

4. Startup Data analysis (프리랜서 오픈마켓 스타트업)

○ Data

- `conversion.head()` → (134244, 19)
 - 웹사이트/모바일 서비스에서 활동한 모든 활동 기록(activity)을 포함한 데이터
 - app event category, eventdatetime, devicetype, osversion, params_campaign 등 컬럼으로 구성
- `funnel.head()` → (53, 6)
 - 사용자가 처음 유입(Acquistion) 되었을 때부터 구매(Revenue)를 할 때 까지의 과정을 구조화 한 데이터
- `category.head()` → (245, 9)
 - 디자인, 번역, 콘텐츠 제작 등 다양한 상품에 대하여 그룹화하고 정리된 카테고리를 포함한 데이터

○ Environment

- Python 3.73 with `jupyter==1.0.0`
- pandas, numpy, matplotlib

➤ Perpose

데이터 클린징(Cleansing)

- 데이터를 분석하기 전 가장 중요한 과정인 데이터 클리닝(Data Cleaning)을 맡아 진행함
- 웹, 아이폰, 안드로이드등 다양한 디바이스와 OS로 서비스를 이용하고 있는 사용자의 웹,앱 행동 데이터를 정리한 뒤, pandas 의 merge나 concat등을 활용해 하나로 묶어주는 작업을 진행함

➤ Solution

- Event datetime Column 연/월/일/시/분/초 추출
- App Os version Column Cleansing : IOS, Android 등
- Device manufacturer Column Cleansing : Samsung, LG, Apple 등
- Event channel Column Cleansing : google, naver, daum 등
- App event label Column Cleansing
- conversion-funnel-category Column 삭제, 변경, 정렬, 병합 등