



DATA ENGINEERING PROJECT

우 리 아 이 놀 이 시 설

전국 어린이 놀이시설 분석 및 사고 예측을  
위한 파이프라인 구축

SUDO 5점조  
()

# Contents

01

INTRO



02

데이터 전처리 과정



03

데이터 정제



04

데이터 대시보드



05

Q&A



# 조 소개 및 주제 선정

- \* 데이터 수집
- \* SPARK로 데이터 정제
- \* DB 설계 및 구축



허정윤

- \* 데이터 수집
- \* AIRFLOW 파이프라인 구축
- \* AWS 관리
- \* DB 설계 및 구축



강중모

FRONT-END



김연우

BACK-END



김화신

## keywords



전국 어린이 놀이시설 현황



놀이 시설에서 발생하는 안전사고



어린이 놀이시설 안전관리 개선

분석에서 시작하는 우리 아이 안전

## 주제 선정 이유

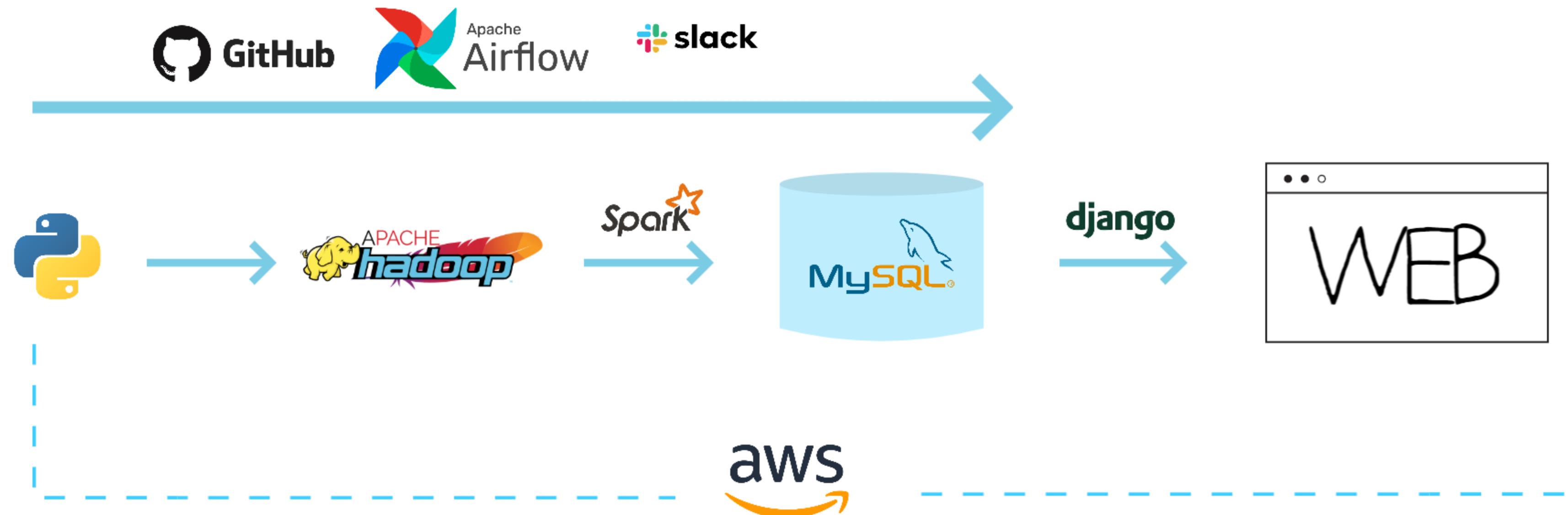
성장기 아이들, 뛰어놀고 싶고 장난도 많이 치는데요. 특히 놀이 시설에서는 더 활동적이라 사고도 많곤 하죠. 그래서 저희 조에서는 어린이 시설에서 발생하는 안전사고와 그에 따른 분석을 위해 데이터 파이프라인을 구축하였습니다.



intro

## 프로젝트 타임라인(WBS)

# 프로젝트 분석방법



# 데이터 전처리 과정



SUDO 5점조

# 데이터 정의서

이름	출처	행	열	담당자
전국어린이놀이시설 안전사고정보서비스.xml	공공데이터포털 행정안전부	약 3000개	7개	허정윤
전국어린이놀이시설 안전검사정보서비스.xml		약 45만개	3개	김연우
전국어린이놀이시설 관리자 정보서비스.xml		약 14만개	3개	허정윤
전국어린이놀이시설 정보서비스.xml		약 8만개	8개	강중모
전국어린이놀이시설기구 정보서비스.xml		약 42만개	7개	김화신
전국 어린이놀이시설 위도, 경도 크롤링	카카오지도, 구글지도	약 8만개	2개	강중모

# 데이터 정제



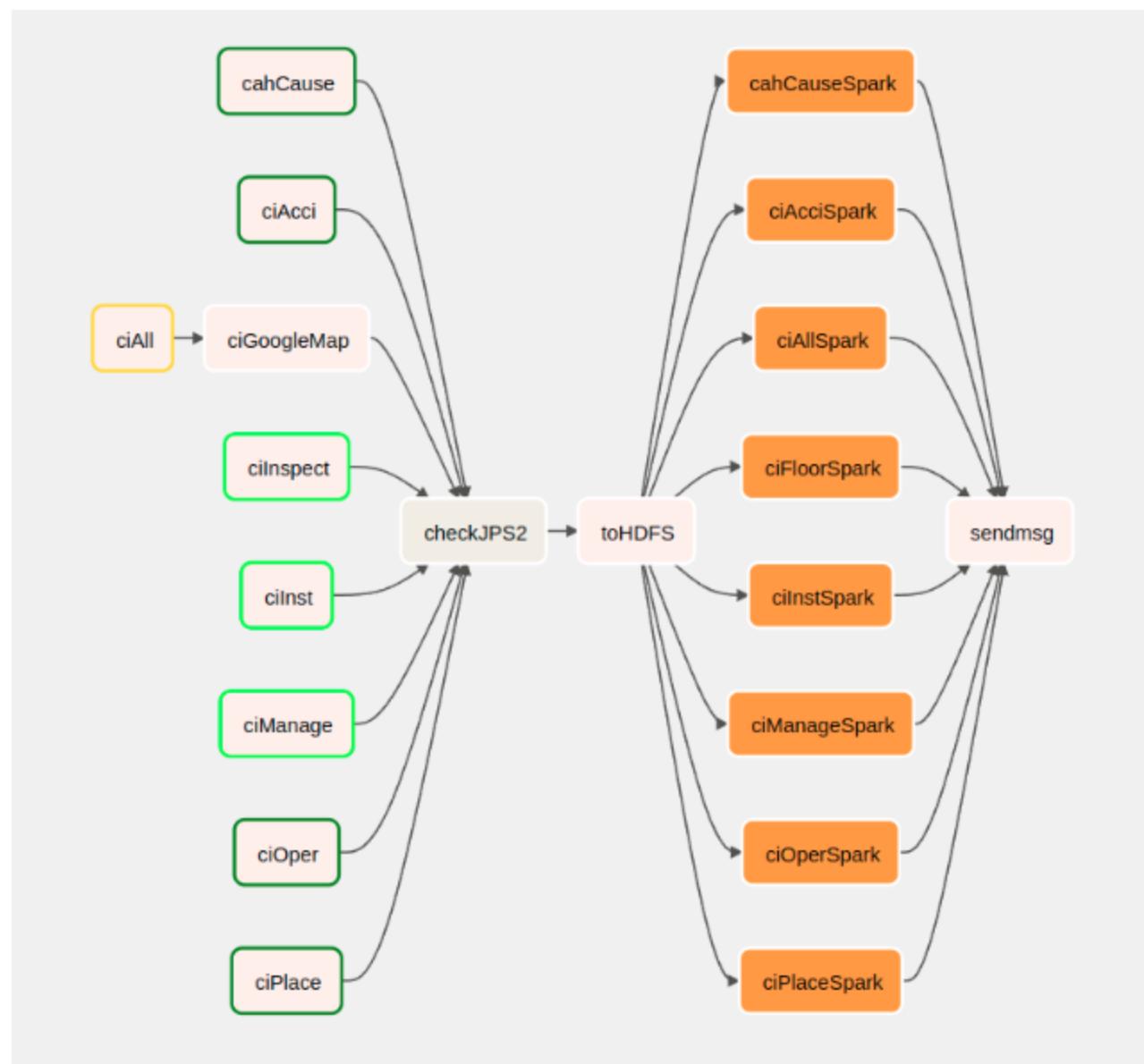
SUDO 5점조

# 스파크



Tables_in_children
ci_accident
ci_all
ci_cah_cause
ci_floor_info
ci_inst
ci_manager
ci_place_info

# Airflow&slack



**child-chatbot** 웹 오후 3:30  
 [03/31 15:30:02] ci\_placeCode DATA LOCAL STORAGE SUCCESS  
 [03/31 15:30:02] ci\_operCode DATA LOCAL STORAGE SUCCESS  
 [03/31 15:30:02] cah\_cause DATA LOCAL STORAGE SUCCESS  
 [03/31 15:30:14] ci\_accident DATA LOCAL STORAGE SUCCESS

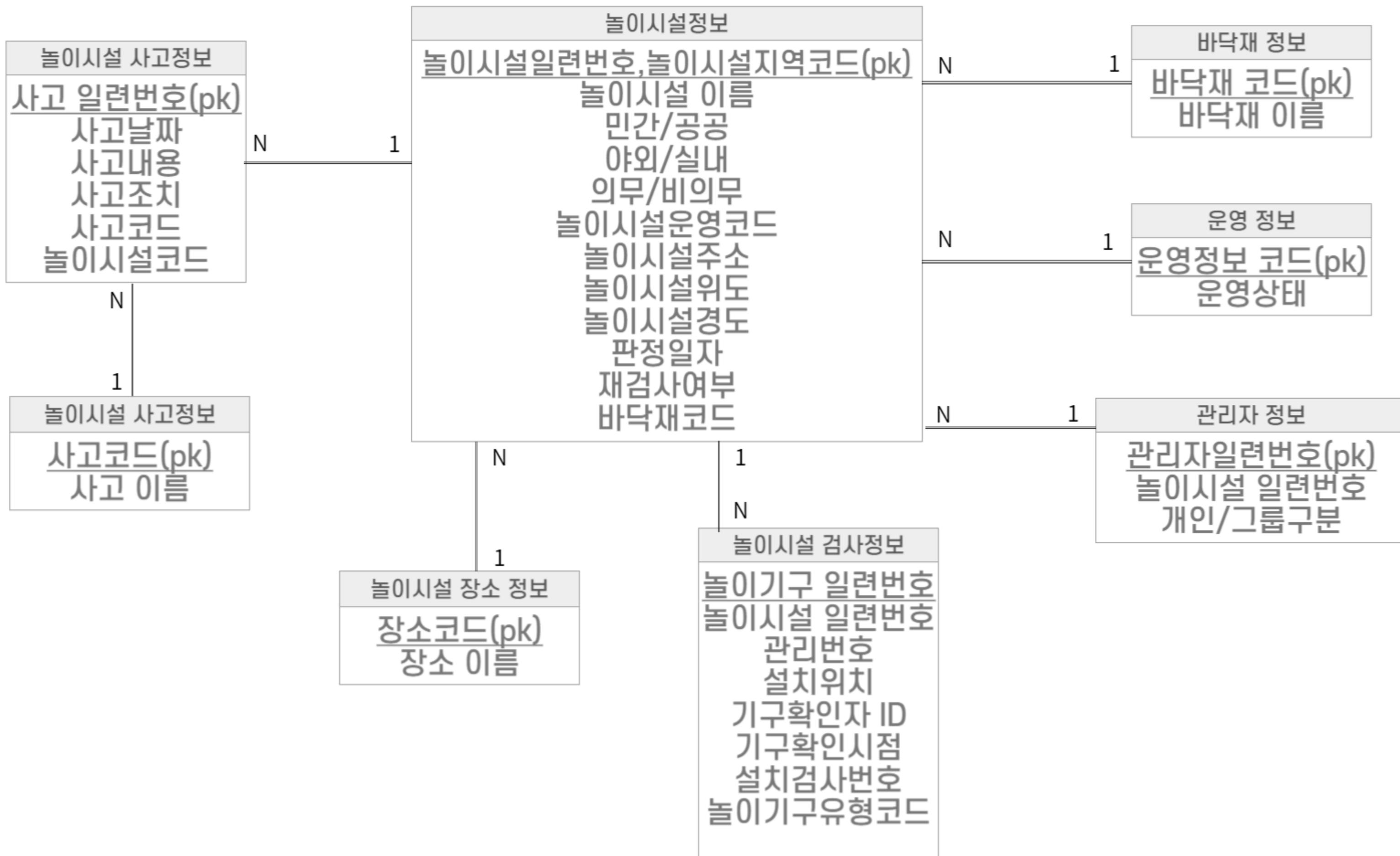
**child-chatbot** 웹 오후 3:35  
 [03/31 15:35:49] ci\_all DATA LOCAL STORAGE SUCCESS  
 [03/31 15:38:26] ci\_manager DATA LOCAL STORAGE SUCCESS

**child-chatbot** 웹 오후 3:53  
 [03/31 15:53:37] ci\_inst DATA LOCAL STORAGE SUCCESS

**child-chatbot** 웹 오후 4:05  
 [03/31 16:05:00] ci\_inspection DATA LOCAL STORAGE SUCCESS  
 [03/31 16:05:06] ALL FILES HDFS STORAGE DONE  
 ✓ 1 😊

[03/31 16:06:42] ci\_manager SPARK-SUBMIT SUCCESS  
 [03/31 16:07:26] ci\_accident SPARK-SUBMIT SUCCESS  
 [03/31 16:09:23] ci\_all SPARK-SUBMIT SUCCESS  
 [03/31 16:09:46] ci\_place SPARK-SUBMIT SUCCESS  
 [03/31 16:10:08] cah\_cause SPARK-SUBMIT SUCCESS  
 [03/31 16:10:31] ci\_oper SPARK-SUBMIT SUCCESS  
 [03/31 16:11:46] ci\_inst SPARK-SUBMIT SUCCESS  
 [03/31 16:12:26] ci\_floor SPARK-SUBMIT SUCCESS  
 [03/31 16:12:27] ALL DONE!  
 ✓ 1 😊



**ERD**

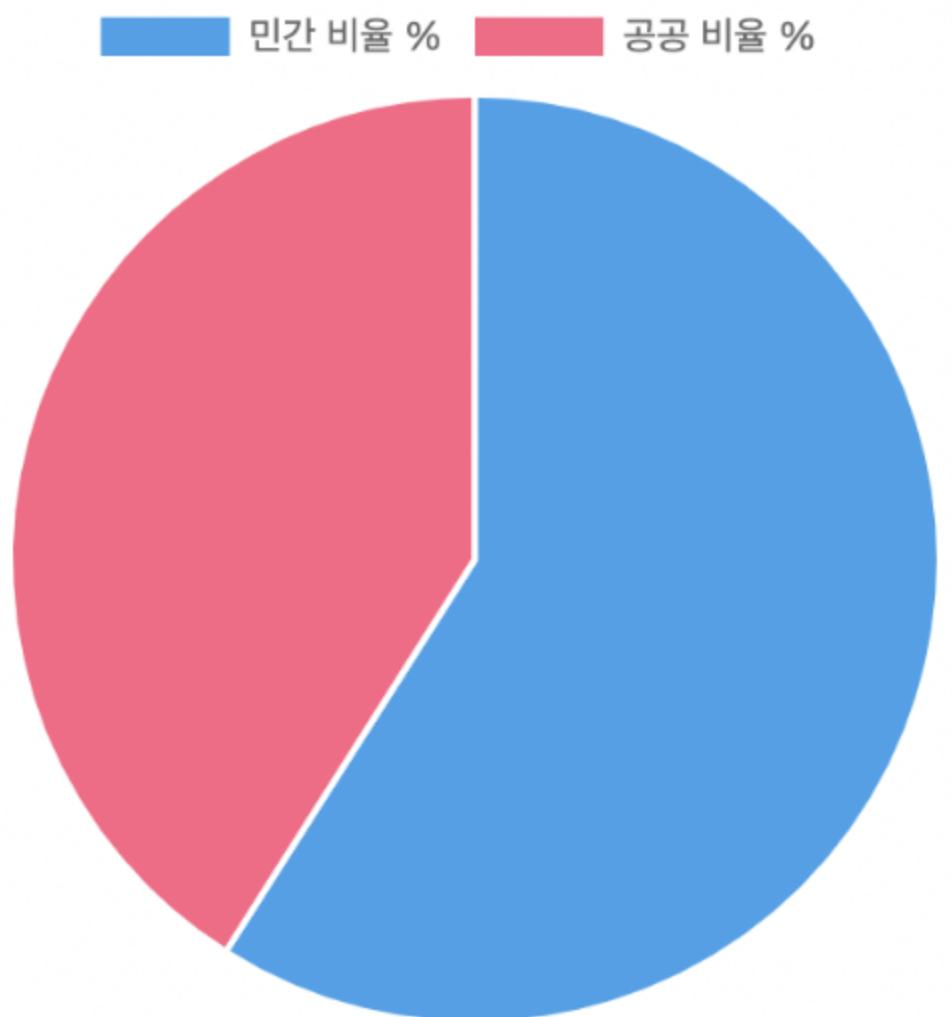
# 데이터 대시보드



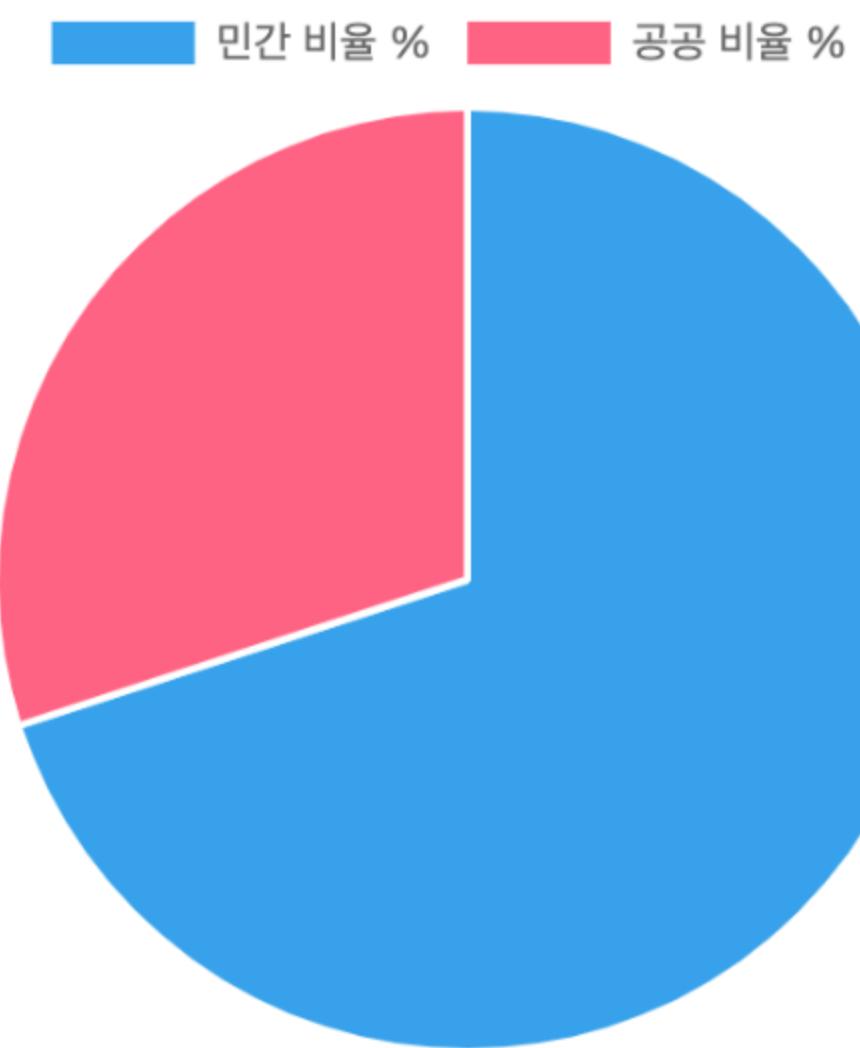
SUDO 5점조

# 민간/공공 분포

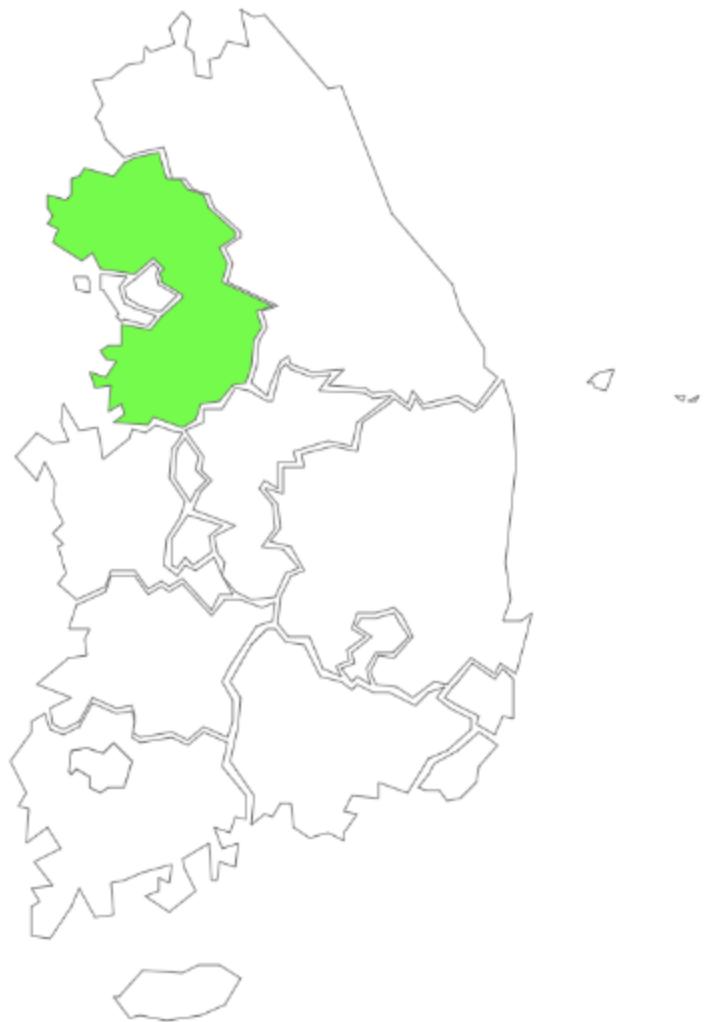
민간, 공공시설분포



민간,공공시설분포



# 지역별 놀이시설 개수

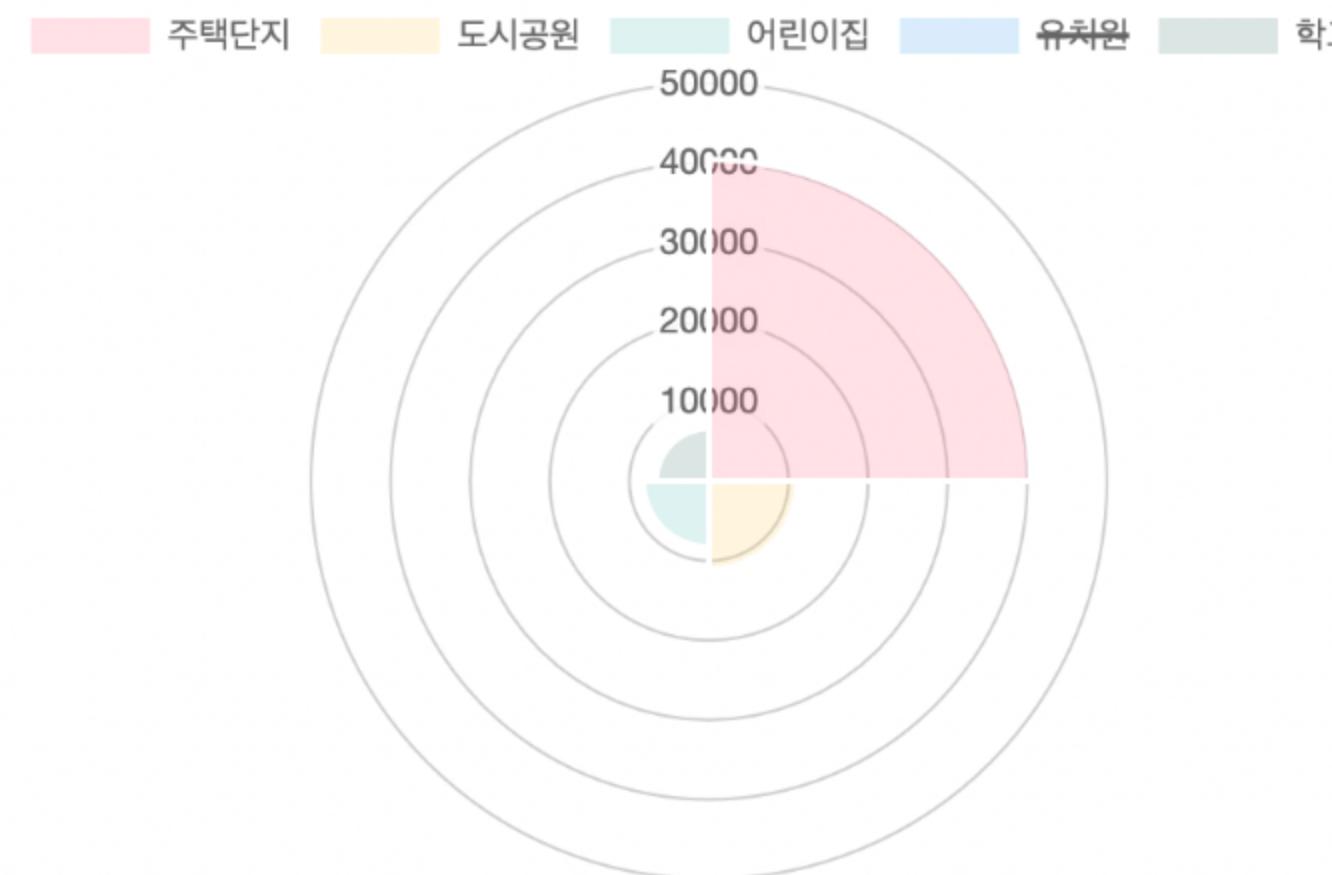


경기도

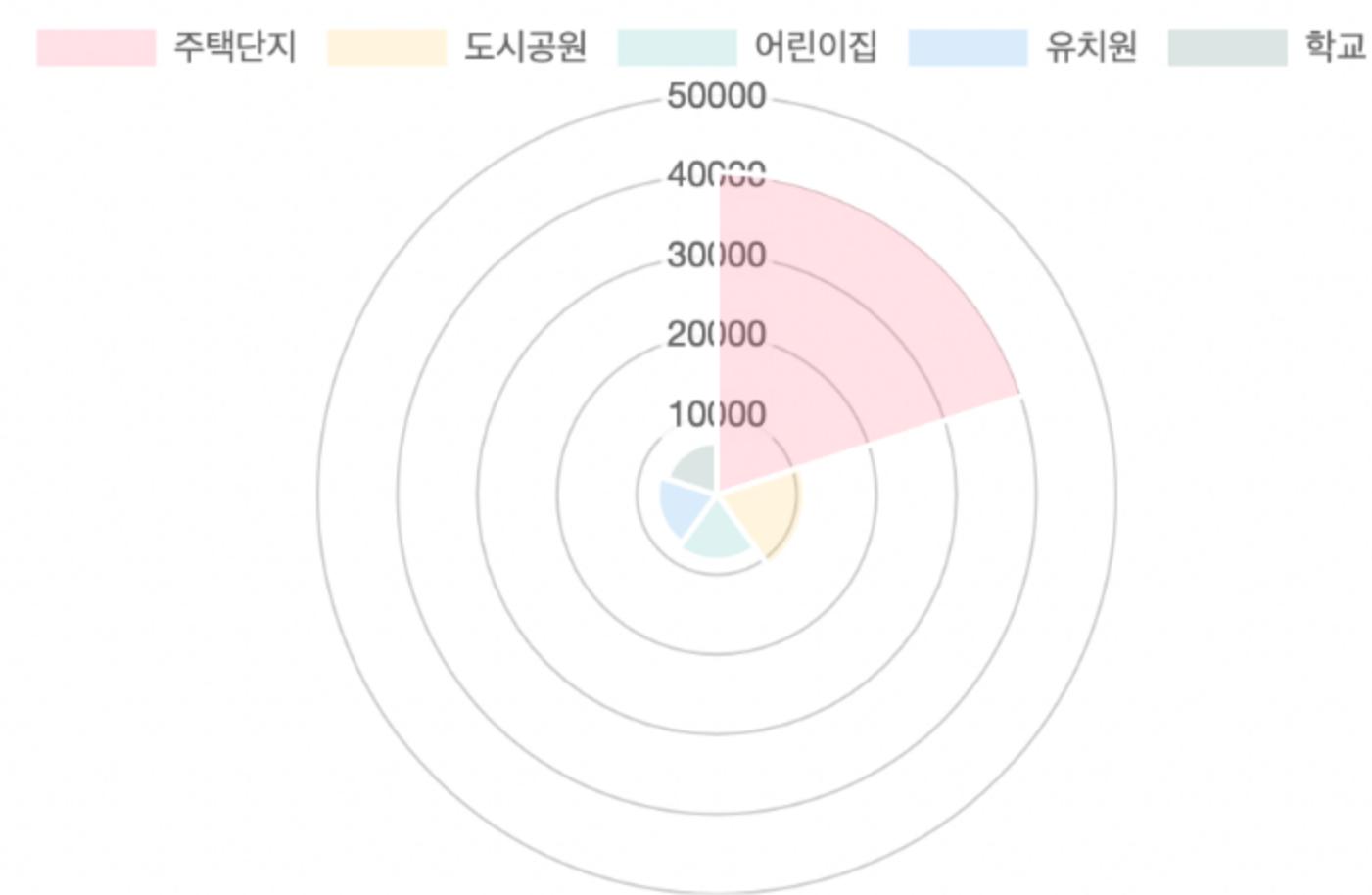
놀이시설 갯수 : 21572

# 놀이시설 설치위치 top5

놀이시설 Top5

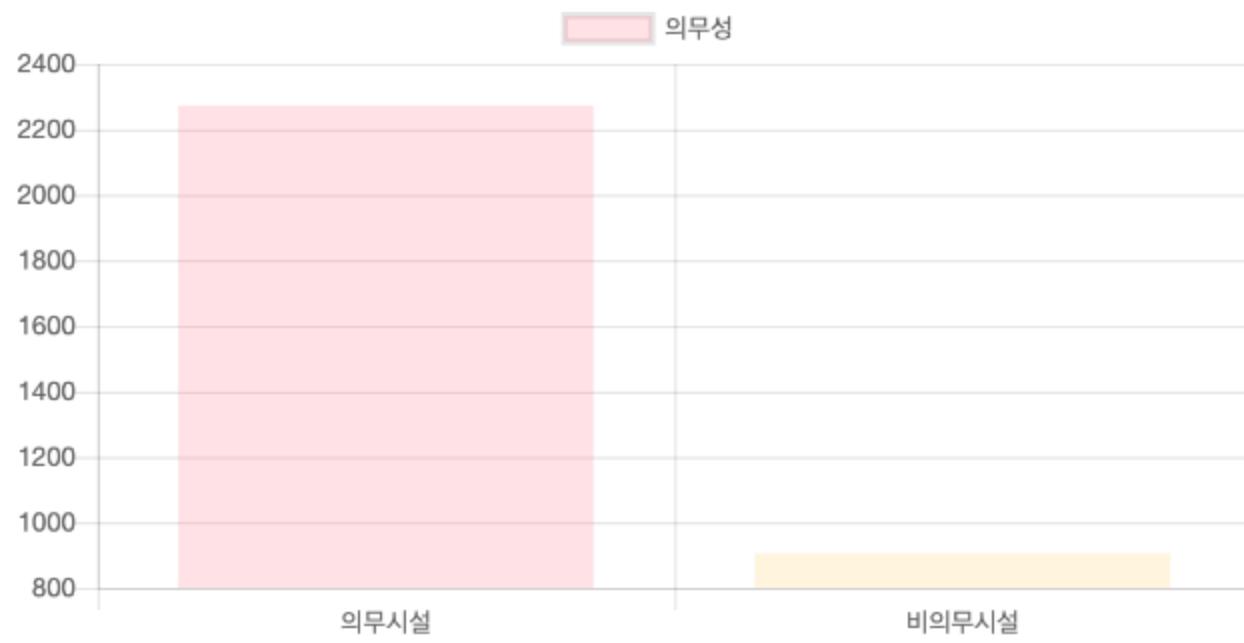


놀이시설 Top5

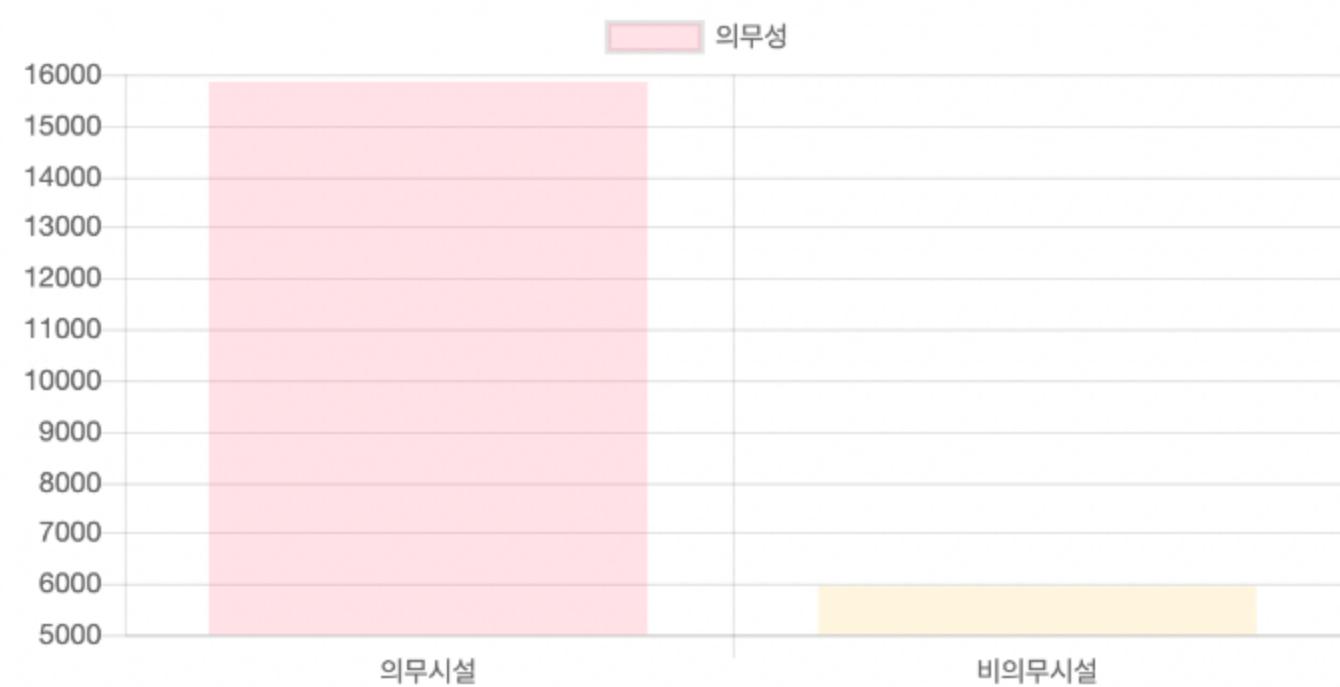


# 지역별 의무시설

의무시설 및 비의무시설 수

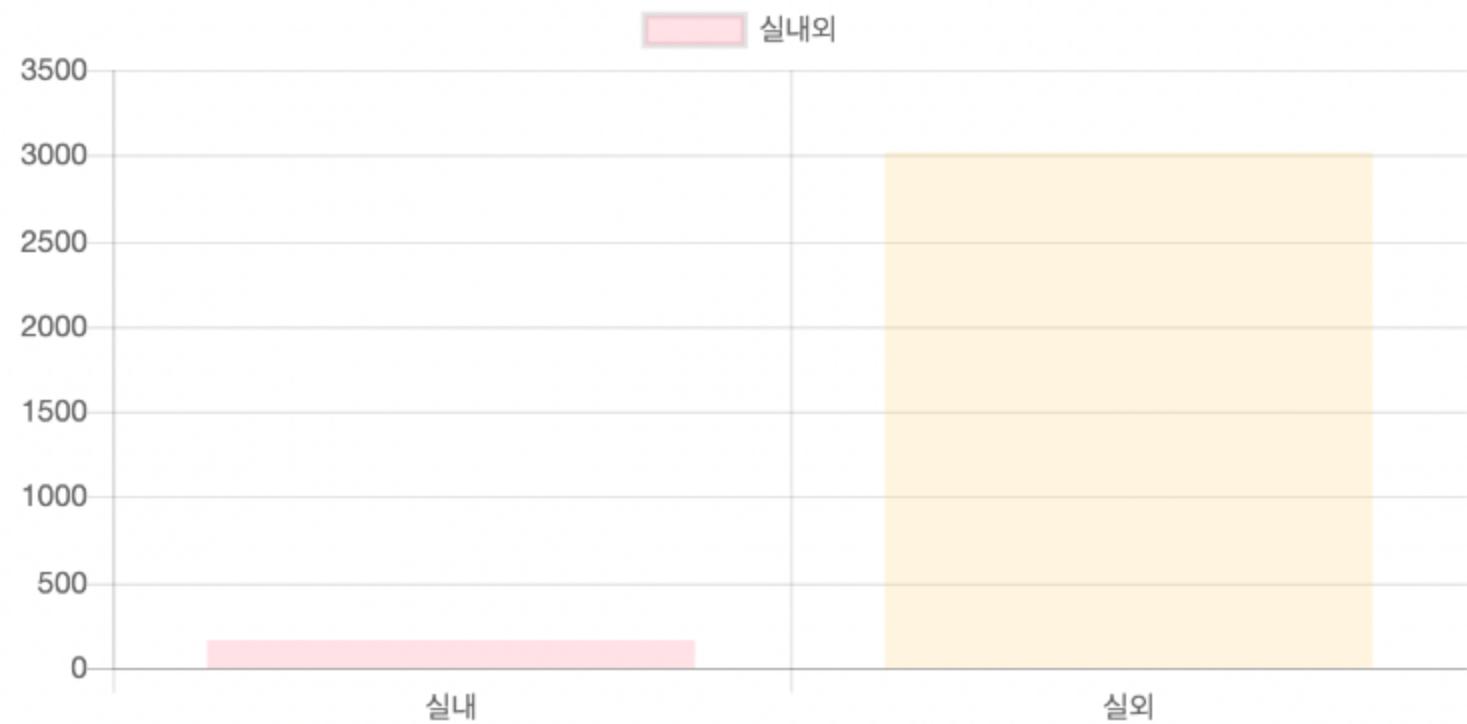


의무시설 및 비의무시설 수

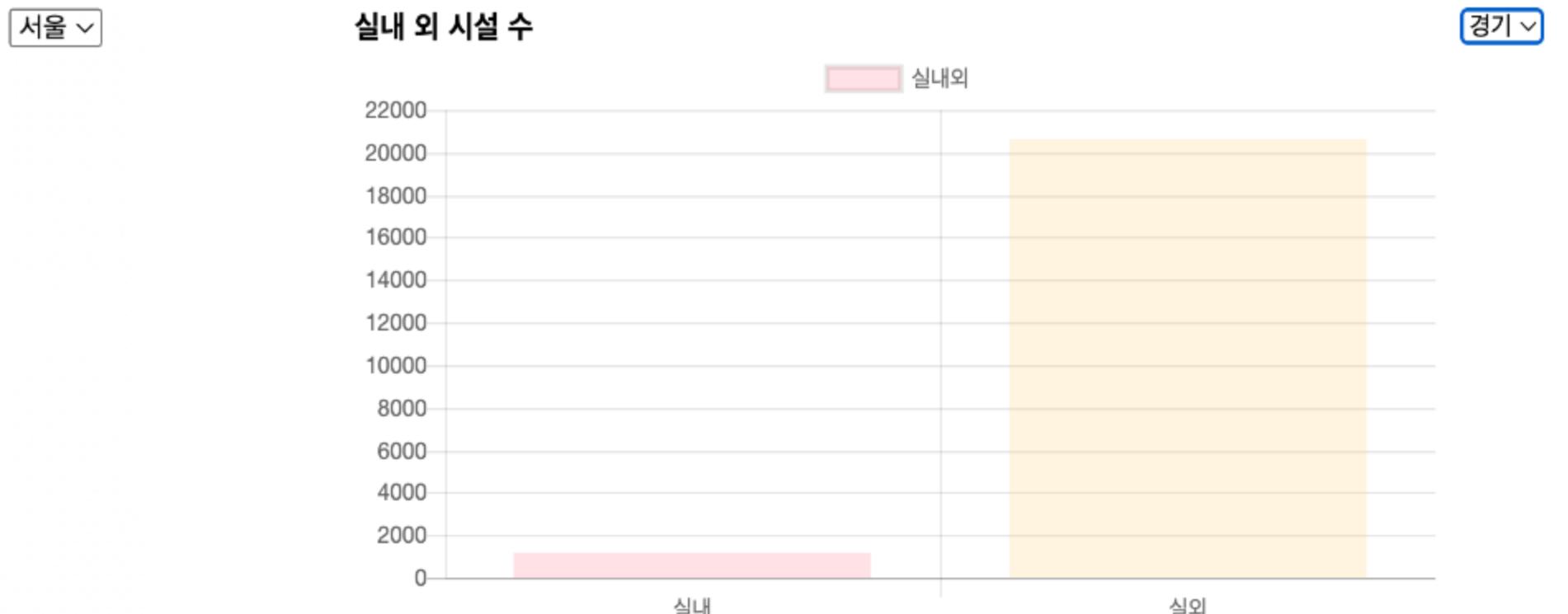


# 놀이시설 실/내외 개수

실내 외 시설 수



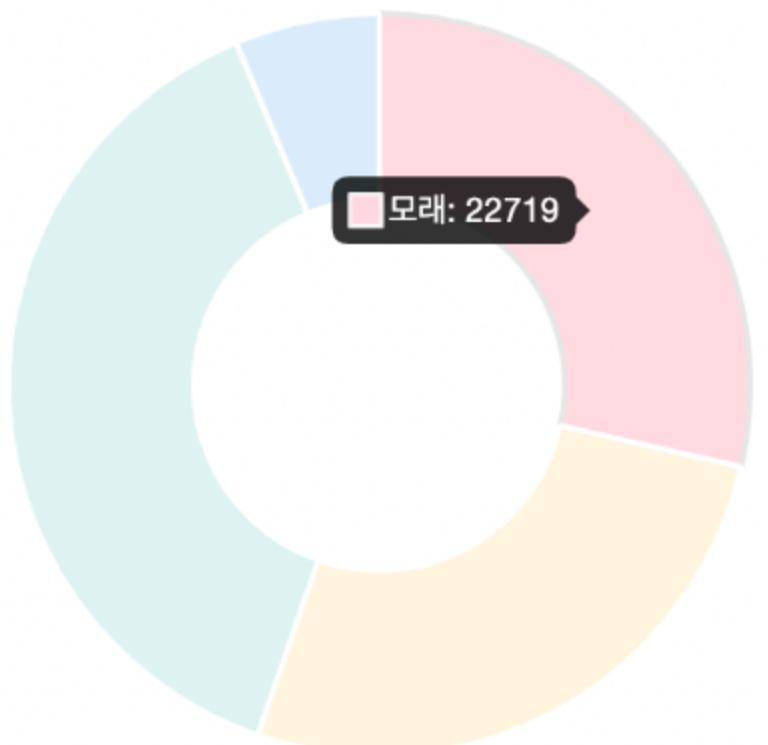
실내 외 시설 수



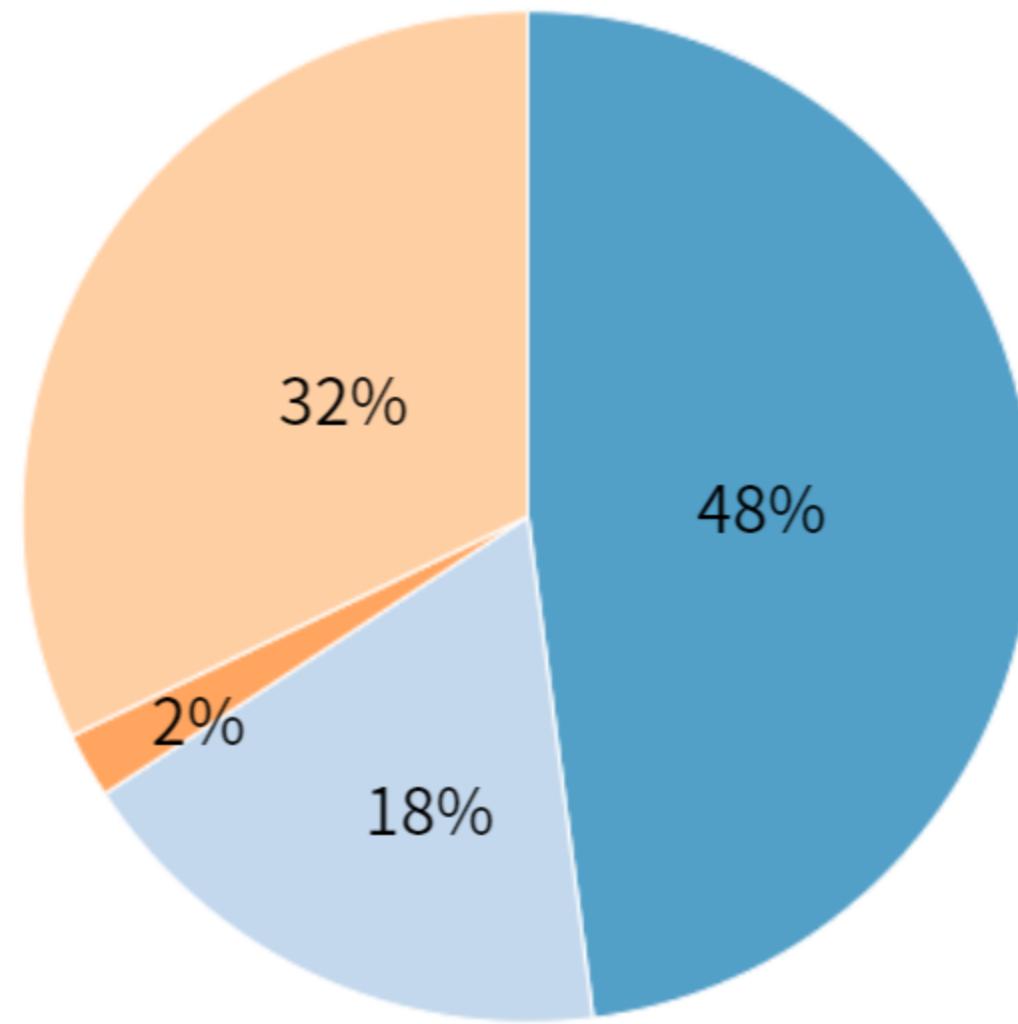
# 놀이시설별 바닥재

종류 별 바닥재 사용 수

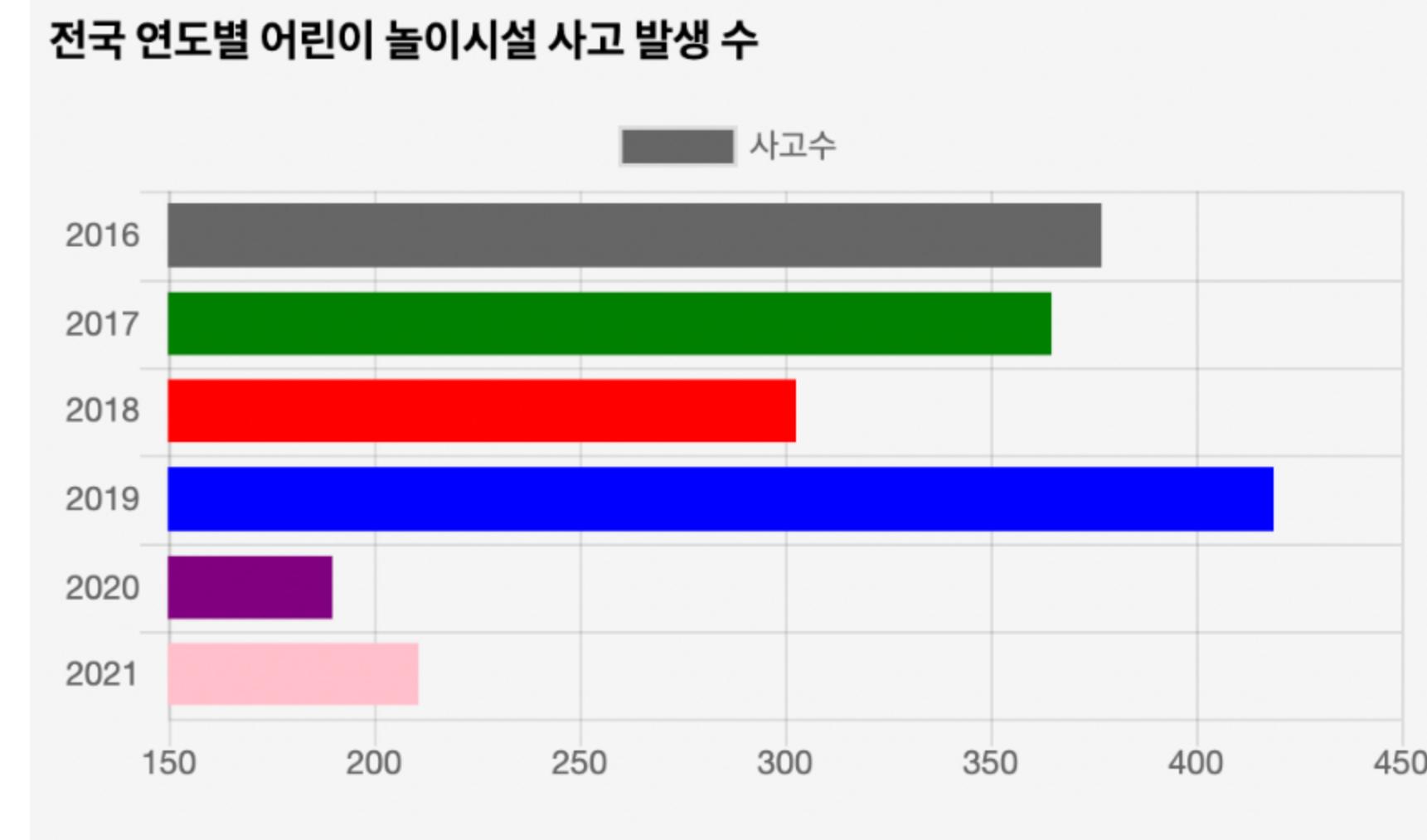
모래 고무바닥재 포설도포바닥재 기타바닥재



● 모래 ● 고무바닥재 ● 기타바닥재 ● 포설도포바닥재



# 연도별 사고 건수



# Q&A



“

끝까지 경청해주셔서 감사합니다

”

SUDO 5점 조

<https://github.com/JngMkk/PipelineProject>