

# 빅데이터 사이언스 실무자 양성과정

## < 파일럿 프로젝트 착수보고서 >

주 제	대한민국 2017.09.07.~2019.09.08. 유기견 현황 조사 및 분석		
팀 명	Fors	일 자	2019년 9월 10일
팀 장	김태웅 <dxodnd@gmail.com>	팀 원	김봉수, 백서연

### 1. 과제 개요

현재 대한민국의 유기견의 개체 수는 반려견 개체 수에 따라 증가하는 경향을 보이고 있다.

하지만 대부분의 사람들은 실질적인 유기견 현황에 대한 데이터의 부재로 인하여 심각성에 대하여 자각하지 못하고 있다. 공급은 수요에 따라서 이루어지므로 대중들에게 유기견의 현황을 알려 무분별한 유기 및 입양을 방지하고자 한다.

전국의 동물보호센터 데이터를 바탕으로 유기견 개체 수 및 특징을 분석하여 대중들에게 알림으로써 반려견 유기에 대한 경각심을 상기시키고 올바른 입양 문화를 형성하려 한다.

### 2. 데이터 수집 전략

- 동물보호관리시스템 홈페이지의 유기견 데이터 크롤링
- 공공데이터 포털에서 제공하는 동물보호관리시스템 유기견 조회 서비스 데이터 활용 신청

### 3. 데이터 분석 전략

- 특징(품종, 성별, 중성화여부, 나이, 체중)에 따른 유기견 개체 수 상관관계 파악
- 시기, 지역에 따른 유기견 개체 수 상관관계 파악
- 상기 정보를 토대로 정확한 수치 제공 및 데이터 시각화

### 4. 결과 보고 예측 및 실행 계획

#### (1) 일정

09.10.	09.11.	09.12.	09.13.	09.14
데이터 수집	데이터 수집	추석	추석	데이터 전처리
09.15.	09.16.	09.17.	09.18.	09.19.
데이터 분석	데이터 시각화, 발표자료 작성	데이터 시각화, 발표자료 작성	발표 자료 제출, 모의 발표	프로젝트 발표

## (2) 역할 분담

- 김태웅(조장): 데이터 수집, 데이터 분석, 데이터 시각화, 소스코드 작성
- 김봉수: 보고서 작성, 데이터 분석, 데이터 시각화
- 백서연: 발표자료 작성, 데이터 분석, 데이터 시각화

## (3) 가설

- 나이가 더 많은 강아지들이 더 많이 버려질 것이다.
- 질병이 있는 강아지들이 더 많이 버려질 것이다.
- 중성화를 안 한 강아지들이 더 많이 버려질 것이다.
- 특정 날짜에 강아지들이 더 많이 버려질 것이다.
- 체중이 높을수록 강아지들이 더 많이 버려질 것이다.
- 지방에서 강아지들이 더 많이 버려질 것이다.
- 년도에 가장 유행했던 강아지들이 더 많이 버려질 것이다.

### (3.1) 가설에 따른 결과 도출을 위한 전략 수립

- 가설에 있는 데이터들을 pandas, numpy를 이용하여 분석하고 matplotlib을 이용한 시각화

## (4) 파일럿 프로젝트를 통해 얻고 싶은 것

- 팀원들과의 협업력 증진
- 수많은 과정들에서 갈등 발생시 원만한 해결-
- 크롤링과 같은 데이터 수집 능력 향상
- 데이터 마이닝을 통한 파이썬 이해
- 데이터 분석을 통한 Pandas를 이용한 Dataframe, Series, table 이해
- 데이터 시각화를 통한 Matplotlib, seaborn 활용 능력 증진
- 데이터 결과를 이용한 사회적 문제 해결 능력 증진

**감사합니다**