

익명화 기술을 이용한 안전한 교통 빅데이터 분석

2021학년도 캡스톤디자인

상명대학교 컴퓨터과학전공 CO-LAB팀

C O N T E N T S

- 01 **배경, 주제 및 목적**
- 02 **활용방안 및 기대효과**
- 03 **프로그램 및 언어**
- 04 **빅데이터 분석 설명**
- 05 **앱 설명**
- 06 **시연 영상**

01

배경, 주제 및 목적

익명화 기술을 이용한 안전한 교통 빅데이터 분석



배경

개인의 이동경로 유출
집 주소 등의 민감한 정보 유출

+



주제

LDP 익명화 기술을 이용한
안전한 교통 빅데이터 분석

+



목적

개인의 프라이버시 보호

02

활용방안 및 기대효과

- > 지하철 혼잡도 분석 및 개선
- > 버스, 택시 시간별 요일별 구간 혼잡도 분석
- > 교통체증 파악
- > 노선 계획

03

프로그램 및 언어

익명화 기술을 이용한 안전한 교통 빅데이터 분석

Jupyter
Notebook

jupyter

Python

Android
Studio



Java
Sqlite



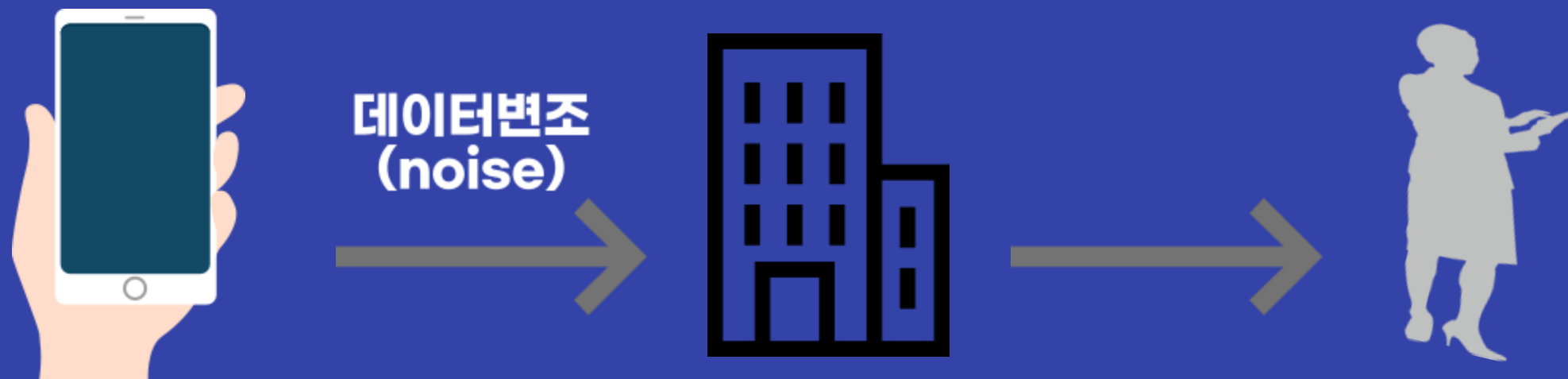
04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

LDP (Local Differential Privacy)

- > 개인 정보 비식별화 기술
- > 개인 프라이버시 보존
- > 랜덤화 알고리즘 이용



04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

$$\Pr(A(Va)=결과값1) \leq e^{\epsilon} * \Pr(A(Vb)=결과값2)$$



privacy

$\epsilon=0.1$

ϵ = privacy budget

$\epsilon=0.5$

$\epsilon=1$



accuracy

04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

출처

서울열린데이터광장
서울시 지하철 호선별
역별시간대별 승하차 인
원정보 2021년6월

출처

서울열린데이터광장
서울시 지하철호선별 역
별 승하차 인원 정보
2021년7월5일(월)
~11일(일)

필요한 데이터

승하차 호선
승차하역
총 승객수

04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

승차 \ 하차	A역	B역	C역	D역
A역	0	1	0	0
B역	0	0	0	0
C역	0	0	0	0
D역	0	0	0	0

A역 승차 B역 하차 가정

0100 0000 0000 0000

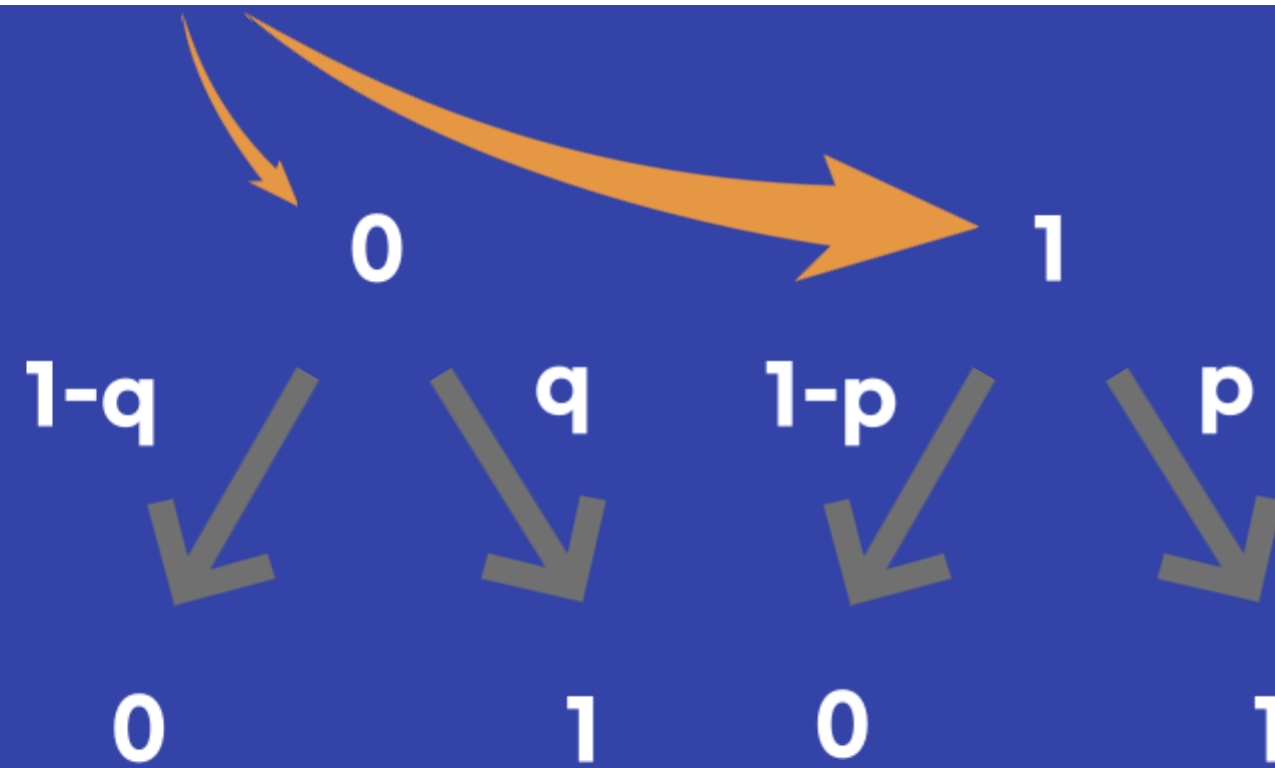
04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

N명의 원본데이터

0100 0000 0000 0000

0100 0000 \ddots 0000 0000
 $p=0.5$ $q=1/(1+e^{\epsilon})$
 ϵ =privacy budget

04

빅데이터 분석 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석

N명의 변형데이터

0000 0000 0110 0000

0000 0100 0000 0000

1의 개수= n'

원본데이터 1의 개수구하기

$$n = (n' - N * q) / (0.5 - q)$$

$$p = 0.5 \quad q = 1 / (1 + e^{\epsilon})$$

 ϵ = privacy budget

$$n' = n * p + (N - n) * q$$

$$n' - N * q = n * (p - q)$$

$$n = (n' - N * q) / (p - q)$$

$$n = (n' - N * q) / (0.5 - q)$$

(n' : 변형 후 1개의 개수,

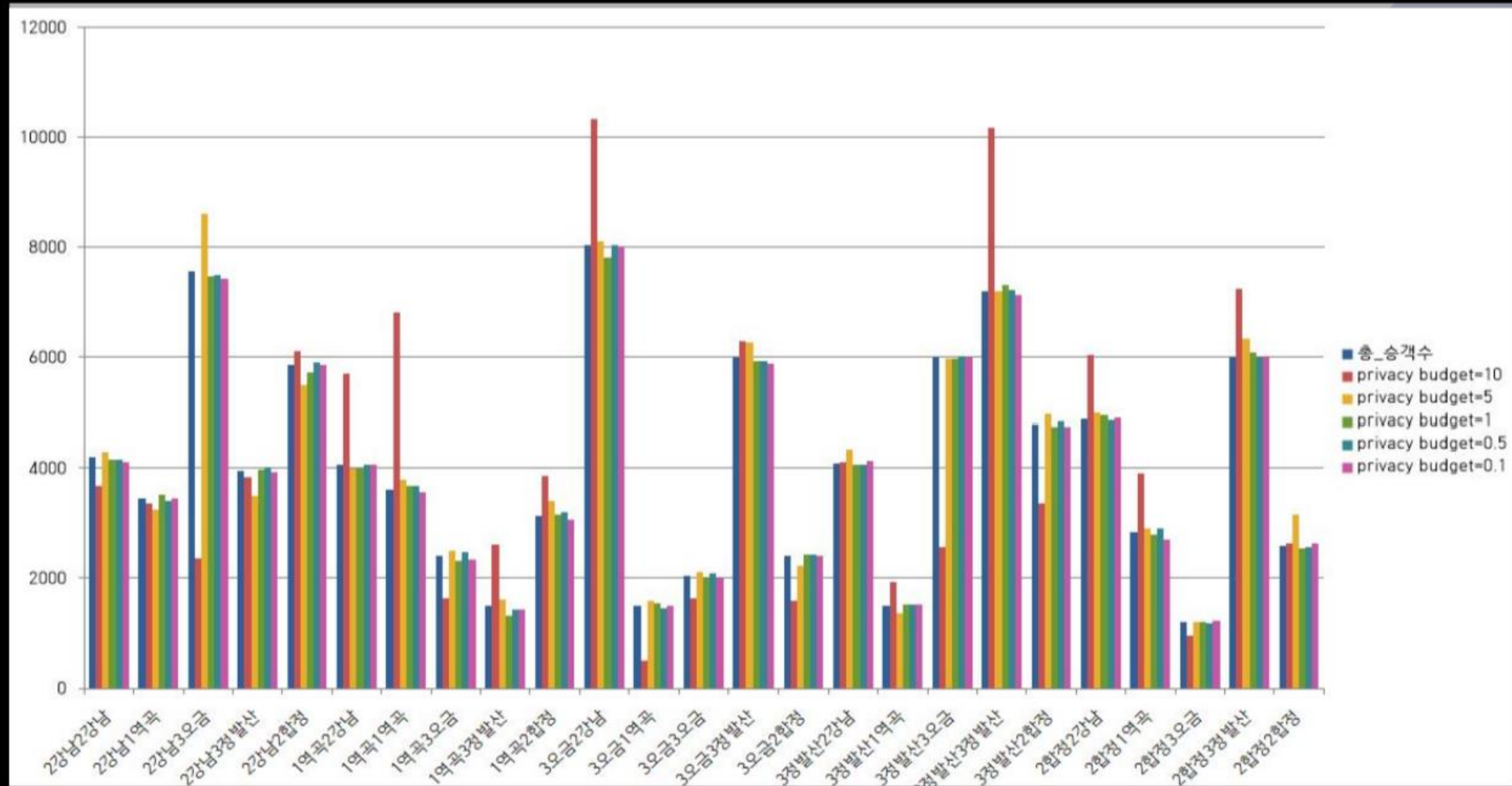
n : 실제 1개의 개수,

N : 인원수)

04

빅데이터 분석 설명

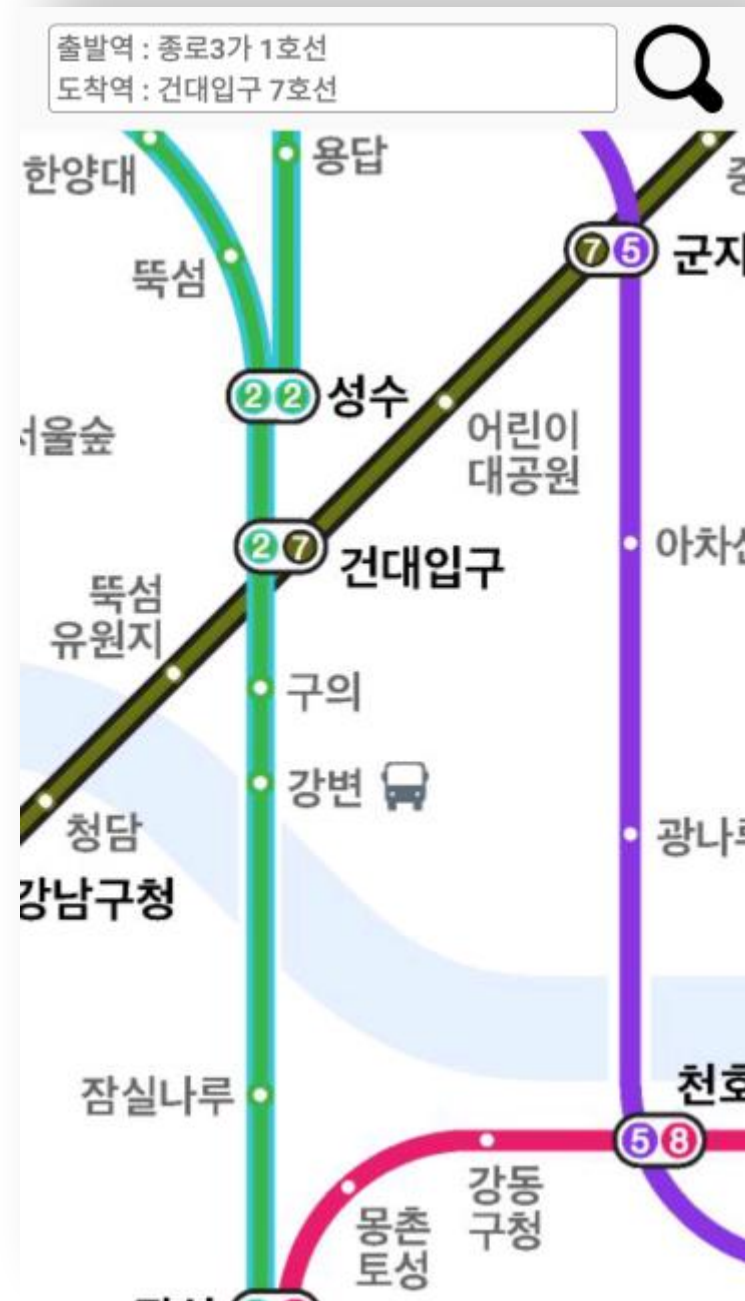
익명화 기술을 이용한 안전한 교통 빅데이터 분석



05

앱 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석



시작화면



지하철역선택화면

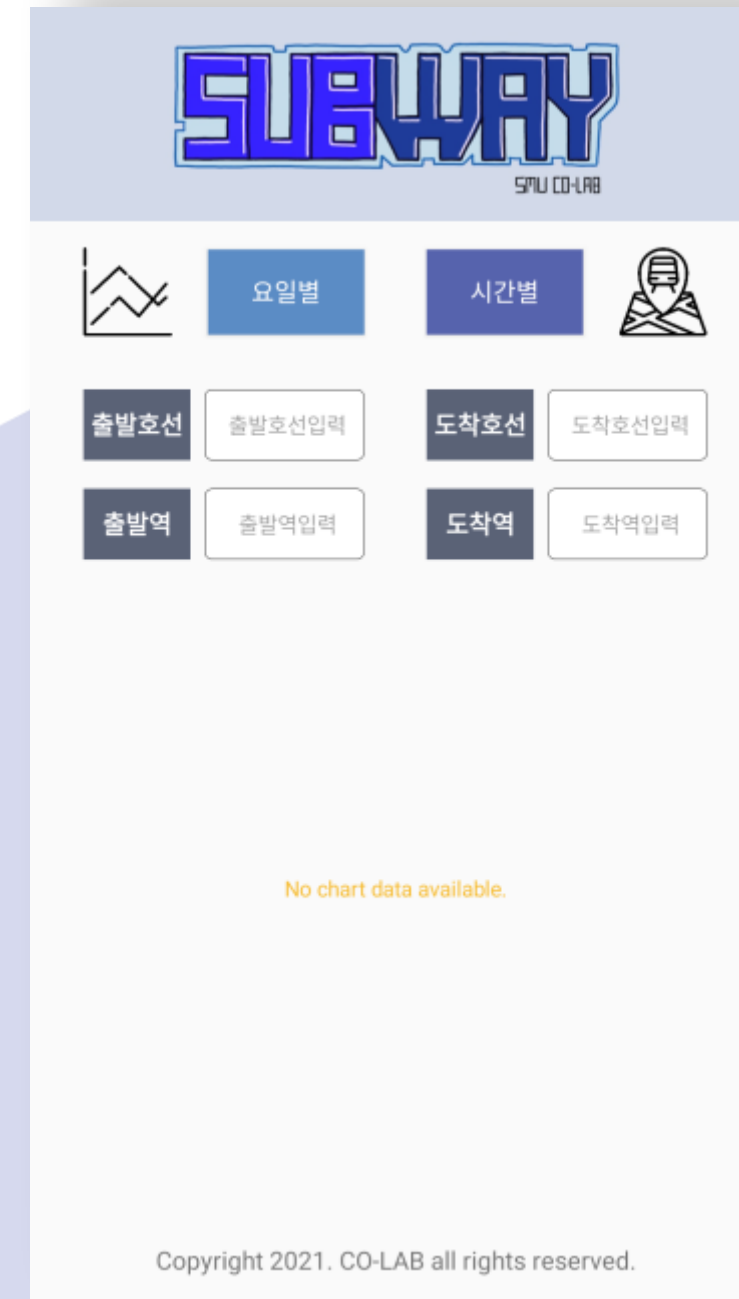
05

앱 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석



사이드메뉴화면



검색화면

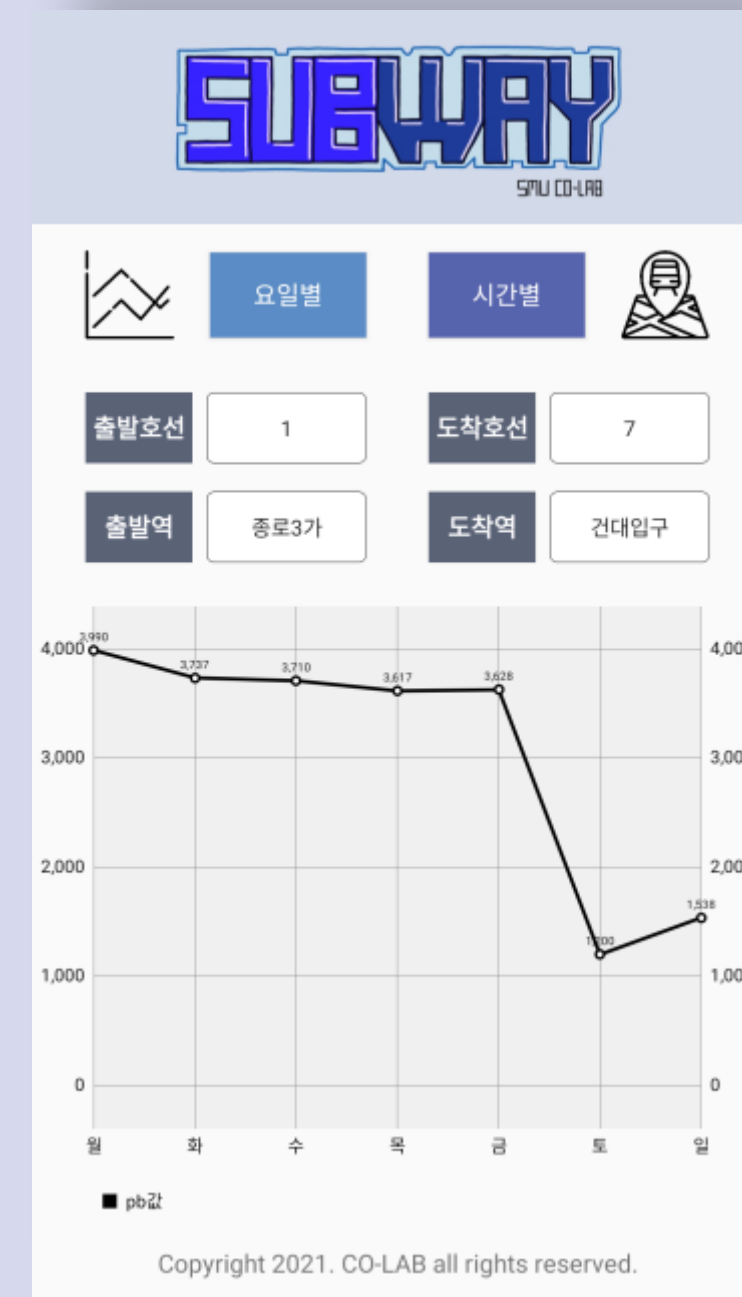
05

앱 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석



TOP3 승하차역화면

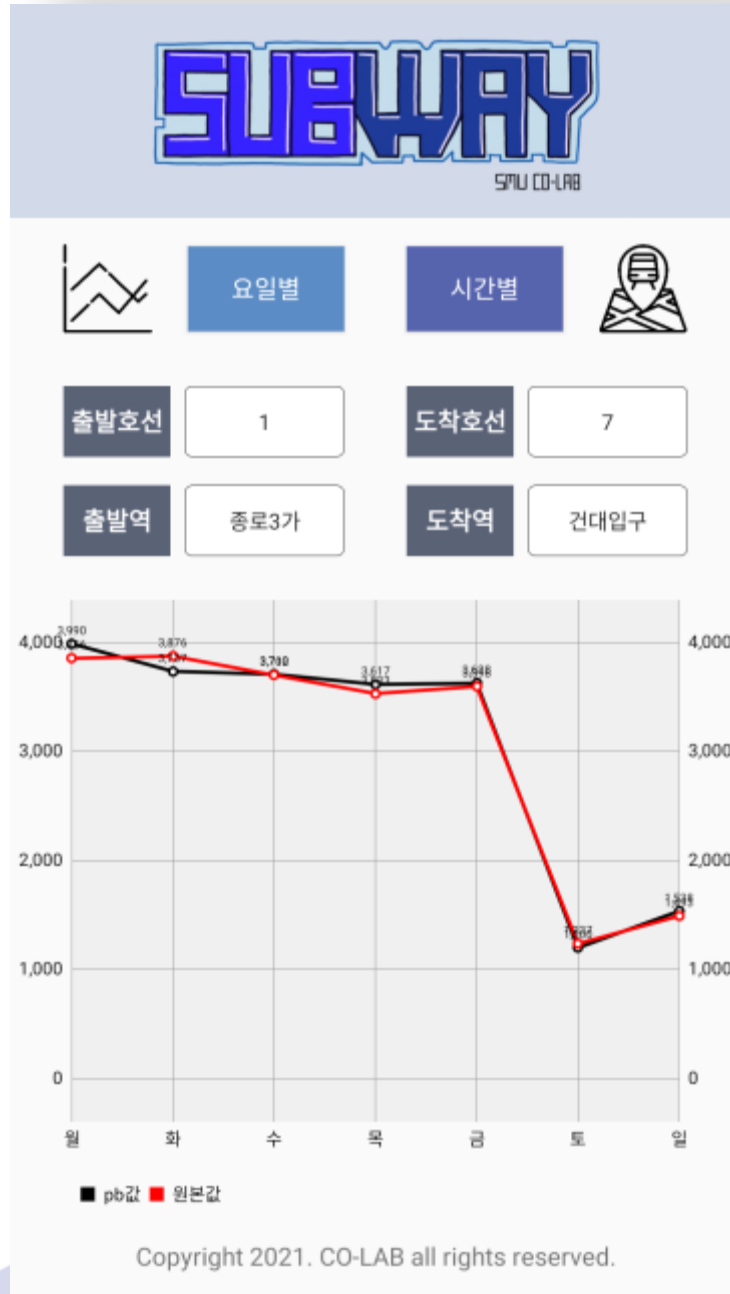


요일, 시간별 그래프화면

05

앱 설명

익명화 기술을 이용한 안전한 교통 빅데이터 분석



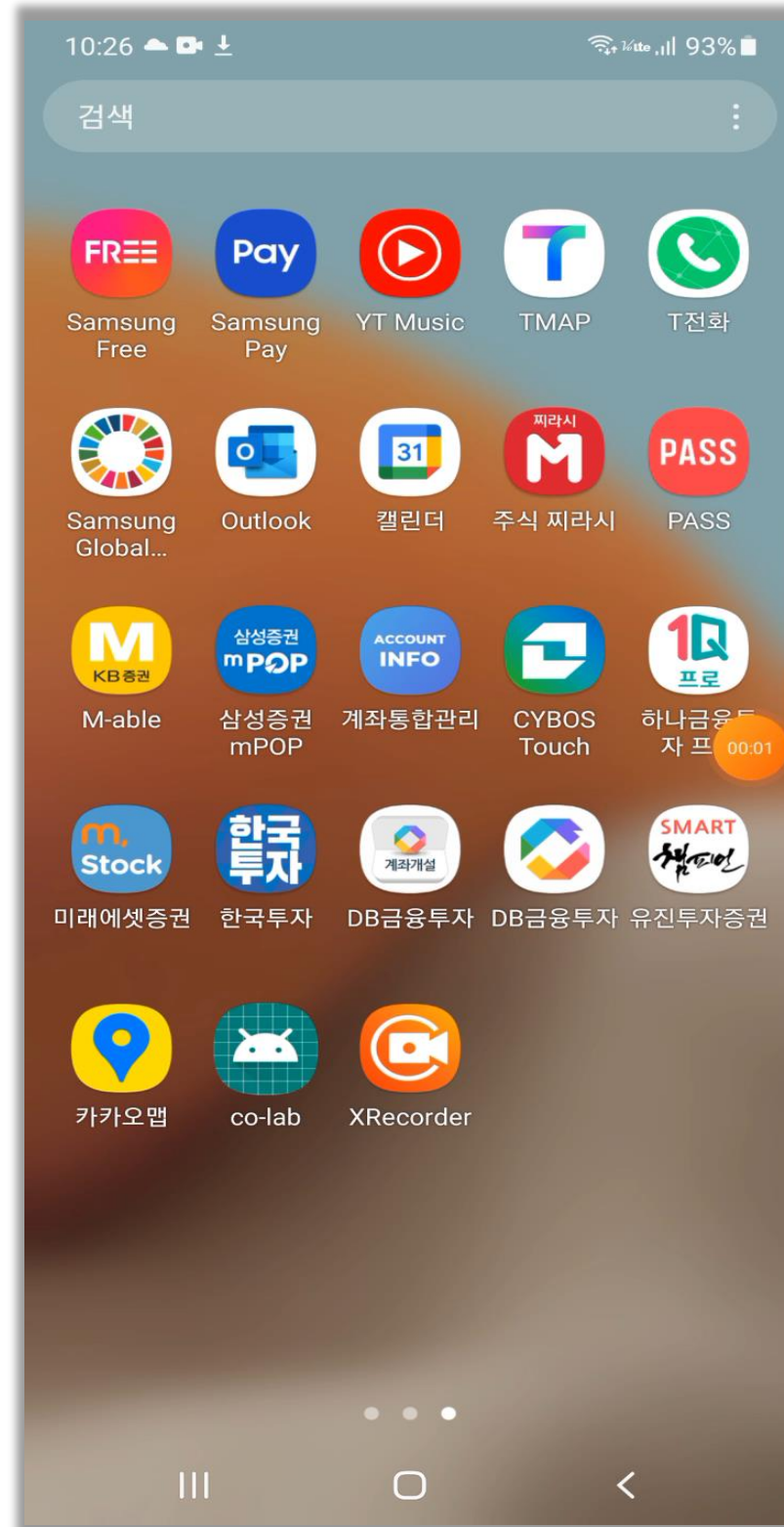
원본데이터 비교그래프화면



역별 비교그래프화면

06

시연 영상



감사합니다.

지도교수

김종욱 교수님

팀장

박예진

팀원

박연주, 백웅현, 배소은, 최수빈
