

# 자유 프로젝트 기획안

Data Analytics of Chai Corp.

2022.09.19. 이진규 @LUCI031

# Table of Contents

- 프로젝트 주제 선정
- 데이터 소개
- 분석 방법
- 프로젝트 진행안

# 자유 주제

점심 메뉴도 고르기 힘든 세상에 자유 주제라니!

# 프로젝트 주제 선정

주제를 정하기 어려울 땐 목표를 다시 상기해보자

- 데이터 분석가
  - 현재 진로를 데이터 분석가로 정한 상태
  - 이미 1개의 포트폴리오를 완성
  - 이력서도 얼추 작성한 상태 -> ‘어랏...포트폴리오 칸이 왜 이렇게 텅하지?’
  - 결론: 포트폴리오를 하나 더 만들어야겠다!

# 프로젝트 주제 선정

지난 여름엔 말이죠...

- Grocery Auction
  - [https://github.com/LUCI031/grocery\\_auction](https://github.com/LUCI031/grocery_auction)
  - <농수산물 별 물가 상승 정도 분석 및 경매 가격 분석 시각화>라는 주제로 프로젝트 진행
  - 데이터 전처리, EDA, 시각화까지 최선을 다했고 결과물도 만족!...스럽다만...뭔가...뭔가 모자르다!
  - 뭐가 모자른 걸까...그건 바로!

# INSIGHT

인사이트가...모자르다...

# 프로젝트 주제 선정

## 보기만 해도 아찔해지는 취업 공고들

<의견 전달 방법하는 방법>  
딜라이트룸 노션 블로그 -> [Apply] 탭 -> 딜라이트룸의 첫인상

### 주요 업무

- 400만 0
- 니다.
- 전사에서
- 데이터 분석
- 사용자에게 높은 이해도와 데이터에 기반한 실험 설계 및 ;
- 주요 지표 정의 및 택사노미 관리를 담당해요.
- 주로 사용하게 되는 툴과 기술: Amplitude, Braze, Firebase, BigQuery

### 채용 상세

#### 사용하시게 될 기술들

- Python
- Airflow
- AWS
- Tableau

#### 주요업무

- 전사적으로 데이터에 기반한 최적의 서비스적인, 사업적인 의사결정을 내릴 수 있도록 하는 일련의 활동들을 lead, guide, and support
- 분석을 통한 데이터에 기반한 intelligence와 action item 도출
- Intelligence와 action item이 실제 서비스에 잘 반영될 수 있도록 사업/서비스기획/디자인/개발팀 등과 적

통해 서비스를 개선해본 경험이 있으신 분

## • A/B 테스트와 인과 추론에 대한 통계적 이해를

- 사용자분들께 정말 좋은 서비스, 놀라운 경험을 제공하고 싶은 마음을 가지신 분
- 데이터와 데이터베이스에 관한 관심과 이해를 갖추신 분
- SQL 구현 역량을 갖추신 분
- 모바일, 온라인 서비스를 좋아하고 꾸준히 사용하시는 분
- 서비스의 사용성과 성장을 위해 빠른 시도로 다양한 가설을 설정하실 수 있고, 실험하실 수 있으신 분
- 데이터를 명확하고, 유연하게 커뮤니케이션하고 설득을 할 수 있는 능력을 보유하신 분

#### 우대사항

- 실제 모바일 서비스 관련 데이터 분석 업무 경험이 있으신 분
- Tableau 사용 경험이 있는 분
- 데이터 관련 전공 (통계학/데이터/컴퓨터공학/산업공학 등) 혹은 데이터 관련 교육 이수자
- 많은 양의 데이터를 효율적으로 핸들링할 수 있으신 분
- AARRR, 퍼널 분석, 등 모바일, 웹 서비스에 필요한 데이터 분석 기법에 대한 이해도와 관심을 갖추신 분

# A/B테스트

그게 뭔데...어떻게 하는건데...



# 프로젝트 주제 선정

## 통계학

- 약점 극복
  - (당연하게도) 데이터분석 포지션을 뽑는 기업들은 대부분 통계학 관련 전공이나 통계학에 대한 지식이 있는 사람을 원한다.
  - 공고 란에도 A/B 테스트, Funnel 분석 등 다양한 데이터 분석 기법을 사용하여 업무를 진행한다고 써있는 경우가 많다.
  - 통계학적 지식이 없는 Insight는 깊이가 얕다.
  - 통계학적 지식을 활용한 프로젝트를 해보자!...우선 공부부터 하고
- ...근데 데이터는?

# 데이터 소개

## 어떤 데이터가 적절할까?

- 데이터 선정 조건
  - 통계학적 지식을 활용할 수 있는 ‘숫자형 데이터’
  - 전처리는 전 프로젝트 때 많이 해봤기 때문에 Null 값 등등 전처리할 값들이 많지 않은 데이터
  - 데이터는 클수록 좋다! 크기가 충분한 데이터
- 이러한 조건을 토대로 각종 공공 데이터를 찾아봤지만 만만치 않다...그러다 문득 생각난 데이터셋이 있었으니...

# 데이터 소개

## Chai

- 데이터 설명
  - 은행 계좌 기반의 간편 결제 서비스 Chai의 데이터
  - 2019.08~2020.03월까지 카드 사용에 따른 할인 금액 데이터
  - 거래 한 건당 고객 정보, 가입 정보, 할인 금액, 사용 날짜, 상업 코드 등을 알 수 있음
  - 데이터 갯수: 100만개 (라고 적혀있었으나 실제 데이터는 56만개 가량)
- 완벽하게 계획과 맞아떨어지는 데이터셋이므로 Chai 데이터를 활용하기로 결정

# 데이터 소개

## 분석 목표

- 문제 설정: Chai는 2019년 5월에 런칭한 서비스이기 때문에 데이터가 제공된 시점 (2019.08~2020.03)의 경우 데이터 부족으로 주요 타겟을 공략하는 전략의 수립이 어려웠을 것이라 예상.
- 목표: 타겟 고객층 분석 및 고객층 분류를 통해 이용자 수 증가 및 재구매율을 높여 충성 고객 확보, 고객층에 맞는 전략 수립
- 분석 방법: 특징 별 고객층 분류, 소비 패턴, 할인 정도, 푸시 알림 설정 여부 등 여러 요소들을 비교하여 타겟 고객층 발견 및 재구매율 상승 포인트 발견

# 분석 방법

## 티끌 모아 태산

- 데이터 전처리
  - 시계열 데이터, Labeling 등 전처리 진행
  - 데이터 확인 결과 56만개의 데이터 모두 결측치는 없다(!)

# 분석 방법

## 티끌 모아 태산

- 데이터 분석
  - Column 별 탐색 및 분석, 시계열 분석
  - Gender, Push\_Permission, is\_foreigner 등 데이터를 이용한 A/B테스트 등 다양한 가설 수립 및 검증
  - 소비패턴 등 다양한 요소들을 이용하여 고객층 분류
  - 숫자형 데이터이므로 분산, 편차, 분포 등을 활용하여 좀 더 깊이있는 Insight 창출 예정
  - +) 상황에 따라 Clustering도 진행

# 분석 방법

## 티끌 모아 태산

- 시각화
  - Tableau를 이용하여 시각화
  - 여러 기법을 활용한 다각적인 전달보단 분석 결과 표현에 초점

# 프로젝트 진행안

## 천릿길도 한걸음부터

- 기획 (9/5~9/19)
  - 통계학 공부 (누워서 읽는 통계학, 한달 공부 데이터 분석 읽기)
  - 기본적인 데이터 탐색 및 전처리
  - 기획안 발표 준비



# 프로젝트 진행안

## 천릿길도 한걸음부터

- 1차 수행 기간 (9/26~10/7)
  - 통계학 공부
  - 전반적인 EDA 진행
  - 여러가지 시점에서 데이터 활용 (시계열, 고객군 분류, 분석, 가설 검정 등)
  - +) 진척도에 따라 Clustering까지 진행

# 프로젝트 진행안

## 천릿길도 한걸음부터

- 2차 수행 기간 (10/10~10/21)
  - EDA 및 분석 마무리
  - Tableau 이용한 시각화
  - PPT 및 Github 정리 통해 포트폴리오 형태로 마무리

# Thank you

성공적인 프로젝트를 기원합니다!