Data Science Group 1

End to End 7process

1. Object Setting

**Data set Motivation**

: **Food waste** has become a **critical issue** in achieving the Sustainable Development Goals (SDGs) in today’s world. This project aims to systematically **analyze food waste** through data, and from this analysis, **propose efficient operational strategies and waste reduction plans**.

**Furthermore**, it explores the potential to implement an incentive system **based on the amount of food waste generated**, enabling restaurants or operational units to receive rewards according to their waste reduction performance, thus providing a practical and scalable management approach.

+

텀 프로젝트에 맞는 조건(nan값 존재, 등)

데이터 분석 전 가설 -> 온도와 습도, 인적 요인과 같은 feature들이 상관 관계를 가지고 food\_waste\_kg에 영향을 줄것이다~.

Ex\_ 높은 습도와 온도 -> 부패 속도 증가 -> 쓰레기량 증가

**Data set info**

Proposal 참조

+plot을 이용해 시각화 진행(카테고리형 데이터의 시각화)

그래프, 라인, 도표, 스크린샷이(가) 표시된 사진

자동 생성된 설명

도표, 스크린샷, 그래프이(가) 표시된 사진

자동 생성된 설명

수치형 데이터의 데이터 분포

전 처리 진행 전 box plot을 이용하여 Outlier 확인 하기

도표, 직사각형, 평면도, 스크린샷이(가) 표시된 사진

자동 생성된 설명

* 전처리 전 ~ featuer에서 outlier ~개 각각 발견

텍스트, 스크린샷, 도표, 평행이(가) 표시된 사진

자동 생성된 설명

히트맵 이용하여 각 feature 간의 상관관계 도출

진행 전 meals\_Server 와 예측하려는 food\_waste\_kg간의 상관 관계가 높다는 것 확인

실험 전 상관 관계를 계산해보니 막상 환경적 요인 인적요인은 그닥 영향 없음.

기존의 개별 feature 간 분석에서는 온도, 습도, 인력 수가 food\_waste와 약한 상관관계를 보였지만, 클러스터링을 통해 **요인 별 구조를** 도출함으로써, 개별 변수의 노이즈를 제거하고

유사한 환경/운영/인력 조건이라는 그룹 차원에서의 결과를 확인하기로 함.

앞의 motivation과 조화 필요