

🔥 Warm Up 🔥

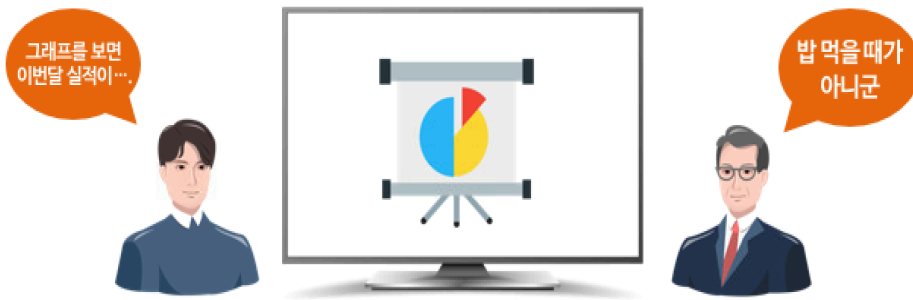
😬 왜 탐색적 데이터 분석(EDA)이 중요해요?

😊 EDA를 통해 데이터를 요약하고 시각화하여 의사결정을 지원합니다.

데이터 위주의 보고



데이터 시각화를 통한 보고



📌 시각화를 통한 보고와 EDA(탐색적 데이터 분석)를 통한 전처리 단계 비교

특징	시각화를 통한 보고	EDA(탐색적 데이터 분석)를 통한 전처리
목적	데이터의 주요 정보를 시각적으로 전달	데이터의 구조와 특성을 이해하고, 분석을 위한 데이터 준비
주요 활동	그래프와 차트 작성, 데이터 패턴과 추세 시각화	데이터 분포 확인, 이상치 탐지, 변수 간 관계 파악
사용 도구	다양한 차트 및 그래프 도구 (예: 막대 그래프, 파이 차트)	통계적 방법 및 다양한 시각화 기법 (예: 히스토그램, 상자 그림)
결과물	인사이트와 결론 도출, 의사결정에 필요한 시각적 자료	데이터 정제 및 변환, 분석에 필요한 데이터셋 구축
주요 대상	경영진, 비즈니스 이해 관계자, 의사결정권자	데이터 분석가, 데이터 과학자
데이터 분석 범위	요약된 데이터나 주요 메트릭 중심	전체 데이터셋을 대상으로 한 상세 분석
시각화 목적	데이터를 이해하기 쉽고 명확하게 전달	데이터의 특성과 문제점을 파악하고, 분석 방향 설정

특징	시각화를 통한 보고	EDA(탐색적 데이터 분석)를 통한 전처리
결과 활용	보고서, 프레젠테이션, 비즈니스 전략 수립	데이터 모델링 의사결정, 추가 분석 작업, 데이터 품질 향상
초점	주요 결과와 KPI(Key Performance Indicator)에 집중	데이터 탐색, 이상치 식별, 데이터 분포 및 상관 관계 분석
예시	매출 현황 보고, 시장 점유율 그래프, 성과 리뷰	결측값 처리, 데이터 변환, 특성 공학(feature engineering)
시간 투자	상대적으로 짧고 주기적인 분석	초기 분석 단계에서 상당한 시간 투자, 데이터 이해를 위한 노력

👁️ Definition

EDA란 데이터분석의 첫 번째 단계로, 데이터의 특징과 데이터에 내재된 관계를 알아내기 위해 통계적 분석과 그래프를 활용해 탐구하는 것

- 분석하고자 하는 데이터의 유형과 범위, 수준을 확인
- 데이터의 분포를 통해 의미를 추출
- 분석 목적을 달성하기 위한 적절한 알고리즘을 선택하기 위한 자료로 사용됨