



코로나 이후 이동량, 정말 줄었을까?

패스트캠퍼스 데이터사이언스 스쿨 17기
서채빈, 이지영

목 차

01 주제 선정

- 개요_ 주제선정 동기
- 프로젝트 목적
- 진행 계획

02 데이터 수집 방법

- 대중교통 이용량 데이터
- 차량 이동량 데이터
- 코로나 확진자 현황 데이터

03 데이터 처리 및 분석

- 빈도 분석
- 상관관계 분석

04 추후 개선 및 느낀점

- 프로젝트 종료 후 소감



01

주제 선정

- 개요 : 주제선정 동기
- 프로젝트 목적
- 진행 계획

주제선정 동기

ABOUT SUBJECT

“ 코로나19로 지난해 주말
대중교통 이용 36.1% '뚝' 떨어졌다 ”

<https://news.naver.com/main/read.nhn?mode=LS2D&mid=shm&sid1=103&sid2=240&oid=421&aid=0005239199>

코로나 이후 '교통량'이 줄었다는 기사는 끊임없이 쓰아지지만, 과연?

2020년은 코로나로 인해 일상생활에 많은 변화가 생겼다. 소위 '비대면 시대'가 막을 올렸다고 말하며, 사회적 거리두기가 1년 넘게 시행되는 중이다.

하지만, 연일 대중교통과 차량 이동량이 대폭 감소했다는 기사와는 달리 집단감염은 꾸준히 발생하며, SNS에서는 '인증샷'이 끊이지 않는다.

과연 정말 사람들이 안 나갈까?

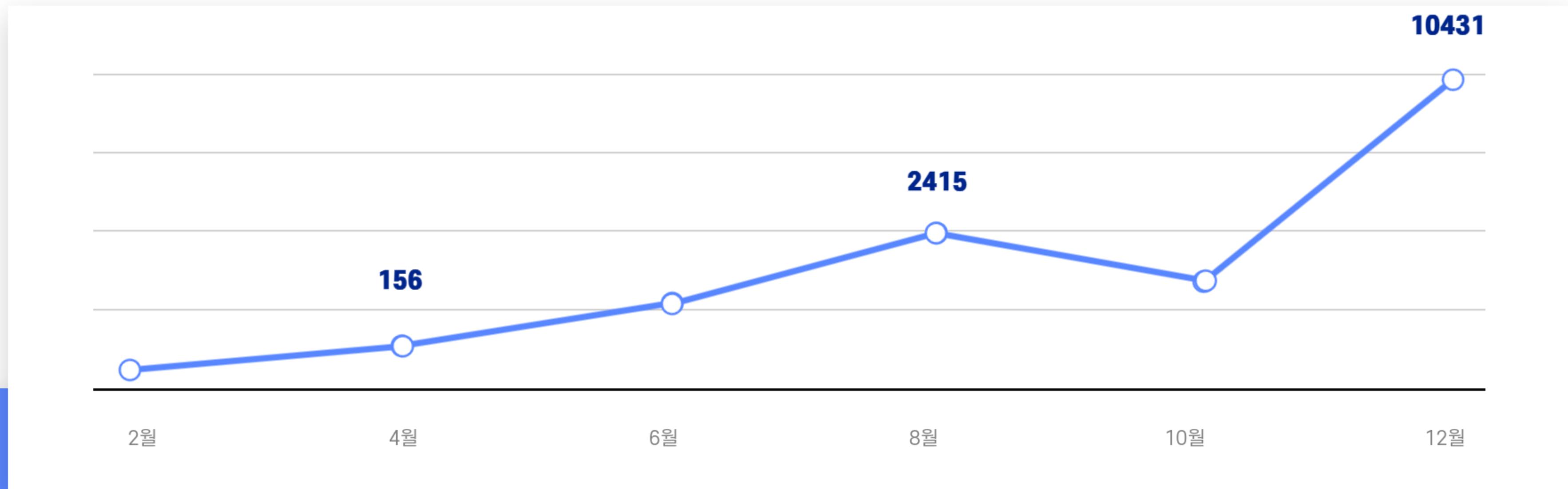
대중교통 이용량의 감소만큼 차량 이용량이 증가하지는 않았을까?



프로젝트 목적

PROJECT PURPOSE

[2020년 서울특별시 확진자 현황]



코로나의 지속적인 감염 폭발 현상 가운데
인구 이동량 추이를 밝혀보자 한다.

대중교통 이용량은 감소한 것이 맞는지,
차량 이동량은 감소하였는지,
코로나 확진자 현황과 인구 이동량의 변화 간 어떤 관계가 존재하는지.

진행 계획

PROGRESS PLAN

2019년 대비
2020년 대중교통
이용량의 변화

2019년 대비
2020년 차량
통행량의 변화

대중교통 이용량
및 차량 통행량 간
비교 분석

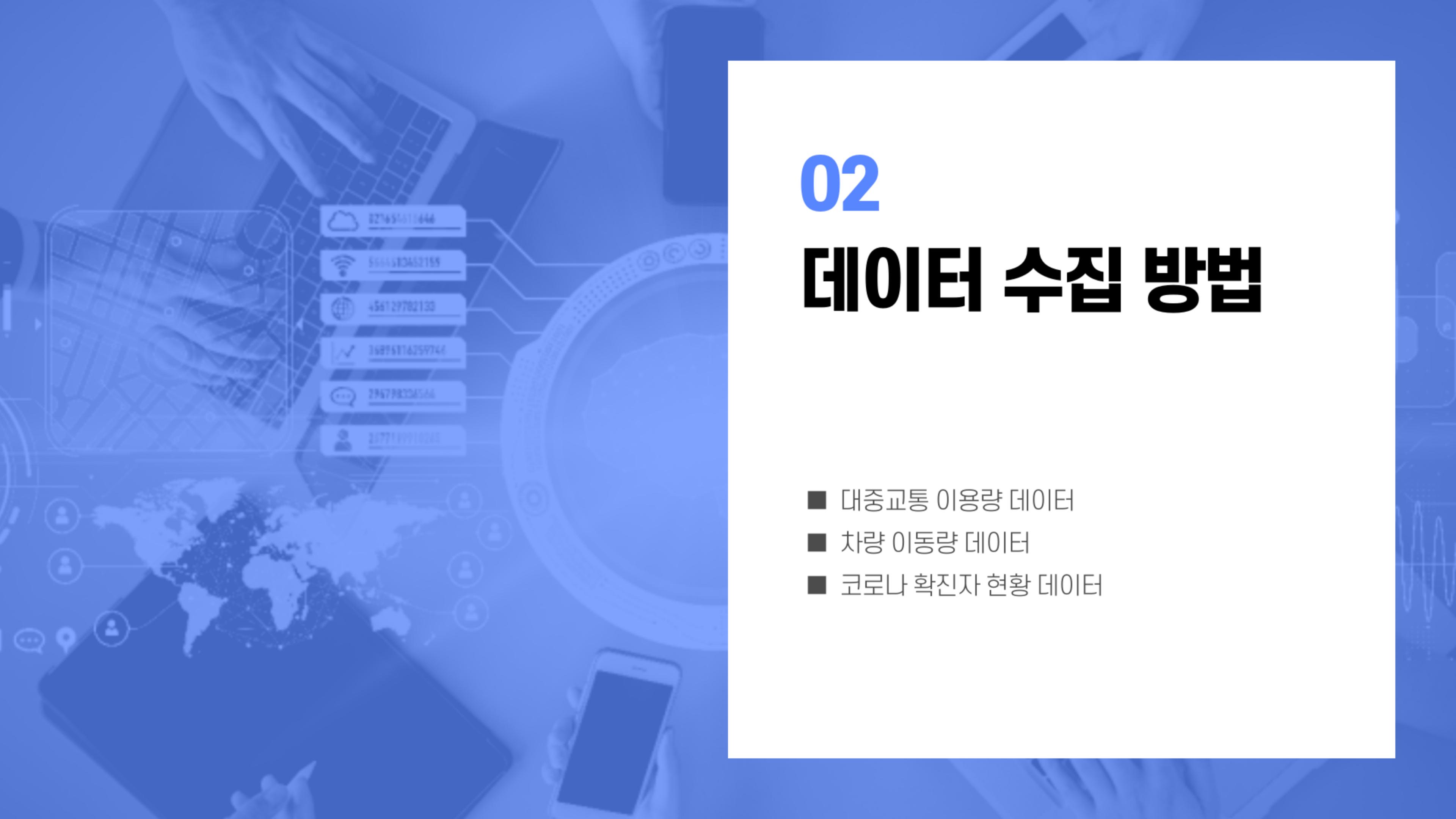
코로나 확진자 수
대비 대중교통 및
차량 이동량 변화

2019, 2020 전체 합계 및 변화,
월별 / 요일별(주말/평일) / 지역별
이용량 년도 별 비교

2019, 2020 전체 합계 및 변화,
월별 / 요일별(주말/평일) / 지역별
이용량 년도 별 비교

월별, 지역별 이동량 데이터 중
수단 간 공통점 및 차이점,
그 외 기타 특이점 비교 분석

코로나 확진자 수에 따른
교통량 비교,
지역별 교통량에 따른
코로나 확진자 수 상관관계



02

데이터 수집 방법

- 대중교통 이용량 데이터
- 차량 이동량 데이터
- 코로나 확진자 현황 데이터

데이터 수집 방법

HOW TO COLLECT DATA



대중교통 이용량

약 2년 간의 서울특별시 내 대중교통 이용량
데이터는 웹크롤링을 이용하여 수집



차량 통행량

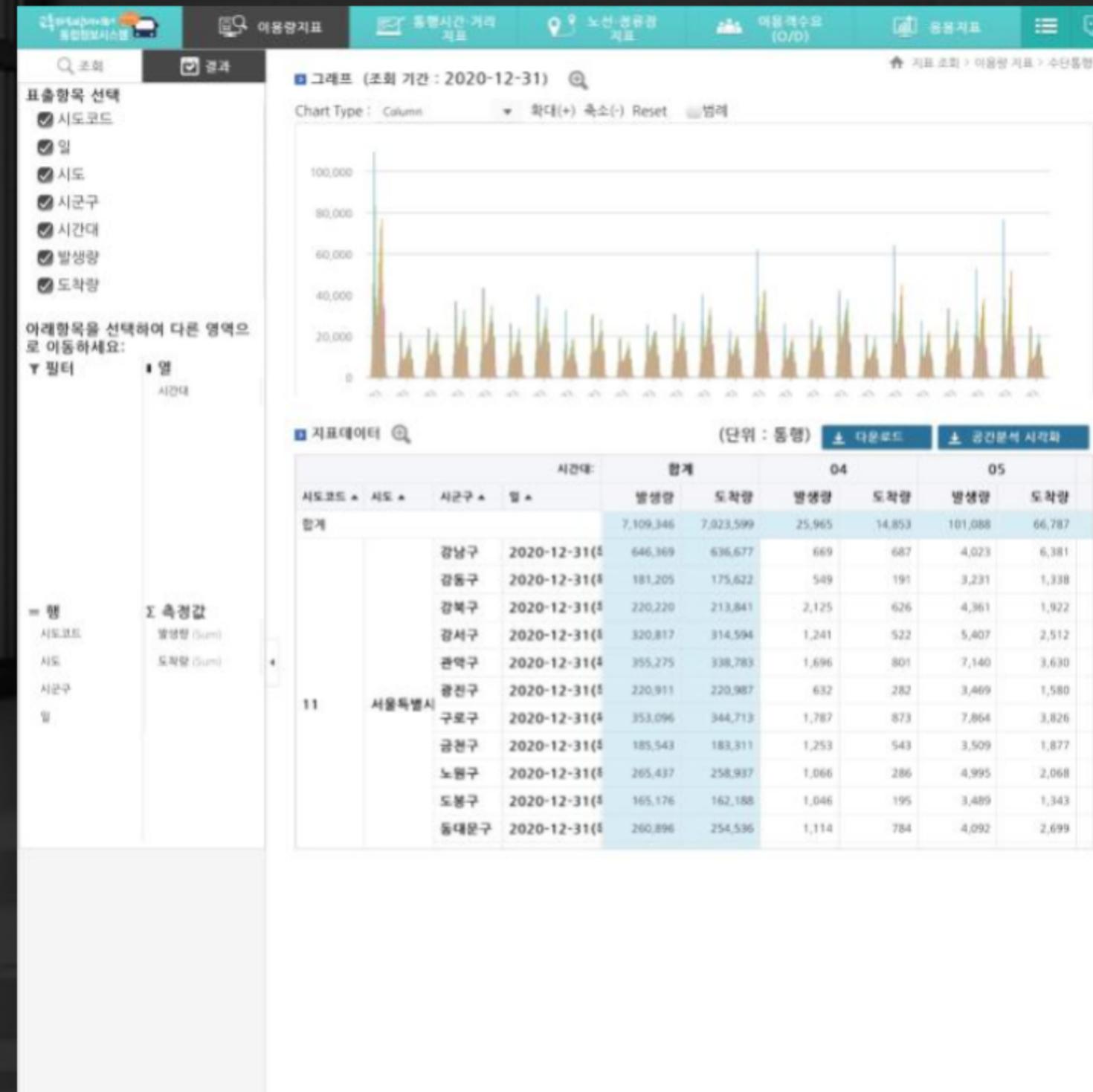
서울특별시 내 차량 통행량은
엑셀 파일을 다운로드하여 CSV로 변환



코로나 확진자 현황

코로나 확진자 현황에 관한 수치는
공공데이터 CSV파일을 다운로드

대중교통 이용량



교통카드 빅데이터 통합정보시스템

홈페이지를 이용한 웹크롤링

<https://stcis.go.kr/pivotIndi/wpsPivotIndicator.do?siteGb=P&indiClss=IC01>라는 교통카드 빅데이터 통합정보시스템 홈페이지에는 공항버스, 광역버스, 마을버스, 시내버스, 공항버스, 도시철도 이용량을 1일 단위별로 검색이 가능하다. 이에 2019년 01월 01일부터 2020년 12월 31일까지의 데이터를 각각 1년치씩 따로 크롤링하여 CSV파일로 변환하였다. 더불어 Pandas의 pivot 함수를 통해 월별-대중교통이용량, 지역별-대중교통이용량을 나타내었다.

* 결측데이터 없음

차량 통행량

The screenshot shows the SEOUL TOPIS website interface. The top navigation bar includes links for '서울소식' (Seoul News), '응답소' (Feedback Center), and '정보공개' (信息公开). The main content area is titled '교통량 정보' (Traffic Volume Information) under the '서울시 교통정보' (Seoul Traffic Information) section. A sub-menu for '교통량 정보' is active. Other menu items include '속도 정보' (Speed Information), '대중교통 정보' (Public Transport Information), '교통이용 통계' (Traffic Use Statistics), '기타 정보' (Other Information), '교통 현황' (Traffic Status), and 'OPEN API'. Below the menu, a section titled '지점별 일자별 교통량' (Point-by-point Daily Traffic Volume) displays download statistics for January through October 2021. The data is summarized in the following table:

월	다운로드 수
1월	616
2월	325
3월	0
4월	0
7월	0
8월	0
9월	0
10월	0

At the bottom, there are links for 'Home', 'Seoul Traffic Information Portal', 'Alert Center', 'TOPIS Introduction', and 'How to Find Us'. A footer note states: '045151 서울특별시 종로구 익수궁길 15(서소문동 37, 서소문정사 1동 14호) | 전화번호 - 대산콜센터 : 120'.

서울특별시 교통량정보 지점별, 일자별 교통량

년도별 월별 데이터 다운로드

https://topis.seoul.go.kr/refRoom/openRefRoom_2.do

서울시 교통량정보 제공 사이트에서는
년도를 지정하면 월별로 각 톨게이트 지점 당 차량 이용량 조회가 가능하다.
이에 2019년 01월 01일부터 2020년 12월 31일까지의 데이터에 관해
각각 1년치 씩 월별로 EXCEL파일을 다운로드하여 CSV파일로 병합하였다.
더불어 Pandas의 pivot 함수를 통해 월별-대중교통이용량,
지역별-대중교통이용량을 나타내었다.

* 결측데이터 없음

코로나 확진자 현황

The screenshot shows the homepage of the Public Data Portal. At the top, there's a search bar with the placeholder "어떤 공공데이터를 찾으시나요?" and a magnifying glass icon. Below the search bar is a section titled "검색조건" (Search Conditions) with dropdown menus for "분류체계" (Classification System), "서비스유형" (Service Type), and "확장자" (File Extension). A large blue banner below these sections is titled "인기검색어" (Popular Search Terms) and lists "1. 코로나". Underneath this banner is a section titled "테마별" (By Theme) with categories like "카테고리별" (Category), "국가중점데이터별" (National Key Data), and "제공기관유형별" (Provider Agency Type). Below this is a row of icons representing various government departments: 교육 (Education), 국토관리 (Land Management), 공공행정 (Public Administration), 재정금융 (Finance), 산업고용 (Industry Employment), 사회복지 (Social Welfare), 식품건강 (Food Safety), 문화관광 (Culture and Tourism), 보건의료 (Healthcare), 재난안전 (Disaster Safety), and 교통 (Transportation). At the bottom of the page, there are sections for "인기 데이터" (Popular Data) and "파일 데이터" (File Data), along with a link to "오픈 API". A footer at the very bottom lists three items: 1. 소상공인시장진흥공단_상가(상권)정보, 2. 도로교통공단_가해운전자 연령층별 월별 교통사고 통계, and 3. 도로교통공단_나고야항별 교통사고 통계.

공공데이터 포털

서울특별시 코로나 확진자 현황

2020년도 전체 데이터 파일 다운로드

<https://www.data.go.kr/>

공공데이터포털 사이트에서는 '코로나 확진자'로 검색하면
여러 관련 데이터와 더불어 서울시 코로나 확진자 현황 데이터를
다운로드 받을 수 있다.

이에 2020년 01월 01일부터 2020년 12월 31일까지의
서울특별시 내 코로나 확진자 데이터를 CSV파일로 다운로드 받았다.

* 결측데이터 없음

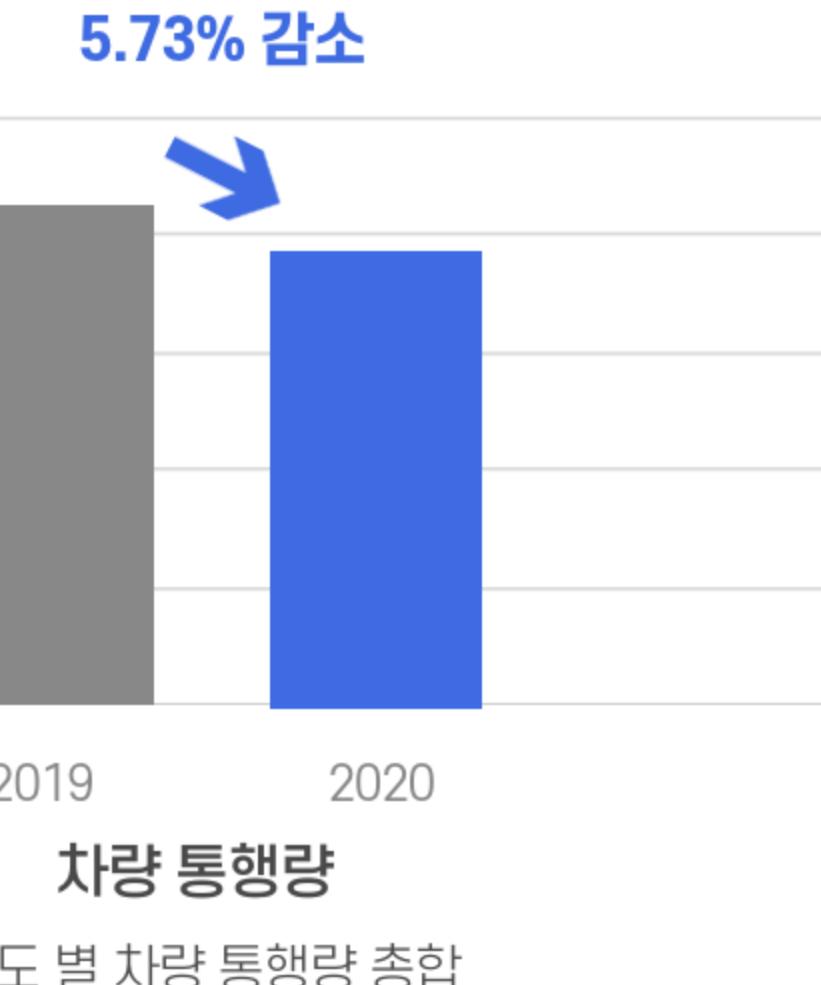
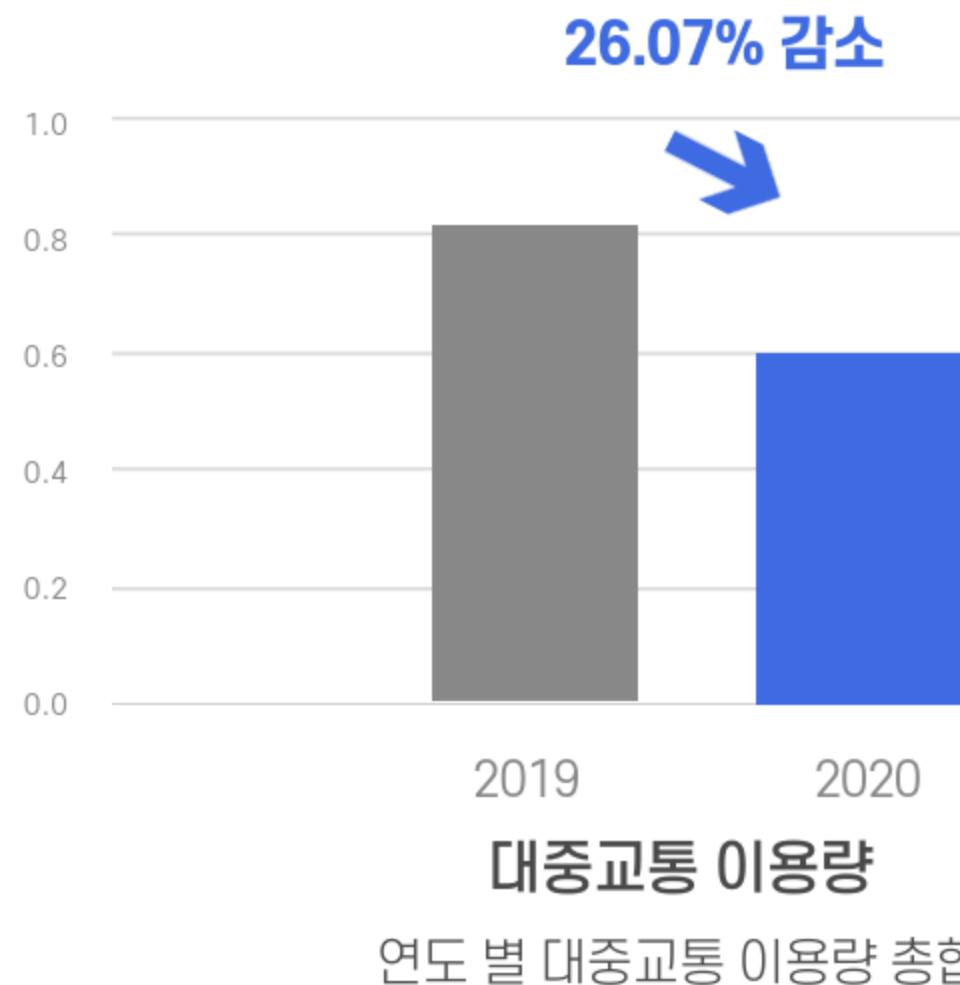


03

데이터 처리 및 분석

- 2019, 2020 서울특별시 내 전체 교통량 비교
- 월별 교통량 비교 : 대중교통 / 차량
- 요일별(주말/평일) 교통량 비교 : 대중교통 / 차량
- 지역별 교통량 비교 : 대중교통 / 차량
- 코로나 확진자 수에 따른 교통량 비교
: 대중교통 / 차량

1) 2019, 2020 전체 교통량 비교

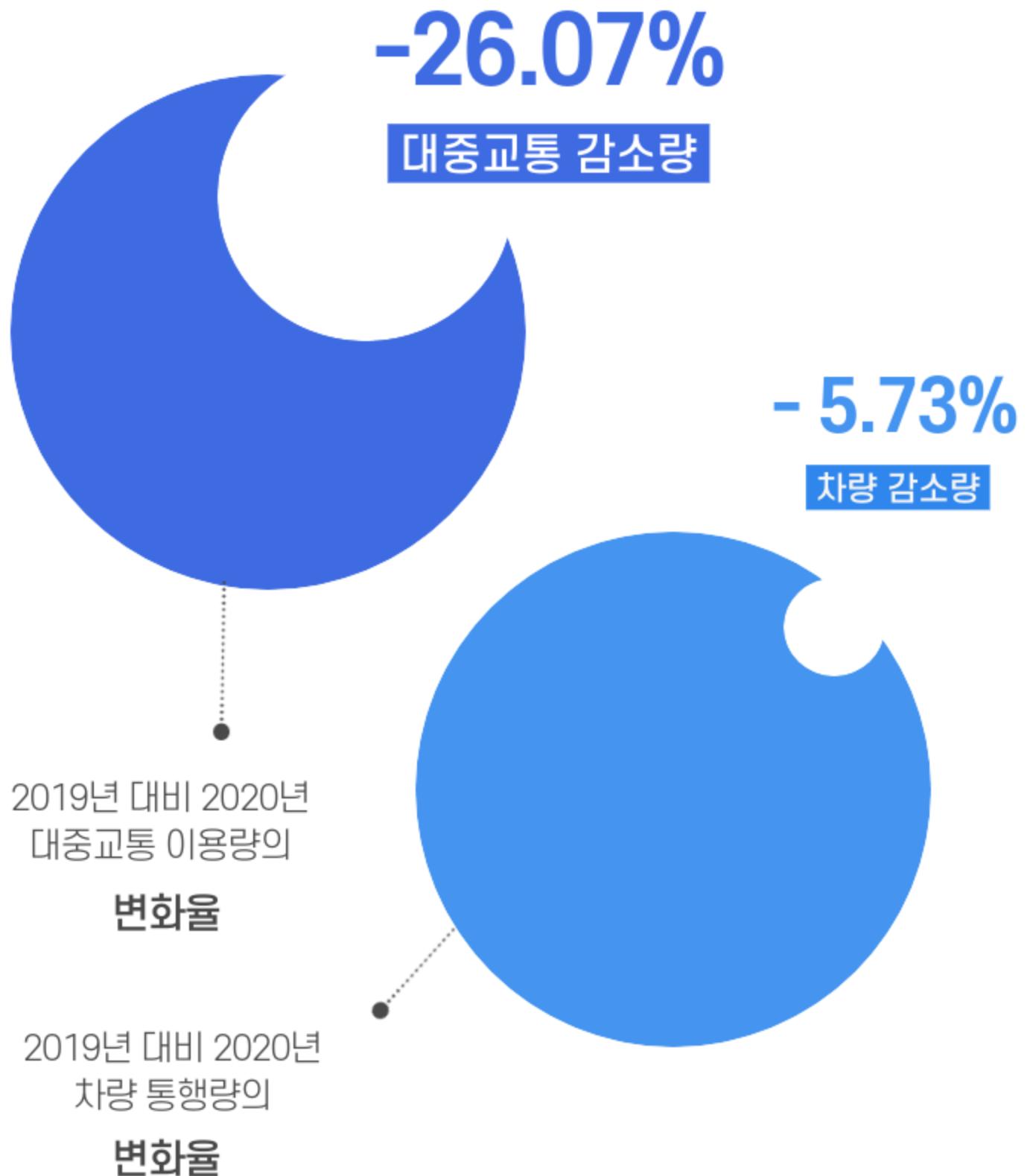


대중교통과 차량 모두 감소,
그러나 감소 폭은 상이함

대중교통에 비해 확연히 적은
감소폭의 차량 통행량

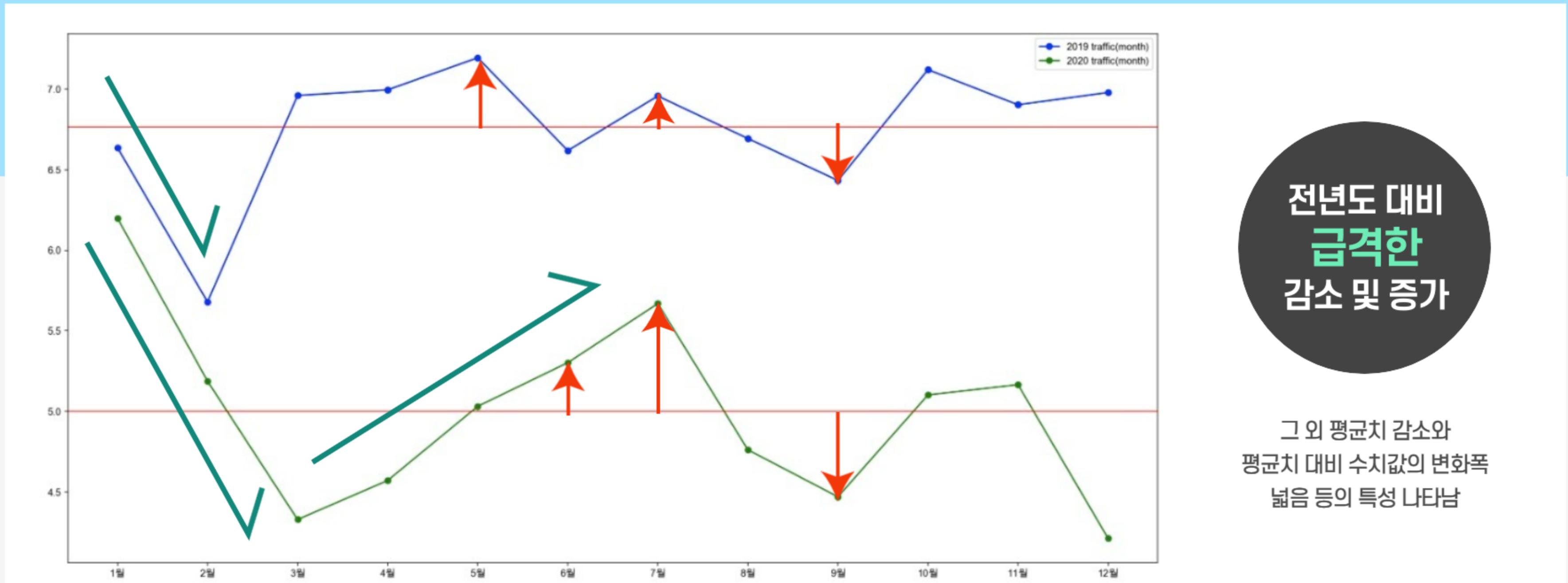
Insight

1. 대중교통과 차량 이용률 모두 전년도 대비 감소하였다.
인구의 이동량 또한 전년도 대비 감소한 것이 사실이다.
2. 따라서 '인구의 이동량은 그대로일 것이며 대중교통 이용량의 감소 폭 만큼 차량 이용량이 증가했을 것'이라는 가설은 설립하지 않는다.
3. 그러나 26.07%라는 대중교통 감소량 대비, 5.73%라는 차량 감소량으로 보았을 때
대중교통 이용률 대비 차량 이용률(선호도)은 비교적 건재하기 때문에
따라서 '차량이 대중교통을 일부분 대체하였다'고 볼 수 있다.



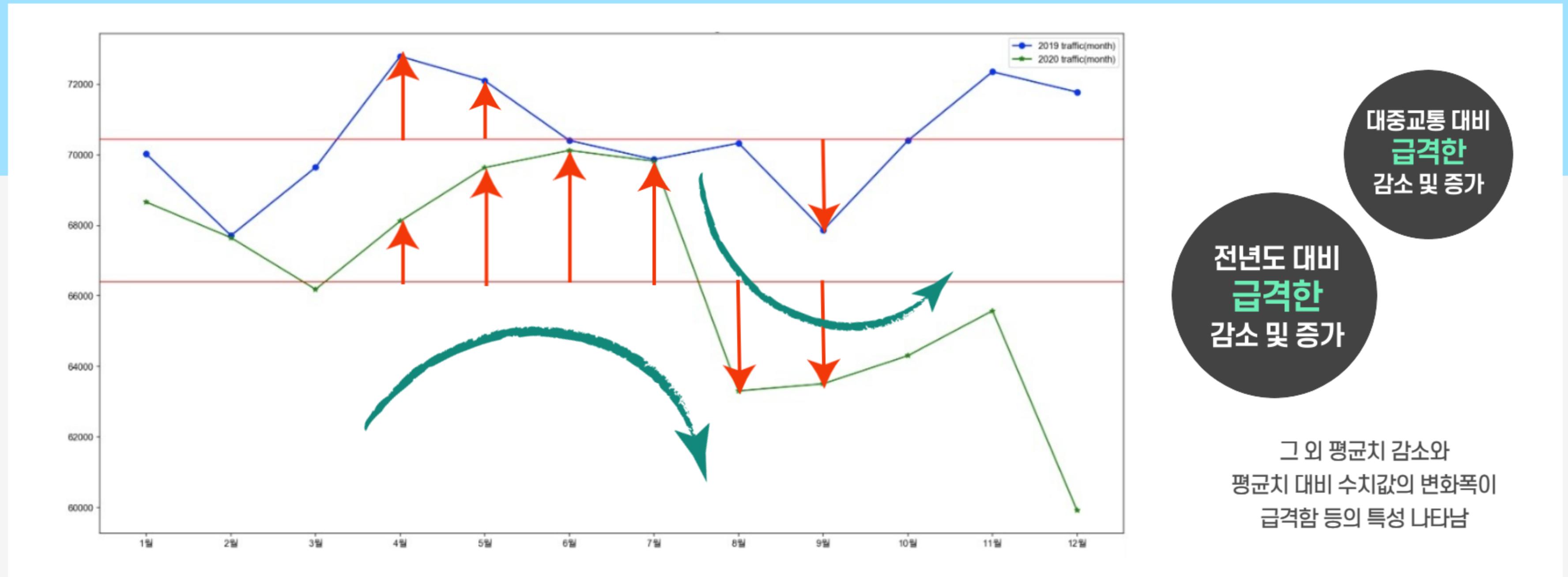
2-1) 2019, 2020년 월별 대중교통량 평균 및 추이

대중교통 이용량에 대해 월별 수치를 각각 연도 별로 나타내었을 때, 2019년에 비해 2020년의 평균치가 낮고
감소 혹은 증가 추세가 더욱 오래 지속됨을 알 수 있으며, 평균치 대비 변화폭이 넓다.



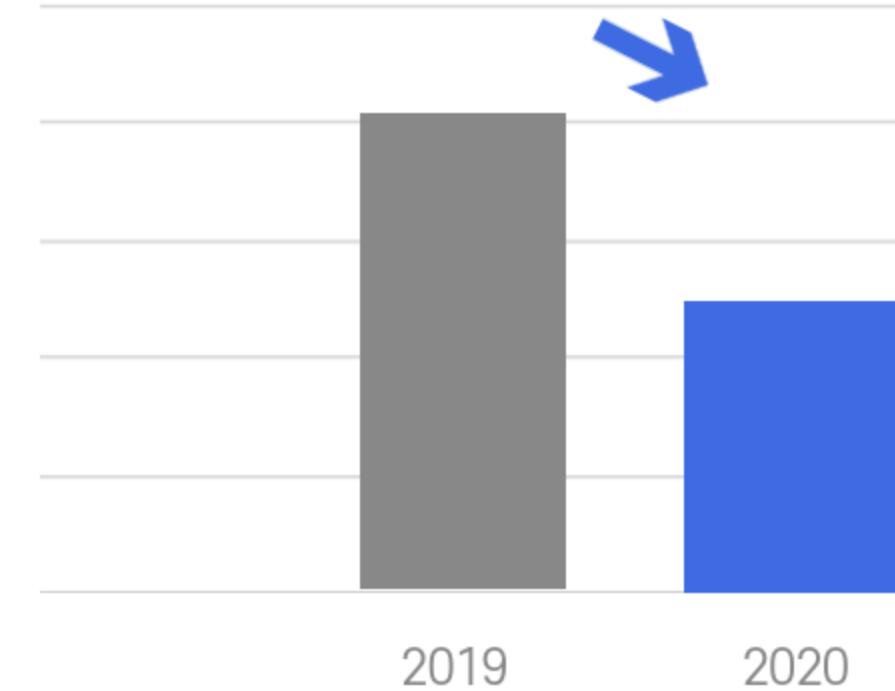
2-2) 2019, 2020년 월별 차량통행량 평균 및 추이

차량 통행량에 대해 월별 수치를 각각 연도 별로 나타내었을 때, 2019년에 비해 2020년의 평균치가 낮고
감소 혹은 증가 추세가 매우 급격하여 더욱 오래 지속됨을 알 수 있으며, 평균치 대비 변화폭이 넓다.

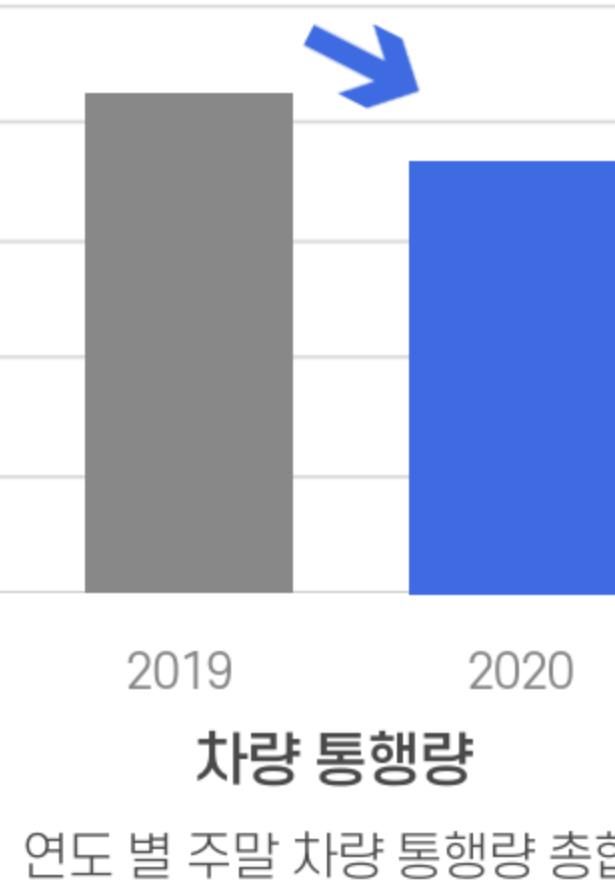


3-1) 2019, 2020 요일별(주말) 교통량 비교

35.35% 감소



10.27% 감소



대중교통과 차량 모두 감소,
그러나 변화폭은 상이함

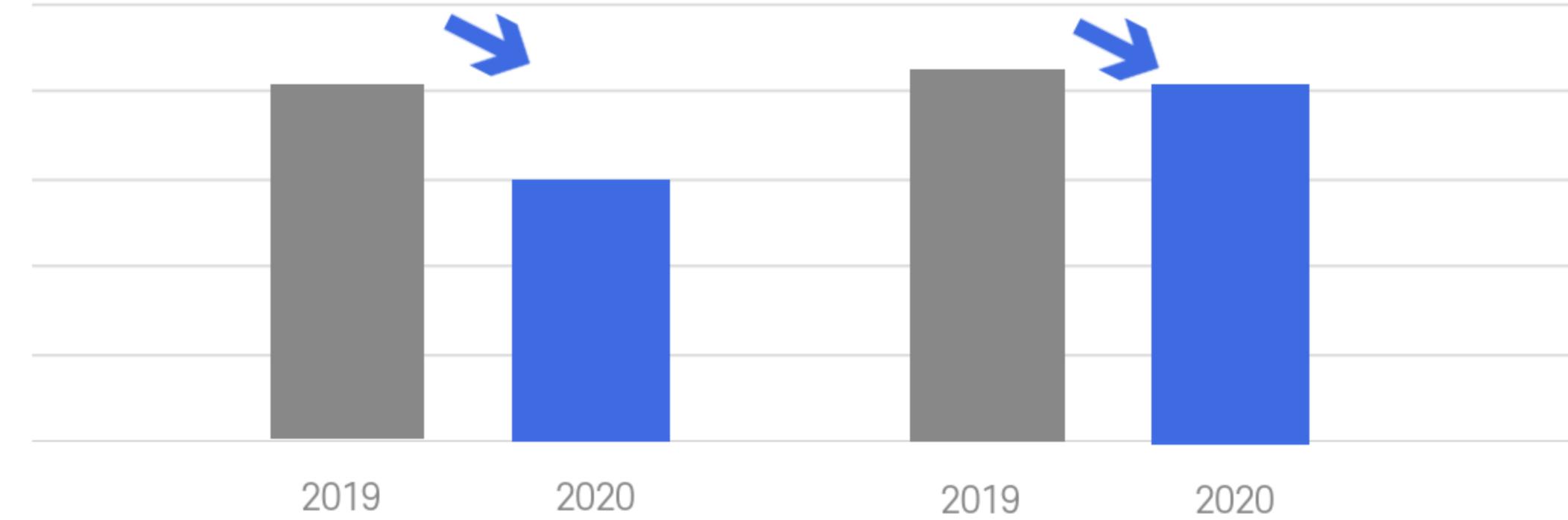
대중교통에 비해 약 1/3 정도
감소폭의 차량 통행량

연도 별 주말 대중교통 이용량 총합

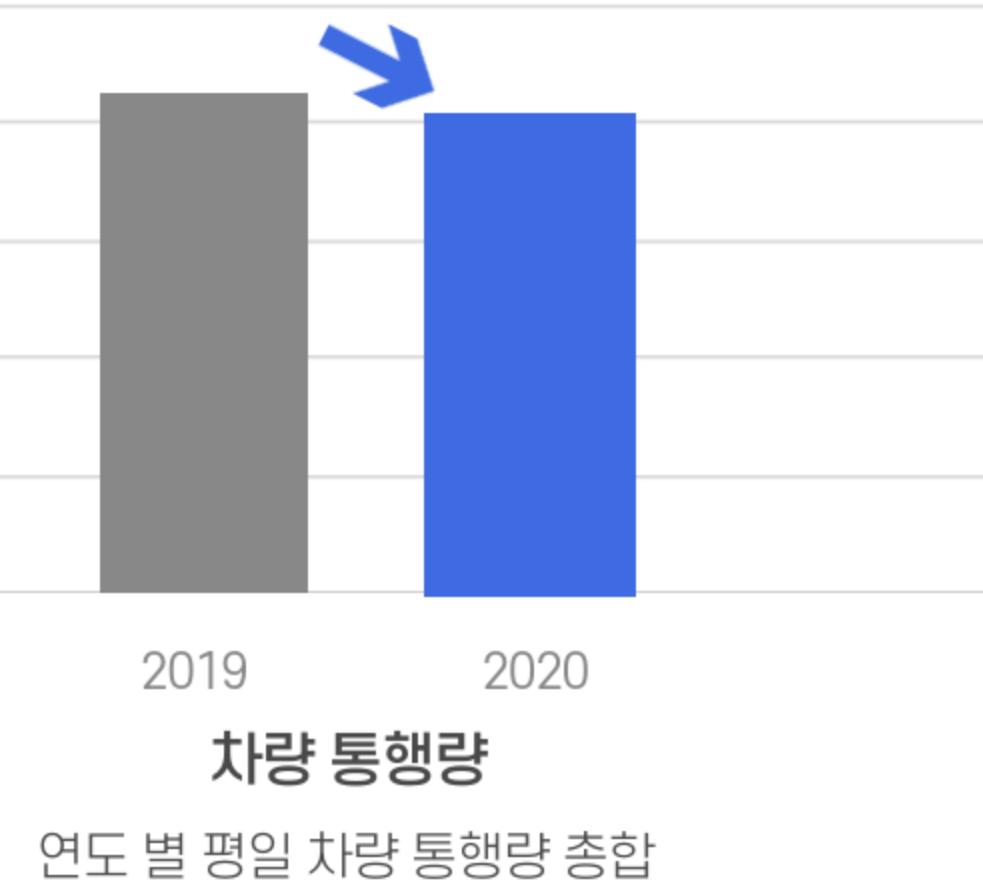
연도 별 주말 차량 통행량 총합

3-2) 2019, 2020 요일별(평일) 교통량 비교

23.57% 감소



3.79% 감소



대중교통과 차량 모두 감소,
그러나 변화폭은 상이함

대중교통에 비해 약 1/7 정도
감소폭의 차량 통행량

Insight

2019년 -> 2020년
평일 대중교통 감소율 대비

차량통행량 변화율

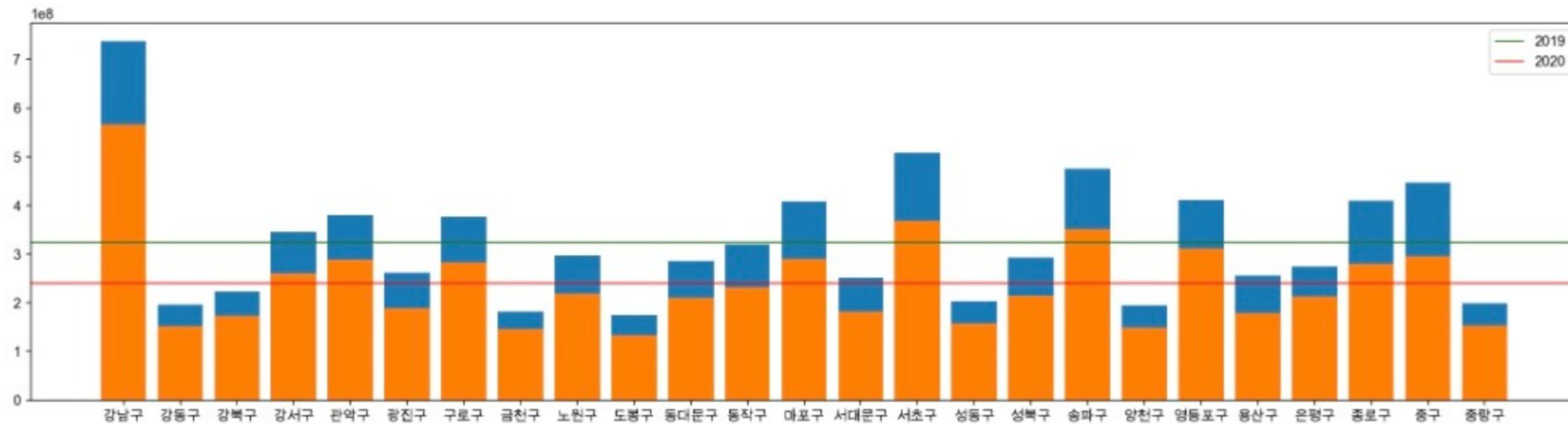
1. 2019년도 대비 2020년의 **주말** 교통량은 **대중교통 35.35%, 차량 10.27%** 감소하였으며 차량 감소량은 대중교통 감소량의 **1/3 정도**임을 알 수 있다.
2. 2019년도 대비 2020년의 **평일** 교통량은 **대중교통 23.57%, 차량 3.79%** 감소하였으며 차량 감소량은 대중교통 감소량의 **1/7 정도**임을 알 수 있다.
3. 즉 요일과 관계없이 **교통량 자체는 전체 감소하였으나** 대중교통과 차량 모두 평일 대비 **주말의 감소량이 높으며** 평일과 주말 모두 **차량 대비 대중교통의 감소량이 월등히 높다.**
4. 이를 종합했을 때, 서울시민은 대체적으로 이동량 자체가 감소하였으나 각자의 이유로(재택근무, 여가 시 대중교통보다 차량 선호 등)로 **차량에 비해 대중교통을 보다 기피하게 된 것을 유추해볼 수 있다.**



2019년 -> 2020년
주말 대중교통 감소율 대비
차량통행량 변화율

4-1) 2019, 2020년 지역별 대중교통량 평균 및 추이

대중교통 이용량에 대해 지역별 수치를 각각 연도 별로 나타내었을 때, 2019년에 비해 2020년의 평균치가 낮고
2019년 대비 2020년의 감소치는 각 지역의 기존 수치에 대체로 비례하는 모습을 보인다.



2019년 지역별 대중교통 이용량의 평균은 324,652,736대이며, 2020년 지역별 대중교통 이용량의 평균은 240,028,655 대이다.

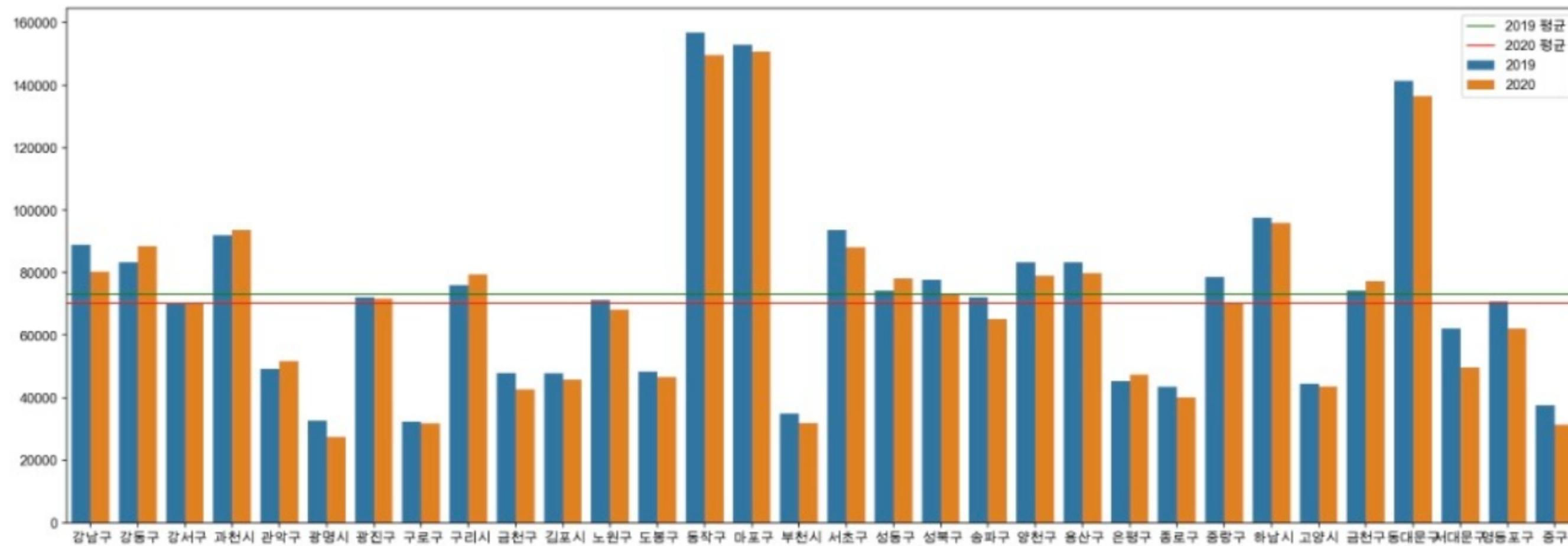
평균이 감소한 것과 같은 맥락으로 각 지역별 대중교통 이용량 또한 모두 감소하였으며

각 지역별 감소량은 해당 지역의 2019년 통행량에 비례하는 모습을 보인다.

EX. 강남구 : 기존 통행량 많음 → 감소량 많음 / 강동구 : 기존 통행량 강남구 대비 적음 → 감소량 적음.

4-2) 2019, 2020년 지역별 차량통행량 평균 및 추이

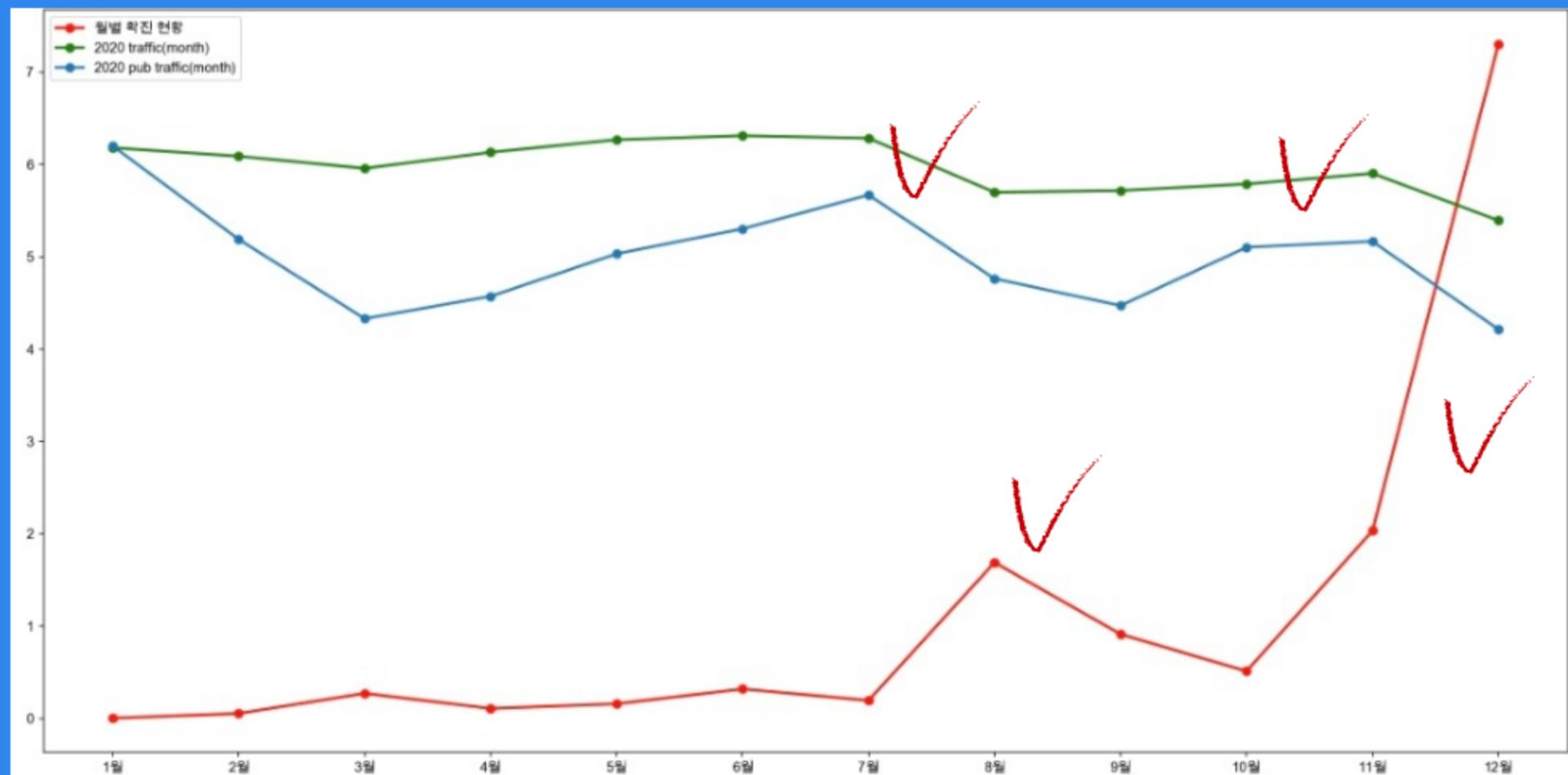
차량통행량에 대해 지역별 수치를 각각 연도 별로 나타내었을 때, 2019년에 비해 2020년의 평균치는 낮지만
2019년 대비 2020년의 차량통행량은 지역별로 감소량이 대체로 적거나 오히려 증가하는 모습을 보인다.



2019년 지역별 차량통행량의 평균은 73,021 대이며, 2020년 지역별 대중교통 이용량의 평균은 70,227 대이다.
평균이 감소한 것과는 별개로 각 지역별 대중교통 이용량은 미세하게 감소하였거나 오히려 다소 증가한 지역도 존재한다.

EX. 강남구 : 전년도 대비 소량 감소 / 강동구 : 전년도 대비 소량 증가

5-1) 교통수단별 추이와 코로나 확진자 추이 간 비교



● 월별 확진자 현황

● 2020년 월별 대중교통량

● 2020년 월별 차량통행량



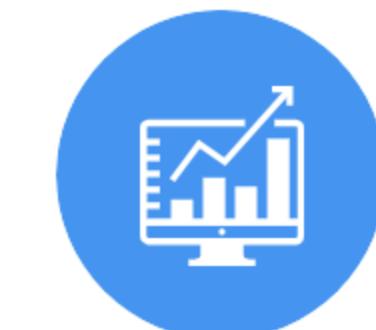
차량통행량 증감폭

차량통행량은 대체로 월별 증가 혹은 감소 정도의 변화가 크지 않다.



대중교통량 증감폭

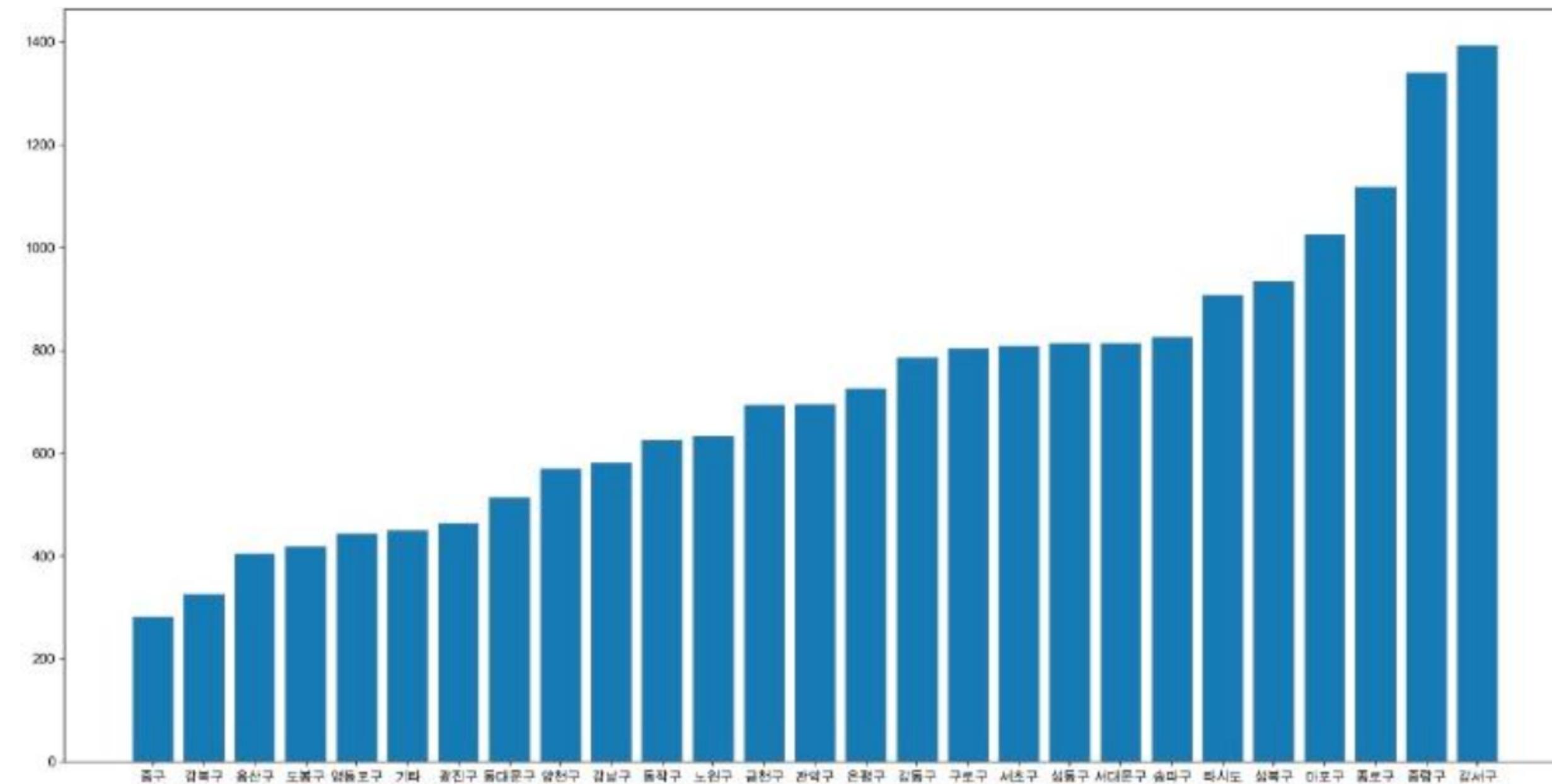
대중교통량은 증가 혹은 감소 정도의 변화 폭이 비교적 역동적이다.



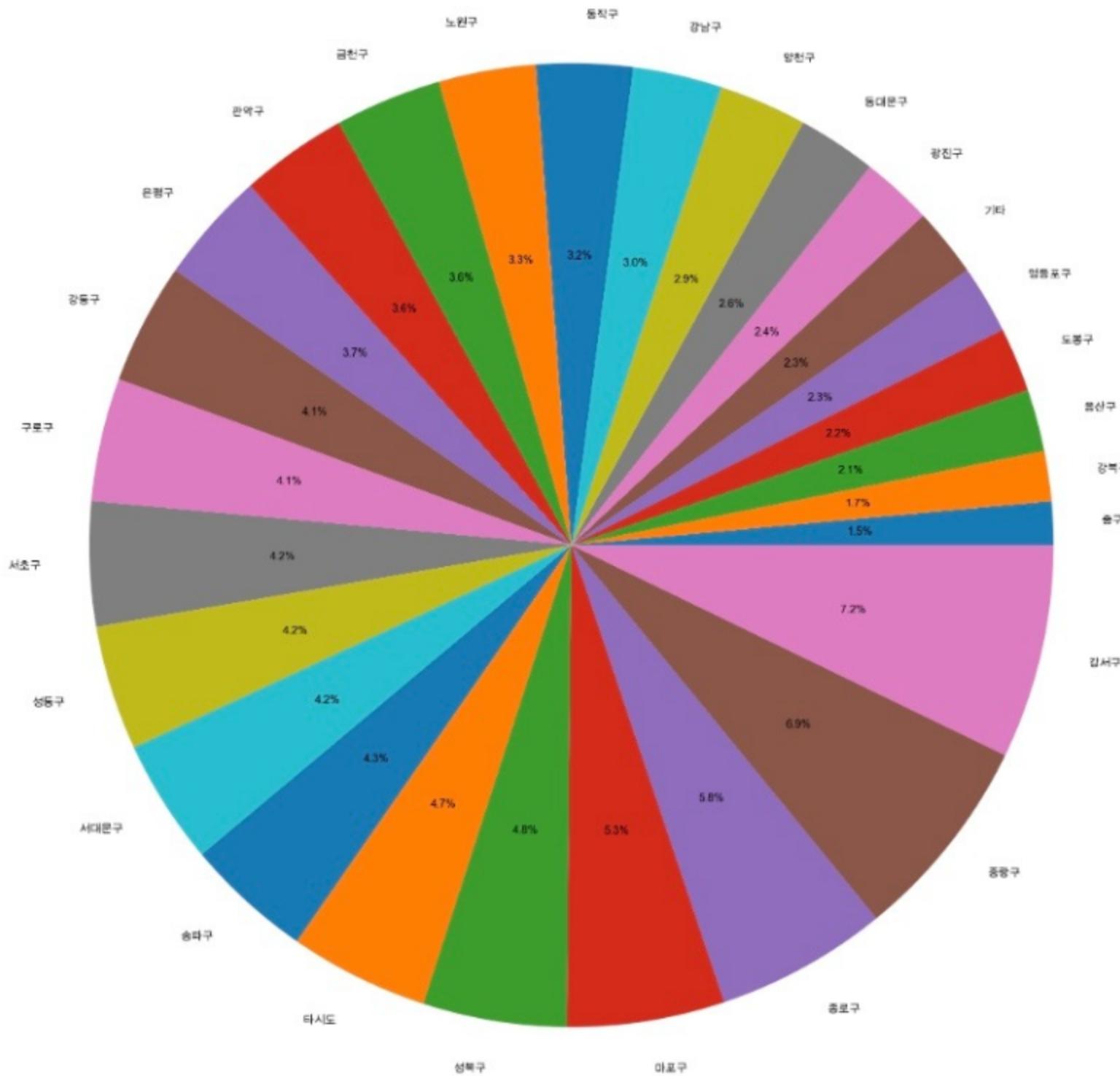
월별 확진자 급증 시기

대중교통량이 증가 추세를 보이면 그 다음달의 월별 확진자가 급증하는 것이 특징

5-2) 지역별 추이와 코로나 확진자 현황



5-3) 서울 내 지역 별 코로나 확진자 분포

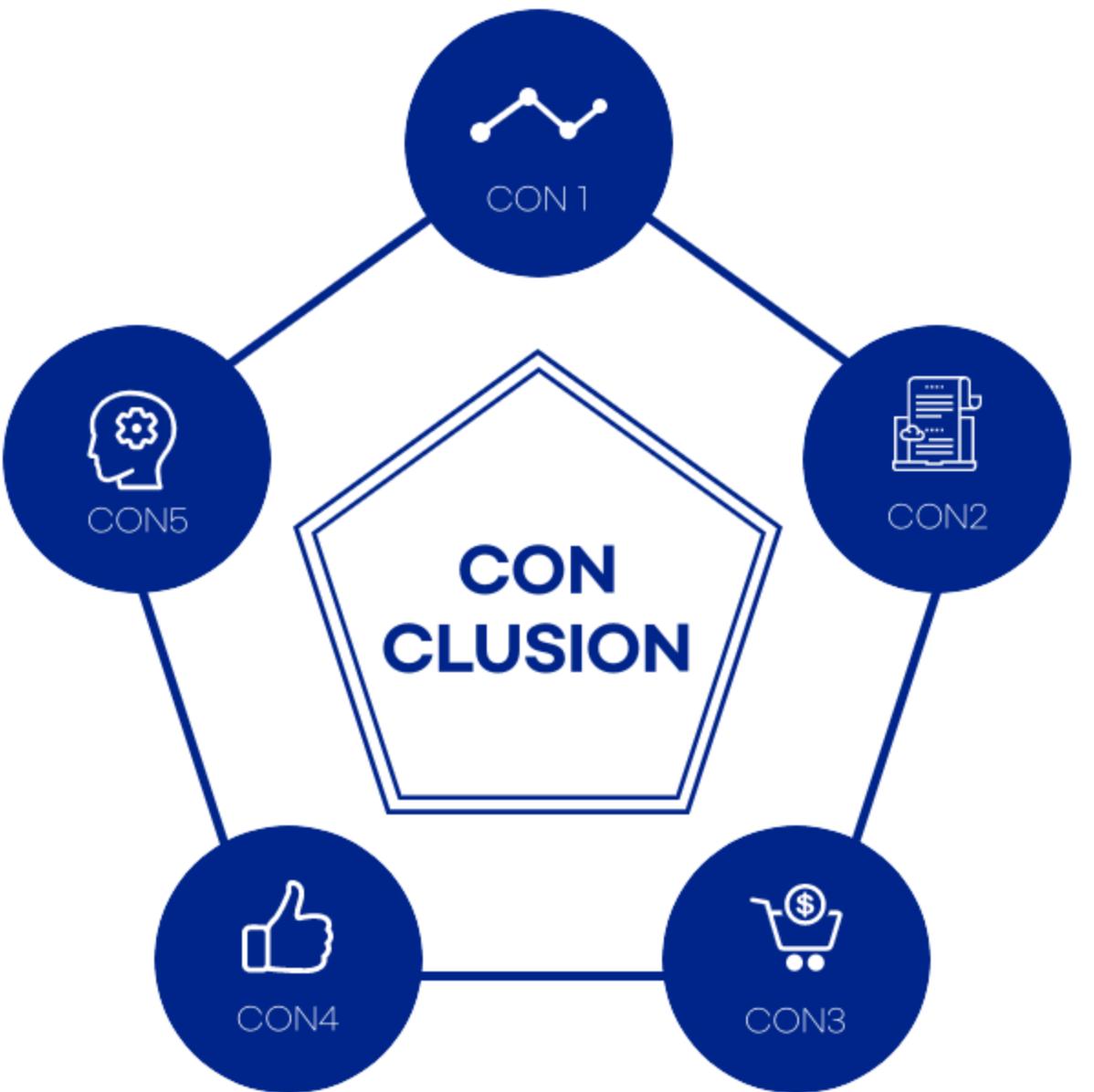


2020년 1년 간 전체 코로나 확진자 발생량 전체 중
각 지역 별 분포 정도를 시각화하면

1.5%로 가장 적은 분포를 보이는 중구부터
7.2%로 가장 많은 분포를 보이는 강서구까지
각 지역의 분포 정도를 한 눈에 볼 수 있다.

이 때 최소 분포지역과 최대 분포지역 간의 차이는
비교적 두드러지는 편이나,
전체적으로 맞닿아있는 지역 간의 차이는
비교적 크지 않은 수치를 보이며 이어지는 것으로 보인다.

결론



01

대중교통과 차량 이용률 모두 전년도 대비 감소하였다.
그러나 26.07%라는 대중교통 감소량 대비, 5.73%라는 차량 감소량으로
보았을 때 **차량이 대중교통을 일부분 대체하였다**고 볼 수 있다.

02

월별 대중교통량 평균 및 추이는 2019년에 비해 2020년의 평균치가 낮고
감소 혹은 증가 추세가 매우 급격하여 **차량이 대중교통보다 더욱 변화폭이 크다**.

03

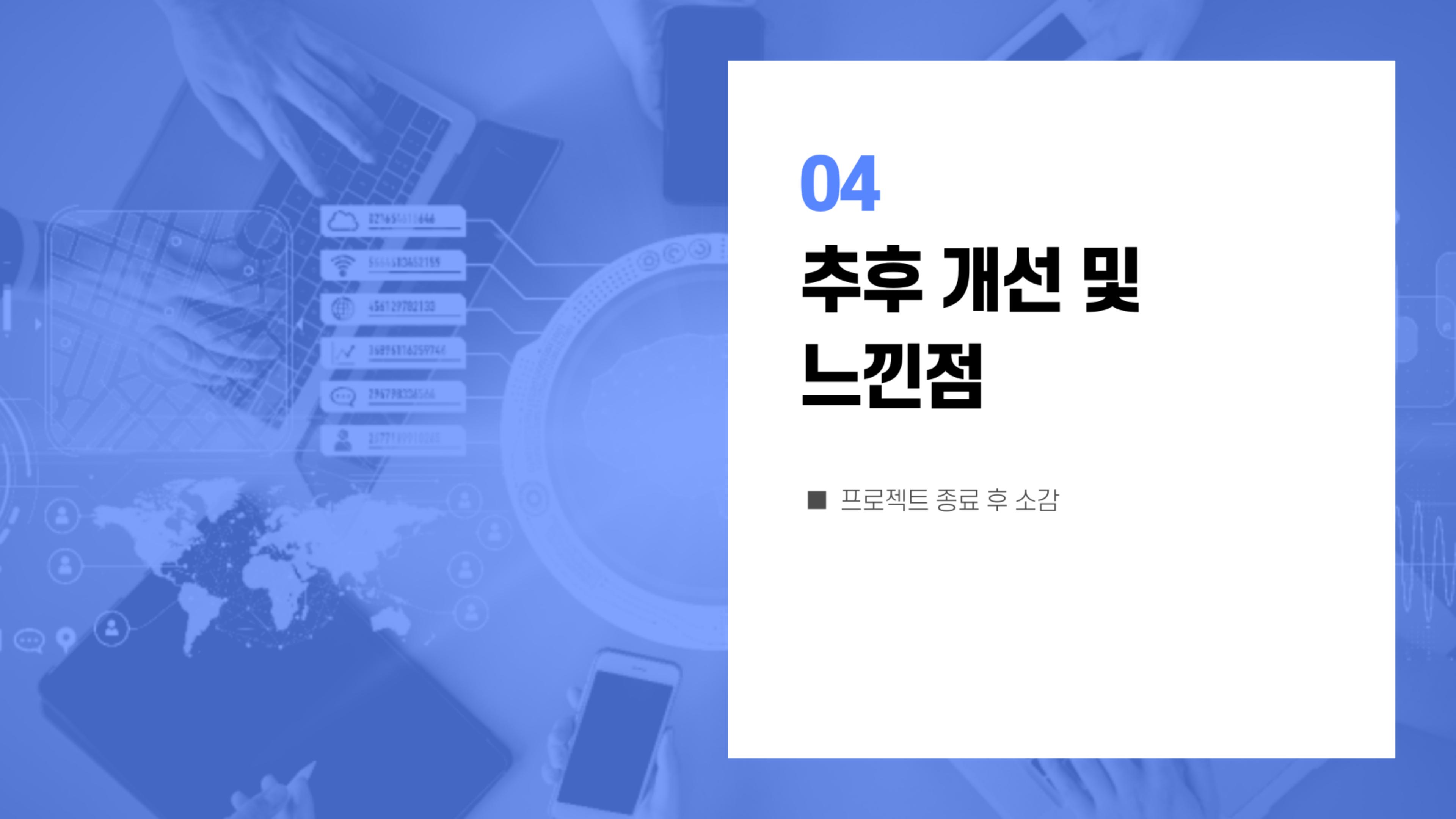
요일별 대중교통량 평균 및 추이는 대중교통과 차량 모두 **평일 대비 주말의**
감소량이 높으며 평일과 주말 모두 차량 대비 대중교통의 감소량이 월등히 높다.

04

지역별 대중교통 이용량은 대중교통의 경우 전년도 대비 전 지역 고르게
감소하였으나, **차량의 경우 '감소, 유지, 증가'로 다양한 변화 추이**를 보였다.

05

교통수단별 추이와 코로나 확진자 추이를 비교해보았을 때 **대중교통량이**
증가 추세를 보이면 그 다음달의 월별 확진자가 급증하는 특징이 나타났다.



04

추후 개선 및 느낀점

■ 프로젝트 종료 후 소감

추후 개선점 및 느낀점

관련 데이터에 대한 갈증과
예상과 다른 상관 관계의 부재



DATA



CORRELATION

대중교통과 차량이동량 데이터 모두 **정류장 위치** 혹은 툴게이트의 **정확한 위도와 경도**를 포함하고 있었기 때문에 지역 별 **사회경제적인 데이터**(EX. 카드 데이터)와 결합한다면, 교통수단을 이용한 인구 외의 인구 이동까지 가능하여 서울시 지도에 지역별 인구 이동 그리고 코로나의 추이를 정밀하게 추출할 수 있었을텐데 카드데이터와 같은 유의미한 데이터를 찾아내는 것에 어려움을 겪어 **원하는 바를 모두 시도해보지 못한 아쉬움**이 남았다.

월별 교통량 - 코로나 확진자 수 혹은 **지역별 교통량 - 코로나 확진자 수** 와 같은 관계에 있어서 당연히 정의 관계(비례)가 성립할 것이라 생각했지만, 2020년 데이터 전체를 기준으로 보았을 때에는 오히려 상관관계가 존재하지 않는 것으로 보였다. 프로젝트 전반적으로 **예상과 다른 측면이 다수 발견되어 다양한 인사이트를 도출**하게 되었다.



THE END

감사합니다.