

AI Challenge for Biodiversity

『 **경연** 』 **Jam Study 발표**

맹꿏이2

2022. 12. 28

목 차

1. 팀 소개
2. 주제 선정 배경
3. 데이터 전처리
4. 시각화
5. 결론
6. Q&A

1. 팀 소개

'맹꽁이2'

: 기후변화 생물지표 100종에 포함된 맹꽁이



팀원

기지원 숙명여자대학교 생명시스템학부

김윤진 숙명여자대학교 통계학과

박지원 숙명여자대학교 통계학과

유영서 숙명여자대학교 영어영문학부

이민지 숙명여자대학교 통계학과

2. 주제 선정 배경

시끄러운 매미 울음소리, 야간 조명 원인으로 밝혀져

A 임재경 | © 입력 2021.10.12 15:40 | ⌂ 댓글 0

사실 매미소리는 인간활동과 기후위기에 영향을 받고 있다. 인간이 지구 위 모든 것을 바꿔놓고 있다고는 하지 만, 매미소리까지 영향을 받는다니 의외다.

도시 장악한 우렁찬 '말매미'

사람들 사이에서 매미소리가 과거에 비해 시끄러워졌다는 평이 많은데 이는 우렁차게 우는 말매미가 도시에서 증가한 것이 원인으로 꼽힌다.

핫한 데서 떼창하는 말매미들

무엇보다 기온 상승은 말매미를 필두로 한 '매미 합창단'이 더 힘차게 울 수 있는 기회를 준다. 매미 연구자 윤기 상 박사는 "높은 기온이 매미가 소리를 더 내기 좋은 환경"이라고 뉴스펍권에 말했다.

매미는 기본적으로 신체구조 상 일정 기온 이상이어야 소리를 낼 수 있고, 또 고온이 되면 소리를 내는 개체가 늘어난다. 말매미의 경우 27도 고온부터 '떼창'할 확률이 크게 늘어난다.

한국의 6대 도시 평균 기온은 인간이 유발한 지구가열화로 인해 1912년부터 2017년까지 1.8°C 상승했다. 도 시가 더워질수록 말매미의 떼창은 빈번해질 가능성이 높다는 것이다. 또 반대로 말하자면 말매미가 마음 놓고 울게 된 지금, 말매미 소리는 인간에 의한 생태계 변화를 일상에서 느낄 수 있는 단서로도 볼 수 있다.

윤기상 박사는 기후변화와 매미소리 변화, 혹은 개체수 증가를 현재 상황에서 직접적으로 연관시키기는 어렵지 만 매미 생태에 기후변화 영향이 있을 것이라고는 생각한다고 말했다.

- 매미 생태는 기후의 영향을 받는다고 알려짐
- 지구사랑탐사대 데이터 중 3번째로 관측 수가 많아 풍부한 자료 활용 가능

" 시각화를 이용한 서식지 특성에 따른

매미 종별 개체수 분석 "

→ 매미의 종다양성에 영향을 미치는 요인 분석

→ 분석한 주요 원인을 기반으로 생태계를 진단

3. 데이터 전처리

POSTING_MASTER

IDX : int(11) - PK
EXPEDITION_NAME: varchar(30)
TITLE : varchar(255)
CATEGORY : varchar(255)
CONTENTS : mediumtext
HOW_MANY : varchar(10)
READ_CNT : int(11)
POS_ADDR : varchar(255)
LATITUDE : varchar(11)
LONGITUDE : varchar(11)
WEATHER : varchar(50)
TEMP : varchar(10)
WIND_SPEED : varchar(10)
HUMIDITY : varchar(10)
PRESSURE : varchar(10)
WRITER_ID : varchar(20) - FK (Refer to MEMBER.ID)
WRITE_TIME : timestamp
DEL_YN : varchar(1)
IS_Q : enum('Y','N')
SUBMIT_YN : varchar(1)



POSTING_EARTH_HISTORY

IDX : int(11) - PK
POSTING_IDX : int(11) - FK (Refer to POSTING_MASTER.IDX)
RESCH_1_Q : varchar(100)
RESCH_2_Q : varchar(100)
RESCH_3_Q : varchar(100)
RESCH_4_Q : varchar(100)
RESCH_5_Q : varchar(100)
RESCH_6_Q : varchar(100)
RESCH_7_Q : varchar(100)
RESCH_8_Q : varchar(100)
RESCH_9_Q : varchar(100)
RESCH_10_Q : varchar(100)
RESCH_11_Q : varchar(100)
RESCH_12_Q : varchar(100)
RESCH_1 : varchar(50)
RESCH_2 : text
RESCH_3 : text
RESCH_4 : text
RESCH_5 : text
RESCH_6 : varchar(25)
RESCH_7 : varchar(25)
RESCH_8 : text
RESCH_9 : varchar(25)
RESCH_10 : varchar(25)
RESCH_11 : varchar(25)
RESCH_12 : text

✓ CATEGORY = '매미'

Q. 매미 종류

A. 참매미, 말매미, 애매미, 털매미, 쓰름매미, 유지매미, 늦털매미, 소요산매미
(중복답변 제거)

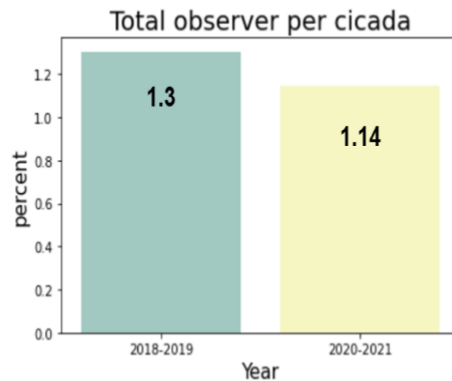
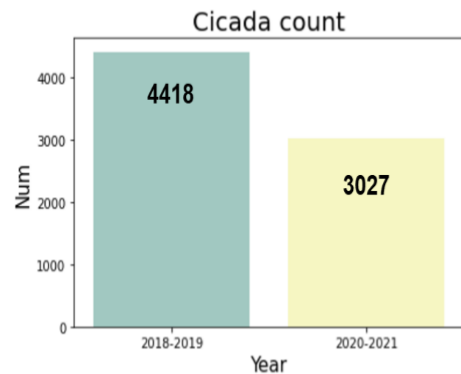
Q. 관찰 장소

- 2020년도를 기준으로, 선택지가 변경됨

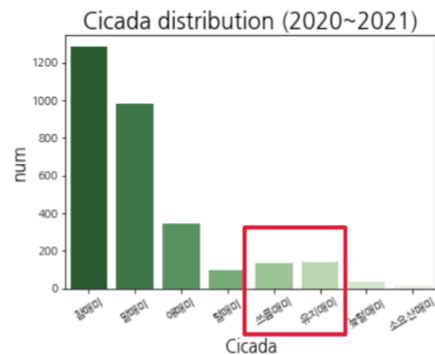
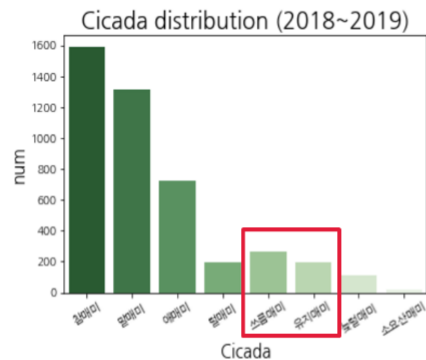
A. 도심, 공원, 숲, 농경지, 습지 - 중복답변 허용 (2018-2019)
주거 단지, 공원 - 중복답변 불가(2020-2021)

⇒ One-Hot Encoding

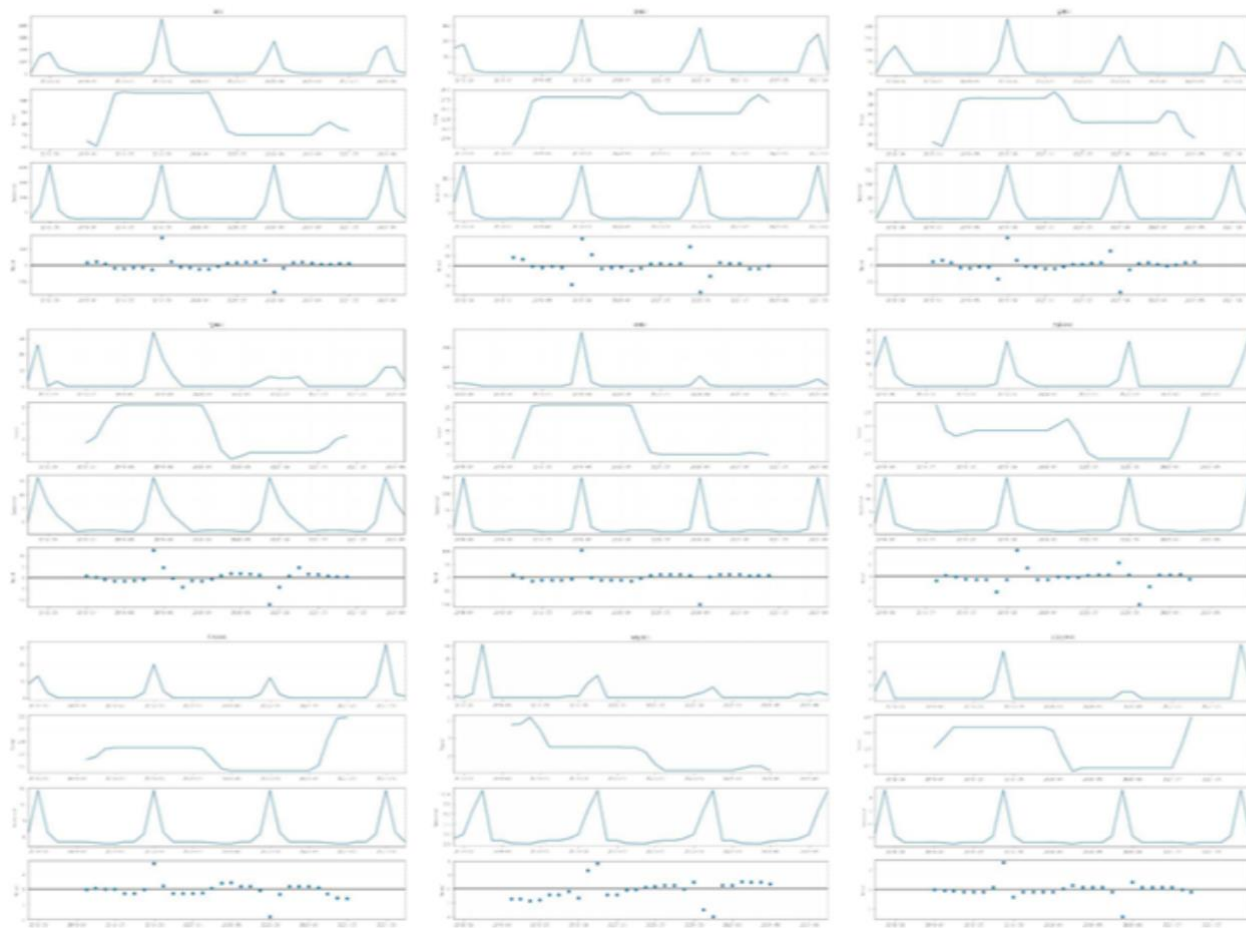
4. 시각화 전체 매미 수 시각화



연도	매미 관찰수	관찰자 수	관찰자당 매미 관찰수
2018-2019	4418	3387	1.3
2020-2021	3027	2648	1.14



4. 시각화 매미 관측 시계열 그래프



✓ 추세

- 2020년도에 종에 관계없이 감소

✓ 계절성

- 참매미, 말매미, 쓰름매미

9월에 피크

- 털매미, 소요산매미

비교적 빠른 시기에 출현, 8월에 피크

- 늦털매미

비교적 늦은 시기에 출현, 10월에 피크

✓ 잔차

여름 초기에 운다고 알려진 털매미, 소요산매미

→ 활동 시기가 늦어짐

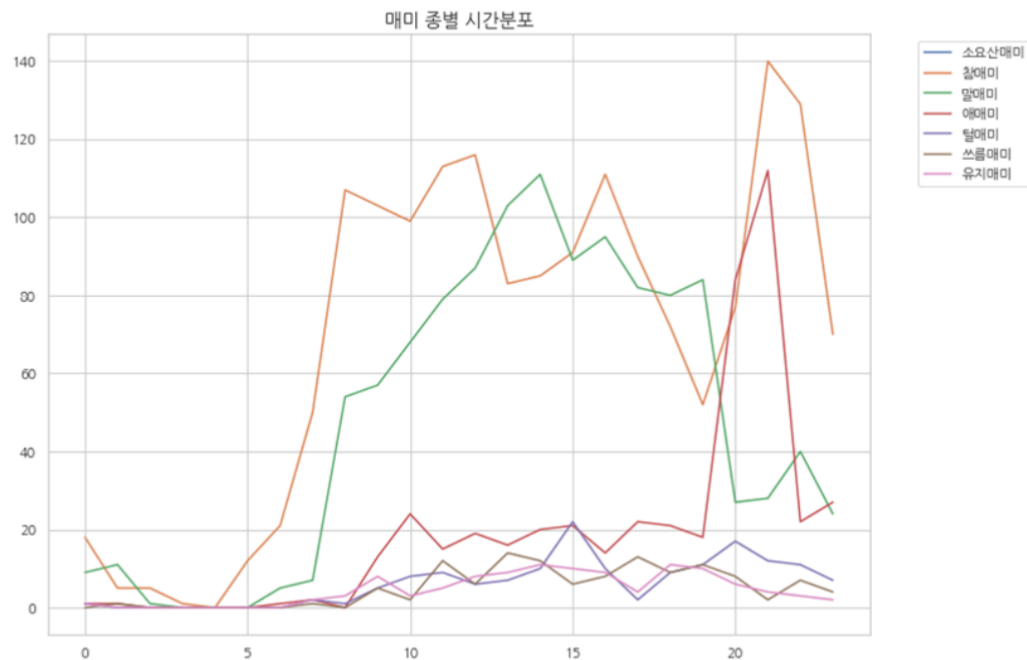
여름 중기에 운다고 알려진 참매미, 말매미, 쓰름매미

여름 말기에 운다고 알려진 늦털매미

→ 활동 시기가 빨라짐

⇒ 매미의 활동기간이 줄어듦

4. 시각화 매미 관측 시계열 그래프

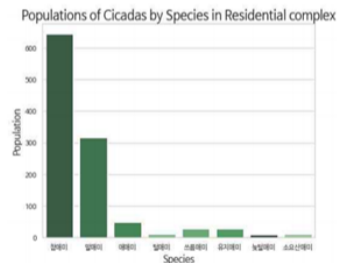
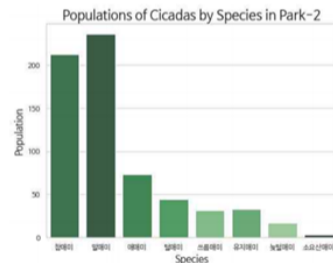
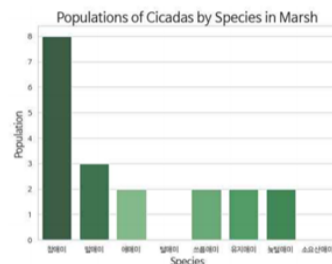
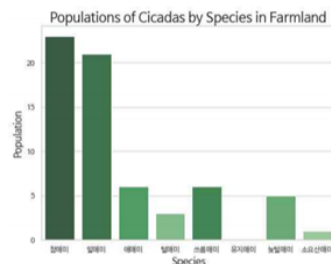
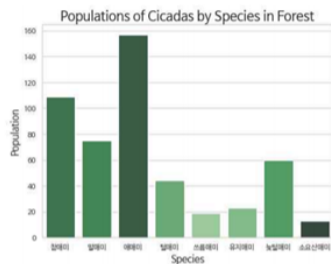
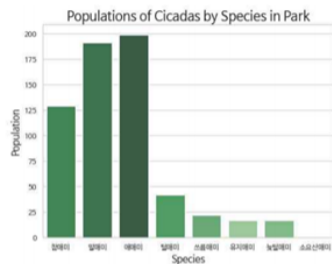
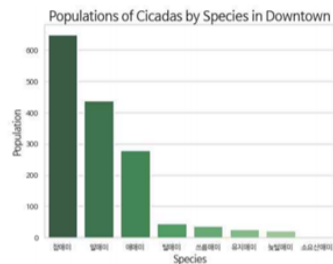


참매미와 말매미가 다른 종들에 비해
이른 시간에 관측되기 시작했고
이어서 유지매미, 쓰름매미, 애매미가 관측됨

→ (김윤재, 기경석. 2018)에서 **도심지역**(서울시)에서
관찰된 매미 일주기와 유사한 형태

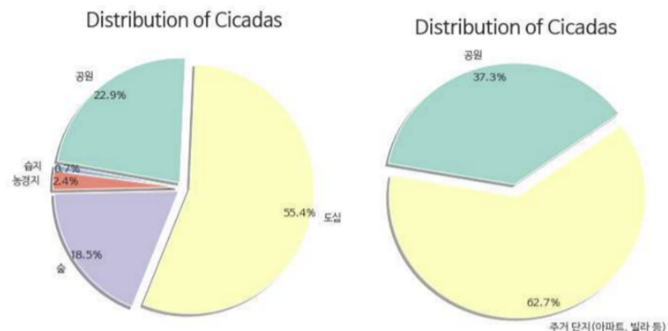
⇒ 서식지별로 나누어 분석할 필요가 있음

4. 시각화 서식지에 따른 매미 종별 분포



도심지일수록 참매미, 말매미의 비율이 우세하고
종별 분포가 균등하지 못함

4. 시각화 서식지에 따른 매미 종별 분포



Simpson 지수

VS

Shannon 지수

종별 개체 수가 얼마나 균등하게 분포하는지,
균등함에 민감

분포하는 종의 수가 얼마나 많은지,
풍성함에 민감

시민과학자 관찰 기록 특성 상, 표본 수가 도심지에 치우쳐져 있어
표본 크기 변화에 상대적으로 영향을 덜 받는

Simpson 지수를 기준으로 분석을 진행

Simpson 지수

: 한 군집으로부터 두 개체를 랜덤하게 추출하였을 때
두 개체가 같은 종에 포함될 확률인 우점도로부터 계산

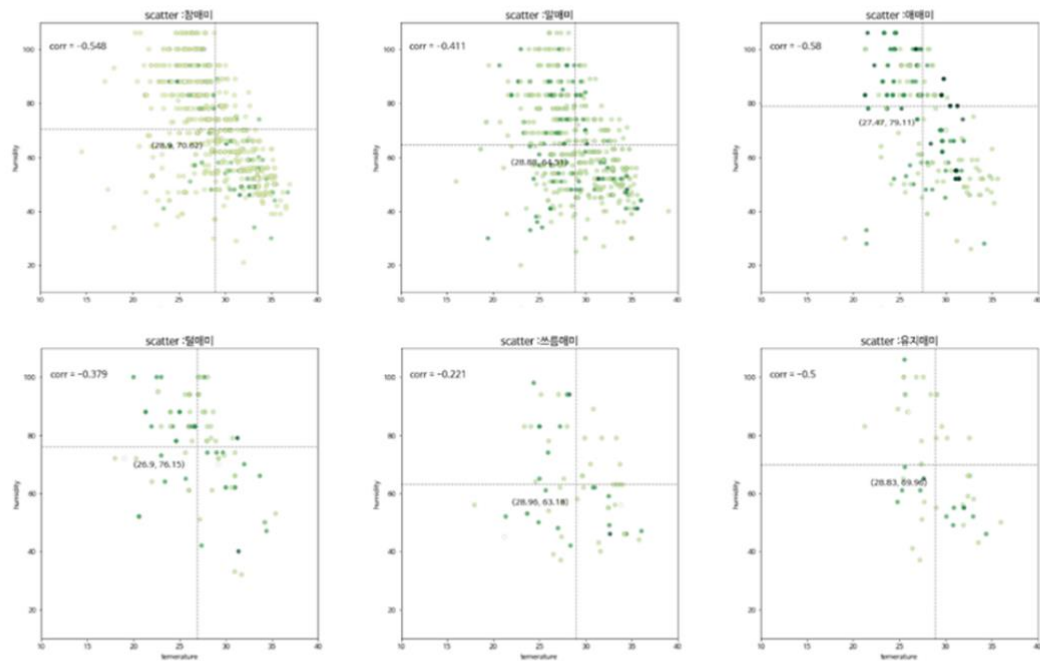
$$C (\text{Simpson의 우점도지수}) = \frac{\sum n_i(n_i-1)}{N(N-1)}$$

(N = total number, n_i : the number of each species)

$$D_c (\text{Simpson의 다양도 지수}) = 1 - C$$

2018-2019					
관측 장소	도시	공원	숲	농경지	습지
Simpson 지수	0.69	0.75	0.81	0.76	0.8
2020-2021					
관측 장소	주거 단지		공원		
Simpson 지수	0.57		0.74		

4. 시각화 기온, 습도에 따른 매미 종별 분포



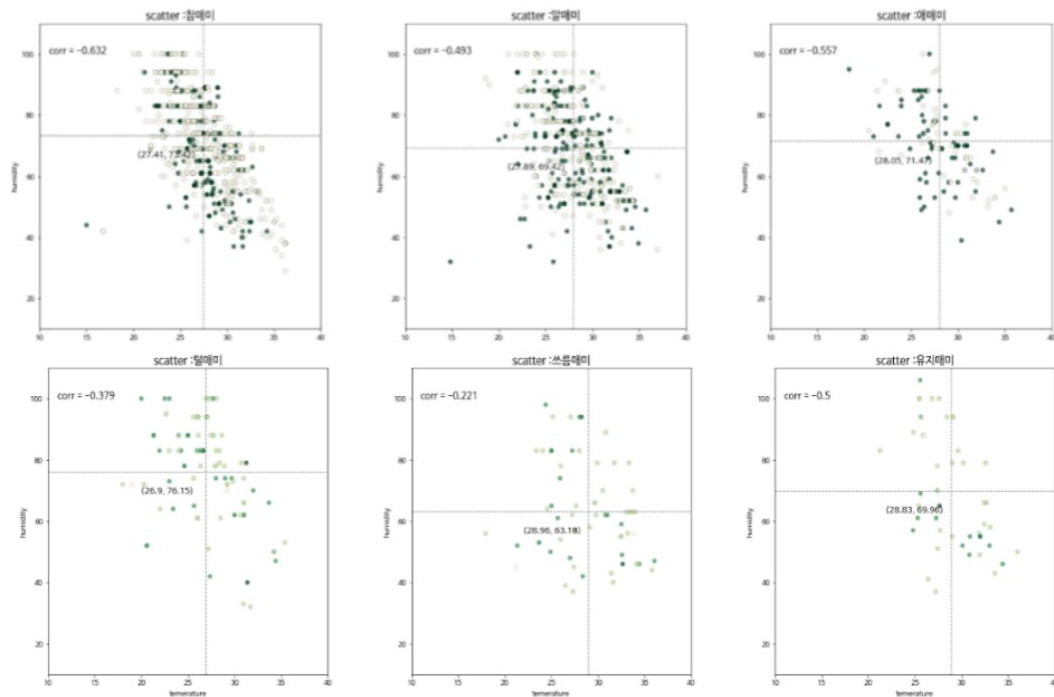
평균 기온

: 쓰름매미 = 참매미 = 말매미 = 유지매미 > 애매미 = 털매미

평균 습도

: 애매미 > 털매미 > 참매미 = 유지매미 > 말매미 = 쓰름매미

4. 시각화 기온, 습도에 따른 매미 종별 분포



평균 기온

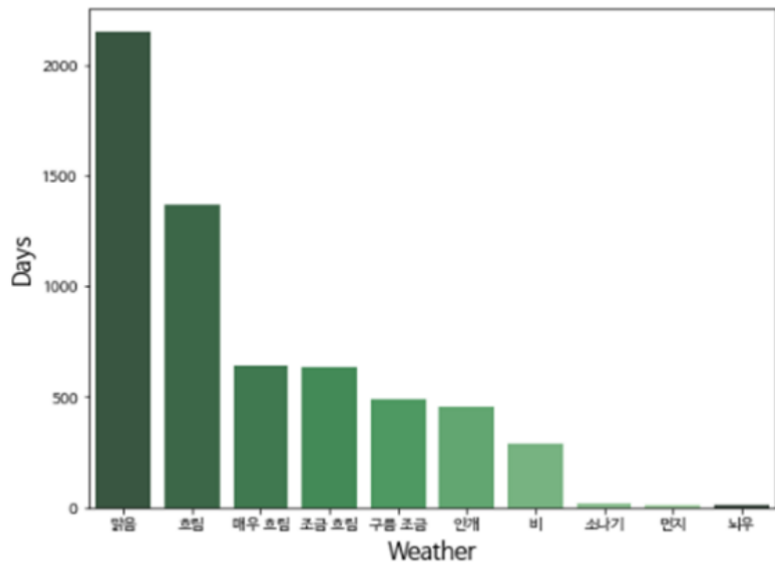
: 쓰름매미 = 유지매미 > 애매미 = 말매미 = 참매미 > 털매미

평균 습도

: 털매미 > 참매미 > 애매미 > 유지매미 = 말매미 > 쓰름매미

→ 서식지별 기온, 습도 차이는 추가적인 분석이 필요

4. 시각화 날씨에 따른 매미 종별 분포



일사량: 맑음>흐림>조금 흐림>매우 흐림

강수량: 비=소나기>맑음

으로 가정할 때

일사량은 매미 개체 수와 양의 상관관계

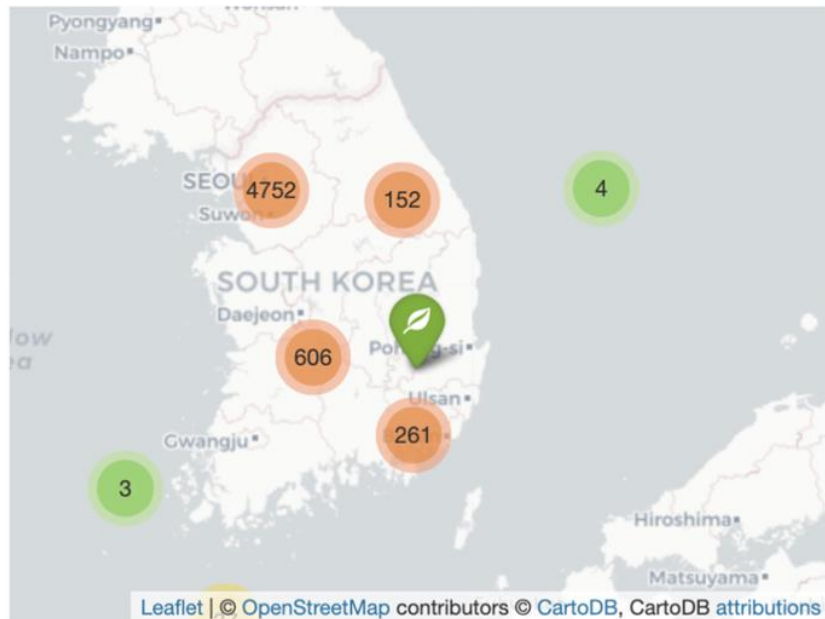
강수량은 매미 개체 수와 음의 상관관계

기상자료개방포털 데이터를 활용한다면 더 정확한 상관관계 확인 가능
각 종별로 일사량, 강수량과의 관계를 따로 비교해볼 필요가 있음

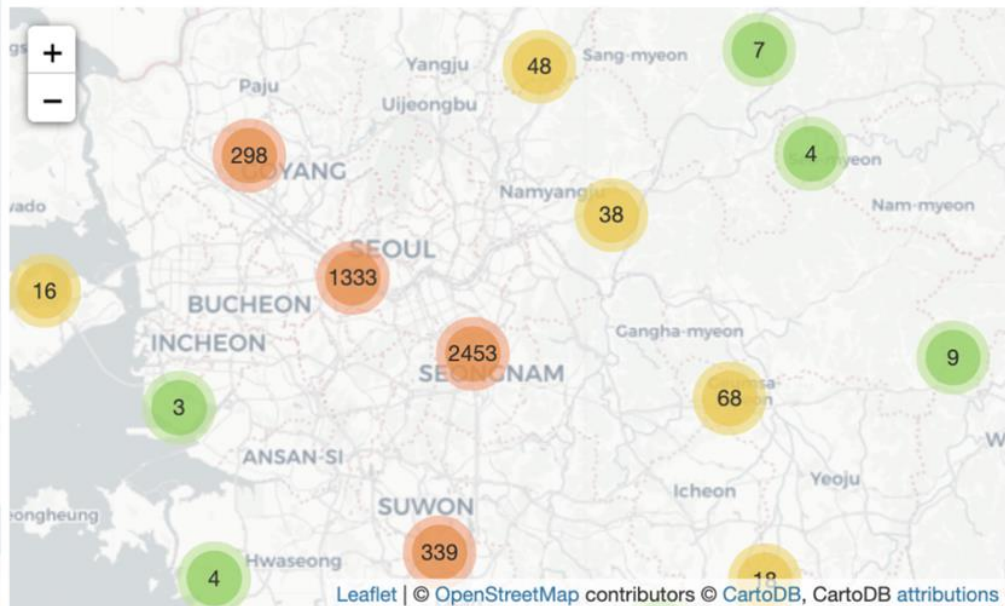
4. 시각화 매미 지도

- 지도 시각화를 위해 Python의 **folium** 라이브러리 사용.
- MarkerCluster Map을 활용해 전체 매미 지도 시각화
- 사용자가 계속해서 확대하는 경우, 숫자가 쪼개지는 형태로 각 매미의 구체적인 종과 서식지를 확인 할 수 있음.
- popup으로 매미종과 관찰 장소를 표기

4. 시각화 매미 지도

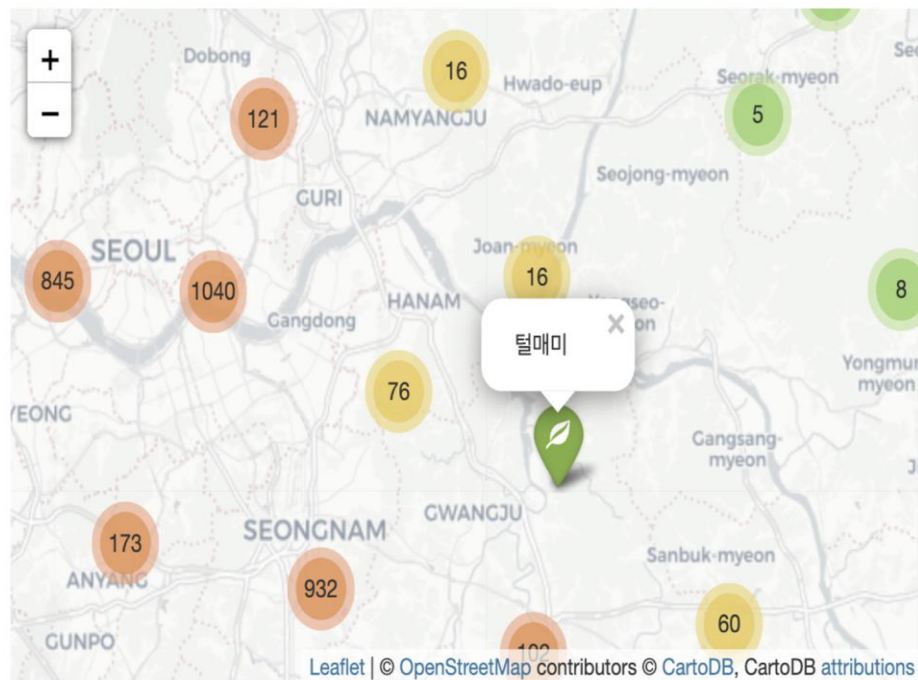


<전체적인 지도>



<확대 했을 때 숫자가 쪼개지는 모습>

4. 시각화 매미 지도



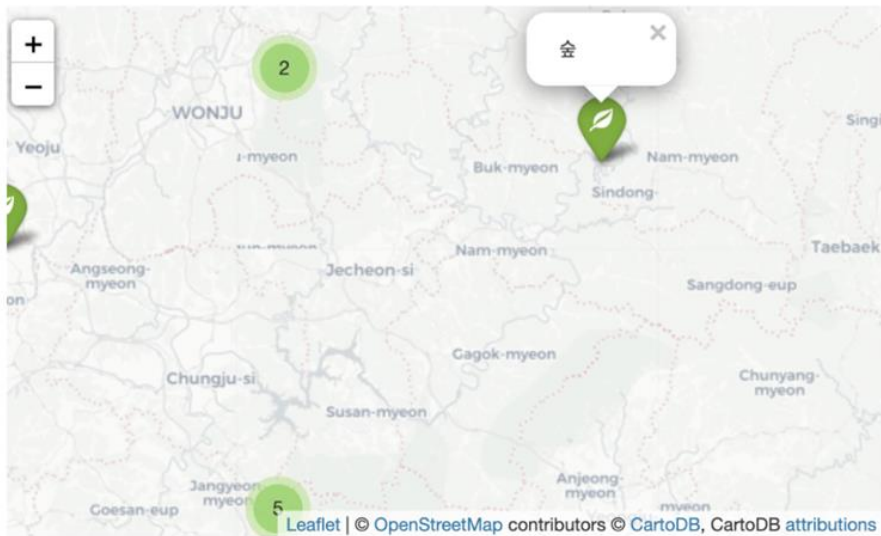
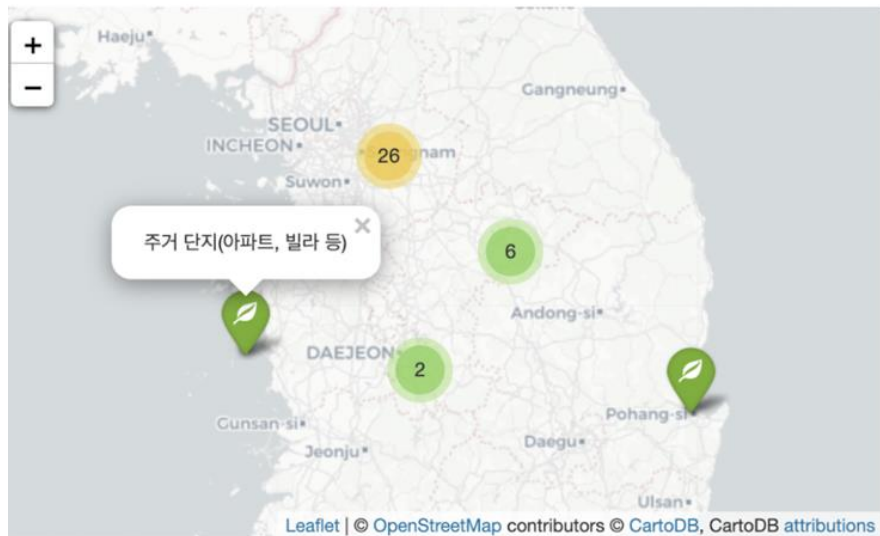
<popup시에 매미 종>

<말매미 지도>



4. 시각화 매미 지도

<소요산매미 지도>



5. 결론

- 매미는 기후에 따라 개체 수, 번식 율음 여부, 활동 기간 등에 영향을 받으며 종별로 영향을 받는 기상요인이 다를 수 있음을 선행연구와 본 분석을 통해 확인하였다.

따라서 각 매미 종의 관측 자료를 통해 해당 관측지역/전 지역의 기후 변화를 확인할 수 있다.

- 종 다양성 지수가 도시화 정도와 반비례하는 경향을 보인다는 점에서, 기존 인구 밀도를 이용한 도시화 지표와는 차별화된 도시화 지표로도 활용 가능성이 있다고 판단된다.

+ 향후 연구 방향

- 매미 생태에 영향을 미치는 요인을 바탕으로 지역 생태계를 진단하고 이를 검정하는 과정이 필요하다.

5. Lessons Learned



기지원

- 이번 대회를 통해서 단순히 분석을 하는 것보다는, “왜” 이 분석을 진행했는지, 분석 결과를 통해서 “우리는 무엇을 할 수 있을지?” 등이 더 중요하는 것을 배웠습니다.
- 심사위원분들의 평가를 통해 부족한 점을 깨닫고 한층 성장할 수 있는 좋은 기회였습니다.

- 생물다양성이라는 주제가 광범위한 만큼 접근하기가 어려웠는데, 이번 공모전을 통해서 매미라는 개체를 통해 주제에 대한 접근성을 높인 것 같습니다.
- 생물 중에 대한 관심을 넓히는 계기가 되었습니다.



박지원

5. Lessons Learned



- 초점을 두었던 매미 종에 대해 데이터 분석적 사고를 가지고 상관관계를 파악하는 것은 매우 흥미로웠습니다.
- 시민과학자의 데이터에서 이제까지 몰랐던 다양한 생물종들까지 접할 수 있어 생물다양성을 필히 보존해야 한다는 생각을 되새길 수 있었습니다.

유영서

- Raw data로 주제 선정부터 데이터 가공, 분석까지 진행하는 것은 처음이었는데 생각보다 좋은 정보들이 도출되어 신기했고 다양한 분석을 적용해볼 수 있어 좋았습니다.
- 구체적인 피드백을 통해 보완점을 생각해볼 수 있는 기회가 되었습니다.



김윤진

Q & A