data.duckdb', read_only=False)
con.execute("CREATE OR REPLACE TABLE company_data AS SELECT * FROM company_df")
con.close()
print("♥ DuckDB 저장 완료 및 연결 종료!")
```

☑ 4분면 분석

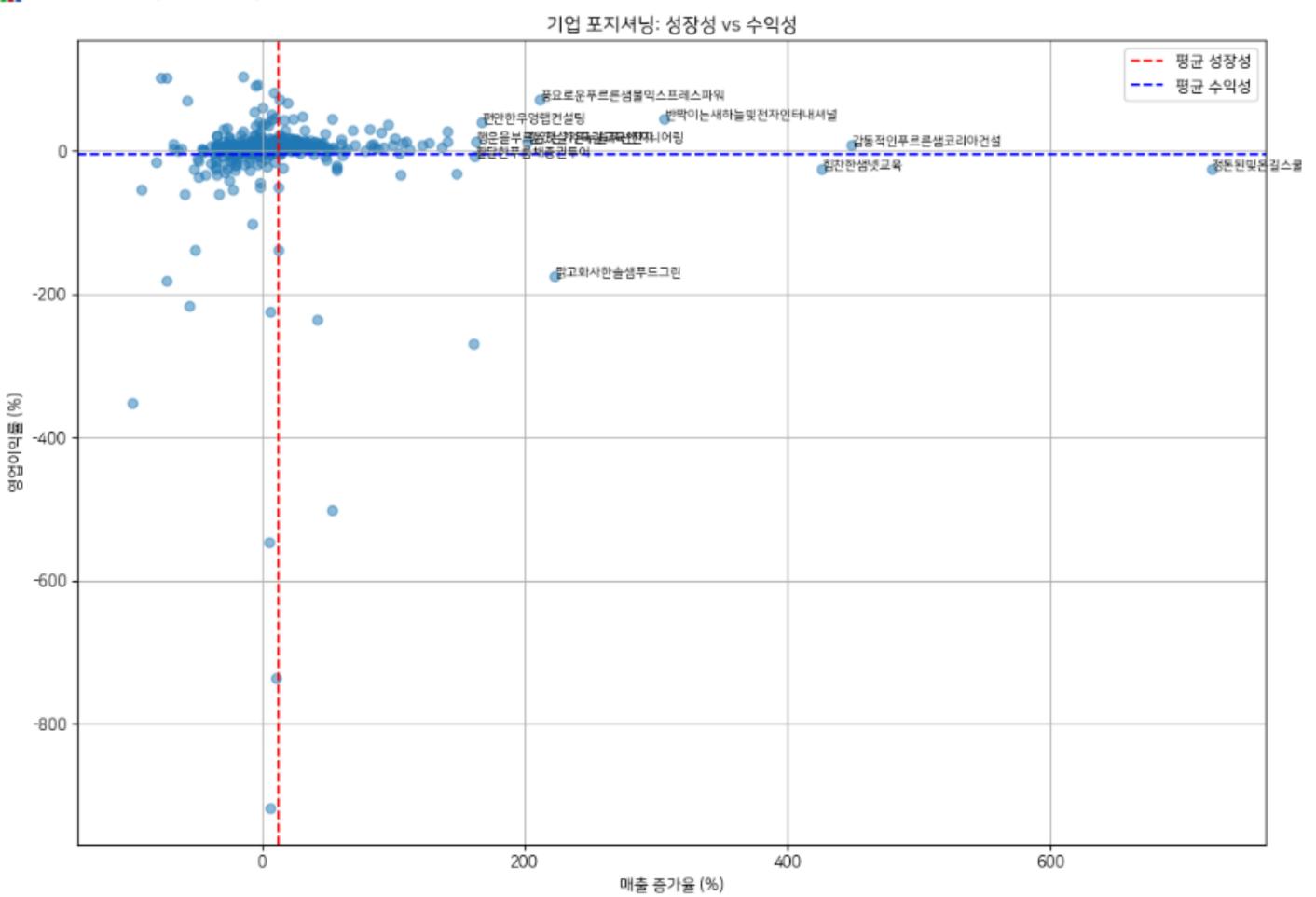
```
# 🔽 DuckDB 재연결 및 최신 연도 데이터 로딩
con = duckdb.connect(database='company_data.duckdb', read_only=False)
query = """
   SELECT *
   FROM company_data
   WHERE 기준연도 = (SELECT MAX(기준연도) FROM company_data)
     AND 매출증가율 IS NOT NULL
     AND 영업이익률 IS NOT NULL
latest_df = con.execute(query).df()
con.close()
print(f" ✓ 최신 연도 데이터 로딩 완료! 대상 기업 수: {len(latest_df)}")
# 🗸 데이터 클렌징 (무한대, NaN, 극단값 제거)
print("\n	 데이터 클렌징 중 (무한대, NaN, 극단값 제거) ...")
latest_df = latest_df.replace([float('inf'), float('-inf')], pd.NA)
latest_df = latest_df.dropna(subset=['매출증가율', '영업이익률'])
latest_df = latest_df[(latest_df['매출증가율'].abs() <= 1000) & (latest_df['영업이익률'].abs() <= 1000)]
print(f"♥️ 데이터 클렌징 완료! 분석 대상 기업 수: {len(latest_df)}")
# 🗸 평균 값 계산 (4분면 기준)
avg_growth = latest_df['매출증가율'].mean()
avg_profitability = latest_df['영업이익률'].mean()
print(f" de real growth:.2f}%")
print(f" ■ 평균 수익성 (영업이익률): {avg_profitability:.2f}%")
# 🗸 시각화
plt.figure(figsize=(12, 8))
# 산점도
plt.scatter(latest_df['매출증가율'], latest_df['영업이익률'], alpha=0.5)
# 4분면 기준선
plt.axvline(x=avg_growth, color='red', linestyle='--', label='평균 성장성')
plt.axhline(y=avg_profitability, color='blue', linestyle='--', label='평균 수익성')
```

```
# 텍스트 라벨 추가 (상위 10개 기업만)
top_companies = latest_df.nlargest(10, '매출증가율')
for _, row in top_companies.iterrows():
    plt.text(row['매출증가율'], row['영업이익률'], row['기업명'], fontsize=8)
plt.title('기업 포지셔닝: 성장성 vs 수익성')
plt.xlabel('매출 증가율 (%)')
plt.ylabel('영업이익률 (%)')
plt.legend()
plt.grid(True)
plt.tight_layout()
plt.show()
print("▼ 4분면 분석 시각화 완료!")
print("\n) 4분면 분석 완료! 다음 단계로 넘어갑니다. (고급 분석 또는 대시보드 구현)")
# 🗸 성장성 수익성이 모두 높은 상위 20개 기업 테이블 출력
high_growth_profit = latest_df[
    (latest_df['매출증가율'] > avg_growth) & (latest_df['영업이익률'] > avg_profitability)
].sort_values(by=['매출증가율', '영업이익률'], ascending=False).head(20)
print("\n  성장성 & 수익성 모두 높은 Top 20 기업:")
from tabulate import tabulate
# ▼ 성장성 & 수익성 모두 높은 Top 20 기업 테이블 출력 (pandas Styler 활용)
top_companies = high_growth_profit.copy().reset_index(drop=True)
top_companies['순위'] = range(1, len(top_companies) + 1)
# 포맷팅: 소수점 2자리, % 추가
top_companies['매출증가율'] = top_companies['매출증가율'].apply(lambda x: f"{x:.2f}%")
top_companies['영업이익률'] = top_companies['영업이익률'].apply(lambda x: f"{x:.2f}%")
# 컬럼 순서 정리
top_companies = top_companies[['순위', '기업명', '기준연도', '매출증가율', '영업이익률']]
# Styler 로 출력
top_companies.style.set_table_attributes("style='display:inline'").set_caption("📊 성장성 & 수익성 모두 높은 Top 20 기업 🚀")
```



💅 성장성 vs 수익성 4분면 분석 시작 (DuckDB 조회)! ▼ 최신 연도 데이터 로딩 완료! 대상 기업 수: 456

✓ 데이터 클렌징 중 (무한대, NaN, 극단값 제거) ...
 ✓ 데이터 클렌징 완료! 분석 대상 기업 수: 449
 Ⅰ 평균 성장성 (매출 증가율): 12.48%
 Ⅰ 평균 수익성 (영업이익률): -4.37%



☑ 기업별 데이터 시각화

```
# 🗸 기업별 성장성 / 수익성 시계열 시각화 (사업자번호 기준 + DuckDB에서 조회)
import matplotlib.pyplot as plt
import duckdb
# 🗸 데이터 로드 및 DuckDB 연결
import pandas as pd
print("\n# DuckDB 연결 및 데이터 로딩 중...")
con = duckdb.connect(database='company_data.duckdb', read_only=False)
# 분석할 기업 사업자번호 선택 (예: '8068088819')
selected_biz_no = '1682284128'
# V DuckDB 에서 데이터 조회
query = f"""
SELECT *
FROM company_data
WHERE 사업자번호 = '{selected_biz_no}'
ORDER BY 기준연도
selected_company_df = con.execute(query).df()
# DuckDB 접속 종료
con.close()
print("♥️ 데이터 로딩 및 DuckDB 연결 종료!")
# 🚺 매출 증가율 시계열
plt.figure(figsize=(10, 4))
plt.plot(selected_company_df['기준연도'], selected_company_df['매출증가율'], marker='o', linestyle='-', label='매출 증가율 (%)', color='skyblue')
plt.title(f"✓ 사업자번호 {selected_biz_no} - 매출 증가율 추이")
plt.xlabel('기준연도')
plt.ylabel('매출 증가율 (%)')
plt.grid(True)
plt.legend()
plt.show()
# 🔽 영업이익률 시계열
plt.figure(figsize=(10, 4))
plt.plot(selected_company_df['기준연도'], selected_company_df['영업이익률'], marker='o', linestyle='-', label='영업이익률 (%)', color='salmon')
plt.title(f"✓ 사업자번호 {selected_biz_no} - 영업이익률 추이")
plt.xlabel('기준연도')
plt.ylabel('영업이익률 (%)')
plt.grid(True)
plt.legend()
plt.show()
print("✓ 기업별 시계열 분석 완료!")
```

☑ 대시보드 app.py

```
app.py X
1 ### app.py
2 import streamlit as st
 3 import pandas as pd
4 import duckdb
 5 import matplotlib.pyplot as plt
6 import matplotlib.font_manager as fm
7 # Streamlit 페이지 설정
8 st.set_page_config(page_title='기업 재무 분석 대시보드', layout='wide')
9 st.title('<mark>▮</mark> 기업 재무 분석 대시보드')
10 st.markdown('사업자번호 또는 기업명을 입력하여 기업을 검색하고 성장성 및 수익성 시계열을 확인하세요.'
11 # 설치된 Nanum 폰트 경로 확인
12 font_dirs = ['/usr/share/fonts/truetype/nanum/']
13 font_files = fm.findSystemFonts(fontpaths=font_dirs)
14
15 # 폰트 매니저에 폰트 추가
16 for font_file in font_files:
      fm.fontManager.addfont(font_file)
18
19 # 폰트 이름 확인
20 nanum_font = fm.FontProperties(fname=font_files[0]).get_name()
22 # 폰트 설정
23 plt.rcParams['font.family'] = nanum_font
24 plt.rcParams['axes.unicode_minus'] = False
26 # ☑ DuckDB 연결 및 데이터 로딩
27 @st.cache_data
28 def load_data():
      con = duckdb.connect(database='company_data.duckdb', read_only=False)
      df = con.execute("SELECT * FROM company data").df()
      con.close()
      return df
34 company_data = load_data()
36 # ✓ 사용자 입력: 사업자번호 또는 기업명 검색
37 search_input = st.text_input('Q 사업자번호 또는 기업명을 입력하세요')
```

```
app.py X
39 if search_input:
      # 🔽 필터링된 데이터
      filtered_data = company_data[
          company_data['사업자번호'].astype(str).str.contains(search_input) |
          company_data['기업명'].str.contains(search_input, na=False)
      ].drop_duplicates(subset=['사업자번호'])
45
      if not filtered_data.empty:
          st.write(f"♪ {len(filtered_data)}개의 기업이 검색되었습니다.")
48
          # 🚺 기업명 리스트로 selectbox 생성
          selected_company_name = st.selectbox('▼ 분석할 기업명을 선택하세요:', filtered_data['기업명'].unique())
49
50
          # 🗸 선택한 기업명에 해당하는 사업자번호 찾기
51
          selected_biz_no = filtered_data[filtered_data['기업명'] == selected_company_name]['사업자번호'].values[0]
52
53
          # 🗸 선택한 기업 데이터 필터링
54
          company_df = company_data[company_data['사업자번호'] == selected_biz_no].sort_values('기준연도')
55
56
          # 🗸 선택된 기업명 및 사업자번호 표시
57
          selected_company_name = company_df['기업명'].iloc[0] if not company_df.empty else '정보 없음'
58
          selected_industry = company_df['업종'].iloc[0] if not company_df.empty else '정보 없음'
59
          st.markdown(f"### ★ 선택된 기업: {selected_company_name} (사업자번호: {selected_biz_no})")
60
          st.markdown(f"#### ◆ 소속 업종: {selected_industry}")
61
          # 🗸 업종 평균 계산용 데이터 필터링
62
          industry_df = company_data[company_data['업종'] == selected_industry].groupby('기준연도').mean(numeric_only=True).reset_index()
63
64
          # 🗸 차트 가로로 배치
65
          col1, col2 = st.columns(2)
67
68
          with col1:
              st.subheader(' / 매출 증가율 추이')
69
70
              fig, ax = plt.subplots(figsize=(6, 4))
71
              ax.plot(company_df['기준연도'], company_df['매출증가율'], marker='o', linestyle='-', color='skyblue', label='선택 기업')
              ax.plot(industry_df['기준연도'], industry_df['매출증가율'], marker='x', linestyle='--', color='gray', label='업종 평균')
72
              ax.set_title("매출 증가율 (%)")
              ax.set xlabel('기준연도')
74
75
              ax.set_ylabel('매출 증가율 (%)')
76
              ax.legend()
77
              ax.grid(True)
              st.pyplot(fig)
78
79
```

☑ 대시보드 app.py

```
with col2:
80
               st.subheader('✓ 영업이익률 추이')
81
82
               fig, ax = plt.subplots(figsize=(6, 4))
83
               ax.plot(company_df['기준연도'], company_df['영업이익률'], marker='o', linestyle='-', color='salmon', label='선택 기업')
               ax.plot(industry_df['기준연도'], industry_df['영업이익률'], marker='x', linestyle='--', color='gray', label='업종 평균')
84
               ax.set_title("영업이익률 (%)")
85
               ax.set_xlabel('기준연도')
86
87
               ax.set_ylabel('영업이익률 (%)')
88
               ax.legend()
89
               ax.grid(True)
               st.pyplot(fig)
90
91
           # # 
추가 지표: ROE 및 부채비율 시계열
92
           col3, col4 = st.columns(2)
93
94
95
           with col3:
               st.subheader('✓ ROE (자기자본이익률) 추이')
96
97
               fig, ax = plt.subplots(figsize=(6, 4))
98
               company_df['ROE'] = (company_df['손익계산서_당기순이익'] / company_df['재무상태표_자본총계']) * 100
99
               industry_df['ROE'] = (industry_df['손익계산서_당기순이익'] / industry_df['재무상태표_자본총계']) * 100
               ax.plot(company_df['기준연도'], company_df['ROE'], marker='o', linestyle='-', color='green', label='선택 기업')
100
               ax.plot(industry_df['기준연도'], industry_df['ROE'], marker='x', linestyle='--', color='gray', label='업종 평균')
101
               ax.set_title("ROE (%)")
102
               ax.set_xlabel('기준연도')
103
104
               ax.set_ylabel('ROE (%)')
105
               ax.legend()
               ax.grid(True)
106
               st.pyplot(fig)
107
102
```

☑ 대시보드 app.py

```
108
109
          with col4:
              st.subheader(' / 부채비율 추이')
110
111
              fig, ax = plt.subplots(figsize=(6, 4))
112
              company_df['부채비율'] = (company_df['재무상태표_부채총계'] / company_df['재무상태표_자본총계']) * 100
              industry_df['부채비율'] = (industry_df['재무상태표_부채총계'] / industry_df['재무상태표_자본총계']) * 100
113
              ax.plot(company_df['기준연도'], company_df['부채비율'], marker='o', linestyle='-', color='purple', label='선택 기업')
114
              ax.plot(industry_df['기준연도'], industry_df['부채비율'], marker='x', linestyle='--', color='gray', label='업종 평균')
115
              ax.set_title("부채비율 (%)")
116
117
              ax.set_xlabel('기준연도')
118
              ax.set_ylabel('부채비율 (%)')
              ax.legend()
119
120
              ax.grid(True)
121
              st.pyplot(fig)
122
          # 🗸 데이터 테이블 표시
123
           st.subheader(' 🗐 기업 데이터 테이블')
124
125
           st.dataframe(company_df)
126
          st.success('V) 대시보드가 성공적으로 로드되었습니다!')
127
128
       else:
           st.warning("검색 결과가 없습니다. 다시 입력해주세요.")
129
130 else:
       st.info("사업자번호 또는 기업명을 입력해주세요.")
132
133
134
```

☑ 서비스 실행

```
# ngrok으로 공개 URL 생성 (Authtoken 필요)
from pyngrok import ngrok

# 여기에 발급받은 ngrok authtoken 입력하세요!
NGROK_AUTH_TOKEN = "2v05JnccD4JQu79tsOyegAHBhvv_29cB4wkhHz1CQpqsbs9DQ"
ngrok.set_auth_token(NGROK_AUTH_TOKEN)

print("[INFO] ngrok 터널 생성 중...")
public_url = ngrok.connect(addr="8501", proto="http")
print(f"[INFO] Streamlit 앱 Public URL: {public_url}")

# Streamlit 앱 실행
!streamlit run app.py --server.port 8501 --server.enableCORS false

print("[INFO] Streamlit 웹 서비스가 실행되었습니다. 위의 Public URL을 클릭하여 접속하세요!")
```

업종 · 기업 · 연도별 재무 지표를 활용한 데이터 기반 분석

재무 분석 실습

☑ 대시보드 완성

☑ 기업 재무 분석 대시보드

사업자번호 또는 기업명을 입력하여 기업을 검색하고 성장성 및 수익성 시계열을 확인하세요. 🔍 사업자번호 또는 기업명을 입력하세요 🔎 19개의 기업이 검색되었습니다. ☑ 분석할 기업명을 선택하세요: 귀중한라온제메카소재 💉 선택된 기업: 귀중한라온제메카소재 (사업자번호: 3168902466) ◆ 소속 업종: 기타 공동 주택 건설업 📈 매출 증가율 추이 🖊 영업이익률 추이 매출 증가율 (%) 영업이익률 (%) ── 선택 기업 400 -×- 업종 평균 -25 300 -50 · 환 200 -75 ΠĮN ਨੂੰ -100 --125 -150 **──** 선택 기업 -175 **│ -×-** 업종 평균 2014 2016 2018 2020 2014 2016 2018 2020 2022 기준연도 기준연도

