[6기]\_데이터 엔지니어링 스타터 키트\_2주차 과제(1)

주제 : 가상의 데이터 인프라 구축

* rwa \_data : collect data
* Analytics : summary table

서비스 : 가자여기(위치기반 직장인을 위한 점심메뉴 추천)

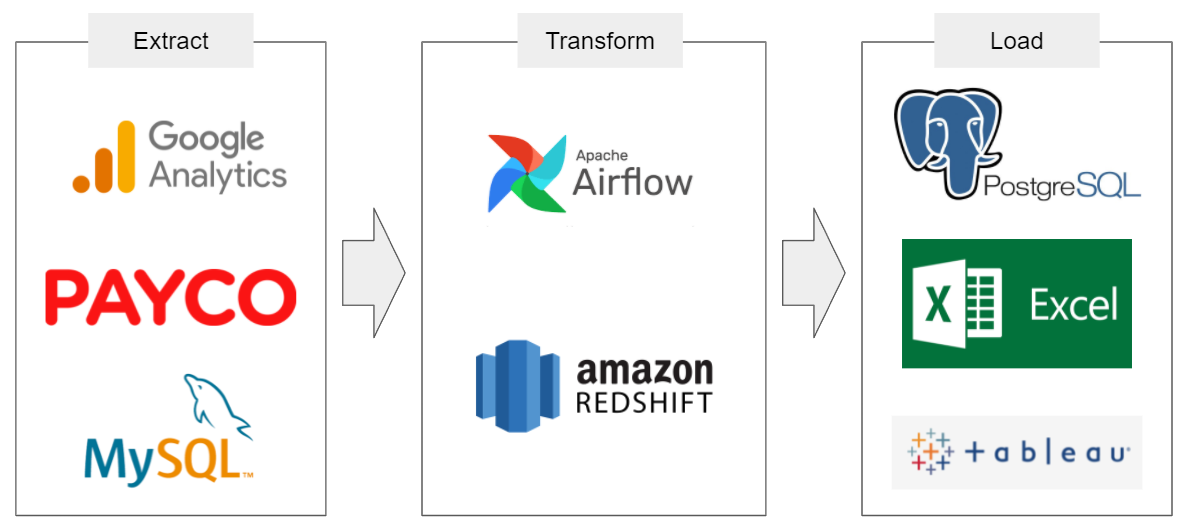
타겟

* 점심시간 전부터 메뉴를 고민하는 회사원
* 여기 주변이 무엇이 맛있는 모르겠지만 실패하고 싶지 않은 외부인
* 음식점을 찾아가는 게 아니라 현재 주변의 맛집을 알고싶은 우리

서비스 이용

* 음식점(고객1) :
  + 등록 : 가게 정보 / 이벤트 및 정보
  + 획득 : 고객 유치 / 주변 상권 정보(상권 분석 및 트렌드 파악)
* 배고픈 회사원(고객2)
  + 등록 : 현재 위치 / 사전 정보 입력(취향, 좋아하는 가게, 메뉴) / 식사 후 정보
  + 획득 : 메뉴 고민 감소 / 식사 만족도 향상

데이터 인프라



[수집]

* Google analytics : 사용자 이벤트 트래킹
* Payco : 페이먼트 솔루션
* ProductionDB : 서비스 제공을 위한 DB(backend)

[처리]

* DATA PIPELINE : 수집, 처리 및 적재 프로세스
* DATAWARE HOUSE : 서비스 제공하며 발생하는 모든 데이터 저장
* AIRFLOW : PIPELINE 스케줄러 \_ 자동화

[적재/분석]

* RDB : 메타 데이터 저장
* EXCEL : 문서 형태로 데이터 제공
* BI : 지표 및 데이터 시각화

[Raw\_data(수집 데이터)]

* 이용자 접속 데이터

접속id / 유저id / 세션id / 채널id

* 이용자 초기 등록 데이터
* 이용자 검색 데이터

검색id / 유저id / 일자 / 키워드

* 이용자 예약 데이터

예약id / 유저id / 일자 / 인원 / 결제id

* 이용자 결제 데이터

결제id / 유저id / 가게id / 메뉴id / 금액 / 결제솔루션id

* 음식점 초기 등록 데이터
* 음식점 운영 데이터

[Analytics\_data(정리 데이터)]

* 시간별/일별/요일별/월별 주요 검색

고유id / 일자 / 키워드 / 횟수

* 시간별/일별/요일별/월별 주요 예약

예약id / 일자 /결제솔루션id / 횟수

* 음식 카테고리 주요 판매 메뉴

고유id / 카테고리 / 음식 / 빈도(%)

* 이용자별 만족도/방문빈도

회원id / 가게id / 방문빈도 / 만족도